

# RIVAGe Feedback during Visual Integration : towards a Generic Architecture

Jean Bullier, Rachid Deriche, Olivier Faugeras, Pascal Girard, Pierre Kornprobst, Théodore Papadopoulo, Thierry Viéville, Simon Thorpe, Denis Fize, Rudy Guyonneau

▶ **To cite this version:**

Jean Bullier, Rachid Deriche, Olivier Faugeras, Pascal Girard, Pierre Kornprobst, et al.. RIVAGe Feedback during Visual Integration : towards a Generic Architecture. [Research Report] RR-5451, INRIA. 2004, pp.74. inria-00070556

**HAL Id: inria-00070556**
**https://inria.hal.science/inria-00070556**

Submitted on 19 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

# *RIVAGe Feedback during Visual Integration : towards a Generic Architecture*

J. Bullier, R. Deriche, O. Faugeras, D. Fize, P. Girard, R. Guyonneau, P. Kornprobst, T. Papadopoulo, S. Thorpe, T. Viéville

**N° 5451**

Décembre 2004

Thème BIO




ISSN 0249-6399

INRIA
SOPHIA ANTIPOLIS

# RIVAGe Feedback during Visual Integration : towards a Generic Architecture

J. Bullier, R. Deriche, O. Faugeras, D. Fize, P. Girard, R. Guyonneau, P. Kornprobst, T. Papadopoulo, S. Thorpe, T. Viéville

Thème BIO — Systèmes biologiques
Projet Odyssee

Rapport de recherche n° 5451 — Décembre 2004 — 74 pages

**Abstract:** We consider the comparative study of visual process integration within either a biological system (i.e. the parieto-ventral and parieto-dorsal pathways of the cortical visual system in the primate) or an artificial system. Both systems deliver an estimation of
- [where], that is to say the motion and structure of the observed scene and of
- [what], i.e. the perceptual grouping and labeling of objects in the scene.

The long term goal is to elaborate a common theory about precise questions in both neurosciences and algorithms and their architecture in artificial vision, including computer vision applications.

Within this framework, the function and behavior of adaptive feedback mechanisms is a key point and on the leading edge of biological studies. The core of this idea is that visual processing is build around:
1) a first computational step allowing to pre-process the input information, provide initial estimates, generate hypotheses about which models to use, ?
2) a refinement step using iterative mechanisms of optimization of the visual perception.

Such mechanisms occur, sometimes implicitly, in artificial vision processes. They are mainly related to such problems as the combination of visual attributes, computed from different sources and then fused for “what” and “where” perceptual

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

tasks; the use of a-priori information, obtained from higher-level visual modules. These modules define models estimated from the data. These models are either given a-priori (e.g. rigidity, shape regularity, ..) or chosen thanks to object labeling obtained during the first computational step.

The present work is a theoretical study in three steps:
(a) a systematic analysis of existing results in neuro-science,
(b) an interpretation of these results from the viewpoint of the variational approach widely used in computer vision
(c) a specification of a simulation tool of parts of the visual cortex.

**Key-words:** Visual cortex, Feedback, Modelization, Simulation

# RIVAGe Rétroaction lors de l'Intégration Visuelle: vers une Architecture Générique.

**Résumé :** On s'intéresse à l'étude comparative de l'intégration des processus visuels au sein de deux systèmes : l'un biologique (précisément les voies pariéto-ventrales et pariéto-dorsales du système visuel cortical chez le primate) et l'autre artificiel. Ce dernier délivre des estimations: « where » du mouvement et de la structure de la scène observée et « what » du groupement perceptuel et de l'identification d'objets de la scène. Dans ce cadre, le rôle et le fonctionnement des mécanismes adaptatifs de rétroaction est au centre des recherches actuelles dans le domaine biologique.

Le but à long terme est d'aboutir à une théorie commune débouchant sur des questions précises dans le champ des neurosciences et des architectures et des algorithmes utilisables en vision artificielle et ses applications.

Au coeur de cette étude est l'idée que le traitement visuel effectue:
- une première vague de calculs qui permet de détourer l'information reçue, de fournir une estimation initiale, de faire des hypothèses sur les objets de la scène, etc..
- un raffinement itératif pour arriver à une perception effective de la scène.

De tels mécanismes adaptatifs sont omniprésents, parfois implicitement, dans les processus de vision artificielle où ils sont bien souvent traités dans le cadre du calcul des variations qui permet, outre de garantir que les problèmes sont mathématiquement bien posés, de répondre à des questions très importantes relatives à la fusion d'informations telles que:
(i) la combinaison d'attributs visuels calculés par des modules différents et utilisés ensuite pour les tâches de « what » et « where » et
(ii) l'utilisation d'information a priori issue de modules visuels de plus haut niveau.

Ce sont des « modèles » des données à traiter, ces modèles étant soit donnés a priori (e.g. rigidité, régularité de forme, etc..), soit issus de l'étiquetage d'objets reconnus par « la première vague » de calculs du système.

Il s'agit ici d'une étude théorique en trois étapes
(a) une analyse systématique de la littérature en neurosciences portant sur les mécanismes adaptatifs de rétroaction au sein du cortex visuel.
(b) une interprétation des résultats de cette analyse à la lumière des apports de l'approche variationnelle utilisée en vision arficielle pour traiter ces questions.

(c) une définition du cahier des charges d'un outil de simulation informatique de certains aspects du fonctionnement du cortex visuel.

**Mots-clés :** Cortex visuel, Rétroaction, Modélisation, Simulation

# Contents

**1 Introduction** ***J. Bullier, O. Faugeras et al.*** **6**

**2 Feedback in regularization mechanisms** ***T. Viéville, O.D. Faugeras et al.*** **18**

**3 Feedbacks in Hebbian learning mechanisms** ***R. Guyonneau*** **45**

**4 Functional properties of area V4 neurons** ***P. Girard*** **58**

**5 Conclusion and perspectives** ***S. Thorpe et al.*** **64**

# 1 Introduction *J. Bullier, O. Faugeras et al.*

The interactions between the communities of researchers studying the biological bases of vision and those interested in developing artificial vision systems have gone up and down during the last twenty years. In the years 1980's, there was a consensus between the two communities highlighted by the publication of the book Vision by David Marr [72]. In those days, processing of visual information was thought to be done in a series of steps, progressing from a local analysis of borders to 3D surface elements and then to the identification of 3D objects. This model was consistent with the hierarchy of cortical areas that has just been proposed by Maunsell and Van Essen [35]: borders were analyzed in area V1 and successively more elaborate degrees of processing were achieved in the higher order areas.

Twenty years later, the situation has changed considerably: biological visual systems can no longer be considered as purely feedforward systems consisting of sets of filters of progressively higher degree of sophistication as one penetrates the different layers of processing. In the field of artificial vision, numerous other models have been proposed to process information. These models work well within their context but a general framework for processing visual information is no longer in sight. The purpose of this work is to examine the possibilities of laying new foundations for interactions between the two communities.

Two of the major advances in the field of biological vision in recent years have been the realization that 1) information does not progress only bottom-up but that there is a very dense network of top-down connections and 2) that there are in the brain internal representations of the external world that are embodied in neuronal networks. This change in perspective has lead biologists to view the visual system no longer as processing information transferred by the retinal ganglion cells (the computer analyzing the output of the camera), but more as processing its own internal representations and checking whether these representations are consistent with the messages sent by the sensory neurons (checking the validity of the models). It is therefore of major importance to evaluate the role of the top-down (also called feedback) connections that are likely to be used to compare the representations of higher levels with those located at lower levels of processing.

The roles attributed to top-down influences are numerous but very few have been tested in detail: feedback connections are thought to be involved in directing attention, in memory retrieval, in comparing internal models with sensory inputs,

in combining global and local processing. Our intention is to examine the role of top-down influences not in isolation but within a general model of the brain such as that proposed by Friston [40]. This model stresses the point that feedback connections are essential here. Although such models are popular among theoreticians, they do not fit with the results of biological experiments: all studies of feedback connections so far [18] converge to conclude that feedback influences act to potentiate the responses of neurons at lower levels. One of the goals of our collaboration will therefore to develop models that incorporate the notions of internal models, non invertible relationship and interactions between top-down and bottom-up influences but are more compatible with results from biological vision, in the hope that this search might reveal some particularly interesting strategies used by biological systems to solve the problems posed by Friston.

Another recent important change in biological vision is the realization that information is not processed in a step-by-step hierarchical fashion but that some higher order areas, that contain sophisticated representations, are activated extremely early in the processing of visual information. In particular, it has been shown that the entire dorsal stream, including the parietal cortex and the frontal cortex are activated only a few milliseconds after area V1 which constitutes the initial processing stage of vision in primates [110]. This new perspective has led to a model in which a first-pass analysis is done by neurons in the dorsal stream and that this first-pass analysis is used, through feedback connections, to optimize the more detailed processing that occurs later (detailed shape, color). One of the goals of our collaboration is to determine whether such a model could be used in the framework of a general model of the visual system to improve and accelerate the processing of visual stimuli.

## Computational neuroscience: a first reading

We consider the topic “integrated model of visual processing”. Regarding the hierarchical organization of the visual cortical areas of the primate (see e.g. [116] for a comparative description of the organization of visual areas in macaque and human cerebral cortex) this corresponds to two main streams: the ventral and dorsal considered as the “what” (i.e. object recognition) and “where” (i.e. object localization and motion analysis) processing streams, respectively. Our goal is to contribute to models of biologically plausible neural computations (see e.g. [16] for details about the biological plausibility of linear and multiplicative computational steps, including

motion detection and short-term memory, while [129] focuses on the biologically plausible implementation of non-linear operation such as minimum computation or comparisons between inputs).

Following [18] we emphasize the fact that cortical visual processing requires information to be exchanged between neurons coding for distant regions in the visual field. Feedback connections from upper-layers are best candidates for such interactions because magnocellular layers of the LGN very rapidly project a "first-pass" information used to guide further processing.

More precisely, [3] demonstrates that the so called "horizontal connections" (i.e. within a cortical, e.g. retinotopic, map) are not fast enough to account for such a transfer, considering the known timings of information transfer [84].

(1) These cortical computations include stereo/binocular disparity. The work of [74] is a recent attempt to explain how complex cells can issue depth percepts for binocular but also monocular (i.e. da Vince stereopsis) cues, including induced effects from contrast changes (i.e. shape from illumination) as simulated in [49]. This also includes disparity tuning: [51] develops a model explaining how the LGN is involved in binocular disparity tuning, matching left and right images with the same contrast polarity but inducing feedback adaptation from signals of opposite polarities. In the present study, we would like to get a step further, considering V3 computations. More precisely, [1] shows the existence of disparity-selective columns including occlusions and proposes that V3 contributes to the processing of stereoscopic depth information and that the parietal areas to which it projects use this information for object depth and 3D shape analysis.

(2) A second aspect is perceptual grouping. For instance, [50] attempts to demonstrate how the known laminar architecture of the V1 and V2 areas of the visual cortex, assuming a functional role for this stratification, is involved in percepts generation, see also [83]. This includes pre-attentive/attentive aspects, as pointed out by [92] describing how the parvocellular stream of the visual cortex performs visual filtering, i.e. attention and perceptual grouping, using feed-forward, feedback and horizontal interactions. One key aspect is the figure / ground segmentation in the brain, as recently reviewed by [100] showing that shape perception depends critically on this cue, while its link to early-vision mechanisms is now relatively well understood [101] e.g. the role of junctions in surface completion and contour matching.

(3) Motion processing is also a key problem. It has been shown in [22] how responses to moving stimuli are derived from transient cell, speed-tuning cell responses of different sizes yielding visual motion perception and speed discrimination, involving both ventral and dorsal visual pathways. At a cognitive level, considering interpolation between prototypes of gestures, [44] attempts to interpret how the cortical "what" and "where" pathways of the visual cortex are involved in complex movements patterns. For instance, a complex spatio-temporal pattern is efficiently represented as a combination of prototypes whose coefficients are estimated using variational estimation methods [43]. Well-known early vision visual cues (e.g. the Koenderink def value easily computed by an affine modelization of the retinal motion field [21]) are other relevant candidates for such a parameterization.

All three previous aspects of visual perception are easily formalized using the "generative model" approach developed in [27] and reviewed in [40]. This allows to relate what is actually known in statistical modelization with existing biological models. A general framework for the role of feedback connections is proposed in these publications. A step further, the link with variational approaches [120] has been already sketched out. Regarding the role of feedback connections, [93] describes a model of visual processing in which feedback connections from a higher- to a lower- order visual cortical area carry predictions of lower-level neural activities, whereas the feed-forward connections carry the residual errors between the predictions and the actual lower-level activities. In the scope of this approach receptive fields are Gabor-like filters with a sigmoid profile output, weights being optimized by a 1st order gradient optimization of a likehood criterion. This is easily generalized to more plausible operators. But the relation with sensory-motor strategies, as discussed in [95] is very important: these authors point out the fact that visual cognition highly depends on the gaze orientation mechanisms and analyze how the appearance information contained in the image is converted into a target position, using saliency maps and separating targeting "what" process and the localization "where" process. This is in close agreement with the dorsal (where) and ventral (what) visual cortical pathways functionalities.

As a conclusion, this first reading of the literature indicates that there are many but rather different frameworks to model the ventral/dorsal visual pathways functionalities, while these functionalities correspond to recent computer vision mechanisms, now formalized in a unifying "variational" framework (e.g. [52, 86] for

recent introductions). It is thus very tempting to investigate whether this unifying framework could also be relevant in modeling biological vision. This is the goal of this project.

## A unifying theoretical framework to analyze visual processes

The adaptive mechanisms described previously are always present -most of the time implicitly- in artificial visual processes. The so-called "variational approach" is a unifying theoretical framework to design and implement this formalism. At a phenomenological level, this framework:

- defines the estimation problem in terms of the optimization of a criterion. This criterion is usually built from two terms:
  (i) One term is related to the data input (e.g. looking for a solution as compatible as possible with this input)
  (ii) One term is related to the a-priori information (i.e. looking for a solution corresponding to plausible properties allowing to regularize the solution)

- implements this global optimization using a local iterative scheme of a parametric or non-parametric map. This scheme arises from the partial differential equations (PDEs) that must be verified by the solution. The architecture of this implementation may correspond to what is processed in a cortical [maxi]column: here is the fundamental motivation to apply this formalism to neuronal computations, since we assume that it provides an alternative to usual artificial "neural-nets" as a basic generic biologically plausible estimation process.

More precisely, such a mechanism allows to model how a visual system, as a whole, solves perceptual tasks not only at a "pixel" level. A key aspect is that, under this assumption, there is a direct link between PDE parameterization and spatio-temporal maps of cortical activity. Therefore we plan to find biological networks of neurons which carry out some of the PDEs computations used in algorithmic vision and conversely, starting from neuro-physiological data, to try to define pertinent PDEs acting on visual inputs.

This should allow the development of not only anatomical but also functional models of activity in cortical areas during a visual task. At a theoretical level, the

variational paradigm is interesting because:
- it provides concise mathematical models of many computer vision problems,
- it allows to study in detail the problems of the existence and the uniqueness of solutions of the resulting equations, including wellposedness and
- to design correct and usually efficient algorithms to calculate approximate solutions of these equations.

Indeed, we are aware that brain modelization has been, over the history, deeply related to "fashionable" metaphors in relation with the different technologies, the "computational" metaphor being the basis of what is called cognitive science [26]. Here, following e.g. [16], we simply consider the biological plausibility of linear and multiplicative computational steps, including motion detection and short-term memory, while [129] details the biologically plausible implementation of the maximum (or minimum, i.e. comparisons between inputs) and follow [40] regarding this concept of biological plausibility. As made explicit in [47] we also will consider not "neuronal networks" but true biological networks of neurons.

## Appendix: hierarchical organization of the visual cortical areas

Following [116] let us very briefly review the key-words of the organization of visual areas in macaque and human cerebral cortex: two main streams the ventral[1] and dorsal[2]

From a "cognitive" point of view they are respectively considered as the "*what*" (i.e. object recognition[3]) and "*where*" (i.e. object localization and motion analysis[4]) processing streams.

The main cortical areas are:

| | |
|---|---|
| V1/V2 | primary visual areas (local 2D computations: edge, color, texture, raw-motion, ..) |
| V3 | visual area (mainly related to binocular disparity/depth computation) |
| V4 | "associative" visual area |
| V5 = MT / MST | visual areas mainly involved in motion computation |
| IT | inferior temporal areas (TE / TEO / STS) |
| VP / VIP / 7a \| PO | visual parietal area \| parieto–occipital |
| FEF | frontal eye field |

[1]**The "ventral" pathway:** Prior to the inferior temporal cortical area, is the so called parietal/ventral pathway (sometimes improperly called "parvocellular" pathway), neurons in the inter-blobs of V1 project to the pale stripes of V2. This pale stripes of V2 project to the inferior temporal cortex. Other feed-forward pathways include the V4 visual area, see [18] for general review. This pathway is composed of feature detectors (simple, complex and hyper-complex cells) (e.g. [59] for an introduction). Neurons in this pathway show a low sensitivity to contrast, high spatial resolution, and low temporal resolution or sustained responses to visual stimuli. See for instance [33], Chap 2 for a discussion.

[2]**The "dorsal" pathway:** Here a second visual cortical pathway, parietal and dorsal, in which the neurons in layer 4B of V1 project to the thick stripes of V2 is considered. Area V2 then projects to V3, V5 (or MT, middle-temporal cortex), and MST (median superior temporal cortex) This pathway is mainly an extension of the magnocellular pathway, but not only.

[3]**About IT:** The inferior temporal cortex is though to consist of three parts: The TEO (the occipital division of the intra-temporal cortex), the TE (the median division), and the STS (superior temporal sulcus).The TEO is used for making discriminations between 2-D patterns which differ in form, color, size, orientation, or brightness. The TE is used for recognition of 3-D objects. Both the TE and STS are thought to be used in facial recognition and in the recognition of familiar objects. The STS may be the place in which the feature maps of objects (which contain separate information about each primitive of an object, such as color, orientation, or form) become object representations.

[4]**About V5:** Simplifying the situation: cells in V5 are particularly sensitive to small moving objects or the moving edge of large objects; cells in dorsal MST respond to the movement of large scenes such as caused by head movements; cells in ventral MST respond to the movement of small objects against their background. See for instance [33], Chapter 10 for a discussion.

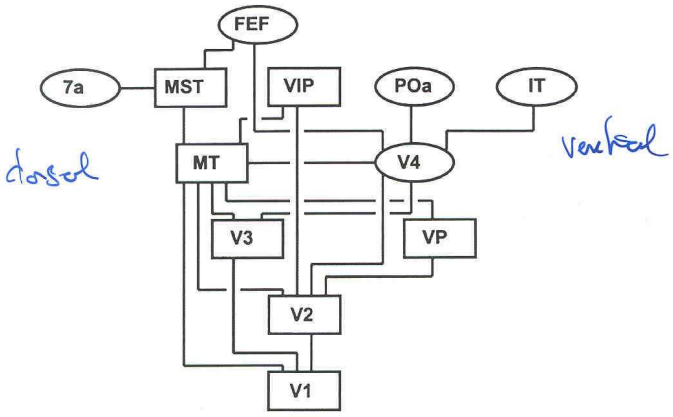



Figure 1: Hierarchical organization of the visual cortical areas, from [18], (see [19] for an introduction). Anatomical localization is given in Fig. 2.

**Feedback in the visual cortex**

If [18] points out the fact that cortical visual processing requires information to be exchanged between neurons coding for distant regions in the visual field, yielding distant neurons within a retinotopic cortical map, the key point is that feedback connections from upper-layers are best candidates for such interactions because mag-

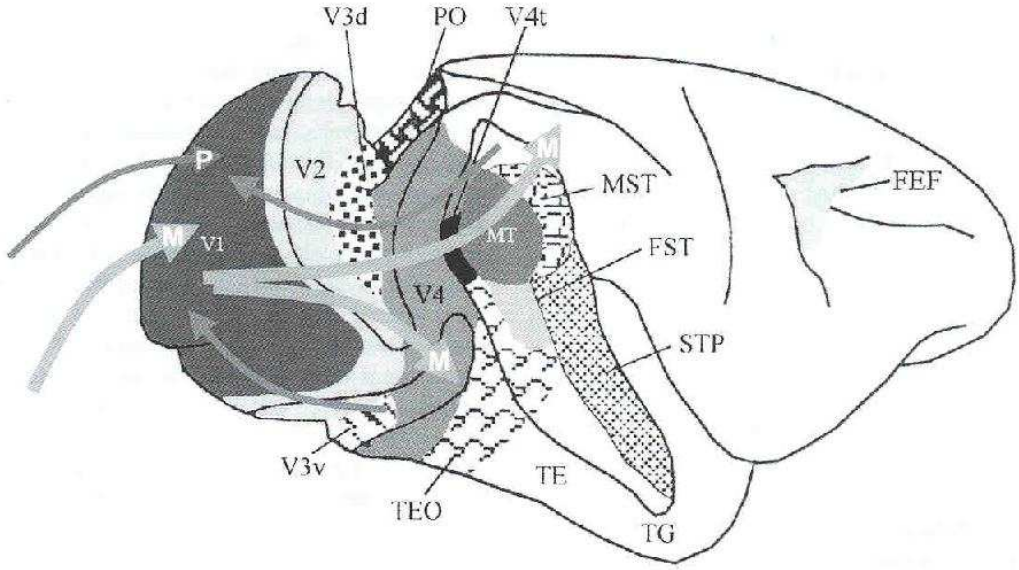



Figure 2: Anatomical localization of visual cortical areas which connectivity is schematized in Fig. 1, (from [18])

nocellular[5] layers of the LGN[6] very rapidly project a "first-pass" information used to guide[7] further processing.

[5]**About (magno/parvo)cellular streams:** There are two classes of cells from the retina and LGN: magnocellular, and parvocellular. These two cell types are contained in different parts of the LGN, and they have different response properties: (i) magnocellular cells receptive fi elds are 2-3 times larger than parvocellular cells receptive fi elds. Parvocellular cells have better acuity and resolution than magnocellular cells which have better sensitivity; magnocellular cells respond well to moving stimuli, whereas parvocellular cells do not; parvocellular cells respond well to color stimuli, whereas magnocellular cells do not.
The magnocellular pathway (phylogenetically older than the parvocellular one) continues the processing of visual detail leading to the perception of shape in area V3 and movement in areas V5 and MST. It has less synaptic relays than the parvocellular pathway, but is faster.

[6]**About LGN:** The Lateral Geniculate Nucleus (LGN) is the neuronal relay from the retina to the cortical visual input; this structure is organized in a laminar architecture with a binocular representation.

[7]**About the use of feedback:** At least two functional reasons for such feedback may be invoked:

[3] demonstrates that so called “horizontal connections” (i.e. within a cortical (e.g., retinotopic) map) are not fast enough to account for such a transfer, considering usual timing of information transfer [84], see Fig. 4 whereas feedback connections are, see Fig. 3.

.

**where to process:** simple cues allow to decide in which part of the retinal map visual processing has to occur:
- for a static image a rough but fast edge detector may be used to determine which contrasted areas (over a given threshold) have to be analyzed, or a large scale orientation detector (i.e. blurred by smoothing, but eliminating noise) may help anisotropic regularization processes to further process some visual cues,
- regarding motion perception, for a rotationally stabilized (though oculomotor reflexes such as vestibulo-ocular or opto-kinetic reflexes) image sequence, non-stationary areas correspond either to (e.g. [122]) mobile objects (i.e. prey/predators) or (because of parallax) close objects (thus in the grasping space or likely to interact with the subject); when the surfaces are not too highly textured, these are in any case reasonable targets to focus attention on;

**what to process:** very fast object recognition, as experimented and modelized by, e.g. [111, 110], allows a “model-based” processing of the visual information with the possibility to choose among memorized previous processing modes, configurations of parameters tuned with respect to this first recognition. This is also related to “focus of attention”, including consciousness [17]

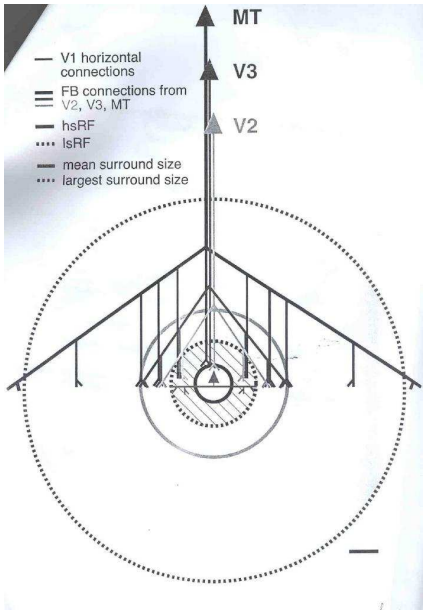



Figure 3: Summary diagram showing the spatial scales of horizontal versus feedback connections (from [3])

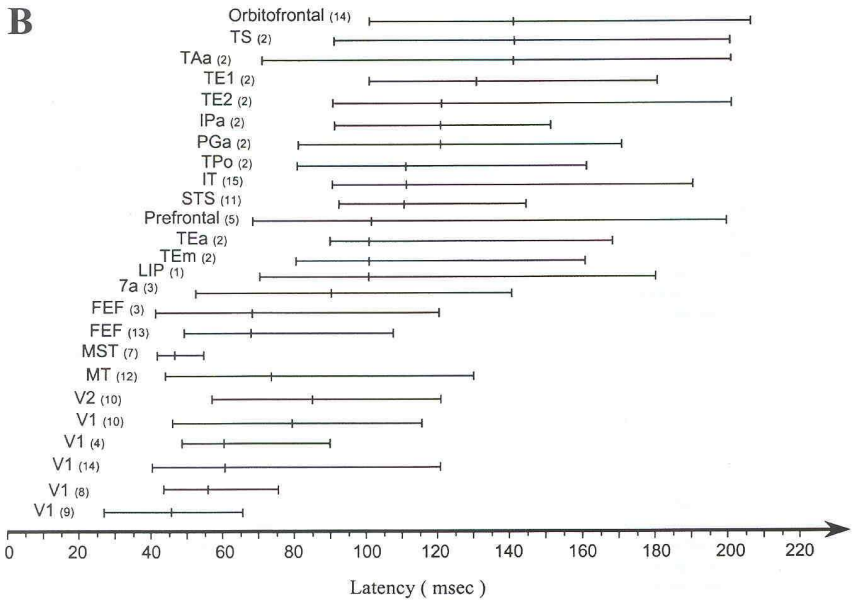



Figure 4: Latencies of visual responses in different cortical areas, from [18]

# 2 Feedback in regularization mechanisms *T. Viéville, O.D. Faugeras et al.*

## Cortical maps

Perceptual processes, in computer or biological vision, require the computation of "maps" of quantitative values. The retinal image itself is a "retinotopic map": for each cell of the retina or each pixel of the image, there is a value corresponding to the image intensity at this location. This is a vector value for color images. One step further, in early-vision, the retinal image contrast is computed at each location, allowing to detect image edges related to boundaries between image areas. Such maps encode not only the contrast magnitude, but several other cues: contrast orientation related to edge orientation, shape curvature, binocular disparity related to the visual depth, color cues, temporal disparity between two consecutive images in relation with visual motion detection, etc.. There are such detectors in both artificial visual systems (see e.g. [38] for a general introduction) and in the brain neuronal structures involved in vision perception (see e.g. [59] for a classical overview). Such maps are not only parameterized by retinotopic locations, but also 3D locations, or even by other parameters such as orientation, retinal velocity, etc..

Both systems have to solve the same perceptual tasks and very likely make the same kind of hypotheses about the observed surroundings: they share the same models. It is thus a relevant challenge to elaborate common theorical tools in both fields, considering neuro-science models and artificial vision algorithms.

## Forward/Backward connections

In [40] the visual cortex is considered as a hierarchy of cortical levels with reciprocal extrinsic cortico-cortical connections among the constituent cortical areas [39]. The notion of a hierarchy depends upon a distinction between forward and backward extrinsic connections. This distinction rests upon different laminar specificity [98, 102].

In this context, *forward connections are driving and backward connections are modulatory*, as suggested by reversible inactivation [103, 45] and functional neuroimaging [15]. In some paradigms, backward connections may also be functionally

driving the signal if their latencies are shorter than forward connections between distant cortical maps [110]. This fact is compatible with the present formalism.

More precisely, *forward* connections are concerned with the promulgation and segregation of sensory information, consistent with: (i) their sparse axonal bifurcation; (ii) patchy axonal terminations; and (iii) topographic projections. On the other hand, *backward* connections are considered to have a role in mediating contextual effects and in the co-ordination of processing channels, consistent with: (i) their frequent bifurcation; (ii) diffuse axonal terminations; and (iii) non-topographic projections [102] (iv) slow time-constants.
Backward connections are more numerous and transcend more levels, e.g. the ratio of forward efferent connections to backward afferents in the lateral geniculate is about 1:10/20. Another example: there are backward connections from TE and TEO to V1 but no monosynaptic connections from V1 to TE or TEO [102]. Backward connections are more divergent than forward connections [131], one point in a given cortical area will connect to a region 5-8mm in diameter in another. For instance: the divergence region of a point in V5 (i.e. the region receiving backward afferents from V5) may include thick and inter-stripes in V2, whereas its convergence region (i.e. the region providing forward afferents to V5) is limited to the thick stripes [131]. They are faster than direct lateral connections [84]. As a consequence, forward connections preserve retinotopy in visual areas, whereas backward connections do not [18].

This is summarized by [40] as reported in table 1.

Following [18], we must emphasize the fact that cortical visual processing requires information to be exchanged between neurons coding for distant regions in the visual field. Feedback connections from upper-layers are best candidates for such interactions because magnocellular layers of the LGN very rapidly project a "first-pass" information used to guide further processing in IT. One step further, [3] demonstrate that the so called "horizontal connections" (i.e. within a cortical, e.g. retinotopic, map) are not fast enough to account for such a transfer, considering the known timings of information transfer [84].

More generally [40], feedback connections are essential when the relationship between retinal inputs and the stimuli that generate them is not invertible. This is the case for practically all situations of vision outside the laboratory because of the interactions between the stimuli and the importance of the context for a given stimulus [40]. In the work of Rao and Ballard [93], the way by which internal representa-

| *Forward connections* | *Backward connections* |
|---|---|
| Sparse axonal bifurcations | Abundant axonal bifurcation |
| Topographically organized | Diffuse topography |
| Originate in supragranular layers | Originate in bilaminar/infragranular layers |
| Terminate largely in layer 4 | Terminate predominantly in supragranular layers |
| Postsynaptic effects through fast | Modulatory afferents activate slow |
| AMPA (1.3-2.4 ms decay) and | (50 ms decay) voltage-sensitive NMDA receptors |
| GABAA (6 ms decay) receptors | |
| *Forward connections are concerned with the promulgation and segregation of sensory information*, consistent with: (i) their sparse axonal bifurcation; (ii) patchy axonal terminations; and (iii) topographic projections. | *Backward connections are considered to have a role in mediating contextual effects and in the co-ordination of processing channels*, consistent with: (i) their frequent bifurcation; (ii) diffuse axonal terminations; and (iii) non-topographic projections [102] (iv) slow time-constants. |

Table 1: Forward/Backward connections connections main properties, from [40]. Postsynaptic/modulatory effect is a conjecture, so is assumptions on synapses mechanisms.

tions and incoming signals are combined is mainly subtractive : only differences are transmitted to higher levels. However, all studies of feedback connections so far [18] converge to conclude that feedback influences act to potentiate the responses of neurons at lower levels. In our framework, this occurs by tuning the cortical map parameters, formalized now.

The question here is to analyze such interactions when considering regularization mechanisms, i.e. study how our model may be used considering several interacting cortical maps.

### The role of feedbacks: a few assumptions

The roles attributed to top-down influences are numerous [18]: feedback connections are thought to be involved in directing attention, in memory retrieval, in comparing internal models with sensory inputs, in combining global and local processing. For instance:

**Where to process (low-level):** simple cues allow to decide in which part of the retinal map visual processing has to occur:
- a rough but fast edge detector may be used to determine which contrasted areas have to analyzed [93],
- large scale orientation detector (smoothing, eliminating noise) may help to tune further processing of visual cues (e.g. figure/background segmentation [100]),
- low-level but efficient focus of attention towards *close*, *mobile* or *textured* objects is derived from fast rotationally stabilized motion perception [122].

More generally, our capacity to dynamically focus our visual perception on a relevant part of the visual scene and to quickly react to its modification, may be related to feedback mechanism. One aspect of this capability (i.e. focus of attention) has been studied in computer vision (e.g. [122, 93]) although limited to low-level perception.

**What to process (high-level):** very fast object recognition, as experimented and modelled by, e.g. [111, 110], allows a "model-based" processing of the visual information with the possibility to choose among memorized previous processing modes, configurations of parameters tuned with respect to this first recognition. This is also related to "focus of attention", including consciousness [17].

**Holistic perception** : naive introspection shows that we are able to almost instantaneously perform the grouping of local and even random tokens or attributes (e.g. the dalmatian picture), yielding a global object perception [83]. This includes "hallucinations" of 2D or 3D shapes from partial information (e.g. the Kanizsa triangle) [50]. Holistic perception may be related to fast-brain object labeling: feedbacks from what has been detected by this "1st stage" may be the key feature of holistic perception [18].

**Opportunism** : One principle which seems to control all visual perception is that "the end justifies the means". This means that in order to elaborate our percepts, our brain has an extraordinary capacity to combine several attributes (color, texture, motion, stereo, etc..), but always choosing those well adapted to a given context or to a given task [93]. This occurs dynamically and without any conscious effort. Feedbacks in the visual cortex seems to be used

to select the relevant attributes, given a task or context [40], as soon as this state has been either detected or input to the system by higher "layers" of the cortex or obtained from a priori information.

**The Partial Differential Equation (PDE) approach**

In computer vision, a relevant and efficient theory is now available regarding the definition and computation of such maps of quantitative values (see e.g. [2] for a didactic introduction to the "axiomatization" of this part of computer vision). This formalism not only provides a clear indication of "what is to be done" (i.e. requirements in order to have coherent and consistent definitions) but also of "how to do it" since the theory is effective in the sense that efficient implementations may be derived (see e.g. [4] for a recent treatise on this subject). The "what is to be done" level is formalized in terms of a criterion to minimize and the "how to do it" level is related to the so-called Euler-Lagrange equations which allow the improvement of an initial guess of the solution to get closer to the optimal solution. At a technical level, as revisited in this paper, these quantities are efficiently computed using implementations of partial differential equations which define regularization processes allowing to obtain well-defined estimations of these quantities. This powerful methodology is also very general in the sense that a large variety of computational problems are solvable within this framework (see, e.g., [31, 30] for a review).

A step further, PDE calculations regularize their input, i.e. provide a well-defined and stable estimation even in the case of noisy or partially defined data. Thanks to this property these algorithms have performances closer to those of biological systems than those of usual artificial neuronal networks.

**Building a link between PDE and NN**

In biological vision modeling, the situation is dominated by the idea that "cortical maps code vector parametric quantities are computed by networks of neurons" (e.g. [47] for a recent review). Such models allow to analyze in details the biological substrate of such computations (e.g. [33]), in other words to propose interesting hypotheses about "how it is done". Although these approaches already bring a good understanding of what is processed in important areas of the brain, and allow the

development of computer simulations [48],there is no clear link with the computer vision algorithms mentioned before.

Establishing such a link would indeed entail two major advantages, by providing 1) a methodology to relate the "what is to be done" and the "how to do it" levels and 2) a very general computational framework for such calculations. This link would furthermore provide a "common framework" between computational and biological vision in order to develop common models and biologically plausible algorithms. However, building such a connection is not obvious because there is yet no clear understanding of how such complex differential nonlinear methods could correspond to neuronal mechanisms although the question has already been raised [34, 90, 66].

**Implementing PDE as NN**

In order to build this link, we must have a set of tools for the implementation of the set of PDEs which are commonly met in Computer Vision with networks of neuronal units. It will appear clearly in the sequel that the difficult point is the anisotropic diffusion operators implementation. Concretely, such mechanisms must integrate the information in a bounded neighborhood so that the cooperation between these units allows a global computation of the quantitative map. Furthermore, to be biologically plausible, such mechanisms must be based on simple local linear feedback [33, 47]. In other words, we must provide a "particle implementation" of such numerical computations. Such a method, used in fluid dynamics computation has been introduced by Leonard [68] followed by Raviat and Mas-Gallic [96] and developed by Degond and Mas-Gallic [28]. It is based on an integral approximation of the diffusion operator used in the regularization mechanism.

The goal of this work is to describe how cthe set of PDEs which are commonly met in Computer Vision can be implemented using biologically plausible networks of neuronal units.

## From regularization to neuronal networks.

As reviewed in the appendix, the regularization mechanisms discussed here are implemented via a PDE of the form:

$$\frac{\partial \mathbf{h}_t}{\partial t} = \left[ \mathbf{\Lambda}\, \mathbf{h}_t - \mathbf{\Lambda}\, \bar{\mathbf{h}}_t - \Delta_{\mathbf{L}} \mathbf{h}_t \right] \tag{1}$$

the key problem being the implementation of the diffusion operator $\Delta_{\mathbf{L}}$ (the other terms $\mathbf{\Lambda}\ \mathbf{h}_t - \mathbf{\Lambda}\ \bar{\mathbf{h}}_t$ corresponds to a simple punctual linear operation).

**Integral form of the diffusion operators.** Let us expand the righthand side of (12) as follows[8]

$$\sum_{j=1}^{m} \text{div}(\mathbf{L}_{ij} \nabla h^j) = \sum_{j=1}^{m} \text{div}(\mathbf{L}_{ij}) \cdot \nabla h^j + \text{Tr}(\mathbf{L}_{ij} \mathbf{H}_j) \tag{2}$$

The divergence of a square matrix $\mathbf{A}$ of size $n$ is the vector of size $n$ whose $i$th component is the divergence of the $i$th row vector of $\mathbf{A}$.

Considering the Dirac distribution $\delta(\mathbf{x})$ (see [106] for an introduction) we may formally write

$$\nabla h^j = \nabla \delta * h^j \quad \mathbf{H}_j = \mathbf{H}_\delta * h^j \ ,$$

where $*$ represents the convolution.

From this follows

$$\text{div}(\mathbf{L}_{ij}) \cdot \nabla h^j = \text{div}(\mathbf{L}_{ij}) \cdot (\nabla \delta * h^j) = (\text{div}(\mathbf{L}_{ij}) \cdot \nabla \delta) * h^j,$$

and

$$\text{Tr}(\mathbf{L}_{ij} \mathbf{H}_j) = \text{Tr}(\mathbf{L}_{ij} \mathbf{H}_\delta * h^j) = \text{Tr}(\mathbf{L}_{ij} \mathbf{H}_\delta) * h^j.$$

Note that $\text{div}(\mathbf{L}_{ij}) \cdot \nabla \delta$ and $\text{Tr}(\mathbf{L}_{ij} \mathbf{H}_\delta)$ are first and second order differential operators:

$$\text{div}(\mathbf{L}_{ij}) \cdot \nabla \delta = \sum_k [\text{div}(\mathbf{L}_{ij})]^k \frac{\partial \delta}{\partial x^k},$$

and

$$\text{Tr}(\mathbf{L}_{ij} \mathbf{H}_\delta) = \sum_{k,l} L_{ij}^{kl} \frac{\partial^2 \delta}{\partial x^k \, \partial x^l}$$

We can now rewrite (2) as

$$[\Delta_{\mathbf{L}} \mathbf{h}]^i = \sum_j \sigma_{ij} * h^j = \int_{\mathbb{R}^n} \sum_j \sigma_{ij}(\mathbf{x} - \mathbf{y})\, h^j(\mathbf{y})\, d\mathbf{y},$$

[8] Using the relation $\text{div}(\mathbf{A}\mathbf{b}) = (\text{div}\mathbf{A}) \cdot \mathbf{b} + \text{Tr}(\mathbf{A}\, D\mathbf{b})$, where $\mathbf{A}$ is a matrix and $\mathbf{b}$ a vector.

where:

$$\sigma_{ij} = \sum_k [\mathrm{div}(\mathbf{L}_{ij})]^k \frac{\partial \delta}{\partial x^k} + \sum_{k,l} L_{ij}^{kl} \frac{\partial^2 \delta}{\partial x^k \, \partial x^l}$$

In words, there is a direct mathematical and canonical link between a differential and an integral operator.

It verifies the conservation property:

$$\int_{\mathbb{R}^n} \Delta_{\mathbf{L}} \mathbf{h}(\mathbf{x}) \, d\mathbf{x} = 0 \tag{3}$$

in coherence with the physical law of conservation property for diffusion processes. In words, we want to balance the $\mathbf{h}$ values in the information diffusion process, i.e. guaranty that if we reduce the value at one location it will increase elsewhere accordingly, as in a fluid which particles are neither created nor deleted. This guarantees the stability of the numerical computations.

Furthermore, this is a "punctual" integral operator in the sense its magnitude is zero except at $\mathbf{x}$. Considering a quadratic semi-norm (i.e. its variance or inertia), it writes:

$$\int_{\mathbb{R}^n - B(\mathbf{x},\varepsilon)} \sigma_{kl} (\mathbf{x} - \mathbf{y})^2 \, d\mathbf{y} = 0 \tag{4}$$

where $B(\mathbf{x}, \varepsilon)$ is a ball around $\mathbf{x}$ of radius $\varepsilon$, which can be as small as possible, but must be excluded, in order (4) to be well-defined because the product of two distributions is not necessarily a distribution.

**Optimal approximation of the diffusion operators.** It is not possible in practice to implement the "punctual" operator $\Delta_{\mathbf{L}} \mathbf{h}$ because what is given, in the real world, is a set of "samples".

More precisely, we have to consider a set of "measures", defined as integral values of the continuous function over the measurement sensor receptive field. Furthermore, since numerical values contains uncertainties, it is a reasonable choice to "average" several values in order to smooth these uncertainties.

This legitimates the fact, at that the implementation level, a differential operator is approximated using an integral operator.

Based on this remark and following [96, 28, 120], we approximate the differential operator by an integral operator of the form:

$$[\Delta^*_{\mathbf{L}}\mathbf{h}(\mathbf{x})]_k = \sum_l \left[\int_{\mathbb{R}^n} \sigma^\epsilon_{kl}(\mathbf{x},\mathbf{y})\,\mathbf{h}^l(\mathbf{y})\,d\mathbf{y}\right] - \bar{\sigma}^\epsilon_{kl}(\mathbf{x})\,\mathbf{h}^l(\mathbf{x}) \tag{5}$$

with $\bar{\sigma}^\epsilon_{kl}(\mathbf{x}) = \int_{\mathbb{R}^n} \sigma^\epsilon_{kl}(\mathbf{x},\mathbf{y})\,d\mathbf{y}$

For such an approximation to be well-defined[9], the integral must be convergent. Here, to obtain this property, we assume that $\sigma^\epsilon(\mathbf{x},\mathbf{y})$ has a bounded support $\mathcal{S}$, *included in a ball* $\mathcal{B}(\mathbf{x},\epsilon)$ *of radius* $\epsilon > \varepsilon$, i.e. $\mathcal{S} \subset \mathcal{B}(\mathbf{x},\epsilon)$.

The 2nd term allows to verifies the conservation property, as soon as the kernel average is symmetric, i.e.

$$\int_{\mathcal{S}} \sigma^\epsilon_{kl}(\mathbf{x},\mathbf{y})\,d\mathbf{x} = \int_{\mathcal{S}} \sigma^\epsilon_{kl}(\mathbf{y},\mathbf{x})\,d\mathbf{x} \tag{6}$$

as derived in [120]. Such kernel are not necessarily symmetric.

It appears as an operator which *computes a weighted mean value around* $\mathbf{x}$ *minus the balanced value at* $\mathbf{x}$.

It is easy to verify that, if $\sigma^\epsilon_{kl}(\mathbf{x},\mathbf{y}) = \sigma_{kl}(\mathbf{x}-\mathbf{y})$ then $[\Delta^*_{\mathbf{L}}\mathbf{h}]_k = [\Delta_{\mathbf{L}}\mathbf{h}]_k$ and the "exact" operator $\Delta_{\mathbf{L}}\mathbf{h}$ is a particular case of the "approximate" operator $\Delta^*_{\mathbf{L}}\mathbf{h}$.

Requiring, $\sigma^\epsilon_{kl}$ to be around each point $\mathbf{x}$ as closed as possible to the "exact" operator $\sigma_{kl}$, we formalize this proximity by the following quadratic criterion:

$$\min_{\sigma^\epsilon} \int_{\mathcal{S}} \sigma^\epsilon_{kl}(\mathbf{x},\mathbf{y})^2 d\mathbf{x}d\mathbf{y} \tag{7}$$

In words, we use a maximal amount of information close to the point: the "sharpest" the operator, the best.

It is easily shown [120] that minimizing (7) is equivalent to choose $\sigma^\epsilon_{kl}(\mathbf{x},\mathbf{y})$ as close as possible to "exact" operator $\sigma_{kl}(\mathbf{x}-\mathbf{y})$ for a quadratic distance, related to the semi-norm specified in (4).

[9] In addition, this operator must belong to a well-defi ned functional space. Here following [28] again, we consider a subset of the distributions: the Sobolev function space $W^{s,\infty}(\mathbb{R}^n)$. In fact, we are going to obtain solutions which are ordinary differentiable functions, so that it is only a formal point.

In order to relate the differential operator $\Delta_{\mathbf{L}}\mathbf{h}$ with its integral approximation $\Delta^*_{\mathbf{L}}\mathbf{h}$. We identify $\Delta_{\mathbf{L}}\mathbf{h}$ with the Taylor expansion of $\Delta^*_{\mathbf{L}}\mathbf{h}$ at some order $r$ [121], yielding a bound on the error for our approximation, as a function of $\epsilon^{r-1}$ i.e.:

$$||\Delta_{\mathbf{L}}\mathbf{h} - \Delta^*_{\mathbf{L}}\mathbf{h}||_{0,\infty} = ||R^r\mathbf{h}||_{0,\infty} \leq C\,\epsilon^{r-1}||\mathbf{h}||_{r+1,\infty}$$

for a fixed constant $C$, since $\mathcal{S} \subset B(\mathbf{x},\epsilon)$. This, for a covering of the parameter space by supports $\mathcal{S} \subset B(\mathbf{x},\epsilon)$ organized in some mesh, allows to conclude that $\lim_{\epsilon w \to 0}\Delta^*_{\mathbf{L}}\mathbf{h} = \Delta_{\mathbf{L}}\mathbf{h}$ and $\lim_{r\to\infty}\Delta^*_{\mathbf{L}}\mathbf{h} = \Delta_{\mathbf{L}}\mathbf{h}$ for small $\epsilon$: in words, this approximation is consistent, with respect to the sampling and the Taylor expansion.

**Integral approximation of a diffusion operator as a NN.** In [120] with a formulation similar to the present one, the integral approximation kernel has been defined using rational functions, i.e. ratio of polynomials, yielding the possibility to defined "poles" but with the risk of numerical instabilities. We consider that, in terms of biological plausibility[10], we must define such approximation without "divisions", avoiding the risk of considering huge values whereas neuronal information has a limited range and using an operation which is not a direct biologically plausible operation.

In fact, we simply assume that a neuron code a value at a given point $\mathbf{p}_u$ of the cortical map. Values at other points are set to zero, yielding:

$$\sigma^{\epsilon}_{kl}(\mathbf{x},\mathbf{y}) = \textstyle\sum_{uv}\sigma^{uv}_{kl}\delta(\mathbf{x}-\mathbf{p}_u,\mathbf{y}-\mathbf{p}_v)$$

where $\mathbf{p}_u$ are the neuron localization in the NN map, while $\delta(\mathbf{x},\mathbf{y})$ is the vector $2\,n$ Dirac distribution. This sampling model, related to a so called "particle process" [28], assumes that each unit computes a discrete set of map values. It is often used in the absence of a priori geometric knowledge on the unit distribution, which is the case for neuronal units. In comparison, in [120], for camera images processing, the sampling model of the continuous equation is based onto a simple but different model, where each unit represents an "average value" around its location which corresponds to the pixelic image formation. A perspective of this work is to consider such alternatives.

[10] See e.g. [16] for details about the biological plausibility of linear and multiplicative computational steps, including motion detection and short-term memory, while [129] focuses on the biologically plausible implementation of nonlinear operation such as minimum computation or comparisons between inputs.

Here, equation (1), is now of the form:

$$\frac{\partial \mathbf{h}_t^k}{\partial t}(\mathbf{p}_u) = -\sum_l \left[ \mathbf{\Lambda}_l^k \left[ \mathbf{h}_t^l - \bar{\mathbf{h}}_t^l \right](\mathbf{p}_u) - \sum_v \sigma_{kl}^{uv} \mathbf{h}_t^l(\mathbf{p}_v) \right] \tag{8}$$

and using a, say, Euler iterative scheme of the form $\mathbf{h}_{t+1}(\mathbf{p}_u) = \mathbf{h}_t(\mathbf{p}_u) + \Delta t \, \frac{\partial \mathbf{h}_t}{\partial t}(\mathbf{p}_u)$ for sufficiently small $\Delta t$ allows the convergence of this 1st order minimization algorithm towards the criterion minimum.

The implementation of (8) has the same architecture and is not more complex than any other neuronal networks NN, including Hopfield networks. Contrary to this basic model, which assumes the symmetry of the weights ($\sigma^{uv} = \sigma^{vu}$) and the absence of reflexivity ($\sigma^{uu} = 0$) we have introduced here weaker assumptions: average symmetry ($\sum_u \sigma^{uv} = \sum_v \sigma^{vu}$) and conservation property, but the real difference is that these assumptions are by construction, always verified and do not introduce additional constraints.

The implementation proposed in (8) allows to consider this computation as a biological plausible mechanism. See e.g. [16] for details about the biological plausibility of linear and multiplicative computational steps.

At a higher scale, in the cortex, the "neuronal unit" is a cortical hyper-column. Our model is thus to be mapped onto usual computational model of cortical columns processes (see [20] for a treatise on the subject), as already proposed by [66]. Such a "processing unit" [59] is discussed in e.g. [90]. Such cortical column architecture is represented in Fig. 5.

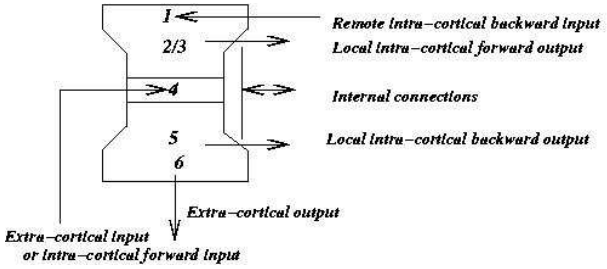



Figure 5: Schematic description of a cortical column, from [20].

We propose to consider the link between the neuronal computation architecture and the structure of a cortical columns as follows:

- *extra cortical input or intra-cortical input from previous layers* (layers IV of the cortex) correspond to the input variable $\bar{\mathbf{h}}$,
- *extra cortical or backward intra-cortical output* (layers V of the cortex) correspond to the output the computed variable $\mathbf{h}$,
- *local intra cortical connections* (layers II/III of the cortex) correspond to inter-parameters interaction, i.e. forward inputs to define $\mathbf{\Lambda}$ and $\mathbf{L}$, thus $\sigma$,
- *backward intra cortical connections* (layers I of the cortex) correspond to inter-parameter interactions, i.e. backward inputs to define $\mathbf{\Lambda}$ and $\mathbf{L}$, thus $\sigma$,
- *internal connections* correspond to the local diffusion mechanism, i.e. correspond to the integration over the operator variable.

**Learning the NN weights in this framework.** Given a cortical map computation unit, as proposed in (8), parameterized by $\mathbf{\Lambda}$ and $\mathbf{L}$ thus $\mathbf{M} = div(\mathbf{L})$, the next step is to describe how to derive $\sigma$. In other words, we have to explain how to proceed to "what is to be done" as stated in (10) to "how to do it" as stated in (8).

Minimizing the quadratic criterion (7) with the linear constraints [C0] [C2] [C1] and (6) yields the following convex quadratic problem, for a given $kl$, with a unique generic solution obvious to derive:

$$\min \sum_{uv} (\sigma_{kl}^{uv})^2 \text{ with } \begin{array}{l} \forall u, \sum_v \sigma_{kl}^{uv} (\mathbf{p}_u - \mathbf{p}_v)^\alpha = \left\{ \begin{array}{rcl} 0 & |\alpha| & > 2 \\ 2\,\mathbf{L}_{kl}^{ij}(\mathbf{p}_u) & \alpha & = \mathbf{e}_i + \mathbf{e}_j \\ \mathbf{M}_{kl}^{j}(\mathbf{p}_u) & \alpha & = \mathbf{e}_j \end{array} \right. \\ \sum_v \sigma_{kl}^{uv} = \sum_u \sigma_{kl}^{uv} \end{array} \tag{9}$$

For a given $kl$, writing $w = (u, v)$ and $\sigma_{kl}^{w}$ the vector indexed by $w$, (9) writes:

$$\min_\sigma ||\sigma||^2 \text{ with } \mathbf{C}\,\sigma = \mathbf{b}$$

where $\mathbf{b} = (0 \cdots 2\,\mathbf{L}_{kl}^{ij}(\mathbf{p}_u) \cdots \mathbf{M}_{kl}^{j}(\mathbf{p}_u) \cdots)$ is the problem input, while the coefficients of the matrix $\mathbf{C}$ are constant and either of the form $(\mathbf{p}_u - \mathbf{p}_v)^\alpha$ or in $\{-1, 0, 1\}$ as directly derived from (9). This matrix $\mathbf{C}$ only depends on the map geometry, neither on its input nor on its values.

This compact form allows to write, in the general case, the explicit solution: $\bar{\sigma} = \mathbf{C}^T (\mathbf{C}\,\mathbf{C}^T)^{-1}\,\mathbf{b}$ with $||\bar{\sigma}||^2 = ||\mathbf{b}||^2_{(\mathbf{C}\,\mathbf{C}^T)^{-1}}$ (see [123] for a derivation and a complete discussion). The key pointis that this solution is easily calculable using an

iterative learning rule of the form:

$$\sigma_{t+1} = \sigma_t - \gamma$$

which converges[11] as soon as $\gamma$ verifies $\gamma^T\,\mathbf{g} > \frac{1}{2}\,||\gamma||^2_{\mathbf{C}^T\,\mathbf{C}}$ with $\mathbf{g} = (\mathbf{C}^T\,\mathbf{C})\,\sigma_t - \mathbf{C}^T\,\mathbf{b}$, i.e. as soon as $||\gamma||$ is small enough and $\gamma$ in the direction of $\mathbf{g}$ (i.e. $\gamma^T\,\mathbf{g} > 0$). When $\mathbf{g} = 0$ we obtain $\sigma = \bar{\sigma}$ and the iteration stops. The vector $\mathbf{g}$ is easily computed as a linear combination of the input $\mathbf{b}$ and the estimate $\sigma$ and the end of the iteration easily detected when $\mathbf{g} = 0$.

Since this learning rule is still valid for approximate values of $\gamma$, biological networks can thus easily implement this rule. This kind of linear learning rule is a particular case of Hebbian learning rule (see [33, 94] for an experimental discussion and [42] for a theoretical development).

What is interesting here, is that the present formalism is related to a tuning mechanism simple enough to be mapped on existing model of biological neuronal networks. More precisely, this derivation is in coherence with [129] who focus on the biologically plausible implementation of nonlinear operation such as minimum computation or comparisons between inputs.

However, at this level of description, how the cortical map parameters are learned is an open question, not addressed here.

## Formalizing the interaction between different cortical maps

Let us now consider, from (10), the definition of several cortical maps, indexed by $i$, i.e.:

$$\mathbf{h}_i = arg\,min\ \mathcal{L}_i \text{ with } \mathcal{L}_i = \frac{1}{2}\int ||\mathbf{h}_i - \bar{\mathbf{h}}_i||^2_{\mathbf{\Lambda}_i(\mathcal{S}*\rho(\mathbf{h_{\bullet i}}))} + ||\nabla\mathbf{h}_i||^2_{\mathbf{L}_i(\mathcal{S}*\rho(\mathbf{h_{\bullet i}}))}$$

[11] This fact is obvious, considering the minimization of $\frac{1}{2}\,||\sigma - \bar{\sigma}||^2_{\mathbf{C}^T\,\mathbf{C}}$ with:

$$\frac{1}{2}\,||\sigma_{t+1} - \bar{\sigma}||^2_{\mathbf{C}^T\,\mathbf{C}} = \frac{1}{2}\,||\sigma_t - \bar{\sigma}||^2_{\mathbf{C}^T\,\mathbf{C}} - \gamma^T\,\mathbf{g} + \frac{1}{2}\,||\gamma||^2_{\mathbf{C}^T\,\mathbf{C}}$$

this convex quadratic criterion being minimum iff $\sigma = \bar{\sigma}$. From the previous equation, $\frac{1}{2}\,||\sigma_{t+1} - \bar{\sigma}||^2_{\mathbf{C}^T\,\mathbf{C}} < \frac{1}{2}\,||\sigma_t - \bar{\sigma}||^2_{\mathbf{C}^T\,\mathbf{C}}$ iff $\gamma^T\,\mathbf{g} > \frac{1}{2}\,||\gamma||^2_{\mathbf{C}^T\,\mathbf{C}}$ so that this bounded criterion decreases as soon as the condition is verified.
The condition $\gamma^T\,\mathbf{g} > \frac{1}{2}\,||\gamma||^2_{\mathbf{C}^T\,\mathbf{C}}$ implies that $\gamma^T\,\mathbf{g} > 0$, i.e. that $\gamma$ and $\mathbf{g}$ are in the same direction (but not necessary aligned). A step further, let us write $\mathbf{u} = \gamma/||\gamma||$. For a fixed $\mathbf{u}$, as soon as $||\gamma|| < 2\,\mathbf{u}^T\,\mathbf{g}/||\mathbf{u}||^2_{\mathbf{C}^T\,\mathbf{C}}$ the condition is verified. Therefore, as soon as $||\gamma||$ is small enough and $\gamma^T\,\mathbf{g} > 0$ the convergence is obtained.
Finally the gradient of this quadratic criterion is precisely equal to $\mathbf{g}$: as a consequence $\mathbf{g} = 0 \Leftrightarrow \sigma = \bar{\sigma}$.

writing $\mathbf{h}_\bullet = (\cdots \mathbf{h}_i, \cdots)$ all cortical maps values and $\mathbf{h}_{\bullet i}$ a sub-vector of $\mathbf{h}_\bullet$ which does not contains $\mathbf{h}_i$. This vector corresponds to the *backward* connections onto the map of index $i$. A spatial smoothing operator $\mathcal{S}$ has been introduced, to take into account the fact that backward connections are divergent, as discussed previously.

A rectification function $\rho()$ has also been introduced, e.g. $\rho(u) = max(u, 0)$ to take into account the fact that only positive quantities is output by cortical maps, as discussed previously. More generally, authors consider profiles of the form $\rho(u) = \gamma(Y(u)u)$, where $Y(u) =$ if $u > 0$ then $u$ else $0$ is the Heaviside function and $\gamma()$ a strictly increasing positive function in $\mathbb{R}^+$ with $\gamma(0) = \gamma'(0) = 0$ (e.g. $\gamma(u) = u^\alpha$ for some $\alpha > 1$, [107]), so that $\rho'(u) = \gamma'(Y(u)u)$, as the reader can easily verify. At this level, there is no need to further specify $\rho()$.

Here, each cortical map value can tune the other map computations, modifying the parameters $\mathbf{\Lambda}_i$ and $\mathbf{L}_i$, as made explicit in the previous equation. The key point is that such interactions are constrained:
(i) cortical value are "averaged", i.e. smoothed in space, before influencing the cortical map parameters,
(ii) a cortical map value does not modify its own parameters,
(iii) a feed-forward connection from a cortical map of index $i$ to a cortical map of index $j$ corresponds to the fact that $\bar{\mathbf{h}}_j = \rho(\mathbf{h}_i)$ and define a lattice (no loops between forward connections).
These constraints correspond to what has been observed in the cortex, as reviewed previously. They are fundamental in the following development.

Forward connections define a lattice of cortical maps and we can say a map of index $i_l$ is "after another" map of index $i_1$, say $i_l \succ i_1$, iff $\bar{\mathbf{h}}_{i_{j+1}} = \rho(\mathbf{h}_{i_j})$ for a sequence of indexes $i_1..i_l$, defining a feed-forward connection. Our understanding of the cortical maps interactions is that this graph has no cycle: we can not have $i \succ i$ for some $i$. Furthermore, with such an acyclic connections and in the *absence* of backward connections, the final result is easy to predict: given some inputs, each iterative computation in a cortical map yield a stable result and from upstream to downstream this stable result propagates. This is the case for very fast brain computation [110] where, due to very short latencies, only feed-forward computations occur.

The situation is very different when backward connections interact: several criteria are to be minimized simultaneously, yielding a apparently very complex dy-

namical system, with the risk of interferences, oscillations, chaotic behavior, etc.. What is the result of such interactions is also to clarify.

The proposed solution of this problem is based on the following *fact* [121] *:*

*minimizing,* $\forall\, i$ *the criteria* $\mathcal{L}_i$ *with respect to* $\mathbf{h}_i$

*is equivalent to minimize with respect to* $\mathbf{h}_\bullet$*, in the general case:*

$$\mathcal{L}_\bullet = \sum_i \nu(||\nabla_i \mathcal{L}_i||)\, \mathcal{L}_i$$

*writing* $\nabla_i = \partial/\partial \mathbf{h}_i$*. Here,* $\nu()$ *is a* $\mathbb{R}^+ \rightarrow \mathbb{R}^+$ *positive strictly increasing profile with* $\nu(u) \geq 0$*,* $\nu'(u) > 0$*,* $\nu(0) = 0$ *and* $\lim_{u \rightarrow 0} \nu'(u)/u = 0$*. For instance* $\nu(u) = u^\alpha$ *with* $\alpha > 2$*.*

This criterion provides a view of what is a common objective for the different cortical maps computations.

As a consequence: *minimizing each criterion without considering the other criteria and parameters also decreases the common criterion*, thanks to the biologically driven assumptions: (i) space smoothing in backward connections, (ii/iii) forward connections graph being a lattice.

This is a crucial fact, because this also means that minimizing each criterion is a convergent process, since it corresponds to a common criterion minimization. As a consequence, we have a formal verification that feedback links in our framework and with the proposed assumptions yields a well-defined process.

More than that, the previous derivation also describes qualitatively how interactions between different cortical maps occurs:

- backward connections have a constant influence in the sense that they can very rapidly tune the processing of a cortical map but do not interfere with the convergence inside a such a map, they propagate information between the cortical maps, in a stable way; very fast propagation can occur in “one step”, i.e. without inducing transient effects
- forward connections act as a “data propagation” though the related lattice and may induce transient effects on downstream layers,
- if a cortical map input is changed (because the cortex inputs vary dynamically) the overall process is still convergent.

## Appendix: Using anisotropic diffusion operators.

Let us briefly revisit the computer vision partial differential equation methodology in the case of Euclidean vector maps (how it can be generalized to non-Euclidean cortical maps[12] computations and how a general class of nonlinear diffusion operators are implemented within the present frameworks is discussed elsewhere [120]). The goal of this section is to introduce this approach and point out to where the key problem is in terms of biological plausibility.

Here, as required in some situations (e.g. [2, 30, 56]), we consider functions whose values are not scalars but vectors.

**Defining a cortical map function from a criterion.** Let us consider a vector map:

$$\mathbf{h} : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^m.$$

from a bounded open set $\Omega$ of $\mathbb{R}^n$ into $\mathbb{R}^m$. $\mathbf{h}$ belongs to a set $\mathcal{F}$ of admissible functions. Generally speaking $\mathcal{F}$ is a dense linear subset of a Hilbert space $H$, the scalar product of which is denoted by $(\cdot,\cdot)_H$. $\bar{\mathbf{h}}(\mathbf{x})$ is a "reference" function related to an input or measure $\mathbf{m}(\mathbf{x})$, as shown in Fig. 6.

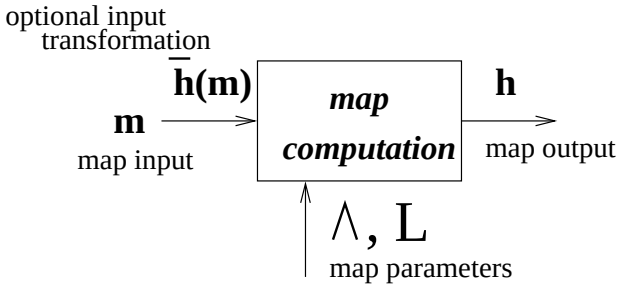



Figure 6: Input/output scheme of regularization process, see text for details.

[12]Generalization from $\mathbb{R}^n$ to nonlinear manifolds is also to be considered because cortical maps are intrinsically "curved space" with the fact that parametric spaces of dimension $n \geq 2$ (e.g. the visual areas which code retinal localization, edges orientation, binocular disparity, color .. in a "interlaced" way) are represented onto cortical maps. This must be represented by a nonlinear space with a variable metric, i.e. a Riemannian manifold (see e.g. [90] for an extended discussion).

In order to define this map, we consider the following optimization problem (as discussed e.g. in [30]):

$$\mathbf{h}^* = arg\ min_{\mathbf{h}} \mathcal{L}(\mathbf{h}) \quad \text{with} \quad \mathcal{L}(\mathbf{h}) = \frac{1}{2} \int \underbrace{||\mathbf{h} - \bar{\mathbf{h}}||^2_{\mathbf{\Lambda}}}_{\text{input}} + \underbrace{\varphi_{\mathbf{L}}(D\mathbf{h})}_{\text{regularization}} \tag{10}$$

where $D(\mathbf{h})$ is the first order derivative of the function $\mathbf{h}$, i.e. its $m \times n$ Jacobian matrix, $\varphi_{\mathbf{L}}$ is a positive definite quadratic form which can be written in terms of the gradients the components of $\mathbf{h}$ as

$$\varphi_{\mathbf{L}}(D(\mathbf{k}) = D\mathbf{h}\, \mathbf{L}\, D\mathbf{h}^T = \sum_{i,j} (\nabla h^i)^T \mathbf{L}_{ij} \nabla h^j,$$

and

$$||\mathbf{h} - \bar{\mathbf{h}}||^2_{\mathbf{\Lambda}} = (\mathbf{h} - \bar{\mathbf{h}})^T \mathbf{\Lambda} (\mathbf{h} - \bar{\mathbf{h}}).$$

In words, the *specification* of this map of values corresponds to an "objective" or a "criterion" to attain. This criterion is built from two terms:
(i) the input term, related to the data input
(i.e. looking for a solution as compatible as possible with the input) and
(ii) the regularization term, related to the a priori information
(i.e. looking for a solution with plausible properties: here which variation is minimal).
Regarding the term related to the input :

- The function $\mathbf{\Lambda} : \mathbb{R}^n \rightarrow S_m^+ \in W^{s,\infty}(\mathbb{R}^n)$, $i, j = 1, \cdots, m$, where $S_m^+$ is the set of square symmetric positive definite matrixes of size $m$, defines a so called, *measurement information metric*.

  This metric allows to represent:
  (i) the *precision of the input function*: the higher this precision in a given direction, the higher the value of $\mathbf{\Lambda}$ in this direction (in a statistical framework, $\mathbf{\Lambda}$ corresponds to the inverse of a covariance matrix)
  (ii) *partial observations*: if the input function $\bar{\mathbf{h}}(\mathbf{x})$ is only defined in some directions, it corresponds to a matrix $\mathbf{\Lambda}$ definite only in these directions (for instance, if $\bar{\mathbf{h}}(\mathbf{x})$ is only defined in the direction $\mathbf{u}$ at a given location we write $\mathbf{\Lambda} = k\, \mathbf{u}\, \mathbf{u}^T$ for some $k$),
  (iii) *missing data*: if the input function $\bar{\mathbf{h}}(\mathbf{x})$ is not defined for some $\mathbf{x}$, we

simply have to state $\mathbf{\Lambda} = 0$ at this location; more generally it represents:
(iv) *linear relations between some measures* $\mathbf{m}$ *and the parameter estimation* $\mathbf{h}$, say $\mathbf{M}\,\mathbf{h} = \mathbf{m}$, which is obviously equivalent to require $||\mathbf{h} - \bar{\mathbf{h}}||^2_{\mathbf{\Lambda}} = 0$ with $\mathbf{\Lambda} = \mathbf{M}^T\,\mathbf{M}$ and $\bar{\mathbf{h}} = \mathbf{M}^T\,\mathbf{m}$ (see e.g. [123] for a development).

On the other hand, regarding the term related to "regularization":

- The functions $\mathbf{L}_{ij} : \mathbb{R}^n \rightarrow S_n \in W^{s,\infty}(\mathbb{R}^n)$, where $S_n$ is the set of square symmetric matrixes define, so called, *diffusion tensor* $\mathbf{L}$, which is :
  - symmetric i.e. $\mathbf{L}^{ij} = \mathbf{L}^{ji}$ and
  - "positive" i.e. so that $\forall \mathbf{M} \in \mathbb{R}^{m\times n}, \mathbf{M}^T\,\mathbf{L}\,\mathbf{M} = \sum_{ijkl} \mathbf{M}^k_i\,\mathbf{L}^{ij}_{kl}\,\mathbf{M}^l_j \geq 0$
  in order the previous definition to be coherent. It defines the *profile* of this regularization.

  Thanks to the *regularization* term $||\nabla\mathbf{h}||^2_{\mathbf{L}}$, since the variation of $\mathbf{h}$ is minimized, a "smoothed" but also well-defined version of $\mathbf{h}$ is obtained [112].

  (a) When the problem is ill-posed, i.e. if there are many (and usually numerically unstable) solutions, the key idea is to choose the solution which variations are minimized: this defines a unique solution but also a well-defined solution.

  (b) When the input function is partially or approximately defined at some points, as discussed previously, the value at such a point is defined using information "around" which diffuses (as discussed now) from well-defined values to undefined or ill-defined values.

**Regularization as a diffusion mechanism.** Our criterion $\mathcal{L}$ being regular, its first variation[13] at $\mathbf{h} \in \mathcal{F}$ in the direction $\mathbf{k} \in H$ is defined by (see, e.g., [24])

$$\delta_{\mathbf{k}}\mathcal{L}(\mathbf{h}) = \lim_{\varepsilon\rightarrow 0} \frac{\mathcal{L}(\mathbf{h}+\varepsilon\mathbf{k}) - \mathcal{L}(\mathbf{h})}{\varepsilon}$$

If the mapping $\mathbf{k} \rightarrow \delta_{\mathbf{k}}\mathcal{L}(\mathbf{h})$ is linear and continuous, the Riesz representation theorem [36] guarantees the existence of a unique vector, denoted by $\nabla_H\mathcal{L}(\mathbf{h})$, and called the gradient of $\mathcal{L}$, which satisfies the equality

$$\delta_{\mathbf{k}}\mathcal{L}(\mathbf{h}) = (\nabla_H\mathcal{L}(\mathbf{h}),\, \mathbf{k})_H$$

[13] Also called its Gâteauxd derivative.

for every $\mathbf{k} \in H$. The gradient depends on the choice of the scalar product $(\cdot,\cdot)_H$ though, a fact which explains our notation. If a minimizer $\mathbf{h}^*$ of $\mathcal{L}$ exists, then the set of equations $\delta_{\mathbf{k}}\mathcal{L}(\mathbf{h}) = 0$ must hold for every $\mathbf{k} \in H$, which is equivalent to $\nabla_H \mathcal{L}(\mathbf{h}) = 0$. These equations are called the Euler-Lagrange equations associated with the energy fucntional $\mathcal{L}$. Here (see e.g. [120]), (10) is minimal only if:

$$\nabla_H \mathcal{L}(\mathbf{h}) = \frac{\partial \mathcal{L}}{\partial \mathbf{h}} = \mathbf{\Lambda}(\mathbf{h} - \bar{\mathbf{h}}) - \Delta_{\mathbf{L}}\mathbf{h} = 0 \tag{11}$$

with a so-called diffusion term:

$$[\Delta_{\mathbf{L}}\mathbf{h}]_i = [\mathbf{div}(D\varphi(D\mathbf{h}))]_i = \mathrm{div}(\sum_{j=1}^{m} \mathbf{L}_{ij} \nabla h^j) \tag{12}$$

This equation simply states that, at an extremum, the criterion is locally "flat" i.e. its 1st order variation vanishes.

The term of diffusion $\Delta_{\mathbf{L}}\mathbf{h}$ allows to "propagate" some information about $\mathbf{h}$ from one point $\mathbf{x}$ to another, because it depends on the variation of $\mathbf{h}$ at $\mathbf{x}$ i.e. on what happens "around" $\mathbf{x}$.

Clearly, without the term of diffusion, the solution is $\mathbf{h} = \bar{\mathbf{h}}$.

Obviously, at the implementation level, the key problem is the calculation of $\Delta_{\mathbf{L}}\mathbf{h}$.

In order to solve (11) which corresponds to the *gradient* of the criterion (10) a 1st order *gradient descent* is usually proposed[14]. Given an initial estimate $\mathbf{h}_0 \in \mathcal{F}$, a time-dependent differentiable function, also noted $\mathbf{h}$, from the interval $[0, +\infty[$ into $H$ is computed as the solution of the following initial value problem:

$$\begin{cases} \frac{d\mathbf{h}}{dt} = -\nabla_H \mathcal{L} = \mathbf{\Lambda}(\bar{\mathbf{h}} - \mathbf{h}) + \Delta_{\mathbf{L}}\mathbf{h} \\ \mathbf{h}(0)(.) = \mathbf{h}_0(\cdot) \end{cases}$$

It leads to a local minimum of the criterion, when iteratively computed using, e.g. an Euler rule of the form: $\mathbf{h}_{t+1} = \mathbf{h}_t + \Delta t \, \frac{d\mathbf{h}}{dt}$.

[14]**About 1st order scheme convergence**. The partial differentiation rule is written $\frac{\partial \mathcal{L}}{\partial t} = \nabla_H \mathcal{L} \frac{\partial \mathbf{h}}{\partial t}$. Considering $\frac{\partial \mathbf{h}}{\partial t} = -\nabla_H \mathcal{L}$ as in (1) yields $\frac{\partial \mathcal{L}}{\partial t} = -||\nabla_{\mathbf{h}} \mathcal{L}||^2 < 0$ so that the criterion is strictly decreasing. Since $\mathcal{L} \geq 0$ it is also bounded and thus converges towards a minimal value. At this local or global minimum $||\nabla_{\mathbf{h}} \mathcal{L}|| = 0$ so that $\mathcal{L}$ is stationary. In fact, the gradient magnitude smoothly decreases with time. In practice, convergence is detected as soon as the gradient is below a given numerical threshold, which always occurs in fi nite time.

Here, since the related quadratic criterion is convex, calculating the *local* partial differential equation (1) at each point leads to the *global* minimization of the criterion (10).

**Relation with linear filtering.** The present mechanism is easily related to a linear filtering operator in the case where $\mathbf{\Lambda}$ and $\mathbf{L}$ are constant. In this particular case, (11) is now a linear differential equation with constant coefficients.

If we consider the Fourier transform $\mathcal{F}[\mathbf{h}(\mathbf{x})] = \mathbf{h}(\mathbf{w})$ of the quantities in (11), using the explicit form given in (12), it is straightforward to obtain in the Fourier domain:

$$\mathbf{h}(\mathbf{w}) = \mathbf{G}_{\mathbf{\Lambda},\mathbf{L}}(\mathbf{w})\,\bar{\mathbf{h}}(\mathbf{w}) \text{ with } \mathbf{G}_{\mathbf{\Lambda},\mathbf{L}}(\mathbf{w}) = [\mathbf{\Lambda}_{kl} + \sum_{ij} \mathbf{L}^{ij}_{kl}\, w_i\, w_j]^{-1}\, \mathbf{\Lambda} \qquad (13)$$

so that we finally obtain the convolution: $\mathbf{h}(\mathbf{x}) = \mathbf{G}_{\mathbf{\Lambda},\mathbf{L}}(\mathbf{x}) * \bar{\mathbf{h}}(\mathbf{x})$.

From (13), it is visible that the corresponding linear filter is defined by the $n^2 + m^2\, n\,(n+1)/2$ coefficients of $\mathbf{\Lambda}$ and $\mathbf{L}$ (the former being a general $n \times n$ matrix, the latter being a symmetric tensor). From obvious linear algebra in the general case, $\mathbf{G}_{\mathbf{\Lambda},\mathbf{L}}(\mathbf{w})$ is a rational fraction with a denominator of degree less than $2\,n$ and a numerator of degree less than $2\,(n-1)$. This corresponds to rather general filters, although not all linear filters can be defined from such a formula.

A step ahead, as shown in [114], if $\mathbf{L}$ is constant and if $\mathbf{\Lambda}$ is negligible, $G_{\mathbf{\Lambda},\mathbf{L}}$ corresponds, at a given time $t$ of the diffusion process defined in (1) to an oriented Gaussian kernel:

$$G_{\mathbf{L},t}(\mathbf{x}) = \frac{1}{\sqrt{(2\,\pi)^n\, 4\,t\,|\mathbf{L}|}}\, e^{-\frac{\mathbf{x}^T\,\mathbf{L}^{-1}\,\mathbf{x}}{4\,t}}$$

However, in both cases, defining the problem from (10) is more informative than defining a filter, since we formally define what is the "objective". Furthermore, implementing it using (1) is much more efficient than explicitly computing a convolution at each point.

A step ahead, if $\mathbf{\Lambda}$ and $\mathbf{L}$ are not constant, the present framework allows to define complex input/output relationships between $\bar{\mathbf{h}}$ and $\mathbf{h}$ which are more general than what is obtained by a linear filter, introducing some "coupling" between each local convolution. This is also called bilateral filtering [4].

**Introducing nonlinear diffusion operators.** Let us now review how nonlinear regularization profiles influence this diffusion of information. More precisely, we consider for this section:

$$\mathbf{h}^* = arg\ min_{\mathbf{h}} \int \frac{1}{2}||\mathbf{h}-\bar{\mathbf{h}}||^2_{\mathbf{\Lambda}} + \Phi\left(\varphi_{\mathbf{L}}(D\mathbf{h})\right) \Rightarrow \mathbf{\Lambda}\left(\mathbf{h}-\bar{\mathbf{h}}\right) - \Delta_{\Phi,\mathbf{L}}\mathbf{h} = 0$$

with

$$\left[\Delta_{\Phi,\mathbf{L}}\mathbf{h}\right]^i = \text{div}\left(\Phi'(\varphi(D\mathbf{h}))D\varphi(D\mathbf{h})\right)) = 2\,\text{div}\left(\Phi'(\varphi_{\mathbf{L}}(D\mathbf{h}))\sum_j \mathbf{L}_{ij}\,\nabla h^j\right).$$

Using standard differential calculus[15], the righthand side is equal to twice

$$\Phi^{''}(\varphi_{\mathbf{L}}(D\mathbf{h}))\nabla\varphi_{\mathbf{L}}(D\mathbf{h})\cdot\sum_j \mathbf{L}_{ij}\,\nabla h^j + \Phi'(\varphi_{\mathbf{L}}(D\mathbf{h}))\text{div}(\sum_j \mathbf{L}_{ij}\,\nabla h^j).$$

The term $\nabla\varphi_{\mathbf{L}}(D\mathbf{h})$ is given by

$$\nabla\varphi_{\mathbf{L}}(D\mathbf{h}) = \sum_{i,j}\nabla\left((\nabla h^j)^T\mathbf{L}_{ij}\nabla h^i\right),$$

and[16]

$$\nabla\left((\nabla h^j)^T\mathbf{L}_{ij}\nabla h^i\right) = \mathbf{H}_j\mathbf{L}_{ij}\nabla h^i + \mathbf{L}_{ij}\mathbf{H}_i\nabla h^j + D\mathbf{L}_{ij}(\nabla h^i,\,\nabla h^j),$$

where $\mathbf{H}_i,\ \mathbf{H}_j$ are the Hessians of $h^i$ and $h^j$ and $D\mathbf{L}_{ij}$ is the derivative of the matrix $\mathbf{L}_{ij}$. Hence we have

$$\nabla\varphi_{\mathbf{L}}(D\mathbf{h}) = \sum_i\left((\mathbf{H}_i\mathbf{L}_{ii}+\mathbf{L}_{ii}\mathbf{H}_i)\nabla h^i + D\mathbf{L}_{ii}(\nabla h^i,\nabla h^i)\right) +$$

$$\frac{1}{2}\sum_{i,\,j,\,i\neq j}(\mathbf{H}_j\mathbf{L}_{ij}+\mathbf{L}_{ij}\mathbf{H}_j)\nabla h^i+(\mathbf{H}_i\mathbf{L}_{ij}+\mathbf{L}_{ij}\mathbf{H}_i)\nabla h^j+D\mathbf{L}_{ij}((\nabla h^i,\nabla h^j)+(\nabla h^j,\nabla h^i))$$

This corresponds to a large number of regularization mechanisms (see e.g. [113] for a recent review).

[15] The fact that $\text{div}(\alpha\mathbf{v}) = \nabla\alpha\cdot\mathbf{v} + \alpha\,\text{div}\,\mathbf{v}$.

[16] Using the fact that $\nabla(\mathbf{u}\cdot\mathbf{v}) = (D\mathbf{u})\,\mathbf{v} + (D\mathbf{v})\,\mathbf{u}$.

In the case where $\mathbf{L} = Id_{mn\times mn}$ so that $\varphi_{\mathbf{L}}(D\mathbf{h}) = ||D\mathbf{h}||^2$ and $\mathbf{L}_{ij} = \delta_{ij} Id_{n\times n}$, we have

$$\nabla\varphi_{\mathbf{L}}(D\mathbf{h}) = 2\sum_j \mathbf{H}_j \nabla h^j$$

and

$$\frac{1}{2}\left[\Delta_{\Phi,Id}\mathbf{h}\right]_i = \Phi'\left(||D\mathbf{h}||^2\right)\,\Delta h^i + 2\,\Phi''\left(||D\mathbf{h}||^2\right)\nabla h^i\,\cdot\sum_j \mathbf{H}_j\nabla h^j \qquad (14)$$

If we assume further that $\Phi(||D\mathbf{h}||^2) = \sum_j \Phi_j(||\nabla h^j||^2)$ things simplify even further:

$$\left[\Delta_{\Phi,Id}\mathbf{h}\right]^i = \text{div}\left(\Phi_i'(||\nabla h^i||^2)\nabla h^i\right)$$

and the resulting equations are *decoupled*:

$$\left[\Delta_{\Phi,Id}\mathbf{h}\right]^i \;=\; \Phi_i'\left(||\nabla h^i||^2\right)\,\Delta h^i \;+\; 2\,\Phi_i''\left(||\nabla h^i||^2\right)||\nabla h^i||^2\;\;\boldsymbol{\eta}_i^T\,\mathbf{H}_i\,\boldsymbol{\eta}_i\;, \qquad (15)$$

where $\boldsymbol{\eta}_i = \frac{\nabla h^i}{||\nabla h^i||}$.

In such a case, we obtain:
(i) a term related to $\Phi'$ which corresponds to an isotropic diffusion (here $\Delta h^i$ is the Laplacian of the $i$-th component of $\mathbf{h}$) and
(ii) a term related to $\Phi''$ which corresponds to an anisotropic diffusion process in the direction $\eta$ of the gradient.

Furthermore, in this case all components $h^i$ of $\mathbf{h}$ are decoupled.

A natural requirement [2, 31] is to:
($\alpha$) obtain isotropic diffusion for small gradient magnitude because we want to maximize the propagation of information in uniform (thus information less) parts of the parametric space and
($\beta$) cancel the propagation of information for high gradient magnitudes in the direction of the gradient in order to preserve large (thus significant) variations of the input function.

In other words, we want to eliminate small, thus likely noisy, variations of the parametric map, but preserve large variations. This is the case, e.g. of $\Phi(u) = \sqrt{u}$ and several other profiles, as reviewed in [4].

Regarding linear regularization, would it be possible to obtain the same property but preserving the linearity of the diffusion equation? Thanks to what has been

derived in e.g. [81] (regarding the estimation of image motion) there is a positive answer to this question.

Following [81], let us consider the initial or previous estimation $\nabla\widehat{\mathbf{h}}$ of $\nabla\mathbf{h}$: either the input function $\bar{\mathbf{h}}$ or the estimation obtained by previous steps of the iterative process. In other words, we *feedback* the previous estimation $\widehat{\mathbf{h}}$ to obtain an improved estimate.

With this specification, (14) writes:

$$[\Delta_{\mathbf{L}}\mathbf{h}]_k^T = \widehat{\lambda}_\perp \, \Delta\mathbf{h}^k + \left[\widehat{\lambda}_{||} - \widehat{\lambda}_\perp\right] \, \widehat{\eta}^{k\,T} \, \nabla^2\mathbf{h}^k \, \widehat{\eta}^k + o(||\nabla\mathbf{h} - \nabla\widehat{\mathbf{h}}||) \qquad (16)$$

with $\widehat{\eta} = \nabla\widehat{\mathbf{h}}/||\nabla\widehat{\mathbf{h}}||$, $\widehat{\lambda}_\perp = \lambda_\perp\left(||\nabla\widehat{\mathbf{h}}||\right)$ and $\widehat{\lambda}_{||} = \lambda_{||}\left(||\nabla\widehat{\mathbf{h}}||\right)$, while:

$$\lambda_\perp(u) = \Phi'(u) \text{ and } \lambda_{||}(u) = \lambda_\perp(u) + 2\,\lambda'_\perp(u)/u^2 \qquad (17)$$

We may even choose a more general form for $\lambda_{||}(u)$ or $\lambda_\perp(u)$, as soon as:

$$\lim_{u\to 0}\lambda_{||}(u) = 0 \text{ and } \lim_{u\to+\infty}\lambda_\perp(u) = 0.$$

As reviewed in [2, 4], when considering a nonlinear profile, at the implementation level, a similar *linear approximation* is always used. It is thus useless to consider nonlinear profiles $\Phi()$ at the implementation level: anisotropic linear diffusion is sufficient.

Thanks to this derivation, we also obtain a suitable form of the tensor $\mathbf{L}$, defined from (16) and eventually (17).

A step further, as discussed in [121] the same mechanism is generalizable to the computation of *harmonic maps on Riemannian manifold*: for a differentiable map $\mathbf{h} : N \to M$ where $M$ is a Riemannian manifold of dimension $m$ which metric is written $\mathbf{f}$, we may consider the energy density of the function:

$$e(\mathbf{h})(\mathbf{x}) = \tfrac{1}{2}\,\mathbf{g}^{ij}(\mathbf{x})\,\mathbf{f}_{ab}(\mathbf{h}(\mathbf{x}))\,\tfrac{\partial\mathbf{h}^a(\mathbf{x})}{\partial x^i}\,\tfrac{\partial\mathbf{h}^b(\mathbf{x})}{\partial x^j}$$

and the function energy:

$$E(\mathbf{h}) = \textstyle\int_N \, e(\mathbf{h})\,\sqrt{g}\,dx^1 \wedge .. \wedge dx^n$$

which related Euler equation is:

$$\Delta\mathbf{h}^k = \textstyle\sum_{ij}\,\frac{1}{\sqrt{g}}\,\frac{\partial}{\partial x^j}\left(\sqrt{g}\,\mathbf{g}^{ij}\,\frac{\partial\mathbf{h}^k}{\partial x^j}\right) + \mathbf{g}^{ij}\,\sum_{ab}\,\mathbf{F}^k_{ab}(\mathbf{h})\,\frac{\partial\mathbf{h}^a}{\partial x^i}\,\frac{\partial\mathbf{h}^b}{\partial x^j} = 0$$

where $\mathbf{F}^a_{bc} = \frac{1}{2}\,\sum_d \mathbf{f}^{ad}\left(\frac{\partial\mathbf{f}_{bd}}{\partial x^c} + \frac{\partial\mathbf{f}_{cd}}{\partial x^b} - \frac{\partial\mathbf{f}_{bc}}{\partial x^d}\right)$ is the Christoffel symbols of the metric $\mathbf{f}$.

Solutions to these equations are called harmonic maps.

If $N = \mathcal{S}^1$ with its standard metric we obtain the geodesic curves of $M$ as solutions, while if $M = \mathbb{R}^m$ we re-obtain the Laplace-Beltrami operator. Intrinsically, this equation means $trace(\nabla d\mathbf{h}) = 0$ where $d\mathbf{h}$ is the differential of $\mathbf{h}$ [63].

Although out of the scope of the present study, it will be an interesting perspective of the present work to generalize what will be developed in the sequel to this kind of equations, the key problem being to choose a relevant linearization of the nonlinear term $\mathbf{g}^{ij} \sum_{ab} \mathbf{F}^k_{ab}(\mathbf{h}) \frac{\partial \mathbf{h}^a}{\partial x^i} \frac{\partial \mathbf{h}^b}{\partial x^j}$, e.g. to consider:

$\mathbf{L}^{ij}_{kl} = \mathbf{g}^{ij} \delta_{kl}$

and

$\mathbf{M}^j_{kl} = \sum_i \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^j} \left(\sqrt{g}\, \mathbf{g}^{ij}\right) + \frac{1}{2} \left[\mathbf{g}^{ij} \sum_a \mathbf{F}^k_{al}(\widehat{\mathbf{h}}) \frac{\partial \widehat{\mathbf{h}}^a}{\partial x^i} + \mathbf{g}^{ji} \sum_a \mathbf{F}^k_{la}(\widehat{\mathbf{h}}) \frac{\partial \widehat{\mathbf{h}}^a}{\partial x^i}\right] \delta_{kl}$

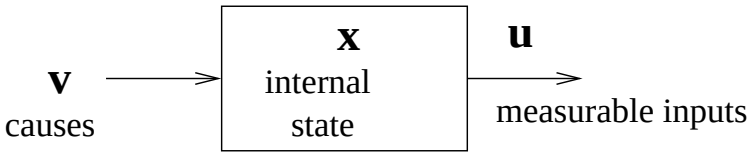
where $\widehat{\mathbf{h}}$ is a previous estimation of $\mathbf{h}$, as discussed when deriving (16).

We will not further develop this point here, but simply wanted to point out such a perspective of the present approach: generalization from $\mathbb{R}^n$ to nonlinear manifolds is also to be considered because cortical maps are intrinsically "curved space" with the fact that parametric spaces of dimension $n \geq 2$ (e.g. the visual areas which code retinal localization, edges orientation, binocular disparity, color .. in a "interlaced" way) are represented onto cortical maps. This must be represented by a nonlinear space with a variable metric, i.e. a Riemannian manifold (see e.g. [90] for an extended discussion).

## Appendix: a view of the Friston framework

Let us briefly review the formalism proposed in [40] with the simple goal to show the direct link with the proposed PDE approach.

Starting with trivial "universal" scheme of an input/output system:




the world to percieved is considered as a deterministic dynamical system :

$$\begin{array}{rcll} \dot{x} & = & f(x,\nu) & x(-\infty)=0 \\ u & = & g(x) & \end{array} \tag{18}$$

the goal of the perception for the brain being to *to find "causes" $\nu$ from inputs* $u$. Here inputs are not the external signals (light) but the neuronal information input in the cortical map. The variable $x$ is the "hidden" deterministic system state, including the system memory.

As a first step, in (18), we can eliminate (in fact "integrate") the influence of $x$: we directly relate the input $\boxed{u = P(\nu,\beta)}$ to the *recent history* of the causes $\nu$, using the following Fliess expansion:

$$u(t) = \underbrace{\int_0^t \kappa_1(\tau)\,\nu(t-\tau)\,d\tau}_{\text{influence from previous causes}} + \underbrace{\int_0^t\int_0^t \kappa_2(\tau,\tau')\,\nu(t-\tau)\,\nu(t-\tau')\,d\tau d\tau'}_{\text{modulatory influence between causes}} + .. \tag{19}$$

i.e. we parameterize this causal relationship with parameter:

$$\beta = [\kappa_1(\tau) = \left.\tfrac{\partial u(t)}{\partial \nu(t-\tau)}\right|_{t=0}, .., \kappa_2(\tau,\tau') = \left.\tfrac{\partial u(t)}{\partial \nu(t-\tau)\,\partial \nu(t-\tau')}\right|_{t=0}, ..]$$

defining Volterra kernels. Here $\kappa_1(\tau)$ defines the cause/input relationship while $\kappa_2(\tau,\tau')$ defines the correlations between different causes.

Using a Baysian apporach, let us find the "maximally probable" estimation of the causes $\nu$, knowing the input $u$ (maximal likehood):

$$\max_\nu log(p(\nu|u)) = \max_\nu \left[ log(p(u|\nu)) + log(p(\nu)) \underbrace{-log(p(u))}_{\text{don't care}} \right]$$

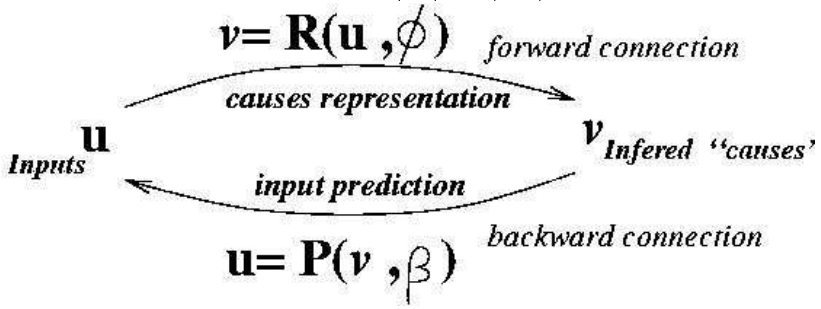
and let us introduce two functions of

- *Expectation:* $\nu = R(u, \Phi)$ which "infers" the causes from the given inputs (which is parameterized by forward connections $\Phi$)
- *estiMation:* $u = P(\nu, \beta)$ which 'predicts" the input from "a priori" causes (which is parameterized by backward connections $\beta$)

yielding:

$$\begin{array}{lrcl} \max_\nu & log(p(u|\nu)) & + & log(p(\nu)) \\ & \text{Conditional information} & & \text{A priori information} \\ & \beta \text{ tuning : } u = P(\nu, \beta) & & \Phi \text{ tuning : } \nu = R(u, \Phi) \\ \max_\nu & log(p(P(\nu, \beta)|\nu)) & + & log(p(R(u, \Phi)) \\ & \text{Estimation} & & \text{Expectation} \end{array}$$

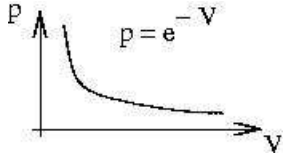
the inference being coherent iff : $u = P(R(u, \Phi), \beta)$ as schematized here:




This means that causes are estimated here using a expectation-minimization algorithm, defining a specific role to forward/backward connections.

A step further, let us choose a probability model, with two assumptions

Additive error, i.e. $u = P(\nu, \beta) + \epsilon$
Gibb's potential $p(\epsilon) = e^{-V(\epsilon,\theta)}$




e.g. using an exponential probality distribution: $V(\epsilon, \theta) = log(c(\theta) + b(\theta)^T a(\epsilon))$ including Gaussian, Poisson, binomial, uniform, .. laws. The maximal likehood approach corresponds now to the combined minimization of two criteria:

$$\min_\beta V(u - P(\nu, \beta), \theta) \qquad \min_{\Phi\ \theta} V(u - P(R(u, \Phi), \beta), \theta) \tag{20}$$

Here the probablity parameter $\theta$ is "expectated" with $\Phi$, i.e. is estimated by forward connections.

The convergence is "guaranteed" towards a sub-optimal (local minimum) solution if the related are not convex, and towards the global minimum otherwise.

One key property is related to the combination of cortical map, modeled within this framework, as illustrated here:

from [40]

when combining several cortical maps, the Baysian framework provides immediatly a formalization of the whole system behavior.

In our framework, (20) criteria are minimized using a quadratic formulation (introducing Riemanian metrics if required):

$$\mathcal{L}_i = \frac{1}{2} \int ||\mathbf{h}_i - \bar{\mathbf{h}}_i||^2_{\mathbf{\Lambda}_i(\mathcal{S}*\mathbf{h}_{\bullet \mathbf{i}})} + ||\nabla \mathbf{h}_i||^2_{\mathbf{L}_i(\mathcal{S}*\mathbf{h}_{\bullet \mathbf{i}})}$$

yielding to a PDE of the form: $\frac{\partial \mathbf{h}_i}{\partial t} = -\nabla \mathcal{L}_i$ as discussed in this section.

# 3 Feedbacks in Hebbian learning mechanisms *R. Guyonneau*

Hebbian learning denotes a class of processes that modify the synapses strength depending on the spiking activity of the corresponding neuron. Since neural activity emerges from the interaction of the organism with its environment, it follows that sensory experience has an important role in refining synaptic connections; thus the belief that Hebbian learning constitues the neural correlates of learning, memory, adaptation... Here, we will report experimental proofs of long-term modifications and the theoretical considerations and shortcomings they suggest for rate-based theories. This will lead us to introduce recent experimental findings on spike-timing-dependent plasticity (STDP), a new local learning rule that essentially favors the first-spike times in shaping a neuron s selectivity. The implication of STDP in defining feedback connections in the visual system is to be considered in future work.

> *When an axon of cell A is near enough to excite a cell B and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased.*
> *[ D. Hebb, The organisation of behaviour, p.62, 1949]*

## Long-term modifications in the brain

An intuition at first, activity-dependent synaptic modifications, as suggested by Donald Hebb, received support from experiments exploring the long-term potentiation (LTP), Hebbian in essence, and its unforeseen, anti-Hebbian counterpart: long-term depression (LTD)[17]

LTP denotes an enduring (>1 hour) increase in synaptic efficacy that results from highfrequency stimulation of an afferent (input) pathway. Bliss and Lomo first

[17]The underlying cellular basis of activity-dependent synaptic plasticity is the subject of rich experimental investigations. For instance, both LTP and LTD seem to critically involve the postsynaptic concentration of calcium ions. However, in the present report, we will only mention biophysical mechanisms scarcely and instead focus on the effects of Hebbian learning at the system level, its impact on neurons and networks on a theoretical ground.

described it in the anesthesized rabbit hippocampus in vivo: brief high-frequency electrical stimulus, i.e. a few seconds at 100 Hz, evoked an increase in synaptic efficacy that lasted for days to weeks ([13]; see Figure 7A). Since then, LTP has been reported for excitatory synapses throughout the brain (reviewed in [9, 64] for an example in the visual system).

On the other hand, LTD stands for a decrease in synaptic weight, caused by lowfrequency stimulation (typically 1-3 Hz; see Figure 7B). Interestingly, homosynaptic LTD depression that happens on the synapse being stimulated - can occur at the same synapses that support LTP and vice-versa ([32, 80, 57]). Synapses can thus be assumed to be bidirectionally modifiable. Hence, how does activity interact with changes in synaptic efficacy? Based on the properties of homosynaptic LTP and LTD, a simple learning rule can be formulated as follows: active synapses are potentiated when their activity consistently correlates with a postsynaptic response greater than some threshold value, $\theta_{LTP}$, and are depressed when their activity correlates with a response less than $\theta_{LTP}$ but greater than some lower threshold, $\theta_{LTD}$ (Figure 7C).

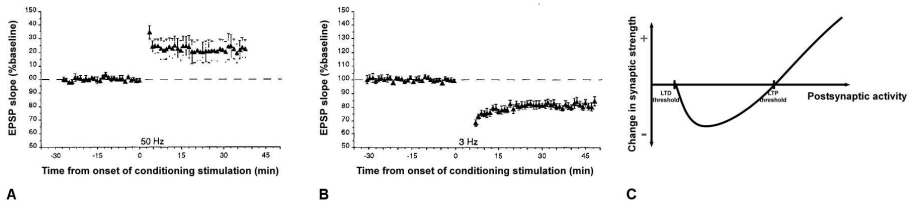



Figure 7: **Long-term modifications** Normalized average ($\pm$SEM) of experiments in which 900 pulses were delivered at different frequencies. Each point represents measures of the initial slope of the population EPSP. (A) 50 Hz Long-term potentiation (LTP) of the stimulated synapse. (B) 3 Hz Long-term depression (LTD) of the stimulated synapse (taken from [32]). (C) A synaptic learning rule based on data such as in A and B.

## Theoretical implications for rate-based theories

Clearly, long-term modifications of synaptic connections appear to depend on firing rates. A basic synaptic learning rule based on LTP, when considered under the

scope of rate-based theories, is sensitive to the spatial correlations in the input [42]. Nonetheless, the absence of an explicit LTD mechanism can be pointed out: synaptic depression is indeed simulated by a weight decay, which happens at a much shorter timescale than long-term depression.

When taking LTD into account, the synaptic learning rule (Figure 7C) is very similar to a class of synaptic modification algorithms introduced by Cooper and colleagues [23]: here the LTP threshold is the modification threshold $\theta_m$ and the LTD threshold a post-synaptic response of zero that corresponds to the mean response of uncorrelated inputs. They showed that this rule leads to the development of distributed input-selective postsynaptic responses in a simple artificial neural network.

Activity-based synaptic modifications also inspired various correlation-based rules, where both the Hebbian ( neurons that fire together wire together ) and anti-Hebbian ( neurons that fire out of sync lose their link ) components can be found. While these studies differ in assumptions, about network architecture for example, and details such as the relationship between pre- and postsynaptic activity, they do account for some aspects of cerebral development, particularly in the visual system and notably for ocular dominance columns development (e.g. [125, 12, 78, 8]).

## Limits of the original Hebb postulate

The Hebb postulate as such - increasing synaptic strength in response to activity - is a positive feedback process. The activity that modifies synapses is reinforced by Hebbian plasticity, which leads to more activity and further modification: if neuron A makes neuron B fire, then the connection between A and B is strengthened; thus, neuron A becomes more likely to make neuron B fire; their connection becomes strengthened again, and again, and again. Without appropriate adjustments of the synaptic plasticity rules or the imposition of constraints, Hebbian modification tends to produce uncontrolled growth of synaptic strengths.

The easiest way to control synaptic strengthening is to bound the values that a synaptic weight can take. Putting an upper limit is biologically supported: one can easily imagine that the presynaptic terminal, where the amount of readily releasable vesicles determines the synapse strength, has a finite capacity. As for a lower limit, plasticity processes do not change an excitatory synapse into an inhibitory synapse and vice-versa. The synaptic saturation constraint thus imposes that all synaptic weights must lie between 0 and a maximal, constant value, $w^{max}$.

But uncontrolled growth is not the only problem associated with Hebbian plasticity. Under this paradigm, synapses are assumed to be modified independantly, which can have deleterious consequences for learning: all the synaptic weights could be driven to their maximum allowed values, causing the post-synaptic neuron to lose selectivity to different input patterns. The development of input selectivity through Hebbian learning would thus require a second principle: competition between synapses, so that some weaken when others become strong [53, 77].

Which neural mechanism could implement such a mechanism? Very little is known experimentally about it, but one simple way to achieve this would be through a global intracellular signal acting as a mirror of many synapses state. Although not biophysically realistic, this idea inspired many models of Hebbian learning [79, 42] that, typically, limit the sum of synaptic strengths received by a cell, or its mean activity.

## Spike timing-dependent plasticity

A recent experimental finding can solve the problem of competition in Hebbian learning while giving interesting insights on what matters at the neuronal level to encode information: spike-timing-dependent plasticity.

It was first discovered that long-term changes in synaptic efficacy could be differentially up- or down-regulated, depending on the rather precise occurence timing of postsynaptic spikes relative to presynaptic ones: when an action potential hits a synapse before the postsynaptic neuron fires, the corresponding synapse is potentiated and inversely, when it does so after the postsynaptic response, the synapse is weakened [71]. Hence, this mechanism is indeed Hebbian (a synapse causing a response is reinforced) as well as anti-Hebbian (a synapse carrying late-arriving inputs is weakened). Further studies gave a significant role to precise spike times as it was discovered that the amount of modification sharply depended on them ([11]; Figure 8).

The theoretical implications of spike timing-dependent synaptic plasticity have since been the subject of numerous investigations. While it may be involved in temporal pattern recognition or coincidence detection, the most interesting feature of STDP lies in its competitive and stabilizing nature: it strengthens correlated inputs while being insensitive to firing rates or the degree of variability of a given synaptic input [108, 42].

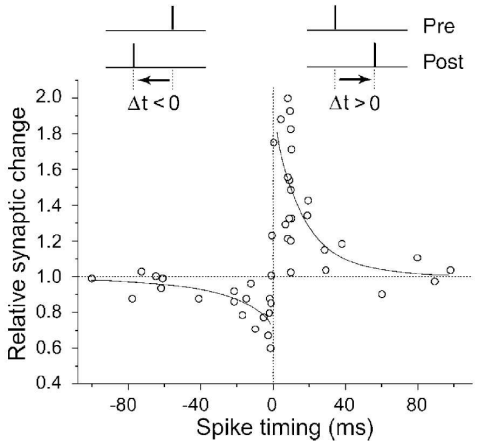



Figure 8: **Spike Timing-Dependent Plasticity** (taken from [11]) The temporal relationship between single pre- and postsynaptic spikes determines how much a synapse is modified. When the postsynaptic spike comes before the presynaptic one (left part of the graph, under the $t < 0$ ), the synapse is weakened. When the presynaptic spike contributes to the postsynaptic one (right part, under the $t > 0$ ), the synapse is reinforced. Remarkably, the amount of modification itself depends on the delay between these two events: maximal when the spikes are close together, the effect gradually decreases and disappears in excess of a few tens of milliseconds.

However, these studies have mostly been conducted using firing rates as the vector for the neural code while STDP is, intrinsically, sensitive to precise spike times. Increasing experimental evidence also goes toward a neural code based on precise spike times[117], which implies a certain degree of reproducibility in stimulus-locked responses. When spike times are taken into account and reproduced identically from one exposition to the other in the form of an asynchronous spike wave, STDP enables a neuron to find inputs carrying the very first spikes, thus considerably shortening the latency of the postsynaptic response ([108, 42]; Figure 9A-B). Remarkably, this trend tuning to the earliest spikes along with a postsynaptic latency reduction - may still arise in biologically realistic conditions as neither jitter nor spontaneous activity seem to hinder it (Figure 9B-D).

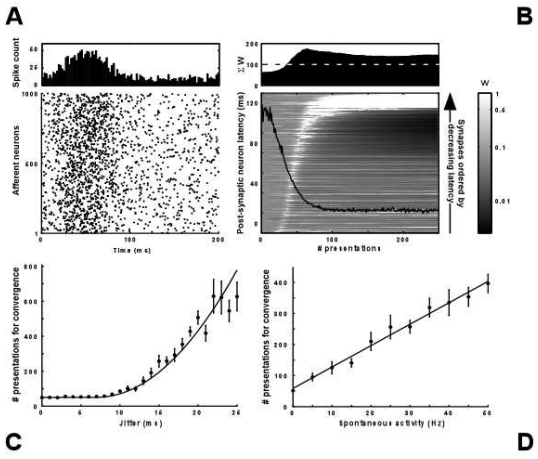



Figure 9: **STDP effects with realistic spike waves** (A) Typical input pattern. A neuron is exposed to a reproducible, jittered (5ms) spike wave amidst 5 Hz spontaneous activity. (B) Learning motion. When repeatedly exposed with more or less the same pattern, STDP is still able to find and focus on the earliest afferents while the postsynaptic neuron responds increasingly fast. The black line refers to the left axis and shows the postsynaptic latency. The image in the background depicts the evolution of synaptic weights (ordered by latency; white is for maximal potentiation, black for maximal depression). (C) Resistance to jitter. From 0 to 10 ms jitter, the effects are negligeable while after that, the convergence is slowed quadratically. (D) Resistance to spontaneous activity. Increasing its rate linearly only slows convergence to the earliest ones.

Why is that so? For one given input pattern presentation, the input spikes elicit a postsynaptic response, triggering the STDP rule. Synapses carrying input spikes just preceding the post-synaptic one are potentiated while later ones are weakened. The next time this input pattern is re-presented, firing threshold will be reached sooner which implies a slight decrease of the post-synaptic spike latency. Consequently the learning process, while depressing some synapses it had previously potentiated, will now reinforce different synapses carrying even earlier spikes than the preceding time. By iteration, it follows that upon repeating presentation of the same input spike pattern, the post-synaptic spike latency will tend to stabilize at a

minimal value while the first synapses become fully potentiated and later ones fully depressed.

Yet, two questions arise. First, regarding the input pattern, it could be argued that STDP favors the first ones only because spike waves are constituted of one spike per afferents. Does the trend still emerge if each input neuron fires more than one reproductible spike? When the input pattern is based on spike trains, that is several reproducible spikes per afferent (Figure 10A), STDP is able to selectively reinforce the earliest ones although, by starting very far from them, it has no way of knowing exactly where they are (Figure 10B). A second question concerns the effect of such a remarkable emerging structure: having a neuron responding fast is not very interesting if it is not also selective to it. The neuron was tested for postsynaptic latency in response to the target pattern compared to 50,000 distractors. It showed that, as learning goes on, the neuron becomes increasingly faster in responding to the target rather than to any distractors (Figure 10C). Conclusively, through spike-timing-dependent Hebbian learning, a neuron becomes selective to a distinct pattern by concentrating potentiation in synapses receiving its earliest, reproduced afferents [54].

## STDP in Vision Feedforward learning

So far, STDP implications have been mostly explored on a theoretical ground: be it based on rates or spike times, simulations rarely adress its function in any particular system. But the late paradigm, where a neuron is presented with an asynchronous spike wave, can be extended to the visual system in a feedforward (FF) manner.

Indeed, reproducible temporal structures can be found in MT [5], and from retinal ganglion cells to the inferotemporal cortex [10, 70, 97, 118, 82]. It thus becomes possible to investigate the effects of STDP learning on an efferent neuron receiving inputs from V1 for example. The repetition of similar input patterns could quite simply be the result of multiple exposures with the same stimulus at different times in life. Alternatively, a single stimulus exposure could result in a sequence of similar processing waves through the rhythmic activity of cortical oscillations [58]. As little as an hundred presentation could be enough to make a structured representation emerge and stabilize that is built upon the earliest afferents of the input spike wave (Figure 11).

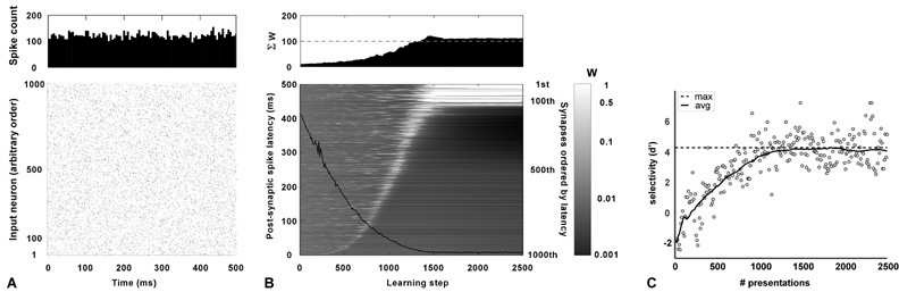



Figure 10: **STDP finds and focuses on the earliest input to be more selective** (A) Typical input pattern. The neuron is exposed to a reproducible, jittered (5ms) spike trains amidst 5 Hz spontaneous activity. The spike trains were generated according to a Poisson process. Hence, information is homogenously spread over time and afferents. (B) Learning motion. The neuron starts to fire at 400 ms after presentation of the input pattern. While it is far from the first spikes of each train, STDP will nonetheless reach for the earliest afferents to eventually focus on them while depressing later ones. (C) Selectivity. At each step in (B), the output neuron was tested for latency to the target pattern against 50,000 distractor ones. While at first the neuron responds significantly later to the target than to any distractor, it becomes steadily more selective to the target pattern used for training: it fires increasingly sooner to the target than to any distractors.

However, neurons are not isolated in the brain; it would thus seem improbable to observe such emergence in a population of neurons. Imagine we extend the preceding simulation to more input images, randomly presented to a population of as many potentially learning cells. Each time one of these cells fires, we will consider that it inhibits its neighborhood - the other cells in the population - through a feedback (FB) inhibition mechanism. Said otherwise, only one cell learns the V1-filtered spiking activity induced by the stimulus at each step. The results are remarkable: while the neurons obviously do not know which image was presented, they do yet retain one stimuli and one only (Figure 12A). When considering the reponses at the population level, one can also witness that learning makes the responses selective by implementing a sparse code (Figure 12B).

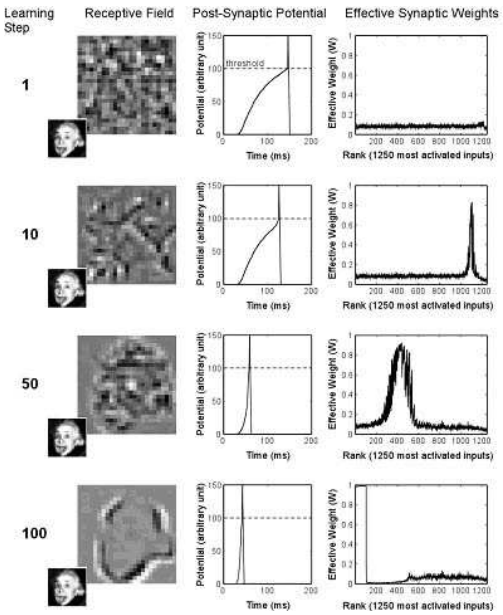



Figure 11: **Einstein: STDP learning of a V1-filtered face** A population of V1-like cells encodes an orientation for each pixel in the image presented to the network (here, Einstein s face); each cell acts as an analog-to-delay converter where the latency of its first spike depends on the strength of the orientation in its receptive field. Time taken to achieve recognition of the stimulus decreases (middle column) while a structured representation emerges and stabilizes (left column) that is built upon the earliest afferents of the input spike wave (right column). Receptive field is linearly reconstructed based on the weight synapse and the selectivity of the afferent neuron. It is thus redundant with the effective weights distribution.

This learning paradigm suggests how neurons in the visual system can be made to be selective to particular aspects of the visual scene: Hebbian learning in the form of a local, autonomous[18] learning rule such as STDP can lead to highly selective

[18]As far as STDP is concerned, the neuron s spiking activity is the main determinant of the synaptic modifi cations. These regulations in turn affect integration of the presynaptic spikes before producing postsynaptic ones. Thus, the learning rule makes the cell autonomous, at least functionally speaking.

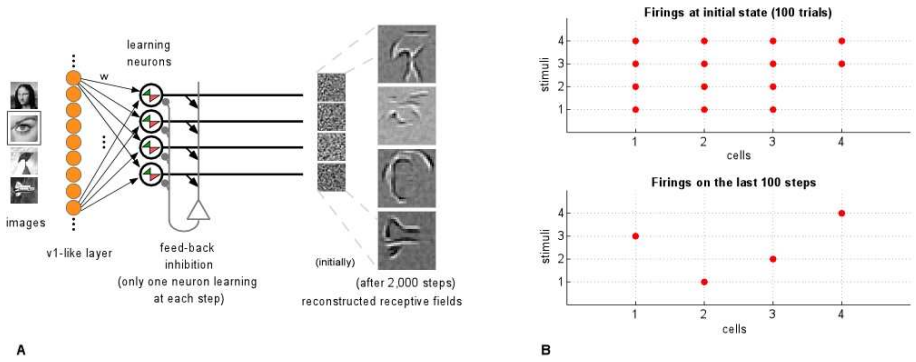



Figure 12: **Population learning: Emergence of selective responses** (A) Architecture of the network and receptive fields. At each step, one of the 4 images on the left is presented to the network. The V1-like layer generates one spike per cell just like in Figure 11. Spike times are then jittered (5 ms gaussian delay) and 5 Hz Poisson-inspired spontaneous activity is added to the spike pattern at each presentation. The changing incoming activity is propagated towards the next layer where 4 neurons integrate it. The first to fire inhibits its neigbours and triggers the STDP learning rule. After 2,000 random presentations, each neuron has learned one stimulus and one only, in the same manner as in Figure 11. (B) Population response. A red dot indicates if the neuron fired first when presented with the corresponding stimulus. (top) Initially, each neuron is likely to respond to any stimuli when tested for first response on 100 trials without STDP learning (plasticity was shut off for the test). (bottom) Finally, any input image is clearly identifiable based on which neuron fires (learning steps #1901 to #2000).

visual responses reminiscent of the notorious grandmother cells . It can be thought of as a case of supervised learning: the input world is indeed chosen by someone. It nonetheless exhibits a degree of autonomy: while the monitor chooses the input, it does not act on the regulatory process affecting the synaptic weights. In fact, such a system just self-organizes according to the world it is connected to.

Unsupervised learning would thus translate in the case where the input is natural , i.e. generated through the presentation of random patches taken from natural scenes images. When the input layer consists of retinal ganglion cells selectivities (ON- and OFF-center receptive field), the learning layer displays the emergence

of independent filters, very similar to those observed in the primary visual cortex ([89]; see also Figure 13A). In itself, this result lends further support to first-spike time coding since it coincides with an experimentally observed, fundamental fact of neuroscience (Hubel and Wiesel, 1962). In that sense, it also supports the feed-forward model of orientation selectivity in V1, based on the ordered alignment of LGN inputs [60, 29]. But the main point resides in this architecture giving an idea of how experience may shape selectivities in feed-forward processing. It becomes thus interesting to reiterate the process to investigate its effects at the next level. Indeed, if selectivities at the bottom (V1) and top (IT) of the visual hierarchy are well-documented, what lies in-between, in V2 for example, remains unclear. In the present framework, intermediate representations would thus consist of more complex structures than in V1, like long orientations, parallel bars, curvatures or vertices (Figure 13B). Building selectivities along the ventral pathway could hence lead to face-selective neurons in the higher stages of the visual hierarchy.

## STDP in Vision - FeedBack connections

The connections between different areas may be classified into 3 different types according to their preferential termination: FF connections typically terminate in layer 4, lateral ones in a columnar way and FB ones prefer a multilaminar pattern excluding layer 4. Followingly, cortical areas could be organized in a hierarchy according to the pattern of reciprocal connections between them [39]. However, it has already been suggested that the rank of a given area in the anatomical hierarchy does not necessarily correspond to that of the order of activation. For instance, area MT may be placed at the top of the anatomical hierarchy while yielding early responses that would qualify it for the earliest stages of visual processing; area MT, but also subsequent area MST and frontal eye field (FEF), would thus be part of the fast brain , where median latencies typically occur between 40 and 80 ms ; as opposed to the slow brain (100-150 ms responses) that groups most areas of the temporal lobe and some areas of the frontal cortex located rostral to the frontal eye field [84].

According to this hypothesis, latencies is seen as a dynamical principle of hierarchical organisation in the visual system, which would have consequences in refining laminar patterns of connection. We also presently shown that STDP tunes neurons to the earliest spikes they receive: latencies basically dictate which affer-

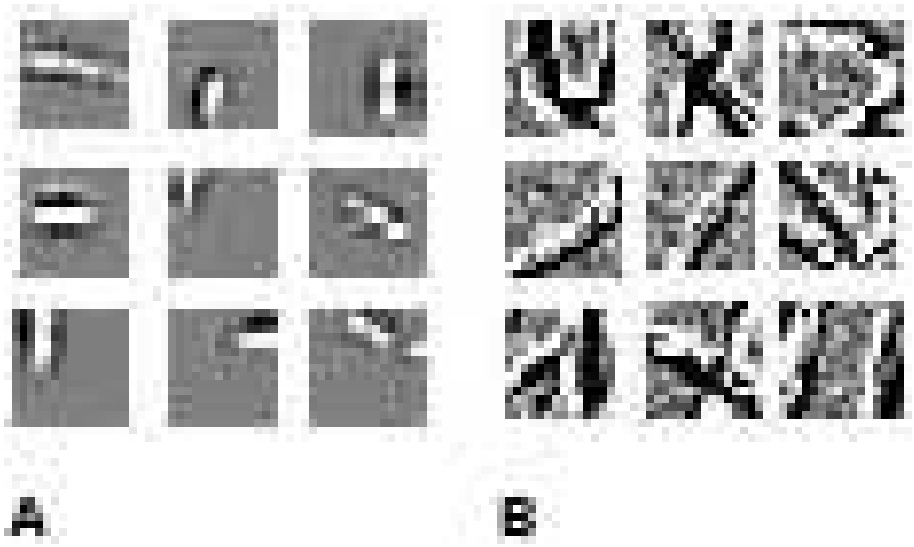



Figure 13: **Natural emergence of receptive fields** (A) When the input layer consists of ON- and OFF-center cells, analog to retinal ganglion cells, the presentation of random patches taken from natural scenes images leads to the emergence of V1-like oriented and located selectivities (sample). (B) Receptive fields emerging when the input consists in images of faces filtered by V1-inspired orientations (sample). This kind of structures may indicate what intermediary visual representations in the ventral pathway look like.

ents are selected. The similarities are such that one can wonder how STDP could influence connections between visual areas. For instance, the first, fastest wave of visual information, channelled by the magnocellular (M) pathways, reaches MT via V1 in a FF manner; from there, the M input goes backwards to V1 in time to meet the slower by 20ms, parvocellular (P) input [85]. It follows that P neurons in the superficial layers would be influenced by M inputs from MT because they arrive sufficiently early to be enhanced by STDP. Asymmetrically, M neurons having already responded on the first pass, their connections receiving MT inputs would be depressed. Hence, the M pathway would indeed act as an ignition device for the P flow in V1 [115] or even, if acting selectivity-wise too, as an active blackboard by

modulating V1 and V2 responses [18]; it would also be temporally decoupled: one M flow wound not influence the next one.

Similarly, an apparent anomaly of the visual system could be explained by STDP. According to the classical view, FEF is located high in the visual hierarchy, in a FB position compared to V4 [39]. However, FEF-V4 connections display a high %SLN - percentage of supragranular layers neurons retroactively marked when injecting in V4 -, indicative of FEF-to-V4 FF connections [7]. From the anatomical point of view, FEF initially possesses neurons in the supragranular and infragranular layers both connecting to V4 layer 4. In terms of latency responses after stimulus presentation, FEF tends to fire before V4. The first wave of activity hitting V4 would thus come from FEF supragranular neurons and elicit the earliest responses which would consequently be strengthened. At the same time, the supragranular activity in FEF propagates locally in its lower layers which would in turn fire slightly later. Inputs from the infragranular layers arriving after the first spikes in FEF, these connections would be depressed. Hence, FEF would hierarchically be placed before V4.

# 4 Functional properties of area V4 neurons *P. Girard*

Synthetic bibliographic work to refine a scientific question that needs to be studied in the framework of RIVAGE

We will only consider here the dorsal part of area V4 of the macaque monkey, which has been much more thoroughly investigated than the ventral part because of stereotaxic accessibility. Several groups studied the basic properties of V4 receptive fields in the seventies and the eighties. Early studies by [130] strongly emphasized that V4 is a colour centre since 100

More recent studies have pointed to the fact that a successive build up of more complex attributes is achieved in the ventral pathway from V2 to anterior IT. [65] have tried to determine the simplest stimuli that drive V4 cells, starting from complex stimuli and simplifying them. The outcome is that V4 neurons seem to prefer elementary shapes looking like complex textons (Figure 14). Other authors have shown selectivity of V4 cells to cartesian and noncartesian gratings [41] (Figure 15), [87]. The main issue of these studies is that visual areas can use critical features that can be more elaborated (curvatures, intersections, ..) than simple bars or gratings that are usually considered. These selectivities seem more elaborated and concern a greater proportion of cells in V4 than in V2 [55] and less elaborated and concern a smaller proportion of cells than in IT.

On the other hand, area V4 should not be considered as a mere midlevel in a mechanistic LEGO. We will first consider here the behavioural effects of lesions or inactivations of V4 followed by the analysis of several studies that have demonstrated highlevel modulation of V4 neuronal responses by attention, learning or planning of eye movements.

## Consequences of lesions or inactivations of area V4

These studies reveal consistent deficits in the perception of shapes rather than in the perception of colour, and are therefore in agreement with the physiological studies mentioned above.

Interestingly, lesion studies by [104, 105] (Figure 16) revealed a much stronger deficit for when the monkey has to detect or discriminate lowsalience objects. Whenever objects are smaller, dimmer or masked, the performances of monkeys drop when V4 is lesioned. [75] V4 lesion studies corroborate this view, since monkeys

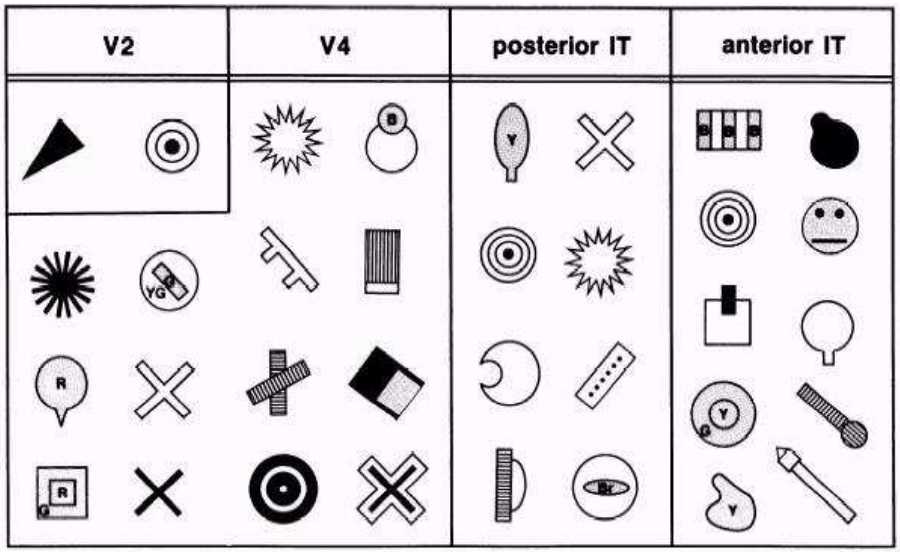



Figure 14: Example of complex critical features in 4 cortical areas. Y B G R stand for yellow, blue, green and red.

ability to perform an orientation popout task was more degraded when distracters have the same salience (more difficult) than the target. He also shows [76] that comparison of shapes that have been rotated in 3D is altered especially if the discriminations are made difficult. [126] showed that the effects of V4 lesions on a variety of discrimination tasks are particularly strong when a stimulus is embedded in distracters that are perceptually similar to the stimulus.

## Higher level functions of area V4

Visual stimuli that are difficult to see may require extra attentional load or learning processes. For instance, [73] have shown that the tuning curve of neurons for orientation is boosted up and broadened when a monkey directs its attention to a grating in a receptive field (whereas it is not when attention is directed outside the receptive field). Attention is a rather ubiquist phenomenon; as far as V4 is concerned, the role of an attention signal in V4 might be linked to perceptual learning. Some authors have demonstrated that V4 plays a role in the generalization of perceptual tasks.

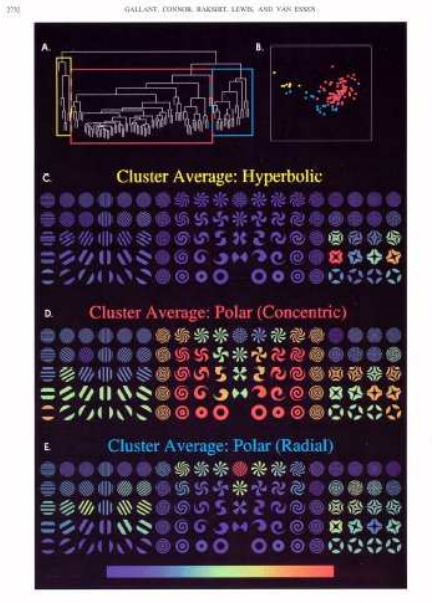



Figure 15: A: Schematic diagram of the original dendrogram with 3 cell groups. B: Metric multidimensional scaling space. C,D,E: Average angular response of the cells of each group.

[127] have shown that monkeys faculty to transfer a learned discrimination to an untrained region of the visual field is impaired by a V4 lesion. Neuronal correlates of perceptual learning have been identified by [128] who have shown that orientation tuning curves of V4 neurons are narrowed around the orientation that has been trained in a discrimination task.

Using more realistic stimuli, [91] (Figure 17 and Figure 18) have recently studied in V4 the neuronal correlates of perceptual learning in macaque monkeys. Macaque monkeys are trained to discriminate pictures of natural objects that are made barely visible because of noise addition. They observed a major effect of learning consisting in an increase of neural information for degraded images, in correlation with better behavioural performances.

In conclusion, V4 seems to play an important role for enhancing behavioural responses to stimuli that may be difficult to see whatever could be the source of

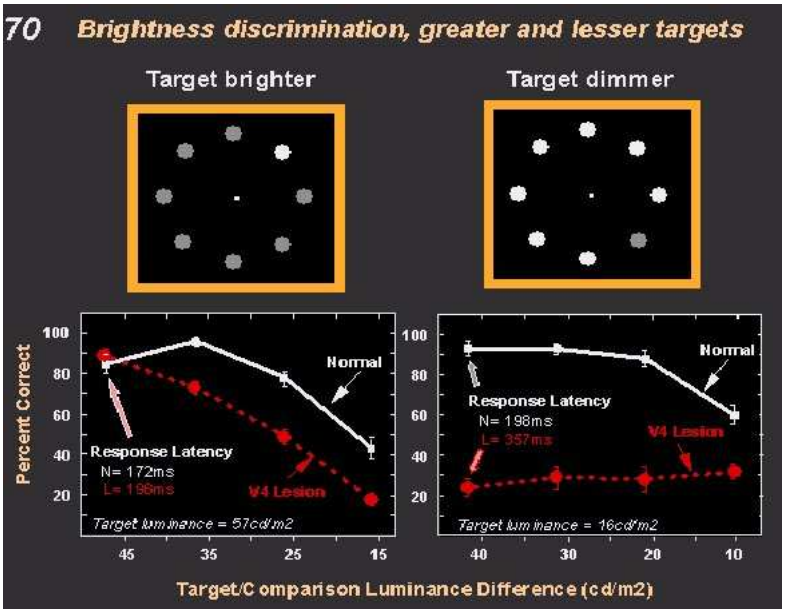



Figure 16: Effect on behavioural performances of macaque monkeys with V4 lesions: When the target is dimmer than the distractors, latencies increase and the performance (saccade to target drops).

low salience. Since the basic low level features that drive V4 neurons are complex (see above), their processing is more likely to be disrupted in a cluttered environnement like natural visual scenes. An attentional signal is required and seems to be present that helps to process such stimuli, in some cases, perceptual learning reinforces the process [6]. Lesions disrupt these processes (although it remains open that reversible inactivation experiments which leave no time for plasticity, reveal stronger deficits, [46]. An analogy can be made with [61] experiments on inactivation of feedback pathways. When feedback pathway from MT to lower extrastriate areas is inactivated, processing of low visibility stimuli is disrupted in those areas that receive the feedback. Hence, corticocortical feedback pathway could act as a gain controller to enhance lowsalience stimuli. Although this set of experiments concerned the dorsoparietal pathway, one can suggest that the same mechanism could be present in the ventroinferotemporal pathway. In that case, IT to V4 feed-

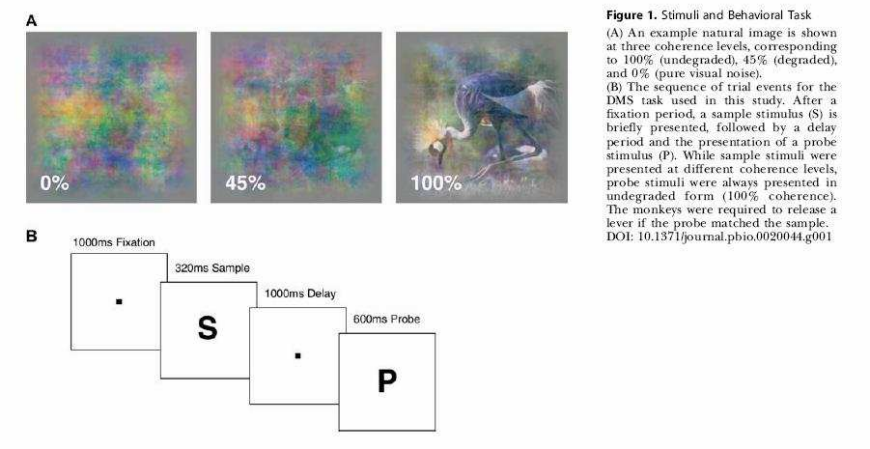



**Figure 1.** Stimuli and Behavioral Task
(A) An example natural image is shown at three coherence levels, corresponding to 100% (undegraded), 45% (degraded), and 0% (pure visual noise).
(B) The sequence of trial events for the DMS task used in this study. After a fixation period, a sample stimulus (S) is briefly presented, followed by a delay period and the presentation of a probe stimulus (P). While sample stimuli were presented at different coherence levels, probe stimuli were always presented in undegraded form (100% coherence). The monkeys were required to release a lever if the probe matched the sample.
DOI: 10.1371/journal.pbio.0020044.g001

Figure 17: A: An example of natural image is shown at three coherence level, corresponding different degradation of the stimulus. B: Sequence of trial events for a DMS task, while the sample stimulus was presented at different level of degradation, the probe stimulus was always degraded.

back would play a role in enhancing behavioural performance. The functional hypotheses about the role of V4 / IT connectivity hence will be as stated above (autre chapitre du projet..) , a first pass of visual processing would consist in bringing up to IT elementary complex descriptors to detect or categorize a visual object. In case of low visibility or active search, a feedback signal should be necessary to segment this object from the visual scene, keep tracking it or even learn to process it more efficiently.

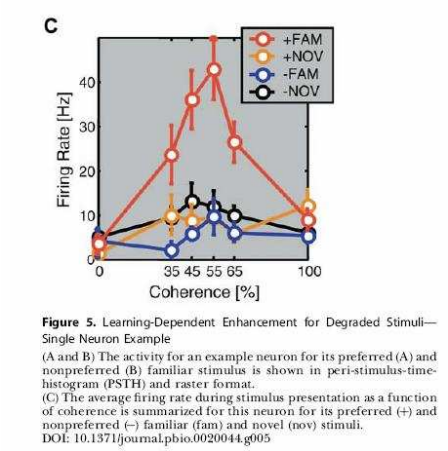



**Figure 5.** Learning-Dependent Enhancement for Degraded Stimuli—Single Neuron Example

(A and B) The activity for an example neuron for its preferred (A) and nonpreferred (B) familiar stimulus is shown in peri-stimulus-time-histogram (PSTH) and raster format.
(C) The average firing rate during stimulus presentation as a function of coherence is summarized for this neuron for its preferred (+) and nonpreferred (–) familiar (fam) and novel (nov) stimuli.
DOI: 10.1371/journal.pbio.0020044.g005

Figure 18: Learning-dependent enhancement for degraded stimuli: single neuron example. The average firing rate during stimulus presentation as a function of coherence, in the prefered (+), non preferred (-) direction of the neuron, for familiar (fam) or novel (nov) stimuli.

# 5 Conclusion and perspectives *S. Thorpe et al.*

These four contributions allow to describe what is the next step regarding "integrated model of visual processing". One goal is to focus on the "what" (i.e. object recognition) stream, more precisely perceptual grouping and object segmentation [67]. For instance, [50] attempts to demonstrate how the known laminar architecture of the V1 and V2 areas of the visual cortex, assuming a functional role for this stratification, is involved in percepts generation. This includes pre-attentive/attentive aspects, as pointed out by [84]. The parvocellular stream of the visual cortex performs visual filtering, i.e. attention and perceptual grouping, using feed-forward, feedback and horizontal interactions. One key aspect is the figure / ground segmentation in the brain, as recently reviewed by [101] showing that shape perception depends critically on this cue, while its link to early-vision mechanisms is now relatively well understood e.g. the role of junctions in surface completion and contour matching. Following this track the present project focuses on subsequent layers of the visual stream : V4/IT.

It is now clarified [67, 18] that, regarding the adaptive mechanisms described previously, the formalization in artificial visual processes using the so-called PDE (Partial Differential Equation) or "variational approach", allows to model how a biological visual system, as a whole, solves perceptual tasks not only at a "pixel" level. A key aspect is that, under this assumption, there is a direct link between PDE parameterization and spatio-temporal maps of cortical activity. Therefore we plan to find biological networks of neurons which carry out some of the PDE-based computations used in algorithmic vision and conversely, starting from neuro-physiological data, to try to define pertinent PDEs acting on visual inputs. This should allow the development of not only anatomical but also functional models of activity in cortical areas during a visual task.

All three previous aspects of visual perception are easily formalized using the "generative model" approach reviewed in [40]. This allows to relate what is actually known in statistical modeling with existing biological models. A general framework for the role of feedback connections is proposed in these publications. A step further, the link with variational approaches has been sketched out (section 2), analyzing how a certain class of PDE vis feedbacks. Regarding the role of feedback connections, [93] describe a model of visual processing in which feedback connections from a higher- to a lower- order visual cortical area carry predictions of

lower-level neural activities, whereas the feed-forward connections carry the residual errors between the predictions and the actual lower-level activities. In the scope of this approach receptive fields are Gabor-like filters with a sigmoid profile output, weights being optimized by a 1st order gradient optimization of a likelihood criterion. This is now generalized to more plausible operators. But the relation with sensory-motor strategies, is very important: these authors point out the fact that visual cognition depends heavily on sensory-motor mechanisms. They also analyze how saliency maps and more generally attentional effects are to be taken into account: this is still to be done in our approach.

Back to the real brain, a functional theory of this connectivity has been proposed by [18, 84]: he proposes that a first wave of computation allows an initial estimate of the global aspect of objects contained in the visual scene. This is made possible because of the fast latencies of feedforward connections (magnocellular pathway). The two kinds of connections termed feedforward and feedback mainly from anatomical considerations about sources and arrival of projections are involved in an iterative refinement of processing is taking place to come up with an efficient perception taking into account local components of the visual scene. To achieve this, one can make the hypothesis of a mechanism based upon feedback connections. Feedback connections could redirect global information after a selection by attention for instance and activate long latency feedforward (parvocellular pathway) that carries more detailed information (color, texture, ..). This is the basis of this common preliminary work.

## Perspectives

Numerous studies have shown that the responses of inferotemporal cortex neurons are selective to complex stimuli like faces [88] or trees [124]. Latencies of these responses can be very short, below 100 ms. [109] have shown that the first part of the response can correspond to stimulus global information (one can crudely identify that a face has been presented). The later part of the response seem to be selective to more local details of the stimulus, like facial expression. Fast activation of inferotemporal cortex is in agreement with psychophysical and electrophysiological studies (EEG) that demonstrate both in human and macaque monkeys, the very fast ability to categorize pictures of natural objects [37].

Such a duality of inferotemporal responses fits with above-mentioned iterative functional model of cortico-cortical connections. A first pass aims at globally or crudely detecting the object ; in turn feedback connections coordinate responses of lower stages to specify local attributes belonging to the object (color, texture, components) which then will be transmitted to inferotemporal cortex (for instance for memory storage, template matching etc). Area V4, which receives important feedback connectivity from IT and contains neurons selective to elementary attributes such as non Cartesian gratings and textures, is an interesting area to test this hypothesis.

The role of feedback connectivity has been studied by [61]. They have shown, in anesthetized monkeys, that feedback connections from area MT can control the gain of visual responses in areas V1, V2 et V3. This gain control helps figure-ground detection, especially if the stimulus has low visibility (low contrast, visual noise). This function of feedback connections remains to be questioned in awake macaque monkeys involved in a behavioural task. Our objectives fall into two broad classes, computational and neurophysiological, as detailed below.

**Modeling and simulating the feedback control between IT and V4 using a large spiking neuronal network.**

We propose to use the PDE formalism in order to model and simulate some of the neuronal computations in V1-V4 and IT for a paradigm for visual segmentation and object recognition (similar ideas have already been applied to V1 and V2, as discussed in e.g. [67] where a preliminary implementation related to the PDE formalism has been proposed and experimented). The functional assumption is that fast-brain mechanisms allow us to recognize objects (in fact to build assumptions about the kind of object(s) present in the scene) as discussed in e.g. [119]. This is based on the hypothesis that object segmentation is driven by such top-down assumptions. More precisely, once the fast-brain pathway to IT has detected the presence of a class of objects, it sends feedback information to some of the lower stages V1-V4 to guide the segmentation process from elementary image features. Similar ideas have been recently used in computer vision [25, 99, 14, 69] to constrain image segmentation with object priors, but have to be developed further in this context. The introduction of the PDE formalism should allow us to formalize this mechanism as a metaphore of a feedback between IT (object “identification” layer)

and V4 (object "attribute grouping" layer) and to propose a biologically plausible implementation. Schematically, this implementation will proceed in two phases. In the first phase the class of object present in the image will be identified by applying the Support Vector Machine formalism [119] to vectors built from early neuronal responses [37, 110]. This will index a data base of silhouette models at different scales that will then be used as priors to drive a PDE based image segmentation process. The conversion between continuous quantities required by the PDE formalism and spike trains can be achieved through various processes, the simplest one being based on such ideas as the elementary cortical column model of Jansen and Rit [62]. This is an open question and the present researchs will focus on it. The validation of this approach will be conducted (a) at the computational level by measuring the performances of the system on real images and by comparing them to published results and (b) at the neurophysiological level by a so-called probing experiment described next.

**Probing the function of feedback from inferotemporal cortex to V4.**

By analogy with HupÃ© et al experiments on feedback and low visibility, one can suggest that IT to V4 feedback plays a role in enhancing behavioral performance in categorization in low visibility. It has been recently shown by [91] that the increase in the amount of information transmitted by V4 cells parallels the behavioral performance enhancement seen when the animal learns to identify natural images that are made difficult to see by noise addition. If our hypothesis is correct, resistance to image degradation and identification performance should drop when IT is inactivated in a categorization task. We will conjointly search for neuronal correlates in V4 during IT inactivation giving special care about the global and local aspects of the responses.

As a conclusion, this first step indicates that there are many but rather different frameworks to model the ventral/dorsal visual pathways functionalities, while these correspond to recent computer vision mechanisms, now formalized in a unifying "variational" framework. It is thus a challenge to carry on investigating to what extend this unifying framework could be relevant in modeling biological vision. This is the goal of the next step of this work.

# References


[1] D. Adams and S. Zeki. Functional organization of macaque V3 for stereoscopic depth. *J. Neurophysiol.*, 86:2195–2203, 2001.

[2] L. Alvarez and J. Morel. Formalization and computational aspects of image analysis. *Acta Numerica*, pages 1–59, 1994.

[3] A. Angelucci and J. Bullier. Reaching beyond the classical receptive field of V1 neurons: horizontal or feedback axons? *J. Physiol. (Paris)*, 2002.

[4] G. Aubert and P. Kornprobst. *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, volume 147 of *Applied Mathematical Sciences*. Springer-Verlag, Jan. 2002.

[5] W. Bair and C. Koch. Temporal precision of spike trains in extrastriate cortex of the behaving macaque monkey. *Neural Comput*, 8:1185–1202, 1996.

[6] M. Bar. A cortical mechanism for triggering top-down facilitation in visual object recognition. *J Cogn Neurosci*, 15:600–609, 2003.

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

[8] M. Bear. The role of ltp and ltd in development and learning. In *Mechanistic relations between development and learning*. John Wiley and Sons Ltd., 1998.

[9] M. Bear and A. Kirkwood. Bidirectional plasticity of cortical synapses. In *Cortical Plasticity: LTP and LTD.*, pages 191–205. Oxford: Bios Scientific, 1996.

[10] M. Berry, D. Warland, and M. Meister. The structure and precision of retinal spike trains. *Proc Natl Acad Sci USA*, 94:5411–5416, 1997.

[11] G. Bi and M. Poo. Activity-induced synaptic modifications in hippocampal culture: dependence on spike timing, synaptic strength and cell type. *J Neurosci*, 18:10464–10472, 1998.

[12] E. Bienenstock, L. Cooper, and P. Munro. Theory for the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex. *J Neurosci*, 2:32–48, 1982.

[13] T. Bliss and T. Lomo. Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *J Physiol.*, 232:331–356, 1973.

[14] E. Borenstein and S. Ullman. Class-specic, top-down segmentation. In *ECCV 2002*. Springer Verlag, Berlin, 2002.

[15] C. Büchel and K. Friston. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fmri. *Cereb. Cortex*, 7:768–778, 1997.

[16] G. Bugmann. Biologically plausible neural computation. *Biosystems*, 40:11–19, 1997.

[17] J. Bullier. Feedback connections and conscious vision. *Trends in Cognitive Science*, 5(9):369–370, 2001.

[18] J. Bullier. Integrated model of visual processing. *Brain Res. Reviews*, 36:96–107, 2001.

[19] J. Bullier. *Neural Basis of Vision*, chapter 1, pages 1–40. Stevens' Handbook of Experimental Psychology. New York: Wiley, 2002.

[20] Y. Burnod. *An adaptive neural network: the cerebral cortex*. Masson, Paris, 1993. 2nd edition.

[21] C. Caudek and N. Rubin. Segmentation in structure from motion: modeling and psychophysics. *Vision Research*, 41:2715–2732, 2001.

[22] J. Chey, S. Grossberg, and E. Mingolla. Neural dynamics of motion processing and speed discrimination. *Vision Res.*, 38:2769–2786, 1997.

[23] L. Cooper. How we learn; how we remember: toward an understanding of brain and neural systems. In *Selected papers of Leon N. Cooper.* London: World scientifi c, 1995.

[24] R. Courant. *Calculus of Variations*. New York, 1946.

[25] D. Cremers, T. Kohlberger, and C. S. orr. Nonlinear shape statistics in mumford shah based segmentation. In *ECCV*, pages 93–108, 2002.

[26] J. G. Daugman. *Brain metaphor and brain theory*, chapter 2. Blackwell Publishers, Oxford, 2001.

[27] P. Dayan and L. F. Abbott. *Theoretical Neuroscience : Computational and Mathematical Modeling of Neural Systems*. MIT Press, 2001.

[28] P. Degond and S. Mas-Gallic. The weighted particle method for convection-diffusion equations. *Mathematics of Computation*, 53(188):485–525, 1989.

[29] A. Delorme, L. Perrinet, and S. Thorpe. Network of integrate-and-fi re neurons using rank order coding b: spike timing dependant plasticity and emergence of orientation selectivity. *Neurocomputing*, 38:539–545, 2001.

[30] R. Deriche and O. Faugeras. Les EDP en Traitement des Images et Vision par Ordinateur. *Traitement du Signal*, 13(6), 1996.

[31] R. Deriche, O. Faugeras, G. Giraudon, T. Papadopoulo, and R. Vaillant. Four applications of differential geometry to computer vision. In G. Orban and H.-H. Nagel, editors, *Artifi cial and Biological Vision Systems*, Basic Research Series, pages 93–141. Springer–Verlag, 1993.

[32] S. Dudek and M. Bear. Homosynaptic long-term depression in area ca1 of hippocampus and effects of n-methyl-d-aspartate receptor blockade. *Proc Natl Acad Sci USA*, 89:4363–4367, 1992.

[33] R. Durbin, C. Miall, and G. Mitchinson, editors. *The computing neuron*. Addison-Wesley, 1989.

[34] C. Eliasmith and C. H. Anderson. *Neural Engineering: Computation, Representation and Dynamics in Neurobiological Systems*. MIT Press, 2003.

[35] D. V. Essen and J. Maunsell. Hierarchical organization and functional streams in the visual cortex. *Trends in Neurosciences*, 6(9), 1983.

[36] L. Evans. *Partial Differential Equations*, volume 19 of *Graduate Studies in Mathematics*. Proceedings of the American Mathematical Society, 1998.

[37] M. Fabre-Thorpe, G. Richard, and S. Thorpe. Rapid categorization of natural images by rhesus monkeys. *Neuroreport*, 9(2):303–308, 1998.

[38] O. Faugeras. *Three-Dimensional Computer Vision: a Geometric Viewpoint*. MIT Press, 1993.

[39] D. Felleman and D. V. Essen. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*, 1:1–47, 1991.

[40] K. Friston. Functional integration and inference in the brain. *Prog Neurobiol*, 68:113–143, 2002.

[41] J. Gallant, C. Connor, S. Rakshit, J. Lewis, and D. V. Essen. Neural responses to polar, hyperbolic, and cartesian gratings in area v4 of the macaque monkey,. *J Neurophysiol*, 76:2718–2757, 1996.

[42] W. Gerstner and W. M. Kistler. Mathematical formulations of hebbian learning. *Biol Cybern*, 87:404–415, 2002.

[43] M. Giese and M. Lappe. Measurement of generalization fi elds for the recognition of biological motion. *Vision Research*, 38:1847–1858, 2002.

[44] M. Giese and T. Poggio. Neural mechanisms for the recognition of biological movements and actions. *Nature Reviews Neuroscience*, 4:179–192, 2003.

[45] P. Girard and J. Bullier. Visual activity in area v2 during reversible inactivation of area 17 in the macaque monkey. *J. Neurophysiol.*, 62(6):1287–1301, 1989.

[46] P. Girard, S. Lomber, and J. Bullier. Shape discrimination defi cits during reversible deactivation of area v4 in the macaque monkey,. *Cereb Cortex*, 12:1146–1156, 2002.

[47] T. Gisiger, S. Dehaene, and J. P. Changeux. Computational models of association cortex. *Curr. Opin. Neurobiol.*, 10:250–259, 2000.

[48] N. H. Goddard, M. Huckab, F. Howell, H. Cornelis, K. Shankar, and D. Beeman. Towards neuroml: Model description methods for collaborative modelling in neuroscience. *Philosophical Transactions of the Royal Society*, 356(1412):1209–1228, 2001.

[49] S. Grossberg and N. McLoughlin. Cortical dynamics of three-dimensional surface perception: binocular and half-occluded scenic images. *Neural Networks*, 10(9):1583–1605, 1997.

[50] S. Grossberg, E. Mingolla, and W. D. Ross. Visual brain and visual perception: how does the cortex do perceptual grouping? *Trends in Neurosciences*, 20(3):106–111, 1997.

[51] A. Grunewald and S. Grossberg. Self-organization of binocular disparity tuning by reciprocal corticogeniculate interactions. *Journal of Cognitive Neuroscience*, 10:199–215, 1998.

[52] F. Guichard and J.-M. Morel. *Image analysis and P.D.E.'s*. Tutorials on Geometrically Based Motion, IPAM, Ucla, Los Angeles, 2001.

[53] R. Guillery. Binocular competition in the control of geniculate cell growth. *J Comp Neurol*, 144:117–146, 1972.

[54] R. Guyonneau, R. vanRullen, and S. Thorpe. Neurons tune to the earliest spikes through stdp. *Neural Computation*, 2004. In review.

[55] J. Hegde and D. V. Essen. Selectivity for complex shapes in primate visual area v2,. *J Neurosci*, 20(RC61), 2000.

[56] G. Hermosillo. *Variational Methods for Multimodal Image Matching*. PhD thesis, INRIA, The document is accessible at ftp://ftp-sop.inria.fr /robotvis/html/Papers /hermosillo:02.ps.gz, 2002.

[57] A. Heynen, W. Abraham, and M. Bear. Bidirectional modification of ca1 synapses in the adult hippocampus in vivo. *Nature*, 381:163–166, 1996.

[58] J. Hopfield. Pattern recognition computation using action potential timing for stimulus representation. *Nature*, 376:33–36, 1995.

[59] D. Hubel. *L'oeil, le cerveau et la vision : les étapes cérébrales du traitement visuel*. L'univers des sciences. Pour la science, 1994.

[60] D. Hubel and T. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat visual cortex. *J Physiol*, 160:106–154, 1962.

[61] J. Hupe, A. James, B. Payne, S. Lomber, P. Girard, and J. Bullier. Cortical feedback improves discrimination between figure and background by v1, v2 and v3 neurons. *Nature*, 394:784–791, 1998.

[62] B. H. Jansen and V. G. Rit. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol. Cybern.*, 73:357–366, 1995.

[63] J. Jost. *Riemannian geometry and geometric analysis*. Springer Verlag, 2002.

[64] A. Kirkwood and M. Bear. Hebbian synapses in the visual cortex. *J Neurosci*, 14:1634–1645, 1994.

[65] E. Kobatake and K. Tanaka. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex,. *J Neurophysiol*, 71:856–923, 1994.

[66] J. J. Koenderink. The brain as a geometry engine. *Psychol. Res.*, 52:122–127, 1990.

[67] I. Kokkinos, R. Deriche, P. Maragos, and O. Faugeras. A biologically motivated and computationally tractable model of low and mid-level vision tasks. In *Proceedings Eighth European Conference on Computer Vision,*, Prague, May 2004.

[68] A. Leonard. Vortex methods for flow simulations. *J. Comput. Phys*, 37:289–335, 1980.

[69] M. Leventon, E. Grimson, O. Faugeras, S. Wells, and R. Kikinis. Knowledge-based segmentation of medical images. In S. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision, and Graphics*. Springer-Verlag, 2003.

[70] R. Liu, S. Tzonev, S. Rebrik, and K. Miller. Variability and information in a neural code of the cat lateral geniculate nucleus. *J Neurophysiol*, 86:2789–2806, 2001.

[71] H. Markram, J. Lübke, and etal. Regulation of synaptic efficiency by coincidence of postsynaptic aps and epsps. *Science*, 275:213–215, 1997.

[72] D. Marr. *Vision*. W.H. Freeman and Co., 1982.

[73] C. McAdams and J. Maunsell. Effects of attention on orientationtuning functions of single neurons in macaque cortical area v4. *J Neurosci*, 19:431–441, 1999.

[74] N. McLoughlin and S. Grossberg. Cortical computation of stereo disparity. *Vision Res*, 38(1):91–99, 1998.

[75] W. Merigan. Basic visual capacities and shape discrimination after lesions of extrastriate area v4 in macaques. *Vis Neurosci*, 13:51–60, 1996.

[76] W. Merigan and H. Pham. V4 lesions in macaques affect both single and multipleviewpoint shape discriminations. *Vis Neurosci*, 15:359–367, 1998.

[77] K. Miller. Synaptic economics: competition and cooperation in synaptic plasticity. *Neuron*, 17:371–374, 1996.

[78] K. Miller, J. Keller, and M. Stryker. Ocular dominance column development: analysis and simulation. *Science*, 245:605–615, 1989.

[79] K. Miller and D. MacKay. The role of constraints in hebbian learning. *Neural Comp*, 6:100–126, 1996.

[80] R. Mulkey, C. Herron, and R. Malenka. An essential role for protein phosphatases in hippocampal long-term depression. *Science*, 261:1051–1055, 1993.

[81] H. Nagel and W. Enkelmann. An investigation of smoothness constraint for the estimation of displacement vector fi elds from image sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8:565–593, 1986.

[82] K. Nakamura. Neural processing in the subsecond time range in the temporal cortex. *Neural Comput*, 10:567–595, 1998.

[83] H. Neumann and E. Mingolla. Computational neural models of spatial integration in perceptual grouping. In T. . P. Kellman, editor, *From Fragments to Objects: Grouping and Segmentation in Vision*, pages 353–400. Amsterdam: Elsevier, 2001.

[84] L. Nowak and J. Bullier. *The Timing of Information Transfer in the Visual System*, volume 12 of *Cerebral Cortex*, chapter 5, pages 205–241. Plenum Press, New York, 1997.

[85] L. Nowak, M. Munk, P. G. P, and J. Bullier. Visual latencies in areas v1 and v2 of the macaque monkey. *Vis Neurosci*, 12:371–387, 1995.

[86] S. Osher and R. P. Fedkiw. *Level set methods : overview and recent results*. Tutorials on Geometrically Based Motion, IPAM, Ucla, Los Angeles, 2001.

[87] A. Pasupathy and C. Connor. Responses to contour features in macaque area v4. *J Neurophysiol*, 82:2490–2502, 1999.

[88] D. Perrett, E. Rolls, and W. Caan. Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Research*, 47:329–342, 1982.

[89] L. Perrinet. Emergence of fi lters from natural scenes in a sparse spike coding scheme. *Neurocomputing*, 58–60:821–826, 2004.

[90] J. Petitot and Y. Tondut. Vers une neuro-géométrie. fi brations corticales, structures de contact et contours subjectifs modaux. *Mathématiques, Informatique et Sciences Humaines*, 145:5–101, 1999.

[91] G. Rainer, H. Lee, and N. K. Logothetis. The effect of learning on the function of monkey extrastriate visual cortex. *PLOS Biology*, 2(2):275–283, 2004.

[92] R. Raizada and S. Grossberg. Towards a theory of the laminar architecture of the cerebral cortex: Computational clues from the visual system. *Cerebral Cortex*, 13:100–113, 2003.

[93] R. Rao and D. Ballard. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-fi eld effects. *Nat Neurosci*, 2(1):79–87, 1999.

[94] R. Rao and T. J. Sejnowski. Spike-timing-dependent hebbian plasticity as temporal difference learning. *Neural Comput.*, 13(10):2221–2237, 2001.

[95] R. Rao, G. Zelinsky, M. M. Hayhoe, and D. Ballard. Eye movements in iconic visual search. *Vision Research*, 42(11):447–1463, 2002.

[96] P. Raviat. An analysis of the particle methods. In F. Brezzi, editor, *Numerical Methods in Fluid Dynamics*, volume 1127 of *Lecture Notes in Math.*, pages 243–324. Springer Verlag, Berlin, 1985.

[97] B. Richmond and L. Optican. Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. ii. information transmission. *J Neurophysiol*, 64:370–380, 1990.

[98] K. Rockland and D. Pandya. Laminar origins and terminations of cortical connections in the occipital lobe in the rhesus monkey. *Brain Res*, 179:3–20, 1979.

[99] M. Rousson and N. Paragios. Shape priors for level set representations. In A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, editors, *Proceedings of the 7th European Conference on Computer Vision*, volume 2, pages 78–92, Copenhagen, Denmark, May 2002. Springer–Verlag.

[100] N. Rubin. Figure and ground in the brain. *Nature Neuroscience*, 4:857–858, 2001.

[101] N. Rubin. The role of junctions in surface completion and contour matching. *Perception*, 30:339–366, 2001.

[102] P. Salin and J. Bullier. Corticocortical connections in the visual system: structure and function. *Psychol. Bull.*, 75:107–154, 1995.

[103] J. Sandell and P. Schiller. Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *J. Neurophysiol.*, 48:38–48, 1982.

[104] P. Schiller. The effects of v4 and middle temporal (mt) area lesions on visual performance in the rhesus monkey. *Vis Neurosci*, 10:717–746, 1993.

[105] P. Schiller. Effect of lesions in visual cortical area v4 on the recognition of transformed objects. *Nature*, 376:342–444, 1995.

[106] L. Schwartz. *Théorie des distributions*. Hermann, 1957.

[107] E. P. Simoncelli and D. Heeger. A model of neuronal responses in visual area mt. *Vision Research*, 38:743–761, 1998.

[108] S. Song, K. D. Miller, and L. F. Abbott. Competitive hebbian learning through spike-timing-dependent synaptic plasticity. *Nature Neurosci*, 3(9):919–926, 2000.

[109] Y. Sugase, S. Yamane, S. Ueno, and K. Kawano. Global and fi ne information coded by single neurons in the temporal visual cortex. *Nature*, 400:869–873, 1999.

[110] S. Thorpe and M. Fabre-Thorpe. Seeking categories in the brain. *Science*, 291:260–263, 2001.

[111] S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381:520–522, 1996.

[112] A. Tikhonov. Regularization of incorrectly posed problems. *Soviet. Math. Dokl.*, 4:1624–1627, 1963.

[113] D. Tschumperlé. *PDE's Based Regularization of Multivalued Images and Applications*. PhD thesis, Université de Nice-Sophia Antipolis, Dec. 2002.

[114] D. Tschumperlé and R. Deriche. Vector-valued image regularization with PDE's : A common framework for different applications. In *IEEE Conference on Computer Vision and Pattern Recognition*, Madison, Wisconsin (United States), June 2003.

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

[116] D. Van-Essen. Organization of visual areas in macaque and human cerebral cortex. In L. Chapula and J. Werner, editors, *The Visual Neurosciences*. MIT Press, 2003.

[117] R. VanRullen, R.Guyonneau, and S. Thorpe. Spike times make sense. *Trends in Neurosciences*, 2004. In review.

[118] J. Victor and K. Purpura. Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J Neurophysiol*, 76:1310–1326, 1996.

[119] T. Viéville and S. Thorpe. A deterministic biologically plausible classifier. In *8th ICCNS*. Boston University, 2004.

[120] T. Viéville. Biologically plausible regularization mechanisms. RR 4625, INRIA, 2002.

[121] T. Viéville. Towards biologically plausible regularization mechanisms. RR 4965, INRIA, 2003.

[122] T. Viéville, E. Clergue, R. Enciso, and H. Mathieu. Experimentating with 3-D vision on a robotic head. *Robotics and Autonomous Systems*, 1995. 14(1).

[123] T. Viéville, D. Lingrand, and F. Gaspard. Implementing a multi-model estimation method. *The International Journal of Computer Vision*, 44(1), 2001.

[124] R. Vogels. Categorization of complex visual images by rhesus monkeys. part 1: behavioural study. *Eur J Neurosci*, 4:1223–1238, 1999.

[125] C. von der Malsburg. Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14:85–100, 1973.

[126] P. D. Weerd, R. Desimone, and L. Ungerleider. Generalized deficits in visual selective attention after v4 and teo lesions in macaques,. *Eur J Neurosci*, 18:1671–1762, 2003.

[127] P. D. Weerd, R. Desimone, and L. Ungerleider. Impairments in spatial generalization of visual skills after v4 and teo lesions in macaques (macaca mulatta). *Behav Neurosci*, 117:1441–1448, 2003.

[128] T. Yang and J. Maunsell. The effect of perceptual learning on neuronal responses in monkey visual area v4. *J Neurosci*, 24:1617–1626, 2004.

[129] A. J. Yu, M. Giese, and T. Poggio. Biophysiologically plausible implementations of maximum operation. *Neural Computation*, 14(12), 2003.

[130] S. Zeki. Colour coding in rhesus monkey prestriate cortex. *Brain Res*, 53:422–429, 1973.

[131] S. Zeki and S. Shipp. The functional logic of cortical connections. *Nature*, 335:311–316, 1988.

*This Odysee / Cerco common work has been realized within the scope of the RIVAGe project (ACI)*



Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Futurs : Parc Club Orsay Université - ZAC des Vignes
4, rue Jacques Monod - 91893 ORSAY Cedex (France)
Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifi que
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38334 Montbonnot Saint-Ismier (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399