

**ABSTRACT**

Modeling visual perception of motion by connectionist networks offers various areas of research for the development of real-time models of dynamic perception-action. In this paper we present the bases of a bio-inspired connectionist approach that is part of our development of neural networks applied to autonomous robotics. Our model of visual perception of motion is based on a causal adaptation of spatiotemporal Gabor filters. We use our causal spatiotemporal filters within a modular and strongly localized architecture that performs a shunting inhibition mechanism. This model has been evaluated on artificial as well as natural image sequences.

**INTRODUCTION**

Processing and interpretation of visual perception of motion has an increasing importance for autonomous robotics. In this domain, connectionist models have brought their power of generalization and their robustness to noisy or incomplete information. Their intrinsic parallelism combined with local processings offers various areas of research for the development of real-time models of perception-action applied to dynamic visual perception of motion.

For a robust, adaptive and embeddable computation paradigm in digital reconfigurable circuits (e.g. FPGA, Field Programmable Gate Arrays) our preliminary developments of neural models focus on highly distributed and local processings, as such implementation devices require. Our claim is that bio-inspiration is a good guideline to satisfy such constraints.

We first propose a rapid survey of the state of the art in neurophysiology and of the algorithms that have been proposed for the estimation of the optical flow. We then present the bases of a bio-inspired connectionist approach. It uses a particular adaptation of the spatiotemporal Gabor filters. It also takes advantage of a modular and strongly localized approach for the visual perception of motion that handles a shunting inhibition mechanism. Finally we describe the evaluation of this model on several image sequences.

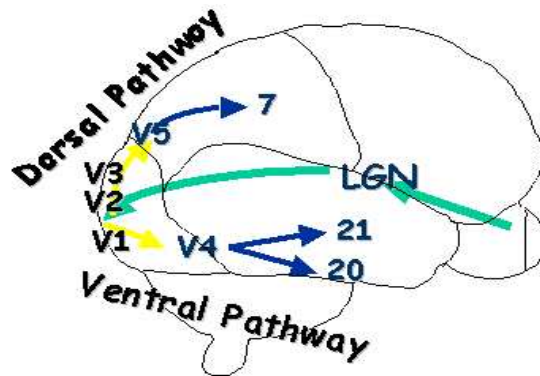
**STATE OF THE ART**

Currently, the most bio-inspired computer models for motion detection use local detections by means of the integration of various directions for various scales and spaces to end in a global answer [40, 1]. Motion detection, spatiotemporal local inhibition, and integration are the main ideas of neurosciences research that will inspire our connectionist conception.

**Biological point of view**

We now know that motion detection and analysis are achieved by means of a cascade of neural operations [33], namely: the

\*Mexican scholarship of the National Council of Science and Technology (CONACyT, 113287/169076), México.



(a) Cerebral Cortex

FIG. 1 – Optical flow pathways

detection of local motion signals within restricted regions of the visual field and their integration into more global descriptions of the direction and speed of object motion.

**The course of signals from the retina**

From the retina up to the cerebral cortex of a human being, seventeen different areas take part in vision processing [36]. The main areas may be organized in four stages: acquisition and compression of light signals in the retina; their relaying in the lateral geniculate nucleus (LGN); their cortical analysis in the primary visual area (V1), and their secondary treatment in areas 20 and 21 of the temporal cerebral cortex<sup>1</sup> and in area 7 of the parietal cerebral cortex<sup>2</sup> (see figure 1).

Two pathways may be discovered in this course of visual signals [32]: the ventral or occipito-temporal pathway<sup>3</sup> and the dorsal or occipito-parietal pathway<sup>4</sup>. The first one mostly consists of parvocellular cells and it is responsible for the perception of the objects and of their shape, and the second one mostly consists of magnocellular cells and it is responsible for motion and space perception. In the present study, we are particularly interested in the dorsal pathway and in areas specialized in motion analysis.

**Direction and orientation detection in V1**

The human brain areas that are specialized in motion perception are [33]: the primary visual area (V1), the middle temporal area (MT)<sup>5</sup>, the middle superior temporal area (MST), the kine-

1. These Brodmann areas approximately correspond to infero-temporal area, IT.

2. This Brodmann area corresponds to the posterior parietal area, PP.

3. It includes visual areas V1, V2, and V4 and the IT.

4. It includes visual areas V1, V2, and V3, the middle temporal area -MT/V5- and the PP.

5. MT+ for human beings

(SIS).

The first cortical analysis is performed in V1 by ensuring contrast sensitivity thanks to extended receptive fields. These neurons mainly send their extensions in the vertical direction and they are tuned to a preferred direction of motion [19] so that they perform a local analysis of motion energy that is called a filter. These “filters” detect the presence of basic features of moving objects within their receptive fields.

These orientation-selective cells may be modeled as spatio-temporal filters [1, 39] and their receptive fields may be modeled as a product of inhibitory and excitatory interactions in space and time.

On the other hand, contrast detection is sufficient for the identification of motion direction, so that the visual mechanisms that extract movement are built from direction-selective primitives [6].

All these cells are selective in terms of preferred direction and spatial frequency. They are not very sensitive to the composition of the wavelength of the motion stimulus, and they are limited to a local motion detection. These cells “see” only a part of the scene. Their response is ambiguous because in every direction there is a family of selective cells that react to the same spatio-temporal changes of brightness as a local cell, and that generate the same response. This problem is known in computer vision as the “aperture problem”<sup>6</sup>.

In short, the local motion of a retinal image is extracted by neurons in V1 that have a receptive field similar to a small spatially bounded window where they can detect the presence of movement in a specific direction. *This strongly localized processing based on lateral interactions is our first source of inspiration towards a bio-inspired model for motion detection and estimation.*

But the visual perception of motion is not completely determined by the local responses in the neural receptive fields. These responses are also handled to obtain speed informations after having collected and combined them from V1 and after having grouped them together in MT. The ambiguity of individual neural responses is resolved by this combination of signals. This backpropagation of information from MT back to V1 is part of our further works (this paper only deals with forward and lateral interactions within V1). The cortical and temporal analysis of motion (from V1 to MT/MST areas) in the brain will be the main source of our bio-inspired connectionist approach.

### Lateral connections and shunting inhibition

In the primary visual cortex, neurons with similar receptive field properties are gathered into functional modules. Within each module, inhibitory and pyramidal cells are densely interconnected in a local neural circuit [38]. At cortical sites where horizontal connections converge towards the same neural populations, neural responses turn away from linear integration due to of the emergence of strong inhibitory post-synaptic potentials (or emergent inhibition). These horizontal connections get such a strong inhibition from local inhibitory networks.

A *shunting inhibition* occurs when the potential of the inhibi-

6. This problem was first identified in [17]: there is a local ambiguity that is due to a linear constraint between the two components of the optical flow, and that results in confusing speed direction with the normal of local contrasts. This local ambiguity was later termed the “aperture problem” [24].

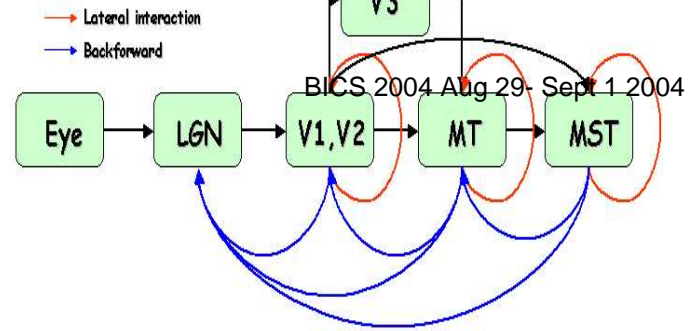


FIG. 2 – Relation between visual cortex areas (see subsection ).

tory synapse is close to the resting potential of the cell. Such an inhibition locally cancels excitation when it lies on the path between the location of the excitatory synapse and the cell body [21].

Although a shunting-inhibition mechanism is theoretically able to produce the localized interactions that are necessary to explain many direction-selective properties, it has not been possible to perform intracellular tests within dendrites to confirm this mechanism [13].

A lateral interaction mechanism of all locally strong inhibitions between different directions of motion in V1 would yield a global direction-selective emergent response. Our approach has been developed so as to take advantage of such a mechanism (see subsection ).

Figure 2 summarizes the different interactions between cortical areas that have been described in this section.

### Mathematical point of view

Motion detection and estimation are low level cognitive tasks, whereas its segmentation and interpretation are high level cognitive tasks. Our approach keeps dealing with low level tasks, that make use of optical flow estimation and shunting inhibition mechanisms.

#### Optical flow

The “optical flow is the distribution of apparent velocities of movement of brightness patterns in an image” [18]. It tries then to match pixels having the same intensity.

According to the brightness conservation principle [18]:

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (1)$$

where  $I(x,t) = I(x + \delta x, y + \delta y, t + \delta t)$  is the brightness at point  $(x,y)$  of the image. Assuming that it remains constant in time  $t$ , the fundamental principles of the various approaches of optical flow estimation [18, 4, 10] are derived: matching (correlation-based), gradient (intensity-based differential), stochastic (relaxation), and spatiotemporal (frequency-based filtering).

The first approach, correlation-based, is based on the local conservation of intensity distribution [2, 28]. Motion vectors are built from local image matching.

The second approach, intensity-based differential, is based on the intensity conservation of an image within a small interval of time [18, 26, 22]. Motion vectors are computed from local intensity temporal derivatives.

concerns motion discontinuities [16, 5]. Intermittent intensities are modelled with the help of stochastic processes or Kalman filters.

But the aperture problem is critical with these three approaches. We are more interested in frequency-based filtering approaches. Among their advantages, such approaches appear as more robust to the aperture problem. In these approaches, the spatiotemporal frequencies are considered along with the speed of motion stimulus, and the optical flow becomes the identification of an energy surface in the space of spatiotemporal frequencies. Movement-sensitive mechanisms that are based on the spatially and temporally oriented energy in the space of Fourier have been found to be able to estimate the motion in places where the other approaches fail. Therefore, motion detection in images means here the extraction of spatiotemporal orientations [1, 15, 11].

In that aim, two sub-approaches are known: phase-based filtering and energy-based filtering. In the first case, the constituents of speed may be defined in terms of immediate movement at phase-level of the outputs of Gabor filters tuned to speed [9, 42]. In the second case, these spatiotemporal oriented filters may be built on separate filters, i.e. the spatial and temporal filterings may be sequentially applied to the input image [1, 41]. Obtaining speed would be based on a codification scheme of populations where speed is expressed as the ratio of the outputs of energy filters tuned to various speeds. Assuming a linearity, the output of each filter changes with various contrasts but their ratio remains constant.

Among the advantages of energy methods one could mention an intrinsic smoothing which reduces the effects of the ‘‘aperture problem’’, a better robustness to noise, and good-quality results for natural image sequences [35]. Drawbacks are a high computational cost and a lack of precision near the borders of the movement.

The principles of spatiotemporal approaches consist in localizing non-null energy surfaces or speed surfaces that correspond to an apparent translation movement on a spatial support bounded by means of spatiotemporal oriented filters.

### Shunting inhibition equation

Recent researches on inhibition mechanisms propose two kinds of models: subtractive inhibition [34] and shunting inhibition [37, 13]. The first one is the most used. The second one uses inhibitory synapses that are applied at given points on dendrites and that divide (rather than subtract from) the potential change which passively propagates from more distant synapses [3].

Torre and Poggio suggested that inhibition acts separately within each branch of the cell’s dendritic tree. This inhibition works through a synapse that causes a local change of the membrane conductance (shunting inhibition) and a slight hyperpolarization.

Considering a patch of membrane that receives excitation ( $g_e(t)$ ) and shunting inhibition ( $g_i(t)$ ), voltage  $V$  obeys

$$C \frac{dV(t)}{dt} + (g_e(t) + g_i(t) + g_l(t))V(t) = g_e(t)E_e + g_l(t)E_l \quad (2)$$

where  $C$  is the membrane capacitance,  $g_l(t)$  is the membrane leak conductance and  $E_e$  (resp.  $E_l$ ) is the reversal potential of  $g_e(t)$  (resp.  $g_l(t)$ ).

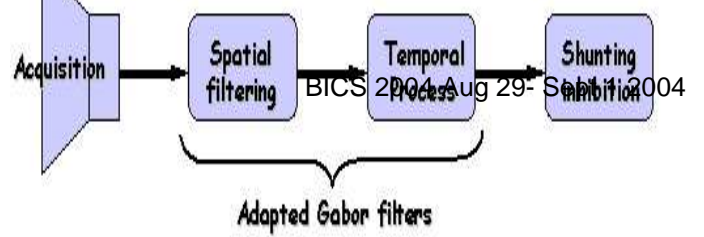


FIG. 3 – Main steps of our connectionist approach (see subsection).

If  $g_i(t) \gg g_e(t)$  then  $V(t)$  tends towards the following equilibrium value

$$V(t) \rightarrow \frac{g_e(t)E_e + g_l(t)E_l}{g_i(t)} \rightarrow 0 \quad (3)$$

With a low contrast, high order linearities may be neglected by an estimate Fourier series. Therefore, this inhibition is division-like rather than subtraction-like.

## A BIO-INSPIRED CONNECTIONIST MODEL

With the aim of realizing a bio-inspired connectionist model based on highly local and distributed computations, and embedded onto a mobile robot that moves about in a dynamic environment with moving targets, we first propose a neural architecture that models the following aspects: lateral interactions in V1 with a shunting inhibition mechanism, and spatiotemporal filtering based on an adaptation of Gabor filters. This filtering is performed in order to model the magnocellular cells<sup>7</sup> that may be seen as motion sensors for the extraction of generalized spatiotemporal gradients of the image intensity and of its temporal derivatives [31].

### Causal spatiotemporal filters

Receptive fields modeled by bidimensional spatial Gabor filters were proposed by Marcelja [23] and they were discovered in biological vision systems [31].

The spatial part of a standard Gabor function approximates well the spatial profile of receptive fields in the cerebral cortex [20]. But its temporal part is non-causal (negative weights are assigned to immediate past images). Several more causal approaches have been proposed. Adelson and Bergen used an asymmetric spatial distribution [1] and Gzrywacz and Yuille made it with Gabor functions in spatiotemporal co-dependence [11, 12]. They concluded that directional selectiveness is equal to orientation in space-time.

Our approach handles causality in a simple and local way with a strong hypothesis that ensures the ability to detect local motions.

Let  $I(x,y,t)$  be an image sequence representing the shape of intensity in the time-varying image, assuming that every point has an invariant brightness. Let us assume that  $I(x,y,t) = I(x - ut, y - vt)$  where  $(u,v)$  is the motion vector of a small region of the image, and where  $I(x,y)$  is the frame of image sampling at time  $t = 0$ . Thanks to the hypothesis of an high enough sampling frequency to ensure local motion detection<sup>8</sup>, we may assume an

7. The spatiotemporal Gabor filters also model the simple cells in the visual cerebral cortex of primates [27, 14]

8. This hypothesis is also made in other works with local detection [30].

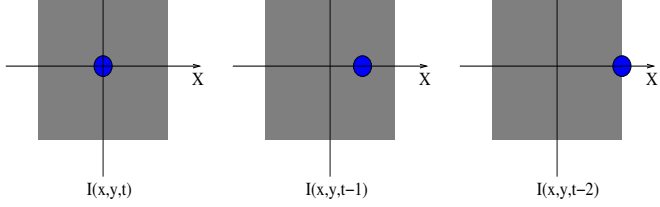


FIG. 4 – Temporal decomposition of our causal spatio-temporal filter for  $\theta = 180^\circ$ ,  $v = 1$  and  $\tau = 3$  (see subsection ).

immediate constant local speed. Therefore, for a given supposed motion direction and speed, we expect to identify a local motion by finding a spatial contrast at expected places and times, as figure 4 shows in the case of an horizontal movement with unit speed.

By applying then an oriented Gabor filter,  $G_\theta(x,y)$  with  $0 \leq \theta < 2\pi$ , in  $I(x,y)$  we obtain (intensity conservation principle):

$$0 = \int \int \frac{dI(x,y,t)}{dt} \Big|_{t=0} G_\theta(\hat{x} - \hat{u}, \hat{y} - \hat{v}) dx dy$$

$$= \frac{d \int \int I(x - ut, y - vt) G_\theta(\hat{x} - \hat{u}, \hat{y} - \hat{v}) dx dy}{dt} \Big|_{t=0} \quad (4)$$

where the rotational equations give

$$\hat{x} = (x - \xi) \cos \theta - (y - \eta) \sin \theta$$

$$\hat{y} = (x - \xi) \sin \theta + (y - \eta) \cos \theta \quad (5)$$

with  $(\xi, \eta) \in \Upsilon$  a small neighbourhood around  $(x, y)$ , and

$$\hat{u} = \frac{t-t'}{\tau-1} v \cos \theta, \quad \hat{v} = \frac{t-t'}{\tau-1} v \sin \theta \quad (6)$$

for  $\tau$  consecutive images,  $t' \leq t$ , and a supposed velocity  $v$  that ranges from  $-\omega$  to  $\omega$  where  $\omega$  is the number of supposed absolute speeds.

$(\hat{x}, \hat{y})$  is the place where the oriented Gabor signal is going to be computed in a standard way:

$$G_\theta(\hat{x}, \hat{y}) = \frac{1}{2\pi\sigma_x\sigma_y} \exp\left(-\frac{\hat{x}^2}{2\sigma_x^2} - \frac{\gamma^2 \hat{y}^2}{2\sigma_y^2}\right) \exp(2\pi i \frac{\hat{x}}{\lambda} + \phi) \quad (7)$$

is the response function to the impulse of the Gabor filter that models the function of the ganglion magnocellular cells, where  $\gamma$  is the eccentricity of the receptive field and  $\sigma_x, \sigma_y$  its dimensions,  $\lambda$  is the wavelength and  $\phi$  is the phase.

By discretizing equation 4, we finally compute the following spatio-temporal filter<sup>9</sup>:

$$f_{\tau, \theta, v}(x, y, t) = \frac{1}{\tau} \sum_{t'=0}^{\tau-1} \sum_{(\hat{x}, \hat{y})} G_\theta(\hat{x} - \frac{t-t'}{\tau-1} v \cos \theta, \hat{y} - \frac{t-t'}{\tau-1} v \sin \theta) \quad (8)$$

The measure that is obtained by a single filter is not able to determine the 2D motion vector. It is necessary to use a set of filters that differ only in orientation. Then they are gathered in a

<sup>9</sup>. This basic motion feature detection is combined with the detection of strong motionless contrasts, so as to suppress any confusion between motion and contrast.

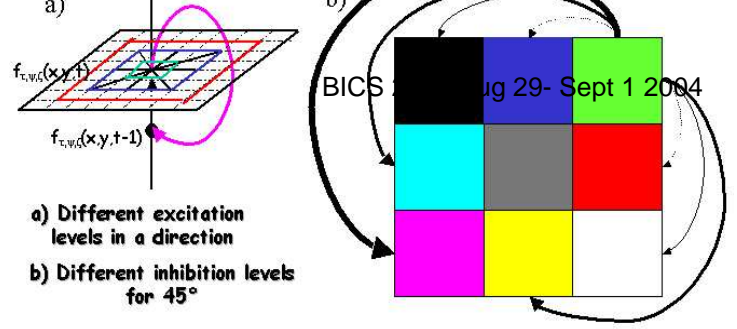


FIG. 5 – The different levels for excitatory and inhibitory connections that depend on influence range  $\kappa$  (see subsection ).

vector called *motion sensor vector* where every orientation is a motion sensor.

### Neural architecture

We propose a fully modular connectionist model that is strongly local and computationally distributed, and that is inspired by neurophysiological results about magnocellular pathways. It uses a set of excitatory and inhibitory lateral interactions inspired by the cortical columns that stem from V1 (see sections and ) and it uses a strongly local scheme of connectivity with an *influence range* that depends on the local response obtained by equation 4 (see the main steps in figure 3). This variable influence range is motivated by lateral interaction models [8, 7], where the concept of adaptive neighbourhood to select a working scale at each point of the image and for each computation level is fundamental.

After the acquisition of images in RGB format, the following steps are performed on greyscale images:

1. Causal spatiotemporal filtering in two phases:
  - (a) Spatial filtering for all regions  $\Upsilon$  in the image.
  - (b) Temporal processing on the image sequence for  $\omega$  different supposed absolute velocities.
2. Shunting inhibition mechanism.
  - (a) Computation of the influence range  $\kappa$ .
  - (b) Excitation-inhibition process.
3. Standard winner-take-all properties of competitive processing.

### Spatiotemporal filtering

We apply equation 8. Parameters  $\gamma, \sigma, \lambda$  and  $\phi$  are fixed according to the suggestions of [29].

### Shunting inhibition mechanism

We adapt equation 2 according to ideas of Moga [25] and Fernandez et al. [8].

We propose for a local time,  $T \ll t$ , and its filtering response at time  $t$ ,  $F(x, y, T = 0) = f_{\tau, \theta, v}(x, y, t)$ , then

$$\frac{\delta F(x, y, T)}{\delta T} = -A \cdot F(x, y, T) + (B - F(x, y, T)) \cdot g_e(x, y, T) + (C - F(x, y, T)) \cdot g_i(x, y, T) \quad (9)$$



$$\begin{aligned}
g_e(x,y,T) &= f_{\tau,\theta,v}(x,y,t) + \sum_{(x',y') \in \Lambda_1} (w_{(x',y')} \cdot F(x',y',T-1)) \\
g_i(x,y,T) &= \sum_{(x',y') \in \Lambda_2} (v_{(x',y')} \cdot F(x',y',T-1))
\end{aligned} \tag{10}$$

with  $A, B$ , and  $C$  being constants.  $g_e(x,y,T)$  is the magnitude of excitation with an influence range  $\kappa_{x,y}$  that is computed as the integer part of

$$\frac{\Gamma |F(x,y,T)|}{\text{saturation}} \tag{11}$$

where  $\Gamma$  is the supposed maximum influence range, and  $\text{saturation} = 2 \max_{x,y} (|f_{\tau,\theta,v}(x,y,t)|)$  is the maximum value of all Gabor filters. This influence range determines the size of the influence blob in sets  $\Lambda_1$  and  $\Lambda_2$ , and as a consequence it determines their cardinality. Weights  $w_{(x',y')}$ , with  $(x',y') \in \Lambda_1$ , are the excitation connection weights.  $g_i(x,y,T)$  is the magnitude of inhibition with an influence range  $\kappa_{x,y}$ , and weights  $v_{(x',y')}$  are the inhibition connection weights<sup>10</sup> (see figure 5).

### Standard WTA properties

For  $\zeta$  different supposed velocities  $v_i$ , for  $\psi$  possible directions, and in order to have only one local response, we apply a standard WTA to all supposed velocities and to all directions. Then

$$f_{\tau,\Theta,V}(x,y,t) = \max_{v \in -\omega, \dots, \omega} \left( \sum_{\theta} f_{\tau,\theta,v}(x,y,t) \cdot \hat{d}_{\theta} \right) \tag{12}$$

where  $V$  is the winner velocity in direction  $\Theta$  of pixel  $(x,y)$  at time  $t$ ,  $\hat{d}_{\theta}$  is the unit vector in direction  $\theta$ .

## RESULTS

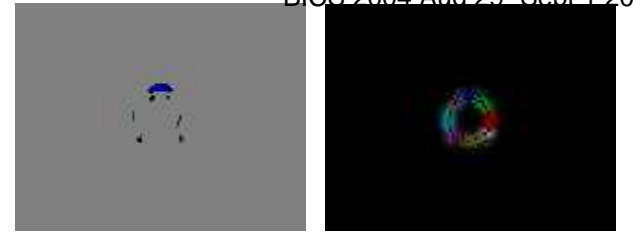
Our model has been tested with synthetic and natural image sequences. First results show that our approach induces some improvements in both motion feature detection and signal enhancement, thanks to our causal spatiotemporal filters and to the shunting inhibition mechanism respectively. Nevertheless, parameterization is still a difficult problem (a balance must be obtained between the enhancement of motion detectors and the spreading of their enhanced outputs).

A simple example of synthetic images corresponds to a ball that goes forward and backward in a corridor. As shown in figure 6, an improvement of basic motion feature detection is given by our spatiotemporal filters, that are essentially based on our strong hypothesis of a high enough sampling rate that makes possible the local detection of an immediate constant speed.

Natural image sequences raise more contrast-linked problems than synthetic ones. An example of real images is the famous Hamburg taxi scene from the University of Hamburg<sup>11</sup>. In this sequence, contrast-based motion detection appears to “forget” some moving objects with weak contrasts. As shown in figure 7, our model takes advantage of a densely interconnected architecture of shunting inhibition to enhance neuron outputs, so that these “forgotten” objects re-emerge.

10. In direction  $\theta$ , a neural response in  $(x,y)$  inhibits  $\frac{100\rho}{\pi}\%$  in direction  $\theta + \rho$  (with  $0 \leq \rho \leq \pi$ )

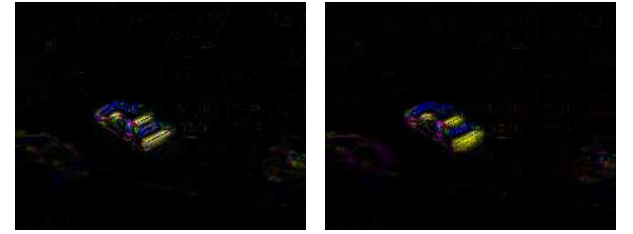
11. In this sequence, there are four moving objects: a taxi turns the corner, a car in the lower left part drives from left to right, a van in the lower right part drives from right to left and a pedestrian walks in the upper right part. In this sequence, there is a slight egomotion from left to right and from up to bottom.



(a) Gabor

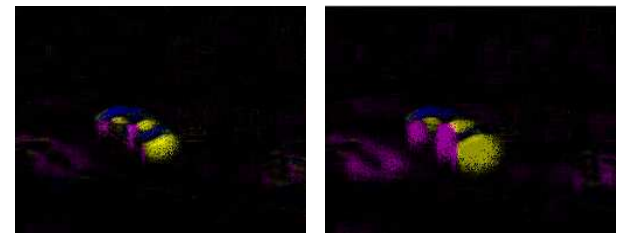
(b) causal

FIG. 6 – Synthetic image sequence with  $\tau = 3$ ,  $\omega = 2$  (ball going forward): (a) motion feature detection with standard Gabor filters, (b) motion feature detection with our causal spatiotemporal filters



(a) input of the shunting inhibition process

(b) after 20 epochs



(c) after 50 epochs

(d) after 100 epochs

FIG. 7 – Shunting inhibition mechanism applied to a real image sequence. Parameters in equation 9 are set according to the ideas proposed by Moga [25], namely:  $A = 1.4f_{sat}$ , and  $B = C = 0.95f_{sat}$ , where  $f_{sat}$  is the neuron saturation value. A maximum influence range  $\Gamma = 7$  is chosen. Parameters are  $\tau = 3$ ,  $\omega = 2$ , and 0.0001 as learning rate.

Perception of motion is vital for living beings. According to neuro-physiological results of primate vision study, we set out that: the retina plays a role of light signals filtering, the lateral geniculate nucleus performs a binocular integration, V1 carries out a spatial analysis, MT and MST carry out a temporal analysis. The strong reduction of information from the retina and the huge amount of lateral interactions in superior areas bring to the foreground the dynamics of our world perception. These various levels of processing and their respective feedback illustrate the cooperations that exist between retinotopically close columns. Our connectionist architecture makes use of such cooperations.

In this paper, we have presented the early stages of a strongly localized bio-inspired connectionist model for the perception of motion. This first part of our work models forward and lateral interactions in V1, thanks to the use of adapted spatiotemporal filters based on Gabor filters, and to a shunting inhibition mechanism. Similar interactions within MT are studied, and their influence will be backpropagated towards V1.

Current efforts focus on the definition of a model of the dorsal pathway for the detection of the movement of one or several objects in various dynamic scenes after egomotion extraction. Further works will aim at integrating the various stages of the dorsal pathway followed by the optical flow. These works will still stress our modular, purely connectionist and strongly localized point of view.

## REFERENCES

[1] E. H. Adelson and J. Bergen. Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, 2(7):284–299, 1985.

[2] P Anandan. A computational framework and an algorithm for the measurement of visual motion. *International Journal Computer Vision*, 2:283–310, 1989.

[3] Michael A. Arbib, editor. *The Handbook of Brain Theory and Neural Networks*. Bradford Book MIT Press, Cambridge, Massachusetts 02142, USA, second edition, 2003.

[4] S. S. Beauchemin and J. L. Barron. The computation of optical flow. Technical report, Department of Computer Science, June 1995.

[5] M. J. Black. *Robust incremental optical flow*. PhD thesis, Yale University, Yale, 1992.

[6] A. M. Derrington and G. B. Henning. Detecting and discriminating the direction of motion of luminance and colour gratings. *Visual Research*, 33:799–811, 1993.

[7] M. A. Fernández, J. Mira, M. T. López, A. Manjarres, and S. Barro. Local Accumulation of persistent activity at synaptic level: application to motion analysis. In J. Mira and F. Sandoval, editors, *From Natural to Artificial Neural Computation, IWANN'95*, volume 930, pages 137–143, Berlin, 1995. Springer.

[8] Antonio Fernández-Caballero, José Mira-Mira, Miguel A. Fernández-Graciani, and Ana E. Delgado. On motion detection through a multi-layer neural network architecture. *Neural Networks*, 16:205–222, 2003.

[9] D. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International Journal Computer Vision*, 5:77–104, 1990.

Flow Algorithms. Technical report, Computer Science Department, 1999.

[11] N. M. Grzywacz and A. L. Yuille. A model for the estimate of local image velocity by cells in the visual cortex. In *Proceedings Royal society London B*, volume 239, pages 129–161, London, 1990.

[12] N. M. Grzywacz and A. L. Yuille. Theories for the visual perception of local velocity and coherent motion. In *Computational models of visual processing*, pages 231–252. M. I. T., 1991.

[13] Norberto M. Grzywacz and David K. Merwine. *Directional Selectivity*, pages 353–358. In Arbib [3], second edition, 2003.

[14] P. Hammond and J. Reck. Influence of velocity on directional tuning of complex cells in cat striate cortex for texture motion. *Neuroscience Letters*, 19:309–314, 1981.

[15] David J. Heeger. A model for the extraction of the image flow. *Journal of the Optical Society of America A*, 4(8):1455–1471, 1987.

[16] F. Heitz and P. Bouthemy. Multimodal estimation of discontinuous optical flow using Markov random fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(12):1217–1232, 1993.

[17] Berthold K. P. Horn and B. G. Schunk. “Determining optical flow”: a retrospective. *Artificial Intelligence*, 59:81–87, 1993.

[18] Berthold K. P. Horn and Brian G. Schunk. Determining optical flow. Technical Report A. I. Memo No. 572, Artificial Intelligence Laboratory, April 1980.

[19] D. H. Hubel and T. N. Weisel. Receptive fields, binocular interaction and functional architecture in the cats visual cortex. *Journal Physiology*, 160:106–154, 1962.

[20] J. Jones and L. Palmer. An evaluation on the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58:1233–1258, 1987.

[21] Christof Koch, Chun-Hui Mo, and William Softky. *Single-Cell Models*, pages 1044–1049. In Arbib [3], second edition, 2003.

[22] B. Lucas and T. Kanade. An iterative image registration technique with an applications to stereo vision. *DARPA IU Workshop*, pages 121–130, 1981.

[23] S. Marcelja. Mathematical description of the reponses of simple cortical cells. *Journal of the Optical Society of America A*, 70/11:1297–1300, 1980.

[24] D. Marr, editor. *Vision*. Freeman, San Francisco, CA, 1982.

[25] Sorin Moga. *Apprendre par imitation: une nouvelle voie d'apprentissage pour les robots autonomes*. PhD thesis, Université de Cergy-Pontoise, Cergy-Pontoise, France, September 2000.

[26] H. H. Nagel. On the estimation of optical flow: Relation between different approaches and some new results. *Artificial Intelligence*, 33:299–324, 1987.

[27] W. T. Newsome, M. S. Gizzi, and J. A. Movshon. Spatial and temporal properties of neurons in macaque MT. *In-*

- [28] M. Ogata and T. Sato. Motion-detection model with two stages: Spatiotemporal filtering and feature matching. *Journal of the Optical Society of America A*, 9(3):377–387, 1992.
- [29] N. Petkov and P.Kruizinga. Computational models of visual neurons specialised in the detection of periodic and aperiodic oriented visual stimuli: bar and grating cells. *Biological Cybernetics*, 76:83–96, 1997.
- [30] Patrick Pirim. Procédé et dispositif fonctionnant en temps réel pour le repérage et la localisation d’une zone en mouvement relatif dans une scène, ainsi que pour la détermination de la vitesse et de la direction du déplacement. Technical Report WO 98/05002, France, February 1998.
- [31] D. Pollen and S. Ronner. Phase relationships between adjacent simple cells in the visual cortex. *Science*, 212:1409–1411, 1981.
- [32] Hengyi Rao, Tiangang Zhou, Yan Zhuo, Silu Fan, and Lin Chen. Spatiotemporal Activation of the Two Visual Pathways in Form Discrimination and Spatial Location: A Brain Mapping Study. *Human Brain Mapping*, 18:79–89, 2003.
- [33] Robert Sekuler, Scott N. J. Watamaniuk, and Randolph Blake. Motion Perception. *Steven’s Handbook of Experimental Psychology*, 1:121–176, 2002.
- [34] Eero P. Simoncelli and David J. Heeger. A Model of Neural Responses in Visual Area MT. *Visual Research*, 38(5):743–761, 1998.
- [35] Adrian Spinei. *Estimation du mouvement par triades de filtres de Gabor: Application au mouvement d’objets transparents*. PhD thesis, Laboratoire des Images et des Signaux, October 1998.
- [36] S. Sunaert, P. van Hecke, G. Marchal, and G. A. Orban. Motion-responsive regions of the human brain. *Experimental Brain Research*, 127:355, 1999.
- [37] V. Torre and T. Poggio. A synaptic mechanism possibly underlying directional selectivity to motion. *Proceedings Royal society London B*, 202:409–416, 1978.
- [38] Thomas R. Tucker and Lawrence C. Katz. Recruitment of Local Inhibitory Networks by Horizontal Connections in Layers 2/3 of Ferret Visual Cortex. *Journal of Neurophysiology*, 89:501–512, 2003.
- [39] J. P. H. van Santen and G. Sperling. Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, 1:451, 1984.
- [40] J. P. H. van Santen and G. Sperling. Elaborated Richardt detectors. *Journal of the Optical Society of America A*, 2:300–321, 1985.
- [41] A. B. Watson and A. J. Ahumada Jr. Model of human visual motion sensing. *Journal of the Optical Society of America A*, 2(2):322–341, 1985.
- [42] A. M. Waxman, J. Wu, and F. Bergholm. Convected activation profiles and receptive fields for real time measurement of short range visual motion. In *IEEE CVPR*, pages 717–723, Ann Arbor, 1988.