



SOMMA: Cortically Inspired Paradigms for Multimodal Processing

Mathieu Lefort, Yann Boniface, Bernard Girau

► To cite this version:

Mathieu Lefort, Yann Boniface, Bernard Girau. SOMMA: Cortically Inspired Paradigms for Multimodal Processing. International Joint Conference on Neural Networks, Aug 2013, Dallas, United States. hal-00859986

HAL Id: hal-00859986

<https://hal.inria.fr/hal-00859986>

Submitted on 9 Sep 2013

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.

SOMMA: Cortically Inspired Paradigms for Multimodal Processing

Mathieu Lefort, Yann Boniface and Bernard Girau

LORIA, Campus Scientifique, BP 239, 54506 Vandœuvre-lès-Nancy Cedex, France
{mathieu.lefort, yann.boniface, bernard.girau}@loria.fr

Abstract—SOMMA (Self Organizing Maps for Multimodal Association) consists on cortically inspired paradigms for multimodal data processing. SOMMA defines generic cortical maps - one for each modality - composed of 3-layers cortical columns. Each column learns a discrimination to a stimulus of the input flow with the BCMu learning rule [26]. These discriminations are self-organized in each map thanks to the coupling with neural fields used as a neighborhood function [25]. Learning and computation in each map is influenced by other modalities thanks to bidirectional topographic connections between all maps. This multimodal influence drives a joint self-organization of maps and multimodal perceptions of stimuli. This work takes place after the design of a self-organizing map [25] and of a modulation mechanism for influencing its self-organization [26] oriented towards a multimodal purpose. In this paper, we introduce a way to connect these self-organizing maps to obtain a multimap multimodal processing, completing our previous work. We also give an overview of the architectural and functional properties of the resulting paradigm SOMMA.

I. INTRODUCTION

SOMMA introduces emergent functional properties of cortical computation - such as adaptability to various and changing data (see for example [15]) or unsupervised and online learning - in the computer science field. These properties are closely linked to the cortical architecture that is transposed to our computational paradigm SOMMA. In our connectionist architecture, cortical column is a basic and generic computational unit. Thus, our architecture has a mesoscopic level of description with keeping a cortical plausibility of the computation. Moreover, our model includes some of the architectural properties of the cortex such as layered cortical columns, cortical areas and topographic connections for an original computational data processing.

SOMMA applies such a cortically inspired computation and architecture to multimodal data processing, which is a key point for an artificial - or biological - agent to interact with its environment. Indeed, an agent measures the state of its surrounding environment thanks to numerous sensors providing data in various senses, each one having its own properties. It has to merge these multiple data to obtain a consistent perception of the environment. Psychological experiments tend to highlight that this multimodal merging is based on the detection of environmental regularities. For example, in the ventriloquist aftereffect, we locate a sound source where a temporally congruent visual flash should be, even if in the training session visual and sound signals were not spatially congruent [11, 32]. By the way, such a

data processing is consistent with sensori-motor theories. In SOMMA, multimodal merging is obtained by learning and detecting spatial regularities of a multimodal input flow.

In the next section, we describe the main architectural and computational paradigms of the cortex that are integrated in our work. In the third section, we give an overview of architecture and dynamic of SOMMA (detailed equations are given in annex I) with a focus on the introduced multimap multimodal processing. In the fourth section, we illustrate this multimodal learning in SOMMA with some artificial data.

II. CORTICAL INSPIRATION

As a reminder, SOMMA does not target a cortical modeling but aims to extract some architectural and computational paradigms of the cortex for a multimodal data processing in computer science. Thus, the following description is simplified from a biological point of view.

A. Cortical columns and areas

The cortex is subdivided into functional cortical areas as, for example, sensory areas. Even if sensory areas are specialized in one modality, they seem to have generic architecture and data processing [18, 30]. A cortical area is composed of maxi columns that are delimited by incoming connections from the thalamus [28]. Each maxi column is composed of cortical columns, which are vertical structures - i.e. perpendicular to the cortical surface - of highly interconnected neurons [27, 33]. A cortical column has six layers [22]. Each layer receives various connexions - usually topographic - that can be classified in two main flows [20]: a feedforward flow that comes from the thalamus and a cortical flow that comes from other cortical columns located either in the same or in distant cortical areas.

Each cortical column constituting a maxi column provides a specific data processing on the feedforward flow defining the maxi column. This processing is modulated by the lateral flow. For example, in the primary visual cortex, this processing consists on detecting one privileged orientation [19, 31]. At the cortical area level, these orientations are self-organized, meaning that two close columns discriminate two close orientations [8]. Such a functional self-organization can also be observed in the auditory cortex for example [35].

These biological evidences are integrated and adapted as computational paradigms in SOMMA:

- Cortical columns are generic and basic computational units of our connectionist architecture. Such a choice was already done in multiple models [1, 9, 17, 29].
- Cortical areas have a generic architecture and their functional specialization depends on their inputs.
- Each cortical column becomes sensible to a specific stimulus and these discriminations are self-organized at the map level as in Kohonen maps [23]. Self-organization provides interesting functional properties - such as dimensionality reduction and generalization to unknown stimuli - by topological projection of a high dimensional space on a low dimension manifold [4, 7].

B. Multimodal processing

In classical view of multimodal computation in the cortex, each modality is processed in some dedicated area and multimodal merging is obtained in associative high level areas. Such a hierarchical processing were used in some computational model of multimodal computation with self-organizing maps [21, 37].

However, more and more biological evidences tend to highlight a multimodal processing in all the cortex even in sensory areas [10]:

- There are direct connections between sensory cortices [12, 36].
- Neural activities in one sensory area may be influenced by stimuli from other modalities [2, 14].
- Sensory areas contain multimodal neurons, i.e. neurons activated by multiple modal stimuli [5, 6].

We introduce all these evidences for an original multimodal data processing in SOMMA. We propose to directly connect all maps so that computation in each map is influenced by the one in other maps. Thus, even though each cortical map processes a specific modality received by its input flow, multimodal processing is distributed in all cortical maps without dedicated associative maps. Such a distributed multimodal processing may provide some interesting functional properties:

- limited number of computational units and connections,
- increased robustness to lesion thanks to completely distributed processing,
- dimensionality reduction by separated processing,
- undifferentiated monomodal and multimodal processing.

III. SOMMA PARADIGMS

A. Overview

In SOMMA, each modal flow, part of the multimodal flow, is processed by one dedicated modal map. All modal maps have a bidimensional generic architecture composed of columns. Each column has three layers named sensory, cortical and perceptive. Each layer receives a specific data flow coming through weighted connections (see figure 1). At the map level, activities in each layer represent a spatial coding of the incoming data.

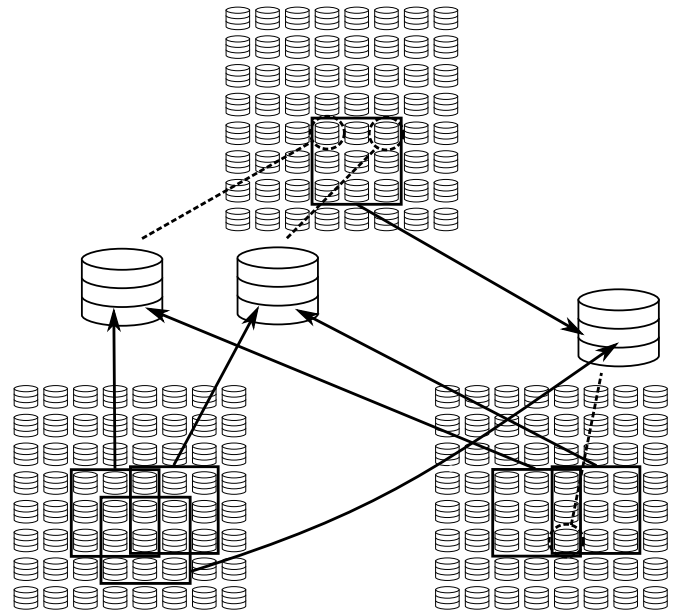


Fig. 2: Multimodal connections between maps are topographic and bidirectional.

The sensory layer uses the BCMu learning rule [26] to provide a tabular coding to a stimulus. This means that the sensory activity is function of a distance between the current stimulus and a discriminated one¹. This learned discrimination depends on the weights of the feedforward connections. After learning, these discriminations are self-organized at the map level leading to a spatial coding of the current stimulus.

All maps are reciprocally connected with topographic multimodal connections (see figure 2). More precisely, the cortical layer of a column located at (i, j) in one map receives incoming connections from the perceptive layer of columns located in a square centered around (i, j) in all others maps. Such spatially organized connections preserve the spatial coding of perceptive activities. Thus, at the map level, cortical activities represent the perception that should be obtained in the map to be consistent with perceptions made in other modalities.

The perceptive activity is computed according to the CNFT (Continuum Neural Field Theory) paradigm proposed in [3] which was already adapted to computational models (see for example [34]). This paradigm leads to the emergence of a stereotyped activity bump driven by the activation of sensory and cortical layers. Thus, localization of the activity bump is the result of a consensus between the local sensation and the other modal perceptions, providing a multimodal spatial coding of the current stimulus.

Sensory and perceptive layers as well as their coupling

¹More precisely sensory layer discriminates a spatial pattern that appears at several time steps in the input data flow. Thus, sensory activity depends on the detection of this spatial pattern in the current stimulus independently of other activations in the stimulus outside of this pattern. For more details on this coding and its advantages comparing to the one provided by prototype neurons of Kohonen maps see [24].

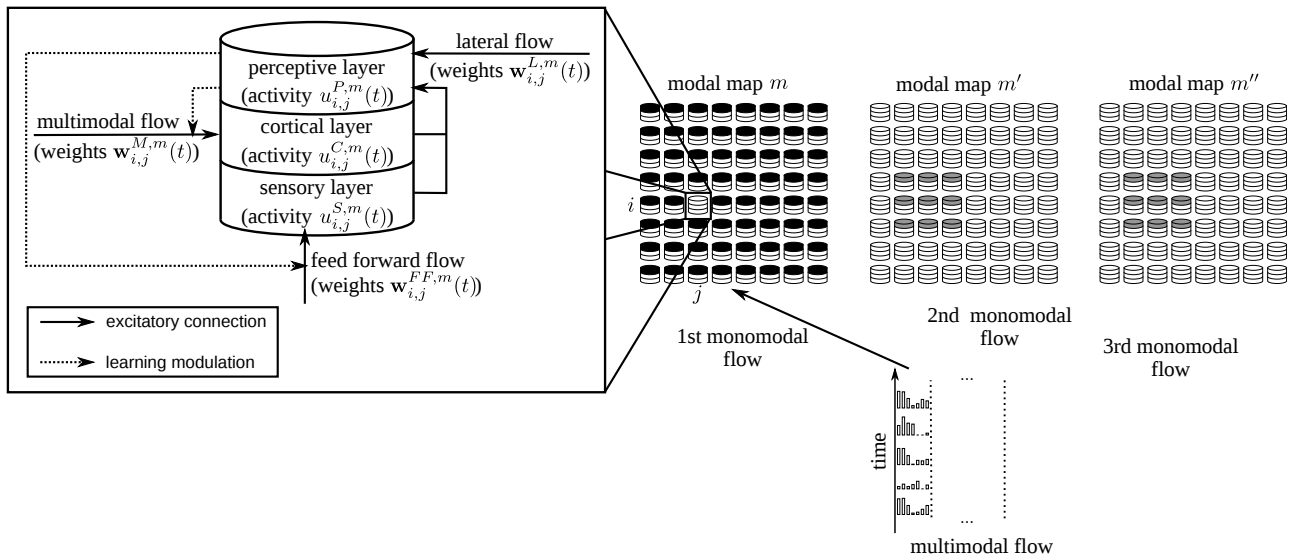


Fig. 1: SOMMA architecture. A multimodal flow is made up of multiple (here three) modal flows. Each one is processed by a modal map with a generic architecture. Each map is an assembly of generic 3-layers columns organized in a bidimensional grid. Each layer receives a specific data flow through weighted connections which are illustrated for the zoomed column. The sensory layer receives the modal stimulus. The cortical layer is connected to perceptive layer of columns in all other modal maps (in gray) through topographic connections. The perceptive layer receives perceptive activities of all other columns of the same map (in black) by means of the lateral flow.

were already presented in [25, 26]. In this article, we introduce the cortical layer and its incoming multimodal topographic connections. It provides an influence on the perception in each map by the one made in the other maps that leads to two linked properties:

- A joint self-organization i.e. a self-organization in each map so that the sensory layer of columns located at the same position in different maps discriminate monomodal stimuli that belong to the same multimodal stimulus.
- Multimodal perception of a stimulus is based on the joint self-organization i.e. that the perceptions of corresponding monomodal stimuli are localized at the same place in all maps.

These properties emerge in SOMMA by continuous computation and learning. This means that at each time step layers activities and plastic connections are updated (see annex I for detailed equations). In the next section, we give an overview of the emergent dynamic in SOMMA that leads to multimodal perception and joint self-organization.

IV. SOMMA PROCESSING

A. Computation

Each reception of a stimulus triggers activities changes in all maps that converge to a multimodal perception. Evolution of the dynamical system defined by SOMMA is explained in successive steps to be clear. However, it has to be kept in mind that in practice this dynamic is much more complex as

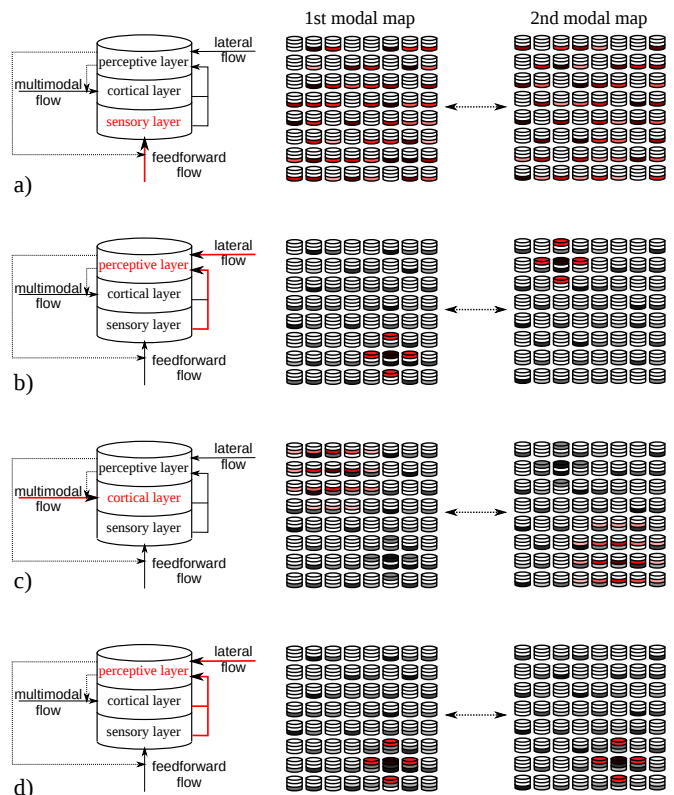


Fig. 3: Reception of a stimulus leads to activities changes in SOMMA. Dynamic of evolution - generated by continuous computation equations - results from interactions between incoming data flows of the column that stabilize on a multimodal perception. For more details please refer to the text.

it emerges progressively by local interactions between units whose activities continuously evolve.

Each sensory layer provides a tabular coding of the current stimulus. As initial feedforward weights are random, initial discriminations of sensory layers are neither sharp nor self-organized. Thus, sensory activities are spread all over the map (see figure 3 a). When a joint self-organization has emerged, sensory activities will be spatially localized in the map as close columns discriminate close stimuli.

In each column, the sensory layer excites the perceptive one. Lateral connections between perceptive layers of the same map have a *Mexican hat* shape that provide competition between columns with local excitation and global inhibition as defined in the CNFT [3]. Interactions between these sensory excitation and lateral competition lead to the emergence of a perceptive activity bump in each map where sensory activities are locally the highest (see figure 3 b).

A perceptive activity bump leads to a cortical activity bump located at the same position in all other maps because of the multimodal topographic connections (see figure 3 c). In each column, the cortical layer modulates the perceptive layer excitation provided by the sensory layer. Thus, a perceptive activity bump tends to appear where sensory and cortical activities are both high. This cortico-perceptive excitatory loop between all maps creates a resonance mechanism similar to the one used in [13, 16, 29]. Thus, all perceptions tend to appear at the same position defining a multimodal perception of the current stimulus (see figure 3 d).

B. Learning

In SOMMA, learning consists on modifying weights of the feedforward connections. This is achieved by using the BCMu learning rule, a modified version of the BCM learning rule that we proposed in [26]. This learning rule has two coupled dynamics that provide its functional properties:

- a temporal competition between stimuli - included in the BCM learning rule - that raises a discrimination to a stimulus of the input flow,
- a modulation and a regulation mechanisms - introduced in BCMu - that influence this discrimination so that a stimulus that is simultaneously present with a modulation signal will be discriminated. The higher is the modulation signal, the higher is the probability, for the corresponding stimulus, to be discriminated.

In SOMMA, perceptive activity is used as a modulation signal for the BCMu learning rule. Because of a lower time constant for learning than for computation, weights are mainly modified when multimodal perception has emerged. Resulting learning of multimodal correlations may be described in three stages:

- Each sensory layer learns autonomously a discrimination which is a fixed point of the BCMu learning rule [26]. By the way, this autonomous emergence improves the coupling with the perceptive layer using CNFT (please refer to [25] for technical arguments).

- Because of the bump shape of perceptive activities, close columns receive close modulation signals. As a result, sensory layer of close columns will have close discriminations. Thus, perceptive activities act as a decentralized neighborhood function, ensuring map self-organization [25].
- Because of the resonance mechanism, perceptive activity bumps are forced to be located at the same position in all maps. The modulation and regulation mechanisms of BCMu constraint the current stimulus to be discriminated at this specific position (see [26] for more details). Consequently, monomodal stimuli of a multimodal stimulus are learned at the same position in all maps leading to a joint self-organization.

V. RESULTS

A. Three modalities, linear learning

In order to illustrate multimodal processing in SOMMA, we use a simple artificial flow with three modalities. Each stimulus of a modal flow has a 10×10 dimension and is processed in a 20×20 map.

To create a multimodal stimulus, we place a Gaussian with fixed variance and amplitude at a random position - picked with an uniform distribution - in a bidimensional space (see figure 4). For this protocol, the space is toric so that to avoid side effects. This space is discretized to obtain a 10×10 matrix of activities. Each of the three monomodal stimuli is equal to this matrix of activities plus an uniform random noise. Such stimuli are presented successively to the model for some time steps, so that activities can converge and stabilize enough time for learning².

Sensory layer of each column learns a discrimination to an input stimulus i.e. a Gaussian at a specific position in the input space. Each discrimination is represented by a colored oriented bar, the orientation (respectively the color) represents the Gaussian position in the first (respectively second) input space dimension. Discriminations are normalized so that to obtain a toric representation with orientations between 0 and π and colors from red to green and back to red (see scales at the middle and the bottom of figures 5 a, b and c).

Self-organization of each map can be observed as there is a spatial continuity of orientation and color all over the map (see figures 5 a, b and c). Multimodal influence provided by topographic connections leads to a joint self-organization of the model as self-organizations in all maps are similar. To quantify this joint self-organization, we compute the average difference between discriminations of columns located at the same position in each map. This difference should be zero for a perfect joint self-organization. In this example, we can clearly see that this difference decreases because of reciprocal influence of the maps and stabilizes close to zero meaning that multimodal learning converged (figure 5 d).

²Parameters used for this experiment are gathered in a configuration file available at www.loria.fr/~lefortma/ijcnn/trimodal.cfg. It can be used with the dedicated SOMMA simulation framework available at www.loria.fr/~lefortma/software/SOMMA.tar.gz

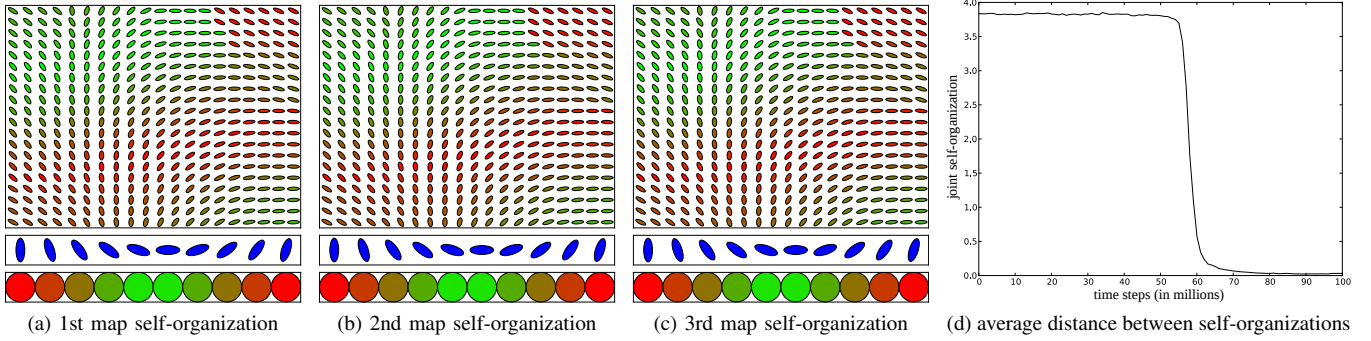


Fig. 5: Joint self-organization of three maps receiving similar noisy inputs. a-b-c) Top: each colored oriented bar represents the discrimination of the sensory layer. Middle and bottom: respectively scale of orientation and color. d) Average distance between discrimination of columns situated at the same position in each map.

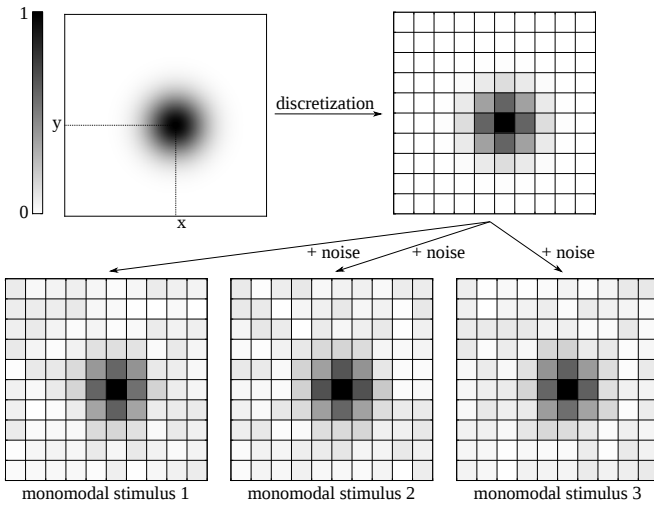


Fig. 4: Artificial multimodal stimuli are composed of noisy discretized Gaussian, one in each modality, located at the same randomly chosen position picked with a uniform distribution.

B. Simple robotic arm, non linear learning

In order to illustrate the ability of SOMMA to learn a non linear relationship between modal stimuli, we use two 20×20 maps receiving simple proprioceptive and visual stimuli from an artificial arm (see figure 6). As in the previous experiment, modal stimuli are a noisy Gaussian in a 10×10 matrix of activities. Gaussian position in a proprioceptive stimulus is picked randomly with a uniform distribution in $[0, \pi] \times [0, \pi]$ and represents arm joints position resulting from a motor babbling. Gaussian position in the corresponding visual stimulus is the hand spatial position resulting from this arm joints configuration. Gaussian positions are normalized so that to be spread in the 10×10 input space³.

Results are presented in figure 7 with the same kind of representation as in section V-A except for two points. Firstly,

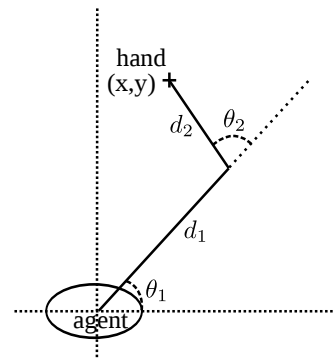


Fig. 6: The arm moves in a bidimensional space. The visual position of the hand (x, y) is determined by the joint positions (θ_1, θ_2) and by fixed arm and forearm lengths respectively denoted d_1 and d_2 . In our experiment, $d_1 = 1$ and $d_2 = 0.5$.

as input spaces are not toric, discriminations are normalized to have orientation in $[0, \pi/2]$ and color from red to green. Secondly, to be able to visualize the joint self-organization, discriminations of both maps are represented in the visual space. This means that for a column in the proprioceptive map, we do not plot the discriminated joints position but the corresponding visual hand position corresponding to this joints configuration. This space transformation is also used to quantify the joint self-organization.

As in the trimodal protocol we can observe the self-organization of each map and the joint self-organization of the model. However, there is more discontinuities in map self-organizations (close columns discriminate distant stimuli) and more differences between self-organizations than in the previous protocol. This is mainly due to the increased complexity of the task that involve non linear relationship between modalities and a non uniform distribution of the visual inputs. Moreover, our aim was not to optimize the joint self-organization quantification (parameters used for both experiments are the same). We want to illustrate the ability of SOMMA to deal with various kind of inputs and

³Parameters used for this experiment are gathered in a usable configuration file available at www.loria.fr/~lefortma/ijcnn/arm.cfg

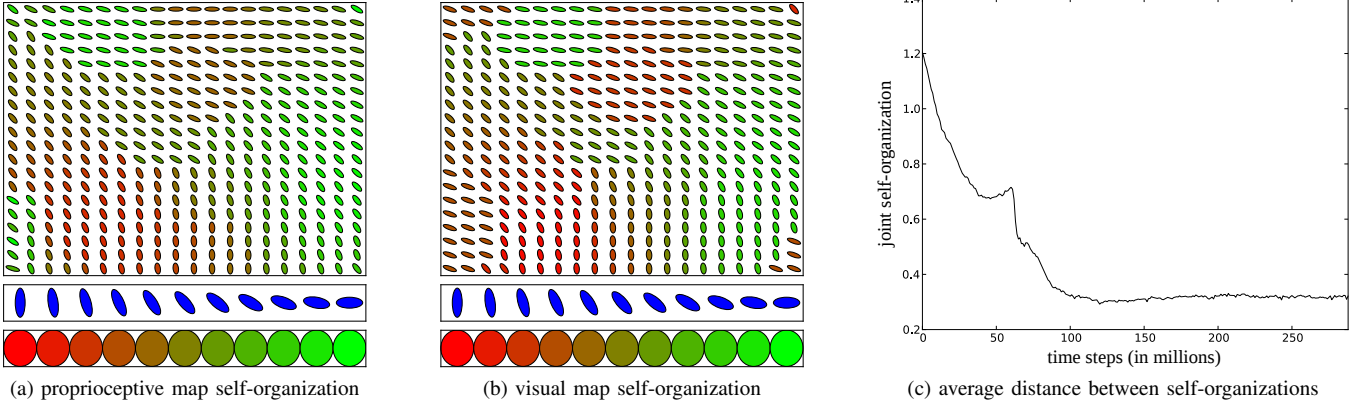


Fig. 7: Joint self-organization of two maps receiving proprioceptive and visual stimuli from an artificial arm. a-b) Top: each colored oriented bar represents the corresponding discrimination of the sensory layer in the visual field. Middle and bottom: respectively scale of orientation and color. c) Average distance between discrimination of columns situated at the same position in each map.

the multimodal influence of each self-organization that can be clearly see with the decrease of the average distance between self-organizations.

VI. CONCLUSION

SOMMA defines cortically inspired cortical maps for multimodal computation. In previous articles [25, 26], we proposed cortical maps for monomodal processing using the coupling of BCMu learning rule with neural fields. In this article, we introduce multimodal processing which is based on the addition of a cortical layer that receives topographic connections between maps to modulate computation in each map. This multimodal interconnection creates a resonance mechanism between maps that leads to two functional properties. Firstly, learned self-organization in each map is influenced by the other maps so that monomodal stimuli composing a multimodal stimulus are learned at the same position in each map, defining a joint self-organization. Secondly, multimodal perception is based on this joint self-organization: each monomodal stimulus is spatially coded by an activity bump in each map and a multimodal stimulus is represented by activity bumps located at the same position in each map.

Multimodal learning of SOMMA was validated with a simple example containing three modalities with a linear relationship. It is also efficient with non linear relationship as we illustrate it with a simple robotic arm. Moreover, SOMMA processing is generic as these two experiments used the same parametrization.

Future work will focus on validating our paradigms with more complex relationships between modalities (redundant, partial, ...) and with real input flows. Preliminary results seem promising. Moreover, as computation in SOMMA is generic, it should be interesting to integrate our cortical maps in a hierarchical architecture to obtain a multilevel multimodal computation.

APPENDIX I EQUATIONS

Computation in SOMMA is generic, meaning that equations of activities evolution and weights learning are the same for each column of each map. In this section, we detail these equations for a column located at position (i, j) in a modal map m . Please refer to figure 1 for detailed denomination of column activities and connections weights.

A. Sensory activity

Sensory layer, whose activity is denoted $u_{i,j}^{S,m}(t)$, receives the current modal stimulus $\mathbf{x}^m(t) = (\mathbf{x}_1^m(t), \dots, \mathbf{x}_n^m(t))$ through feedforward connections with weights $\mathbf{w}_{i,j}^{FF,m}(t) = (w_{i,j,1}^{FF,m}(t), \dots, w_{i,j,n}^{FF,m}(t))$. Sensory activity is computed as the weighted sum of the current modal stimulus:

$$u_{i,j}^{S,m}(t) = \sum_{k=1}^n w_{i,j,k}^{FF,m}(t) x_k^m(t)$$

B. Cortical activity

Cortical activity $u_{i,j}^{C,m}(t)$ is computed as the weighted sum of the perceptive activities that are received through topographic multimodal connections (see figure 2):

$$u_{i,j}^{C,m}(t) = \sum_{m' \neq m, (i', j')} w_{i,j,i',j'}^{M,m,m'}(t) u_{i',j'}^{P,m'}(t)$$

$w_{i,j,i',j'}^{M,m,m'}(t)$ stands for the weight at time t of the multimodal connection coming from the cortical column situated at (i', j') in map m' to the cortical column situated at (i, j) in map m and $u_{i',j'}^{P,m'}(t)$ is the perceptive activity of column situated at (i', j') in map m' .

C. Perceptive activity

Perceptive activity $u_{i,j}^{P,m}(t)$ is computed by using neural field paradigm:

$$u_{i,j}^{P,m}(t) = f(u_{i,j}^{P,m}(t-dt) + \underbrace{\gamma[-u_{i,j}^{P,m}(t-dt) + \sum_{(i',j')} w_{i,j,i',j'}^{L,m} u_{i',j'}^{P,m}(t)}_{\text{lateral term}} + \underbrace{\alpha u_{i,j}^{S,m}(t)(u_{i,j}^{C,m}(t) + \delta)}_{\text{afferent term}}] + \underbrace{\gamma[-u_{i,j}^{P,m}(t-dt) + \sum_{(i',j')} w_{i,j,i',j'}^{L,m} u_{i',j'}^{P,m}(t)}_{\text{lateral term}} + \underbrace{\alpha u_{i,j}^{S,m}(t)(u_{i,j}^{C,m}(t) + \delta)}_{\text{afferent term}}])}_{\text{decay term}}$$

with f a sigmoid function to limit the perceptive activity, γ the time constant, α the afferent force, δ the cortical preactivation and $w_{i,j,i',j'}^{L,m}$ the lateral connection from column (i', j') to (i, j) in map m . Lateral connections have a fixed *Mexican hat* shape for all columns and all maps: $w_{i,j,i',j'}^{L,m} = Ae^{-\frac{\|(i,j),(i',j')\|^2}{a^2}} - Be^{-\frac{\|(i,j),(i',j')\|^2}{b^2}}$ with $\|\cdot, \cdot\|$ the euclidean distance and A, B, a, b constants so that $A > B > 0$ and $b > a > 0$.

The afferent term excites the neural field, whose activity is shaped to a bump thanks to the lateral term. The decay term ensures disappearing of the activity bump when no more afferent excitation is received.

D. Learning

Feedforward weights $\mathbf{w}_{i,j}^{FF,m}(t) = (w_{i,j,1}^{FF,m}(t), \dots, w_{i,j,n}^{FF,m}(t))$ are modified by the BCMu learning rule using the perceptive activity as a modulation signal. Equations are the following:

$$\begin{aligned} u_{i,j}^{S',m}(t) &= u_{i,j}^{S,m}(t) + \underbrace{\lambda u_{i,j}^{P,m}(t)}_{\text{modulation term}} \\ \theta_{i,j}^m(t) &= \frac{1}{\tau} u_{i,j}^{S',m}(t)^2 + (1 - \frac{1}{\tau}) \theta_{i,j}^m(t-dt) \\ \beta_{i,j}^m(t) &= \beta_0 (1 + \underbrace{\theta_{i,j}^m(t-dt)^2}_{\text{competitive term}}) \\ \forall k, \Delta w_{i,j,k}^{FF,m}(t) &= \underbrace{\eta^{FF} x_k^m(t) [u_{i,j}^{S',m}(t)(u_{i,j}^{S',m}(t) - \theta_{i,j}^m(t)) - \beta_{i,j}^m(t) N(u_{i,j}^{P,m}(t)) u_{i,j}^{S',m}(t)^2 w_{i,j,k}^{FF,m}(t-dt)]}_{\text{regulation term}} \end{aligned}$$

with $u_{i,j}^{S',m}(t)$ the sensory modulated activity, λ the strength of the modulation, τ the temporal window of the exponential filter used to compute the sliding threshold $\theta_{i,j}^m(t)$, $\beta_{i,j}^m(t)$ the adaptive strength of the regulation term, η^{FF} the constant learning rate, $\mathbf{x}^m(t) = (x_1^m(t), \dots, x_n^m(t))$ the current modal stimulus and N a function that detects the lack of modulation, i.e. $N(u_{i,j}^{P,m}(t)) = 0$ (resp. $N(u_{i,j}^{P,m}(t)) = 1$) if $u_{i,j}^{P,m}(t) > 0$ (resp. $u_{i,j}^{P,m}(t) = 0$).

The competitive term, by means of the sliding threshold, generates a temporal competition between stimuli that raises a discrimination to a stimulus. The modulation terms favors modulated stimuli in the temporal competition between stimuli. The regulation mechanism ensures the modulation to be efficient, i.e. that the discriminated stimulus is a modulated one. For more details, please refer to [26].

REFERENCES

- [1] F. Alexandre, F. Guyot, J.P. Haton, and Y. Burnod. The cortical column: a new processing unit for multilayered networks. *Neural networks*, 4(1):15–25, 1991.
- [2] B.L. Allman, L.P. Keniston, and M.A. Meredith. Not just for bimodal neurons anymore: the contribution of unimodal neurons to cortical multisensory processing. *Brain topography*, 21(3):157–167, 2009.
- [3] S. Amari. Dynamics of pattern formation in lateral-inhibition type neural fields. *BiolCyb*, 27:77–87, 1977.
- [4] D.H. Ballard. Cortical connections and parallel processing: Structure and function. *Behavioral and Brain Sciences*, 9(01):67–90, 1986.
- [5] D.S. Barth, N. Goldberg, B. Brett, and S. Di. The spatiotemporal organization of auditory, visual, and auditory-visual evoked potentials in rat cortex. *Brain Research*, 678(1-2):177–190, 1995.
- [6] J.K. Bizley and A.J. King. Visual-auditory spatial processing in auditory cortical neurons. *Brain Research*, 1242:24–36, 2008.
- [7] F. Blayo and P. Demartines. Data analysis: How to compare kohonen neural networks to other techniques? *Artificial neural networks*, pages 469–476, 1991.
- [8] W.H. Bosking, Y. Zhang, B. Schofield, and D. Fitzpatrick. Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, 17(6):2112, 1997.
- [9] Y. Burnod. *An adaptive neural network: the cerebral cortex*. Masson editeur, 1990.
- [10] G. Calvert, C. Spence, and B.E. Stein. *The handbook of multisensory processes*. The MIT Press, 2004.
- [11] L.K. Canon. Intermodality inconsistency of input and directed attention as determinants of the nature of adaptation. *Journal of Experimental Psychology*, 84(1):141, 1970.
- [12] C. Cappe, E.M. Rouiller, and P. Barone. Multisensory anatomical pathways. *Hearing research*, 258(1-2):28–36, 2009.
- [13] G.A. Carpenter, S. Grossberg, N. Markuzon, J.H. Reynolds, and D.B. Rosen. Fuzzy artmap: A neural network architecture for incremental supervised learning of analog multidimensional maps. *Neural Networks, IEEE Transactions on*, 3(5):698–713, 1992.
- [14] L.R. Dehner, L.P. Keniston, H.R. Clemo, and M.A. Meredith. Cross-modal circuitry between auditory and somatosensory areas of the cat anterior ectosylvian sulcal cortex: a new inhibitory form of multisensory convergence. *Cerebral Cortex*, 14(4):387–403, 2004.
- [15] T. Elbert and B. Rockstroh. Reorganization of human cerebral cortex: the range of changes following use and injury. *The Neuroscientist*, 10(2):129–141, 2004.
- [16] S. Grossberg. Adaptive pattern classification and universal recoding: I. parallel development and coding of neural feature detectors. *Biological cybernetics*, 23(3):121–134, 1976.
- [17] E. Guigon, B. Dorizzi, Y. Burnod, and W. Schultz.

- Neural correlates of learning in the prefrontal cortex of the monkey: A predictive model. *Cerebral Cortex*, 5(2):135–147, 1995.
- [18] K. Holthoff, E. Sagnak, and O.W. Witte. Functional mapping of cortical areas with optical imaging. *NeuroImage*, 37(2):440–448, 2007.
- [19] D.H. Hubel and T.N. Wiesel. Shape and arrangement of columns in cat’s striate cortex. *The Journal of physiology*, 165(3):559–568, 1963.
- [20] C. Johansson and A. Lansner. Towards cortex sized artificial neural systems. *Neural Networks*, 20(1):48–61, 2007.
- [21] M. Johnsson, C. Balkenius, and G. Hesslow. Associative self-organizing map. In *International Joint Conference on Computational Intelligence (IJCCI)*, pages 363–370. Citeseer, 2009.
- [22] E.R. Kandel, J.H. Schwartz, T.M. Jessell, S.A. Siegelbaum, and AJ Hudspeth. *Principles of neural science*, volume 3. Elsevier New York, 1991.
- [23] T. Kohonen. Self-organized formation of topologically correct feature maps. *Biological cybernetics*, 43(1):59–69, 1982.
- [24] M. Lefort. *Apprentissage spatial de corrélations multimodales par des mécanismes d’inspiration corticale*. PhD thesis, Université Nancy II, 2012.
- [25] M. Lefort, Y. Boniface, and B. Girau. Coupling bcm and neural fields for the emergence of self-organization consensus. *From Brains to Systems*, pages 41–56, 2011.
- [26] M. Lefort, Y. Boniface, and B. Girau. Unlearning in the bcm learning rule for plastic self-organization in a multi-modal architecture. *Artificial Neural Networks and Machine Learning–ICANN 2011*, pages 93–100, 2011.
- [27] R. Lorente de Nó. Cerebral cortex: architecture, intracortical connections, motor projections. *Physiology of the nervous system*, pages 288–330, 1949.
- [28] J.S. Lund, A. Angelucci, and P.C. Bressloff. Anatomical substrates for functional columns in macaque monkey primary visual cortex. *Cerebral Cortex*, 13(1):15–24, 2003.
- [29] O. Ménard and H. Frezza-Buet. Model of multi-modal cortical processing: Coherent learning in self-organizing modules. *Neural Networks*, 18(5):646–655, 2005.
- [30] K.D. Miller, D.J. Pinto, D.J. Simons, and J. Daniel. Processing in layer 4 of the neocortical circuit: new insights from visual and somatosensory cortex. *Current opinion in neurobiology*, 11(4):488–497, 2001.
- [31] V.B. Mountcastle. Modality and topographic properties of single neurons of cats somatic sensory cortex. *J. Neurophysiol*, 20(4):408–434, 1957.
- [32] M. Radeau and P. Bertelson. The after-effects of ventriloquism. *The Quarterly journal of experimental psychology*, 26(1):63, 1974.
- [33] P. Rakic. Guidance of neurons migrating to the fetal monkey neocortex. *Brain research*, 33(2):471, 1971.
- [34] N.P. Rougier and J. Vitay. Emergence of attention within a neural population. *Neural Networks*, 19(5):573–581, 2006.
- [35] C.E. Schreiner. Order and disorder in auditory cortical maps. *Current Opinion in Neurobiology*, 5(4):489–496, 1995.
- [36] C.E. Schroeder and J. Foxe. Multisensory contributions to low-level, unisensory processing. *Current Opinion in Neurobiology*, 15(4):454–458, 2005.
- [37] M. Vavrečka and I. Farkaš. Unsupervised model for grounding multimodal representations. 2010.