



Surfing a spike wave down the ventral stream

Rufin VanRullen^{a,b,*}, Simon J. Thorpe^b

^a Division of Biology, California Institute of Technology, MC 139-74, Pasadena, CA 91125, USA

^b Centre de Recherche Cerveau et Cognition, 133 Rte de Narbonne, 31062 Toulouse Cedex, France

Received 28 January 2002; received in revised form 27 June 2002

Abstract

Numerous theories of neural processing, often motivated by experimental observations, have explored the computational properties of neural codes based on the absolute or relative timing of spikes in spike trains. Spiking neuron models and theories however, as well as their experimental counterparts, have generally been limited to the simulation or observation of isolated neurons, isolated spike trains, or reduced neural populations. Such theories would therefore seem inappropriate to capture the properties of a neural code relying on temporal spike patterns distributed across large neuronal populations. Here we report a range of computer simulations and theoretical considerations that were designed to explore the possibilities of one such code and its relevance for visual processing. In a unified framework where the relation between stimulus saliency and spike relative timing plays the central role, we describe how the ventral stream of the visual system could process natural input scenes and extract meaningful information, both rapidly and reliably. The first wave of spikes generated in the retina in response to a visual stimulation carries information explicitly in its spatio-temporal structure: the most salient information is represented by the first spikes over the population. This spike wave, propagating through a hierarchy of visual areas, is regenerated at each processing stage, where its temporal structure can be modified by (i) the selectivity of the cortical neurons, (ii) lateral interactions and (iii) top-down attentional influences from higher order cortical areas. The resulting model could account for the remarkable efficiency and rapidity of processing observed in the primate visual system.

© 2002 Elsevier Science Ltd. All rights reserved.

Keywords: Neural coding; Visual processing; Object recognition; Attention; Spike timing; Rank order coding

1. Introduction

Over 30 years ago (Perkel & Bullock, 1968; Perkel, Gerstein, & Moore, 1967a,b), it was suggested that relevant aspects of the neural code may be carried by single spikes or temporal patterns of spikes across populations of neurons, rather than by a simple mean spike firing rate (Adrian, 1926). Although codes based on firing rate have dominated both theoretical and experimental neuroscience, in recent years, this alternative conception of neural processing has gained more and more ground (Bialek & Rieke, 1992; Bialek, Rieke, de Ruyter van Steveninck, & Warland, 1991; Rieke, Warland, de Ruyter van Steveninck, & Bialek, 1997), and a number of ideas have been proposed as to how information

carried by unitary firing events could be generated, extracted, and used. Among other options, it has been proposed that information could be carried by the neurons' precise firing latencies relative to a given reference event (Celebrini, Thorpe, Trotter, & Imbert, 1993; Hopfield, 1995; Maass, 1997; McClurkin, Gawne, Optican, & Richmond, 1991; Richmond, Optican, Podell, & Spitzer, 1987, 1990; Thorpe, 1990); by the occurrence or repetition of specific firing sequences such as triplets (Lestienne, 1996) or more complex patterns (Abeles, Bergman, Margalit, & Vaadia, 1993; Diesmann, Gewaltig, & Aertsen, 1999; Prut et al., 1998; Vaadia et al., 1995); by synchronous firing events among different neurons or populations (Singer, 1993; Singer & Gray, 1995), or by phase differences during periods of oscillatory firing (König, Engel, Roelfsema, & Singer, 1995a).

One justification for the growing interest in codes that use temporal information comes from the realization that visual processing in primates is extremely fast (Keysers, Xiao, Foldiak, & Perrett, 2001; Oram & Perrett, 1992; Perrett, Rolls, & Caan, 1982; Potter & Levy,

* Corresponding author. Address: Division of Biology, California Institute of Technology, MC 139-74, Pasadena, CA 91125, USA. Tel.: +1-626-395-2879; fax: +1-626-796-8876.

E-mail addresses: rufin@klab.caltech.edu (R. VanRullen), thorpe@cerco.ups-tlse.fr (S.J. Thorpe).

1969; Thorpe, Fize, & Marlot, 1996; VanRullen & Thorpe, 2001a). It has been argued that classical neural codes based on the neurons' mean firing rates would have difficulty operating at such high speeds—hence the need for new alternatives (Bair, 1999; Gautrais & Thorpe, 1998; Maass, 1997; Rieke et al., 1997; Thorpe & Imbert, 1989; VanRullen & Thorpe, 2001b).

Interestingly, most hypotheses that have been made for potential neural codes based on spiking neurons have found at least some experimental support. Temporal precision in response to fluctuating inputs is much higher than might be predicted by the response to continuous inputs (Bryant & Segundo, 1976; Mainen & Sejnowski, 1995), and the temporal structure of neural responses is often reported to show millisecond precision in the retina (Meister & Berry, 1999), the LGN (Reich, Victor, Knight, Ozaki, & Kaplan, 1997; Reinagel & Reid, 2000), or various areas of the visual cortex (Bair & Koch, 1996; Buracas, Zador, DeWeese, & Albright, 1998). Note, however, that most of the experimental data concerns dynamic stimuli and it is still an open question whether temporal aspects of the spike train can code for spatial aspects of the stimulus (Singer, 1999; Stevens & Zador, 1995). So far, the most compelling experimental evidence for a temporal code spatially distributed across large neural populations has been obtained for coding schemes based on firing synchrony (Singer & Gray, 1995). This comes as no surprise however, when most of the experimental research on temporal neural codes in the past 10 years has focused specifically on this hypothesis (Alonso, Usrey, & Reid, 1996; Eckhorn et al., 1988; Engel, König, & Singer, 1991a,b; Gray, Engel, König, & Singer, 1992; König & Engel, 1995; Kreiter & Singer, 1996). The limitations inherent to neural recording methods (in particular the number of neurons that can be simultaneously recorded) and data analysis methods make it difficult to explore other possibilities in a systematic way (deCharms & Zador, 2000). Information theory, which can describe the information content of neural responses, requires assumptions to be made a priori about the neural code (Panzeri, Treves, Schultz, & Rolls, 1999; Victor, 2000). Loosely speaking, this means that if one is not looking specifically for a given neural code in the population response, one is very unlikely to find it, even though this code might constitute the only meaningful part of the response. This limitation is further enhanced by the limitation in the number of concurrently recorded cells: if the neural code is indeed distributed among many neurons, then important aspects of the neural response could be missing because the relevant neurons are simply not being recorded. In short, even after 100 years of experimental and theoretical research in neuroscience, it is still not too late to formulate new hypotheses about the neural code. Only then can these ideas be tested experimentally, and validated or rejected.

Here we describe such a neural code in which information is distributed across large populations of neurons, and represented by relative spike firing times in a single wave of action potentials. The present article is meant to integrate in a unified framework the results of various recent theoretical studies and simulations. In addition, it presents a number of new observations and unpublished material. We show how this code can be applied to model information processing in the ventral pathway of the primate visual system (e.g. object recognition, categorization), and how such a model could account for the efficacy and rapidity of processing in natural systems. Throughout the paper, we place ourselves in a situation where an image is presented to our model visual system at time zero, and only a limited amount of time (e.g. 150 ms) is available to activate a first high-level representation of the scene and its component objects (Thorpe et al., 1996). The neural coding scheme that we use here is a version of the Rank Order Coding proposed by Thorpe (1990) and Thorpe and Gautrais (1997, 1998).

In Section 2 we present one of the most remarkable features of this code, a direct relationship linking visual saliency (or more generally, input contrast) and spike asynchrony, which will constitute the theoretical basis for the following sections. In Section 3 we describe how the properties of retinal ganglion cells can be used to implement such a code, and demonstrate its efficiency in the context of information transmission between the retina and the visual cortex. We then illustrate in Sections 4 and 5 how this information embedded in the first wave of spikes generated in the retina can be decoded by post-synaptic neurons, and how it can propagate in a feed-forward way through a simple hierarchical model of the visual system, to implement fast and reliable object recognition. In Sections 6 and 7, we show that the asynchronous nature of this neural coding scheme allows feed-forward processing to be refined by using both lateral interactions and top-down attentional modulation, without significant slowing down of the system, and without involving any recurrent computational loops. Specifically, we expand on previous work and propose that the visual system can be thought of in terms of a dynamic functional hierarchy. In addition, we present a new illustration of the effects of attention and spike timing from the point of view of the post-synaptic neuron. All of these properties are summarized in Section 8, where we attempt for the first time to sketch a theory of rapid visual processing. Finally, we suggest in Section 9 that such a system could be extended to work under more generic conditions of visual stimulation, i.e. with a continuous flow of visual information, rather than a single visual scene appearing at a given reference time.

At this point it is necessary to clarify the purpose of the present article. Rank Order Coding, as well as the

type of model used here, are still at this stage only a working hypothesis. The work described here is intended to demonstrate the possibilities of such a framework, and how efficient it can prove when applied to real-world problems such as high-level vision or object recognition in natural images. Until supportive experimental evidence is obtained, we are not claiming here that the human visual system necessarily uses such a scheme. Therefore, and for reasons mentioned before, we do not specifically focus here on the biological plausibility of this hypothesis (until Section 10 which will be entirely devoted to this question), although relevant experimental work will be mentioned when available. Instead, we hope that if we are able to demonstrate here that such a coding scheme could be both natural (in terms of its implementation) and efficient (in terms of its performance and rapidity of processing) for real-world visual systems, then we may stimulate interest in experimental neuroscientists, and draw their attention to specific questions that they would not address otherwise.

2. Building the foundations—A general principle: saliency translates into latency

A widely used simplification considers real neurons as “integrate-and-fire” devices. A neuron integrates its inputs over time until it reaches a threshold, and fires a single action potential. The neuron is then reset and, after a certain refractory period, starts integrating information again. This property has in general been used to support the idea that the firing rate of a neuron is a monotonous function of the strength of its input (i.e. the more a neuron is activated, the more it fires). But this property also has a strong implication in the temporal domain. The time at which a neuron reaches its threshold is also a monotonous (decreasing) function of its activation (i.e. the more a neuron is activated, the sooner it fires). This means that the latency of firing of a neuron, just as much as its firing rate, will reflect the strength of its input.

Now consider a population of neurons which, at a given time, is stimulated with a particular input stimulus or intensity pattern (Fig. 1). In theory, if we know the exact firing rate for each neuron of the population, we can describe this input pattern with arbitrary precision. However, as will be argued later, this would require a certain amount of time before each neuron emits a sufficient number of spikes to determine its firing rate reliably. In contrast, simply knowing the time of emission of the first spike of each neuron (i.e. its latency of firing) can provide the same information much faster. Yet another option would be to rely on the specific order in which these first spikes were generated over the whole population. Indeed, the first spike of the population

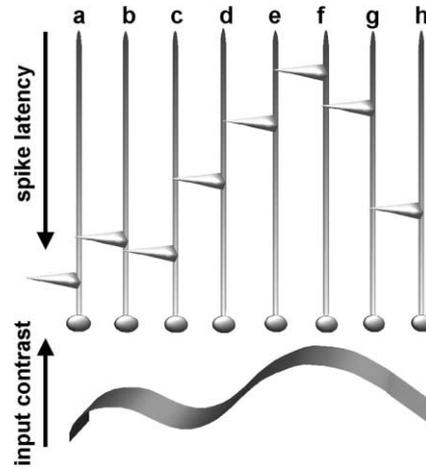


Fig. 1. When a given input pattern is presented to a neural population, the early part of the population response can be described as a spatio-temporal wave of spikes. Within such a wave, a simple consequence of the basic properties of integrate-and-fire neurons is that the most activated cells will have the shortest latencies. Input contrast thus translates into temporal asynchrony. The specific order in which different cells fire can also be used as a code. With 8 input neurons, 8! (more than 40,000) different input patterns can be distinguished.

corresponds to the most activated neuron, the second spike to the next most activated neuron, etc. This idea is the basis of the Rank Order Coding Scheme (Thorpe, 1990; Thorpe & Gautrais, 1997, 1998). The amount of information that can be transmitted with such a code grows with the factorial of the number of neurons in the population. For relatively large neural populations, the information transmission power of this code can meet the requirements of virtually any visual task.

The reasons for using a relative timing code (i.e. order of firing) rather than an absolute one (i.e. based on the exact latency of firing) are twofold. On the one hand, the implementation of an exact latency coding scheme requires specific additional circuitry, such as delay lines (Hopfield, 1995), that does not seem compatible with the pattern of connectivity usually reported in visual cortical neurons. On the other hand, the use of relative timing information is very intuitive and natural. A system using such a scheme would certainly be unable to make precise judgments about exact input intensity values, but it is actually a well-known fact that the human visual system performs much better at directly comparing stimulus features (e.g. luminance, hue, contrast, etc.) than at reporting their exact values. For example, it is almost impossible to differentiate between two stimuli whose luminances are separated by less than 10%, when viewed in isolation; however, this difference becomes striking when the stimuli are presented side by side. Furthermore, coding with ranks naturally provides the system with an invariance to changes in mean stimulus intensity or contrast that can only be achieved with a high computational cost in most other artificial visual systems. It

will be demonstrated in Section 4 how a target neuron can be made sensitive to the particular order of firing of its inputs.

Under the conditions described above, the first spikes in a wave of action potentials correspond to the most activated neurons. Whereas these neurons could be thought of as simply representing the highest intensity values of the stimulus, it is better to think of them as carrying the *most salient information*. In the retina for example, the stimulus property that determines the activation level, and thus the latency of firing of a neuron, is not luminance per se, but rather luminance contrast. It is well known to experimenters in psychophysics and electrophysiology (Reynolds, Chelazzi, & Desimone, 1999) that the primate visual system interprets stimulus contrast as a primary determinant of stimulus saliency, and there is plenty of data showing that latency varies with stimulus contrast (Gawne, Kjaer, & Richmond, 1996). Similarly, the hierarchical organization of the visual system in different processing levels of increasing complexity suggests that each level is responsible for extracting more and more complex features of the stimulus (Barlow, 1972). Yet another, more uniform way to interpret this organization could be that each level represents stimulus saliency in a comparable way (in our case, the most salient information is always carried by the first neurons to fire), but it is the concept of saliency itself that is refined at each stage: roughly equivalent to mere stimulus contrast in the retina or LGN, gradually incorporating information about edges and contours in V1 and V2, object identity or category in the temporal cortex, and possibly the behavioral or task relevance of the stimulus in higher-level areas such as the parahippocampal complex (Miyashita & Hayashi, 2000; Suzuki, 1996) or the prefrontal cortex (Crick & Koch, 1990; Freedman, Riesenhuber, Poggio, & Miller, 2001; Thorpe & Fabre-Thorpe, 2001).

At this stage this tentative distinction between a visual system specialized in feature extraction and one where visual saliency would be the primary object of all computations may be only a matter of terminology. However this statement will play a central role in the next sections, and the importance of such a distinction will hopefully be made clearer by the end of this article.

It is worthwhile to stress here that this type of framework does not apply equally well to different coding schemes. While it is perfectly reasonable to argue that the response of maximal amplitude in a firing-rate based network could also represent the most salient information, one must keep in mind that this information could only be available after all neurons in the population have fired a sufficient number of spikes to determine their firing rates reliably. By contrast, in our framework this information is available as soon as the first neuron fires (by definition, conveying the most salient information).

Other authors have pointed out the possibility of using variations in firing latency as a source of information about the visual stimulus. For example, Nakamura (1998) illustrated how the visual system can select the most salient information for object recognition, by using the first spikes of a given cortical level to trigger lateral inhibition and “prune” irrelevant (i.e. less salient) information. Another possibility, proposed by Wörgötter, Opara, Funke, and Eysel (1996) (see also Opara & Wörgötter, 1996) is to make use of the differences in firing latencies introduced by differences in input contrast among distinct objects to facilitate grouping and segmentation by neuronal synchronisation mechanisms. The specificity of our framework is that these latency differences directly constitute the neuronal code, and are used as such by post-synaptic neurons.

3. Spike wave initiation: estimating information from retinal ganglion cells

A classical argument to reconcile firing rate-based codes with the astonishing speed of visual processing is that multiple spikes from redundant neurons can be used to encode a mean firing rate, averaged over large neuronal populations rather than long time windows (Gerstner, 2000; Shadlen & Newsome, 1994, 1998). By using a few hundred neurons to encode a particular value, it becomes possible to transmit information within an acceptable level of precision in only a few milliseconds (Gautrais & Thorpe, 1998). However, this costly strategy cannot be used in the retina for example, where the very limited number of ganglion cells available (on the order of 1 million in primates) precludes any extensive use of neuronal redundancy. As our retinas constitute the only source of visual input to our brain, it is critical to be able to transmit retinal information reliably and rapidly to the visual cortex.

We have briefly described in Section 2 the basic mechanism by which an input pattern to a population of neurons is converted into a spike wave representing visual information in its spatio-temporal structure. We demonstrate here that this mechanism can be applied to real-world situations where the neural population constitutes a simple model of the retina (Fig. 2) and the input pattern presented to this population is a grayscale natural image. The model and results presented in this section are reproduced from VanRullen and Thorpe (2001b).

When a natural image is presented to our model retina (Fig. 2A), each neuron integrates the luminance information inside its receptive field to determine the contrast in the input image at a particular position, polarity (ON- or OFF-center) and spatial frequency (Fig. 2B). The level of activation of the receptive field in turn determines the time at which the neuron will fire.

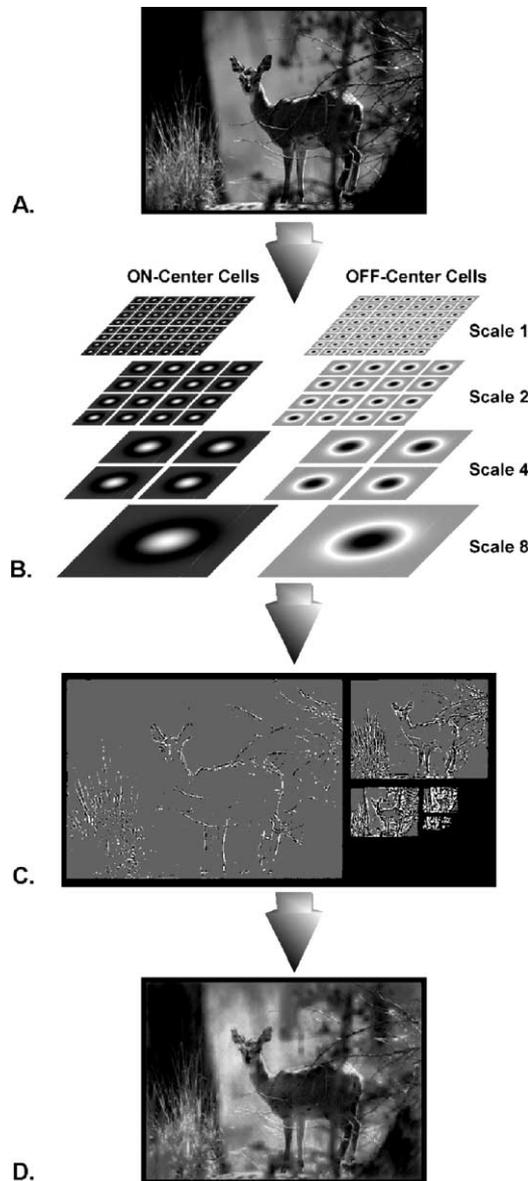


Fig. 2. The approach taken by VanRullen and Thorpe (2001b) to explore the relevance of temporal coding in the retina. An input image (A) is presented to a simple model of the retina (B). The model ganglion cells respond to luminance contrast (with a receptive field described by a difference of Gaussians) at different polarities (ON- and OFF-center cells) and spatial scales (for clarity, only four different scales are represented here, whereas the model used eight different scales). These ganglion cells are organized retinotopically (i.e. encode stimulus contrast at different positions in the original image), and information is subsampled so that the number of neurons at each scale is inversely proportional to the square of the spatial frequency. The spiking output of the ganglion cells can be represented by retinotopic maps as in (C), one map for each spatial scale. The different map sizes reflect the subsampling at low spatial frequencies. Each pixel in these maps corresponds to a pair of ON- and OFF-center ganglion cells encoding input contrast at the same location and spatial scale. White and black pixels represent ON- and OFF-center cells respectively that have generated a single spike. This snapshot was taken after 5% of the ganglion cells had fired. We evaluate the information transmitted by these output spikes by using them to reconstruct an estimate of the input image (D).

Over the whole population of ganglion cells, a spatio-temporal wave of spikes is initiated gradually (Fig. 2C). To evaluate the information transmission power of our coding scheme, one can stop this spike wave initiation process at any time (i.e. when a given percentage of ganglion cells have fired), and calculate how much information has been transmitted about the input image. A simple way of doing so is to use the temporal structure of this spike wave to reconstruct an estimate (Fig. 2D) of the input image (Stanley, Li, & Dan, 1999). The mutual information between the original image and its reconstruction is a direct measure of the information transmitted by the model retina.

Stimulus reconstruction is computed as follows. We place ourselves as an imaginary observer at the output of the retina, collecting the spikes as they are generated. The first spike received, which represents the most salient (i.e. highest contrast) information, is given a maximal weight, and the following spikes progressively lower weights, with a decreasing function that reflects the average statistics of natural images (for further details see VanRullen & Thorpe, 2001b). Apart from this a priori knowledge of the visual environment, the only information used for this reconstruction is the specific order in which ganglion cells have fired. Fig. 3 shows three examples of stimulus reconstruction obtained at different moments of the spike wave initiation process. Less than 1% of discharges in the retina appears sufficient to report most of the contents of the input stimulus. In fact, more than 50% of the maximum information that can potentially be transmitted by this model retina is carried by the first 1% of the spikes.

Further simulations (VanRullen & Thorpe, 2001b) demonstrated that this form of rank-based coding outperforms classical coding schemes relying on mean firing rates over a Poisson spike train, even when 20% random noise is applied to each neuron's firing latency. The impressive encoding power of rank order coding was also stressed by an earlier theoretical analysis that addressed the same issue (Gautrais & Thorpe, 1998): a "count code" relying on how many spikes are generated by each ganglion cell in a given time window performs very poorly when each cell only has time to generate at most one spike; even a rate code relying on the mean inter-spike interval needs at least two spikes to give a first estimate of the firing rate; in contrast, Rank Order Coding only needs one spike per neuron, and can transmit a lot of information even when only a small percentage of neurons has actually fired.

To summarize, the first wave of action potentials generated in the retina in response to a visual stimulation can carry enough information in its spatio-temporal structure to allow further processing. Inside this wave, the most salient information is represented by the very first spikes, so that only a very small proportion of them (e.g. approximately 1%) are needed to transmit the most

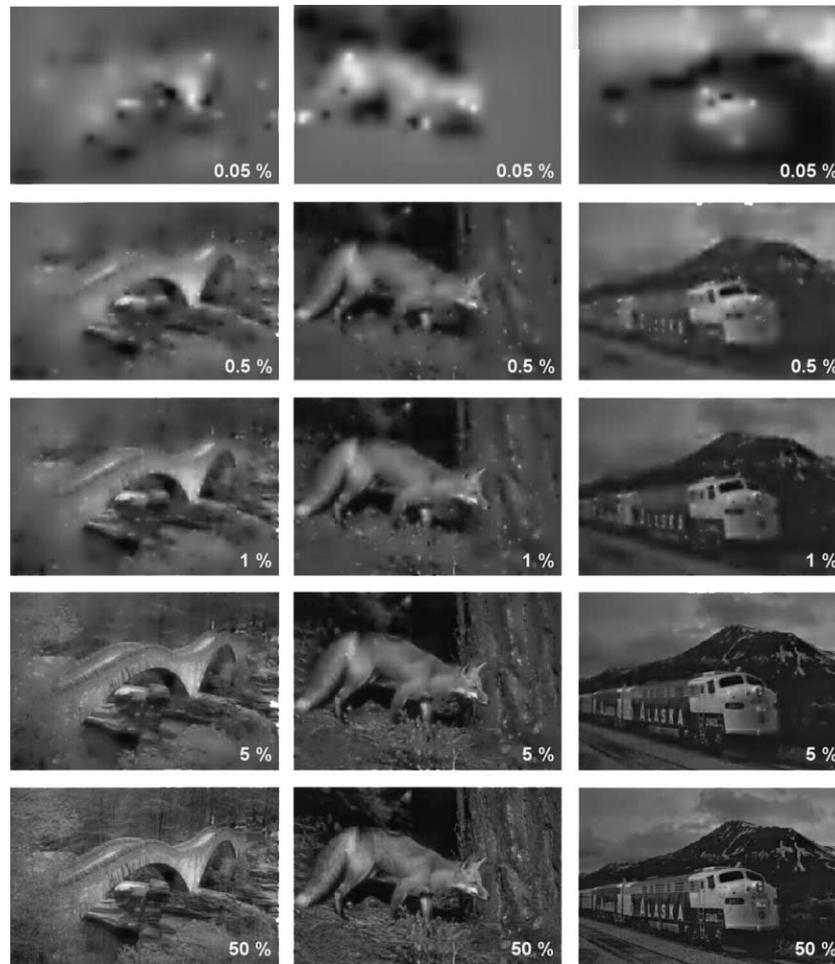


Fig. 3. Examples of stimulus reconstructed from the first wave of spikes initiated in the retina. The percentage of ganglion cells that have generated a single spike is indicated for each reconstruction. It appears that 1% or less is already sufficient to obtain a clear idea of the contents of the input image.

useful information. A visual processing system receiving this spike wave can perform most of its computations at the very beginning of the information transmission process. We will describe in Section 5 how a simple hierarchical model of the ventral stream of the visual cortex can rapidly extract visual information from this first spike wave and make efficient use of this information to implement reliable object recognition tasks. We will focus on the example of face detection in natural images, as it is a well-studied domain, with numerous computational models and algorithms that can be used as a benchmark. Before this however, we must address one further question: how the relative times of spikes in a spike wave can be decoded by a post-synaptic neuron.

4. Spike wave decoding: reading out relative spike times

The temporal structure of the first wave of spikes generated in the retina carries virtually all the information in the input image needed for recognition. To make

use of this information, neurons in the visual cortex that receive this spike wave need to be sensitive to temporal structure in the incoming spikes. Specifically, they should respond selectively to a particular sequence of activation of their afferents, and not to the activation of the same afferents in a relatively different order.

Classical models of neural processing (e.g. Rolls & Treves, 1998; Rumelhart & McClelland, 1986) consider that the activation level of a target neuron reflects the scalar product of its input matrix (the activities, or firing rates of afferent neurons) by its weight matrix (the strength of each afferent's connection to the target neuron). This product is maximal when the input matrix matches the pattern of synaptic weights, and this determines the neuron's selectivity. In our case however, this scheme would not be particularly efficient, because if each input fires one spike, the resulting activity level would be identical whatever the input firing order. Indeed, the input pattern is represented in the temporal domain, and it is in the temporal domain that it must be decoded. One simple way of doing so is to desensitize the

target neuron each time an input spike is received (Thorpe & Gautrais, 1998). Just as we gradually decreased the contribution of later spikes to our stimulus reconstruction from retinal activity (Section 3), a target neuron should give maximal weight to the first inputs it receives, and later spikes should have progressively less and less influence on this neuron's activity. Under these conditions, the resulting activity level for a target neuron will be the scalar product of its synaptic weight matrix with the decreasing desensitization function distributed over the matrix of afferent spikes. With an appropriate threshold, the neuron can then be made selective to a particular order of firing of its inputs. In fact, this approach is functionally equivalent to the classical weight–intensity vectors product approach, with an automatic normalization of inputs (implemented by the desensitization function) in each receptive field.

This desensitization could be achieved by using a rapid form of shunting inhibition (Thorpe, Delorme, & VanRullen, 2001), as illustrated in Fig. 4. A group of input neurons projects to a target cell and to a population of inhibitory interneurons, which are in turn connected to the same target cell. When the input neurons start to fire, they progressively activate the target neuron, but at the same time trigger a form of fast inhibition through the action of the inhibitory interneurons, essentially implementing a desensitization process. Such circuits have been observed for example at the entry point of the visual cortex, in layer IV of the primary visual area V1 of macaque monkeys, where inputs from the LGN make excitatory connections with both target pyramidal cells and fast-spiking inhibitory interneurons,

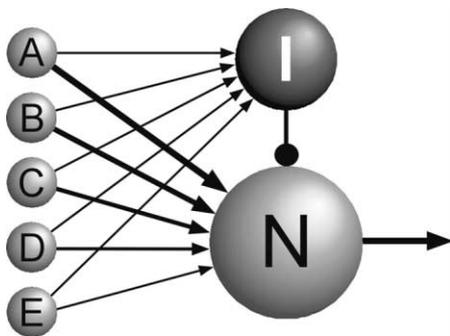


Fig. 4. A circuit involving shunting inhibition can render a neuron selective to the order of firing of its afferents. The neuron N receives excitatory inputs from each of the inputs A–E. The strength of these connections is represented by the width of the corresponding arrow. In addition, the neuron receives shunting inhibition from a pool of inhibitory neurons (I) whose activity increases every time one of the inputs fires. As a result, only the first input to fire will be unaffected by the shunting inhibition, and the amount of inhibition will build up progressively while the neuron receives a wave of spikes. The final activation of this neuron will thus be maximal only when the inputs are activated in the order of their weights. The neuron N is effectively selective to the order of firing of the inputs A–E.

which in turn make contacts on the soma of the same target pyramidal neurons (Callaway, 1998). Intracellular recordings in cat V1 have shown that this type of shunting inhibition can act during the first milliseconds of the target neuron's response (Borg-Graham, Monier, & Fregnac, 1998), making it possible for the desensitization process to function within a single input spike wave.

One interesting prediction of this hypothesis is that interneurons in the inhibitory circuit would not need to be particularly selective, since their principal role is just to “count” how many inputs have already fired. There is recent experimental evidence from the somatosensory barrel cortex showing that fast-spiking inhibitory interneurons receive strong convergent inputs from thalamic afferents having very distinct selectivities, with the result that these interneurons are considerably less tuned than their inputs (Swadlow & Gusev, 2002). Similarly unstructured receptive fields may be seen in at least some fast-spiking interneurons in visual cortex (J. Hirsch, personal communication). These interneurons are unusual in that they can fire very quickly in response to thalamic inputs, and at rates of up to 600 spikes per second, meaning that the effect of the inhibition can be almost instantaneous. Furthermore, there is evidence that these cells are electrically coupled via gap junctions (within a radius of less than 200 μm), meaning that entire populations of inhibitory interneurons will tend to respond together (Amitai et al., 2002; Galarreta & Hestrin, 1999). Together, these properties mean that the very first responses during a processing wave will be the only ones not affected by intra-cortical inhibition and that the responses to later arriving inputs will be progressively attenuated. This is precisely what would be required for rank-order decoding.

Note however that shunting inhibition is not necessarily the only way that rank order decoding could be implemented at the neuronal level (although it is the only one so far that has found experimental support). Cortical neurons display such impressive complexity, of which we know as yet so little (Koch, 1997), that it is conceivable that other candidate desensitization mechanisms (e.g. rapid cross-synaptic adaptation) could be unraveled in the near future, and we wish to leave this question open for further investigation.

5. Spike wave propagation: feed-forward hierarchical object recognition

At the level of neural populations, the neurons that receive input spikes in an order that is closely matched to their pattern of weights will have the strongest activity levels and will thus tend to fire early. Even with one output spike for each neuron, the temporal structure of the population response will therefore reflect the

pattern of activity over the whole population. In other words, this encoding–decoding procedure is cascadable. A neural population can read out the precise temporal structure inside an incoming spike wave, and in turn generate an output spike wave, reflecting the population's selectivity. Such a mechanism therefore seems appropriate for large-scale implementation.

We designed a simple hierarchical model of the visual system, and applied it to the specific task of detecting faces in natural images (VanRullen, Gautrais, Delorme, & Thorpe, 1998). The system is composed of four layers, through which information flows in a purely feed-forward mode. The first level implements a model retina, similar to the one presented in Section 3, although with only one spatial scale. ON- and OFF-center cells respond, respectively, to positive and negative luminance contrast in their receptive fields. The latency of firing of each model ganglion cell reflects the input contrast strength. The output of this first layer is therefore a spatio-temporal spike wave, with all of the underlying properties described earlier. This spike wave is received at the second level by neural populations selective to edges of different orientations inside their receptive field. This layer thus corresponds to a simple model of the primary visual cortex. The selectivity for edges is obtained by using an oriented Gabor function as the neurons' weight matrix, and the order desensitization function ensures that the neurons will respond only when the orientation inside their receptive field matches their selectivity. At the next level, neurons were trained (using a supervised learning procedure, not described here) to respond selectively to the firing order characteristic of the presence of a mouth, left or right eye within the receptive field. Finally, this information is combined at the last level, corresponding to a rough model of the infero-temporal cortex, where neurons would respond only to the simultaneous presence of a mouth, left eye and right eye, with the correct spatial layout, i.e. to the presence of a face.

As illustrated in Fig. 5, the model is able to reliably detect and localize faces in natural images. Tests were performed on large image databases. When compared to the performance of other classical models of face processing (Moghaddam & Pentland, 1995; Rowley, Baluja, & Kanade, 1998; Sung & Poggio, 1994; Turk & Pentland, 1991; Valentin, Abdi, O'Toole, & Cottrell, 1994), this model displays impressive computational power (VanRullen et al., 1998). Detection rates were at least as good as alternative systems described in the literature, with false alarm rates that were much smaller. Furthermore, processing speeds could be one or two orders of magnitudes faster. The principal reason for this remarkable rapidity is that all the computation in our model is event driven: no calculation is required when a neuron remains silent; neurons start computing as soon as they receive inputs, and most of the time will respond

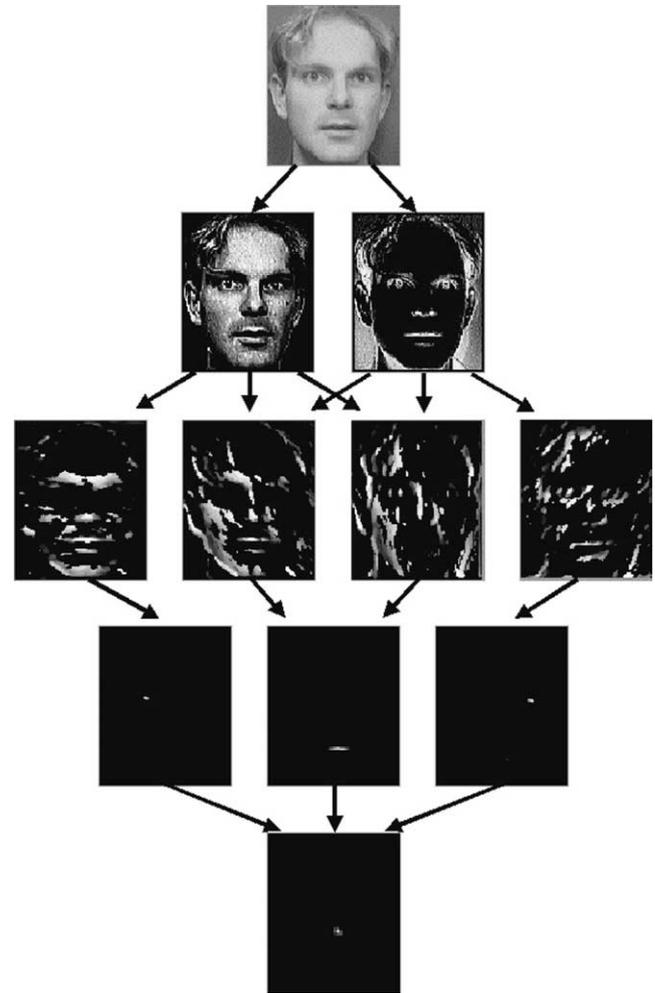


Fig. 5. Architecture and typical response of the face detection model (VanRullen et al., 1998). The system is composed of four layers. Each layer contains a number of retinotopic feature maps. Each pixel in these maps represents a neuron. The position of the pixel in the map corresponds to the position of the neuron's receptive field center in the input image. Gray pixels represent neurons which have fired a single spike, with the gray level reflecting the order of firing of the neuron in the corresponding layer (the first neuron to fire is represented by a white pixel). The input image is decomposed in a model retina by ON- and OFF-center ganglion cells. The spatio-temporal spike wave initiated in the retina is propagated through the system in a feed-forward mode. In the second layer, neurons respond to an edge of a particular orientation (eight orientations separated by 45°; only four orientations shown here). Neurons in the third layer were trained to respond to the firing order pattern specific to the presence of a left eye, mouth or right eye in their receptive field. Outputs from these maps are combined at the last level, where the neurons respond selectively to the presence of a face centered in their receptive field. The position of the firing neuron(s) in this layer reflects the location of the face(s) in the input image. Quantitative results reveal that this model outperforms most classical computational approaches of face processing.

after only a small percentage of their afferents have fired. As can be inferred from the analysis in Section 3, in most circumstances between 1% and 5% of discharges in the retina can be sufficient to detect a face in the input image.

The type of model described here is in essence very simple. A single spike wave propagates in a purely feed-forward way through a hierarchical arrangement of neuronal layers of increasing complexity. This feed-forward template-matching approach is similar to other classical models of object processing (e.g. Fukushima & Miyake, 1982; Riesenhuber & Poggio, 1999), and can be extended to even more demanding processing tasks including face identification (Delorme & Thorpe, 2001). The question that we address however is not whether this particular functional algorithm or architecture is valid, but rather how the specific neural code used here affects the rapidity and efficiency of processing. Provided that the spatio-temporal structure of this spike wave is taken into account, only one spike per neuron is sufficient to perform all of the underlying computations. In contradiction to most classical views of neural processing, this demonstrates that mean firing rates integrated over relatively large time windows, and computational loops or iterative feed-back mechanisms, are not necessary for implementing sophisticated visual processing. Taken together, the model's simplicity and performance make it a very good candidate to explain how the primate visual system can achieve high-level visual processing tasks in a very limited time (100–150 ms).

6. Feed-forward lateral interactions and the distinction between anatomical and functional feed-back

We have seen that a pure feed-forward hierarchical model of the visual system seems sufficient to perform at least some “high-level” visual object recognition tasks. However a range of visual mechanisms and algorithms, such as contour integration, or perceptual filling-in, are known to require lateral interactions among neurons (Field, Hayes, & Hess, 1993; Gilbert, Das, Ito, Kapadia, & Westheimer, 1996; Kapadia, Ito, Gilbert, & Westheimer, 1995). Most models of contour integration make extensive use of recurrent loops or feed-back iterative mechanisms (Gove, Grossberg, & Mingolla, 1995; Grossberg & Mingolla, 1985; Heitger & von der Heydt, 1993; Li, 1998; Shashua & Ullman, 1988), and it might be thought that using such mechanisms would be incompatible with very rapid visual processing. It could be that the kind of rapid visual processing that we are modeling here does not involve much contour integration. However, in a recent study (VanRullen, Delorme, & Thorpe, 2001), we presented a model based on relative spike timing which was able to perform reliable contour integration without the need for recurrent processing.

The model relies on the temporal structure in a spike wave constituting the input to a population of orientation-selective cells, a crude model of the primary visual cortex V1. The core principle is that the first firing cells of the population, which represent the most salient

contours, will start to influence their neighbors through lateral connections before those neighbors fire, i.e. while they are still integrating incoming information. This is compatible with electrophysiological recordings in cats (Volgushev, Vidyasagar, & Pei, 1995) showing that the delay between the onset of a post-synaptic potential in a V1 cell and the emission of its first spike leaves enough time for lateral feed-forward interactions to occur. The response of these neurons will thus not only reflect the specific orientation falling inside their receptive field, but also the degree of alignment of this orientation with the surrounding contours. By limiting the number of spikes per neuron to zero or one, we effectively ensure that there can be no recurrent loop in the system: a neuron which fires and in doing so influences neighboring cells, cannot be influenced in return.

The pattern of lateral connections that we used was similar to the one described by Li (1998). Hence our model had the same functionality as this clearly iterative one, yet was able to display comparable performance (Fig. 6) without the use of computational loops. As noted previously, the present model does not aim to demonstrate that the particular algorithm chosen for the simulations (adapted from Li, 1998) is the most appropriate, but simply that our temporal coding scheme provides important computational advantages. In short, powerful image processing algorithms such as contour integration, classically modeled with recurrent feedback loops, can be obtained in one single pass through a neural network. In the same way, perceptual filling-in or border ownership coding (Zhou, Friedman, & von der Heydt, 2000) could be obtained by the fast recruitment of lateral interactions. The implementation of feed-forward lateral interactions only requires that we take into account the temporal asynchrony inherent in a wave of spikes propagating through the system. Because it is the first neurons to fire that initiate this lateral wave of activity modulation, the most salient stimulus features will determine the entire course of processing.

One might argue, on the other hand, that because neurons in our model are locally mutually interconnected, there is in fact some sort of recurrent processing taking place. This cannot be true however when each cell can only generate at most one spike. Thus determining whether or not a neural algorithm involves recurrent processing depends not only on the anatomical direction of connections in the circuit (i.e. feed-forward, lateral, feed-back), but also critically on the order in which specific neurons or populations are activated. We therefore propose to distinguish between a classical, static anatomically-defined hierarchy and a more functional hierarchy, dynamically defined by the relative times at which neurons in the circuit are activated (see also VanRullen et al., 2001). This dynamic functional hierarchy centers on the following intuitive definition of feed-back: there is feed-back in a neural circuit if the

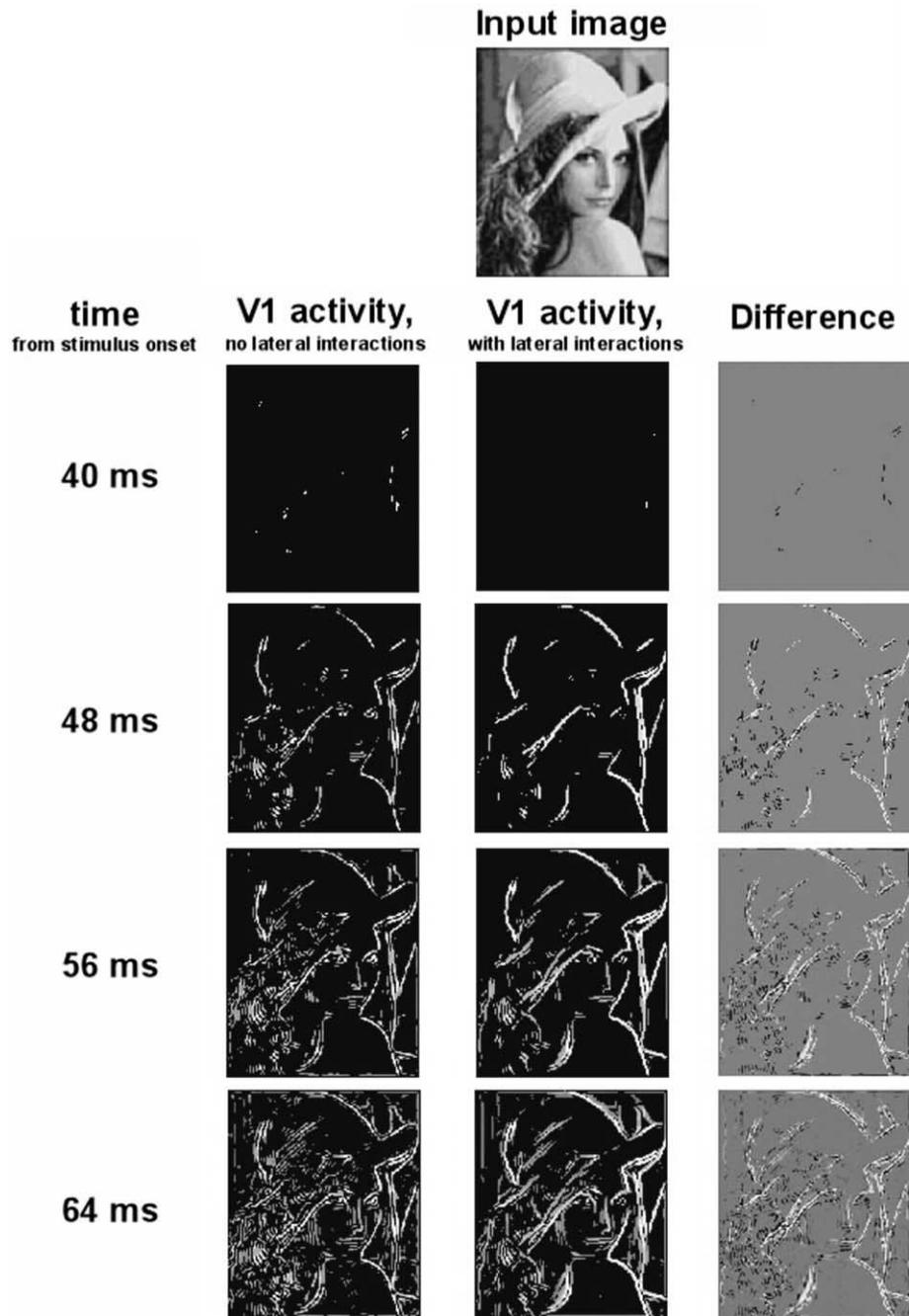


Fig. 6. Estimation of the time course of our feed-forward contour integration mechanism. The first column represents cortical activity in our model of the primary visual cortex, without lateral interactions, the second shows cortical activity building up when feed-forward lateral interactions are used. The number of firing cells is identical in the two conditions. The last column represents the difference between the two situations. Dark spots (resp. bright) correspond to places where activity is decreased (resp. increased) by lateral modulation. Activity appears to leave the places with little or no structure (e.g. feathers on the hat) and is recruited at the locations of well-aligned contours (e.g. hairline, contours of the face and hat). Because all computations rely on the first spike of each cortical cell, the effects of contour integration can be observed in the first few milliseconds of processing. This estimation is based on the assumption of a uniform distribution of discharges in the LGN, and a 40 ms delay between image presentation and the beginning of cortical activity.

firing of a neuron A influences the firing of another neuron B, which in turn (possibly through a set of intermediate connections) modulates the firing of neuron A (Treves, Rolls, & Tovee, 1996). An “anatomically-defined” feed-back (respectively, feed-forward) connec-

tion between two mutually interconnected neurons can in fact act as a “functionally-defined” feed-forward (respectively, feed-back) connection when, for a particular input stimulation, the neuron located higher up in the anatomical hierarchy responds before the one lower

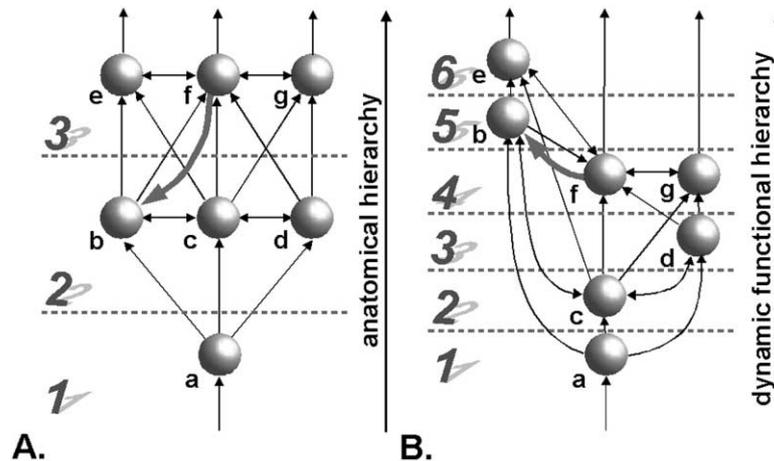


Fig. 7. (A) Example of a neural circuit composed of seven neurons distributed in three different anatomical levels. In this static anatomically-defined hierarchy, thin black arrows represent feed-forward (or lateral for double-headed arrows) connections, and the thick gray arrow depicts a feed-back connection from level 3 to level 2. (B) After stimulus presentation, the propagation of a spike wave in this circuit has resulted in the firing sequence {a; c; d; f-g; b; e}. In a dynamic, functional hierarchy where the first neurons to fire occupy the first “early” levels, the gray arrow actually represents a feed-forward connection. At the time *f* fires, *b* is a potential target for neuron *f*, because it has not yet been activated: a spike flowing from *f* to *b* constitutes a “forward” transmission of information. Conversely, at the time *b* fires, *f* is no longer a potential target to neuron *b*: a spike propagating from *b* to *f* would constitute functional “feed-back” of information.

down (Fig. 7). Experimental studies indeed reveal that the relative latencies of activation of neurons in different visual cortical areas overlap enormously, with in many cases activation occurring simultaneously across separate, anatomically “successive” visual areas (Bullier & Nowak, 1995; Nowak & Bullier, 1998; Schmolesky et al., 1998). Under these conditions, the respective anatomical locations of neuronal subpopulations constitute a poor indicator of the type of communication (feed-forward, feedback) taking place between them. The relevant issue is in the temporal domain: which group of neurons acts as input to which.

The significance of this statement for neural processing is worth underlining. When a spike wave propagates through a hierarchical arrangement of neuronal layers, part of the information can actually be transmitted through anatomically feed-back connections to modulate the activity of cells or populations that are still integrating inputs, without any significant increase in propagation time. This could explain why feed-back has been found to affect the earliest part of neuronal responses (Hupe et al., 2001). In contrast, classical feed-back loops and recurrent iterations would cost a great deal of computational time, with virtually equivalent functional properties. Specifically, the neuronal refractory periods will put a physical limit on the speed with which processing can occur.

7. Top-down modulation and attention

The model of the visual system that we have described so far is not realistic for at least two main rea-

sons. First, it is static, i.e. computation does not depend on the task being performed. Two identical stimulations will always yield the same sequence of firing in the system, hence the same response. In contrast, real visual systems are highly adaptive, so that the same neural populations can perform a wide range of different computations, depending not only on intrinsic properties of the visual input, but also on internal states of the system, even when it operates in an ultra-rapid visual processing mode (VanRullen & Thorpe, 2001c). Second, our model is expensive in terms of the number of neurons required: for example, to achieve position invariance, we have so far used one neuron for each object or feature at each possible location. In contrast, it is well known that neurons in the primate visual system can often respond to their preferred stimulus over a wide range of locations in their receptive field. Neurons in the monkey infero-temporal cortex have receptive fields that can sometimes cover the majority of the visual field. This type of organization however makes the system more vulnerable to error, as two or more different objects might fall inside the same receptive field, and cause illusory conjunctions (Treisman & Schmidt, 1982).

Visual attention appears to represent biology’s response to these limitations. Attention can modulate neural responses in accordance to central, higher-level goals, and in doing so allows the different component features or objects of the visual scene to be parsed into meaningful chunks, even when these features or objects fall inside the same receptive fields (Desimone & Duncan, 1995; Mozer & Sittin, 1998). At the cellular level, a range of experiments (see Reynolds & Desimone, 1999) tend to show that the response of a neuron to two

stimuli falling inside its receptive field is indeed intermediate between the responses to each stimulus alone, a phenomenon that could be responsible for illusory conjunctions. Directing attention to one of them draws the neuron's response to the response elicited by this stimulus alone, thus resolving the ambiguity. Interestingly, it appears that visual attention and saliency (here defined as input contrast) draw on the same resources, so that an increase in the relative saliency of one of the two stimuli will have virtually the same effect as the attentional modulation just described (Reynolds et al., 1999, 2000).

In the context of a neuronal processing framework relying on relative spike timing as a vector of visual saliency, what could be the ideal substrate for visual attention? In our model the early firing neurons have the strongest influence on their target cells. As a consequence, a simple and straightforward way to enhance the significance of a particular feature, property or region of the visual scene is to let the neurons coding for that feature, property or region be among the first to fire. Their decreased onset latency will be interpreted by the system as increased saliency. This modulation can be achieved through a localized lowering of thresholds, or equivalently, an increase in the neurons membrane potentials (VanRullen & Thorpe, 1999).

To illustrate the properties of this mechanism, consider a population of LGN cells transmitting local contrast information to the visual cortex (Fig. 8). Under “normal” conditions of processing (i.e. without attention), just as in the retina (Section 3), the most activated cells will reach their threshold early and thus be among the first to fire. A target cell receiving the spikes from this population will rapidly gather information about the most salient features of the stimulus. Suppose now that the resting state of cells in our population is biased towards a certain region of the visual field, so that neurons in this attended region have an increased tendency to fire earlier. The target cell will now interpret stimulus features at this particular location as the most salient ones. In addition, because of the desensitization function (resulting in an automatic normalization of inputs in each receptive field), it will attribute less importance to the rest of its afferents, even though they might fire at the same exact time as in the “non-attended” condition. This decrease in relative saliency for visual information outside the attended region, well known to psychologists (Cave, 1999; Steinman, Steinman, & Lehmkuhle, 1995) and physiologists (Kastner, De Weerd, Desimone, & Ungerleider, 1998; Smith, Singh, & Greenlee, 2000; Tootell et al., 1998; Vanduffel, Tootell, & Orban, 2000), appears here as a simple consequence



Fig. 8. Simulated reconstructions of the information transmitted by a population of LGN cells after 1% have fired a single spike, under different conditions of attentional modulation. Top, left, with no attentional bias, the first information transmitted corresponds to the most salient (i.e. contrasted) regions. When attention is drawn to a particular location (as indicated by the bright circle in the other three reduced images), it biases information flow, letting information propagate faster at this location. The first information received at the next processing level, interpreted as the most relevant, will thus correspond to the most salient locations in the absence of a top-down attentional bias, and to the attended stimulus features when attention is present. This attentional mechanism appears to increase the relative saliency of the attended stimulus (Reynolds et al., 1999). It biases the population's response towards the response that would be elicited by the attended stimulus presented in isolation (Reynolds & Desimone, 1999). As a result, a target neuron receiving this information at the next level would behave as if its receptive field had “shrunk” around the attended location (Moran & Desimone, 1985). The algorithm used for these reconstructions is similar to the one described in Section 3, and in more detail in VanRullen and Thorpe (2001b). The attentional focus is implemented as a localized lowering of firing thresholds, equivalent to an increase in membrane potential, for neurons whose receptive fields center fall inside the attended region.

of the increase in relative saliency for the attended region.

Although there is supportive evidence that spatial attention can modulate neural activities at the level of the LGN (Crick, 1984; Koch & Ullman, 1985; Vanduffel et al., 2000), it is certainly not the only area where attentional modulation is observed. At virtually every stage in the ventral processing stream of the visual system, neural responses have been found to depend on the attentional state (although there has been some controversy about the effects of attention in the primary visual cortex; e.g. Luck, Chelazzi, Hillyard, & Desimone, 1997; Motter, 1993). This attentional mechanism, illustrated here in the case of the LGN, can act at every processing level of the visual system. The temporal bias towards the attended location increases gradually at each stage, so that the first information to reach high-level neurons, which will determine their response, represents specifically the attended object. This feature is particularly important in systems where “biased competition” (Desimone & Duncan, 1995) is needed to resolve the ambiguities induced by the simultaneous presence of more than one object in a neuron’s receptive field.

To demonstrate this property, we designed a hierarchical model of object recognition in the ventral pathway of the visual system, in which the size of neuron’s receptive fields increases at each level of the hierarchy (VanRullen & Thorpe, 1999). Under these conditions, a fair degree of position invariance could be achieved. However, neurons at higher levels of the hierarchy were likely to respond wrongly when more than one object was present in their receptive field.

The system was designed so that neuronal selectivity became increasingly complex across the different levels, ranging from contrast or orientation in the retina and V1 to more abstract features such as terminations, T or L junctions in higher levels. Each neuron at the last level, corresponding to the infero-temporal cortex, was trained to respond specifically to a particular view of a given object. Nine different objects were used for these simulations (Fig. 9). Whenever an object was presented in isolation, regardless of its retinotopic position, only the neurons selective to this object were activated at the last level. There was no need for attentional modulation in this case. However, when two different objects appeared in an object-selective neuron’s receptive field, one of them being its “preferred” stimulus, the probability of activation of the neuron was typically less than 45% (i.e. intermediate between the responses that would be elicited by each stimulus alone). In addition, each neuron at the last level had a 5% probability of responding to the simultaneous presentation of any two objects to which it was not selective (i.e. an illusory conjunction could occur between the two stimuli).

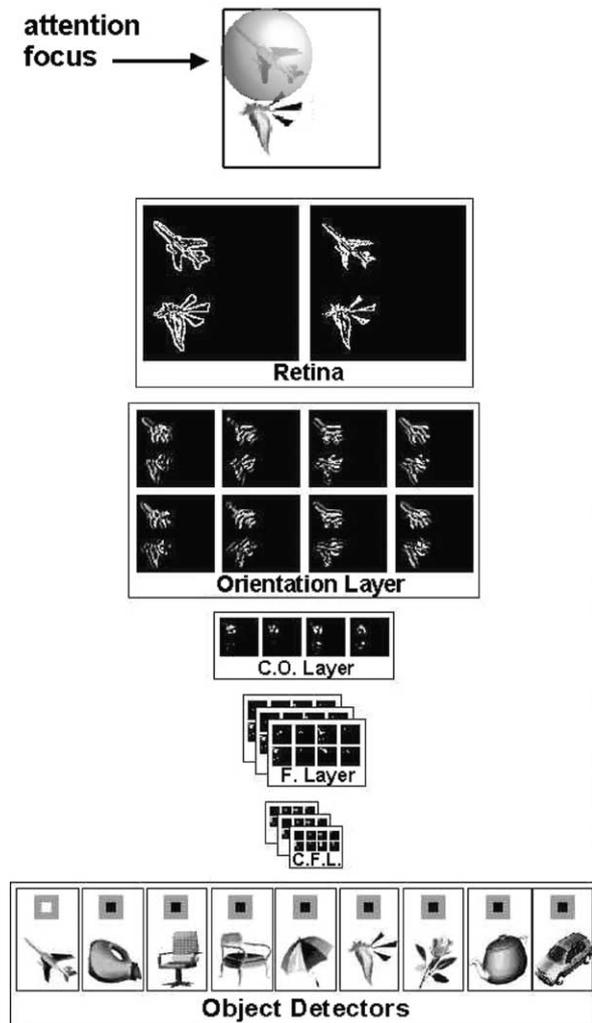


Fig. 9. Result of the propagation of an artificial scene containing two different stimuli through an object detection model equipped with our attentional mechanism. The system is composed of a hierarchy of neuronal layers with increasingly complex selectivities (from top to bottom, Retina: contrast-sensitive cells with two polarities; Orientation layer: selectivity to four different orientations with two different polarities; Complex orientation layer: four orientations, invariance to polarity; Features layer: selectivity to T- or L-junctions and terminations; Complex features layer: similar to the previous layer, with a further degree of position invariance) and increasing receptive fields sizes. At the last level, corresponding to a simplification of the infero-temporal cortex, object-specific cells have receptive fields wide enough to include both input stimuli. They respond specifically to their preferred stimulus when presented in isolation. Without attentional modulation however, these cells have only less than a 50% chance of responding when their preferred stimulus is presented simultaneously with another stimulus. When attention is drawn to one of the two stimuli (as indicated by the bright spot in the input image), the thresholds of neurons with receptive field centers at the attended location are lowered at different levels of the system (orientation, complex orientation, features and complex features layer), so that the corresponding neurons will have a tendency to respond earlier. This information will thus be interpreted as the most salient by the next level of processing. The first information to reach the object-selective cells will represent the attended object, and these cells will respond as if this object had been presented in isolation. Under these conditions, the attended object is correctly recognized in 96% of cases. Adapted from VanRullen and Thorpe (1999).

When an attentional bias towards one of the stimuli was applied to this system, in the form of a 5–20% threshold decrease for neurons with receptive fields falling in the attended region, at different levels of the system (orientation, complex orientation, features and complex features layers), the selectivity of the neurons was restored. An object-selective neuron would respond to its preferred stimulus with a probability of 96% when it was the attended object, and with a probability of around 2% when attention was drawn to the other object of the pair. In addition, the probability of responding to a pair of non-preferred objects (illusory conjunction) was less than 0.5%.

This attentional effect is achieved in a remarkably straightforward way: when the pair of stimuli is presented, neurons at the early levels of the system start integrating information at the same time; but because of the attentional bias, the neurons coding for the attended stimulus will have a tendency to respond before the ones coding for the unattended object; this bias can increase to a significant extent at each level, so that the first information to reach higher-level neurons will represent the attended stimulus only. For an object-selective neuron receiving this input, the effect would be virtually identical to that seen by presenting the attended stimulus in isolation. When the information corresponding to the unattended object finally reaches the highest levels, the system's response has already been generated. In addition, because of the progressive desensitization responsible for the order decoding process, this information will have a fairly limited effect on the neuron's activity: the neuron behaves as if its receptive field had shrunk around the attended location (Moran & Desimone, 1985). Note however that in this system, there is no change whatsoever in the way unattended inputs are processed. The latency of firing of neurons coding for unattended locations can be identical to the one that would be observed without attentional modulation. It is only when this unattended information enters a receptive field that has been touched by the attention focus that a suppressive effect will be observed. In other words, the suppression of unattended inputs in this model is not a direct, explicit inhibitory mechanism (as would be the case in a segmentation process), but arises as a consequence of the intrinsic properties of asynchronous propagation and temporal order selectivity inside a single receptive field. This implies that under conditions where the attentional focus is directed to an object outside a neuron's receptive field, the selectivity of this neuron to an unattended stimulus will be unchanged (Moran & Desimone, 1985): spatially separated receptive fields act as independent processing channels. The degree of independence might in fact be a function of the degree of overlap between the receptive fields under consideration, as well as the receptive fields of afferent neurons at earlier levels. This could explain why some

authors have found that neuronal selectivity can sometimes be modified by attention directed outside their receptive field, and that the extent of this modulation is a function of the distance between the attention focus and the receptive field center (Connor, Gallant, Preddie, & Van Essen, 1996, 1997).

It should be noted that the model of attention described here has one obvious limitation: we assume the existence of some form of attentional control signal to direct the focus of attention, but do not explicitly model the process by which this focus is selected. This would require additional mechanisms (e.g. saliency map, decision processes, etc.) that go beyond the scope of our model.

There is in fact some recent experimental data that strongly supports the idea that attention can shorten the onset latency of visual neurons. EEG studies show that the latencies of event-related potentials can be reliably shortened by attention (Di Russo & Spinelli, 1999a,b), in particular for early ERP components such as the N60 or P100. Such findings fit with data from experimental psychology (Pashler, 1998) showing that precueing target location can decrease reaction times in rapid discrimination or go/no-go tasks (Kingstone, 1992; Neumann, Esselmann, & Klotz, 1993; Posner, Snyder, & Davidson, 1980; Proverbio & Mangun, 1994). Shiu and Pashler (1993) report for a "letter/digit" discrimination task a 30 ms decrease in response time, which would be compatible with a 20% attention-induced temporal bias under conditions where visual processing is done in around 150 ms (Thorpe et al., 1996). Similarly, the line-motion illusion (Hikosaka, Miyauchi, & Shimojo, 1991, 1993a,b) can easily be explained by an attention-related decrease of firing latencies. This illusion occurs when a line is presented 50 ms after a brief flash, and aligned with this flash. Under these conditions, subjects perceive the line as appearing gradually, away from the position of the flash. The interpretation advanced by the authors is that the transient flash captures attention, which in turn diminishes firing latencies of neurons with receptive fields near the location of the flash, generating a temporal firing order similar to what would be obtained with an actual movement of the line. This observation recently led Kirschfeld and Kammer (2000) to propose a mechanism for attentional selection similar to the one we had described in which attention produces subthreshold modulations of neuronal resting states that decrease firing latencies in the attended region. Note however that many studies have failed to demonstrate an effect of attention on neuronal latencies, whether using single-cell recordings (Reynolds, Pasternak, & Desimone, 2000) or event-related potentials (e.g. Mangun, 1995). The reason for this discrepancy in results is unknown at the time, and could constitute a potential weakness for our hypothesis.

To conclude this section, it is worth insisting on the main characteristic of the model that we described: stimulus saliency is here the common basis for feature based, bottom-up visual processing and top-down attentional selection. This unified framework, compatible with biological timing constraints, can account for a wide range of experimental observations, while displaying most of the computational advantages of various other models of visual processing and attention.

In this framework, propagation of information through the visual system is guided by the most salient features and locations, implicitly and automatically suppressing competing unattended information, resulting in a dynamic restructuring of receptive fields (Ghazanfar & Nicolelis, 2001; Wörgötter et al., 1998). As pointed out by Salinas and Abbott (1997), this type of saliency-based attentional mechanism is perfectly suited to implement “shifter circuits” or “dynamic routing” processes: a realignment of high-level neurons receptive fields with the attentional focus, proposed to be responsible for position or size invariance (Anderson & Van Essen, 1987; Olshausen, Anderson, & Van Essen, 1993). In our case however, this realignment can be obtained on the basis of a single wave of action potentials propagating in a feed-forward mode. No feedback (Hamker, 2000), no routing-dedicated neurons, and no saliency map (Itti & Koch, 2001) are needed to control this rapid gating process. Similarly, the active selection of the most salient information (first inputs having the strongest impact) coupled with the implicit suppression of unattended or non-salient inputs in our model implements a center-surround organization of the attentional window, without the explicit inhibitory connections usually required for this purpose (Itti & Koch, 2000; Koch & Ullman, 1985). Alternatively, the result of such a scheme can be considered as roughly equivalent to a softer version of a non-linear winner-take-all mechanism (Lee, Itti, Koch, & Braun, 1999) or MAX-like operation (Riesenhuber & Poggio, 1999), which has been recently shown to provide a way of implementing routing of relevant information in a feed-forward model of object recognition. But here again, these computational properties can be obtained as a simple consequence of the temporal code proposed, and do not need to be implemented explicitly.

8. A theory of rapid visual processing

At this point the results and propositions described in the previous sections can be formalized in a tentative theory of rapid visual processing.

- When a new image is presented to the retina, or when a saccade occurs, leading to a new fixation, the early

part of the response of ganglion cells can be considered as a wave of action potentials, carrying visual information in its spatio-temporal structure.

- At the next level, neurons are selective to the spatio-temporal structure of the pattern of spikes reaching their receptive fields. Here again, the most activated neurons fire earlier. The output of this next level is therefore another spike wave, in which the first spikes represent the most salient information.
- As soon as the first neuron in a given population emits its first output spike, efferent neurons at the next level will start computing. In many cases computation can thus be performed when only a few (1–5%) input neurons have fired.
- At each stage, lateral (or backwards) interactions can take place to clean up or enhance the signal selectively. These interactions modify the spatio-temporal structure of the wave, but do not question its feed-forward propagation in a dynamic functional (rather than anatomical) hierarchy.
- At the system level, this mechanism is reproduced in a cascade: a wave of retinal spikes propagates through the system, and is regenerated at each level, with the most salient information always represented by the first spikes of the wave. But as this wave propagates through the visual system, the definition of saliency itself is refined at each level.
- Top-down mechanisms can act as a temporal bias on such a spike wave. Attention gives a temporal precedence to the neurons representing relevant information, and in doing so increases their relative saliency for the next processing stages.
- In the highest levels of the visual hierarchy (e.g. infero-temporal cortex), a neuron or group of neurons selective to a particular object will be rapidly activated, after a single pass through the system, if this object is presented alone in the input image. If the object is presented in a cluttered environment, competition will take place between the different stimuli. The neuron’s selectivity will be preserved if this object or its component features are particularly salient, or if attention is drawn selectively to this object or its component features.

In short, the scheme that we propose here relies on the propagation of a single spike wave carrying information in its spatio-temporal structure, which can be modified at each stage by (i) the neurons feature selectivities, (ii) lateral feed-forward interactions, and (iii) top-down attentional influences. This powerful computational scheme, based on the tight relationship between relative spike timing and visual saliency, is probably one of the best candidate theories to explain how the human visual system can access a high-level representation of the objects constituting the visual scene in only 150 ms.

9. Spike tide: a continuous flow of spike waves?

The previous sections present a theory of neural coding and information transmission in the visual system that can account for the speed of visual processing observed in various experiments.

Biological vision however, is in general quite different from laboratory situations where a stimulus can appear for only a few milliseconds, followed by a dark screen. The retinal image is in fact constantly changing, replaced after each saccade (roughly every 300 ms in humans) or micro-saccade (roughly an order of magnitude more frequent) by a totally new image, or at best one that is vaguely correlated with the previous one. How could our framework cope with this continuous flow of information? How can a neuron tell the trailing edge of the input spike wave corresponding to one stimulation from the leading edge of the spike wave representing the next one?

One option might be to implement some form of reset mechanism to separate processing of successive inputs. It might be that saccades and micro-saccades themselves might constitute the trigger for such a reset process (Martinez-Conde, Macknik, & Hubel, 2000). For example, the transmission of visual information by LGN neurons in cats has been found to be facilitated immediately after passive eye movements, and inhibited during fixation (Lal & Friedlander, 1989, 1990a,b). The resetting of temporal integration mechanisms would thus be obtained as a result of visual transients, either externally generated (as part of the stimulation) or internally initiated by eye movements. Indeed it has long been known that a total absence of change in the retinal stimulation results in the complete disappearance of the visual scene (Coppola & Purves, 1996; Ditchburn & Ginsborg, 1952; Riggs & Ratcliff, 1952).

Yet another possibility might be that rhythmic oscillations, for example in the LGN, either intrinsically generated or mediated by projections from the visual cortex, could serve as a basis for this reset mechanism. Cortico-thalamic influences have been found to modify LGN oscillatory rhythms in cats both in the alpha (Contreras, Destexhe, Sejnowski, & Steriade, 1996) and gamma range (Castelo-Branco, Neuenschwander, & Singer, 1998). In the same way, cortical subthreshold oscillations, for example in the gamma range (Engel et al., 1991b; Frien, Eckhorn, Bauer, Woelbern, & Kehr, 1994; Gray, König, Engel, & Singer, 1989; König, Engel, & Singer, 1995b) could be used to mediate temporal coding mechanisms (Fries, Neuenschwander, Engel, Goebel, & Singer, 2001; Lampl & Yarom, 1993; Nowak, Sanchez-Vives, & McCormick, 1997; Salinas & Sejnowski, 2001; Volgushev, Chistiakova, & Singer, 1998). In fact, this active reshaping of temporal response profiles could take place simultaneously in or among dif-

ferent subcortical and cortical areas or neuronal subpopulations.

One interesting issue in this context is the notion of a temporal perceptual frame: the maximum interval between two successive stimuli for which they are still perceived as one single event (Lichtenstein, 1961). This time window appears just large enough (given the range of latencies in the retina) to allow the propagation of a single spike wave: approximately 40 ms or less. Somehow the visual system seems to know that two spike patterns temporally separated by more than one average spike wave duration can potentially refer to two distinct events in the outside world, whereas two spike patterns separated by less than one spike wave duration are more likely to carry information about different properties of the same perceptual event. Although this could be interpreted as a simple limitation of perceptual temporal resolution due to neuronal integration time constants, this explanation is unlikely, in particular because in many situations neurons can show much higher precision, on the order of a few milliseconds or less (Reinagel & Reid, 2000).

Interestingly, the size of this temporal perceptual window has been found to be closely related to cortical rhythms (Callaway & Layne, 1964; Gho & Varela, 1988; Varela, Toro, John, & Schwartz, 1981). For a given delay between two flashes of light, the probability of perceived simultaneity depends on the phase of the ongoing alpha rhythm at which the stimuli are presented. In other words, these rhythmic oscillations are able to separate perceptual events, and could therefore differentiate between the spike wave corresponding to one visual stimulation and the spike wave representing the next one.

The hypothesis that we put forward here is reminiscent of an idea that attracted considerable interest more than 30 years ago (Harter, 1967). Based on the observation that visual perception is “discrete” in nature, a number of authors have suggested that cortical oscillatory rhythms might act as cortical “excitability cycles” (Callaway & Alexander, 1960) or perceptual “scanning moments” (Pitts & McCulloch, 1947; Shallice, 1964; Stroud, 1949), implementing a sort of “neuronic shutter” (Lindsley, 1952), rather like the shutter of a camera, or the aperture of a movie projector. These theories however were unclear regarding the type of neural coding that might take place inside such a cycle or moment (Harter, 1967). A code relying on the temporal structure of a single spike wave fits naturally in this context. Temporal differences in such a wave, because they do not participate in the perceptual sequencing of visual events, can indeed be used for representing spatial information and other visual properties such as stimulus saliency. Accordingly, temporal differences on the order of 10 ms or less have been found to significantly affect such perceptual spatial judgments as texture (Leonards,

Singer, & Fahle, 1996) or figure ground (Fahle, 1993; Kandil & Fahle, 2001) segmentation and grouping (Usher & Donnelly, 1998), although these effects might depend on the order of onset of the figure and ground stimuli (Beaudot, 2002). On the other hand, numerous studies have demonstrated that temporal differences of 40 ms or less are generally not experienced in the temporal domain (i.e. are not perceived as successive; Allport, 1968; Anstis, 1979; Gho & Varela, 1988; Lichtenstein, 1961).

Note that the idea of a perceptual frame should not be understood as an absolute process. The duration of the perceptual integration window can depend critically on various factors intrinsic to the type of stimulus, the task in which the system is engaged or the cortical area involved in this task. Motion perception, for example, can involve longer time constants than form or color perception (e.g. Moutoussis & Zeki, 1997a,b). Different types of oscillatory rhythms, in different frequency bands, might serve to underly these different aspects of visual processing. We focus here on those mechanisms that could constitute a general basis for these more specific processes.

Under continuous conditions of stimulation, propagating spike waves could be actively reshaped at each level by the action of oscillatory rhythms. This internal regulation of the spatio-temporal structure of population responses might for example be used to facilitate response generation at a given time, or prevent it for another period, keeping spike waves reasonably separated to avoid “temporal jamming”. Fig. 10 illustrates this “spike tide” hypothesis with a simple example of a population constituted by 20 neurons. The input stimulation, either static or in movement, is continuous. However the spike wave initiation process, coupled with this active temporal framing mechanism, convert this continuous input into a discrete succession of spike waves: in other words, a spike tide (Fig. 10A). Each of these spike waves contains in their spatio-temporal structure all the relevant information gathered about the stimulation applied during a certain time window. However, the temporal structure alone does not convey much information by itself, and an experimenter recording spike trains from this population without any knowledge of the spatial arrangement of these neurons would find no systematic temporal structure in individual responses (Fig. 10B). In contrast, the temporal separation between successive spike waves would result in observed periodic oscillations of the population firing rate (Multi-unit activity, Fig. 10C). Finally, two neurons participating in such a succession of spike waves would also display a very strong temporal correlation (Fig. 10D), because spikes of both neurons are actively distributed in separated spike waves of a temporally limited extent. In this context, subthreshold oscillations serve as a mechanism for temporal regulation, while oscillations

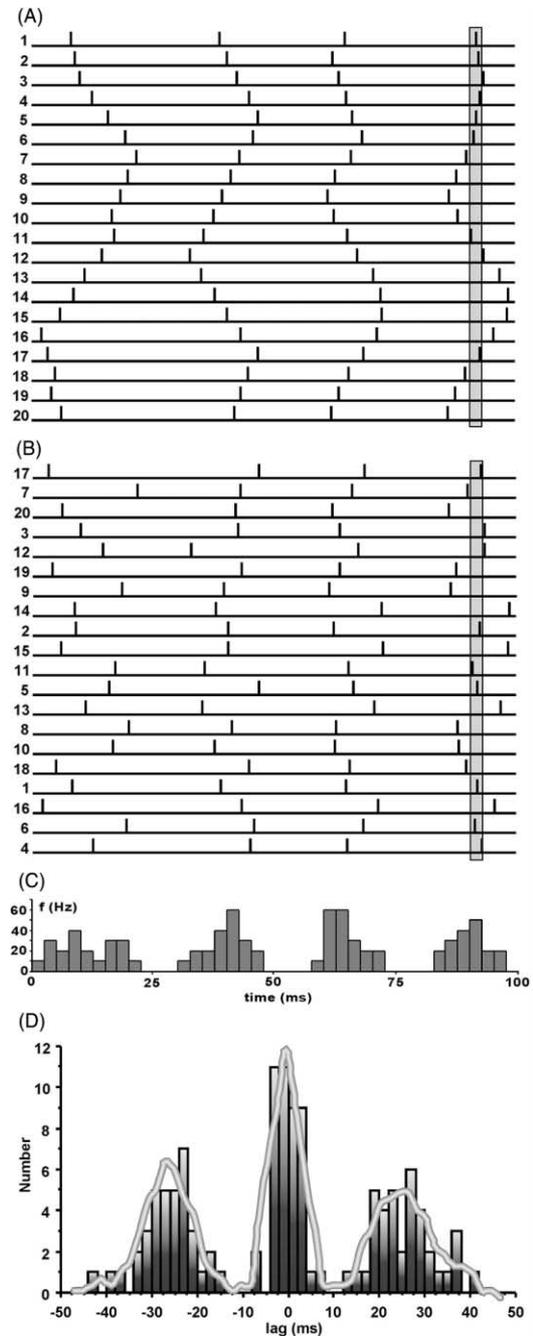


Fig. 10. (A) The hypothesis of neural information coding by spike temporal asynchrony over a neural population suggests that, under normal conditions of stimulation, a spatially organized population (here 20 neurons) responds by a succession of spike waves: a “spike tide”. (B) When this same population is organized randomly, no systematic spatio-temporal structure is apparent. This corresponds to what an experimenter would record using a multi-electrode array. (C) Multi-unit activity (MUA), calculated here with a time bin of 2.5 ms (gray window in A and B). With the parameters used here (on average 25 ms between two successive spike waves), this activity oscillates at a 40 Hz frequency. (D) Cross-correlogram between two model neurons participating in the same spike waves. Each neuron’s response is here obtained by concatenating the spike trains from odd vs. even neurons in A (an approximation of a situation where the same stimulation would be presented repeatedly, slowly drifting in space). A strong correlation with zero phase lag is observed in this example.

of the average firing rate, or firing coincidences, arise as a consequence of the temporal succession of spike waves. This proposition is at odds with many current views of neuronal temporal coding, assuming that the neural code is embedded in periodic firing oscillations or synchronous firings, rather than the opposite (Engel & Singer, 2001; Singer & Gray, 1995). In fact, the idea that the required reset mechanism for a temporal coding scheme might be the origin of gamma-band cortical oscillations has already been suggested by Parodi, Combe, and Ducom (1996). A recent experimental study of neuronal responses in the visual cortex of cats and monkeys provides strong support for this view (Fries et al., 2001). These authors found that firing latencies of visual neurons coding for similar stimulus properties (position, orientation) tend to vary together on a trial by trial basis, independent of the fluctuations of the mean firing rates, and that this variation in latency is primarily determined by the phase of the ongoing gamma oscillations of the local field potentials. If one assumes that such shifts in latencies with cortical oscillations do not perturb but generally preserve the relative timing (or the order) of spikes within a population of interest, this experimental observation is exactly what one would predict on the basis of the present hypothesis.

It should be underlined that this hypothesis of cortical and subcortical oscillatory rhythms acting as a temporal framing mechanism, actively reshaping population responses, even if very attractive in the context of rapid visual processing based on the propagation of spatio-temporal spike waves, is only one amongst many other possibilities. For example, it could be that this spike wave propagation is only initiated with global visual transients, and that the system would switch into a more classic firing-rate-based code for later processing. Only further experimental investigations could provide the necessary evidence supporting one or the other suggestions. But in any case, the generic framework presented here, stressing the relationship between visual saliency and relative spike timing, and the advantages of taking into account the precise spatio-temporal structure of the first spike wave induced by a visual stimulation, should be considered as one of the first serious attempts to propose computational strategies compatible with the severe temporal constraints imposed by the speed of visual processing in primates. The model makes a number of clear experimental predictions which will hopefully be tested in the near future, opening the way to an increased understanding of visual function.

10. Biological plausibility

Although to date there has been no experimental observation to directly confirm our theory, there is also no direct experimental evidence of the contrary. A quick

overview of the literature shows that many experimental results tend to favor our hypothesis. First, it is on the grounds of experimental observations that this theory has emerged: the speed of processing in biological visual systems indeed calls for new, non rate-based neural codes (Keyzers et al., 2001; Perrett et al., 1982; Thorpe et al., 1996). The precise temporal structure of neuronal spike trains is consistently reported to carry more information than the mean firing rate alone (Bialek et al., 1991; Borst & Theunissen, 1999; McClurkin et al., 1991; Richmond et al., 1987, 1990). The temporal precision of spike trains in response to transient stimuli has been shown to be on the order of 1 ms or smaller in various preparations *in vitro* (Bryant & Segundo, 1976; Mainen & Sejnowski, 1995) and *in vivo* in the retina (Meister & Berry, 1999), the LGN (Reich et al., 1997; Reinagel & Reid, 2000) and various areas of the visual cortex (Bair & Koch, 1996; Buracas et al., 1998). Moreover, this remarkable precision is not only found in response to purely dynamic stimuli. In particular, the latency of the first spike emitted in response to a given stimulation can depend largely on stimulus contrast (e.g. Gawne et al., 1996), but also on other stimulus parameters such as orientation (Celebrini et al., 1993). Even when firing latency does not appear to be directly correlated to the time of external stimulation, recent evidence (Fries et al., 2001) suggests that it could in fact vary in reference to other intrinsic rhythms of the visual system such as gamma oscillations. If these results were consistently reproduced, and more observations of temporal spike patterns and first-spike latency (or spike order) coding for non-temporal aspects of the stimulus were collected in the visual system as they seem to be found in other modalities (Carr, 1993; Panzeri, Petersen, Schultz, Lebedev, & Diamond, 2001; Petersen, Panzeri, & Diamond, 2001), a consensus might finally emerge as to whether or not the absolute or relative latency of neuronal first spikes can constitute a viable carrier of visual information. All that it would really require is that few visionary or simply open-minded electrophysiologists start to systematically consider neuronal latencies or spike times as another relevant variable, either in reference to external events (stimulation), to firing events in other neurons (when multi-electrode recordings are available) or to the oscillatory phase of local field potentials. We hope that this article, our theory and its specific predictions will help stimulate such an experimental effort.

Acknowledgements

The authors wish to thank F. Hamker, G. Kreiman, C. Rashe for helpful comments, and C. Koch for his support.

References

- Abeles, M., Bergman, H., Margalit, E., & Vaadia, E. (1993). Spatiotemporal firing patterns in the frontal cortex of behaving monkeys. *Journal of Neurophysiology*, 70(4), 1629–1638.
- Adrian, E. D. (1926). The impulses produced by sensory nerve endings: Part I. *Journal of Physiology (London)*, 61, 49–72.
- Allport, D. A. (1968). Phenomenal simultaneity and the perceptual moment hypothesis. *British Journal of Psychology*, 59(4), 395–406.
- Alonso, J. M., Usrey, W. M., & Reid, R. C. (1996). Precisely correlated firing in cells of the lateral geniculate nucleus. *Nature*, 383(6603), 815–819.
- Amitai, Y., Gibson, J. R., Beierlein, M., Patrick, S. L., Ho, A. M., Connors, B. W., & Golomb, D. (2002). The spatial dimensions of electrically coupled networks of interneurons in the neocortex. *Journal of Neuroscience*, 22(10), 4142–4152.
- Anderson, C. H., & Van Essen, D. C. (1987). Shifter circuits: a computational strategy for dynamic aspects of visual processing. *Proceedings of the National Academy of Science, USA*, 84(17), 6297–6301.
- Anstis, S. (1979). Apparent movement. *Handbook of sensory physiology (Vol. VIII)*. New York: Springer Verlag.
- Bair, W., & Koch, C. (1996). Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Computation*, 8(6), 1185–1202.
- Bair, W. (1999). Spike timing in the mammalian visual system. *Current Opinion in Neurobiology*, 9, 447–453.
- Barlow, H. B. (1972). Single units and sensation: a neuron doctrine for perceptual psychology? *Perception*, 1(4), 371–394.
- Beaudot, W. H. (2002). Role of onset asynchrony in contour integration. *Vision Research*, 42(1), 1–9.
- Bialek, W., & Rieke, F. (1992). Reliability and information transmission in spiking neurons. *Trends in Neurosciences*, 15(11), 428–434.
- Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R., & Warland, D. (1991). Reading a neural code. *Science*, 252(5014), 1854–1857.
- Borg-Graham, L. J., Monier, C., & Fregnac, Y. (1998). Visual input evokes transient and strong shunting inhibition in visual cortical neurons. *Nature*, 393(6683), 369–373.
- Borst, A., & Theunissen, F. E. (1999). Information theory and neural coding. *Natural Neuroscience*, 2(11), 947–957.
- Bryant, H. L., & Segundo, J. P. (1976). Spike initiation by transmembrane current: a white-noise analysis. *Journal of Physiology*, 260(2), 279–314.
- Bullier, J., & Nowak, L. G. (1995). Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Current Opinion in Neurobiology*, 5(4), 497–503.
- Buracas, G. T., Zador, A. M., DeWeese, M. R., & Albright, T. D. (1998). Efficient discrimination of temporal patterns by motion-sensitive neurons in primate visual cortex. *Neuron*, 20(5), 959–969.
- Callaway, E. I., & Alexander, J. D. J. (1960). The temporal coding of sensory data: an investigation of two theories. *Journal of General Psychology*, 62, 293–309.
- Callaway, E. I., & Layne, R. S. (1964). Interaction between the visual evoked response and two spontaneous biological rhythms: the EEG alpha cycle and the cardiac arousal cycle. *Annals of the New York Academy of Sciences*, 112, 421–431.
- Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annual Review of Neuroscience*, 21, 47–74.
- Carr, C. E. (1993). Processing of temporal information in the brain. *Annual Review of Neuroscience*, 16, 223–243.
- Castelo-Branco, M., Neuenschwander, S., & Singer, W. (1998). Synchronization of visual responses between the cortex, lateral geniculate nucleus, and retina in the anesthetized cat. *Journal of Neuroscience*, 18(16), 6395–6410.
- Cave, K. R. (1999). The FeatureGate model of visual selection. *Psychological Research*, 62(2–3), 182–194.
- Celebrini, S., Thorpe, S., Trotter, Y., & Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake monkey. *Visual Neuroscience*, 10, 811–825.
- Connor, C. E., Gallant, J. L., Preddie, D. C., & Van Essen, D. C. (1996). Responses in area V4 depend on the spatial relationship between stimulus and attention. *Journal of Neurophysiology*, 75(3), 1306–1308.
- Connor, C. E., Preddie, D. C., Gallant, J. L., & Van Essen, D. C. (1997). Spatial attention effects in macaque area V4. *Journal of Neuroscience*, 17(9), 3201–3214.
- Contreras, D., Destexhe, A., Sejnowski, T. J., & Steriade, M. (1996). Control of spatiotemporal coherence of a thalamic oscillation by corticothalamic feedback. *Science*, 274(5288), 771–774.
- Coppola, D., & Purves, D. (1996). The extraordinarily rapid disappearance of entopic images. *Proceedings of the National Academy of Science, USA*, 93(15), 8001–8004.
- Crick, F. (1984). Function of the thalamic reticular complex: the searchlight hypothesis. *Proceedings of the National Academy of Science, USA*, 81(14), 4586–4590.
- Crick, F., & Koch, C. (1990). Towards a neurobiological theory of consciousness. *Seminars in Neuroscience*, 2, 263–275.
- deCharms, R. C., & Zador, A. (2000). Neural representation and the cortical code. *Annual Review in Neuroscience*, 23, 613–647.
- Delorme, A., & Thorpe, S. J. (2001). Face identification using one spike per neuron: resistance to image degradations. *Neural Networks*, 14(6–7), 795–803.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, 18, 193–222.
- Diesmann, M., Gewaltig, M. O., & Aertsen, A. (1999). Stable propagation of synchronous spiking in cortical neural networks. *Nature*, 402(6761), 529–533.
- Russo, F., & Spinelli, D. (1999a). Electrophysiological evidence for an early attentional mechanism in visual processing in humans. *Vision Research*, 39(18), 2975–2985.
- Di Russo, F., & Spinelli, D. (1999b). Spatial attention has different effects on the magno- and parvocellular pathways. *Neuroreport*, 10(13), 2755–2762.
- Ditchburn, R. W., & Ginsborg, B. L. (1952). Vision with a stabilized retinal image. *Nature*, 170, 36–37.
- 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? Multiple electrode and correlation analyses in the cat. *Biological Cybernetics*, 60(2), 121–130.
- Engel, A. K., König, P., & Singer, W. (1991a). Direct physiological evidence for scene segmentation by temporal coding. *Proceedings of the National Academy of Science, USA*, 88(20), 9136–9140.
- Engel, A. K., Kreiter, A. K., König, P., & Singer, W. (1991b). Synchronization of oscillatory neuronal responses between striate and extrastriate visual cortical areas of the cat. *Proceedings of the National Academy of Science, USA*, 88(14), 6048–6052.
- Engel, A. K., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences*, 5(1), 16–25.
- Fahle, M. (1993). Figure-ground discrimination from temporal information. *Proceedings of the Royal Society of London Series B Biological Sciences*, 254(1341), 199–203.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local “association field”. *Vision Research*, 33(2), 173–193.
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, 291(5502), 312–316.
- Frien, A., Eckhorn, R., Bauer, R., Woelbern, T., & Kehr, H. (1994). Stimulus-specific fast oscillations at zero phase between visual areas V1 and V2 of awake monkey. *Neuroreport*, 5(17), 2273–2277.

- Fries, P., Neuenschwander, S., Engel, A. K., Goebel, R., & Singer, W. (2001). Rapid feature selective neuronal synchronization through correlated latency shifting. *Natural Neuroscience*, 4(2), 194–200.
- Fukushima, K., & Miyake, S. (1982). Neocognitron: a new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern Recognition*, 15, 455–469.
- Galarreta, M., & Hestrin, S. (1999). A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature*, 402(6757), 72–75.
- Gautrais, J., & Thorpe, S. J. (1998). Rate coding vs temporal order coding: a theoretical approach. *Biosystems*, 48(1–3), 57–65.
- Gawne, T. J., Kjaer, T. W., & Richmond, B. J. (1996). Latency: another potential code for feature binding in striate cortex. *Journal of Neurophysiology*, 76(2), 1356–1360.
- Gerstner, W. (2000). Population dynamics of spiking neurons: fast transients, asynchronous states, and locking. *Neural Computation*, 12(1), 43–89.
- Ghazanfar, A. A., & Nicolelis, M. A. (2001). Feature article: The structure and function of dynamic cortical and thalamic receptive fields. *Cerebral Cortex*, 11(3), 183–193.
- Gho, M., & Varela, F. J. (1988). A quantitative assessment of the dependency of the visual temporal frame upon the cortical rhythm. *Journal of Physiology*, 83(2), 95–101.
- Gilbert, C. D., Das, A., Ito, M., Kapadia, M., & Westheimer, G. (1996). Spatial integration and cortical dynamics. *Proceedings of the National Academy of Science, USA*, 93(2), 615–622.
- Gove, A., Grossberg, S., & Mingolla, E. (1995). Brightness perception, illusory contours, and corticogeniculate feedback. *Visual Neuroscience*, 12(6), 1027–1052.
- Gray, C. M., König, P., Engel, A. K., & Singer, W. (1989). Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties. *Nature*, 338(6213), 334–337.
- Gray, C. M., Engel, A. K., König, P., & Singer, W. (1992). Synchronization of oscillatory neuronal responses in cat striate cortex: temporal properties. *Visual Neuroscience*, 8(4), 337–347.
- Grossberg, S., & Mingolla, E. (1985). Neural dynamics of perceptual grouping: textures, boundaries, and emergent segmentations. *Perception and Psychophysics*, 38(2), 141–171.
- Hamker, F. H. (2000). Distributed competition in directed attention. In V. G. Barattoff & H. Neuman (Eds.), *Proceedings in Artificial Intelligence (Vol. 9)*, pp. 39–44). Berlin, AKA: Akademische Verlagsgesellschaft.
- Harter, M. R. (1967). Excitability cycles and cortical scanning: a review of two hypotheses of central intermittency in perception. *Psychological Bulletin*, 68(1), 47–58.
- Heitger, F., & von der Heydt, R. (1993). *A computational model of neural contour processing: Figure-ground segregation and illusory contours*. Proceedings of the 4th International Conference on Computer Vision (pp. 32–40).
- Hikosaka, O., Miyauchi, S., & Shimojo, S. (1991). Focal visual attention produces motion sensation in lines. *Investigative Ophthalmology and Visual Science*, 32(Suppl.), 716.
- Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993a). Focal visual attention produces illusory temporal order and motion sensation. *Vision Research*, 33(9), 1219–1240.
- Hikosaka, O., Miyauchi, S., & Shimojo, S. (1993b). Visual attention revealed by an illusion of motion. *Neurosci Research*, 18(1), 11–18.
- Hopfield, J. J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376, 33–36.
- Hupe, J. M., James, A. C., Girard, P., Lomber, S. G., Payne, B. R., & Bullier, J. (2001). Feedback connections act on the early part of the responses in monkey visual cortex. *Journal of Neurophysiology*, 85(1), 134–145.
- Itti, L., & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention. *Vision Research*, 40(10–12), 1489–1506.
- Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Natural Review of Neuroscience*, 2(3), 194–203.
- Kandil, F. I., & Fahle, M. (2001). Purely temporal figure-ground segregation. *European Journal of Neuroscience*, 13(10), 2004–2008.
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys. *Neuron*, 15(4), 843–856.
- Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, 282(5386), 108–111.
- Keyser, C., Xiao, D.-K., Foldiak, P., & Perrett, D. I. (2001). The speed of sight. *Journal of Cognitive Neuroscience*, 13(1), 1–12.
- Kingstone, A. (1992). Combining expectancies. *Quarterly Journal of Experimental Psychology*, 44, 69–104.
- Kirschfeld, K., & Kammer, T. (2000). Visual attention and metacontrast modify latency to perception in opposite directions. *Vision Research*, 40(9), 1027–1033.
- Koch, C., & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry. *Human Neurobiology*, 4(4), 219–227.
- Koch, C. (1997). Computation and the single neuron. *Nature*, 385, 207–210.
- König, P., Engel, A. K., Roelfsema, P. R., & Singer, W. (1995a). How precise is neuronal synchronization? *Neural Computation*, 7(3), 469–485.
- König, P., Engel, A. K., & Singer, W. (1995b). Relation between oscillatory activity and long-range synchronization in cat visual cortex. *Proceedings of the National Academy of Sciences, USA*, 92(1), 290–294.
- König, P., & Engel, A. K. (1995). Correlated firing in sensory-motor systems. *Current Opinion in Neurobiology*, 5(4), 511–519.
- Kreiter, A. K., & Singer, W. (1996). Stimulus-dependent synchronization of neuronal responses in the visual cortex of the awake macaque monkey. *Journal of Neuroscience*, 16(7), 2381–2396.
- Lal, R., & Friedlander, M. J. (1989). Gating of retinal transmission by afferent eye position and movement signals. *Science*, 243(4887), 93–96.
- Lal, R., & Friedlander, M. J. (1990a). Effect of passive eye movement on retinogeniculate transmission in the cat. *Journal of Neurophysiology*, 63(3), 523–538.
- Lal, R., & Friedlander, M. J. (1990b). Effect of passive eye position changes on retinogeniculate transmission in the cat. *Journal of Neurophysiology*, 63(3), 502–522.
- Lampl, I., & Yarom, Y. (1993). Subthreshold oscillations of the membrane potential: a functional synchronizing and timing device. *Journal of Neurophysiology*, 70(5), 2181–2186.
- Lee, D. K., Itti, L., Koch, C., & Braun, J. (1999). Attention activates winner-take-all competition among visual filters. *Natural Neuroscience*, 2(4), 375–381.
- Leonards, U., Singer, W., & Fahle, M. (1996). The influence of temporal phase differences on texture segmentation. *Vision Research*, 36(17), 2689–2697.
- Lestienne, R. (1996). Determination of the precision of spike timing in the visual cortex of anaesthetised cats. *Biological Cybernetics*, 74(1), 55–61.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10(4), 903–940.
- Lichtenstein, M. (1961). Phenomenal simultaneity with irregular timing of components of the visual stimulus. *Perceptual and Motor Skills*, 12, 47–60.
- Lindsley, D. B. (1952). Psychological phenomena and the electroencephalogram. *Electroencephalography and Clinical Neurophysiology*, 4, 443–456.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2,

- and V4 of macaque visual cortex. *Journal of Neurophysiology*, 77(1), 24–42.
- Maass, W. (1997). Fast sigmoidal networks via spiking neurons. *Neural Computation*, 9(2), 279–304.
- Mainen, Z. F., & Sejnowski, T. J. (1995). Reliability of spike timing in neocortical neurons. *Science*, 268, 1503–1506.
- Mangun, G. R. (1995). Neural mechanisms of visual selective attention. *Psychophysiology*, 32(1), 4–18.
- Martinez-Conde, S., Macknik, S. L., & Hubel, D. H. (2000). Microsaccadic eye movements and firing of single cells in the striate cortex of macaque monkeys. *Natural Neuroscience*, 3(3), 251–258.
- McClurkin, J. W., Gawne, T. J., Optican, L. M., & Richmond, B. J. (1991). Lateral geniculate neurons in behaving primates. II. Encoding of visual information in the temporal shape of the response. *Journal of Neurophysiology*, 66(3), 794–808.
- Meister, M., & Berry, II, M. J. (1999). The neural code of the retina. *Neuron*, 22(3), 435–450.
- Miyashita, Y., & Hayashi, T. (2000). Neural representation of visual objects: encoding and top-down activation. *Current Opinion in Neurobiology*, 10(2), 187–194.
- Moghaddam, B., & Pentland, A. 1995. *Probabilistic visual learning for object detection*, Paper presented at the Fifth International Conference on Computer Vision, Cambridge, MA.
- Moran, J., & Desimone, R. (1985). Selective attention gates visual processing in the extrastriate cortex. *Science*, 229, 782–784.
- Motter, B. C. (1993). Focal attention produces spatially selective processing in visual cortical areas V1, V2, and V4 in the presence of competing stimuli. *Journal of Neurophysiology*, 70, 909–919.
- Moutoussis, K., & Zeki, S. (1997a). A direct demonstration of perceptual asynchrony in vision. *Proceedings of the Royal Society of London Series B, Biological Sciences*, 264(1380), 393–399.
- Moutoussis, K., & Zeki, S. (1997b). Functional segregation and temporal hierarchy of the visual perceptive systems. *Proceedings of the Royal Society of London Series B, Biological Sciences*, 264(1387), 1407–1414.
- Mozer, M. C., & Sittin, M. (1998). Computational modeling of spatial attention. In H. Pashler (Ed.), *Attention* (pp. 341–393). New York: Psychology Press.
- Nakamura, K. (1998). Neural processing in the subsecond time range in the temporal cortex. *Neural Computation*, 10(3), 567–595.
- Neumann, O., Esselmann, U., & Klotz, W. (1993). Differential effects of visual-spatial attention on response latency and temporal-order judgment. *Psychological Research*, 56(1), 26–34.
- Nowak, L. G., Sanchez-Vives, M. V., & McCormick, D. A. (1997). Influence of low and high frequency inputs on spike timing in visual cortical neurons. *Cerebral Cortex*, 7(6), 487–501.
- Nowak, L., & Bullier, J. (1998). The timing of information transfer in the visual system. In J. H. Kaas, K. Rockland, & A. Peters (Eds.), *Cerebral cortex* (pp. 205–241). New York: Plenum Press.
- Olshausen, B. A., Anderson, C. H., & Van Essen, D. C. (1993). A neurobiological model of visual attention and invariant pattern recognition based on dynamic routing of information. *Journal of Neuroscience*, 13(11), 4700–4719.
- Opara, R., & Wörgötter, F. (1996). Using visual latencies to improve image segmentation. *Neural Computation*, 8(7), 1493–1520.
- Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *Journal of Neurophysiology*, 68(1), 70–84.
- Panzeri, S., Treves, A., Schultz, S., & Rolls, E. T. (1999). On decoding the responses of a population of neurons from short time windows. *Neural Computation*, 11(7), 1553–1577.
- Panzeri, S., Petersen, R. S., Schultz, S. R., Lebedev, M., & Diamond, M. E. (2001). The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron*, 29(3), 769–777.
- Parodi, O., Combe, P., & Ducom, J. C. (1996). Temporal coding in vision: coding by the spike arrival times leads to oscillations in the case of moving targets. *Biological Cybernetics*, 74(6), 497–509.
- Pashler, H. (1998). *The psychology of attention*. Cambridge, MA: MIT Press.
- Perkel, D. H., & Bullock, T. H. (1968). Neural coding. *Neurosciences Research Program Bulletin*, 6(3), 221–348.
- Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967a). Neuronal spike trains and stochastic point processes I. The single spike train. *Biophysical Journal*, 7, 391–418.
- Perkel, D. H., Gerstein, G. L., & Moore, G. P. (1967b). Neuronal spike trains and stochastic point processes II. Simultaneous spike trains. *Biophysical Journal*, 7, 419–440.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329–342.
- Petersen, R. S., Panzeri, S., & Diamond, M. E. (2001). Population coding of stimulus location in rat somatosensory cortex. *Neuron*, 32(3), 503–514.
- Pitts, W., & McCulloch, W. W. (1947). How we know universals: the perception of auditory and visual forms. *Bulletin of Mathematical Biophysics*, 9, 127–147.
- Posner, M. I., Snyder, C. R. R., & Davidson, B. J. (1980). Attention and the detection of signals. *Journal of Experimental Psychology: General*, 109, 160–174.
- Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology*, 81(1), 10–15.
- Proverbio, A. M., & Mangun, G. R. (1994). Electrophysiological and behavioral “costs” and “benefits” during sustained visual-spatial attention. *International Journal of Neurosciences*, 79(3–4), 221–233.
- Prut, Y., Vaadia, E., Bergman, H., Haalman, I., Slovlin, H., & Abeles, M. (1998). Spatiotemporal structure of cortical activity: properties and behavioral relevance. *Journal of Neurophysiology*, 79(6), 2857–2874.
- Reich, D. S., Victor, J. D., Knight, B. W., Ozaki, T., & Kaplan, E. (1997). Response variability and timing precision of neuronal spike trains in vivo. *Journal of Neurophysiology*, 77(5), 2836–2841.
- Reinagel, P., & Reid, R. C. (2000). Temporal coding of visual information in the thalamus. *Journal of Neuroscience*, 20(14), 5392–5400.
- Reynolds, J. H., & Desimone, R. (1999). The role of neural mechanisms of attention in solving the binding problem. *Neuron*, 24(1), 19–29, 111–125.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, 19(5), 1736–1753.
- Reynolds, J. H., Pasternak, T., & Desimone, R. (2000). Attention increases sensitivity of V4 neurons. *Neuron*, 26(3), 703–714.
- Richmond, B. J., Optican, L. M., Podell, M., & Spitzer, H. (1987). Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. *Journal of Neurophysiology*, 57(1), 132–146.
- Richmond, B. J., Optican, L. M., & Spitzer, H. (1990). Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. I. Stimulus-response relations. *Journal of Neurophysiology*, 64(2), 351–369.
- Rieke, F., Warland, D., de Ruyter van Steveninck, R. R., & Bialek, W. (1997). *Spikes: exploring the neural code*. Cambridge, MA: MIT.
- Riesenhuber, M., & Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Natural Neuroscience*, 2(11), 1019–1025.
- Riggs, L. A., & Ratcliff, F. (1952). The effects of counteracting the normal movements of the eye. *Journal of Optical Society of America*, 42, 872–873.
- Rolls, E., & Treves, A. (1998). *Neural networks and brain function*. New York: Oxford University Press.

- Rowley, H. A., Baluja, S., & Kanade, T. (1998). Neural network-based face detection. *IEEE Transactions on Pattern Analog Machine Intelligence*, 20, 23–38.
- Rumelhart, D. E., & McClelland, J. L. (1986). *Parallel distributed processing explorations in the microstructure of cognition*. Foundations, Vol. 1. Cambridge: MIT Press/Bradford Books.
- Salinas, E., & Abbott, L. F. (1997). Invariant visual responses from attentional gain fields. *Journal of Neurophysiology*, 77(6), 3267–3272.
- Salinas, E., & Sejnowski, T. J. (2001). Correlated neuronal activity and the flow of neural information. *Natural Review of Neuroscience*, 2(8), 539–550.
- Schmolesky, M. T., Wang, Y., Hanes, D. P., Thompson, K. G., Leutgeb, S., Schall, J. D., & Leventhal, A. G. (1998). Signal timing across the macaque visual system. *Journal of Neurophysiology*, 79(6), 3272–3278.
- Shadlen, M. N., & Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Current Opinion in Neurobiology*, 4(4), 569–579.
- Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *Journal of Neuroscience*, 18(10), 3870–3896.
- Shallice, T. (1964). The detection of change and the perceptual moment hypothesis. *British Journal of Statistical Psychology*, 17, 113–135.
- Shashua, A., & Ullman, S. (1988). *Structural saliency*. Proceedings of the international conference on computer vision (pp. 482–488).
- Shiu, L.-P., & Pashler, H. (1993). *Spatial precuing in single-element displays: noise-reduction or signal enhancement?* Paper presented at the 34th Annual Meeting of the Psychonomic Society, Washington, DC.
- Singer, W. (1993). Synchronization of cortical activity and its putative role in information processing and learning. *Annual Review of Physiology*, 55, 349–374.
- Singer, W., & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555–586.
- Singer, W. (1999). Time as coding space? *Current Opinion in Neurobiology*, 9(2), 189–194.
- Smith, A. T., Singh, K. D., & Greenlee, M. W. (2000). Attentional suppression of activity in the human visual cortex. *Neuroreport*, 11(2), 271–277.
- Stanley, G. B., Li, F. F., & Dan, Y. (1999). Reconstruction of natural scenes from ensemble responses in the lateral geniculate nucleus. *Journal of Neuroscience*, 19(18), 8036–8042.
- Steinman, B. A., Steinman, S. B., & Lehmkuhle, S. (1995). Visual attention mechanisms show a center-surround organization. *Vision Research*, 35(13), 1859–1869.
- Stevens, C. F., & Zador, A. (1995). Neural coding: The enigma of the brain. *Current Biology*, 5(12), 1370–1371.
- Stroud, J. (1949). Psychological moment in perception. In H.V. Foerster (Ed.), *Conference on cybernetics* (pp. 27–63). New York: Josiah Macy Jr. Foundation.
- Sung, K., & Poggio, T. (1994). *Example-based learning for view-based human face detection*. Proceedings Image Understanding Workshop II (pp. 843–850).
- Suzuki, W. A. (1996). The anatomy, physiology and functions of the perirhinal cortex. *Current Opinion in Neurobiology*, 6(2), 179–186.
- Swadlow, H. A., & Gusev, A. G. (2002). Receptive-field construction in cortical inhibitory interneurons. *Natural Neuroscience*, 5(5), 403–404.
- Thorpe, S. J., & Imbert, M. (1989). Biological constraints on connectionist models. In R. Pfeifer, Z. Schreier, F. Fogelman-Soulié, & L. Steels (Eds.), *Connectionism in perspective* (pp. 63–92). Amsterdam: Elsevier.
- Thorpe, S. J. (1990). Spike arrival times: a highly efficient coding scheme for neural networks. In R. Eckmiller, G. Hartman, & G. Hauske (Eds.), *Parallel processing in neural systems* (pp. 91–94). North-Holland: Elsevier.
- Thorpe, S. J., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, 381, 520–522.
- Thorpe, S. J., & Gautrais, J. (1997). Rapid visual processing using spike asynchrony. In M. C. Mozer, M. Jordan, & T. Petsche (Eds.), *Advances in neural information processing systems* (Vol. 9, pp. 901–907). Cambridge: MIT Press.
- Thorpe, S. J., & Gautrais, J. (1998). Rank order coding: a new coding scheme for rapid processing in neural networks. In J. Bower (Ed.), *Computational neuroscience: trends in research*. New York: Plenum Press.
- Thorpe, S., Delorme, A., & VanRullen, R. (2001). Spike-based strategies for rapid processing. *Neural Networks*, 14(6–7), 715–725.
- Thorpe, S. J., & Fabre-Thorpe, M. (2001). Seeking categories in the brain. *Science*, 291(5502), 260–263.
- Tootell, R. B., Hadjikhani, N., Hall, E. K., Marrett, S., Vanduffel, W., Vaughan, J. T., & Dale, A. M. (1998). The retinotopy of visual spatial attention. *Neuron*, 21(6), 1409–1422.
- Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14(1), 107–141.
- Treves, A., Rolls, E. T., & Tovee, M. J. (1996). On the time required for recurrent processing in the brain. In V. Torre & F. Conti (Eds.), *Neurobiology: Proceedings of the International School of Biophysics, XXIII Course, May 1995* (pp. 371–382). New York: Plenum Press.
- Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86.
- Usher, M., & Donnelly, N. (1998). Visual synchrony affects binding and segmentation in perception. *Nature*, 394(6689), 179–182.
- Vaadia, E., Haalman, I., Abeles, M., Bergman, H., Prut, Y., Slovlin, H., & Aertsen, A. (1995). Dynamics of neuronal interactions in monkey cortex in relation to behavioural events. *Nature*, 373(6514), 515–518.
- Valentin, D., Abdi, H., O’Toole, A., & Cottrell, G. W. (1994). Connectionist models of face processing: a survey. *Pattern Recognition*, 27, 1209–1230.
- Vanduffel, W., Tootell, R. B., & Orban, G. A. (2000). Attention-dependent suppression of metabolic activity in the early stages of the macaque visual system. *Cerebral Cortex*, 10(2), 109–126.
- VanRullen, R., Gautrais, J., Delorme, A., & Thorpe, S. J. (1998). Face processing using one spike per neuron. *Biosystems*, 48(1–3), 229–239.
- VanRullen, R., & Thorpe, S. J. (1999). Spatial attention in asynchronous neural networks. *NeuroComputing*, 26–27, 911–918.
- VanRullen, R., Delorme, A., & Thorpe, S. (2001). Feed-forward contour integration in primary visual cortex based on asynchronous spike propagation. *Neurocomputing*, 38–40(1–4), 1003–1009.
- VanRullen, R., & Thorpe, S. J. (2001a). The time course of visual processing: from early perception to decision-making. *Journal of Cognitive Neuroscience*, 13(4), 454–461.
- VanRullen, R., & Thorpe, S. J. (2001b). Rate coding vs temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Computation*, 13(6), 1255–1283.
- VanRullen, R., & Thorpe, S. J. (2001c). Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artificial objects. *Perception*, 30(6), 655–668.
- Varela, F. J., Toro, A., John, E. R., & Schwartz, E. L. (1981). Perceptual framing and cortical alpha rhythm. *Neuropsychologia*, 19(5), 675–686.
- Victor, J. D. (2000). How the brain uses time to represent and process visual information. *Brain Research*, 886(1–2), 33–46.
- Volgushev, M., Vidyasagar, T. R., & Pei, X. (1995). Dynamics of the orientation tuning of postsynaptic potentials in the cat visual cortex. *Visual Neuroscience*, 12(4), 621–628.

- Volgushev, M., Chistiakova, M., & Singer, W. (1998). Modification of discharge patterns of neocortical neurons by induced oscillations of the membrane potential. *Neuroscience*, *83*(1), 15–25.
- Wörgötter, F., Opara, R., Funke, K., & Eysel, U. (1996). Utilizing latency for object recognition in real and artificial neural networks. *Neuroreport*, *7*(3), 741–744.
- Wörgötter, F., Suder, K., Zhao, Y., Kerscher, N., Eysel, U. T., & Funke, K. (1998). State-dependent receptive-field restructuring in the visual cortex. *Nature*, *396*(6707), 165–168.
- Zhou, H., Friedman, H. S., & von der Heydt, R. (2000). Coding of border ownership in monkey visual cortex. *Journal of Neuroscience*, *20*(17), 6594–6611.