Title
Construction and evaluation of an integrated dynamical model of visual motion perception

Authors
Tlapale, Émilien
Dosher, Barbara Anne
Lu, Zhong-Lin

Publication Date
2015-07-01

DOI
10.1016/j.neunet.2015.03.011

Peer reviewed

HHS Public Access

Author manuscript

Neural Netw. Author manuscript; available in PMC 2016 July 01.

Published in final edited form as:

Neural Netw. 2015 July ; 67: 110–120. doi:10.1016/j.neunet.2015.03.011.

Construction and evaluation of an integrated dynamical model of visual motion perception

Émilien Tlapale[a], **Barbara Anne Dosher**[a], and **Zhong-Lin Lu**[b]

Émilien Tlapale: etlapale@uci.edu; Barbara Anne Dosher: bdosher@uci.edu; Zhong-Lin Lu: lu.535@osu.edu

[a]Memory Attention and Perception Laboratory, Cognitive Sciences Department, University of California, Irvine, CA, USA

[b]Laboratory of Brain Processes, Department of Psychology, Ohio State University, Columbus, OH, USA

Abstract

Although numerous models describe the individual neural mechanisms that may be involved in the perception of visual motion, few of them have been constructed to take arbitrary stimuli and map them to a motion percept. Here, we propose an integrated dynamical motion model (IDM), which is sufficiently general to handle diverse moving stimuli, yet sufficiently precise to account for a wide-ranging set of empirical observations made on a family of random dot kinematograms. In particular, we constructed models of the cortical areas involved in motion detection, motion integration and perceptual decision. We analyzed their parameters through dynamical simulations and numerical continuation to constrain their proper ranges. Then, empirical data from a family of random dot kinematograms experiments with systematically varying direction distribution, presentation duration and stimulus size, were used to evaluate our model and estimate corresponding model parameters. The resulting model provides an excellent account of a demanding set of parametrically varied behavioral effects on motion perception, providing both quantitative and qualitative elements of evaluation.

Keywords

visual motion perception; random dot kinematograms; systematic parameter variations; threshold estimation; spatialized model

1. Introduction

Although our understanding of the underlying neural mechanisms of motion perception is incomplete, a range of motion processing models have been proposed to account for various

Correspondence to: Émilien Tlapale, `etlapale@uci.edu`.

A website associated with the present work, including source code, can be found at http://emilien.tlapale.com/motion-model

properties of visual motion perception. Among those models, perceptual models provide a high-level characterization of stimuli and perception, and neural models suggest biologically plausible mechanisms to process motion information. Of course, perceptual models often propose possible neural mechanisms, and neural model properties often show direct connections to known perceptual phenomena. Yet, the links between the two levels of analysis can be sketchy: Sometimes, the connections between the neural components are not well understood; sometimes studying the resulting complex systems requires the application of a wide range of analytic procedures and significant computational power.

With increasing precision of experimental measurements and computational power for simulations, it has become possible to design and implement visual motion models encompassing both neurophysiology and psychophysics, further explicating the links between perception and its neural substrate. In doing so, a successful visual motion processing model will incorporate and implement various neural components associated with different neural populations at multiple levels of visual motion processing, including components for motion detection, motion integration, perception, and decision.

In the literature, neural motion models typically consider two-stages of motion processing where a local motion detector typifies processing in primate cortical area V1 (van Santen and Sperling, 1984; Adelson and Bergen, 1985) and feeds motion information to a large scale integrator associated with area MT (Heeger et al., 1996; Simoncelli and Heeger, 1998; Rust et al., 2006)

There have been a number of attempts to extend the traditional two-stage model. The models of Chey et al. (1997, 1998) include variants to account for motion detection and motion integration, yet provide only qualitative comparison of model predictions to experimental results. Similar models have been proposed to investigate the role of feedback between cortical areas V1 and MT (Bayerl and Neumann, 2004), to include form-motion interactions (Berzhanskaya et al., 2007; Bayerl and Neumann, 2007; Beck and Neumann, 2010), to further study the dynamics of motion integration (Tlapale et al., 2010), or to consider rotations and expansions (Raudies et al., 2011). Those models extend the traditional two-stage motion models (Heeger et al., 1996; Simoncelli and Heeger, 1998; Rust et al., 2006), effectively solving the motion aperture problem through non-linear normalization, and include various initial motion detection stages based on specific neural computations (Chey et al., 1997, 1998), phenomenological models (Bayerl and Neumann, 2004), or motion energy models (Tlapale et al., 2010).

Bayesian models have been used to model motion perception at a higher level (Weiss and Adelson, 1998; Weiss et al., 2002) or eye movements (Montagnini et al., 2007; Bogadhi et al., 2011). Typically those models assume probabilistic inputs and outputs, with various levels of abstraction, such as segregation between 1D and 2D components (Montagnini et al., 2007), and define some quantitative value to be maximized. Although general neural implementation strategies for Bayesian mechanisms have been proposed (Rao, 2004), as is the case for other optimization methods such as variational approaches (Viéville et al., 2007), the link to the effective neural computations is often not well specified. Indeed the

focus and power of those approaches lie in the high level description of the task performed by the neural system, with respect to its a priori knowledge.

Physiologists have applied standard models of decision-making to random dot motion and other stimuli. Perceptual decisions in two-alternative forced choice are represented in drift diffusion models in which positive or negative evidence accumulate until a threshold is reached (Stone, 1960; Ratcliff, 1978). More recently, Wang (2002) constructed a spiking neural network accounting for a range of perceptual decision making experiments in random dot motion, and later provided a neural field approximation of their model (Wong and Wang, 2006; Wong et al., 2007), while Machens et al. (2005) proposed a model for two interval forced choice decision paradigms. To make computational and mathematical analysis tractable, most of those approaches assume static or minimum representations where the input typically represents two subpopulations of MT neurons corresponding to the two choices in a forced choice paradigm, with an average activity given by a linear function of the motion coherence level in the stimulus (Mazurek et al., 2003; Wong and Wang, 2006; Wong et al., 2007).

Finally, the influence of internal and external noise on perception is largely ignored in multi-scale models. Yet, noise is an important component of both the stimuli and the processing at every level of the visual system. In the motion domain, the classic random dot kinematograms are defined by the large amount of (external) noise they contain. Manipulating stimulus noise is also a particularly useful tool to analyze a system, allowing researchers to distinguish between several types of internal noise (Lu and Dosher, 1999, 2008) and to investigate the influence of top-down signals (Lu and Dosher, 1998).

We propose an integrated dynamical motion (IDM) model of motion perception that incorporates biologically plausible motion detection and motion integration mechanisms, as well as a decision mechanism to account for reaction times in perceiving motion from random dot kinematograms. Our model includes temporal dynamics that allow us to consider the systematic stimulus variations and corresponding empirical results such as those described in Watamaniuk et al. (1989); Watamaniuk and Sekuler (1992). Such parametric variations, including systematic changes of random dot distribution, presentation duration and stimulus size, remain uncommon in research on motion perception, but provide an empirically grounded test bed for models of motion perception.

In Section 2 we detail the design of the multi-scale model of motion perception based on the known architecture of the visual system. We also make use of dynamical simulations and numerical continuations to identify stable regions of the parameter space for model implementation. In Section 3 we focus on modeling the set of random dot experiments in Watamaniuk et al. (1989); Watamaniuk and Sekuler (1992). We show that the model is able to account for the experimental data and specify the corresponding parameters. In Section 4 we discuss the biological plausibility of the model, compare it to alternative approaches, and conclude with future extensions of the framework.

2. Model

To reproduce, explain and predict motion perception and perceptual decisions from specific motion stimuli, we devised a detailed integrated dynamical motion (IDM) model of visual motion perception. The model is grounded in the current anatomical and electrophysiological knowledge of the human and primate visual system. As such it follows the typical multi-stage view of motion integration, where a first stage detects motion and is linked to the activity of V1 or MT component cells (Section 2.1), while a second stage integrates the activity to extract global motion and is linked to the activity of MT pattern cells (Section 2.2). The model is able to reproduce a variety of MT neural responses, and explain the corresponding motion percepts with additional decision mechanisms linked to neural processing in area LIP (Section 2.3). The interactions between the three layers are represented in Figure 1.

2.1. Motion detection

We follow a standard approach to define directional V1 cells by combining the responses of two non-directional V1 subpopulations. Formally, we start by representing stimuli as varying luminance values, noted $I(t, x, y) \in \mathbb{R}$ where t, x, and y, are the temporal and visual field positions. The difference between this approach to direction selectivity in V1 and prior approaches (Adelson and Bergen, 1985; Escobar et al., 2009) is the use of temporally monophasic filters matching primate cell recordings (De Valois and Cottaris, 1998; De Valois et al., 2000), in addition to the temporally biphasic filters.

The response of directional cells tuned to direction θ is defined by

$$D(t,x,y,\theta)=(M_\theta+B)\overset{t,x,y}{*}I(t,x,y), \quad (1)$$

where $\overset{t,x,y}{*}$ denotes the spatiotemporal convolution operator, M_θ is the kernel of temporally monophasic, spatially odd neurons (the *lower left* component in Figure 1 and Figure 2), and B is the kernel of temporally biphasic, spatially even neurons (the *lower right* component in the same figures). The kernel of monophasic neurons is defined as

$$M(t,x,y)=\Gamma_{n_M,\tau_M}(G_{\sigma_M}(x-a\cos\theta,y-a\sin\theta)-G_{\sigma_M}(x+a\cos\theta,y+a\sin\theta)), \quad (2)$$

where $\Gamma_{n,\tau}$ and G_σ are temporal and spatial localization functions respectively, and defined as

$$\Gamma_{n,\tau}(t)=(nt)^n\frac{\exp(-nt/\tau)}{(n-1)!\tau^{n+1}}, \quad (3)$$

which is a Gamma function (de Vries and Principe, 1991), and

$$G_\sigma(x,y)=\frac{1}{2\pi\sigma^2}\exp\left(-\frac{x^2+y^2}{2\sigma^2}\right), \quad (4)$$

which is a Gaussian function with standard deviation (spread) σ.

The second non-directional subpopulation combines a biphasic temporal part with two Gamma components and an even spatial profile represented as a difference of Gaussians:

$$B(t,x,y)=(\Gamma_{n_B,\tau_B}-\Gamma_{n'_B,\tau'_B})(t)(G_{\sigma_B}-bG_{\sigma'_B})(x,y). \quad (5)$$

As shown in Figure 1 and Figure 2, the spatial and temporal lobes of the non-directional subpopulations are in quadrature, their combination producing a slanted spatiotemporal receptive field typical of directional neurons (Adelson and Bergen, 1985). These monophasic and biphasic receptive fields, and the resulting directional response units, correspond to the experimental data of De Valois and Cottaris (1998); De Valois et al. (2000).

The spatial and temporal filters modeling V1 populations can be efficiently approximated by infinite impulse response filters: for the Gaussian filters of (2) and (5) we combine the recursive filters proposed by Deriche (1990), whereas for the Gamma filters we use a cascade of exponential recursive filters (Principe et al., 1993; Wohrer, 2008).

2.2. Motion integration

The activity $A(t, x, y, \theta)$ of MT neurons with receptive fields centered at (x, y) and a preferred motion direction θ is defined by the differential equation

$$\frac{\partial A}{\partial t}=-\tau_A A+S\circ(F_A+L_A), \quad (6)$$

The dynamics of the neuron populations are controlled by the time constant τ_A, with S a sigmoid non-linearity defined by

$$S(u)=\frac{\alpha}{\beta+\exp(-\lambda u+\gamma)}, \quad (7)$$

and applied to both the feedforward and lateral components of the activity, F_A and L_A, respectively.

Motion information is integrated across the spatial receptive field of MT neurons in the feedforward component, by pooling information from V1 neurons with matching preferences:

$$F_A(t,x,y,\theta)=cG_{\sigma_A}\overset{x,y}{*}S(D(t,x,y,\theta)-dD(t,x,y,\theta+180^\circ)). \quad (8)$$

The spatial receptive field of MT neurons has a Gaussian shape with standard deviation σ; $\overset{x,y}{*}$ is the spatial convolution operator. The value of d is set sufficiently small: as d approaches 1 the receptive field of the motion detector becomes less and less slanted, until it is no longer sensitive to motion. The sigmoid nonlinearity limits the spatial range of opponency to a level comparable to V1 receptive fields (Qian and Andersen, 1994; Heeger et al., 1999).

The lateral component disambiguates motion information by selecting directions with the largest responses:

$$L_A(t,x,y,\theta)=(G_\sigma - eG_{\sigma'}) \overset{\theta}{*} A(t,x,y,\theta) \quad (9)$$

where $\overset{\theta}{*}$ is the convolution operator in the direction domain. The difference of Gaussians inhibits directions more than 45° apart (Snowden et al., 1991), consistent with the selection pattern of inhibitory connections in layer MT (see Figure 1).

2.3. Perceptual decision making

We model perceptual decision as a drift diffusion process driven by the activity of two populations that we identify with subpopulations of area LIP. The general form is given by the stochastic equation

$$\mathrm{d}C_i(t)=(F_{C_i}(t)+L_{C_i}(t))\mathrm{d}t+\mathrm{d}W \quad (10)$$

where C_i defines the activity of the population encoding decision i and dW is a Wiener process. Although not used in the present model, the activity of the LIP population could be extracted by applying a sigmoid activation function. A decision is made when the difference between the two activities reaches a certain threshold ϱ, or after a fixed maximal duration $t_{\max}$.

The feedforward component that weights MT activity and provides input to the perceptual decision units in LIP is described by:

$$F_{C_i}(t)=\iiint_{X\times Y\times\Theta} W_i(\theta)I(t,x,y,\theta)\mathrm{d}x\mathrm{d}y\mathrm{d}\theta, \quad (11)$$

where W_i is the feedforward connectivity from MT to decision unit i. The lateral connectivity within the decision units in LIP includes both self-excitation and cross-inhibition:

$$L_{C_i}(t)=fC_i(t) - gC_{\bar{i}}(t), \quad (12)$$

where $\bar{\imath}$ denotes the alternative decision. Such a self-excitation/cross-inhibition connectivity is often employed in decision models and has been used to successfully approximate

patterns of decision accuracy and reaction time histograms in several motion experiments (Wong and Wang, 2006; Wong et al., 2007; Deco and Martí, 2007).

We first consider some qualitative predictions of the model in coherent motion perception, one of the most popular motion perception tasks used in both psychophysics and physiology. In particular, we ran the experiment of Kayser et al. (2009), a human equivalent of Britten et al. (1993) or Roitman and Shadlen (2002), in which observers were asked to identify the motion direction of random dot kinematograms in a left/right judgment, while the coherence of the display is varied. Here coherence refers to the proportion of dots moving in the global direction while noise dots move in random directions. For this task, the model results are consistent with the patterns observed in behavioral experiments, as illustrated in Figure 3.

In Figure 4 we show model predictions of the average response times as a function of motion coherence, as well as the response time histograms at two coherence levels. The model predictions match qualitatively with the shape of empirical response time distributions, which has a long tail for longer latencies. The predicted response time distributions are also consistent with the observation that parametric variations in reaction time experiments often affect the long tail of response time distributions while having little impact on the short tail (Ratcliff, 1978; Roitman and Shadlen, 2002). As coherence increases and the motion direction is easier to identify, the accuracy increases, and the response times decrease with tightening of the long tail. Further development of the response time aspects model, possibly with addition of baseline encoding and response distributional parameters (often used in diffusion models, for example), would be necessary to improve the quantitative fits. Our goal here was to document the ability of the model to generally account for the observed data patterns.

2.4. Parameterization

The equations defining the model components described in the previous sections are inspired by the known architecture of the visual system, in particular from the connectivity between different stages of motion processing. The exact mapping between each of the components and physiology often remains loosely specified in motion models. In this section, we consider physiological recordings, numerical model analysis, and perceptual responses to refine our model and provide constraints, where available, on the plausible values of some of its parameters.

For the lower sensory neural populations, some parameters can be constrained by the physiology, in particular by receptive field mappings. This is the case for the temporal and spatial (retinotopic) extents of the motion detectors, Equations (1)–(5), whose parameters we generalized from the neural recordings of De Valois et al. (2000). Although we can reproduce the original receptive fields to precisely match the particular cells recorded in the original experiment, compare for instance our Figure 2 with Figure 7 of De Valois et al. (2000), the cell recordings are often sparse, describing only a few cells or investigating responses in a limited range of stimulus variations. The physiological estimates are used to broadly constrain the ranges of the parameter values, allowing for further fine-tuning.

The dynamics of neural populations modeled as a dynamical system are tuned by varying their time constant, in Equations (6) and (10). In the model, the temporal properties of the MT population (6) were set to approximate the dynamics of MT neurons solving a motion aperture problem in the macaque (Pack and Born, 2001). The dynamics of the perceptual decision component (10) were initially set from the experimental data of a similar stimulus described in Lorenceau et al. (1993), and later fine tuned to fit the experimental data of Watamaniuk et al. (1989); Watamaniuk and Sekuler (1992). These datasets are challenging because the model must account for the effects of systematic variations of stimulus parameters.

The selection of some of the parameters is relatively easy, either because they have clear biological mapping, as for the motion detectors defined by Equations (1)–(5), or because they are directly linked to observable outputs, such as for the perceptual decision equations (10). However, some of the parameters have an elusive biological meaning or association with physiological recordings. This is particularly the case for the parameters defining motion integration. For those parameters we analyze the effects of systematic parameter variations on the model responses. This analysis allows us to better understand and characterize the model and to reduce the plausible range of parameter values, and so to constrain the search for exact values to fit observed empirical data from human/animal observers.

We applied dynamic systems techniques to analyze individual parameters by assessing the model for plausible conditions of operation. For example, we assume that the motion integration component of the model solves, at its scale, the aperture problem. This observation has been made in the visual cortex (Pack and Born, 2001; Pack et al., 2004), and is at the basis of a number of motion models. We selected stimuli, such as gratings moving behind different apertures and studied how the model with different parameter values behaves for such stimuli.

When a grating is moving behind a rectangular aperture, it induces a motion percept whose global direction follows the longer edges (Figure 5). In the intermediate state where all edges have equal length, multiple percepts can occur, either following one of the edges or their average direction (Castet et al., 1999; Fisher and Zanker, 2001; Meso and Masson, 2015). The literature suggests that such a multi-stable perception is the result of the motion integration mechanisms that select 2D local motion information coming from the aperture edges (Tlapale et al., 2010; Rankin et al., 2013, 2014). As such, we argue that a similar stimulus can ensure that our model properly disambiguates motion information.

Figure 5 illustrates possible local motion interpretations of one such stimulus, a moving grating viewed behind a diamond aperture. For this kind of stimulus, we have multiple 1D motion cues, and two sets of 2D motion cues. For a small aperture, MT neurons should be able to solve the aperture problem by selecting one of the two possible 2D directions.

We ran our motion integration model on a synthetic stimulus with its size matched to that of MT receptive fields until it reached a stable solution, for a given value of a model parameter such as γ which defines the non-linearity in Equation (7). Because the bifurcation analysis is

facilitated by time-invariant inputs, we bypassed the motion detection stage, directly providing a synthetic, yet plausible, directional response to the MT neurons. We obtained solutions for systematic variations of the parameters through numerical continuation, an efficient family of techniques to follow solutions across parameter changes and for which multiple techniques and software have been developed (Doedel, 1981; Salinger et al., 2001; Henderson, 2002; Dhooge et al., 2003).

Figure 6 illustrates the influence of parameter γ in the motion selection stage. For each value of the parameter, a number of solutions can exist, attainable by different initial conditions. Some solutions are unstable, in that a small perturbation in the system will switch to another stable solution, and are denoted by *dashed lines* on the figure. The stable solutions, more likely to be attained by temporal integration and observed in cell recordings, are represented by *solid lines.* In the vertical axis we show the ℓ^2-norm of the solutions, as a way to summarize their overall activity.

For large γ values, we obtained a single solution that consists of incoherent motion signals and is represented by the *red solid line* in Figure 6, and has a low overall activity (ℓ^2-norm). As γ decreases, this incoherent solution becomes unstable (*red dashed line*), and the two solutions based on the 2D motion information appear, indicating that the motion integration stage has been able to select them. Those 2D solutions are shown by the *blue lines* in the figure. Note that the two 2D solutions have the same activity in this symmetric stimulus, and thus overlap in overall activity scores.

The continuation analysis depicted in Figure 6 improves our understanding of the effect of the non-linearity defined in Equation (7) and used in Equation (6), and allows us to specify the range of parameter values for which the model would function appropriately. In this case, the analysis suggests that γ should be less than about 4.12 to avoid a region of unstable incoherent solutions that do not comport with observed behavior and physiology. Studying the other parameters gives us a set of parameter ranges for the model. In Figure 7 we performed a numerical continuation on parameter α in Equation (7), which suggested both a lower and upper bound for its parameter value range.

Having defined a set of parameters and parameter value ranges in which the model exhibits appropriate properties, we further refined the model by optimizing its output to best fit the experimental data of Watamaniuk et al. (1989); Watamaniuk and Sekuler (1992). Note that the comparison to the results of Kayser et al. (2009) shown in Figure 3 and Figure 4 were performed after the optimization stage. In particular, the following set of parameters were allowed to vary: a and σ in Equation (8), τ in Equation (10), W_i in Equation (11), and f and g in Equation (12). We summarize the parameter value ranges and the selected parameter values for our model in Table 1.

3. Results

We assess the model through comparison with several demanding published data sets, chosen because the parametric variations of stimulus variables that challenge the predictive range of the model. The experiments of Watamaniuk and Sekuler (1992) provide several threshold curves that allow us to test predictions of our model on motion percepts from

random dot kinematograms in three different stimulus domains: temporal, spatial and directional. The basic stimulus consists of a random dot kinematogram in which dot positions are updated with directions randomly selected from Gaussian distributions with given dispersions. Formally, the *n*-th position of the *i*-th dot is defined by

$$p_{i,n}=p_{i,n-1}+s\overrightarrow{\theta}_{i,n}, \mathbb{P}(\theta_{i,n})=G_{\sigma}(\theta_0), \quad (13)$$

where *s* is the speed of the random dots and $\mathbb{P}(\theta_{i,n})$ defines the probability density function of the random motion directions $\theta_{i,n}$.

During the experiments of Watamaniuk and Sekuler (1992), motion direction thresholds were estimated using a staircase procedure. For each trial, two stimuli were presented: one in which the random dots were moving upward (90°) on average, and one in which the random dots were moving slightly leftward (> 90°). The order of presentation of those upward and leftward stimuli was randomly chosen for each trial. Observers had to determine whether the second stimulus was moving leftward or rightward relative to the first one in the trial. Although the model could be extended to consider differences between cardinal and non-cardinal motion directions, here we rotated the stimuli to have true leftward and rightward stimuli.

In the first experiment of Watamaniuk and Sekuler (1992), the duration of the stimulus was varied. Observers exhibit a steep decrease in motion direction discrimination threshold as presentation time increases over the first 500 ms, after which the threshold approaches an asymptote. The effects of varying stimulus duration are shown for a wider (σ of 25.5°) and a narrower (σ of 4.3°) distribution of dot motion directions. The data provide strong constraints on the dynamics of neural populations in area LIP in the model. With the proper parameters, the model predictions replicate the behavioral data (Figure 8).

The joint modulations of motion distributions and temporal properties of the stimulus place strong joint constraints on the model, which was largely able to fit both qualitative and quantitative aspects of the data. There is some indication of over-prediction by the model in the 250–500 ms range of the 25.5° data. While it is possible that further specialization of the model parameters might accommodate these data, several individual data points in this range appear at variance with the wide range of asymptotic performance level and may represent behavioral noise; it is also possible that specialized heuristics may emerge in conditions with such large distributions of dot motion directions (see discussion of Figure 10 below).

In the second experiment of Watamaniuk and Sekuler (1992), motion direction discrimination threshold curves were measured by varying the diameter of the stimuli, which provides a test of the spatial extent of the receptive fields in the model. In Figure 9 we show that our model provides an excellent fit to the data for both wider (σ of 25.5°) and narrower (σ of 4.3°) motion direction distributions, correctly predicting both the respective threshold curves and the greater sensitivity to aperture size in the presence of wider range of random motion directions. The overall fit of the model for conditions jointly varying aperture diameter and breadth of the dot motion direction distributions is very close. The single point

of deviation at the largest aperture and 25.5° may again reflect noise in the data, or possible heuristic responses to the broad motion distributions.

In the third experiment of Watamaniuk and Sekuler (1992), the standard deviation of the random dot motion direction distribution is varied systematically, with σ taking eight possible values between 0° and 34°. In Figure 10 we show the predicted motion direction discrimination threshold of our model as a function of the standard deviation of the distribution of the random dot motion direction. These predictions are consistent with the data from one of the observers in the original experiment. The other two observers showed a peculiar trend not discussed in the original article: for the broadest random dot motion direction distributions, the threshold decreased from its maximum. Although our model is unable to reproduce this irregularity with the default feedforward connectivity, manually setting the decision units to be more sensitive to the extreme directions allows the model to reproduce this anomalous result. The exact mechanism underlying this pattern of behavior remains unknown, but our results suggest that it might correspond to a strategy that emphasize motion signals near the extreme values of the direction distributions.

In an experiment reported in an earlier paper (Watamaniuk et al., 1989), the same authors estimated motion direction discrimination threshold curves for random dot kinematograms in which the direction of each dot was constant during each trial, i.e. $\forall i \forall n \theta_{i,n} = \theta_{i,0}$, in Equation (13) for a set of standard deviations of the random dot motion direction distributions and display durations (from 150 ms to 1250 ms). Taking into account the other changes in the stimulus, our model is able to closely reproduce their results (Figure 11 and Figure 12). Figure 11 plots the motion direction discrimination threshold for different display durations as a function of the standard deviation of random dot motion direction, while Figure 12 plots the same data for different standard deviations of random dot motion direction distributions as a function of display duration. Although there may be small remaining deviations from the data, the model provides a remarkable quantitative account of the demanding interplay between duration and the breadth of the motion direction distributions.

4. Discussion

4.1. Motion detection

Motion integration models have taken a variety of approaches to represent local motion detection. At one extreme the input to later motion processing stages is defined abstractly based on plausible assumptions. For instance, in the Bayesian model of Weiss et al. (2002) the distribution of the response from motion detectors conserves luminance and is normal. The Bayesian models of Montagnini et al. (2007); Bogadhi et al. (2011) segregate motion information between 1D and 2D channels, each having normally distributed responses. As input to their neural fields model, Rankin et al. (2013) uses motion detector responses that are discrete in 1D space and direction domains. The neural fields decision models of Wang (2002); Wong and Wang (2006); Wong et al. (2007) also use a fixed static input.

Neural circuits for motion detection have been the focus of intense research. Among the first models, correlation detectors originally proposed to model the weevil visual system

(Hassenstein and Reichardt, 1956) have been generalized to other insects and proposed as a basis of human visual motion perception (van Santen and Sperling, 1984). Correlation detectors have the advantage of an intuitive description and a simple implementation. They also have been shown to be equivalent or similar to other commonly used motion energy models (van Santen and Sperling, 1985). There have been successful attempts at using correlation detectors as the input to larger motion integration models (Bayerl and Neumann, 2004, 2007; Tlapale et al., 2011). Yet, although the correlation models match the anatomical structure of other invertebrates, such a flies (Clark et al., 2011), their neural substrate in primates remains unclear.

To some extent equivalent to the correlation detectors, motion energy models (Adelson and Bergen, 1985) are a family of models inspired by linear systems. Compared to correlation detectors, or other neural models (Chey et al., 1997, 1998), several of the basic elements of motion energy models have been found in the vertebrate visual system. In particular, the recordings of De Valois and Cottaris (1998); De Valois et al. (2000) describe two V1 populations with monophasic and biphasic temporal responses, having properties similar to the motion filters in Adelson and Bergen (1985). Several large-scale integration models use motion energy as the initial stage (Simoncelli and Heeger, 1998; Tlapale et al., 2010), but some of their traditional features do not have a strong biological interpretation. For instance, motion detectors traditionally include a motion opponency term that inhibits neurons tuned to opposite motion directions (Reichardt, 1957). In motion energy models, motion opponency was added to account for purely perceptual recordings (Adelson and Bergen, 1985), and has since then been included in several V1 models (Bayerl and Neumann, 2004; Tlapale et al., 2011). Although motion opponency definitively exists in the primate visual system, evidence suggests that it does not occur at the level of V1, but more probably between V1 and MT (Qian and Andersen, 1994; Heeger et al., 1999). In our model, we considered motion opponency as a dendritical computation in MT (Equation (8)). Also, the proposed motion detection filters are causal (Equation (3)), i.e. they are not influenced by stimuli from the future, unlike some non-causal energy models that use Gaussian derivative temporal profiles (Simoncelli and Heeger, 1998; Tlapale et al., 2010).

In this article, we developed and tested the integrated dynamic motion (IDM) model, a multi-scale motion system for motion perception inspired by known architecture of the visual system, and used it to model the behavioral responses to one family of stimuli, namely high-contrast random dot kinematograms. Our intention is to apply the model to a wide range of visual motion stimuli. Some additional work is necessary. In particular, contrast normalization has proven to be an important factor in reproducing neural activity (Carandini et al., 1997; Lu and Sperling, 1996) and is still lacking in our model.

4.2. Dynamics

Although visual motion includes, by definition, a temporal component, a number of models in the motion integration literature do not incorporate dynamics in the predictions of their models. This simplification of the motion detection stage can facilitate the mathematical analysis of the models, where the stimulus is sometimes fixed or simplified (Rankin et al., 2013). Classical models of activity in areas V1 and MT, using either energy filters

(Simoncelli and Heeger, 1998) or simplified static inputs (Rust et al., 2006) also lack consideration of the underlying dynamics. At higher levels, Bayesian models have been used to model perception or decisions (Chalk et al., 2010; Sotiropoulos et al., 2011), but these models do not integrate dynamics in the perceptual process. One notable exception is the Bayesian models of Montagnini et al. (2007); Bogadhi et al. (2011), which modeled smooth pursuit eye movements from basic motion information.

Here, we proposed a dynamical model where the activity of neural populations is allowed to vary in time, similar to the neural field formulation (Wilson and Cowan, 1972, 1973; Amari, 1977). Allowing our system to be dynamic is a more realistic reflection of brain activities in which there is no final state, and there are temporal variations in neural activities (Chey et al., 1997, 1998; Bayerl and Neumann, 2004). It allows the definition of appropriate readouts of eye movements, perception (Tlapale et al., 2010) or response times and accuracies (Roitman and Shadlen, 2002). In this article, we described the effects of stimulus coherence on response times (see Figure 3 Figure 4) as determined by the dynamical decision aspect of the model. We also showed how the temporal dynamics of the decision stage are consistent with empirical investigations (Watamaniuk et al., 1989; Watamaniuk and Sekuler, 1992) (see Figure 8, Figure 11 and Figure 12), which was achieved by the high-level decision models of Wang (2002); Wong and Wang (2006); Wong et al. (2007), but not by classical motion integration models lacking a decision stage.

Since our model includes neural dynamics, it is also able to disambiguate motion information through subtractive inhibition Equation (9), as in Tlapale et al. (2010), whereas motion integration models lacking dynamics incorporate a divisive inhibition to account for the selection processes (Nowlan and Sejnowski, 1995; Simoncelli and Heeger, 1998; Rust et al., 2006; Busse et al., 2009). Although subtraction is a simple operation readily mapped to biological mechanisms in neurons, some properties of the divisive inhibition, such as contrast gain control, remain to be investigated in the current framework (see Carandini and Heeger (2012) for a review).

4.3. Motion selection

Subtractive inhibition has previously been incorporated into motion integration models and used to accommodate ambiguous signals generated by display apertures. For instance, with translating line stimuli, line endings generate unambiguous motion signals that reinforce the correct direction in MT receptive fields through feedforward connections, and ultimately inhibit the other directions, through lateral inhibition. Based on the same properties, our model is also able to solve the aperture problem, and reproduce the dynamics shown in Tlapale et al. (2010) in its account of smooth pursuit eye movements or perception.

In the case of noisy random dot kinematograms, high levels of noise from broad random dot motion direction distributions often lead the model to make incorrect decisions. This can be seen in Figure 10, where the average threshold and its standard deviation for random dot motion direction distributions with standard deviations greater than 20 degrees increases substantially. The important parameters that allow us to reproduce the empirical data are the strength of the lateral inhibition and the direction sensitivity in V1 and MT.

4.4. Motion integration

Neither models of neural activity in MT (Simoncelli and Heeger, 1998; Rust et al., 2006) nor high-level models (Wang, 2002; Wong and Wang, 2006; Wong et al., 2007) include a spatial dimension to model the limited receptive fields of their neurons. Although some spatial effects can be simulated, for instance by varying the number of local motion components in Montagnini et al. (2007); Bogadhi et al. (2011), they remain largely unable to handle spatially windowed stimuli. Models that consider spatial effects typically integrate motion across space using larger receptive fields in higher cortical areas, primarily in MT (Chey et al., 1997, 1998; Bayerl and Neumann, 2004). Here, we followed this standard implementation. Unlike Tlapale et al. (2010), we did not include a feedback mechanism from MT to V1. A subsampling of the number of neurons between V1 and MT is used to mimic the cortical magnification factor.

Our model represents the activity in neural populations, but we do not explicitly link the activity to the average firing rate, which is generally achieved by considering what is inside the non-linearity as currents, and using the sigmoid activation function to perform the transformation from current to firing rate, with mechanisms such as gain modulation directly incorporated into the non-linearity (Abbott and Chance, 2005). We use an abstract activity measure in this paper; further comparison to neurophysiological recordings would be facilitated by a direct mapping to firing rates.

4.5. Decision

In its formalism, our model follows existing decision models that represent accumulation of evidence by stochastic differential equations (Ratcliff, 1978; Usher and McClelland, 2001; Wong and Wang, 2006; Wong et al., 2007). Although those models have been shown to be closely related (Bogacz et al., 2006), our contribution is to integrate a general motion integration model with those decision mechanisms.

The inclusion of a decision mechanism also allows us to make predictions about response times on a wide range of stimulus variations. Accounting for accuracy, average response times, and response time distributions jointly is demanding. While we were able to show that the current model implementation is broadly consistent with the data, certain additional distributional assumptions may be needed to achieve full joint quantitative fits to data, similar to those used in stochastic models of decision (random drift models such as Ratcliff (1978)), where they can be necessary even when the models do not include predictive specifications from the stimulus but instead estimate evidence parameters. In its current form, the model only allows behavioral decisions between two possible choices, without a continued ongoing process for multiple choices. The model currently does not implement a mechanism that would cause alternation or switching between the solutions of the bifurcation diagram over time, as often occurs in the multistable stimulus presented in Figure 5 (Meso and Masson, 2015). Often, the switch between multistable perceptual states is considered a consequence of adaptation. Accounting for such phenomena would require a model elaboration that included an adaptation process.

4.6. Parameterization

Numerical continuations provide an efficient way to analyze the effect of parameters on a model. Here, we apply this method to study and constrain parameter value ranges, based on expected properties of the system. This allows us to offer a more detailed explanation of the set of parameters than is typical in the motion integration literature. Numerical continuations are already employed in computational neuroscience, in particular for models with a low dimensionality such as decision models (Machens et al., 2005; Wong et al., 2007). The computing power now available allowed us to apply it to larger models.

There are several challenges in using numerical methods to explore large neural models. First, numerical continuations naturally work on temporally invariant stimuli that have to be selected by the experimenter. Then, they are only able to continue solution branches from known points in the solution space, making it easy to miss disconnected solutions. Finally, the analysis of a large number of parameters simultaneously remains problematic.

Despite these limitations numerical continuations offer a useful descriptive analysis of the parameter value ranges, allow better model definitions, and give quantitative and qualitative descriptions of changes in the model, such as the ones induced by learning mechanisms.

Here, we showed that numerical and qualitative methods can be used symbiotically with standard methods of quantitative parameter estimation and fitting to provide very good to excellent fits of performance accuracy or threshold data in demanding experiments with parametric variations in the stimuli. Remaining small deviations between the predictions and the data may reflect noise or, in certain key stimulus conditions, use of simplifying heuristics. At the same time, any model is to some degree a simplification, and further developments may improve the ability of the model to account for other aspects of behavioral performance.

5. Conclusion

In this article we proposed an integrated dynamical motion (IDM) model of motion perception including several novel or classical mechanisms selected for their biological plausibility. We were able to test the resulting model on the exact stimuli defined by a family of random dot kinematograms in several behavioral experiments, showing that: 1) the selected mechanisms correctly interact together; 2) they are able to explain the experimental observations; 3) a precise validation of motion models can come directly from experimental data with sufficient number of stimulus variations. Moreover, the generality of our model facilitates further study of additional neural mechanisms, such as perceptual learning, and provides a structure within which future work may help to discriminate possible cortical locations.

Acknowledgments

This work was supported by the National Eye Institute Grant # EY-17491.

References

Abbott LF, Chance FS. Drivers and modulators from push-pull and balanced synaptic input. Progress in Brain Research. 2005; 149:147–55. [PubMed: 16226582]

Adelson EH, Bergen JR. Spatiotemporal energy models for the perception of motion. Journal of the Optical Society of America A. 1985; 2(2):284–99.

Amari S. Dynamics of pattern formation in lateral-inhibition type neural fields. Biological Cybernetics. 1977; 27(2):77–87. [PubMed: 911931]

Bayerl P, Neumann H. Disambiguating Visual Motion Through Contextual Feedback Modulation. Neural Computation. 2004; 16(10):2041–2066. [PubMed: 15333206]

Bayerl P, Neumann H. Disambiguating Visual Motion by Form-Motion Interaction—a Computational Model. International Journal of Computer Vision. 2007; 72(1):27–45.

Beck C, Neumann H. Interactions of motion and form in visual cortex - A neural model. Journal of Physiology – Paris. 2010; 104(1-2):61–70.

Berzhanskaya J, Grossberg S, Mingolla E. Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. Spatial Vision. 2007; 20(4):337–395. [PubMed: 17594799]

Bogacz R, Brown E, Moehlis J, Holmes P, Cohen JD. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. Psychological Review. 2006; 113(4):700–65. [PubMed: 17014301]

Bogadhi AR, Montagnini A, Mamassian P, Perrinet LU, Masson GS. Pursuing motion illusions: a realistic oculomotor framework for Bayesian inference. Vision Research. 2011; 51(8):867–80. [PubMed: 20974165]

Britten KH, Shadlen MN, Newsome WT, Movshon JA. Responses of neurons in macaque MT to stochastic motion signals. Visual Neuroscience. 1993; 10(6):1157–1169. [PubMed: 8257671]

Busse L, Wade AR, Carandini M. Representation of concurrent stimuli by population activity in visual cortex. Neuron. 2009; 64(6):931–42. [PubMed: 20064398]

Carandini M, Heeger DJ. Normalization as a canonical neural computation. Nature Reviews Neuroscience. 2012; 13(1):51–62.

Carandini M, Heeger DJ, Movshon JA. Linearity and normalization in simple cells of the macaque primary visual cortex. The Journal of Neuroscience. 1997; 17(21):8621–44. [PubMed: 9334433]

Castet E, Charton V, Dufour A. The extrinsic/intrinsic classification of two-dimensional motion signals with barber-pole stimuli. Vision Research. Mar; 1999 39(5):915–32. [PubMed: 10341945]

Chalk M, Seitz AR, Seriès P. Rapidly learned stimulus expectations alter perception of motion. Journal of Vision. 2010; 10(8):1–18.

Chey J, Grossberg S, Mingolla E. Neural dynamics of motion grouping : from aperture ambiguity to object speed and direction. Journal of the Optical Society of America A. 1997; 14(10):2570–2594.

Chey J, Grossberg S, Mingolla E. Neural dynamics of motion processing and speed discrimination. Vision Research. 1998; 38(18):2769–86. [PubMed: 9775325]

Clark DA, Bursztyn L, Horowitz MA, Schnitzer MJ, Clandinin TR. Defining the computational structure of the motion detector in Drosophila. Neuron. 2011; 70(6):1165–77. [PubMed: 21689602]

De Valois RL, Cottaris NP. Inputs to directionally selective simple cells in macaque striate cortex. Proceedings of the National Academy of Sciences of the United States of America. 1998; 95(24): 14488–93. [PubMed: 9826727]

De Valois RL, Cottaris NP, Mahon LE, Elfar SD, Wilson JA. Spatial and temporal receptive fields of geniculate and cortical cells and directional selectivity. Vision Research. 2000; 40(27):3685–702. [PubMed: 11090662]

de Vries B, Principe JC. A Theory for Neural Networks with Time Delays. Advances in Neural Information Processing Systems. 1991; 3

Deco G, Martí D. Deterministic analysis of stochastic bifurcations in multi-stable neurodynamical systems. Biological Cybernetics. 2007; 96(5):487–96. [PubMed: 17387505]

Deriche R. Fast algorithms for low-level vision. IEEE Transactions on Pattern Analysis and Machine Intelligence. 1990; 12(1):78–87.

Dhooge A, Govaerts W, Kuznetsov YA. MATCONT: A MATLAB package for numerical bifurcation analysis of ODEs. ACM Transactions on Mathematical Software. 2003; 29(2):141–164.

Doedel E. AUTO: A program for the automatic bifurcation analysis of autonomous systems. Congressus Numerantium. 1981; 30:265–284.

Escobar MJ, Masson GS, Viéville T, Kornprobst P. Action Recognition Using a Bio-Inspired Feedforward Spiking Network. International Journal of Computer Vision. 2009; 82(3):284–301.

Fisher N, Zanker JM. The directional tuning of the barber-pole illusion. Perception. 2001; 30:1321–1336. [PubMed: 11768487]

Hassenstein B, Reichardt W. Systemtheoretische Analyse der Zeit-, Reihenfolgen-, und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers Chlorophanus. Zeitschrift für Naturforschung. 1956; 11:513–524.

Heeger DJ, Boynton GM, Demb JB, Seidemann E, Newsome WT. Motion opponency in visual cortex. The Journal of Neuroscience. 1999; 19(16):7162–7174. [PubMed: 10436069]

Heeger DJ, Simoncelli EP, Movshon JA. Computational models of cortical visual processing. Proceedings of the National Academy of Sciences of the United States of America. 1996; 93(2): 623–627. [PubMed: 8570605]

Henderson ME. Multiple parameter continuation: Computing implicitly defined k-manifolds. International Journal of Bifurcation and Chaos. 2002; 12(3):451–476.

Kayser AS, Buchsbaum BR, Erickson DT, D'Esposito M. The Functional Anatomy of a Perceptual Decision in the Human Brain. Journal of Neurophysiology. 2009; 103(3):1179–1194. [PubMed: 20032247]

Lorenceau J, Shiffrar M, Wells N, Castet É. Different motion sensitive units are involved in recovering the direction of moving lines. Vision Research. 1993; 33(9):1207–17. [PubMed: 8333170]

Lu ZL, Dosher BA. External noise distinguishes attention mechanisms. Vision Research. 1998; 38(9): 1183–1198. [PubMed: 9666987]

Lu ZL, Dosher BA. Characterizing human perceptual inefficiencies with equivalent internal noise. Journal of the Optical Society of America A. 1999; 16(3):764–778.

Lu ZL, Dosher BA. Characterizing observers using external noise and observer models: assessing internal representations with external noise. Psychological Review. 2008; 115(1):44–82. [PubMed: 18211184]

Lu ZL, Sperling G. Contrast gain control in first- and second-order motion perception. Journal of the Optical Society of America A. 1996; 13(12):2305–2318.

Machens CK, Romo R, Brody CD. Flexible control of mutual inhibition: a neural model of two-interval discrimination. Science. 2005; 307(5712):1121–1124. [PubMed: 15718474]

Mazurek ME, Roitman JD, Ditterich J, Shadlen MN. A Role for Neural Integrators in Perceptual Decision Making. Cerebral Cortex. 2003; 13(11):1257–1269. [PubMed: 14576217]

Meso AI, Masson GS. Dynamic resolution of ambiguity during tri-stable motion perception. Vision Research. 2015; 107:113–123. [PubMed: 25555566]

Montagnini A, Mamassian P, Perrinet L, Castet É, Masson GS. Bayesian modeling of dynamic motion integration. Journal of Physiology – Paris. 2007; 101(1-3):64–77.

Nowlan SJ, Sejnowski TJ. A selection model for motion processing in area MT of primates. The Journal of Neuroscience. 1995; 15(2):1195–1214. [PubMed: 7869094]

Pack CC, Born RT. Temporal dynamics of a neural solution to the aperture problem in visual area MT of macaque brain. Nature. 2001; 409(6823):1040–2. [PubMed: 11234012]

Pack CC, Gartland AJ, Born RT. Integration of Contour and Terminator Signals in Visual Area MT of Alert Macaque. The Journal of Neuroscience. Mar; 2004 24(13):3268–80. [PubMed: 15056706]

Principe JC, de Vries B, Guedes de Oliveira P. The Gamma Filter – A new class of adaptive IIR filters with restricted feedback. IEEE Transactions on SIgnal Processing. 1993; 41(2):649–656.

Qian N, Andersen RA. Transparent motion perception as detection of unbalanced motion signals. II. Physiology. The Journal of Neuroscience. 1994; 14(12):7367–7380. [PubMed: 7996182]

Rankin J, Meso AI, Masson GS, Faugeras O, Kornprobst P. Bifurcation study of a neural field competition model with an application to perceptual switching in motion integration. Journal of Computational Neuroscience. 2014; 36:193–213. [PubMed: 24014258]

Rankin J, Tlapale E, Veltz R, Faugeras O, Kornprobst P. Bifurcation analysis applied to a model of motion integration with a multistable stimulus. Journal of Computational Neuroscience. 2013; 34(1):103–24. [PubMed: 22870848]

Rao RPN. Bayesian computation in recurrent neural circuits. Neural Computation. 2004; 16(1):1–38. [PubMed: 15006021]

Ratcliff R. A theory of memory retrieval. Psychological Review. 1978; 85(2):59–108.

Raudies F, Mingolla E, Neumann H. A model of motion transparency processing with local center-surround interactions and feedback. Neural Computation. 2011; 23(11):2868–914. [PubMed: 21851277]

Reichardt W. Autokorrelations-auswertung als Funktionsprinzip des Zentral-nervensystems. Zeitschrift für Naturforschung. 1957; 12(b):448–457.

Roitman JD, Shadlen MN. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. The Journal of Neuroscience. 2002; 22(21):9475–9489. [PubMed: 12417672]

Rust NC, Mante V, Simoncelli EP, Movshon JA. How MT cells analyze the motion of visual patterns. Nature Neuroscience. 2006; 9(11):1421–31.

Salinger AG, Pawlowski RP, Romero LA. Scalable bifurcation analysis algorithms for large parallel applications. Computational Fluid and Solid Mechanics. 2001:1647–1650.

Simoncelli EP, Heeger DJ. A model of neuronal responses in visual area MT. Vision Research. 1998; 38(5):743–761. [PubMed: 9604103]

Snowden RJ, Treue S, Erickson RG, Andersen RA. The response of area MT and V1 neurons to transparent motion. The Journal of Neuroscience. 1991; 11(9):2768–85. [PubMed: 1880548]

Sotiropoulos G, Seitz AR, Seriès P. Perceptual learning in visual hyperacuity: A reweighting model. Vision Research. 2011; 51(6):585–99. [PubMed: 21316384]

Stone M. Models for choice-reaction time. Psychometrika. 1960; 25(3):251–260.

Tlapale, E.; Kornprobst, P.; Masson, GS.; Faugeras, O. A Neural Field Model for Motion Estimation. In: Bergounioux, M., editor. Mathematical Image Processing Vol 5 of Springer Proceedings in Mathematics. Vol. Ch. 9. Springer Berlin; Heidelberg: 2011. p. 159-179.

Tlapale E, Masson GS, Kornprobst P. Modelling the dynamics of motion integration with a new luminance-gated diffusion mechanism. Vision Research. 2010; 50(17):1676–1692. [PubMed: 20553965]

Usher M, McClelland JL. The time course of perceptual choice: The leaky, competing accumulator model. Psychological Review. 2001; 108(3):550–592. [PubMed: 11488378]

van Santen JP, Sperling G. Temporal covariance model of human motion perception. Journal of the Optical Society of America A. 1984; 1(5):451–73.

van Santen JP, Sperling G. Elaborated Reichardt detectors. Journal of the Optical Society of America A. 1985; 2(2):300–321.

Viéville T, Chemla S, Kornprobst P. How do high-level specifications of the brain relate to variational approaches? Journal of Physiology – Paris. 2007; 101(1-3):118–35.

Wang XJ. Probabilistic decision making by slow reverberation in cortical circuits. Neuron. 2002; 36(5):955–968. [PubMed: 12467598]

Watamaniuk SN, Sekuler R. Temporal and spatial integration in dynamic random-dot stimuli. Vision Research. 1992; 32(12):2341–2347. [PubMed: 1288010]

Watamaniuk SN, Sekuler R, Williams DW. Direction perception in complex dynamic displays: the integration of direction information. Vision Research. 1989; 29(1):47–59. [PubMed: 2773336]

Weiss, Y.; Adelson, EH. Tech Rep 1624. Massachusetts Institute of Technology; 1998. Slow and Smooth : a Bayesian theory for the combination of local motion signals in human vision.

Weiss Y, Simoncelli EP, Adelson EH. Motion illusions as optimal percepts. Nature Neuroscience. 2002; 5(6):598–604.

Wilson HR, Cowan JD. Excitatory and inhibitory interactions in localized populations of model neurons. Biophysical Journal. 1972; 12(1):1–24. [PubMed: 4332108]

Wilson HR, Cowan JD. A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. Kybernetik. 1973; 13(2):55–80. [PubMed: 4767470]

Wohrer, A. Ph D thesis. Université Nice Sophia Antipolis; 2008. Model and large-scale simulator of a biological retina, with contrast gain control.

Wong KF, Huk AC, Shadlen MN, Wang XJ. Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. Frontiers in Computational Neuroscience. 2007; 1(6):1–11. [PubMed: 18946523]

Wong KF, Wang XJ. A recurrent network mechanism of time integration in perceptual decisions. The Journal of Neuroscience. 2006; 26(4):1314–1328. [PubMed: 16436619]

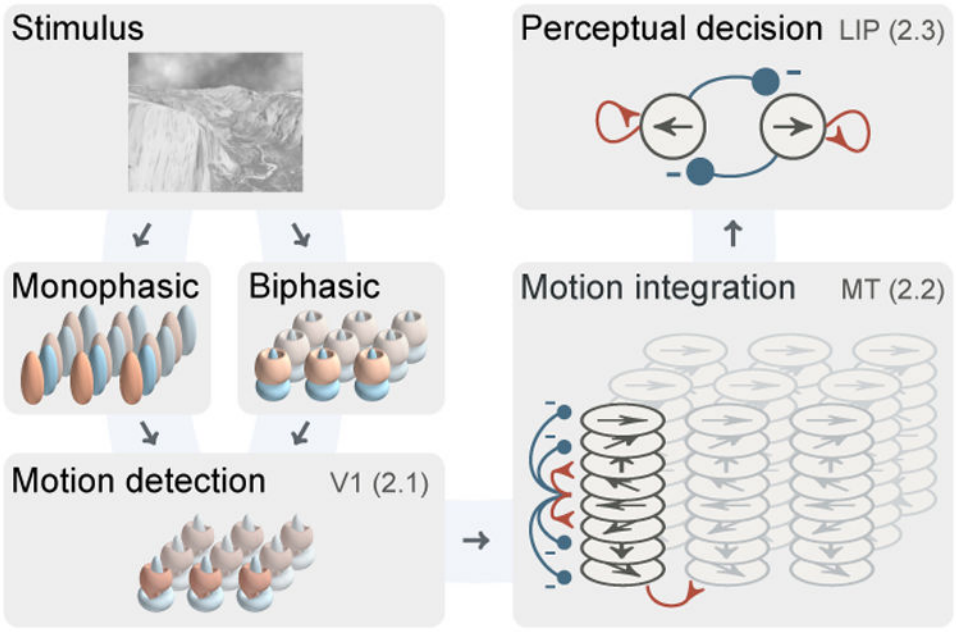

Figure 1.

Overview of the mechanisms of the integrated dynamical motion (IDM) model and their corresponding putative cortical locations. The motion detection stage includes two non-directional subpopulations, with biphasic and monophasic temporal responses, that are combined by directional neurons. Neurons in the motion integration stage implement motion opponency in their dendrites and pool motion information across larger portions of the visual field to solve the aperture problem at the scale of MT receptive fields. The perceptual decision stage aggregates motion information from different units over time to discriminate between possible percepts. All the populations except the decision stage are spatialized.

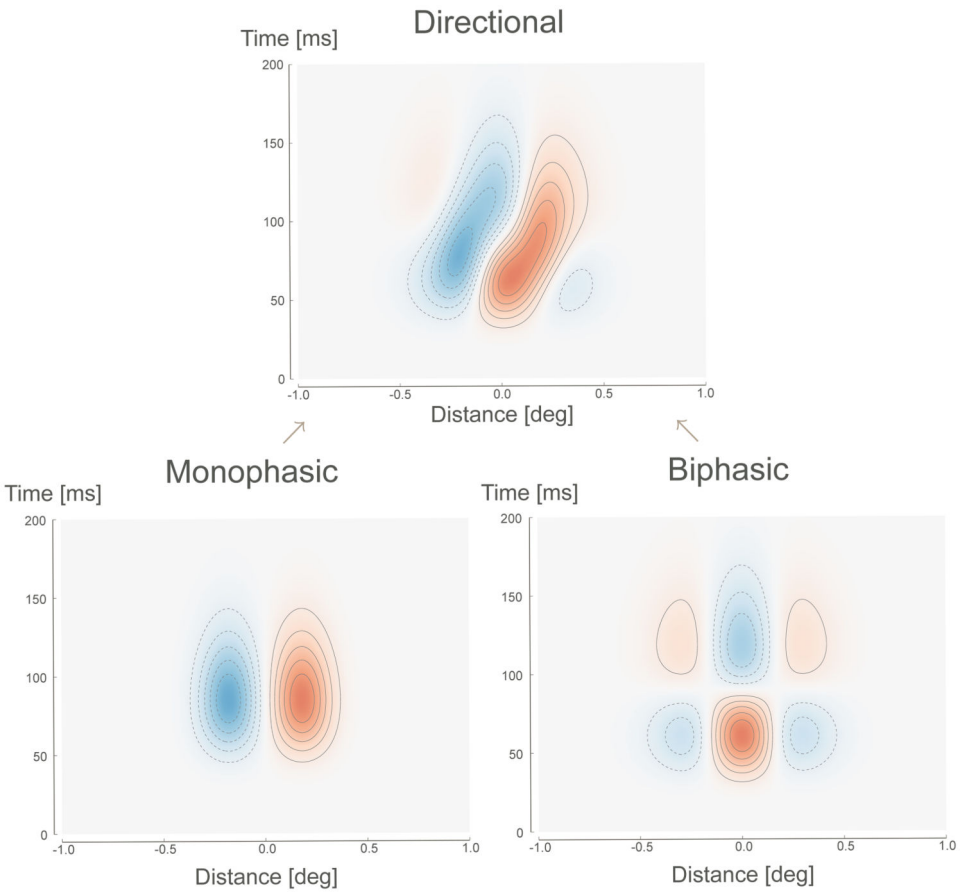

Figure 2.

2D cross-sections of the spatiotemporal receptive fields of V1 neurons. The model combines information from two non-directional subpopulations (*lower plots*) to generate directional response (*upper plot*). Dotted and solid contour lines denote inhibitory (*blue*) and excitatory (*red*) regions respectively.

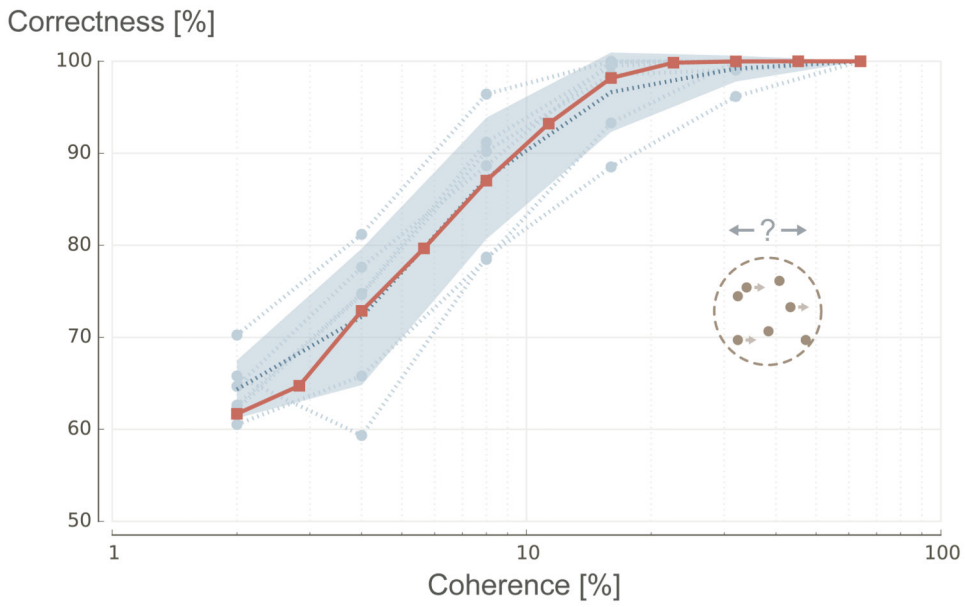

Figure 3.

Accuracy of the IDM model on the left/right discrimination task of Kayser et al. (2009). The *red squares* represent the average percent correct of the model over 5120 trials, varying the coherence level of the RDK stimulus. The original data for each subject as well as the average and standard deviation across subjects are reproduced here by the *blue circles*.

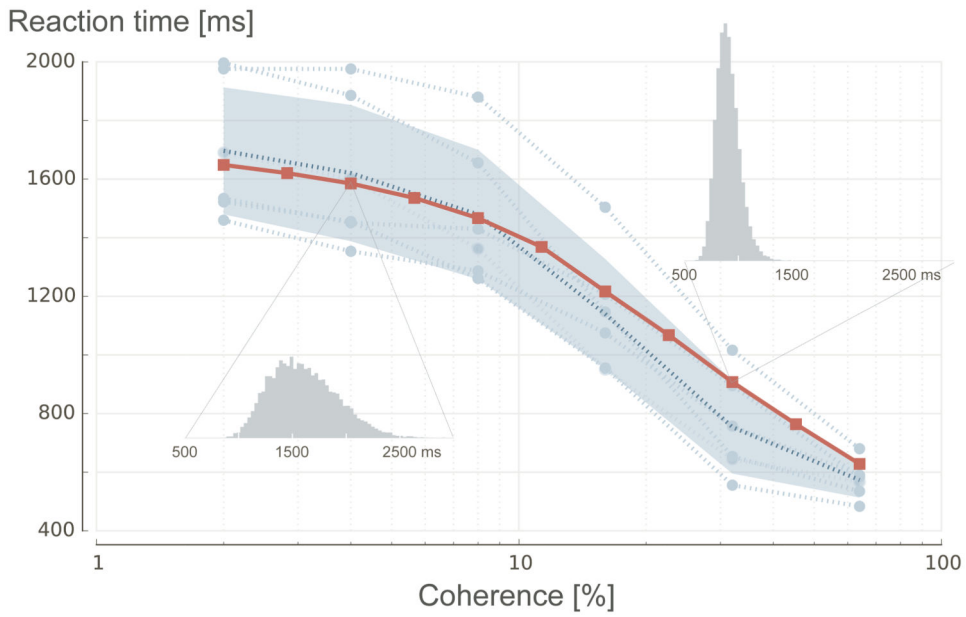

Figure 4.

Reaction time averages of the IDM model on the left/right discrimination task of Kayser et al. (2009). The reaction times of the model are represented by the *filled red squares*, while the subjects data, average and standard deviation are shown by the *blue circles*. Sampled reaction time histograms are shown for the 4% and 32% coherence levels.

Figure 5.

Ambiguities generated by local motion information. When a grating is moving behind a rectangular aperture, the perceived motion follows the elongated edge (the barberpole illusion). When both edges have the same length, the stimulus is ambiguous (Castet et al., 1999; Fisher and Zanker, 2001; Meso and Masson, 2015) and the perceived motion direction can either follow one of the edges or their average. Considering the local motion cues generated by such a stimulus, we can distinguish two zones: on the border the (2D) motion information is unambiguous and follow the edges (*green and orange regions*); in the center the (1D) motion information is ambiguous and can match any direction on a half circle (*yellow region*).

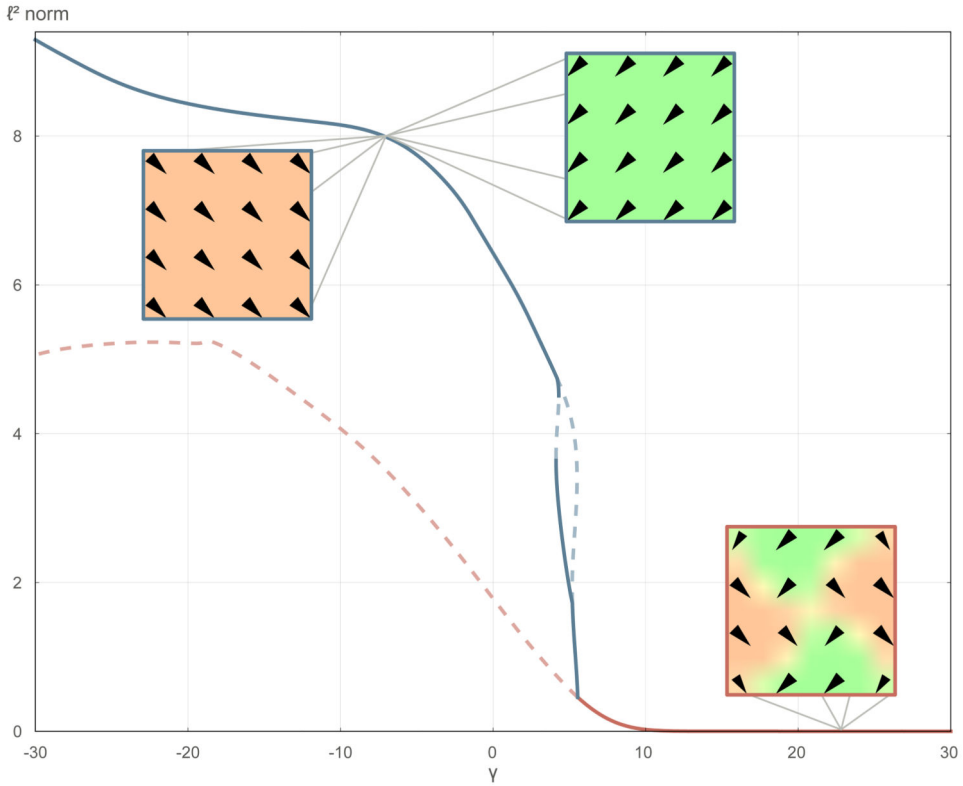

Figure 6.

Numerical continuation of the parameter γ in Equation (7) on the stimulus illustrated schematically in Figure 5. In the *solid and dashed blue* solutions, with larger activities, one of the two 2D motion directions is selected. In the *red* solutions, an incoherent aggregate of motion directions is obtained. For the motion selection mechanisms in MT to work correctly, γ must either be a small positive or a negative number. The *dashed lines* denote unstable solutions, differentiating them from the stable solutions represented by *solid lines*. At each single point on the curves, the solution is a neuron population where each cell is tuned for a given direction, and has a receptive field centered at a specific retinotopic position. We show a spatialized representation for three different solutions in the figures with arrows.

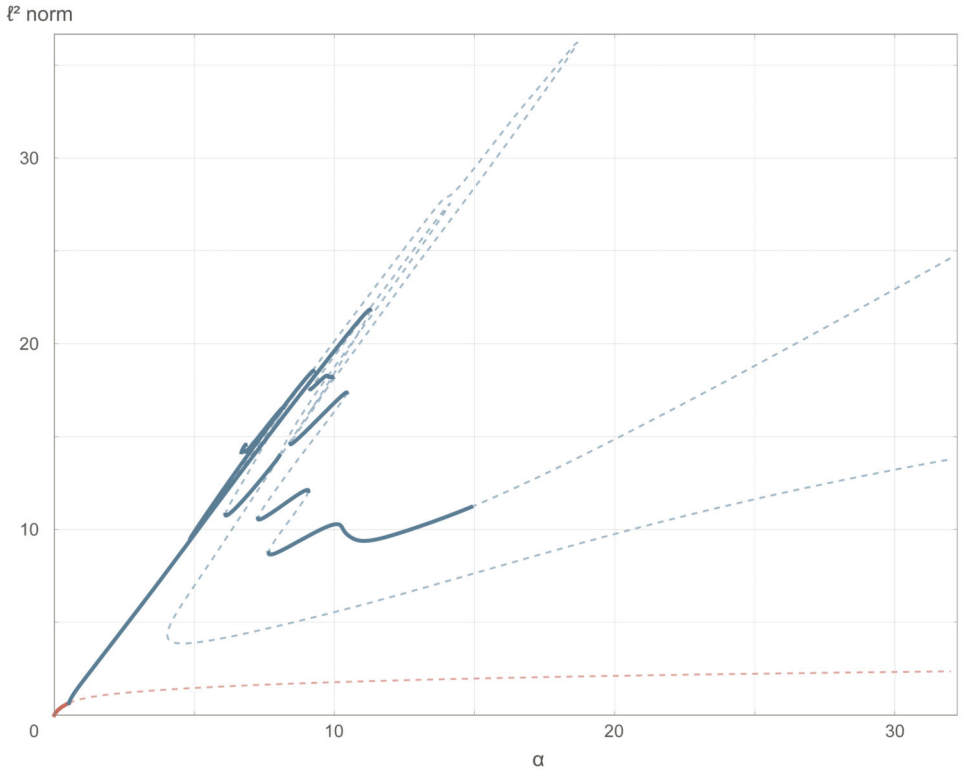

Figure 7.

Numerical analysis of varying parameter a in Equation (7), computed on the stimulus schematically illustrated in Figure 5. There are stable solutions only within a restricted range of a values, allowing us to define the lower and upper bounds for the parameter.

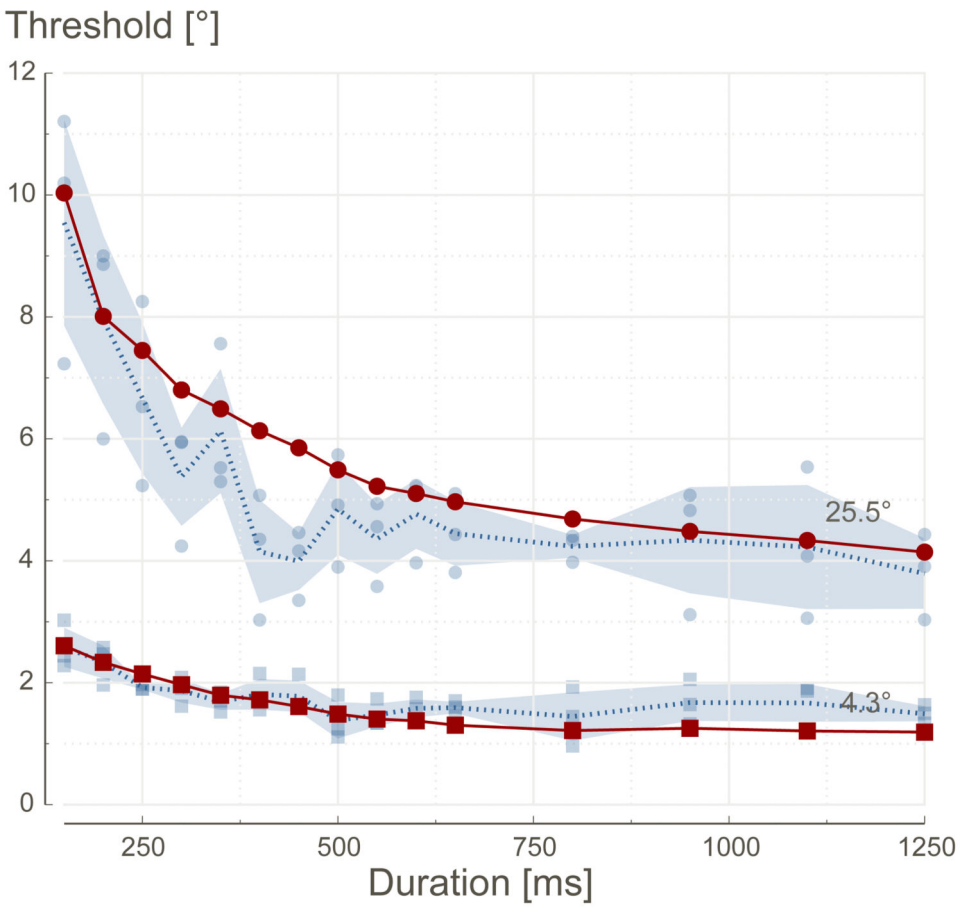

Figure 8.

Direction threshold [°] as a function of stimulus presentation time for two random dot motion direction distributions with σ of 25.5 and 4.3°, respectively, plotted as disks and rectangles. The predictions of the IDM model (*solid red*) are compared to the average data of Experiment 1 in Watamaniuk and Sekuler (1992), represented by the *empty blue markers* with standard deviation shown as a *blue ribbon*. Individual subject data of the original experiment are shown as *light blue markers*. Root mean square (RMS) errors of the model predictions for the two conditions are 0.91 and 0.23°, respectively.

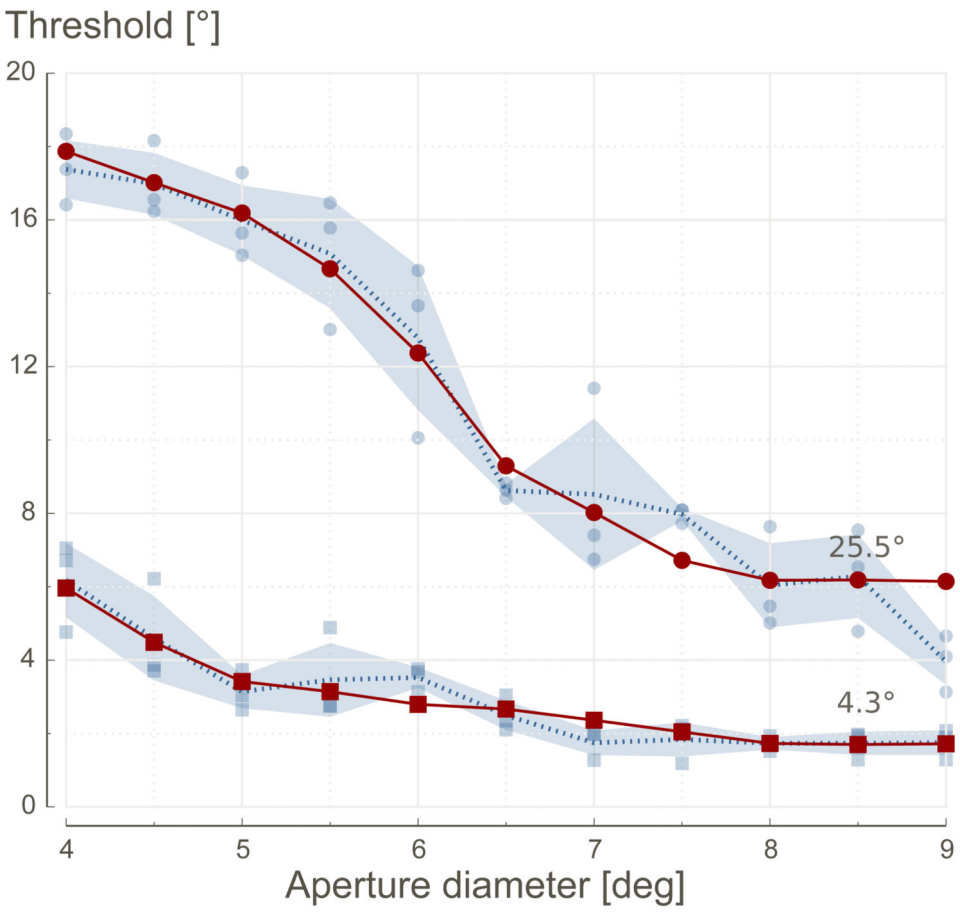

Figure 9.

Motion direction discrimination threshold [°] as a function of stimulus aperture diameter for two random dot motion direction distributions. Markers and symbols follow the convention of Figure 8. Data are from Experiment 2 of Watamaniuk and Sekuler (1992). RMS errors of the model predictions for the two conditions are 0.84° and 0.33°, respectively.

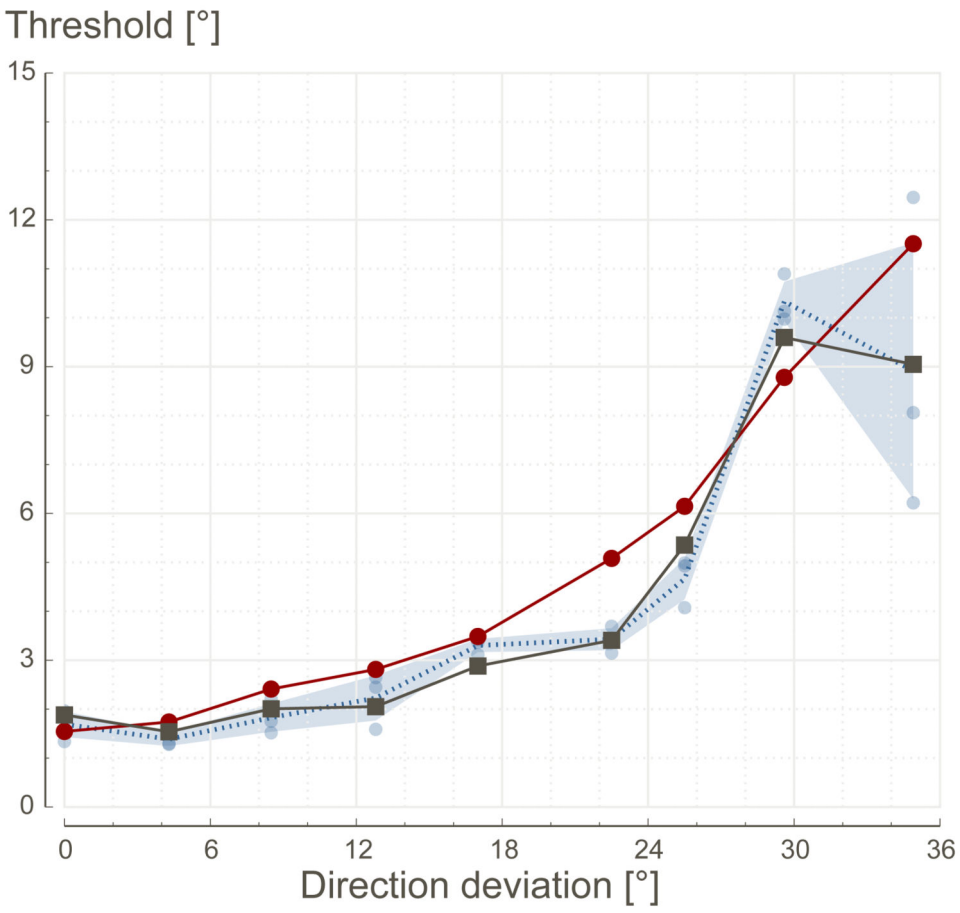

Figure 10.

Motion direction discrimination threshold [°] as a function of the standard deviation of the random dot motion direction distribution. We identified several mechanisms that are able to reproduce the data. For highly tuned direction bandwidths, mostly in the motion detectors, the model (*solid red circles*) globally follows pattern of data from the original experiment in Watamaniuk and Sekuler (1992) (*empty blue circles*). Replacing the constraints on direction tuning by a stronger selection mechanism in MT, we obtain a steep slope for larger standard deviations or random dot motion direction distributions, closer to the original experiment, but without the decrease for the largest $\sigma = 34.9°$. We are able to generate this characteristic by favoring the outlier directions of the broadest motion direction discrimination in the feedforward connectivity to the decision unit (*solid dark squares*). Respective RMS errors between the model and data are 1.29° and 0.39°.

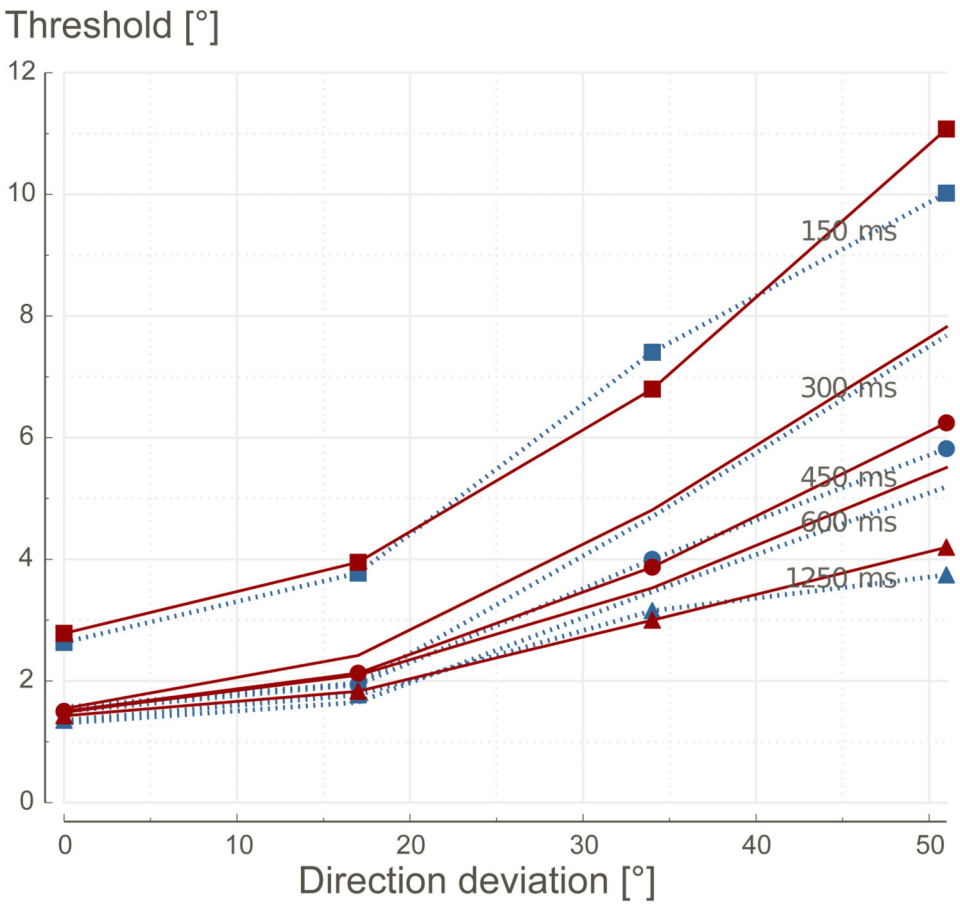

Figure 11.

Motion direction discrimination threshold [°] as a function of the standard deviation of random dot motion direction distribution for five different presentation times. Model predictions in red are compared to the original data from the experiment (*blue markers*) in Watamaniuk et al. (1989). The average root mean square error between the model and data across all conditions is 0.33°.

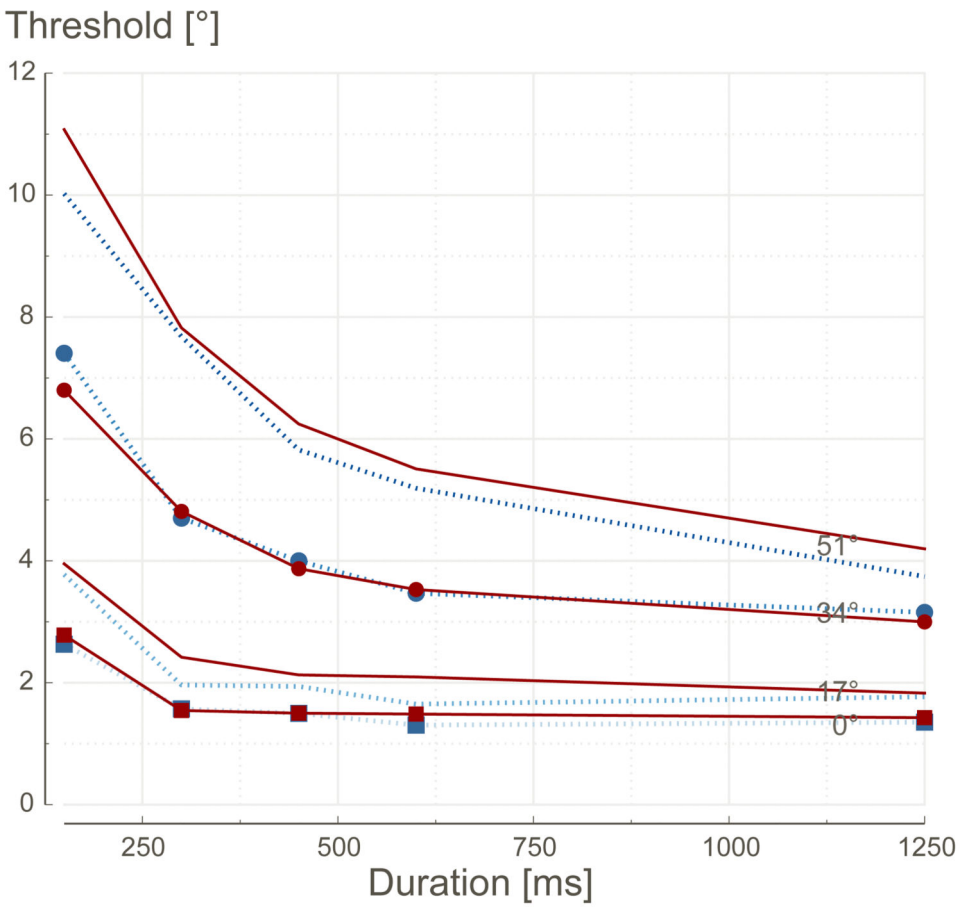

Figure 12.

An alternative representation of Figure 11 inverting direction distribution and presentation time. The data correspond to those in Figure 2 of Watamaniuk et al. (1989).

Table 1

Values defining the model with associated acceptable parameter ranges, as constrained by physiology and numerical analysis.

Eq.	Parameters	Description
(2)	$n_M = 11$ $\tau_M = 85$ms $a = 0.18°$ $\sigma_M = 0.1°$	From De Valois et al. (2000). This is a single population parameterization that is varied to generate responses for different directions, multiple speeds and scales.
(5)	$n_B = 8$ $\tau_B = 85$ms $\sigma_B = 0.15°$ $b = 0.75°$ $n'_B = 10$ $\tau'_B = 95$ms $\sigma'_B = 0.2°$	(idem)
(7)	$\alpha_A = 3$ $\beta_A = 2$ $\gamma_A = 2$	Constrained by a continuation analysis on selected set of stimuli to bound parameter values. Constrained parameter ranges are $0.60 < \alpha < 5.00$ (see Figure 7), $\beta > 1.15$, and $\gamma < 4.12$ (see Figure 6).
(8)	$c = 6.74$ $d = 0$ $\sigma_A = 3.11°$	Continuation analysis on α for motion selectivity, and then by a fit to the experimental data of Watamaniuk et al. (1989); Watamaniuk and Sekuler (1992). Motion opponency defined by d was not considered in the current article (and S is set to be the identity). The larger receptive fields of MT neurons were set to fit the experimental data.
(9)	e=100	Constrained by continuation analysis
(11)	$W_i(\theta) = \max(0, \sin(\theta - \theta_i))$	Where θ_i is the preferred direction corresponding to decision i. This parameter set is modified for the alternative results of Figure 10. Set so that the simulation results on random dot kinematograms of Section 3 fit the experimental data.
(12)	$f = 0.92$ $g = 0.13$	Set so that the simulation results on random dot kinematograms of Section 3 fit the experimental data.