PERGAMON

Neural Networks 14 (2001) 715–725

Neural Networks

www.elsevier.com/locate/neunet

2001 Special issue

Spike-based strategies for rapid processing

Simon Thorpe*, Arnaud Delorme, Rufin Van Rullen

Centre de Recherche Cerveau & Cognition UMR 5549,133 route de Narbonne, 31062 Toulouse, France

Received 29 January 2001; accepted 29 January 2001

Abstract

Most experimental and theoretical studies of brain function assume that neurons transmit information as a rate code, but recent studies on the speed of visual processing impose temporal constraints that appear incompatible with such a coding scheme. Other coding schemes that use the pattern of spikes across a population a neurons may be much more efficient. For example, since strongly activated neurons tend to fire first, one can use the order of firing as a code. We argue that Rank Order Coding is not only very efficient, but also easy to implement in biological hardware: neurons can be made sensitive to the order of activation of their inputs by including a feed-forward shunting inhibition mechanism that progressively desensitizes the neuronal population during a wave of afferent activity. In such a case, maximum activation will only be produced when the afferent inputs are activated in the order of their synaptic weights. © 2001 Published by Elsevier Science Ltd.

Keywords: Rank Order Coding; Latency; Rapid visual processing; Spikes; Retina; Information

1. Introduction

Most models of neural systems implicitly assume that information is transmitted by neurons in the form of a firing rate code. For example, the vast majority of Artificial Neural Network and Connectionist models use an approach that can be summarized as follows: take a large number of neuron-like processing units, connect them together with variable weight connections that are the rough equivalent of synapses, and use a rule in which the activation level of each unit is some function of the weighed sum of all the inputs to each neuron. It is a strategy that seems obviously 'biologically inspired', but there is one feature of real biological neural networks that is missing from the vast majority of artificial systems. In nearly all artificial systems, each unit sends its activation level to all the targets as a continuous value, often a floating point number between 0 and 1 (sometimes between -1 and $+1$). In contrast, biological neurons send information in the form of a sequence of spikes. The notion that one can summarize a spike train as a single continuous variable is one that is very firmly entrenched, not just in artificial neural networks, but also throughout neuroscience. Indeed, it goes back to the very start of experimental neurophysiology in the 1920s when the first recordings of the electrical activity of sensory fibers by Adrian (1928) showed that firing rate increased with increasing stimulus intensity.

Even today, neurophysiologists often assume that all the useful information that can be learned about neural coding can be summarized in the form of a Post-Stimulus Time Histogram (PSTH) that plots firing rate as a function of time. Given this state of affairs, it is hardly surprising that few in the artificial neural network community have felt the need to look at alternative coding schemes. However, in the last few years, an increasing number of scientists has begun to take seriously the possibility that the use of spikes opens up a whole range of alternative coding options, some of which have profound implications for the nature of neural computation (Maass & Bishop, 1999; Rieke, Warland, Ruyter van Steveninck & Bialek, 1996). One of the motivations behind such work has been the realization that there are situations where processing is too fast to be compatible with a conventional rate based code. We will review such evidence and argue that other alternative spike based coding schemes can be considerably more efficient. In particular, we will discuss the merits of a coding scheme that encodes information in the relative timing of spikes across a population of neurons, or more specifically, in the order in which neurons fire. We will argue that such coding schemes have a number of features that make them ideally suited for certain types of rapid processing tasks. These features include speed, robustness, and ease of implementation, and make such schemes particularly attractive for designing artificial processing systems.

---

* Corresponding author. Tel.: +33-5-62-17-28-03; fax: +33-5-62-17-28-09.

E-mail address: thorpe@cerco.ups-tlse.fr (S. Thorpe).

PII: S0893-6080(01)00083-1

2. The processing speed constraint

In 1989, Thorpe and Imbert argued that the existence of neurons in the primate brain that could respond selectively to complex visual stimuli such as faces, food and familiar 3D objects only 100–150 ms after stimulus onset imposes a major constraint on models of visual processing (Thorpe & Imbert, 1989). They argued that to reach such neurons, information about the stimulus would need to cross something like 10 layers of neurons on the way from the retinal photoreceptors. This means that each individual processing stage would need to operate in not much more that 10 ms. Given that cortical neurons rarely fire at rates much above 100 Hz, this seems to imply that such processing can be accomplished under conditions where each individual neuron only gets to fire either none or one spike. This seriously limits the precision with which individual neurons could send information using a firing rate code because it largely excludes codes that make use of the interspike interval between two spikes as a way of estimating instantaneous firing rate.

In the last decade, the generality of the processing speed constraint has been reinforced. Initially, it was possible to argue that the short response latency of inferotemporal neurons with selective visual responses to stimuli such as faces could constitute a special case. However, more recent work has shown that even complex natural scenes that have never been seen previously can be successfully categorized on the basis of only 150 ms of processing in humans (Antal, Keri, Kovacs, Janka & Benedek, 2000; Thorpe, Fize & Marlot, 1996), and even more rapidly in monkeys (FabreThorpe, Richard & Thorpe, 1998). This type of Ultra-Rapid Visual Categorization (URVC) has a number of interesting features. First, it seems to be largely colorblind, since monochromatic images are processed very efficiently (Delorme, Richard & Fabre-Thorpe, 2000), a result consistent with the idea that it relies primarily on rapid magnocellular visual pathways. Second, categorization can be as rapid for totally novel images as it is for highly familiar images, implying that contextual information is largely unnecessary (Fabre-Thorpe, Delorme, Marlot & Thorpe, 2001). Third, it is not specific for biologically important categories like animals, because it is equally fast and accurate for a completely artifactual category, namely, means of transport (Van Rullen & Thorpe, 2001b). Finally, it does not require direct fixation of the object, since it works very well for stimuli presented in parafoveal vision (Fabre-Thorpe, Fize, Richard & Thorpe, 1998), and has even been demonstrated for images presented in extreme peripheral vision (Thorpe, Gegenfurtner, Fabre-Thorpe & Bülthoff, 1999).

The existence of this very rapid and automatic processing mode provides strong evidence for a largely feed-forward processing mechanism capable of operating very quickly. Other recent evidence also points in the same direction. In 1992, Oram and Perrett demonstrated that even the very start of the neuronal response of neurons in inferotemporal cortex could be highly selective for particular stimuli, a hallmark of feed-forward processing (Oram & Perrett, 1992). Similar early selectivity was also reported for orientation selective neurons in V1 (Celebrini, Thorpe, Trotter & Imbert, 1993). More recently, there have been reports that the selectivity of neurons in high level visual areas such as inferotemporal cortex could withstand very rapid changes in the input, resulting either from masking (Kovacs, Vogels & Orban, 1995; Rolls & Tovee, 1994; Rolls, Tovee & Panzeri, 1999), or by rapid serial visual presentation (Keysers, Xiao, Foldiak & Perrett, 2001).

Together, these various strands of experimental evidence argue strongly in favor of the view that at least some forms of visual processing can be performed very rapidly—so rapidly, that few if any of the neurons at each level of the processing hierarchy will have enough time to emit more than one spike before those in the next layer have to respond.

While vision is fast, processing in other sensory pathways can in many cases be even faster, imposing even more serious constraints. For example, neurons in the bat auditory cortex can respond just 8 ms after stimulus onset, which, given the number of intervening subcortical processing stages leaves only a couple of milliseconds at each level (Jen, Sun & Lin, 1989), and similarly strong constraints will apply in the somatosensory system. Likewise, there are numerous cases in invertebrate sensory systems where the input–output timing constraints are particularly severe (Carr, 1993).

3. Is rate coding fast enough?

Most neurophysiologists believe that a Poisson-like rate code is, to a first approximation, a reasonable description of the way that neurons transmit information. Describing the spike generation process as Poisson is clearly a simplification, because it ignores the fact that real neurons have refractory periods that prevent them from generating a large number of spiking events in a short period. Nevertheless, a Poisson model is a reasonable starting point. Gautrais and Thorpe (1998) looked at the efficiency of such a model as a means of transmitting information. They argued that Poisson rate codes are probably too inefficient to account for the rapid information transmission required for sensory processing.

Suppose that an observer is listening to the output of a single neuron and, during a 10 ms observation window, the neuron emits one spike. If we assume a Poisson process, what can we conclude about the true firing rate of the neuron? It turns out that the most that we can say is that there is a 90% chance that the true firing rate lies somewhere in the range 5–474 Hz (Gautrais & Thorpe, 1998). Clearly, this is unlikely to be of much use for transmitting detailed information about the level of excitation in a sensory receptor. There are two ways in which the accuracy of the measurement can be improved. One is to use a longer observation window, the other to use a population of neurons to transmit the information. In a sense, these are equivalent since observing a single Poisson process for t milliseconds provides the same result as observing n Poisson processes for t/n ms. Thus, by using a population of 30 neurons, and an observed firing rate of 30 spikes in 10 ms, the 90% confidence interval for the population firing rate can be reduced to roughly 100 ± 30 Hz. However, to obtain a precision of 100 ± 10 Hz would require no less than 281 redundant and independent neurons. This seems to be a very expensive way to transmit one analog value with only limited precision.

While it is clear that, given enough neurons, one can obtain whatever level of precision you could want, there are numerous situations where there may simply not be enough neurons available. Take the example of the optic nerve. There are roughly 100 million photoreceptors in the primate retina, and information from these receptors needs to be compressed into the activity of roughly 1 million optic nerve fibers. Since we know that this figure includes both ON- and OFF-center receptive fields as well as transient and sustained channels, we can assume that for any particular type of information there will be no more than a few hundred thousand ganglion cells to cover the entire visual field. Is it really conceivable that one could allow 281 such fibers to be used to transmit the image intensity at each point in the image? If this were really the case, one would only be able to sample about 30 by 30 points in the image per channel, a value that seems totally incompatible with the highly detailed information provided by the retina. Furthermore, although there may be a small degree of redundancy between neighboring retinal ganglion cells, the overlap is relatively small (Meister & Berry, 1999), and in general the available evidence indicates that coding in the retina is designed to eliminate redundancy as much as possible.

While population rate coding seems incompatible with the bandwidth of the optic nerve, it could be argued that population rate coding could be used to transmit information between processing stages further on in the visual system. After all, while there may only be 1 million retinal ganglion cells projecting to the LGN, there are probably hundreds of millions projecting from V1 to extra-striate cortical areas such as V2 and MT. While this might seem like a large number, it needs to be remembered that although there are only a relatively small number of different types of retinal ganglion cells, the number of different image parameters that need to be encoded by the activity of cells in V1 is way higher. Neurons in V1 need to encode a large number of parameters that include orientation, spatial frequency, stereoscopic disparity, color, and motion, but will also include other more complex characteristics. As yet, we have virtually no idea of how much bandwidth is required to transmit information between cortical areas. Nevertheless, it seems likely that the conventional population rate code approach will be hard pushed to cope with the bandwidth requirements of rapid intra-cortical information transfer.

There are other problems facing the traditional rate coding view. One has emerged from a number of recent studies that have examined the way in which synapses respond to repeated activation, for example during a burst of afferent activity. The results are complex, because different results have been reported for different types of synaptic connections (Abbott, Varela, Sen & Nelson, 1997; Thomson, 2000; Thomson, Deuchars & West, 1996; Tsodyks & Markram, 1997). In some cases the response to a second spike arriving via a particular synapse can be transiently enhanced (facilitation). However, in many cases, the effect of a second pulse is significantly attenuated and there may even be total failure if one attempts to reactivate the same synapse without leaving a delay of 50–100 ms for the synapse to recover. Remarkably, this even seems to be true for putative thalamic excitatory inputs to cortical pyramidal cells (Stratford, Tarczyhornoch, Martin, Bannister & Jack, 1996). If such results are confirmed, it would imply that cortical cells are effectively blind to firing rates in the lateral geniculate nucleus above about 10–20 spike.s^{-1}! At the very least such results mean that using the firing rates of individual cells to transmit accurate analog information will be of limited use. Again, the data strongly indicate that we need to think about how information can be coded across a population of cells, rather than thinking about each cell in isolation.

Finally, yet another problem for rate coding stems from the fact that the firing rate distribution of real neurons is not flat, but rather heavily skewed towards low firing rates. Consider again the case of an observer listening to the output of a neuron who detects one spike during a 10 ms time window. What would be the best estimate of the firing rate if we assume a Poisson process? One might think that the correct answer should be 100 spike. s^{-1}. However, this would only be the case if all firing rates were equally likely. Suppose that the distribution of firing rates was actually fitted by a roughly exponential function as suggested by recent experimental data (Baddeley et al., 1997), and that the mean firing rate was 30 spike. s^{-1}. Gautrais and Thorpe (1998) pointed out that under these conditions, the best estimate of the true firing rate given one spike in 10 ms would in fact be 23 spike. s^{-1}. To complicate matters even more, the best estimate of the true firing rate will depend on the length of the observation window, but even when 10 spikes have been observed during 100 ms, the best estimate would still only be 75 spike. s^{-1}. Indeed, in order to obtain an estimated firing rate of 100 spike. s^{-1} with a 10 ms window, the window would need to contain at least four spikes (i.e. an observed firing rate of 400 spike. s^{-1}).

4. Alternative coding schemes

Fortunately, rate coding is by no means the only option available. Over 30 years ago, a meeting on Neural Coding

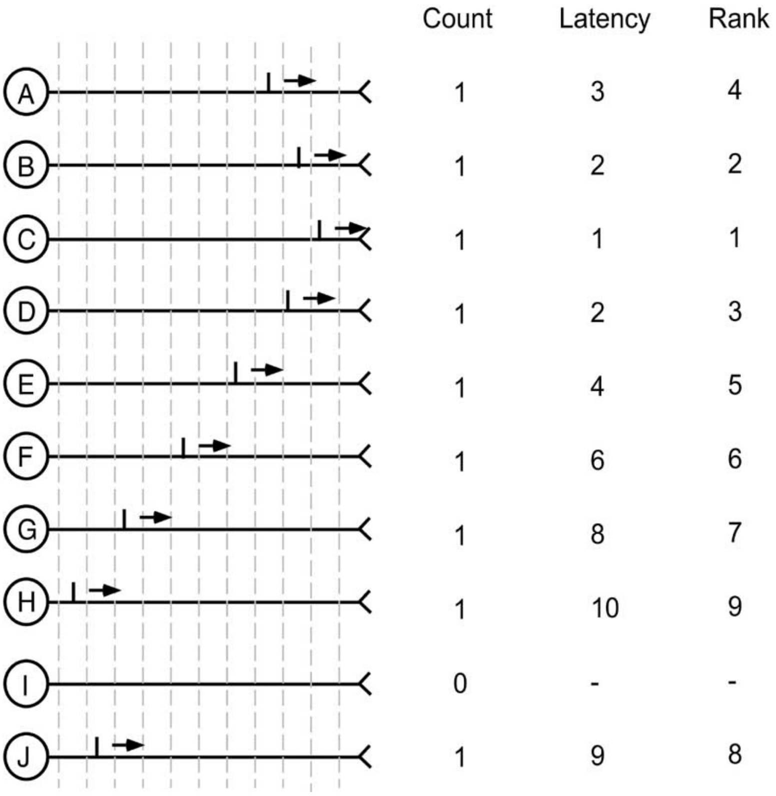

Fig. 1. Comparison between three coding schemes that can operate in a short time window. The 10 neurons A–J emit spikes at different times. By using a count code, corresponding to a population rate code, there are only 10 + 1 states of the system. If the latency of each spike can be determined with millisecond precision, there are 10^{10} possible states. Finally, with a rank code, there are 10! possible states.

assessed the plausibility of a wide range of different coding schemes (Perkel & Bullock, 1968). Many of these alternative schemes are still perfectly viable. Furthermore, in recent years, a number of other coding schemes have been proposed, many of which make use of the fact that real neurons use spikes. Indeed, the fact that neurons use spikes to transmit information opens up a whole new range of coding options, many of which are largely unexplored.

Let us return to the problem posed by the very rapid processing demonstrated by a number of recent studies. Such studies indicate that sophisticated processing can be achieved under conditions where each neuron only gets to fire either none or one spike, considerably reducing the number of possible coding schemes that we need to take into consideration. What are the options? In the following section, we will consider just a few of the more obvious possibilities.

To make the differences as clear as possible, consider the 10 neurons illustrated in Fig. 1. Let us suppose that a stimulus has been presented, and we can record the responses of each neuron during a 10 ms observation window. How might one decode information concerning the stimulus?

4.1. Count code

The first option is simply to count the number of neurons that have spiked during a particular time window. This amounts to using the neurons to implement a population rate code of the type described in the previous section. Of the 10 neurons in Fig. 1, nine have fired a spike during the observation window, which corresponds to a population firing rate of nine/10 spikes per 10 ms, or 90 spikes s^{-1}. With such a coding scheme, it is clear that the maximum amount of information that can be transmitted is equal to $\log_2(N + 1)$ bits, where N is the number of neurons, since there are only $N + 1 = 11$ possible states of the system. This sets the upper limit on the amount of information that can be transmitted using 10 neurons at 3.46 bits.

4.2. Binary code

A more efficient way of using the same 10 neurons would be to use them as a binary code. It would be a bit like using the 10 neurons as if they were lines on a parallel printer port. The pattern corresponding to the situation in Fig. 1 would be the binary pattern 1111111101, one of 1024 possible patterns. In the case of such a binary code, the maximum amount of information that can be transmitted with N neurons will simply be $\log_2(2^N) = 10$ bits. This is clearly a much more efficient code than the simple count code. Note however, that the information provided by the binary coding scheme depends critically on the length of the observation window. If the window had only been 5 ms long, we would have obtained the value 111110000. Binary codes of this sort clearly have the problem that they are not stable over time.

4.3. Timing code

The maximum amount of information is provided when the decoding mechanism can determine the precise time of each spike on each input line. In this case, the total amount of information that can be transmitted will simply depend on the number of channels, and the precision with which the latency of each spike can be determined. If we suppose that spikes can be timed with a precision of 1 ms, the maximum amount of information that could be transmitted in t ms will be $N^*\log_2(t)$ bits, namely over 33 bits. Such timing based codes are clearly potentially extremely powerful, but have the drawback that the decoding mechanism required to determine the precise latency of each spike may be prohibitively complicated to implement when using real neurons.

4.4. Rank order code

Yet another possibility is to look not at the precise timing of spikes for each input, but rather the order in which the neurons fire (Thorpe & Gautrais, 1998). In this case the neurons in Fig. 1 could be thought of as transmitting the order $C > B > D > A > E > F > G > J > H > I$. This particular order is only one of the 10! orders that can be obtained with 10 neurons—more than 3.6 million possibilities. Rank Order Codes can in principle be used to transmit up to $\log_2(N!)$ bits of information, which in the case of 10 neurons is over 21 bits.

4.5. Codes using synchrony

There are numerous other codes that can be used even when each neuron is only allowed to fire a maximum of one spike. One popular option is to use synchrony to link particular groups of neurons as in Fig. 2. Rather than treating the 10 neurons as a binary code (0111011001), the neurons are grouped into two blocks (0111022002), greatly increasing the number of possible combinations. In this case, the number of possible codes that can be transmitted in a particular observation window will depend on the number of different subgroups that can be differentiated. Thus, if the number of possible phases that can be distinguished in a 10 ms time frame was arbitrarily fixed at 3, one could in theory transmit up to 4^{10} patterns this way, allowing a maximum of $\log_2(4^{10})$ bits of information to be transferred, namely 20 bits.

Hopefully it will be obvious from this very rapid and incomplete overview that even with one spike per neuron and a short observation window, there are a large number of potential coding schemes that need to be examined. It is interesting to note that the population rate code, which in this case corresponds to a simple count code, is by far the least efficient of those considered here. Thus, as in the previous section, where we argued that a Poisson like rate code appears too weak to account for the speed and efficiency of information transmission, we again see that the conventional rate based coding schemes fall well behind other alternative strategies.

Temporal codes clearly have a big advantage, but before we can build a model that can take advantage of the potential bandwidth associated with temporal codes, there are two vital issues that need to be addressed. The first concerns the question of how the temporal information might get into the spike patterns in the first place. The second issue concerns how neurons in later processing stages might decode the information.

5. The origin of temporal information

In the last section, we introduced a number of alternative coding strategies that make use of the temporal structure of the spikes produced by a population of neurons. We showed that if it was possible to determine the precise firing time of spikes on each channel, the total amount of information that can be transmitted can be very large. Alternatively, just using the rank order of spikes in different neurons can also be very effective. The question now is, where might such differences in spike timing originate?

One obvious source is the sensory stimulus itself. There are many sensory systems in which the relative time of arrival of stimuli at different receptors is used for processing. Sound localization is an obvious example where the fact that a sound pulse reaches the left ear before the right one is a cue that the sound source lies on the left of the animal. Motion processing in the visual system also depends on differences in the time of arrival of the stimulus on different receptors. The temporal precision of such mechanisms can be very impressive (Carr, 1993). For example, central neurons in the electric fish are sensitive to timing differences of less than a millisecond (Kawasaki, Rose & Heiligenberg, 1988), and in bats the echolocation system relies on timing differences that can be even smaller (Edamatsu & Suga, 1993). There is thus ample evidence that sensory systems can indeed make use of very small differences in the timing of spikes in different populations of neurons.

Note, however, that in nearly all these cases, the timing difference was already present in the sensory stimulus itself. In 1990, we argued that timing differences will emerge at almost every stage of the sensory pathways, simply as a result of the basic integrate and fire properties of neurons and that these timing differences can also be used for processing (Thorpe, 1990).

To explain the principle, note that the basic rate-coding model is based on the notion that sensory neurons can be considered as analog-to-frequency converters—as stimulus intensity increases, firing rate increases as well. But an alternative view would be to consider the neuron as an analog-to-

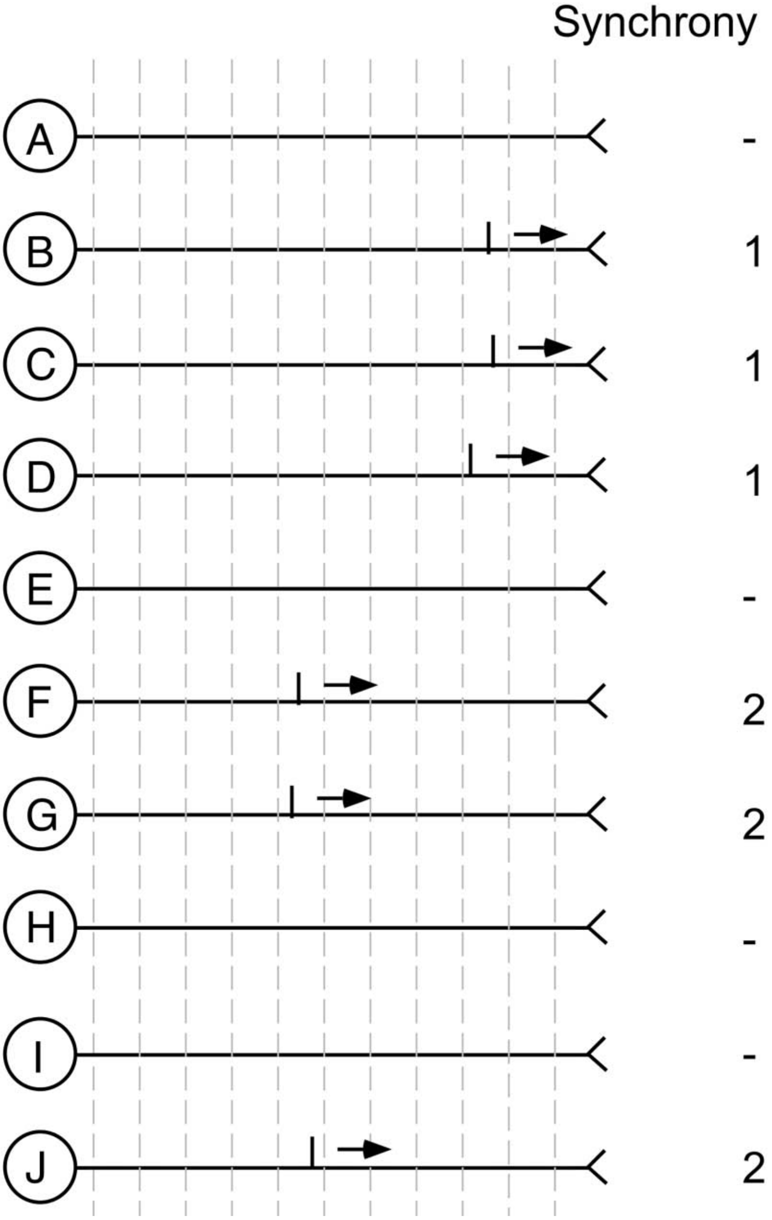

Fig. 2. Diagram illustrating how the use of synchrony can be used to increase the bandwidth of the same group of 10 cells. In this case, neurons are labeled as belonging to one of two groups of synchronized cells.

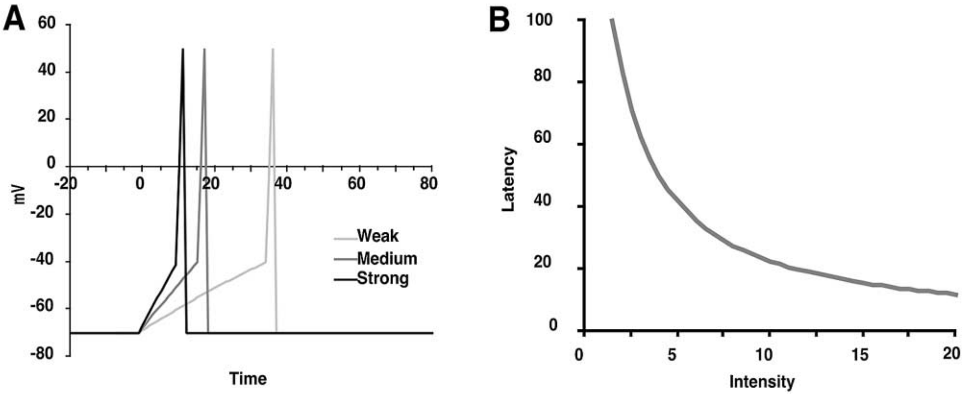

Fig. 3. (A) Illustration of how a simple integrate-and-fire model of a neuron will produce spikes whose latency depends on the intensity of the stimulation. (B) A typical latency-intensity curve showing how the latency of the first spike will vary as a function of input strength.

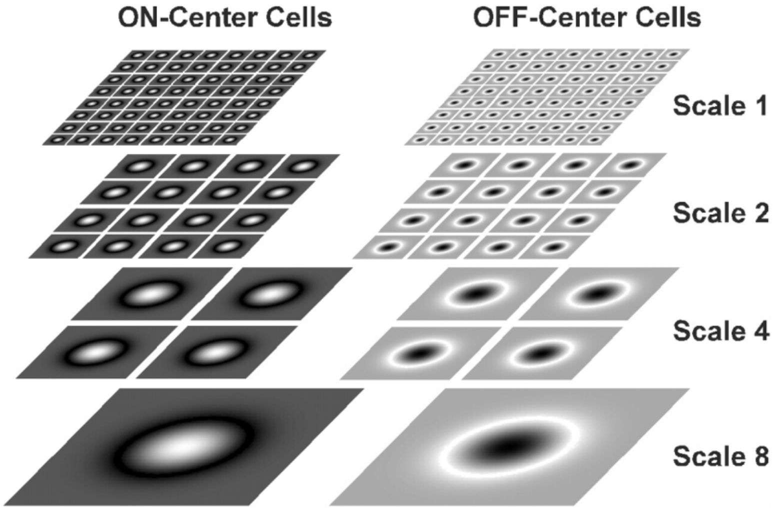

Fig. 4. The simple multi-scale model used in Van Rullen and Thorpe's study of coding in the optic nerve. The model has ON- and OFF-center receptive fields at a range of spatial scales (Van Rullen & Thorpe, 2001a). Sampling density is highest for the cells with the highest spatial resolution.

delay converter, a perfectly reasonable suggestion given the basic integrate-and-fire characteristics of neurons. When a stimulus is presented, the time for the neuron to reach threshold will depend on the strength of the input—with strong stimuli, the neuron will reach threshold rapidly, whereas weaker stimuli will take an increasingly long time to fire (see Fig. 3(A)). Indeed, if the neuron is a leaky integrator, particularly weak stimuli may never manage to get the neuron's membrane potential over threshold. The result is that effectively all sensory neurons will show a characteristic intensity-latency function in which there is a progressive decrease in latency with increasing intensity (see Fig. 3(B)).

Thus, when we look at the pattern of firing illustrated in Fig. 1, it is not unreasonable to suppose that this particular order of firing might result from a situation in which the strength of the sensory stimulus applied to each of the input neurons differs. In this particular case, it is likely that input C receives the strongest input, input B the next strongest and so on. The idea seems very obvious, and yet surprisingly few models in computational neuroscience seem to make use of this simple feature.

Note that the intensity to delay conversion means that both the precise timing of each spike and the order code become viable options. However, in the following sections we will concentrate on the Rank Order Code, because it has a number of interesting computational properties and because it can be implemented very easily. In the next section, we will return to the question of coding in the optic nerve to address the question of how effective rank order coding could be as a means to transmit information from the retina to the brain.

6. Rank order coding in the retina

The integrate and fire properties of retinal ganglion cells mean that, in response to a flashed stimulus, the neurons will tend to fire in an order that reflects the spatial characteristics of the image. The well known center-surround organization of receptive fields in the retina means that local contrast rather than the physical intensity of the stimulus will be most important for determining the responsiveness of retinal ganglion cells. Thus, one could in principle use the order of firing of cells in the optic nerve as a way of encoding the image.

This possibility has recently been explored by Rufin Van Rullen (Van Rullen & Thorpe, 2001a) who examined whether a rank-order coding scheme could be used to efficiently transmit information about an image from the retina to the cortex. The study used a very simple model of the retina in which the image was represented by ganglion cells with ON-center and OFF-center receptive fields at different scales (see Fig. 4).

In response to the presentation of an image, the activation strength of each ganglion cell will effectively reflect the result of a local convolution of the image. The problem was then to investigate how well this activation strength could be transmitted to the brain using different coding strategies. The conventional view is that retinal ganglion cells transmit information about their activation levels in the form of a firing rate code. This assumes that the brain can determine reasonably accurately the firing rates of all the neurons in the optic nerve. But there is an alternative scheme in which one simply needs to determine the order in which the cells in the optic nerve fire. This information can be used to reconstruct the image by plugging in the receptive field of each neuron that fires with a weight that depends on the order with which the cell fired—those cells that fire first are given a high weighting, whereas those that fire later on are given less and less importance. Examples of how this reconstruction scheme operates are illustrated in Fig. 5(A) which illustrates the fact that, when using this rank order scheme, the identity of many of the objects in natural images

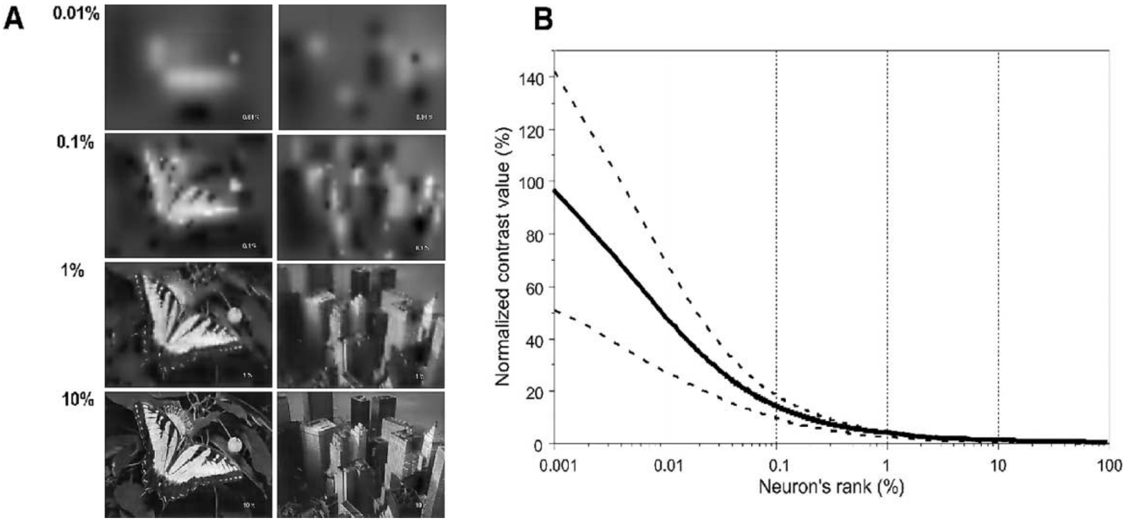

Fig. 5. (A) Progressive reconstruction of two images based on the order of firing of retinal ganglion cells. Results are shown as a function of the number of cells that have fired one spike (adapted from Van Rullen & Thorpe, 2001a).(B) Mean contrast values as a function of the cell's firing rank (as a percentage of the total number of neurons) averaged over more than 3000 natural images.

can be determined when only 1–2% of the cells have fired one spike. Although there is little direct data concerning this question, it seems likely that when a natural image is presented to an array of retinal ganglion cells, 1–2% of cells may well fire a spike during an observation window sufficiently short to be compatible with the temporal constraints mentioned earlier.

Note that in order to obtain this result, the weighting of each retinal spike was adjusted using a form of Look-Up-Table that varies the impact of each spike as a function of its order. In order to determine which values to use, a large number of natural images were tested to determine how the typical contrast values varied with rank. As shown in Fig. 5(B), contrast decreases with rank in a very systematic way. If we take a model retina with 100 000 neurons, and give a maximal weight to the first neurons to fire, once 10 cells have fired, we should reduce the weighting to around 50%, when 100 cells have fired we should use an effective weight of only 15% and when 1000 cells have fired we only need to use a weighting of around 5%. This very rapid drop in importance explains why the first 1% of cells is so good at allowing reconstruction to occur, because it effectively means that we can largely ignore cells that fire later on. Indeed, if the aim of the visual system was to produce an energy efficient code, one that used the smallest number of spikes to transmit the image, one could use leaky integrators in the retina to ensure that, in response to a typical scene, only 1–2% of cells actually reach threshold.

It is interesting to note the parallels between this form of rank order based coding scheme and some of the recent image compression schemes developed by the graphics industry. JPEG compression applies a series of convolutions to an image and then throws away all the components below a certain threshold value—to get higher compression, one simply sets the threshold at a higher level. This is effectively what occurs in the rank order coding model with the interesting twist that the data is sent to the next stage in a temporally ordered sequence, with the most salient data being sent first.

7. Decoding rank order

The rapid decrease in weighting that we use in the image reconstruction can in fact be used more generally as a decoding mechanism for rank based information. The idea is a simple one, and can be implemented in a feed-forward network that includes a population of inhibitory interneurons. Consider the situation in Fig. 6 in which a neuron *N* receives excitatory inputs from five input neurons in the previous layer, but that in addition, each of the input neurons also connects to a population of interneurons (*I*) that produce shunting inhibition in the target cell. Because of the shunting inhibitory circuit, if the input cells fire in a particular order, the first input to fire will produce a maximal effect on the target cell, but the effectiveness of later firing inputs will become progressively attenuated by the build up in inhibition. Suppose that, in addition, the excitatory connection strengths of the five input neurons vary, so that input A has the strongest connection, input B the second strongest connection and so forth. In such a case, it is easy to see that the maximum excitation in the target neuron will be produced when the input neurons fire in the order of their connection strengths, starting with the highest weight (Thorpe & Gautrais, 1998).

To see why, consider the case where the weights of the excitatory synapses from the input neurons A–E are, respectively, 5, 4, 3, 2 and 1. Initially, the modulatory effect of the shunting inhibition is null and so each input is maximally effective. However, every time one of the inputs fires, the

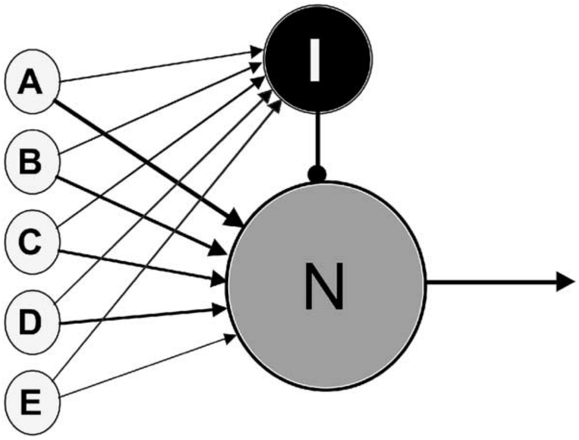

Fig. 6. A simple circuit sensitive to the order of activation of inputs A–E. The neuron *N* receives excitatory inputs from each of the inputs. In addition it receives shunting inhibition from a pool of inhibitory interneurons whose activity increases every time one of the inputs fires. As a result, only the first input to fire is unaffected by the shunting inhibition, and the inhibition builds up progressively during the processing of a wave of spikes. The final activation of the output neuron will be maximal only when the inputs are activated in the order of their weights.

shunting inhibition attenuates the effectiveness of the inputs. A simple model would be to say that after each input spike, the sensitivity of the target cell decreases by 50%. In this case, if the inputs fire in the order $A > B > C > D > E$, the total excitatory input to the target cell will be $(5*0.5^0)+(4*0.5^1) + (3*0.5^2) + (2*0.5^3) + (1*0.5^4) = 8.06$. Any other pattern of firing will produce a lower level of activation with the weakest response being produced when the inputs fire in the opposite order, in which case the final activation would be 3.56. By setting the threshold of the target neuron to an appropriate value, the neuron can be made arbitrarily selective. For example, with the threshold set at 8.0, only the order ABCDE would be capable of driving the neuron over threshold, an impressive result given that there are $5! = 720$ possible orders with five inputs.

One of the appealing features of such a decoding scheme is that it is simple to implement in biological hardware. Most decoding schemes that involve precise temporal information end up involving large numbers of neurons and are thus expensive to implement. For example, sound localization makes use of differences in the time of arrival of spikes originating in the left and right ears, but to decode this information requires expensive delay lines (Carr, 1993). In the case of rank order decoding, on the other hand, a simple feed-forward network with shunting inhibition is sufficient. Interestingly, such an arrangement appears to be remarkably frequent in sensory processing pathways. For example, in the visual system, thalamic afferents originating in the lateral geniculate nucleus will make direct excitatory connections onto the dendrites of pyramidal cells in layers IV and V, but at the same time, they also contact fast-spiking inhibitory interneurons which make contacts on the soma of the same pyramidal cells (Callaway, 1998). It seems plausible to suppose that such connections could produce shunting inhibition of the target cells and could very rapidly reduce the sensitivity of these cells to excitatory inputs. Indeed, recent intracellular recording studies have shown that in response to a visual input, shunting inhibition builds up very rapidly during the first milliseconds of the response (Borg-Graham, Monier & Fregnac, 1998). This is exactly what would be needed to implement the rank decoding scheme proposed here.

8. Learning and rank order coding

A further advantage of the rank order coding scheme is that it is relatively straightforward to implement learning in such a network. As noted in the previous section, a target neuron can be made sensitive to the order of its inputs by using a desensitization mechanism such as shunting inhibition to progressively decrease the effectiveness of inputs arriving later on. In order to make a neuron sensitive to a particular temporal sequence of activation, it is sufficient to use a learning rule that increases synaptic weights for inputs that fire early, and reduces them for inputs that fire later on. Interestingly, this sort of timing dependent synaptic plasticity is precisely what has been described recently in a number of studies (Bi & Poo, 1999; Markram, Lübke, Frotscher & Sakmann, 1997; Song, Miller & Abbott, 2000). Excitatory Post-Synaptic Potentials (EPSPs) that occur before a post-synaptic spike are strengthened, whereas those that fire after the post-synaptic spike become depressed. Further work will be required to see precisely how such rules could relate to Rank Order Coding, but it seems likely that the net result of such a mechanism will be that inputs that always fire in advance of the post-synaptic neuron will receive maximum reinforcement—and this is precisely what we are looking for.

Using SpikeNet, a software program designed for simulating large networks of asynchronously firing integrate-and-fire neurons (Delorme, Van Rullen, Gautrais & Thorpe, 1999) we have already shown that simple feed-forward architectures are capable of performing non-trivial tasks that include the detection of faces in natural images (Van Rullen, Gautrais, Delorme & Thorpe, 1998). More recently, this work has been extended to include the view-independent identification of faces (Delorme & Thorpe, 2001, this volume). Both models have the same basic features. First, the input layer is composed of neurons with ON- and OFF-center receptive fields similar to those of the retinal ganglion cells in Van Rullen's study of retinal coding described earlier. These cells perform a local convolution of the image and then emit one spike at a time that depends on the strength of the signal. These spikes are then sent to arrays of neurons in the next layer that is the rough equivalent of V1. By connecting the ON- and OFF-center cells appropriately, it is relatively simple to produce neurons that are selective to edges with a range of orientations. In both simulations we used eight orientation maps, corresponding to steps of 45°, although there is no particular reason why one should not use a much more detailed set of orientation maps. In the next stage, we use the outputs of the orientation maps to drive feature detection maps that are trained with a supervised learning rule to respond when the neurons in the orientation maps fire in the appropriate order.

9. Extensions to the basic model

Although the face-detection and face-identification models described in the previous section use an entirely feed-forward processing architecture, other work in our group has shown that other types of architecture can be incorporated into the same basic scheme. For example, in a recent study, we showed that horizontal connections between orientation selective maps can be used to implement contour integration even under conditions where each neuron only gets to fire one spike (Van Rullen et al., 2001a, 2001b). The pattern of connectivity used is very much the sort that has been used in a number of recent models of contour processing (Li, 1998). However, because in our simulations each neuron only fires one spike we can effectively prevent the system from using a true iterative mechanism. The result is that the model can perform contour integration very rapidly, in a way that is compatible with the severe timing constraints on visual processing imposed by studies on ultra-rapid visual categorization.

Another extension to the basic model is to allow top-down influences to modulate processing in earlier stages. This can be used to implement a simple mechanism for spatial attention in which processing of a particular region of the visual scene can be given higher priority by effectively lowering the threshold of neurons with receptive fields in that area (Van Rullen & Thorpe, 1999). Note that this effective threshold reduction could easily be obtained by providing an increase in tonic background excitation to neurons in a particular region. The net result is that information concerning the attended region of visual space will reach later stages in the processing hierarchy earlier, thus giving them a competitive advantage.

10. Concluding comments

Rate coding has dominated almost all theoretical and experimental work on neural function for more than half a century. The idea that the output of a neuron can be distilled into a single number is certainly an appealing simplification, and one that has proved useful in a great deal of theoretical work. However, real neurons transmit information as spikes, and as soon as one tries to implement even the simplest rate coding model with real neurons that produce real spikes, things start to get very complicated. As we have tried to argue in this paper, the remarkable speed of sensory processing, together with the anatomical and physiological constraints mean that a simple rate coding scheme is almost certainly inadequate. Processing is too fast to allow the firing rates of individual neurons to be measured with any precision, and although one certainly can attempt to measure the firing rate of a population of neurons, it turns out that this is an extremely inefficient way of transmitting information.

Fortunately, the fact that neurons use spikes opens up a huge range of other potential codes, many of which have received little or no attention. In this paper we have stressed the possibility that sensory systems could easily make use of the fact that the timing of the first spike in response to a stimulus provides information about the strength of that stimulus. This simple fact is one that should be obvious to anyone with even the most rudimentary knowledge of neurophysiology. And yet, the possibility that the order in which neurons fire could be used to encode information is one that has hardly been tested experimentally. Our simulations have shown that networks of asynchronously firing neurons of this sort can indeed perform sophisticated tasks that include face detection and identification. The fact that such simulations can work even under conditions where each neuron only gets to fire one spike is an important one, because it is one of only a very few mechanisms that are compatible with the extremely severe temporal constraints imposed by visual processing.

Clearly, a great deal of further work will be required if we are to fully understand the computational implications of spiking neurons. This work will require close interactions between experimentalists, theoreticians and modelers. But it is an area that holds a great deal of promise and may well provide a key to our understanding of the brain in decades to come.

References

Abbott, L. F., Varela, J. A., Sen, K., & Nelson, S. B. (1997). Synaptic depression and cortical gain control. *Science*, *275* (5297), 220–224.

Adrian, E. D. (1928). *The basis of sensation*, New York: Norton.

Antal, A., Keri, S., Kovacs, G., Janka, Z., & Benedek, G. (2000). Early and late components of visual categorization: an event-related potential study. *Brain Res. Cogn. Brain Res.*, *9* (1), 117–119.

Baddeley, R., Abbott, L. F., Booth, M. C., Sengpiel, F., Freeman, T., Wakeman, E. A., & Rolls, E. T. (1997). Responses of neurons in primary and inferior temporal visual cortices to natural scenes. *Proc. R. Soc. Lond. B. Biol. Sci.*, *264* (1389), 1775–1783.

Bi, G., & Poo, M. (1999). Distributed synaptic modification in neural networks induced by patterned stimulation. *Nature*, *401* (6755), 792–796.

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.

Callaway, E. M. (1998). Local circuits in primary visual cortex of the macaque monkey. *Annu. Rev. Neurosci.*, *21*, 47–74.

Carr, C. E. (1993). Processing of temporal information in the brain. *Annu. Rev. Neurosci.*, *16*, 223–243.

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.

Delorme, A., Richard, G., & Fabre-Thorpe, M. (2000). Ultra-rapid categorisation of natural scenes does not rely on colour cues: A study in monkeys and humans. *Vision Res.*, *40* (16), 2187–2200.

Delorme, A., & Thorpe, S. J. (2001). Face identification with one spike per neuron: resistance to image degrations. *J. Neural Networks*, *14*, 795–803.

Delorme, A., Van Rullen, R., Gautrais, J., & Thorpe, S. (1999). SpikeNET: A simulator for modeling large networks of integrate and fire neurones. *Neurocomputing*, *26-27*, 989–996.

Edamatsu, H., & Suga, N. (1993). Differences in response properties of neurons between two delay-tuned areas in the auditory cortex of the mustached bat. *J. Neurophysiol.*, *69* (5), 1700–1712.

Fabre-Thorpe, M., Delorme, A., Marlot, C., & Thorpe, S. (2001). A limit to the speed of processing in ultra-rapid visual categorization of novel natural scenes. *J. Cogn. Neurosci.*, *13* (2), 171–180.

Fabre-Thorpe, M., Fize, D., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of extrafoveal natural images: Implications for biological models. In J. Bower, *Computational neuroscience: Trends in research 1998* (pp. 7–12). New York: Plenum Press.

FabreThorpe, M., Richard, G., & Thorpe, S. J. (1998). Rapid categorization of natural images by rhesus monkeys. *Neuroreport*, *9* (2), 303–308.

Gautrais, J., & Thorpe, S. (1998). Rate coding versus temporal order coding: A theoretical approach. *Biosystems*, *48* (1–3), 57–65.

Jen, P. H., Sun, X. D., & Lin, P. J. (1989). Frequency and space representation in the primary auditory cortex of the frequency modulating bat Eptesicus fuscus. *J. Comp. Physiol. [A]*, *165* (1), 1–14.

Kawasaki, M., Rose, G., & Heiligenberg, W. (1988). Temporal hyperacuity in single neurons of electric fish. *Nature*, *336*, 173–176.

Keysers, C., Xiao, D. K., Foldiak, P., & Perrett, D. I. (2001). The speed of sight. . *J. Cogn. Neuroscience*, *13*, 90–101.

Kovacs, G., Vogels, R., & Orban, G. A. (1995). Cortical correlate of pattern backward masking. *Proc. Natl. Acad. Sci. USA*, *92* (12), 5587–5591.

Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Comput.*, *10* (4), 903–940.

Maass, W., & Bishop, C. M. (1999). *Pulsed neural networks*, Cambridge, MA: MIT Press.

Markram, H., Lübke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Nature*, *275*, 213–215.

Meister, M., & Berry II, M. J. (1999). The neural code of the retina. *Neuron*, *22* (3), 435–450.

Oram, M. W., & Perrett, D. I. (1992). Time course of neural responses discriminating different views of the face and head. *J. Neurophysiol.*, *68* (1), 70–84.

Perkel, D. H., & Bullock, T. H. (1968). Neural coding. *Neuroscience Research Program Bulletin*, *63*, 221–348.

Rieke, F., Warlan, D., Ruyter van Steveninck, R., & Bialek, W. (1996). *Spikes: Exploring the neural code*, Cambridge, MA: MIT Press.

Rolls, E. T., & Tovee, M. J. (1994). Processing speed in the cerebral cortex and the neurophysiology of visual masking. *Proc. R. Soc. Lond. B. Biol. Sci.*, *257* (1348), 9–15.

Rolls, E. T., Tovee, M. J., & Panzeri, S. (1999). The neurophysiology of backward visual masking: information analysis. *J. Cogn. Neurosci.*, *11* (3), 300–311.

Song, S., Miller, K. D., & Abbott, L. F. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity [In Process Citation]. *Nat. Neurosci.*, *3* (9), 919–926.

Stratford, K. J., Tarczyhornoch, K., Martin, K. A. C., Bannister, N. J., & Jack, J. J. B. (1996). Excitatory synaptic inputs to spiny stellate cells in cat visual cortex. *Nature*, *382* (6588), 258–261.

Thomson, A. M. (2000). Facilitation, augmentation and potentiation at central synapses. *Trends Neurosci.*, *23* (7), 305–312.

Thomson, A. M., Deuchars, J., & West, D. C. (1996). Neocortical local synaptic circuitry revealed with dual intracellular recordings and biocytin-filling. *J. Physiology—Paris*, *9* (3–4), 211–215.

Thorpe, S. J. (1990). Spike arrival times: A highly efficient coding scheme for neural networks. In R. Eckmiller, G. Hartman & G. Hauske, *Parallel processing in neural systems* (pp. 91–94). North-Holland: Elsevier.

Thorpe, S., Fize, D., & Marlot, C. (1996). Speed of processing in the human visual system. *Nature*, *381* (6582), 520–522.

Thorpe, S. J., & Gautrais, J. (1998). Rank order coding. In J. Bower, *Computational neuroscience: Trends in research 1998* (pp. 113–118). New York: Plenum Press.

Thorpe, S., Gegenfurtner, K., Fabre-Thorpe, M., & Bülthoff, H. H. (1999). Categorisation of complex natural images in extreme peripheral vision. *Perception*, *28*, 61.

Thorpe, S. J., & Imbert, M. (1989). Biological constraints on connectionist models. In R. Pfeifer, Z. Schreter & F. Fogelman-Soulié, *Connectionism in perspective* (pp. 63–92). Amsterdam: Elsevier.

Tsodyks, M. V., & Markram, H. (1997). The neural code between neocortical pyramidal neurons depends on neurotransmitter release probability. *Proc. Natl. Acad. Sci. USA*, *94* (2), 719–723.

Van Rullen, R., Gautrais, J., Delorme, A., & Thorpe, S. (1998). Face processing using one spike per neurone. *Biosystems*, *48* (1–3), 229–239.

Van Rullen, R., & Thorpe, S. (1999). Spatial attention in asynchronous neural networks. *Neurocomputing*, *26–27*, 911–918.

Van Rullen, R., & Thorpe, S. J. (2001a). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput.*, *13*, 1255–1283.

Van Rullen, R., & Thorpe, S. J. (2001b). The time course of visual processing: from early perception to decision-making. *J. Cogn. Neurosci.*, *13*, 454–461.