LETTER Communicated by Hugh Wilson

Suprathreshold Intrinsic Dynamics of the Human Visual System

Gopathy Purushothaman
gopathy@uchicago.edu
Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77204-4005, U.S.A.

Haluk Öğmen
ogmen@uh.edu
Department of Electrical and Computer Engineering and Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-4005, U.S.A.

Harold E. Bedell
hbedell@optometry.uh.edu
College of Optometry and Center for Neuro-Engineering and Cognitive Science, University of Houston, Houston, TX 77204-4005, U.S.A.

Intrinsic high-frequency neural activities have been observed in the visual system of several species, but their functional significance for visual perception remains a fundamental puzzle in cognitive neuroscience. Spatiotemporal integration in the human visual system acts as a low-pass filter and makes the psychophysical observation of high-frequency activities very difficult. A computational model of retino-cortical dynamics (RECOD) is used to derive experimental paradigms that allow psychophysical studies of high-frequency neural activities. A reduced-parameter version of the model is used to quantitatively relate psychophysical data collected in two of these experimental paradigms. Statistical analysis shows that the model's account of the variance in the data is, in general, highly significant. We suggest that psychophysically measured oscillations reflect intrinsic neuronal oscillations observed in the visual cortex.

1 Introduction

In 1912, Fröhlich made the "assumption that the sense organs . . . respond to stimuli with rhythmical excitations" (cited in Ansbacher, 1944). He distinguished between rhythmic responses to nonrhythmic stimuli (intrinsic oscillations) and rhythmic responses to rhythmic stimuli (entrained oscillations). Since then, electrophysiological recordings have revealed oscillatory activities of both types in the visual system of several species (Başar & Bul-

lock, 1992). The role of intrinsic oscillations in visual perception remains a fundamental puzzle in cognitive neuroscience. It has been suggested that temporal coordination of rhythmic activity can serve as a neural code underlying visual awareness, focused arousal, attention, and feature binding (Milner, 1974; von der Malsburg, 1981; Sheer, 1984; Eckhorn et al., 1988; Gray & Singer, 1989; Crick & Koch, 1990; Llinas, 1992; Engel, Fries, Konig, Brecht, & Singer, 1999). Tests of these proposals yielded controversial results (Fahle & Koch, 1995; Kiper, Gegenfurtner, & Movshon, 1996; Leonards, Singer, & Fahle, 1996; Blake & Yang, 1997; Elliott & Müller, 1998; Alais, Blake, & Lee, 1998; Usher & Donnelly, 1998). One common aspect of these studies was the use of entrained oscillations. However, the limitation of entrained oscillations in probing the functional roles of intrinsic oscillations has been highlighted (Fahle & Koch, 1995; Leonards et al., 1996; Blake & Yang, 1997).

How can the intrinsic dynamics of the visual system be studied psychophysically? The complete dynamics of a linear time-invariant (LTI) system can be determined by its impulse response in the time domain or, equivalently, by its transfer function in the frequency domain. This LTI methodology has been applied to the human visual system, with the finding that the morphology of the estimated impulse response depends strongly on stimulus parameters, such as size, luminance, and adaptation level (Kelly, 1961; Ikeda, 1986; Georgeson, 1987; Manahilov, 1995). This is expected from the fact that the human visual system is nonlinear and time variant. Early phenomenal observations in response to static or moving stimuli indicated a repetitive aspect of perception for stimuli well above detection threshold, possibly reflecting high-frequency oscillations (Bidwell, 1899; McDougall, 1904). However, the human visual system is known to integrate signals over space and time (Davson, 1990). This integration would act as a "low-pass" filter and make the psychophysical observation of high-frequency oscillations difficult.

A first approach to bypass the integration stage would be to obtain an independent estimate of the integration dynamics and to process the data with an inverse filter of the integration stage. However, the nonlinear and time-varying nature of the visual system, coupled with the difficulties in isolating the integration process, make the estimation of this inverse filter a challenging problem. Moreover, because the inverse filter would be essentially a high-pass filter, it would be extremely sensitive to noise.

A second approach would be to design a stimulus paradigm that leads to interactions between the neural responses before the integration stage. If these interactions result in modified activities that reflect the intrinsic oscillations even after the integration stage, then one could relate the psychophysically measured performance to the underlying intrinsic neural activity. In this study, we follow the second approach.

2 General Architecture of the RECOD Model

A model of retino-cortical dynamics (RECOD) (Öğmen, 1993), whose general structure is shown in Figure 1, constitutes the theoretical basis of our study.

The primate retina contains two major populations of ganglion cells: one population with fast phasic (transient) responses and a second population with slower tonic (sustained) responses (e.g., Gouras, 1968; De Monasterio & Gouras, 1975; Croner & Kaplan, 1995; Kaplan & Benardete, 2001). The lower two ellipses in Figure 1 represent these two populations of retinal ganglion cells in our model. The typical responses of these neurons to a pulse input are shown in the figure (denoted by R). These populations project to distinct layers of the lateral geniculate nucleus (LGN), forming two parallel afferent pathways (the magnocellular and the parvocellular), as shown in the figure. In turn, magnocellular and parvocellular pathways form the primary inputs of different visual areas subserving various functions, such as the computation of motion, form, and brightness (Livingstone & Hubel, 1988). However, at the cortical level, these two pathways interact (Van Essen, Anderson, & Felleman, 1992) and the loci/degree of their interactions are not fully established (Sincich & Horton, 2002).

The model uses a lumped representation for the cortical targets of magnocellular and parvocellular pathways. The cortical targets of the magnocellular pathway represent the areas that play a major role in the computation of motion and temporal change. The cortical targets of the parvocellular pathway represent the areas that play a major role in the computation of dynamic form and brightness (see the upper ellipses in Figure 1). In terms of cortical interactions between these pathways, the model postulates reciprocal inhibition, as shown by the arrows between the upper ellipses in the figure. The lumped representation for the areas involved in the computation of dynamic form and brightness contains recurrent connections (not shown in the figure) to represent the extensive feedback observed in postretinal areas (Van Essen et al., 1992). This recurrent circuit produces oscillatory responses (denoted by P in the figure). As noted, oscillatory responses have been observed in the visual system of several species, ranging from invertebrates to vertebrates (Başar & Bullock, 1992). Oscillatory activities are ubiquitous in human electroencephalograms (EEGs) (Başar, 1998). Oscillatory activities have also been recorded intracerebrally in the primate visual cortex (Doty & Kimura, 1963; Maunsell & Gibson, 1992; Livingstone, 1996; Friedman-Hill, Maldonado, & Gray, 2000; Maldonado, Friedman-Hill, & Gray, 2000; Fries, Reynolds, Rorie, & Desimone, 2001; Rols, Tallon-Baudry, Girard, Bertrand, & Bullier, 2001).

To link neural responses to visual perception, the model postulates that the oscillatory activity is temporally integrated (ΣP in Figure 1) according to the psychophysically determined temporal-integration characteristics of the

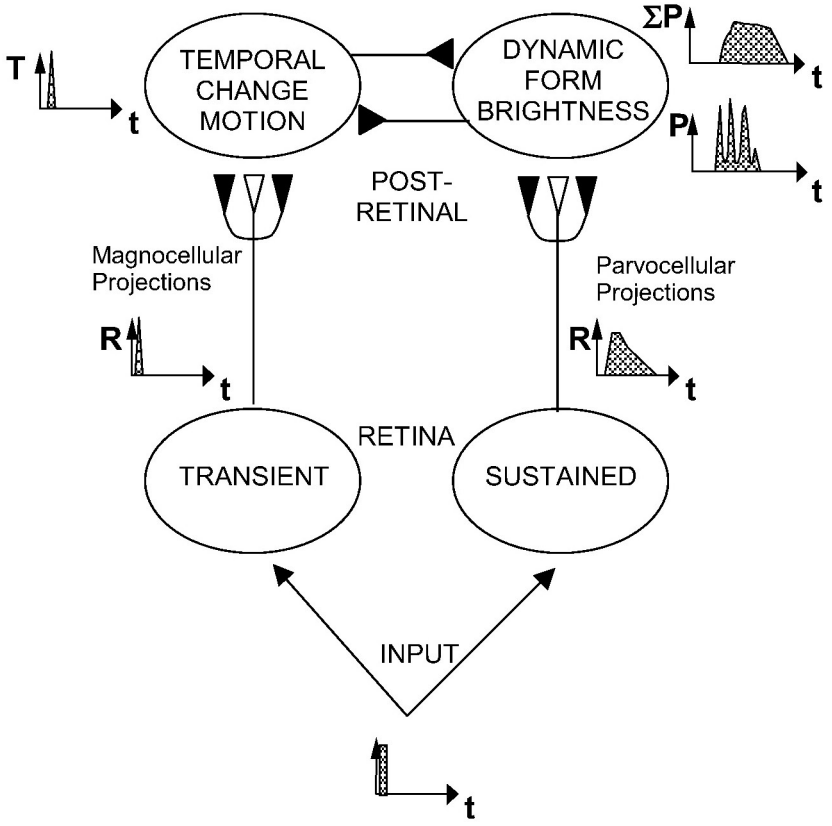


Figure 1: A schematic representation of the pathways and the neural populations in the retino-cortical dynamics (RECOD) model. The open and filled triangular symbols depict excitatory and inhibitory interactions, respectively. The lower two ellipses correspond to populations of retinal ganglion cells with fast-phasic (transient) and slow-tonic (sustained) response properties. The typical responses (denoted by R) of these neurons to a pulse input are shown in the figure. The upper-right ellipse represents the postretinal areas that play a major role in determining the perceived form and brightness of stimuli. It contains internal recurrent connections (not shown in the figure) and produces oscillatory responses (denoted by P in the figure). According to the RECOD model, the major excitatory drive for these neurons comes from sustained retinal ganglion cells. The upper-left ellipse represents a population of neurons that encode temporal change and motion characteristics of stimuli. Their response is transient (denoted by T in the figure). The model includes reciprocal inhibitory interactions between neurons in the transient and the sustained pathways. In particular, it is suggested that inhibitory signals from the afferent transient pathway are used to inhibit or reset the activities in the recurrent loops of the sustained pathway.

human visual system[1] and that this temporally integrated activity underlies the perception of brightness and form.

3 Prediction of the RECOD Model for Metacontrast Masking

Metacontrast masking is an experimental paradigm that is widely used to probe the dynamics of the visual system, typically within the context of nonlinear models (Breitmeyer, 1984; Bachmann, 1994; Breitmeyer & Öğmen, 2000). In this paradigm, the visibility of a target stimulus is measured in the presence of a spatially neighboring but nonoverlapping mask stimulus. The mask stimulus is turned on after the target stimulus with a delay between the onset of mask and target, called the stimulus onset asynchrony (SOA) (see Figure 2). Due to the interactions between the responses generated by the target and mask stimuli, the visibility of the target is affected by the mask stimulus. Based on extensive experimental studies, metacontrast masking functions (visibility of the target as a function of SOA) have been categorized into two general types: type A (monotonic), where peak masking occurs at SOA $= 0$, and type B (nonmonotonic or U-shaped), where peak masking occurs at a positive value of SOA (Breitmeyer, 1984; Bachmann, 1994). An analysis of the RECOD model in the context of metacontrast masking produced a novel prediction: that spatiotemporally localized (briefly flashed dots) suprathreshold stimuli should produce a new type of metacontrast masking function that exhibits gamma range (30–70 Hz) oscillations (Purushothaman, Öğmen, & Bedell, 2000). The metacontrast stimulus is shown in Figure 2, and the prediction of the model is illustrated in Figure 3A.

The plots at the top of Figure 3A depict the postretinal transient activities that the target (T) and the mask (M) would generate if they were presented in isolation. The triangle with the dashed pattern depicts the fast transient response in postretinal areas driven by the magnocellular pathway and the sawtooth-shaped activity depicts the oscillatory response in postretinal areas driven by the parvocellular pathway. The arrows between activities

Figure 1: *continued*. The model also postulates that the oscillatory activity is temporally integrated (ΣP) according to the temporal integration characteristics of the human visual system prior to the perception of brightness and form and that this temporally integrated activity underlies the perception of brightness and form. A detailed description and formal equations of the model can be found in Purushothaman, Öğmen, Chen, and Bedell (1998) and Purushothaman, Öğmen, and Bedell (2000).

[1] ΣP in Figure 1 represents the stage of integration that occurs after the generation of oscillatory activities in postretinal areas that determine perceived brightness and form. Integration also occurs at earlier levels, for example, as indicated in Figure 1 by the sluggish response of retinal-sustained neurons to a pulse input.

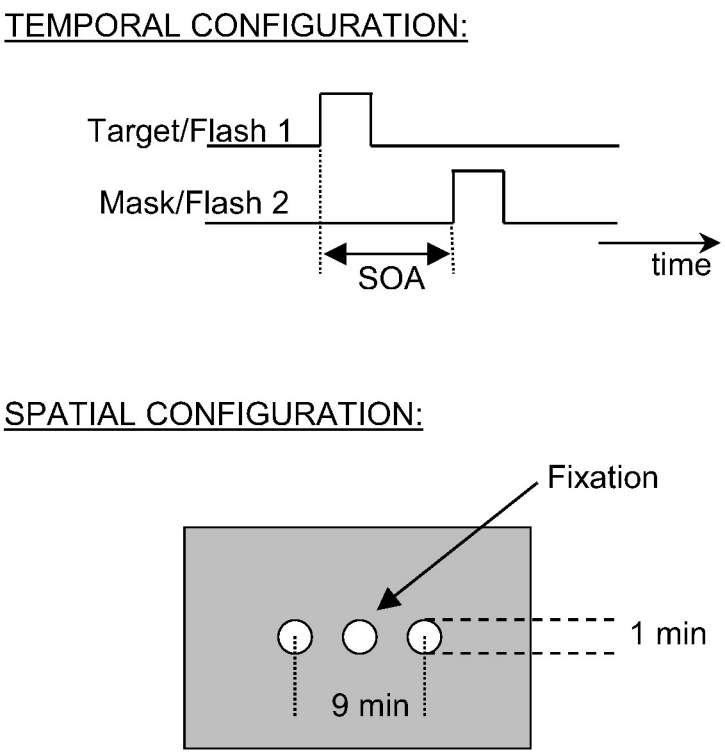


Figure 2: Stimulus for the metacontrast experiment. On each trial, the target was presented 4.5 min arc to the left or the right of a fixation dot, and the mask was presented on the opposite side. Note that in each trial, the fixation dot was turned off 200 ms before the first stimulus was presented.

illustrate the inhibitory interactions between the transient and oscillatory activities. When the SOA is such that the mask-generated transient activity overlaps in time with the bursting phase of the target-generated oscillatory activity (SOA_1, on the left), this burst of activity will be suppressed due to the inhibition exerted by the transient activity. This will cause a decrease in the integrated activity generated by the target. As a result, the visibility of the target will decrease for SOA1, as shown in the metacontrast function at the bottom of the figure. When the SOA is such that the mask-generated transient activity occurs between two bursting phases of the oscillatory activity (SOA_2, on the right), inhibition will not suppress significantly the oscillatory activity in the sustained channel. This will result in a relatively high level of perceived target brightness, as shown in the metacontrast function at the bottom of the figure.

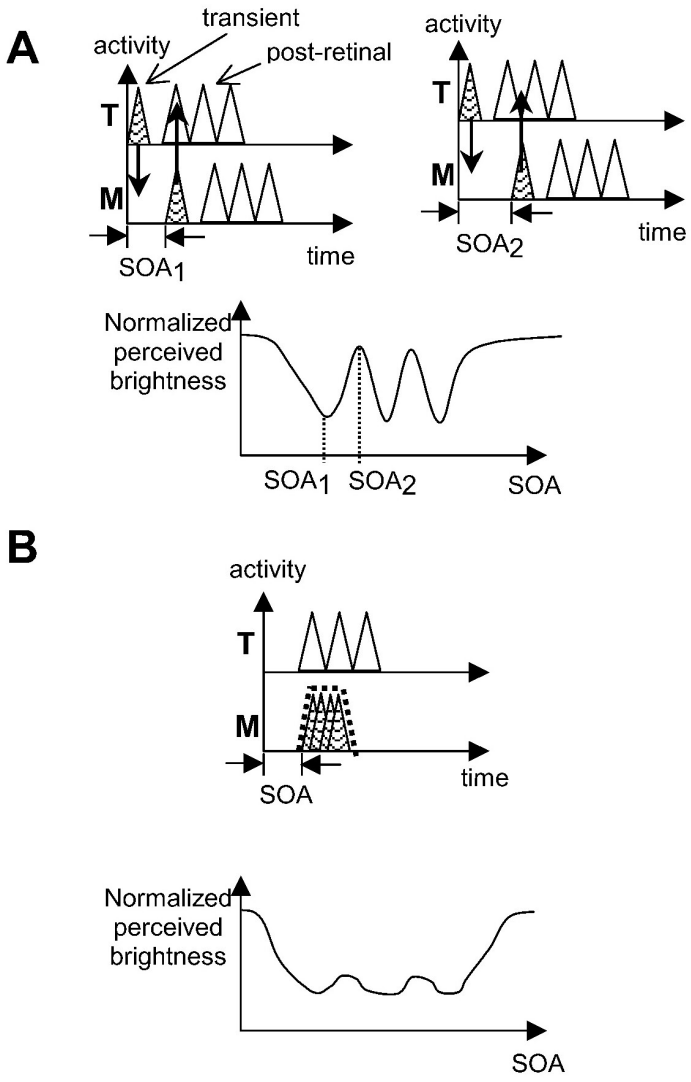


Figure 3: Graphical illustration of the model predictions for metacontrast masking. (A) Schematic representation of activities generated by the target (T on the y-axis) and the mask (M on the y-axis) for two different values of SOA. The arrow indicates inhibitory interactions. The bottom panel depicts the predicted metacontrast function. (B) Depiction of the oscillatory activity generated by the target (T) and the temporally spread transient activity generated by the mask (M) for spatiotemporally extended stimuli. The transient activity generated by the mask is spread in time due to the temporal combination of transient responses generated at the temporal and/or spatial edges of the mask stimulus. The bottom panel illustrates the predicted metacontrast function (from Purushothaman et al., 2000).

When a range of SOA values is considered, one can see that the shape of the metacontrast function will reflect the shape of the oscillatory activity. In order to suppress selectively the different bursting phases of the oscillations, the mask-generated transient activity should be relatively brief with respect to the period of the bursts. We suggest that this can be achieved by using a spatiotemporally localized stimulus—a small, brief flash. Using a temporally extended stimulus would generate a more temporally spread-out transient activity due to the temporal spread of the response components generated at the onset and offset of the stimulus. Similarly, using a spatially extended stimulus would generate a more temporally spread-out transient activity due to response components generated at the edges of the mask stimulus when they combine temporally with a delay that depends on their distance from the target. As a result, spatiotemporally extended stimuli would generate a masking function with significantly less prominent oscillations, as shown in Figure 3B. It is also important to note that the sampling of the SOA axis in the experiment should be high enough to reveal rapid fluctuations in the metacontrast function. The three columns in Figure 4 show psychophysical metacontrast data from three observers. The top three rows correspond to three individual runs, and the bottom row plots the average of these three runs. The magnitude of the metacontrast functions is variable across observers ranging from about two log units to one-half log unit. However, as the comparison of the three runs for each subject shows, the magnitude of the metacontrast function is relatively stable within each subject. Moreover, all the observers show oscillatory masking functions. The location of the dips across different runs remains consistent for observers GP and HO but is less so for observer BG. Statistical analysis of the average data using the Duncan multiple-range grouping test supports oscillations in the masking functions of all observers. For observer GP, there are three statistically significant ($p = 0.05$) dips, located at 45–55, 65–75, and 85–90 ms. For observer HO, there are four significant dips, located at 40, 50, 60–75, and 85–100 ms. For observer BG, there are two significant dips, located at 40 and 80 ms.

4 Prediction of the RECOD Model for the Perceived Number of Flashes

An investigation using the metacontrast masking paradigm was the first step in our approach to study the intrinsic dynamics of the human visual system. In this article, we use a related but independent experimental paradigm in order to establish more firmly and to cross-validate the psychophysical observation of intrinsic oscillations. All experimental parameters were identical to those used in the previous metacontrast experiment except that the target and mask dots were flashed successively at the location of a previously viewed fixation stimulus, shortly after it was turned off. This paradigm was shown to produce U-shaped masking functions for stimuli of larger dimensions when the SOA was coarsely sampled (Breitmeyer & Horman, 1981).

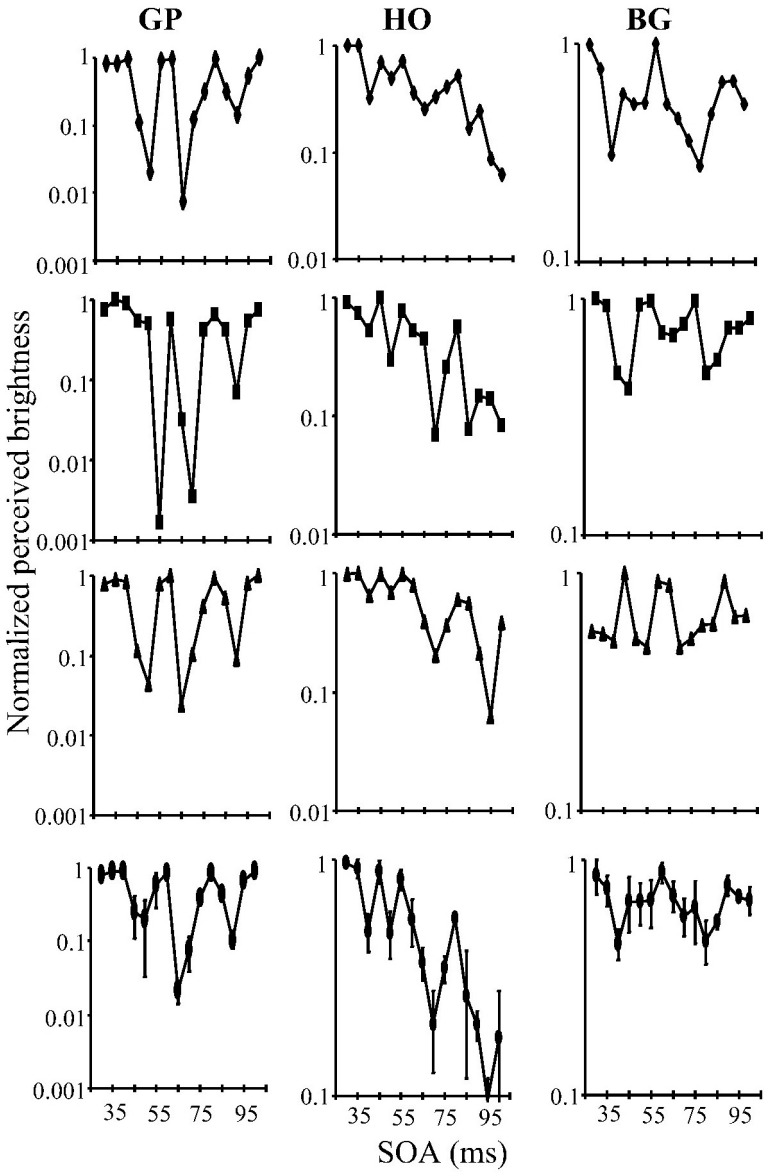
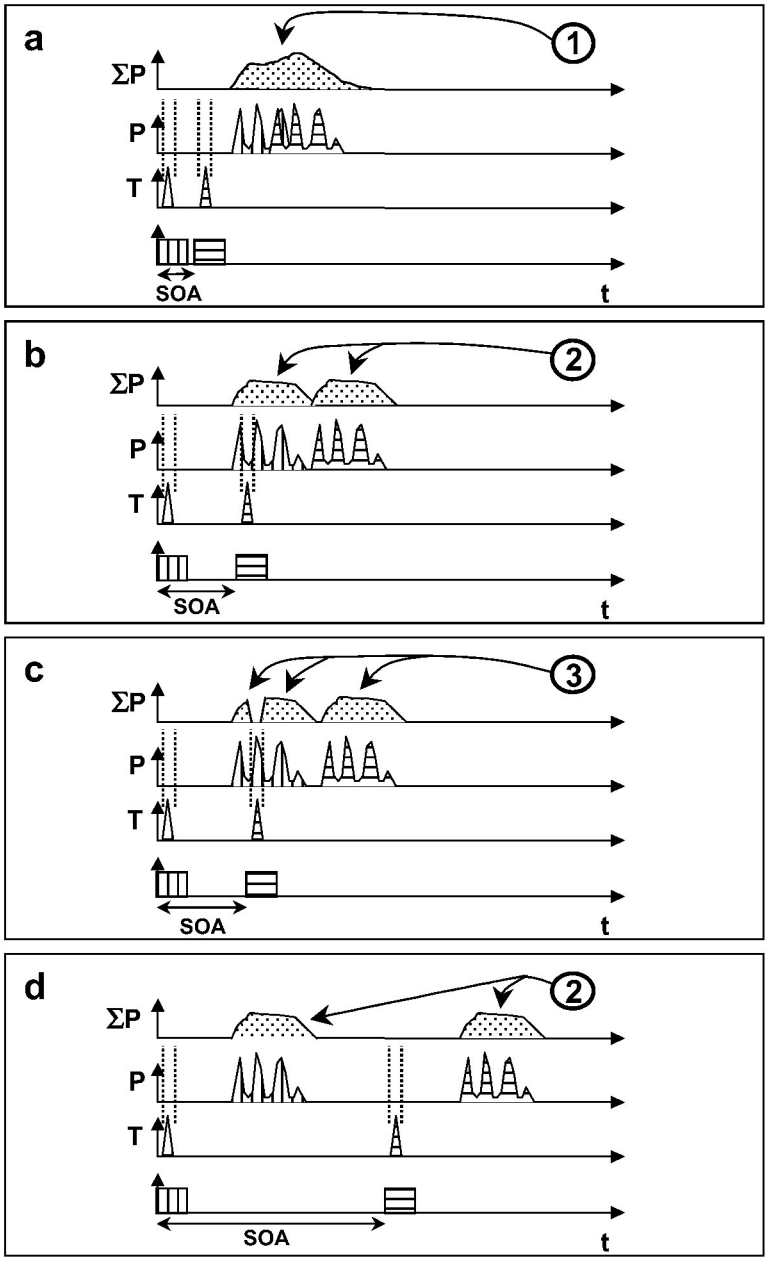


Figure 4: Metacontrast functions for observers GP (left column), HO (middle column), and BG (right column). The top three rows show individual runs, and the bottom row plots the average +/− SEM of these three runs. The data in the left two columns are from Purushothaman et al. (2000).

To obtain empirical evidence that is independent of our previous study, we changed the observers' task from brightness matching to a report of the perceived number of flashes. Previously, Bowen (1989) showed, with larger targets and coarse temporal sampling, that for short SOAs, the two flashes are integrated, and the observers report seeing a single flash. For large SOAs, there is little interaction between the flashes, and the observers report seeing two flashes. However, for intermediate SOAs, observers reported seeing three (or more) flashes even though physically only two flashes were presented. Bowen (1989) showed that this surprising finding can be explained if one assumes multiphasic (oscillatory) responses in the visual system. By applying the LTI approach and using a single channel conceptualization of the visual system, he derived an impulse response in the form of an exponentially damped sine wave with a 10 Hz frequency. Here, we extend Bowen's approach using finer sampling of the SOA axis to prevent potential aliasing effects that would limit the estimate of the oscillation frequency. In addition, the multichannel structure of the RECOD model produces a novel prediction for this paradigm, which is illustrated in Figure 5.

The first and the second flash and their resulting activities are depicted by vertical and horizontal dashing, respectively. The encircled numbers in each panel show the predicted number of perceived flashes. When the SOA is short (see Figure 5a), there is no temporal overlap between transient and sustained activities, and the summed activity of the two flashes yields a (temporally) single integrated postretinal response. Consequently, the postretinal responses to the two flashes are fused together, resulting in the perception of a single flash. As SOA increases, the postretinal activities of the two flashes gradually become temporally separated (see Figures 5b–5d). Figures 5b shows an intermediate SOA such that the transient inhibition from the second flash falls between two bursts in the postretinal response

Figure 5: *Facing page*. Graphical illustration of the model predictions for the number-of-flashes experiment. The first and the second flash and their resulting activities are depicted by vertical and horizontal dashing, respectively. The circled numbers in each panel show the predicted number of perceived flashes. The panels depict increasing values of SOA, with a corresponding to the shortest and d to the longest SOA. The vertical dashed lines from transient activities to postretinal oscillatory activities illustrate the time interval during which transient activities exert their inhibitory influence. When this time interval falls on an "off-phase" of the oscillation, as in b, the temporal profile of the oscillatory activity is not significantly modified. When this time interval falls during an on-phase of the oscillation, as in c, the inhibition suppresses the corresponding burst and generates a longer temporal gap between the previous and the following bursts. Because of this temporal gap, a decrease in activity persists even after the integration stage producing a response profile with three temporally segregated regions.


a
ΣP
P
T
SOA
t
1
b
ΣP
P
T
SOA
t
2
c
ΣP
P
T
SOA
t
3
d
ΣP
P
T
SOA
t
2

to the first flash. Hence, the postretinal response to the first flash remains relatively unaffected, resulting in two distinct regions in the integrated activity and a percept of two flashes. When the SOA is further increased such that the transient response to the second flash occurs during the on-phase of the oscillation (see Figure 5c), the corresponding burst is suppressed. This increases the time interval between the temporally adjacent bursts. This extended off-phase is preserved in the integrated activity from the first flash, resulting in two temporally distinct epochs of postretinal activities. Because the second flash generates an additional temporally separated activity, observers are predicted to report seeing three flashes. It can be inferred that if the SOA is gradually increased, the probability of reporting two and three flashes will alternate temporally following the oscillation pattern of the postretinal activity. When the SOA becomes large (see Figure 5d), the two flashes generate transient and sustained activities that no longer overlap temporally, and the responses to the two flashes will remain distinctly separate. Therefore, for large SOAs, the model predicts the perception of two distinct flashes.

Thus, in addition to agreeing with Bowen's (1989) findings mentioned above, the model makes the prediction that the shape of the probability curves for perceiving one, two, and three flashes should depend on the oscillations observed in the metacontrast masking functions. Note that because the observersí responses are distributed in three distinct categories in the number of dots paradigm, oscillations are expected to be less prominent in each of the probability functions. However, the shape of the probability curves should depend on the oscillatory profiles of the metacontrast functions. To test this hypothesis, we examine whether the model, whose response profiles are derived from our subjects' metacontrast data, provides a statistically significant account of the probabilities of perceiving different numbers of flashes. While the graphical depiction of the model's prediction is intended to convey the main idea, the specific quantitative prediction of when the observers are expected to report seeing one, two, or three flashes will depend on the relative strength of each region of the response and how much a region is temporally separated from other regions. The strength and separation of the regions, in turn, depend on several parameters, such as the duration of neural responses, the oscillation profiles, and the latencies of the signals. In section 6.2, we introduce the metrics used to quantify signal strength and segregation in order to translate model-generated neural activities into response probabilities. Furthermore, by incorporating the empirical metacontrast data of the observers directly to the model, we carry out the quantitative analysis of the data using only four free parameters.

5 Parametrically Reduced Version of the RECOD Model

The RECOD model is defined by a system of nonlinear differential equations whose parameters reflect neurophysiological properties of the retinocortical

pathways. A detailed description and formal equations of the model can be found in Purushothaman et al. (2000) and Purushothaman, Öğmen, Chen, and Bedell (1998). Because in this study we are considering two psychophysical measures obtained using identical stimulus parameters, with the only exception of spatial position of the stimuli, we were able to minimize the number of free parameters involved in curve fitting by using a simplified version of the model. According to our model, metacontrast masking functions reflect the temporal profiles of postretinal responses for spatiotemporally localized stimuli. Consequently, we approximated the postretinal response to a 10 ms flash of intensity 2.5 log units above detection threshold by using the metacontrast function of each observer. Thus, if we denote by $\rho_T(t)$ the postretinal oscillatory response to the target, this response was set to

$$\rho_T(t) = 1 - M(\text{SOA} = t),$$

where M(SOA) is the normalized metacontrast masking function, so that the dips in the metacontrast masking correspond to peaks (bursts) in the postretinal response.[2] In the number-of-flashes experiment, the target and the mask stimuli are flashed at the same spatial location, and the composite postretinal response to the target and mask stimuli, which were flashed at the same spatial location, at a given SOA, $\rho(t, \text{SOA})$, was computed as

$$\rho(t, \text{SOA}) = \rho_T(t) + \rho_M(t),$$

with $\rho_M(t) = \rho_T(t - \text{SOA} + \phi)$. The latency of signals in the visual pathways depends on stimulus parameters (e.g., Mansfield, 1973; Purushothaman, Patel, Bedell, & Öğmen, 1998; Maunsell et al., 1999). The parameter ϕ reflects the decrease in the latency of the second flash due to the facilitation provided by the first flash. When the transient activity (T in Figure 1) temporally coincided with the postretinal activity in the sustained pathway, the net postretinal activity, $P(t, \text{SOA})$, including the inhibitory effect of the transient activity, was approximated by a complete suppression of activity at that time instant:

$$P(t, \text{SOA}) = \rho(t, \text{SOA})\{1 - \delta(t - \text{SOA} + \theta)\},$$

where $\delta(t)$ is the Kronecker delta function. The parameter θ reflects the reduction in latency for the transient activity generated by the second flash.

[2] Note that this formulation is not intended to be general but constitutes an approximation for the experimental conditions and results in hand. For example, in the absence of masking, the formula yields 0 as the postretinal response and does not correspond to an uninhibited postretinal response. However, the formula is a valid approximation for our data, where all observers showed masking effects.

6 Methods

6.1 Psychophysical Methods. The experiment was performed on the same set-up used in our previous metacontrast masking experiments (Purushothaman et al., 2000). All experimental parameters and procedures were the same except that the two flashes (Target and Mask) were spatially overlapping. The stimuli consisted of 1 arc min diameter dots, set to a luminance of 2.5 log units above their detection threshold. The dots were flashed foveally at the center of a homogeneous background field of luminance 13 cd/m^2. The duration of the dot stimuli was 10 ms. The onset of the second dot was delayed by SOA values that ranged from 30 to 100 ms in 5 ms steps. To minimize order effects and potential observer biases, the SOA values were interleaved randomly within each run. The probability of perceiving one, two, and three or more flashes for each SOA was computed based on a total of 90 responses collected in three sessions. Data were collected on three observers (GP, HO, and BG, an observer naive to the purpose of the experiment at the time of data collection). For observers GP and HO, we used metacontrast data from a previous study (Purushothaman et al., 2000) and collected only data for the number-of-flashes experiment. For observer BG, we collected in each session both metacontrast and the number-of-flashes data. The order of the experiments was randomized across three sessions run on three different days. In the number-of-flashes experiment, the observers were asked to report, through a joystick, whether they saw one, two, or three flashes. The details of the metacontrast experiment are given in Purushothaman et al. (2000).

6.2 Modeling Methods. As explained schematically in Figure 5, the number of perceived flashes depends on the number of distinct regions in the integrated postretinal response. In order to quantify this and to convert it to probability measures, we used two indices reflecting the relative energy and temporal segregation of each region. The postretinal activity generated by an isolated presentation of a single flash, $\rho_T(t)$, produces the perception of a single flash. It served as a reference in the computation of these indices. The first index is given by

$$M_1(\mathrm{SOA}) = \Sigma_{t\in T1} P(t, \mathrm{SOA}) / \Sigma_{t\in D} \rho_T(t, \mathrm{SOA}),$$

where $T1$ corresponds to the temporal span of the region under consideration and D is the duration of the response to a single flash. The second index is given by

$$M_2(\mathrm{SOA}) = \Sigma_{t\in T1} \{P(t, \mathrm{SOA}) - \rho_T(t, \mathrm{SOA})\}^2 / \Sigma_{t\in D} \{\rho_T(t, \mathrm{SOA})\}^2.$$

This provides a normalized measure of the difference between the composite response in the region under consideration and the reference response (single flash) and thus provides an indication of the temporal separation

between regions. The product of these two measures was used as an estimate of that region's likelihood to produce the percept of a separate flash. We normalized this product by

$$L = \{M_1M_2 - \min(M_1M_2)\}/\max(M_1M_2).$$

At short SOAs, the combined response to the two flashes was considered in two regions: the first region running from the beginning of the response to the point at which the transient inhibition had suppressed the response, and the second region running from that point onward to the end of the response. For each region, L was calculated, and the probability of reporting two flashes was computed by $G(S + L_1L_2)$, where G and S are linear scaling parameters. Because at short SOAs, the response consisted of only two regions, the probability of reporting three flashes was 0. Thus, the probability of reporting one flash was (1 − probability-of-reporting-two-flashes). After a critical SOA, the combined response to the two flashes starts showing either two or three distinct regions. This critical SOA depends on the duration D, the location of bursts in $\rho_T(t)$, and the latency difference between sustained and transient activities. After that critical SOA, the probability of reporting one flash became 0. The probability of reporting two flashes was computed using the same procedure described above, and the probability of reporting three or more flashes was set to (1 − probability-of-reporting-two-flashes). The four free model parameters that were customized to fit each observer's data are G, S, ϕ, and θ. Their values are given in Table 1. Note that all calculations were done in discrete time steps of 5 ms (imposed by the sampling rate of the metacontrast masking function) and are limited by this resolution.

7 Results

Psychophysical results for the three observers are shown by the filled squares in Figures 6, 7, and 8. The averaged probability of perceiving one, two, and three flashes is shown in separate panels as functions of SOA. For all observers, the probability of perceiving a single flash decays by an SOA value of 70 ms. The probability of seeing two flashes starts to increase for SOA

Table 1: Values of the Free Parameters Obtained by Curve Fitting.

Observer	G	S	ϕ (ms)	θ (ms)
GP	2.3	0.06	15	0
HO	8.2	0.01	10	5
BG: Averaged data (Fig. 8) and runs 1 and 2 in Fig. 10.	4.5	0.04	15	0
BG: Run 3 in Fig. 10.	3.7	0.03	15	0

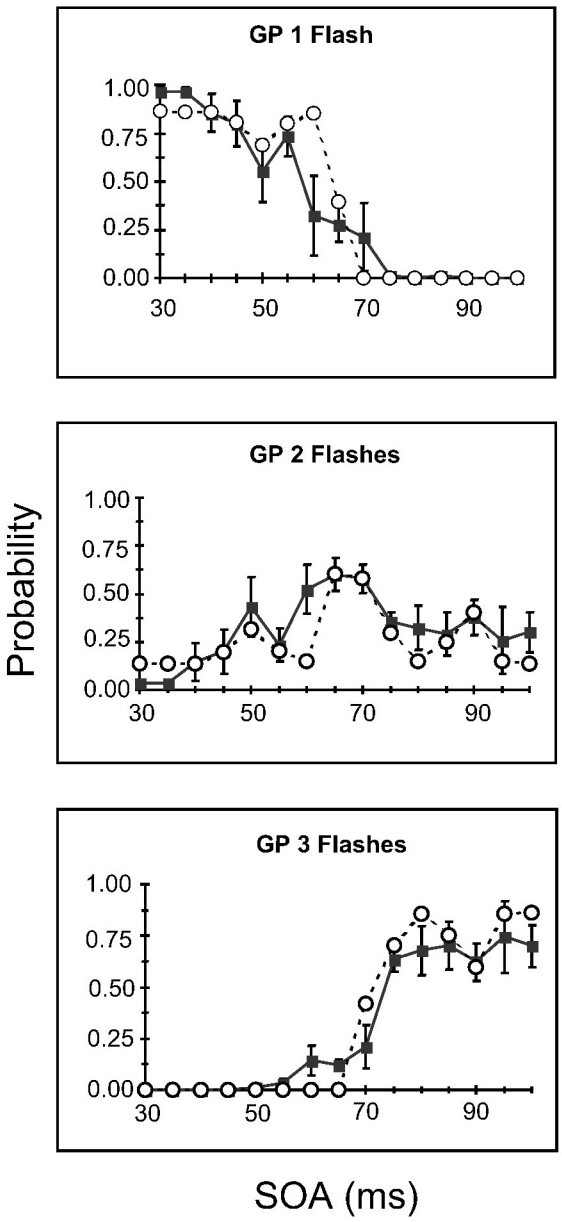


Figure 6: The probability of perceiving one (top panel), two (middle panel), and three or more (bottom panel) flashes as a function of SOA (filled squares) along with corresponding model predictions (open circles) for observer GP. The data represent the mean +/− SEM of the three runs.

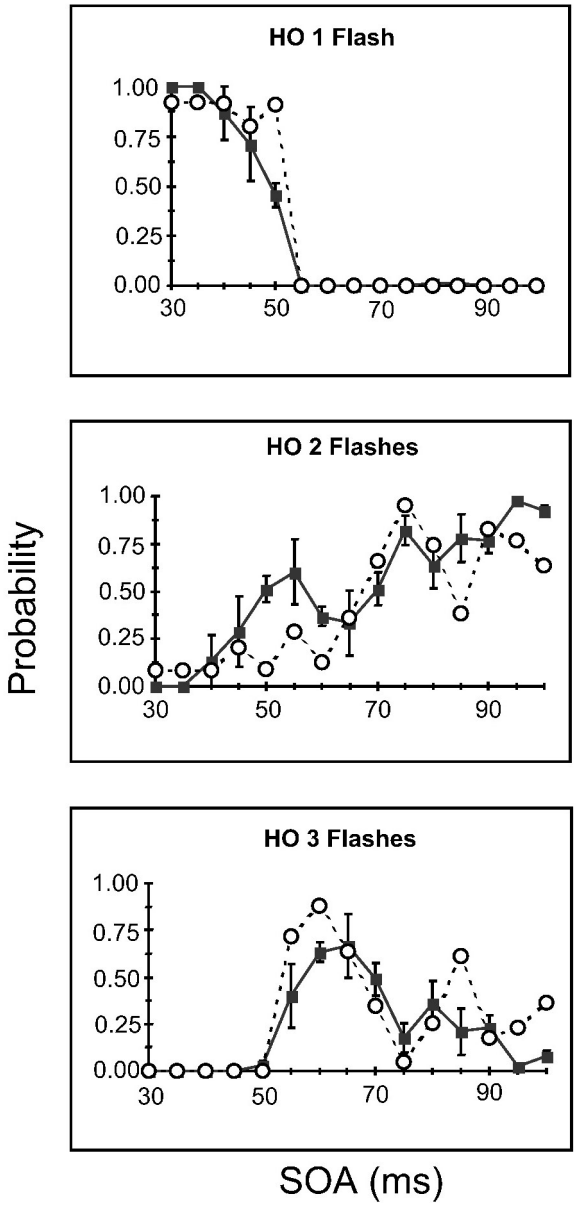


Figure 7: Same as Figure 6, for observer HO.

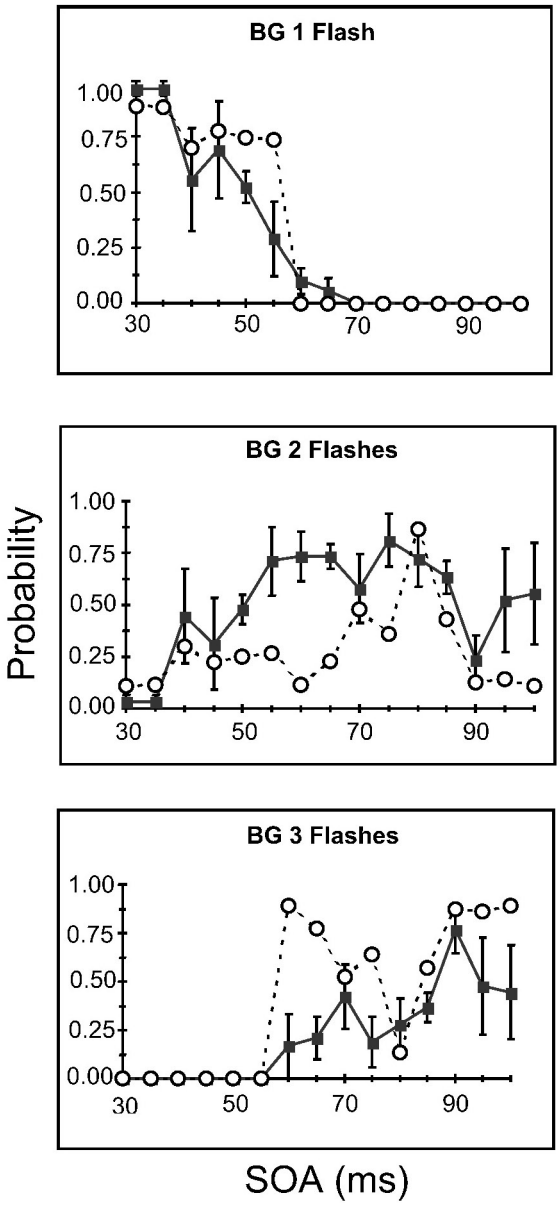


Figure 8: Same as Figure 6, for observer BG.

values of 40 ms. For SOA values larger than 50 ms, observers' responses comprise one-, two-, and three-flash categories. Moreover, the probability functions of individual observers are not smooth and exhibit an oscillatory character. The data averaged across the three observers are shown in Figure 9. The curves are now much smoother, indicating that the oscillatory profiles of the three observers are not in phase. For comparison, the data averaged from the three observers from Bowen's study are superimposed on our averaged data. Bowen's study used a lower sampling frequency of the SOA axis but covered a wider range of SOA values. The data points shown in the figure are those that fall into the range used in our study. Quantitatively, the probability of perceiving one flash decays faster in our study, which results in a relatively higher probability of two-flash responses in our data. Three-flash responses are similar. Overall, and notwithstanding the parametric differences in the two studies, one can see that the two data sets are qualitatively similar.

Are the oscillations in the probability functions random fluctuations due to experimental noise, or do they reflect intrinsic oscillations of visual processes, as suggested by the model? By using two spatially adjacent dots of identical characteristics compared to the stimuli used here, we found that the perceived brightness as a function of SOA exhibits an oscillatory character (Purushothaman et al., 2000). If the oscillations were due to experimental noise, one would expect very little correlation between the oscillations in the two experimental paradigms, in particular given that the observers' task was quite different in the two experiments. On the other hand, if the oscillations reflect a common underlying mechanism, such as intrinsic neuronal oscillations, as suggested by the RECOD model, there should be a significant correlation between the two data sets. To test this hypothesis, we established a link between the two data sets: perceived brightness of the first dot as a function of SOA and the perceived number of dots as a function of SOA, based on the RECOD model. Model predictions were fitted to the data by using each observer's metacontrast masking data and by adjusting four free parameters (see section 6). The results are shown by the open circles superimposed to the data for each observer in Figures 6, 7, and 8. For observer GP, the model is in good agreement with data in both the general shape and the locations of the peaks.

For one-flash responses, both the data and the model dip at SOA = 50 ms. The data peak at SOA = 55 ms and decay thereafter to a probability of 0%. The model peaks at SOA = 60 ms and decays thereafter. Note that all computations were carried out with a 5 ms resolution limit imposed by the data. For two-flash responses, both the data and the model show peaks at SOA values of 50 ms, 65 ms, and 90 ms. For three-flash responses, the model predicts a dip at SOA = 90 ms flanked by two peaks. In the data, although there is a dip at SOA = 90 ms, the magnitudes of the peaks are smaller than those in the model. The probability function starts to rise at an SOA of approximately 60 ms—10 ms earlier than the model.

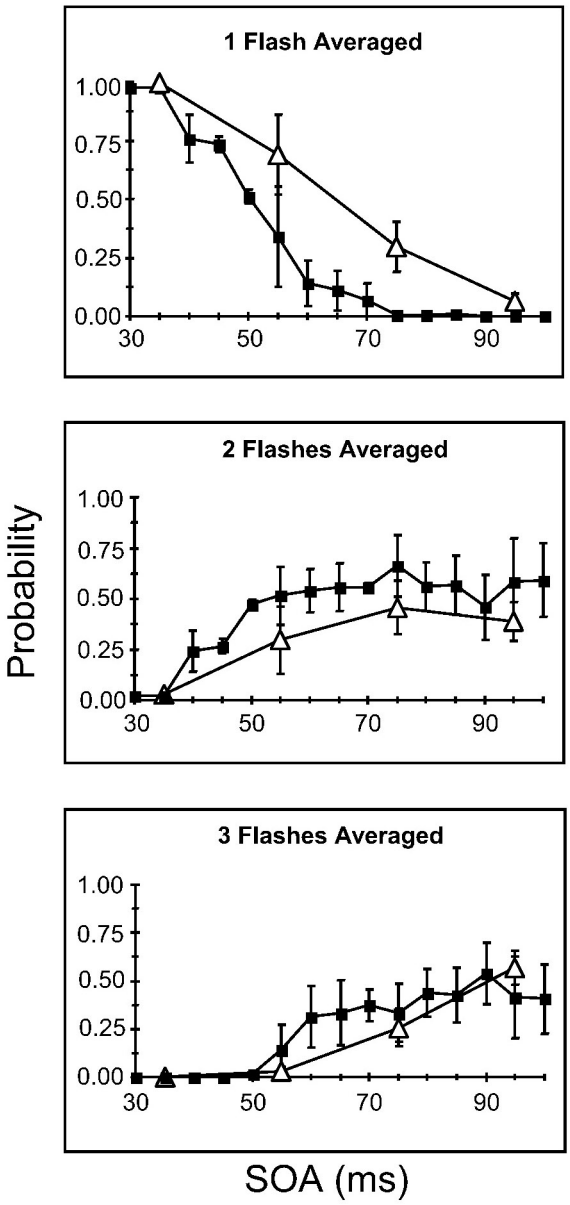


Figure 9: The data of Figures 6, 7, and 8 averaged across the three observers (filled squares) +/− SEM. Open triangles are the data averaged across the three observers +/− SEM from Bowen (1989).

For observer HO, there is good overall agreement between the model and the data, although some discrepancies can be noted. For one-flash responses, the model predicts a peak at SOA = 50 ms, which is not present in the data. For two-flash responses, the model predicts two smaller peaks instead of the larger peak at SOA = 55 ms. There is a good agreement for the location and magnitude of the peak at 75 ms. For larger SOAs, the model response drops below the data, and the next peak occurs 5 ms earlier in the model. For three-flash responses, there is a relatively good agreement for the location of the peaks, although quantitatively, the model predictions are higher than the data. For observer BG, while the model predicts well the location of the peak at SOA = 45 ms for the one-flash responses, it does not capture well the two- and three-flash responses. Quantitatively, the coefficients of determination for observers GP, HO, and BG are $r^2 = 0.82$, 0.73, and 0.20, respectively. The model's account of the variance in the data is highly significant for observers GP and HO ($F_{3,26} = 38.67$, $p = 9.9 \times 10^{-10}$ and $F_{3,26} = 23.63$, $p = 1.35 \times 10^{-7}$, respectively) but not significant for observer BG ($F_{3,26} = 2.23$, $p = 0.11$).

In order to investigate further the poor performance of the model for observer BG, we studied whether the unaccounted variance in his data was attributable to day-to-day variations in his responses. This hypothesis is suggested by the fact that neuronal oscillations can be modulated, for example, by arousal or attention (Molotchnikoff & Shumikhina, 1996; Munk, Roelfsema, König, Engel, & Singer, 1996; Steinmetz et al., 2000; Fries et al., 2001) and by our finding that sizable day-to-day variations in the metacontrast masking function occur in some observers (see Figure 4; see also Purushothaman et al., 2000). We considered the data collected on three different days and computed standard errors of the mean (SEM) for each SOA. As Figures 6, 7, and 8 suggest, the average (across all SOAs) SEM for observer BG is larger than those for the other two observers (average SEMs: GP: 0.074, HO: 0.058, BG: 0.096).

We took advantage of the fact that for observer BG, the metacontrast and the number-of-flashes data were collected during each session in an interleaved fashion. We estimated how much day-to-day variability contributes to the poorer fit for this observer by fitting the number-of-flashes data collected on a given day using the metacontrast masking data collected on the same day. As shown in Figure 10, this procedure yields better fits. The model gives a very good fit to data for runs 1 and 2, especially if we take into consideration the 5 ms resolution limit of the computations. For run 1, one-flash responses decay monotonically for both the model and the data. For two-flash responses, both the model and the data show peaks at SOA = 55 ms and 80 ms. The data show an additional peak at SOA = 100 ms, which is not present in the model. For three-flash responses, both the model and the data show a peak at SOA = 60 ms and a dip at SOA = 80 ms. Because the sum of probabilities across one-, two-, and three-flash responses is 1, the inability of the model to predict the peak at SOA = 100

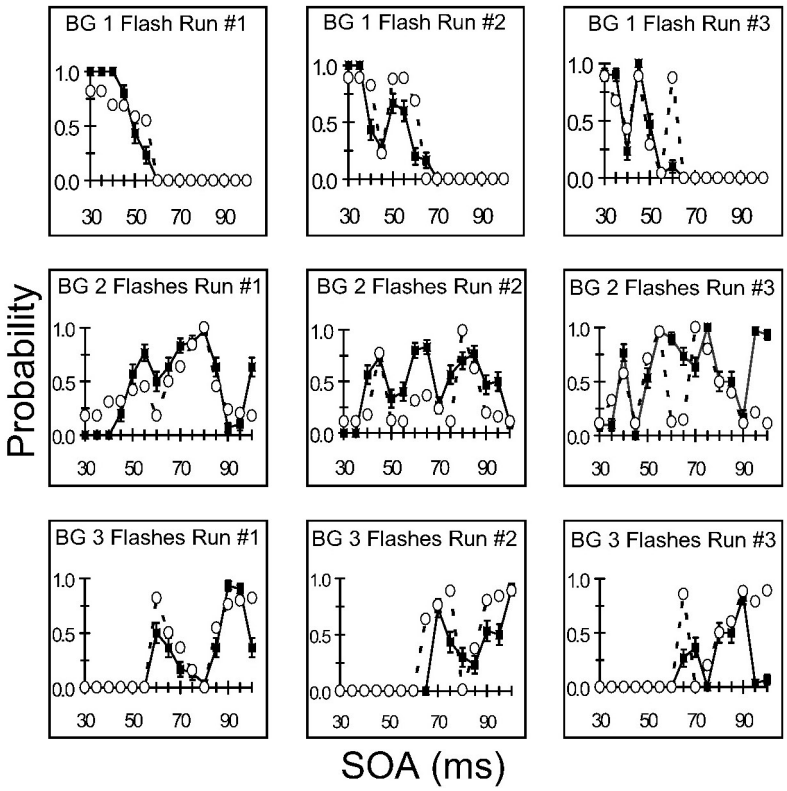


Figure 10: The probability of perceiving one (top panels), two (middle panels), and three or more (bottom panels) flashes as a function of SOA (filled squares) and the corresponding model predictions (open circles) for observer BG. The three columns represent the results of three different sessions. The error bars correspond to estimates based on the binomial distribution. Evaluation of the fit of the model to the data was conducted by calculating the pooled coefficient of determination ($r^2 = 0.46$) across the data sets collected on all three days.

ms for two-flash responses causes it to miss the dip at the same SOA for three-flash responses.

For run 2, the one-flash responses show a dip at SOA = 45 ms and a peak at 50 ms for both the model and the data. For two-flash responses, both the model and the data show peaks at SOA = 45 ms and 65 ms. The magnitude of the second peak is smaller in the model. The data show a third peak at SOA = 85 ms, which occurs 5 ms earlier in the model. For three-flash responses, the data show peaks at SOA = 70 ms and 100 ms. The model has peaks at SOA = 75 ms and 100 ms. For one-flash responses of run 3, while

the model predicts well the peak at SOA = 45 ms, it predicts an additional peak at 60 ms, which is not present in the data. For two-flash responses, the model predicts well the peaks at SOA = 40 ms and 55 ms and the dips at SOA = 45 ms and 90 ms. The peak at SOA = 75 ms occurs 5 ms earlier in the model. The model misses the peak at SOA = 95 ms, and the magnitude of the drop between the peaks at SOA = 55 ms and 75 ms is much larger in the model than in the data. For three-flash responses, the data show one peak at SOA = 70 ms and a second at 90 ms. The model has a peak at 65 ms and a second at 90 ms.

Although the timings of the peaks are close to those observed in the data, the model does not predict well the magnitude of the peak at 65 to 70 ms. In addition, the inability of the model to predict the peak at SOA = 95 ms for two-flash responses causes it to miss the dip at the same SOA for three-flash responses. All three runs taken together yield a coefficient of determination of $r^2 = 0.46$, and the model's account of the variance in the data is now highly significant ($F_{11,78} = 5.98, p = 5.57 \times 10^{-7}$). Overall, the model provides a good fit to all our data with the exception of run 3 for observer BG.

It is interesting to note that for runs 1 and 2, where we found a good match between the model and the data, the curve fitting used the same set of parameters. For run 3, where the match was poorer, the fit required a slightly different set of parameters. However, changing the parameters did not produce a match comparable to runs 1 and 2. We interpret this failure of the model as an indication that the model is not a universal approximator that can fit any set of data. Rather, when the model produces acceptable fits, it generates a consistent set of parameters.

8 Discussion

By using a computational model of retinocortical dynamics, a statistically significant correlation is found between oscillatory data obtained in two similar but independent experimental paradigms. The two authors who participated in the experiments (GP and HO) were not naive to the purpose of the experiment; however, their prior knowledge of the hypotheses is not likely to have influenced their performance, for the order of the SOA values was completely randomized during the experiment. Although the observer could clearly distinguish between a "short" and a "long" SOA, it is highly unlikely that the observer could modulate his performance according to a complex function of SOA that correlates with metacontrast masking functions collected in a previous study.

Under normal viewing conditions, observers do not perceive high-frequency oscillations that would result from intrinsic oscillations in the visual system. As shown in Figures 1 and 5, we suggest that the visual system filters these intrinsic oscillations through spatiotemporal integration before a conscious percept is formed. The specific experimental paradigms and

stimulus conditions used in our studies allow us to reduce the effect of this spatiotemporal integration and make the putative intrinsic oscillations measurable through psychophysical methods. Previously, the predictions of the RECOD model were shown to be in agreement with neurophysiological oscillatory data (Purushothaman et al., 2000), suggesting a possible link between the psychophysically measured oscillations and those observed in the visual cortex. Combined psychophysical and neurophysiological experiments in behaving animals could assess this link more directly. Our study does not directly address the question of whether intrinsic oscillations play a functional role in visual perception. However, methods like those described here can be applied to test psychophysically the suggested functional roles of intrinsic oscillatory neural activity, such as in feature binding.

Acknowledgments

This work was supported by grants R01-MH49892, R01-EY05068, and P30-EY07551 from the National Institutes of Health. We thank B. G. Breitmeyer, S. S. Patel, and E. L. Smith for helpful suggestions.

References

Alais, D., Blake, R., & Lee, S-H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience, 1*, 160–164.

Ansbacher, H. L. (1944). Distortion in the perception of real movement. *J. Exp. Psychology, 34*, 1–23.

Bachmann, T. (1994). *Psychophysiology of visual masking: The fine structure of conscious experience*. New York: Nova Science Publishers.

Başar, E. (1998). *Brain function and oscillations*. Berlin: Springer-Verlag.

Başar, E., & Bullock, T. H. (1992). *Induced rhythms in the brain*. Boston: Birkhäuser.

Bidwell, S. (1899). *Curiosities of light and sight*. London: Swan Sonnenschein and Co.

Blake, R., & Yang, Y. (1997). Spatial and temporal coherence in perceptual binding. *Proc. Natl. Acad. Sci. USA, 94*, 7115–7119.

Bowen, R. W. (1989). Two pulses seen as three flashes: A superposition analysis. *Vision Res., 29*, 409–417.

Breitmeyer, B. G. (1984). *Visual masking: An integrative approach*. New York: Oxford University Press.

Breitmeyer, B. G., & Horman, K. (1981). On the role of stroboscopic motion in metacontrast. *Bull. Psychonomic Soc., 17*, 29–32.

Breitmeyer, B. G., & Öğmen, H. (2000). Recent models and findings in visual backward masking: A comparison, review, and update. *Perception and Psychophysics, 62*, 1572–1595.

Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Semin. Neurosci., 2*, 263–275.

Croner, L. J., & Kaplan, E. (1995). Receptive fields of P and M ganglion cells across the primate retina. *Vision Res., 35*, 7–24.

Davson, H. (1990). *Physiology of the eye* (5th ed.). New York: Pergamon.
De Monasterio, F. M., & Gouras, P. (1975). Functional properties of ganglion cells of the rhesus monkey retina. *Journal of Physiology, 251*, 167–195.
Doty, R. W., & Kimura, D. S. (1963). Oscillatory potentials in the visual system of cats and monkeys. *Journal of Physiology, 168*, 205–218.
Eckhorn, R., Bauer, R., Jordan, W., Brosch, M., Kruse, W., Munk, M., & Reitboeck, H. J. (1988). Coherent oscillations: A mechanism of feature linking in the visual cortex? *Biol. Cybern., 60*, 121–130.
Elliott, M. A., & Müller, H. J. (1998). Synchronous information presented in 40-Hz flicker enhances visual feature binding. *Psychological Science, 9*, 277–283.
Engel, A. K., Fries, P., Konig, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, & consciousness. *Conscious Cogn., 8*, 128–151.
Fahle, M., & Koch, C. (1995). Spatial displacement, but not temporal asynchrony, destroys figural binding. *Vision Res., 35*, 491–494.
Friedman-Hill S., Maldonado, P. E., & Gray C. M. (2000). Dynamics of striate cortical activity in the alert macaque: I. Incidence and stimulus-dependence of gamma-band neuronal oscillations. *Cerebral Cortex, 10*, 1105–1116.
Fries, P., Reynolds, J. H., Rorie, A. E., & Desimone, R. (2001). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science, 291*, 1560–1563.
Georgeson, M. A. (1987). Temporal properties of spatial contrast vision. *Vision Res., 27*, 765–780.
Gouras, P. (1968). Identification of cone mechanisms in monkey ganglion cells. *Journal of Physiology, 199*, 533–547.
Gray, C. M., & Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. USA, 86*, 1698–1702.
Ikeda, M. (1986). Temporal impulse response. *Vision Res., 26*, 1431–1440.
Kaplan, E., & Benardete, E. (2001). The dynamics of primate retinal ganglion cells. *Progress in Brain Research, 134*, 1–19.
Kelly, D. H. (1961). Visual responses to time-dependent stimuli, II. Single channel model of the photopic visual system. *J. Opt. Soc. Am., 51*, 747–754.
Kiper, D. C., Gegenfurtner, K. R., & Movshon, J. A. (1996). Cortical oscillatory responses do not affect visual segmentation. *Vision Res., 36*, 539–544.
Leonards, U., Singer, W. & Fahle, M. (1996). The influence of temporal phase differences on texture segmentation. *Vision Res., 17*, 2689–2697.
Livingstone, M. S. (1996). Oscillatory firing and interneuronal correlations in squirrel monkey striate cortex. *Journal of Neurophysiology, 75*, 2467–2485.
Livingstone, M., & Hubel, D. (1998). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science, 240*, 740–749.
Llinas, R. R. (1992). Oscillations in CNS neurons: A possible role in the generation of 40-Hz oscillations. In E. Başar & T. H. Bullock (Eds.), *Induced rhythms in the brain* (pp. 269–283). Boston: Birkhäuser.
Maldonado, P. E., Friedman-Hill, S., & Gray C. M. (2000). Dynamics of striate cortical activity in the alert macaque: II. Fast time scale synchronization. *Cerebral Cortex, 10*, 1117–1131.
Manahilov, V. (1995). Spatiotemporal visual response to suprathreshold stimuli. *Vision Res., 35*, 227–237.

Mansfield, R. J. W. (1973). Latency functions in human vision. *Vision Res., 13*, 2219–2234.

Maunsell, J. H. R., Ghose, G. M., Assad, J. A., McAdams, C. J., Boudreau, C. E., & Noerager, B. D. (1999). Visual response latencies of magnocellular and parvocellular LGN neurons in macaque monkeys. *Visual Neuroscience, 16*, 1–14.

Maunsell, J. H. R., & Gibson, J. R. (1992). Visual response latencies in striate cortex of the macaque monkey. *Journal of Neurophysiology, 68*, 1332–1344.

McDougall, W. (1904). The sensations excited by a single momentary stimulation of the eye. *Br. J. Psychol., 1*, 78–113.

Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review, 81*, 521–535.

Molotchnikoff, S., & Shumikhina, S. (1996). The lateral posterior-pulvinar complex modulation of stimulus-dependent oscillations in the cat visual cortex. *Vision Res., 36*, 2037–2046.

Munk, M. H. J., Roelfsema, P. R., König, P., Engel, A. K., & Singer, W. (1996). Role of reticular activation in the modulation of intracortical synchronization. *Science, 272*, 271–274.

Öğmen, H. (1993). A neural theory of retino-cortical dynamics. *Neural Networks, 6*, 245–273.

Purushothaman, G., Öğmen, H., & Bedell, H. E. (2000). Gamma-range oscillations in backward-masking functions and their putative neural correlates. *Psychological Review, 107*, 556–577.

Purushothaman, G., Öğmen, H., Chen, S., & Bedell, H. E. (1998). Motion deblurring in a neural network model of retino-cortical dynamics. *Vision Research, 38*, 1827–1842.

Purushothaman, G., Patel, S. S., Bedell, H. E., & Öğmen, H. (1998). Moving ahead through differential visual latency. *Nature, 396*, 424.

Rols, G., Tallon-Baudry, C., Girard, P., Bertrand, O., & Bullier, J. (2001). Cortical mapping of gamma oscillations in areas V1 and V4 of the macaque monkey. *Visual Neuroscience, 18*, 527–540.

Sheer, D. E. (1984). Focused arousal, 40-Hz EEG, and dysfunction. In: T. Elbert, B. Rockstroh, W. Lutzenberger, & N. Birbaumer (Eds.), *Self-regulation of the brain and behavior* (pp. 64–84). Berlin: Springer-Verlag.

Sincich, L. C., & Horton, J. C. (2002). Divided by cytochrome oxidase: A map of the projections from V1 to V2 in macaques. *Science, 295*, 1734–1737.

Steinmetz, P. N., Roy, A., Fitzgerald, P. J., Hsiao, S. S., Johnson, K. O., & Niebur, E. (2000). Attention modulates synchronized neuronal firing in primate somatosensory cortex. *Nature, 404*, 187–190.

Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature, 394*, 179–182.

Van Essen, D. C., Anderson, C. H., & Felleman, D. J. (1992). Information processing in the primate visual system: An integrated systems perspective. *Science, 255*, 419–423.

von der Malsburg, C. (1981). *The correlation theory of brain function* (Int. Rep. No. 81-2). Göttingen: Max Planck Institute for Biophysical Chemistry.

Received October 8, 2002; accepted June 2, 2003.