

Rapid Reweighting of Sensory Inputs and Predictions in Visual Perception

William Turner

wft@stanford.edu

Stanford University, Stanford, CA 94305, USA; and Queensland University of Technology, Kelvin Grove 4059, QLD, Australia

Oh-Sang Kwon

oskwon@unist.ac.kr

Minwoo J.B. Kim

minu333@unist.ac.kr

Ulsan National Institute of Science and Technology, Ulsan, South Korea

Hinze Hogendoorn

hinze.hogendoorn@qut.edu

Queensland University of Technology, Kelvin Grove 4059, QLD, Australia

A striking perceptual phenomenon has recently been described wherein people report seeing abrupt jumps in the location of a smoothly moving object (“position resets”). Here, we show that this phenomenon can be understood within the framework of recursive Bayesian estimation as arising from transient gain changes, temporarily prioritizing sensory input over predictive beliefs. From this perspective, position resets reveal a capacity for rapid adaptive precision weighting in human visual perception and offer a possible test bed within which to study the timing and flexibility of sensory gain control.

1 Introduction ---

Accurate visual localization of objects is critical for adaptive behavior, both evolutionarily (e.g., targeting prey while hunting) and in modern life (e.g., navigating through traffic). However, there are instances where localization goes awry. One of the most striking examples is the double-drift illusion, which occurs when a moving object contains internal motion in a direction orthogonal to its global trajectory (see Figure 1A; Lisi & Cavanagh, 2015). Under such conditions, if the object is viewed peripherally, extreme mislocalizations can occur.

Recently it has been shown that the double-drift illusion can be reset, such that people report seeing the object abruptly jump back toward its true position (Nakayama & Holcombe, 2020). These position resets can be

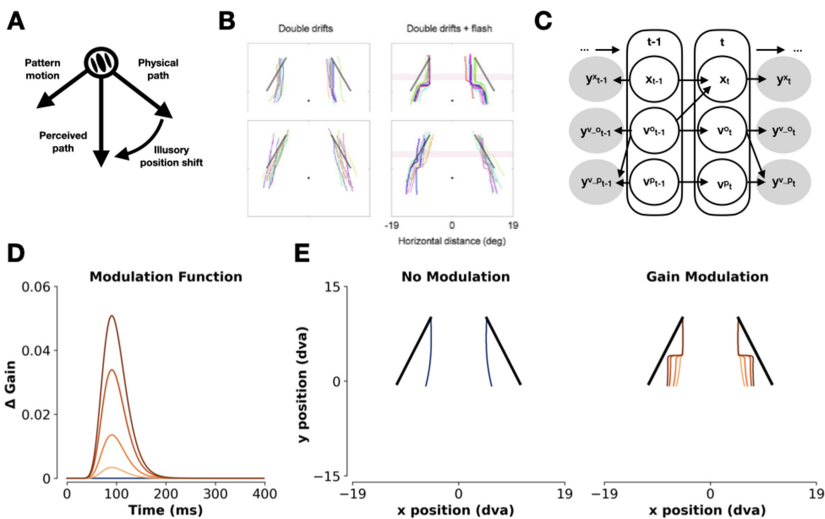


Figure 1: Accounting for position resets in the double-drift illusion. (A) When viewed peripherally (under high sensory uncertainty), a diagonally moving Gabor stimulus can appear to travel along a vertical path if its internal pattern (phase) moves in an orthogonal direction. (B) With a brief flash of white patches near the Gabor, participants report seeing abrupt position resets, such that the stimulus appears to move along a zigzag-shaped path. Panels show hand-drawn stimulus trajectories from two participants viewing double-drift stimuli with (right) and without (left) a transient flash (Nakayama & Holcombe, 2020). With flashed patches (presented in the region highlighted by the red strip), an abrupt kink can be seen midway through the trajectories, indicating that participants saw the stimulus jump back toward its true position. (C) A generative model of the Bayesian observer, adapted from Kwon et al. (2015). White nodes indicate hidden variables, and gray nodes indicate observable variables, which are noisy measurements of the connected hidden variables. Arrows indicate causal links. The superscripts o and p refer to object and pattern velocities, respectively. (D) A Bayesian object-tracking model (Kwon et al., 2015) gives a principled account of position resets. Transient gain modulations, modeled via a log-normal function with varying amplitude (see panel E), drive the rapid reweighting of inputs and predictions within the model, resulting in a zigzag-shaped trajectory estimate. Black diagonal lines indicate the true stimulus trajectories, blue lines show estimated trajectories under no gain modulation, and orange lines show model-estimated trajectories under gain modulations of varying amplitude.

triggered by visual transients near the object or may occur spontaneously ('t Hart et al., 2022). When asked to draw the (linearly moving) object's trajectory, participants draw a zigzag-shaped path (see Figure 1B).

Currently, the cause of position resets is unknown. One high-level account suggests that shifts of attention may somehow reset the perceived position of the object back to its veridical position (Nakayama & Holcombe, 2020). However, a computationally rigorous account of this phenomenon is lacking, and the question of why attention might drive resets has not been addressed.

Here, we consider this phenomenon from the perspective of recursive Bayesian estimation, where perception is viewed as an unfolding inference process in which sensory inputs are combined with internally generated predictions to derive more precise estimates of world states (e.g., an object's position and speed). From this perspective, we show that position resets can arise from transient gain changes that temporarily prioritize sensory input over predictive beliefs.

2 Model and Results

We simulated an object-tracking model previously used to account for a variant of the double-drift illusion (the curveball illusion; for full details, see the linked code and Kwon et al., 2015). At its core is a generative model of motion dynamics (i.e., the physical laws of motion) used to derive predictions about current world states. From this, predictions are made about the current position and velocity of an object given its past state (see the prediction model, equations 2.1–2.3 in section 2.1). These are integrated with noisy sensory inputs (see the measurement model, equations 2.4–2.6 in section 2.1) to derive more precise state estimates (see Figure 1C for a graphical depiction of the generative model).

2.1. Modeling Motion and Position Perception via Recursive Bayesian Estimation. We can specify the model in discrete form with the following time update equations:

Prediction Model

$$x_t = x_{t-1} + \Delta t \cdot v_{t-1}^{\text{object}}, \quad (2.1)$$

$$v_t^{\text{object}} = \alpha v_{t-1}^{\text{object}} + \sigma_{v_{t-1}^{\text{object}}} \Omega_t^{vo}, \quad (2.2)$$

$$v_t^{\text{pattern}} = \beta v_{t-1}^{\text{pattern}} + \sigma_{v_{t-1}^{\text{pattern}}} \Omega_t^{vp}. \quad (2.3)$$

These equations specify predictions about the object's current state at time t , given its previous state at time $t - 1$. Equation 2.1 updates the object's position (x) by integrating its global velocity (v^{object}) over time. Equations 2.2 and 2.3 are stationary gaussian processes that characterize global object (v^{object}) and internal pattern velocities (v^{pattern}), with α and β determining velocity correlations over time (see Kwon et al., 2015, for a full description), $\sigma_{v_{t-1}^{\text{object}}}$ and $\sigma_{v_{t-1}^{\text{pattern}}}$ specifying standard deviations (SDs) of changes in

object and pattern velocities, and Ω_t representing unit variance gaussian noise.

Measurement Model

$$y_t^x = x_t + \eta^x \Omega_t^{yx}, \quad (2.4)$$

$$y_t^{object} = v_t^{object} + \eta^{v-object} \Omega_t^{vo}, \quad (2.5)$$

$$y_t^{pattern} = v_t^{object} + v_t^{pattern} + \eta^{v-pattern} \Omega_t^{vp}. \quad (2.6)$$

These equations specify sensory measurements of retinal stimulus position (y^x), retinal object velocity (y^{object}), and the retinal velocity of the pattern within the object ($y^{pattern}$). These are corrupted by gaussian noise with SDs given by η^x , $\eta^{v-object}$, and $\eta^{v-pattern}$, respectively.

To build an intuitive understanding of how the model works, we will walk through one time step. First, predictions about the current world state are generated. That is, predictions are made as to the current position and velocity of the object, as well as the velocity of its internal pattern, via the previous estimates of these variables and the internal model of motion dynamics (equations 2.1–2.3). These predictions can be written in matrix form $\hat{x}$, representing a set of position and velocity estimates along the vertical and horizontal dimensions, with corresponding uncertainties (P). Simultaneously, new sensory measurements are received (z), with corresponding uncertainties (R), according to the measurement model (equations 2.4–2.6). Crucially, the predicted and measured states will likely differ, and thus a final state estimate can be determined by combining these two sources of information. The goal here is to find the optimal state estimate given the noisy predictions and measurements. Under the assumption that both follow gaussian distributions, the optimal combination of these sources is afforded by the Kalman filter (Kalman, 1960). This provides the optimal solution to our inference problem, balancing the relative precision of each information source. The Kalman gain K (equation 2.7) determines the weighting that minimizes the posterior variance as follows:

$$K = \frac{P}{P + R} + M. \quad (2.7)$$

Here, P represents the uncertainty of the predictions, R represents the uncertainty of sensory inputs, and M is a gain modulation term that we have added to the base model.¹ Examining equation 2.7, we can see that gain is determined by the relative uncertainty of predictions and inputs (the ratio of P and R). With more precise inputs or less certain predictions, gain is high. Conversely, with more precise predictions or less certain inputs, gain

¹For simplicity we omit the observation matrix (H) from all equations.

is low. The term M represents an additive gain modulation parameter that we have introduced to the standard Kalman filter framework. Its additive nature is somewhat arbitrary, as we currently don't have the empirical basis to differentiate between distinct functional forms of modulation (see section 3 for further consideration). By adjusting the gain modulation matrix (M) in a time-varying manner, we can simulate transient changes in gain while remaining agnostic to their underlying cause (see below and section 3). That is, we can effectively modulate the weighting between sensory inputs and predictions, with increased gain shifting the weight toward incoming sensory information.

Having specified the gain, a weighted combination of the predictions ($\hat{x}$) and inputs (z) can be taken. That is, we can derive a precision-weighted combination of our predictions and measurements, generating an optimal state estimate:

$$\hat{x}' = \hat{x} + K(z - \hat{x}). \quad (2.8)$$

Thus, the final state estimate ($\hat{x}'$) is determined by the current predictions ($\hat{x}$) plus the precision-weighted prediction error—the difference between sensory inputs and predictions, multiplied by the Kalman gain. Consequently, with more precise sensory input, the gain will be high, and thus the input will play a greater role in determining the final state estimate. Conversely, with more uncertain input, or precise predictive beliefs, sensory input will be down-weighted, and the final estimate will be largely determined by the current predictions.

To summarize, this model casts perception as an unfolding Bayesian inference process. At a given time point, predictions are generated and sensory inputs are received. A final perceptual estimate is then derived via a weighted combination of these two noisy information sources according to their relative precision. This base model (Kwon et al., 2015) has been used to give a unifying account of a range of different motion and position phenomena, neatly explaining why illusory effects often arise during peripheral viewing, when sensory uncertainty is high. In the following section, we show that this model can also explain the puzzling phenomenon of position resets via the addition of transient gain modulations. These temporarily shift the weight given to inputs and predictions in the unfolding inference process, leading to abrupt resets in the object's trajectory estimate.

2.2. Simulation Results: Transient Gain Modulations Can Drive Position Resets. In our simulations, we examined the effect of transient gain modulations on model-derived object trajectory estimates. That is, we examined how brief changes in the weight given to sensory inputs and predictions affected the perceived trajectory of the stimulus.

To simulate gain modulations, from 900 ms after stimulus onset, we varied the diagonal of the additive gain modulation matrix (term M in

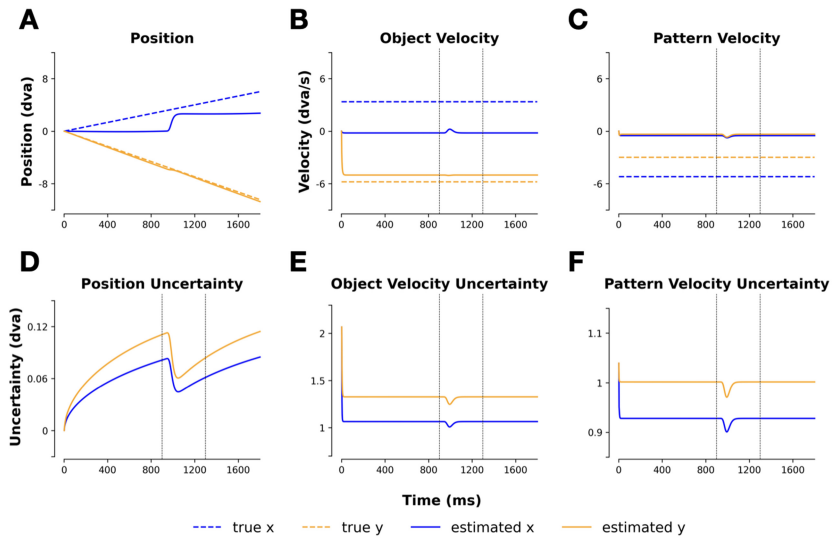


Figure 2: Examining the accuracy and uncertainty of state estimates. Panels A to C show the true and model estimated states of object position, velocity, and pattern velocity, respectively. Estimates along the horizontal and vertical dimensions are shown in blue and orange, respectively. Panels D to F show the standard deviation (square root of the diagonal elements of matrix P in equation 2.7) of these state predictions over time. Note that these simulations are run within the highest gain modulation condition.

equation 2.7) according to a log-normal function with varying amplitude (see Figure 1D). This allowed us to simulate time-varying, transient changes in gain. Simulating modulations of varying strength, we found that transient increases in gain reliably induced position resets. Stronger modulations triggered more abrupt resets, with model trajectory estimates qualitatively mirroring those from Nakayama and Holcombe (2020). These resets occur because the gain modulations drive a rapid reweighting of predictions and inputs, temporarily up-weighting input and down-weighting predictions and shifting the object's estimated position back toward its true location.

To get a sense of the internal dynamics of the model during reweighting, we plot the true and model estimated states (position, velocity, and pattern velocity), as well as the uncertainty (SD) of these estimates over time (see Figure 2), for the strongest gain modulation condition. Following gain modulation, the model estimates are transiently more accurate and more certain. This suggests that participants should also briefly perceive the object, and its internal pattern, to be moving more quickly in their veridical direction,

a prediction that to our knowledge has yet to be tested and thus warrants further investigation.

MATLAB and Python scripts for recreating these simulations are available at <https://github.com/bootstrapbill/position-reset-model>.

3 Discussion

We have given a computational account of position resets in visual localization (Nakayama & Holcombe, 2020). Through the lens of recursive Bayesian estimation, we have shown that resets can arise from gain modulations that temporarily prioritize sensory inputs over predictive beliefs. From this perspective, position resets reveal a capacity for rapid adaptive precision weighting (Yon & Frith, 2021) in human visual perception.

Importantly, the current model is agnostic to the cause of gain changes, with modulations occurring via a general additive gain parameter (equation 2.7). At least two hypotheses warrant investigation. First, attention shifts (either bottom-up or top-down) may sharpen incoming sensory information (Nakayama & Holcombe, 2020). With increased precision, these inputs will be upweighted relative to predictions, triggering a reset. Second, abrupt visual transients may trigger a reduction in the precision of internally generated predictions (e.g., “Something has changed, so my predictions may no longer hold”; see Press et al., 2020). This too would lead to the prioritization of inputs over predictions, and thus a reset. For proof-of-principle simulations of these accounts, see the linked code.

These accounts may not be mutually exclusive and could be further tested by examining whether resets are preceded by a sharpening of neural object position representations (Turner et al., 2023; Turner et al., 2025; Yon et al., 2018). As a general test of this overarching perspective, studies could examine whether manipulations that increase sensory uncertainty (visual noise) can prevent resets. If this is possible, this would support the core idea that the relative precision of inputs and predictions ultimately determines their perceptual influence. More targeted manipulations of stimulus parameters may also be examined. For example, the model predicts that with less precise stimulus position information (e.g., a wider gaussian window), there should be a greater double-drift effect and more abrupt resets (i.e., sharper kinks in the trajectory estimate). Conversely, the same effects are predicted with a more precise pattern motion (e.g., a higher contrast grating/pattern; see linked code for auxiliary simulations). With further support for this model, studies may then investigate the neurochemical basis of gain modulations—for example, by testing whether specific pharmacological interventions can inhibit or disrupt resets (Moran et al., 2013), or by studying pupil-linked arousal dynamics during spontaneous resets (‘t Hart et al., 2022).

It is important to note that this model effectively assumes that internal (model) time is equivalent to external (physical) time. However, this is at

odds with the well-established fact of neural signalling delays within the visual system (see Turner et al., 2024 for a recent review). This is not a pressing problem here, since the dynamics of the simulated stimulus are constant, making the simulated behavior invariant across sensory lags. Nevertheless, future studies should consider introducing sensory delays and converting the existing model into a Bayesian smoother (for related work see Grush, 2005; Rao et al., 2001; Perrinet et al., 2014; Khoei et al., 2017). This would allow for theoretical examination of possible interactions between changes in stimulus dynamics and sensory gain that can then be tested empirically.

In general, we encourage further empirical study of position resets. The goal of this note has been relatively modest: to show that the framework of recursive Bayesian estimation can qualitatively capture this puzzling perceptual phenomenon. We have restricted our focus to qualitative assessments because participants' hand-drawn trajectories will likely have been influenced by a range of nonperceptual factors (e.g., motoric costs, memory effects). As we have noted, the additive nature of the assumed gain modulation (M) is somewhat arbitrary. Given the lack of finely grained empirical measurements of position resets, we do not yet have the empirical basis for arbitrating between competing functional forms of gain modulation. A limitation of the additive formulation is that bounds must be placed on the value of this term to ensure model stability. Nevertheless, within these bounds, this assumption is a reasonable first approximation. Future work may refine this and explicitly test between differing functional forms of gain modulation, once the prerequisite finely grained measurements of position resets have been collected. Crucially, since a static positional reference has been shown to markedly diminish the double-drift effect (Saleki et al., 2021), future studies should avoid relative judgment tasks and may instead adopt alternative nonreferential measurement approaches (Liu et al., 2021). Moreover, given that eye movements have been shown to temporarily reset the double-drift effect (Lisi & Cavanagh, 2015; Nakayama & Holcombe, 2020), with illusory influence slowly reemerging as a function of saccade latency (Ueda et al., 2018; Massendari et al., 2018), future studies should control eye movement behavior. These studies could also examine whether the temporal dynamics of illusion reemergence can constrain the temporal dynamics of gain modulations within the current model.

If resets reflect rapid precision weighting, then individual variability in this phenomenon should be examined. For example, studies could test whether certain clinical populations are resistant to experiencing resets, as this may be indicative of reduced perceptual flexibility—that is, a stubborn perceptual experience. Relatedly, individual differences in the double-drift illusion have been observed to correlate with individual differences in the twinkle-goes and flash-grab illusions, suggesting they may share underlying mechanisms (Cottier et al., 2023). Future studies may therefore attempt to account for these illusions within the existing framework. As a speculative example, in the twinkle-goes illusion, an increased reliance

on predictions in the face of highly uncertain input (due to dynamic noise) may drive the lingering percept of the stimulus.

In sum, we have shown that position resets in the double-drift effect (Nakayama & Holcombe, 2020) can arise from transient gain changes in a recursive Bayesian estimation framework. This suggests a capacity for rapid dynamic precision weighting in human visual perception. Future studies may further examine this phenomenon to better understand the temporal dynamics of sensory gain control.

Acknowledgments

This work was supported by a QUT ECRIS Grant awarded to W.T. and Australian Research Council grant FT200100246 awarded to H.H.

Author Contributions

All of us contributed to conception and design. W.T. developed and ran the simulations. O.K. and M.K. reviewed the code. W.T. drafted the article, and all of us reviewed the manuscript.

Competing Interests

We declare no competing interests.

References

- Cottier, T. V., Turner, W., Holcombe, A. O., & Hogendoorn, H. (2023). Exploring the extent to which shared mechanisms contribute to motion-position illusions. *Journal of Vision*, 23(10), 8. 10.1167/jov.23.10.8
- Grush, R. (2005). Internal models and the construction of time: Generalizing from state estimation to trajectory estimation to address temporal features of perception, including temporal illusions. *Journal of Neural Engineering*, 2(3), S209. 10.1088/1741-2560/2/3/S05
- Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME, Journal of Basic Engineering*, 82, 34–45. 10.1115/1.3662552
- Khoei, M. A., Masson, G. S., & Perrinet, L. U. (2017). The flash-lag effect as a motion-based predictive shift. *PLOS Computational Biology*, 13(1), e1005068. 10.1371/journal.pcbi.1005068
- Kwon, O. S., Tadin, D., & Knill, D. C. (2015). Unifying account of visual motion and position perception. *Proceedings of the National Academy of Sciences*, 112(26), 8142–8147. 10.1073/pnas.1500361112
- Lisi, M., & Cavanagh, P. (2015). Dissociation between the perceptual and saccadic localization of moving objects. *Current Biology*, 25(19), 2535–2540. 10.1016/j.cub.2015.08.021
- Liu, S., Tse, P. U., & Cavanagh, P. (2021). *The perceived position of a moving object is reset by temporal, not spatial limits*. bioRxiv. 10.1101/2021.12.14.472615

- Massendari, D., Lisi, M., Collins, T., & Cavanagh, P. (2018). Memory-guided saccades show effect of a perceptual illusion whereas visually guided saccades do not. *Journal of Neurophysiology*, 119(1), 62–72. 10.1152/jn.00229.2017
- Moran, R. J., Campo, P., Symmonds, M., Stephan, K. E., Dolan, R. J., & Friston, K. J. (2013). Free energy, precision and learning: The role of cholinergic neuromodulation. *Journal of Neuroscience*, 33(19), 8227–8236. 10.1523/JNEUROSCI.4255-12.2013
- Nakayama, R., & Holcombe, A. O. (2020). Attention updates the perceived position of moving objects. *Journal of Vision*, 20(4), 21. 10.1167/jov.20.4.21
- Perrinet, L. U., Adams, R. A., & Friston, K. J. (2014). Active inference, eye movements and oculomotor delays. *Biological Cybernetics*, 108(6), 777–801. 10.1007/s00422-014-0620-8
- Press, C., Kok, P., & Yon, D. (2020). The perceptual prediction paradox. *Trends in Cognitive Sciences*, 24(1), 13–24. 10.1016/j.tics.2019.11.003
- Rao, R. P., Eagleman, D. M., & Sejnowski, T. J. (2001). Optimal smoothing in visual motion perception. *Neural Computation*, 13(6), 1243–1253. 10.1162/08997660152002843
- Saleki, S., Cavanagh, P., & Peter, U. T. (2021). A position anchor sinks the double-drift illusion. *Journal of Vision*, 21(6), 3. 10.1167/jov.21.6.3
- ‘t Hart, B. Marius, Henriques, D. Y., & Cavanagh, P. (2022). Measuring the double-drift illusion and its resets with hand trajectories. *Journal of Vision*, 22(2), 16. 10.1167/jov.22.2.16
- Turner, W., Blom, T., & Hogendoorn, H. (2023). Visual information is predictively encoded in occipital alpha/low-beta oscillations. *Journal of Neuroscience*, 43(30), 55375545. 10.1523/JNEUROSCI.0135-23.2023
- Turner, W., Sexton, C., & Hogendoorn, H. (2024). Neural mechanisms of visual motion extrapolation. *Neuroscience and Biobehavioral Reviews*, 156, 105484. 10.1016/j.neubiorev.2023.105484
- Turner, W., Sexton, C., Johnson, P. A., Wilson, E., & Hogendoorn, H. (2025). Predictable motion is progressively extrapolated across temporally distinct processing stages in the human visual cortex. *PLOS Biology*, 23(5), e3003189.
- Ueda, H., Abekawa, N., & Gomi, H. (2018). The faster you decide, the more accurate localization is possible: Position representation of “curveball illusion” in perception and eye movements. *PLOS One*, 13(8), e0201610. 10.1371/journal.pone.0201610
- Yon, D., & Frith, C. D. (2021). Precision and the Bayesian brain. *Current Biology*, 31(17), R1026R1032.
- Yon, D., Gilbert, S. J., de Lange, F. P., & Press, C. (2018). Action sharpens sensory representations of expected outcomes. *Nature Communications*, 9(1), 4288. 10.1038/s41467-018-06752-7