Temporal codes and sparse representations:
A key to understanding rapid processing in the visual system

Rudy Guyonneau *, Rufin VanRullen, Simon J. Thorpe [1]

Centre de Recherche "Cerveau et Cognition", UMR 5547, Faculté de Médecine, 133 route de Narbonne, 31000 Toulouse, France

Abstract

Where neural information processing is concerned, there is no debate about the fact that spikes are the basic currency for transmitting information between neurons. *How* the brain actually uses them to encode information remains more controversial. It is commonly assumed that neuronal firing rate is the key variable, but the speed with which images can be analysed by the visual system poses a major challenge for rate-based approaches. We will thus expose here the possibility that the brain makes use of the spatio-temporal structure of spike patterns to encode information. We then consider how such rapid selective neural responses can be generated rapidly through spike-timing-dependent plasticity (STDP) and how these selectivities can be used for visual representation and recognition. Finally, we show how temporal codes and sparse representations may very well arise one from another and explain some of the remarkable features of processing in the visual system.
© 2005 Elsevier Ltd. All rights reserved.

Keywords: Visual processing; Temporal codes; Sparse representations; Learning, STDP

1. Introduction

Ever since Adrian demonstrated in the 1920s that information about the outside world enters the nervous system in the form of series of action potentials, spiking activity has been thought of as the primary form of neural communication. His observations indicated that the nervous system uses firing rates to encode information because as the intensity of the stimulus was increased, sensory neurons increased the number of action potentials they emitted [3]. Since then, rate-based codes have, for example, been demonstrated to play a role in the perception of motion direction [51]; they can also constitute the relevant variable used in observing brain phenomena (see [48] for an experimental example in visual attention) and investigating theoretical hypotheses [56]. But, more importantly, rate codes dominate the way in which we think about the brain [4]. The major role attributed to firing rates has a direct empirical reason: reporting them is the most simple and practical means to measure neural activity—a complex phenomenon happening at very small spatial and temporal scales in a (supposedly) noisy environment. More fundamentally too, individual spikes did not seem to be temporally precise enough to carry information other than through their mere number [43,55,53].

Yet, a growing number of experimental studies have questioned this postulate, at least where sensory modalities are concerned: spike times can indeed be reproducible, meaning that presenting the organism with the same stimulus can elicit spikes with the same precise timing, in the order of the millisecond (see [34] for a biophysical in vitro study; [8] for experimental evidence; a review in [71]). It thus seems that reliable information can be encoded in the timing of individual spikes, possibly using the degree of synchrony across subsets of neurons [2,54]. Another possibility has to do with the time-to-first-spike

* Corresponding author. Tel.: +335 62 17 28 39; fax: +335 62 17 28 00.
  E-mail addresses: rudy.guyonneau@cerco.ups-tlse.fr (R. Guyonneau), rufin.vanrullen@cerco.ups-tlse.fr (R. VanRullen), simon.thorpe@cerco.ups-tlse.fr (S.J. Thorpe).
  [1] Tel.: +335 62 17 28 03.

[62], a theoretical proposal which has recently received experimental support in humans: Johansson and Birznieks reported that on the basis of the timing of the first spikes generated in the median nerve within the upper arm, one can infer with confidence which stimulus was present in the environment, in that case the direction of the force applied to the fingertip [29]. Although it still remains to be experimentally proven that the brain actually uses a coding scheme based on the fine spatio-temporal structure of activity, spike times—and especially *first* spike times—nonetheless encode more information than firing rates (see [20,45] for an experimental proof).

There are also cases where rate-based models simply have a hard time explaining some remarkable aspects of brain performances, namely those in which speed is of the essence. Experimental studies of neurons in various parts of the monkey brain have demonstrated selective responses to complex visual stimuli such as faces, food and familiar 3D objects only 100–150 ms after stimulus onset [44,49,50]. Similarly, ultra-rapid visual categorisation (URVC) experiments show that a complex natural scene can be categorized based on the presence or absence of a target object (animal, vehicle) in around 150 ms [60,65,66] and even more rapidly in monkeys [16]. Such temporal constraints pose a major challenge to theories of coding if we are to account for rapid processing in the visual system. In order to reach the inferior temporal cortex (IT) where high-level visual descriptions are represented, information about the stimulus needs to cross something like 10 neuronal processing stages on the way from retinal photoreceptors. That leaves only about 10 ms of processing at each stage, a small temporal window which comprises synaptic delays ($\sim$1 ms), post-synaptic integration, spike generation and propagation to the next stage. These propagation times may be considerable because intra-cortical conduction velocities may often be limited to only 1–2 m s^{-1} [12]. Furthermore, firing rates in the visual system typically do not exceed 100 Hz implying that, on average, few cells can fire more than 1 spike in the critical 10 ms window.

It follows that neurons only have time to fire a single spike to achieve rapid visual processing tasks where complex categories such as animals can be detected, effectively ruling out any codes based on more than 1 spike [59].

Under this constraint of one-spike-per-cell coding, how could the brain possibly encode information fast enough to perform such speed-demanding tasks? And how are these selective responses generated? Here, we will discuss the possibility that information resides in the spatio-temporal structure of spike waves (one-spike-per-afferent asynchronous patterns) and argue that, coupled to a biological Hebbian learning rule which is intrinsically suited for temporal codes (spike-timing-dependent plasticity, STDP), neurons can be made to be highly selective to particular aspects of the visual scene. In addition to the implications of STDP for feedforward visual processing, we will also raise hypotheses about its role in establishing connectivities between different areas of the visual cortex, especially where feedback connections are concerned.

2. Spike times codes

What does a neuron tell us when emitting action potentials? Usually, we say that its firing rate reflects an analogue value such as stimulus intensity. Following the trail blazed by Barlow, we can even take a step further in reflexion and state that a single unit not just codes for the presence of a certain feature but for the *probability* of its presence—or degree of certainty [4]. The answer thus looks very simple: the higher the firing rate, the higher the stimulus intensity.

But at the same time, the rate-only solution tends to overlook an even simpler fact: a neuron, even according to the simplistic integrate-and-fire model, reaches its threshold fastest when the incoming stimulus matches its selectivity. This is the basic idea at the heart of asynchronous coding schemes that propose that, effectively, neurons can be thought of as acting as analog-to-delay converters [59,61] (Fig. 1a). In other terms, the activation, or stimulus
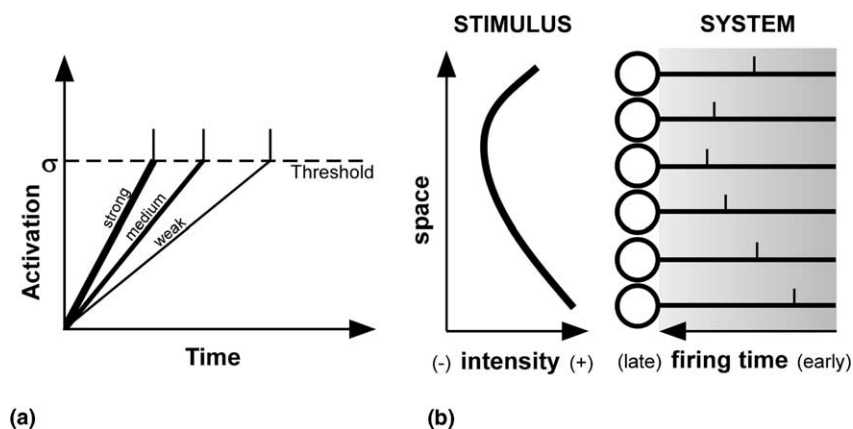

Fig. 1. *Latency encoding of stimulus.* (a) For a single neuron, the weaker the stimulus, the longer the time-to-first-spike. (b) When presented to a population of neurons, the stimulus evokes a spike wave, within which asynchrony encodes the information.

intensity, would determine the firing latency, with respect to a reference signal which can be either external—the stimulus onset [14,42,29]; or internal—possibly determined by local-field potentials (LFP) at a particular frequency [27]. For example of the latter case, the location of a rat within its environment can be retrieved from the precise phase relationship of action potentials fired by "place cells" with respect to the hippocampal EEG theta rhythm [40,37].

Interestingly, retinal ganglion cells can be seen as analog-to-delay units [64] and display the favourable feature of emitting precisely timed spikes [47,8]. Considering that the retina tiles the visual scene with an array of such cells, it follows that what is visually presented to the organism can be translated into a wave of first spikes whose times are more or less precisely repeated from one presentation of the same stimulus to the other. Said otherwise, the order of firing within a population of cells, induced by the relative spike times, can be used to encode information with a single spike per neuron (Fig. 1b). VanRullen and Thorpe provided theoretical evidence supporting this view by demonstrating that, when applied to retinal ganglion cells, such a coding scheme allowed identification of many stimuli when as few as 1.0% of the cells have fired [67]. The main demonstration relied on precise latencies but in real conditions, information transmission using the fine spatio-temporal structure of spike waves could be impaired by intrinsic sources of perturbations, from vesicle release to spike generation and propagation, via post-synaptic receptor activation and excitatory post-synaptic potential (EPSP) production. When considering spikes as information-carriers, the consequence is their times being jittered, meaning that in vivo, stimulus-elicited action potentials tend to appear *around* the same time, with a precision typically in the order of one to several milliseconds [8,34]. However, VanRullen and Thorpe showed that stimulus identification was still sharp when jitter was simulated [67]. And, as we will see later in Section 3, jitter and spontaneous activity seem to have little to no effect on temporal learning [24].

At this point, it is worth emphasizing that the sparse level (∼1%) of diagnostic activation in the input layer corresponds to the very first spikes, meaning that after stimulus onset the information needed to achieve identification can very rapidly reach the cortex for subsequent categorisation in the ventral pathway. If that small amount of information is sufficient, and given the high metabolic cost of spike generation [31,32], it could be useful to restrict firing in the input layer to economise energy. How might this be achieved? Here again, temporal coding can provide a simple solution in the form of a feedback, disynaptic inhibitory mechanism acting as a *k*-Winner-Take-All (*k*-WTA) operation on the input: depending on the threshold for its activation, inhibition could shut down the entire population once the first *k* cells have fired in the input layer. The kind of time-locked recurrent circuit proposed here exists in CA1 of the hippocampus, that produces inhibition in response to the onset of a series of spikes [46]. Specifically,

Pouille and Scanziani showed the ability of such circuits to differentially process the first action potential in a spike train from its rate, at least as far as feedback inhibition is concerned.

But now, if information is encoded in the rank order, how could an efferent neuron decode it? A simple feedforward shunting inhibition mechanism could in principle be used to produce a selective response (Fig. 2). Progressively desensitizing the output neuron with each incoming spike allows the first inputs to be fully effective while later ones produce less and less activation. The total amount of activation will thus depend on how well the order of firing within the input pattern matches the corresponding synaptic weights, maximal activation happening when they superimpose precisely. Hence, if the synaptic weights are set according to the stimulus to be identified (see Section 3), the corresponding efferent neuron would only respond optimally in the presence of that particular stimulus [61]. This idea has recently been tested within a biological model of the lateral geniculate nucleus; it successfully demonstrated that asynchronous excitatory spike waves and disynaptic feedforward inhibition generate orientation selectivity in the primary visual cortex while providing contrast invariance through automatic gain control [15].

How plausible is the suggestion that shunting inhibition could be used to implement rank-order decoding? To answer this, let us take a look at two requirements of shunting inhibition in the context of rank-order coding. The first requirement is that the inhibition needs to be very rapid if it is to be able to follow the temporal details of the wave of incoming activity. The second is that the inhibitory mechanisms should not be selective—they should effectively respond whenever an input fires, irrespective of the information it bears. Recent experimental work on fast
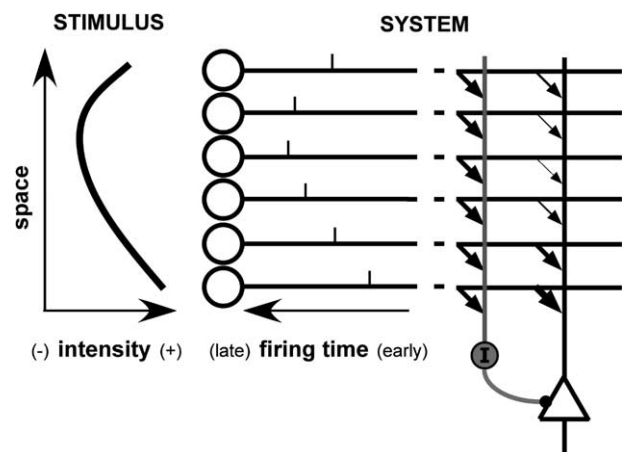

Fig. 2. *Shunting inhibition for rank-order decoding.* A pyramidal neuron (triangle cell) receives excitatory input from the afferent units through synapses with variable weights. Unit I receives an excitatory input from the same population but through synapses with equal weights. This inhibitory cell generates shunting inhibition that progressively desensitizes the pyramidal cell as more and more of the inputs fire. First inputs are effective while later ones produce less and less activation.

spiking interneurons in the somatosensory cortex supports the existence of such circuits in the brain. Indeed, Swadlow and Gusev characterized the fast spiking inhibitory interneurons as having very small somas that allow for a very rapid reaction to the activation of their inputs. Furthermore, they do receive strong inputs and show essentially no stimulus selectivity [58]. Moreover, these cells are known to be electrically coupled [23]: these connections will tend to make the population fire together, thus considerably reducing any selectivity its cells could possess [19]. Similar reports of inhibitory neurons with poor or even absent selectivity have also been seen in the visual cortex ([26]; Nowak, personal communication). Finally, evidence in the visual system shows that shunting inhibition targets neurons of cat primary visual cortex as evidenced by the threefold increase in soma conductance recorded very rapidly after presentation of a visual stimulus [11].

The system we propose for fast information processing between two neuronal "layers" thus consists of three circuits: one for information propagation based on a FF excitatory drive, a FB inhibitory circuit for controlling the amount of activity in the first layer and one FF inhibitory circuit for the decoding of information in the second layer through shunting inhibition. Note that the URVC experiments at the core of this theoretical proposal use brief transient stimuli whereas our everyday visual experience is one of continuous exposition to changing images. From there, one generally assumes that perception involves a continuous flow of information (but see [70] for a challenging view on this topic). Translated at the neuronal level, spikes are continuously propagated: if neurons are shunted after receiving their first few spikes, they would miss information carried by later ones, when something changes in the visual field for example. Within our scheme, one would need some reset mechanism to have the output neurons recover from the (decaying) shunting effect. This is precisely what the FB inhibition circuit could provide: it could block the FF excitatory drive to the second layer but it might also affect the interneuron population that consequently would stop inhibiting the second layer. Neurons there would no longer be subject to FF inhibition and thus recover their decoding power to read more incoming information.

A temporal neural code based on latency rank-order coding thus provides a theoretical explanation of the speed with which the brain is able to categorize complex visual scenes. All along the ventral pathway, a wave of spikes propagates in a feedforward manner, initiated at the level of the retinal ganglion cells, going through areas V1, V2 and V4 to reach IT where complex shapes are selectively encoded. At each stage of this hierarchy, information is contained within the asynchronous firing pattern, with the first spikes corresponding to the most salient features for the next stage [68,69]. The question remains as to how these "features" are shaped. In the next section, we will explore the possibility that neurons in the visual system can be made to be selective to particular aspects of the visual scene using the temporal code described so far, coupled with spike-timing-dependent plasticity (STDP).

3. Temporal learning with STDP

How can neurons be made selective? This question directly refers to the neural correlates of learning and synaptic plasticity (see a review in [1]). And since we are concerned with the way neurons encode information in their activity, it specifically relates to the field of activity-driven synaptic modifications (also termed "Hebbian" learning in reference to Donald O. Hebb's original postulate [25]).

The rules for long-term potentiation (LTP) and depression (LTD) were initially described in terms of rate-based approaches (see [10,7] for reviews). Thus, after having been stimulated at high firing rates, an excitatory synapse can show potentiation (LTP). Interestingly, recent experimental studies have brought these rules into the temporal domain as it has been found that LTP and LTD could be induced depending on the temporal relationship of a pre-synaptic spike relative to a post-synaptic one: when the former precedes the latter, the synapse is reinforced; it is depressed when the post-synaptic spike is emitted before a spike hits the synapse [35]. STDP became even more "time-friendly" when it was later discovered that the amount of modification depended on the precise delay between the two spikes: maximal when the pre- and post-synaptic spikes are close together, the effects gradually decrease and disappear with intervals in excess of a few tens of milliseconds [9,72,17]. This discovery not only served to revive the debate on timing in neural codes in a simple yet elegant way but more importantly has the important characteristic of allowing weight modifications to be regulated locally. This can be seen in the role of back-propagating action potentials in signalling the occurrence of a post-synaptic potential at the synaptic locus [33].

By providing an experimental basis for synaptic learning rules based on spikes rather than firing rates, spike-timing-dependent synaptic plasticity has become the subject of numerous theoretical investigations. While it may be involved in the formation, as well as the refinement, of cortical maps [57], temporal pattern recognition or coincidence detection (see [30] for a review), its most interesting feature resides in its competitive and stabilizing nature: it strengthens correlated inputs while being insensitive to firing rates or the degree of variability of a given synaptic input [56,21].

However, these modelling studies have mostly been conducted using firing rates as the vector for the neural code whereas STDP is, intrinsically, sensitive to precise spike times. What of STDP effects with temporal codes? We will first consider that reproducibility in stimulus-locked neural responses implies that neurons could potentially be repeatedly exposed to the same spatio-temporal input patterns (Fig. 3a). This might occur as the result of multiple exposures with the same stimulus at different times in life, or
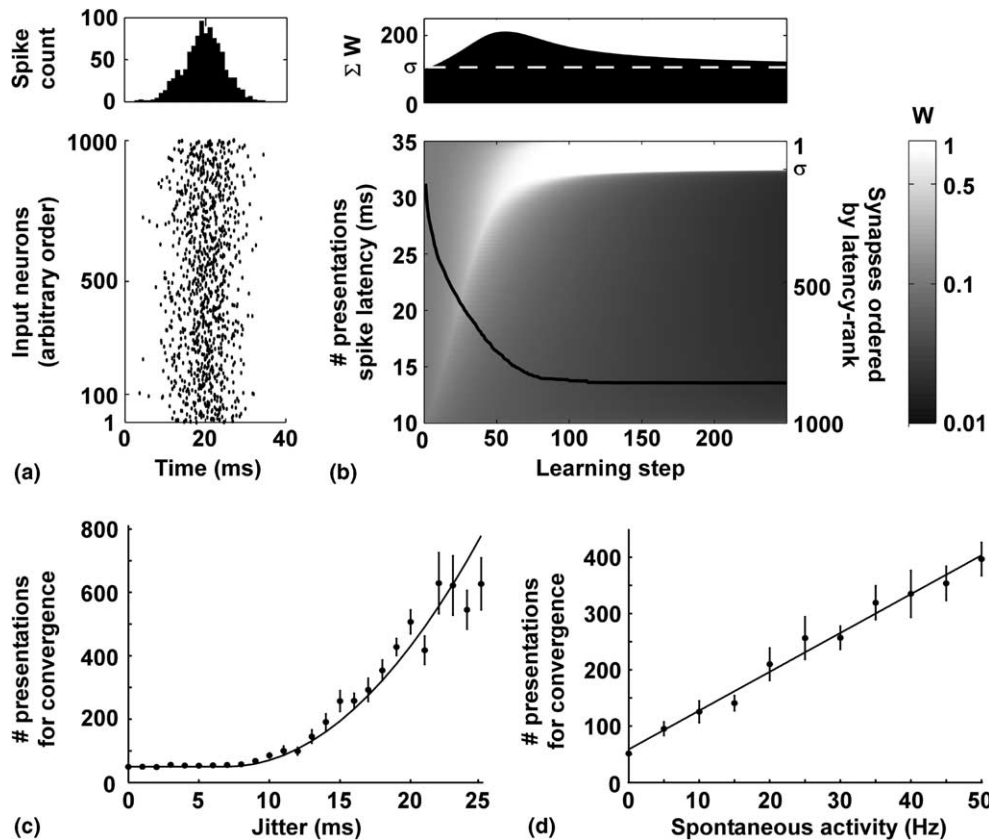
Fig. 3. *Temporal learning for an asynchronous spike wave.* With an identically repeated input spike wave, the neuron learns to react faster to its target. Synaptic weights converge onto the earliest firing afferents. (a) *Typical incoming activity.* (Bottom) Raster plot of an asynchronous spike wave (mean 20 ms, 5 ms standard deviation) being presented as such from one step to the other; when the post-synaptic neuron spikes, the STDP learning rule is applied and its potential reset to 0 before going to the next presentation. Prior to presentation, the pre-synaptic neurons do not fire any spikes. (Top) The corresponding post-stimulus time histogram (PSTH) showing the induced Gaussian form corresponding to the reproduced spike wave. (b) *Dynamics of repeated STDP.* (Top) Sum of all synaptic weights at each presentation (the dashed line represents the output neuron threshold). The sum of the synaptic weights stored in the afferents stabilizes at threshold value. (Bottom) The horizontal axis corresponds to the number of presentations (i.e. the learning step). The black line refers to the left axis and shows the reduction of post-synaptic latency during the course of learning. The background image refers to the right axis where each synapse weight is mapped by a gray-level index (see the corresponding bar on the right). Synapses are ordered by spiking latency of the corresponding neuron within the reproducible input pattern. During learning, earliest synapses become fully potentiated and later ones are weakened. (c) *Effect of jitter.* Jitter is generated by a Gaussian distribution. Increasing its standard deviation does not affect convergence until about 10 ms. From there, it slows the system roughly quadratically. (d) *Effect of spontaneous firing rate.* Increasing background activity slows convergence approximatively linearly.

alternatively, a single stimulus exposure could also potentially result in a sequence of similar processing waves through the rhythmic activity of cortical oscillations [27]. In either of these conditions, theoretical simulations have shown that two concomitant phenomena will take place: a reduction in the latency of the post-synaptic response along with the selective reinforcement of the synapses receiving the earliest repeated inputs [56,22,24] (Fig. 3b). The dynamical consequence of the asymmetrical, retrograde form of STDP (retrograde because potentiation affects what happened *before* the post-synaptic spike, thus favoring a ''back-in-time'' motion), this trend is simplistically explained as follows: for one given input pattern presentation, the input spikes elicit a post-synaptic response, triggering the STDP rule: synapses carrying input spikes just preceding the post-synaptic one are potentiated while later ones are weakened. The next time this input pattern

is re-presented, firing threshold will be reached sooner which implies a slight decrease of the post-synaptic spike latency. Consequently the learning process, while depressing some synapses it had previously potentiated, will now reinforce different synapses carrying even earlier spikes than the preceding time. By iteration, it follows that upon repeated presentation of the same input spike pattern, the post-synaptic spike latency will tend to stabilize at a minimal value while the first synapses become fully potentiated and later ones fully depressed.

Remarkably, stabilization occurs when enough of the earliest inputs are potentiated so as to evoke a post-synaptic spike with a single spike per input. Hence the output neuron threshold (σ) determines the number of excitatory synapses actually selected (Fig. 3b, top), which can be evaluated in real cortical neurons to 10–40—out of as many as 10,000—according to electrophysiological measures [53].

This trend still arises in biologically realistic conditions as neither jitter (Fig. 3c) nor spontaneous activity (Fig. 3d) can prevent it from occurring [24]. Simply put, additional jitter makes a given spike more likely to fall out of the LTP/LTD window of STDP, thus slowing down the convergence.

Spontaneous activity means that additional spikes occur randomly in either the LTP part of the STDP window, or in its LTD part, in proportion with the respective time-span of these modification windows (respectively 20 ms for LTP and 22 ms for LTD—see [24] for details). It follows that in the course of learning, the noise-induced total potentiation affecting a given synapse is slightly inferior to the noise-induced total depression because, statistically, more spikes will have triggered LTD than LTP. Thus with more spontaneous activity, there is a imbalance in favour of the depression process. But here again, spontaneous activity only slows down the convergence process.

Conclusively, as long as reproducible inputs are preserved from one presentation to the other, STDP is able to track them and concentrate high weights on the earliest firing ones. This naturally leads a neuron to respond rapidly to precisely repeating patterns on the basis of the first few afferent spikes, even in a "noisy" system. Additionally, if we are ready to accept that selectivity to a stimulus can be expressed in the temporal domain if a neuron responds earlier to a given stimulus than to any other, then learning through STDP does make the neuron more selective [24]. This theoretical result thus supports the idea that information can potentially be encoded in a single spike wave. But how would the phenomenon apply in the case of the visual system? What can be learnt in terms of visual representations in the ventral pathway when mimicking experience-dependent synaptic shaping?

4. STDP in vision: feedforward processing

To make the situation more realistic, imagine how an array of retinal ganglion cells would respond to a flashed bar (Fig. 4a). Using the fact that the most strongly activated cells reach threshold first, presenting an oriented bar would evoke a wave of spikes whose first action potentials corresponds to the *ordered* borders of the stimulus. Now suppose that a neuron in visual cortex receives initially weak inputs from all the retinal afferents within a particular receptive field area (Fig. 4b—step #0), and that in response to the flash, it finally goes over threshold and fires a spike. According to the trend described in Section 3, repeated presentation of the same stimulus would progressively refine its set of afferents until the high weights are all on the earliest firing inputs. Followingly, we can infer that orientation selectivity in V1 could be explained by an ordered alignment of LGN inputs [28] (Fig. 4b—step #100) using latency rank-order coding as demonstrated in [15].

As proposed in Section 2, rapid visual processing could be achieved by using an asynchronous wave of spikes, initi-
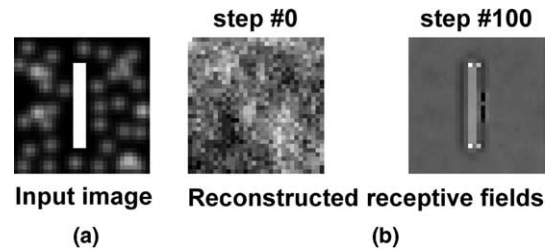
Fig. 4. *Shaping of orientation selectivity.* A model of the retina is repeatedly presented with an oriented bar (a). The evoked activity is propagated towards a neuron in visual cortex via the LGN. Initially its inputs are very weak (b—step #0) but the spike-timing-dependent plasticity process refines them until selectively reinforcing the first ones (b—step #100). As these correspond to the ordered borders of the stimulus, the selectivity of the visual neuron is thus composed of an ordered alignment of LGN inputs. The receptive field is linearly reconstructed based on the synaptic weights and the selectivity of the afferent neuron (ON- and OFF-center receptive fields).

ated in the retina and regenerated at each stage of the hierarchy to be propagated throughout the ventral pathway [68,69]. Hence, neurons could be exposed to reproducible spike waves at each level of the hierarchy. Interestingly, experimental observations showed that V1 cells display a degree of selectivity in their latencies, respective to stimulus onset [14]. Then, one can extend the present architecture to take V1 cells responses and selectivities into account, generating spikes at times that depend on the local orientation, similarly to RGCs responses in the former model.

Presenting such a population with the same image repeatedly exposes a neuron in a later stage of the hierarchy (say, V2 or V4) to a characteristic temporally organized wave of spikes at the output from V1. As expected, the post-synaptic potential progressively becomes steeper, and stabilizes with a minimal latency (Fig. 5, middle column) while at the same time the earliest inputs are finally selected (Fig. 5, right column). But, unlike the results in the earlier theoretical section (Section 3), here we can get a glimpse of the representations used through the linear reconstruction of the receptive field, based on the set of synaptic weights and the selectivity of the corresponding afferent neurons (Fig. 5, left column): in the course of learning, the reconstructed "optimal stimulus" for the neuron does not seem to make much sense. But as the neuron "matures", a structured representation emerges, that is built upon the earliest afferents of the input spike wave. As a consequence, with as few as a hundred presentations, the neuron can "learn" to recognize a particular image, in this case, a well known image of Einstein.

Of course, normally there is not just one neuron receiving the output of V1. The next question thus addresses the effects of STDP in population learning where the same set of inputs projects to a number of cells in the next layer. We will then extend this simple situation to one in which several different input patterns are randomly presented to a population of as many potentially learning cells. Suppose now that each time one of the cells fires, it inhibits its neighbourhood—the other cells in the population—through a
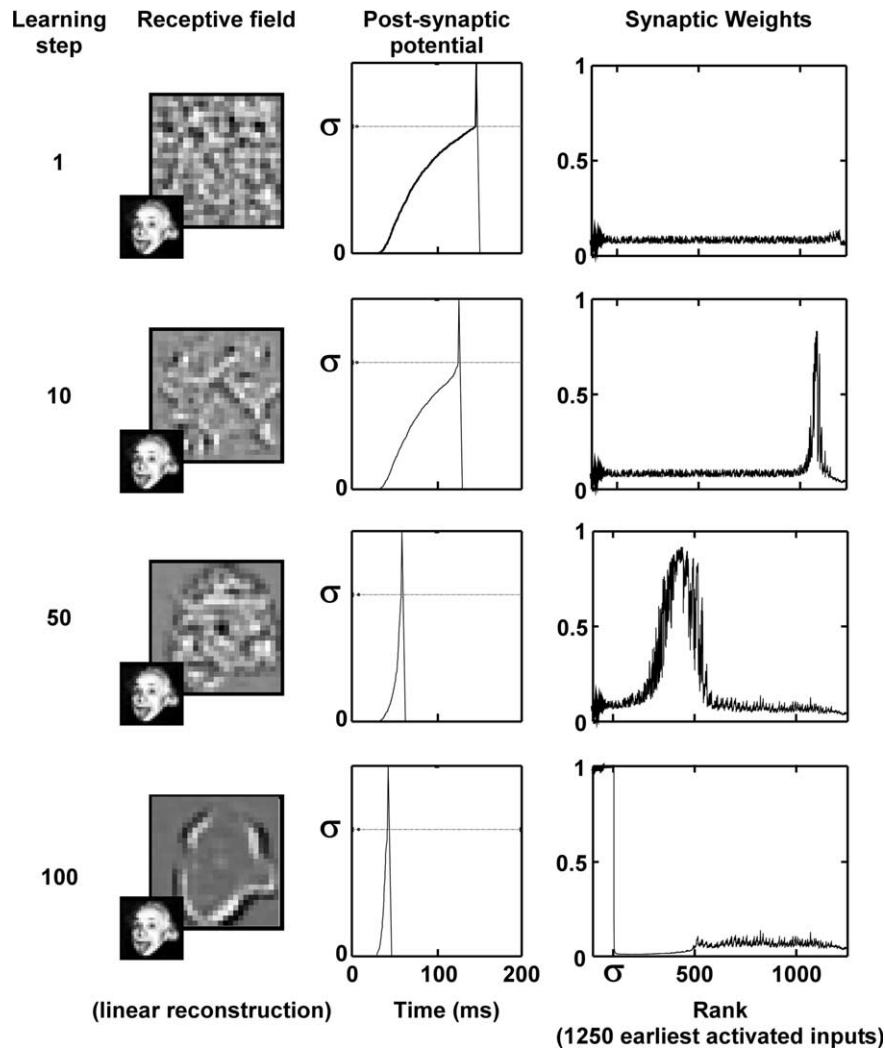
Fig. 5. *Einstein*: *STDP learning of a V1-filtered face.* A population of V1-like cells encodes an orientation for each pixel in the image presented to the network (here, Einstein's face); each cell acts as an analog-to-delay converter where the latency of its first spike depends on the strength of the orientation in its receptive field. Time taken to achieve recognition of the stimulus decreases (middle column) while a structured representation emerges and stabilizes (left column) that is built upon the earliest afferents of the input spike wave (right column). Note that the information concerning the evolution of the synaptic weights in the course of learning is represented twice on this figure. First in the distribution of synaptic weights on the right. It is also present in the receptive field on the left, that is linearly reconstructed based on the synaptic weights and the selectivity of the corresponding afferent neurons (here, orientation selective filters).

feedback (FB) mechanism. As a result of this competition, only one cell learns the V1-filtered spiking activity induced by the stimulus at each step. The results are remarkable: while the neurons obviously do not know which image was presented, each output neuron nevertheless retains one stimulus and one only, by blindly separating the sources (Fig. 6a). When considering the responses at the population level, one can also note that learning makes the responses selective by implementing a sparse code [41] (Fig. 6b). Indeed, the rapid responses displayed by the neuron when presented with the stimulus used for learning [24] translates within the present architecture into a situation where the neuron that fires first inhibits its neighbors, and thus becomes the only one to fire. The local, autonomous STDP learning rule can then lead to rapid, highly selective visual responses reminiscent of the notorious

"grandmother cells" [4]. The population would thus be able to encode, in its first spike, the very limited world it was exposed to, but not other features except for very similar ones (different views of Mona Lisa for example).

Within these feed-forward visual architectures, neural selectivities at a given level can be shaped using STDP from the responses of neurons at the preceding stage (Fig. 4), according to the statistics of the external world, just like letters (ON- and OFF-center inputs for example) compose words (orientations). In that sense, the fact that STDP is a purely local Hebbian mechanism allows the whole procedure to be recurrently applied, ad libitum, provided "forms" at the initial level are precisely known and the architecture respected. Hence, arbitrarily complex selectivities, for visual representation and recognition, can be autonomously generated in later stages, through exposition
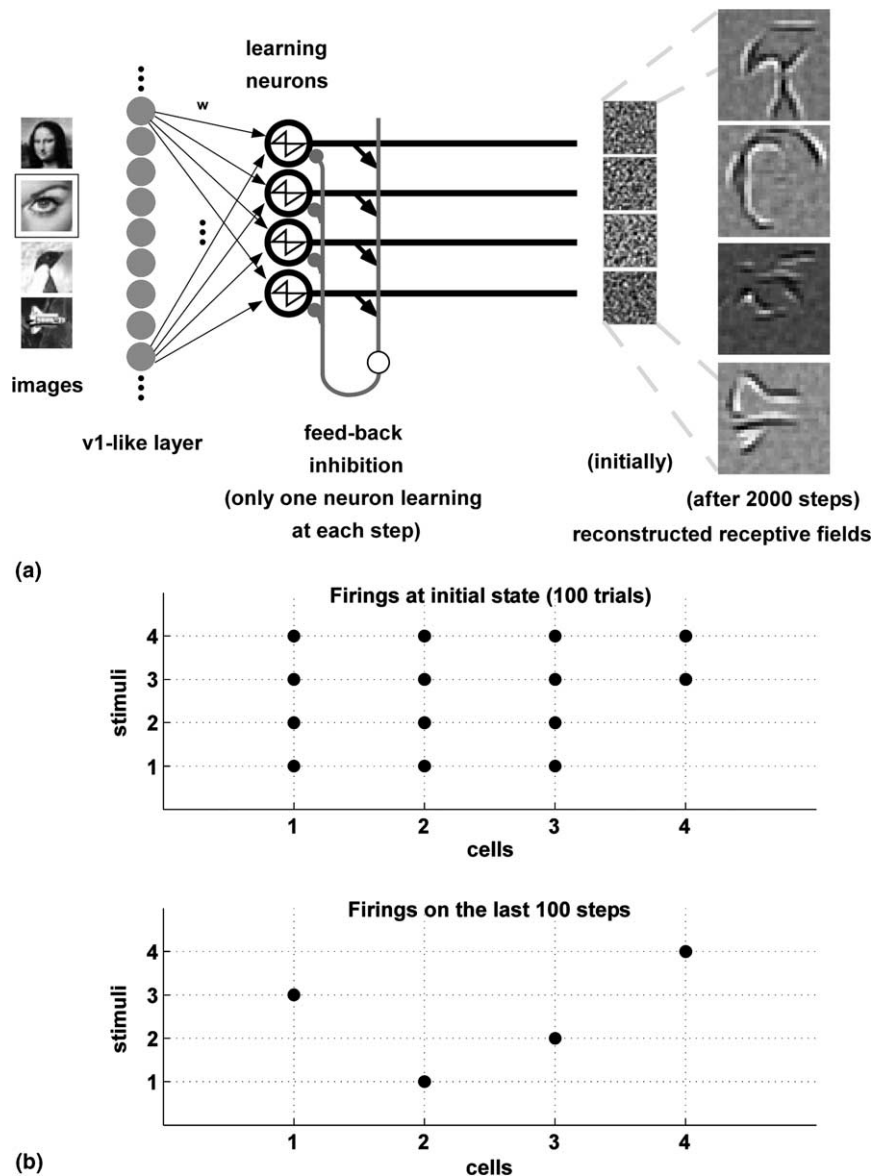
Fig. 6. *Population learning*: *Emergence of selective responses.* (a) *Architecture of the network and receptive fields.* At each step, one of the four images on the left is presented to the network. The V1-like layer generates one-spike-per-cell just like in Fig. 4. Spike times are then jittered (5 ms Gaussian delay) and 5 Hz Poisson-inspired spontaneous activity is added to the spike pattern at each presentation. The changing incoming activity is propagated towards the next layer where four neurons integrate it. The first to fire inhibits its neighbours and triggers the STDP learning rule. All four neurons start with initially random weights. After 2000 random presentations, each neuron has learned one stimulus and one only, in the same manner as in Fig. 4. (b) *Population response.* A thick dot indicates if the neuron ever fired first when presented with the corresponding stimulus. (Top) Initially, each neuron is likely to respond to any of the stimuli when tested for first response on 100 trials without STDP learning (plasticity was shut off for the test). (Bottom) At the end of training, any input image is clearly identifiable based on which neuron fires (learning steps #1901 to #2000).

to patches of natural scenes for example. Our future explorations of "supervised" and "unsupervised" learning architectures based on this idea but on a larger scale should hopefully lend empirical support to this theoretical claim.

5. STDP in vision: information flow and feedback connectivity

So far, the architectures that we have been considering are very simple ones, with essentially only feed-forward connections—the exception being the feed-back inhibition used to perform the k-WTA operation. But real cortical architecture is more diversified in terms of connectivity as connections between different areas may be classified into three different types according to their preferential termination: FF connections typically terminate in layer 4, lateral ones in a columnar way and FB ones prefer a multilaminar pattern excluding layer 4. As a consequence, cortical areas can be organized in a hierarchy according to the pattern of reciprocal connections between them [18]. However, it has already been suggested that the rank of a given area in the anatomical hierarchy does not necessarily correspond to that of the order of activation. For instance, area MT

may be placed at the top of the anatomical hierarchy while yielding early responses that would qualify it for the earliest stages of visual processing; area MT, but also subsequent area MST and frontal eye field (FEF), would thus be part of the "fast brain", where median latencies typically occur between 40 and 80 ms; as opposed to the "slow brain" (100–150 ms responses) that groups most areas of the temporal lobe and some areas of the frontal cortex located rostral to the frontal eye field [39].

According to this hypothesis, latency can be seen as a dynamical principle of hierarchical organisation in the visual system, one based on the visual information flow (as opposed to the classical, static connectivity defined by the anatomy). We have shown here that STDP tunes neurons to the earliest spikes they receive: latencies effectively dictate which afferents are selected. The similarities are such that one can wonder what would happen if the STDP learning rule is applied in the case of architectures with a richer set of connections. Activity-dependent mechanisms cannot be invoked to influence synaptic targeting and the consequent structural changes (see [52] for a review). In the case of V4 for example, Batardiere and colleagues recently showed that the hierarchical ranks of its connected areas are clearly established in the immature cortex, even if some remodelling may take place between fetal life and adulthood (not for FEF though [5]). But, on the basis of existing connections, and depending on the latencies observed in different areas of the cortex, STDP could influence where the result of a computation in a given area is mainly used, as well as shaping the selectivities of the neurons at this location.

For instance, the magnocellular (M) pathway could act as an ignition device for the parvocellular (P) pathway in V1 as suggested by Ullman [63]. Bullier proposed that V1 and V2 act as 'active blackboards', recipients of the computation performed at higher levels [13]. A first, fastest wave of visual information, channelled by the magnocellular (M) pathways, reaches MT via V1 in a FF manner [36]; from there, the M input goes "backwards" to V1 in time to meet the slower by 20 ms, parvocellular (P) input [38]. It follows that the two asynchronous flows meet in the striate cortex; neurons would receive both inputs from the feedback magnocellular flow of MT and the feedforward parvocellular flow from the LGN in a temporally limited window; STDP could thus shape their selectivity based on correlated activity among both FF and FB information flows—retaining relevant characteristics of both a crude, global analysis (M) and a more precise local analysis (P).

6. Conclusion: sensory information processing

When we aim at understanding how the brain works, we are faced with the puzzling challenge of penetrating its essential complexity. We can however single out its most significant performances, and provided we can explain them easily, unveil a piece of the mind game.

In that sense, previous work on ultra-rapid visual processing has imposed major constraints that can help improve our understanding of the neural code. Given the physiological properties of nerve cells, the known anatomy of the visual system and the primate performances in rapid visual categorisation tasks, it is hypothesized that fast neural information processing can be achieved using a single spike per neuron, where the most activated neurons encode the most important information.

Separately, the spiking activity of a neuron is known to have an influence on the strength of its synapses. Given that neural activity emerges from the interaction of the organism with its environment, it follows that sensory experience has an important role in refining synaptic connections. The question of learning, of how one's experience affects one's way to be in the world, can, arguably, be addressed at the neuronal level.

As we showed in this paper, we have used the idea of a temporally structured neural information in the context of learning through STDP and the results display a striking coherence. On one side—rank-order coding—neurons might emit the most salient information first. On the other side—STDP—neurons seem to learn to preferentially listen to what arrives first. Although physically separated and having no means of communicating other than through spikes, neural populations seem to agree on what makes sense to efficiently achieve tasks where time is of the essence.

Specifically, neural representations could be built on a very small subset of afferents, characterized by their correlated repetition and most of all their early arrival, effectively implementing sparse codes in the process [41]. Said otherwise, experience at the neuronal level selects inputs characterized by their ability to transmit relevant information in the fastest way possible. Precisely, expertise, or maturation, would correlate with the development of quickness as displayed by the latency reduction exposed here and in ERPs of visual categorisation during childhood [6]. One usually says that practice makes perfect; in fact, repetition may very well make swift.

Note that the temporal approach presented here does in no way discard the evident role of firing rates. In fact, the brain may very well use these two different strategies that need not be opposed but rather considered in a complementary way; as experimental evidence shows us, neurons can make the distinction between these [47]. Rates-based models nonetheless get most of the neuroscientific attention, mostly because temporal coding still lacks a clean-cut experimental demonstration; and at the same time, temporality presents us with a simple suggestion: sensory systems tend to be as reliable and economic as possible.

References

[1] L.F. Abbott, S.B. Nelson, Synaptic plasticity: taming the beast, Nat. Neurosci. 3 (Suppl.) (2000) 1178–1183.

[2] M. Abeles, Corticonics: Neural Circuits of the Cerebral Cortex, Cambridge University Press, Cambridge, 1991.

[3] E. Adrian, The Basis of Sensation, Londres, Christophers, 1928.

[4] H.B. Barlow, Single units and sensation: a neuron doctrine for perceptual psychology? Perception 1 (1972) 371–394.

[5] A. Batardiere, P. Barone, K. Knoblauch, P. Giroud, M. Berland, A.M. Dumas, H. Kennedy, Early specification of the hierarchical organization of visual cortical areas in the macaque monkey, Cereb. Cortex 12 (2002) 453–465.

[6] M. Batty, M.J. Taylor, Visual categorisation during childhood: an ERP study, Psychophysiology 39 (2002) 482–490.

[7] M.F. Bear, Bidirectional synaptic plasticity: from theory to reality, Philos. Trans. R. Soc. Lond. B Biol. Sci. 358 (2003) 649–655.

[8] M.J. Berry, D.K. Warland, M. Meister, The structure and precision of retinal spike trains, Proc. Natl. Acad. Sci. USA 94 (1997) 5411–5416.

[9] G.Q. Bi, M.M. Poo, Synaptic modifications in cultured hippocampal neurons: dependence on spike timing, synaptic strength, and post-synaptic cell type, J. Neurosci. 18 (1998) 10464–10472.

[10] G.Q. Bi, M.M. Poo, Synaptic modification by correlated activity: Hebb's postulate revisited, Annu. Rev. Neurosci. 24 (2001) 139–166.

[11] L.J. Borg-Graham, C. Monier, Y. Fregnac, Visual input evokes transient and strong shunting inhibition in visual cortical neurons, Nature 393 (1998) 369–373.

[12] J. Bullier, M.E. McCourt, G.H. Henry, Physiological studies on the feedback connection to the striate cortex from cortical areas 18 and 19 of the cat, Exp. Brain Res. 70 (1988) 90–98.

[13] J. Bullier, Integrated model of visual processing, Brain Res. Rev. 36 (2001) 96–107.

[14] S. Celebrini, S. Thorpe, Y. Trotter, M. Imbert, Dynamics of orientation coding in area V1 of the awake primate, Vis. Neurosci. 10 (1993) 811–825.

[15] A. Delorme, Early cortical orientation selectivity: how fast inhibition decodes the order of spike latencies, J. Comput. Neurosci. 15 (2003) 357–365.

[16] M. Fabre-Thorpe, G. Richard, S.J. Thorpe, Rapid categorisation of natural images by rhesus monkeys, Neuroreport 9 (1998) 303–308.

[17] D.E. Feldman, Timing-based LTP and LTD at vertical inputs to layer II/III pyramidal cells in rat barrel cortex, Neuron 27 (2000) 45–56.

[18] D.J. Felleman, D.C. Van Essen, Distributed hierarchical processing in the primate cerebral cortex, Cereb. Cortex 1 (1991) 1–47.

[19] M. Galarreta, S. Hestrin, Electrical synapses between gaba-releasing interneurons, Nat. Rev. Neurosci. 2 (2001) 425–433.

[20] J. Gautrais, S. Thorpe, Rate coding versus temporal order coding: a theoretical approach, Biosystems 48 (1998) 57–65.

[21] W. Gerstner, W.M. Kistler, Mathematical formulations of Hebbian learning, Biol. Cybern. 87 (2002) 404–415.

[22] W. Gerstner, W.M. Kistler, Learning to be fast, in: Spiking Neuron Models, Cambridge University Press, Cambridge, 2002, pp. 421–432.

[23] J.R. Gibson, M. Beierlein, B.W. Connors, Two networks of electrically coupled inhibitory neurons in neocortex, Nature 402 (1999) 75–79.

[24] R. Guyonneau, R. VanRullen, S.J. Thorpe, Neurons tune to the earliest spikes through STDP, Neural Comput. 17 (2005) 859–879.

[25] D.O. Hebb, The Organisation of Behavior: A Neuropsychological Theory, NY, 1949.

[26] J.A. Hirsch, L.M. Martinez, C. Pillai, J.M. Alonso, Q. Wang, F.T. Sommer, Functionally distinct inhibitory neurons at the first stage of visual cortical processing, Nat. Neurosci. 6 (2003) 1300–1308.

[27] J.J. Hopfield, Pattern recognition computation using action potential timing for stimulus representation, Nature 376 (1995) 33–36.

[28] D.H. Hubel, T.N. Wiesel, Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, J. Physiol. 160 (1962) 106–154.

[29] R.S. Johansson, I. Birznieks, First spikes in ensembles of human tactile afferents code complex spatial fingertip events, Nat. Neurosci. 7 (2004) 170–177.

[30] A. Kepecs, M.C. van Rossum, S. Song, J. Tegner, Spike-timing-dependent plasticity: common themes and divergent vistas, Biol. Cybern. 87 (2002) 446–458.

[31] P. Lennie, The cost of cortical computation, Curr. Biol. 13 (2003) 493–497.

[32] W.B. Levy, R.A. Baxter, Energy efficient neural codes, Neural Comput. 8 (1996) 531–543.

[33] J.C. Magee, D. Johnston, A synaptically controlled, associative signal for Hebbian plasticity in hippocampal neurons, Science 275 (1997) 209–213.

[34] Z.F. Mainen, T.J. Sejnowski, Reliability of spike timing in neocortical neurons, Science 268 (1995) 1503–1506.

[35] H. Markram, J. Lubke, M. Frotscher, B. Sakmann, Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs, Science 275 (1997) 213–215.

[36] J.H. Maunsell, T.A. Nealey, D.D. DePriest, Magnocellular and parvocellular contributions to responses in the middle temporal visual area (MT) of the macaque monkey, J. Neurosci. 10 (1990) 3323–3334.

[37] M.R. Mehta, A.K. Lee, M.A. Wilson, Role of experience and oscillations in transforming a rate code into a temporal code, Nature 417 (2002) 741–746.

[38] L.G. Nowak, M.H. Munk, P. Girard, J. Bullier, Visual latencies in areas V1 and V2 of the macaque monkey, Vis. Neurosci. 12 (1995) 371–384.

[39] L.G. Nowak, J. Bullier, The timing of information transfer in the visual system, in: K.S. Rockland, J.H. Kaas, A. Peters (Eds.), Extrastriate Visual Cortex in Primates, Plenum Press, New York, 1997, pp. 205–241.

[40] J. O'Keefe, M.L. Recce, Phase relationship between hippocampal place units and the EEG theta rhythm, Hippocampus 3 (1993) 317–330.

[41] B.A. Olshausen, Sparse codes and spikes, in: R.P. Rao, B.A. Olshausen, M.S. Lewicki (Eds.), Probabilistic Models of the Brain: Perception and Neural Function, MIT Press, 2002, pp. 257–272.

[42] S. Panzeri, R.S. Petersen, S.R. Schultz, M. Lebedev, M.E. Diamond, The role of spike timing in the coding of stimulus location in rat somatosensory cortex, Neuron 29 (2001) 769–777.

[43] D.H. Perkel, G.L. Gerstein, G.P. Moore, Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains, Biophys. J. 7 (1967) 419–440.

[44] D.I. Perrett, E.T. Rolls, W. Caan, Visual neurones responsive to faces in the monkey temporal cortex, Exp. Brain Res. 47 (1982) 329–342.

[45] R.S. Petersen, S. Petersen, M.E. Petersen, Population coding of stimulus location in rat somatosensory cortex, Neuron 32 (2001) 503–514.

[46] F. Pouille, M. Scanziani, Routing of spike series by dynamic circuits in the hippocampus, Nature 429 (2004) 717–723.

[47] D.S. Reich, J.D. Victor, B.W. Victor, T. Victor, E. Kaplan, Response variability and timing precision of neuronal spike trains in vivo, J. Neurophysiol. 77 (1997) 2836–2841.

[48] P.R. Roelfsema, V.A. Lamme, H. Spekreijse, Object-based attention in the primary visual cortex of the macaque monkey, Nature 395 (1998) 376–381.

[49] E.T. Rolls, M.K. Sanghera, A. Roper-Hall, The latency of activation of neurones in the lateral hypothalamus and substantia innominata during feeding in the monkey, Brain Res. 164 (1979) 121–135.

[50] E.T. Rolls, D.I. Perrett, A.W. Caan, F.A. Wilson, Neuronal responses related to visual recognition, Brain 105 (1982) 611–646.

[51] C.D. Salzman, K.H. Britten, W.T. Newsome, Cortical microstimulation influences perceptual judgements of motion direction, Nature 346 (1990) 174–177.

[52] F. Sengpiel, P.C. Kind, The role of activity in development of the visual system, Curr. Biol. 12 (2002) R818–R826.

[53] M.N. Shadlen, W.T. Newsome, Noise, neural codes and cortical organization, Curr. Opin. Neurobiol. 4 (1994) 569–579.

[54] W. Singer, Time as coding space? Curr Opin. Neurobiol. 9 (1999) 189–194.

[55] W.R. Softky, C. Koch, The highly irregular firing of cortical cells is inconsistent with temporal integration of random EPSPs, J. Neurosci. 13 (1993) 334–350.

[56] S. Song, K.D. Miller, L.F. Abbott, Competitive Hebbian learning through spike-timing-dependent synaptic plasticity, Nat. Neurosci. 3 (2000) 919–926.

[57] S. Song, L.F. Abbott, Cortical development and remapping through spike timing-dependent plasticity, Neuron 32 (2001) 339–350.

[58] H.A. Swadlow, A.G. Gusev, Receptive-field construction in cortical inhibitory interneurons, Nat. Neurosci. 5 (2002) 403–404.

[59] S. Thorpe, M. Imbert, Biological constraints on connectionist modelling, in: R. Pfeifer, Z. Schreter, F. Fogelman-Soulié, L. Steels (Eds.), Connectionism in Perspective, Elsevier, Amsterdam, 1989, pp. 63–92.

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

[61] S. Thorpe, A. Delorme, R. Van Rullen, Spike-based strategies for rapid processing, Neural Networks 14 (2001) 715–725.

[62] S. Thorpe, Spike arrival times: a highly efficient coding scheme for neural networks, in: R. Eckmiller, G. Hartman, G. Hauske (Eds.), Parallel Processing in Neural Systems and Computers, Elsevier, Amsterdam, 1990, pp. 91–94.

[63] S. Ullman, Sequence seeking and counter streams: a computational model for bidirectional information flow in the visual cortex, Cereb. Cortex 5 (1995) 1–11.

[64] V.J. Uzzell, E.J. Chichilnisky, Precision of spike trains in primate retinal ganglion cells, J. Neurophysiol. 92 (2004) 780–789.

[65] R. VanRullen, S.J. Thorpe, Is it a bird? Is it a plane? Ultra-rapid visual categorisation of natural and artifactual objects, Perception 30 (2001) 655–668.

[66] R. VanRullen, S.J. Thorpe, The time course of visual processing: from early perception to decision-making, J. Cogn. Neurosci. 13 (2001) 454–461.

[67] R. VanRullen, S.J. Thorpe, Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex, Neural Comput. 13 (2001) 1255–1283.

[68] R. VanRullen, S.J. Thorpe, Surfing a spike wave down the ventral stream, Vision Res. 42 (2002) 2593–2615.

[69] R. VanRullen, Visual saliency and spike timing in the ventral visual pathway, J. Physiol. Paris 97 (2003) 365–377.

[70] R. VanRullen, C. Koch, Is perception discrete or continuous? Trends Cogn. Sci. 7 (2003) 207–213.

[71] R. VanRullen, R. Guyonneau, S.J. Thorpe, Spike times make sense, Trends Neurosci. 28 (2005) 10–14.

[72] L.I. Zhang, H.W. Tao, C.E. Holt, W.A. Harris, M. Poo, A critical window for cooperation and competition among developing retino-tectal synapses, Nature 395 (1998) 37–44.