Biol Cybern (2008) 98:33–48
DOI 10.1007/s00422-007-0194-9

ORIGINAL PAPER

Biological Cybernetics

The symplectic structure of the primary visual cortex

Alessandro Sarti · Giovanna Citti · Jean Petitot

Received: 22 March 2007 / Accepted: 1 October 2007 / Published online: 16 November 2007

Abstract We propose to model the functional architecture of the primary visual cortex V1 as a principal fiber bundle where the two-dimensional retinal plane is the base manifold and the secondary variables of orientation and scale constitute the vertical fibers over each point as a rotation–dilation group. The total space is endowed with a natural symplectic structure neurally implemented by long range horizontal connections. The model shows what could be the deep structure for both boundary and figure completion and for morphological structures, such as the medial axis of a shape.

1 Introduction

There have recently been new applications of differential geometry to neurophysiology of visual perception. In particular, it has been shown in Petitot and Tondut (1999), Petitot (2003) and Citti and Sarti (2006) that the functional architecture of the area V1 of the primary visual cortex can be geometrically modeled as a three-dimensional structure $V = M \times P$, where the two-dimensional retinal plane M is the base manifold and the secondary variable of orientation constitutes the fiber P over each point.

Even if V1 is concretely a two-dimensional neural layer, it implements $n > 2$ degrees of freedom and is thus abstractly of dimension $n > 2$. This difference between the physical concrete and the geometrical abstract dimensions correspond to what Hubel called an "engrafting" of variables:

> "What the cortex does is map not just two but many variables on its 2-dimensional surface. It does so by selecting as the basic parameters the two variables that specify the visual field coordinates (distance out and up or down from the fovea), and on this map it engrafts other variables, such as orientation and eye preference, by finer subdivisions" (Hubel 1988, p. 131).

The total space V of the fiber bundle modeling V1 is endowed with a natural *contact* structure neurally implemented by long range horizontal connections. It is possible to hypothesize that this fundamental contact structure (making V1 an implementation of the space of *1-jets* of curves in M) is at the basis of the *contour integration* process.

Now it is well known that contact structures can be extended to *symplectic* structures by adding a new dimension corresponding to the possibility of changing *scale* at every point of M, that is treating scale as a gauge.

In Citti and Sarti (2006) the secondary variable of orientation constitutes the vertical fibers over each point as a rotation group. Thus, we will propose to add the scale dimension to the model Citti and Sarti (2006) and to model the functional geometry of V1 as a principal fiber bundle G where the 2-dimensional retinal plane M is again the base manifold and the secondary variables of orientation and scale constitute now the fibers over each point as a rotation–dilation group. In terms of Hubel's engrafted variables our model engrafts therefore two variables: orientation and scale. The total space will be endowed with a natural symplectic structure also neurally implemented by long range horizontal connections.

A. Sarti (✉)
Department of Electronics, Information and Systems,
University of Bologna, Bologna, Italy
e-mail: asarti@deis.unibo.it

G. Citti
Department of Mathematics,
University of Bologna, Bologna, Italy

J. Petitot
EHESS and CREA, Ecole Polytechnique, Paris, France

Moreover we will show how this symplectic structure can be naturally deduced from the filtering of the optical signal by the receptive fields (RFs) of visual neurons and their functional architecture.

Our model naturally integrates several features of the visual cortex observed in experiments of neurophysiology, psychophysics and neuroimaging.

We start from the well known observation that the set of simple cells receptive profiles (RPs) is generated by the action of the affine group of the visual plane on a Gabor mother function.

Each hypercolumn of simple cells RPs, defined at a retinal point (x, y), forms a 2-dimensional subgroup of rotations–dilations. This structure is identically repeated for every point of the retina, and can therefore be considered as a fiber of the 4-dimensional principal bundle.

Then we show that the neural process of *maximal selectivity* due to intracortical short range inhibitory connections internal to hypercolumns selects in each case a section of the fibration. This mechanism will be described as the lifting of boundaries in M to curves in V1 and whole figures in M to surfaces in V1.

From the neurophysiological point of view, there is evidence of connections between simple cells belonging to different hypercolumns. These neural connections are what neurophysiologists call long range horizontal connections. They implement what geometers call a *connection* operating a *parallel transport* between fibers of a bundle.[1] This latter connection (in the geometric sense) can be expressed in terms of a symplectic form on the 4-dimensional principal bundle G, introducing on it what is called a *complex structure*.

It was already shown in Petitot and Tondut (1999) and Citti and Sarti (2006) that the integral curves of the natural contact structure in the 3-dimensional fiber bundle V is a mathematical representation of the *association field* of Field et al. (1993). In our model here, the integral curves of the symplectic structure model more finely the connectivity pattern between simple cells in V1, as observed by electrophysiological experiments. In other words, the symplectic structure introduces a system of natural coordinates in the 4-dimensional space $\mathbb{C}^2$, which is implemented by neural connectivity.[2] Notice that many of the anatomical and functional features we will discuss and model are common to most mammals, nevertheless we will refer particularly to features experimentally found in tree shrew.

We also refer to the following results, for a theoretical description of the functional architecture of the striate cortex (Bressloff et al. 2001; Bressloff and Cowan 2003; Swindale 2004, 2000; Ben 2003; Ben and Zucker 2004).

In this paper we model the cortex as a cotangent fiber bundle that can be interpreted as a phase space from the physical point of view. This is the natural space for stochastic dynamical systems. The idealized model we presented can be thought as the deterministic counterpart of these stochastic models.

The paper is organized as follows:

- In Sect. 2 we consider the basic properties of simple cells RPs and outline the group structure underlying the set of cells. We show also how 1-forms can be naturally associated to RPs.
- In Sect. 3 we deepen the group structure and introduce the hypercolumnar structure and the related geometrical concept of a principal fiber bundle.
- In Sect. 4 we prove that maximal selectivity in the hypercolumns explains and justifies the association between RPs and 1-forms.
- In Sect. 5 the "horizontal" connectivity between hypercolumns is considered and the notion of symplectic structure on the fiber bundle is introduced. We define also the associated complex structure.
- In Sect. 6 we present numerical simulations of integral curves of the symplectic structure and compare them with neurophysiological experimental results about long range horizontal connections. We show also that the lifting of images yields Lagrangian manifolds, which explains the well known but enigmatic perceptive relevance of the medial axis of a shape.
- For reader convenience we collected in Appendix A the proof of the main theorem, and in Appendix B the basic definitions of differential geometry instruments used in the paper.

2 The receptive profiles of simple cells

In this section we define the RP of neural cells, make explicit the effects on them of a change of frame in the visual plane and show how to associate to a RP a differential 1-form selecting its preferred orientation.

2.1 Receptive fields and receptive profiles

A great variety of cells respond to a stimulus on the retinal plane $M \subset \mathbb{R}^2$. The RF of a visual neuron is classically defined as the domain of the retina to which it is connected through the neural connections of the retino-geniculo-cortical pathways (projecting from the retina to the cortex through the lateral geniculate nucleus along the thalamic way) and whose stimulation elicits a spike response.

[1] It must be emphasized that the technical lexicons of neurophysiology and geometry use the same terms "fibers", "connections", etc. in completely different ways. In general the context will eliminate any ambiguity.

[2] $\mathbb{C}^2$ is 2-dimensional over $\mathbb{C}$ but 4-dimensional over $\mathbb{R}$.

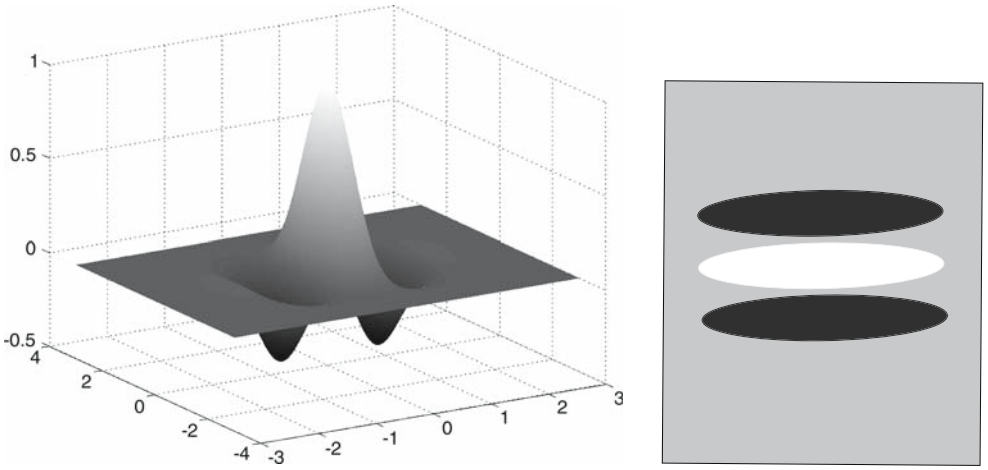

Fig. 1 A Gabor filter (*left*) and a schematic representation of its sign (*right*). Positive sign is in *white* and negative in *black*

In the sequel, we will restrict our models to the RF in this narrow sense. Classically, a RF is decomposed into ON (positive contrast) and OFF (negative contrast) zones according to the type of response to light and dark luminance Dirac stimulations. There exists, therefore, a RP of the visual neuron, which is simply its impulse response as a filter kernel. It is a function $\Psi(x, y)$ (where x, y are retinal coordinates) $\Psi : M \to \mathbb{R}$ which is defined on the retinal plane M and measures the response (ON / OFF) of the neuron to stimulations at the point (x, y). Sophisticated techniques enable the recording of the level curves of the RPs (see, e.g., De Angelis et al. 1995). A light and dark spot or bar is switched on and off at different positions of the RF and the mean response is measured. One uses for instance random sequences of flashes (of 50 ms) distributed over a lattice of 20×20 positions, with 100 ms − 1 s for each response after each flash, and takes the mean value on 10 flashes at each position (white noise analysis). The correlation of the inputs (flashes) with the outputs (spikes) yields the transfer function of the neuron.

It is a classical result of neurophysiology, already strongly emphasized by Marr (1982) in the late 1970s, that the RPs of the retinal ganglion cells are like Laplacians of Gaussians. On the contrary, the simple cells of V1 are strongly oriented, and are often interpreted as Gabor patches (trigonometric functions modulated by a Gaussian).

2.2 The set of receptive profiles

The visual plane M will be identified with the plane $\mathbb{R}^2$ endowed with coordinates (x, y) and a fixed frame (∂_x, ∂_y). This is of course a wild idealization and presupposes the choice of an arbitrary global origin $O = (0, 0)$.

When a visual stimulus I of intensity $I(x, y) : M \subset \mathbb{R}^2 \to \mathbb{R}^+$ activates the retinal layer of photoreceptors (identified with the visual field) $M \subset \mathbb{R}^2$, the cells centered at every point (x, y) of M process in parallel the retinal stimulus with their RP $\Psi_{(x,y)}$, which is defined on M.

Each RP acting on a point (x, y) depends upon a preferred direction θ and a scale σ, and it has been observed experimentally that *the set of simple cells RPs is obtained via translations* to the point (x, y), *rotations of angle* θ, *and dilations of scale* $s = \mathrm{e}^{\sigma}$ *from a unique profile, of Gabor type*. This means that there does exist a "mother" profile $\Psi_0(\xi, \eta)$ from which all the observed profiles can be deduced by transformation.

A good formula for Ψ_0 seems to be

$$\Psi_0(\xi, \eta) = \mathrm{e}^{-(\xi^2+\eta^2)}\mathrm{e}^{2i\eta}.$$

In Fig. 1 we represent its even part $\mathrm{e}^{-(\xi^2+\eta^2)}\cos(2\eta)$, which will be studied in this paper.

It has to be noted that the RPs $\Psi_{(x,y)}$ are localized in a neighborhood of the point (x, y). Hence, they will be expressed in local coordinates, around the point (x, y), obtained by rotation and dilation of the initial ones. Any vector will then have two representations: coordinates $(\tilde{\xi}, \tilde{\eta})$ in the global frame centered at $(0, 0)$, and (ξ, η) in the local frame centered at (x, y) :

$$\begin{cases} \xi = \mathrm{e}^{-\sigma}\Big((\tilde{\xi} - x)\cos\theta + (\tilde{\eta} - y)\sin\theta\Big), \\ \eta = \mathrm{e}^{-\sigma}\Big(-(\tilde{\xi} - x)\sin\theta + (\tilde{\eta} - y)\cos\theta\Big). \end{cases} \tag{1}$$

Therefore all the observed profiles can be modeled as

$$\Psi_{x,y,\theta,\sigma}(\tilde{\xi}, \tilde{\eta}) = \mathrm{e}^{-2\sigma}\Psi_0(\xi, \eta).$$

In this model we consider a uniform distribution of orientations and scales, even if there is a neurophysiological evidence of covariation of scale and orientation (Issa et al. 2000).

We recall that a change of frames can act on the space M or, by duality, on functions defined on M. Precisely, if A is a

transformation of M, it transforms also functions ψ through the formula

$$A\psi(\tilde{\xi},\tilde{\eta}) = \frac{1}{\det(A)}\psi(A^{-1}(\tilde{\xi},\tilde{\eta})).$$

Accordingly we will denote $A^{-1}_{x,y,\theta,\sigma}$ the transformation defined in (1), and obtain

$$\Psi_{x,y,\theta,\sigma}(\tilde{\xi},\tilde{\eta}) = (A_{x,y,\theta,\sigma}\Psi_0)(\tilde{\xi},\tilde{\eta}).$$

The expression of A thus becomes

$$A_{x,y,\theta,\sigma}(\xi,\eta) = \begin{pmatrix} x \\ y \end{pmatrix} + \mathrm{e}^{\sigma}\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}\begin{pmatrix} \xi \\ \eta \end{pmatrix}.$$

Notice that for simplicity we identify translations in the retinal plane and translations in the cortical layers, neglecting the conformal log-polar retino-cortical mapping. Anyway, the mapping can easily be taken into account by introducing a suitable metric on cortical layers.

2.3 The output of receptive profiles

The overall output O of the parallel filtering is given by the integral of the signal $I(x, y)$ times the bank of filters:

$$\begin{aligned} O_{(\theta,\sigma)}(x,y) &= \int_M I(\tilde{\xi},\tilde{\eta})\Psi_{(x,y,\theta,\sigma)}(\tilde{\xi},\tilde{\eta})\mathrm{d}\tilde{\xi}\mathrm{d}\tilde{\eta} \\ &= \int_M I(\tilde{\xi},\tilde{\eta})\Psi_0\left(A^{-1}_{x,y,\theta,\sigma}(\tilde{\xi},\tilde{\eta})\right)\frac{\mathrm{d}\tilde{\xi}\mathrm{d}\tilde{\eta}}{\mathrm{e}^{2\sigma}}. \end{aligned} \quad (2)$$

Consequently we can perform a change of variable with respect to the transformation A. We get:

$$O_{(\theta,\sigma)}(x,y) = \int \Psi_0(\xi,\eta)I\left(A_{(x,y,\theta,\sigma)}(\xi,\eta)\right)\mathrm{d}\xi\mathrm{d}\eta.$$

We explicitly note that the linear transformation A acts on the retinal plane M, argument of I. On the contrary the inverse transformation A^{-1} acts on the domain of Ψ_0. This allows us to interpret the domain of Ψ_0 as dual plane, with respect to the retinal plane.

3 Functional geometry of hypercolumns

3.1 The hypercolumnar structure

The hypercolumnar structure organizes the cells of V1 in hypercolumns covering a small chart of the visual field (their RF) and corresponding to parameters such as orientation, scale, ocular dominance, direction of movement or color for a fixed retinal position (x, y). The hypercolumnar organization means, therefore, essentially that to each position (x, y) of the retina M is associated a full exemplar $P_{(x,y)}$ of the space of such "secondary" variables. We restrict here ourselves to orientation θ and scale σ. As we have seen in the introduction, even if the cortical space is a bidimensional layer it "engrafts" secondary variables. Petitot and Tondut (1999) and Citti and Sarti (2006) proposed a first approximated model with orientation. In Bosking et al. (1997) the orientation θ is described as a single engrafted secondary variable, in terms of the principal *fiber bundle* of the roto-translation group.[3] Here we will use the same geometrical instruments to propose a new model, which takes into account not only orientation θ but also scale σ, and leads, therefore, to a 4-dimensional fiber bundle.

3.2 The fiber $\mathrm{RP}_{(x,y)}$ over each retinal point (x, y)

To every fixed point (x, y) of the retinal plane M is associated the complete set of values for orientation θ and scale σ. In other words, for (x, y) fixed, we consider all filters obtained via rotations of an angle θ and dilations of scale e^{σ} from a fixed one $\Psi_{(x,y,0,0)}$. so that the set of filters over the point (x, y) becomes:

$$\begin{aligned} \mathrm{RP}_{(x,y)} &= \left\{\Psi_{(x,y,\theta,\sigma)} : (x, y) \text{ fixed}, (\theta, \sigma) \text{ variable}\right\} \\ &= \left\{A_{(\theta,\sigma)}\Psi_{(x,y,0,0)} : (\theta,\sigma) \in S^1 \times \mathbb{R}\right\}, \end{aligned}$$

where we have denoted $A_{(\theta,\sigma)} = A_{(0,0,\theta,\sigma)}$.

3.3 The group of orientation and scale operating in a hypercolumn

The set of filters at the point (x, y) is generated as image of

$$G_2 = \left\{A_{(\theta,\sigma)} : (\theta,\sigma) \in S^1 \times \mathbb{R}\right\}.$$

The set G_2 is a trivial commutative group (topologically a cylinder), since, if we apply in sequence two of its elements, we get a new element of it:

$$A_{(\theta_1,\sigma_1)}A_{(\theta_0,\sigma_0)}(\xi,\eta) = A_{(\theta_0+\theta_1,\sigma_0+\sigma_1)}(\xi,\eta).$$

Formally, we have a map

$$\begin{aligned} &\left(A_{(\theta_1,\sigma_1)}, \Psi_{(x,y,\theta,\sigma)}\right) \in G_2 \times \mathrm{RP}_{(x,y)} \\ &\quad\mapsto \Psi_{(x,y,\theta+\theta_1,\sigma+\sigma_1)} \in \mathrm{RP}_{(x,y)} \end{aligned}$$

which formally expresses that the group G_2 acts on the set of all filters RP respecting the fibers $\mathrm{RP}_{(x,y)}$. Each filter is obtained from the fixed one via one of these transformations: $\mathrm{RP}_{(x,y)}$ is the orbit $G_2\left(\Psi_{(x,y,0,0)}\right)$.

[3] The fiber bundle is a mathematical structure, locally described as a cartesian product $M \times G$, where M models the retinal plane, and G the group of engrafted variables. In the model this expresses the fact that that at every retinal point $(x, y) \in M$ is associated a whole hypercolumn. (We refer to Definition 4 in Appendix B for the precise definition of a fiber bundle).

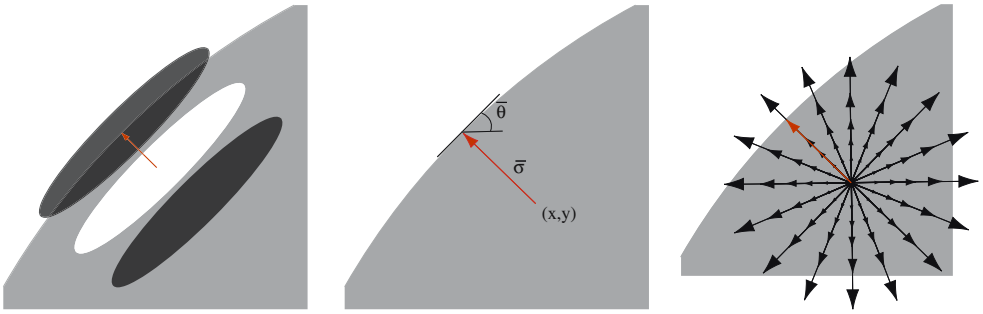

Fig. 2 The simple cell centered in (x, y) takes the maximal activity $O(x, y, \bar{\theta}, \bar{\sigma}) = \max_{(\theta,\sigma)} O(x, y, \theta, \sigma)$ at a point $\bar{\sigma}$ which is, up to a constant, the logarithm of the distance to the nearest boundary and $\bar{\theta}$ is the direction of this boundary

The change of variable associated to the transformation A acts on the RFs $\Psi_{(x,y,\theta,\sigma)}$, as well as on the fixed frame ∂_x, ∂_y, which will be rotated to the vectors

$$X_1 = e^{\sigma}\left(\cos(\theta)\partial_x + \sin(\theta)\partial_y\right),$$
$$X_3 = e^{\sigma}\left(-\sin(\theta)\partial_x + \cos(\theta)\partial_y\right).$$

4 Maximal selectivity and geometrical interpretation of lifting

We come now to one of our central points. Up to now we have described the hypercolumnar structure in terms of differential geometry. Now we introduce the functionality of hypercolumns and particularly the orientation and scale *selectivity*.

In the past years several models have been presented to explain the emergence of orientation and scale selectivity in the primary visual cortex. These models use different combinations of feedforward (thalamic) and feedback (intracortical) signals and consider different involvements of excitatory and inhibitory short range connections (Miller et al. 2001; Carandini and Ringach 1997; Bar et al. 1995; Shelley et al. 2000). Even if the basic mechanism producing strong orientation selectivity is controversial; nevertheless it is evident that the intracortical circuitry is able to filter out all the spurious directions and to strictly keep the direction of maximum response of the simple cells.

4.1 The maximization procedure

Neurophysiologically, orientation and scale selectivity is the action of intracortical short range connections to select the maximum response from the outputs:

$$O_{(\theta,\sigma)}(x, y) = \int_M I(\tilde{\xi}, \tilde{\eta})\Psi_{(x,y,\theta,\sigma)}(\tilde{\xi}, \tilde{\eta})\mathrm{d}\tilde{\xi}\mathrm{d}\tilde{\eta}$$
$$= \int_M \Psi_0(\xi_1, \eta_1) I(A_{(x,y,\theta,\sigma)}(\xi_1, \eta_1))\mathrm{d}\xi_1\mathrm{d}\eta_1, \tag{3}$$

where the last integral is obtained with a change of variables.

This maximal selectivity is the simplest mechanism to accomplish the selection from among all different cells responses to effect a lift in the cotangent space. Given an input I, the neural processing associates to each point (x, y) of the retina M a point $(x, y, \bar{\theta}, \bar{\sigma})$ in the cortex, we interpret this mechanism as a lifting into the fiber of the parameter space $\mathbb{R}^2(x, y) \times S^1(\theta) \times \mathbb{R}(\sigma)$ over (x, y). Precisely, the odd part of the filters $\Psi_{(x,y,\theta,\sigma)}$ lifts the boundaries of the image and the even part of the filters lifts the interior of the objects.[4] The odd part has already been studied in Petitot and Tondut (1999) and Citti and Sarti (2006). Hence we will focus here on the even part of the filter.

We will denote $(\bar{\theta}, \bar{\sigma})$ the values of the maximal response:

$$O(x, y, \bar{\theta}, \bar{\sigma}) = \max_{(\theta,\sigma)} O(x, y, \theta, \sigma).$$

This maximality condition can be mathematically expressed requiring that the gradient of O with respect to the variables (θ, σ) vanishes at the point $(x, y, \bar{\theta}, \bar{\sigma})$:

$$\nabla^{(\theta,\sigma)} O(x, y, \bar{\theta}, \bar{\sigma}) = 0.$$

We will also require that at the maximum point the Hessian is strictly negatively definite:

$$\mathrm{Hess}(O) < 0.$$

For simplicity we will now *reduce to contours*, that is to domains E in M with regular boundaries ∂E, the image I being a cartoon image, formally expressed as a piecewise constant function, which takes only the values 0 or 1. As we will see with the following theorem, a geometrical interpretation of the values of $\bar{\sigma}$ and $\bar{\theta}$ selected by the maximization procedure just described is that they detect respectively *the distance and the direction of the nearest boundary* (see Fig. 2).

Theorem 1 *At first-order approximation, for any fixed value of (x, y), the output function $O(x, y, \theta, \sigma)$ reaches a local maximum at the point $\bar{\theta}, \bar{\sigma}$, where $d(\bar{\sigma}) = \frac{1}{\sqrt{2}} e^{\bar{\sigma}}$ denotes the*

[4] Recall that $\Psi_0(\xi, \eta) = e^{-(\xi^2+\eta^2)}e^{2i\eta} = e^{-(\xi^2+\eta^2)}(\cos(2\eta) + i\sin(2\eta))$. The odd part is the sin part and the even part the cos part.

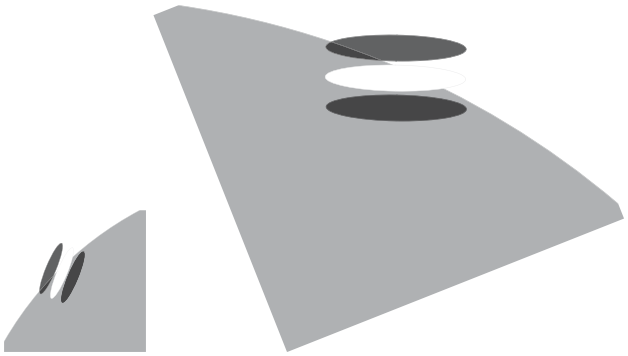

Fig. 3 The output of a simple cell is equal to the output of the mother filter on an appropriately translated, rotated, and dilated image. This simple cell can, therefore, be associated with the cotangent vector (the 1-form) deduced from $\mathrm{d}y$ by the same transformation

distance of (x, y) *from the nearest boundary of the image* I*, and* $\bar{\theta}$ *denotes the orientation of this boundary at the point where the distance is achieved.*

The proof is given in the Appendix A.

This theorem ensures that the simple cell selects the couple $(\bar{\theta}, \bar{\sigma})$ in such a way that, X_3 being the vector

$$X_3 = \mathrm{e}^{\bar{\sigma}} \left(-\sin(\bar{\theta})\partial_x + \cos(\bar{\theta})\partial_y\right), \tag{4}$$

then $(x, y) + \frac{1}{\sqrt{2}} X_3$ belongs to the nearest boundary of I.

4.2 Simple cells and associated 1-forms

Let us provide a new interpretation of the action of RPs as filters. A simple cell $\Psi_{\theta,\sigma}(\xi, \eta)$ gives the maximal response for contours with direction θ. We can, therefore, consider that it *selects* such a direction. But a traditional way to select orientations at points of M is to consider *differential 1-forms* on M. (We refer to Appendix B for the definition of 1-form).

One form is a linear function, defined on the tangent space, with real values. The form $\mathrm{d}y$, dual of ∂_y, selects the y-component of a tangent vector. Analogously the form obtained by $\mathrm{d}y$, rotating in the inverse direction with respect to A, selects the coefficient of X_3 on $T_{(x,y)}M$, i.e., on the tangent plane of M at the point (x, y):

$$\left(\mathrm{e}^{-\sigma}(-\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y)\right)(\xi X_{1|(x,y)} + \eta X_{3|(x,y)}) = \eta.$$

We will, therefore, associate to any filter $\Psi_{\xi,\eta,\theta,\sigma}$ *which is able to select the direction* X_3*, the 1-form*

$$\mathrm{e}^{-\sigma}(-\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y)$$

which selects the same direction (see Figs. 3, 4).

Due to this correspondence between simple cells and 1-forms selecting their preferential orientation, for every point (x, y) the fiber $\mathrm{RP}_{(x,y)}$ of the set of filters $\Psi_{(x,y,\theta,\sigma)}$ is made out of elements of the form:

$$\left\{\omega_{(\theta,\sigma)} = \mathrm{e}^{-\sigma}(-\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y) : \theta \in [0, \pi], \sigma \in \mathbb{R}\right\}. \tag{5}$$

Using this particular representation, we can say that the coefficients (θ, σ) provide a representation of a co-vector in *log-polar coordinates*, and the elements of this type completely fill the 2-dimensional cotangent space $T^*_{(x,y)}M$ at (x, y). Hence the fiber over the point (x, y) (the hypercolumn

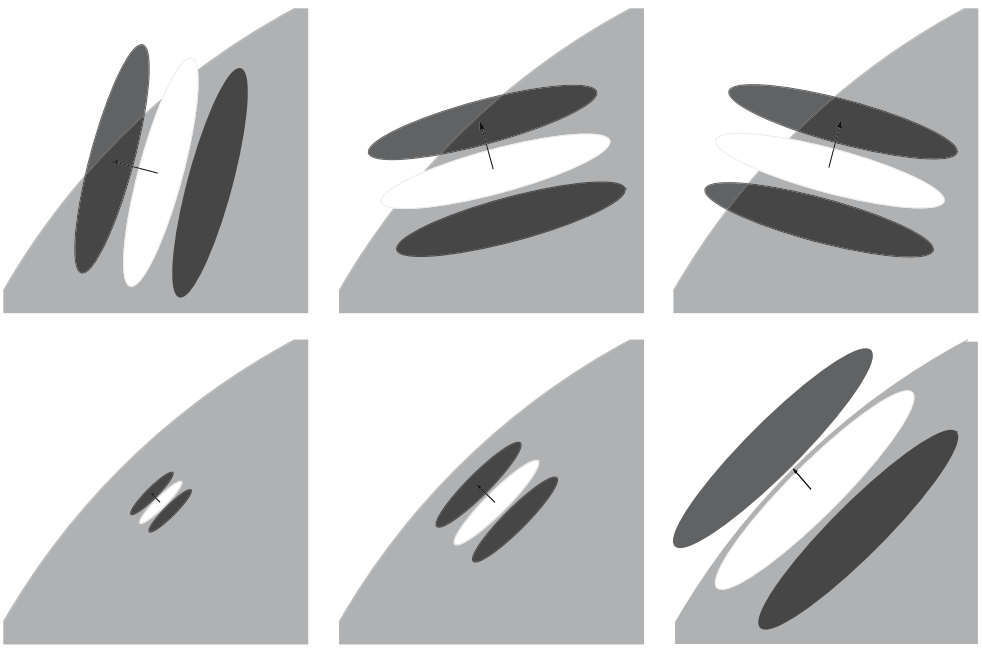

Fig. 4 Simple cells act on the image as 1-forms (represented in the figure by *black arrows*), i.e., they select the tangent vector of optimal orientation and scale, where optimality is given by selecting the maximally spiking cell over the fiber

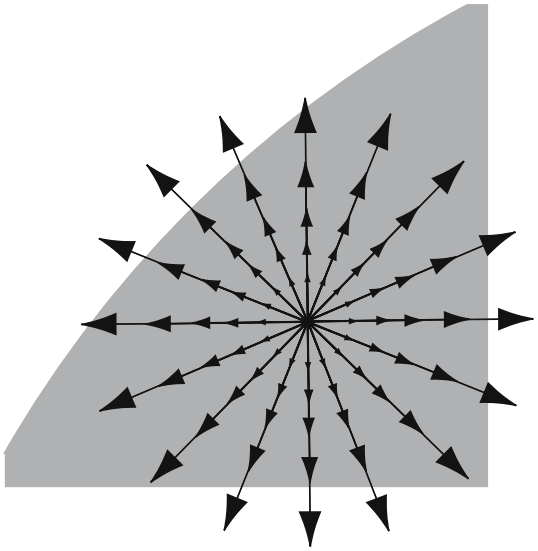

Fig. 5 A hypercolumn centered in (x, y) acts on the image as a fiber of 1-forms

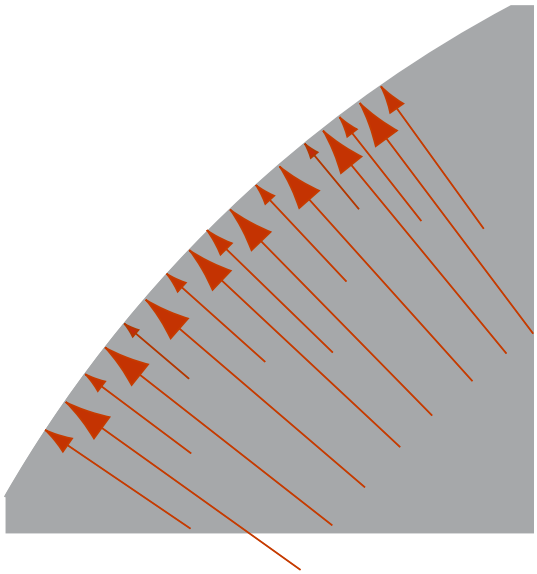

Fig. 6 The set of maximal responding cells selects a field of 1-forms which lifts the image into V1

RP$_{(x,y)}$) can be identified with the cotangent space $T^*_{(x,y)}M$ at this point (see Fig. 5).

4.3 A field of 1-forms

We see that the selection of $\bar{\theta}$ and $\bar{\sigma}$ by maximal response of the internal local neural circuitry of the hypercolumn RP$_{(x,y)}$ defines a function

$$\begin{aligned} \bar{\Sigma} : M &\to RP \\ (x, y) &\mapsto (x, y, \bar{\theta}, \bar{\sigma}) \end{aligned}$$

Identifying RP with T^*M, $\bar{\Sigma}$ becomes a field of 1-forms (see Fig. 6). We can interpret this function as the lift the image I in M in the fiber bundle $\pi : RP \to M$.

5 Long range neural connections and symplectic structure

In the previous section we have analyzed the selection of an orientation $\bar{\theta}$ and a scale $\bar{\sigma}$ induced in each hypercolumn by its internal circuitry. We will now take into account the relationships between different hypercolumns. We know that these relations are neurophysiologically implemented in long range excitatory intracortical intercolumnar "horizontal" connections. We have seen on the other hand what are the geometrical transformations which relate the fibers RP$_{(x,y)}$ in the fiber bundle RP.[5] In this section we will focus on this double-neurophysiological and geometrical-functional architecture and emphasize the action of the group G of roto-translations and scaling.

[5] See Appendix B for the definition of a fiber bundle and a fiber.

5.1 The cotangent bundle associated to the set of receptive fields

We will now make more precise the geometrical structure of the set of all hypercolumns that is implemented in the functional architecture of V1. For every fixed point (x, y), the set RP$_{(x,y)}$ of RPs based at (x, y) has been endowed with a structure of a homogeneous space under the action of a 2-dimensional group G_2.

The set of all hypercolumns is the 4-dimensional fiber bundle

$$\mathrm{RP} = \bigcup_{(x,y)} \mathrm{RP}_{(x,y)}.$$

If we consider the hypercolumns as the cotangent space at each point (x, y) of M, we associate to the bundle RP the cotangent space T^*M with its system of log-polar coordinates:

$$T^*M = \bigcup_{(x,y)} T^*_{(x,y)}M = \bigcup_{(x,y)} \Big\{(x, y, \theta, \sigma) : \mathrm{e}^{-\sigma}(-\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y) \in T^*_{(x,y)}M\Big\}.$$

It is clear by definition that any element in the cotangent bundle is a 4-dimensional vector (x, y, θ, σ) where $(x, y) \in M$, while $(\theta, \sigma) \in T^*_{x,y}(M)$ are the coordinates in the cotangent plane at the point (x, y).

The group law G_2 defined on RP$_{(x,y)}$ together with the group law defined on the base space M define a group law on the whole bundle.

This group G is

$$G \simeq \left\{ A_{(x,y,\theta,\sigma)} : (x, y, \theta, \sigma) \in \mathbb{R}^2 \times S^1 \times \mathbb{R} \right\}.$$

For reader convenience we compute here the group law. Calling

$$r_\theta = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix},$$

and applying in sequence two transformations we get:

$$\begin{aligned} &A_{(x,y,\theta,\sigma)} A_{(x_1,y_1,\theta_1,\sigma_1)}(\xi, \eta) \\ &\quad = A_{(x,y,\theta,\sigma)} \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \mathrm{e}^{\sigma_1} r_{\theta_1} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right) \\ &\quad = \begin{pmatrix} x \\ y \end{pmatrix} + \mathrm{e}^{\sigma} r_\theta \left(\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \mathrm{e}^{\sigma_1} r_{\theta_1} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \right) \\ &\quad = \begin{pmatrix} x \\ y \end{pmatrix} + \mathrm{e}^{\sigma} r_\theta \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + \mathrm{e}^{\sigma+\sigma_1} r_{\theta+\theta_1} \begin{pmatrix} \xi \\ \eta \end{pmatrix} \\ &\quad = A_{(x_2,y_2,\theta_2,\sigma_2)}(\xi, \eta) \end{aligned}$$

with

$$\begin{aligned} &\sigma_2 = \sigma + \sigma_1, \quad \theta_2 = \theta + \theta_1, \\ &\begin{pmatrix} x_2 \\ y_2 \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix} + \mathrm{e}^{\sigma} r_\theta \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}. \end{aligned} \tag{6}$$

We emphasize again that the composition is *not* commutative.

5.2 Horizontal long range connections and the tangent space to the cotangent plane

The "vertical" structure of hypercolumns lifts retinal points to cortical ones but it is largely non-sufficient to implement a *global coherence* , for which the visual system must be able to *compare* two retinotopically neighboring fibers $\mathrm{RP}_{(x_a,y_a)}$ and $\mathrm{RP}_{(x_b,y_b)}$ over two neighboring points a and b of M. This problem has been solved at the empirical level by the discovery of "horizontal" cortico–cortical connections (see, e.g., Das and Gilbert 1995). Horizontal connections are long ranged (up to 6–8 mm) and connect cells of approximately *the same orientation*. In distant hypercolumns this allows curved contents of connection fields that have been modeled quantitatively by Ben (2003). To detect them (see, e.g., Ts'o et al. 1986), one can (i) measure the correlations between cells belonging to different hypercolumns; (ii) compare the preferred orientation of a reference cell with the preferred orientation of other cells; (iii) compute cross-correlograms. One verify that cells with neighboring orientations are strongly correlated while cells with sufficiently different orientations are decorrelated.

The key experimental fact is that, while short range connections inside a hypercolumn are isotropic, long range horizontal connections are on the contrary highly *anisotropic* and restricted to cells sharing essentially the same orientation. These two different types of connections implement two different levels of structure: (i) the short range connections implement the local triviality of the fibration $\pi : \mathrm{RP} \to M$, while (ii) the long range connections implement a richer structure.

Long-range horizontal cortico–cortical connections insure the large scale coherence of retinotopy. Without them, neighboring hypercolumns would become functionally independent and retinotopy would lose any immanent reality for the system itself. That cortico–cortical connections connect neurons of essentially the same orientation in different hypercolumns means that the system is able to know, for (x', y') different from (x, y), if an orientation θ at (x, y) is approximatively the same as an orientation θ' at (x', y').

The direction of the motion between points in the cotangent bundle T^*M belongs to the tangent space to the cotangent space. Since (x, y, θ, σ) are coordinates for T^*M, then $(\partial_x, \partial_y, \partial_\theta, \partial_\sigma)$ is a tangent frame at the origin. On the other side the group G acts on T^*M, and the differential of the left translation L_g of G by g (that is the map $h \mapsto gh$ of G onto itself) allows one to transport the standard basis of T_0G to a *left invariant* basis at any other point of the space. A direct computation shows that this left invariant basis is precisely:

$$\begin{cases} X_1 = \mathrm{e}^{\sigma} \left(\cos(\theta)\partial_x + \sin(\theta)\partial_y \right), \\ X_2 = \partial_\theta, \\ X_3 = \mathrm{e}^{\sigma} \left(-\sin(\theta)\partial_x + \cos(\theta)\partial_y \right), \\ X_4 = \partial_\sigma. \end{cases}$$

The fact that the standard basis $(\partial_x, \partial_y, \partial_\theta, \partial_\sigma)$ is *not* left invariant manifests the crucial phenomenon of *nonholonomy*.

5.3 The symplectic structure and the scale as a gauge field

We will now consider another classical structure defined on the cotangent bundle T^*M, namely its *symplectic structure*.[6] Indeed for every σ we have selected as fundamental 1-form the 1-form defined in Citti and Sarti (2006) and Petitot (2003):

$$\omega = \mathrm{e}^{-\sigma} \left(-\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y \right).$$

The subspace $V = \mathbb{R}^2(x, y) \times S^1(\theta) \times \{\sigma\}$, associated to the 1-form for a scale σ, is a *contact structure*. We take then as symplectic form on T^*M the 2-form $\mathrm{d}\omega$ obtained by differentiating ω with respect to all its variables. The symplectic 2-form $\mathrm{d}\omega$ is written as

$$\begin{aligned} \mathrm{d}\omega &= \left(\mathrm{e}^{-\sigma} \cos(\theta)\mathrm{d}x + \mathrm{e}^{-\sigma} \sin(\theta)\mathrm{d}y \right) \wedge \mathrm{d}\theta \\ &\quad \left(-\mathrm{e}^{-\sigma} \sin(\theta)\mathrm{d}x + \mathrm{e}^{-\sigma} \cos(\theta)\mathrm{d}y \right) \wedge \mathrm{d}\sigma \\ &= \omega_1 \wedge \omega_2 + \omega_3 \wedge \omega_4, \end{aligned}$$

[6] See Appendix B for the definition.

where ω_i is the dual form of X_i. This form can be identified with the *left invariant* 2-form deduced by left translations from the standard 2-form on G at 0:

$$\mathrm{d}x \wedge \mathrm{d}\theta + \mathrm{d}y \wedge \mathrm{d}\sigma.$$

5.4 The complex structure

It is well known that every symplectic structure induces a *complex structure*. Let B be the matrix associated to the symplectic form $\mathrm{d}\omega$ defined by $\mathrm{d}\omega(X, X') = \langle BX, X' \rangle$[7] for every pair (X, X') belonging to the tangent plane $T_{(x,y,\theta,\sigma)}G$:

$$\begin{cases} X = \xi\partial_x + \eta\partial_y + \vartheta\partial_\theta + \varsigma\partial_\sigma, \\ X' = \xi'\partial_x + \eta'\partial_y + \vartheta'\partial_\theta + \varsigma'\partial_\sigma. \end{cases}$$

We have:

$$B = \mathrm{e}^{-\sigma} \begin{pmatrix} 0 & 0 & -\cos(\theta) & \sin(\theta) \\ 0 & 0 & -\sin(\theta) & -\cos(\theta) \\ \cos(\theta) & \sin(\theta) & 0 & 0 \\ -\sin(\theta) & \cos(\theta) & 0 & 0 \end{pmatrix}. \quad (7)$$

By definition of a 2-form we have $\mathrm{d}\omega(X, X') = -\mathrm{d}\omega(X', X)$, which implies that

$$B^* = -B$$

(where B^* is the transpose of B) since

$$\begin{aligned} \mathrm{d}\omega(X, X') &= \langle BX, X' \rangle = \langle X, B^*X' \rangle = \langle B^*X', X \rangle \\ &= -\mathrm{d}\omega(X', X) = -\langle BX', X \rangle. \end{aligned}$$

From this it is clear that $-B^2 = BB^*$ is nonnegative definite. In our case

$$-B^2 = \mathrm{e}^{-2\sigma} I,$$

where I is the identity map. Hence, if we write

$$P = \sqrt{-B^2} = \mathrm{e}^{-\sigma} I,$$

and put

$$J = BP^{-1} = \mathrm{e}^{\sigma} B,$$

it is clear that $J^2 = -I$, which means that J defines a complex structure. In our case

$$J = \begin{pmatrix} 0 & 0 & -\cos(\theta) & \sin(\theta) \\ 0 & 0 & -\sin(\theta) & -\cos(\theta) \\ \cos(\theta) & \sin(\theta) & 0 & 0 \\ -\sin(\theta) & \cos(\theta) & 0 & 0 \end{pmatrix}. \quad (8)$$

Through J, all the tangent spaces of $G = \mathbb{R}^2(x, y) \times S^1(\theta) \times \mathbb{R}(\sigma)$ can be identified with $\mathbb{C}^2$.

The complex/symplectic structure naturally associates the vector fields X_1 and $JX_1 = X_2$, and X_3 and $JX_3 = X_4$. The planes $\{X_1, X_2\}$ and $\{X_3, X_4\}$ spanned by X_1, X_2 and X_3, X_4 at a point (x, y, θ, σ) are complex lines in $\left(T_{(x,y,\theta,\sigma)}G, J\right) \simeq \mathbb{C}^2$.

[7] $\langle A, B \rangle$ is the scalar product on the tangent planes.

6 Filtering, symplectic geometry and shape analysis

6.1 Integral curves of special vector fields

Viewed in V with $\sigma = \sigma_0$, $\{X_1, X_2\}$ is the *contact plane* $C_{(x,y,\theta)}V$. The integral curves of the contact structure C starting at a fixed point $(x_0, y_0, \theta_0, \sigma_0)$ are tangent to the contact plane at every point. They are called *contact curves* and write

$$\begin{aligned} &\gamma'(t) = X_1(\gamma(t)) + k(t)X_2(\gamma(t)), \\ &\gamma(0) = (x_0, y_0, \theta_0, \sigma_0), \end{aligned} \quad (9)$$

k being a parameter varying in $\mathbb{R}$. In particular for k constant, the solution is

$$\begin{cases} x = \frac{1}{k}\left(\sin(kt + \theta_0) - \sin(\theta_0) + kx_0\right), \\ y = \frac{1}{k}\left(-\cos(kt + \theta_0) + \cos(\theta_0) + ky_0\right), \\ \theta = kt + \theta_0, \\ \sigma = \sigma_0. \end{cases}$$

For $k = 0$ its projection on the (x, y) plane is the x axis and for $k \neq 0$ it is a circle of radius $1/k$ tangent to the x axis (see Fig. 7 left). The emergence of curvature in this context is strictly related to the curve detection model based on curvature in Parent and Zucker (1989).

If we take into account the scale σ, x and y are scaled by e^{σ} while θ remains the same. Analogously, we can consider the integral curves of the vector fields $X_3 + kX_4$,

$$\begin{aligned} &\gamma'(t) = X_3(\gamma(t)) + k(t)X_4(\gamma(t)), \\ &\gamma(0) = (x_0, y_0, \theta_0, \sigma_0). \end{aligned} \quad (10)$$

The solution is

$$\begin{cases} x = -\frac{\sin(\theta_0)}{k}\mathrm{e}^{\sigma_0}\left(\mathrm{e}^{kt} - 1\right) + x_0, \\ y = \frac{\cos(\theta_0)}{k}\mathrm{e}^{\sigma_0}\left(\mathrm{e}^{kt} - 1\right) + y_0, \\ \theta = \theta_0, \\ \sigma = kt + \sigma_0. \end{cases}$$

Its projection on the (x, y) plane is *independent* of k and orthogonal to the direction θ_0. For k variable the integral curve is the line of slope k in the fixed "vertical" plane $\{X_3, X_4\}$ (see Fig. 7 right).

The projection of these two classes of integral curves on the base plane (x, y) is plotted in Fig. 8. Their pattern is in good agreement with the pattern of long range connections found both in neurophysiological and psychophysical experiments. In fact, excitatory connections are confined to two regions, one flaring out along the axis of orientation of the

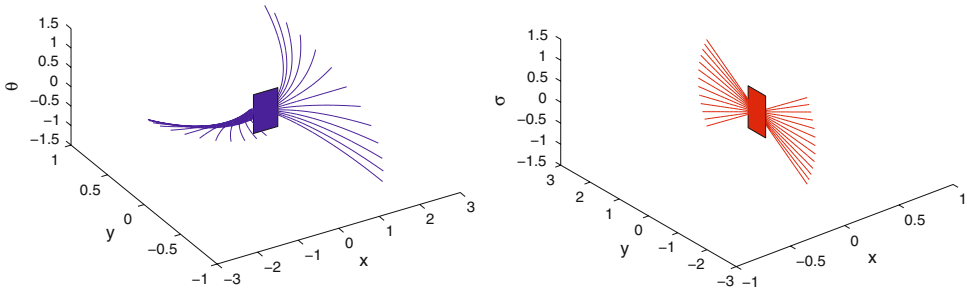

Fig. 7 Integral curves respectively of the (X_1, X_2) and (X_3, X_4) vector fields

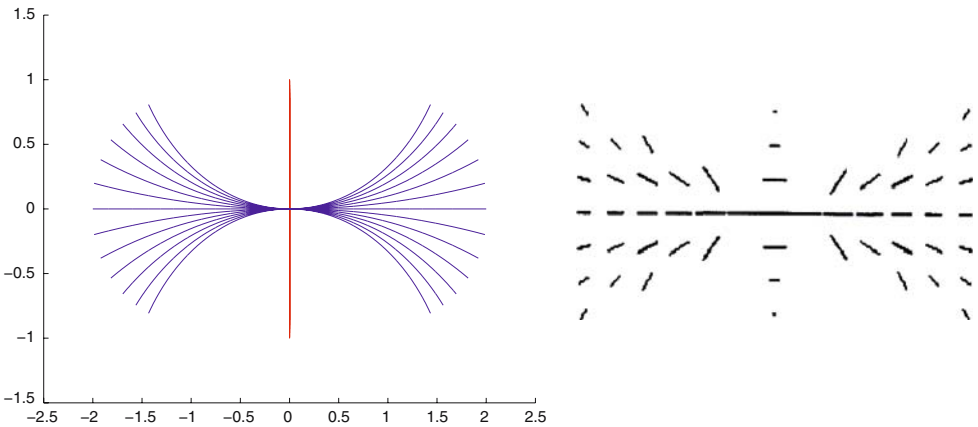

Fig. 8 The projection of the integral curves of the symplectic structure on the retinal plane (*left*) reveals the pattern of co-axial and trans-axial connections found by neurophysiological experiments (*right*, from Yen and Finkel 1998)

cell (co-axial, in blu), and another confined to a narrow zone extending orthogonally to the axis of orientation (trans-axial, in red). The co-axial connections are similar to the "association field" proposed by Field et al. (1993) and are represented here as integral curves of the fields X_1, X_2 (see also Citti and Sarti 2006). The symplectic model adds *a second set of trans-axial excitatory connections* which extends orthogonally from the orientation axis of the cell. There is anatomical evidence consistent with the existence of this orthogonal connections (Fitzpatrick 1996; Lund et al. 1985; Mitchinson and Crick 1982; Rockland and Lund 1982, 1993). The trans-axial connections are represented here as integral curves of the fields X_3, X_4. The reason why co-axial connections spread out in a fan, while trans-axial connections are more spatially focused is that the roto-translation fields X_1, X_2 not only do not commute, but their commutator is linearly independent from them:

$$[X_1, X_2] = -X_3,$$

while the vectors X_3, X_4 do not commute, their commutator is linearly dependent upon them:

$$[X_3, X_4] = -X_3.$$

Then integral curves of roto-translation fields are not planar (Fig. 8, left) while integral curves of X_3, X_4 belong to a plane (Fig. 8, right) whose projection on the (x, y) plane is just a line (red line in Fig. 8, right).

6.2 Liftings and Lagrangian manifolds

Let us now return to the lifted sets described in Sect. 3 and see how we can interpret them relatively to the symplectic structure we have just defined.

We associated to each point (x, y) of the retinal plane, values $\bar{\theta}$ and $\bar{\sigma}$ which we have interpreted as the direction of the nearest boundary and (a function of) the distance from the center of the filter to this boundary. Let us consider the example of Fig. 9, that of a white ellipse E on a black boundary. The boundary ∂E is lifted as in the contact structure and gives the boundary of the down left ellipse. The gray levels code the values of θ periodically from 0 (black) to π (white). But due to the scale factor σ, we can lift not only the boundary ∂E but *every point* (x, y) of the base. We generate that way a *surface* in G: $\Sigma = \{(x, y, \bar{\theta}, \bar{\sigma})$. As we have shown in Sect. 4, Σ is a part of the extremal surface where the gradient of the output $O(x, y, \bar{\theta}, \bar{\sigma})$ vanishes:

$$\bar{\Sigma} = \big\{(x, y, \bar{\theta}, \bar{\sigma}) : \partial_\theta O(x, y, \bar{\theta}, \bar{\sigma}) = 0, \\ \partial_\sigma O(x, y, \bar{\theta}, \bar{\sigma}) = 0, \mathrm{Hess}(O) < 0\big\},$$

where Hess denote the Hessian matrix of the function O.

The condition on the Hessian ensures that $\bar{\Sigma}$ is a regular manifold. This locally defines $\bar{\theta}$ and $\bar{\sigma}$ as two functions $\bar{\theta}(x, y)$ and $\bar{\sigma}(x, y)$ on the base plane.

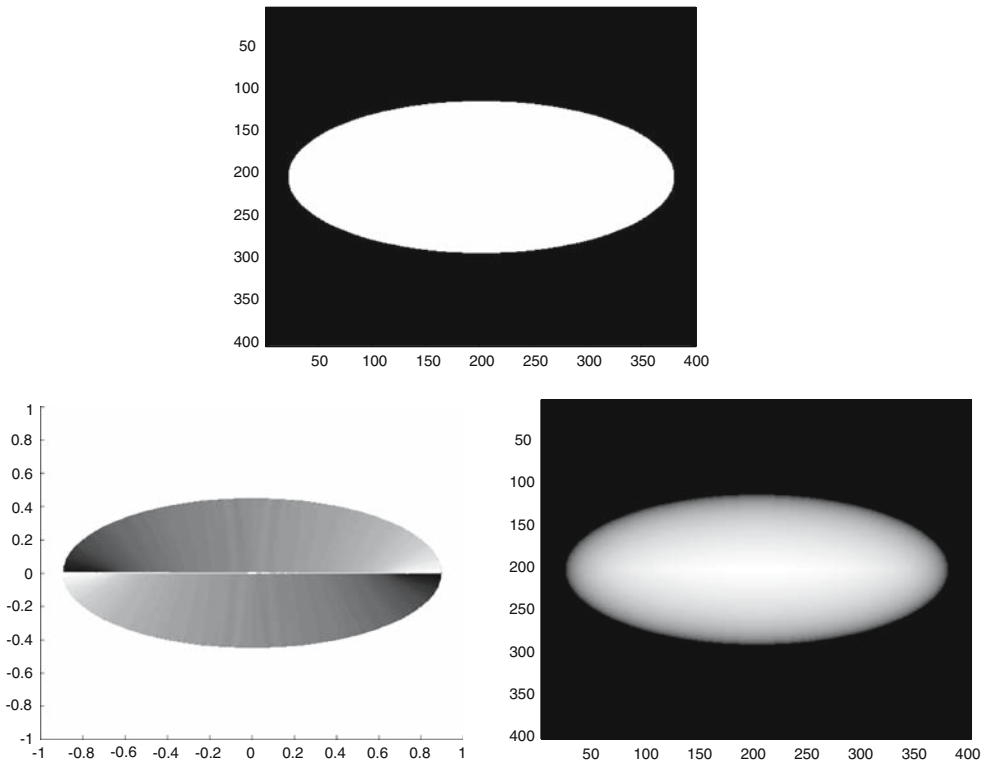

Fig. 9 *Up center*: an example of lifting of an ellipse. *Down*: the corresponding fields of selected orientation $\bar{\theta}$ (*left*) and scale $\bar{\sigma}$ (*right*). The exterior structure has been cancelled for graphical reasons

From the properties of the two functions $\bar{\theta}, \bar{\sigma}$, we can deduce interesting properties of Σ relative to the symplectic structure, at least under the assumptions of Theorem 1 of Sect. 4. Indeed, according to this theorem $\bar{\theta}(x, y)$ is the direction of the nearest boundary ∂E and $\bar{\sigma}(x, y)$ (up to a factor) the distance of (x, y) to this boundary. This implies that the projections of the level curves of $\bar{\theta}(x, y)$ are orthogonal to the boundary ∂E and dually that the level curves of $\bar{\sigma}(x, y)$ are parallel to ∂E (see Fig. 10). In other words, *the filtering of the* indicatrix function of the set E *by the family of filters* $\Psi_{(x,y,\theta,\sigma)}$ *realizes the well-known propagation of the boundary by the eikonal equation of geometrical optics*: the boundary propagates as parallel wave fronts and its points generate rays orthogonal to the fronts. We can call it a Huyghens model. The singularities of the propagation generate the so-called *"cut locus"* or *symmetry axis* of the shape. The symmetry axis of a shape, which is so fundamental for its morphological analysis, segmentation, and recognition as was shown since Harry Blum's pioneering works (Blum 1973) by René Thom, David Marr, David Mumford, Steve Zucker, James Damon and many others (see Petitot and Zucker 1989; Kimia 2003), *has therefore a neurophysiological relevance in our model*. It is not surprising since symplectic structures constitute the framework of Hamiltonian mechanics and the Huyghens model

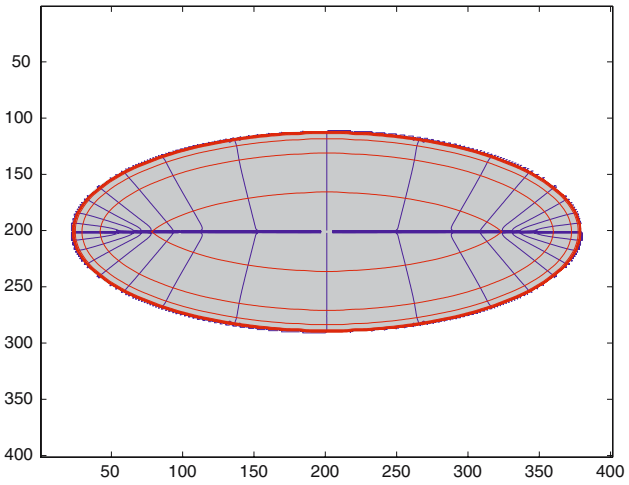

Fig. 10 Level curves of $\bar{\theta}(x, y)$ (*blue*) and $\bar{\sigma}(x, y)$ (*red*)

(propagation of rays and wave fronts) is universal in Hamiltonian systems. As far as a symplectic structure is neurally implemented in V1 it is natural to observe Hamiltonian structures such as symmetry axes.

Consider now a point (x, y) in the base space M. $\bar{\theta}(x, y)$ and $\bar{\sigma}(x, y)$ are as in Fig. 10:

$$X_1 = e^{\bar{\sigma}} \left(\cos(\bar{\theta})\partial_x + \sin(\bar{\theta})\partial_y\right)$$

is parallel to the orientation of the boundary ∂E at the point of minimal distance $\bar{\sigma}$ and therefore

$$X_1(\bar{\sigma}) = 0.$$

Analogously,

$$X_3 = \mathrm{e}^{\bar{\sigma}}\left(-\sin(\bar{\theta})\partial_x + \cos(\bar{\theta})\partial_y\right)$$

is orthogonal to the orientation of this nearest boundary and it is constant along this direction $\bar{\theta}$. Therefore,

$$X_3\left(\bar{\theta}\right) = 0.$$

From these relations we can deduce that

Theorem 2 *The tangent plane to the surface* Σ *belongs to the kernel of the symplectic form* $\mathrm{d}\omega$ *at every point* $(x, y, \bar{\theta}, \bar{\sigma})$.

Proof We compute first the tangent plane to the lifted surface Σ at $(x, y, \bar{\theta}, \bar{\sigma})$. In the basis $\{X_1, \ldots, X_4\}$ it is spanned by the two vectors

$$\begin{cases} X_1 + X_1\bar{\theta}X_2 + X_1\bar{\sigma}X_4, \\ X_3 + X_3\bar{\theta}X_2 + X_3\bar{\sigma}X_4. \end{cases}$$

On the other side the symplectic 2-form $\mathrm{d}\omega$ can be represented as

$$\mathrm{d}\omega = \omega_1 \wedge \omega_2 + \omega_3 \wedge \omega_4$$

with $\{\omega_1, \ldots, \omega_4\}$ the dual basis of $\{X_1, \ldots, X_4\}$. Hence, applying the two form to the tangent vectors we get

$$(\omega_1 \wedge \omega_2 + \omega_3 \wedge \omega_4)\left(X_1 + X_1\bar{\theta}X_2 + X_1\bar{\sigma}X_4, X_3 + X_3\bar{\theta}X_2 + X_3\bar{\sigma}X_4\right) = X_3\bar{\theta} - X_1\bar{\sigma}.$$

But as $X_1\bar{\sigma} = X_3\bar{\theta} = 0$, we see that the tangent vector fields annihilate the symplectic form □

Definition 1 Let (G, Ω) be a symplectic 4-dimensional manifold. A (smooth) surface Σ of G is called *Lagrangian* if $\Omega\,|_{T\Sigma} \equiv 0$.

Corollary 1 *The lifted surface* Σ *is* Lagrangian.

6.3 Boundaries and objects: 1- and 2-forms

If we fix the value of the scale σ, we obtain a simplified model, only dependent on orientation, but independent of the scale. The lifted surface simply reduces to a curve Γ lifting the boundary ∂E in the roto-translation group and called the *Legendrian lift* of ∂E, and we find again the situation studied in Citti and Sarti (2006). In this paper it was proved that the tangents to the lifted curve Γ all lie in the so-called contact planes which turn out to be the kernels of the 1-form

$$\omega = -\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y.$$

In previous paper (Petitot and Tondut 1999; Petitot 2003) it has been proposed a linearized version of the same form.

The analogous result for surfaces is contained in the previous theorem, since the lifted surface is Lagrangian with respect to the 2-form $\mathrm{d}\omega$.

7 Conclusion

We have shown how to filter a contour using a family of filters derived from a mother filter whose RP is that of a simple cell of V1. This family is an homogeneous space under the action of the group $G(x, y, \theta, \sigma)$ of translations, rotations, and scaling of the retinal plane M. The horizontal cortico–cortical connections implement the natural G-invariant symplectic structure of this space.

We have then shown that the maximal response of the filters to an input contour ∂E computes at every point (x, y) of M the nearest distance $d(\bar{\sigma})$ of (x, y) to ∂E and the orientation $\bar{\theta}$ of ∂E at its nearest point to (x, y). The lifted surface $\Sigma = \{(x, y, \bar{\theta}, \bar{\sigma})\}$ is a Lagrangian surface and realizes the Huyghens model (rays and wave fronts propagation) of the contour ∂E, propagation whose singularities constitute the cut locus (the symmetry axis) of the shape E.

Appendix A: Proof of theorem 1

We give a simple sketch of the proof using the fact that I is the indicatrix function of a set E with regular boundary ∂E, i.e., is 1 inside the set E and -1 outside. Hence the gradient ∇I vanishes out of the boundary ∂E.

It is not restrictive to consider a fixed cell $\Psi_{(0,\sigma)}(\xi, \eta)$. As $\Psi_{(0,\sigma)}(\xi, \eta)$ is considered to be defined on the tangent plane $T_O M$, we compute the result up to the first-order approximation. Then

$$\begin{aligned}\Psi_{(0,\sigma)}(\xi, \eta) &= \frac{1}{\mathrm{e}^{2\sigma}}\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}}\cos(2\eta/\mathrm{e}^{\sigma}) \\ &\simeq \frac{1}{\mathrm{e}^{2\sigma}}\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}}\left(1 - 2\frac{\eta^2}{\mathrm{e}^{2\sigma}}\right).\end{aligned}$$

But it is trivial to verify that

$$\frac{1}{\mathrm{e}^{2\sigma}}\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}}\left(1 - 2\frac{\eta^2}{\mathrm{e}^{2\sigma}}\right) = \partial_\eta\left(\frac{1}{\mathrm{e}^{2\sigma}}\eta\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}}\right)$$

and therefore, up to first-order approximation,

$$\Psi_{(0,\sigma)}(\xi, \eta) \simeq \partial_\eta\left(\frac{\eta}{\mathrm{e}^{2\sigma}}\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}}\right). \tag{11}$$

In our case the tangent vector $X_3 = \mathrm{e}^{\sigma}\left(-\sin(\theta)\partial_x + \cos(\theta)\partial_y\right)$ reduces to $X_3 = \mathrm{e}^{\sigma}\partial_\eta$ at O. Hence, if we denote by $K_{(0,\sigma)}(\xi, \eta)$ the function on $T_O M$

$$K_{(0,\sigma)}(\xi, \eta) = \frac{\eta}{\mathrm{e}^{3\sigma}}\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}},$$

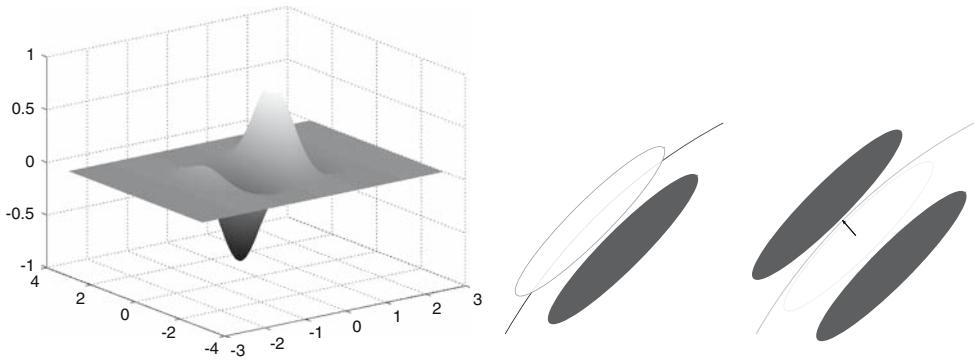

Fig. 11 The shape of the function $K_{(0,\sigma)}$. Note that, since the derivative of $K_{(x,y,\theta,\sigma)}$ is the even filter $\Psi_{(x,y,\theta,\sigma)}$, then $K_{(x,y,\theta,\sigma)}$ is an odd filter. *Left*: it reaches its maximum at a distance $d(\sigma) = \mathrm{e}^{\sigma}/\sqrt{2}$ from the center. *Center*: the choice of $K_{(x,y,\theta,\sigma)}$ which maximizes the output O. Recall that the function K has been chosen in such a way that its derivative is the filter $\Psi_{(x,y,\theta,\sigma)}$. *Right*: the corresponding shape of $\Psi_{(x,y,\theta,\sigma)}$, which vanishes where $K_{(x,y,\theta,\sigma)}$ is maximum

we get (Fig. 11)

$$\begin{aligned}\Psi_{(0,\sigma)}(\xi,\eta) &= \mathrm{e}^{\sigma}\partial_{\eta}\left(\frac{\eta}{\mathrm{e}^{3\sigma}}\mathrm{e}^{-(\xi^2+\eta^2)/\mathrm{e}^{2\sigma}}\right)\\ &= X_3\left(K_{(0,\sigma)}(\xi,\eta)\right).\end{aligned}$$

Hence the output $O(0,0,0,\sigma)$ is given by

$$O(0,0,0,\sigma) = \int_M X_3\left(K_{(0,\sigma)}(\xi,\eta)\right) I(\xi,\eta)\mathrm{d}\xi\,\mathrm{d}\eta$$

and, integrating by parts, we get

$$O(0,0,0,\sigma) = -\int_M K_{(0,\sigma)}(\xi,\eta) X_3\left(I(\xi,\eta)\right)\mathrm{d}\xi\,\mathrm{d}\eta$$

since $K_{(0,\sigma)}(\xi,\eta)I(\xi,\eta)$ vanishes at infinity.

Since I is the indicatrix function of E, its gradient ∇I vanishes inside and outside E, and can be represented, as a Dirac mass concentrated on ∂E, with the direction of the outer normal ν of E: $\nabla I = \nu\delta_{|\partial E}$. Then, if $\langle\bullet \mid \bullet\rangle$ is the scalar product, we have $X_3(I)\,\mathrm{d}\xi\,\mathrm{d}\eta = \langle X_3 \mid \nabla I\rangle = \langle X_3 \mid \nu\rangle\,\delta_{|\partial E} =$ (since X_3 and ω are dual one of each other) $= \langle\omega,\nu\rangle = \omega(\nu)$. We get finally that

$$O(0,0,0,\sigma) = -\int_{\partial E} K_{(0,\sigma)}(\xi,\eta)\omega(\nu)\delta_{|\partial E}.$$

Since ω is the dual form of X_3, ω vanishes for the direction X_1, reaches it maximum $\mathrm{e}^{-\sigma}$ for the orthogonal direction X_3 and its minimum $-\mathrm{e}^{-\sigma}$ for the opposite direction $-X_3$. The value $\omega(\nu)$ is, therefore, extremal when the boundary ∂E is of orientation $\theta = 0$.

As for $K_{(0,\sigma)}(\xi,\eta)$, we have

$$K'_{(0,\sigma)}(0,\eta) = \left(-2\eta^2\mathrm{e}^{-2\sigma}+1\right)\mathrm{e}^{-3\sigma-\eta^2/\mathrm{e}^{2\sigma}}$$

and $K'_{(0,\sigma)}(0,\eta)$ vanishes for

$$\eta = \frac{1}{\sqrt{2}}\mathrm{e}^{\sigma}.$$

Up to first-order approximation, we can suppose that the regular boundary ∂E is an horizontal line $\eta = \mathrm{cst}$. Then

$$\begin{aligned}O(0,0,0,\sigma) &= -\int_{\partial E} K_{(0,\sigma)}(\xi,\eta)\omega(\nu)\delta_{|\partial E}\\ &= -\int_{\mathbb{R}} K_{(0,\sigma)}(\xi,\eta)\mathrm{e}^{-\sigma}\mathrm{d}\xi\\ &= -\sqrt{\pi}\frac{\eta}{\mathrm{e}^{3\sigma}}\mathrm{e}^{-\eta^2/\mathrm{e}^{2\sigma}}.\end{aligned}$$

It is proportional to $K_{(0,\sigma)}(0,\eta) = \frac{\eta}{\mathrm{e}^{3\sigma}}\mathrm{e}^{-\eta^2/\mathrm{e}^{2\sigma}}$ and reaches it maximum

$$\sqrt{\pi}\frac{1}{\sqrt{2}}\mathrm{e}^{\sigma}\frac{1}{\mathrm{e}^{3\sigma}}\mathrm{e}^{-\left(\frac{1}{\sqrt{2}}\mathrm{e}^{\sigma}\right)^2/\mathrm{e}^{2\sigma}} = \sqrt{\frac{\pi}{2}}\frac{1}{\mathrm{e}^{(2\sigma+\frac{1}{2})}}$$

when $K_{(0,\sigma)}(0,\eta)$ reaches it minimum.

The integral $O(0,0,0,\sigma)$ will therefore reach its maximum

$$\sqrt{\frac{\pi}{2}}\frac{1}{\mathrm{e}^{(2\sigma+\frac{1}{2})}}$$

when $K_{(0,\sigma)}$ attains its maximum on ∂E, and $-X_3$ and ν coincide. The orientation $\theta = 0$ is the orientation of the boundary, while $K_{(0,\sigma)}$ is maximum at the point (ξ,η) which have distance $d(\sigma) = \frac{1}{\sqrt{2}}\mathrm{e}^{\sigma}$ to the boundary ∂E. This argument is local, but in order to see that the maximum is reached at the nearest boundary, we fix the direction of Ψ and we vary the scale. Since Ψ has null mean, the integral defining O tends to 0 when the scale tends to 0 and the filter Ψ tends to concentrate on the support of the image I. Hence increasing the value of the scale, the integral increases until the first zero of Ψ cross the boundary of I. At this point the integral start

decreasing again. Hence the maximum is reached when θ is the orientation of the nearest boundary. □

Appendix B: Differential geometry instruments

Let us briefly recall the definition of some mathematical concepts used in the paper.

By simplicity we give all definitions in $\mathbb{R}^2$. The definitions in vector spaces of higher dimension is analogous. If the local parametric equation of a curve γ is $(x(t), y(t))$ then its tangent vector at a point $(x(t), y(t))$ will be

$$\mathbf{X} = (\alpha_1(t), \alpha_2(t)) = (x'(t), y'(t)).$$

If we identify the curve γ with the trajectory of a point, the tangent vector $\mathbf{X}$ describes the direction of the motion at every point.

Definition 2 Tangent space. The set of tangent vectors is denoted $T_{(x,y)}M$, and it is a plane, called tangent plane at M at the point (x, y).

At every tangent vector $\mathbf{X} = (\alpha_1, \alpha_2)$, we can associate the directional derivative along the vector $\mathbf{X}$, and we denote it with a similar symbol (where ∂_x means $\frac{\partial}{\partial x}$):

$$X = \alpha_1 \partial_x + \alpha_2 \partial_y.$$

With this identification the basis of the tangent space $T_{(x,y)}M$ at (x, y) is (∂_x, ∂_y).

Definition 3 Cotangent space. Linear functions defined on the tangent space at a point (x, y) are called 1-forms, or cotangent vectors. The set of 1-forms at a point (x, y) is denoted $T^*_{(x,y)}$ and it is a vector space of dimension 2. Its basis is denoted $(\mathrm{d}x, \mathrm{d}y)$. This means that a general 1-form can be expressed as

$$\omega = \omega_1 \mathrm{d}x + \omega_2 \mathrm{d}y.$$

By definition, a form is a function defined on the tangent space, and, being linear the action is formally analogous to a scalar product:

$$\omega(X) = \omega_1 \alpha_1 + \omega_2 \alpha_2.$$

However this operation is called duality, instead of scalar product, since it acts between different spaces: ω_1, ω_2, are coefficients of a 1-form and α_1, α_2 coefficients of a tangent vector.

Example 1 The simplest example of a 1-form is the differential of a regular function f, which can be represented as

$$\mathrm{d}f = \frac{\partial f}{\partial x}\mathrm{d}x + \frac{\partial f}{\partial y}\mathrm{d}y.$$

Hence it is expressed in terms of the standard basis $\mathrm{d}x, \mathrm{d}y$.

Applying the 1-form $\mathrm{d}f$ to a tangent vector $X = (\alpha_1, \alpha_2)$, we obtain the directional derivative of f, which can be considered the projection of the differential in the direction X.

$$\mathrm{d}f(X) = \alpha_1 \partial_x f + \alpha_2 \partial_y f.$$

We can as well give a notion which generalize the idea of orthogonal and parallel vectors, using the duality in place of the scalar product. Given a form $\omega = \omega_1 \mathrm{d}x + \omega_2 \mathrm{d}y$, its dual vector field is the vector field which has the same components: $X = \omega_1 \partial_x + \omega_2 \partial_y$, formally being parallel to ω. The kernel of the 1-form ω is the set of vector X such that

$$\omega(X) = 0.$$

In the model of odd cells, we considered the 1-form

$$-\sin(\theta)\mathrm{d}x + \cos(\theta)\mathrm{d}y.$$

It is easy to see that the vector fields

$$X_1 = \cos(\theta)\partial_x + \sin(\theta)\partial_y, \quad X_2 = \partial_\theta$$

are orthogonal to ω, so that they belong to the kernel of ω.

The kernel of a 1-form is a subset of the tangent plane. For this reason the action of simple cells RPs on the image can be modeled as the selection of tangent vector (to the level lines) by a 1-form. In particular, the vector X_1, which describes the direction of the level lines of the image, belongs to the kernel of the form ω.

Up to now we have fixed the point (x, y) and considered the tangent space $T_{(x,y)}$ to M at the point (x, y). We can now vary the point (x, y). The union of all tangent spaces is called the tangent bundle of M:

$$TM = \bigcup_{(x,y)\in M} T_{(x,y)}.$$

A point of the tangent bundle is denoted

$$(x, y, \alpha_1, \alpha_2),$$

where $(x, y) \in M$ and (α_1, α_2) belongs to the tangent space at the point (x, y). Then we have a natural projection

$$\pi : TM \to M \quad \pi(x, y, \alpha_1, \alpha_2) = (x, y).$$

More generally let us now define a fiber bundle, which is the mathematical structure proposed to describe the hypercolumnar structure.

Definition 4 Fiber bundle. A fiber bundle is defined by two differentiable manifolds M and C, a group G, and a projection π. C and M are called, respectively, base space and total space.

The total space is locally described as a cartesian product

$$C = M \times G,$$

meaning that at every point $(x, y) \in M$ is associated a whole copy of the group G, called the fiber. The function π is a surjective differential map, which locally acts as follows

$$\pi : M \times G \to M \quad \pi(x, y, g) = (x, y),$$

where g is an element of G. In the paper the base space is implemented in the retinal space and the total space in the cortical space. Particularly the group G of rotations and scales to the point (x, y) is implement in an hypercolumn over the same point.

Definition 5 Section of a fiber bundle. A function

$$\Sigma : (x, y) \to (x, y, \bar{g})$$

defined on the base space M with values in the fiber bundle is called a section of the fiber bundle. In other words a section is the selection of a point on a fiber.

In the model introduced here we associate to each point (x, y) of the retinal plane an orientation and scale value, i.e., a point $(\bar{\theta}, \bar{\sigma})$ in the feature space, defining a section of the cortical bundle

$$\Sigma : (x, y) \to (x, y, \bar{\theta}, \bar{\sigma}).$$

The geometrical tools and bundles described up to now can be used to define complex geometrical structures.

Contact geometry is the study of a geometric structure on smooth manifolds given by a hyperplane distribution in the tangent bundle. This hyperplane can be specified by an orthogonal vector, or, more precisely by an orthogonal 1-form:

Definition 6 Contact structure. A contact structure is an odd dimensional differentiable structure with a 1 form ω, which satisfies a 'maximum nondegeneracy' condition

$$\omega \wedge (\mathrm{d}\omega)^n \neq 0.$$

Starting with a contact structure ω a classical way to define a symplectic structure is to multiply ω by a real variable $\omega_1 = \alpha\omega$ and differentiating ω_1. Note that $\Omega = \mathrm{d}\omega_1$ is a 2-form such that $\mathrm{d}\Omega = 0$. It can be expressed in terms of the standard basis

$$\Omega = \Omega_{ij}\mathrm{d}x_i \wedge \mathrm{d}x_j.$$

The space is now even-dimensional, since we add the extra variable α. Hence the formal definition of a symplectic structure is the following:

Definition 7 Symplectic structure. A symplectic form Ω on an even manifold M is a nondegenerate, closed two-form

$$\Omega = \Omega_{ij}\mathrm{d}x_i \wedge \mathrm{d}x_j.$$

Explicitly, nondegeneracy of the form means that the matrix Ω_{ij} is a skew-symmetric and nonsingular matrix, i.e., with nonvanishing determinant. The requirement that Ω is closed, means that

$$\mathrm{d}\Omega = 0,$$

where d is the exterior derivative.

The simplest example of a symplectic form is

$$\sum_i \mathrm{d}x_i \wedge \mathrm{d}y_i,$$

in a complex space $\{(x_1 + iy_1 \ldots x_n + iy_n)\}$. In this case the 2-form expresses the coupling of the variables x_i with the correspondent variables y_i in the complex structure.

References

Bar O, Sompolinsky H, Ben-Yishai R (1995) Theory of orientation tuning in visual cortex. PNAS 92:3844–3848

Ben S (2003) Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. Neural Comp 16(3):445–476

Ben S, Zucker S (2004) Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. Neural Comput 16:445–476

Blum H (1973) Biological shape and visual science. J Theor Biol 38:205–287

Bosking W, Zhang Y, Schoenfield B, Fitzpatrick D (1997) Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J Neurosci 17(6):2112–2127

Bressloff PC, Cowan JD, Golubitsky M, Thomas PJ, Wiener M (2001) Geometric visual hallucinations, Euclidean symmetry and the functional architecture of striate cortex. Phil Trans Roy Soc B 40:299–330

Bressloff PC, Cowan JD (2003) Functional geometry of local and horizontal connections in a model of V1. J Physiol (Paris) 97:221–236

Carandini M, Ringach DL (1997) Predictions of a recurrent model of orientation selectivity. Vision Res 37:3061–3071

Citti G, Sarti A (2006) A cortical based model of perceptual completion in the roto-translation space. J Math Imaging Vision 24(3):307–326

Das A, Gilbert CD (1995) Long range horizontal connections and their role in cortical reorganization revealed by optical recording of cat primary visual cortex. Nature 375:780–784

De Angelis GC, Ozhawa I, Freeman RD (1995) Receptive-field dynamics in the central visual pathways. Trends Neurosci 18(10):451–458

Field DJ, Hayes A, Hess RF (1993) Contour integration by the human visual system: evidence for a local association Field. Vision Res 33:173–193

Fitzpatrick D (1996) The functional organization of local circuits in visual cortex: insights from the study of tree shrew striate cortex. Cerebral Cortex 6:329–341

Hubel DH (1988) Eye, brain and vision. Scientific American Library

Issa NP, Trepel C, Stryker MP (2000) Spatial frequency maps in cat visual cortex. J Neurosci 22(22):8504–8514

Kimia BB (2003) On the role of medial geometry in human vision. In: Petitot J, Lorenceau J (eds) Neurogeometry and visual perception. J Physiol Paris 97(2–3):155–190

Lund J, Fitzpatrick D, Humphrey AL (1985) The striate visual cortex of the tree shrew. In: Jones EG, Peters A (eds) Cerebral cortex. Plenum, New York, pp 157–205

Marr D (1982) Vision. Freeman, San Francisco

Mitchinson G, Crick F (1982) Long axons within the striate cortex: their distribution, orientation, and patterns of connections. PNAS 79:3661–3665

Miller KD, Kayser A, Priebe NJ (2001) Contrast-dependent nonlinearities arise locally in a model of contrast-invariant orientation tuning. J Neurophysiol 85:2130–2149

Parent P, Zucker SW (1989) Trace interference, curvature consistency and curve detection. IEEE Trans Pattern Anal Machine Intell 11:823–839

Petitot J (1989) Modèles morphodynamiques pour la grammaire cognitive et la sémiotique modale. RSSI Can Semiotic Assoc 9(1-2-3): 17–51

Petitot J, Tondut Y (1999) Vers une Neurogeometrie. Fibrations corticales, structures de contact et contours subjectifs modaux, Mathematiques, Informatique et Sciences Humaines, vol 145. EHESS, CAMS, Paris, pp 5–101

Petitot J (2003) The neurogeometry of pinwheels as a sub-Riemannian contact structure. In: Petitot J, Lorenceau J (eds) Neurogeometry and visual perception. J Physiol Paris 97(2–3):265–309

Rockland KS, Lund JS (1982) Widespread periodic intrinsic connections in the tree shrew visual cortex. Science 215:532–534

Rockland KS, Lund JS (1993) Intrinsic laminar lattice connections in the primate visual cortex. Journal Comp Neurol 216(3):303–318

Shelley M, Wielaard DJ, McLaughlin D, Shapley R (2000) A neuronal network model of macaque primary visual cortex (V1): orientation selectivity and dynamics in the input layer $4C_\alpha$. PNAS 97:8087–8092

Yen SC, Finkel LH (1998) Extraction of perceptually salient contours by striate cortical networks. Vision Res 38(5):719–741

Swindale NV (2004) How different feature spaces may be represented in cortical maps. Network 15:217–242

Swindale NV (2000) How many maps are there in visual cortex. Cerebral Cortex 7:633–643

Ts'o D, Gilbert CD, Wiesel TN (1986) Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. J Neurosc 6(4):1160–1170