A Simple Mechanism to Reproduce the Neural Solution of the Aperture Problem in Monkey Area MT

Maria-Jose Escobar, Guillaume S. Masson, Pierre Kornprobst

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

A Simple Mechanism to Reproduce the Neural Solution of the Aperture Problem in Monkey Area MT

Maria-Jose Escobar — Guillaume S. Masson — Pierre Kornprobst

N° ????

June 2008

Thème BIO

ISSN 0249-6399 ISRN INRIA/RR--????--FR+ENG

A Simple Mechanism to Reproduce the Neural Solution of the Aperture Problem in Monkey Area MT

Maria-Jose Escobar* , Guillaume S. Masson† , Pierre Kornprobst‡

Thème BIO — Systèmes biologiques
Projet Odyssée

Rapport de recherche n° ???? — June 2008 — 12 pages

Abstract: We propose a simple mechanism to reproduce the neural solution of the aperture problem in monkey area MT. More precisely, our goal is to propose a model able to reproduce the dynamical change of the preferred direction (PD) of a MT cell depending on the motion information contained in the input stimulus. The PD of a MT cell measured through drifting gratings differs of the one measured using a barberpole, which is highly related with its aspect ratio. For a barberpole, the PD evolves from the perpendicular direction of the drifting grating to a PD shifted according to the aspect ratio of the barberpole. The mechanisms underlying this dynamic are unknown (lateral connections, surround suppression, feed-backs from higher layers). Here, we show that a simple mechanism such as surround-inhibition in V1 neurons can produce a significant shift in the PD of MT neurons as observed with barberpoles of different aspect ratios.

Key-words: Aperture problem, surround inhibition, MT, barberpole

* Maria-Jose.Escobar@sophia.inria.fr
† Guillaume.Masson@incm.cnrs-mrs.fr
‡ Pierre.Kornprobst@sophia.inria.fr

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

A Simple Mechanism to Reproduce the Neural Solution of the Aperture Problem in Monkey Area MT

Résumé : Pas de résumé

Mots-clés : Aperture problem, surround inhibition, MT, barberpole

Contents

1 Introduction

The preferred direction (PD) of a MT cell has been generally measured through a drifting grating, where most of the times the cell shows a clear direction selectivity. Studies, as the ones done by Pack et al. [13, 14, 6] show that the PD can be modified depending on the input stimulus. Specifically, [14] showed that the PD measured using barberpoles instead of grating is biased toward perception, i.e., the side of the barberpole with the longest side. This PD deviation, compared to the one measured drifting grating, depends on the aspect ratio of the barber pole.

Evidence of microelectrode recordings in MT of alert monkey reveal that during the first 80ms after the onset stimulus the response is strongly biased by 1D motion, i.e., the direction defined by the orthogonal direction to the contours, but lately the 2D motion direction is encoded. These experiments suggest that the aperture problem is solved within the first 100ms of the onset stimulus [13].

The mechanisms underlying the PD deviation of a MT cell are not at all defined. It looks like that the primate visual system initially considers all the information available (ambiguous and unambiguous), and that along time, it refines it in order to solve the aperture problem. This convergence in time can be associated to different and complex neural networks which convey information coming from other areas of the visual system as feedbacks [5] or horizontal connections.This phenomenon is also associated to the contribution of terminators or end-points in different areas of the visual field such as V2 or V1 [5, 4, 12] which should require slightly longer latencies.

We propose that a simple mechanism such as surround-inhibition in V1 neurons can produce a significant shift in the PD of MT neurons as observed with barberpoles of different aspect ratios. The surround-inhibition, modeled by an isotropic difference of gaussians (DoG), acts like an end-stopping cell [16, 10] enhancing the responses of the V1 cells located at the border of the barber poles, and inhibiting the activation of the V1 cells located at the center. This article is organized as follows. In Section 2, we describe our model and the experimental protocol. In Section 3, we present the results obtained. Finally, conclusion is given in Section 4.

2 Methods

2.1 The model

The motion processing architecture is mainly divided into two stages: V1 and MT. V1 neurons integrate in time the motion information contained in the input stimulus together with the interactions between V1 cells. After a nonlinear stage, the V1 output feed the MT neurons. The MT neurons integrate the V1 information along time and space so that the output is the evolution in time of the membrane potential of a MT neuron. After a second nonlinear stage, this membrane potential is converted to an estimation of the mean firing rate of the neuron. The values of the mean firing rate are the ones used to validate our model with physiological data shown in, e.g., [13, 14, 6].

2.1.1 Dynamic V1 layer

Each V1 cell is defined by a dynamic neuron entity, integrating in time the local motion information contained in the input stimulus together with the responses of neighboring V1 cells. The motion information processing is done through energy filters [1] which feeds the V1 neuron as an external input current. So, a V1 neuron will be ruled by the following equation

$$\frac{du_i^{V1}(t)}{dt} = \frac{R}{\tau}\left[I_i^{inh}(t) + g_L\left(E^{rest} - u_i^{V1}(t)\right) + I_i^{ext}(t)\right],$$

where $u_i^{V1}(t)$ represents the membrane potential of the V1 neuron i, $I_i^{inh}(t)$ the inhibitory current given by the response of neighboring V1 cells, $I_i^{ext}(t)$ is the excitatory input current given by the output of the energy filter bank over the input stimulus. E^{rest} is resting potential of the V1 neuron set as $0mV$. Parameters related with the dynamic of the neuron are g_L (leak coefficient), R and $\tau = RC$.

External input current $I_i^{ext}(t)$: The external input current $I_i^{ext}(t)$ is an excitatory input corresponding to the response of cells tuned in a given direction and for a range of speeds. This response is classically modeled by a bank of energy filters modeling a V1 complex cell (see [7] for a V1 cell classification).

In this paper, the local motion information is calculated from the luminosity information of the input video sequence through Reichardt detectors (see, e.g., [3] for implementation details). Given a velocity space $\mathcal{V}$ of possible velocities, the Reichardt detectors returns for each pixel of the input sequence, a velocity map with the 'probability' associated for each velocity, denoted by $p(t, x, v)$ for all $v \in \mathcal{V}$. Examples of velocity maps are shown in Figure 1.

Then, one need to extract the response for a given motion direction from the distribution $p(t, x, v)_{v\in\mathcal{V}}$. To do this, the velocity map was divided into 8 angular sectors $S_k = [\frac{k\pi}{4} - \frac{\pi}{8}, \frac{k\pi}{4} + \frac{\pi}{8}]$, $(k = 0, .., 7)$, each of them representing the motion orientation $\theta_k = \frac{k\pi}{4}$. Within each sector k, we used the maximal probability to set the current of the ith V1 cell, i.e.,

$$I_i^{ext}(t) = \max\{p(t, x, v)|v \in S_k\}.$$

Inhibitory input current $I_i^{inh}(t)$: The value of the inhibitory input current $I_i^{inh}(t)$ is given by the interactions between V1 cells located inside a neighborhood. So, $I_i^{inh}(t)$ can be written as

$$I_i^{inh}(t) = \frac{1}{R}\sum_{j\in\Omega} w_{ij}^{inh}(|\mathbf{x}_i - \mathbf{x}_j|)u_j^{V1}(t - \delta), \tag{1}$$

where Ω is the neighborhood defined as 2.2 times the size of the V1 receptive field [16]. Equation (1) pools the membrane potentials $u_j^{V1}(\cdot)$ of all neurons j in the spatial neighborhood of neuron i, independently of the direction tuning of the cells and with a time delay δ. Weights w_{ij}^{inh} represent the locality: w_{ij}^{inh} is defined by a gaussian function of the distance between the neurons positions i and j.

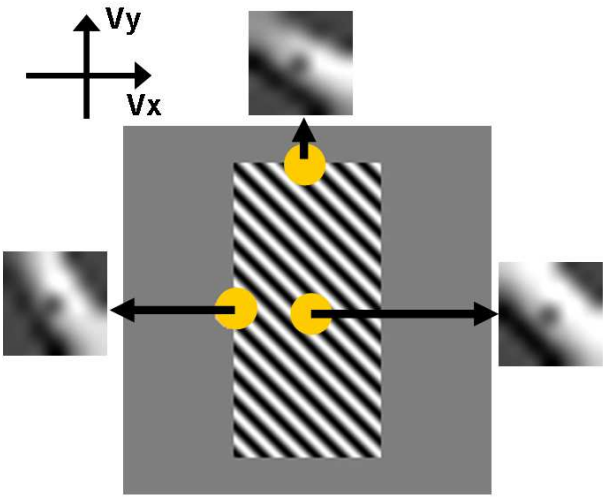

Figure 1: Velocity maps obtained using Reichardt motion detectors for three different locations in the Barberpole: center, vertical border and horizontal border.

Nonlinearity at the end of V1 stage: It is well known in biology that the V1 output shows several nonlinearities due to: response saturation, response rectification, or contrast gain control (see [2]). In order to obtain a nonlinear saturation in the V1 response, the V1 output is passed through a sigmoid function defined by $S(x) = [1 + \exp(-(x - ab)/b)]^{-1}$, where the parameters a and b were tuned to have a suitable response in the case of drifting gratings as inputs. So, the V1 output $r_i^{V1}(t)$, representing an estimation of the mean firing rate for a membrane potential $u_i^{V1}(t)$, is given by

$$r_i^{V1}(t) = S\left(u_i^{V1}(t)\right). \tag{2}$$

2.1.2 Higher order motion analysis by MT

The dynamics of the MT neurons are modeled by a simplified conductance based neuron without input currents [8]. Considering a neuron i, defined by its membrane potential $u_i^{MT}(t)$, the conductance-driven equation is given by

$$\tau \frac{du_i^{MT}(t)}{dt} = G_i^{exc}(t)\left(E^{exc} - u_i^{MT}(t)\right) + g^L\left(E^{rest} - u_i^{MT}(t)\right),$$

where E^{exc} and E^{rest} are constant which typical values can be 70mV and 0mV, respectively. According to (3), $u_i^{MT}(t)$ will belong to the interval $[E^{rest}, E^{exc}]$ and it will be mainly driven by input pre-synaptic neurons which will push the membrane potential $u_i^{MT}(t)$ towards E^{exc}, with a strength defined by $G_i^{exc}(t)$. The last term will drive $u_i^{MT}(t)$ towards the resting potential E^L with a con-

stant strength given by g^L. In this paper, we do not consider inhibitory effects in the neurons, as e.g., surround interactions.

The MT neuron i is part of a neural network where the excitatory input conductances $G_i^{exc}(t)$ is obtained pooling the activity of all the pre-synaptic neurons connected to it. Each MT cell has a receptive field built from the convergence of afferent V1 complex cells, with a size of about 4-6 times bigger than the V1 receptive field [11], which is normally called the classical-receptive-field (CRF). So, $G_i^{exc}(t)$ can be defined as

$$G_i^{exc}(t) = \max\left(0, \sum_{j\in\Phi_i} w_{ij}^{exc} r_j^{V1} - \sum_{j\in\Phi_i'} w_{ij}^{exc} r_j^{V1}\right),$$

where $\Phi_i = \{j \in \text{CRF} \mid \varphi_{ij} < \pi/2\}$, $\Phi_i' = \{j \in \text{CRF} \mid \varphi_{ij} > \pi/2\}$, and where the connection weight w_{ij}^{exc} is the efficacy of the synapse from neuron j to neuron i, which is proportional to the angle φ_{ij} between the two preferred motion direction-selectivity of the V1 and MT cell. It is important to remark that the value of the conductance will always be greater or equal to zero.

The connections weights w_{ij}^{exc} will be given by

$$w_{ij}^{exc} = k_c w_{cs}(|\mathbf{x}_i - \mathbf{x}_j|)\cos(\varphi_{ij}), \quad 0 \le \varphi_{ij} \le \pi, \tag{3}$$

where k_c is an amplification factor, φ_{ij} is the absolute angle between the preferred ith MT cell direction and the preferred jth V1 cell direction. Weight w_{cs} is a gaussian function depending on the distance between the MT cell i and the V1 cell j (their mapping on the same image space).

Similarly to the V1 stage, the membrane potential of the MT neurons $u_i^{MT}(t)$ is passed through a nonlinearity in order to obtain an estimation of the mean firing rate $r_i^{MT}(t)$ (see equation (2)).

2.2 Experimental protocol

Our experiments were carried out over two types of stimuli: drifting grating and barberpoles. We used a drifting grating of fixed spatial and temporal frequencies (0.1 cycles/pixel and 1 cycles/sec, respectively), of size 200×200 pixels, and a drifting orientation of $\theta = 45°$. Starting from this drifting grating we created two barberpole stimuli, with 4:1 and 2:1 aspect ratios, respectively (some snapshots of these stimuli can be seen in Figure 2). We placed over these stimulus an array of V1 cells, with a radius of 90 pixels and a constant density of 0.2 cells/pixel. Instead of studying the behavior of one cell and rotate the input stimulus for each of the 8 orientations, we equivalently study the behavior of 8 cells tuned for the 8 orientations for the same stimulus.

We ran the system during 1.2sec recording V1 and MT cells' outputs. For the three stimuli we put special emphasis in V1 cells located at the center of the stimulus and at the center of the longest border (for barberpoles only). The MT cell that we consider was placed for the three stimuli at the center of the image. The positions of V1 cells of interest are shown in Figure 2.

The results were divided into two stages: *early* and *late* responses. The *early* measurement is obtained without the surround inhibition effect in V1 cells ($I_i^{inh}(t) = 0$), while for the *late* response the $I_i^{inh}(t)$ is branched and the surround effect is delayed in $\delta = 30ms$.

(a) (b) (c)

Figure 2: Snapshots of the three stimuli. (a) For the drifting grating sequence, the V1 cell of interest is marked by an orange circle at the center. (b)-(c) barberpole sequences with aspect ratios 4:1 and 2:1. Together with the V1 cell of interest located at the center of the sequence, there is a second one placed at the center of the longest border. In the three cases the MT cell of interest is placed at the center of the images

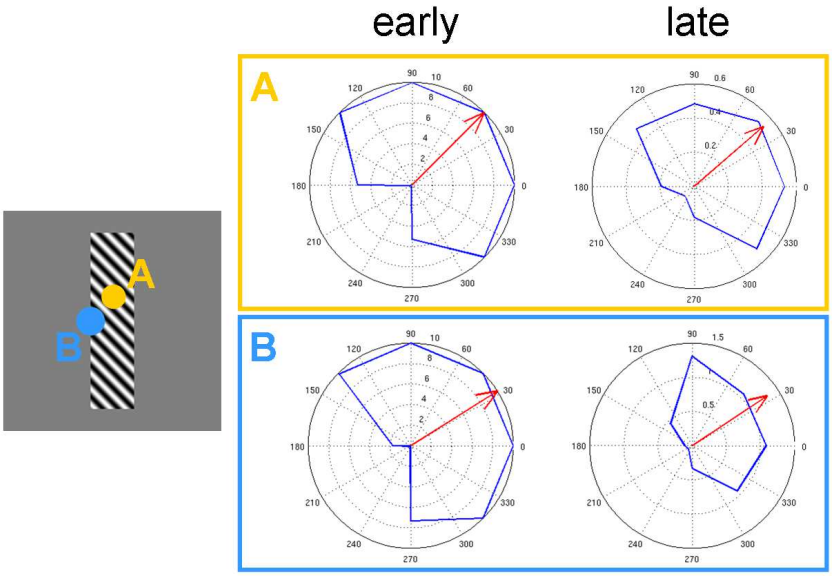

Figure 3: Polar diagrams obtained for two V1 cells placed at the center of the barberpole and at the center of its longer side. For each V1 cell the graphs for the early and late response are also shown. The red arrow denotes the vector average corresponding to the PD of the cells.

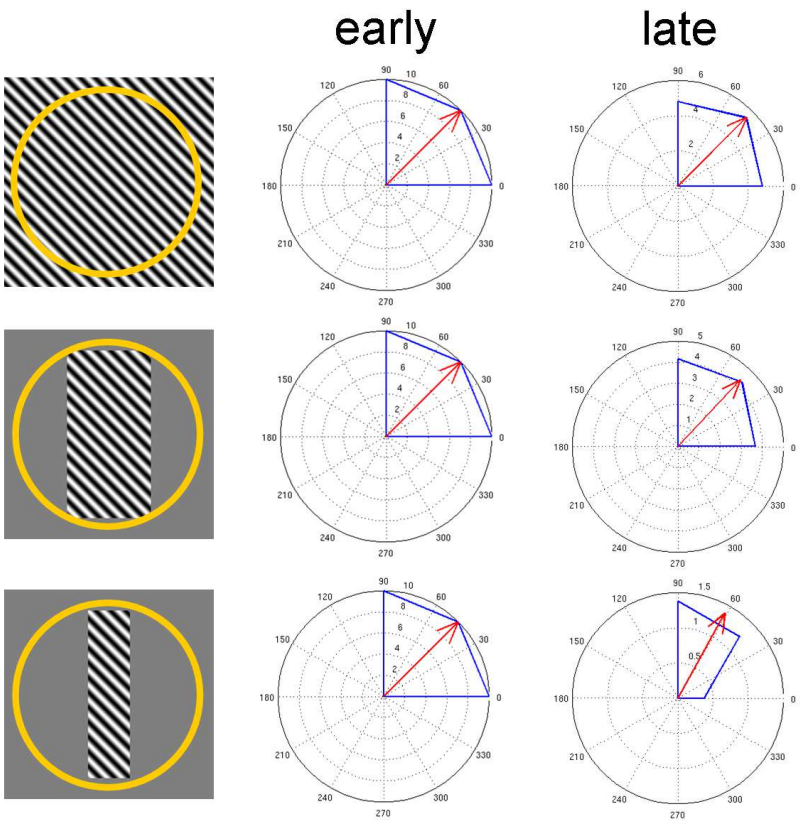

Figure 4: Responses of the MT cells located at the center of the input sequence for the three stimuli tested. In the *late* response it is possible to see that the PD (noted as a red arrow) shifting is clearly marked in the case of barberpole with aspect ratio of 4:1. The barberpole with aspect ratio 2:1 also shows a shifting in its PD towards the vertical direction but it is not so emphasized as the case of aspect ratio 4:1.

3 Results

The outputs of the different stages of our model are represented by polar diagrams, showing in each case, the response of a certain cell for different stimulus orientation.

We show the responses of two V1 cells (see Figure 3), one located at the center of the barberpole (B) and a second one located at the center of the longest border (A). The diagrams were obtained considering the mean firing rate of each V1 cell ($\bar{r}_i^{V1}$) within the time windows defining the *early* and *late stages*. For these two cells we obtain the *early* and *late* response where the effect of the surround inhibition can be seen. We did not show the response obtained for the drifting grating.

Regarding MT cells, for the three input stimuli we placed eight MT cells at the center of the stimuli for each of the 8 directions we wanted to test. We measured the response of each cell for the *early* and *late* stages, obtaining in both cases, the mean firing rate of each cell ($\bar{r}_i^{MT}$). Figure 4 shows the responses of MT cells putting special emphasis in the *late* response of the barberpole compared to the *late* response of the drifting grating. Although both barberpole's responses show a PD shifting towards the longest border, the effect is really clear in the third case, showing that the aspect ratio is an important issue.

4 Conclusion

Here we presented a simple mechanism to reproduce the PD shifting observed in MT cells [13, 14, 6]. This mechanism is based on the surround suppression done by V1 neurons, where only a local information is still considered. The surround suppression of V1 neurons here operates as a motion contour detection. The global integration and final decision is taken by MT cell, which can be better improved adding additional mechanisms such as isotropic/anisotropic surround interactions (either suppression or facilitation, see, e.g., [9]).

We also showed the relevance of the aspect ratio of barberpoles in the PD shifting of MT cells, which is highly related with perception and physiology measurements. In our case the PD shifting effect is due to the way that the pooling of V1 neurons is carried out. The MT neuron receptive field is modeled as a Gaussian, i.e. the cells located more on the border have a lower connection weight which conveys less influence in the PD shifting mechanisms.

The deviation observed in the preferred direction of MT neurons in our model (about 15°) is smaller than found by Pack et al. (2004) using barber-poles (about $25 - 30^\circ$ in about 150ms). This indicates that the weight given to 2D motion information is still not sufficient to indicate the true global motion in our model. To reach a satisfying solution of the aperture problem, several solutions will be explored. First, change in PD in MT neurons depend on several parameters of the non-linear stage at the end of V1 processing. This is consistent with the finding of Rust et al. [15] that pattern-selectivity in MT neurons depends critically of the nonlinear processing (i.e., divisive normalisation) at both V1 and MT level. Nonlinearities seem to play a fundamental role which needs to be clarified. Second, different types of center-surround interactions working at different spatial scales and with different relative orientation might be at work for extracting 2D features motion. Third, a better diffusion process of 2D information can be achieved through anisotropic interactions

between different locations (see [17]). In future work, we will explore the contributions of these different mechanisms to be able to fully render the dynamics of a population of MT neurons.

Acknowledgments

This work was partially supported by the EC IP project FP6-015879, FACETS and CONICYT Chile. We also thank Emilien Tlapale who kindly cooperated with its implementation of Reichardt detectors.

References

[1] E.H. Adelson and J.R. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2:284–299, 1985.

[2] D.G. Albrecht, W.S. Geisler, and A.M. Crane. *Nonlinear properties of visual cortex neurons: Temporal dynamics, stimulus selectivity, neural performance.*, pages 747–764. MIT Press, 2003.

[3] P. Bayerl and H. Neumann. Disambiguating visual motion through contextual feedback modulation. *Neural Computation*, 16(10):2041–2066, 2004.

[4] P. Bayerl and H. Neumann. Disambiguating visual motion by form–motion interaction – a computational model. *International Journal of Computer Vision*, 72(1):27–45, 2007.

[5] J. Berzhanskaya, S. Grossberg, and E Mingolla. Laminar cortical dynamics of visual form and motion interactions during coherent object motion perception. *Spatial Vision*, 20(4):337–395, 2007.

[6] Richard T. Born, Christopher C. Pack, Carlos Ponce, and Si Yi. Temporal evolution of 2–dimensional direction signals used to guide eye movements. *Journal of Neurophysiology*, 95:284–300, 2006.

[7] M. Carandini, J. B. Demb, V. Mante, D. J. Tollhurst, Y. Dan, B. A. Olshausen, J. L. Gallant, and N. C. Rust. Do we know what the early visual system does? *Journal of Neuroscience*, 25(46):10577–10597, November 2005.

[8] A. Destexhe, M. Rudolph, and D. Paré. The high-conductance state of neocortical neurons in vivo. *Nature Reviews Neuroscience*, 4:739–751, 2003.

[9] X. Huang, T. D. Albright, and G. R. Stoner. Adaptive surround modulation in cortical area mt. *Neuron*, 53:761–770, March 2007.

[10] H.E. Jones, K.L. Grieve, W. Wang, and A.M. Sillito. Surround suppression in primate v1. *Journal of Neurophysiology*, 86:2011–2028, 2001.

[11] D. R. Mestre, G. S. Masson, and L. S. Stone. Spatial scale of motion segmentation from speed cues. *Vision Research*, 41(21):2697–2713, September 2001.

[12] C. C. Pack, M. S. Livingstone, K. R. Duffy, and R. T. Born. End-stopping and the aperture problem: Two-dimensional motion signals in macaque v1. *Neuron*, 39(4):671–680, 2003.

[13] Christopher Pack and Richard Born. Temporal dynamics of a neural solution to the aperture problem in visual area mt of macaque brain. *Nature*, 409:1040–1042, feb 2001.

[14] Christopher Pack, Andrew Gartland, and Richard Born. Integration of contour and terminator signals in visual area mt of alert macaque. *The Journal of Neuroscience*, 24(13):3268–3280, 2004.

[15] NC Rust, V Mante, EP Simoncelli, and JA Movshon. How mt cells analyze the motion of visual patterns. *Nature Neuroscience*, (11):1421–1431, 2006.

[16] M.P. Sceniak, M.J. Hawken, and R. Shapley. Visual spatial characterization of macaque v1 neurons. *Journal of Neurophysiology*, 85:1873–1887, 2001.

[17] É. Tlapale, G. Masson, and P. Kornprobst. Biological model of motion integration and segmentation based on form cues. Technical Report RR-6293, INRIA, sep 2007.

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 scientifique
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