PERGAMON

Vision Research 42 (2002) 2781–2797

Vision Research

www.elsevier.com/locate/visres

Orientation dependent modulation of apparent speed: a model based on the dynamics of feed-forward and horizontal connectivity in V1 cortex

Peggy Seriès [a,b,*], Sébastien Georges [a,b], Jean Lorenceau [a,b], Yves Frégnac [a]

[a] *Unité de Neurosciences Intégratives et Computationnelles UPR 2191 CNRS, 1 Av. de la Terrasse, Gif sur Yvette, 91198 Cedex, France*
[b] *Laboratoire de la Physiologie de la Perception et de l'Action, UMR 9950 CNRS—Collège de France, 11 Place Marcelin Berthelot, Paris, 75231 Cedex, France*

Received 10 April 2002; received in revised form 24 July 2002

Abstract

Psychophysical and physiological studies suggest that long-range horizontal connections in primary visual cortex participate in spatial integration and contour processing. Until recently, little attention has been paid to their intrinsic temporal properties. Recent physiological studies indicate, however, that the propagation of activity through long-range horizontal connections is slow, with time scales comparable to the perceptual scales involved in motion processing. Using a simple model of V1 connectivity, we explore some of the implications of this slow dynamics. The model predicts that V1 responses to a stimulus in the receptive field can be modulated by a previous stimulation, a few milliseconds to a few tens of milliseconds before, in the surround. We analyze this phenomenon and its possible consequences on speed perception, as a function of the spatio-temporal configuration of the visual inputs (relative orientation, spatial separation, temporal interval between the elements, sequence speed). We show that the dynamical interactions between feed-forward and horizontal signals in V1 can explain why the perceived speed of fast apparent motion sequences strongly depends on the orientation of their elements relative to the motion axis and can account for the range of speed for which this perceptual effect occurs (Georges, Seriès, Frégnac and Lorenceau, this issue).

Keywords: Striate cortex; Lateral interactions; Long-range horizontal connections; Apparent speed

1. Introduction

Our understanding of how the brain processes visual inputs has long relied on the basic concept of neurons with spatially limited receptive fields (RFs), "blind" to remote influences. This view has recently been challenged by physiological studies showing that the responses of V1 neurons to oriented stimuli presented within their RF can be markedly modulated by stimuli falling in surrounding "silent" regions which by themselves fail to activate the cell (review in Fitzpatrick, 2000; Frégnac & Bringuier, 1996). Whether this contextual influence is facilitatory or suppressive depends on the contrast and on the spatial configuration (orientation, alignment) of the pattern elements inside and outside the RF. At the anatomical level, these influences are supposed to be mediated by feedback projections from higher cortical areas, and by long-range horizontal (LH) connections within V1 (review in Gilbert, Das, Kapadia, & Westheimer, 1996). These connections link regions over several millimeters, tend to connect cells with similar orientation preferences, and more specifically, cells whose RFs are topographically aligned along an axis of collinearity (in cat: Schmidt, Goebel, Löwel, & Singer, 1997; tree shrew: Bosking, Zhang, Schofield, & Fitzpatrick, 1997; monkey: Sincish & Blasdel, 2001).

The highly specific architecture of LH connections and the activity they relay over large regions of the visual field suggested that they may be important for the processing of visual contours. Psychophysical studies (review in Hess & Field, 1999) have reported strong facilitatory interactions among iso-oriented collinear

* Corresponding author. Present address: Department of Brain and Cognitive Science, University of Rochester, Meliora Hall, Rochester, NY 14627, USA.
E-mail addresses: peggy.series@iaf.cnrs-gif.fr (P. Seriès), yves.fregnac@iaf.cnrs-gif.fr (Y. Frégnac).

0042-6989/02/$ - see front matter
PII: S0042-6989(02)00302-4

elements, whereas weak facilitation or suppressive interactions were found for iso-oriented parallel configurations. These interactions decrease with the distance or the orientation difference between the inducing elements. Altogether, these findings yielded the notion of a perceptual "association field" (Field, Hayes, & Hess, 1993) whose characteristics closely resemble the physiological and anatomical properties of LH connections, suggesting that both are related.

To date, most psychophysical and physiological studies on center/surround interactions have used displays where center and surround stimuli were presented simultaneously. Not much is known about the temporal characteristics of contour integration or about the dynamics of center/surround modulations of V1 responses. However, recent imaging studies in monkey V1 (Grinvald, Lieke, Frostig, & Hildesheim, 1994) and intracellular recordings in cat area 17 (Bringuier, Chavane, Glaeser, & Frégnac, 1999) have shown that propagation of activity though LH connections is much slower (0.05–0.5 m/s) than that observed along feed-forward (FF) and feedback connections (3–20 m/s). Bringuier et al. (1999) reported that a focal pulse-like visual stimulation outside the RF elicits a depolarization of the neuron's membrane potential whose onset occurs after a temporal delay that depends linearly on the distance between the focal stimulation and the RF, and can be as long as 50 ms. These delays are comparable to perceptual time scales, raising the possibility that the dynamics of center/surround modulations may have perceptual counterparts.

In our companion paper (Georges, Seriès, Frégnac, & Lorenceau, this issue), we reported that apparent motion sequences appear faster when the visual elements they contain are aligned with the motion path (collinear sequences) than when they are at an angle with it (parallel sequences). This effect is particularly large for speeds in range [40–96°/s] and peaks around 64°/s.

We here investigate whether this perceptual bias could reflect the dynamics of center/surround modulations in V1 cortex. We reason that spreading activity through LH connections evoked by a first stimulus may modulate the dynamics—and in particular the latency—of the neuronal responses to a second stimulus, presented from a few milliseconds to a few tens of milliseconds later, at neighboring positions in visual field. Because LH projections tend to connect iso-oriented iso-aligned RFs, these modulations are expected to primarily affect sequences of collinear elements. The differential latency modulations of successively activated V1 cortical units may then bias the response of their MT target neurons, resulting in an overestimation of sequence speed for shorter delays.

To investigate this possibility and provide a conceptual framework that links physiology and perception, we have developed a simple two-stages model. The first stage captures the basic dynamical properties of V1 cortical cells' responses to FF inputs and of activity through LH connections. A simplified MT-like stage processes speed by "reading-out" the spatio-temporal correlation of V1 responses.

This paper is organized as follows. We first analyze the behavior of the V1 stage, in response to sequences of brief and non-overlapping oriented stimuli. This model exhibits latency modulations that are selective to particular spatio-temporal configurations (orientation, speed) of the visual inputs. We analyze the dependency of these configurations on model parameters, when sequence speed is varied by controlling the temporal interval or the spatial interval between the subsequent stimuli. We show that these two versions of the model require different assumptions and lead to different predictions (Section 3). We then investigate how V1 latency modulations could affect speed processing at the MT stage (Section 4) and compare the predictions of the full model with the psychophysical results presented in our companion paper. We show that such a simple "input summation" mechanism is sufficient to account for our data (Section 5). We finally propose further psychophysical and physiological experiments that might be used to test or further extend the validity of the model.

2. The model

The V1 model (Fig. 1) was designed to be as simple as possible while capturing the basic dynamical properties of cortical cell responses and of activity through LH connections. It contains an array of N visual cells regularly spaced in cortex, with the same preferred orientation and non-overlapping RFs (i.e. which belong to distinct hypercolumns). Their RFs are either aligned along their orientation preference axis or orthogonal to it. These cells interact via LH connections, characterized by a slow speed of propagation. Each cell is described as a low-pass linear filter (RC circuit), with a membrane potential $v(t)$ obeying:

$$C\frac{\mathrm{d}v(t)}{\mathrm{d}t} = -\frac{v(t)}{R} + I(t) \tag{1}$$

where C is the membrane capacitance, R is the membrane resistance and $I(t)$ are the synaptic currents arriving at the cell's soma. These are described by the linear summation of FF and horizontal synaptic inputs:

$$I(t) = I_{\mathrm{f}}(t) + I_{\mathrm{h}}(t) \tag{2}$$

$I_{\mathrm{f}}(t)$ represents the compound synaptic current evoked by the activation of the FF pathway, when an optimal

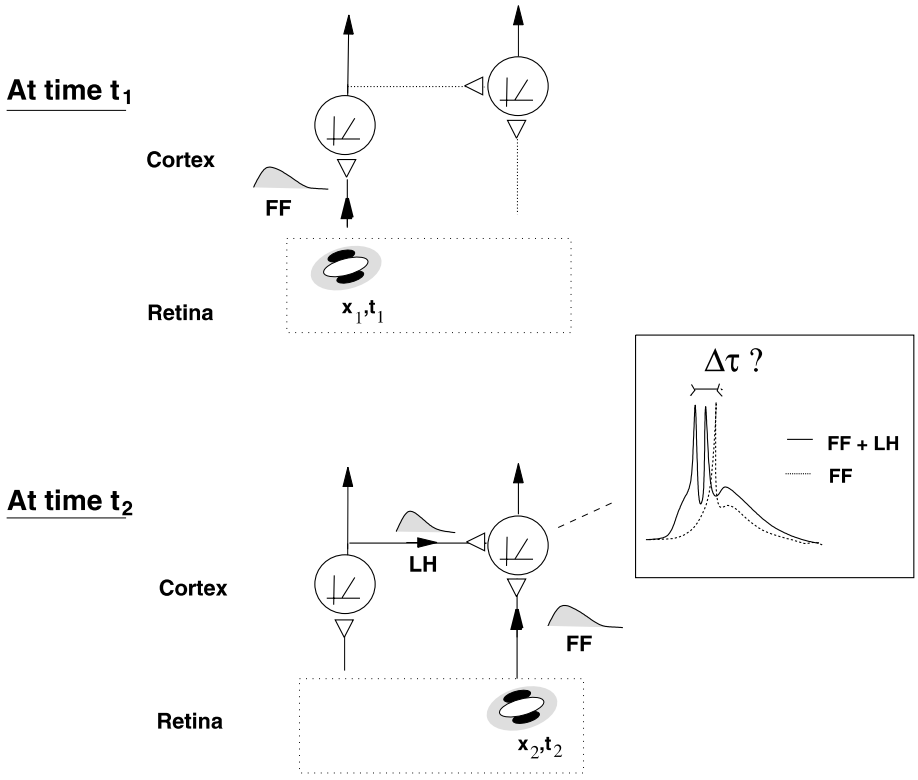

Fig. 1. Cartoon of the V1 model, which represents an array of cortical units (linear low pass filters followed by a rectification) that have the same preferred orientation and non-overlapping RFs. Units that have collinear RFs interact through LH connections. The response of each unit evokes a wave of sub-threshold horizontal activity that slowly propagates in cortex. Our work is based on the hypothesis that, for particular spatio-temporal configurations of the visual inputs, LH and FF inputs temporally overlap, which results in a modulation of response latency (inset).

stimulus is presented within the cell's RF. $I_h(t)$ represents the compound synaptic current evoked by the activation of the LH pathway. $I_h(t)$ conveys visual information from outside the classical RF and only exerts a modulatory influence on the target cell. As the model is used to simulate responses to very brief stimuli, these signals are modeled by two single α-functions, triggered at time ξ_f and ξ_h, and defined by $I_f(t) = A_f f(t - \xi_f, \tau_f)$ and $I_h(t) = A_h f(t - \xi_h, \tau_h)$ where:

$$f(t,\tau) = \begin{cases} \frac{t}{\tau} e^{-(t/\tau)} & \text{if } t \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

A_f and A_h denote the amplitude of the input currents and τ_f and τ_h their time-constant. Parameters ξ_f and ξ_h depend on the spatio-temporal configuration of the visual inputs.

The transformation from the membrane potential to the spike rate $R(t)$ was modeled by a rectification function that is zero for membrane potentials below a threshold v_T then grows linearly:

$$R(t) \propto [v(t) - v_T]_+ \quad (4)$$

The time-delay required for each unit to cross its firing threshold after the onset of an afferent FF signal was considered to be a measure of the latency of the cell's response. When a cell crosses threshold, it emits a LH signal, [1] that reaches the soma of the post-synaptic target neuron after a delay equal to the ratio of the traveled distance to the speed of LH propagation.

2.1. Neuron model

Parameters R, C and v_T control the membrane time-constant of the modeled neuron. These parameters were chosen to account for the fact that V1 response latencies decrease with increasing contrast (Gawne, Kjaer, & Richmond, 1996). We assumed that, due to integration time within V1, response latencies decrease by about 30 ms when the stimulus contrast increases by one log-scale unit. Note that although the modeled neuron was developed in analogy with the known properties of single cell's dynamics, its behaviour is assumed to represent that of a pool of locally interacting cells tuned to similar stimulus characteristics.

[1] This choice was elected because it allowed a simple control of the amplitude and time-course of the LH signal independently of the pre-synaptic supra-threshold response dynamics (which were not fully modeled).

2.2. Feed-forward inputs

The amplitude A_f of the FF input was taken to be a linear function of the stimulus contrast. The dynamics of this input current, controlled by the time-constant τ_f, was constrained by intracellular data showing that the membrane potential response to an oriented element flashed during 16 ms inside the RF lasts about 100 ms (Baudot et al., 2000).

2.3. Horizontal inputs

The model LH connections mediate a subliminar excitatory signal, and only exist between iso-oriented RFs. Their efficacy increases linearly with the degree of alignment of the pre- and post-synaptic RFs (Schmidt et al., 1997). In physiology, LH connections are thought to link neurons that are at least about one hypercolumn apart, the local connectivity ($\lesssim 500$ μm) being isotropic (Das & Gibert, 1999). Their anatomical density (e.g. Bosking et al., 1997) and functional strength (e.g. Bringuier et al., 1999) are also known to decrease with distance. To account for these results, we assume that the efficacy of LH connections is zero between RFs separated by less d_{min}, peaks for a separation of d_{opt} and then decreases linearly with distance with a slope α.

The retino-cortical magnification factor M—the distance separating two units in cortex divided by the separation of the centers of their RFs in visual field—is a critical parameter to take into account in the description of the influence of separation on lateral interactions. M is here described as a constant parameter, equal to the average magnification factor over the perifoveal range of eccentricities at which the stimuli were presented. In the following, the speed of propagation through LH connections (in cortical space) is denoted v_c (m/s). The ratio of v_c to M, denoted ϖ (°/s), describes the horizontal propagation speed mapped in retinal space.

2.4. Model inputs

Model inputs (Fig. 2) were sequences composed of two or four flashed oriented elements. Their speeds v_s was defined as $v_s = \Delta x_s/\Delta t_s$, where Δx_s and Δt_s denote the spatial and temporal interval between the elements. We analyzed independently the cases where speed was varied by:

- Varying the temporal interval Δt_s while the spatial separation Δx_s is fixed ("Model FX", Fig. 2(A)). In that case, the populations of neurons responding to the different elements of the sequence are independent of sequence's speed.
- Varying the spatial separation Δx_s while the temporal interval Δt_s is fixed ("Model FT", Fig. 2(B)). In that case, the populations of neurons responding to the different elements of the sequence differ when sequence speed varies, and their cortical separation increases with sequence speed.

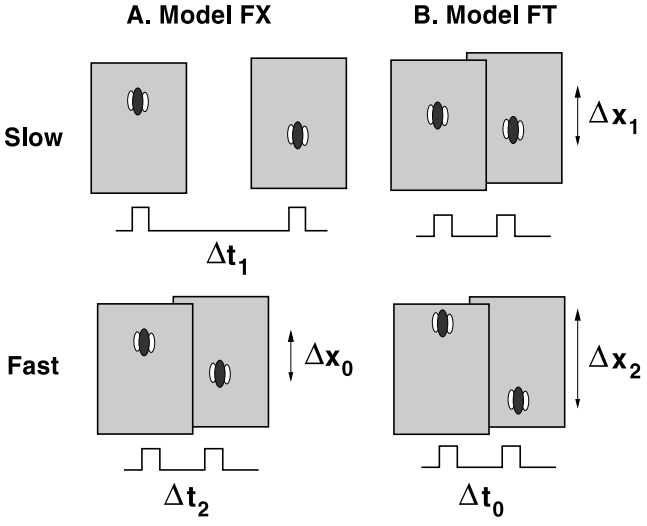

Fig. 2. Model inputs are apparent motion sequences of brief and non-overlapping oriented stimuli. The speed of these sequences is controlled by varying either the temporal interval (Model FX), or the spatial interval between the elements (Model FT).

The fixed spatial or temporal intervals were chosen to be comparable to the mean length of the sequences used experimentally (2°) and to multiples of 16 ms (inter-frame interval).

3. Dynamics of V1 center/surround interactions

Most experimental and theoretical studies on center/surround interactions have used simultaneous center and surround stimuli, presented for various durations. They have then commonly focused on the steady-state amplitude modulation of the responses to the center stimulus. By contrast, we here consider brief and asynchronous events and we investigate the influence of a first event (which can be assimilated to a "surround") on the dynamics of the responses to the next ("center").

3.1. Latency modulations as a function of the temporal overlap between FF and horizontal signals

The model was first used to analyze the evolution of the response latency of a single cell that receives a brief supra-threshold FF signal at time ξ_f and a brief sub-threshold LH signal at time ξ_h, when their relative timing $\Delta\xi_{(f-h)} = \xi_f - \xi_h$ is varied.

It is clear qualitatively that if the LH signal reaches the soma while the cell is integrating the FF signal but is still below threshold, the summation of both signals will result in a modulation of the response latency. If, on the contrary, LH inputs arrive "too late" (the cell has

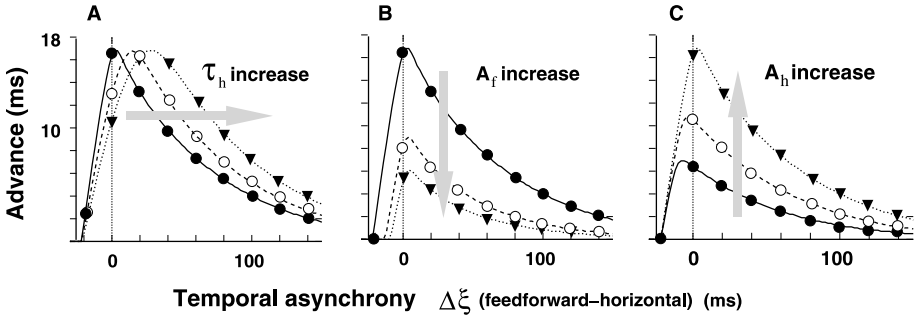

Fig. 3. Latency modulation of one model cell as a function of the relative timing of the afferent FF and the LH signals. (A) The time-course of the LH signal is varied: $\tau_h = 1.5$ ms (●); 5 ms (○); or 10 ms (▼). (B) The amplitude of the FF signal is varied, $A_f = 2$ (●); 2.85 (○) or 4 (▼). (C) The amplitude of the LH signal is varied: $A_h = 1.5$ (●); 3 (○) or 6 (▼). Other parameters: $C = 1$ nF; $R = 50$ MΩ; $V_T = 10$ mV; when not stated otherwise: $\tau_f = 8$ ms; $\tau_h = 1.5$ ms; $A_f = 2$; $A_h = 6$.

already crossed threshold—condition I), or "too early" (they no longer influence the cell's membrane potential at the onset of the FF input—condition II), no latency modulation can occur.

Latency modulations can be shown to appear in a temporal window whose duration is equal to the duration of the LH signal, and whose lower and upper bounds are given by conditions I and II, i.e. if:

$$-t_0 < \Delta\xi_{(f-h)} < -t_0 + d_h \quad (5)$$

where d_h denotes the duration of the LH signal, and t_0 is the response latency to the FF signal alone. In this range, latency modulations are maximal if the LH signal peaks when the cell crosses threshold, as a result of the integration of both signals (Appendix A). This implies that the optimal interval $\Delta\xi_{(f-h)}$ decreases if the time-constant of the LH signal is shortened (Fig. 3(A)).

Latency modulations increase if the amplitude of the LH signal A_h is increased, while they decrease, in a more pronounced way, if A_f is increased (Fig. 3(B) and (C)). If we thus assume that the decrease of the LH signals is slower than or equal to that of FF signals when contrast is decreased, [2] the model suggests that latency modulations should increase when the contrast of the stimulus is decreased. Finally, the amplitude of LH signals being dependent on the alignment of the pre- and post-synaptic RFs, response latencies are maximally advanced for collinear configurations and unchanged for parallel configurations. In the chosen range of parameters, the summation of FF and LH inputs induces a shortening of response latency of up to 17 ms (for plausibility of this, see Section 6), the latter value being observed when FF and LH inputs are approximately synchronous.

Note that, with this model, the dynamics and amplitude of the supra-threshold response also vary when the relative timing of the two signals varies (not shown). Because a more detailed model would be required for a realistic description of the complex non-linear properties of supra-threshold modulations observed experimentally (Baudot et al., 2000), we leave their description for further studies.

3.2. Latency modulations as a function of speed

To determine how these temporal constraints translate in terms of sequence speed, we next consider a two-units network, that receives as inputs a sequence of two collinear elements. We assume that the first cell is maximally responsive to the first element of the sequence while the second cell responds to the second element. In this situation, the relative timing $\Delta\xi_{(f-h)}$ of the FF and LH inputs received by the second unit varies with sequence speed. It is given by:

$$\Delta\xi_{(f-h)} = \Delta t_s - t_0 - \varpi^{-1}\Delta x_s \quad (6)$$

Using Eqs. (5) and (6), we can predict the range of sequence speeds for which latency modulations are possible (speed range, SR). Depending on which of the spatial (FX) or the temporal (FT) interval between the elements is fixed, we have:

[2] While the dependence of FF amplitude on stimulus contrast is well documented experimentally, that of LH signals is not clear. There are, however, some indications that LH signals remain effective at low stimulus contrasts. Physiological recordings in vitro have shown that the horizontally evoked post-synaptic responses to electric shocks of increasing amplitude often comprised a disynaptic IPSP that truncated or even dominated the response (Hirsh & Gilbert, 1991). In in vivo studies, facilitatory interactions are mostly observed at low contrasts of the *center* stimulus (e.g. Kapadia, Westheimer, & Gilbert, 2000; Polat, Mizobe, Pettet, Kasamatsu, & Norcia, 1998). These studies have commonly used a fixed high contrast level for the surround stimulus. However, one study reported that facilitatory interactions remain strong when both center and surround stimuli are presented at low contrasts (Kapadia et al., 2000). In psychophysics, it was also shown that the contrast of the masks was not a critical parameter for the decrease of the contrast threshold of the target (Polat, 1999). In other cases, it was shown that low contrast surrounds were more likely to induce facilitatory interactions than high contrast surrounds (Xing & Heeger, 2001).

$$\begin{aligned}\text{FX: SR} &= \left[\left(\frac{d_h}{\Delta x_s}+\frac{1}{\varpi}\right)^{-1};\varpi\right]\\ \text{FT: SR} &= \begin{cases}\left[\varpi\left(1-\frac{d_h}{\Delta t_s}\right);\varpi\right] & \text{if } \Delta t_v > d_h\\ [0,\varpi] & \text{otherwise}\end{cases}\end{aligned} \tag{7}$$

The lower bound of this interval depends on the propagation speed ϖ and on the duration of the horizontal signal relative to the fixed spatial or temporal interval, while its upper bound is strictly equal to ϖ. From Eq. (6), we can also derive the sequence speed $\widehat{V_s^{\phi}}$ which induces a maximal latency modulation (optimal speed), as a function of the optimal timing $\widehat{\Delta\xi_{(f-h)}}$:

$$\begin{aligned}\text{FX: } \widehat{v_s^{\phi}} &= \left(\frac{\widehat{\Delta\xi_{(f-h)}}+t_0}{\Delta x_s}+\frac{1}{\varpi}\right)^{-1}\\ \text{FT: } \widehat{v_s^{\phi}} &= \begin{cases}\varpi\left(1-\frac{\widehat{\Delta\xi_{(f-h)}}+t_0}{\Delta t_s}\right) & \text{if } \Delta t_s > \widehat{\Delta\xi_{(f-h)}}+t_0\\ 0 & \text{otherwise}\end{cases}\end{aligned} \tag{8}$$

The optimal speed increases when the (fixed) spatial (FX) or temporal (FT) separation is increased. It also increases with the propagation speed ϖ. Because ϖ is inversely proportional to the magnification factor, the optimal speed should increase with visual eccentricity. Finally, the optimal speed increases when the response latency to the FF signal alone decreases, which occurs when the contrast of the visual elements is increased, and when $\widehat{\Delta\xi_{(f-h)}}$ decreases i.e. when the time-constant of the LH signal is reduced. Note that Eqs. (5)–(8) are independent of the chosen model implementation and parameters.

3.2.1. Model FX

Fig. 4 illustrates the variations of the response latency of the second unit, when sequence speed is increased from 1°/s to 250°/s by varying Δt_s. The spatial separation [3] Δx_s was fixed at either 1° or 2° and four different values of the propagation speed ϖ were chosen (66°/s, 166°/s, 333°/s, 1000°/s). Except for the latter, these values are all compatible with the known values of the horizontal propagation speed, for perifoveal stimuli. [4]

In these simulations, the lower limit of the speed range is comprised between 0 and 10°/s and its upper limit is equal to ϖ. When ϖ increases, the optimal speed increases from 25°/s to 39°/s for $\Delta x_s = 1°$, and from 36.5°/s to 74°/s for $\Delta x_s = 2°$.

3.2.2. Model FT

When speed is controlled by varying Δx_s, the cortical separation between the activated units increases with sequence speed. In that case, the variations of response latency reflect the intersection of two constraints: the relative timing of FF and LH signals arriving on the second unit (temporal constraints), and the dependency of the horizontal connections' efficacy on cortical separation (spatial constraints).

The *temporal constraints* can be examined by first considering that the efficacy of LH connections is independent of cortical separation. The speed range and optimal speed are then given by Eqs. (7) and (8). A striking feature of the corresponding simulations (Fig. 4(B)) is that latency modulations are still observed for the lowest speeds. This is due to the fact that the duration of the LH signal is longer than Δt_s (Eq. (7), LH signal is never "too early"). Similarly, when Δt_s is shorter than $\widehat{\Delta\xi} + t_0$ (e.g. $\Delta t_s = 16$ ms, Fig. 4(B) top), latency modulations decrease monotonically with sequence speed (Eq. (8), LH signal is always "too late"). When Δt_s increases above $\widehat{\Delta\xi} + t_0$ (e.g. $\Delta t_s = 48$ ms, Fig. 4(B) bottom), the effect becomes "band-bass" as a function of speed, and the optimal speed increases linearly with ϖ. In all cases, the speed range is equal to $[0,\varpi]$.

If we now assume that the efficacy of LH connections is zero below a separation of d_{min} (in visual space), peaks for a separation of d_{opt} and then decreases linearly with distance with a slope α, these *spatial constraints* imply that LH signals only affect a specific range of sequence speed:

$$\text{SR}' = \left[\frac{d_{min}}{\Delta t_s};\frac{d_{opt}+\frac{4}{\alpha}}{\Delta t_s}\right] \tag{9}$$

The particular speed for which the efficacy of the LH signal is maximal is:

$$\widehat{v_s}' = \frac{d_{opt}}{\Delta t_s} \tag{10}$$

Depending on the chosen values of d_{min}, d_{opt}, and α, these spatial constraints can dramatically affect how latency modulations vary with speed (Fig. 4(C)). In particular, the latency modulations at the lowest speeds (smallest spatial separations) predicted on the basis of the temporal constraints now vanish, making the effect "band-pass" as a function of speed. When the spatial constraints are more stringent than the temporal constraints, as in Fig. 4(C) bottom, the position of the

[3] In this case, we suppose that the strength of horizontal connections is identical between cells whose RF centers are separated by 1° or 2°.

[4] The estimation of ϖ depends on the estimation of the magnification factor M in perifovea, which varies with eccentricity and between species, and of the propagation speed v_c. Measured values of the apparent speed of propagation along horizontal connections range between 0.1 and 1 m/s (Bringuier et al., 1999; Girard, Hupé, & Bullier, 2001; Grinvald et al., 1994). If we take an estimate of M between 2 and 5 mm/deg, corresponding to an eccentricity of 2–6° for human retinocortical projection (Dow, Snyder, Vautin, & Bauer, 1981; Sereno et al., 1995), ϖ varies between 20°/s and 500°/s.

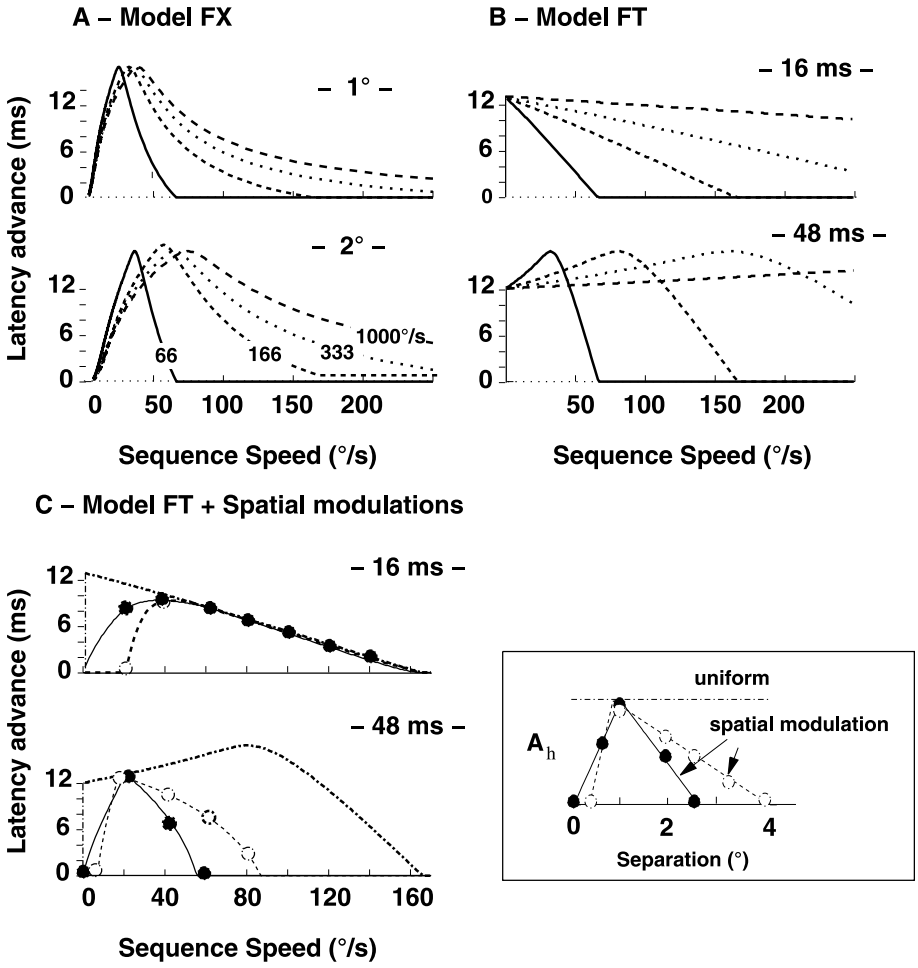

Fig. 4. Simulations of the response latency of the cortical unit activated by the second element of the motion sequence, as a function of sequence speed. (A) Model FX. Four different horizontal propagation speed values are used: $\varpi = 66°/s$, 166°/s, 333°/s, 1000°/s. The spatial separation between sequence elements is fixed at either $\Delta x_s = 1°$ (top) or $\Delta x_s = 2°$ (bottom). (B) Model FT. LH efficacies are independent of the traveled distance. The temporal interval between the visual elements is fixed at either $\Delta t_s = 16$ ms (top), or $\Delta t_s = 48$ ms (bottom). (C) LH efficacies vary with the separation of the RFs, with $d_{min} = 0°$, $d_{opt} = 1°$, $\alpha = -60\%$/deg (●) or $d_{min} = 0.3°$, $d_{opt} = 0.8°$, $\alpha = -30\%$/deg (○). $\varpi = 166°/s$. Assuming $M = 3$ mm/deg, this corresponds to $d_{min} = 0$ mm, $d_{opt} = 3$ mm, $\alpha = -20\%$/mm (●) or $d_{min} = 1$ mm, $d_{opt} = 2.5$ mm, $\alpha = -10\%$/mm (○), $\varpi = 0.5$ m/s.

peak reflects the spatial separation for which LH connections are most efficient, and the limits of the speed range for which latency modulations are observed correspond to the minimal and maximal extent of LH interactions.

3.2.3. Conclusion

A simplified model of V1 intracortical connectivity predicts that the subliminar horizontal activity evoked by an element of a motion sequence can modulate the latency of the response to a subsequent element. This effect decreases when the strength of the FF signal increases, and when the strength of the LH signal decreases. It is sensitive to the precise timing of FF and LH signals and to the alignment of sequence's elements. Depending on how speed is varied, the modulations of response latency reflect the temporal (FX) or the spatio-temporal (FT) constraints imposed on the summation of FF and LH signals. The temporal constraints are primarily controlled by the speed of propagation through LH connections and the duration of the LH signal relative to the temporal separation between sequence's elements, while the spatial constraints are dictated by the spatial architecture of LH connections. The range of sequence speeds for which latency modulations are expected was shown to be always bounded by the speed of horizontal propagation ϖ (°/s) and to correspond to fast motion on the retina. It is commonly believed that relative differences in neural latencies could influence the processing of visual motion, potentially explaining a variety of illusions (e.g. Hikosaka, Miyauchi, & Shimojo, 1993; Mateeff, Bohdanecky, Hohnsbein, Ehrenstein, & Yakimoff, 1991; Whitney, Murakami, & Cavanagh, 2000). In the following, we investigate how the predicted modulations of V1 response dynamics could bias the apparent speed of motion sequences.

4. From the dynamics of V1 center/surround modulations to apparent speed

4.1. MT-like stage and apparent speed

Motion processing and speed discrimination are largely performed by motion selective visual neurons in the MT/MST complex (Mikami, Newsome, & Wurtz, 1986; Newsome, Britten, & Movshon, 1989). These areas predominantly receive direct inputs from V1 magnocellular cells (Maunsell & Newsome, 1987). Moreover, MT neurons respond to apparent motion sequences in a spatio-temporal range similar to that observed psychophysically in humans (Mikami et al., 1986) which suggests that the perception of apparent motion may be mediated by the spatio-temporal correlation of V1 inputs in MT neurons. We hypothesize that differential latency modulations of V1 cortical units could bias the response of their MT target neurons to higher speeds for shorter delays. This bias can be expected to affect behavioral responses as these have been shown to be correlated with the responses of MT units (Newsome et al., 1989).

To explore this hypothesis, a second processing stage is introduced. We assume that the MT-like stage evaluates sequences' speed on the basis of the spatio-temporal correlation between the activation onsets of the sequentially stimulated V1 units. For the purpose of the model, we assume that the output signals of the modeled V1 units are high pass filtered and rectified before converging in MT neurons. This operation, here simply modeled as a temporal derivation, enhances the initial transient of the V1 responses and suppresses the portions that vary more slowly in time. The resulting V1 model is comparable to the "sandwich model" proposed by Carandini, Mehler, Leonard, and Movshon (1996) to describe the dynamics of the spike-encoding properties of V1 cells. The MT stage (Fig. 5) is composed of an idealized population of Reichardt detectors (e.g. Borst & Eghelaaf, 1989; Zanker, 1999) that perform a multiplication of two V1 unit's responses, one of which is delayed by $\Delta\tau$. For all pairs of V1 units separated by a distance Δx_v, we assume that there exists a large population of correlators with a sampling base equal to Δx_v, and smoothly varying delays $\Delta\tau$. The response of correlator j in this population is then given by:

$$C_j = \int [\dot{r}_1(t+\Delta\tau_j)]_+ [\dot{r}_2(t)]_+ \,\mathrm{d}t \tag{11}$$

where $\dot{r}_1(t)$ and $\dot{r}_2(t)$ are the first time derivative of the V1 responses $r_1(t)$ and $r_2(t)$, and $[\cdot]_+$ denotes a rectification operation. The correlator that is maximally activated is selected through a "winner-take-all" mechanism. Its read-out (temporal delay $\Delta\tau_{max}$, sampling base Δx_{max}) is used as a measure of the perceived speed ($\Delta x_{max}/\Delta\tau_{max}$) of the sequence. Because of the high-pass filtering, this is equivalent to considering that the apparent speed of a two element sequence is processed on the basis of the separation of the sequentially stimulated V1 units (Δx_s) and the time-delay between their activation onsets ($\Delta t_s - \delta\tau(v_s^\phi)$). The apparent speed v_s^Ψ of a two-elements sequence moving at a physical speed v_s^ϕ thus obeys:

$$v_s^\Psi = \frac{\Delta x_{max}}{\Delta\tau_{max}} = \frac{\Delta x_s}{\Delta t_s - \delta\tau(v_s^\phi)} \tag{12}$$

where $\delta\tau(v_s^\phi)$ is the latency advance of the second cell's response. When the spatio-temporal configuration of the visual inputs is such that there is no modulation of the response latency of the second unit ($\delta\tau(v_s^\phi) = 0$, inappropriate alignment, distance, or speed), the sequence's apparent speed v_s^Ψ is equal to its physical speed. On the contrary, a latency reduction biases the estimation of speed towards higher speeds. It can be shown that the

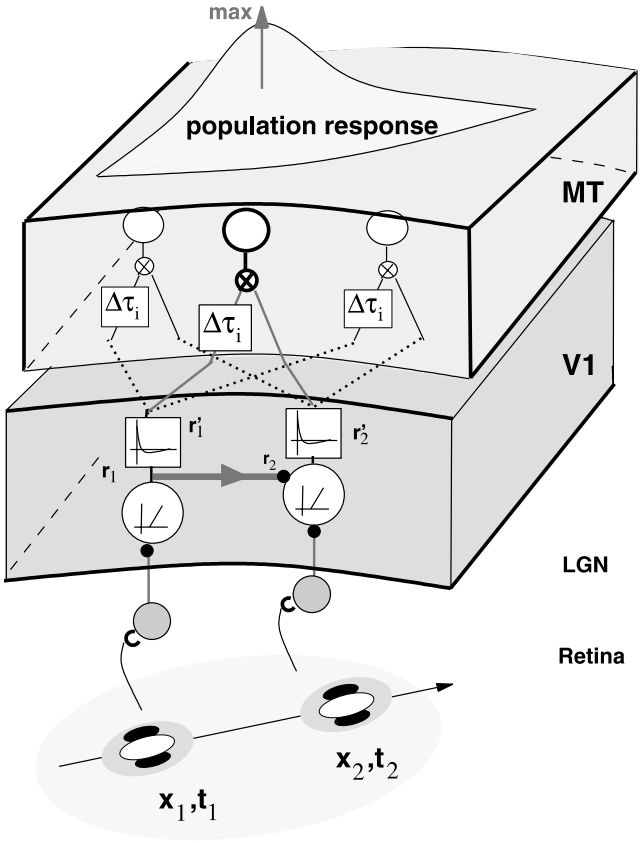

Fig. 5. Cartoon of the equivalent circuit of the full model. Model inputs are visual sequences of varying speeds. Each element of these sequences is first processed at the V1 stage by a RC circuit followed by a rectification and a high pass filter ("sandwich model"). V1 outputs then converge to the MT stage, which consists of a large population of Reichardt correlators. The apparent speed of the sequence is given by the read-out of the correlator that is maximally active. The reduction in response latency resulting from the summation of FF and LH signals in V1 biases the spatio-temporal correlation performed by the MT detectors towards higher speeds.

perceived speed is always bounded by the horizontal propagation speed ϖ (Appendix B).

4.2. Apparent speed modulations

It is clear that the range of speeds that are misjudged corresponds to that for which V1 response latencies are modulated (SR, Eq. (7)). To quantify the resulting bias in speed estimation, we used the ratio (Gain, $G(v_s^\phi)$) of the apparent speed to the physical speed:

$$G(v_s^\phi) = \frac{v_s^\Psi}{v_s^\phi} \tag{13}$$

Depending on which of the spatial (FX) or temporal (FT) interval between the elements is fixed, $G(v_s^\phi)$ can be expressed as:

$$\begin{aligned} \text{FX:}\ & G(v_s^\phi) = 1 + \left(\frac{\Delta x_s}{\delta\tau(v_s^\phi) v_s^\phi} - 1 \right)^{-1} \\ \text{FT:}\ & G(v_s^\phi) = 1 + \left(\frac{\Delta t_s}{\delta\tau(v_s^\phi)} - 1 \right)^{-1} \end{aligned} \tag{14}$$

This implies that the amplitude of the speed bias depends on the amplitude of the V1 latency modulations relative to the temporal separation between the sequence's elements.

4.2.1. Model FX

Because the temporal separation is not fixed in model FX, the optimal speed is *not* equal to the sequence speed inducing maximal latency modulations in V1 (Eq. (8)). It is always slightly faster, when the product $\delta\tau(v_s^\phi)v_s^\phi$ is maximal. Similarly, for a constant latency modulation, the gain increases when Δt_s decreases, i.e. when the fixed Δx_s is decreased (e.g. 1° vs 2°) or when speed increases. As a consequence, increasing the horizontal propagation speed ϖ, by shifting the curve toward higher speeds, produces higher gains. Fig. 6(A) illustrates these dependencies. When ϖ is increased from 66°/s to 1000°/s, the maximal gain varies between 1.8 and 4.3 for $\Delta x_s = 1°$ and between 1.4 and 3.6 for $\Delta x_s = 2°$. The optimal speed varies between 27.7°/s and 66.2°/s for $\Delta x_s = 1°$, and between 38.4°/s and 113.7°/s for $\Delta x_s = 2°$.

4.2.2. Model FT

Eq. (14) captures the variations of the gain in model FT when only the temporal constraints are considered. As before, the maximal gain decreases when Δt_s is increased (e.g. 16 vs 48 ms, Fig. 6(B)), but because Δt_s does not vary with speed, the optimal speed is equal to the speed which induces maximal latency modulations (Eq. (8)) and the amplitude of the maximal gain is independent of the position of the optimal speed. The variations of the gain are therefore a much more faithful image of the variations of response latency than under Model FX. The addition of the spatial constraints can easily be deduced from these results and Fig. 4(C). It is illustrated in Fig. 6(C) using the spatial architectures described previously.

4.2.3. Conclusion

If the estimation of sequence speed is based on a spatio-temporal correlation of V1 activities, this simple model predicts an alignment-dependent perceptual bias in the estimation of speed. This effect is predicted to appear for a particular range of sequence speeds, defined by the spatio-temporal constraints imposed on the summation of LH and FF inputs. Its magnitude depends on the amplitude of the V1 latency modulations relative to the temporal separation between sequence's elements. Model simulations predict that the overestimation of the speed of collinear sequence relative to that of parallel sequence is large, and peaks for speeds values that do not exceed 100°/s. In Section 5, we explore whether such a simple mechanism is sufficient to quantitatively account

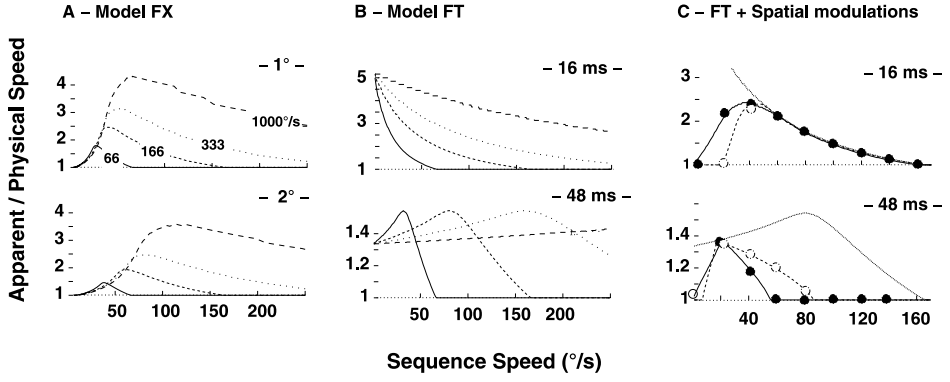

Fig. 6. Apparent speed gain (apparent/physical speed) as a function of sequence speed. (A) Model FX. The gain peak increases when Δx_s is decreased (1° (top) vs 2° (bottom)), and when it corresponds to higher sequence speed (i.e. when ϖ is increased). (B) Model FT. The LH connection strength is independent of the cortical separation. The gain peak decreases when Δt_v is increased (16 vs 48 ms). (C) The LH connection strength varies with separation according to the architectures shown in Fig. 4(C). $\varpi = 166$°/s.

for the pattern of results shown in Experiments 1, 2 and 4 of our companion paper.

5. Orientation-dependent bias in apparent speed: simulation of the psychophysical results

The psychophysical study presented in Georges et al. (this issue) aimed at measuring speed discrimination of apparent motion sequences, using elongated stimuli of different orientations relative to the motion path. It revealed that apparent motion sequences appear faster when the visual elements they contain are aligned with the motion path (collinear sequences) than when they are at an angle with it. This effect is large for high speeds (40–64–96°/s), peaks at 64°/s, decreases for intermediate speeds (12–24°/s) and disappears at low speeds (4°/s) (Experiment 1). This speed bias decreases as the angle between the motion axis and the Gabor patch increases (Experiment 3). When compared with sequences made of non-oriented elements, the speed of collinear sequences is overestimated while the speed of parallel sequences is underestimated (Experiment 4).

5.1. Methods

5.1.1. Model extensions

In order to replicate the psychophysical paradigm (3 to 5 elements, $\Delta t_s = 16.6$ ms), our "FT" model was first extended so as to describe a network of 4 units, processing input sequences composed of 4 elements. For simplicity, we assume that LH connections are only influent between units that are selective to two subsequent elements of the visual sequence (nearest neighbor connectivity). As before, their strength is dependent on the spatial separation between the connected units. We then considered that speed estimation is based on the spatio-temporal correlation of the V1 responses to the first and last elements of a visual sequence. [5]

In psychophysics, we found that the speed of collinear (resp. parallel) sequences is overestimated (resp. underestimated) compared to that of sequences made of non-oriented elements. If non-oriented sequences are "neutral" in the orientation domain, this suggests that the speed bias involves a relative facilitation for collinear sequences and a relative suppression for parallel sequences. It may then appear that long-range inhibitory connections between parallel RFs are necessary to account for our data. Although such an implementation is theoretically possible and can lead to a reasonable fit of the data (not shown), it has yet only weak experimental support. [6] An alternative possibility is that non-oriented stimuli elicit a response from cells tuned to all orientations, which in turn propagates in the network of horizontal connections. The observed over/under-estimation can therefore be interpreted solely with long-distance excitation, assuming no orientation-dependent bias for parallel sequences, a weak over-estimation for non-oriented sequences, and a strong over-estimation for collinear sequences. The following simulations were performed under these assumptions.

5.1.2. Discrimination stage and variability

The model was further extended so as to include a speed discrimination processing stage and to account for the observed variability in this procedure. We consider that, as a result of all possible sources of variability in V1 or at higher processing stages, the apparent speed can be described as a Gaussian random variable V_s^Ψ with a mean μ_s equal to the deterministic value v_s^Ψ (Eq. (12)), and a variance σ_s^2 chosen to be a function of the mean: $\sigma_s^2 = \rho(v_s^\Psi)^\beta$. The discrimination performance in the simulated forced choice experiment, involving a reference sequence with physical speed v_{ref}^ϕ and a comparison sequence with physical speed v_{comp}^ϕ, is given by the probability that the reference speed is perceived as being faster than the comparison speed:

$$P(V_{ref}^\Psi > V_{comp}^\Psi) = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{v_{ref}^\Psi - v_{comp}^\Psi}{\sqrt{2[\sigma_{ref}^2 + \sigma_{comp}^2]}}\right)\right] \quad (15)$$

where erf is the normal error function ($\mathrm{erf}(x) = (2/\sqrt{\pi})\int_0^x e^{-t^2}\,dt$). The parameters ρ (slope) and β (dependency on speed) were derived from the psychophysical data, in the situations where no perceptual bias was

[5] This assumption was elected for its simplicity. In reality, it is not clear which elements of the visual sequence are used by human observers to estimate its speed. Note that if we alternately consider that speed estimation of an n-elements sequence is based on all subsequent responses, and if the latency modulation of the nth unit is not equal to $(n-1)$ times that of the second unit, the model predicts that sequence's speed should be perceived as being non-constant along its trajectory (i.e. decelerating in the collinear configuration). A more accurate description should probably be based on some temporal averaging of the speed information present in the stimulus (Watamaniuk & Duchon, 1992). Preliminary simulations show that similar results are obtained under this assumption.

[6] LH axons are thought to be collaterals of excitatory pyramidal neuron. Of these axons at least 80% of their synapses are made with other pyramidal neurons (McGuire, Gilbert, Rivlin, & Wiesel, 1991). Long-range suppression mediated by the remaining 20% synapses on interneurons, being disynaptic, can be thought to show different dynamics than monosynaptic excitation, and possibly to affect the responses only after their very onset (Hirsh & Gilbert, 1991). Moreover, most psychophysical (e.g. Polat & Sagi, 1993) and physiological studies which reported lateral interactions in a parallel configuration indicated that they were of the same sign (but weaker) than in the collinear configuration (but see Kapadia et al., 2000).

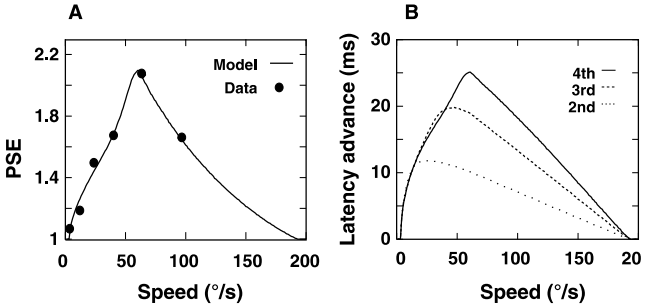

Fig. 7. (A) Best fit of the experimental PSE. The total fit error was computed as the root of the squared errors summed over all data points: rms $= \sqrt{(1/n)\sum(y_i - \hat{y}_i)^2}$ where y represent the experimental data points and $\hat{y}$ the model predictions. rms $= 0.036$. (B) Latency modulations of the model units that responded to the second, third and fourth element of the collinear sequence as a function of sequence speed. No latency advance is observed for parallel sequences. Response latency to the FF signal alone is equal to 28.4 ms with this set of parameters (Table 1).

observed. Satisfactory fits were obtained when $\beta = 2.1$ and $\rho = 0.1$.

5.2. Results

All model parameters (except for the neuron's parameters which were unchanged from previous sections) were optimized using the downhill simplex algorithm (Press, Teukolsky, Vetterling, & Flannery, 1992) to fit the points of subjective equality (PSE, Fig. 3 in companion paper), which indicate, for each collinear speed, the value of the parallel sequence's speed that is perceived, on average, as moving at the same speed. Fig. 7(A) illustrates the result of this procedure. The best fitting parameters are presented in Table 1. Because we assume that parallel sequences are perceived at their veridical speed, the definition of the PSEs strictly corresponds to that of the gain used in the previous section. Fig. 7(B) presents the corresponding latency advance of the second, third and fourth units that are activated by the sequence. Note that the shape and maximal amplitude (−25.1 ms) of the modulation of the fourth unit response, could have been directly inferred from Eq. (8).

Model performances were then tested in the comparison between sequences composed of collinear elements and sequences composed of parallel elements

Table 1
Model parameters that best fit the experimental data

	Parameter	Value
Neuron	Resistance: R	50 MΩ (fixed)
	Capacitance: C	1 nF (fixed)
	Threshold: V_T	10 mV (fixed)
FF input	Amplitude: A_f (contrast)	2.1
	Time-constant: τ_f	8.29 ms
Horizontal input	Amplitude: A_h	3.08
	Time-constant: τ_h	1.3 ms
Speed of horizontal propagation	ϖ	194°/s
Spatial architecture of long-range connections	Minimal distance: d_{min}	0.05°
	Optimal distance: d_{max}	0.97°
	Slope: α	−43%/°
Variability	β	2.1
	ρ	0.1
Connectivity between non-oriented elements	A_b	1.5

These parameters were optimized using the downhill simplex algorithm (Press et al., 1992).

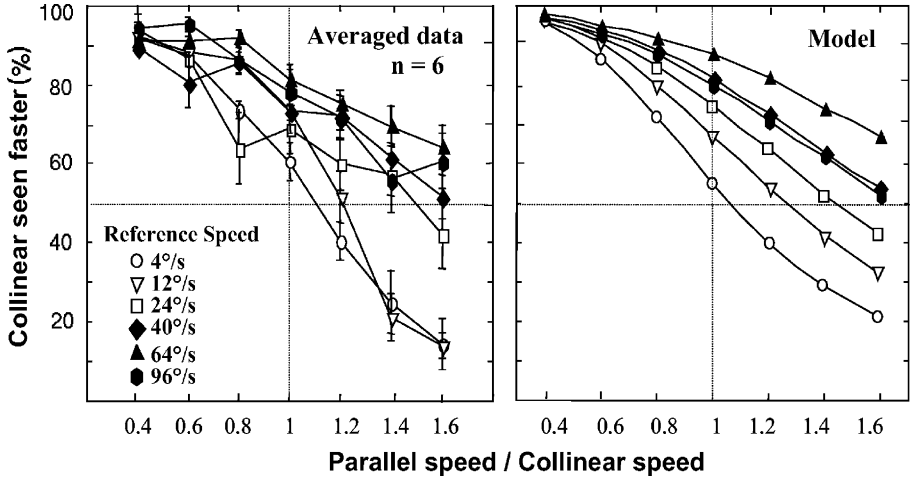

Fig. 8. Comparison between collinear and parallel sequences. (A) Psychophysical results. Average proportion of the trials in which the collinear sequence is perceived as being faster than the parallel sequence. Six reference speeds were used for the collinear sequence (4°/s, 12°/s, 24°/s, 40°/s, 64°/s, 96°/s) and each of these were compared to parallel sequences whose speeds vary between −60% and +60% of the collinear sequence. (B) Model results.

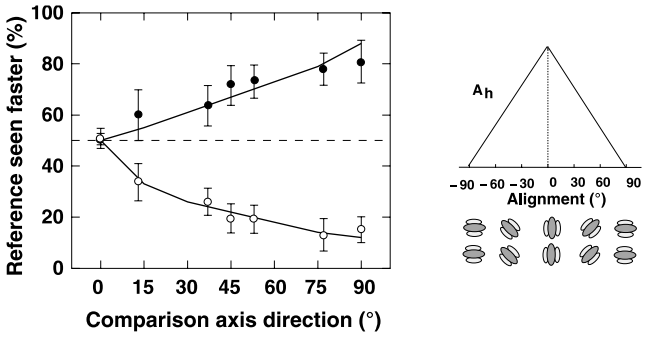

Fig. 9. Influence of alignment. The reference sequence was collinear (●) (resp. parallel (○)) and the comparison sequence was made to gradually deviate for the reference sequence configuration, becoming more and more parallel (resp. collinear). The speed of all sequences was equal to 64°/s. The left graph presents the experimental points (●/○) and the model probability (—) for the reference sequence to be perceived as being faster than the comparison sequence. The LH connection strength between iso-oriented units increases linearly with pre- and post-synaptic RFs' alignment (right).

(Experiment 1). Fig. 8 shows that the model quantitatively accounts for the over-estimation of the speed of collinear sequences at high speeds (40–96°/s).

The model was then used to simulate Experiment 3, in which the reference sequence was either collinear or parallel, and the comparison sequence was made to gradually deviate from the reference sequence configuration. To simulate this experiment, we assumed that the efficacy of horizontal connections increases linearly with the alignment of the connected RFs. This simple hypothesis is sufficient to account for the observed sensitivity of the speed bias to alignment (Fig. 9).

We finally investigated the ability of the model to account for the comparison between collinear or parallel sequences (reference) and sequences composed of non-oriented elements (comparison, Experiment 4). As discussed before, we assumed that non-oriented elements interact through (weak) facilitatory connections. The strength of these connections was described by a new parameter (A_b, cf. Table 1), optimized to fit the data. All other parameters were identical to those used in Figs. 7–9. Fig. 10 shows that the model satisfactorily accounts for the overestimation (resp. underestimation) of the speed of collinear (resp. parallel) sequences relative to that of non-oriented sequences.

The set of parameters resulting from the fit of the psychophysical data correspond to the spatial and temporal constraints imposed on the summation of FF and LH signals that best account for the observed bias under our hypotheses. These describe a spatio-temporal map of interactions illustrated in Fig. 11, which can be compared at the physiological level with the synaptic "integration field" of V1 neurons (e.g. Bringuier et al., 1999), and which, at the perceptual level, is reminiscent of the "association field" described by Field et al. (1993).

The model first requires that facilitatory interactions are specific to collinear configurations and exist between stimuli separated by up to 3° of visual angle, which is consistent with psychophysical findings (e.g. Polat & Sagi, 1993). Our analysis and simulations also show that a satisfactory fit of the psychophysical data is obtained for ϖ in range (150°/s, 200°/s). If we take an estimate [7] of the magnification factor $M \simeq 3$ mm/deg, the predicted cortical speed of propagation v_c is a little higher (0.45–0.6 m/s) than measured in cat and monkey V1 (0.05–0.5 m/s, Bringuier et al., 1999). The required maximal extent of LH connections is about 9 mm of cortical tissue, which is roughly consistent with observations in cat and monkey V1 (e.g. Angelucci, Levitt, & Lund, 2002; Bringuier et al., 1999; Gilbert et al., 1996). The model robustness and dependency on each parameter (contrast, time-constants, spatial structure of connectivity, propagation speed etc.) can easily be derived from the analysis we have carried in the previous sections for model FT.

6. Discussion

We have suggested that the orientation-dependent bias in speed discrimination reported in our companion paper could be interpreted as a perceptual correlate of the spatio-temporal dynamics of V1 center/surround modulations. The present study aimed at investigating this hypothesis by providing a conceptual framework that links physiology and perception. We have shown that a simple mechanism based on the summation of FF and LH signals within V1 is successful in fitting our data.

6.1. Model simplifications

A number of simplifications were made. Some were due to a lack of detailed experimental data, as for example in the description of the amplitude and time-course of supra- and sub-threshold responses to flashed (16 ms) oriented or non-oriented stimuli, or to a difficulty to integrate existing data into a simple description. We have tried to constrain our choices with plausible assumptions that we intend to refine with further theoretical and experimental investigation.

Our simple V1 stage, for example, does not account for the emergence of orientation selectivity (Ferster & Miller, 2000), nor for the complex characteristics of

[7] This is difficult because data is sparse for human V1 and the eccentricity was not fixed in our psychophysical experiments (sequences were rectilinear in general, and eye movements were not recorded).

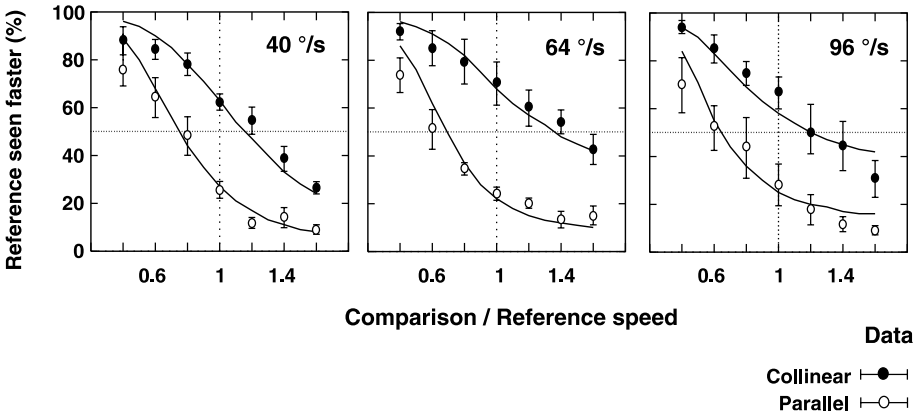

Fig. 10. Comparison between collinear or parallel sequences with sequences made of non-oriented elements. Percentage of the trials for which the reference sequence, which is either collinear (●) or parallel (○), is perceived as being faster than the comparison sequence, which is made of non-oriented elements. The experimental points (●/○) and model probabilities (—) are superimposed. Three reference speeds are used: 40°/s, 64°/s, 96°/s.

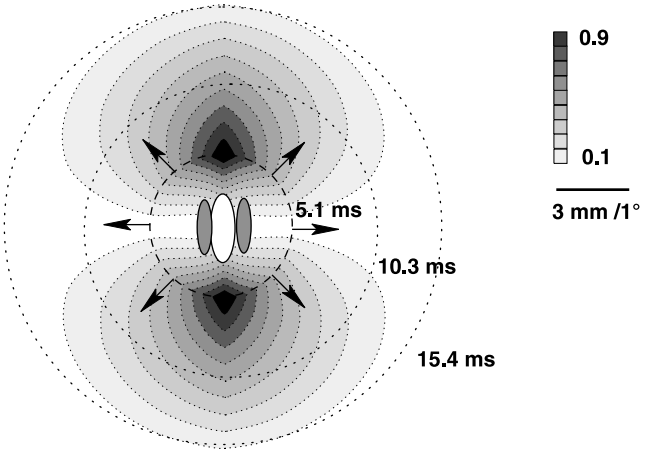

Fig. 11. Spatio-temporal subliminar influence of the neuronal response to a vertical visual element on iso-oriented neighboring neurons. The gray patterns represent the normalized strength of LH connections running between the central pool of cells and cells located at each position of the map. The concentric circles illustrate the time required for LH signals to propagate from the center to the periphery, once it is triggered at the central pre-synaptic site. We here assume a magnification factor of $M = 3$ mm/deg.

center/surround modulations (Dragoi & Sur, 2000; Somers et al., 1998). The "inputs summation" mechanism that we propose may also have a much richer biological implementation. For example, the summation of horizontal and FF inputs is known to become supralinear when the post-synaptic membrane is depolarized above a certain level (Baudot et al., 2000; Hirsh & Gilbert, 1991; Yoshimura, Sato, Imamura, & Watanabe, 2000). Similarly, background activity and the precise timing and dynamics of individual PSPs should probably be taken into account in order to fully characterize the potential influence of LH inputs on V1 cells' synaptic integration. Finally, although we here suggest that LH connections are responsible for the dynamics of V1 center/surround modulations, it remains possible that feedback projections from higher cortical areas (e.g. V2, MT) participate in the phenomenon that we report. The formalism used here could easily be adapted to include such a mechanism.

Other simplifications concern the apparent speed processing stage and the decision stage. These were voluntarily kept to a minimal description, as we considered that the modeling of their underlying physiological mechanisms were out of the scope of the present study. It would of course be important to investigate how they could be conciliated with more detailed models of MT cortex (e.g. Nowlan & Sejnowski, 1995; Simoncelli & Heeger, 1998), apparent motion and speed processing (e.g. Chey, Grossberg, & Mingolla, 1997; Francis & Grossberg, 1996).

6.2. Predictions

Despite its simplicity, our model provides a number of predictions that can be used to test its validity and refine its level of description. As such, we think that it may be useful in providing a framework for more detailed models of long-range interactions and spatio-temporal modulations in visual processing.

The key assumption of the model is that V1 cells' response latency can be strongly modulated by the summation of horizontal and FF inputs. This prediction is supported by recent intracellular recordings performed in our laboratory in cat area 17 using from-periphery-to-center sequences of optimally oriented Gabors flashed across the RF width or length (Baudot et al., 2000). These experiments revealed that fast collinear apparent motion sequences often result in a shortening of visually evoked sub-threshold and spiking latencies by 5–15 ms, a range of values that is consistent with the model's assumptions.

Our model also generates a set of predictions that can be tested in psychophysics (Georges, Seriès, &

Lorenceau, 2000). These derive essentially from the analysis presented in Sections 3 and 4.

- The model first predicts a decrease of the speed bias for higher speeds than were tested experimentally, and a disappearance of the effect above a critical speed equal to the minimum between the propagation speed ϖ and $(d_{\max}/\Delta t_v)$, where $d_{\max}$ denotes the maximal extent of LH interactions in visual space (Eq. (7)).
- The amplitude of the effect should strongly depend on the contrast of the stimuli (Fig. 3), with stronger bias at low contrast: for example, a low contrast collinear sequence should appear as being faster than a high contrast collinear sequence of the same speed. Psychophysical data (Georges, Seriés, & Lorenceau, 2000) indicate that this prediction is valid, although it may seem to contradict previously documented contrast effects on perceived speed [8] (Blakemore & Snowden, 1999; Stone & Thomson, 1992).
- When the visual sequence comprises an increasing number of elements separated by a constant time interval, latency modulations of the sequentially activated units saturate to a constant level, after a few elements (cf. e.g. Fig. 7(B)). If observers base their judgment on the whole duration of the sequence, the perceptual bias should thus decrease as the number of frames increases. [9]
- The optimal speed and range of misjudged speeds should increase when increasing the eccentricity of the visual stimulation (Eq. (8)), provided facilitatory interactions still occur (Hess & Dakin, 1997; Xing & Heeger, 2000).
- The observed perceptual bias should be highly dependent on how sequence speed is varied in the experimental settings. For example, when speed is controlled by a variation of spatial separation Δx_s (resp. temporal interval Δt_s), decreasing Δt_s (resp. Δx_s) should produce stronger effects (Eq. (14)). Similarly, as speed was controlled by varying Δx_s in our experimental study, the observed bias is expected to reflect both the spatial and the temporal constraints imposed on the summation of FF and LH inputs. Complementary experiments in which speed will be controlled by varying Δt_s (FX) will help isolate their relative influence.
- Finally, psychophysical evidence indicate that the architecture of long-range interactions depends on the spatial frequency ($f = \lambda^{-1}$) of the test stimulus (e.g. Polat & Sagi, 1993): long-range interactions were shown to peak at a distance of $\sim 3\lambda$, and then to decrease linearly with distance, up to $\sim 10\lambda$. There is, to our knowledge, no physiological explanation for these findings. However, if V1 horizontal connections are responsible for the spatial interactions reported in these experiments and are specific to spatial frequency, this could suggest that they cover greater cortical distances between cells that are selective to lower spatial frequencies, than between cells selective to higher spatial frequencies. Under this hypothesis, the spatial constraints of the connectivity should translate into a shift of the effect toward lower speeds if spatial frequency is increased.

Other predictions can be made. First, if our model is correct, it suggests that lateral interactions exist between non-oriented elements. Consistent with this prediction, preliminary data indicate that the contrast threshold of a gaussian blob decreases when it is surrounded by other blobs (data collected in our lab by D. Alais).

Second, although the analysis we have performed focused on the influence of horizontal inputs on response latency (which was here identified to the first-spike latency), we expect the summation of FF and LH inputs to affect other dimensions of V1 responses, and in particular their supra-threshold amplitude and dynamics. We cannot exclude that these modulations could provide alternative or complementary explanations for the psychophysical phenomena that we report. It would also be interesting to investigate whether they could be detected in other experimental paradigms, involving, for example, contrast detection or perceptual saliency. Interestingly, under our model's assumptions, if latency advances are accompanied by increases in response amplitude, these would correspond to the emergence—at high speeds—of a preference for a motion axis aligned with the preferred orientation, which could be comparable with that found experimentally (Geisler, Albrecht, Crane, & Stern, 2001; Wörgotter & Eysel, 1989).

More generally, our results suggest that both the properties of V1 RFs and visual perception are influenced not only by the spatial context but also by the temporal context in which an object is presented. This contributes to recent findings (Arieli, Sterkin, Grinvald, & Aertsen, 1996) showing that the notion of what is "noise" in cortical activity may have to be revised. Ongoing activity following the presentation of a visual object could partly reflect the subliminar propagation of

[8] Stone and Thomson (1992), for example, reported that when two gratings moving at the same speed (4°/s) are presented simultaneously, the lower-contrast grating appears slower (a phenomenon which our model does not account for). However, these effects were shown to depend on the temporal presentation of the stimuli: when the two gratings are presented sequentially (like in our experiments) instead of simultaneously, the contrast effect decreases. They are also known to decrease when speed increases above ~8°/s (Blakemore & Snowden, 1999).

[9] In psychophysics, different numbers of frames (from 3 to 5) were used to generate the apparent motion sequences. The predicted influence of sequence length on the speed discrimination bias may thus partly explain the large variability (reflected by the slope of the psychometric curves) and the high discrimination thresholds found in the experimental data.

a family of contours or trajectories in which this object could be embedded. Whether this mechanism could participate in other perceptual "illusions" (line motion effect (Hikosaka et al., 1993), flash-lag (Whitney et al., 2000), "motion streaks" (Geisler, 1999)), and/or be related to other studies showing that motion processing is facilitated when the motion signals are extended in the direction of motion (e.g. Anstis & Ramachandran, 1987; Vreven & Verghese, in press; Watamaniuk, McKee, & Grzywacz, 1995) will be the focus of our future research.

Acknowledgements

This work has been supported by grants from CNRS, INSERM (PROGRES) to Y.F. and J.L. and HSFP RG0103-1998-B to Y.F.P.S. was supported by fellowships from FRM and Singer Polignac fundations. We thank Marc Potters for help with Appendix A and Adrien Rebollo for helpful discussions.

Appendix A

It can be shown analytically that the membrane potential $v(t)$ resulting from the low-pass filtering of the two synaptic currents (Eq. (1)) obeys:

$$v(t) = A_f g(t - \xi_f, \tau_f) + A_h g(t - \xi_h, \tau_h) \tag{A.1}$$

with:

$$g(t,\tau) = \begin{cases} \displaystyle\int_0^t \frac{t}{\tau} e^{-(t/\tau)} e^{-((t-u)/RC)}\, du \\ \quad = \dfrac{\tau e^{-(t/RC)} - \left(\tau + \left(1 - \dfrac{\tau}{RC}\right)t\right) e^{-(t/\tau)}}{\left(1 - \dfrac{\tau}{RC}\right)^2} & \text{if } t \geqslant 0 \\ 0 & \text{otherwise} \end{cases} \tag{A.2}$$

By definition, the cell's response latency t_T obeys $v(t_T) = V_T$, where V_T is the activation threshold. There is no close-form analytical solution for the response latency as a function of all other parameters. However, when the amplitude of the LH signal is small compared to that of the FF signal ($A_h \ll A_f$), a first order approximation of the cell's response latency t_T is given by:

$$t_T \simeq t_0 - \frac{A_h g(t_0 - \xi_h, \tau_h)}{A_f g'(t_0 - \xi_f, \tau_f)} \tag{A.3}$$

where t_0 denotes the latency of the response to the FF signal alone and g' is the first temporal derivative of g. Eq. (A.3) indicates that latency modulations: (i) appear in a temporal window $\Delta\xi_{(f-h)}$ whose duration is equal to the duration of the LH signal; (ii) increase linearly with A_h; (iii) are inversely proportional to A_f; (iv) are maximal when $g'(t_0, \xi_f, \tau_f)$ is minimal, i.e. when the slope of the FF signal is low at t_0; (v) are maximal when $g(t_0, \xi_h, \tau_h)$ is maximal, i.e. when the FF signal peaks at t_0.

When $A_h \sim A_f$, an extremum analysis can be used. If we denote by $\delta\tau$ the latency modulation and consider that ξ_f is fixed, $\delta\tau$ obeys:

$$A_f g(t_0 + \delta\tau - \xi_f) + A_h g(t_0 + \delta\tau - \xi_h) = V_T \tag{A.4}$$

We consider $\delta\tau$ as a regular function of ξ_h, and we aim at determining the value of ξ_h for which $\delta\tau$ is maximum. We differentiate Eq. (A.4):

$$A_f g'(t_0 + \delta\tau - \xi_f,) \frac{d(\delta\tau)}{d\xi_h} + A_h g'(t_0 + \delta\tau - \xi_h)\left(\frac{d(\delta\tau)}{d\xi_h} - 1\right) = 0 \tag{A.5}$$

This gives:

$$\frac{d(\delta\tau)}{d\xi_h} = \frac{A_h g'(t_0 + \delta\tau - \xi_h)}{A_f g'(t_0 + \delta\tau - \xi_f,) + A_h g'(t_0 + \delta\tau - \xi_h)} \tag{A.6}$$

which implies that $\delta\tau$ is maximum when the spatio-temporal combination of FF and LH signals is such that the horizontal signal is at its peak when the cell crosses activation threshold.

Although we focus on excitatory LH signals, these results are also valid for weak inhibitory interactions. For strong inhibitory horizontal signals, the above analysis breaks down as the latency modulation can vary discontinuously with $\xi_f - \xi_h$. In these cases, the optimal spatio-temporal configuration corresponds to the case where the inhibitory horizontal signal arrives just before t_0 ($\xi_h \simeq t_{0-}$).

Appendix B

If we consider that the first FF signal is triggered at $t = 0$, the first cell crosses threshold at time t_0 and the LH signal arrives at the second cell at time $\xi_h = t_0 + \varpi^{-1}\Delta x_s$. If no LH signal were present, the second unit would fire at time $\Delta t_s + t_0$. A response latency modulation can only exist if the LH signal reaches the target cell before it has crossed threshold, which gives:

$$t_0 + \varpi^{-1}\Delta x_s < \Delta t_s + t_0 \tag{A.7}$$

$$V_s^{\phi} = \frac{\Delta x_s}{\Delta t_s} < \varpi$$

i.e. only sequence speeds below ϖ can induce latency modulations.

Appendix C

By definition, a "modulated" response latency occurs after both the FF (condition 1) and the LH signals (condition 2) have arrived. Condition 1 implies that:

$$\Delta t_s + t_0 - \delta\tau > \Delta t_s \iff \delta\tau < t_0 \quad \text{(A.8)}$$

Condition 2 implies that:

$$\Delta t_s + t_0 - \delta\tau > t_0 + \varpi^{-1}\Delta x_s \iff \delta\tau < \Delta t_s - \varpi^{-1}\Delta x_s \quad \text{(A.9)}$$

This gives:

$$v_s^{\Psi} = \frac{\Delta x_s}{\Delta t_s - \delta\tau(v_s^{\phi})} < \varpi$$

$$G(v_s^{\phi}) < \frac{\varpi}{v_s^{\phi}}$$

The amplitude of the response latency advance can never exceed the temporal interval Δt_s. For sequence speeds below ϖ, the perceived speed is always bounded by ϖ.

References

Angelucci, A., Levitt, J. B., & Lund, J. (2002). Anatomical origins of the classical receptive field and modulatory surround field of single neurons in Macaque Visual Cortical Area V1. *Progress in Brain Research, 136*, 373–388.

Anstis, S., & Ramachandran, V. S. (1987). Visual inertia in apparent motion. *Vision Research, 27*, 755–764.

Arieli, A., Sterkin, A., Grinvald, A., & Aertsen, A. (1996). Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science, 273*, 1868–1871.

Baudot, P., Chavane, F., Pananceau, M., Edet, V., Gutkin, B., Lorenceau, J., Grant, K., & Frégnac, Y. (2000). Cellular correlates of apparent motion in the association field of cat area 17 neurons. *Society for Neuroscience Abstracts,* 162.2.

Blakemore, M., & Snowden, R. (1999). The effect of contrast upon perceived speed: a general phenomenon? *Perception, 28*, 33–48.

Borst, A., & Eghelaaf, M. (1989). Principles of visual motion detection. *Trends in Neurosciences, 12*, 297–306.

Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience, 17*, 2112–2127.

Bringuier, V., Chavane, F., Glaeser, L., & Frégnac, Y. (1999). Horizontal propagation of visual activity revealed in the synaptic integration field of area 17 neurons. *Science, 283*, 695–699.

Carandini, M., Mehler, F., Leonard, C., & Movshon, J. (1996). Spike train encoding by regular spiking cells of the visual cortex. *Journal of Neurophysiology, 76*, 3425–3441.

Chey, J., Grossberg, S., & Mingolla, E. (1997). Neural dynamics of motion processing and speed discrimination. *Vision Research, 38*, 2769–2786.

Das, A., & Gibert, C. (1999). Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature, 399*, 655–661.

Dow, B. M., Snyder, A. Z., Vautin, R. G., & Bauer, R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. *Experimental Brain Research, 44*, 213–228.

Dragoi, V., & Sur, M. (2000). Dynamic properties of recurrent inhibition in primary visual cortex: Contrast and orientation dependence of contextual effects. *Journal of Neurophysiology, 83*, 1019–1030.

Ferster, D., & Miller, K. D. (2000). Neural mechanisms of orientation selectivity in the visual cortex. *Annual Review of Neuroscience, 23*, 441–471.

Field, D. J., Hayes, A., & Hess, R. (1993). Contour integration by the human visual system: evidence for a local association field. *Vision Research, 33*, 173–193.

Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology, 10*, 438–443.

Francis, G., & Grossberg, S. (1996). Cortical dynamics of form and motion integration: persistence, apparent motion and illusory contours. *Vision Research, 36*, 149–173.

Frégnac, Y., & Bringuier, V. (1996). Spatio-temporal dynamics of synaptic integration in cat visual cortical receptive fields. In A. Aertsen & V. Braitenberg (Eds.), *Brain Theory—Biological and Computational Principles* (pp. 143–199). Elsevier.

Gawne, T. J., Kjaer, T. W., & Richmond, B. J. (1996). Latency: Another potential code for feature binding in striate cortex. *Journal of Neurophysiology, 76*, 1356–1360.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature, 400*, 65–69.

Geisler, W. S., Albrecht, D. G., Crane, A. M., & Stern, L. (2001). Motion direction signals in the primary visual cortex of cat and monkey. *Visual Neuroscience, 18*, 501–516.

Georges, S., Seriès, P., & Lorenceau, J. (2000). Contrast dependency of high speed apparent motion. *Perception* (suppl.), 29(96d).

Georges, S., Seriès, P., Frégnac, Y., & Lorenceau, J. (this issue). Orientation dependent modulation of apparent speed: Psychophysical evidence. *Vision Research*.

Gilbert, C. D., Das, A., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Sciences of the United States of America, 93*, 615–622.

Girard, P., Hupé, J. M., & Bullier, J. (2001). Feedforward and feedback connections between areas V1 and V2 of the monkey have similar conduction velocities. *Journal of Neurophysiology, 85*, 1328–1331.

Grinvald, A., Lieke, E. E., Frostig, R. D., & Hildesheim, R. (1994). Cortical point-spread function and long-range lateral interactions revealed by real-time optical imaging of macaque monkey primary visual cortex. *The Journal of Neuroscience, 14*, 2545–2568.

Hess, R. F., & Dakin, S. C. (1997). Absence of contour linking in peripheral vision. *Nature, 390*, 602–604.

Hess, R., & Field, D. (1999). Integration of contours: New insights. *Trends in Cognitive Sciences, 3*, 480–486.

Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research, 33*, 1219–1240.

Hirsh, J. A., & Gilbert, C. D. (1991). Synaptic physiology of horizontal connections in the cat's visual cortex. *The Journal of Neuroscience, 11*, 1800–1809.

Kapadia, M. K., Westheimer, G., & Gilbert, C. (2000). Spatial distribution of contexual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology, 84*, 2048–2062.

Mateeff, S., Bohdanecky, Z., Hohnsbein, J., Ehrenstein, W. H., & Yakimoff, N. (1991). A constant latency difference determines directional anisotropy in visual motion perception. *Vision Research, 31*, 2235–2237.

Maunsell, J., & Newsome, W. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience, 10*, 363–401.

McGuire, B., Gilbert, C., Rivlin, P., & Wiesel, T. (1991). Target of horizontal connections in macaque primary visual cortex. *Journal of Comparative Neurology, 305*, 370–392.

Mikami, A., Newsome, W., & Wurtz, R. (1986). Motion selectivity in macaque visual cortex. i. mechanisms of diretion and speed

selectivity in extrastriate area MT. *Journal of Neurophysiology, 55*, 1308–1327.

Newsome, W. T., Britten, K. H., & Movshon, J. A. (1989). Neural correlates of a perceptual decision. *Nature, 341*, 52–54.

Nowlan, S., & Sejnowski, T. (1995). A selection model for motion processing in area MT of primates. *The Journal of Neuroscience, 15*, 1195–1214.

Polat, U. (1999). Functional architecture of long-range perceptual interactions. *Spatial Vision, 12*, 143–162.

Polat, U., Mizobe, K., Pettet, M., Kasamatsu, T., & Norcia, A. M. (1998). Collinear stimuli regulate visual responses depending on cell's contrast threshold. *Nature, 391*, 580–584.

Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: suppression and facilitation revealed by lateral masking experiments. *Vision Research, 33*, 993–999.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1992). *Numerical Recipes in C: The Art of Scientific Computing* (Second ed.). Cambridge University Press.

Schmidt, K., Goebel, R., Löwel, S., & Singer, W. (1997). The perceptual grouping criterion of collinearity is reflected by anisotropies of connections in the primary visual cortex. *European Journal of Neuroscience, 9*, 1083–1089.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., Rosen, B. R., & Tootell, R. B. H. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science, 268*, 889–893.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research, 38*, 743–761.

Sincish, L. C., & Blasdel, G. G. (2001). Oriented axon projections in primary visual cortex of the monkey. *The Journal of Neuroscience, 21*, 4416–4426.

Somers, D. C., Todorov, E., Siapas, A. G., Toth, L. J., Kim, D., & Sur, M. (1998). A local circuit approach to understanding integration of long-range inputs in primary visual cortex. *Cerebral Cortex, 8*, 204–217.

Stone, L., & Thomson, P. (1992). Human speed perception is contrast dependent. *Vision Research, 32*, 1535–1549.

Vreven, D., & Verghese, P. (in press). Integration of speed signals in the direction of motion. *Perception and Psychophysics.*

Watamaniuk, S. N., McKee, S. P., & Grzywacz, N. M. (1995). Detecting a trajectory embedded in random-direction motion noise. *Vision Research, 35*, 65–77.

Watamaniuk, S., & Duchon, A. (1992). The human visual system averages speed information. *Vision Research, 5*, 931–941.

Whitney, D., Murakami, I., & Cavanagh, P. (2000). Illusory spatial offset of a flash relative to a moving stimulus is caused by differential latencies for moving and flashed stimuli. *Vision Research, 40*, 137–149.

Wörgotter, F., & Eysel, U. T. (1989). Axis of preferred motion is a function of bar length in visual cortical receptive fields. *Experimental Brain Research, 76*, 307–314.

Xing, J., & Heeger, D. J. (2000). Center-surround interactions in foveal and peripheral vision. *Vision Research, 40*, 3065–3072.

Xing, J., & Heeger, D. J. (2001). Measurement and modeling of center-surround suppression and enhancement. *Vision Research, 41*, 571–583.

Yoshimura, Y., Sato, H., Imamura, K., & Watanabe, Y. (2000). Properties of horizontal and vertical inputs to pyramidal cells in the superficial layers of the cat visual cortex. *The Journal of Neuroscience, 20*, 1931–1940.

Zanker, J. M. (1999). Speed tuning in elementary motion detectors of the correlation type. *Biological Cybernetics, 90*, 109–119.