Downing College

University of Cambridge

Physiology, Development and Neuroscience

Part II
Experimental Research Project

Modulation of Visual Search Patterns After Introduction of Effective Motor Noise

Author:

Chris Barnes

Email Address:

cb619@cam.ac.uk

Supervisors:
Dr. Gergő Órban
Prof. Daniel Wolpert

Computational and Biological
Learning Laboratory

Department of Engineering

I confirm that the material in this dissertation is not copied from any published material, unless it is clearly identified as such and a full source reference given.

I confirm that the material in this dissertation is not a paraphrase or abstract of any published material, unless it is clearly identified as such and a full source reference given.

I understand that the examiners may use any means available to test for plagiarism within my dissertation.

I confirm that, other than where indicated, this dissertation is my own work.

I confirm that the word count is ....

Signature ....

Name: Christopher Lloyd Barnes

April 26, 2012

Summary

Visual search is one process used in building up an internal representation of a scene in order to localise a target (in order to search effectively, information must be integrated across a number of saccades). This is usually achieved by collecting visual data from fixations on a number of discrete points in the scene in what is suggested to be a Bayesian fashion: the inhomogeneity of spatial acuity across the visual field (due to foveation and physical aberrations) is taken into account and each fixation point is selected to add the most information to the internal representation. The search strategy is optimised to minimise costs such as time spent and energy expended.

Eye movements are almost entirely ballistic in nature. In order to be able to slot the information received from a new fixation point straight into the current estimation of the scene, the direction and amplitude of each eye movement needs to correspond closely to the motor system's plan: otherwise, integration of perisaccadic information and the formation of an eye position-independent representation of the scene is not possible. If these eye movements were inaccurate and unreliable, valuable time and neural resources would need to be spent aligning the new data with the current internal image (particularly in early stages of the search), by lining up the edges; like a jigsaw puzzle.

In this research project, the degree to motor errors contribute to altered search patterns was investigated. A high-speed computer monitor was used to present a randomised visual scene comprised of many homogeneous distractors and one target stimulus. Gaze position was recorded using a 1000Hz infra-red camera. To produce the illusion of a defective occulomotor system, we introduced a gaze-contingent display-shift paradigm during the periods of saccadic suppression- effective blindness- associated with rapid eye movements. Using this technique, saccades made in certain directions (in this case, any saccade made in a rightward direction, i.e. a bearing of 0 to 180°) could be 'penalised' by triggering a display offset. In post-hoc processing, various features of the visual search were assessed. We were interested in whether the effective motor noise had any impact on the statistics of the search; firstly in terms of saccade timing, and secondly in whether global orientation patterns reflected the statistics of the effective motor noise.

Analysis of the results suggested that the introduction of effective motor noise did indeed produce changes in the visual search: we demonstrated that the display shift increases the duration of the fixation immediately following the offset ($p < 0.01$). Further analysis was performed in order to assess whether this change was a learned change in strategy or an immediate response (it seems to be immediate or at least very fast), and to see whether it was the product of an error signal

which, over time, could change the distribution of saccade directions in order to avoid the error. We did not find evidence for changes in the orientation statistics of eye movements.

Contents

1   Background

1.1   Saccades

The visual system is very slow. Even the fastest human photoreceptors take on the order of 20ms to recover from firing an action potential. On the other hand, the occulomotor system is capable of being very fast, able to produce accelerations in the region of $10000°s^{-2}$. As such, rather than constrain the speed of the eye movement to the effective 'refresh rate' of the retina to reduce motion blurring, it seems adaptively beneficial to flick the gaze back and forth as quickly as possible, and ignore the intolerably blurred visual information accrued during the movement itself; perception of the visual world is built up from the snapshots of data accrued between such movements. These high-speed, highly-accurate, ballistic and discrete eye movements are called saccades, and give rise to a number of phenomena.

Firstly, because there is virtually no visual feedback available as the gaze traverses the visual field, the movement is almost entirely ballistic- executed exactly as planned rather than modulated by afferent signals. This means that any motor error is only discovered when the gaze arrives at the fixation point, when the expected visual input (taken fom the sum of low-reliability information gained from other fixations) is compared to the actual input. This error signal, and that generated by proprioceptors, can induce oculomotor adaptation to recover normal eye function if, for example, one muscle is damaged; and is the basis of saccade adaptation (Snow et al., 1985).

Because saccades are made constantly (a rate of around 2-3 in every waking second), they are highly fatigue-resistant. This, along with their speed and characteristic start-stop profile, makes visual search a good experimental paradigm for investigating the motor system as a whole.

1.1.1   Saccadic Suppression

Despite making constant saccades in our exploration of the visual world, our perception of that world is not marred by the blurring which should accompany each eye movement.

This is because the oculomotor system suppresses the inflow of optical information during a saccade, masking any stimuli presented in this time. This makes the eye a sampled-data system. However, neither do we perceive our own blindness for the duration of the saccade. The brain fills in the visual memory gap using information from before and after the saccade. This gives rise to the illusion of chronostasis– making a saccade to the second hand of a clock often results in the perception that the hand is stationary for longer than a second. It has been suggested that the brain fills in the memory for the time spent making a saccade with the scene perceived at the end of it (Yarrow et al., 2001). Similarly, positioning yourself close to a mirror and looking from one eye to the other, it is very difficult to see any evidence of your own eyes moving. We do not notice much of a time lag between the two fixations because stimuli with reduced visibility are perceived as being shorter in duration than those which are easily visible (Terao et al., 2008); by a factor of around two for perisaccadic stimuli (Morrone et al., 2005).

This means that transient stimuli applied during the saccade go largely unnoticed (Dodge, 1900), so long as they are stimuli produced by a change in luminance rather than equiluminant hue and thus processed mainly by the magnocellular pathway (Burr et al., 1994)- this mechanism ensures that some changes in detail can still be made out, but that the vast majority of 'motion blur' is suppressed. It follows that any changes made to the scene during saccadic suppression are only recognised at the end of the saccade. Therefore, an unanticipated shift of the entire scene during a saccade would be indistinguishable from any other event causing a disparity between the predicted and actual visual input at the end of a saccade; for example, unpredicted movement of the stimulus or a motor error. In experiments on the adaptation of individual saccades, Rolfs et al. (2010) showed that displacing a visual scene during a saccade was perceived as a motor error, insofar as the oculomotor system adapted its commands to prevent the disparity between predicted and actual scene at the end of the saccade– they trained the motor system to 'believe' that it was overshooting targets by displacing the target towards the origin of the saccade, and over time saccades made to fixate the same targets from the same origin became smaller and smaller.

1.2   Bayesian Inference

All sensory systems, particularly biological, are fallible. There are countless sources of error, including limiting factors such as receptor size, regeneration time and spontaneous

firing rate; imperfections in the accessory structures such as spherical aberrations; and assumptions made by the higher processing centres. Therefore, any internal representation of the world built up by such sensory systems are noisy versions of the 'true' data: any inference made about the state of the world should thus take into account this noise. In the classical biological example of fight-or-flight, possibly catching a glimpse in the corner of your eye of something which may be large and hairy is not enough information to assume that there is a bear in the vicinity and to abandon your dinner; even though the potential bear has, to some extent, been 'seen'. Bayesian inference provides a tool for such weighting of certainty, and by implication produces an internal representation based on a probability density rather than a single unqualified estimate. Bayes' Theorem defines probability as the belief in a particular event, which depends on a prior level of belief, some new information concerning the belief, and the reliability of that information (Bayes, 1763). This is a much better model than the frequentist view of probability as the number of occurences of that event given the same starting conditions: the Bayesian model accepts the chaotic principle that in fact, any given starting condition can only produce one outcome. The ability to predict that outcome comes only from the information one has about the starting conditions: stochasticity is merely the effects of a low level of information regarding these conditions, often at the atomic level. Take a cheat coin: upon immediately seeing the coin, one expects it to be like all other coins and have an obverse and reverse. At this point, a frequentist would assume a 50% chance of a 'heads' result and continue to do so even as tails came up again and again; the Bayesian observer uses the information from each flip to inform his estimation of the odds. In increasing levels of complexity, the same could be true for predicting the roll of a loaded die, or the weather based on whether a butterfly flaps its wings. A full knowledge of the starting conditions makes prediction trivial, but this is usually impossible; as such it is vital to take into account the information you do have.

The theorem can be derived from the principle of conditional probability, and is used to work out the odds of an A's occurrence given B's observation, $P(A|B)$; using a prior assumption of how probable A is, $P(A)$, and the likelihood of observing B given A, $P(B|A)$.

$$P(A|B) \;=\; \frac{P(A \cap B)}{P(B)} \tag{1}$$

$$P(B|A) \;=\; \frac{P(A \cap B)}{P(A)} \tag{2}$$

$$P(A \cap B) \;=\; P(A|B) \cdot P(B) = P(B|A) \cdot P(A) \tag{3}$$

$$P(A|B) \;=\; \frac{P(B|A) \cdot P(A)}{P(B)} \tag{4}$$

Applied in a visual setting, A could be an illumination and B a photoreceptor firing in response to that illumination. The visual cortex uses its prior knowledge of the odds of A is, including contextual clues (likely during the day, for example; less likely indoors)- and how likely B is both with and without the stimulus (in this case, it's assumed that stimuli are relatively infrequent, and thus that $P(B|1-A) \approx P(B)$)– if B was a cone, which has a high level of eigengrau, an action potential without A could be fairly likely.

There are several ways to increase the strength of the inference of A; that is, how sure the observer is that A has occurred. Firstly, more information can be accrued over time: this effectively increases the belief in an iterative fashion. Alternatively, for some stimuli information can be gathered in parallel- for example, an object's position can be gauged by both looking at it and reaching out to touch it, bringing new information from the arm's proprioceptors and mechanoreceptors in the fingers. Finally, the function $P(B|A)/P(B)$ can be increased by using a more accurate or less noisy sensor: localising a point in space is easier using the eyes than it is using an arm because the arm has less accurate proprioceptors and more degrees of freedom.

In this context, an understanding of Bayesian inference is necessary because it is likely that such processing informs our planning of saccades in visual search (Najemnik and Geisler, 2005). Furthermore, the well-demonstrated ability of the sensorimotor system to weight sensory information and movement selection by their reliability (an example of the latter is the automatic choice to perform fine motor tasks with the more capable of the hands, given the unequal distribution of dexterity) is based in a Bayesian interpretation of the world.

1.3   Visual Search

The visual system is heterogeneous in its reliability: foveation, the dense clustering of cones in the centre of the retina, leads to a much greater spatial and chromatic acuity in some areas of the visual field than in others (Wertheim, 1891, translated by Dunsky, 1980). As such, a stimulus received in the fovea, at the centre of the visual field, is much more reliable than one received in the periphery. Information received from the periphery can be reinforced either by holding the gaze steady and integrating inputs over a longer time period, or making a saccade to a different location, which can bring the fovea to bear on an interesting spot or just provide another sample from a different area of the peripheral retina.

Najemnik and Geisler (2005) suggested that humans utilise an understanding of this heterogeneity in a Bayesian fashion to influence their search strategy: information received at each fixation point is weighted by its likelihood depending on where in the visual field it came from; rather than build up an 'image' of the scene, the internal representation is an odds distribution. Humans behave in a very similar fashion to computer programs designed to minimise the number of fixations required to localise a target: Najemnik and Geisler's ideal searcher attempted to maximise the gain of information with each new fixation, not taking into account the distance of that fixation from the current one. This suggests that each saccade is targetted to a specific location rather than being effectively randomised to cover a reasonable area of the scene.

2   Hypothesis

Given that humans optimise their search strategy to minimise costs (Najemnik and Geisler, 2005), that gross motor systems can adapt to changing environmental and internal conditions in order to maintain efficacy (Snow et al., 1985; Rolfs et al., 2010), and that intrasaccadic displacement of a visual scene is sufficient to stimulate such adaptation (Rolfs et al., 2010), it was hypothesised that the strategy could be perturbed by simulating motor unreliability with movements in certain directions using an intrasaccade display shift; if the perturbation was significant, the new optimal strategy could show a change in the distribution of saccade frequencies in the penalised versus the non-penalised directions.

3 Materials and Methods

3.1 Method Outline

Subjects were seated at an eye tracker fixed to both a chin and forehead rest, and a desk, 35cm away from a computer monitor. The eye tracker was calibrated to best detect their pupil, taking into account their eye shape and the color of their iris and sclera. The eye tracker was also set up to accurately detect the location of their gaze on the screen using a 16-point calibration. The subject was presented with a central fixation stimulus at the beginning of each trial. After a short fixation on this target, a field of 70-100 pseudorandomly distributed distractors (each in the form of an 'L', and henceforth referred to as such) and one target (in the form of and henceforth referred to as a 'T'); the average distance between them was 2.2°. The stimuli were distributed as a hexagonal lattice with a random rotation applied to each hexagon, ensuring reasonably uniform coverage of the scene. The trial lasted 4 seconds, during which time the subject had to find and fixate the T. Successful trials were terminated once fixation on the T had been detected.

Figure 1: A sample scene from the visual search task. This scene measures $\approx$ 17cm ×17cm, and is presented on a screen measuring 40cm × 30cm. Note the target T on the right hand side of the scene, $^1/_3$ of the way from the top.

Each trial was followed immediately by the central fixation stimulus to start the next trial. In total, each subject completed 900 trials in < 90 minutes. Every 150 trials (< 15 minutes), the subject had an opportunity to briefly rest and recalibrate the eye tracker if necessary. For the purposes of analysis, the trials were split into blocks of 50 (< 5 minutes).

In the first 100 trials, the subject was allowed to perform the task normally. From the 101st trial (the 3rd block) and for the remainder of the experiment, saccades made

with a positive rightward component caused a display shift of random length, $\mu = 0.5$cm (normally distributed with $\sigma = 0.005$cm), in a random direction. In post-hoc analysis, saccades were counted, and their direction and magnitude calculated: this count was split into those with positive leftward and rightward components. Various other statistics were also collected, including fixation duration, saccade frequency and trial success rate.

3.1.1 Subjects

Analysable data were obtained from four naïve subjects. Each had good vision or vision corrected with contact lenses. Subjects diagnosed as dyslexic were prohibited from participating as the task required rapid processing of graphical symbols. For the same reason, subjects whose native language does not use the Roman alphabet were rejected. Finally, all subjects' native language read left-to-right in order to eliminate possible variation in saccade direction bias due to the task's reasonable similarity to reading (although no such bias was detected in control trials using a t-test, $p = 0.421$). A further one subject's experiment had to be terminated due to having particularly long and dark eyelashes, which interfered with the gaze detection in certain regions of the screen.

3.1.2 Equipment

3.1.2.1 Eye Tracker The eye tracker used was SR Research's EyeLink 1000; a 1kHz infra-red camera which uses both pupil and coronal measurements to infer gaze position. This makes it relatively robust to small accidental head movements, allowing the use of a padded head rest rather than bite bar to maintain head position. This is far more comfortable for the subject and thus allows longer trial periods.

3.1.2.2 Computers The monitor used as a 100Hz CRT, allowing fast and easily controllable refreshes of the visual scene. The experiment program used was run on a separate computer to the slave machine running the EyeLink, connected by ethernet in order to pass commands and data between them at high rates. The display shown to the subject was duplicated for the benefit of the experimenter on a separate screen.

3.2 Event Detection

A major difficulty in eye tracking experiments is that the movements are so fast and so small that the position data is reasonably noisy. As eye velocity and acceleration is calculated, this noise is compounded: this can make it difficult to cleanly separate saccades from noise like small head movements, detector noise, blinks and so on. Figure 2 demonstrates the difficulty in trying to get an accurate saccade count using only an eye speed threshold.
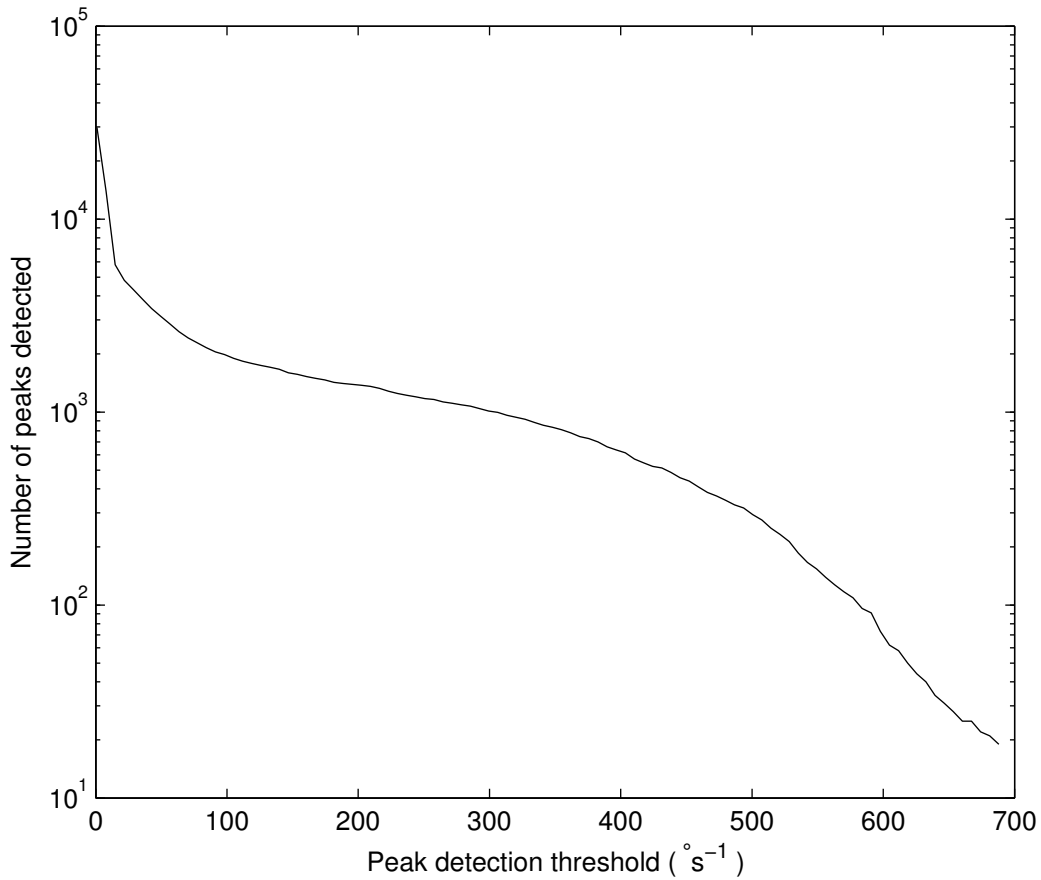

Figure 2: Many saccade detection algorithms simply use an eye speed threshold. However, if this were to provide a good solution, one would expect a clear and sharp elbow in a graph of number of saccades detected against speed threshold used when counting saccades, for the same data set. Even with a 5-point moving-average filter this is clearly not the case.

There are a number of solutions to this, each applicable in different situations. Furthermore, the problem is compounded when it is necessary to detect the saccade while it is being made (as was the case in this paradigm, where the saccade had to be identified and its orientation estimated relatively long before fixation was achieved).

14

3.2.1 Online Detection

For this experiment, saccades had to be detected early enough that calculations could be performed on the data preceding detection to estimate their trajectory, and the display could be offset (including refreshing the monitor, a time-consuming process on the millisecond scale). Therefore, too restrictive a detection algorithm would detect the saccade very late and so the display offset would likely occur after the saccade had ended, and not be masked. However, too permissive an algorithm would produce a high number of false positives (due to the noisy nature of the readings), causing a display offset when there was no saccadic suppression at all. As seen in Figure 2, there is no perfect threshold to use for detecting saccades 100% accurately. Furthermore, the noise level of instantaneous speed and acceleration recordings (i.e. displacement between readings divided by time between readings is very high. Options for filtering the data are fairly sparse given that the choice must be computationally light and very fast: in this instance a moving average filter was used.

Given that this moving average necessarily brings in lag of half the size of the sliding window; it was necessary to balance a window size large enough to prevent significant false-positive display offsets against one small enough to prevent the delay causing post-saccade offsets. Experimentation showed that a window size of 5 items ($\approx$ 5ms) was sufficient for this purpose- various window sizes were tested with the total number of display offsets per trial constrained; the 5-item window had the lowest rate of perception reported by the subject at $\approx 3.5\%$.

Attempts to optimise the thresholds used proved time-consuming and ineffective, due to the requirements and multifactorial nature of the detection: it was necessary to maximise the number of saccade-like movements detected, while ensuring that a high percentage were longer than the maximum time necessary to perform the calculations and refresh the monitor ($\approx$ 12ms). Results were highly dependent on a number of assumptions and found to be inferior to the EyeLink's default parameters in terms of the subject perceiving display offsets: as such, the online saccade detection algorithm was triggered when the eye had exceeded a speed of $40°s^{-1}$, an acceleration of $6000°s^{-2}$ and $0.6°$ displacement from the last stable fixation point.

While this is an acceptable online method for detecting movements which look reasonably like saccades, the relatively high thresholds mean that smaller saccades are only

detected very late in their trajectory. Furthermore, the saccade event as detected by this process is not useful for analysing the form of the saccades themselves (for example, the duration of the saccades, and thus the inter-saccadic fixations), because by definition the saccade must be in well on its way before the EyeLink is able to detect it. Furthermore, blinks were only detected when the EyeLink completely lost track of the pupil; the movement of the eyelid in the milliseconds preceding this produced unphysiologically high speed and acceleration values, which sometimes caused the detection of a 'saccade'.

3.2.2 Post-hoc Detection

With the time afforded by after-the-fact processing, it is possible to use a much better detection algorithm. The noise level could be further reduced by using a much larger window size for the moving average, but this would also extinguish any high-frequency characteristics of the saccade profile, particularly at the end of the movement. It would also significantly flatten the saccade itself- as the window size approached and exceeded the duration of small saccades they would be virtually undetectable.

Post-hoc analysis for this experiment was performed using the smoothing and event detection algorithm described by Nyström and Holmqvist (2010). Firstly, position data was smoothed using a simplified least-squares filter (Savitzky and Golay, 1964) of window size 25 (two times the minimum saccade duration, plus one) and order 2. For each position bin X_i and Y_i, the bins $i - 12$ to $i + 12$ were fitted to a quadratic curve; bin i was then replaced with the filter's estimation of where it 'should' be. This step actually has very little effect on the data as taken from the eye tracker. However, it allows a very fast, easy and low-noise production of first and second order derivates (simply differentiating the fitted quadratic function).
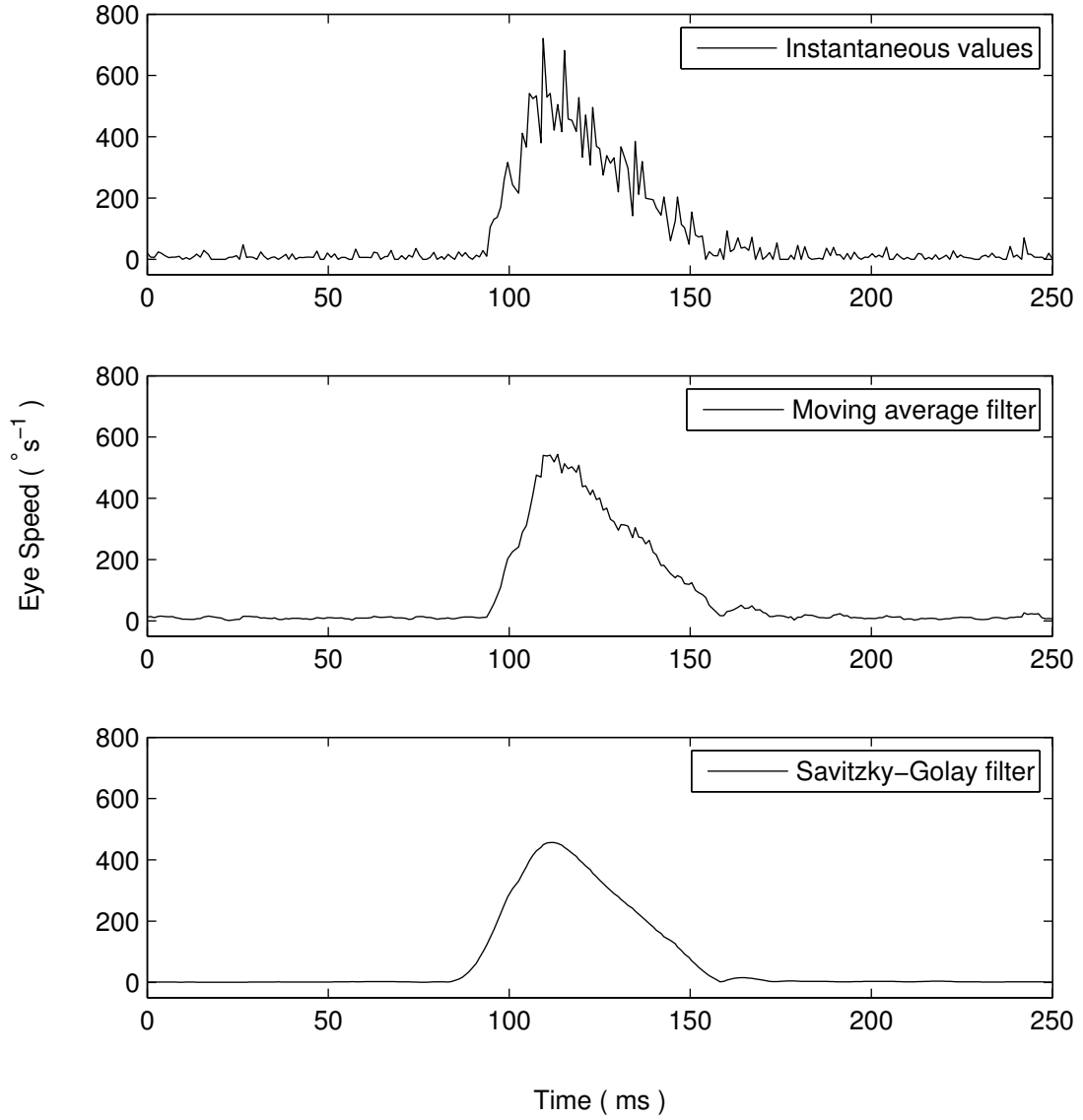
Figure 3: These plots show the speed data generated by the same saccade under different filters. The moving average filter used here is effectively the speed data 'seen' by the EyeLink. Note the practically noiseless fixations, characteristic shape and lack of spurious local maxima and minima of the Savitzky-Golay filtered plot. While there is some gross flattening of the plot, the high-frequency elements are still conserved to a reasonable degree.

Speed thresholds for saccade detection were found in an entirely data-driven manner: Nyström and Holmqvist point out that the noise level can vary not only between individuals, but even between trials, and that as such the optimal detection threshold should also change. Due to the nature of this experiment, thresholds were generated per calibration block of 150 trials. Saccade peaks were identified as any period where the

17

eye's speed was greater than six times the calibration block's noise level ($\mu = 22.4^{\circ}s^{-1}$, $\sigma = 8.80^{\circ}s^{-1}$)- a margin accurate enough to detect microsaccades (Engbert and Kliegl, 2003). Saccade onset was found by tracking backwards through the data from the peak to find the first local minimum where the speed was below three times the noise level ($\mu = 6.83^{\circ}s^{-1}$, $\sigma = 0.916^{\circ}s^{-1}$). Saccade offset was found using a threshold constructed from a weighted combination of three times the noise level, and a measure of the noise level immediately preceding the individual saccade in question (see Nyström and Holmqvist, 2010): $\mu = 6.47^{\circ}s^{-1}$, $\sigma = 2.73^{\circ}s^{-1}$. Tracking forward through the data from the peak, the saccade was judged as finished at the first local minimum where the eye's speed was below this threshold.

Finally, the system avoids detecting post-saccade elastic oscillations as saccades by incorporating a term for the minimum accepted fixation length- saccades detected as initiating less than 40ms after the offset of a previous saccade are struck from the record. Even corrective saccades, the shortest-latency saccades that one would expect to see, have a latency of ≈130ms (Henson, 1978); as such, this assumption should not introduce any non-detection errors. Regions of the position data judged by the EyeLink to be blinks were removed (practically, replaced with the Not-a-Number or 'NaN' arithmetic representation), along with 30ms either side of them to prevent errant detection of saccades.

This confers a number of advantages compared to using the EyeLink's estimation of saccade duration, demonstrated by Figure 4. Firstly, the Nyström-Holmqvist algorithm gets a much more accurate estimation of the time and location of the saccade's onset and offset; vital for performing calculations based on saccade and fixation duration and displacement. The avoidance of 'rebound detections' also allows for a much more accurate estimation of total number of saccades made (between 1.5 and 5% of saccades produce rebounds detected by the EyeLink as extra saccades).
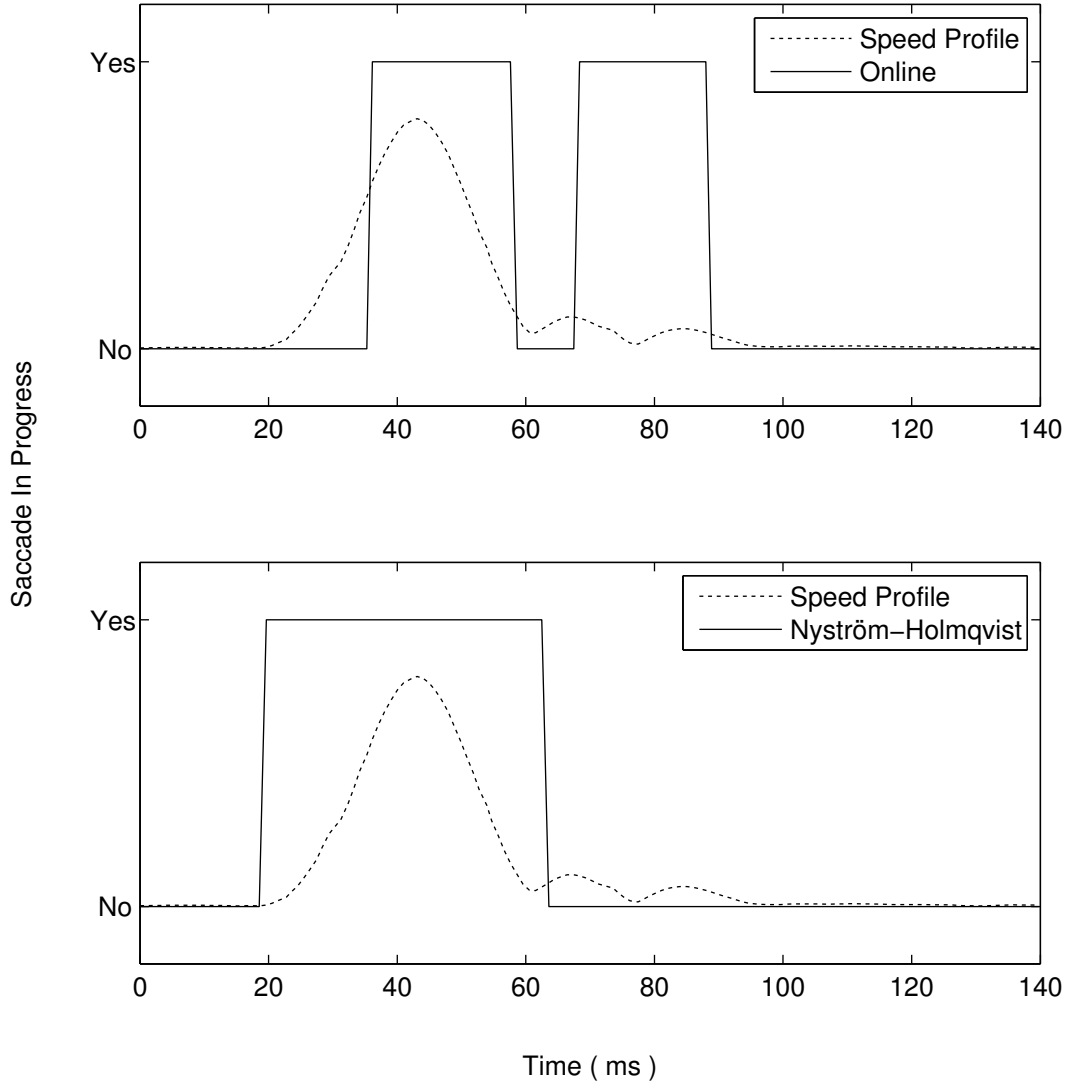
Figure 4: A comparison of the saccade recognition abilities of the online versus the Nyström-Holmqvist event detection algorithm. Note the lag in the online detection and the false positive detection due to the elastic oscillations.

3.3 Direction Prediction

Due to their ballistic nature, saccades are basically made in straight lines. As such, it is reasonably trivial to predict the direction of the next fixation point almost as soon as a saccade is underway. However, given that the position data is fairly noisy, investigations were made into whether averaging position points could arrive at a more robust prediction. This analysis was performed on an early data set.

Data examined for each saccade were the last stable gaze location (i.e. the position before the speed and acceleration thresholds were exceeded), the location where the saccade is confirmed (i.e. when the displacement threshold is also exceeded), and the location where the gaze once again becomes stable. The correlation between initial bearing (i.e. the direction of movement between last stable location and saccade confirmation location) and total bearing (the direction of movement between last stable location and post-saccade stable location) showed, unsurprisingly, a gradient within 1% of 1 (95% confidence interval 1.003 to 1.011). The R^2 was 0.969.

The same analysis was performed using average locations in a window of 5ms before gaze stability was lost, the first 5ms after the detection of a saccade, and the 5ms after gaze stability was regained. The gradient remained within 1% of 1 (95% confidence interval 0.9995 to 1.006). The R^2 was 0.975. This small gain in robustness introduced a significant delay in the prediction of direction, and thus in the display offset, so the simpler method was used in the final experiment.

To account for the margin for error in the direction prediction, the penalised direction range was reduced slightly in order to prevent the display offsets 'spilling over' onto saccades intended to be made in a leftward direction. It was found that leaving a tolerance of 5.7°was sufficient to ensure that 95% of saccades detected as being made towards the penalised hemifield did, in fact, end further right than the point of origin.

4   Observations

4.0.1   Identifying an Error Signal

Firstly, it was vital to determine whether the display offsets actually had any impact on the search. The search strategy is supposed to be optimised to minimise costs, but if no cost can be identified in the introduction of the display offsets, then the strategy cannot be expected to change. The crudest measure of the task's difficulty is simply the mean failure rate of the control trials versus the earliest test trials- i.e. for how many trials in the blocks C1 and C2, and T3 and T4, the subject failed to fixate upon the T within

the given time period. There was no significant difference in the failure rate between the control blocks and the earliest test blocks. This is somewhat expected, as the subject can also be expected to improve at the task as time goes on, and for relatively small trial blocks and this small a sample, the data is not particularly internally consistent.

Neither did the average frequency of saccades across all subjects and all trials in a block seem to change significantly with the introduction of the display offsets. However, there was a significant change in the length of fixations directly after display offsets (see Figure 5), which is apparent from the very first trial block that the offsets were introduced. This is an important finding as it shows that the display offset does have a detectable impact on the search: therefore, the brain is aware something is amiss after these offsets. The fact that the impact is negative (a delay in making the next saccade) suggests that, should the search pattern be plastic and able to adopt a new optimal solution, this paradigm may be able to stimulate such an adaptation.
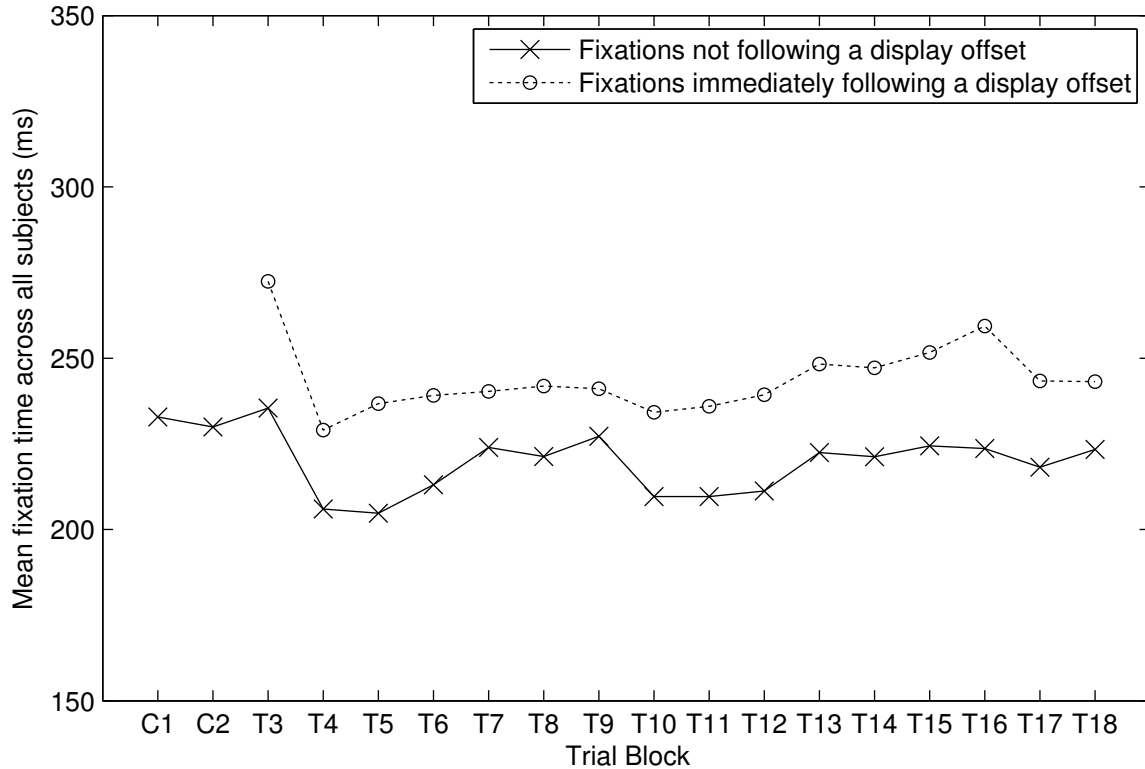

Figure 5: Display offsets cause a significant increase in duration of fixations immediately following the offset ($p < 0.01$ using a rank sum test). Across all trials and all subjects the mean difference in fixation time between post-offset and not post-offset was around 25ms. Note that the control trials C1 and C2 had no offsets, and thus no post-offset fixations.

There are several possible mechanisms which may cause this delay, to be discussed later.

This increase in fixation length did not appear to change as the experiment progressed: as such, it seems to be an immediate response to the offset rather than a learned strategy to cope with it.

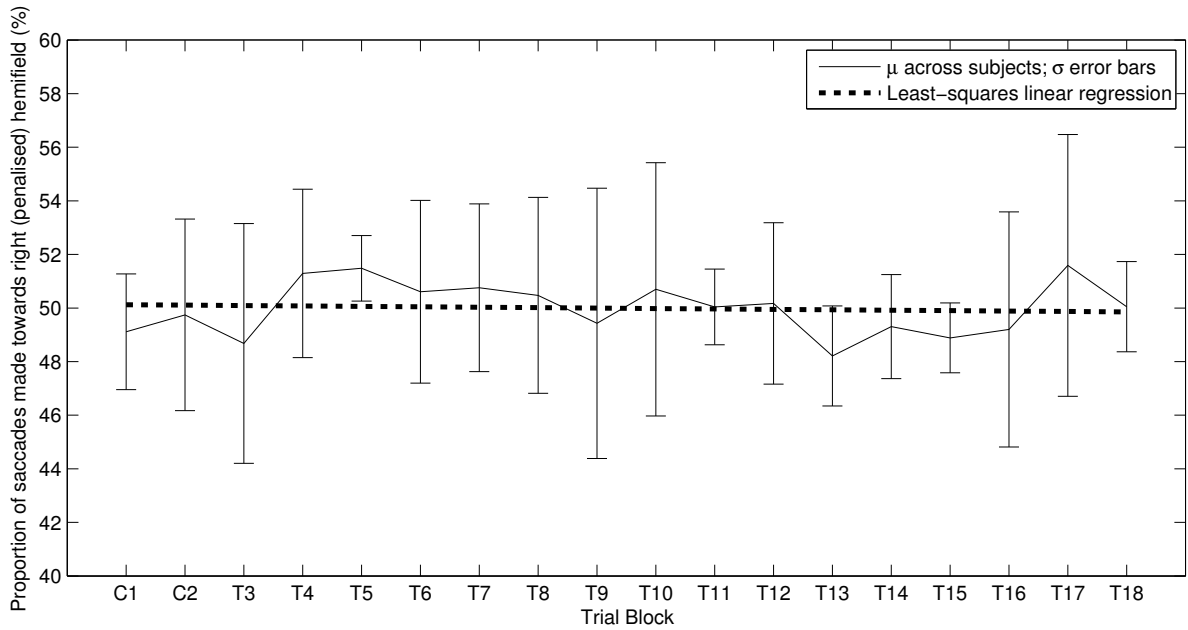### 4.0.2 Identifying Changes in the Search Strategy

Figure 6: No change in the distribution of saccade directions throughout the training period was found. A rank sum test comparing the proportion of saccades made into the right hemifield during the control trials C1 and C2 compared with the last test trials T17 and T18 returns $p = 0.458$; this does not allow the null hypothesis, that they are from distributions of equal median, to be rejected.

Clearly Figure 6 shows that, on this time scale, the search pattern does not adapt to avoid saccades prone to noise. Neither does there seem to be any significant change in saccade size or frequency. There is no evidence that microsaccades are made to correct for the display offset, or that fixations immdiately preceding saccades made towards the penalised hemifield are any longer than those preceding other saccades (which might be expected, in order that the visual system can collect more information about its next

target region through the current peripheral visual field and so have a better idea of the 'jigsaw edges' that will be required to slot the new information into the internal representation).
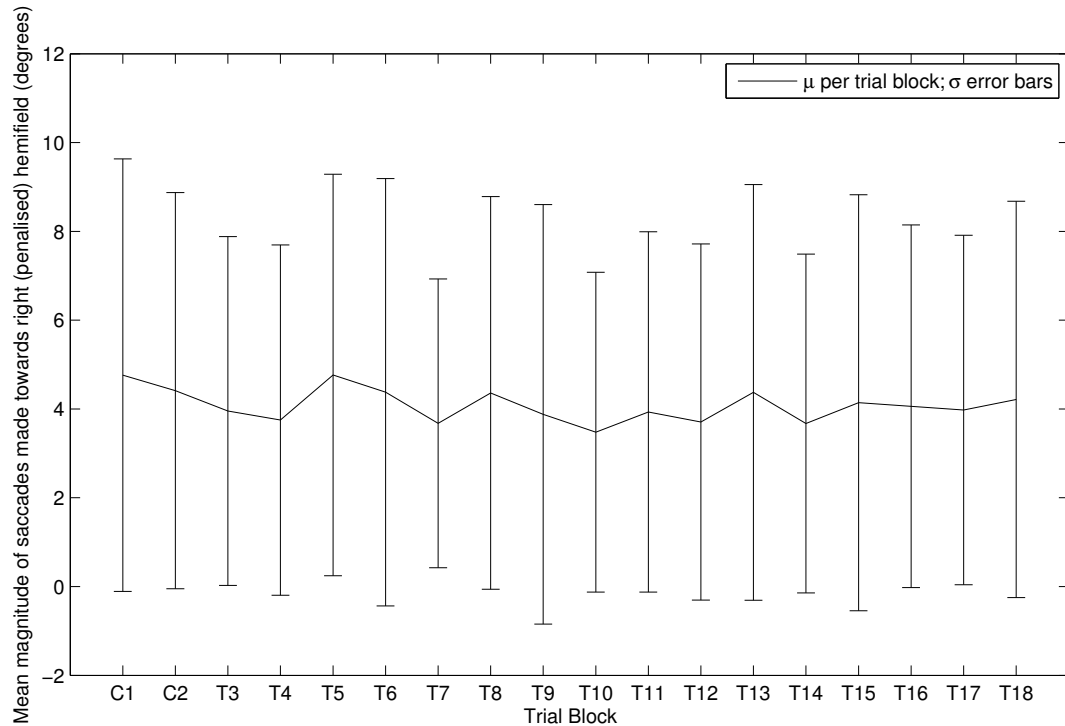

Figure 7: There is no significant change in the magnitude of saccades made into the penalised hemifield during the task.

Figure 7 shows that the magnitude of saccades does not change throughout the course of the task, another possible adaptation to the added noise.

5 Discussion

5.1 Results

5.1.1 Why does the display shift cause increased fixation duration?

A number of processes determine the length of a fixation.

1. Saccadic suppression must be relieved. This is effectively instantaneous: saccadic suppression is switched off even when the relative location of the gaze on the scene is only stationary for an instant- for example, if the entire scene is moving fast enough to induce visual blurring, a saccade which causes the gaze to move at the same speed as the scene even for a millisecond is enough for masking to end (Deubel et al., 1987).

2. The visual system must process the first post-suppression visual snapshot; this takes around 25ms, a fairly inelastic delay.

3. If the searcher is building up an internal representation of the scene, the information from the new fixation must be fitted into the existing representation. In order to save on visual processing and short-term memory, information seems to be integrated only over 4-5 fixations (Najemnik and Geisler, 2005).

4. A decision is made as to the usefulness of integrating information over a lengthy fixation against making another saccade to a new fixation point in order to collect more information there.

5. Fixation is maintained for as long as is deemed useful by the searcher, then (according to Najemnik and Geisler) the distribution of information across the internal representation of the scene is assessed, the fixation point which would yield the greatest amount of information is selected, and the motor commands to make the saccade here are passed.

The step most likely to be affected by the display offset seems to be the time taken to integrate the new visual field into the existing internal representation of the scene. Yarrow et al. (2001) showed that there was a qualitative difference in the handling of new visual scene data if it differed from the expectation (in their experiments, by the absence of chronostasis). It is not unreasonable to suggest that if chronostasis does not occur, the time taken to slot the new information into the current representation is increased due to having to match peripheral data in a jigsaw fashion. However, it is also possible that the fixation is lengthened by design, in order to accrue better peripheral data for such edge matching and to ensure that it was, indeed, an apparent unpredicted displacement of the whole scene, rather than movement of certain elements within the scene.

5.1.2   Why does this seem not to stimulate adaptation of the search strategy?

There are a number of reasons that the expected changes to search pattern are not seen. Either the training period used in this experiment is not long enough for the adaptation to occur, or it is not possible for this type of modulation to occur in response to this stimulus.

In simple motor adaptation paradigms like the one described in Snow et al. (1985), a single motor unit can easily be identified as the source of a predictable and sizeable error, and that unit can simply have its gain increased or decreased by a higher circuit to account for the change. Even in Rolfs et al.'s global adaptation paradigm, the adaptation applied to single movements and simply corrected an error, rather than having to recognise that the error could not be corrected for and to adjust motor planning, rather than command execution, circuits. A reasonable analogy here is that of a recent amputee. As searching saccades are almost entirely subconscious, even once the brain has developed a sense that particular movements do not achieve the desired result it may continue producing commands for those movements: there are reports of amputees reaching for a ringing phone with an arm which no longer exists, many weeks or even months after surgery.

There is also the question of necessity. Analysis of the success rate of the task, inversely proportional to the time required to localise the target in the scene and thus proportional with the searching ability of the searcher, suggested that the introduction of direction-contingent noise did not, in fact, decrease the subject's ability to localise the target to any significant degree. If visual search is entirely uninhibited by the noise added by this paradigm, it stands to reason that adaptation of the search strategy is not necessary and thus does not happen. This suggests that, while optimised to a certain extent, the exact position of each fixation is not crucial to the optimality of the search model. After all, given that the Bayesian data are collected from the whole visual field, a displacement of the fovea by less than 1°will not significantly impact the information gain from each fixation.

This, in turn, suggests that a saccade producing an error signal does not on its own stimulate adaptation of a strategy. The error signal may well provide a substrate for such adaptation if it is necessary, but given the minimal (or even nonexistent) decrease

in search efficacy the adaptation may not need to happen at all. Compare this to Rolfs et al.'s paradigm, where adaptation was entirely necessary to fixate the target quickly.

Finally, a lack of contingency of the noise with the saccade may have been a confounding factor. A nontrivial proportion of saccades were followed by oscillations of amplitude large enough to cause the detection of another saccade. Such rebound detections would have one of two effects: either the displacement would be entirely out of the period of saccadic masking and so harm the illusion of the displacements only occurring as motor errors, or the displacement would occur following a saccade made into the non-penalised hemifield (because the rebound would be in the opposite direction). Late detections in small saccades also caused a number of display offsets to occur outside of the saccade (between 3 and 7% of all shifts), further decaying the saccade-offset contingency by increasing the perception of display offsets. Lastly, the display offset magnitude was relatively constant. Given that the size of offset magnitude which can be masked by a saccade depends on the magnitude of the saccade (Bridgeman et al., 1975), for small saccades the offset magnitude may have been too large and, again, increased the perception of the offsets. While perception of the offsets is not in itself a particular flaw in the experiment (significant motor errors would, after all, be presented to the conscious mind just as these over-large offsets are), decreased perception would help maintain the illusion of motor error rather than being able to consciously attribute error signals to the scene's movement (and possibly ignore them). Using a saccade magnitude-dependent display offset magnitude would also present a more physiological error stimulus. This does introduce the further step in saccade detection of predicting its magnitude, which can be achieved using the saccade's peak speed (Carpenter, 1988), but can be made more accurate by using peak speed $\times$ acceleration time (Lin et al., 2004), which adds even more of a delay before the display offset can be produced.

5.2 Further work

5.2.1 Possible improvements on the paradigm

As was stressed in 3.2.1 and in Nyström and Holmqvist (2010), the event detection algorithm is all-important when it comes to experiments based on detecting saccades,

and different filters can yield quite different results. Although attempts were made to improve the detection parameters, preliminary experiments suggested that they were not successful. A pilot study could be devised where a subject had to make a series of small saccades to targets presented sequentially around the screen. So long as the targets were not too far apart, the number of saccades should be equal to the number of targets presented. A range of moving average window sizes, and speed, acceleration and displacement thresholds could be iteratively applied to the raw position data in order to produce a set of parameters which detect the correct number of saccades as early as possible in each saccade. In order to save iterations, rather than vary the speed and acceleration independently, the relationship between them in the initial phase of the saccade could be found and they could be varied in proportion– this is possible because saccades have such a characteristic shape. A very simple improvement to the detector would be to introduce a minimum fixation length, as used in the Nyström-Holmqvist detector: this would prevent the rebound detections which seem to cause many post-saccade offsets.

One of the major delays in display offset is, in fact, the refresh rate of the monitor. Even with a reasonably fast monitor, the refresh rate is slow enough that many saccades are over by the time that the new scene is presented. With new technology making monitors faster, cheaper and more available, a pilot experiment could be designed to ascertain whether there is a significant decrease in the perception of offsets using a higher-frequency monitor (while higher-frequency monoscopic monitors are not required in the consumer market, the advent of 3DTV has effectively doubled all monitor refresh rate requirements and some high-end screens already use up to 240Hz). Furthermore, it would be possible to incorporate into a detector an algorithm to make a decision as to whether a display offset is feasible in the time remaining. It is possible to keep track of how long it was since the last monitor refresh onset and offset, and to predict the duration of a saccade using the same process as that used to predict its magnitude– the amplitude over peak speed has a reasonable correlation with saccade duration (Lin et al., 2004). These data, although they may introduce a very small delay for the calculations, could prevent many of the post-saccade offsets.

Incorporating a reliable saccade magnitude predictor would also allow an increase in the gain of the display offset for larger saccades. Offsets up to $^1/_3$ the size of a saccade are masked almost entirely (Bridgeman et al., 1975), so using this as an absolute upper bound (once the tolerance of the prediction has been found) would allow a greater level

of noise to be used without risking a high rate of perception of the offsets. It is possible that the noise level in the current paradigm was just too low to force the search strategy to adapt to it.

After analysis of preliminary results, we tried to simulate the effect of larger display offsets by increasing the density of stimuli in the scene. However, an idiosyncrasy in the program caused there to be a cap on the number of stimuli which could appear on the screen without causing problems in their placement. As such, the size of the scene had to be decreased in order to increase the density. It became apparent that to have the average distance between stimuli comparable to the noise level (at which point we would expect a serious deficit in the searcher's ability, particularly if the stimuli were arranged in a more regular pattern), the scene would have to be so small that the program would falsely recognise that the subject had fixated upon the target (noise in the calibration process required there to be a tolerance of $\approx 1.6°$ around the location of the target; for small scenes, this may be a significant portion of the scene). Furthermore, as the scene gets smaller, a greater proportion of the scene is within the fovea and so saccade-driven search is less necessary. If the stimulus cap could be increased, so could the density of the distractors (possibly with a concommittant increase in stimulus brightness in order to offset the increase in difficulty of the task) to bring the distance between them closer to the noise level.

A final alteration would be to more thoroughly investigate the robustness of the Nyström-Holmqvist detector. In some early experiments, where the peak saccade threshold was generated per block of 50 trials rather than 150, the iterative algorithm failed to converge on a low threshold, instead increasing exponentially until it was larger than most of the saccades. Increasing the threshold-generation block size solved this problem and thus it was not investigated further.

5.2.2 Expansion

5.2.2.1 Characterising the change in fixation time

In order to determine whether the change in fixation time was truly immediate, or a very fast learned adaptation, a new paradigm could be introduced whereby display

offsets occurred randomly for 50% of saccades, regardless of their direction. The task could be set up to be very difficult, if not impossible (by using a low stimulus contrast, or instructing the subject to find a target which is not there), in order to ensure a long, unbroken series of saccades (although if it went on for too long, the subject would likely start to increase the length of their fixations as part of their search strategy: the duration would need to be limited to prevent this). Fixation time following each saccade could be calculated in post-processing. From the start of the trial, the data for post-offset and not post-offset could be collected and analysed to investigate whether the fixation lengths differed from very early in the task or developed more gradually. Again, a number of control trials (with no noise) could be run before the test trial in order to establish an average fixation length to compare the post-offset fixations to.

The offset gain could also be varied, in order to investigate whether the increase in fixation time scaled with the size of the offset. Finally, if the display offset was to be linked to a direction such that it could be predicted by the searcher (rather than occurring randomly), and ample time was given to learn of this contingency, the extinction of the learning could be tested by returning to the control condition. If indeed it was a learned and designed alteration to the search pattern in order to increase the time for integration of visual data at this unexpected fixation point, extinction would be gradual rather than instantaneous.

5.2.2.2 Testing the Nyström-Holmqvist detector

Thresholds are determined by removing all speed data above an arbitrary threshold T_i, then finding the mean and standard deviation of the data which remain.
$T_{i+1} = \mu_i + 6\sigma_i$. This process is repeated until T converges on a reasonably low value. Given that, after smoothing, the speed data at fixation have very little noise, the major contributor to the σ of the data is the up- and down-strokes of the saccade's speed profile. As such, in trials with a particularly high frequency of saccades, which will have more of these 'noise'-introducing up- and down-strokes, this measure of 6 times the noise level may, in fact, produce a new threshold which is higher than the previous one (this is possible regardless of the initial threshold). This would cause the threshold to diverge instead of converge, and so the majority of saccades in that trial would be missed entirely. Increasing the block size would marginalise these high-frequency trials. However, the fact that this is possible in the physiological range of saccade frequencies is a cause for further investigation.

Using a very similar paradigm to that used in this experiment, the Nyström-Holmqvist saccade peak speed thresholds could be found on a per-trial basis. Saccades could be counted using some other, constant, detection threshold. A simple test for significant difference in mean would show whether convergent trials are of a lower saccade frequency than divergent trials.

6 Acknowledgements

References

Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances. Communicated by Mr. Price, in a letter to John Canton, M.A. and F.R.S. Published posthumously.

Bridgeman, B., Hendry, D. and Stark, L. (1975). Failure to detect displacement of the visual world during saccadic eye movements. Vision Research *15*, 719 – 722.

Burr, D. C., Morrone, M. C. and Ross, J. (1994). Selective suppression of the magnocellular visual pathway during saccadic eye movements. Nature *371*, 511–513.

Carpenter, R. H. S. (1988). Movements of the Eyes chapter 4, p. 72. Pion Ltd.

Deubel, H., Elsner, T. and Hauske, G. (1987). Saccadic Eye Movements and the Detection of Fast-Moving Gratings. Biological Cybernetics , *57*, 37–45.

Dodge, R. (1900). Visual perception during eye movement. Psychological Review , *7*, 454–465.

Dunsky, L. I. (1980). Peripheral Visual Acuity. American Journal of Optometry and Physiological Optics , *57*, 915–924.

Engbert, R. and Kliegl, R. (2003). Microsaccades uncover the orientation of covert attention. Vision Research , *43*, 1035–1045.

Henson, D. (1978). Corrective saccades: Effects of altering visual feedback. Vision Research , *18*, 63–67.

Lin, H.-H., Chen, Y.-F., Chen, T., Tsai, T.-T. and Huang, K.-H. (2004). Temporal analysis of the acceleration and deceleration phases for visual saccades. Biomedical Engineering- Applications, Basis & Communications , *16*, 355–362.

Morrone, M. C., Ross, J. and Burr, D. (2005). Saccadic eye movements cause compression of time as well as space. Nature Neuroscience , *8*, 950–954.

Najemnik, J. and Geisler, W. S. (2005). Optimal eye movement strategies in visual search. Nature , *434*, 387–391.

Nyström, M. and Holmqvist, K. (2010). An adaptive algorithm for fixation, saccade, and glissade detection in eyetracking data. Behaviour Research Methods , *42*, 188–204.

Rolfs, M., Knapen, T. and Cavanagh, P. (2010). Global Saccadic Adaptation. Vision Research , *50*, 1882–1890.

Savitzky, A. and Golay, M. J. E. (1964). Smoothing and Differentiation of Data by Simplified Least Squares Procedures. Analytical Chemistry , *36*, 1627–1639.

Snow, R., Hore, J. and Vilis, T. (1985). Adaptation of saccadic and vestibulo-ocular systems after extraocular muscle tenectomy. Investigative Ophthalmology and Visual Science , *26*, 924–931.

Terao, M., Watanabe, J., Yagi, A. and Nishida, S. (2008). Reduction of stimulus visibility compresses apparent time intervals. Nature Neuroscience , *11*, 541–542.

Wertheim, T. (1891). Über die indirekte Sehscharfe. Zeitschrift für Psychologie und Physiologie der Sinnesorgane , *7*, 172–187.

Yarrow, K., Haggard, P., Heal, R., Brown, P. and Rothwell, J. C. (2001). Illusory perceptions of space and time preserve cross-saccadic perceptual continuity. Nature , *414*, 302–305.