From Following Edges to Pursuing Objects

GUILLAUME S. MASSON1 AND LELAND S. STONE2
1Centre de Recherche en Neurosciences Cognitives, Centre National de la Recherche Scientifique, F13402 Marseille, France; and 2National Aeronautics and Space Administration, Ames Research Center, Human Factors Research and Technology Division, Moffett Field, California 94035-1000

Masson, Guillaume S. and Leland S. Stone. From following edges to pursuing objects. J Neurophysiol 88: 2869–2873, 2002; 10.1152/jn.00987.2001. Primates can generate accurate, smooth eye-movement responses to moving target objects of arbitrary shape and size, even in the presence of complex backgrounds and/or the extraneous motion of non-target objects. Most previous studies of pursuit have simply used a spot moving over a featureless background as the target and have thus neglected critical issues associated with the general problem of recovering object motion. Visual psychophysicists and theoreticians have shown that, for arbitrary objects with multiple features at multiple orientations, object-motion estimation for perception is a complex, multi-staged, time-consuming process. To examine the temporal evolution of the motion signal driving pursuit, we recorded the tracking eye movements of human observers to moving line-figure diamonds. We found that pursuit is initially biased in the direction of the vector average of the motions of the diamond’s line segments and gradually converges to the true object-motion direction with a time constant of approximately 90 ms. Furthermore, transient blanking of the target during steady-state pursuit induces a decrease in tracking speed, which, unlike pursuit initiation, is subsequently corrected without an initial direction bias. These results are inconsistent with current models in which pursuit is driven by retinal-slip error correction. They demonstrate that pursuit models must be revised to include a more complete visual afferent pathway, which computes, and to some extent latches on to, an accurate estimate of object direction over the first hundred milliseconds or so of motion.

INTRODUCTION

In a complex changing visual world, tracking eye movements provide critical support for vision by stabilizing the image of a moving object of interest on the retina and keeping it close to the fovea. The signal driving the smooth component of tracking (pursuit) is dominated by visual motion (Rashbass 1961). Current pursuit models reflect this fact by including a motion-control loop in which the output (eye velocity) attempts to minimize the input (retinal slip) through negative feedback. However, a control strategy based exclusively on the delayed feedback of retinal slip cannot achieve stable steady-state pursuit. The dominant view is that an internal positive-feedback loop within the brain stem and cerebellum acts to enhance the stability of the retinal-slip negative-feedback controller by providing eye-velocity memory at the efferent end (see Lisberger et al. 1987). This theory can explain why pursuit is sustained even when the target is transiently occluded (Becker and Fuchs 1985) or retinally stabilized (Newsome et al. 1988), although eye speed can slowly decay during such interruptions of normal vision.

The visual motion signal driving pursuit has been probed by examining the initial so-called “open-loop” response, which occurs before feedback can have an effect (Lisberger and Westbrook 1985). This and similar studies led to the conclusion that pursuit is driven by the linear sum of derivatives of retinal slip with static nonlinear scaling (Lisberger et al. 1987). However, the standard spot stimulus used in these and most other previous pursuit studies is trivial, in that, unlike natural objects, the “local” motion of the dot-as-image-feature and the “global” motion of the dot-as-object are identical. Real-world moving objects have more complex shapes containing multiple local edges, with different orientations as well as unambiguous features. In processing such objects, the visual system faces the “aperture problem”; the motion of a single local edge is ambiguous, and alone, cannot signal the object’s direction or speed. A motion-integration process is necessary to combine local measurements and to recover the velocity vector of an object or pattern (see Movshon et al. 1985). In this study, we measured the onset of pursuit to line-figure objects to examine the dynamics of the visual motion integration underlying pursuit and to challenge current pursuit models.

METHODS

Eye movements were recorded from four observers (2 naïve), using methods described in detail elsewhere (Masson et al. 2000). Briefly, a PC running REX controlled stimulus presentation and data acquisition. Stimuli were generated with an SGI Octane workstation and back-projected along with the red fixation point onto a large translucent screen (80° × 60°) using a video-projector (1280 × 1024 pixels at 76 Hz). The position of the right eye was sampled at 1 kHz using the scleral search-coil technique. We used a conjoint velocity and acceleration threshold to detect and remove saccades (Krauzlis and Miles 1996).

The stimuli were line-figures (line luminance: 60 cd/m²) of equal area (71°2) but of different shapes and tilts (Beutter and Stone 2000), moving in one of the cardinal directions at 5°/s over a black (<0.1 cd/m²) untextured background (Fig. 1, top). The three shapes were a square diamond (main axes: 11.9°) and clockwise (CW) or counterclockwise (CCW) elongated and tilted diamonds (28.3° by 5° parallelograms tilted ±45° with respect to vertical with internal angles of 14.5° and 165.5°). These square and elongated diamonds have properties similar to those of type I and type II moving plaids, respectively.

Received 4 December 2001; accepted in final form 15 July 2002.

Address for reprint requests: L. Stone, NASA ARC, MS 262–2, Moffett Field, CA 94035-1000 (E-mail: lstone@mail.arc.nasa.gov).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
(Wilson et al. 1992). For the square diamond, the vector average of the edge motions (i.e., the motions orthogonal to the edges) is co-linear with the object-motion direction. For the elongated diamonds, it is biased $\pm 44^\circ$ away from the object's direction (see Beutter and Stone 2000). In experiment 1, all three object types moved for $450 \pm 50$ ms. In experiment 2, only the square and CCW tilted diamonds were used, moving for $700 \pm 50$ ms. Randomly interleaved on one-half the trials, the object was blanked (rendered invisible) 300 ms after stimulus onset and reappeared 91 ms later at the location consistent with continued constant-velocity motion.

Observers had their head stabilized by chin and forehead rests. Each trial started with the presentation of a stationary fixation point. Observers were required to fixate for $300 \pm 150$ ms within a $1^\circ \times 1^\circ$ window. The fixation point was then extinguished, and a moving object was presented. Observers were instructed to track the object center, and trials were aborted if eye position did not stay within $2^\circ$ of the object center (approximately 2% of trials, independent of condition). All conditions were randomly interleaved to minimize cognitive expectations and anticipatory pursuit. We collected approximately 60 trials per condition and observer over several days.

RESULTS
Pursuit onset transients (experiment 1)

Figure 1 illustrates, for one observer, the mean horizontal and vertical speed profiles of the pursuit response to $5^\circ$/s object motion for all object types and motion directions. Results were similar for the other observers. The square diamond elicited pure horizontal or vertical eye movements with negligible off-axis responses (Fig. 1, B and E). However, the tilted diamonds (Fig. 1, A, C, D, and F) transiently elicited oblique pursuit. For example, an upward moving CCW tilted diamond generated rightward as well as upward pursuit (orange traces in Fig. 1, A and D). The off-axis rightward response (Fig. 1A) was transient, started decreasing approximately 100 ms after pursuit onset, and essentially disappeared approximately 200 ms later. Conversely, a CW diamond moving upward generated a transient leftward response (orange trace in Fig. 1C). The initial pursuit direction therefore systematically depended on the object’s tilt, as well as its motion direction.

Figure 2 plots the mean tilt-induced direction error (the pursuit direction for the tilted diamond $-$ the pursuit direction for the square diamond) as a function of time for both object tilts. For all four stimulus directions, in the initial analysis interval (150–190 ms after motion onset), the pursuit direction deviated by approximately $30^\circ$ toward the direction perpendicular to the tilt. Analysis of the performance of individual observers (Table 1) indicates that the differential error (one-half the difference between the pursuit direction for the CW tilt and that for the CCW tilt) decayed approximately exponentially with a mean ($\pm$SD across observers) time constant of $91.9 \pm 3.5$ ms to an insignificant level in the steady state ($0.4^\circ \pm 1.9^\circ$). Furthermore, the mean initial error extrapolated back to the time of the earliest pursuit response (125 ms after motion onset) is $48.6^\circ \pm 6.4^\circ$, which is indistinguishable from that predicted by vector averaging. Similar results were obtained for one observer (GM), tested with both higher ($10^\circ$/s) and lower ($2.5^\circ$/s) object speeds, showing that the finding is robust across a range of absolute speeds.

Pursuit correction transients (experiment 2)

The time course of pursuit direction during initiation could reflect the temporal evolution of the visual input (i.e., from average local motion to object motion) or could result from negative feedback driving the minimization of edge motion (i.e., as retinal motion is driven to 0, eye motion will converge to object motion). To distinguish between these possibilities,
we performed a follow-up experiment in which we transiently blanked the target during ongoing pursuit.

Figure 3A plots the average pursuit response of a naive observer to vertical motion of a CCW tilted diamond. As above, pursuit showed an initial bias toward the vector-average direction. Target blanking resulted in a decrease in vertical eye speed (dashed lines). After the reappearance of the target, this tracking error was corrected by a purely vertical eye acceleration with the approximately 100- to 150-ms delay expected for a visual process. Therefore the initial open-loop, visually driven correction starts off in the object-motion direction before feedback is available. The fact that retinal slip after blanking was only approximately 2°/s cannot explain the absence of a transient direction bias in the postblanking correction, as similarly low retinal speeds generate large transient direction biases during pursuit onset. Figure 3, B–E, shows that large errors occurred after initial target appearance (mean absolute error ± SD across observers: 30.9° ± 5.3°) but not after target reappearance (2.0° ± 0.6°).

**DISCUSSION**

Our data show that pursuit is initiated in the direction of the average of the local edge motions, but converges toward that of object motion over a period of approximately 200 ms. The blanking experiment shows that this process need not repeat itself once the direction of object motion has been established, even when large visually driven corrective eye movements are needed and performed. This demonstrates that the object-motion signal driving steady-state pursuit is not simply the result of negative feedback working to minimize retinal edge motion.

Our findings are consistent with those of previous primate neurophysiological and human behavioral studies. Visual motion processing is hierarchical, proceeding through multiple

### TABLE 1. Best-fitting exponential-decay parameters to the differential direction-error data

<table>
<thead>
<tr>
<th>Observer</th>
<th>Right</th>
<th></th>
<th></th>
<th>Left</th>
<th></th>
<th></th>
<th></th>
<th>Up</th>
<th></th>
<th></th>
<th></th>
<th>Down</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A (°)</td>
<td>τ (ms)</td>
<td>B (°)</td>
<td>A (°)</td>
<td>τ (ms)</td>
<td>B (°)</td>
<td>A (°)</td>
<td>τ (ms)</td>
<td>B (°)</td>
<td>A (°)</td>
<td>τ (ms)</td>
<td>B (°)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GM</td>
<td>190</td>
<td>104</td>
<td>-4.5</td>
<td>207</td>
<td>105</td>
<td>-7.5</td>
<td>214</td>
<td>87</td>
<td>-1.5</td>
<td>243</td>
<td>87</td>
<td>-0.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AK</td>
<td>268</td>
<td>83</td>
<td>1.6</td>
<td>281</td>
<td>88</td>
<td>1.8</td>
<td>50</td>
<td>93</td>
<td>7.0</td>
<td>95</td>
<td>101</td>
<td>5.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LS</td>
<td>105</td>
<td>114</td>
<td>-6.5</td>
<td>109</td>
<td>98</td>
<td>-3.9</td>
<td>392</td>
<td>72</td>
<td>-2.8</td>
<td>358</td>
<td>87</td>
<td>-4.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IB</td>
<td>163</td>
<td>90</td>
<td>-3.3</td>
<td>226</td>
<td>82</td>
<td>-3.9</td>
<td>111</td>
<td>107</td>
<td>2.3</td>
<td>257</td>
<td>71</td>
<td>8.5</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ E = A \cdot \exp(-t/\tau) + B \] with \( E \) being half the differential direction error (i.e., the difference between the direction errors observed for the CW and CCW diamonds), \( t \) is time after motion onset, \( \tau \) is the time constant of the exponential decay, and \( A \) and \( B \) are the scaling and offset parameters, respectively. Data were fit using the Matlab simplex algorithm.
cortical tiers starting in primary visual cortex (V1), moving through the middle temporal area (MT), and then onto a number of areas including the medial superior temporal area (MST), consistent with theoretical analyses that identified the need for a sequence of computations to derive object motion (e.g., Movshon et al. 1985; Newsome et al. 1988). V1 neurons respond to local edge motion, while some MT neurons respond better to the global stimulus velocity (Movshon et al. 1985), with their direction selectivity appearing to shift from edge-motion to object-motion direction over time (Pack and Born 2001). It has previously been shown that the smooth eye-movement response to moving patterns or lines reflects a similar dynamic process (Masson et al. 2000; Pack and Born 2001). Our results extend these findings to the pursuit of stimuli with multiple edges and features as found in real objects, and quantify the time course in humans.

Our blanking results show that the dynamics of pursuit direction are not simply a manifestation of negative feedback indirectly driving eye velocity toward object velocity through minimization of retinal edge motion. This is consistent with the earlier observation of significantly different steady-state pursuit responses to two objects with identical edge motions, but different perceived object motions (Stone et al. 2000). Our results also parallel the observation that the perceived direction of moving type II plaids (Wilson et al. 1992) or tilted lines (Lorenceau et al. 1993) is initially biased toward the edge-motion direction(s), but converges to the true motion direction in approximately 100 ms. Pursuit models must therefore be modified to incorporate more complex visual processing, which provides memory within a time-varying, straight-through object-motion pathway (Newsome et al. 1988; Stone et al. 2000), as opposed to current models, which have a time-invariant straight-through retinal-motion pathway and provide all memory through feedback at the efferent end (Lisberger et al. 1987). Furthermore, although studies have established links between the visual signals used for steady-state perception and pursuit (see Stone et al. 2000), future studies are needed to determine if the dynamic visual processing for pursuit is shared with perception and to establish the behavioral limits and neural locus of the observed object-motion memory.

The authors thank Drs. B. Beutterm and A. Krukowski for comments on earlier drafts of the manuscript.

This study was supported by the Centre National de la Recherche Scientifique, the Ministère de la Recherche (ACI-2000-5052), and National Aeronautics and Space Administration (RTOPs 711-51-12 and 131-20-30).

REFERENCES


FIG. 3. Error correction during ongoing pursuit. A: average horizontal and vertical pursuit speeds for the upward and downward motion of a CCW tilted diamond with (—) and without (——) blanking (shaded interval). B–E: mean ± SD (across observers) direction error in 2 intervals (black bars) within the open-loop period: 150–190 ms after target appearance and after the time of target re-appearance.


