Sensory and Motor Systems

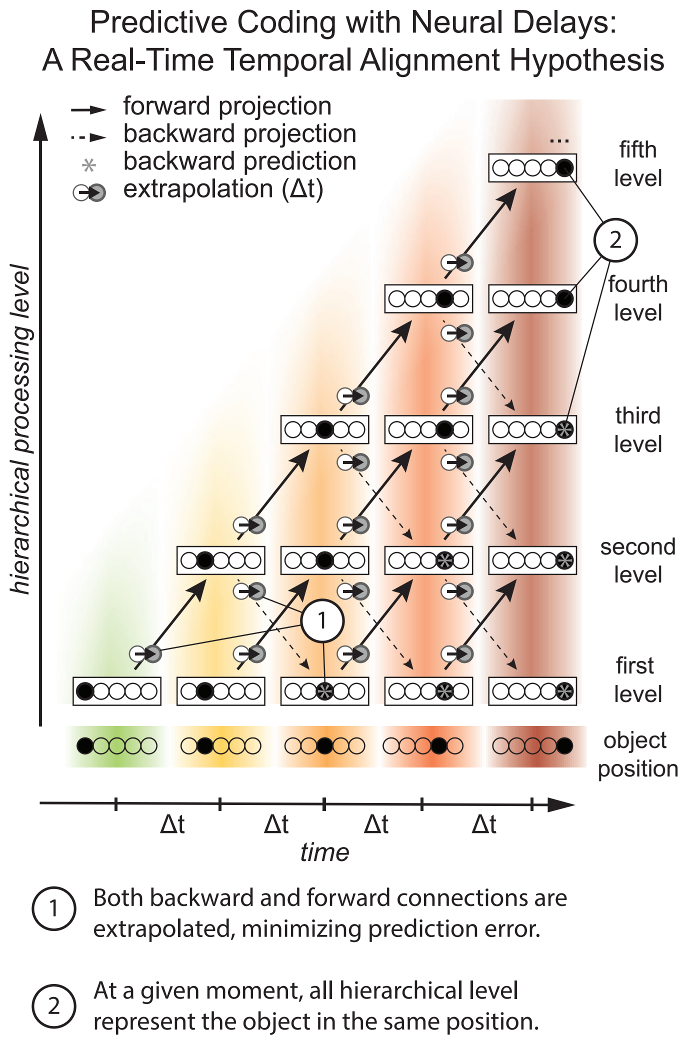
Predictive Coding with Neural Transmission Delays: A Real-Time Temporal Alignment Hypothesis

Hinze Hogendoorn,[1,2] and Anthony N. Burkitt[3]

https://doi.org/10.1523/ENEURO.0412-18.2019

[1]Melbourne School of Psychological Sciences, University of Melbourne, Melbourne, Victoria 3010, Australia, [2]Helmholtz Institute, Department of Experimental Psychology, Utrecht University, 3512 JE, Utrecht, The Netherlands, and [3]NeuroEngineering Laboratory, Department of Biomedical Engineering, University of Melbourne, Melbourne, Victoria 3010, Australia

Visual Abstract



Significance Statement

Despite the enormous scientific interest in predictive coding as a model of cortical processing, most models of predictive coding do not take into account that neural processing itself takes time. We show that when the framework is extended to allow for neural delays, a model emerges that provides a natural, parsimonious explanation for a wide range of experimental findings. It is consistent with neurophysiological data from animals, predicts a range of well known visual illusions in humans, and provides a principled solution to the temporal binding problem. Altogether, it explains how predictive coding mechanisms cause different brain areas to align in time despite their different delays, and in so doing it explains how cortical hierarchies function in real time.

Hierarchical predictive coding is an influential model of cortical organization, in which sequential hierarchical levels are connected by backward connections carrying predictions, as well as forward connections carrying prediction errors. To date, however, predictive coding models have largely neglected to take into account that neural transmission itself takes time. For a time-varying stimulus, such as a moving object, this means that backward predictions become misaligned with new sensory input. We present an extended model implementing both forward and backward extrapolation mechanisms that realigns backward predictions to minimize prediction error. This realignment has the consequence that neural representations across all hierarchical levels become aligned in real time. Using visual motion as an example, we show that the model is neurally plausible, that it is consistent with evidence of extrapolation mechanisms throughout the visual hierarchy, that it predicts several known motion–position illusions in human observers, and that it provides a solution to the temporal binding problem.

Key words: alignment; extrapolation; neural delays; prediction; predictive coding; temporal

Introduction

Predictive coding is a model of neural organization that originates from a long history of proposals that the brain infers, or predicts, about the state of the world on the basis of sensory input (von Helmholtz, 1867; Gregory, 1980; Dayan et al., 1995). It has been particularly influential in the domain of visual perception (Srinivasan et al., 1982; Mumford, 1992; Rao and Ballard, 1999; Spratling, 2008, 2012), but has also been extensively applied in audition (for review, see Garrido et al., 2009; Bendixen et al., 2012), the somatosensory system (van Ede et al., 2011), motor control (Blakemore et al., 1998; Adams et al., 2013), and decision science (Schultz, 1998; Summerfield and de Lange, 2014), where it accounts for a range of subtle response properties and accords with physiology and neuroanatomy. Although it has been criticized for being insufficiently articulated (Kogo and Trengove, 2015), it has been developed further into a general theory of cortical organization (Bastos et al., 2012; Spratling, 2017) and even been advocated as a realization of the free-energy principle, a principle that might apply to all self-organizing biological systems (Friston, 2005, 2010, 2018; Friston and Kiebel, 2009).

An essential principle of predictive coding is a functional organization in which higher organizational units “predict” the activation of lower units. Those lower units then compare their afferent input to this backward prediction, and feed forward the difference: a prediction error (Rao and Ballard, 1999; where the term “prediction” is used in the strictly hierarchical sense, rather than the everyday temporal sense of predicting the future). This interaction of backward predictions and forward prediction errors characterizes each subsequent level of the processing hierarchy (Fig. 1*a*).

In this review, we argue that neural transmission delays cause the classical model of hierarchical predictive coding (Rao and Ballard, 1999) to break down when input to the hierarchy changes on a timescale comparable with that of the neural processing. Using visual motion as an example, we present two models that extend the classical model with extrapolation mechanisms to minimize prediction error for time-varying sensory input. One of these models, in which extrapolation mechanisms operate on both forward and back mechanisms, has as a consequence that all levels in the hierarchy become aligned in real time. We argue that this model not only minimizes prediction error, but also parsimoniously explains a range of spatiotemporal phenomena, including motion-induced position shifts such as the flash-lag effect and related illusions, and provides a natural solution to the question of

Received October 25, 2018; accepted March 20, 2019; First published March 28, 2019.

The authors declare no competing financial interests.

H.H. and A.N.B. designed research; H.H. and A.N.B. performed research; H.H. and A.N.B. wrote the paper.

H.H. was supported by the Australian Government through the Australian Research Council's Discovery Projects funding scheme (project DP180102268).

Correspondence should be addressed to Hinze Hogendoorn at Hhogendoorn@unimelb.edu.au.

https://doi.org/10.1523/ENEURO.0412-18.2019

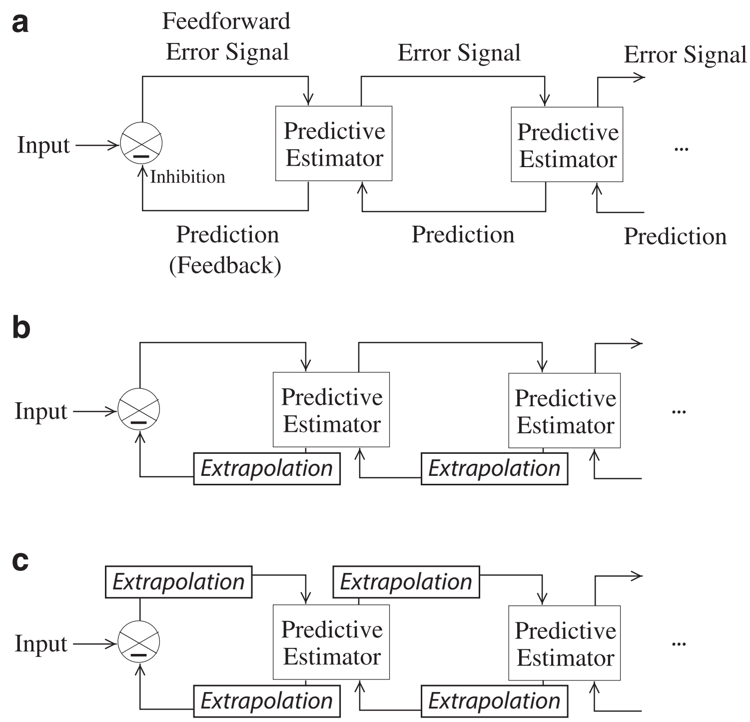


Figure 1. The classical hierarchical predictive coding model and two possible extensions. ***a***, The classical predictive model (model A). This model consists of hierarchically organized loops of forward and backward connections. Backward signals carry predictions, and forward signals carry prediction errors (Rao and Ballard, 1999). ***b***, The predictive model with extrapolated feedback (model B). To handle time-varying stimuli such as motion, the classical model can be expanded to include an extrapolation mechanism on the backward step. This would be one way to minimize prediction error for time-varying stimuli. ***c***, The predictive model with real-time alignment (model C). In this model, extrapolation mechanisms work on both forward and backward steps. Like model B, this would minimize prediction error, but has the additional consequence that it aligns the content of neural representations across the hierarchy. Diagram labels in ***b*** and ***c*** are as in ***a***, but are omitted for clarity.

how asynchronous processing of visual features nevertheless leads to a synchronous conscious experience (the temporal binding problem).

Minimizing prediction error

The core principle behind the pattern of functional connectivity in hierarchical predictive coding is the minimization of total prediction error. This is considered to be the driving principle on multiple biological timescales (Friston, 2018).

At long timescales, the minimization of prediction error drives evolution and phylogenesis. Neural signaling is metabolically expensive, and there is therefore evolutionary pressure for an organism to evolve a system of neural representation that allows for complex patterns of information to be represented with minimal neural firing. A sparse, higher-order representation that inhibits the costly, redundant activity of lower levels would provide a metabolic, and therefore evolutionary, advantage. Theoretically, the imperative to minimize complexity and metabolic cost is an inherent part of free-energy (i.e., prediction error) minimization. This follows because the free energy provides a proxy for model evidence, and model evidence is accuracy minus complexity (Friston, 2010). This means that minimizing free energy is effectively the same as self-evidencing (Hohwy, 2016), and both entail a minimization of complexity or computational costs, and their thermodynamic concomitants.

At the level of individual organisms, at timescales relevant to decision-making and behavior, the minimization of prediction error drives learning (den Ouden et al., 2012; Friston, 2018). Predictions are made on the basis of an internal model of the world, with the brain essentially using sensory input to predict the underlying causes of that input. The better this internal model fits the world, the better the prediction that can be made, and the lower the prediction error. Minimizing prediction error therefore drives a neural circuit to improve its representation of the world: in other words, learning.

Finally, at subsecond timescales relevant to sensory processing, the minimization of prediction error drives the generation of stable perceptual representations. A given pattern of sensory input feeds in to forward and backward mechanisms that iteratively projects to each other until a dynamic equilibrium is reached between higher-order predictions and (local) deviations from those predictions. Because this equilibrium is the most efficient representation of the incoming sensory input, the principle of prediction error minimization works to maintain this representation as long as the stimulus is present. This results in a perceptual representation that remains stable over time. Interestingly, because there may be local minima in the function determining total prediction error, a given stimulus might have multiple stable interpretations, as is, for example, the case for ambiguous stimuli such as the famous Necker cube (Necker, 1832).

Hierarchical versus temporal prediction

In discussing predictive coding, there is an important distinction to be made regarding the sense in which predictive coding models predict.

Descriptively, predictive coding models are typically considered (either implicitly or explicitly) to reflect some kind of expectation about the future. For example, in perception, preactivation of the neural representation of an expected sensory event ahead of the actual occurrence of that event reflects the nervous system predicting a future event (Garrido et al., 2009; Kok et al., 2017; Hogendoorn and Burkitt, 2018). In a decision-making context, a prediction might likewise be a belief about the future consequences of a particular choice (Sterzer et al., 2018). These are predictions in the temporal sense of predicting future patterns of neural activity.

However, conventional models of predictive coding such as the one first proposed by Rao and Ballard (1999) are not predictive in this same sense. Rather than predicting the future, these models are predictive in the hierarchical sense of higher areas predicting the activity of lower areas (Rao and Ballard, 1999; Bastos et al., 2012; Spratling, 2012, 2017). These models do not predict what is going to happen: rather, by converging on a configuration of neural activity that optimizes the representation of a stable sensory input, they hierarchically predict what is happening. The use of the word “prediction” in this context is therefore somewhat unfortunate, as conventional models of predictive coding do not actually present a mechanism that predicts in the temporal way that prediction is typically used in ordinary discourse.

Predicting the future

To date, the temporal dimension has been nearly absent from computational work on predictive coding. With the exception of generalized formulations (see below), it is generally implicitly assumed that sensory input is a stationary process (i.e., that it remains unchanged until the system converges on a minimal prediction error solution). The available studies of dynamic stimuli in a predictive coding context consider only autocorrelations within the time series of a given neuron: in other words, the tendency of sensory input to remain the same from moment to moment (van Hateren and Ruderman, 1998; Rao, 1999; Huang and Rao, 2011). Computationally, this prediction is easily implemented as a biphasic temporal impulse response function (Dong and Atick, 1995; Dan et al., 1996; Huang and Rao, 2011), which is consistent with the known properties of neurons in the early visual pathway, including the lateral geniculate nucleus (LGN; Dong and Atick, 1995; Dan et al., 1996). However, this is only the most minimal temporal prediction that a neural system might make: the prediction that its input remains unchanged over time.

A more general framework is provided by generalized (Bayesian) filtering (Friston et al., 2010). The key aspect of generalized predictive coding is that instead of just predicting the current state of sensory impressions, there is an explicit representation of velocity, acceleration, and other higher orders of motion. This sort of representation has been used to model oculomotor delays during active vision (Perrinet et al., 2014). However, neither this framework nor predictive coding approaches implicit in Kalman filters (that include a velocity-based prediction) explicitly account for transmission delays.

As such, what is missing from conventional models of predictive coding is the fact that neural communication itself takes time. Perhaps because the delays involved in *in silico* simulations are negligible, or perhaps because the model was first articulated for stationary processes, namely static images and stable neural representations, models of neural circuitry have entirely neglected to take into account that neural transmission incurs significant delays. These delays mean that forward and backward signals are misaligned in time. For an event at a given moment, the sensory representation at the first level of the hierarchy needs to be fed forward to the next hierarchical level, where a prediction is formulated, which is then fed back to the original level. In the case of a dynamic stimulus, however, by the time that the prediction arrives back at the first hierarchical level, the stimulus will have changed, and the first level will be representing the new state of that stimulus. As a result, backward predictions will be compared against more recent sensory information than the information on which they were originally based, and which they were sent to suppress. If this temporal misalignment between forward and backward signals would not somehow be compensated for, under the classical hierarchical predictive coding model any time-varying stimulus would generate very large prediction errors at each level of representation, which is typically not seen in electrophysiological recordings of responses to stimuli with constant motion (as opposed to unexpected changes in the trajectory; Schwartz et al., 2007; McMillan and Gray, 2012; Dick and Gray, 2014).

Because of neural transmission delays, prediction error is minimized not when a backward signal represents the sensory information that originally generated it, but when it represents the sensory information that is going to be available at the lower level by the time the backward signal arrives. In other words, prediction error is minimized when the backward signal anticipates the future state of the lower hierarchical level. For stimuli that are changing at a constant rate, estimating that future state requires only rate-of-change information about the relevant feature, and it follows that if such information is available at a given level, it will be recruited to minimize prediction error. When allowing for transmission delays, hierarchical predictions therefore need to become temporal predictions: they need to predict the future.

Two extended models

A clear example of common, time-varying sensory input is visual motion. Here, we use visual motion to illustrate the limitations of the classical predictive coding model when input is time varying, and present two extensions to the classical model that would solve these limitations. In

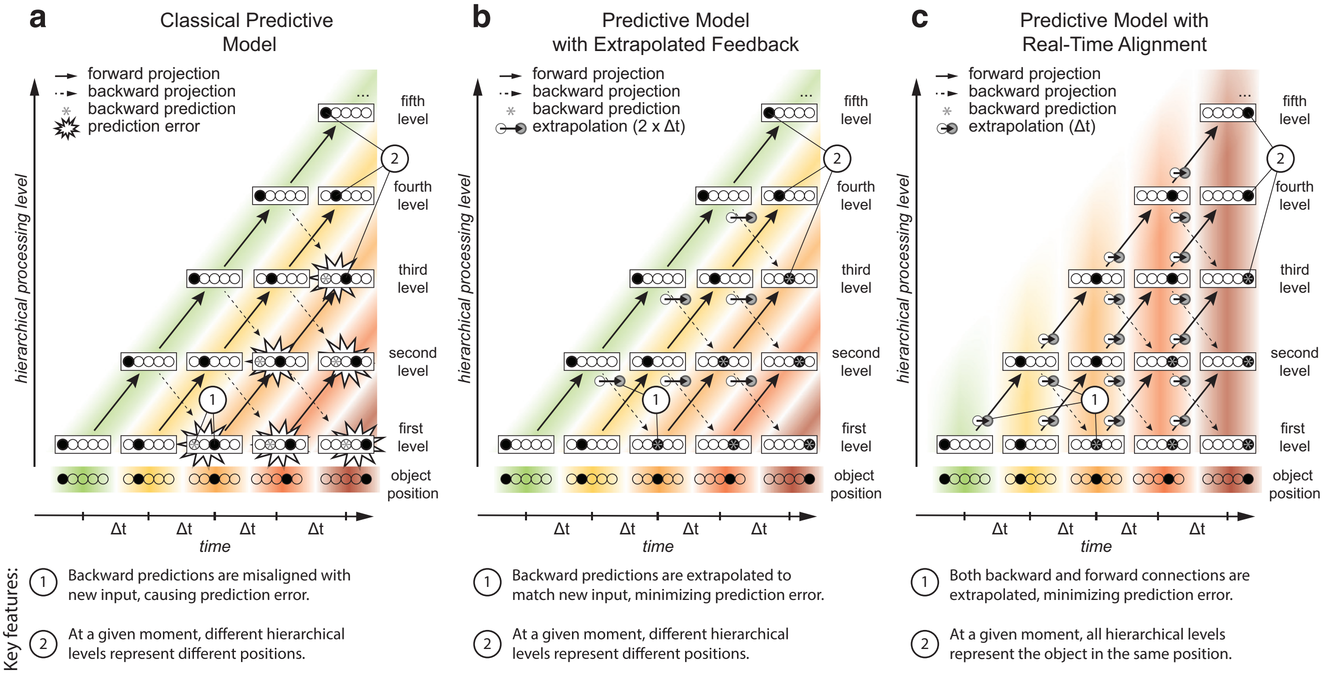


Figure 2. Three simplified models of the representation of the position of a moving object throughout the visual processing hierarchy under the predictive coding framework. Each rectangle denotes the neural representation of the position of the object at a given hierarchical level and at a given time, with the filled circle indicating the object in one of five possible positions. In this simplified representation, all connections are modeled as incurring an equal transmission delay (Δt). Colored bands link corresponding representations, and numbered circles highlight the core features of each model. ***a***, The classical predictive model (model A) of predictive coding comprises forward connections from one hierarchical level to the next level (solid lines), and backward connections to the previous level (dashed lines). No allowance is made for neural transmission delays, such that backward connections carry a position representation (asterisks) that is outdated by the time that signal arrives. The resulting mismatch with more recent sensory input generates large errors, which subsequently propagate through the hierarchy (emphasized with starbursts). ***b***, In the predictive model with extrapolated feedback (model B), an extrapolation mechanism operates on predictive backward projections, anticipating the future position of the object. This mechanism on the backward projections compensates for the total time-cost incurred during both the forward and the backward portion of the loop. This would minimize total prediction error in this simplified model. However, the mechanism would rapidly become more complex when one considers that individual areas tend to send and receive signals to and from multiple levels of the hierarchy. ***c***, In the predictive model with real-time alignment (model C), extrapolation mechanisms compensate for neural delays at both forward and backward steps. This parsimoniously minimizes total prediction error, even for more complex connectivity patterns. Additionally, the model differs from the first two models in that at any given time, all hierarchical levels represent the same position. Conversely, in the first two models, at any given time neural transmission delays mean that all hierarchical levels represent a different position. This crucial difference is evident as vertical, rather than diagonal, colored bands linking matching representations across the hierarchy. The consequence of this hypothesis is therefore that the entire visual hierarchy becomes temporally aligned. This provides an automatic and elegant solution to the computational challenge of establishing which neural signals belong together in time: the temporal binding problem. It is also consistent with demonstrated extrapolation mechanisms in forward pathways and provides a parsimonious explanation for a range of motion-induced position shifts.

this illustration, we consider neural populations at various levels of the visual hierarchy that represent position, for instance as a Gaussian population code. Additionally, we will argue that to effectively represent position despite transmission delays, those neural populations will additionally represent velocity. Consequently, each neural population representing a particular position would consist of subpopulations representing the range of stimulus velocities.

The classical predictive model (model A)

In the classical hierarchical predictive coding model (Rao and Ballard, 1999; Fig. 1*a*), neural transmission delays mean that backward predictions arrive at a level in the processing pathway a significant interval of time after that level fed forward the signal that led to those predictions. Because the input to this level has changed during the elapsed time, this results in prediction error (Fig. 2*a*).

The predictive model with extrapolated feedback (model B)

In this extension to the classical model, each complete forward/backward loop uses any available rate-of-change information to minimize prediction error. In the case of visual motion, this means that the circuit will use concurrent velocity information to anticipate the representation at the lower level by the time the backward signal arrives at that level. In this model, an extrapolation mechanism is implemented only at the backward step of each loop (Fig. 1*b*) to minimize predictive error while leaving the forward sweep of information unchanged (Fig. 2*b*).

The predictive model with real-time alignment (model C)

In the classical model, prediction error results from the cumulative delay of both forward and backward transmission. Prediction error is minimized when this cumulative delay is compensated at any point in the forward–backward loop. Evidence from both perception (Nijhawan, 1994, 2008) and action (Soechting et al., 2009; Zago et al., 2009) strongly suggests that at least part of the total delay incurred is compensated by extrapolation at the forward step. Accordingly, in this model, we propose that extrapolation mechanisms work on both forward and backward signals: any signal that is sent from one level to another, whether forward or backward, is extrapolated into the future that precisely compensates for the delay incurred by that signal while it is in transit (Fig. 1*c*). In addition to minimizing prediction error, this model has the remarkable consequence of synchronizing representations throughout the hierarchy: under this model, all levels represent a moving object in the same position at the same moment (Fig. 2*c*), independent of where in the hierarchy each level lies.

Evaluating the evidence

We have argued above that the classical model of predictive coding (Rao and Ballard, 1999; Huang and Rao, 2011; Bastos et al., 2012; Spratling, 2017) will consistently produce prediction errors when stimuli are time varying and neural transmission delays are taken into consideration (Fig. 2*a*). We have proposed two possible extensions to the classical model, each of which would minimize prediction error. Here, we evaluate the evidence for and against each of these two models.

Neural plausibility

Models B and C are both neurally plausible. First, both models share key features with existing computational models of motion–position interaction. For example, Kwon et al. (2015) recently advanced an elegant computational model uniting motion and position perception. An interaction between motion and position signals is a central premise of their model, such that instantaneous position judgments are biased by concurrent motion stimuli in a way that is qualitatively similar to the model we propose here. In fact, they demonstrate that this interaction predicts a number of properties of a visual effect known as the curveball illusion, in which the internal motion of an object causes its position (and trajectory) to be misperceived. Below, we argue that our extension of hierarchical predictive coding similarly predicts a number of other motion-induced position shifts. One such illusion is the flash-lag effect (Nijhawan, 1994), in which a stationary bar is flashed in alignment with a moving bar, but is perceived to lag behind it. Importantly, Khoei et al. (2017) recently presented a detailed argument that the flash-lag effect could be explained by a Bayesian integration of motion and position signals, and postulated that such mechanisms might compensate for neural delays. This interpretation of the interaction between motion and position signals is entirely consistent with the mechanisms we propose in models B and C. Finally, we have previously argued that the prediction errors that would necessarily arise under these models when objects unexpectedly change direction would lead to another perceptual effect known as the flash-grab illusion (Cavanagh and Anstis, 2013; van Heusden et al., 2019; Fig. 1). Consequently, the models we propose are compatible with existing empirical and computational work.

Second, both extensions to the classical model incorporate neural extrapolation mechanisms at each level of the visual hierarchy. This requires first that information about the rate of change be represented and available at each level. In the example of visual motion, it is well established that velocity is explicitly represented throughout the early visual hierarchy, including the retina (Barlow and Levick, 1965), LGN (Marshel et al., 2012; Cheong et al., 2013; Cruz-Martín et al., 2014; Zaltsman et al., 2015), and primary visual cortex (V1; Hubel and Wiesel, 1968). As required by both models, velocity information is therefore available throughout the hierarchy. Indeed, it follows from the prediction error minimization principle that if velocity is available at each level, it can and will be used at each level to optimize the accuracy of backward predictions involving that level.

Finally, the extrapolation processes posited in both models should cause activity ahead of the position of a moving object, preactivating (i.e., priming the neural representation of) the area of space into which the object is expected to move. This activation ahead of a moving object is consistent with the reported preactivation of predicted stimulus position in cat (Jancke et al., 2004) and macaque (Subramaniyan et al., 2018) V1, as well as in human EEG (Hogendoorn and Burkitt, 2018). Importantly, if the object unexpectedly vanishes, such extrapolation would preactivate areas of space in which the object ultimately never appeared. This is consistent with psychophysical experiments where an object is perceived to move into areas of the retina where no actual stimulus energy can be detected (Maus and Nijhawan, 2008; Shi and Nijhawan, 2012), as well as with recent fMRI work showing activation in retinotopic areas of the visual field beyond the end of the motion trajectory (Schellekens et al., 2016). Although the models presented here are by no means the only models that would be able to explain these results, these results do demonstrate properties predicted by models B and C.

Hierarchical complexity

Model C is more robust to hierarchical complexity than model B. Figure 2 shows extremely simplified hierarchies: it omits connections that span multiple hierarchical levels, it represents all time delays as equal and constant, and it shows only a single projection from each area. In reality, of course, any given area is connected to numerous other areas (Felleman and Van Essen, 1991), each with different receptive field properties. Transmission delays will inevitably differ depending on where forward signals are sent to and where backward signals originate from. Importantly, a given area might receive input about the same moment through different pathways with different lags. A purely backward extrapolation process (model B) would not naturally compensate for these different lags, because it would not be able to differentiate between the forward signals that led to the prediction. Conversely, a model with extrapolation mechanisms at both forward and backward connections (model C) would be able to compensate for any degree of hierarchical complexity: for a feedback loop with a large transmission delay, prediction error would be minimized by extrapolating further along a trajectory than for a feedback loop with a smaller transmission delay. In this way, each connection would essentially compensate for its own delay.

Motion-induced position shifts

Model C predicts motion-induced position shifts, namely visual illusions in which motion signals affect the perceived positions of objects. In model C, the representation in higher hierarchical levels does not represent the properties of the stimulus as it was when it was initially detected, but rather the expected properties of that stimulus at a certain extent into the future. In the case of visual motion, this means that the forward signal represents the extrapolated position rather than the detected position of the moving object. One consequence is that at higher hierarchical levels, a moving object is never represented in the position at which it first appears (Fig. 2*c*, top diagonal). Instead, it is first represented at a certain extent along its (expected) trajectory. Model C therefore neatly predicts a visual phenomenon known as the Fröhlich effect, in which the initial position of an object that suddenly appears in motion is perceived as being shifted in the direction of its motion (Fröhlich, 1924; Kirschfeld and Kammer, 1999). Conversely, model B cannot explain these phenomena, as the feedforward position representation reflects the veridical location of the stimulus, rather than its extrapolated position (Fig. 2*b*).

Model C is also consistent with the much studied flash-lag phenomenon (Nijhawan, 1994), in which a briefly flashed stimulus that is presented in alignment with a moving stimulus appears to lag behind that stimulus. Although alternative explanations have been proposed (Eagleman and Sejnowski, 2000), a prominent interpretation of this effect is that it reflects motion extrapolation, specifically implemented to compensate for neural delays (Nijhawan, 1994, 2002, 2008). Model C is not only compatible with that interpretation, but it provides a principled argument (prediction error minimization) for why such mechanisms might develop.

Neural computations

There are two key computations that the neural-processing hierarchy needs to carry out to implement the proposed real-time alignment of model C. First, it is necessary that the position representation at each successive level incorporates the effect of the neural transmission delay to compensate for the motion, rather than simply reproducing the position representation at the preceding level. Second, this delay-and-motion-dependent shift of the position representation must occur on both the forward and backward connections in the hierarchy. Moreover, these computations must be learned through a process that is plausible in terms of synaptic plasticity, namely that the change in a synaptic strength is activity dependent and local. The locality constraint of synaptic plasticity requires that changes in the synaptic strength (i.e., the weight of a connection) can only depend on the activity of the presynaptic neuron and the postsynaptic neuron. Consequently, the spatial distribution of the synaptic changes in response to a stimulus are confined to the spatial distribution of the position representation of the stimulus, which has important consequences for the structure of the network that emerges as a result of learning.

By incorporating the velocity of the stimuli into the prediction error minimization principle, it becomes possible for a learning process to selectively potentiate the synapses that lead to the hierarchically organized position representations in both models B and C. This requires that the appropriate velocity subpopulation of each position representation is activated to generate the neural representations illustrated in Figure 2, *b* and *c*. The prediction error is generated by the extent of misalignment of the actual input and the predicted input position representation on the feedback path, which involves the combined forward and backward paths, as illustrated in Figure 2. Consequently, the prediction error minimization principle results in changes to the weights on both the forward and backward paths. What distinguishes models B and C is that the probability of potentiation depends on the extent of the spatial overlap of the position representations at successive times. In model C, this spatial overlap between the forward and backward paths is the same, namely the position representation has changed by the same extent during both the forward and backward transmission delays, assuming that the neural delay time is the same for both paths. However, in model B the forward and backward paths are quite different: the forward path has complete congruence in the position representation between adjacent levels, whereas the backward path has a position representation that corresponds to the sum of the forward and backward delays, so that the position representations between the two levels on the backward path are much further apart. Since the position representations are local with a distribution that falls off from the center (e.g., an exponential or power-law fall off), the probability of potentiation of the backward pathway in model B is correspondingly lower (by an exponential or power-law factor). As a result, a local learning rule that implements the prediction error principle would tend to favor the more equal distribution of delays between the forward and backward paths of model C. Since the same learning principle applies to weights between each successive level of the hierarchical structure, it is capable of providing the basis for the emergence of this structure during development.

Both models B and C posit interactions between motion and position signals at multiple levels in the hierarchy. This is compatible with a number of theoretical and computational models of motion and position perception. For example, Eagleman and Sejnowski (2007) have argued that for a whole class of motion-induced position shifts, a local motion signal biases perceived position, precisely the local interactions between velocity and position signals that we propose. Furthermore, these instantaneous velocity signals have been shown to affect not only the perceptual localization of concurrently presented targets, but also the planning of saccadic eye movements aimed at those targets (Quinet and Goffart, 2015; van Heusden et al., 2018). We argue that, under the hierarchical predictive coding framework, delays in neural processing necessarily lead to the evolution of motion–position interactions.

Furthermore, both models B and C posit such interactions at multiple levels, including very early levels in the hierarchy. This is consistent with recent work on the flash-grab effect, a motion-induced position shift in which a target briefly flashed on a moving background that reverses direction is perceived to be shifted away from its true position (Cavanagh and Anstis, 2013). Using EEG, the interaction between motion and position signals that generates the illusion has been shown to have already occurred within the first 80 ms following stimulus presentation, indicating a very early locus of interaction. A follow-up study using dichoptic presentation revealed that, even within this narrow time frame, extrapolation took place in at least two dissociable processing levels (van Heusden et al., 2019).

Temporal alignment

A defining feature of model C is that, due to extrapolation at each forward step, all hierarchical areas become aligned in time. Although neural transmission delays mean that it takes successively longer for a new stimulus to be represented at successive levels of the hierarchy (Lamme et al., 1998), the fact that the neural signal is extrapolated into the future at each level means that the representational content of each consecutive level now runs aligned with the first hierarchical level, and potentially with the world, in real time. In the case of motion, we perceive moving objects where they are, rather than where they were, because the visual system extrapolates the position of those objects to compensate for the delays incurred during processing. Of course, the proposal that the brain compensates neural delays by extrapolation is not new (for review, see Nijhawan, 2008). Rather, what is new here is the recognition that this mechanism has not developed for the purpose of compensating for delays at the behavioral level, but that it follows necessarily from the fundamental principles of predictive coding.

The temporal alignment characterizing model C also provides a natural solution to the problem of temporal binding. Different visual features (e.g., contours, color, motion) are processed in different, specialized brain areas (Livingstone and Hubel, 1988; Felleman and Van Essen, 1991). Due to anatomic and physiologic differences, these areas process their information with different latencies, which leads to asynchrony in visual processing. For example, the extraction of color has been argued to lead the extraction of motion (Moutoussis and Zeki, 1997b; Zeki and Bartels, 1998; Arnold et al., 2001). The question therefore arises of how these asynchronously processed features are nevertheless experienced as a coherent stream of visual awareness, with components of that experience processed in different places at different times (Zeki and Bartels, 1999). Model C intrinsically solves this problem. Because the representation in each area is extrapolated to compensate for the delay incurred in connecting to that area, representations across areas (and therefore features) become aligned in time.

Local time perception

Model C eliminates the need for a central timing mechanism. Because temporal alignment in this model is an automatic consequence of prediction error minimization at the level of local circuits, no central timing mechanism (e.g., internal clock; Gibbon, 1977) is required to carry out this alignment. Indeed, under this model, if a temporal property of the prediction loop changed for a particular part of the visual field, then this would be expected to result in localized changes in temporal alignment. Although this is at odds with our intuitive experience of the unified passage of time, local adaptation to flicker has been found to distort time perception in a specific location in the visual field (Johnston et al., 2006; Hogendoorn et al., 2010). Indeed, these spatially localized distortions have been argued to result from adaptive changes to the temporal impulse response function in the LGN (Johnston, 2010, 2014), which would disrupt the calibration of any neuron extrapolating a given amount of time into the future. By pointing to adaptation in specific LGN populations, this model has emphasized the importance of local circuitry in explaining the spatially localized nature of this temporal illusion. The counterintuitive empirical finding that the perceived timing of events in specific positions in the visual field can be distorted by local adaptation is therefore consistent with the local compensation mechanisms that form part of model C.

Violated predictions

Under the assumption that the observer's percept corresponds to the current stable representation of the stimulus at a given level of representation, models B and C produce stable percepts when objects move on predictable trajectories. In model B, the represented location at any given moment will lag behind the outside world by an amount dependent on the accumulated delay in the hierarchy, and in model C the represented location will be synchronous with the outside world. In both cases, however, the representation at the appropriate level, and hence the conscious percept, will continue to stably evolve. In contrast, when events do not unfold predictably, such as when objects unexpectedly appear, disappear, or change their trajectory, this introduces new forward information, which is, of course, at odds with backward predictions at each level. This gives rise to prediction errors at each level of the hierarchy, with progressively larger prediction errors further up the hierarchy (as it takes longer for sensory information signaling a violation of the status quo to arrive at higher levels). During the intervening time, those areas will continue to (erroneously) extrapolate into a future that never comes to pass.

The breakdown of prediction error minimization (and therefore accurate perception) in situations where events unfold unpredictably fits with empirical studies showing visual illusions in these situations. As noted above, the flash-lag effect is one such visual illusion. Under model C, this effect occurs because the position of the moving object can be extrapolated due to its predictability, whereas the flashed object, which is unpredictable, cannot (Nijhawan, 1994). As the sensory information pertaining to the position of the moving object ascends the visual hierarchy, it is extrapolated at each level. Conversely, the representation of the flash is passed on "as-is," such that a mismatch accumulates between the two. This mismatch between the predictable motion and the unpredictable flash has a parallel in the flash-grab effect, a visual illusion whereby a flash is presented on a moving background that reverses direction (Cavanagh and Anstis, 2013). The result is that the perceived position of the flash is shifted away from the first motion sequence (Blom et al., 2019) and in the direction of the second motion sequence. Importantly, a recent study parametrically varied the predictability of the flash and showed that the strength of the resulting illusion decreased as the flash became more predictable (Adamian and Cavanagh, 2016). Likewise, in studies of temporal binding, the asynchrony in the processing of color and motion is evident only when rapidly moving stimuli abruptly change direction (Moutoussis and Zeki, 1997a). These illusions reveal that accurate perception breaks down when prediction is not possible, consistent with model C, but not with the classical predictive model A.

Conclusions and future directions

Altogether, multiple lines of evidence converge in support of model C: an extension of the hierarchical predictive coding framework in which extrapolation mechanisms work on both forward and backward connections to minimize prediction error. In this model, minimal prediction error is achieved by local extrapolation mechanisms compensating for the specific delays incurred by individual connections at each level of the processing hierarchy. As a result, neural representations across the hierarchy become aligned in real time. This model provides an extension to classical predictive coding models that is necessary to account for neural transmission delays. In addition, the model predicts and explains a wide range of spatiotemporal phenomena, including motion-induced position shifts, the temporal binding problem, and localized distortions of perceived time.

Neural implementation

We have taken error minimization as the organizing principle of predictive coding, but now extended to incorporate the local velocity of the stimuli. This is in keeping with the approach of previous authors, who have successfully modeled hierarchical prediction error minimization (Rao and Ballard, 1997, 1999; Friston, 2010; Huang and Rao, 2011; Bastos et al., 2012; Spratling, 2017) and motion–position interactions (Kwon et al., 2015) as a Kalman filter (Kalman, 1960). The extension proposed here can, in principle, be implemented straightforwardly by incorporating the local velocity as an additional state variable in a manner analogous to that proposed for motion–position interactions (Kwon et al., 2015). Such an approach, in which the prediction error minimization incorporates the expected changes that occur due to both the motion of the stimulus and the propagation of the neural signals, is possible when the local velocity is one of the state variables that each level of the hierarchy has access to. This could be seen as a special case of generalized predictive coding, where there is an explicit leveraging of (sometimes higher-order) motion representations (Friston et al., 2010). This framework even goes beyond the first-order velocity implicit in a Kalman filter to include higher orders of motion. Mathematically, this makes it easy to extrapolate forward and backward in time by simply taking linear mixtures of different higher order motion signals using Taylor expansions (Perrinet et al., 2014). This would be one neurally plausible way in which the extrapolation in Figure 2 could be implemented.

Furthermore, it is important to note that by minimizing prediction error, the system will automatically self-organize as if it "knows" what neural transmission delays are incurred at each step. However, models B and C do not require these delays to be explicitly represented at all. Rather, if position and velocity are corepresented as state variables, and the system is exposed to time-variant input (e.g., a moving object), then selective Hebbian potentiation would suffice to strengthen those combinations of connections that cause the backward projection to accurately intercept the new sensory input. This would be one possible implementation of how extrapolation mechanisms calibrate to rate-of-change signals.

It is important to note that the proposed model extends, rather than replaces, the conventional formulation of predictive coding as first posited by Rao and Ballard (1999). We have not discussed how this model would function in the spatial domain to develop different receptive field properties at each level, as this has been discussed at length by other authors (Rao, 1999; Rao and Ballard, 1999; Jehee et al., 2006; Huang and Rao, 2011), and we intend our model to inherit these characteristics. Our model only extends the conventional model by providing for the situation when input is time variant. When it is not (i.e., for static stimuli), our model reduces to the conventional model.

Altogether, the details of how cortical circuits implement extrapolation processing, how those circuits interact with receptive field properties at different levels, and what synaptic plasticity mechanisms underlie the formation of these circuits still remain to be elaborated, and these are key areas for future research.

Prediction in the retina

The proposed model follows from the principle of error minimization within predictive feedback loops. These feedback loops are ubiquitous throughout the visual pathway, including backward connections from V1 to LGN. Although there are no backward connections to the retina, extrapolation mechanisms have nevertheless been reported in the retina (Berry et al., 1999), and these mechanisms have even been found to produce a specific response to reversals of motion direction, much akin to a prediction error (Holy, 2007; Schwartz et al., 2007; Chen et al., 2014). Indeed, the retina has been argued to implement its own, essentially self-contained predictive coding mechanisms that adaptively adjust spatiotemporal receptive fields (Hosoya et al., 2005). In the absence of backward connections to the retina, our model does not directly predict these mechanisms, instead predicting only extrapolation mechanisms in the rest of the visual pathway where backward connections are ubiquitous. On the long timescale, compensation for neural delays in the retina provides a behavioral, and therefore evolutionary, advantage, but more research will be necessary to address whether any short-term learning mechanisms play a role in the development of these circuits.

The ubiquity of velocity signals

We have emphasized hierarchical mechanisms at early levels of visual processing, consistent with extrapolation in monocular channels (van Heusden et al., 2019), early EEG correlates of extrapolation (Hogendoorn et al., 2015), and evidence that extrapolation mechanisms are shared for both perceptual localization and saccadic targeting (van Heusden et al., 2018). However, an influential body of literature has proposed that the human visual system is organized into two (partly) dissociable pathways: a ventral "what" pathway for object recognition and identification, and a dorsal "where" pathway for localization and motion perception (Mishkin and Ungerleider, 1982; Goodale et al., 1991; Goodale and Milner, 1992; Aglioti et al., 1995). Two decades later, the distinction is more nuanced (Milner and Goodale, 2008; Gilaie-Dotan, 2016), but the question remains whether velocity signals are truly ubiquitous throughout the visual hierarchy. However, given that the identity of a moving object typically changes much more slowly than its position, it may be that the neuronal machinery that underwrites (generalized) predictive coding in the ventral stream does not have to accommodate transmission delays. This may provide an interesting interpretation of the characteristic physiologic time constants associated with the magnocellular stream (directed toward the dorsal stream), compared with the parvocellular stream implicated in object recognition (Zeki and Shipp, 1988). Consequently, whether real-time temporal alignment is restricted to the "where" pathway or is a general feature of cortical processing that remains to be elucidated.

The functional role of prediction error

Throughout this perspective, we have considered prediction error as something that an effective predictive coding model should minimize. From the perspective of the free energy principle, the imperative to minimize prediction error can be argued from first principles. The sum of squared prediction error is effectively variational free energy, and any system that minimizes free energy will therefore self-organize to its preferred physiologic and perceptual states. However, we have not addressed the functional role of the prediction error signal itself. As noted by several authors, this signal might serve an alerting or surprise function (for review, see den Ouden et al., 2012). In the model proposed here, the signal might have the additional function of correcting a faulty extrapolation (Nijhawan, 2002, 2008; Shi and Nijhawan, 2012). In this role, prediction error signals would work to eliminate lingering neural traces of predictions that were unsubstantiated by sensory input: expected events that did not end up happening. This corrective function has been modeled as a "catch-up" effect for trajectory reversals in the retina (Holy, 2007; Schwartz et al., 2007), and as an increase in position uncertainty (and therefore an increase in relative reliance on sensory information) when objects change trajectory in a recently proposed Bayesian model of visual motion and position perception (Kwon et al., 2015). Further identifying the functional role of these signals is an exciting avenue for future research.

References

Adamian N, Cavanagh P (2016) Localization of flash grab targets is improved with sustained spatial attention. J Vis 16:1266.

Adams RA, Shipp S, Friston KJ (2013) Predictions not commands: active inference in the motor system. Brain Struct Funct 218:611–643.

Aglioti S, DeSouza JF, Goodale MA (1995) Size-contrast illusions deceive the eye but not the hand. Curr Biol 5:679–685.

Arnold DH, Clifford CW, Wenderoth P (2001) Asynchronous processing in vision: color leads motion. Curr Biol 11:596–600.

Barlow HB, Levick WR (1965) The mechanism of directionally selective units in rabbit's retina. J Physiol 178:477–504.

Bastos AM, Usrey WM, Adams RA, Mangun GR, Fries P, Friston KJ (2012) Canonical microcircuits for predictive coding. Neuron 21: 695–711.

Bendixen A, SanMiguel I, Schröger E (2012) Early electrophysiological indicators for predictive processing in audition: a review. Int J Psychophysiol 83:120–131.

Berry MJ, Brivanlou IH, Jordan TA, Meister M (1999) Anticipation of moving stimuli by the retina. Nature 398:334–338.

Blakemore S-J, Goodbody SJ, Wolpert DM (1998) Predicting the consequences of our own actions: the role of sensorimotor context estimation. J Neurosci 18:7511–7518.

Blom T, Liang Q, Hogendoorn H (2019) When predictions fail: correction for extrapolation in the flash-grab effect. J Vis 19:3.

Cavanagh P, Anstis SM (2013) The flash grab effect. Vis Res 91:8–20.

Chen EY, Chou J, Park J, Schwartz G, Berry MJ (2014) The neural circuit mechanisms underlying the retinal response to motion reversal. J Neurosci 34:15557–15575.

Cheong SK, Tailby C, Solomon SG, Martin PR (2013) Cortical-like receptive fields in the lateral geniculate nucleus of marmoset monkeys. J Neurosci 33:6864–6876.

Cruz-Martín A, El-Danaf RN, Osakada F, Sriram B, Dhande OS, Nguyen PL, Callaway EM, Ghosh A, Huberman AD (2014) A dedicated circuit linking direction selective retinal ganglion cells to primary visual cortex. Nature 507:358.

Dan Y, Atick JJ, Reid RC (1996) Efficient coding of natural scenes in the lateral geniculate nucleus: experimental test of a computational theory. J Neurosci 16:3351–3362.

Dayan P, Hinton GE, Neal RM, Zemel RS (1995) The Helmholtz machine. Neural Comput 7:889–904.

den Ouden HEM, Kok P, de Lange FP (2012) How prediction errors shape perception, attention, and motivation. Front Psychol 3:548.

Dick PC, Gray JR (2014) Spatiotemporal stimulus properties modulate responses to trajectory changes in a locust looming-sensitive pathway. J Neurophysiol 111:1736–1745.

Dong DW, Atick JJ (1995) Temporal decorrelation: a theory of lagged and nonlagged responses in the lateral geniculate nucleus. Network 6:159–178.

Eagleman DM, Sejnowski TJ (2000) Motion integration and postdiction in visual awareness. Science 287:2036–2038.

Eagleman DM, Sejnowski TJ (2007) Motion signals bias localization judgments: a unified explanation for the flash-lag, flash-drag, flash-jump, and Frohlich illusions. J Vis 7:3.

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1:1–47.

Friston K (2005) A theory of cortical responses. Philos Trans R Soc Lond B Biol Sci 360:815–836.
Friston K (2010) The free-energy principle: a unified brain theory? Nat Rev Neurosci 11:127–138.
Friston K (2018) Does predictive coding have a future? Nat Neurosci 21:1019–1021.
Friston K, Kiebel S (2009) Predictive coding under the free-energy principle. Philos Trans R Soc Lond B Biol Sci 364:1211–1221.
Friston K, Stephan K, Li B, Daunizeau J (2010) Generalised filtering. Math Probl Eng 2010:621670.
Fröhlich FW (1924) Über die messung der empfindungszeit. Pflügers Archiv Die Gesamte Physiol Menschen Tiere 202:566–572.
Garrido MI, Kilner JM, Stephan KE, Friston KJ (2009) The mismatch negativity: a review of underlying mechanisms. Clin Neurophysiol 120:453–463.
Gibbon J (1977) Scalar expectancy theory and Weber's law in animal timing. Psychol Rev 84:279–325.
Gilaie-Dotan S (2016) Visual motion serves but is not under the purview of the dorsal pathway. Neuropsychologia 89:378–392.
Goodale MA, Milner AD (1992) Separate visual pathways for perception and action. Trends Neurosci 15:20–25.
Goodale MA, Milner AD, Jakobson LS, Carey DP (1991) A neurological dissociation between perceiving objects and grasping them. Nature 349:154–156.
Gregory RL (1980) Perceptions as hypotheses. Philos Trans R Soc Lond B Biol Sci 290:181–197.
Hogendoorn H, Burkitt AN (2018) Predictive coding of visual object position ahead of moving objects revealed by time-resolved EEG decoding. Neuroimage 171:55–61.
Hogendoorn H, Verstraten FAJ, Johnston A (2010) Spatially localized time shifts of the perceptual stream. Front Psychol 1:181.
Hogendoorn H, Verstraten FAJ, Cavanagh P (2015) Strikingly rapid neural basis of motion-induced position shifts revealed by high temporal-resolution EEG pattern classification. Vision Res 113:1–10.
Hohwy J (2016) The self-evidencing brain. Noûs 50:259–285.
Holy TE (2007) A public confession: the retina trumpets its failed predictions. Neuron 55:831–832.
Hosoya T, Baccus SA, Meister M (2005) Dynamic predictive coding by the retina. Nature 436:71–77.
Huang Y, Rao RPN (2011) Predictive coding. Wiley Interdiscip Rev Cogn Sci 2:580–593.
Hubel DH, Wiesel T (1968) Receptive fields and functional architecture of monkey striate cortex. J Physiol 195:215–243.
Jancke D, Erlhagen W, Schöner G, Dinse HR (2004) Shorter latencies for motion trajectories than for flashes in population responses of cat primary visual cortex. J Physiol 556:971–982.
Jehee JFM, Rothkopf C, Beck JM, Ballard DH (2006) Learning receptive fields using predictive feedback. J Physiol Paris 100: 125–132.
Johnston A (2010) Modulation of time perception by visual adaptation. In: Attention and time (Nobre K, Coull JT, eds), pp 187–200. Oxford, UK: Oxford UP.
Johnston A (2014) Visual time perception. In: The new visual neurosciences (Werner JS, Chalupa LM, eds), pp 749–762. Cambridge, MA: MIT.
Johnston A, Arnold DH, Nishida S (2006) Spatially localized distortions of event time. Curr Biol 16:472–479.
Kalman RE (1960) A new approach to linear filtering and prediction problems. J Basic Eng 82:35.
Khoei MA, Masson GS, Perrinet LU (2017) The flash-lag effect as a motion-based predictive shift. PLoS Comput Biol 13:e1005068.
Kirschfeld K, Kammer T (1999) The Fröhlich effect: a consequence of the interaction of visual focal attention and metacontrast. Vision Res 39:3702–3709.
Kogo N, Trengove C (2015) Is predictive coding theory articulated enough to be testable? Front Comput Neurosci 9:111. .
Kok P, Mostert P, de Lange FP (2017) Prior expectations induce prestimulus sensory templates. Proc Natl Acad Sci U S A 114: 10473–10478.
Kwon O-S, Tadin D, Knill DC (2015) Unifying account of visual motion and position perception. Proc Natl Acad Sci U S A 112:8142–8147.
Lamme VAF, Supèr H, Spekreijse H (1998) Feedforward, horizontal, and feedback processing in the visual cortex. Curr Opin Neurobiol 8:529–535.
Livingstone M, Hubel DH (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 240:740–749.
Marshel JH, Kaye AP, Nauhaus I, Callaway EM (2012) Anterior-posterior direction opponency in the superficial mouse lateral geniculate nucleus. Neuron 76:713–720.
Maus GW, Nijhawan R (2008) Motion extrapolation into the blind spot. Psychol Sci 19:1087–1091.
McMillan GA, Gray JR (2012) A looming-sensitive pathway responds to changes in the trajectory of object motion. J Neurophysiol 108:1052–1068.
Milner AD, Goodale MA (2008) Two visual systems re-viewed. Neuropsychologia 46:774–785.
Mishkin M, Ungerleider LG (1982) Contribution of striate inputs to the visuospatial functions of parieto-preoccipital cortex in monkeys. Behav Brain Res 6:57–77.
Moutoussis K, Zeki S (1997a) A direct demonstration of perceptual asynchrony in vision. Proc Biol Sci 264:393–399.
Moutoussis K, Zeki S (1997b) Functional segregation and temporal hierarchy of the visual perceptive systems. Proc Biol Sci 264: 1407–1414.
Mumford D (1992) On the computational architecture of the neocortex. II. The role of cortico-cortical loops. Biol Cybern 66:241–251.
Necker LA (1832) LXI. Observations on some remarkable optical phænomena seen in Switzerland; and on an optical phænomenon which occurs on viewing a figure of a crystal or geometrical solid. Lond Edinburgh Dublin Philos Mag J Sci 1:329–337.
Nijhawan R (1994) Motion extrapolation in catching. Nature 370:256–257.
Nijhawan R (2002) Neural delays, visual motion and the flash-lag effect. Trends Cogn Sci 6:387.
Nijhawan R (2008) Visual prediction: psychophysics and neurophysiology of compensation for time delays. Behav Brain Sci 31:179–239.
Perrinet LU, Adams RA, Friston KJ (2014) Active inference, eye movements and oculomotor delays. Biol Cybern 108:777–801.
Quinet J, Goffart L (2015) Does the brain extrapolate the position of a transient moving target? J Neurosci 35:11780–11790.
Rao RPN (1999) An optimal estimation approach to visual perception and learning. Vision Res 39:1963–1989.
Rao RPN, Ballard DH (1997) Dynamic model of visual recognition predicts neural response properties in the visual cortex. Neural Comput 9:721–763.
Rao RPN, Ballard DH (1999) Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nat Neurosci 2:79–87.
Schellekens W, van Wezel RJA, Petridou N, Ramsey NF, Raemaekers M (2016) Predictive coding for motion stimuli in human early visual cortex. Brain Struct Funct 221:879–890.
Schultz W (1998) Predictive reward signal of dopamine neurons. J Neurophysiol 80:1–27.
Schwartz G, Taylor S, Fisher C, Harris R, Berry MJ (2007) Synchronized firing among retinal ganglion cells signals motion reversal. Neuron 55:958–969.
Shi Z, Nijhawan R (2012) Motion extrapolation in the central fovea. PLoS One 7:33651.
Soechting JF, Juveli JZ, Rao HM (2009) Models for the extrapolation of target motion for manual interception. J Neurophysiol 102: 1491–1502.
Spratling MW (2008) Reconciling predictive coding and biased competition models of cortical function. Front Comput Neurosci 2:4.
Spratling MW (2012) Predictive coding accounts for V1 response properties recorded using reverse correlation. Biol Cybern 106:37–49.
Spratling MW (2017) A review of predictive coding algorithms. Brain Cogn 112:92–97.

Srinivasan MV, Laughlin SB, Dubs A (1982) Predictive coding: a fresh view of inhibition in the retina. Proc R Soc Lond B Biol Sci 216:427–459.

Sterzer P, Voss M, Schlagenhauf F, Heinz A (2018) Decision-making in schizophrenia: a predictive-coding perspective. Neuroimage. Advance online publication. Retrieved March 27, 2019. doi: 10.1016/j.neuroimage.2018.05.074.

Subramaniyan M, Ecker AS, Patel SS, Cotton RJ, Bethge M, Pitkow X, Berens P, Tolias AS (2018) Faster processing of moving compared with flashed bars in awake macaque V1 provides a neural correlate of the flash lag illusion. J Neurophysiol 120:2430–2452.

Summerfield C, de Lange FP (2014) Expectation in perceptual decision making: neural and computational mechanisms. Nat Rev Neurosci 15:745–756.

van Ede F, de Lange F, Jensen O, Maris E (2011) Orienting attention to an upcoming tactile event involves a spatially and temporally specific modulation of sensorimotor alpha- and beta-band oscillations. J Neurosci 31:2016–2024.

van Hateren JH, Ruderman DL (1998) Independent component analysis of natural image sequences yields spatio-temporal filters similar to simple cells in primary visual cortex. Proc R Soc Lond B Biol Sci 265:2315–2320.

van Heusden E, Rolfs M, Cavanagh P, Hogendoorn H (2018) Motion extrapolation for eye movements predicts perceived motion-induced position shifts. J Neurosci 38:8243–8250.

van Heusden E, Harris AM, Garrido MI, Hogendoorn H (2019) Predictive coding of visual motion in both monocular and binocular human visual processing. J Vis 19:3.

von Helmholtz H (1867) Handbuch der physiologischen optik. Leipzig: Voss.

Zago M, McIntyre J, Senot P, Lacquaniti F (2009) Visuo-motor coordination and internal models for object interception. Exp Brain Res 192:571–604.

Zaltsman JB, Heimel JA, Van Hooser SD (2015) Weak orientation and direction selectivity in lateral geniculate nucleus representing central vision in the gray squirrel Sciurus carolinensis. J Neurophysiol 113:2987–2997.

Zeki S, Bartels A (1998) The asynchrony of consciousness. Proc R Soc Lond B Biol Sci 265:1583–1585.

Zeki S, Bartels A (1999) Toward a theory of visual consciousness. Conscious Cogn 8:225–259.

Zeki S, Shipp S (1988) The functional logic of cortical connections. Nature 335:311–317.