Motion perception during sinusoidal smooth pursuit eye movements: Signal latencies and non-linearities

Jan L. Souman
Max Planck Institute for Biological Cybernetics, Tübingen, Germany

Tom C. A. Freeman
School of Psychology, Cardiff University, Cardiff, United Kingdom

Smooth pursuit eye movements add motion to the retinal image. To compensate, the visual system can combine estimates of pursuit velocity and retinal motion to recover motion with respect to the head. Little attention has been paid to the temporal characteristics of this compensation process. Here, we describe how the latency difference between the eye movement signal and the retinal signal can be measured for motion perception during sinusoidal pursuit. In two experiments, observers compared the peak velocity of a motion stimulus presented in pursuit and fixation intervals. Both the pursuit target and the motion stimulus moved with a sinusoidal profile. The phase and amplitude of the motion stimulus were varied systematically in different conditions, along with the amplitude of pursuit. The latency difference between the eye movement signal and the retinal signal was measured by fitting the standard linear model and a non-linear variant to the observed velocity matches. We found that the eye movement signal lagged the retinal signal by a small amount. The non-linear model fitted the velocity matches better than the linear one and this difference increased with pursuit amplitude. The results support previous claims that the visual system estimates eye movement velocity and retinal velocity in a non-linear fashion and that the latency difference between the two signals is small.

Keywords: smooth pursuit, sinusoidal motion, signal latency, extra-retinal, eye movements


Introduction

The ability to judge visual motion is critical in such daily activities as driving a car, playing ball sports, or walking on a crowded pavement. As we move through the world, our head and eyes rotate to track objects of interest, helping us to avoid obstacles and move towards our goal. The retinal input to the visual system therefore consists of a complex pattern of motion that arises from a number of different sources. Nevertheless, observers are able to judge visual motion and move about effectively in most situations. This suggests that the visual system somehow compensates for the retinal effects of eye movements and self-motion.

In this study, we concentrate on how visual motion is perceived during smooth pursuit eye movements. Smooth pursuit adds motion to the retinal image and thereby complicates the relationship between retinal motion and motion in the world. One way the visual system overcomes this problem is to add estimates of eye velocity and retinal velocity in order to recover motion with respect to the head (for physiological evidence, see Ilg, Schumann, & Thier, 2004; Newsome, Wurtz, & Komatsu, 1988; Tanaka, 2005). Figure 1 shows a general framework for this type of compensation process and suggests three important issues. The first concerns the nature of the signal that estimates eye velocity (denoted \( \hat{E} \) in the figure). The traditional view is that this signal is “extra-retinal,” depending on an efference copy of the oculomotor command (Sperry, 1950; Von Holst, 1954; Von Holst & Mittelstaedt, 1950) and/or proprioceptive feedback from the eye muscles (Gauthier, Nommay, & Velcher, 1990a, 1990b; Skavenski, 1972; Wang, Zhang, Cohen, & Goldberg, 2007). However, some authors have suggested that retinal and extra-retinal signals interact to produce a composite estimate of eye velocity, as indicated by the oblique dashed line in Figure 1 (Brenner & Van den Berg, 1994; Crowell & Andersen, 2001; Goltz, DeSouza, Menon, Tweed, & Vilis, 2003; Haarmeier & Thier, 1996; Harris, 1994; Post & Leibowitz, 1985; Turano & Massof, 2001; Wertheim, 1990, 1994).

A second issue suggested by Figure 1 concerns the transducer functions that relate the estimated eye velocity \( \hat{E} \) and retinal velocity \( \hat{R} \) to the physical velocities \( E \) and \( R \) (indicated by \( f \) and \( g \) in the figure). According to the standard linear model, the perceived head-centered motion \( \hat{H} \) is the sum of two linear velocity estimates (e.g., Freeman & Banks, 1998; Souman, Hooge, & Wertheim, 2005a):

\[
\hat{H} = r\hat{R} + e\hat{E},
\]

where \( r \) and \( e \) are the gains of the retinal and eye movement signals, respectively. Previous studies have
shown that many examples of motion perception during smooth pursuit are well approximated by this model. In particular, the linear model is able to quantify how the Aubert–Fleischl phenomenon (moving objects appear slower when pursued) and the Filehne illusion (stationary objects appear to move during pursuit) vary with eye velocity (Freeman, 2001). It also describes perceived motion direction during pursuit quite well (Souman et al., 2005a). However, the linear model is less able to describe other instances of velocity perception, such as general velocity matching tasks in which observers are asked to estimate the speed of stimuli that are neither stationary nor moving at the speed of the pursuit target (Freeman, 2001; Turano & Massof, 2001). In these cases, models with non-linear transducers fit the data better, although the improvement is quite small (Souman, Hooge, & Wertheim, 2006). A specific example of a non-linear model is discussed in more detail below.

A third issue suggested by Figure 1 concerns the temporal relationship between the two signals. This has received little attention in the literature. Both signals are, of course, neural in origin, and are therefore likely to be subject to transmission delays. Crucially, these delays may not be equal to each other. This would not affect visual motion perception if eye velocity and retinal motion were always constant because the sum of the two signals would produce the same estimate of head-centered motion, regardless of which signal lagged the other. However, eye velocity is rarely constant, so transmission delays could present a real problem to an observer trying to judge motion during pursuit.

Similar problems in timing arise when judging position rather than motion during pursuit. In this case, different transmission delays affect location judgments whether pursuit velocity is constant or time varying. Previous work has shown that observers make localization errors compatible with an eye position signal that leads the retinal signal by about 100 ms (Brenner, Smeets, & Van den Berg, 2001; Mateeff, Yakimoff, & Dimitrov, 1981; Schlag & Schlag-Rey, 2002; Ward, 1976). In comparison, little work has investigated signal latencies in head-centered motion perception, partly because nearly all previous studies combined constant pursuit target motion and stimulus velocity. An exception is the study by Freeman, Banks, and Crowell (2000), who presented observers with a ground-plane simulating forward self-motion in a constant direction. When tracking a sinusoidally moving pursuit target, observers experienced a “slalom illusion,” as though they were slalom skiing through the scene. Freeman et al. had observers null the slalom illusion (and a comparison Filehne condition) by adjusting the amplitude and the phase of a simulated eye rotation. They then applied the linear model (Equation 1) to their data, in order to estimate the gain ratios and the latency differences between eye movement signal and retinal signal. They concluded that latency differences only played a minor role. Another study investigating motion perception during sinusoidal pursuit was reported by Mergner, Rottler, Kimmig, and Becker (1992). In their experiment, observers had to null the perceived motion of the pursuit target in real time by using a joystick, or, in a separate condition, indicated the current direction of the pursuit target with an unseen pointer. Both measures showed a small but increasing lag relative to the pursuit target for higher frequencies. From their results, it is not clear to what extent this lag represents perceptual latencies. They did not report pursuit accuracy, so the reported lag might have been simply due to the eyes lagging the pursuit target.

In the current study, we investigated signal latencies using a more straightforward two-interval speed matching task. We first describe how the linear model applies to this task, highlighting the way differences in signal size and latency impact on the subsequent motion percept. In

![Figure 1](https://jov.arvojournals.org/pdfsaccess.ashx?url=/data/journals/jov/933528/)
Models

According to the standard linear model, perceived head-centered velocity $\mathbf{H}$ is a linear combination of retinal image velocity $\mathbf{R}$ and eye velocity $\mathbf{E}$ with gains $r$ and $e$, respectively (Equation 1). If $\mathbf{R}$ and $\mathbf{E}$ vary in time and the signals have different latencies, this will affect the perceived velocity $\mathbf{H}(t)$ at time $t$. As we used sinusoidal movements of one single frequency $f$ in our experiment, these latencies translate into phase shifts:

$$\dot{\mathbf{H}}(t) = rR\sin(2\pi ft + \varphi + \rho) + eE\sin(2\pi ft + \theta + \epsilon), \tag{2}$$

where $R$ and $E$ now represent movement amplitudes, $\varphi$ is the phase of the retinal image motion with respect to the pursuit target, $\rho$ is the phase shift of the retinal signal, $\theta$ is the phase of the eye movement with respect to the pursuit target, and $\epsilon$ represents the phase shift of the eye movement signal. Figure 2 illustrates this equation in a phasor plot. Sinusoidal motion is represented as a vector, with the angle with respect to the positive $x$-axis indicating phase and the distance to the origin representing amplitude. Adding sinusoids of the same frequency is equivalent to adding vectors in the phasor plot. By simple trigonometry, the amplitude $\dot{\mathbf{H}}$ of the perceived motion $\mathbf{H}$ satisfies the equation:

$$\dot{\mathbf{H}}^2 = (rR)^2 + (eE)^2 + 2reR\cos(\theta - \varphi + \epsilon - \rho). \tag{3}$$

In our experiment, observers judged the motion of a sinusoidally moving random dot pattern in two separate intervals. In the pursuit interval, they tracked a moving target with their eyes while making the judgment, whereas in the fixation interval the target was stationary. A staircase procedure was used to determine the point where the dot pattern appeared to have the same peak velocity in the two intervals (the point of subjective equality). At this point, the perceived motion amplitude $\dot{\mathbf{H}}_p$ in the pursuit interval equals the perceived amplitude during fixation $\dot{\mathbf{H}}_f$.

This gives:

$$\dot{\mathbf{H}}_f = \dot{\mathbf{H}}_p \Leftrightarrow (rR_p)^2 = (rR_p)^2 + (eE)^2 + 2reR_p\cos(\theta - \varphi + \epsilon - \rho), \tag{4}$$

![Figure 2](https://jov.arvojournals.org/pdfaccess.ashx?url=/data/journals/jov/933528/)

Figure 2. Applying the linear model to the peak velocity judgment task used in the current experiments. Circles represent sinusoidal motion in polar coordinates. The angle with the positive horizontal axis indicates the phase of a sinusoidal movement with respect to the pursuit target $\mathbf{T}$. The distance to the origin indicates the amplitude of the movement. The left-hand figure shows the model applied to the pursuit interval. $\mathbf{T}$ represents the motion of the pursuit target, which by definition had zero phase. $\mathbf{E}$ is the actual eye movement, with phase $\theta$. $\mathbf{H}_p$ is the head-centered motion of the stimulus shown during pursuit, and $\mathbf{R}_p$ is the resulting retinal motion with phase $\varphi$. $\mathbf{E}$ and $\mathbf{R}_p$ represent the estimates of eye movement and retinal image motion made by the visual system, with phase lags $\epsilon$ and $\rho$, respectively. $\mathbf{H}_p$ is the sum of these two and represents the estimated head-centered velocity of the stimulus during pursuit. The right-hand figure shows the same for the fixation interval, where both the fixation target motion $\mathbf{T}$ and the eye movement $\mathbf{E}$ equal zero. Consequently, the retinal image motion $\mathbf{R}_f$ of the stimulus equals the head-centered motion $\mathbf{H}_f$ and the estimated head-centered velocity $\dot{\mathbf{H}}_f$ equals the estimated retinal image velocity $\mathbf{R}_f$. The amplitude of the head-centered motion $\mathbf{H}_f$ in the fixation interval was varied according to a staircase procedure, while its phase was randomly chosen in every trial.
where subscripts $f$ and $p$ refer to fixation interval and pursuit interval, respectively. Dividing both sides by $r^2$ gives

$$R_f^2 = R_p^2 + \left(\frac{e}{r}E\right)^2 + 2\frac{e}{r}R_pE\cos(\theta - \varphi + \varepsilon - \rho).$$

According to the linear model, therefore, the velocity matches are determined by two free parameters: the gain ratio $e/r$ and the phase difference $\varepsilon - \rho$. Note that the individual gains $e$ and $r$ and the individual phases $\varepsilon$ and $\rho$ cannot be resolved (see Freeman, 2001; Freeman & Banks, 1998; Souman et al., 2006).

We compared the linear model to the non-linear model of Freeman (2001). The models differ in the type of speed transducers used to convert input speed into output signal. Freeman (2001) assumed a power law relationship (see also Turano & Massof, 2001):

$$\hat{R} = (R + 1)^e - 1$$

and

$$\hat{E} = (E + 1)^e - 1,$$

with power coefficients $r$ and $e$. Replacing the linear relationships in Equation 4 by these non-linear transducers of motion amplitude and solving for the retinal amplitude during fixation at the point of subjective equality gives

$$R_f = \left(1 + \sqrt{R_p^2 + \hat{E}^2 + 2R_p\hat{E}\cos(\theta - \varphi + \varepsilon - \rho)}\right)^{\frac{1}{e}} - 1,$$

with $R_p$ and $\hat{E}$ defined as in Equations 6 and 7, respectively. Consequently, the non-linear model has three free parameters: the two power coefficients $r$ and $e$ and the phase difference $\varepsilon - \rho$. As with the linear model, the individual transducer functions (Equations 6 and 7) can only be estimated up to an arbitrary scale factor (for further discussion, see Freeman, 2001).

**Model predictions**

Figure 3 shows the predicted amplitude matches for the linear model over a range of gain ratios $e/r$, with the phase difference between the signals set to zero ($\varepsilon - \rho = 0$). We distinguish between cases with accurate pursuit (Figures 3A–3B) and with a 15° pursuit lag (Figures 3C–3D). In Figure 3A, the amplitude matches are shown as a function of the retinal phase in the pursuit interval, with retinal amplitude held constant at 1°. This corresponds to one set of conditions in our experiment. The linear model predicts that the squared amplitude matches should lie on a sinusoid. Veridical amplitude matches would equal the head-centered motion amplitude of the motion on the screen and fall on the dotted line. This corresponds to a gain ratio $e/r = 1$. Reducing the gain ratio produces similarly shaped curves, but with smaller amplitudes and different offsets from zero. If an observer completely fails to compensate for the effects of the eye movements at all (i.e., $e/r = 0$), the amplitude matches will be equivalent to the retinal motion amplitude (horizontal dashed line).

The model predictions show an important counter-intuitive feature when eye movements are only partially compensated for. As the colored curves in Figure 3A show, the predicted amplitude matches do not always lie between the head-centered and retinal amplitudes but can actually fall below the retinal amplitude for higher retinal phases. Not shown is how a phase difference between the
two signals ($\epsilon - \rho$) affects the predictions. This is more straightforward than the effect of the gain ratio. A phase difference causes the curves to shift horizontally (see Equation 5).

In other conditions of the experiment, we kept the phase relationship between the motion stimulus and the pursuit target constant (90°) and varied the relative amplitude. As shown in Figure 3B, the predicted amplitude matches now lie on a parabola. Different gain ratios $e/r$ produce curves with different heights and slopes. However, the amplitude matches only fall below the retinal amplitude (dashed line) if there is a non-zero phase difference $\epsilon - \rho$ between the two signals.

The range of retinal phases and amplitudes shown in Figures 3A–3B assume pursuit is accurate. However, in our experiment we found that the eyes lagged the pursuit target systematically by about 15°. The predictions of the linear model when this pursuit lag is taken into account are shown in Figures 3C–3D. The range of retinal phases in Figure 3C is compressed relative to that with accurate pursuit and retinal amplitude during pursuit is no longer constant. Figure 4 illustrates why this is the case. The open circles represent conditions in which the phase of the motion stimulus with respect to the pursuit target was varied, while relative amplitude was kept constant. Because of the pursuit lag (compare the filled square with the open one), the actual retinal image motion is shifted away from the origin (filled circles). This shift both compresses the range of retinal phases and makes the retinal amplitude non-constant in Figure 3C. In Figure 3D, the pursuit lag causes the curves to shift rightward.

Predictions for the non-linear model exhibit roughly the same pattern as those described above. The main difference is that the local shapes of the curves alter (see Results, below). The hypothetical phase difference $\epsilon - \rho$ has a similar effect in both models.

### Experiment 1

#### Methods

**Participants**

The first author and five paid volunteers participated in the experiment (4 males, 2 females; median age 26 years). They all had normal or corrected-to-normal vision. All naïve participants gave their written informed consent and the experiment was conducted in agreement with the 1964 Declaration of Helsinki.

**Apparatus and stimuli**

Visual stimuli were presented on a large cylindrical screen (radius 350 cm; field of view 240° horizontally, ~45° vertically). This made the vertical edges of the screen invisible to the participant (in pilot studies we found that head-centered motion judgments were quite inconsistent when using a smaller stimulus surrounded by a window moving with the pursuit target). The stimuli were projected onto the screen by three projectors at a frame rate of 60 Hz. Alignment of the images of the three projectors was carried out using a custom built system. The participant was seated with his or her head in the center of the screen, supported by a chin rest. The position of both eyes was registered at 250 Hz with an infrared video-based eye tracker (Eyelink I, SMI Sensomotoric Instruments).

The motion stimulus that had to be judged by the observers consisted of two horizontal bands of random dot pattern placed above and below a fixation target (see Figure 5). Each dot subtended ~0.5° and had a luminance of ~2.5 cd/m² (the fixation target had a small black hole at its center to improve fixation). The density of the dot patterns was 0.1 dot/deg². The bands were 10° high, separated by 10° and covered the entire screen horizontally. Both dot pattern and target could be made to move sinusoidally with an independent amplitude and phase. The frequency of oscillation was 0.5 Hz throughout.
Procedure

Each trial consisted of a pursuit interval followed by a fixation interval (Figure 5). In the pursuit interval, observers tracked a fixation target moving horizontally at eye height with an amplitude of 3°. The sinusoidal motion was presented for 2 periods (i.e., 4 s). At the beginning of the second period, the dot pattern appeared, moving with an amplitude and phase determined by the condition being tested.

In the fixation interval, the target remained stationary at the center of the screen. After 2 s, the dot pattern appeared, moving with an amplitude determined by two randomly interleaved 1 up/1 down staircases. The phase of the dot pattern with respect to fixation target onset was selected at random from trial to trial for this interval. Observers were instructed to compare the head-centered motion of the dot pattern in the two intervals and choose which had the greater peak velocity with respect to their head. The staircases made logarithmic adjustments to the motion amplitude in the fixation interval, homing in on the point at which the perceived peak velocities in the two intervals matched.

The phase and amplitude of the random dot pattern motion in the pursuit interval were varied in 11 separate conditions. Phase and amplitude were defined with respect to the motion of the fixation target. The conditions consisted of two cross-sections of this 2D space of amplitude and phase. In one cross-section, the relative motion amplitude was 1°, while the phase ranged from 0° to 180° in 30° steps (see open circles in Figure 4). In the other cross-section, the amplitude was 0°, 0.5°, 1°, 2°, or 3° with a relative phase of 90° (omitted from Figure 4 for clarity). All conditions were replicated five times. Within each replication, conditions were tested in random order. The first replication was used as practice and the data were not included in the final analysis. The total experiment took approximately ten hours per observer.

Data-analysis

Eye movement data were analyzed off-line. To simplify model fitting, trials were excluded from further analysis if (1) they contained saccades coinciding with presentation of the random dot pattern; (2) pursuit gain differed more than 15% from unity; or (3) phase lag was higher than 25°. Saccades were detected using a velocity criterion of 35°/s or higher. To further reduce the variation in retinal motion, we also discarded trials in which pursuit gain or phase differed more than 2 standard deviations from their respective means. Fifty percent of the trials were excluded.
this way. These strict criteria were used to isolate a collection of trials in which both eye movements and the associated retinal velocities were approximately uniform. After removing trials with inaccurate pursuit, the eye movements in the remaining trials had a pursuit gain close to unity and a phase of about $-15^\circ$. This small pursuit lag is commonly found with sinusoidal pursuit at the frequency we used (Barnes, Barnes, & Chakraborti, 2000; Freeman et al., 2000; Leigh & Zee, 1999; Lisberger, Evinger, Johanson, & Fuchs, 1981).

For all observers, we computed for each condition the average retinal amplitude and phase in the pursuit interval (note that the circular mean was used for phases, which avoids problems of periodicity; see Batschelet, 1981). In order to estimate the amplitude matches, the point of subjective equality was determined by fitting a cumulative Gaussian to all the trials aggregated per condition for each observer. We used the maximum-likelihood procedure described by Wichmann and Hill (2001), with lapse rate included as a free parameter. Model fitting was implemented in MatLab (version 7.3) using least-squares minimization and the \texttt{fminsearch} function. The fitting was performed on the mean squared amplitude matches, averaged across observers. Local minima were avoided by repeating the fitting procedure with different initial values for the parameters.

In order to evaluate model performance, we used Akaike’s Information Criterion (Akaike, 1974; Burnham & Anderson, 2004). This criterion is based on the mean squared prediction errors and also takes into account the number of free parameters, which differed between the linear and non-linear models. It is computed as

$$AIC_c = n \log \left( \frac{\sum \hat{e}_i^2}{n} \right) + 2k + \frac{2k(k+1)}{n-k-1},$$  \hspace{1cm} (9)

where $n$ is the number of observations, $k$ represents the number of free parameters, and $e_i$ are the errors in model predictions. The last term in Equation 9 corrects for the small number of observations (Hurvich & Tsai, 1989; Shono, 2000). Lower values of $AIC_c$ indicate better model performance. Below, we report the differences in $AIC_c$ between the two models, as the absolute values are not interpretable (Burnham & Anderson, 2004).

**Results**

Figure 6 shows the mean squared amplitude matches, averaged across the six observers. In Figure 6A, the amplitude matches are shown for the conditions in which the phase of the motion stimulus with respect to the pursuit target varied, while the relative amplitude was constant ($1^\circ$). As in Figure 3, the dotted lines define veridical performance, while the dashed lines show what the amplitude matches would be for an observer judging retinal image motion alone. For small retinal phases, the observed amplitude matches fell between these two curves. For larger retinal phases, the amplitude matches...
were lower than the retinal amplitudes during pursuit, exactly as predicted by the linear model. This pattern was exhibited by all six observers. In the conditions in which amplitude varied and relative phase was constant (90°), the amplitude matches increased with retinal amplitude (Figure 6B). For the highest retinal amplitudes, they were lower than the retinal amplitude. The results therefore suggest that observers partially accounted for the effects of the eye movements on retinal image motion.

Figure 6 also shows the best fitting model curves for the linear model (blue) and the non-linear model (red). The linear model captures the overall trends in the data quite well. In particular, it is able to explain how the amplitude matches can cross the dashed curve defining retinal motion amplitude. A better fit was obtained with the non-linear model (AICc was 1.91 lower for this model than for the linear one). More specifically, the non-linear model was better able to capture the squared amplitude matches for low retinal phases (Figure 6A) and low retinal amplitudes (Figure 6B).

The best fitting parameter values are listed in Table 1. For the linear model, the gain ratio $e/r$ was close to 0.6, which is similar to previous estimates reported in the literature (Freeman, 2001; Freeman & Banks, 1998; Freeman et al., 2000; Souman et al., 2005a, 2006). The estimated phase difference $\phi - \theta$ between the two signals was about $-12^\circ$, indicating that the eye movement signal lagged the retinal signal by about 67 ms. These results are similar to those obtained by Freeman et al. (2000) using a nulling task. For the non-linear model, the ratio of the two power coefficients $e/r$ was 1.45/1.81, which again is very similar to the values reported previously (Freeman, 2001; Souman et al., 2006). The estimated phase difference $\phi - \theta$ for the non-linear model was $-19^\circ$, which is slightly larger than that of the linear model, but in the same direction.

The findings suggest that while the two motion signals differ in gain, the latency difference is rather small. Moreover, the better fit obtained with the non-linear model suggests the visual system uses non-linear speed transducers to estimate eye velocity and retinal velocity. Similar conclusions have been arrived at by Freeman (2001), Souman et al. (2006), and Turano and Massof (2001). However, Experiment 1 only investigated a single pursuit amplitude, making it difficult to discriminate clearly between linear and non-linear speed transducers in the case of the eye movement signal. Experiment 2 therefore examined velocity matching over a larger number of different pursuit amplitudes.

### Experiment 2

#### Methods

**Participants**

The first author and three paid volunteers participated in the experiment (all males; median age 25 years). They all had normal or corrected-to-normal vision. The naïve participants gave their written informed consent and the experiment was conducted in agreement with the 1964 Declaration of Helsinki. None of the naïve participants had participated in Experiment 1.

**Design and procedure**

The experiment followed the same procedure as Experiment 1. Three different pursuit target amplitudes were used (2°, 3°, and 4°) and the number of conditions was reduced compared to those in Experiment 1. Relative phases of 0°, 90°, and 180°, combined with 1° amplitude and amplitudes of 0° and 3° combined with 90° phase were used. All five conditions were combined with the three pursuit amplitudes, creating 15 different conditions in total. Again, each condition was replicated five times, with the first replication being considered practice. In total, the experiment took about 15 hours per observer.

#### Results

After removing 30% of the trials using the eye-movement exclusion criteria defined above, the remaining trials had a pursuit gain close to unity and a phase lag of $\sim15^\circ$. These values are similar to those found in Experiment 1 and did not depend on pursuit amplitude.

Figure 7 shows the average amplitude matches as a function of retinal phase and amplitude. The pattern of results was similar to that obtained in Experiment 1. The amplitude matches decreased for larger retinal phases and increased for larger retinal amplitudes. The linear model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Linear model</th>
<th>Non-linear model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ratio $e/r$</td>
<td>Experiment 1: 0.56</td>
<td>Experiment 1: 0.56</td>
</tr>
<tr>
<td>Phase difference $\phi - \theta$</td>
<td>Experiment 1: -11.80</td>
<td>Experiment 1: -18.89</td>
</tr>
<tr>
<td></td>
<td>Experiment 2: 0.68</td>
<td>Experiment 2: 0.85</td>
</tr>
<tr>
<td></td>
<td>Experiment 2: -6.97</td>
<td>Experiment 2: -6.95</td>
</tr>
</tbody>
</table>

Table 1. Best fitting parameter values for the linear model and the non-linear model. For the linear model, the ratio $e/r$ represents a gain ratio; for the non-linear model it is the ratio between the two power coefficients (see Equation 8).
Figure 7. Average amplitude matches (±1 SEM) and best fitting model curves (Experiment 2). The different rows show the data for the 2°, 3°, and 4° pursuit amplitudes, respectively. The left-hand panels show the squared velocity matches as a function of retinal phase in the pursuit interval. The right-hand panels show the velocity matches as a function of retinal amplitude. The best fitting model curves are shown for the linear model (blue) and the non-linear model (red).
transducers and signal independence. lag as well as examining evidence in support of non-linear retinal signal. Below, we discuss possible reasons for this model. However, both models revealed a small but model or as a ratio of power exponents for the non-linear model (expressed either as a gain ratio in the case of the linear model). Again, both power coefficients were above unity ($e = 1.48$; $r = 1.74$). Both models estimated the phase difference $\varepsilon - \rho$ to be about $-7^\circ$, indicating that the eye movement signal lagged the retinal signal by about 39 ms.

### Discussion

Most studies on motion perception during smooth pursuit use constant speeds for pursuit and retinal motion. Consequently, the issue of the temporal relationship between the eye movement signal and the retinal signal has received little attention. In the present study, we used sinusoidal motion for both the pursuit target and motion stimulus to investigate this issue. To model the results, we assumed that eye movement velocity and retinal velocity are coded by two independent signals, each implementing specific transducer relationships between input velocity and output estimate. Non-linear transducers produced better fits. Our results also showed that observers only partially compensated for the effects of the eye movements on retinal image motion. This agrees with previous findings (De Graaf & Wertheim, 1988; Freeman, 2001; Freeman & Banks, 1998; Freeman et al., 2000; Mack & Herman, 1973, 1978; Souman et al., 2005a, 2006; Souman, Hooge, & Wertheim, 2005b; Turano & Heidenreich, 1999; Turano & Massof, 2001; Wertheim, 1994). Compensation was partial mainly because the two signals differed in size (expressed either as a gain ratio in the case of the linear model or as a ratio of power exponents for the non-linear model). However, both models revealed a small but systematic lag of the eye movement signal relative to the retinal signal. Below, we discuss possible reasons for this lag as well as examining evidence in support of non-linear transducers and signal independence.

### Signal latencies

Both the linear and non-linear model suggest that the eye movement signal lagged the retinal signal by around 10° (~55 ms). This is similar to the latency difference found by Freeman et al. (2000) using a nulling task. The estimated difference was smaller in our second experiment than in the first one, but this may have been because we sampled the amplitude/phase space more coarsely or because different observers participated in the two experiments.

The latency difference we found is different to that reported by studies of perceived position during pursuit. Errors in localization have been taken to imply that the eye position signal leads the retinal signal by ~100 ms (Brenner et al., 2001; Mateeff et al., 1981; Schlag & Schlag-Rey, 2002; Ward, 1976). Not only is this latency difference almost twice as large as we find here, it is also in the opposite direction. It is possible that this reflects differences in the signals used to encode position and motion in the brain. Area 7a (in monkeys) plays an important role in integrating eye position with retinal location (Andersen, 1989; Andersen, Essick, & Siegel, 1985, 1987; Andersen & Mountcastle, 1983). Conversely, area MST is crucial for motion perception during smooth pursuit (Barton et al., 1996; Bradley, Maxwell, Andersen, Banks, & Shenoy, 1996; Ilg & Thier, 2003; Newsome et al., 1988; Pack, Grossberg, & Mingolla, 2001; Shenoy, Bradley, & Andersen, 1999). The separation of motion and position processing in the cortex may be a simple, if somewhat uninformative, reason why we find different signal latencies for motion perception than those reported for localization.

An additional reason may be that most localization studies use flashed stimuli. The flash-lag literature suggests that these may have a longer processing latency than a moving stimulus like the one we used (Nijhawan, 1994, 2001; Ogmen, Patel, Bedell, & Camuz, 2004; Whitney & Murakami, 1998; Whitney, Murakami, & Cavanagh, 2000; however, see Eagleman & Sejnowski, 2000, 2007; Krekelberg & Lappe, 1999, for different explanations). This implies that the latency difference between the eye movement signal and the retinal signal is smaller for moving stimuli. However, the magnitude of the flash-lag phenomenon (about 40 to 80 ms; see Nijhawan, 1994, 2001; Whitney & Murakami, 1998; Whitney et al., 2000) is too small to explain the difference between our results and those from localization studies entirely.

Some authors have proposed that in localization eye position signals lead retinal signals because the estimate of eye position is based on an efference copy of the oculomotor command (Brenner et al., 2001; Mateeff et al., 1981). According to this account, the efferent eye signal is combined with an afferent retinal signal, without correcting for the latency difference. Our data show that the eye movement signal lags the retinal signal, suggesting that in motion perception the eye movement signal is not (exclusively) based on an efference copy. Rather, an afferent signal, such as proprioceptive feedback from the eye muscles and eye lids, might be used (Gauthier et al., 1990a, 1990b; Skavenski, 1972; Wang et al., 2007).
Linear vs. non-linear model

In previous studies, the standard linear model of motion perception during smooth pursuit eye movements (Equation 1) has been shown to describe motion perception quite well for a variety of tasks and stimuli (Freeman & Banks, 1998; Freeman et al., 2000; Souman et al., 2005a; Wertheim, 1987). It also captured the general trend of the observed amplitude matches in our experiments. However, the linear model did not fit our data at small retinal phases and low retinal amplitudes. In these instances, the non-linear model adapted from Freeman (2001) performed better. Although the non-linear model had to be fitted with an additional free parameter, the results of Experiment 2 indicate that this was not the only reason for the better goodness-of-fit of the non-linear model. The difference between the two models systematically increased for larger pursuit amplitudes. Analysis of model performance in terms of AIC showed that the increased goodness-of-fit was not just due to the additional free parameter in the non-linear model. This suggests that the visual system uses non-linear transducers for estimating both retinal velocity and eye velocity. Previous studies have arrived at the same conclusion (Freeman, 2001; Souman et al., 2006; Turano & Massof, 2001).

The best fitting power coefficients $e$ and $r$ were $\sim 1.5$ and 1.8, respectively. This suggests that the speed transducers as defined in Equations 6 and 7 are expansive. The values are similar to those reported by Freeman (2001) in a study that used constant pursuit and retinal speed. The nature of speed transducers underlying motion perception has received little attention in the literature. In principle, the shape of the speed transducers impacts on magnitude estimates of visual speed and on discrimination performance. The few studies that have employed magnitude estimation tend to report linear or compressive relationships (Kennedy, Hettinger, Harm, Ordy, & Dunlap, 1996; Kennedy, Yessenow, & Wendt, 1972). However, magnitude estimation is susceptible to a number of factors that are not easy to control (Poulton, 1979). Speed discrimination data, on the other hand, cannot be directly related to the transducer functions either. Discrimination thresholds are not only determined by the transducer function but also by the relationship between signal level and noise (Georgeson & Meese, 2006). It is therefore difficult to devise an alternative test for the shape of the non-linear transducers revealed by our experiments.

Independence of the signals

Several authors have suggested that the eye velocity signal might not only be based on an efference copy or proprioceptive feedback but also on retinal image characteristics (Brenner & Van den Berg, 1994; Crowell & Andersen, 2001; Goltz et al., 2003; Haarmeier, Bunjes, Lindner, Berret, & Thier, 2001; Haarmeier & Thier, 1996; Harris, 1994; Post & Leibowitz, 1985; Turano & Massof, 2001; Wertheim, 1994; Wertheim & Van Gelder, 1990). According to Wertheim (1994), the visual input to the eye movement signal will be largest when the visual stimulus covers a large part of the visual field, contains low spatial frequencies, moves slowly on the retina, and is presented for a sufficiently long time. In our experiments, observers were presented with a very large motion stimulus that covered the entire field of view horizontally and $30^\circ$ vertically. The dots of the motion stimulus were Gaussian blurred, sparsely distributed and so contained considerable energy at low spatial frequencies. Moreover, they were presented for a full 2 s and their retinal peak speed during pursuit was $13^\circ$/s at most (far less in most conditions). We therefore maximized the chances of finding an interaction between the retinal signal and the eye movement signal. However, our results show that the non-linear model with two independent signals describes our data very well. Parsimony suggests that a visual component in the eye movement signal is not needed to explain our data.

One reason why some authors have claimed retinal inputs into the eye movement signal is the finding that adapting simultaneously to retinal motion and pursuit affects head-centered motion perception (Crowell & Andersen, 2001; Haarmeier et al., 2001; Haarmeier & Thier, 1996). In a recent study, however, one of us has shown that this type of adaptation affects both retinal and eye-motion signals prior to the site of combination (Freeman, 2007). This suggests that the effect of simultaneous adaptation on head-centered motion could also be explained by the type of independence defined by the models described in the current paper.

In the present study, we also failed to find evidence for another type of interaction between the signals. Several studies have shown that compensation for the effects of eye movements on retinal motion may depend on the relative motion direction of the two (Brenner & Van den Berg, 1994; Crowell & Andersen, 2001; Tong, Aydin, & Bedell, 2007; Tong, Patel, & Bedell, 2005, 2006; Turano & Heidenreich, 1999). Compensation has been found to be better when retinal motion is in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006). In our experiments, each pursuit interval contained retinal motion in the same as well as in the opposite direction to pursuit, with the proportion of retinal motion in the same as well as in the opposite direction to pursuit (for an exception, see Souman et al., 2005a, 2006).
Acknowledgments

The authors wish to thank Katja Mayer and Manish Sreenivasa for help with data collection and Hans-Günther Nusseck for technical assistance. Marc Ernst contributed in helpful discussions. Author JLS was financially supported by NWO travel grant R 56-485 and by the EU sixth framework research project Cyberwalk (FP6-511092). TCAF received funding from the Wellcome trust.

Commercial relationships: none.
Corresponding author: Tom C. A. Freeman.
Email: freemant@cardiff.ac.uk.
Address: School of Psychology, Cardiff University, United Kingdom.

References


