



# The timing of vision – how neural processing links to different temporal dynamics

Timothée Masquelier<sup>1\*</sup>, Larissa Albantakis<sup>1</sup> and Gustavo Deco<sup>1,2</sup>

<sup>1</sup> Unit for Brain and Cognition, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Barcelona, Spain

<sup>2</sup> Institut Català de Recerca i Estudis Avançats, Universitat Pompeu Fabra, Barcelona, Spain

## Edited by:

Gabriel Kreiman, Harvard, USA

## Reviewed by:

Dietmar Heinke, University of Birmingham, UK

Eirini Mavritsaki, University of Birmingham, UK

Jedediah Miller Singer, Children's Hospital Boston, USA

## \*Correspondence:

Timothée Masquelier, Unit for Brain and Cognition, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Roc Boronat, 138, 08018 Barcelona, Spain.

e-mail: timothee.masquelier@alum.mit.edu

In this review, we describe our recent attempts to model the neural correlates of visual perception with biologically inspired networks of spiking neurons, emphasizing the dynamical aspects. Experimental evidence suggests distinct processing modes depending on the type of task the visual system is engaged in. A first mode, crucial for object recognition, deals with rapidly extracting the glimpse of a visual scene in the first 100 ms after its presentation. The promptness of this process points to mainly feedforward processing, which relies on latency coding, and may be shaped by spike timing-dependent plasticity (STDP). Our simulations confirm the plausibility and efficiency of such a scheme. A second mode can be engaged whenever one needs to perform finer perceptual discrimination through evidence accumulation on the order of 400 ms and above. Here, our simulations, together with theoretical considerations, show how predominantly local recurrent connections and long neural time-constants enable the integration and build-up of firing rates on this timescale. In particular, we review how a non-linear model with attractor states induced by strong recurrent connectivity provides straightforward explanations for several recent experimental observations. A third mode, involving additional top-down attentional signals, is relevant for more complex visual scene processing. In the model, as in the brain, these top-down attentional signals shape visual processing by biasing the competition between different pools of neurons. The winning pools may not only have a higher firing rate, but also more synchronous oscillatory activity. This fourth mode, oscillatory activity, leads to faster reaction times and enhanced information transfers in the model. This has indeed been observed experimentally. Moreover, oscillatory activity can format spike times and encode information in the spike phases with respect to the oscillatory cycle. This phenomenon is referred to as “phase-of-firing coding,” and experimental evidence for it is accumulating in the visual system. Simulations show that this code can again be efficiently decoded by STDP. Future work should focus on continuous natural vision, bio-inspired hardware vision systems, and novel experimental paradigms to further distinguish current modeling approaches.

**Keywords:** vision, attention, spiking neurons, neurodynamics, oscillations, STDP, neural coding, decision making

## INTRODUCTION

Our visual system is continuously challenged with various types of tasks, such as recognizing other people or objects, searching for a friend in a crowd, or determining the direction and speed of other cars while driving. To solve these different tasks, it has been hypothesized that visual information arriving in the primary visual area (V1) of the cortex is further processed via two specialized pathways: first, the ventral stream associated with forms and colors mostly involved in “what” tasks like object recognition, and, second, the dorsal stream which is mostly processing “where” information and motions. In general, however, the visual areas form a complex network, and the two main processing pathways are strongly interconnected. It is therefore hardly possible to derive anatomically the neural dynamics – that is, neural activity evolution over time – underlying visual processing. Nevertheless, different visual tasks such as recognition, search, and motion detection, not only vary with respect to “what” has to be processed, they also differ in “how fast” and “how accurate or detailed” the respective perception can (or must) be accomplished. The diverse temporal dynamics of

visual processing can motivate and help distinguishing feedforward, feedback, and top-down influences during specific visual tasks. In turn, computational models that account for reaction times, and the time course of neural activity during visual tasks, can offer mechanistic explanations for internal cortical processes such as activity accumulation, attentional effects, and information transfer.

Here, we review some of our recent models based on spiking neural networks (SNN) that describe neuronal correlates of several visual tasks at multiple timescales. These models are all biologically plausible, reproduce a broad range of experimental observations, and predict others. They help to understand the neural dynamics underlying visual processing, and in particular visual processing times. More specifically, feedforward models can account for the phenomenal speed of object recognition (see Fast Feedforward Processing, Latency Coding, and STDP). This type of rapid processing presumably depends on the ability of the visual system to learn how to recognize familiar visual primitives in an unsupervised manner. Spike timing-dependent plasticity (STDP) may play a key role here. Feedforward processing is usually sufficient

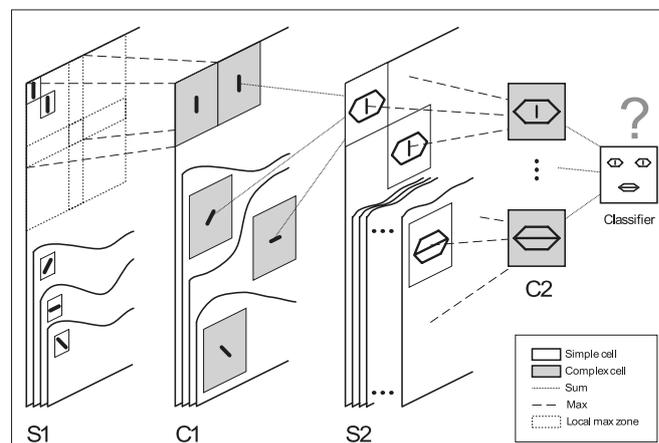
to extract the glimpse of a visual scene in 100–200 ms. Recurrent connectivity, however, allows accumulating evidence over longer timescales (several hundreds of milliseconds) whenever a finer visual discrimination is needed (see Slower Visual Decision Making). Such recurrent connections, in combination with bottom-up and top-down connections between brain areas, are also crucial to mediate attentional mechanisms through biased-competition (see Top-Down Attention), and can account for both “pop out” and serial modes in visual search. Attention not only up-modulates the firing rates of the neurons encoding the attended features, but also enhances their synchrony, enabling faster reaction times, dynamic information routing, and phase-of-firing coding (PoFC; see Oscillations Format Visual Processing). The phase patterns may be decoded thanks to STDP. Finally, we evoke important unsolved questions and future directions in Section “Unsolved Questions and Future Directions.”

### FAST FEEDFORWARD PROCESSING, LATENCY CODING, AND STDP

Vision can be extremely fast. There is now considerable behavioral and electrophysiological evidence showing that the primate visual system can achieve high-level object recognition in just 80–100 ms after stimulus onset (see Thorpe’s review in this Special Topic). This phenomenal speed imposes severe constraints on the underlying neural processes. Given that about 10 neuronal layers are involved in that sort of processing, the time window available for each neuron to perform its computation is only of about 10 ms. As the firing rates in the visual system are barely above 100 Hz, such a small window will consequently contain at most one spike (Thorpe and Imbert, 1989). A classical rate coding scheme, where individual neurons encode information in their mean firing rate, is thus ruled out. Instead, the information has to be encoded by which of the afferents were recruited, and possibly additionally by the relative recruiting times. This scheme is referred to as “rank order coding” (Thorpe and Gautrais, 1998). Note that if computation is restricted to one spike per neuron, the use of feedback loops is also ruled out. This implies that the first spike wave after stimulus onset probably does much more than conventionally assumed (VanRullen and Thorpe, 2002). Simulations have confirmed that it is indeed possible to perform fast and robust object recognition even in cluttered natural images, using only one spike per neuron, and feedforward connectivity (VanRullen et al., 1998; Delorme and Thorpe, 2001; Masquelier and Thorpe, 2007; Weidenbacher and Neumann, 2008).

In this section, we focus on how STDP may shape this kind of processing. STDP is a physiological mechanism of activity-driven synaptic regulation, where an excitatory synapse is reinforced when it receives a spike before a postsynaptic one is emitted (long-term potentiation, LTP). In the opposite case, its strength is weakened (long-term depression, LTD), when the postsynaptic spike precedes the presynaptic one. STDP has been observed both *in vivo* and *in vitro* in many species (from insects to mammals) and in many brain areas, including the visual cortex (see Caporale and Dan, 2008 for a review). Note that STDP is in agreement with Hebb’s (1949) postulate because it reinforces the connections with those presynaptic neurons that fired slightly before the postsynaptic neuron, which are the ones that “took part in firing it.”

What happens if such a rule is at work in a hierarchical neuronal network crossed by waves of spikes generated by visual stimuli? In Masquelier and Thorpe (2007), we assessed this question using a model inspired by HMAX (HMAX stands for “Hierarchical Model And X” – where X is a highly non-linear maXimum operation; Riesenhuber and Poggio, 1999; Serre et al., 2007). In an attempt to model the increasing complexity and invariance observed along the ventral pathway, we used a four-layer hierarchy (S1–C1–S2–C2) in which simple cells (S) gained their selectivity from a linear sum operation, while complex cells (C) gained invariance from a non-linear max pooling operation (see **Figure 1**). However, our network operates in the temporal domain: when presented with an image, the first layer’s S1 cells, emulating V1 simple cells, detect edges with four preferred orientations, and the more strongly a cell is activated, the earlier it fires a first spike. There is evidence for this so-called “intensity-to-latency conversion” in V1, where response latency decreases with stimulus contrast (Gawne et al., 1996; Albrecht et al., 2002), and also with the proximity between the stimulus orientation and the cell’s preferred orientation (Celebrini et al., 1993). These S1 spikes are then propagated asynchronously through the subsequent layers, where STDP takes place. Interestingly, within this time-to-first-spike coding framework, the maximum operation of complex cells simply consists of



**FIGURE 1 | Overview of the five layer feed forward spiking neural network used in Masquelier and Thorpe, 2007.** As in HMAX (Riesenhuber and Poggio, 1999; Serre et al., 2007), we alternate simple cells that gain selectivity through a sum operation, and complex cells that gain shift and scale invariance through a max operation (which in our framework simply consists of propagating the first received spike). Cells are organized in retinotopic maps until the S2 layer (inclusive). S1 cells detect edges. C1-maps subsample S1-maps by taking the maximum response over a square neighborhood. S2 cells are selective to intermediate complexity visual features, defined as a combination of oriented edges (here, we symbolically represented an eye detector and a mouth detector). There is one S1–C1–S2 pathway for each processing scale (not represented). Then C2 cells take the maximum response of S2 cells over all positions and scales, and are thus shift- and scale-invariant. Finally, a classification is done based on the C2 cells’ responses (here we symbolically represented a face/non-face classifier). In the brain, equivalents of S1 cells may be in V1, S2 cells in V1–V2, S2 cells in V4–PIT, C2 cells in AIT, and the final classifier in PFC. Here STDP shapes the C1-to-S2 connectivity. **Figure 2** shows an example of resulting selectivities after exposing the network to face images. Figure modified from Masquelier and Thorpe (2007).

propagating the first spike emitted by a given group of afferents (Rousselle et al., 2003). This can be achieved efficiently by one spiking neuron with low threshold that has synaptic connections from all neurons in the group [such “low threshold” relay cells are found in both the lateral geniculate nucleus (LGN), Rathbun et al., 2010 and the cortex, Swadlow and Gusev, 2002].

When we exposed the network to natural images, we observed that the neurons equipped with STDP gradually became selective to prototypical patterns that were both salient, and consistently present in the images. During the convergence process, synapses compete with each other (Song et al., 2000), and the winning synapses are those through which the earliest spikes arrive (on average; Song et al., 2000; Guyonneau et al., 2005). Interestingly, these earliest spikes, which correspond to the most salient regions of an image, are typically the most informative (VanRullen and Thorpe, 2001). Furthermore, the resulting effect of this “early input selection” is to make the postsynaptic neuron respond more quickly (Song et al., 2000; Gerstner and Kistler, 2002; Guyonneau et al., 2005).

**Figure 2** shows an example, in which we exposed the network to face images, and where the STDP neurons indeed became selective to face features. Note that we used unsegmented images, but the background was not learned since backgrounds are too different from one image to another for the STDP process to converge. It is important to note that up to this point, the learning was fully unsupervised. No external teacher’s signal or previous knowledge was given to the model. For example, in **Figure 2**, the system obviously had no idea it was going to see faces. The features were only learned due to statistical regularities in the training dataset. However, the output of the STDP neurons can be fed into a supervised classifier, leading to robust object categorization, even with few (~10) STDP-learned features (Masquelier and Thorpe, 2007).

It is well known that the visual system is plastic and can learn frequently encountered visual features or feature contingencies (Jiang and Chun, 2001). The model predicts that frequently occurring features are not only more likely to be learned, but will also be processed and recognized faster than unfamiliar ones (recall that postsynaptic latencies decrease with training). Consistent with this, psychophysical experiments show that familiar categories such as faces are processed faster (Crouzet et al., 2010), and that processing times can be speeded up with experience (Masquelier et al., 2008).

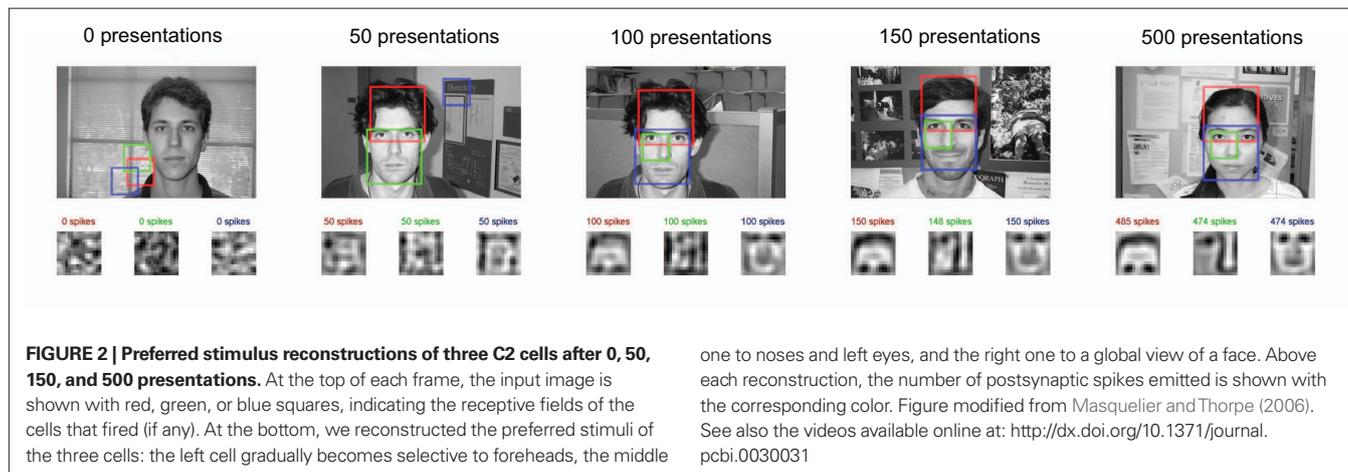
One important limitation in our study is that we used a noise-free deterministic model, while real neuronal responses are known to be variable. Future work will assess its robustness to neuronal noise. One can distinguish two kinds of response variability, or lack thereof: reliability and precision (Tiesinga et al., 2008). When a neuron fires approximately the same number of spikes on each trial, it is said to be reliable, whereas, when the spikes occur almost at the same time across trials, it is said to be precise. We have recently demonstrated that STDP-based pattern learning needs a precision of 10–20 ms, when in fact it is relatively insensitive to a lack of reliability, providing the input patterns involves a sufficient number of afferents (Gilson et al., unpublished observation). It would be interesting to quantify this number for the kind of rapid visual processing exposed in this section.

Finally, it is worth mentioning that STDP-based unsupervised learning is not restricted to natural image statistics. In fact, any arbitrary spike pattern that consistently repeats in the input can be learned (Masquelier et al., 2008, 2009a).

## SLOWER VISUAL DECISION MAKING

As we have seen in the previous section, feedforward processing of the first spike wave can be sufficient to rapidly extract the glimpse of a visual scene. Being so reactive is obviously advantageous in numerous emergency situations, such as obstacle/projectile avoidance or prey/predator/friend identification. But when reactivity is less crucial, integrating the visual information over time will generally improve perception, especially when visual evidence is noisy, moving, and ambiguous.

A psychophysical paradigm, designed to study the time course of slow perceptual decision making, is the random-dot motion (RDM) discrimination task (Roitman and Shadlen, 2002; Palmer et al., 2005; Churchland et al., 2008). Subjects performing this task have to decide on the net direction of motion in a patch of randomly moving dots. The quantity of the sensory evidence and thus, the task difficulty, is controlled by the amount of coherent motion. In the free response version, as soon as the subjects have gathered enough evidence to make a choice, they usually indicate their decision by a saccade to a target located in the corresponding direction. Reaction times in the RDM task are typically long, in the order of several 100 ms, with faster responses to more coherent

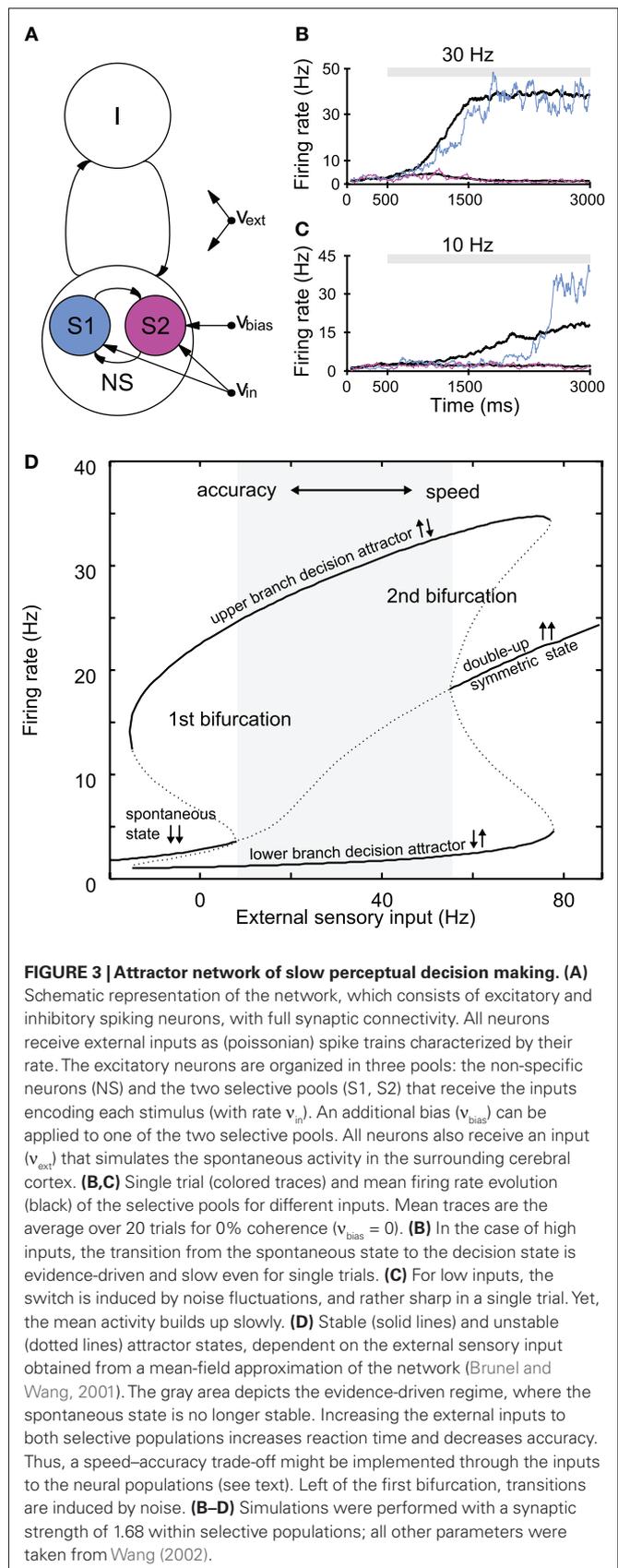


motions. A decision criterion is needed to determine how much evidence is “enough” to terminate the accumulation process, and to initiate the corresponding saccade. In theory, there are several possible decision criteria, such as relative or absolute thresholds (or “bounds”). Neurophysiological evidence from different cortical areas so far suggests a fixed firing rate threshold independent of reaction times (see below; Roitman and Shadlen, 2002; Schall et al., 2002; Churchland et al., 2008).

To identify possible neural correlates of this accumulation-to-bound concept, the psychophysical RDM task was combined with simultaneous recordings of decision-related activity from several brain areas along the dorsal visual stream [middle temporal (MT) and lateral intra-parietal (LIP) area, prefrontal cortex (PFC), and the superior colliculus (SC)]. All of them form part of the cognitive link between visual sensation and saccadic movement (reviewed in Schall, 2003; Smith and Ratcliff, 2004; Opris and Bruce, 2005). Particularly single neuron activity in area LIP of behaving monkeys has been found to increase gradually during motion viewing, dependent on task difficulty and according to choice behavior (Shadlen and Newsome, 2001; Roitman and Shadlen, 2002), while upstream of LIP in area MT neurons fire monotonically as a function of motion coherence (Britten et al., 1993). Area MT might thus provide the sensory evidence that is passed on to LIP for integration. Besides, the recorded LIP activity suggests a fixed firing rate threshold, as it reaches a uniform level, independent of response time or difficulty (about 40–80 ms prior to the saccade). Apart from these 40–80 ms for motor preparation, the rather long latency between signal onset and the onset of build-up activity in LIP (~190 ms) has to be subtracted from the measured reaction times to arrive at an estimate of the pure decision time, i.e., the actual time during which evidence is accumulated (Roitman and Shadlen, 2002; Churchland et al., 2008).

Decision-related activity build-up was also found downstream of LIP, in the dorsolateral PFC (Kim and Shadlen, 1999) and SC (Horwitz and Newsome, 1999). Interestingly, all neurons that exhibit ramping activity characteristically show persistent neural firing in delayed memory or decision tasks (Gnadt and Andersen, 1988; Shadlen and Newsome, 2001). This observation has inspired the application of a biophysically based model of working memory (Brunel and Wang, 2001) on decision making (Wang, 2002). In the model, strong recurrent connections generate attractor states, which facilitate sustained spiking activity in excitatory subpopulations of the neural network (Figure 3A, S1 and S2), while global inhibitory feedback leads to competition between these subgroups, and thus enables categorical decision making. In the following, the basic network shown in Figure 3A serves as a building block to model particular brain regions that participate in the processing of competitive features. The spiking neuron models of visual-attention mechanisms and information transfer, which are described in the subsequent sections, involve multiple cortical areas, and hence consist of several of these basic decision units. Here, in the case of the RDM task, the network can be viewed as a representation of a local microcircuit in area LIP, where one neural subpopulation is selective for each of the possible motion directions (see Albantakis and Deco, 2009 for multiple choices).

Decision formation corresponds to the transition from the spontaneous state of the network (where all neurons fire at low firing rates) to a decision state [where one selective population (the “winner”)



fires at high rates (**Figures 3B,C**]). If the connection strengths are fixed, the input strength determines whether particular attractor states are stable or not (**Figure 3D**). For sufficient external inputs, the spontaneous state becomes unstable ( $>10$  Hz in **Figure 3D**), and the system “relaxes” into one of the two possible decision states driven by sensory evidence. The transition time increases the closer the system is to this bifurcation point. In addition to the attractor configuration, the network’s long synaptic time-constant, generated by a high NMDA to AMPA receptor ratio, is crucial for a slow transition and for the model’s ability to accumulate inputs.

Note that, although individual spiking neurons are simulated, the decision outcome is determined by the pooled activity of the selective neural populations, consistent with a rate-code, and not by individual spikes, as opposed to the feedforward network for object recognition described in Section “Fast Feedforward Processing, Latency Coding, and STDP.” Also, in contrast to the feedforward model, the decision making model is inherently stochastic, as every neuron in the network receives its own individual background inputs in the form of Poisson spike trains. As there are a finite number of neurons in the network, the resulting output spike rate of each neural population also fluctuates in time around the noise-free value, or, equivalently, the firing rate obtained for an infinite number of neurons. The neural noise plays an important role in the model’s decision making. First, it is responsible for the probabilistic outcome of the decision process when faced with ambiguous evidence for both alternatives (as in **Figures 3B,C**). Moreover, we showed that in the case of low sensory input, where the spontaneous state is still stable (left to the bifurcation), fluctuations due to the network’s finite size noise can cause transitions to the decision state (Martí et al., 2008). Without noise (corresponding to an infinite amount of neurons), the network would stay indefinitely in the spontaneous state for small external inputs. If the number of neurons in the network is small, fluctuations that are large enough to induce a transition to the decision state are more probable. These noise-driven decisions exhibit rather sharp switches in activity on single trials with long, exponentially distributed decision times. Nevertheless, averaging across trials with different decision times in the noise-driven condition results in a gradual build-up of activity (**Figure 3C**), consistent with the experimentally observed neural firing rates, which are trial-averaged single neuron activity.

As mentioned above in the evidence-driven regime, the transition times of the model depend on the common external input to the selective populations, with faster transients and lower accuracy for higher sensory inputs. Even if both selective populations receive the same input (no bias), the average (chance level) decision time will thus be shorter if this common input is higher (**Figure 3D**). This model characteristic arises through the non-linearity of the attractor landscape, and offers an interesting alternative mechanism to control the speed–accuracy trade-off (Roxin and Ledberg, 2008), apart from adapting the decision threshold, as suggested by conceptual models of decision making (Ratcliff and Smith, 2004; Palmer et al., 2005). In this context, we recently showed that the attractor model is capable of reproducing changes of mind that emerged through speed–pressure in a slightly altered RDM task (Resulaj et al., 2009), if the decision threshold is set low and, in addition, the external inputs applied to both selective populations

are high (Albantakis and Deco, 2011). Specifically, changes of mind in the model became more frequent, the closer the system was to the second bifurcation, where the symmetric state returns to be stable.

Another implication of the non-linearity inherent to the attractor model is the violation of the so-called time-shift invariance: evidence occurring earlier during the accumulation process will have a greater effect on the decision outcome than later evidence, which happens only when the transient is already converging toward one of the decision attractors (Wong et al., 2007). This prediction was indeed observed in a RDM experiment, where brief pulses of motion added to the random-dot stimulus affected the final choice more at earlier onset times (Huk and Shadlen, 2005). To produce this effect with a linear accumulator model, additional time-dependent features like collapsing decision bounds or an urgency signal need to be superimposed on the conceptual model.

In general, most models of perceptual decision making so far focused exclusively on sensory evidence accumulation. In that sense, the non-linear attractor model is a notable exception, as it is further able to account for other modalities of decision making neurons, like persistent activity and their responses to visual target signals (Wong and Huk, 2008). Nevertheless, not much is yet known about the physiological mechanisms of the various internal states, which can play a significant role in the decision making process, such as speed–accuracy trade-off, urgency, reward expectation, or attention.

## TOP-DOWN ATTENTION

Another situation where pure fast feedforward processing of spiking information is insufficient to perform the required computation, arises when a task demands the evaluation of a crowded and/or complex visual scene. In this case, the visual system is unable to simultaneously evaluate the immense amount of information conveyed in a complex scene just by the initial fast feedforward sweep of information transfer. Precisely to cope with this problem, attentional mechanisms are required to account for the selection of relevant scene information. In addition to the local recurrent connections treated in the previous section, intercortical recurrent connections between different brain areas shape the focus of attention.

## BIASED-COMPETITION MECHANISMS CAN ACCOUNT FOR THE ATTENTIONAL SPOTLIGHT

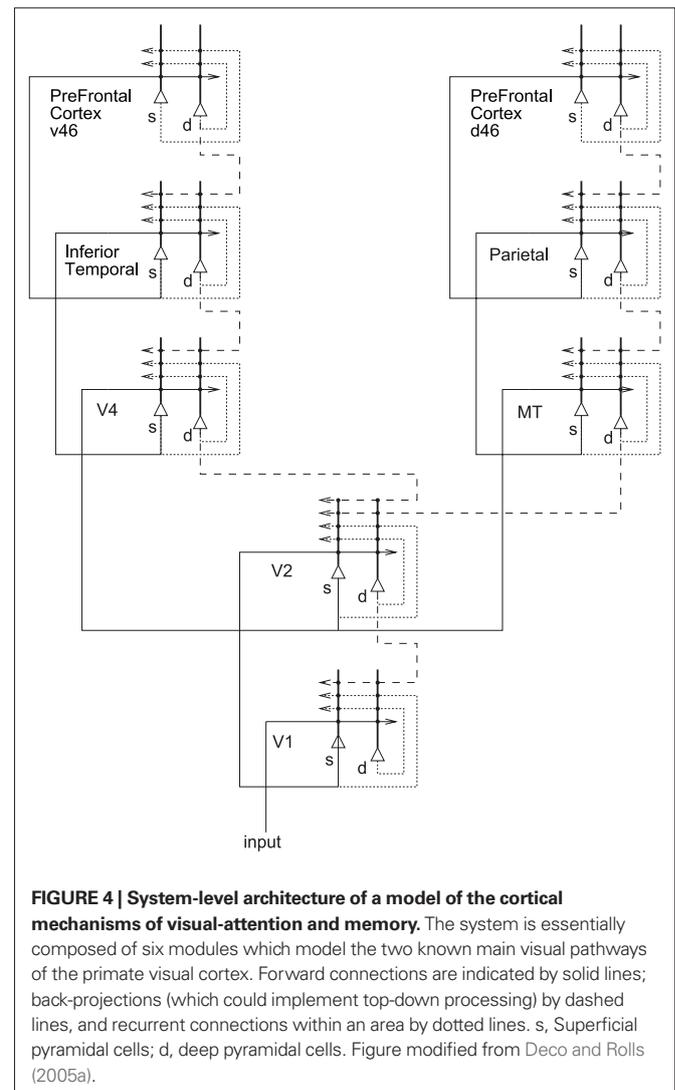
Attentional mechanisms optimize the processing of bottom-up relevant aspects of the sensory signal by adding top-down influences. These top-down signals bias the system to concentrate on only a small proportion of the incoming information relevant for the behavioral task under consideration. Top-down and bottom-up processing result from intercortical connections between the different brain areas. Indeed, one quarter of all possible connections between areas is realized in the human brain, most of which being of recurrent nature (Salin and Bullier, 1995). Thus, partial representations held in different cortical areas might be integrated by mutual cross communication, mediated by the inter-area neuronal fibers. The role of recurrent processing is central to modern perspectives on hierarchical inference in the brain. Modern accounts (e.g., predictive coding) see the brain as actively constructing predictions of its sensorium that are mediated by top-down connections, and tested against sensory evidence to provide a prediction error

(Spratling, 2008; Friston and Kiebel, 2009; Hesselmann et al., 2010). This error is then propagated through the system, and accumulated to optimize representations of the causes of sensory input. This view is based upon Helmholtzian ideas, and regards the brain as testing hypotheses about the causes of sensations. In this spirit, perception could be handled as an inverse inference problem, whose goal is to estimate the factors that have generated the particular percept. Indeed, this can be formalized in the framework of Bayesian Decision Theory (Friston and Kiebel, 2009; Hesselmann et al., 2010).

Further neurophysiological evidence gives rise to the assumption that each cortical area is capable of representing a set of alternative hypotheses encoded in the activities of different cell assemblies [similar to the selective populations (S1, S2) in the decision making network (Figure 3A)]. Representations of different conflicting hypotheses inside each area compete with each other for activity and representation (Desimone and Duncan, 1995). However, each area represents only a part of the environment and/or internal state. In order to achieve a coherent global representation, different cortical areas bias each other's internal representations by communicating their current states to other areas through inter-area connections. They favor thereby certain sets of local hypotheses over others. For example, different objects present in the visual field could compete for representation in one brain area (Wolfe, 1994). This competition might be resolved by a bias given to one of them from another area, as obtained from this other area's local view-encoding. For example, it could favor the behaviorally relevant location in the visual field, and thus the object corresponding to that location to be represented in the first area (Rolls and Deco, 2002, 2010). Each brain area might thus act like the decision network described in Figure 3A, with multiple competing alternatives. By recurrently biasing each other's competitive internal dynamics, the global neocortical system dynamically achieves a global representation in which each area's state is maximally consistent with those of the other areas. This view has been referred to as the "biased-competition" hypothesis (Desimone and Duncan, 1995).

In parallel to this competition-centered view, a cooperation-centered picture of brain operation has been formulated, where global representations find their neuronal correlate in assemblies of co-activated neurons (Hebb, 1949). Co-activation of neurons induces stronger mutual synaptic connections between themselves, which leads to assembly formation. Reverberatory communication between assembly members then results in persistent neuronal activation, and gives rise to a representation extended in time, as described in Section "Slower Visual Decision Making" for visual decision making. The concept of neuronal assemblies was later formalized in the framework of statistical physics (Hopfield, 1982; Amit and Brunel, 1997; Brunel and Wang, 2001), where assemblies of co-activated neurons form attractors in the phase space of the recurrent neuronal dynamics (patterns of co-activation can represent fixed points from which the dynamical system evolves). In summary, the formalism of attractor dynamics, including biased competition and cooperation, offers a unifying principle for the "slow" recurrent integration and segregation of information in multi-area neurocognitive modeling of brain functions (Deco and Rolls, 2005a,b; Deco et al., 2009; Mavritsaki et al., 2011).

A cortical architecture that implements the principles of visual-attention described above is shown in Figure 4 (see Deco and Rolls, 2005a,b) for more details). The figure shows how the dorsal "where" visual stream (reaching the posterior parietal cortex, PP) and the ventral "what" visual stream (via V4 to the inferior temporal cortex, IT) interact through early visual cortical areas (such as V1 and V2) to account for many aspects of visual-attention. The system is composed of six modules [V1 (the primary visual cortex), V2–V4, IT, PP, ventral PFC v46, and dorsal PFC d46], reciprocally connected according to anatomical data. This multi-area neurodynamical model implements the principle of biased-competition (presented above) at the local and global brain area level. Information from the retina reaches V1 via the LGN. The attentional top-down signal biasing the intra- and intercortical competition is assumed to come from PFC area 46 (modules d46 and v46). In particular, feedback connections from area v46 with the IT module could specify the target object in a visual search task. The feedback connections from area d46 with the PP module generate the bias to a targeted spatial location in an object recognition task given a spatial attentional cue. Each brain area consists of mutually coupled neuronal populations,



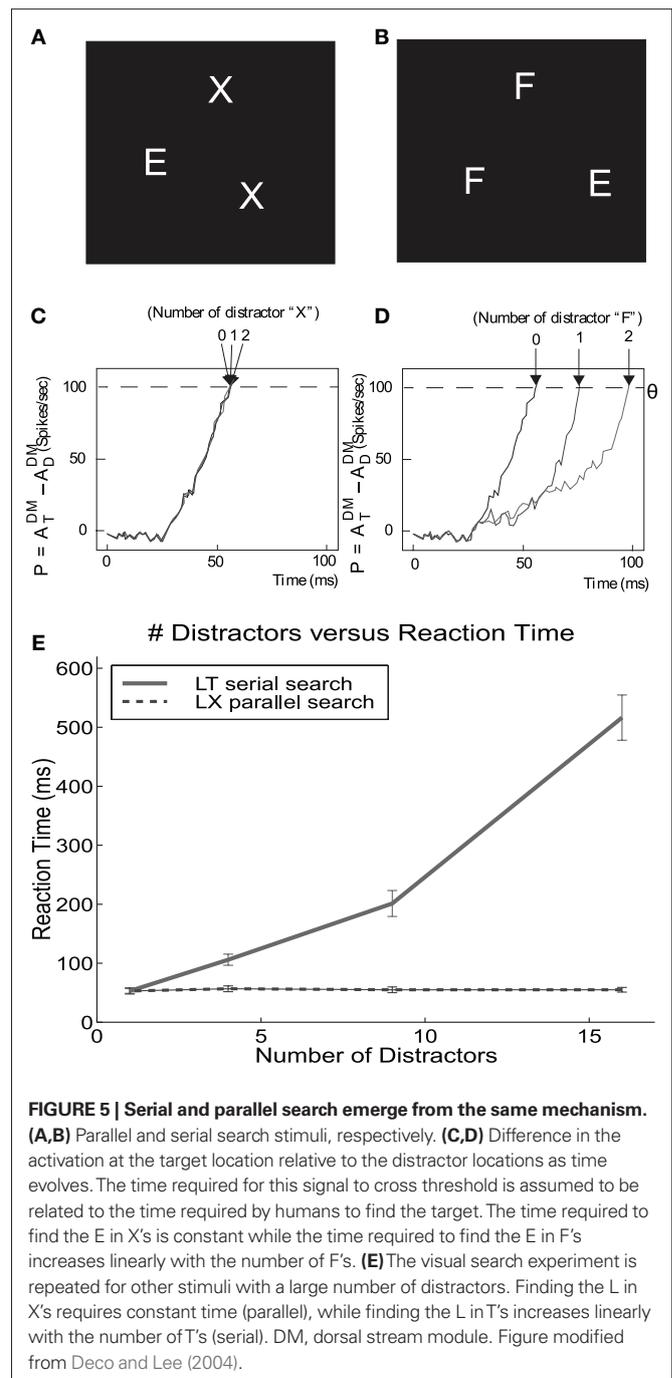
whose dynamics are described by conductance-based synaptic and spiking neuronal models. The equations describing the detailed neuronal dynamics can be further reduced using mean-field techniques. The mean-field approximation consists of replacing the temporally averaged discharge rate of a neuron with the instantaneous ensemble average of the activity of the neuronal population (see Rolls and Deco, 2010). The dynamical evolution of activity at the level of a cortical area can be simulated in the framework of the present model by integrating the population activity in a given area over space and time. An explicit spiking neuron simulation of two coupled brain regions (V2 and V4) engaged in biased-competition, with each population acting according to the network shown in **Figure 3A**, is described in (Deco and Rolls, 2005b), and revealed further insights into the non-linear interactions between bottom-up and attentional top-down effects.

### ATTENTION IN VISUAL SEARCH

One source of evidence for attentional mechanisms in visual processing comes from psychophysical experiments using visual search tasks. This was proposed by Treisman and Gelade (1980); see also (Pashler, 1998) for other types of experiments evidencing attention. There, subjects examine a display containing randomly positioned items in order to detect a previously defined target. All other items in the display, which are different from the target, play the role of distractors. The main phenomenology can be understood from the dependence of the measured reaction time as a function of the number of items in the display. There are two main types of searching displays, namely: feature search or “pop out,” and conjunction or serial search. In a feature search task, the target differs from the distractors in a single feature, (e.g., only in its color). In this case, search times are independent of the number of distractors. In a conjunction search task, the target is defined by a conjunction of features, and each distractor shares at least one of those features with the target. The conjunction search experiments show that search time increases linearly with the number of distractors, implying a serial process.

The computation of a visual search works as follows. An external top-down bias from prefrontal area v46 to the IT module drives the competition in IT in favor of the population encoding the target object. Then, the intermodular back-projected attentional modulation IT–V4–V1 enhances the activity of the populations in V4 and V1, which encode the component features of the target. Only the locations in V1 matching the back-projected target features are up-regulated. The enhanced firing of the neuronal populations encoding the particular location of the target in V1 lead to increased activity in the spatially mapped forward pathway from V1 to V2–V4 to PP. This results in an increased firing in the PP module in the location that corresponds to the target. Consequently, these cascades of biased-competitions compute the location of the target, and are made explicit by the enhanced firing activity of neuronal populations at the location of the target in the spatially organized PP module. (Deco and Lee, 2004) showed that the properties of feature and conjunction search are both reproduced by this attentional architecture, as shown in **Figure 5**.

The implication of these computational results is that, while the network searches the visual field in parallel, there are differences in the latencies of the neural responses in the different conditions,



**FIGURE 5 | Serial and parallel search emerge from the same mechanism.**

(A,B) Parallel and serial search stimuli, respectively. (C,D) Difference in the activation at the target location relative to the distractor locations as time evolves. The time required for this signal to cross threshold is assumed to be related to the time required by humans to find the target. The time required to find the E in X's is constant while the time required to find the E in F's increases linearly with the number of F's. (E) The visual search experiment is repeated for other stimuli with a large number of distractors. Finding the L in X's requires constant time (parallel), while finding the L in T's increases linearly with the number of T's (serial). DM, dorsal stream module. Figure modified from Deco and Lee (2004).

related to how easily the dynamical system can perform the constraint satisfaction for the different conditions (see also Heinke and Backhaus, 2011).

### OSCILLATIONS FORMAT VISUAL PROCESSING

As we have seen above, communication between higher and lower level brain areas is crucial to direct attention in visual search or complex visual scenes. Information transfer mediated by local and intercortical recurrent connections is generally associated with oscillatory activity. Particularly in the visual system, oscillatory

activity has been widely reported experimentally, especially in the gamma frequency range. Yet, whether oscillations have a major functional role or, instead, would only be a by-product of neuronal information processing, is still debated. In this section, we argue that some of our recent modeling studies suggest at least three main functions for oscillatory activity:

### OSCILLATIONS AND ATTENTION

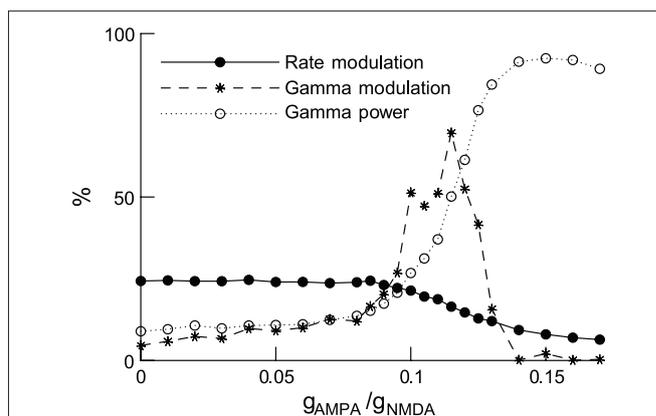
The biased-competition theory claims that the neuronal response – in terms of firing rate – to simultaneously presented stimuli is a weighted average of the response to isolated stimuli, and that attention biases the weights in favor of the attended stimulus (Desimone and Duncan, 1995). Thus, a neuron's firing rate increases when its preferred stimulus is attended, but decreases when the non-preferred one is attended. More recently, it has been shown that attention has also an effect on synchrony: selective attention to a visual stimulus specifically enhances the gamma-band synchronization among neurons in monkey's extrastriate visual cortex driven by that stimulus (Fries et al., 2001b, 2008; Bichot et al., 2005; Taylor et al., 2005; Womelsdorf et al., 2006). In humans, several EEG and MEG studies have found similar effects (Jensen et al., 2007; Tallon-Baudry, 2009). Although rate and gamma synchrony modulations occur simultaneously, it is not clear if and how they are mechanistically related.

To investigate this issue, we recently extended the analysis of the above-mentioned model (Deco and Rolls, 2005b), in which biased-competition is implemented in a network of excitatory and inhibitory spiking neurons (as in **Figure 3A**), and attention is modeled as an additional input to the neurons encoding the attended stimulus. We looked at the effect of this input on both firing rates and gamma synchronization (Buehlmann and Deco, 2008). In order to allow oscillations; we increased the ratio of excitatory synaptic conductivities  $g_{\text{AMPA}}/g_{\text{NMDA}}$ . Indeed, when the shorter AMPA latencies dominate over the long-lasting NMDA ones, the latency of the excitatory components is smaller than the one of the inhibitory GABA components, resulting in the generation of oscillations (Brunel and Wang, 2003).

In accordance with the experiments, a stimulus generates correlated neural activity in the gamma frequency band, and its power is stronger for the neurons encoding the attended stimulus than for the neurons encoding the unattended stimulus. As the  $g_{\text{AMPA}}/g_{\text{NMDA}}$  conductance ratio increases, the attentional rate modulation decreases monotonically but the gamma modulation first increases up to a maximum and then decreases (**Figure 6**). These results imply that rate and gamma modulations can occur independently of each other, and are therefore not concomitant effects. Furthermore, gamma modulations are desirable because they were found to decrease the reaction times, in line with experimentation in monkeys (Womelsdorf et al., 2006). This suggests an optimal  $g_{\text{AMPA}}/g_{\text{NMDA}}$  conductance ratio.

### COMMUNICATION THROUGH COHERENCE

Another desirable effect of rhythmic synchronization is that it allows the flexible routing of information between neuron pools. Consider two pools, A and B, oscillating at the same frequency. A projects on B, but A's spikes will significantly influence B if, and only if, they arrive during a critical period of excitability. Thus, by shifting the



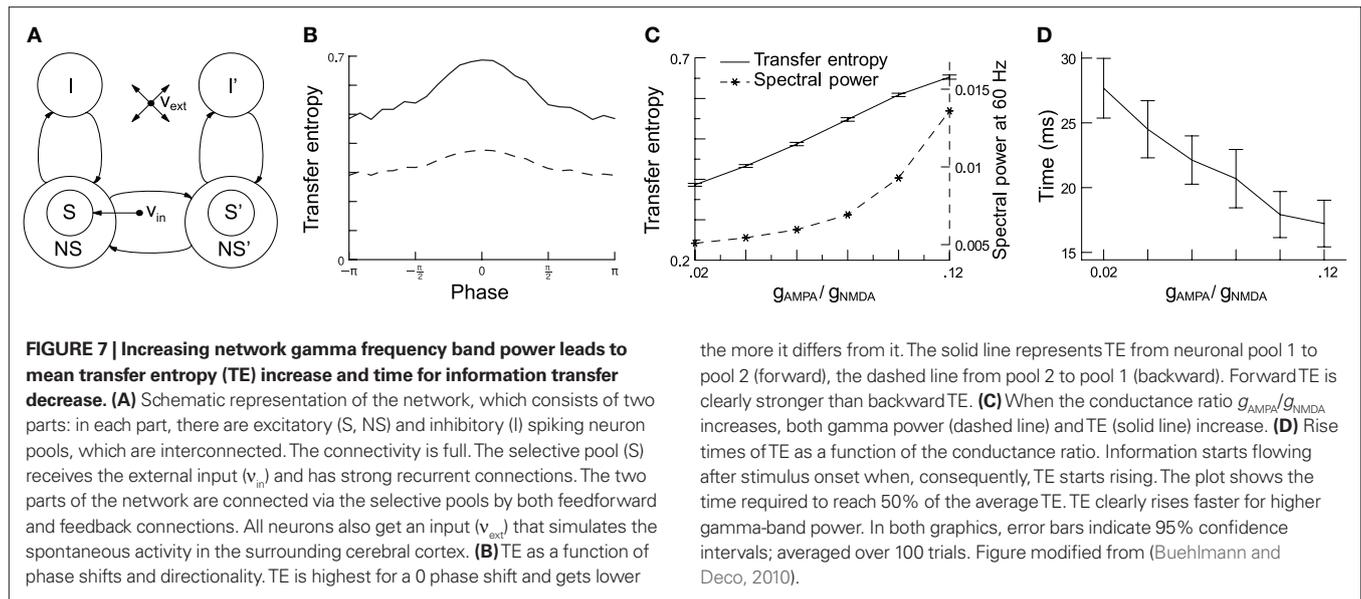
**FIGURE 6 | In attention, rate and gamma modulations are not concomitant effects.** Rate modulation (solid curve) and gamma modulation (dashed curve) as a function of the excitatory synaptic conductance ratio  $g_{\text{AMPA}}/g_{\text{NMDA}}$ . Increasing this ratio increases rhythmic gamma-band power (dotted curve), decreases the rate modulation monotonically while gamma modulation has a peak. Either of the two modulations can be dominant, depending on the gamma power. Figure modified from (Deco et al., 2011).

phase between the pools, one can virtually activate or deactivate the communication link between the pools. This is known as the “communication through coherence” (CTC) hypothesis (Fries, 2005). Direct physiological evidence for it is found in cat and monkey visual systems (Womelsdorf et al., 2007). In humans, the fact that a near-threshold visual stimulus can be perceived or not, depending on the phase of ongoing EEG oscillations at stimulus onset (Busch et al., 2009; Mathewson et al., 2009), is consistent with CTC.

Recently, we quantified the effect of phase shifting on the communication between two oscillating neuronal pools (**Figure 7A**) using transfer entropy (TE; Buehlmann and Deco, 2010). TE is an information theoretical measure that quantifies the statistical coherence between systems, and is able to distinguish between shared and transported information (Schreiber, 2000). In accordance with the experiments, we found that (i) there is an optimal phase relation at which TE is highest between the two groups of neurons (**Figure 7B**), that (ii) TE increases as a function of the gamma power (**Figure 7C**), and (iii) the speed of information transfer increases as a function of the gamma power, measured from the time required to reach 50% of the TE after stimulus onset (**Figure 7D**). Taken together, these findings support the CTC hypothesis and, as rhythmic neuronal synchronization makes information transport more efficient and flexible, they suggest that it has an important functional role.

### PHASE-OF-FIRING CODING, AND STDP-BASED DECODING

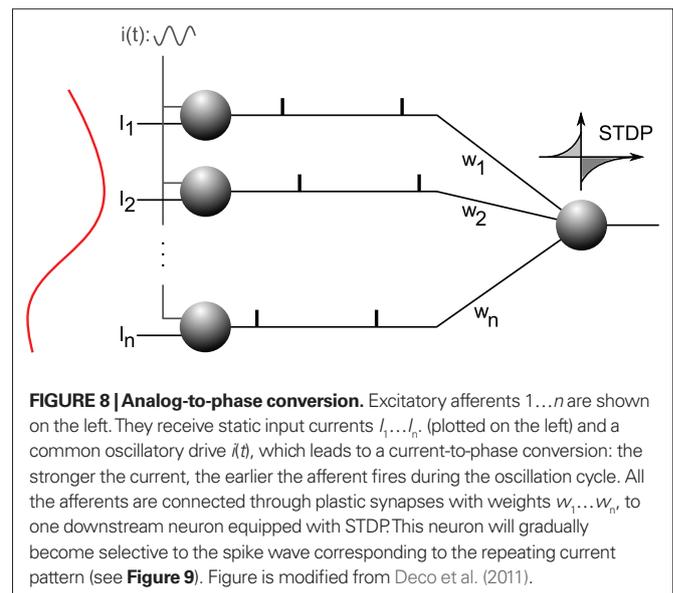
Communication through coherence suggests that there is an optimal time window for a neuron pool A to send spikes to another pool B, so that they have a significant impact on B. But how can information be encoded in those spikes? Recent experiments have established that information can be encoded in the spike phases with respect to a background oscillation in the local field potential (LFP)—a phenomenon referred to as PoFC. Evidence for such coding has been seen in the visual system, in particular in V1 (König et al., 1995; Fries et al., 2001a; Montemurro et al., 2008; Vinck et al., 2010) and V4 (Lee et al., 2005). These firing phase preferences could result from combining



an oscillatory drive with a stimulus-dependent current that would produce the variations in preferred phases (Hopfield, 1995). This mechanism is supported by direct physiological evidence *in vitro* (Schaefer et al., 2006; McLelland and Paulsen, 2009). However, it remains unknown if such a firing activity can be decoded, that is if downstream neurons can respond selectively to patterns of phases in their inputs, and if this behavior can be learned.

We have shown recently that STDP can solve the problem efficiently (Masquelier et al., 2009b). Specifically, a single neuron equipped with STDP (Figure 8) can robustly detect a hidden pattern repeating at random intervals, which involves only a subset of its afferents, and is automatically encoded in their firing phases (Figure 9). The oscillatory drive improves the spike time precision by decreasing their sensitivity to initial conditions, and avoiding jitter accumulation, so that they depend mainly on the current input values (Brette and Guigon, 2003; Hasenstaub et al., 2005; Schaefer et al., 2006; Markowitz et al., 2008). The ability of STDP to detect repeating spike patterns had been noted before in continuous activity (Masquelier et al., 2008, 2009a), but it turns out that oscillations greatly facilitate learning, which is possible even when only a small fraction of the afferents (~10%) exhibits PoFC. A benchmark with more conventional rate-based codes demonstrated the superiority of oscillations and PoFC for both STDP-based learning and speed of decoding, which only takes one oscillatory cycle.

The oscillatory drive formats the spike times into waves (Figure 9A) that are similar to the first spike waves after visual stimulus onset described in Section “Fast Feedforward Processing, Latency Coding, and STDP.” It is thus not so surprising that neurons equipped with STDP can also detect and learn repeating patterns in the spike waves caused by the oscillatory drive. This new oscillation-based scheme, however, can account for continuous vision, when no external time reference such as a stimulus onset is available. The scheme is particularly appealing for the processing of static, or slowly changing visual stimuli, which, without oscillations, would not generate precisely timed spikes (eye movements may be an alternative, see Continuous Vision). Consistent with our

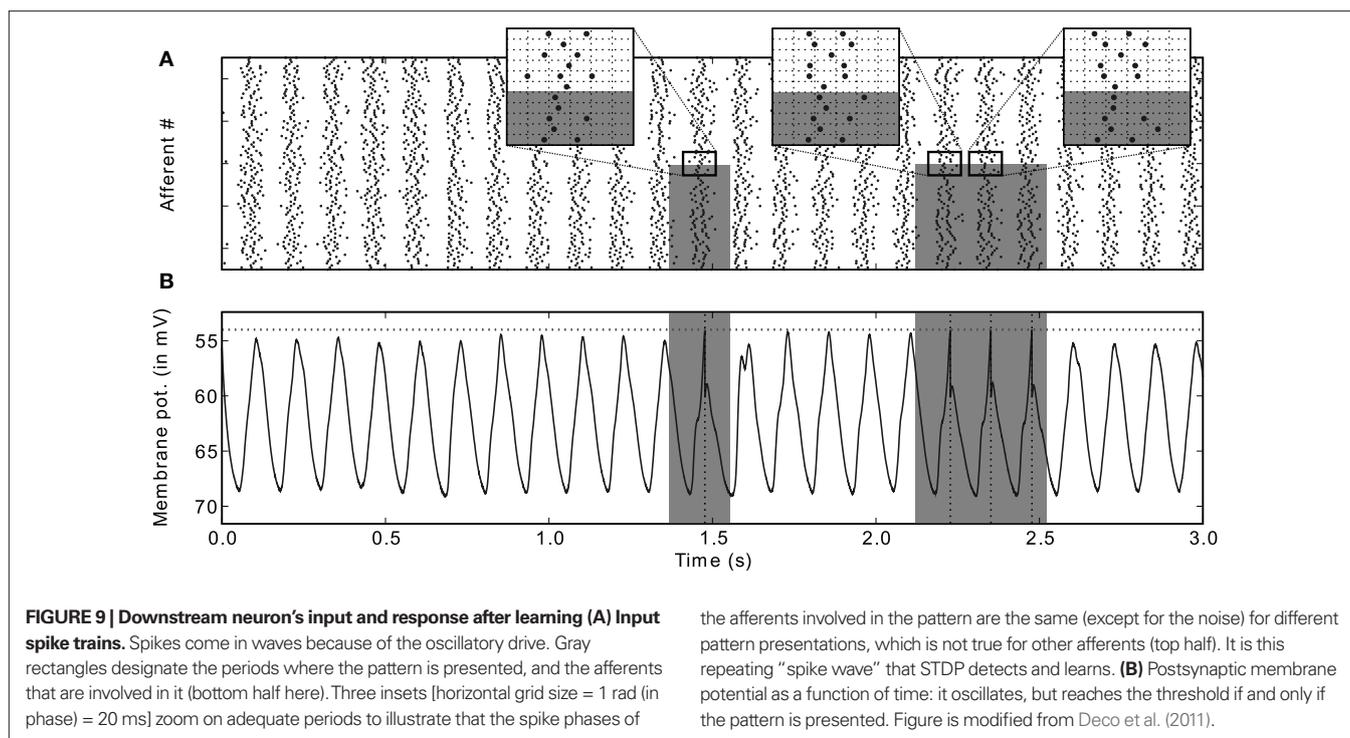


proposal, a growing body of experimental evidence in animals and humans demonstrates that successful long-term memory encoding correlates with increased oscillatory activity across a broad range of frequencies (from theta to gamma), in particular in the visual modality (Jensen et al., 2007; Klimesch et al., 2008; Tallon-Baudry, 2009). Interestingly, beyond mere oscillation power, what seems to be a prerequisite for successful visual memory formation is that single units should be phase-locked to the oscillation (Rutishauser et al., 2010) – a result consistent with our model.

## UNSOLVED QUESTIONS AND FUTURE DIRECTIONS

### CONTINUOUS VISION

In Section “Fast Feedforward Processing, Latency Coding, and STDP,” we focused on the transient activity generated when a stimulus suddenly appears at a given time from the dark, a paradigm



extensively studied in the lab but rather unnatural. A more natural situation is that an image is formed on the retina at  $t = t_0$  after a body or head movement, a saccade, or a micro-saccade (all of these are referred to as “movement” below). In that case, the “intensity-to-latency conversion” hypothesis we made is questionable for several reasons, in particular in the retina. First, the input current to a retinal ganglion cell (RGC) is a spatiotemporally filtered version of the luminance signal, as opposed to a spatially filtered version [among other things the surround signal is delayed (Enroth-Cugell et al., 1983; Cai et al., 1997)], and this spatiotemporal filtering does not stop during the movements. This means that the RGC input currents at  $t = t_0$  and slightly after, depend not only on the current image, but also on what happened during the movement, and possibly even before. Furthermore, these currents are integrated and converted into spikes. This introduces another dependence on history (the same input current does not lead to the same spike latencies, depending on when the last spike was emitted). For all these reasons, the times-to-first-spikes with respect to  $t_0$  are probably poor encoders of the current image. However, because the history of neighboring cells is likely to be similar, it seems reasonable to assume that this history will typically have a similar effect on their spike times, and thus a weak effect on their *relative* spike times – but this should be confirmed by simulations. Consistent with this idea, relative latencies are found to be more reliable than absolute ones in the retina (Gollisch and Meister, 2008).

We feel it is time to build models able to deal not only with “stimulus onset paradigms,” as the ones reviewed in Section “Fast Feedforward Processing, Latency Coding, and STDP,” but also with continuous vision, including body, head, and eye movements and moving stimuli. Such models could also simulate, unlike the current ones, the experimental protocols of rapid serial visual presentation

(RSVP) and visual masking. The timescales involved in natural continuous vision processing are fast (~10 ms; Butts et al., 2007), and individual neurons’ firing rates are not well defined at such a fine temporal resolution. Spiking neuron models should be preferred. STDP, which is able to detect consistently repeating spike patterns even in continuous activity (Masquelier et al., 2008, 2009a), probably plays a key role in continuous vision as well. Last but not least, continuous vision involves feedback loops, which should thus be included in those models. These should generate – among other things – self-sustained oscillations (Gray and Singer, 1989), and their desirable consequences reviewed in Section “Oscillations Format Visual Processing.”

#### HARDWARE IMPLEMENTATIONS

As we have seen in this review, encoding and processing information with spike patterns is an efficient strategy which is probably extensively used in the visual system. Software simulation of these mechanisms is time consuming though, which can reduce their relevance for technology. Silicon hardware implementations, however, could be several orders of magnitude faster than the biological hardware (which is incredibly slow: neurons cannot fire more than a few hundred spikes per second, and those impulses propagate on axons between neurons with a velocity of 1–2 m/s). This means that an artificial vision system based on biological algorithms implemented on silicon hardware could, in principle, clearly outperform animals including humans.

One appealing technology to implement spike-based processing is the so-called address event representation (AER), where the spikes are carried as addresses of sending or receiving neurons on a digital bus. Time “represents itself” as the asynchronous occurrence of the event. AER was first proposed in 1991 by

Mead's Lab at the California Institute of Technology (Sivilotti, 1991), and has been used since then by a wide community of hardware engineers. Furthermore, the recently discovered memristive nanoscale devices (Strukov et al., 2008) provide an appealing implementation of the STDP functionality (Linares-Barranco and Serrano-Gotarredona, 2009).

Together with Linares' group, we are building hardware self-learning models of the visual cortex, which combine both AER and memristor technologies. In a first attempt to simulate the early visual system, we used a simple set up combining an AER artificial retina (Lichtsteiner et al., 2007) and a SNN mimicking V1 (the LGN was ignored). The artificial retina sensed the external world in a continuous (frame-free) manner, and generated spikes that were asynchronously propagated, as they flowed in, until they reach the V1 SNN. In this network, neurons were equipped with memristor-based STDP (for now simulated). This enabled them to gradually become orientation selective, as the system was exposed to natural stimuli (Zamarreño-Ramos et al., 2011). These results are still preliminary, but very encouraging. We speculate that this line of research will yield revolutionary results in the next decade.

#### **DISTINGUISHING DECISION MAKING MODEL APPROACHES**

Models on the accumulation of noisy evidence, as for instance during continuous motion viewing, come in a huge variety of flavors, which may be very difficult to distinguish on the basis of just behavioral data or even mean firing rates. Finding new analytical methods and intelligently designed experiments to distinguish the different approaches is thus a major future challenge in the field of perceptual decision making. Several recent studies have acknowledged this objective with a particular emphasis on multiple alternatives (Ditterich, 2010; Leite and Ratcliff, 2010; Purcell et al., 2010; Churchland et al., 2011).

Analyzing higher-order statistical properties (i.e., a variance and within-trial correlation measure) of neurophysiological data from a two- and four-alternative RDM task, (Churchland et al., 2011) could help distinguish between models categorized by their different sources of variability. Models with just one source of variability [either with a randomly varying slope but no within-trial noise (Carpenter and Williams, 1995), or a fixed slope with a random distribution of firing rates at each time-step (Cisek et al., 2009)] failed to account for the higher-order measures, although they agreed with behavior and mean firing rates. On the other hand, all different implementations of a stochastic accumulation to threshold, the drift-diffusion model (Ratcliff and Rouder, 1998) – a model based on probabilistic population codes (Beck et al., 2008) – and a recurrent attractor model (Wong et al., 2007) – a reduction of the model described in Section “Slower Visual Decision Making” – also matched the experimental data in variance and correlation.

Based on human behavioral data from a RDM task with three alternatives and three motion components, (Ditterich, 2010) intended to distinguish more detailed aspects of conceptual accumulation-to-bound models with regard to their goodness of fit and their neurophysiological predictions. Perfect integrators were compared to leaky, saturating integrators, with either feedback or feedforward inhibition. Note that most of the discussed models were found equivalent for certain parameter ranges (Bogacz et al.,

2006). Hence, it might not be too surprising that none of the models could be excluded based only on the fits to the behavioral data. However, they differ substantially in their neurophysiological predictions on how the integrator states should evolve over time (see Table 2 in Ditterich, 2010). Invasive neural recordings from monkeys performing the same task will hopefully soon settle the dispute. Moreover, feedforward and feedback inhibition respectively suggest either negative or positive correlation between the integrator units, which might be tested with multi-electrode recordings.

Finally, for equal coherences in all three motion directions, (Niwa and Ditterich, 2008) measured faster mean reaction times for higher coherence levels, consistent with the predictions from the non-linear recurrent attractor network for increasing external inputs to all selective populations (see Slower Visual Decision Making). While models with feedforward inhibition require a scaling of the variance of the sensory signals in order to account for this effect, conceptual models with feedback inhibition could explain the result just with a change of the mean input (Ditterich, 2010). In that context, the predictions of the biophysically based attractor model on reaction times and changes of mind could also be tested more rigorously in a change of mind RDM experiment with two directionally opposite motion components (see Albantakis and Deco, 2011).

#### **CONCLUSION**

With this review we aimed to outline, within the frame of SNNs, the various ways in which different processing timescales imply and connect to different neural dynamics in the visual system. For object recognition, the high processing speed excludes extensive crosstalk between neural populations, and feedforward connectivity seems sufficient to explain experimental observations. However, recurrent connections are crucial for any non-linear operation, such as data integration or shaping the focus of attention, in tasks where higher level processing is beneficial in spite of consequently longer reaction times. Moreover, oscillatory activity might act as a higher-order mechanism for routing and encoding the exchanged information. Depending on which particular task the visual system is currently engaged in, the amount of information that is transmitted back and forth within and between the relevant brain areas thus varies substantially. Nonetheless, not only the amount of information exchanged between neural populations is task-dependent, the way the information is encoded also differs for the different processing modes. In fact, with the four modes of temporal processing, we have presented four distinct ways of how information might spread through the visual pathways: during object recognition for the fast feedforward sweep of activity along the ventral pathway, which consists of only one or a few spikes at each processing level, “rank order coding” (Thorpe and Gautrais, 1998) allows to convey information despite the low number of spikes, which excludes the classic rate coding scheme. Rate coding does still play the dominant role in visual discrimination tasks, where information is accumulated in decision-related brain areas along the dorsal visual stream. If the interplay between top-down and bottom-up signaling contributes to solving task-specific challenges to the visual system (such as directing attention or visual search), information may be routed via oscillatory activity, as described in the CTC theory (Fries, 2005). Finally, background oscillations in the LFP could serve as an

internal substitute for an external temporal reference frame, which allows temporal encoding of information through the spike phases. This PoFC (Montemurro et al., 2008) provides a possible temporal code that is applicable even in the absence of external time frames, as in continuous vision or for long-lasting stimuli.

To conclude, we have shown that all these different coding schemes can be implemented in biologically inspired spiking neuron models with the associated neural dynamics determined by their network connectivity. The connection weights in the respective models were assumed to have formed according to Hebbian rules. Synapses that implement STDP can further shape the spiking network to perform temporal coding, and also to decode the information again. It remains to be investigated how robust the

temporal codes are when faced with real, noisy sensory inputs, and to what extent the brain actually takes functional advantage of the various hypothetical neural codes. Yet, with our review, we want to emphasize the complementary nature and common basic principles of the different encoding schemes and neural dynamics that might operate alternatively in the visual system through a switch procedure, or even simultaneously, through multiplexed temporal scales (Victor, 2000; Panzeri et al., 2010).

## ACKNOWLEDGMENTS

The authors were supported by the Fyssen Foundation, the FP7 European Project Coronet, and the CONSOLIDER-INGENIO 2010 Programme CSD2007-00012.

## REFERENCES

- Albantakis, L., and Deco, G. (2009). The encoding of alternatives in multiple-choice decision making. *Proc. Natl. Acad. Sci. U.S.A.* 106, 10308–10313.
- Albantakis, L., and Deco, G. (2011). Changes of mind in an attractor network of decision-making. *PLoS Comput. Biol.* 7, e1002086. doi:10.1371/journal.pcbi.1002086
- Albrecht, D. G., Geisler, W. S., Frazor, R. A., and Crane, A. M. (2002). Visual cortex neurons of monkeys and cats: temporal dynamics of the contrast response function. *J. Neurophysiol.* 88, 888–913.
- Amit, D. J., and Brunel, N. (1997). Model of global spontaneous activity and local structured activity during delay periods in the cerebral cortex. *Cereb. Cortex* 7, 237–252.
- Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., and Pouget, A. (2008). Probabilistic population codes for bayesian decision making. *Neuron* 60, 1142–1152.
- Bichot, N. P., Rossi, A. F., and Desimone, R. (2005). Parallel and serial neural mechanisms for visual search in macaque area v4. *Science* 308, 529–534.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., and Cohen, J. D. (2006). The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* 113, 700–765.
- Brette, R., and Guigon, E. (2003). Reliability of spike timing is a general property of spiking model neurons. *Neural Comput.* 15, 279–308.
- Britten, K., Shadlen, M., Newsome, W., and Movshon, J. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis. Neurosci.* 10, 1157–1169.
- Brunel, N., and Wang, X. J. (2001). Effects of neuromodulation in a cortical network model of object working memory dominated by recurrent inhibition. *J. Comput. Neurosci.* 11, 63–85.
- Brunel, N., and Wang, X.-J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* 90, 415–430.
- Buehlmann, A., and Deco, G. (2008). The neuronal basis of attention: rate versus synchronization modulation. *J. Neurosci.* 28, 7679–7686.
- Buehlmann, A., and Deco, G. (2010). Optimal information transfer in the cortex through synchronization. *PLoS Comput. Biol.* 6, e1000934. doi: 10.1371/journal.pcbi.1000934
- Busch, N. A., Dubois, J., and VanRullen, R. (2009). The phase of ongoing eeg oscillations predicts visual perception. *J. Neurosci.* 29, 7869–7876.
- Butts, D. A., Weng, C., Jin, J., Yeh, C.-I., Lesica, N. A., Alonso, J.-M., and Stanley, G. B. (2007). Temporal precision in the neural code and the timescales of natural vision. *Nature* 449, 92–95.
- Cai, D., DeAngelis, G. C., and Freeman, R. D. (1997). Spatiotemporal receptive field organization in the lateral geniculate nucleus of cats and kittens. *J. Neurophysiol.* 78, 1045–1061.
- Caporale, N., and Dan, Y. (2008). Spike timing-dependent plasticity: a hebbian learning rule. *Annu. Rev. Neurosci.* 31, 25–46.
- Carpenter, R. H., and Williams, M. L. (1995). Neural computation of log likelihood in control of saccadic eye movements. *Nature* 377, 59–62.
- Celebrini, S., Thorpe, S., Trotter, Y., and Imbert, M. (1993). Dynamics of orientation coding in area V1 of the awake primate. *Vis. Neurosci.* 10, 811–825.
- Churchland, A. K., Kiani, R., Chaudhuri, R., Wang, X.-J., Pouget, A., and Shadlen, M. N. (2011). Variance as a signature of neural computations during decision making. *Neuron* 69, 818–831.
- Churchland, A. K., Kiani, R., and Shadlen, M. N. (2008). Decision-making with multiple alternatives. *Nat. Neurosci.* 11, 693–702.
- Cisek, P., Puskas, G. A., and El-Murr, S. (2009). Decisions in changing conditions: the urgency-gating model. *J. Neurosci.* 29, 11560–11571.
- Crouzet, S. M., Kirchner, H., and Thorpe, S. J. (2010). Fast saccades toward faces: face detection in just 100 ms. *J. Vis.* 10, 16.1–16.17.
- Deco, G., Buehlmann, A., Masquelier, T., and Hugues, E. (2011). The role of rhythmic neural synchronization in rest and task conditions. *Front. Hum. Neurosci.* 5:4. doi: 10.3389/fnhum.2011.00004
- Deco, G., and Lee, T. S. (2004). The role of early visual cortex in visual integration: a neural model of recurrent interaction. *Eur. J. Neurosci.* 20, 1089–1100.
- Deco, G., and Rolls, E. T. (2005a). Attention, short-term memory, and action selection: a unifying theory. *Prog. Neurobiol.* 76, 236–256.
- Deco, G., and Rolls, E. T. (2005b). Neurodynamics of biased competition and cooperation for attention: a model with spiking neurons. *J. Neurophysiol.* 94, 295–313.
- Deco, G., Rolls, E. T., and Romo, R. (2009). Stochastic dynamics as a principle of brain function. *Prog. Neurobiol.* 88, 1–16.
- Delorme, A., and Thorpe, S. J. (2001). Face identification using one spike per neuron: resistance to image degradations. *Neural Netw.* 14, 795–803.
- Desimone, R., and Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* 18, 193–222.
- Ditterich, J. (2010). A comparison between mechanisms of multi-alternative perceptual decision making: ability to explain human behavior, predictions for neurophysiology, and relationship with decision theory. *Front. Neurosci.* 4:184. doi: 10.3389/fnins.2010.00184
- Enroth-Cugell, C., Robson, J. G., Schweitzer-Tong, D. E., and Watson, A. B. (1983). Spatio-temporal interactions in cat retinal ganglion cells showing linear spatial summation. *J. Physiol. (Lond.)* 341, 279–307.
- Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci. (Regul. Ed.)* 9, 474–480.
- Fries, P., Neuenschwander, S., Engel, A. K., Goebel, R., and Singer, W. (2001a). Rapid feature selective neuronal synchronization through correlated latency shifting. *Nat. Neurosci.* 4, 194–200.
- Fries, P., Reynolds, J. H., Rorie, A. E., and Desimone, R. (2001b). Modulation of oscillatory neuronal synchronization by selective visual attention. *Science* 291, 1560–1563.
- Fries, P., Womelsdorf, T., Oostenveld, R., and Desimone, R. (2008). The effects of visual stimulation and selective visual attention on rhythmic neuronal synchronization in macaque area v4. *J. Neurosci.* 28, 4823–4835.
- Friston, K., and Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 364, 1211–1221.
- Gawne, T., Kjaer, T., and Richmond, B. (1996). Latency: another potential code for feature binding in striate cortex. *J. Neurophysiol.* 76, 1356–1360.
- Gerstner, W., and Kistler, W. (2002). *Spiking Neuron Models*. Cambridge, MA: Cambridge University Press.
- Gnadt, J. W., and Andersen, R. A. (1988). Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* 70, 216–220.
- Gollisch, T., and Meister, M. (2008). Rapid neural coding in the retina with relative spike latencies. *Science* 319, 1108–1111.
- Gray, C. M., and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 86, 1698–1702.

- Guyonneau, R., VanRullen, R., and Thorpe, S. (2005). Neurons tune to the earliest spikes through STDP. *Neural Comput.* 17, 859–879.
- Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., and McCormick, D. A. (2005). Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* 47, 423–435.
- Hebb, D. O. (1949). *The Organization of Behavior*. New York: Wiley.
- Heinke, D., and Backhaus, A. (2011). Modelling visual search with the selective attention for identification model (vs-saim): a novel explanation for visual search asymmetries. *Cogn. Comput.* 3, 185–205.
- Hesselmann, G., Sadaghiani, S., Friston, K. J., and Kleinschmidt, A. (2010). Predictive coding or evidence accumulation? false inference and neuronal fluctuations. *PLoS ONE* 5, e9926. doi: 10.1371/journal.pone.0009926
- Hopfield, J. (1995). Pattern recognition computation using action potential timing for stimulus representation. *Nature* 376, 33–36.
- Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U.S.A.* 79, 2554–2558.
- Horwitz, G. D., and Newsome, W. T. (1999). Separate signals for target selection and movement specification in the superior colliculus. *Science* 284, 1158–1161.
- Huk, A. C., and Shadlen, M. N. (2005). Neural activity in macaque parietal cortex reflects temporal integration of visual motion signals during perceptual decision making. *J. Neurosci.* 25, 10420–10436.
- Jensen, O., Kaiser, J., and Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci.* 30, 317–324.
- Jiang, Y., and Chun, M. M. (2001). Selective attention modulates implicit learning. *Q. J. Exp. Psychol. A* 54, 1105–1124.
- Kim, J. N., and Shadlen, M. N. (1999). Neural correlates of a decision in the dorsolateral prefrontal cortex of the macaque. *Nat. Neurosci.* 2, 176–185.
- Klimesch, W., Freunberger, R., Sauseng, P., and Gruber, W. (2008). A short review of slow phase synchronization and memory: evidence for control processes in different memory systems? *Brain Res.* 1235, 31–44.
- König, P., Engel, A. K., Roelfsema, P. R., and Singer, W. (1995). How precise is neuronal synchronization? *Neural Comput.* 7, 469–485.
- Lee, H., Simpson, G. V., Logothetis, N. K., and Rainer, G. (2005). Phase locking of single neuron activity to theta oscillations during working memory in monkey extrastriate visual cortex. *Neuron* 45, 147–156.
- Leite, F. P., and Ratcliff, R. (2010). Modeling reaction time and accuracy of multiple-alternative decisions. *Atten. Percept. Psychophys.* 72, 246–273.
- Lichtsteiner, P., Posch, C., and Delbruck, T. (2007). An 128x128 120db 15us-latency temporal contrast vision sensor. *IEEE J. Solid State Circuits* 43, 566–576.
- Linares-Barranco, B., and Serrano-Gotarredona, T. (2009). Memristance can explain spike-time-dependent-plasticity in neural synapses. *Nat. Precedings*. Available at <http://hdl.handle.net/10101/npre.2009.3010.1>
- Markowitz, D. A., Collman, F., Brody, C. D., Hopfield, J. J., and Tank, D. W. (2008). Rate-specific synchrony: using noisy oscillations to detect equally active neurons. *Proc. Natl. Acad. Sci. U.S.A.* 105, 8422–8427.
- Martí, D., Deco, G., Mattia, M., Gigante, G., and Giudice, P. D. (2008). A fluctuation-driven mechanism for slow decision processes in reverberant networks. *PLoS ONE* 3, e2534. doi: 10.1371/journal.pone.0002534
- Masquelier, T., Guyonneau, R., and Thorpe, S. J. (2008). Spike timing dependent plasticity finds the start of repeating patterns in continuous spike trains. *PLoS ONE* 3, e1377. doi: 10.1371/journal.pone.0001377
- Masquelier, T., Guyonneau, R., and Thorpe, S. J. (2009a). Competitive STDP-based spike pattern learning. *Neural Comput.* 21, 1259–1276.
- Masquelier, T., Hugues, E., Deco, G., and Thorpe, S. J. (2009b). Oscillations, phase-of-firing coding, and spike timing-dependent plasticity: an efficient learning scheme. *J. Neurosci.* 29, 13484–13493.
- Masquelier, T., and Thorpe, S. J. (2006). “Face feature learning with spike timing dependent plasticity,” in *Proceedings of the 1st French Conference on Computational Neuroscience (NeuroComp)*, Pont-à-Mousson.
- Masquelier, T., and Thorpe, S. J. (2007). Unsupervised learning of visual features through spike timing dependent plasticity. *PLoS Comput. Biol.* 3, e31. doi: 10.1371/journal.pcbi.0030031
- Mathewson, K. E., Gratton, G., Fabiani, M., Beck, D. M., and Ro, T. (2009). To see or not to see: prestimulus alpha phase predicts visual awareness. *J. Neurosci.* 29, 2725–2732.
- Mavritsaki, E., Heinke, D., Allen, H., Deco, G., and Humphreys, G. W. (2011). Bridging the gap between physiology and behavior: evidence from the sots model of human visual attention. *Psychol. Rev.* 118, 3–41.
- McLelland, D., and Paulsen, O. (2009). Neuronal oscillations and the rate-to-phase transform: mechanism, model and mutual information. *J. Physiol. (Lond.)* 587, 769–785.
- Montemurro, M. A., Rasch, M. J., Murayama, Y., Logothetis, N. K., and Panzeri, S. (2008). Phase-of-firing coding of natural visual stimuli in primary visual cortex. *Curr. Biol.* 18, 375–380.
- Niwa, M., and Ditterich, J. (2008). Perceptual decisions between multiple directions of visual motion. *J. Neurosci.* 28, 4435.
- Opris, I., and Bruce, C. J. (2005). Neural circuitry of judgment and decision mechanisms. *Brain Res. Brain Res. Rev.* 48, 509–526.
- Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* 5, 376–404.
- Panzeri, S., Brunel, N., Logothetis, N. K., and Kayser, C. (2010). Sensory neural codes using multiplexed temporal scales. *Trends Neurosci.* 33, 111–120.
- Pashler, H. E. (1998). *The Psychology of Attention*. Cambridge, MA: MIT Press.
- Purcell, B. A., Heitz, R. P., Cohen, J. Y., Schall, J. D., Logan, G. D., and Palmeri, T. J. (2010). Neurally constrained modeling of perceptual decision making. *Psychol. Rev.* 117, 1113–1143.
- Ratcliff, R., and Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychol. Sci.* 9, 347–356.
- Ratcliff, R., and Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychol. Rev.* 111, 333–367.
- Rathbun, D. L., Warland, D. K., and Usrey, W. M. (2010). Spike timing and information transmission at retinogeniculate synapses. *J. Neurosci.* 30, 13558–13566.
- Resulaj, A., Kiani, R., Wolpert, D. M., and Shadlen, M. N. (2009). Changes of mind in decision-making. *Nature* 461, 263–266.
- Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025.
- Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489.
- Rolls, E., and Deco, G. (2002). *Computational Neuroscience of Vision*. Oxford: Oxford University Press.
- Rolls, E., and Deco, G. (2010). *The Noisy Brain*. Oxford: Oxford University Press.
- Rousset, G., Thorpe, S., and Fabre-Thorpe, M. (2003). Taking the max from neuronal responses. *Trends Cogn. Sci. (Regul. Ed.)* 7, 99–102.
- Roxin, A., and Ledberg, A. (2008). Neurobiological models of two-choice decision making can be reduced to a one-dimensional nonlinear diffusion equation. *PLoS Comput. Biol.* 4, e1000046. doi: 10.1371/journal.pcbi.1000046
- Rutishauser, U., Ross, I. B., Mamelak, A. N., and Schuman, E. M. (2010). Human memory strength is predicted by theta-frequency phase-locking of single neurons. *Nature* 464, 903–907.
- Salin, P. A., and Bullier, J. (1995). Corticocortical connections in the visual system: structure and function. *Physiol. Rev.* 75, 107–154.
- Schaefer, A. T., Angelo, K., Spors, H., and Margrie, T. W. (2006). Neuronal oscillations enhance stimulus discrimination by ensuring action potential precision. *PLoS Biol.* 4, e163. doi: 10.1371/journal.pbio.0040163
- Schall, J. D. (2003). Neural correlates of decision processes: neural and mental chronometry. *Curr. Opin. Neurobiol.* 13, 182–186.
- Schall, J. D., Stuphorn, V., and Brown, J. W. (2002). Monitoring and control of action by the frontal lobes. *Neuron* 36, 309–322.
- Schreiber, T. (2000). Measuring information transfer. *Phys. Rev. Lett.* 85, 461–464.
- Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., and Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 411–426.
- Shadlen, M. N., and Newsome, W. T. (2001). Neural basis of a perceptual decision in the parietal cortex (area lip) of the rhesus monkey. *J. Neurophysiol.* 86, 1916–1936.
- Sivilotti, M. (1991). *Wiring Considerations in Analog VLSI Systems With Application To Field-Programmable Networks*. Ph.D. thesis, Comput. Sci. Div., California Inst. Technol., Pasadena, CA.
- Smith, P. L., and Ratcliff, R. (2004). Psychology and neurobiology of simple decisions. *Trends Neurosci.* 27, 161–168.
- Song, S., Miller, K., and Abbott, L. (2000). Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nat. Neurosci.* 3, 919–926.
- Spratling, M. W. (2008). Predictive coding as a model of biased competition in visual attention. *Vision Res.* 48, 1391–1408.

- Strukov, D. B., Snider, G. S., Stewart, D. R., and Williams, R. S. (2008). The missing memristor found. *Nature* 453, 80–83.
- Swadlow, H. A., and Gusev, A. G. (2002). Receptive-field construction in cortical inhibitory interneurons. *Nat. Neurosci.* 5, 403–404.
- Tallon-Baudry, C. (2009). The roles of gamma-band oscillatory synchrony in human visual cognition. *Front. Biosci.* 14, 321–332.
- Taylor, K., Mandon, S., Freiwald, W. A., and Kreiter, A. K. (2005). Coherent oscillatory activity in monkey area v4 predicts successful allocation of attention. *Cereb. Cortex* 15, 1424–1437.
- Thorpe, S., and Gautrais, J. (1998). “Rank order coding,” in *Computational Neuroscience: Trends in Research*, ed. J.M. Bower (New York: Plenum Press), 113–118.
- Thorpe, S., and Imbert, M. (1989). “Biological constraints on connectionist modeling,” in *Connectionism in Perspective*, Eds Rolf Pfeifer, F. Fogelman-Soulie, Luc Steels, and Z. Schreter (Amsterdam: Elsevier), 63–92.
- Tiesinga, P., Fellous, J.-M., and Sejnowski, T.J. (2008). Regulation of spike timing in visual cortical circuits. *Nat. Rev. Neurosci.* 9, 97–107.
- Treisman, A. M., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136.
- VanRullen, R., Gautrais, J., Delorme, A., and Thorpe, S. (1998). Face processing using one spike per neurone. *BioSystems* 48, 229–239.
- VanRullen, R., and Thorpe, S. (2001). Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural Comput.* 13, 1255–1283.
- VanRullen, R., and Thorpe, S. (2002). Surfing a spike wave down the ventral stream. *Vision Res.* 42, 2593–2615.
- Victor, J. D. (2000). How the brain uses time to represent and process visual information(1). *Brain Res.* 886, 33–46.
- Vinck, M., Lima, B., Womelsdorf, T., Oostenveld, R., Singer, W., Neuenschwander, S., and Fries, P. (2010). Gamma-phase shifting in awake monkey visual cortex. *J. Neurosci.* 30, 1250–1257.
- Wang, X.-J. (2002). Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* 36, 955–968.
- Weidenbacher, U., and Neumann, H. (2008). “Unsupervised learning of head pose through spike-timing dependent plasticity,” in *Perception in Multimodal Dialogue Systems, volume 5078/2008 of Lecture Notes in Computer Science*, Eds E. André, L. Dybkjær, W. Minker, H. Neumann, R. Pieraccini, and M. Weber (Berlin: Springer), 123–131.
- Wolfe, J. (1994). Guided search 2.0: a revised model of visual search. *Psychon. Bull. Rev.* 1, 202–238.
- Womelsdorf, T., Fries, P., Mitra, P. P., and Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature* 439, 733–736.
- Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., and Fries, P. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609–1612.
- Wong, K.-F., and Huk, A. C. (2008). Temporal dynamics underlying perceptual decision making: insights from the interplay between an attractor model and parietal neurophysiology. *Front. Neurosci.* 2:245–254. doi: 10.3389/neuro.01.028.2008
- Wong, K.-F., Huk, A. C., Shadlen, M. N., and Wang, X.-J. (2007). Neural circuit dynamics underlying accumulation of time-varying evidence during perceptual decision making. *Front. Comput. Neurosci.* 1:6. doi: 10.3389/neuro.10/006.2007
- Zamarreño-Ramos, C., Camuñas-Mesa, L., Perez-Carrasco, J. A., Masquelier, T., Serrano-Gotarredona, T., and Linares-Barranco, B. (2011). On spike-timing-dependent-plasticity, memristive devices, and building a self-learning visual cortex. *Front. Neurosci. Neuromorph. Eng.* 5:26. doi: 10.3389/fnins.2011.00026

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 17 March 2011; accepted: 20 June 2011; published online: 30 June 2011.

Citation: Masquelier T, Albantakis L and Deco G (2011) The timing of vision – how neural processing links to different temporal dynamics. *Front. Psychology* 2:151. doi: 10.3389/fpsyg.2011.00151

This article was submitted to *Frontiers in Perception Science*, a specialty of *Frontiers in Psychology*.

Copyright © 2011 Masquelier, Albantakis and Deco. This is an open-access article subject to a non-exclusive license between the authors and *Frontiers Media SA*, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other *Frontiers* conditions are complied with.