

The Coherence Theory: A Unifying Framework for Visual Processing

Raul C. Mureşan

*Nivis Research, Gh. Bilaşcu 85, Cluj-Napoca, Romania, Europe
raulmuresan@personal.ro, <http://www.raulmuresan.home.ro>*

Abstract

This paper presents the preliminary results of an extensive research that aims at finding the powerful mechanisms used by the brain to achieve complex visual processing tasks. We mainly focus on a completely new theory stating that neural processes detect, enhance and infer coherence of the stimulus and of the neural activity by the means of two mechanisms (correlation and decorrelation). We backup our theory by using the new paradigm of temporal coding and show that simple, yet very powerful effects like synchrony and asynchrony can be estimated in an ultra-rapid fashion. Also we state that our coherence principle can account for Gestalt effects, perceptual grouping, detection of illusory contours, contour integration, border ownership, visual attention and maybe object recognition.

Keywords: Coherence; Temporal Coding; Correlation; Decorrelation; Spiking neurons; Shunting inhibition.

1. Introduction

In recent years, numerous studies showed that our nervous system, in particular the visual system, is able to perform complex tasks in an ultra-rapid fashion [36]. Experimental evidence supports this statement as perceptual grouping and border-ownership effects emerge on a very short time scale (border-ownership effects can emerge in only 25 ms [46]). However, classical rate-coding is faced with serious problems in accounting for these findings, since the neurodynamics are rather slow.

Not only the speed of the visual system is amazing but also the complexity of tasks it achieves on such a short time scale. For example, perceptual grouping and figure ground processes occur at low levels in the visual system [21]. There is a great challenge in trying to explain such evidence. A strong mechanism should be found that could group together features and solve "the binding problem" as stated by von der Malsburg [40,41,42,43] and Milner [24]. "The Temporal Correlation Hypothesis" emerged as a strong candidate for solving these complex tasks [12]. The main idea behind this hypothesis is the synchronization of neural activity that would eventually lead to the formation of assemblies. The binding of features and

perceptual grouping is based on the neural assemblies that dynamically emerge during visual processing. However, the hypothesis has been widely criticized and the main argument for rejecting it has been the difficulty of replicating the experiments that favored it. We state in the present paper that the correlation hypothesis is indeed a good candidate for solving the complex tasks mentioned. We also show that experiments were divergent and correlation was frequently hard to find because experimentalists did not look in the right direction (correlation in spike timing rather than average firing rate) [12].

The extensive research in the recent years emphasized very complex visual processes like contour integration [22], illusory contour approximation, border-ownership coding [46], curvature estimation [28], just to name a few. Researchers strove to create models to account for these complex processes. One extensively studied phenomenon had been the contour integration [13,22]. However these models are very complex and dedicated synaptic connection must exist just to account for one effect. Eventually, neurons, as well as the associated synapses that are assigned to carry out the tasks mentioned would lead, if not to an exponential explosion, at least to a great redundancy when binding all processes together. It seems we need a unifying framework and a more general mechanism to achieve these functions.

We believe that grouping as well as binding, figure-ground segregation, border-ownership, contour integration, illusory contour perception, visual attention and the Gestalt effects are all the side-effects of a more general principle that we call the "Coherence Principle". The temporal correlation is only one of the mechanisms that contribute to the achievement and estimation of coherence in the nervous system.

2. The "Coherence Theory" framework

In the following sections we present the basics of coherence as a "driving force" in the visual nervous system. We briefly present the main idea of coherence that we define as the "Coherence Principle". The means of detecting and inferring coherence are presented in the form of

correlation, estimated by synchrony, and decorrelation, estimated using asynchrony. Finally we describe a very simple, yet biologically inspired, neural architecture that can account for some effects predicted by our theory.

2.1. The Coherence Principle

When trying to analyze a complex system like the brain, the usual method is to present the known input, then measure the output and finally infer a possible structure and internal representation. We deal here with a typical case of "black-box" system (even though modern techniques offer more and more details about the internal structure of the brain). Such an approach leads often to speculations and probably a general tendency of reconstructing the complexity in an incremental way (model some cells for contour integration, others for scale estimation, etc). But such an approach is subject to numerous pitfalls because some particular mechanism achieving a given function might conflict with other functions. The only solution to achieve all functions would be the multiplication of specialized units [30], which eventually leads to a high redundancy, unlikely in the nervous system.

A better way to address the analysis of such a system would start by explaining "**why**" and subsequently explaining "**how**". The Coherence Principle is exactly the "**why**" part of our theory.

So, what is the "Coherence Principle" about? To make a clear explanation we shall start by finding its origins. All around in nature we see, hear, understand, things that differ in some way from a random, hazardous organization. We separate entities like objects, words, ideas, based on their intrinsic properties, on the homogeneity or quite differently by their heterogeneous nature. However noise and hazard are heterogeneous aspects of the real world. But in turn, there are heterogeneous aspects of the world that make sense. Why? Because variation also follows some rules and we might say "their heterogeneity is homogenous".

To be more exact we come up with the example of space, speed and acceleration (this example is essential for our purpose and will be used frequently through this paper). For a simple space estimator, a uniformly moving object is incoherent. But for a speed estimator, the constant

movement is coherent. A constant variation in speed generates incoherence at the speed estimator level but coherence at the acceleration estimation level. Thus, coherence is not just a simple property that relates to equilibrium states or to the constancy of some parameters. It is much like a context dependent property. An Italian speaker might seem incoherent to a Japanese speaker. Roughly speaking, coherence is a very general principle that could be related to inference and rules. We expect that every coherent thing can be described by one or several rules and that the subjective aspect of the coherence lies in the ability of the subject to infer and use these rules (another Italian speaker knows the rules of Italian language). Inferring the rules is learning, while using them is understanding. We used the term "inferring" because learning in this framework means adapting internal coherence under external coherent supervision (which is quite different from memorizing).

We might question the link between neuroscience, in particular vision neuroscience, and the "Coherence Principle". In our opinion the link is essential. Because the brain is a machine for estimating coherence and generating coherence. When trying to understand how the brain works we have to keep in mind what it tries to achieve. Then, maybe, we can find out how it works.

Let us focus now on a narrower field of neuroscience: the visual system. At this level, there is a great deal of evidence that the cortical areas achieve coherence when a stimulus is presented. Border-ownership cells are always coherent, orientation selective cells are also coherent, they always implement the same function and there is a rule that can be inferred from their behavior.

What about correlation? What about synchrony? It has been proved that synchrony as a way of encoding correlation can account for a wide range of phenomena that otherwise might be unexplainable (perceptual grouping, feature linking, etc). In fact, correlation is only an aspect of coherence. Coherence is more general. For example two correlated neural firings indicate that there is coherence, if those firings can be grouped in the same context (they encode parts of the same object, for example). However, uncorrelated firings do not necessarily mean there is no coherence. Take for example the image of a square filled with a smooth gradient pattern. The

square is very salient, even though two neurons encoding different parts of the gradient, fire uncorrelated. But the neurons are coherent since they represent the same object. They "just don't know it". A subsequent level of neurons could detect the constancy in variation and bind together the previous level neurons. This is exactly the context problem formulated earlier (space and speed). As a conclusion, we state that correlation is only a mechanism for estimating or achieving coherence.

We know that hebbian update is also a way of learning correlation between a stimulation pattern and another pattern or internal representation. Thus, hebbian update ultimately means achieving coherence in the internal representation. Predictive coding [31] also fits this framework because error detection and correction is a way of achieving coherence between the internal representation and the external stimulus.

Hierarchical organization is not surprising at all if we accept the coherence principle. For a system that estimates coherence the perfect architecture is a hierarchic one. Each level encodes a coherent aspect of the world. The first levels of the visual system might encode and detect local features while the upper levels bind these features together in a coherent manner and encode object parts and finally entire objects. The whole process is driven by coherence.

2.2. Formulation of the "Coherence Principle"

Given the background presented earlier, we can apply these observations to formulate a simple yet powerful principle that tries to explain "why" the visual system acts like it does.

The coherence principle in the visual nervous system's context is the following: *the nervous system is a coherence detector that extracts, processes, enhances and infers coherent aspects of the real world by the means of two mechanisms: correlation and decorrelation.*

2.3. Consequences of the "Coherence Principle"

In the next sections we will describe the basic mechanisms that can be used in the "coherence" framework and show a simple architecture that achieves illusory contour detection with no special tuning for orientation (unlike the other models [13,22]).

Previous studies [26] proved that networks of neurons, laterally interconnected, tend to form assemblies based on synchronous firing. We stated that synchronism is a very powerful way of encoding local coherence of the stimulus. Also, because the synchronous firings would eventually dominate the global activity of a pulse-coupled network, one can even estimate the intrinsic properties of the object by only analyzing the global pulse signal [2,26].

Given these observations, two major questions come into our attention: How can visual synchrony emerge on a very short time scale? and How does the visual system estimate the degree of synchrony among a certain population of neurons? Answering these questions is in fact a "missing link" in neuroscience.

We were witnesses in the last decade of a strong conflict between "The Temporal Correlation Hypothesis" [11,24,40,41,42,43] and "Hierarchical Modelling" [1,9,15,30,32]. We will try to prove that synchrony as a way of estimating correlation and hierarchic organization coexist with no conflict at all and both temporal correlation and hierarchical models could be unified under a more general principle: the "Coherence Principle". Furthermore, we will show that spike-timing based neural codes are reliable estimators of synchrony with ultra-rapid dynamics.

2.4. Synchronization in 10 ms?

There is a great deal of evidence that neurons are able to select their synaptic inputs on the basis of timing. Let us consider just inhibition as a neural mechanism to achieve this. Shadlen and Newsome [33,34] treat inhibition as a balancing and stabilizing function in a neural pool. Although this is probably an important function of inhibition, it can also generate shunting and hyperpolarizing effects that shape the temporal structure of neural activity on a millisecond time scale [3,4,10,23]. Thus, shunting inhibition in general and fast shunting inhibition in particular could offer a strong mechanism for temporal coding rather than firing rate coding.

Since the development of neuroscience, the primary question behind all neural systems is how they encode information. At this time, there is widely accepted that neurons encode

information by the means of their firing rate. In this framework, we speak about stability and fixed points in the oscillatory processes. However, numerous recent studies seriously question such a neural code because it is unable to account for ultra-rapid brain processes. It is true however that the firing rate of a neuron is directly linked to the stimulus and excitatory as well as inhibitory afferent activity. But we can understand firing rate as a side effect of millisecond time scale processes. It is unlikely that the nervous system effectively "loses time" and waits for stabilization. We can put forward a simple yet unanswerable question to those that sustain firing rate coding: "If the neural system encodes information by means of stabilized, fixed point firing rates, then what happens during the stabilization process, before the system reaches its equilibrium point? Do we hallucinate during stimulus onset?". The answer to the second question is clearly no!

Taking into account the new evidence and the temporal codes previously studied by Thorpe [35,37] we can consider rank-order coding as a strong candidate for neural coding [38,39]. The neural mechanism behind such a coding is fast shunting inhibition. In principle, considering a target neuron and a pool of inhibitory interneurons, each time an afferent spike arrives, the inhibitory pool shunts down the efficacy of spike integration of the target neuron. Simple implementations of such a neural coding proved their power in dealing with large amounts of information [5,6,25,27].

Exactly describing the millisecond time scale processes is essential for our theory. Therefore we concentrate our efforts in showing how shunting inhibition shapes neural activity and how synchrony can fit in this framework.

Figure 1 shows the schematic organization that achieves the shunting inhibition effect.

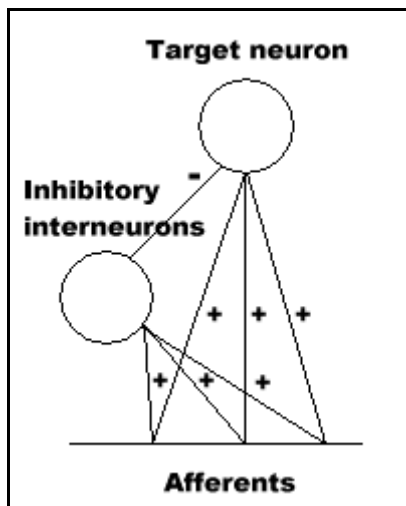


Fig. 1. Schematic organization of neurons and synapses that achieves the shunting inhibition effect. "+" denotes excitatory synapses whereas "-" denotes inhibitory synapses.

Each time an afferent neuron spikes, the target neuron is excited along with the associated inhibitory interneuron pool. When the next afferent spike arrives, the inhibitory shunting effect decreases the efficacy of spike integration. Thus, two side effects occur in this neural system. The first effect is the automatic normalization of activity. The target neuron doesn't enter a saturation state because shunting effectively decreases sensitivity to incoming spikes and, as afferent activity increases, the shunting effect also increases. The second, and maybe most important effect, is the capability of the target neuron to encode spike timing and ordering of the afferents. Arnaud Delorme stated in his PhD thesis that choosing a monothonical function for the value of synaptic weights of the target neuron leads eventually to the ability of the neuron to detect a specific afferent spike order. The enormous encoding power of this scheme has been outlined by S.J. Thorpe [35], showing that rank order codes can be used, in principle, to transmit up to $\log_2(N!)$ bits of information (where N is the number of neurons).

Looking more closely to the dynamics of the neural system presented, we can observe that if more afferents spike at the same time, the delay on the inhibitory interneuron path would lead to a decrease of synaptic efficacy only after the spikes have been integrated by the target neuron. Thus, spike-timing synchrony over the afferents generates a much stronger effect on the target neuron than asynchronous afferent spiking. Considering that all the excitatory synaptic weights of the target neuron are equal, the activity of the neuron (which we now model in terms of spike

timing) encodes the synchrony of afferent spiking. Synchrony of afferent neurons leads the target neuron into a strong activation state, determining it to fire earlier. This mechanism could be propagated up several levels and the homogeneity degree (in terms of spike timing) of two neural pools can be reliably estimated and propagated in a neural hierarchy.

The timing required for detecting spike synchrony is much reduced, because the events are generated on millisecond time scale. In our particular model (**Fig. 3**), one spike / level is enough and taking into account the biological neural dynamics, we can estimate a delay from stimulus onset to synchrony detection of no more than 10 or less milliseconds / level in the 60 Hz oscillatory band (taking into account the timing required for spike integration and emission of the first spike, when the neuron is strongly excited). Of course the total delay depends on the number of cortical levels.

We have not only shown how synchrony can be understood in the temporal coding framework but also provided a method of estimating synchronous firing in ultra-rapid manner. A final observation needs to be outlined. The strength of shunting inhibition determines the resolution of synchrony detection. A strong shunting inhibition would lead to sharp correlation detection (high degree of coherence) whereas a weak shunting inhibition relaxes the synchrony detection. Also, we have to mention that synchrony is a way of estimating correlation even though correlation and coherence are much general principles.

2.5. What about decorrelation and asynchrony?

There are two ways of understanding asynchrony and more generally decorrelation. Decorrelation either means that there is no link between two phenomena or it can be viewed as the braking of correlation due to some changes in the characteristics of the phenomena analyzed. In terms of visual processing, occlusion that determines the formation of "T" junctions induces decorrelation between the intersection area and the surrounding parts that belong to the objects.

Although researchers always emphasized the importance of synchrony and correlation, we believe that decorrelation and asynchrony is at least equally important. Our opinion has its

origins in the "Theory of Information" which tells us that variation generates much more information than constancy. Variation determines the breaking of correlation between synchronous areas and the asynchronous ones. Visual attention is an important example in this direction. We believe that visual attention, at least the bottom-up aspect of it, is guided by the quantity of information extractable from a natural image. Visual focus is attracted by highly conspicuous areas where the amount of information is higher than in the surrounding regions [16,17].

Estimating asynchrony is straightforward because it could be implemented by adding a neuron that receives strong inhibition from synchrony detectors and strong excitation from stimulus afferents. If the afferents are asynchronous the amount of inhibition is low and the final neuron's activity is high because of strong afferent excitation. Also we have to mention that the detection of synchrony is a faster process than asynchrony estimation (Raul C. Mureşan - unpublished data).

2.6. Fitting synchrony and asynchrony with coherence

In the preceding paragraphs we have stated that synchrony and asynchrony can be reliably estimated on a very short time scale. We have also outlined that coherence is a general and subjective principle and that correlation (estimated by means of synchrony) and decorrelation (estimated as asynchrony) can contribute to coherence detection and enhancement. We will try to show how. For explanation purpose, we choose a concrete example: the Gestalt principle for illusory contour detection (**Fig. 2**).

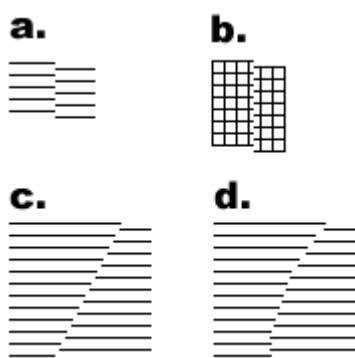


Fig. 2. Stimulus images for testing the illusory contour effect. (a) A clear vertical line is visible between the two gratings. (b) Even though synchrony breaks in the same fashion as with the horizontal lines, the line is no more visible due to the lack of coherence. (c) An oblique line is visible. (d) The separation area is not homogenous, the upper side being less coherent.

What are the origins of the Gestalt effect for illusory contour detection? The "Coherence Principle" offers us an essential clue. The nervous system tries to maintain and enhance coherence in the areas where coherence breaks. As we have emphasized, the correlated areas are also coherent (from a subject's point of view that associates those areas into the same context - object). Initially, uncorrelated areas are incoherent but the nervous system aims at finding out coherence in that area. Yes, a coherence in decorrelation. This situation can be compared to the space, distance and acceleration example. A simple way to infer coherence is the competition between correlation and decorrelation and also the reciprocal influence between the correlated areas and uncorrelated areas. Since an ideal noisy environment does not contain any correlation at all, there is no competition between correlation and decorrelation, nor reciprocal influence and thus no coherence. This means that the coherence principle also tells us why the neural system resists so remarkably to noise and why processing stages can eliminate noise.

Based on our previous studies [25,27] we implemented a simple neural architecture to test the way correlation and decorrelation can be estimated and how they can be used in the "Coherence Theory" framework. The basic idea behind the proof is that there are separate populations of neurons that estimate either synchrony or asynchrony in the stimulus. Synchrony builds up pretty much in the same way as outlined by Mureşan [26], by the means of lateral interaction. Predominantly, lateral interaction is inhibitory, the receptive field (RF) profile resembling with a center-surround structure (narrow center excitation and wide, silent surround suppression). Such a RF profile fits very well biological observations and maybe asymmetric suppression could also play an important part in achieving various functions and detecting coherence [44].

We used V1, V2 and V4 cells with no orientation preference (a further study should investigate how orientation columns interact). A sketch of the model is presented in **Fig 3**.

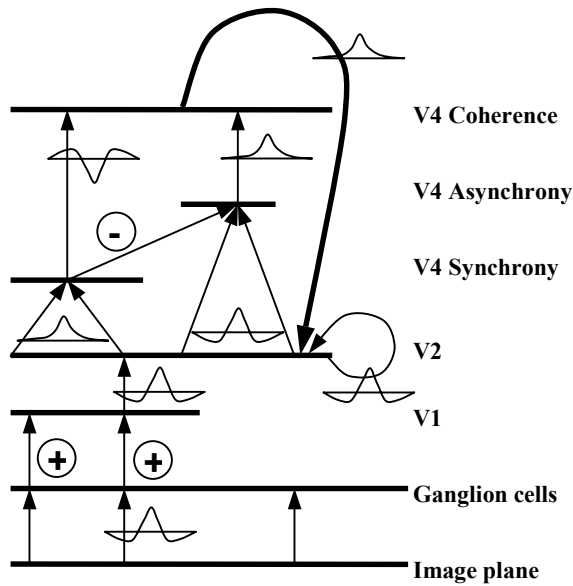


Fig. 3. The proof architecture for the illusory contour detection. The receptive field profiles are shown on connections between layers. "+" and "-" denote simple constant excitation or inhibition. No directional preference is included in the model.

We excluded orientation preference intentionally to prove that interaction between orientation columns is not necessary when trying to explain illusory contour detection, "T" junction detection and primitive perceptual grouping. However, orientation preference might help a lot to achieving these functions.

Basically, the architecture has six layers, each of them having a well-defined function. The V4 layers, in our model, are responsible with synchrony and asynchrony detection as well as coherence detection. Curvature, as shown by Pasupathy and Connor [28,29] plays an essential role in influencing V4 activity. Curvature estimation, in the "Coherence Theory" framework is basically similar to the estimation of acceleration in our previous example. We only need two levels that estimate the "correlation of decorrelation". Contour conformation [29] detection could be achieved by using an additional level. As seen in our model, we used only one level, for simplicity reasons. Further study is necessary to show how the described effects can be achieved in the "Coherence Theory" framework.

The highest level of the architecture detects asynchronous areas that are modulated by synchronous regions. In other words, the last map detects the way uncorrelated areas should be

modulated to regain correlation. This effort of the visual system eventually has the effect of discovering the correlation between decorrelated parts (if it exists). In **Fig. 2** we can see that while in **(a)** the decorrelated region possesses some correlation with respect to the surrounding region, in **(b)** this correlation cannot be found. Thus **(a)** generates a coherent line while **(b)** does not (subjective evaluation). **Fig. 4** presents the V2 firing pattern modulated by top-down coherence enhancement synapses, after just 15 ms (equivalent biological timing).

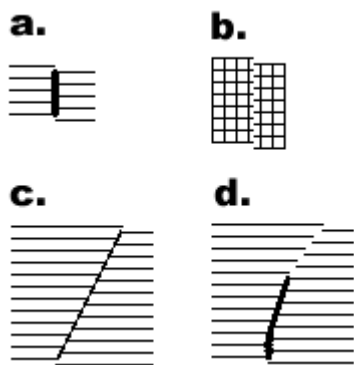


Fig. 4. V2 firings after 15 ms of simulation. Stimulus images used are presented in **Fig. 2**.

As stated before, coherence is subjective. Subjectivity can also derive from the size of receptive fields. One can even see that when zooming the image in **Fig. 2 (b)**, as texture leaves the fovea centralis area, coherence starts to emerge and the vertical illusory line becomes visible.

3. Discussion

We showed that synchrony and asynchrony, as measures of correlation and decorrelation are powerful mechanisms that can be used in the "Coherence Theory" framework. However, our study mentioned so far only qualitative aspects. A discussion over the quantitative aspects is necessary to reveal further aspects of our theory.

3.1. The amount of correlation and decorrelation

Personal observations led us to the conclusion that the strength of connection between the V4 synchrony, asynchrony and coherence levels has a great impact on the effect top-down modulation achieves. A stronger balance over synchronous synapses leads to the fragmentation

of the modulatory effect and eventually to perceptual grouping. In contrast, a balance towards asynchronous inputs leads to discontinuity detection. A certain compromise can be found so that the system acts like a coherence discovery process. **Fig. 5** shows different effects when adjusting the final map's synapse strengths with the synchronous and asynchronous V4 maps.

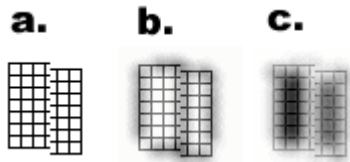


Fig. 5. Modulatory effects in V2 (black is the strongest). (a) Stimulus image. (b) Connections with asynchronous components dominate leading to discontinuity detection. (c) Connections with synchronous components dominate, leading to perceptual fragmentation (a key aspect for perceptual grouping).

Other effects are also remarkable when presenting an environment that contains occlusion (**Fig. 6**).

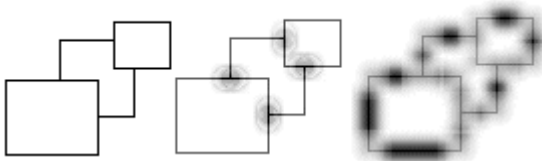


Fig. 6. Modulatory effects in V2 (black is the strongest). (a) Stimulus image. (b) "T" junction detection when asynchronous component dominates. (c) Connections with synchronous components dominate, leading to perceptual fragmentation.

3.2. How experimental evidence fits our theory

Although our model is quite simple and does not include many well-known functions of the visual system, it is able to reproduce in an ultra-rapid way, some important experimental and psychological evidence. However, we will discuss some aspects of visual processing that we expect would be explainable in the "Coherence Theory's" context.

We begin by explaining how Gestalt principles are explainable in "coherence" terms. It is well known that visual features tend to group together according to some properties of visual features [18,19,20,45] like: proximity, similarity, closure and collinearity, common fate, etc. Proximity as we emphasized earlier is a source of subjectiveness in coherence terms. Two

features might be grouped together if the distance between them related to the overall feature size doesn't exceed a certain threshold. The context of coherence, in this case, is the size of the features while the coherence aspect is the distance. Similarity is probably best explained in the coherence framework. When the visual system tries to enhance coherence it achieves grouping of the coherent areas of activity. The example in **Fig. 5 (c)** shows how the interior features of the two fragments are coherent and grouped together, although the intersection of lines generates high asynchrony. Collinearity is just a simple case of coherence, when orientation information is available, by pooling over simple V1 cells. Common fate, when analyzed more profoundly can be reduced to coherence. All Gestalt principles are ultimately reduced to and unified by the "Coherence Principle".

Contour integration can also be regarded as a side effect of the coherence enhancement process. But this time, it is not the result of some predefined synaptic connections and blind enhancement techniques that would eventually generate the increase in contour length. The process, in coherence context, is controlled by top-down modulation after the system detected coherence in some features of the image and tries to maximize and discover coherence where it the previous coherent context brakes up. To be clearer, let us consider a simple, isolated line. A blind butterfly synaptic mechanism would increase activity of neighboring collinear neurons. The classical concept is that such an undesired effect is eliminated by competition between contour integration and end-stopping processes. But it seems less likely that the visual system is subject to such a high waste of processing resources. The coherence theory is not subject to such a flaw since contour integration would occur only if coherence can be enhanced (the case when a collinear line is in the neighboring of the target line). End-stopping in this context is also the side effect of the "Coherence Principle".

Top down modulation can be responsible also for border ownership. The ultra-rapid mechanisms presented could offer reliable explanation to the 25 millisecond delay before border ownership manifests [46].

Bottom-up visual attention mechanisms are clearly very easy to fit in our theory. Asynchrony and decorrelation could be interpreted at different levels (remember the space, speed and acceleration example) and the attentional modulation can be controlled by a variety of clues. Maximal information principle is respected and reinforced in our framework as **Fig. 6 (b)** shows and the levels at which analysis is done represent the different contexts in which information is relevant (and maximal).

What about object recognition? As our previous studies show [26], the characteristics of the stimulus can be also encoded in other ways than by saliency matching. The evidence that the intermediate stages of processing, like V4 encode curvature, contour conformation, border ownership, etc, are explainable in terms of coherence analysis rather than component detection (like in the RBC model [1]). Further research will come to confirm our assumption that object recognition mechanisms are largely based on coherence analysis as a way of detecting and binding object characteristics. It is less likely that the visual system wastes resources by learning objects in an isomorphic fashion or creating complex objects by reconstructing them from parts and then matching against a memorized pattern. It is more likely that our visual system is capable of learning the characteristics of objects and our capability of abstraction would not be at random.

Although our theory might be highly controversial we are only at the beginning in exploring its full power. We have shown that simple yet very efficient mechanisms could be constructed in the coherence framework and that every aspect of visual processing can be thought of as different sides of the same unifying principle: the "Coherence Principle".

4. Conclusions

The present paper described briefly a novel theory that could explain experimental and psychological evidence of visual processing in the nervous system. The principle of coherence

as a driving process of neural systems is very general and can be applied to all kinds of neural systems (auditory and motor processing, etc).

We believe that coherence and its subjective aspects are the key to understand the flexibility of our brain. While most of the other models cannot account for different interpretations of a visual feature, restricting themselves to its detection and classification, the coherence framework provides a different approach. Using the two basic mechanisms of correlation and decorrelation, the nervous system can infer coherence under a wide variety of contexts. We believe that the hierarchic organization of the visual system is in fact a context hierarchy and that our nervous system analyses coherence both in morphological and spatial dimensions. As the processing converges to later stages, concepts (like lines, textures, curves) are integrated in more and more complex contexts (like objects, object groups, etc) under the supervision of coherence detection mechanisms.

The very different approach of our theory also comes from the way it deals with primary visual processing evidence like illusory contour detection, Gestalt principles, contour integration, etc. Unlike the recent trends to consider these effects as the "*modus operandi*" of the visual system, the coherence theory states that they are side-effects of a simple, yet unifying mechanism of detection, enhancement and inference of coherence. Furthermore, such a mechanism is naturally flexible and context sensitive, its power of expression, as we expect, is enormous.

The path for a wide range of research for our theory is now opened. Just to name a few aspects for further research, we outline:

- studying how the coherence framework can account for partially unexplained curvature estimation and contour conformation [28,29];

- showing the way border-ownership effects emerge in an ultra-rapid fashion (unexplained at this time) [46];

- understanding how contour integration emerges in context sensitive processes (not just simple geometrical mechanisms [13,22] that are very unlikely to account for many pshycological evidence);
- studying feature binding and figure-ground segregation with the coherence and context sensitive mechanisms;
- discovering a new way of recognizing objects based on their characteristics, infered using coherence mechanisms.

5. Experimental setup

As a proof of concept we used the simple architecture presented in **Fig. 3**. Each layer consists of a neural map having the same size as the input image. The input images used were 102x102 pixel grayscale bitmaps with 256 levels of gray.

5.1. Modeling the neurons

We used simple, spiking, integrate-and-fire neurons [5,6,25,27]. For simplicity reasons no leakage has been included in the model. The amount of current leak, in the short period the neuron's state is pooled (simulation of ultra-rapid visual tasks), can be neglected (no rate based coding is present). Neurons receive input spikes and increase their internal activity, until they reach a fixed threshold. The parameters of the neuron fit the biological evidence. The resting potential is -65 mV and the threshold -45 mV. No refractory period included. We use rank order coding as neural code and every neuron has 3 parameters: an activation level (**U**), a synaptic modulation (**M**) and an instantaneous sensitivity (**S**). The activation level determines the moment of spike generation while the synaptic modulation and the sensitivity, both model the fast shunting inhibition. The update rule states that for every incoming spike, the activation level of the neuron is updated with the synapse weight modulated by the instantaneous sensitivity of the neuron. The sensitivity is then decreased using the synaptic modulation factor [25] ($M \in (0..1)$).

$$U^{(t+1)} \leftarrow U^{(t)} + S^{(t)} \cdot W[\text{current_synapse}] \quad (1)$$

$$\mathbf{s}^{(t+1)} \leftarrow \mathbf{s}^{(t)} \cdot \mathbf{M} \quad (2)$$

The value of the synaptic modulation factor (\mathbf{M}) determines the strength of shunting inhibition (as \mathbf{M} approaches 0, shunting efficacy increases).

5.2. Modelling the synapses

The architecture implemented makes extensive use of the retinotopic organization of some parts of the visual system. Synapses are modeled as connection kernels, with variable size, that characterize the entire connection between two neural maps. Three types of kernels were used to implement the architecture in **Fig. 3**: center-surround kernels (**C-S**), gaussian kernels (**G**) and simple, constant kernels (**K**).

$$\mathbf{C-S}(x, y) = \text{onGain} \cdot e^{-\frac{(x^2+y^2)}{2\sigma_1}} + \text{offGain} \cdot e^{-\frac{(x^2+y^2)}{2\sigma_2}} \quad (3)$$

$$\mathbf{G}(x, y) = \text{Gain} \cdot e^{-\frac{(x^2+y^2)}{2\sigma}} \quad (4)$$

$$\mathbf{K}(x, y) = \text{Gain} \quad (5)$$

Usually the **onGain** and **offGain** have different signs, depending on the type of center-surround effect (center-ON-surround-OFF or center-OFF-surround-ON). Depending on the value of **Gain**, which stands for the synapse weight (5) or at least the magnitude of the gaussian (4), excitatory or inhibitory effects can be modeled.

5.3. Parameter values for simulation

Finally, we present the exact parameters used for modeling the system in **Fig. 3**.

Ganglion cells were modeled as 3x3 unit center-ON-surround-OFF kernel, with $\sigma_1 = 0.18$, $\sigma_2 = 0.53$, and normalized to 0.

Unless specified, modulation factors (\mathbf{M}) were set to 1 by default. Constant gains for connecting ganglion cells with V1 cells were 21 mV / spike.

V1-V2 connections are modeled by a center-surround kernel, with $\sigma_1 = 0.33$, $\sigma_2 = 1.66$, **onGain** = 65 mV, **offGain** = -90 mV. Lateral interactions in V2 are modeled by using the same kernel. Synaptic modulatory effects are weak in V2 ($\mathbf{M} = 0.99$) and have normalization purpose.

V4 synchrony detectors are connected with V2 by using a gaussian kernel, with $\sigma = 1.66$ and **Gain** = 70 mV. Moderated synaptic modulatory effects are included, **M** = 0.7.

V4 asynchrony detectors are connected to V2 by means of a center-surround kernel, with $\sigma_1 = 0.33$, $\sigma_2 = 1.66$, **onGain** = 40 mV, **offGain** = -100 mV. Also, there is a strong inhibitory connection with V4 synchrony detectors, modeled as a constant kernel, with hyperinhibition of -100 mV / spike. Modulatory effects at synaptic level induce very strong shunting inhibition, **M** = 0.2.

The final map is connected to the synchrony and asynchrony detectors by the means of a center-surround kernel, with $\sigma_1 = 1$, $\sigma_2 = 2.33$, **onGain** = -100 mV, **offGain** = 10 mV, and a gaussian kernel with $\sigma = 1$, **Gain** = 80 mV respectively.

Top-down modulation from the final map to V2 is achieved by using a center-surround kernel, with $\sigma_1 = 5$, $\sigma_2 = 5.66$, **onGain** = 300 mV, **offGain** = -10 mV.

Acknowledgements

We thank to SC. NIVIS SRL (www.nivis.com) for supporting this research. Also, we would like to thank Arnaud Delorme for his good advice in neuroscience and writing scientific papers.

References

1. I. Biederman, Recognition by components: A theory of human image understanding. *Psychol. Rev.* 94 (1987), 115–147.
2. E.L. Brown, B.M. Wilamowski, Spatial to Temporal Conversion of Images Using A Pulse-Coupled Neural Network, *IJCNN'99*, Washington, DC, July 10-16, 1999, #521, Session 7.1.
3. P. Bush, T. Sejnowski, Inhibition synchronizes sparsely connected cortical neurons within and between columns in realistic network models. *J. Comput. Neurosci.* 3 (1996), 91–110.
4. J.J. Chrobak, G. Buzsaki, Gamma oscillations in the entorhinal cortex of the freely behaving rat. *J. Neurosci.* 18 (1998), 388–398.
5. A. Delorme, L. Perrinet, S.J. Thorpe, Networks of Integrate-and-Fire Neurons using Rank Order Coding B: Spike Timing Dependant Plasticity and Emergence of Orientation Selectivity, *Neurocomputing* 38-40(1-4) (2001), 539-545.
6. A. Delorme, S.J. Thorpe, Face recognition using one spike per neuron: resistance to image degradation. *J. Neural Networks* 14 (2001), 795-803.

7. R. Eckhorn, H.J. Reitboeck, M. Arndt, Feature Linking via Synchronization among Distributed Assemblies: Simulations of Results from Cat Visual Cortex. *Neural Comp.* 2 (1990), 293-307.
8. R. Eckhorn, H.J. Reitboeck, M. Arndt, P. Dicke, Feature Linking via Stimulus-Evoked Oscillations: Experimental results from Cat Visual Cortex and Functional Implications from a Network Model. *Proc. Int JCNN'89* 1 (1989), 723-730.
9. D.J. Felleman, D.C. van Essen, Distributed hierarchical processing in the primate cerebral cortex, *Cereb. Cortex* 1 (1991), 1-47.
10. W.J. Freeman, Mass Action in the Nervous System, *New York: Academic Press*, (1975).
11. A.K. Gray, C.M. Konig, P. Engel, W. Singer, Stimulus-Dependant Neural Oscillations in the Cat Visual Cortex Exhibit Inter-Columnar Synchronization which Reflects Global Stimulus Properties. *Nature* 338 (1989), 334-337.
12. C.M. Gray, The Temporal Correlation Hypothesis of Visual Feature Integration: Still Alive and Well. *Neuron* 24 (1999), 31-47.
13. F. Heitger, R. von der Heydt, E. Peterhans, E. Rosenthaler, O. Kübler, Simulation of neural contour mechanisms: representing anomalous contours. *Image and Vision Computing* 16 (1998), 407-421.
14. J.J. Hopfield, Neural Networks and Physical Systems with Emergent Collective Computational Abilities. *Proc. Nat. Academy of Sciences USA* 79 (1982), 2554-2558.
15. D. Hubel, T. Wiesel, Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat, *J. Neurophysiol.* 28 (1965), 229-289.
16. L. Itti, C. Koch, A saliency-based search mechanism for overt and covert shifts of visual attention. *Vis. Res.* 40 (2000), 1489-1506.
17. L. Itti, C. Koch, Computational modelling of visual attention. *Nat. Rev. Neurosci.* 2 (2001), 1-10.
18. G. Kanisza, The Organization of Vision, Praeger, New York, (1979).
19. K. Koffka, Principles of Gestalt Psychology. Harcourt, Brace and World, New York, (1935).
20. W. Köhler, Gestalt Psychology. Bell and Sons, London, (1930).
21. V.A.F. Lamme, The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci.* 15 (1995), 1605-1615.
22. Z. Li, A Neural Model of Contour Integration in the Primary Visual Cortex, *Neural Comput.* 10 (1998), 903-40.
23. W.W. Lytton, T.J. Sejnowski, Simulations of cortical pyramidal neurons synchronized by inhibitory interneurons. *J. Neurophysiol.* 66 (1991), 1059-1079.
24. P. Milner, A model for visual shape recognition. *Psychol. Rev.* 81 (1974), 521-535.

25. R.C. Mureşan, Complex Object Recognition Using a Biologically Plausible Neural Model, by R. Muresan, *Proceedings of the 2nd WSEAS International Conference on Robotics, Distance Learning and Intelligent Communication Systems*, in press, (2002).
26. R.C. Mureşan, Pattern Recognition Using Pulse-Coupled Neural Networks and Discrete Fourier Transforms, *Neurocomp.*, in press, (2003).
27. R.C. Mureşan, Visual Scale Independence in a Network of Spiking Neurons, *ICONIP '02 Proceedings*, 4 (2002), 1739-1743.
28. A. Pasupathy, C.E. Connor, Responses to Contour Features in Macaque Area V4. *J. Neurophysiol* 82 (1999), 2490-2502.
29. A. Pasupathy, C.E. Connor, Shape Representation in Area V4: Position-Specific Tuning for Boundary Conformation. *J. Neurophysiol* 86 (2001), 2505-2519.
30. D. Perrett, M. Oram, Neurophysiology of shape processing. *Imaging Vis. Comput.* 11 (1993), 317-333.
31. R.P.N. Rao, D.H. Ballard, Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects, *Nat. Neurosci.* 2 (1999), 79-87.
32. M. Reisenhuber, T. Poggio, Are Cortical Models Really Bound by the "Binding Problem"? *Neuron* 24 (1999), 87-93.
33. M.N. Shadlen, W.T. Newsome, Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* 4 (1994), 569-579.
34. M.N. Shadlen, W.T. Newsome, The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J. Neurosci.* 18 (1998), 3870-3896.
35. S.J. Thorpe, A. Delorme, R. van Rullen, Spike-based Strategies for Rapid Processing. *Neural networks* 14 (2001), 715-725.
36. S. Thorpe, D. Fize, C. Marlot, Speed of processing in the human visual system, *Nature*, 381 (1996), 520-522.
37. S.J. Thorpe, J. Gautrais, Rank order coding, In J. Bower, *Computational neuroscience: Trends in research*. Plenum Press, New-York, 113-118, (1998).
38. R. van Rullen, S.J. Thorpe, Rate coding versus temporal order coding: what the retinal ganglion cells tell the visual cortex. *Neural. Comput.* 13 (2001), 1255-1283.
39. R. van Rullen, S.J. Thorpe, The time course of visual processing: from early perception to decision-making. *J. Cogn. Neurosci.* 13 (2001), 454-461.
40. C. von der Malsburg, Am I thinking assemblies? In *Proceedings of the Trieste Meeting on Brain Theory*, G. Palm and A. Aertsen, eds. (Springer: Berlin), (1986).
41. C. von der Malsburg, Nervous structures with dynamical links. *Ber. Bunsenges. Phys. Chem.* 89 (1985), 703-710.
42. C. von der Malsburg, The correlation theory of brain function. *MPI Biophysical Chemistry, Internal Report* 81-2 (1981).

43. C. von der Malsburg, The What and Why of Binding: The Modeler's Perspective, *Neuron*. 24 (1999), 95-104.
44. G.A. Walker, I. Ohzawa, R.D. Freeman, Asymmetric Suppression Outside the Classical Receptive Field of the Visual Cortex. *J. Neurosci.* 19 (1999), 10536-10553.
45. M. Wertheimer, Laws of organization in perceptual forms. In *A Source Book of Gestalt Psychology*, W.D. Ellis, ed. Routledge and Kegan Paul, London, (1955).
46. H. Zhou, H.S. Friedman, R. von der Heydt, Coding of Border Ownership in Monkey Visual Cortex, *J. Neurosci.* 20 (2000), 6594-6611.