

A Bayesian Model of Perceived Head-Centered Velocity during Smooth Pursuit Eye Movement

Tom C.A. Freeman,^{1,*} Rebecca A. Champion,¹ and Paul A. Warren²

¹School of Psychology, Cardiff University, Tower Building, Park Place, Cardiff CF10 3AT, UK

²School of Psychological Sciences, University of Manchester, Zochonis Building, Brunswick Street, Manchester M13 9PL, UK

Summary

During smooth pursuit eye movement, observers often misperceive velocity. Pursued stimuli appear slower (Aubert-Fleischl phenomenon [1, 2]), stationary objects appear to move (Filehne illusion [3]), the perceived direction of moving objects is distorted (trajectory misperception [4]), and self-motion veers away from its true path (e.g., the slalom illusion [5]). Each illusion demonstrates that eye speed is underestimated with respect to image speed, a finding that has been taken as evidence of early sensory signals that differ in accuracy [4, 6–11]. Here we present an alternative Bayesian account, based on the idea that perceptual estimates are increasingly influenced by prior expectations as signals become more uncertain [12–15]. We show that the speeds of pursued stimuli are more difficult to discriminate than fixated stimuli. Observers are therefore less certain about motion signals encoding the speed of pursued stimuli, a finding we use to quantify the Aubert-Fleischl phenomenon based on the assumption that the prior for motion is centered on zero [16–20]. In doing so, we reveal an important property currently overlooked by Bayesian models of motion perception. Two Bayes estimates are needed at a relatively early stage in processing, one for pursued targets and one for image motion.

Results

Figure 1A demonstrates a consistent finding across a number of studies that eye velocity is often underestimated with respect to image velocity. This finding could reflect differences in the accuracy of underlying motion signals, especially given that sensory information encoding eye speed is likely to be based on motor commands [21] in the situations described in the figure. However, differences in accuracy imply a visual system that has failed to remove systematic errors between early sensory signals, despite evidence that adaptation and calibration lie at the heart of most visual function [22]. The alternative view is that sensory evidence is unbiased (i.e., accurate) but can vary in uncertainty (i.e., precision). In a Bayesian framework, the uncertain sensory evidence is combined with prior expectations about particular properties of the world [12–14]. For motion perception, a plausible prior is that objects are largely at rest [16–20]. The prior is therefore centered on 0, in which case perceived speed decreases as uncertainty rises (Figure 1B). To investigate whether this idea

could account for the illusions described in Figure 1A, we measured thresholds for discriminating the speed of pursued and fixated stimuli. At the same time we also measured the size of the Aubert-Fleischl phenomenon in order to see whether the discrimination data could predict the underestimation of eye speed.

In the main experiment, observers judged which of two stimulus intervals (Figure 2A) appeared faster over a series of trials. Fixation intervals (F) consisted of a moving stimulus and a stationary fixation target. The stimulus moved behind a static circular window to ensure that approximately the same region of retina was stimulated during fixation and pursuit. Eye pursuit intervals (E) consisted of the target, window, and stimulus all moving together. The two types of interval were combined in three different ways. Discrimination trials contained F-F or E-E intervals. These two types of trial therefore allowed us to evaluate the precision of signals accompanying fixated and pursued stimuli, respectively. Perceived-speed trials contained E-F or F-E intervals. Trials of this third type allowed us to quantify the size of the Aubert-Fleischl phenomenon because they determined the relative difference in perceived speed between fixated and pursued stimuli.

Figure 2B shows the results of the experiment (symbols) and the model fitting (lines). Each row corresponds to one of the five observers who took part. The discrimination data in the left column show that thresholds were higher for pursued stimuli, meaning that the speed of pursued stimuli was harder to discriminate than the speed of fixated stimuli. Observers were therefore less certain about the sensory evidence defining pursued stimuli, and so their judgments were less precise. The poorer discrimination could have potentially been due to the absence of relative motion, because unlike the fixation intervals, the dot pattern, target, and window all moved together in the pursuit intervals. However, relative motion only influences thresholds at slow speeds [23, 24], a finding we confirmed in a control experiment reported in Figure S1 available online. We therefore conclude that the presence or absence of eye movement is paramount in driving the differences in precision.

According to a Bayesian explanation, greater thresholds during pursuit mean that pursued stimuli should appear slower than fixated stimuli because of the greater influence of the zero-motion prior (Figure 1B). Analysis of E-F and F-E trials supported this prediction (Figure 2B, right column). For all observers, fixated motion needed to be slowed by around 50% to achieve the perceived-speed match. Moreover, the accuracy of eye movements during fixation or pursuit could not explain the data (Figure 2C). The results therefore provide excellent qualitative agreement with the requirements of a Bayesian explanation.

Model

To determine the extent to which the discrimination data predicts the perceived-speed data, we implemented the Bayesian model shown in Figure 3 (see Supplemental Experimental Procedures for details). The model is designed to account for the type of perceptual errors represented by the four illusions identified in Figure 1A. It therefore recovers perceived

*Correspondence: freemant@cardiff.ac.uk

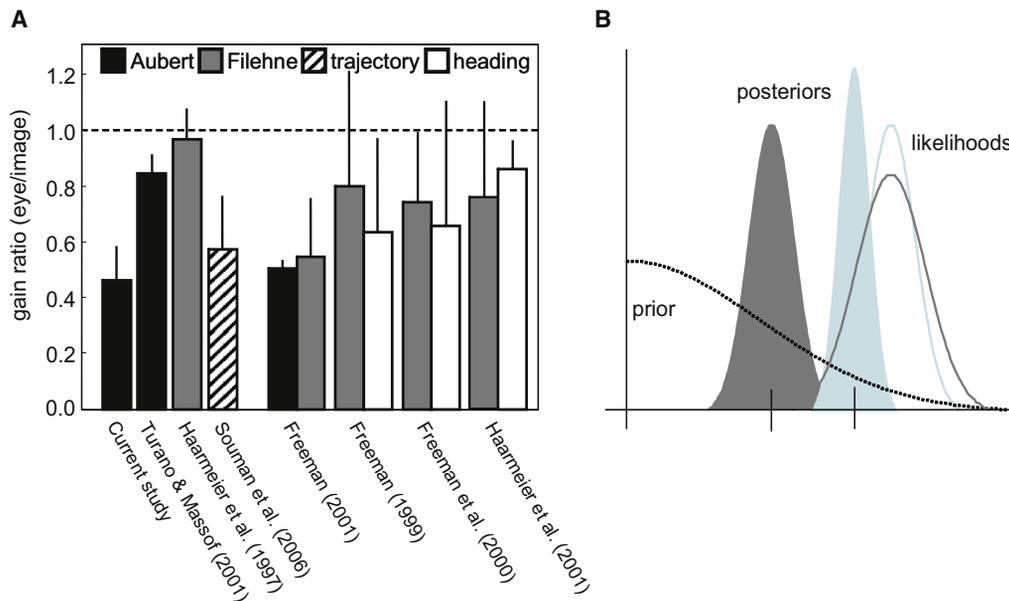


Figure 1. Pursuit-Based Velocity Illusions and Bayesian Inference

(A) Summary of recent studies of the Aubert-Fleischl phenomenon, Filehne illusion, trajectory misperception, and perceived heading. The psychophysical data have been converted into a “gain ratio,” which expresses the magnitude of the signals encoding eye velocity with respect to image motion [4, 6, 8, 41]. The four studies on the right directly compared pairs of illusions. Data were taken from figure 3 of [8]; figure 2 of [42]; pp. 69–70 of [4]; figures 4 and 5 of [6]; figure 3 of [41]; figure 7 of [5]; and figure 4 (and T. Haarmeier, personal communication) of [43]. Error bars correspond to 95% confidence intervals.

(B) Bayes law applied to motion perception. Perceived speed is determined by the location of the peak of the posterior probability distribution (the maximum a posteriori estimate; short vertical lines). The posterior is the product of a likelihood function (representing sensory evidence) and a prior (representing expectations). A plausible prior for motion perception is that objects are at rest, hence the prior is centered on 0. Noisier sensory signals are less precise and so yield wider likelihood functions (dark curve compared to light curve). The posterior in this case therefore shifts closer to the prior (solid dark curve compared to solid light curve).

motion with respect to the head (as opposed to motion with respect to the scene; see [25]). In keeping with the consistency between illusions, the model does not differentiate between how image motion and eye motion are arranged in time. It therefore does not care whether the two motion types are compared consecutively (the Aubert-Fleischl phenomenon) or simultaneously (the Filehne illusion, trajectory misperception, slalom illusion).

The model differs from traditional accounts of head-centered motion perception in two important ways. First, it treats motion measurements and motion estimates separately. In comparison, traditional accounts assume that motion estimates are based on sensory signals alone. Second, it emphasizes the role of relative image motion (between background object and target) and pursuit target motion. Traditional accounts emphasize absolute retinal motion and eye velocity. The eye movements were extremely accurate in our experiments, so relative motion was approximately equal to absolute image motion, and target motion was equal to eye velocity. But not all observers do so well—for instance, pursuit slows by around 10%–20% between 20 and 60+ years of age [26]. The model is therefore designed to account for situations in which pursuit is inaccurate and imprecise (see Discussion).

The model consists of a measurement stage (Figure 3, bottom left) and an estimation stage (Figure 3, top left). The measurement stage contains separate internal noise sources, one for relative motion (R) and one for pursuit target motion (T). The standard deviation of the internal noise was defined as $\sigma(v) = av^p + c$ (see [27, 28] and Supplemental Experimental Procedures). The estimation stage implements separate Bayes

estimators for R and T via the same principle described earlier in Figure 1B. The model assumes two priors, both centered on a speed of 0. The priors represent the observer’s expectations about target motion and relative motion. They are both based on the assumption that objects tend to be at rest. In the absence of reliable sensory information, the observer expects target objects to not move and hence for there to be no relative motion between one object (the target) and another.

The inputs to the estimation stage (the sensory measurements R_m and T_m) vary over intervals and trials as a result of internal noise. Hence, the Bayes estimates R' and T' vary as well, as does their sum H' , which yields the observer’s estimate of head-centered motion. For a two-interval task, two distributions of H' are produced, one for each interval (Figure 3, right). Straightforward application of signal detection theory links the model to data. It does so by combining the two intervals into a single “decision” variable. The probability of choosing one interval over the other can then be determined. For discrimination, the two intervals consist of image motion or eye motion (plus the small contribution of the noise term “c” when the input is 0). For the perceived-speed condition, the decision variable combines the two different types of motion. The combination is therefore independent of temporal order, reiterating the fact that how the different types of motion are arranged in time does not matter to the model.

Figure 2B shows that thresholds were approximately constant for faster speeds when expressed as a proportion of the standard. This is known as Weber’s law and is thought to reflect an early nonlinearity in the coding of speed, combined with fixed internal noise (i.e., noise independent of speed) [17, 29, 30]. However, as analogous work on contrast

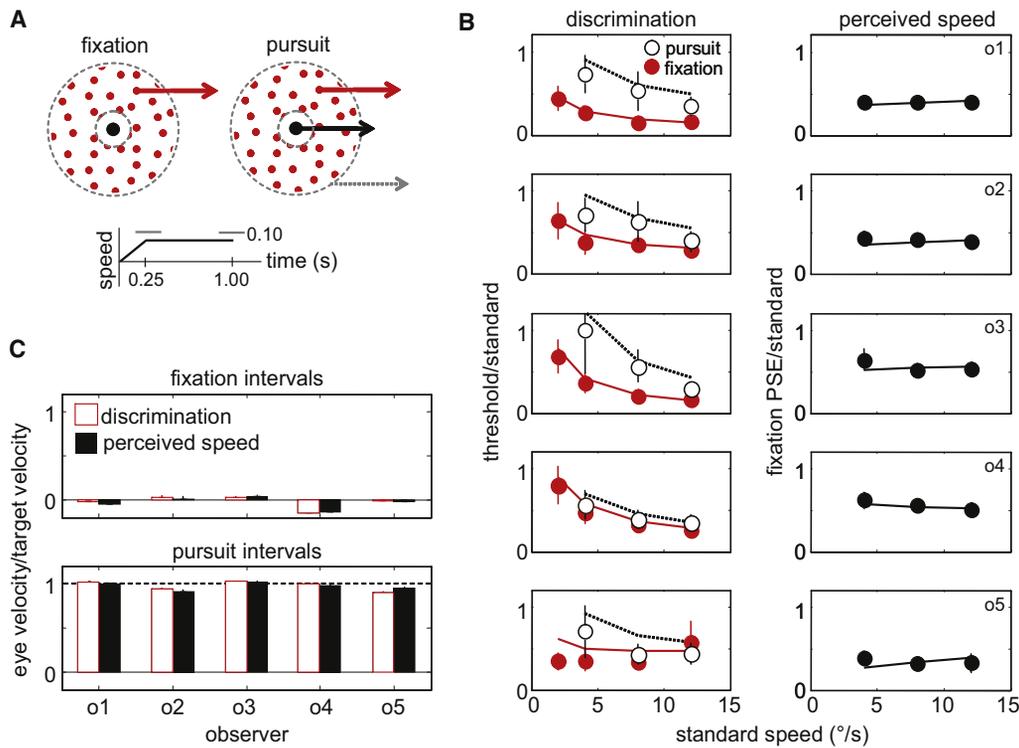


Figure 2. Methods and Main Results

(A) Random dot stimuli ($0.64 \text{ dots/degree}^2$, dot diameter = 0.1° , red dots on black background) were viewed in a completely darkened room on a standard cathode ray tube (100 Hz) from a distance of 70 cm. Fixation intervals contained a static fixation target (diameter = 0.2°) and static window (outer diameter = 10° , inner diameter = 2°). Pursuit intervals contained the stimulus, target, and window all moving together, with the start position randomly perturbed by 2° . Motion was ramped over the first 0.25 s and continued at a constant speed for a total duration of 1.0 s. Ramped and constant motions were randomly perturbed by 0.1 s. Discrimination trials and perceived-speed trials were randomly interleaved in the same session. Psychometric functions were obtained via a method of constant stimuli. Cumulative Gaussians were fit to data via maximum likelihood estimation: discrimination thresholds were defined as the standard deviation of the Gaussian, and perceived speed was defined as its mean (point of subjective equality [PSE] for matching fixated tests to pursued standards). (B) Results for five observers. Thresholds and perceived speeds are reported as a fraction of the standard speed. Pursued stimuli were less easy to discriminate than fixated stimuli (left column, open symbols versus closed symbols). Perceived speed slowed during pursuit (right column). Lines show thresholds and PSEs determined by a Bayesian model fit to the raw psychometric data (i.e., they were not fit to the thresholds and PSEs shown in the figure; see [17] for similar strategy). Error bars are 95% confidence intervals, obtained via a bootstrapping technique. (C) Eye movements for fixation intervals (top) and pursuit intervals (bottom). Open bars correspond to discrimination trials, and closed bars correspond to perceived-speed trials. Eye movements were measured with an Eyelink 1000 eye tracker, sampling at 1000 Hz. Trials containing saccades ($\sim 6\%$) were discarded from eye movement analysis.

discrimination has shown [31, 32], similar results can be obtained if an early nonlinearity is combined with variable noise, a point we confirmed in earlier implementations of our model. Our data also show that Weber's law breaks down at slow speeds, a finding well known for image-motion processing [33] and one that our results now extend to pursued stimuli. To account for this latter behavior, the early nonlinearity can be augmented in a number of ways (see [17] for an example). However, from a Bayesian perspective, it turns out that early nonlinearities may not be necessary. Using a combination of variable internal noise and a zero-motion prior, we were able to reproduce the thresholds we found very well, as can be seen by comparing the model and data in Figure 2B. Indeed, invoking early nonlinearities presents problems for Bayesian accounts of head-centered motion perception. If different nonlinearities for R and T were used, the Bayes estimation stage is unnecessary—changes in velocity estimates would be captured by differences in signal accuracy that result from the separate nonlinearities at the measurement stage (see [6, 8] for demonstration). The alternative is to enforce identical nonlinearities. But this is unlikely, given that the

measurements of image motion and eye velocity are based on different types of motion signal.

The second way in which the model differs from previous Bayesian accounts of motion perception is the use of two estimators based on separate likelihoods for R and T. An alternative is to sum signals at the measurement stage and so yield a single head-centered likelihood. However, this could never produce a Filehne illusion, assuming that the signals are unbiased. When smooth pursuit is made over a stationary background, observers typically report that the background appears to move against the eye movement. In this situation, pursuit produces equal and opposite motion in the image (Figure 4A). Hence, the sum of unbiased measurements of eye motion (T_m) and image motion (R_m) must be centered on 0. Importantly, the sum defines the location of the putative head-centered likelihood. Given that the prior is also centered on 0, the posterior distribution defining the observer's estimate must be, too, so no Filehne illusion can result. Similar reasoning shows why trajectory misperception cannot occur (Figure 4B). Of course, one way to fix this alternative Bayesian account is to introduce biases into the initial sensory

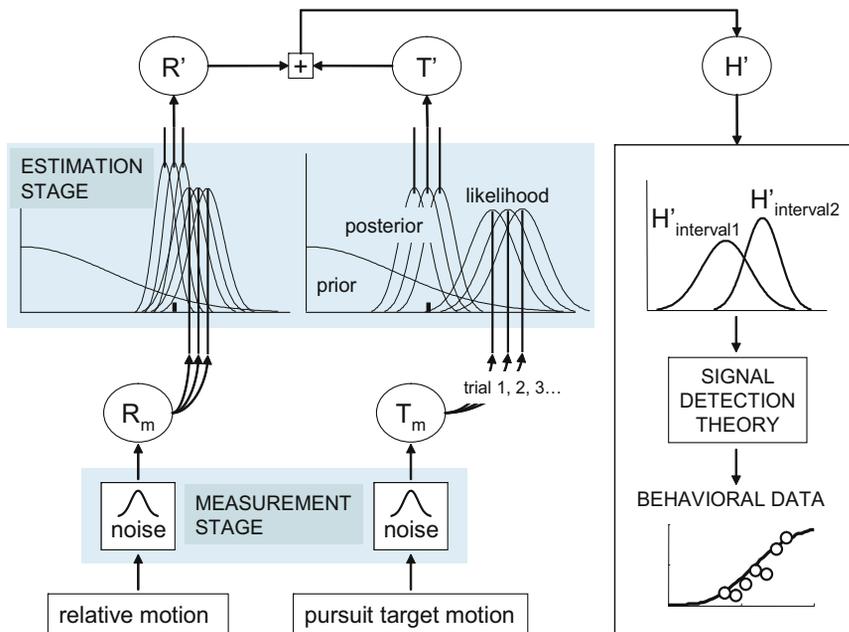


Figure 3. Schematic of the Bayesian Model

Measurement stage: separate sources of internal noise are added to unbiased measurements (sensory signals) of relative motion (R_m) or pursuit-target motion (T_m). The noise varies as a function of speed and sets the spread of likelihoods. Estimation stage: R' and T' are the estimated speeds obtained by multiplying the prior with the likelihood. Their sum yields an estimate of head-centered motion (right). Signal detection theory is used to map the output of the model (e.g., for a two-interval task) onto behavioral performance (i.e., psychometric function).

Allowing the priors to have different standard deviations would not substantially improve the fit.

Discussion

The new Bayesian model of head-centered motion perception presented here is able to explain a range of pursuit-based velocity illusions. The model is based on the idea that sensory signals encoding the speed of eye motion and image motion differ in precision, not accuracy. The model raises a number of issues that need to be considered when applying Bayes theory to motion perception. First, the model emphasizes the role of pursuit-target motion and relative motion, in part because this formulation guarantees that the priors are properties of the world. Second, the Bayes estimates must be made before information about eye motion and image motion is combined, otherwise no illusions can result. Third, separate Bayes estimates for relative motion and pursuit-target motion are required to explain these illusions. Finally, the combination of unbiased signals, variable internal noise, and a prior centered on zero is sufficient to predict the discrimination performance we found.

measurements. But as discussed above, this eliminates the need for a Bayes estimation stage. We conclude that in order to account for pursuit-based velocity illusions, two Bayes estimates are needed, one for R and one for T .

For modeling purposes, we assumed that the two priors had the same standard deviation, a justifiable assumption given that both depend on the idea that objects tend to be at rest. The estimation stage therefore consisted of a single parameter (the standard deviation of the prior), whereas the measurement stage consisted of six (one set each of three internal noise parameters for R and T). The seven-parameter model was fit simultaneously to all ten psychometric functions of each observer via a maximum-likelihood technique. The lines in Figure 2B show the model's discrimination thresholds and perceived speeds. The model predicts the data extremely well.

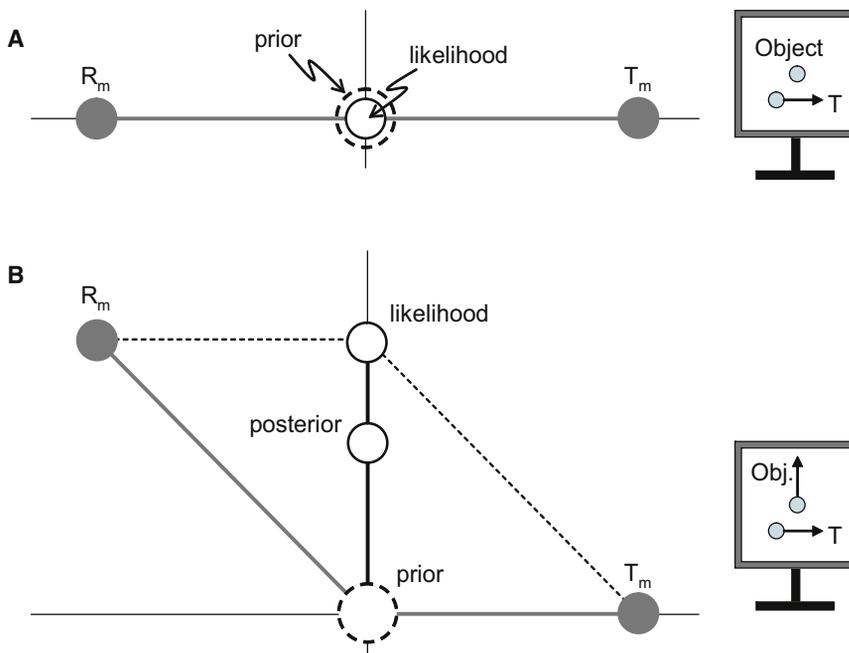


Figure 4. Unbiased Sensory Measurements Cannot Lead to Pursuit-Based Velocity Illusions

(A) No Filehne illusion could result if sensory measurements were biased because pursuit over a stationary background produces relative motion that is equal and opposite to the eye movement. The unbiased sensory signals R_m and T_m must therefore be equal and opposite, too. Their sum defines the location of the putative head-centered likelihood and is centered on 0. Given that the prior is centered on 0, too, the posterior (not shown) must be situated there as well. Stationary objects would always appear stationary.

(B) Using similar reasoning, no misperception of trajectory could result because the putative head-centered likelihood is located in the true direction. In this example, pursuit to a target T is made over a vertically moving object. This produces relative motion that is oblique. The sum of unbiased measurements of R and T defines the location of putative head-centered likelihood and lies in the correct direction (upward, black vertical line). If the prior is centered at 0, the posterior can only ever be located in the correct direction (vertical for this example), at a speed dependent on the spread of the likelihood.

The Aubert-Fleischl phenomenon is known to decline as the contrast of fixated stimuli is lowered [34]. Because contrast does not affect the perceived speed of pursued stimuli, the change to the Aubert-Fleischl phenomenon must be driven solely by the lower estimates of image motion brought about by decreasing contrast. This effect is easy to capture in the model because manipulating contrast is one way of influencing the precision of image motion measurements—indeed, the effect of contrast provides the bulk of the evidence supporting Bayesian models of motion perception [17, 19, 20] (although see [35] for evidence that the relationship between perceived speed and contrast may be more complex than often described; see also [36]). At present it is unknown whether contrast influences the Filehne illusion, trajectory misperception, or slalom illusion in a similar way. It may be that variations in stimulus dimensions like contrast help explain why the degree of underestimation of eye speed varies across the studies shown in Figure 1A.

Unlike traditional accounts of head-centered motion perception, our new model emphasizes the role of relative motion and pursuit-target motion. There are good empirical reasons for doing so. Recent evidence suggests that relative motion overrides the use of absolute retinal motion during pursuit, even when trial-by-trial feedback is given on the latter [37]. Relative motion therefore appears paramount for these types of motion judgment. The use of pursuit-target motion may stem from the fact that pursuit eye movements are not always accurate and also vary over time [26, 38]. Hence, estimating target motion solely on the basis of extraretinal eye velocity information [21] would be subject to the same inaccuracies and temporal variability. These are easily offset by adding localized image motion information related to the movement of the pursuit in the image (retinal slip), an idea supported by the recent finding that discriminating the motion of pursued targets is best predicted by the combination of retinal slip and eye velocity [39] (see also evidence of cells in the medial superior temporal area that respond to eye velocity and retinal slip [40]). Indeed, the combination of slip and eye velocity may explain why the sensory measurement of pursuit-target motion is less precise than corresponding measures of relative motion. By summing eye velocity information with retinal slip, the internal noise related to pursuit-target motion originates from two disparate sources. The internal noise related to relative motion, however, derives from a single source, namely the retinal image. On many occasions, this single noise source can be reduced further by integrating over larger areas.

Supplemental Information

Supplemental Information includes Supplemental Experimental Procedures, three figures, and one table and can be found with this article online at doi:10.1016/j.cub.2010.02.059.

Acknowledgments

The authors would like to thank M. Banks, M. Georgeson, S. Rushton, P. Sumner, K. Langley, J. Solomon, and two anonymous reviewers for helpful discussion, and also T. Haarmeier for providing data for Figure 1. The research was supported by a grant from the Wellcome Trust.

Received: October 9, 2009
Revised: February 18, 2010
Accepted: February 18, 2010
Published online: April 15, 2010

References

1. Fleischl, E.V. (1882). Physiologisch-optische Notizen, 2. Mitteilung. Sitzung Wiener Bereich der Akademie der Wissenschaften 3, 7–25.
2. Aubert, H. (1886). Die Bewegungsempfindung. Pflügers Arch. 39, 347–370.
3. Filehne, W. (1922). Über das optische Wahrnehmen von Bewegungen. Zeitschrift für Sinnephysiologie 53, 134–145.
4. Souman, J.L., Hooge, I.T.C., and Wertheim, A.H. (2006). Frame of reference transformations in motion perception during smooth pursuit eye movements. J. Comput. Neurosci. 20, 61–76.
5. Freeman, T.C.A., Banks, M.S., and Crowell, J.A. (2000). Extraretinal and retinal amplitude and phase errors during Filehne illusion and path perception. Percept. Psychophys. 62, 900–909.
6. Freeman, T.C.A. (2001). Transducer models of head-centred motion perception. Vision Res. 41, 2741–2755.
7. Wertheim, A.H. (1994). Motion perception during self-motion: The direct versus inferential controversy revisited. Behav. Brain Sci. 17, 293–311.
8. Turano, K.A., and Massof, R.W. (2001). Nonlinear contribution of eye velocity to motion perception. Vision Res. 41, 385–395.
9. Beintema, J.A., and van den Berg, A.V. (2001). Pursuit affects precision of perceived heading for small viewing apertures. Vision Res. 41, 2375–2391.
10. Perrone, J.A., and Krauzlis, R.J. (2008). Vector subtraction using visual and extraretinal motion signals: A new look at efference copy and corollary discharge theories. J. Vis. 8, 1–14.
11. Souman, J.L., and Freeman, T.C.A. (2008). Motion perception during sinusoidal smooth pursuit eye movements: Signal latencies and nonlinearities. J. Vis. 8, 1–14.
12. Ernst, M.O., and Bühlhoff, H.H. (2004). Merging the senses into a robust percept. Trends Cogn. Sci. 8, 162–169.
13. MacNeilage, P.R., Banks, M.S., Berger, D.R., and Bühlhoff, H.H. (2007). A Bayesian model of the disambiguation of gravito-inertial force by visual cues. Exp. Brain Res. 179, 263–290.
14. Ma, W.J., Beck, J.M., Latham, P.E., and Pouget, A. (2006). Bayesian inference with probabilistic population codes. Nat. Neurosci. 9, 1432–1438.
15. Clark, J.J., and Yuille, A.L. (1990). Data Fusion for Sensory Information Processing Systems (Boston: Kluwer Academic).
16. Welchman, A.E., Lam, J.M., and Bühlhoff, H.H. (2008). Bayesian motion estimation accounts for a surprising bias in 3D vision. Proc. Natl. Acad. Sci. USA 105, 12087–12092.
17. Stocker, A.A., and Simoncelli, E.P. (2006). Noise characteristics and prior expectations in human visual speed perception. Nat. Neurosci. 9, 578–585.
18. Weiss, Y., Simoncelli, E.P., and Adelson, E.H. (2002). Motion illusions as optimal percepts. Nat. Neurosci. 5, 598–604.
19. Hürlimann, F., Kiper, D.C., and Carandini, M. (2002). Testing the Bayesian model of perceived speed. Vision Res. 42, 2253–2257.
20. Ascher, D., and Grzywacz, N.M. (2000). A Bayesian model for the measurement of visual velocity. Vision Res. 40, 3427–3434.
21. von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. Br. J. Anim. Behav. 2, 89–94.
22. Burge, J., Ernst, M.O., and Banks, M.S. (2008). The statistical determinants of adaptation rate in human reaching. J. Vis. 8, 1–19.
23. Snowden, R.J. (1992). Sensitivity to relative and absolute motion. Perception 21, 563–568.
24. Baker, C.L., Jr., and Braddick, O.J. (1982). Does segregation of differently moving areas depend on relative or absolute displacement? Vision Res. 22, 851–856.
25. Warren, P.A., and Rushton, S.K. (2009). Optic flow processing for the assessment of object movement during ego movement. Curr. Biol. 19, 1555–1560.
26. Kolarik, A.J., Margrain, T.H., and Freeman, T.C.A. (2010). Precision and accuracy of ocular following: Influence of age and type of eye movement. Exp. Brain Res. 201, 271–282.
27. Britten, K.H., Shadlen, M.N., Newsome, W.T., and Movshon, J.A. (1993). Responses of neurons in macaque MT to stochastic motion signals. Vis. Neurosci. 10, 1157–1169.
28. Dakin, S.C., Mareschal, I., and Bex, P.J. (2005). Local and global limitations on direction integration assessed using equivalent noise analysis. Vision Res. 45, 3027–3049.
29. Nover, H., Anderson, C.H., and DeAngelis, G.C. (2005). A logarithmic, scale-invariant representation of speed in macaque middle temporal

- area accounts for speed discrimination performance. *J. Neurosci.* 25, 10049–10060.
30. Zanker, J.M. (1995). Does motion perception follow Weber's law? *Perception* 24, 363–372.
 31. Kontsevich, L.L., Chen, C.-C., and Tyler, C.W. (2002). Separating the effects of response nonlinearity and internal noise psychophysically. *Vision Res.* 42, 1771–1784.
 32. Georgeson, M.A., and Meese, T.S. (2006). Fixed or variable noise in contrast discrimination? The jury's still out. *Vision Res.* 46, 4294–4303.
 33. De Bruyn, B., and Orban, G.A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Res.* 28, 1323–1335.
 34. Sumnall, J.H., Freeman, T.C.A., and Snowden, R.J. (2003). Optokinetic potential and the perception of head-centred speed. *Vision Res.* 43, 1709–1718.
 35. Thompson, P., Brooks, K.R., and Hammett, S.T. (2006). Speed can go up as well as down at low contrast: Implications for models of motion perception. *Vision Res.* 46, 782–786.
 36. Hammett, S.T., Champion, R.A., Thompson, P.G., and Morland, A.B. (2007). Perceptual distortions of speed at low luminance: Evidence inconsistent with a Bayesian account of speed encoding. *Vision Res.* 47, 564–568.
 37. Freeman, T.C.A., Champion, R.A., Sumnall, J.H., and Snowden, R.J. (2009). Do we have direct access to retinal image motion during smooth pursuit eye movements? *J. Vis.* 9, 1–11.
 38. Rasche, C., and Gegenfurtner, K.R. (2009). Precision of speed discrimination and smooth pursuit eye movements. *Vision Res.* 49, 514–523.
 39. Welchman, A.E., Harris, J.M., and Brenner, E. (2009). Extra-retinal signals support the estimation of 3D motion. *Vision Res.* 49, 782–789.
 40. Ilg, U.J., Schumann, S., and Thier, P. (2004). Posterior parietal cortex neurons encode target motion in world-centered coordinates. *Neuron* 43, 145–151.
 41. Freeman, T.C.A. (1999). Path perception and Filehne illusion compared: Model and data. *Vision Res.* 39, 2659–2667.
 42. Haarmeier, T., Thier, P., Reppow, M., and Petersen, D. (1997). False perception of motion in a patient who cannot compensate for eye movements. *Nature* 389, 849–852.
 43. Haarmeier, T., Bunjes, F., Lindner, A., Berret, E., and Thier, P. (2001). Optimizing visual motion perception during eye movements. *Neuron* 32, 527–535.

Supplemental Information

A Bayesian Model

of Perceived Head-Centered Velocity

during Smooth Pursuit Eye Movement

Tom C.A. Freeman, Rebecca A. Champion, and Paul A. Warren

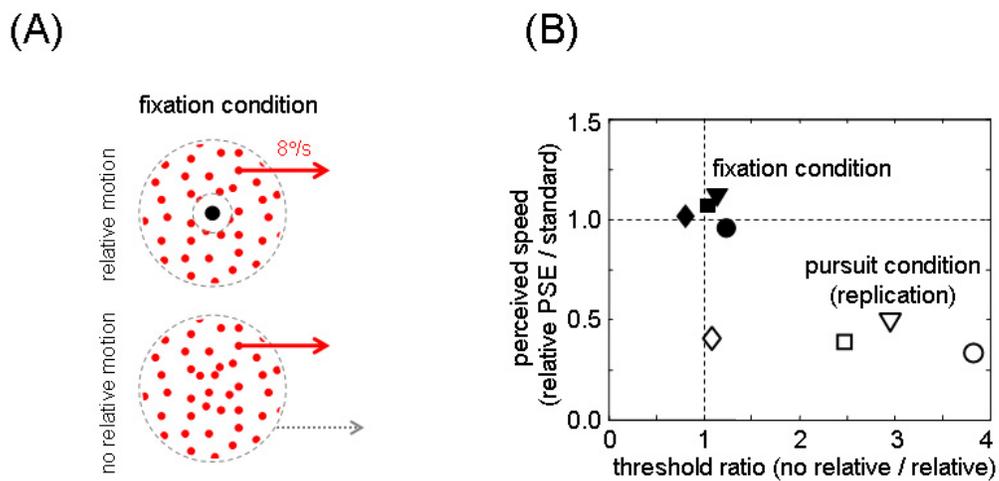


Figure S1.

Results of a control experiment showing that the presence or absence of relative motion is not the reason why pursued stimuli in the main experiment appeared slower than fixated stimuli. Different symbol shapes correspond to different observers. See text for details.

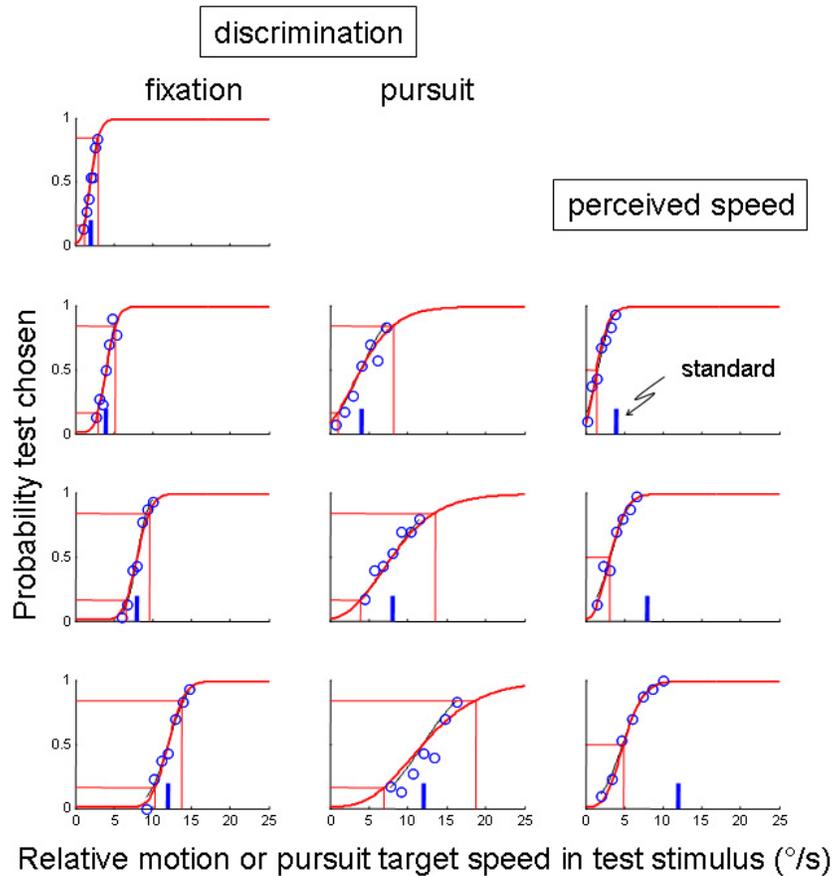


Figure S2.

Representative observer's raw psychometric data with associated model fit (red lines). The model has 7 parameters and these were fit simultaneously across all 10 graphs (see text for details). First two columns correspond to discrimination trials and third column perceived-speed trials. Black curves show standard psychometric functions fit to individual panels. In most cases these are hidden by the model curves.

Table S1. Best-Fitting Parameter Values for Each of 5 Observers

Observer	Internal noise						Prior
	Relative motion (R)			Pursuit target motion (T)			σ
	a	b	c	a	b	c	
o1	0.14635	0.31107	0.13808	0.04384	0.57662	0.40757	0.01186
o2	0.10223	0.34886	0.05690	0.13364	0.30253	0.08618	0.00432
o3	0.36452	0.08794	0.20993	0.31443	0.06663	0.48642	0.06553
o4	0.10415	0.53404	0.17715	0.06329	0.56682	0.40804	0.02205
o5	0.12062	0.22516	0.06325	0.20534	0.19806	0.00460	0.01451

Supplemental Results

Control Experiment

To investigate the influence of relative motion in the results of our main experiment, we compared discrimination and perceived speed for two types of fixation interval shown in Figure S1A. One contained relative motion as before and one contained no relative motion. The results in Figure S1B (closed symbols) showed that the ability to discriminate the speed in the two fixation intervals was the same: the threshold ratio clusters around 1. In fact, this turns out to be a particularly strong test of Bayesian accounts of motion perception, because the absence of a discrimination difference should lead to equal perceived speeds. Again, the closed symbols show that the perceived speed ratio for these two intervals was also close to 1.

To achieve fixation without relative motion, we reduced interval duration to 0.25s. This prevented any significant pursuit eye movement, while simultaneously allowing the presentation of moving stimuli without any stationary references (such as a stationary fixation point and static window). The comparison fixation intervals, those that contained relative motion, had a similarly short duration. Unfortunately, displaying stimuli this briefly meant that short fixation intervals could not be compared to short pursuit intervals – for the same reason that short durations prevent eye movements during fixation, they also prevent pursuit. The closest comparison possible was to repeat the equivalent long-duration conditions of the experiment in the main paper. In doing so, we replicated our original finding (open symbols).

Supplemental Experimental Procedures

Derivation of Model

In the main paper, two velocities (relative motion R and pursuit target motion T) serve as input to the Bayesian model. Here we formalise the model by considering first how a generic velocity (\mathbf{v}) is estimated by the Bayesian observer. We then show how this generic estimate is used to estimate head-centred motion, that is, the sum of separate Bayes estimates of R and T. Finally we show how the collection of estimates maps on to behavioural data (ie. a psychometric function).

Measurement Stage

Velocity \mathbf{v} is measured by an unbiased sensory signal with a mean of \mathbf{v} . The signal is corrupted by Gaussian internal noise with a mean of 0 and standard deviation given by:

$$\sigma_{\text{int}}(\mathbf{v}) = a\mathbf{v}^b + c \quad (1)$$

The nonlinear portion of this equation is similar to that used previously to describe the response variance of cortical cells. Britten et al (1993) investigated the response of cells in motion area MT and found their variance proportional to the mean (i.e. Poisson noise). Hence standard deviation varies as the square-root (i.e. $b=0.5$ – see also Dakin et al, 2005). In our model fits we found values for b in reasonable agreement (mean = 0.32: see Table S1). Equation (1) also defines a parameter ‘ c ’ that represents the internal noise when $\mathbf{v}=0$.

Estimation Stage

Internal noise makes the sensory signal vary from trial to trial. Bayes law allows the observer to draw sensible inferences about the probable state of the world (e.g. the degree of relative motion or pursuit target motion). Uncertain sensory evidence is combined with prior expectations to yield a posterior probability equal to a likelihood function multiplied by the prior (divided by a normalisation factor). The likelihood defines (in part) the variability of the sensory evidence. The location of the peak of the posterior is typically taken as the perceptual estimate (the MAP estimate). Using the symbol ‘ \sim ’ to denote ‘is distributed as’, then assuming that the prior $P \sim N(\mu_p, \sigma_p^2)$ and the likelihood $L \sim N(\mu_L(v), \sigma_L^2)$, it can be shown that the posterior S is also normally distributed with mean $\mu_s(\mathbf{v})$ and variance $\sigma_s^2(\mathbf{v})$ (see Ma et al, 2006):

$$\mu_s(\mathbf{v}) = \frac{\mu_L(\mathbf{v})\sigma_p^2 + \mu_p\sigma_L^2(\mathbf{v})}{\sigma_p^2 + \sigma_L^2(\mathbf{v})} \quad (2)$$

$$\sigma_s^2(\mathbf{v}) = \frac{\sigma_p^2\sigma_L^2(\mathbf{v})}{\sigma_p^2 + \sigma_L^2(\mathbf{v})} \quad (3)$$

Note that for an unbiased estimator, $\mu_L(v) = v$. We assume that stochastic internal noise is the only reason that perceived speed varies across intervals and trials. Each time the observer estimates velocity \mathbf{v} , the mean of the likelihood in Equation (2) is therefore shifted by the internal noise of the measurement (Stocker & Simoncelli, 2006). Consequently we assume that the mean of the likelihood is actually a random variable $\mu'_L(\mathbf{v})$ corrupted by Gaussian noise $n \sim N(0, \sigma_{\text{int}}^2(\mathbf{v}))$. In other words:

$$\mu'_L(\mathbf{v}) \sim N(\mu_L(\mathbf{v}), \sigma_{\text{int}}^2(\mathbf{v})) \quad (4)$$

Note that in general, the standard deviation of the likelihood $\sigma_L(\mathbf{v})$ incorporates both internal noise and external uncertainty. The latter term corresponds to factors such as ambiguities over the projection of three dimensions to two. In the model fits we assumed there was no external uncertainty (ie. $\sigma_{\text{ext}} = 0$) but, for completeness, we derive the model with external uncertainty included:

$$\sigma_L^2(\mathbf{v}) = \sigma_{\text{int}}^2(\mathbf{v}) + \sigma_{\text{ext}}^2 \quad (5)$$

Speed discrimination is determined by the way the mean of the posterior μ_s changes across intervals and trials. More formally, due to its dependence on $\mu'_L(\mathbf{v})$, the mean of the posterior is the random variable μ'_s defined by:

$$\mu'_s(\mathbf{v}) = \frac{\mu'_L(\mathbf{v})\sigma_p^2 + \mu_p\sigma_L^2(\mathbf{v})}{\sigma_p^2 + \sigma_L^2(\mathbf{v})} \quad (6)$$

Consequently, the individual MAP estimates form their own probability distribution. Crucially, we can derive the mean and variance of this probability distribution as follows. We first note that Stocker & Simoncelli (2006) reverse engineered the motion prior from psychophysical data and provided good support for the assumption that it has a mean of zero. Setting μ_p to zero in Equation (6) we obtain:

$$\mu'_s(\mathbf{v}) = \frac{\mu'_L(\mathbf{v})\sigma_p^2}{\sigma_p^2 + \sigma_L^2(\mathbf{v})} \quad (7)$$

Next we make use of standard results from mathematical statistics. For a random variable X and scalar r , the expected value and variance of the random variable rX are given by:

$$\begin{aligned} E[rX] &= rE[X] \\ Var[rX] &= r^2Var[X] \end{aligned}$$

Since Equation (7) tells us that $\mu'_s(\mathbf{v})$ is a linear function of the random variable $\mu'_L(\mathbf{v})$ of which we know the mean and variance (see Equation (4)), we use the above results and substitute Equation (5) to obtain:

$$E[\mu'_s(\mathbf{v})] = \frac{\mu_L(\mathbf{v})\sigma_p^2}{\sigma_p^2 + \sigma_{int}^2(\mathbf{v}) + \sigma_{ext}^2} \quad (8)$$

$$Var[\mu'_s(\mathbf{v})] = \left(\frac{\sigma_p^2}{\sigma_p^2 + \sigma_{int}^2(\mathbf{v}) + \sigma_{ext}^2} \right)^2 \times \sigma_{int}^2(\mathbf{v}) \quad (9)$$

Estimating Head-Centred Motion

Head-centred motion is the sum of eye velocity and retinal motion or, equivalently, pursuit target motion (T) and relative motion (R). Arguments in favour of the latter are presented in the main paper, but note this only becomes an important issue when eye movements and/or fixation are inaccurate. The Bayesian observer modelled here therefore estimates head-centred motion by adding separate estimates of T and R with means and standard deviations defined by Equations (8) and (9). Recall that for an unbiased estimator, $\mu_L(v) = v$. In this case, the posterior for head-centred motion (H) has a mean and variance:

$$E[\mu'_H(\mathbf{T}, \mathbf{R})] = \frac{\mathbf{T}\sigma_p^2}{\sigma_p^2 + \sigma_{int}^2(\mathbf{T}) + \sigma_{ext}^2} + \frac{\mathbf{R}\sigma_p^2}{\sigma_p^2 + \sigma_{int}^2(\mathbf{R}) + \sigma_{ext}^2} \quad (10)$$

$$\begin{aligned} Var[\mu'_H(T, R)] = & \left(\frac{\sigma_P^2}{\sigma_P^2 + \sigma_{int}^2(T) + \sigma_{ext}^2} \right)^2 \times \sigma_{int}^2(T) \\ & + \left(\frac{\sigma_P^2}{\sigma_P^2 + \sigma_{int}^2(R) + \sigma_{ext}^2} \right)^2 \times \sigma_{int}^2(R) \end{aligned} \quad (11)$$

where $\sigma_{int}^2(T)$ and $\sigma_{int}^2(R)$ are the internal noises associated with T and R, respectively. They are defined by two different sets of values for a, b, c (see Equation (1)).

Linking Model to Psychometric Function

Using standard signal detection theory, performance in a two interval task is governed by:

$$d' = \sqrt{2} \frac{E[\mu'_H(T_1, R_1)] - E[\mu'_H(T_2, R_2)]}{[Var[\mu'_H(T_1, R_1)] + Var[\mu'_H(T_2, R_2)]]^{1/2}} \quad (12)$$

where (T_1, R_1) and (T_2, R_2) are the velocities in the two intervals. Note that in our experiments, $R_i = 0$ in pursuit intervals and $T_i = 0$ in fixation intervals. Hence the expectation $E[\mu'_H(T, R)]$ includes one likelihood only. However, this is not the case for the associated variances because Equation (1) defines a constant 'c' corresponding to the noise at $R_i = 0$ or $T_i = 0$. The variance of each interval therefore combines internal noise from both R and T estimates.

For a range of velocities \mathbf{v}_1 (the 'test') about \mathbf{v}_2 (the 'standard'), the probability of choosing one interval over the other is determined by:

$$P = \Phi(d' / \sqrt{2}) \quad (13)$$

where Φ denotes the cumulative distribution function of the standard normal distribution. In the model fitting we included a lapse rate parameter g as suggested by Wichman & Hill (2001). The lapse rate captures those instances when the observer overlooks one or other interval and is forced to guess:

$$P_g = (1 - g)P + 0.5g \quad (14)$$

The parameter g was fixed for all observers at 0.02.

Model Fitting

The model was implemented in MatLab (Mathworks Inc.) using their simplex search utility `fminsearch`. Figure S2 shows the raw data (symbols) for one observer, together with model's psychometric functions (thick red lines) as well as the fitted psychometric functions (thin black lines) that summarise observer thresholds and perceived speeds in main paper (symbols in Figure 3B of main text). The thin black lines are mostly hidden by the fitted model. There are four

standard speeds for fixation discrimination (left column), three standards for pursuit discrimination (middle) and three standards for perceived-speed (right). Best-fitting values for the model's parameters are given in Table S1. The model consists of seven parameters: three internal noise parameters for R and three for T as defined by Equation (1), together with one standard deviation for the prior. Maximum likelihood estimation was used to fit the seven parameters simultaneously across all 10 graphs. This was done separately for each observer. Model thresholds and PSEs were obtained using spline interpolation. Because the psychometric functions for the model are not symmetric about the PSE, thresholds were defined as $0.5[v(0.84) - v(0.16)]$, where $v(P)$ is the test speed interpolated at the probability P. The asymmetry arises in part because of the use variable internal noise defined in Equation (1).