

A Physicist Approach to First-order Analysis of Optic Flow Fields in Extrastriate Cortical Areas

Silvio P. Sabatini, Fabio Solari, Roberto Carmeli,
Paolo Cavalleri, and Giacomo M. Bisio

Department of Biophysical and Electronic Engineering, University of Genoa
Via Opera Pia 11a - 16145 Genova - ITALY silvio@dibe.unige.it

Abstract

A useful representation of visual motion information can be based on first-order elementary components of the optic flow. Medial superior temporal (MST) cells are probably better suited for the analysis of visual motion in terms of such elementary components. This paper discusses the processing stage in the visual pathway from middle temporal (MT) to MST cells. The processing of MST cells is formulated as a physicalist operation, exploiting a log-polar representation of stimulus velocity performed by MT cells. A wave-like spatiotemporal activation function is introduced to describe the global properties of optic flow at MT level. MST cells perform oriented filtering operations on this activation function by approximating a template matching between the optic flow stimulus and the cell preferences for gaze-centered elementary flow components (EFCs).

1 Introduction

Optic flow refers to the changes in the optic array induced by relative motion between an observer and the surrounding world and plays a relevant role in the perception of three-dimensional motion [1]. Information about this overall motion can be extracted from the optic flow at different stages of cortical visual processing, from the primary visual cortex (V1) through the middle temporal area (MT) to the medial superior temporal area (MST). Hierarchical projections among these areas represent the major stream of motion information processing in the primate visual system. Ascension towards extrastriate areas is characterized by

a systematic increase in receptive field size and complexity. Specifically, the sensitivity to large-field complex motion stimuli such as rotation and expansion/contraction [2] [3] [4] [5], in the dorsal portion of the medial superior temporal area (MSTd) supports the view that MST is suited for the analysis of wide-field egomotion induced by head and body movements.

Although the brain uses some kind of optic flow information when looking at moving stimuli, it is not clear how the brain represents and analyzes optic flow fields. Theoretical studies have shown that a smooth optic flow field can be locally approximated by its first-order spatial derivatives or elementary flow components EFCs which have been identified with usual vector operators *div*, *curl*, and *def*, i.e., rotation, expansion, and two components of shear [6], respectively. When the optic flow is characterized by a constant rate of change, as it occurs for egomotion, its first-order spatial properties represent “differential invariants” (i.e., they are almost piecewise constant over rather large regions of the field of view) that almost completely characterize the whole vector field. Such differential invariants are likely candidates for the representation of motion information in extrastriate area MST, where cells with large receptive fields integrate over wide regions of the visual field. This view has been partially supported by recent experimental studies reporting the observation that the preference of many MST neurons for EFCs is often independent of the position of the stimulus across the receptive field [5] [4] [3] [7].

Behavioural properties of MST cells, as postulated by theoretical vector calculus, would require high degree of structural specificity, which might be more than what

can be determined genetically or learned by experience. However, it is not necessary for any physiological system to be able to compute pure vectorial quantities, but it is sufficient to approximate them through spatiotemporal operators, taking advantage of retinocortical transformation and of a proper *physicalist* representation of the optic flow.

In this paper we propose a physicalist model of extrastriate cortical motion processing where optic flow invariants are mapped as spatiotemporal patterns of activation on a log-polar MT representation, and MST processing is modeled by spatiotemporal filtering operators. The characteristic features of our MST-like units, that resemble those observed in real cells, result as emerging properties of the geometrical organization of the model.

2 Cortical representation of optic flow

2.1 Log-polar mapping

Retinotopic organization of early cortical visual areas was first demonstrated by Hubel and Wiesel in V1 [8] but a similar topography is also observed in area MT [9]. Formally, a complex-log transformation provides an accepted model of the mapping from the retina to the (primary) visual cortex in primates [10]. This transformation is described as a conformal mapping of the polar (retinal) plane $\mathbf{z} = re^{i\theta}$ onto a cartesian (cortical) plane $\mathbf{x} = x + iy$:

$$\mathbf{x} = \log(\mathbf{z} + \alpha) \quad (1)$$

where α is a small real constant introduced to avoid a singularity in the mapping.

According to such polar mapping fovea-centered circles are mapped to vertical lines, and radial lines to horizontal lines. This has the effect of transforming dilations and rotations in fovea-centered retinal images to translations in cortical images. The central column of Fig. 1 shows the log-polar mappings of different optic flow fields $\mathbf{v}(x, y)$ obtained by resampling the retinal optic flow fields according to Eq. (1).

2.2 The cortical flow

Following a Gestalt approach, we assume that the cortical code of the optic flow field should represent morphologically the triggering perceptual event (i.e., the spatiotemporal structure of the optic array, cf. [1]). This problem can be embodied by the *isomorphism principle* [11] asserting that perception takes place when systems reactions occur isomorphically to the impinging pattern of activation, i.e., the reacting brain process have the same structure as the perceptual event it gave rise to. Accordingly, we formulate the hypothesis that, at MT level, a “motion event” in the visual space is isomorphically associated to a “motion event” on the cortical surface. More specifically, the activation of a pool of MT cells by a coherent motion pattern yields, at a macroscopic scale, a true *flow* of cortical activity, as a *physicalist* representation of the vectorial information.

Formally, to each vector $\mathbf{v}(\mathbf{x}')$ of the optic flow, we associate a localized traveling wave (wavefront) centered around the cortical location $\mathbf{x}' = (x', y')$:

$$g(\mathbf{x} - \mathbf{x}', t) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{|\mathbf{x} - \mathbf{x}'|^2}{2\sigma^2}\right) \cos(\mathbf{k}(\mathbf{x}') \cdot [(\mathbf{x} - \mathbf{x}') - \mathbf{v}(\mathbf{x}')t]) \quad (2)$$

where \mathbf{k} is a two-dimensional spatial frequency vector oriented as \mathbf{v} with a modulus $|\mathbf{k}|$ independent of \mathbf{x}' . The spatial extension of each wave is limited to a cortical neighborhood characterized by σ . It is worth noting that the cortical flow hypothesis starts from a speculation, partially supported by experimental evidence that when a subject is stimulated by motion stimuli, ordered patterns of dynamic activations arise within retinotopically organized visual areas. These activations should be intended as a macroscopic (i.e, averaged) activity that can be compared to spatiotemporal flow of excitation revealed by current source density analysis, local field potentials, optical imaging, or, at higher scales of observations by fMRI.

The resulting flow derives from the interference of all the local waves generated across a spatial domain S :

$$f(\mathbf{x}, t) = \int_S g(\mathbf{x} - \mathbf{x}', t) d\mathbf{x}' \quad (3)$$

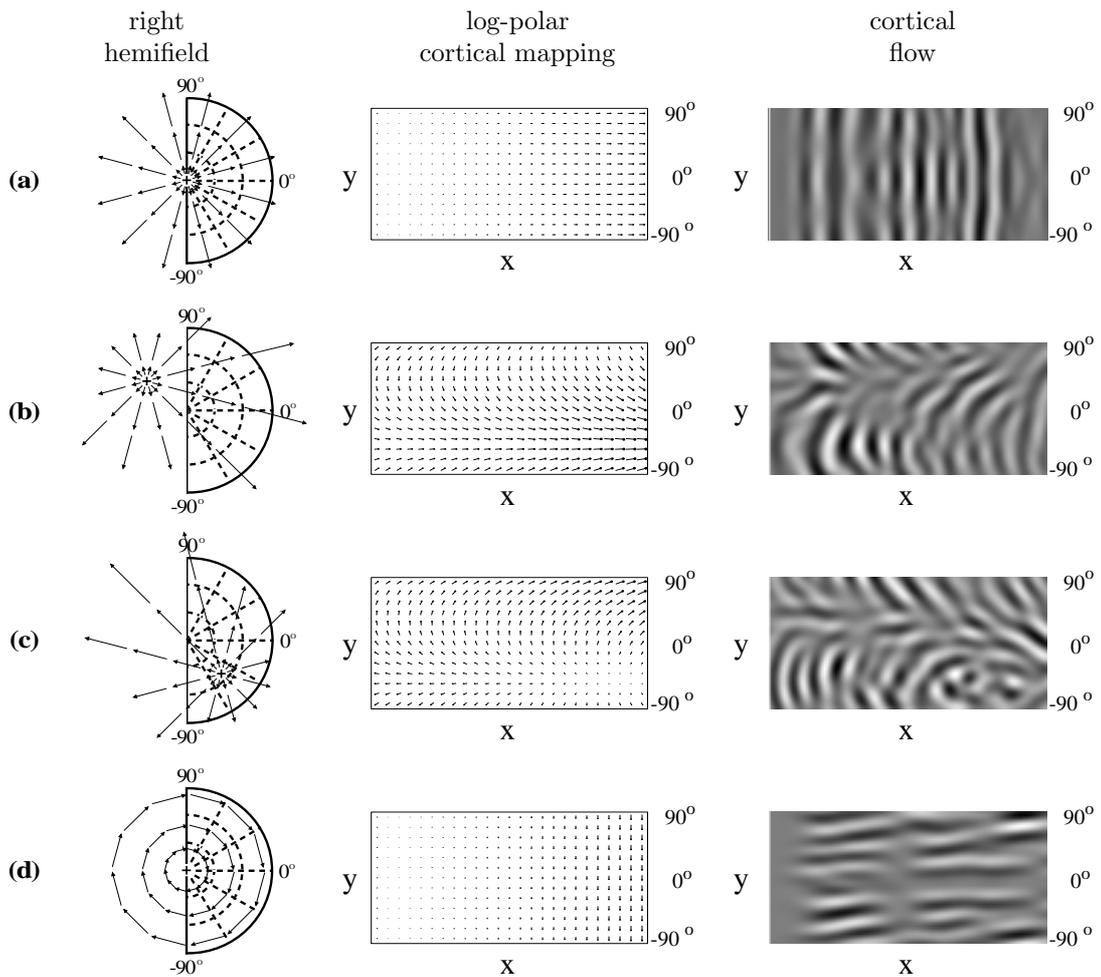


Figure 1: Examples of optic flow representations obtained by our model. *left column*: different optic flow stimuli for (a) a centered expansion, (b) an up-left expansion, (c) a down-right expansion, and (d) a centered rotation; *center column*: their corresponding log-polar mapping; *right column*: “snapshots” of the related cortical flow representations in area MT. The gray scale codes the level of activation on the cortical layer: light gray corresponds to an excitatory level, whereas dark gray corresponds to an inhibitory level.

Eq. (3) can be interpreted as a continuous spatiotemporal interpolation of the input vector field. The degree of smoothness introduced is represented by the geometrical parameters $(|\mathbf{k}|, \sigma)$ of single wavefronts. On an intuitive basis such a representation preserves qualitative properties of the optic flow. Indeed, ordered spatial configurations of the velocity field \mathbf{v} correspond to ordered spatiotemporal configurations of the cortical flow $f(\mathbf{x}, t)$, (see the right column of Fig. 1).

To assess qualitatively robustness and reliability of the cortical representation, we calculated the velocity field from the cortical flow and compared the resulting es-

timations with the original optic flow field. Following the systematic methodology proposed by [12], the evaluation of the optic flow was performed by a comparative use of different techniques proposed in the literature. Fig. 2 shows the original optic flow and its estimate calculated from the cortical flow in the case of an expansion field shifted from its center. In general, we observed that the errors are scattered and comparable in magnitude with the measured performances of the computational technique used to estimate the optic flow (cf. [12]), thus indicating that only a negligible error bias ($< 0.5\%$) is introduced by our representation.

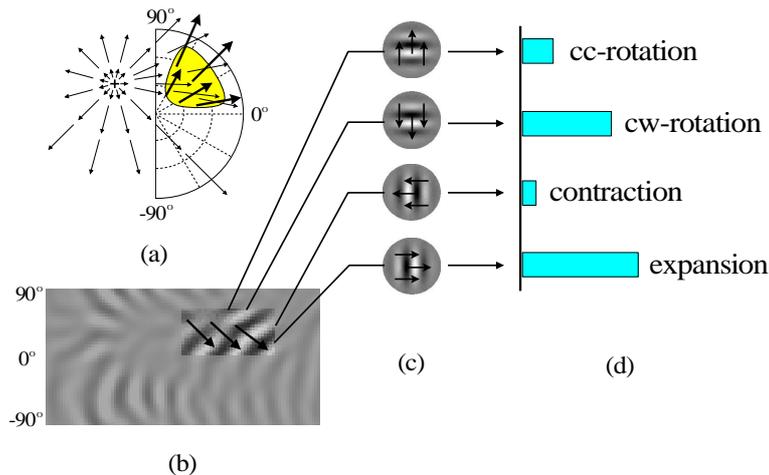


Figure 3: Operations performed by MST units on the cortical flow. (a) A pictorial view of the template matching between the MST cell and the optic stimulus. The shaded region indicates the cell's receptive field and the thick vectors indicate its preference for expansion. The cortical flow corresponding to the optic stimulus is shown in (b) where the region "viewed" by the receptive field is highlighted. The four types of MST filters associated to each EFCs and their average responses are shown in (c) and (d), respectively. Both expansion and clockwise rotation cells have significant responses to the off-centered expansion stimulus

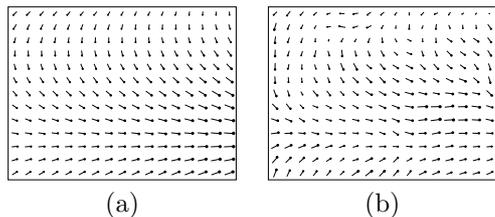


Figure 2: A comparison between the original optic flow (a) for an off-center expansion (cf. Fig. 1b) and its estimate (b) calculated from the cortical flow

3 Spatiotemporal filtering of the cortical flow

MST cells analyze complex motion by integrating local motion signals sent forward from area MT into a more unified representation. To this end, taking advantage of log-polar mapping geometry, we adopt a gaze-centered representation, where the optic flow field is analyzed in terms of how it is related to the fixation point. Each MST model cell is selective to an EFC (rotation or expansion/contraction) referred to the fixation point (i.e., the preferred center of motion of the cell is located on the fovea) and is characterized by a specific tuning to the mean speed of motion in optic

flow stimuli (cf. [13]). From a behavioural point of view, MST cells act as weight vector fields performing a template matching of cell preferences over a portion of the optic flow stimuli. Fig. 3a shows a pictorial representation of the *modus operandi* of an expansion-selective cell on an optic flow field with a shifted center of motion. The model is similar to other template models [14] [15], but it differs from them by the speed tuning and the position of the template, which is always centered on the fovea. From an architectural point of view, the geometry of the log-polar mapping and the cortical flow representation suggest an interpretation of the template vector matching in terms of spatiotemporal filtering operations with kernels $h(\mathbf{x}, t)$ tuned to up/down and left/right directions to detect horizontally- or vertically-directed translational cortical motion. Clockwise and counter-clockwise rotations correspond to downward and upward translations, whereas expansion and contraction result in rightward and leftward translations, respectively (see Fig. 3). The level of activity of an MST cell is given by:

$$e_{mst}(\mathbf{x}, t) = \int f(\mathbf{x} - \mathbf{x}', t - t') h(\mathbf{x}', t') d\mathbf{x}' dt' \quad (4)$$

The matching filters used in the model can be written as localized wavefronts damped

in time:

$$h(\mathbf{x}, t) = \frac{1}{2\pi\sigma_h^2} \exp\left(-\frac{|\mathbf{x}|^2}{2\sigma_h^2} - \frac{t}{\tau}\right) \cos[\mathbf{k}_h \cdot (\mathbf{x} - \mathbf{v}_h t)] \quad (5)$$

where τ is the decay constant, \mathbf{k}_h is the two-dimensional spatial frequency vector, \mathbf{v}_h is related to the speed sensitivity of the cell, and σ_h specifies the spatial extension of the integration area over MT representation which is ultimately related to the size of MST receptive fields. For expansion/contraction filters $\mathbf{k}_h = (k_{hx}, 0)$, $\mathbf{v}_h = (v_{hx}, 0)$ with $v_{hx} > 0$ for expansion and $v_{hx} < 0$ for contraction. For rotation filters $\mathbf{k}_h = (0, k_{hy})$, $\mathbf{v}_h = (0, v_{hy})$ with $v_{hy} > 0$ for counter-clockwise rotation and $v_{hy} < 0$ for clockwise rotation.

A quantitative analysis of the cortical flow would require banks of component-selective units with a range of sensitivities to speed and with multiple sizes of receptive field. However, a qualitative analysis of the optic flow and predictions of MST cell properties are still possible with a single set of component-selective units (see Fig. 4). In this figure, the output of cells selective to EFCs is shown for the five types of motion stimuli illustrated on the left. For the sake of representation, we arranged the elementary component cells in four homogeneous layers topologically organized. The gray level indicates the temporally averaged activity of each cell in the corresponding location. The distribution of activities shown in Fig. 4 qualifies the specificity of response of the cells to single and multiple component optic stimuli.

3.1 Implications to MST cell properties

The primary property of our cells is related to the fact that in the projection from MT to MST cells' sensitivity to motion patterns changes from a local to a global character. Differently from MT cells, our MST model cells are not selective to local translations specified by their optimal direction and uniform speed, but to spatial configurations of the velocity field with first-order differential properties with respect to the fovea. The variety of response properties reported by numerous investigators [2] [3] [4] [5] is often due to heterogeneous samples of cells from

many different anatomic layers of MST, suggesting that it could be the results of successive levels of processing within MST. In a hierarchical perspective, our MST cells can be considered as basic units whose responses can be combined to obtain more powerful descriptors characterized by larger field aggregate selectivity for circular, radial, or planar motion. Basically, the role of lower-level MST neurons in interpreting optic flow fields would be one of responding to visual motion according to the degree of match between the visual input and the preferred optic flow field of the neuron, whereas a qualitative analysis of the fundamental features of optic flow (e.g., the detection of differential invariants) would be the role of higher-level MST neurons.

Higher variations of the activity distribution observed in our model when the center of motion is moved off the fovea (see Fig. 4), with respect to variations associated to comparable shifts in the periphery are a direct consequence of our gaze-centered hypothesis. This also agrees with experiments concerning cells with preferred motion field centered on the fovea, whose tuning is narrower compared to that of cells with peripheral preference tolerating larger displacements [5]. For each model cell the best stimulus should be centered on the fovea (see Fig. 4), though, for a generic stimulus, MST cells respond to circular and radial components about the fixation point. Therefore, different stimulus configurations can trigger a cell response provided that the subpattern of the optic flow locally matches the cell preferences for motion component and speed. By example, if the stimulus speed is not adequate to the speed-tuning of the cell, a stronger response could be obtained by shifting the center of motion out of the fovea. These examples suggest that the numerous and sometimes contradictory physiological cell properties relating MST cell response to the placement of optic flow stimuli in the visual field, including position invariance, preferred location for the center of motion, direction reversals [4] [3] [7] [5], could be revisited from this new (gaze-centered) perspective.

Acknowledgments

This work was partially supported by the *MURST Project "Visual Perception Mi-*

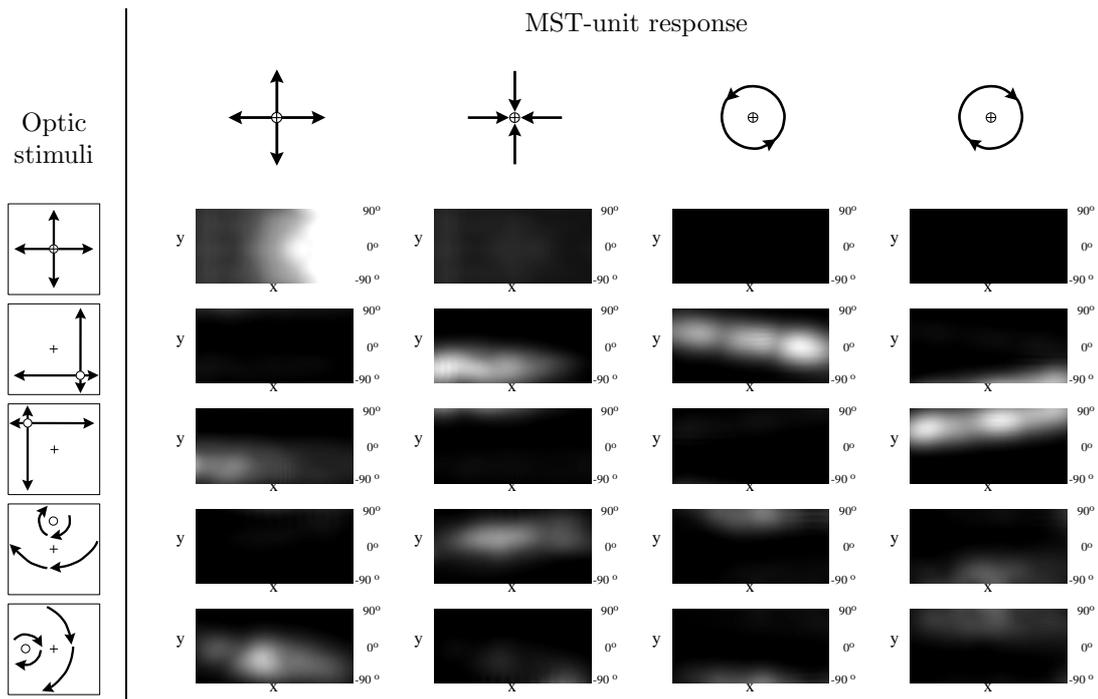


Figure 4: Typical responses of MST units. Four types of units, topologically organized in homogeneous layers are considered. Each type of unit is selective to an EFC (dilation, contraction, counter-clockwise rotation, clockwise rotation). The output of each layer is shown for five types of motion stimuli

crossystems based on Structural Computational Paradigms” and by the ASI Project “Optical Microsystems for Robotic Applications in Space”.

References

- [1] J.J. Gibson. *The perception of the visual world*. Riverside Press, Cambridge, 1950.
- [2] C. J. Duffy and R. H. Wurtz. *J. Neurophysiol.*, 65:1329–1345, 1991.
- [3] G.A. Orban, L. Lagae, A. Verri, S. Raiguel, D. Xiao, H. Maes, and V. Torre. *Proc. Natl. Acad. Sci.*, 89:2595–2599, 1992.
- [4] M.S.A. Graziano, R.A. Andersen, and R.J. Snowden. *J. Neuroscience*, 14:54–67, 1994.
- [5] C. J. Duffy and R. H. Wurtz. *J. Neuroscience*, 15:5192–5208, 1995.
- [6] J.J. Koenderink. *Vision Res.*, 26:161–180, 1986.
- [7] L. Lagae, H. Maes, S. Raiguel, D.K. Xiao, and G.A. Orban. *J. Neurophysiol.*, 71:1597–1626, 1994.
- [8] D.H. Hubel and T.N. Wiesel. *J. Comp. Neurol.*, 158:295–302, 1974.
- [9] T.D. Albright and R. Desimone. *Exp. Brain Res.*, 65:582–592, 1987.
- [10] E.L. Schwartz. *Biol. Cybern.*, 25:181–194, 1977.
- [11] W. Kohler. *Gestalt Psychology*. Liverright Press, 1947.
- [12] J.L. Barron, D.J. Fleet, and Beauchemin S.S. *Intern. J. Comput. Vis.*, 12:43–76, 1994.
- [13] C. J. Duffy and R. H. Wurtz. *J. Neuroscience*, 17:2839–2851, 1997.
- [14] J.A. Perrone and L.S. Stone. *J. Neuroscience*, 18:5958–5975, 1998.
- [15] S. Grossberg, E. Mingolla, and C. Pack. Technical Report CAS/CNS-TR-97-015, Dept. of Cognitive and Neural Systems, Boston, MA, 1998.