Spatial integration in human smooth pursuit

Stephen J. Heinen *, Scott N.J. Watamaniuk

The Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, USA

Received 24 December 1996; received in revised form 23 September 1997

Abstract

When viewing a moving object, details may appear blurred if the object’s motion is not compensated for by the eyes. Smooth pursuit is a voluntary eye movement that is used to stabilize a moving object. Most studies of smooth pursuit have used small, foveal targets as stimuli (e.g. Lisberger SG and Westbrook LE. J Neurosci 1985:5:1662-1673.). However, in the laboratory, smooth pursuit is poorer when a small object is tracked across a background, presumably due to a conflict between the primitive optokinetic reflex and smooth pursuit. Functionally, this could occur if the motion signal arising from the target and its surroundings were averaged, resulting in a smaller net motion signal. We asked if the smooth pursuit system could spatially summate coherent motion, i.e. if its response would improve when motion in the peripheral retina was in the same direction as motion in the fovea. Observers tracked random-dot cinematograms (RDC) which were devoid of consistent position cues to isolate the motion response. Either the height or the density of the display was systematically varied. Eye speed at the end of the open-loop period was greater for cinematograms than for a single spot. In addition, eye acceleration increased and latency decreased as the size of the aperture increased. Changes in the density produced similar but smaller effects on both acceleration and latency. The improved pursuit for larger motion stimuli suggests that neuronal mechanisms subserving smooth pursuit spatially average motion information to obtain a stronger motion signal. © 1998 Elsevier Science Ltd. All rights reserved.

Keywords: Human smooth-pursuit; Eye movements; Spatial integration; Summation

1. Introduction

The smooth pursuit eye movement system evolved to stabilize moving images on the retina and thereby prevent excessive image slip which can degrade acuity [1]. Traditionally, work on the smooth pursuit system has focused on how a small (usually < 1°) moving spot of light drives pursuit (e.g. [2-4]). However, smooth eye movements are impaired when the target spot moves across a textured background [5-7]. The result is not simply due to decreased salience of the moving spot, since the impairment occurs even if the texture immediately surrounding the path of the spot is removed [8]. This is surprising, given that the pursuit system normally operates in a richly textured environment. However, the retinal-images of many objects that we pursue are much larger than the traditional small spot target. For example, the retinal image of a 6 ft tall person 50 ft away subtends 7° of visual angle. If the smooth pursuit system has the capacity to follow larger objects, one might expect that it could sum motion information over the retinal-image of the object, despite the fact that in many cases the image extends well beyond the fovea.

Spatial summation occurs in the human motion processing system. When the size of a constant density random-dot cinematogram (RDC) is increased, both motion discrimination [9,10] and detection (e.g. [11-14]) improve. The improvement is thought to occur because the responses of a greater number of motion detectors are summed or integrated, and hence the larger stimulus produces a stronger motion signal. If the pursuit system uses a motion signal that has been processed similarly, it would seem reasonable that motion information from the parafovea and peripheral retina would facilitate pursuit of a large stimulus, and thereby lead to enhanced image stability during normal tracking. 
Areas of the brain in the temporal lobe are thought to be involved in both motion processing and smooth pursuit, and neurons there have properties that could facilitate pursuit of large stimuli (for reviews see refs. [3,4]). For example, neurons in the middle temporal area (MT) have receptive fields that extend well beyond the fovea [15,16]. Area MT projects to the medial superior temporal lobe (MST), which has receptive fields that are even larger, and where more cells respond during pursuit eye movements [17]. The processing in MT/MST is thought to support the conversion of sensory information into a smooth pursuit motor command, creating a sensorimotor interface [17]. We report here that motion stimuli which produce better psychophysical discrimination also produce higher-gain smooth pursuit. While other researchers have reported that larger targets can improve pursuit [5,18–20], this is the first systematic analysis of the effect of spatial summation on the pursuit response. More importantly, our results imply that the enhanced pursuit is caused by a stronger velocity signal, and is not related to the increase in position information that a larger object would afford, since the stimuli we used were free of consistent local position cues.

2. Methods

2.1. Subjects

Eye movements were recorded from the two authors (SH and SW) both of whom were experienced in smooth pursuit tasks. Both subjects had corrected to normal vision.

2.2. Stimuli

Observers tracked a 0.1° bright spot (the traditional stimulus for pursuit) or a small RDC (up to $10 \times 10^9$) composed of up to 200 moving spots all of which moved in the same direction, and at the same speed on a given trial (see Fig. 1). Each spot remained illuminated throughout the trial. Stimuli moved either left or right across a dimly illuminated homogeneous background, and were presented on an X-Y CRT with P4 phosphor at a viewing distance of 57 cm and at a frame rate of 50 Hz. Stimulus duration was always 600 ms.

An advantage of RDCs is that they are devoid of consistent spatial position cues [21]. Since the dots appeared to drift behind an aperture, there was no real object for the pursuit system to follow in our studies,
merely the global motion of the stimulus. In fact, saccades typically seen while pursuing small spots, and which are thought to be driven by position cues [22], were rarely seen in the eye movement records of our observers when tracking cinematograms. Since position cues have been shown to affect smooth eye movements [23], RDCs are particularly well suited for studying how motion alone influences pursuit.

2.3. Experimental procedure

In the first experiment, we compared the smooth pursuit responses to a single spot with those generated by tracking a 10 x 10° RDC (Fig. 1(A)). The stimulus moved at one of five speeds (4, 5, 6, 7, or 8°/s) which were chosen randomly each trial. These speeds were chosen because they fall in the range of speeds used for human smooth pursuit tracking in the past (e.g. [18,24-28]) and the narrow range of speeds allowed us to measure the precision of the pursuit response by computing oculometric speed discrimination thresholds [see ref. [27]]. Briefly, this procedure involves three steps: (a) computing a z-score for each target speed by taking the difference between the mean eye speed for each target speed and that to the middle speed (6°/s in our experiments) and dividing by the S.D. of the eye speed distributions; (b) plotting these z-scores as a function of target speed and fitting a regression line through the data; and (c) evaluating the function for a z-score of 0.67 (which is equivalent to 75% performance). The aperture surrounding the RDC was circular with a fixed diameter of 10°. The other two experiments were done to determine differential responsiveness of pursuit to changes in either the density or the area of RDCs (Fig. 1(B and C)). For these studies, the cinematogram was visible through a rectangular aperture (width constant at 10°) and the target speed was fixed at 8°/s. When stimulus area was the variable of interest, the height of the aperture was varied (0.5-10°), and the density was fixed at 2.0 dots/degree². When stimulus density was the variable of interest, dot density was varied (0.5-10 dots/degree²) and the aperture height was fixed at 2°. These values were selected so that the number of dots presented across aperture and density conditions would span a similar range and allow performance in the aperture and density conditions to be compared.

A trial was initiated with a button press. The observer then fixated a central point for a random amount of time after the button press (up to 1 s) at which time the fixation point was extinguished and the moving stimulus appeared. In the first experiment, observers tracked stimuli in blocks of 100 trials in which speed and direction were selected randomly for each trial. In the other two experiments, observers tracked in blocks of at least 50 trials, within each of which aperture size and density were fixed. Trials were run in blocks to minimize possible pursuit biases towards the response which would be seen to the mean stimulus size or density. The order of blocks was randomized with the constraint that every condition was seen once before any condition was repeated.

2.4. Eye movement recording and analysis

Horizontal and vertical eye position were measured with a Generation V dual-Purkinje-image eyetracker. The overall noise of our eyetracker system was determined to be less than 1 min arc in tests with an artificial eye. Eye position was calibrated to tracker output before each session by having the observer fixate several times at each of four 5° eccentric cardinal positions and at the center while the experimenter adjusted offsets and gains. The eye position signal was differentiated with analog hardware to obtain eye velocity; both were passed through an anti-aliasing filter and sampled at 500 Hz by a Pentium PC, and the data was stored on disk for processing later. Eye position was measured monocularly, and a patch was placed over the other eye.

Eye velocity signals were filtered digitally off-line using a non-causal Butterworth filter (2 pole, cutoff = 50 Hz). Eye acceleration was obtained by digital differentiation of eye velocity records. A second filter (cutoff = 25 Hz) was then used to smooth acceleration traces. Although the filters were non-causal, it should be noted that any smoothing process by definition distorts a trace. Latency is the pursuit parameter most susceptible to smoothing. Therefore, filter parameters were chosen carefully to minimize latency distortion, but to still remove most of the 60 Hz contamination of the signal. Next, the records were desaccaded using an algorithm that finds epochs of time where first eye acceleration exceeds and then falls below a threshold (2000°/s²). However, for the movement to be considered a saccade, the absolute value of eye deceleration then has to rise above and fall below the same threshold. The saccade epoch was excised from the velocity record, and replaced with a line that connected the point preceding the saccade to the point following it in a fashion similar to algorithms used in the past (e.g. [7,29]). All filtering and other data processing were done using Matlab (Mathworks) software.

In the final analysis, the pursuit records were aligned on latency as determined by visual inspection. Since anticipatory smooth eye movements are almost impossible to eliminate, even with the most rigorous randomization of trial parameters [30], records where clear anticipation (> 2°/s) was present were rejected. In most blocks, less than 5% of the trials were rejected for this reason. We made our final measures in the open-loop period of smooth pursuit initiation, since during this period the response is uncontaminated by motion of the
eye itself (e.g. [7,19]). The duration of the open-loop period (which is directly related to latency) has been in the past set at 100 ms, since most of the work establishing this period has been done in monkeys, which have a pursuit latency that approximates this value [19]. The latency of the pursuit response in humans is considerably longer, on the order of 130 ms [31]. Therefore, we have chosen to define the duration of the human open-loop period as 130 ms.

Eye dynamics were characterized in two ways. Since the sole aim of the first set of experiments was to characterize smooth pursuit to larger stimuli, the velocity gain at the end of the open-loop period was quantified to allow comparison of the results with the vast amount of previous literature which has characterized pursuit gain. The open-loop gain was calculated as mean eye velocity in a 20 ms bin centered on 130 ms (the end of the open-loop period) divided by target velocity. The steady-state gain was computed as mean eye velocity between 290 and 350 ms, again divided by target velocity. This particular steady-state interval was chosen in an attempt to capture pursuit gain after the initial overshoot had subsided and before eye velocity started to decline in anticipation of the stimulus stopping.

The other experiments were done to assess the pursuit response to fine-grain manipulations of aperture size and density. Acceleration during the open-loop period is a more sensitive indicator of the pursuit response [25], therefore this measure was used in the second set of experiments. For these experiments, we analyzed mean eye acceleration in a 20 ms bin centered around 90 ms. This specific choice was made because the open-loop period is thought to be composed of two intervals, only the second of which is responsive to the characteristics of the motion [25,29]. In the monkey, this period is 40–100 ms after pursuit onset. When adjusted for human latency, the center of this period is 90 ms, where we placed our analysis bin [25].

3. Results

3.1. Pursuit of RDCs

The open-loop gain of pursuit responses to the RDCs was enhanced relative to that measured during pursuit of the single spot (Fig. 2). The visual impression from the records is that of a larger response, yet a smoother or more precise encoding of the difference in speeds. The enhanced gain was characterized by measuring eye velocity 130 ms after the eye started to move and then comparing it to target speed (Fig. 3). Gain was greater during tracking of all the cinematograms that we tested and for both observers (one-factor ANOVA results: $F_{1,38} = 30.28$, $P < 0.05$), and at all speeds the cinematograms produced a gain greater than one. This is not surprising, since eye velocity during pursuit initiation commonly overshoots target velocity, as seen here when tracking the 4°/s spot. The cinematogram is apparently bringing the eyes into the overshoot phase earlier. However, steady-state pursuit gain to the spot and the RDCs was not different. For subject SW, the gain during pursuit of the spot and RDC were 1.02 (S.E. = ±0.04) and 0.99 (S.E. = ±0.02) respectively. For subject SH the spot gain was 0.86 (S.E. = ±0.03) and the RDC gain was 0.91 (S.E. = ±0.02).

We have argued that RDCs are free of local position cues. Evidence supporting this is a virtual absence of saccades during pursuit initiation to these stimuli. Although saccades appeared later in the trial as the stimulus took the eye towards the edge of the screen, saccades were present in only 8% of trials during the first 130 ms of tracking the RDCs. Saccades were more frequent during the first 130 ms of tracking the single spot and occurred in 70% of those trials.

Smooth pursuit responses to RDCs whose speeds differed by only 1°/s diverged early, while the responses to single spots with the same speed differences were never clearly different throughout the period of pursuit initiation. Even at the end of the open-loop period (130 ms), eye speeds to different single-spot speeds were not always ordered correctly, i.e. the eye speed to a faster stimulus was sometimes slower than that to a slower stimulus. As evidence that the cinematograms produced more precise pursuit performance as well, we computed speed discrimination thresholds for pursuit eye movements to the 10° RDCs and single spot targets using eye velocities obtained at the end of the open-loop period. For observer SW, the difference in speed needed for threshold discrimination ($z$-score = 0.67) for the single spot targets (35.6%, S.E. ± 5.4%) and RDCs (21.7%, S.E. ± 2.2%) was significantly different ($t(6) = 2.37, P < 0.05$). For observer SH, the difference in speed discrimination threshold was even larger (single spot targets: 185.5%; RDCs: 16.3%, S.E. ± 1.4%). However, the considerable extrapolation for SH's single spot threshold made it impossible to determine an error for the estimate and thus a $t$-test could not be performed. These data show that RDCs generated a more precise pursuit response than the single spots.

3.2. Effects of aperture size

The single spot and the cinematogram are qualitatively different stimuli. At least the absence of local position cues, and possibly other aspects of the cinematogram may have encouraged the observers to adopt different tracking strategies for the two stimuli. We reasoned that if the pursuit system performed spatial summation, a systematic increase in pursuit gain should
be present when the spatial extent was varied by increasing the size of the RDCs. To do this, the horizontal dimension of the aperture was held constant at 10° so that the eye would not reach the end of the display at different times, and the vertical dimension was varied.

Eye acceleration measured at 90 ms after pursuit onset was found to increase as a function of aperture size. We also analyzed pursuit latency during tracking of different size RDCs, and found that it decreased in a rather striking fashion as aperture size was increased (Fig. 4). Interestingly, the absolute value of the latency measured while tracking RDCs was always higher than that seen during single-spot pursuit, with single-spot mean latencies being 126.7 ms (S.E. = ± 4.0 ms) for SH and 124.4 ms (S.E. = ± 6.1 ms) for SW.

3.3. Effects of dot density

In the previous experiment, aperture size was found to affect both the latency and initial eye acceleration of pursuit, suggesting that the smooth pursuit system integrates motion information over the visual field. Therefore one might conclude that a larger stimulus produces higher-gain pursuit. However, interpretation of the result is confounded by the fact that as stimulus size was increased, so were the number of dots. To dissociate the effects of stimulus size and dot number, observers pursued RDCs in which dot density was varied while aperture size was fixed.

Eye acceleration measured at 90 ms after pursuit onset was found to also increase as a function of dot density. Latency decreased as a function of density for
observer SH, but remained unchanged for subject SW (Fig. 5).

In order to compare the influence of changes in spatial extent and dot density on smooth pursuit, eye acceleration data from both experiments were plotted on the same axes as a function of the number of dots (Fig. 6). For one of the observers (SH), increasing the number of dots by increasing the aperture size had a greater effect on eye acceleration than increasing dot number by increasing density. For the other observer (SW), both manipulations increased eye acceleration. However, both observers showed a decrease in latency as the number of dots increases. Thus at these values of aperture size and densities, increasing the number of dots appears to be responsible for the systematic changes in the pursuit response.

4. Discussion

Our results suggest that the pursuit system can utilize motion information presented over a large region of the retina to follow a stimulus. We found that small RDCs elicited pursuit responses with higher gain and greater precision than those produced by the traditional single spot. In addition, eye acceleration in the second phase of pursuit initiation increased monotonically with increases in the spatial extent of the cinematogram. Latency decreased monotonically as spatial extent was increased, suggesting that a stronger motion signal can generate an earlier response from the pursuit system. Increasing the dot density in a fixed-size cinematogram also increased acceleration and decreased latency, although the size of the effects varied somewhat between the two observers.

4.1. What is the object of pursuit?

Large stimuli in our experiment enhanced the gain of pursuit, consistent with what other researchers have found [18,20]. This makes sense, given the objects which we normally pursue in the world such as people, dogs or busses usually subtend relatively large visual angles. What is it about a larger stimulus that makes it generate a stronger pursuit response? Since moving objects have both position and velocity cues, the larger stimulus could activate either more motion detectors or more position-sensitive units. Previous studies which showed enhanced pursuit gain with larger stimuli did not differentiate between these alternatives. However, our results suggest that increasing the number of motion detectors activated and hence increasing the strength of the velocity signal itself can facilitate pursuit. In the present set of experiments, there was no object to pursue, rather observers tracked the global motion [32]. The idea that they were pursuing single elements within the array seems unlikely because systematic differences in pursuit dynamics were found as the density or size of the stimulus was varied. In addition, far fewer saccades were generated while pursuing the RDCs than the single spots, and saccades during pursuit are thought to be driven by position cues.

So, what does position information do for the pursuit system, if anything? The latency difference that we saw when observers pursued the RDCs may provide a clue.
As RDC size was increased, the latency of the response decreased. Interestingly, even with our largest stimulus, the latency was still above that seen when observers tracked the single spot. This situation creates a paradox, given that the dynamics of the eye movement were facilitated by larger stimuli. One possible resolution of this paradox is that position cues may be important for triggering smooth pursuit. An argument that supports this is based upon the extraordinarily low latency of normal pursuit. In humans, pursuit starts ~130 ms after the target starts to move (e.g. [31]). This is shorter than normal saccade latencies which are on the order of 180–200 ms (e.g. [33]). It could be that the pursuit system receives privileged information about object motion from brainstem structures that rely on position cues to generate a signal. Cortical areas may then later extract or refine higher-order stimulus properties, thereby allowing more precise image stabilization, but at the expense of longer processing time.

Our experiments with changing the density make a stronger argument, which is that the pursuit system may be better suited to pursue large objects than smaller ones. When the density of our displays was low, the latency was still high. However, only a few dots (i.e. 10) were present in the lowest-density displays used, so any one dot could have been an easy target for the pursuit system to follow. But, even in this case, the system apparently responded as if it was following a field of dots, as opposed to just a single one, since both latency and acceleration in this condition followed the trends established by manipulating the density of the fields. These results are consistent with recent work which demonstrated that when presented with two targets the response of the pursuit system is a vector average of the two motions presented separately [34]. While single-spot pursuit is truly possible, it may not be the modus operandi of the pursuit system.

4.2. Summation or averaging?

Our data imply that the pursuit system uses motion information which is summed across the visual field; however, the results of other studies suggest that it is not just a simple summation. Keller and Khan [7] showed that pursuit of a small spot over a stationary textured background was impaired. Although salience of the spot was a concern in their study, the results have been replicated in a similar situation, but with texture removed in a 4° strip centered on the trajectory of the spot [8]. In that study, the spot was clearly visible while moving, discounting the salience explanation. The results in these studies might be explained by considering not just simple summation but that the motion signal used for pursuit may be a weighted average of the elements in the visual field. In this model, stationary spots would have a zero motion signal, which when
averaged with the signal from the moving spot would produce a signal smaller than that produced by the spot alone. Of course, this idea also has its limitations since people can pursue one set of dots that transparently overlap another set of dots moving in the opposite direction [35]. This suggests that segregation of a scene achieved by assigning different motions to different objects must take place before pursuit in this situation is possible. If segregation is required whenever spatially distributed motion signals are encountered, this process could also contribute to the longer latencies to RDC stimuli.

4.3. Pursuit or OKN?

One important issue that these experiments raise regards the involvement of the optokinetic (OK) system. The OK system is driven by large or full-field motion and evolved to stabilize the world on the retina, presumably to supplement the vestibulo-ocular reflex during self-motion [36]. Although pathways in the accessory optic tract of the pretectum seem to be specialized for the generation of optokinetic eye movements in the primate [37], traditional OK stimuli (subtending 30° visual angle or larger) have been shown to excite neurons in MT and MST [16], the dorsolateral pontine nucleus [38–40], and regions of the cerebellum [41–44], all areas of the brain that are also thought to be involved in smooth pursuit control [3,4]. Furthermore, some characteristics of behavioral pursuit and OK responses can be quite similar. In fact, the initial phase of OK eye movements are thought to be generated by the smooth pursuit system [45] and has been modeled as such [46]. While foveate animals primarily use the OK system to stabilize retinal motion, animals with foveas (most notably primates) can voluntarily pursue an object with smooth eye movements, even if the object’s motion conflicts with self-induced surround motion. It may be that evolutionary forces modified and expanded on the OK system to produce a system that could accommodate voluntary pursuit. If so, the human pursuit system has apparently taken advantage of its evolutionary history, and can utilize motion signals arising from the peripheral retina to increase the stability of a moving image.

4.4. Implications for visual motion processing

Our results complement those of researchers that have studied large-field motion processing with psychophysical techniques (e.g. [10,47]). For instance, Watanuki and Sekuler [10] found that direction discrimination thresholds for RDCs with directional noise improved systematically as the size of the circular aperture increased from 4° to 9°. In addition, Vergheese and Stone [47] found that adding multiple small moving
stimuli over the visual field improved speed discrimination. These past results as well as our own are consistent with a motion processing system that averages motion information across the visual field. To account for the present data, such a system would require subunits that are smaller than the area that is being averaged across to feed into a summing junction. These results are consistent with current computational models of motion processing that incorporate spatial summation (e.g. [48–53]), and suggest that the behavior of the smooth pursuit system can be used to explore characteristics of human motion processing.

Acknowledgements

This work was supported by the National Eye Institute (ROI EY10838), and the Smith-Kettlewell Eye Research Institute. We would like to thank Min Liu for help in data analysis and preparation of the manuscript.

References


