Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

BIOINGEGNERIA

Ciclo **XXVI**

Settore Concorsuale di afferenza: 09/G2

Settore Scientifico disciplinare: ING-INF/06

TITOLO TESI

SPATIO-TEMPORAL MODELS OF THE FUNCTIONAL ARCHITECTURE OF THE VISUAL CORTEX

Presentata da: **GIACOMO COCCI**

Coordinatore Dottorato	**Relatore**
Prof.ssa ELISA MAGOSSO	**Prof. MAURO URSINO**
_____	_____
	Correlatori
	Prof. ALESSANDRO SARTI
	Prof.ssa GIOVANNA CITTI
	_____

Esame finale anno 2014

Spatio-temporal models of the functional architecture of the visual cortex.

Giacomo Cocci

A mio padre, Luciano.

Contents

Conclusions

Bibliography

Introduction

Understanding the functional architecture of the primary visual cortex (V1) in mammals has become a major subject within the neuroscience research field. Since the first publications of Hubel and Wiesel in the 1960s, many scientists from very different backgrounds have tried to make their contribution to the general objective of writing a mathematical model for the processes of visual perception and cognition [51, 52].

As soon as electrophysiology permitted the recording of the activity of single cortical neurons, the concept of receptive fields (RFs) emerged. The concept of a receptive field has been widely spread in the academical scenario since the beginning of the century, when Sherrington first coined the term to describe the area of the body surface where a stimulus could elicit are reflex [110]. In 1938, Hartline refined the concept and extended it to sensory neurons, stating that "responses can be obtained in a given optic nerve fiber only upon illumination of a certain restricted region of the retina, termed the receptive field of the fiber" [45]. This hasn't been a trivial discovery, as the first electrophysiological experiments made on ganglion and bipolar cells were made using spatially unlimited stimuli, as the brightening of a whole screen, causing small or irrelevant changes in the neural response. Later, the term was also extended to other neurons in the visual pathway, defining the limited area within the visual space where a luminous stimulus could drive electrical responses in a given visual neuron. This electrical response

can be defined as the firing rate or take into account also sub-threshold activity —
depolarization and hyperpolarization of the membrane potential that do not give
rise to an action potential. While the latter choice is obviously the only possible
for studying bipolar cells, as their activity do not generate spikes, the RFs of cells
belonging to higher cortical areas are generally recovered by the elaboration of
their firing rate behavior.

After the pionierisitc work of Hubel and Wiesel, it is widely known in the aca-
demic environment that simple and complex cells in V1 are characterized by elon-
gated receptive fields, and consequently respond best to elongated stimuli, namely
bars, edges, boundaries or contours. With the development of advanced RF re-
construction methodologies, it was demonstrated that their different sensitivity
regions could be modeled with plane waves modulated in amplitude by a Gaussian
envelope. These bi-dimensional Gabor functions lead to the discovery that cortical
neurons are designed to efficiently solve the uncertainty between space/frequency
localization of the visual stimulus [53, 54, 24]. Further studies discovered the main
role of the temporal dimension in the description of the elaboration that the brain
performs on the visual signal coming from the retina [17]. Particularly interesting
spatio-temporal behaviors were found in the response properties of neurons be-
longing to the lateral geniculate nucleus (LGN) and to V1, with excitatory and
inhibitory sub-regions translating through time from the stimulus onset [26, 93].
Indeed, A large class of simple cells and the vast majority of complex cells in V1
shows a very specific space-time behavior in which the spatial phase of their RF
changes gradually as a function of time. This results in RF profiles that tilt along
an oblique axis in the space-time domain.

In the first chapter I will resume the findings oobtained by analyzing a data set
of this kind of spatio-temporal RFs reconstructed by electrophysiological record-
ings via the method of reverse correlation. The data that I reconstructed refers

to 93 spatio-temporal RFs completely describing the response dynamics of V1 simple and complex cells belonging to adult mammals (cats), using stimuli comprising both spatio-temporal white noise distributions and natural scenes. The RFs were fitted using a 3-dimensional Gabor model, which minimizes uncertainties between localization the classical and the Fourier domain. Approximating the raw data with this model and analyzing the results, I found relevant physiological constraints operating within the parameters characterizing the cortical cellular behavior. In fact, the modeled spatio-temporal RFs do not span the whole parameter space, but define a particular sub-space that is sufficient to describe all simple and complex cells in V1. In this thesis I identify this sub-space and I will propose an additional constraint operating within cortical cells: a minimization of the uncertainty over local velocity measurement. This study showed that spatio-temporal receptive profiles can be well approximated by the weighted sum of three-dimensional Gabors, and their unique organization optimizes the spatio-temporal resolution, i.e. the precision with which a neuron can locate stimulus velocities.

Thus, the discovered characteristics of the cells in the primary visual cortex suggest to regard an image as a set of points defined on an extended domain, where every point belonging to the image plane $\mathbb{R}^2$ is lifted to a point in an extended n-dimensional space. Indeed, we know that these cells are spatially organized in such a way that for every point (x, y) of the retinal plane there is an entire set of cells, each one sensitive to a particular instance of the considered feature, that could be the one of local orientation or velocity, giving rise to the so-called hypercolumnar organization. Hypercolumnar organization and neural connectivity between hypercolumns constitute the functional architecture of the visual cortex, that is the cortical structure underlying the processing of visual stimulus.

The mathematical modelling of the functional architecture of the visual cortex in terms of differential geometry was introduced with the seminal works of Koenderink [57] and Hoffmann [48, 49]. While the first author pointed out the differential action of perceptual mechanisms, in particular with respect to jet spaces arising from linear filters, the second author proposed to model the hypercolumnar organization in terms of a fiber bundle structure and pointed out the central role of symmetries in perception expressing them in terms of Lie groups and Lie algebras.

The problems of perception can also be addressed by a purely psychophysical approach. The study of Field Hayes and Hess [34] introduced the notion of association field, as path of information integration along images that can quantitatively satisfy the assumptions of the Gestalt principle of good continuation. The perceptual role of this mechanism was indeed that of contour integration, that typically occurs along field lines associated to locally coherent directions in images.

Almost simultaneously Mumford [73] proposed a variational approach to describe smooth edges, in terms of the elastica functional, that could be implemented with stochastic processes defining curves with random curvature at any point. Indeed, they produce probability distributions in the space $\mathbb{R}^2 \times S^1$ of positions and orientations whose probability peaks follow elastica curves. Williams and Jacobs [136] used such stochastic processes to implement a mechanism of stochastic completion, and interpreted the probability kernel they obtained as tensors representing geometric connections on the space of positions and orientations associated to the neural representation of images due to simple cells.

Many of such results dealing with differential geometry were given a unified framework under the new name of *neurogeometry* by Petitot and Tondud [89], who related the association fields of Field Hayess and Hess with the contact geometry introduced by Hoffmann and the elastica of Mumford. The problem of edge organization in images was then addressed in terms of a stochastic process of

the type of Mumford, introducing nonlinearities in order to take into account the role of curvature, by August and Zucker [6, 7], while the variational approach of Nitzberg, Mumford and Shiota was extended from edges to level lines of images by Ambriosio and Masnou [3].

Then, in [19], Citti and Sarti showed how the functional architecture could be described in terms of Lie groups structures. They interpreted the geometric action of receptive profiles as a lifting of level lines into the space $\mathbb{R}^2 \times S^1$ of positions and orientations, and addressed the problem of occlusion with a nonlinear diffusion-concentration process in such a space of liftings. In particular, this approach allows to introduce orientation, instead of depth, as a third dimension for the disentanglement of crossing level lines. In their model, then, contour completion is justified as a propagation in the sub-Riemannian setting, and the integral curves of the vector fields that generate the Lie algebra can be considered as a mathematical representation of the association fields of Field, Hayes and Hess, hence proving the relation between neural mechanisms and image completion. This method was then concretely implemented in [105]. The problem of boundary completion was also addressed from a slightly different point of view by Zucker [141], who showed the role of Frenet frames.

Exact solutions to the Fokker-Planck equation associated to Mumford stochastic process were provided by Duits and van Almsick [31], and later Duits and Franken [28, 29] unified such stochastic approach with nonlinear mechanisms of the type of August and Zucker, keeping left invariance with respect to the Lie symmetry of the Euclidean motion group and yet allowing the invertibility of the whole process. Their result was applied to the problem of contour enhancement and contour completion, working on the whole Lie group by means of a representation via suitably defined linear filters. This approach was then extended to

different geometric setting by Duits and Führ [30], again with applications to the processing of images.

In the second chapter of this thesis I describe an extension of the Citti-Sarti model of neurogeometry [19]. In the proposed geometrical setting, a base variable (time) and a fiber variable (local velocity) are added to the $\mathbb{R}^2 \times S^1$ contact structure of visual position/local orientation: this defines a fiber bundle, a generalization of the cartesian product $(x, y, t) \times (v, \theta)$, where to every spatio-temporal point (x, y, t) of the base space is associated the full fiber of possible values of orientation and velocities (v, θ). To each point in the space (x, y, t, v, θ), thus, is naturally associated the hyper-plane (called horizontal hyper-plane) whose $x - y$ projection is the line with orientation θ. Curves passing through a point (x, y, t, v, θ), with orientation θ, and velocity v are called admissible or horizontal curves of the structure, since their tangent vector always belong to the horizontal plane. Due to these admissibility conditions, a local stimulus measurement represented by a point $m_0 = (x_0, y_0, t_0, v_0, \theta_0)$ of the manifold will have a certain probability of pertaining to affine stimuli occupying given admissible regions on the space. Considering the resulting probability density function as the probability of the activity of two cells in the visual cortex of being mutually facilitated by long-range horizontal connections, we can model the propagation of the measurements performed by the visual cortex by first convolving a stimulus with a set of 3D Gabor functions, and then convolving the output with the stochastic kernel generated by the proposed contact structure. In particular, I will show through two numerical simulations that this connectivity model can correctly reproduce some non-linear spatio-temporal behaviors of the cells in the primary visual cortex, allowing me to propose a possible mechanism underlying different non-trivial effects found in the literature by means of phenomenological and psychophysiological experiments.

As previously mentioned it is well understood since the beginning of the century, when the psychological concepts of the Gestalt began to be defined, the importance of key global properties of the visual stimulus (neighboring, good continuation, etc.) in the execution of visual cognitive tasks as image segmentation and grouping [134, 124]. Field et al. made some experiments where observers were presented with an ensemble of Gabor patches, a subset of which was consistently aligned along a continuous path. Analyzing whether observers perceived the embedded path or not, they drew position/orientation perceptual association fields, showing that stimulus co-linearity and co-circularity play an important role for feature grouping. Their study showed how chances of perceiving the curvilinear path were high, if the orientation of its features was the one tangent at that point, collapsed as their relative orientation deviated from being tangent, and became significant again when the elements were set orthogonally to the path [34].

We know from many phenomenological findings that the grouping properties obtained by spatial collinearity can easily be broken if one associates a speed and an orthogonal direction of movement to each oriented segment, where the orthogonality is a constraint suggested by the spatio-temporal RPs in the visual cortex. Limits have been found on the maximum rate of change of local speed along a contour, so that based on that visual information the perception of boundaries and shapes is possible. Indeed, a random speed distribution over a dashed line would completely destroy the perception of a single unit as a whole, while enhancing the impression of different segments pertaining to the random background field [92].

Following these findings, it has been introduced by Rainville the concept of motion contour [92]: carrying out some psychophysiological experiments he showed how the brain groups features together also relying on the local speed perpendicular to their orientation axis, with coherent velocities being represented by velocity fields that vary smoothly over space. The former study expanded the already

known notions that local stimulus velocity is discernible (thus determinant for grouping purposes) only when it is orthogonal to the perceived contour or it is not part of a trajectory [60, 46, 121]. Coherently, the analysis carried out in Chapter 1 over a data set of cortical neurons in the primary visual cortex showed how the spatio-temporal shape of their RFs is biased to optimally measure the local stimulus velocity. Thus, it may be inferred that stimulus local direction of movement and speed are additional features driving the spatial integration involved in the perception of shapes and contours.

Another kind of grouping is the one that we perform in space-time, for example when a distinct moving or deforming shape disappears and reappears in the visual field because of the occlusion caused by another moving object. Similarly to what happens for the integration of spatial visual information, the brain is capable to easily predict stimulus trajectories, and to group together elements having similar motion or apparent motion paths. The facilitation in detecting stimuli in motion given a previous cue with coherent trajectory is found to be significantly high, and it cannot rely just on the temporal response summation given by the onset and offset dynamics of classical RFs [118]. One possible explanation for these non-linear effects could be the existence of a specialized facilitatory network linking cells anisotropically and coherently with their axis of motion direction. Ledgeway and Hess studied the perception of spatial contours defined by non-oriented stimuli moving coherently and tangentially along a path, finding rules similar to the ones driving facilitation in position/orientation and inferring a possible role played by a trajectory-specialized network [59]. Another possible evidence of a trajectory-driven connectivity comes from a recent study of the dynamics of neural population response to sudden change of motion direction, where it is shown that for low angular changes a non-linear part of the response provides a sort of spatio-temporal interpolation [138].

Unfortunately, most of the literature on the phenomenology of motion perception put the focus on moving dots. With this particular experimental setting, the presence of a facilitation governing trajectory prediction could also be explained by the classical orientation-driven horizontal connections in response to fast motion streaks [38, 40]. Indeed, the important role played by V1 collinear horizontal connectivity in motion perception has been already hypothesized by Georges et al, who showed that perception of speed is biased by the orientation of a feature's axis of motion: in particular, flashing orientated Gabor patches are perceived as moving faster when they are tangential to the path of apparent motion, while placing them orthogonally makes the effect vanish. Their subsequent analyses outlined that the paths of apparent motion can also be curvilinear, coherently with the position/orientation association fields, and the effect is maximum when the apparent speed matches the conduction velocity of the long-range horizontal connections between orientation-selective cells in V1 [112]. Nevertheless, a recent study showed how trajectories of oriented segments are significantly more detectable for orientations orthogonal to the path of motion, thus once more proposing the existence of two different facilitatory mechanisms [86].

In the third and final chapter I will carefully test the visual spatio-temporal grouping properties of the connectivity geometrical model proposed in Chapter 2. To simulate this grouping effect, I use the a priori geometrical knowledge underlying the definition of the contact structure described in Chapter 2, by applying it to a revised version of a popular method for dimensionality reduction, that is, spectral clustering. I will show how, in the case of spatial grouping of contours in motion, even if the visual perceptual units can be correctly retrieved and distinguished by using just the information on position and local orientation, the spatio-temporal information of local velocity greatly improves the quality of the grouping, assigning to the main objects much less outlier elements, such as noise

or background. In the end of the chapter I will apply the same methodology to full shapes and contours moving in time, so that the perceptual units are defined over the whole $\mathbb{R}^2 \times \mathbb{R} \times S^1 \times \mathbb{R}^+$ spatio-temporal domain. In this case, the affinity matrix can be built using as affinities the values of the 5-dimensional stochastic kernel described in Chapter 2. I will test two different methods of spatio-temporal grouping of shapes, relying on the cellular mechanisms of cortical functional areas V1 and V5/MT, and basing on the results I will discuss and propose the possibility of having many different cortical areas cohoperating in order to carry out advanced tasks of visual grouping.

Chapter 1

The perception and representation of motion in V1

1.1 Introduction

In this study, 93 spatio-temporal RF profiles completely describing the response dynamics of V1 simple and complex cells belonging to adult cats were reconstructed via reverse correlation and spike-triggered covariance analysis, using stimuli taken from both spatio-temporal white noise distributions and natural scenes. Appropriate expedients were used in order to attenuate the effects that derive from using data recorded with different experimental procedures.

The aim of this chapter is to show that while spatio-temporal receptive profiles can be accurately approximated by the weighted sum of three-dimensional Gabors, their unique organization optimizes the spatio-temporal resolution, the precision with which a neuron can locate stimulus velocities. The 3D Gabor model minimizes uncertainties between localization in classical domains (the two spatial dimensions and the temporal dimension) and localization in Fourier domains (three-dimensional Fourier space), dictated by the general information theory's un-

certainty principle [24, 36]. Approximating the raw data with this model and analyzing the results, I found relevant physiological constraints operating within the parameters characterizing cortical cell behavior. In fact, modeled spatio-temporal RFs do not span the whole parameter space and define a particular sub-space that is sufficient to describe all simple and complex cells in V1. This sub-space was identified and an additional constraint operating within cortical cells has been proposed, a minimization of the uncertainty over velocity measurement.

In section 2A I briefly describe the reconstruction of the data set used for the analysis. In section 2B I present the 3D Gabor model used in this paper to approximate the raw RFs data, with an analysis of its main advantages and drawbacks. Section 3A shows the distributions of some of the parameters extrapolated from the fitting process, while in section 3B I define a new kind of uncertainty over velocity measurement and calculate it for each RF in the data set. Finally, in section 4 I conclude these analyses and suggest possible future developments.

1.2   Methods

1.2.1   RF reconstruction

Primary visual cortex simple cells can be seen as a linear systems with multiple inputs, each referring to a different point within the RF, and a single output, followed by a non-linear stage involving saturation and the generation of action potentials. The linear stage of these systems is fully characterized by a 3D spatio-temporal impulse response $h(x, y, t)$ [97]. Intuitively, one can imagine this impulse response as a temporal sequence of spatial maps over the X-Y plane, one for every time instant after stimulation. This sequence can be played as a movie and is a convenient visualization that describes the spatio-temporal dynamics of the neural sensitivity of a given cortical cell (Fig. 1.1).
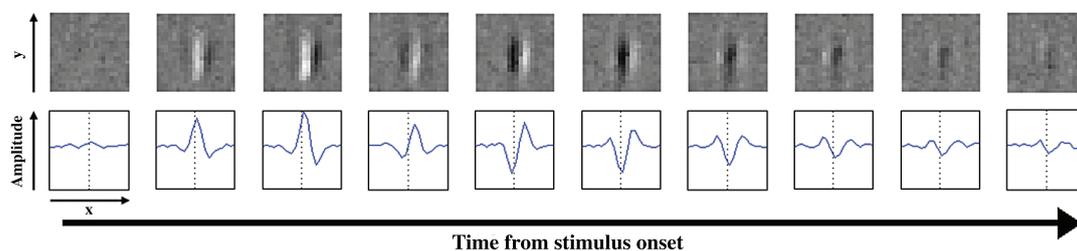
Figure 1.1: In the first row is shown the time course of a simple cell's RF. In the second row are plotted, for each frame, one-dimensional profiles obtained by summing RF data along the axis parallel to the cells preferred orientation. The middle vertical line in these plots highlights how the RF subregions shift in space through time. This cell is direction-selective and this kind of profile is said to be space-time inseparable. The time interval between each frame is 15 msec. (Two columns)

The first data set analyzed in this paper was downloaded from the Visiome Network, a web-based database system with a variety of digital research resources for vision science [122]. The package contained raw spike data file samples coming from electrophysiological recordings made on adult cats for a total of 8 binocular simple cells. Stimuli were dark and bright oriented bars projected at random positions within the RF of the cell being recorded. As these cells were responding to visual information coming from both eyes, two receptive profiles could be extracted from a single cell, representing the linear spatio-temporal behavior of the neuron with stimuli coming from the right or the left eye. The data was shared by Ohzawa who, with DeAngelis and other authors, used it to make a series of studies on cat simple cells, dealing with cell development, binocular disparity and general cortical organization of simple cells in both adult cats and kittens [26, 25, 79]. The reverse correlation process was carried out by a software application released by Ohzawa together with the data.

I recovered a second data set by analyzing neural data recorded from adult cats by Tim Blanche in the laboratory of Nicholas Swindale, University of British Columbia. Data were made available on the NSF-funded CRCNS Data Sharing website [23]. The stimuli utilized were natural scenes, thus I utilized a modified version of the reverse correlation process for the reconstruction to avoid biased RF estimates. The details of this method can be found in [116]. I utilized only six RFs out of ten for the analysis in this work, all of them belonging to simple cells. As recordings were made using polytrodes through various cortical layers [12], there was no prior assurance about the homogeneity of the cell types. In fact, I discarded four of the reconstructed profiles as they presented much more complicated behavior and a high level of noise.

Another data set was downloaded from the NSF-funded CRCNS Data Sharing website [23], obtained by extra-cellular recordings from V1 of anesthetized adult cats. Visual stimuli were one-dimensional white noise (random bars) aligned to the preferred orientation of each cell. These experiments were specifically performed to measure the spatio-temporal RFs of cortical complex cells, and I utilized a spike-triggered covariance method, whose detailed description can be found in [117], to reconstruct 71 spatio-temporal sub-units (those features in the visual stimuli affecting the firing probability of a complex cell) driving the neural responses of 48 cells.

1.2.2 RFs fitting with the 3D Gabor model

The method that I utilized for modeling spatio-temporal simple cell profiles and complex cell sub-units is based on the well-established 2D Gabor model that is conventionally used in describing cellular sensitivity upon the two-dimensional visual space [53, 54, 66]. This model is particularly significant, as Gabor himself in his ground-breaking paper on information theory showed that the family of

functions that take his name is capable of locating an event, both in position and momentum, with the maximum precision possible, thus reaching the theoretical limit over the uncertainty of the measurement of two conjugate variables given by the uncertainty principle [36]:

$$(\Delta x)(\Delta y)(\Delta \xi)(\Delta \eta) \geq \frac{1}{4}\pi^2, \tag{1.1}$$

where Δx and Δy are the uncertainty values over the localization along the x- and y-axis, while $\Delta \xi$ and $\Delta \eta$ are the uncertainty values relative to the localization within the bi-dimensional Fourier domain.

It is known that the notion of a bi-dimensional Gabor wavelet can be generalized from two spatial dimensions to three spatio-temporal dimensions (2D space and time) [1, 90]. This model can be simplified if one decides to sacrifice the information relative to the axis parallel to the elongated sub-regions direction in favor of a simpler description of a cell's spatio-temporal behavior: this is possible by summing each frame along the cell's preferred direction in order to reduce the model to a plane wave within an opportune spatio-temporal window, similarly to what is conventionally seen for profiles over two spatial dimensions:

$$\begin{aligned}
\check{g}(x,t) &= \exp\left\{-\frac{(x-x_0)^2}{2\Delta x^2} - \frac{(t-t_0)^2}{2\Delta t^2}\right\} \\
&\times \exp\left\{-2\pi i\left[\xi_0(x-x_0) + \omega_0(t-t_0)\right]\right\},
\end{aligned} \tag{1.2}$$

where ω_0 is the temporal frequency determining, together with the spatial frequency perpendicular to the cell's preferred direction $(-\eta_0, \xi_0)$, the velocity at which the excitatory and inhibitory sub-regions move within the area subtended by the RF.

It is worth noting that in doing this we lose information about the RF behavior along the y axis, but we am not precluding the model from minimizing the previously formulated uncertainty principle, as long as I suppose the RF has

a Gaussian amplitude response along that direction. This simplified Gabor kernel was utilized to accomplish motion perception tasks even before the reverse correlation technique permitted the reconstruction of spatio-temporal RF profiles [1, 129].

Thus, the spatio-temporal model results are very similar to the 2D spatial RFs, one of the principal differences being the Gabor kernel orientation,

$$\phi = \arctan \frac{\omega_0}{\xi_0}. \tag{1.3}$$

When the 2D domain was merely spatial, this property referred to the orientation that a stimulus had to present to get the maximum response from the cell. Orientation in the spatio-temporal domain represents instead the optimal stimulus velocity ω_0/ξ_0: the more the kernel is skewed, the higher the stimulus velocity for which the cell is tuned. This is valid for both inseparable and separable profiles, with the distinction that separable profiles respond maximally to both movement directions. The stationary oscillations that are typical of separable profiles can, indeed, easily be obtained by inseparable members of the Gabor family function, leading to the final model I used for the fitting process:

$$\check{g}_s(x,t) = \exp\left[i(\xi_0 x - \omega_0 t] + C \exp\left[i(\xi_0 x + \omega_0 t] = 2\exp(i\xi_0 x)\cos(\omega_0 t), \tag{1.4}$$

where C is a constant going from 0 to 1. This parameter is called the separability index, used also by DeAngelis in his model as being the sum of two separable components [27]. The fitting results showed how cortical RFs rarely present purely separable or purely inseparable profiles.

The scope of the modeling process is to evaluate the capacity of a cell in V1 to measure stimulus velocity. The proposed model is not causal nor physically realizable as the Gaussian function modulating the RF temporal amplitude is neither of these. Still, the fitting results show that it can be regarded as an approximation of a causal filter as the Gaussian amplitudes at time $t = 0$ are

always at the order of magnitude of the noise. The non-causal Gabor model allows us to appreciably maintain the RFs main responsive properties while studying their behavior by seeing them as optimal measurement systems and including as internal parameters directly relevant features of the physiological visual processing stage (especially temporal frequency).

The downside of this model is the almost systematic incapability of the temporal Gabor to capture the ending low-amplitude slow dynamics, which sometimes appear in many cortical spatio-temporal RFs. After a quick spectral analysis, one could infer that the role of these slow dynamics could be to approximately maintain a null-mean temporal response without making the velocity-sensitive main dynamics less effective. This observation is clear when one observes Fig. 1.2: along the horizontal axis relative to the temporal null frequency the response amplitude of the RF is always very low, thus not significantly responding to non-moving stimuli.

A quick review of the available literature shows how some previous models, like the DeAngelis one, use a Gaussian envelope on a deformed time axis, thus still ignoring the causality of the filter [27]. Some other models use a causal approach to model the behaviour of cortical RFs, but in doing this one would sacrifice the ease of the extrapolating meaningful parameters [1, 128, 33].

Spatio-temporal reconstructed raw data have been fitted using this model. A total of 93 RFs were studied, 22 reconstructed from recordings made on simple cells and 71 made on complex cells. Although the experiments in which the data were recorded show some variations, as do the final assumptions they were thought to show, I utilized various reconstruction methods in order to reduce the variance due to these differences as much as possible, and the coherence found in between elements belonging to different data sets suggests the success of the proposed methodology. The accuracy of the approximations, the significant correlations found within the parameter distributions, and the concordance of these correlations

Cell #2 (Inseparable profile)

Cell #86 (Separable profile)

Figure 1.2: Plots showing reconstructed data relative to two RFs, both in spatio-temporal classical and Fourier domains. Separable profiles can be obtained by the product of two separate real functions defined over space and time (respectively), and they respond to both stimulus movement directions, while inseparable profiles are structurally direction-selective. The white lines on the spectral plots mark the horizontal axis relative to the zero temporal frequency. It can be noted that response amplitudes over this line remain consistently low, which is an effect of the late slow dynamics of the RFs.

with previously published studies allow us to say that the analysis carried out in this thesis is in some way more robust and comprehensive than the previously cited studies, investigating whether general constraints or relationships are present within the cortex when it's seen as a whole.

I have discarded four of the studied receptive profiles which had only one temporal semi-cycle, thus presenting the same spatial regions of sensitivities throughout the whole duration of their impulse response. The reason is that the approximation with the Gabor model of these non-dynamical types of cells could give false temporal frequency values. Overall, these analyses are focused on RF velocity measuring capabilities, thus I am not interested in non-dynamical cells.

1.3 Results

The model fitted well the spatio-temporal profiles of both complex and simple cells, and could robustly describe some temporal properties that were hardly or only partially captured by other models. The results of the fit for some of the cells in the data set are shown in Fig. 1.3. I calculated a percentage error for each approximated RF by normalizing both the raw data and the fit so that the L2-norm of their difference, divided by two, lies between 0 and 1. The mean fit error was $0, 34$, a relatively low value considering the model utilized and the results obtained in previous studies [27]. Each of the 93 receptive profiles can be labeled with the four parameters that characterize the corresponding 3D Gabor filter: spatial frequency ξ_0, temporal frequency ω_0, spatial width Δx, and temporal width Δt. In the sections that follow, I will refer to frequencies as the absolute value of the corresponding parameters found by the fitting process. The next section analyzes the distribution of the filter parameters, discussing some of the correlations found among them. Most of them are identified in the literature, but here the novelty is

that they appear to be consistent within the three different datasets, disregarding the distinction between simple and complex cells. The second part of the section provides a remarkable new feature of this distribution in terms of optimality of the family of filters with respect to velocity measurements.

1.3.1   Parameters distribution

One of the most studied indices found in the literature is the number of ON and OFF spatial sub-regions within the RF area, driving the excitatory or inhibitory effect that a visual stimulus has on the response of a cortical neuron. Using 1.5, I can compute the sub-region index (DSI) as the number of half-wavelengths, or semi-cycles, covered within two standard deviations:

$$DSI = 8\Delta x \xi_0. \tag{1.5}$$

Fig. 1.4a shows the joint distribution of these two parameters. The totality of the cells studied is shown. The distributions are plotted in logarithmic scale to show clearly the inverse relation between the parameters, that in this representation are distributed along a straight line. It can be seen from the fluctuations that are visible in the histogram that both simple cell profiles and complex cell subunits share the same approximate inverse proportionality relation. This fact has a physiological significance, showing that RF profiles generally have two or three spatial ON and OFF alternating sub-regions, which suggests that they could be thought of as many scaled versions of some general basic structure. These results confirm the conclusions previously made by other authors [25, 95].

Modeling spatio-temporal RF profiles with bi-dimensional Gabor functions allowed us to make the sub-region analysis with respect to the temporal behavior. The product between the impulse response duration Δt and the preferred temporal frequency ω_0 can be used to count the number of alternating excitatory and

Figure 1.3: Plots showing four different example RFs from the data set - one for each row. From left to right, original spatio-temporal data, the fitting obtained using the Gabor model and the function relating velocity uncertainty and RF shape. The function shown in the right plot always presents an optimal value of ϵ where the uncertainty over stimulus velocity measurement is minimized. Crossed circles are plotted in correspondence to the value relative to the true spatio-temporal shape of the plotted RFs, always placing themselves near to the theoretical minimum value. This is true for all the RFs in the data set. The fit error defined at the beginning of section 3 for the RFs in the figure is, from top to bottom, 0, 23,0, 26,0, 19,0, 37.

inhibitory sub-regions subtended by the profiles. Fig. 1.4b shows the joint distribution of the two temporal parameters, and a regularity in following the inverse proportionality rule is also evident here.

(a)

(b)

Figure 1.4: (a) In the left plot is the distribution of the number of sub-region index (DSI) of the profiles and sub-units belonging to the cells studied: a Gaussian function representing distribution mean and standard deviation is plotted with a dashed line. The right-hand plot illustrates the inverse proportionality governing size vs. spatial frequency distribution with a continuous line corresponding to the mean value of their product and dashed lines corresponding to the mean $\pm 2SD$. (b) The same type of plots show relations between temporal frequency and RF duration.

Once again, I have chosen to analyze only cells which present dynamic behavior, discarding those whose receptive profile was not oscillating with time. With this in mind, we can appreciate how in this case the DSI distribution is narrower than in the spatial case, with temporal dynamics generally lying between 1,6 and 2,4

semi-cycles. This index distribution seems to infer that cortical neurons do not need more than one oscillation to accomplish the task for which their temporal behavior is designed. From the point of view of velocity estimation, this seems to be adequate to measure a motion that is approximately linear and uniform within the time window, providing a more rough characterization than the one given by spatial resolution. It is also worth noting that, dividing the mean values of the two distributions we obtain the following:

$$\frac{\Delta t \omega_0}{\Delta x \xi_0} = v \frac{\Delta t}{\Delta x} \simeq \frac{2}{3},$$ (1.6)

from which it is possible to infer that RFs in the data set have plane waves that travel through approximately the 2/3 of the total RF size in the time Δt in which they are active.

These results confirm the conclusions previously made by other authors that the number of spatial excitatory and inhibitory sub-regions of a cell RF is constrained to stay within a small range of values [25, 95]. Sasaki and Ohzawa, performing a study comparing simple and complex cells, stated that complex cell sub-units generally present significantly more sub-regions than simple cell profiles [108]. This fact was not confirmed by the present work, although there were significantly fewer recovered simple cell profiles than complex cell sub-units. Furthermore, recent studies have pointed out that the distinction between simple and complex cells cannot be made as easily as previously thought, and parameter bimodal distributions might be directly caused by this classification [69].

One of the most widely researched relationships between RF profile parameters of visual neurons is the one that compares the tuning for spatial frequencies with the preference for temporal frequencies. Many authors in the last two decades have noticed how a negative correlation between these two variables is generally appreciable, that is, cells tuned for higher spatial frequencies will most probably

present slower temporal dynamics in their responses to visual stimuli and vice versa. This trend was found in independent studies involving electrophysiological recordings from both LGN cells and cortical neurons, regarding both simple and complex cells [67]. Nevertheless, the nature of this relationship, whether the parameters are directly or inversely proportional, has never been precisely defined, mainly because of the wide variance that the distribution presents. Intuitively, when working with variables related to physiological response spectral characteristics one could expect the values to present asymptotes corresponding to their physical constraints, yielding to an inversely proportional relationship. Recently, Tan and Yao asserted that the spatio-temporal characteristic distribution of visual neurons in the LGN is the one that provides the most efficient way to represent stimulus information [115]. In doing this, they showed that the joint distribution of preferred spatio-temporal frequencies seems to follow a strict inversely proportional rule, the product of the two parameters being always lower than a certain value.

Simple cell RF spatio-temporal dynamics were studied by DeAngelis [26]. He also investigated the changes in the parameter distribution during the cortical development by making recordings from visual neurons of adult cats and kittens aged four- and eight-weeks. As a result, he found a significant negative correlation between spatial and temporal frequency tuning in adult cats, and he noticed how this relationship does not occur innately, but slowly emerges along with the brain development. Again, the joint parameter distribution was so sparse that a clear relationship was indistinguishable, even if a general trade-off between the two selectivities was fairly evident.

Fig. 1.5a shows the joint spatio-temporal frequency distribution of the 93 studied profiles. It is possible to see that a fairly significant negative correlation is present, although the great distribution variance does not make it possible to infer

that this is the exact general rule followed by cortical neurons. This distribution agrees with the findings of DeAngelis, whose joint parameter distribution is very similar to the one shown here.

Another constraint that seems to limit the possible Gabor filters used by V1 to process visual stimuli is shown in Fig. 1.5b, where the joint distribution of the uncertainties relative to the two classical domains, Δx and Δt, is plotted. A significant negative correlation is present here, even if the distribution is very sparse.

Physiologically, this fact can be interpreted by recalling the fact that cells that have larger RFs receive input from a greater number of retinal photoreceptors in the first place, which concur in raising the membrane potential of the output synapsed ganglion cells more rapidly. It is possible, then, to imagine these retinal cells starting to spike and exhaust the sensitivity of their impulse response, for example, because of early inhibition before other cells with smaller RFs. This temporal advantage could be propagated (and, perhaps, amplified) as visual information travels through the neural visual pathway. This reflects, in some way, the explanation that Weng, Yeh and Alonso gave to explain the fact that RF size and response latency are directly correlated within the cat LGN [133]. Furthermore, this would not contradict the well-established fact that pre-cortical axon conduction velocities favor the cells that process the lowest spatial frequencies [67].

The sparsity of this distribution is justified if one regards the RFs as stimulus velocity measurement filters: in fact, as we will see in the next section, the sensibility range of stimulus velocity is not only a function of these two variances, but also, in a more complex way, of the position in the spatio-temporal spectral plane. From a signal processing point of view, it therefore makes sense for this distribution correlation to be fairly small.

(a)

(b)

(c)

(d)

Figure 1.5: (a) Joint spatio-temporal frequency distribution. Complex cell sub-units are indicated as black squares, while simple cell profiles are drawn as circles. A dashed line is drawn over a robust linear regression performed over the distribution: it is not intended to fit the data, but just to graphically render the parameters' correlation. This notation is followed within all next subfigures. (b) Joint spatio-temporal uncertainties distribution. (c) RF sizes vs. preferred stimulus velocity. A strong direct proportionality seems to drive these two parameters' distribution. (d) RF profile duration vs. preferred stimulus velocity. A significant negative correlation is found. (Two columns)

Dividing temporal frequency (ω_0) by spatial frequency (ξ_0), one can determine a cell's preferred stimulus velocity, which I will call v_0. This is true for both separable and inseparable profiles, the only difference being in spatio-temporally separable profiles responding with approximately the same strength for the two possible stimulus movement directions. In fact, if a stimulus were a dark or a bright bar moving in time, in the same representation used to create X-T plots this would be viewed as skewed stripes, which would have width and orientation related to its spatial and temporal frequency. In the spatio-temporal Fourier plane, the preferred velocity of a cell corresponds exactly to the slope of the line passing through the origin and intersecting the point that has coordinates given by the neuron's preferred spatio-temporal frequencies.

The relationship between preferred velocities and spatial uncertainty is the one that presented the strongest correlation coefficient. Fig. 1.5c illustrates a regular direct proportionality between RF sizes (along the axis perpendicular to the cell's preferred direction) and preferred stimulus velocity. This means that cortical cells that have the largest RFs are the ones whose task is to process the visual events that have the highest stimulus velocity, and this scheme appears to stand for every kind of cell, complex or simple, that has separable or inseparable profiles, in V1. Fig.1.5d, shows velocity tuning plotted against profile duration: in this case the correlation is not as strong as before, although it still presents a significant value. Nevertheless, a simple assertion can be made: there were no cells in the data sets studied that were tuned for detecting high stimulus velocities and whose response duration was within the highest values found.

1.3.2   Spatio-temporal uncertainty

In the spatio-temporal Fourier plane shown in Fig. 1.6, it is possible to define velocities as the ensemble of points in the plane satisfying the relation

$$v_0 = \frac{\omega_0}{\xi_0}. \tag{1.7}$$

This is true for every line passing through the origin, which thus can be associated with velocity values. For simplicity, I will refer to them as iso-velocity lines. As the 3D Gabor model I have used to approximate the receptive profiles has spatial and temporal frequency axes as principal axes, it is not possible to calculate a velocity bandwidth value by just combining the parameters defining spatial and temporal variances. Nevertheless, some tentative predictions about the shape and dimensions of the RFs in the data set can be made: the parameter distribution that has been previously examined limits the possible shapes and positions over the spectral spatio-temporal plane through which a primary visual cortex neuron's receptive profile can be modeled. While the spatial versus temporal frequency distribution seen in Fig. 1.5a constrains center position, Figg. 1.5c and 1.5d show that a higher iso-velocity line angular coefficient (a greater velocity value) corresponds to wider temporal frequency and narrower spatial frequency variances (uncertainties).

Another way to visualize the spatio-temporal organization of cortical cells is to see the RFs as filters covering the spatio-temporal frequency space. Fig. 1.7 shows the scatter plot of the responses of the cells belonging to the data set studied. Here we can clearly see how the spatio-temporal Fourier plane is far from being completely spanned by the RFs. Instead, a triangular shaped pattern emerges so that stimuli carrying both high spatial and high temporal frequencies, together with those that have extremely high or extremely slow velocities, are lacking "ded-
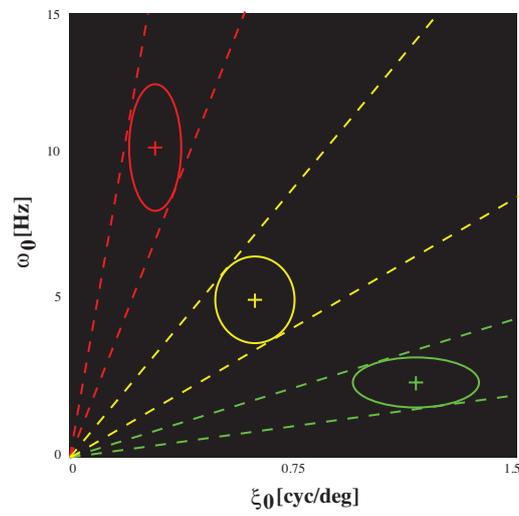
Figure 1.6: Sensitivity ellipses representing three sample Gabor functions in the spatio-temporal Fourier plane. Uncertainty over stimulus velocity measure can be defined as the difference between the angular coefficient of the two iso-velocity lines tangent to these ellipses.

icated" sensory neurons and cannot therefore be perceived. This fact will be discussed further in the next section.

As with spatial and temporal frequencies, one can also calculate the bandwidth for a cell's velocity tuning, that is, uncertainty over velocity measurement. To find velocity measurement uncertainties, I construct an ellipse for every cell, centered at the coordinates (ξ_0, ω_0) indicated by its spatio-temporal frequency sensitivity having $\Delta\xi$ and $\Delta\omega$ as its semi-axes. This would be nothing else but a contour line of the Gaussian function representing the cell's frequency response amplitude. Then, I define the spatio-temporal uncertainty Δv of a given neuron as the difference between the velocity values associated to the iso-velocity lines tangent to this sensitivity ellipse. Stimuli that are capable of significantly influencing a neuron's firing rate should be moving at a velocity within $v_0 \pm \Delta v/2$: this quantity measures the precision with which a neuron locates stimulus velocities. It is important to note that this definition of uncertainty is not formally related to an uncertainty principle, like Δx and Δt. Here I define the uncertainty on the measured velocity as the fluctuations that are associated to this derived quantity when the parameters ξ_0 and ω_0 are allowed to have fluctuations within one standard deviation.

It is worth considering the effects on cortical information processing that these relations imply. For every modeled cell in the data set, I take the four Gabor kernel parameters ξ_0, ω_0, Δx, Δt, that uniquely identify its receptive profile's position over the spectral spatio-temporal plane. Then for every receptive profile, I maintain its spatial and temporal frequency values and define two new variables, the spectral sensitivity area

$$S_A = \Delta\xi\Delta\omega, \tag{1.8}$$

and a positive shape or eccentricity index for the Gaussian sensitivity function:

$$\epsilon = \frac{\Delta\xi}{\Delta\omega}. \tag{1.9}$$

The velocity uncertainty value Δv, following the previous definitions, will be:

$$\Delta v = 2\sqrt{S_A} \frac{\sqrt{\frac{1}{\epsilon}\xi_0^2 + \epsilon\omega_0^2 - S_A}}{\xi_0^2 - \epsilon S_A}. \tag{1.10}$$

If for every receptive profile I fix spatial frequency, temporal frequency, and spectral sensitivity area, it is possible to show how its uncertainty over velocity measurement varies with eccentricity. In this way, I aim to perform a shape analysis over the ellipses, as opposed to a dimensional analysis that could be made by fixing the eccentricities and allowing the areas to vary, which is not interesting for the purposes of this analysis.

Figure 1.7: A scatter plot of the spatio-temporal sensitivity ellipses associated to the RFs in the data set. The responses seem to be concentrated at mid-range frequency values. However, it is possible to note how their shapes are not randomly distributed: vertically elongated ones clearly prefer higher velocity values.

The results for some of the cells are shown in Fig. 1.3, where a given function is associated to each profile (the original data and the Gabor fit are shown). These functions always present an "optimal shape" where the uncertainty over stimulus velocity measurement reaches a minimum value. Circles are plotted in correspondence to the eccentricity value indicating the actual spatio-temporal shape of the

profile. These graphs clearly show receptive profiles always placing themselves near to the theoretical minimum value, if not exactly over it, seemingly presenting an impulse response of optimal size and duration - given their position in the spatio-temporal spectral plane - and thus capable of detecting moving stimuli's velocities with the maximum precision possible.

This analysis shows V1 as having an efficient organization exploiting as far as possible all of the studied degrees of freedom. Spatio-temporal uncertainty is maintained at a low level within the visual cortex by assigning the detection of higher velocities to the cells that have the largest RFs, and smaller spatial frequency bandwidths, while lower velocities are generally detected by cells with a longer response duration and a narrower temporal frequency bandwidth. This confirms the fact that V1 tends to resolve physiologically, in the most efficient way, some kind of uncertainty regarding both spatial and temporal domains so that the possibility of drawing a general spatio-temporal cortical organization of visual neurons can begin to be appreciated.

Indeed, the keyword that seems to best describe the efforts of the brain in processing visual information is "efficiency": from the Gabor-like spatio-temporal impulse responses to the physiological trade-offs between spatial and temporal measured frequencies and spatial and temporal sizes, V1 is designed to offer the most accurate representation possible of the time-varying sequences of retinal images to the next stages in the visual pathway, allowing the brain a "window" through which to interact with the surrounding physical world. The fact that this is done by taking into account simultaneously both spatial and temporal retinal information indicates that a primary raw work of movement detection is operated by V1, whose conventionally accepted main role in visual processing regards resolving stimulus complexity over the spatial domain; and that this is worked out with the maximum precision possible.

Finally, recalling Figure 1.7 it can be observed how the spatio-temporal frequency space is only partially spanned by the responses of the cortical RFs. In particular, the triangular responsive shape found analyzing the data set is consistent with the "blunted diamond" shape of the spatio-temporal window of visibility found in psychophysical experiments and described by Watson et al [130]. The results suggest that the parceling of this window with spatio-temporally oriented Gabor kernels is accomplished and driven by accuracy-based spatio-temporal organization.

Chapter 2

Models of functional architecture in the visual cortex

2.1 Introduction

The modular structure of the mammalian visual cortex has been discovered in the seventhies by the pioneeristic work of Hubel and Wiesel [52, 51]. Many families of cells contribute to sampling and coding the stimulus image. Each family is sensible to a specific feature of the image: position, orientation, scale, color, curvature, velocity, stereo.

The main behaviour of simple cells is that of detection of positions and local orientations via linear filtering of the stimulus, and the linear filter associated to a given cell is called its receptive profile. Daugman [24] proposed a criterion of minimal uncertainty for the shape receptive profile, which results in two dimensional spatial Gabor filters, that was later confirmed by Jones and Palmer [54] and Ringach [94]. Simple cells also possess a temporal behaviour, studied by De Angelis et al. [25], and spatio-temporal receptive profiles could again be interpreted in Chapter 1 in terms of minimal uncertainty, resulting in three dimensional Gabor

filters. However, not all cell activity in response to stumuli can be justified in terms of linear filters, and nonlinearities can sometimes become relevant to model their behaviours [41].

In this chapter I propose a mathematical model of cortical functional architecture for the processing of spatio-temporal visual information, that is compatible with both phenomenological experiments and neurophysiological findings. The adopted theoretical framework follows the outlined path of mathematical models of the activity of the visual cortex, and in particular it continues the geometric approach of Citti and Sarti [19] to a space of higher dimension, since it takes into account time and velocity of stimuli. I will extend the stochastic process of Mumford [73] to this setting, working in a space of liftings arising from the filtering with spatio-temporal receptive profiles, and make use of assumptions like the ones of Ermentrout and Cowan [32] in order to construct a population dynamics able to provide a new form of association fields adapted to the problem of motion integration and motion completion under occlusion. Moreover, the resulting kernels will be comparable to measured neural activities in the presence of stimuli characterized by their direction of motion.

The starting point in Section 3.2 is a process of detection resulting from linear filtering with three dimensional Gabor functions with two spatial and one temporal dimensions, which have been proposed as a model of spatio-temporal receptive profiles of primary visual cortex simple cells (see [25], [20]). Spatio-temporal Gabor filters extend simple cells Gabor behavior as spatial filters [24, 94], that proved its usefulness for the classical task of edge detection, while its role for motion detection was already pointed out in [90]. The Gabor transform takes in input a moving image $f(x, t)$, where $x \in \mathbb{R}^2$ are spatial coordinates and $t \in \mathbb{R}$ is time and provides in output a representation of the signal in the phase space

$$f(x, t) \to F(q, s; p, \nu),$$

with $q \in \mathbb{R}^2$ and $s \in \mathbb{R}$ representing spatio-temporal position and (p, ν) representing, respectively, spatial and temporal frequency.

Since I am mainly interested in spatio-temporal dynamical aspects, I will assign to temporal frequency the meaning of velocity v on the (x, t) coordinates. I will also select the subset of all detected features corresponding to a fixed value of $|p|$. This will end up to be a 5D manifold, with a contact structure, induced by a normalization of the Liuoville form.

In Section 2.3 I will outline that this constraint carries a notion of admissible curves [37] in a deterministic and a stochastic setting, allowing to compute the kernels connecting filters in the 5D manifold. The stochastic processes are completely determined by the described structures of the tangent space of the 5D manifold and they turn out to be described by the fundamental solution of a Fokker Planck equation [80]. In fact the processes contain diffusion in the fiber variables and transport along the remaining generators of admissible tangent directions, in the spirit of [73, 136]. I will compute these kernels in two limit cases: the motion of a contour at a fixed time instant, and the motion of a point moving in time. The first one reduces to the geometry of contours, with a notion of instantaneous velocity, the second one corresponds to point trajectories in time, and can be related to the subset of the Galilei group [114] on the plane.

In Section 2.4 I will discuss the compatibility of the previously calculated kernels with psychophysiological findings reported in the recent literature. Then I will insert the connectivity kernels computed in Section 2.3 in a neural population activity model [32], regarding them as cortical facilitation patterns.

In section 2.5 I will use this population activity model equipped with the suitable connectivity kernels in two numerical simulations, comparing the results to recent phenomenological findings of the perception of contours in motion [92], and

to fMRI measurements of cortical neural activity related to motion perception [138].

2.2   The geometry of spatio-temporal dynamics

In this section I extend an approach introduced in [19] that amounts to model each V1 simple cell in terms of its receptive profile, to interpret its action as a Gabor filtering, and to introduce a geometry of the space compatible with the properties of the output. In this chapter I will consider each cell as sensitive to a local orientation and apparent velocity, that is the velocity orthogonal to a moving stimulus. The collected data lies on a five dimensional manifold $\mathcal{M}$ of space, time, orientation and velocity, a single cell being represented as a point on $\mathcal{M}$. The geometric structure of this manifold will be described in terms of a contact structure, that provides a constraint on the tangent space and on admissible connectivity among cells.

2.2.1   Spatio-temporal receptive profiles

It is known that the visual cortex decomposes the visual stimulus by measuring its local features. Local orientation and direction of movement have been the first visual features of neurons in V1 that have been studied. Receptive profiles (RPs) are descriptors of the linear filtering behavior of a cell and they can be reconstructed by processing electrophysiological recordings [94]. It has been shown that the spatial characteristics of these RPs can be modeled by 2-dimensional Gabor functions [24, 54].

However a large class of cells shows a very specific space-time behavior in which the spatial phase of the RP changes gradually as a function of time [25]. Although many models have been proposed to reproduce these dynamics [25, 1], in [20] it has been shown that a 3-dimensional sum-of-inseparable Gabor model
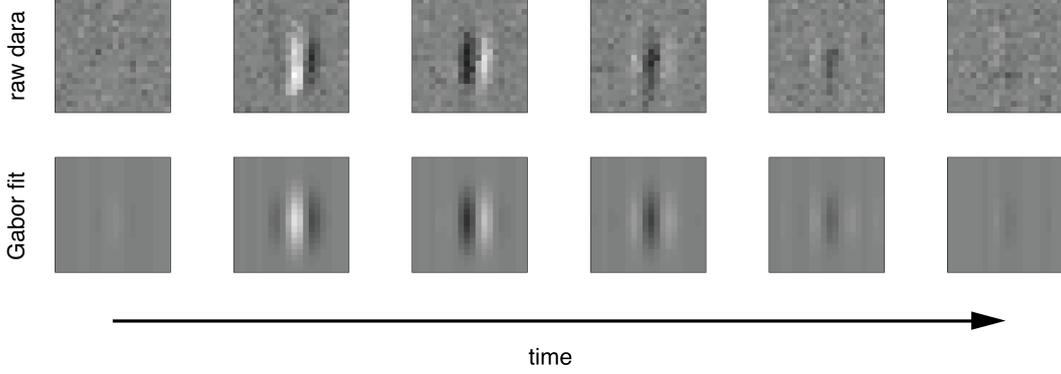
Figure 2.1: A spatio-temporal receptive profile taken from the data set studied in [25] and [20] (cell n.15 in the latter reference). In the top row, the sequence relative to the raw data, reconstructed using reverse correlation from electrophysiological recordings. In the bottom row, the 3-dimensional Gabor fit, visualized as 2D slices.

can fit very well experimental data of both separable and inseparable RPs (Fig. 2.1). Following this approach, I choose Gaussian Gabor filters centered at position $q = (q_1, q_2) \in \mathbb{R}^2$ on the image plane, activated around time $s \in \mathbb{R}$, with spatial frequency $p = (p_1, p_2) \in \mathbb{R}^2$, temporal frequency $\nu \in \mathbb{R}$, spatial width $\sigma_x \in \mathbb{R}^+$ (circular gaussians) and temporal width $\sigma_t \in \mathbb{R}^+$

$$\psi_z^\sigma(x, t) = e^{i(p \cdot (x-q) - \nu(t-s))} e^{-\frac{|x-q|^2}{\sigma_x^2} - \frac{(t-s)^2}{\sigma_t^2}} \tag{2.1}$$

where I have used the abbreviations $z = (q_1, q_2, s) + i(p_1, p_2, \nu) \in \mathbb{C}^3$ and $\sigma = (\sigma_x, \sigma_t) \in \mathbb{R}^+ \times \mathbb{R}^+$. Also note that the functions (2.1) correspond to the propagation of a two dimensional plane wave within the activating window with phase velocity
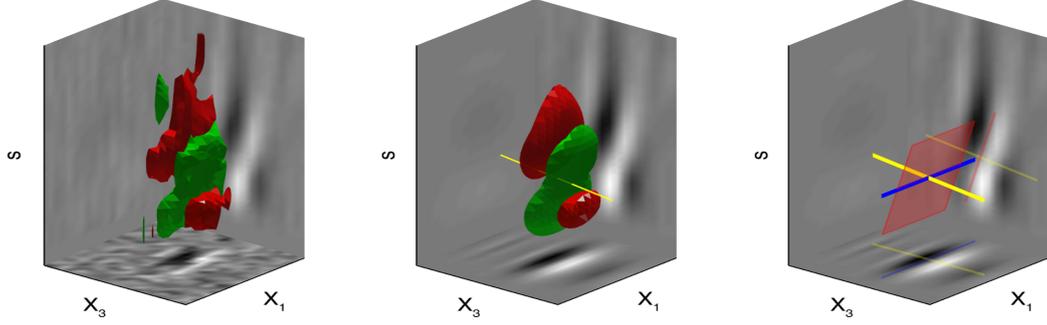
$$v = \frac{\nu}{|p|}. \tag{2.2}$$

Figure 2.2: Visualization of the one-form ω with respect to a cell RP. In the left and center subplots are shown the isosurfaces of the RP already seen in Fig. 2.1. In the left plot I visualize the reconstructed data, while in the middle plot I visualize its Gabor fit. The orientation of the yellow line within the fitted data represents the orientation of the wavefront of the Gabor filter and coincides with the one-form $\omega = \cos\theta dq_1 + \sin\theta dq_2 - vds$, or the dual vector $\vec{X}$. In the right plot the semi-opaque red plane represents the horizontal tangent plane in the R^3 spatio-temporal space, generated by $\vec{X_5}$ and $\vec{X_1}$ (blue line).

The variable $z = (q_1, q_2, s) + i(p_1, p_2, \nu)$ is canonically associated [35] to the phase space

$$\mathbb{R}^6 = \{(q_1, q_2, s, p_1, p_2, \nu)\} \tag{2.3}$$

endowed with the symplectic structure compatible with the complex structure of $\mathbb{C}^3$ of variable z, that is

$$\Omega = d\lambda = dp_1 \wedge dq_1 + dp_2 \wedge dq_2 - d\nu \wedge ds$$

where I have denoted with λ the Liouville form

$$\lambda = p_1 dq_1 + p_2 dq_2 - \nu ds. \tag{2.4}$$

Since I am mainly interested in the geometry of level lines, the analysis can be restricted to the information captured by filters possessing a fixed central spatial

frequency $|p|$. This amounts to fixing a scale of oscillations, hence disregarding the harmonic content of the filtering by focusing on the features of orientations and velocity. We then obtain the reduced 1-form

$$\omega = \frac{\lambda}{|p|} = \cos\theta dq_1 + \sin\theta dq_2 - vds \tag{2.5}$$

where $p = |p|(\cos\theta, \sin\theta)$. This form is defined on the spatio-temporal phase space with fixed frequency, that is the 5-dimensional manifold

$$\mathcal{M} = \mathbb{R}^2 \times \mathbb{R}^+ \times S^1 \times \mathbb{R}^+ = \{\eta = (q_1, q_2, s, \theta, v)\},$$

and it is associated to every Gabor filter as shown in Fig. 3.1. At this level, time s is introduced as a base variable with the same role of $\binom{q_1}{q_2}$. The dual variable of s is v, which has the same role of θ, both being the engrafted variables with respect to time and space. To every point of $\mathcal{M}$ corresponds univocally a Gabor filter whose parameters are the coordinates of the point itself.

2.2.2 Admissible tangent space as constraint on the connectivity on $\mathcal{M}$

I will model the connectivity between points in the space $\mathcal{M}$ in terms of admissible tangent directions of $\mathcal{M}$ itself. From the geometric point of view, the presence of the 1-form (2.5) is equivalent to the choice of a vector field with the same coefficients as ω with respect to the canonical basis $\{\frac{\partial}{\partial q_1}, \frac{\partial}{\partial q_2}, \frac{\partial}{\partial s}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial v}\}$, that is its dual vector (see Fig. 3.1):

$$\vec{X} = (\cos\theta, \sin\theta, -v, 0, 0).$$

The kernel of ω, denoted $\ker\omega$ (or $\omega = 0$) is the space of vectors orthogonal to $\vec{X}$. A basis of this space is constituted of the so-called horizontal or admissible vectors

$$\begin{aligned} \vec{X}_1 &= (-\sin\theta, \cos\theta, 0, 0, 0)\,, \ \ \vec{X}_2 = (0, 0, 0, 1, 0) \\ \vec{X}_4 &= (0, 0, 0, 0, 1)\,, \ \ \vec{X}_5 = (v\cos\theta, v\sin\theta, 1, 0, 0) \end{aligned} \tag{2.6}$$

so that

$$\ker\omega = \mathrm{span}\{\vec{X}_1, \vec{X}_2, \vec{X}_4, \vec{X}_5\} \tag{2.7}$$

that defines the *horizontal tangent space*. It is worth noting that the Euclidean metric on the horizontal planes makes the vector fields X_i orthogonal.

The only manifolds (curves or surfaces) admissible in this space are the ones whose tangent vectors are linear combinations of the horizontal vectors (3.3). The neural connectivity between the receptive fields at different points in the $\mathcal{M}$ space will be defined in terms of these vectors in Section 2.3.

For reader convenience I compute here explicitly all the non-zero commutation relations between the vector fields [1]:

$$\begin{aligned}
[\vec{X}_1, \vec{X}_2] &= \vec{X}_3 &&= (\cos\theta, \sin\theta, 0, 0, 0) \tag{2.8}\\
[\vec{X}_2, \vec{X}_3] &= \vec{X}_1 \\
[\vec{X}_4, \vec{X}_5] &= \vec{X}_3 \tag{2.9}\\
[\vec{X}_2, \vec{X}_5] &= v\vec{X}_1.
\end{aligned}$$

In Fig. 2.3 I have depicted the structure of the tangent fields (3.3). For visualization purposes, I show the fields in two different representations. In Fig. 2.3a

---

[1] To every vector field $\vec{X} = (a_1, a_2, a_3, a_4, a_5)$ I can associate a directional derivative

$$X = a_1\partial_{q_1} + a_2\partial_{q_1} + a_3\partial_s + a_4\partial_\theta + a_5\partial_v$$

with the same coefficients. Then I will call commutator of X and Y

$$[X, Y] = XY - YX.$$

I say that X and Y commute if $[X, Y] = 0$. Note that partial derivatives always commute, while directional derivatives in general do not. It is important to note that, even though the commutator is expressed formally as a second derivative, it is indeed a first derivative, so that it has an associated vector. Hence I can as well define the commutator between the vectors $\vec{X}$ and $\vec{Y}$ as the vector associated to $XY - YX$.

Figure 2.3: Visualization of the admissible tangent space of $\mathcal{M}$, generated by the vectors $\vec{X}_1, \vec{X}_2, \vec{X}_4, \vec{X}_5$. In a) I show the projection in the (q, θ) space. Red planes are the sections of the horizontal planes, generated by $\vec{X}_1$ and $\vec{X}_2$. Red arrows indicate the $\vec{X}_1$ direction, black arrows the $v\vec{X}_3$ one. In b) I depict the projection in the (x_3, s, v) space. The blue planes are the sections of the horizontal planes, spanned by $\vec{X}_4$ and $\vec{X}_5$ The black arrows represents the $v\vec{X}_3$ as before.

I visualize the structure restricted to the spatial and engrafted variables (q, θ, v), spanned by $\vec{X}_1$ and $\vec{X}_2$. The tilting of the planes is due to the non commutative relation (2.8). In Fig 2.3b I show the spatio-temporal structure restricted to the variables (x_3, s, v), where $x_3 = q_1 \cos\theta + q_2 \sin\theta$. Also in this figure the tilting of the planes spanned by $\vec{X}_4$ and $\vec{X}_5$ is due to the non vanishing commutator condition (2.9). It is worth noting that in this setting, the propagation is permitted only along the horizontal planes, while it is forbidden in their orthogonal direction $\vec{X}$. Let us note that $\vec{X}_3 = [\vec{X}_1, \vec{X}_2]$ is linearly independent of the horizontal tangent space at every point. Hence $\{\vec{X}_1, \vec{X}_2, \vec{X}_4, \vec{X}_5\}$ together with their commutators span the whole space at every point. This is the so called Hörmander condition, which will have a crucial role in studying the property of the space. The same condition can be expressed in terms of properties of the form ω, which is called *contact form* [37] since $\omega \wedge d\omega$ is never zero being the volume form of the space.

I also note that I can define on $\mathcal{M}$ a smooth composition law

$$
\begin{aligned}
&(q, s, \theta, v) \odot (q', s', \theta', v') \\
&= (R_\theta(q' + \tbinom{v}{0}s') + q, s' + s, \theta' + \theta, v' + v)
\end{aligned}
\tag{2.10}
$$

where R_θ is a counterclockwise rotation of an angle θ, that is such that the vector fields $\{\vec{X}_1, \vec{X}_2, \vec{X}_3, \vec{X}_4, \vec{X}_5\}$ are left invariant with respect to this law. This implies that also the admissible curves of the structrue and its kernels will be invariant. The manifold $\mathcal{M}$ together with this composition law can also be identified with a subset of the Galilei group (see Appendix).

2.2.3 The output of the receptive profiles

The output of the Gabor filters selects a set of points in $\mathcal{M}$ corresponding to specific features of the image. Since the Gabor filters are always connected in terms of the vectors (3.3), also their output will inherit this structure, and will be concentrated around an admissible surface.

The energy output of a cell with receptive profile ψ_z^σ in presence of a spatio-temporal stimulus $f(x,t)$ is given by

$$
\begin{aligned}
F(q,s,\theta,v) &\doteq |\langle \psi_z^\sigma, f \rangle|^2 \\
&= \left| \int_{\mathbb{R}^3} dx\, dt\, e^{-i|p|((x-q)_1 \cos\theta + (x-q)_2 \sin\theta - v(t-s))} \right. \\
&\qquad \left. e^{-\frac{|x-q|^2}{\sigma_x^2} - \frac{(t-s)^2}{\sigma_t^2}} f(x,t) \right|^2 .
\end{aligned}
\tag{2.11}
$$

I note that taking the square modulus in (3.2) disregards the phase of the corresponding linear filtering, which for many applications (see e.g. [30]) is crucial. In this case however all the relevant information is encoded in this energy model. In particular the geometric quantities θ and v are encoded in the points of maxima of the energy. To see why, I recall that the analogous of formula (3.2) with purely spatial Gabor filters was studied in [19], where the filtering output F_0 was a function of the variables $\binom{q_1}{q_2}$ and θ alone (see also [47]). In order to outline the goemetric meaning of the lifting, I consider its action on level lines which are smooth. In that case, the output F_0 to a stimulus $f_0(x)$ takes its maximum around a value $\theta^* = \theta^*(\binom{q_1}{q_2})$

$$
\max_\theta F_0(\binom{q_1}{q_2}, \theta) = F_0(\binom{q_1}{q_2}, \theta^*)
$$

where $(\cos\theta^*, \sin\theta^*)$ identifies the orientation orthogonal to the level lines of $f_0(x)$. As a consequence, in [19] it is proved that the level lines of f_0 are lifted to curves admissible in the sense that their tangent vector lies in the kernel of the form

$$
\omega_3 = \cos\theta dq_1 + \sin\theta dq_2
$$

which is a contact form that can be obtained by restricting (2.5) to the space $(\binom{q_1}{q_2}, \theta)$.

Since here I take into account time, then the output F is a function defined on $\mathcal{M}$. In perfect analogy with the lower dimensional case, and denoting with $(\binom{q_1}{q_2}, s(\binom{q_1}{q_2}))$ a level set of the stimulus $f(x,t)$, the output F takes its maximum

around the values $\theta^* = \theta^*(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right), s(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right)))$ and $v^* = v^*(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right), s(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right)))$

$$\max_{\theta, v} F(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right), s(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right)), \theta, v) = F(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right), s(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right)), \theta^*, v^*)$$

such that the vector $(\cos\theta^*, \sin\theta^*, v^*)$ is orthogonal to the level set of f. The vector $(\cos\theta^*, \sin\theta^*)$ is orthogonal to the spatial level line, and the scalar v^* represents the apparent velocity in this direction.

The surface

$$\Sigma = \left\{ \left(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right), s(\left(\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}\right)), \theta^*, v^*\right) \right\}$$

is the 5D lifting of the level set of f, and the orthogonality condition implies that it is admissible, in the sense that its tangent vectors lie in the horizontal space, kernel of the form ω (see Fig. 2.4).

This shows that such level sets of f define a contact structure on the manifold, so the study of such geometric features induces to endow $\mathcal{M}$ with the corresponding constraint on the tangent bundle that can be equivalently interpreted as a sub-Riemannian constraint [72, 19, 10].

2.3   Curves and kernels of connectivity

The lifted points of the spatio-temporal stimulus are connected by admissible integral curves. However I will see that not all admissible curves can be considered lifted ones, and I will describe which ones have this property. These curves will represent the association fields in space-time, analogously to the association fields of Field Hayes and Hess[34] in the pure spatial case, and are reminiscent of the classical Gestalt concept of *good continuation*. I will discuss in Section 2.4 the compatibility of this structure with the cortical functional architecture. A probabilistic version of the connectivity field will be provided on the basis of Fokker Planck equations first introduced by D. Mumford in the spatial case [73], and interpreted as
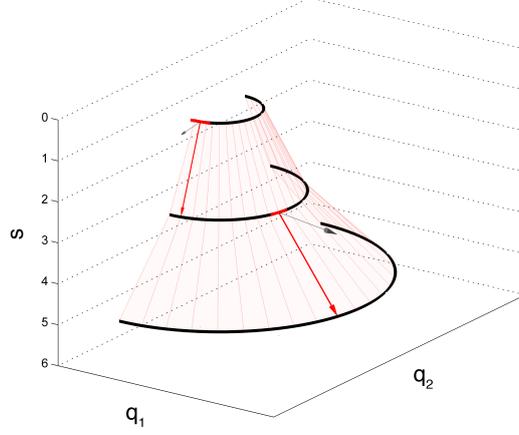
Figure 2.4: A schematical representation of the level set of a function f (in this case, an expanding contour) lifted as an admissible surface Σ on the 5D structure. Red and black arrows indicate, respectively, the local direction of the vector fields $\vec{X}_5$ and $\vec{X}_3$ over two points in space-time.

model of connectivity in Lie groups in several works such as [89, 142, 7, 28, 29, 106]. Here I will extend this stochastic approach to the space-time contact structure.

2.3.1   Generators of lifted curves

We have already seen that the output of RP filtering is concentrated around admissible submanifolds. Nevertheless not all admissible submanifolds can be lifting of images, since lifting are graphs of the functions θ^* and v^*. For example the plane generated by the vectors $\vec{X}_2$ and $\vec{X}_4$ cannot be recovered by lifting. Hence I will study the possible linear combinations of vector fields $\{\vec{X}_i\}$, which can be tangent to lifted level lines. An ω-admissible curve $\gamma \subset \mathcal{M}$ is characterized by $\dot{\gamma} \in \ker \omega$, i.e.

$$\dot{\gamma} = a_1 \vec{X}_1(\gamma) + a_2 \vec{X}_2(\gamma) + a_4 \vec{X}_4(\gamma) + a_5 \vec{X}_5(\gamma)$$

with a_i not necessarily constant. Lifted curves depend on the vectors $\vec{X}_1$ or $\vec{X}_5$, which are tangent to the base space (q, s), and their linear combination. To simplify

the problem, I will consider separately two special cases of particlar interest for the model: the limit cases of contour motion detected at a fixed time ($a_5 = 0$), and motion of a point in time ($a_1 = 0$). The first one is described by integral curves of the vector $\vec{X}_1$ and of the generators of the engrafted variables $\vec{X}_2$ and $\vec{X}_4$:

$$\dot{\gamma}_3 = \vec{X}_1(\gamma) + k\vec{X}_2(\gamma) + c\vec{X}_4(\gamma). \tag{2.12}$$

These curves lie in the section of the contact structure depicted in Figure 2.3a, where k represents Euclidean curvature and c is the rate of change of local velocity along the curve. On the other hand, the motion of a point in time can be described as

$$\dot{\gamma}_T = \vec{X}_5(\gamma) + w\vec{X}_2(\gamma) + a\vec{X}_4(\gamma) \tag{2.13}$$

that are suitable to describe spatio-temporal trajectories of points. The coefficient w of the direction $\vec{X}_2$ is the angular velocity, and the coefficient a of $\vec{X}_4$ is the tangential acceleration.

2.3.2   Curves and kernels for contours in motion

Moving contours as deterministic integral curves

I consider now the geometry of moving contours at a fixed time. These are generated by the integral curves (2.12) so to satisfy the system of equations

$$\begin{cases} \dot{\gamma}(t) &= \vec{X}_1(\gamma) + k\vec{X}_2(\gamma) + c\vec{X}_4(\gamma) \\ \gamma(0) &= \xi_0 \end{cases} \tag{2.14}$$

with variable coefficients k, c. In particular, due to condition (2.8) the set of points reacheable by piecewise constant integral curves of this type (for an explicit expression see e.g. [19]) starting from a fixed point $\xi_0 = (q_0, s_0, \theta_0, v_0)$ is the space

$$\mathcal{N} = \{((\begin{smallmatrix} q_1 \\ q_2 \end{smallmatrix}), s_0, \theta, v)\} \tag{2.15}$$
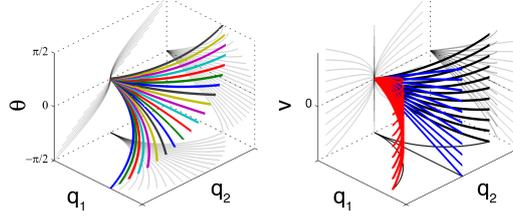
Figure 2.5: The subset of horizontal curves with constant coefficients pertaining to $\Sigma_0(\xi_0)$ by varying parameters k and c, with initial velocity set to 0. Left plot: projections over the variables (q_1, q_2, θ) with different values for k, and same non-null value for c. Right plot: projections over the variables (q_1, q_2, v) relative to some negative and positive values of c. The curves with $k = 0$ (blue ones in the right plot) are straight lines, as the direction of spatial propagation never changes.

where s_0 is fixed, so the space $\mathcal{N}$ can be identified with the space of points $\xi = (\binom{q_1}{q_2}, \theta, v)$.

I also denote with $\Sigma_0(\xi_0)$ the set of points reached by the fan of curves solution to the system (2.14) with constant coefficients (k, c), depicted in Fig. 2.5.

As I did with (2.10) I can define on $\mathcal{N}$ a smooth composition law

$$(q, \theta, v) \boxdot (q', \theta', v') = (R_\theta q' + q, \theta' + \theta, v' + v) \tag{2.16}$$

where R_θ is a counterclockwise rotation of an angle θ, that is such that the vector fields $\{\vec{X}_1, \vec{X}_2, \vec{X}_3, \vec{X}_4\}$ are left invariant with respect to this law. Since this is a group law, the manifold $\mathcal{N}$ together with this (2.16) provides the Lie group $\mathbb{R}^2 \ltimes S^1 \times \mathbb{R}$, that is the direct product of the $SE(2)$ group and the group of the reals with addition.

Stochastic kernel

Let us consider a probabilistic counter part of trajectories (2.14). The vector field X_1 expresses a derivative in the direction of a variable coded on the retina, while

the vectors X_2, X_4 express derivatives in the direction of an eingrafted variable. Due to their different role, I will consider the following vector-valued stochastic process

$$\begin{cases} d\gamma & = & \vec{X}_1(\gamma)dt + \vec{X}_2(\gamma)dW_1 + \vec{X}_4(\gamma)dW_2 \\ \gamma(0) & = & \xi_0 \end{cases} \qquad (2.17)$$

where $W = (W_1, W_2)$ is a two dimensional Brownian motion. The distribution of these stochastic curves is mostly concentrated around the surface $\Sigma_0(\xi_0)$. The density of points reached by this stochastic kernel is then a candidate to implement the mechanism of association fields (see Fig. 2.6). This approach generalizes the approach of random paths introduced in [73, 136] for the $SE(2)$ problem.

If I call $\rho_{\mathfrak{g}^v}(\xi, t | \xi_0, 0)$ the density of points of $\mathcal{N}$ reached at the value t of the evolution parameter by the sample paths of the process (2.17), then $\rho_{\mathfrak{g}^v}$ can be obtained as the fundamental solution

$$(\partial_t + L_{\mathfrak{g}_v})\rho_{\mathfrak{g}^v}(\xi, t | \xi_0, 0) = \delta(\xi - \xi_0)\delta(t)$$

where $L_{\mathfrak{g}_v}$ is the Fokker-Planck operator

$$L_{\mathfrak{g}_v} \doteq X_1 - \kappa^2 X_2^2 - \alpha^2 X_4^2 \qquad (2.18)$$

containing a diffusion over the fiber variables θ and v and a drift in the base variables, and with X_i and X_i^2 I denote first and second order derivatives in direction $\vec{X}_i$. Since $L_{\mathfrak{g}_v}$ contains a set of vector fields that generates the whole tangent space of $\mathcal{N}$, by Hörmander theorem this operator is hypoelliptic [50]. This means that $\rho_{\mathfrak{g}^v}$ is non-null in all $\mathcal{N}$ for any $t > 0$, even if the operator $L_{\mathfrak{g}_v}$ contains only 3 linearly independent fields. Moreover, since it defines the Fokker-Planck equation for the stochastic process (2.17), $\rho_{\mathfrak{g}^v}$ is indeed a (conditional) probability density on $\mathcal{N}$, evolving with the parameter t.

In order to characterize each point of $\mathcal{N}$ in terms of the density of paths (2.17) that reach it, independently of the value of the evolution parameter, we need a
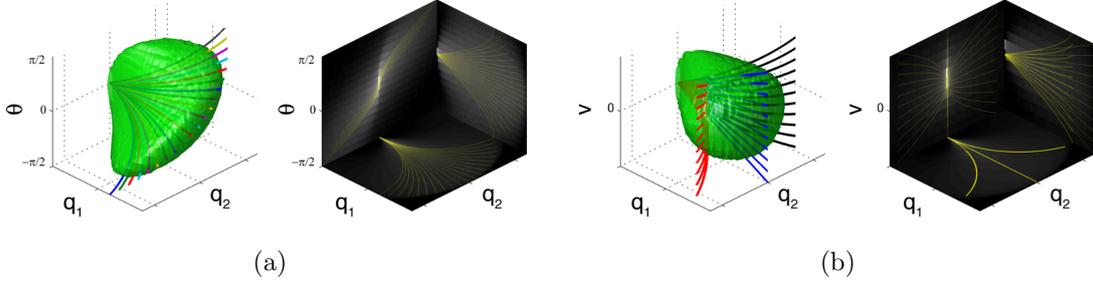
Figure 2.6: Horizontal curves and stochastic kernels for the integration of contours in motion. Left: the isosurface plot of the kernel $\Gamma_{\mathfrak{g}^v}$ (isovalue: $0.002 \max(\Gamma_{\mathfrak{g}^v})$) is superimposed over the subsets of curves (2.14) with constant coefficients already shown in Fig. 2.5. Right: the kernel projections of $\Gamma_{\mathfrak{g}^v}$ relative to the variables a) (q_1, q_2, θ) and b) (q_1, q_2, v) are plotted in gray under the projections of the curves (in yellow). This kernel was obtained by calculating the evolution of 10^6 stochastic paths.

notion corresponding to the fan $\Sigma_0(\xi_0)$, expressed in terms of $\rho_{\mathfrak{g}^v}$. The density of points reached *at any value of the evolution parameter* by the stochastic dynamics (2.17) is given by

$$\Gamma_{\mathfrak{g}^v}(\xi|\xi_0) \doteq \int_{\mathbb{R}} \rho_{\mathfrak{g}^v}(\xi, t|\xi_0, 0)dt \ . \tag{2.19}$$

This derived quantity is actually the fundamental solution of the operator $L_{\mathfrak{g}_v}$, so explicitly we have

$$\left(X_1 - \left(\kappa^2 \partial_\theta^2 + \alpha^2 \partial_v^2 \right) \right) \Gamma_{\mathfrak{g}^v}(\xi|\xi_0) = \delta(\xi - \xi_0) \tag{2.20}$$

and since the vector fields involved in equation (2.18) are left invariant with respect to the group law (2.16), the solution (2.19) possesses the symmetry

$$\Gamma_{\mathfrak{g}^v}(\xi|\xi_0) = \Gamma_{\mathfrak{g}^v}(\xi_0^{-1} \boxdot \xi|0) \quad \forall \, \xi, \xi_0 \in \mathcal{N} \ . \tag{2.21}$$

I recall that in [31] several representations of the exact solution to a problem analogous to (2.20) were presented. In this work I will deal with the numerical

implementation of the fundamental solution $\Gamma_{\mathfrak{g}^v}(\xi|\xi_0)$, developed with standard Markov Chain Monte Carlo methods (MCMC) [101]. This is done by generating random paths obtained from numerical solutions of the system (2.17) and averaging their passages over discrete volume elements, and appears suitable to treat also the more involved case of subsequent equation (2.26). An example is shown in Fig. 2.6, where an isosurface plot of the kernel is plotted over the integral curves (2.14) with constant coefficients, already depicted in Fig. 2.5. By comparing such numerical approximations with [31], I can confirm the accuracy of this method. Moreover, from the figure it can be seen that the probability density is concentrated around the surface $\Sigma_0(\xi_0)$ and decays rapidly away from it. This is reasonable since for this kind of hypoelliptic operators one can generally obtain estimates for the fundamental solution in terms of exponential decay with respect to a geodesic distance computed with respect to the minimal set of vector fields that, together with their commutators, span the whole tangent space [100]. Such a distance is anisotropic, and its balls are squeezed in the directions of the commutators [75]. Here the vector fields involved are $\{X_1, X_2, X_4\}$, hence the concentration of the fundamental solution around the set Σ_0 defined by their integral curves follows by the sharper decay in the direction of the commutators. I also note that in this particular case, due to the availability of exact solutions and estimates from [31], this behaviour can also be checked directly.

2.3.3 Curves and kernels for point trajectories

Point trajectories as deterministic integral curves

The trajectory-type curves introduced in (2.13) are solutions $\gamma : \mathbb{R} \to \mathcal{M}$ to the system of ordinary differential equations

$$
\begin{cases}
\dot{\gamma}(t) & = & \vec{X}_5(\gamma) + w\vec{X}_2(\gamma) + a\vec{X}_4(\gamma) \\
\gamma(0) & = & \eta_0
\end{cases}
\tag{2.22}
$$

for given initial point $\eta_0 = (q_{10}, q_{20}, s_0, \theta_0, v_0)$. In general the coefficients w, a need not to be constant, but we can have a local approximation of any curve in a neighborhood of the starting point if we consider the case of constant coefficients. In this model case when w and a are not zero, (2.22) is explicitly solved by

$$
\begin{cases}
q_1(t) & = & q_{10} + \varrho(\cos(\theta(t) - \phi) - \cos(\theta_0 - \phi)) + \dfrac{a}{w}t\sin\theta(t) \\[2mm]
q_2(t) & = & q_{20} + \varrho(\sin(\theta(t) - \phi) - \sin(\theta_0 - \phi)) - \dfrac{a}{w}t\cos\theta(t) \\[2mm]
s(t) & = & s_0 + t \\[2mm]
\theta(t) & = & \theta_0 + wt \\[2mm]
v(t) & = & v_0 + at
\end{cases}
$$

where $\varrho = \frac{\sqrt{a^2 + v_0^2 w^2}}{w^2}$ and $\phi = \arctan\frac{v_0 w}{a}$, and a reasonable choice is to set $s_0 = 0$, in order to synchronize the evolution parameter t with the time parameter s.

The fan $\Sigma(\eta_0)$ of such curves is depicted in Fig. 2.7. Each of them describes a motion on the plane $\binom{q_1}{q_2}$ that, for small times, consists approximately of arcs of circles with radius ϱ, while for sufficiently large times corresponds to enlarging spirals around a slightly moving center. Their instantaneous acceleration is given by $\binom{\ddot{q}_1}{\ddot{q}_2} = R_\theta\binom{a}{vw}$, so that a corresponds to the tangential acceleration[2] along

---

[2]By direct computation, a is the time derivative of the modulus of the velocity

$$
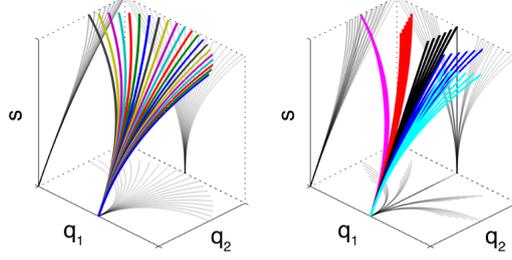a = \frac{d}{dt}\sqrt{\dot{q}_1(t)^2 + \dot{q}_2(t)^2} \ .
$$

Figure 2.7: Projections over the variables (q_1, q_2, s) of the fan of curves with constant coefficients, for different values of parameters w and a and a non-null initial velocity. Different values of w are associated to different curve colors, while gray is used for projections over two dimensional planes. Left plot: curves with different values of w, and a fixed non-null value for a. Right plot: curves with different values of a.

direction X_3:

$$a = \ddot{q}_1 \cos\theta + \ddot{q}_2 \sin\theta$$

while their curvature is

$$\frac{\dot{q}_1 \ddot{q}_2 - \dot{q}_2 \ddot{q}_1}{(\dot{q}_1{}^2 + \dot{q}_2{}^2)^{\frac{3}{2}}} = \frac{w}{v} \ .$$

In particular, when $w = 0$ I obtain straight lines along the θ_0 direction, while for $a = 0$ I obtain circular trajectories of radius $\dfrac{v_0}{w}$.

Stochastic kernel

Similarly to what I have done in section 2.3.2, let us consider the vector-valued stochastic process

$$\begin{cases} d\gamma & = & X_5(\gamma)dt + X_2(\gamma)dW_1 + X_4(\gamma)dW_2 \\ \gamma(0) & = & \eta_0 \end{cases} \qquad (2.23)$$

where $W = (W_1, W_2)$ is a two dimensional brownian motion.

The density $\rho_{\mathfrak{g}_0}(\eta, t | \eta_0, 0)$ of points of $\mathcal{M}$ reached at the value t of the evolution parameter by the sample paths of the process (2.23), is the fundamental solution of the equation

$$(\partial_t + L_{\mathfrak{g}_0})\rho_{\mathfrak{g}_0}(\eta, t | \eta_0, 0) = \delta(\eta - \eta_0)\delta(t) \tag{2.24}$$

where $L_{\mathfrak{g}_0}$ is the Fokker-Planck operator

$$L_{\mathfrak{g}_0} \doteq -\alpha^2 X_4^2 - \kappa^2 X_2^2 + X_5. \tag{2.25}$$

Equation (2.24) is not hypoelliptic, indeed its fundamental solution ρ is concentrated on a submanifold of codimension 1 defined by the equation $t = s$, from the system (2.23). However it is still a Fokker-Planck equation, hence ρ is nonnegative, and integrating the density $\rho_{\mathfrak{g}_0}$ with respect to the evolution parameter t I derive the fundamental solution of the operator $L_{\mathfrak{g}_0}$. Explicitly I have

$$\left(\partial_s - \left(\kappa^2 \partial_\theta^2 + \alpha^2 \partial_v^2 - v X_3\right)\right)\Gamma_{\mathfrak{g}_0}(\eta | \eta_0) = \delta(\eta - \eta_0) \tag{2.26}$$

and I note in particular that this equation is now a hypoelliptic equation that is also the Fokker-Planck equation of a stochastic process defining a propagation in the physical time s. Moreover, since the vector fields that consitute the operator (2.25) are left invariant with respect to the composition law (2.10), the solution (2.19) possesses the symmetry

$$\Gamma_{\mathfrak{g}_0}(\eta | \eta_0) = \Gamma_{\mathfrak{g}_0}(\eta_0^{-1_L} \odot \eta | 0) \quad \forall \, \eta, \eta_0 \in \mathcal{M} \tag{2.27}$$

where $\eta_0^{-1_L}$ stands for the left inverse with respect to (2.10), see also Appendix.

Similarly to what I have done in the previous section, I can obtain $\Gamma_{\mathfrak{g}_0}(\eta | \eta_0)$ at each point by solving numerically the system of stochastic differential equations (2.23) and applying standard Markov Chain Monte Carlo methods. An isosurface plot of the kernel is shown in Fig. 2.8, over the same integral curves of Fig. 2.7 and again we see the concentration around the fan $\Sigma(\eta_0)$.

The fundamental solution (2.26) will be concretely applied in Section 2.5 as a facilitation field, toward the task of motion integration of trajectories.

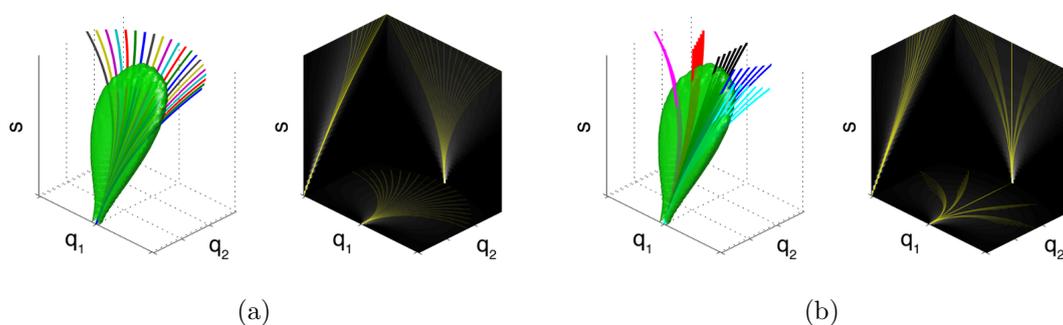(a)   (b)

Figure 2.8: Horizontal curves and stochastic kernels for the integration of motion, with a non-null initial velocity value. As before, different values of k are associated to different curve colors. Left: the isosurface plot of the kernel $\Gamma_{\mathfrak{g}_0}$ (isovalue: $0.001 \max(\Gamma_{\mathfrak{g}_0})$) is superimposed over the curves (2.22) with constant coefficients already shown in Fig. 2.7. Right: the projections of $\Gamma_{\mathfrak{g}_0}$ relative to the variables (q_1, q_2, s) are plotted under the projections of the curves (in yellow). This kernel was obtained by calculating the evolution of 10^6 stochastic paths.

2.4 Neural propagation of boundaries and trajectories

The curves (2.14) and (2.22) can be related to well-defined perceptual mechanisms. More precisely, I propose to consider them as association fields, in the sense of [34], devoted to the integration of contour in motion and trajectories. The perceptual tasks of subjective boundary completion and motion integration have been widely studied by both psychologists and physiologists, and their underlying physiological explanation continues to be an open issue of discussion. The perceptual bias towards collinear stimuli has classically been associated to the long-range horizontal connections linking cells in V1 sharing similar preferences in stimulus orientation. This specialized form of intra-striate connectivity pattern is found across many species, including cats, tree shrews and macaques [56, 15], the main difference being the specificity and the spatial extent of the connections. Furthermore, axons seem to follow the retinotopic cortical map anisotropically, with the axis of anisotropy being related to the orientation tuning of the originating cell [13]. These connections have already been modeled in [19] by means of a sub-Riemannian diffusion process over the orientation space $\mathbb{R}^2 \times S^1$.

Similarly to what happens for the integration of merely spatial visual information, the brain is also capable to easily predict stimulus trajectories [120], and to group together boundary elements sharing similar motion or apparent motion paths [60, 92]. One possible explanation for these effects could be the existence of specialized facilitatory networks linking cells anisotropically and coherently with their preference in velocity and axis of motion direction. Supporting these assumptions, it has been found that the neural preferences in direction of movement are also structurally mapped in the cortical surface, with nearby neurons being tuned for similar motion direction [132], and it has been shown that excita-

tory horizontal connections in the V1 of the ferret are strictly iso-direction tuned [99]. Furthermore, it is known that also extra-striate area MT/V5 is retinotopically organized, its horizontal connectivity pattern being highly structured, with connections reaching columns of cells tuned for similar orientation and direction preference anisotropically and asymmetrically [65, 2]. Moreover, striate and extra-striate cortical areas seem to cooperate, and surround modulation in V1 can be given by the connectivity patterns implemented in both areas by means of fast feedforward and feedback inter-areal projections [4].

Regarding the physiological correlates of motion integration, it was shown that the motion of an occluded object trajectory is significantly represented in the human brain by the same visual areas that process real motion [81]. Moreover, a recent study based on electrophysiological recordings of the V1 of tree shrews showed some non-linear neural behaviors that are coherent with the phenomenological dynamics of motion integration [138]. It is indeed a general assumption that some cortical area in the visual cortex is responsible for predicting future motion, a possible implementation being a specialized connectivity for spatio-temporal trajectory facilitation, that is different from the one in V1 responsible for contour integration [43, 119, 125, 135].

Here, I do not speculate on the exact physiological origin of these psychophysiological findings. However, I want to show that the connectivity kernels $\Gamma_{\mathfrak{g}^v}$ and $\Gamma_{\mathfrak{g}_0}$ arising from the geometry defined in Section 3.2 is capable to reproduce qualitatively some of the effects reported on the works that have been previously cited. I will embed this connectivity in the neural population activity model described in the next paragraph. Then, in Section 2.5 I will simulate the response of cortical visual neurons to artificial stimuli, comparing the results with some psychophysiological findings reported in the recent literature.

2.4.1 Modeling neural activity

The state of a population of cells can be characterized by a real-valued activity variable, which depends on the interaction of the feedforward input at different points, due to the cortical connectivity. The first population activity models are due to Wilson and Cowan [137], and Ermentrout and Cowan [32]:

$$\frac{\mathrm{d}a(\eta, t)}{\mathrm{d}t} = -a(\eta, t) + S\Big(c_f \int_{\mathcal{M}} \Gamma(\eta|\zeta)a(\zeta, t)d\zeta + F(\eta)\Big) \qquad (2.28)$$

where a is the neural activity of the population, F is the feedforward input (3.2), Γ is the facilitation kernel, c_f is the facilitation strength and S is the sigmoidal function

$$S(\tau) = \frac{1}{1 + e^{-\mu(\tau - \beta)}} \ . \qquad (2.29)$$

In the stationary case a first order approximation of the solution of (2.28) is

$$F_0(\eta) = S\Big(c_f \int_{\mathcal{M}} \Gamma(\eta|\zeta)S(F(\zeta))d\zeta + F(\eta)\Big) \qquad (2.30)$$

that is the activity formula that I will use in the experiments of Section 2.5. Let's explicitly note that the term

$$F_T(\eta) = S(F(\eta)) \qquad (2.31)$$

represents the mean neural extra-cellular activity in response to a stimulus.

The geometry of the functional architectures is contained in the kernel Γ, taking into account the deep structures of the connectivity space. Then the overall probability of activation can be obtained by convolution of the activity with the kernel Γ, and the term

$$P(\eta) = \int_{\mathcal{M}} \Gamma(\eta|\zeta)F_T(\zeta)d\zeta \qquad (2.32)$$

is the facilitation pattern resulting from the contribution of horizontal cortico-cortical or feedback inter-areal connectivities.

I note that, since the Lebesgue measure $d\zeta$ on $\mathcal{M}$ is left invariant with respect to the composition law (2.10), and due to the symmetry (2.27), then (2.32) has the structure of a convolution.

In the model of Ermentrout-Cowan only position and orientation were considered and symmetry properties were imposed to the facilitation patterns. In [106], for the same features it was proposed to choose as a kernel Γ the fundamental solution of a Fokker Planck equation deduced from the Euclidean symmetries. In the next section, two numerical simulations will be performed using equation (2.30). In the first one, I will model the propagation of boundaries in motion using the connectivity kernel (2.19). In the second experiment I will model instead the propagation of trajectories, using the 5D kernel (2.26).

2.5   Numerical simulations

In this section some numerical simulations will be performed to test the reliability of the kernels computed in Section 2.3 when introduced in the activity model (2.30) and the results will be compared to some psychophysical and physiological experiments.

2.5.1   The feedforward and extracellular activity in response to a stimulus

For the subsequent numerical simulations, I measured in every discrete point $z_j = (x_j, y_j, t_j)$ of a stimulus f of size $n_x \times n_y \times n_t$, its local energy of orientation and speed by convolving the input image sequence with a pre-determined bank of Gabor filters of fixed spatial frequency centered at the points $\eta_i = (x_i, y_i, t_i, \theta_i, v_i)$,

thus discretizing Eq. (3.2) so to have

$$F(\eta_i) = \left| \sum_{j=1}^{N} e^{-i(p_1(x_j-x_i)+p_2(y_j-y_i)-\nu_i(t_j-t_i))} \right. $$
$$\left. e^{-\frac{|x_j-x_i|^2+|y_j-y_i|^2}{\sigma_x^2} - \frac{(t_j-t_i)^2}{\sigma_t^2}} f(z_j) \right|^2 \tag{2.33}$$

where p_1 and p_2 are the component frequencies $|p|\cos\theta_i$ and $|p|\sin\theta_i$.

I chose the spatio-temporal frequencies and the maximum velocity value v_m represented in the Gabor filter set depending on the stimulus, so that given the couple $(|p|, v_m)$ the filters have a maximum temporal frequency of $\nu_m = |p|v_m$. I normalized the filters so that the response could range from 0 (in regions with no changes in luminance) to 1 (square plane waves sharing the filter parameters $|p|$ and ν_i), corresponding to a normalization of the stimulus contrast. The Gabor scale parameters will follow the relations

$$\begin{cases} \sigma_x & = & \frac{2.5\pi}{4|p|} \\ \sigma_t & = & \frac{\pi}{2\nu_m} \end{cases} \tag{2.34}$$

whose meaning is to approximately have 2.5 spatial subregions under the Gabor Gaussian envelope, and a variable ν_i that allows the filter with maximum velocity to cover one wavelength within the Gabor active time interval (see [20, 25, 94] for the physiological justifications).

The thresholded feedforward input F_T is then computed following Eq. (2.29), representing the mean neural extra-cellular activity in response to the stimulus. In the subsequent analyses, the values to which I set the parameters μ and β in the above stated formulas will be explicitly specified. It is worth noting that these parameters are not due to physiological data, due to the difficulty in finding some reference in the literature, but are chosen as to have a computationally reasonable number of non-null measurements F_T.
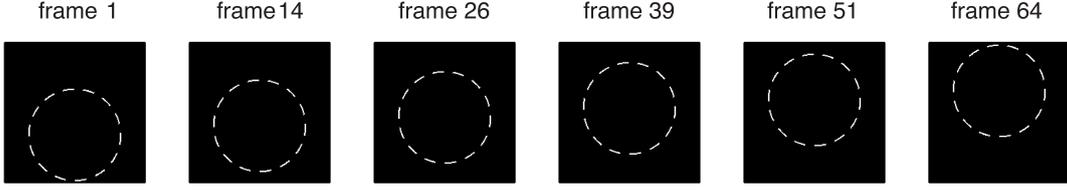
Figure 2.9: The stimulus used in Experiment 1.

2.5.2   Experiment 1 - Contours in motion

For the first simulation I use a dashed circle in linear motion moving upwards within the visual space. The image size is 200×200 pixels and the image sequence is composed of 64 frames. The circle segments are approximately 2 pixels wide and the circle is moving with a uniform speed of 0.5 pixels per frame (see Fig. 2.9). To get a measurement of local spatio-temporal features, I convolve the stimulus with a set of 3-dimensional Gabor functions with (3.2). The spatial frequency parameter is set to $|p| = \pi/2$, so that the Gabor moving subregions match the width of the segments. The maximum local velocity represented in the filter set is $v_m = 1$ pixel per frame, while spatial and temporal scale parameters are calculated accordingly with (2.34). After having convolved the stimulus with the Gabors, I model the neural extra-cellular activity by using (2.29) with parameters $\mu = 10$, $\beta = 0.5$, obtaining F_T.

I select the 4-dimensional subset corresponding to the 32nd frame $F_{T_v}(\xi_i) = F_T(x_i, y_i, 32, \theta_i, v_i)$, where $\xi_i = (x_i, y_i, \theta_i, v_i)$, thus neglecting behaviors over time. This is because the model given by the stochastic model of connectivity distribution $\Gamma_{\mathfrak{g}^v}$ is stationary, even if it is relying on spatio-temporal information, and is compatible with the dimensionality of $\mathcal{N}$ in (2.15).

Following the assumptions that the measurements $F_{T_v}(\xi)$ can model the output of a cortical direction- and speed-selective cell (or of a neural population that are

selective to the same visual features), the overall continuation probability can be obtained with a discretized version of the cortico-cortical facilitation pattern (2.32) with the kernel $\Gamma_{\mathfrak{g}^v}$:

$$P(\xi_i) = \sum_{\xi_i' \in \mathcal{N}} \Gamma_{\mathfrak{g}^v}(\xi_i | \xi_i') F_{T_v}(\xi_i') \tag{2.35}$$

where $\Gamma_{\mathfrak{g}^v}(\xi_i | \xi_i')$ is calculated for every fiber vector (θ_i, v_i) on the same discretized domain $\mathcal{N}$ of the lifted stimulus, using the stochastic approach described in Section 2.3.3. The structure of (2.35) is that of a discrete version of a group convolution with respect to the composition law (2.16), due to the symmetry (2.21) of the kernel and the left invariance of the Lebesgue measure on $\mathcal{N}$.

It is worth noting that the parameters κ and α governing the diffusion process are related to the maximum spatio-temporal curvature of an illusory contour. Due to the lack of reference physiological data, the values for the diffusion coefficients were chosen in such a way that at the final value taken by the evolution parameter of the stochastic paths, the mean square displacements of the fiber variables (θ, v) are $(\pi, v_m/2)$, that means $\kappa = 2$ and $\alpha = 1$.

Finally, the population activity is computed by

$$F_0(\xi_i) = S(F(\xi_i) + c_f P(\xi_i)) \tag{2.36}$$

that is a discretized version of (2.30) where c_f is a coefficient governing the total strength of the excitatory connections. In Fig. 2.10 are shown the results of the simulation. In the top row I have plotted a iso-level surface of the normalized and thresholded output measurements F_{T_v}. The process correctly lifts the stimulus around its theoretical values in the domain $\mathcal{N}$. In the bottom row, the iso-level surface of the global cortical activity F_0 is shown. Note that continuation is performed within the whole manifold $\mathcal{N}$, propagating the local speed and direction cues at the end of the segments, and interpolating the data to define a 4-dimensional set that carries information even in the visual space between the segments, where no
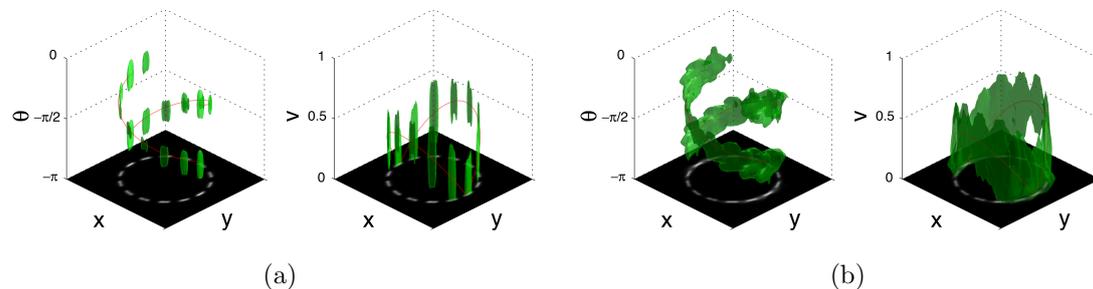
Figure 2.10: Results of Experiment 1: in the left subplots, measurements of local direction of motion θ; in the right subplots, measurements of local velocity v. The theoretical values for a circle translating at a speed of 0.5 pixels per frame are shown with a red line. a) Isosurface plot of F_T (isovalue: 0.2). The information about local orientation and velocity is correctly retrieved, but activity remains clustered within disjoint regions of the domain. b) The modeled extra-cellular activity F_0 after horizontal propagation on the submanifold $\mathcal{N}$ (facilitation strength: $c_f = 40$).

changes in contrast were originally present in the stimulus. These results are coherent with recent psychophysiological findings, where it is stated that a global shape in motion is better perceived if local velocity changes smoothly along its contour [92], and that the integration of the motion of a partially occluded object is facilitated when its visible contours define closed configurations [62].

2.5.3 Experiment 2 - Motion integration

The stimuli that I will use in this simulation will be several instances of an object in motion along a certain direction, that disappears at a given time position t_1, and reappears at t_2 with a direction of motion changed by $\Delta\theta$ in a position that is coherent to a piecewise continuous trajectory (Fig. 2.11). It is a well documented fact that humans tend to perceive the two trajectories as pertaining to a single unit just for small values of $\Delta T = t_2 - t_1$ and $\Delta\theta$, an effect that is commonly referred
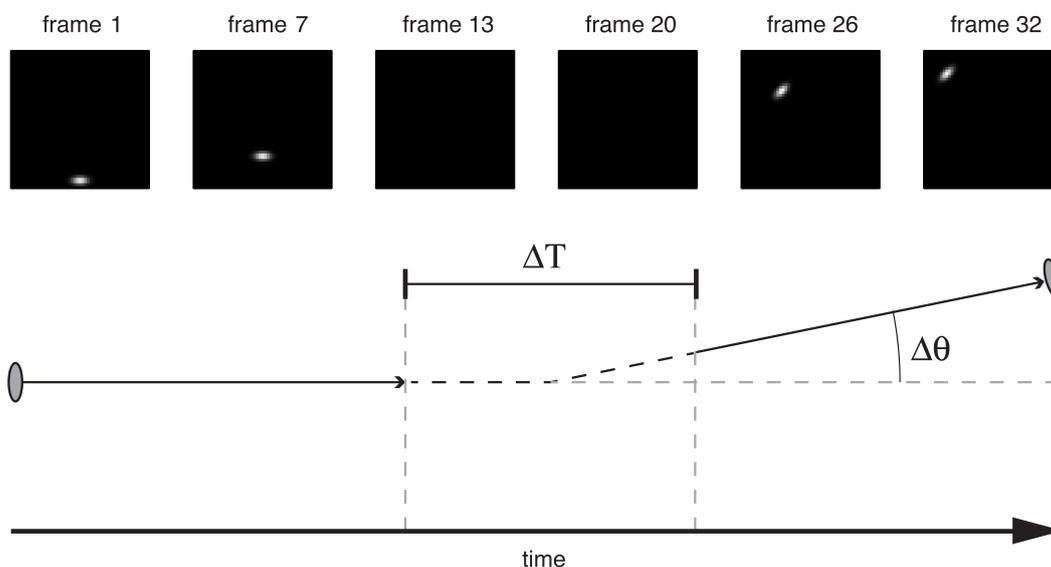
Figure 2.11: An example stimulus sequence, and the paradigm used to generate the stimuli in Experiment 2.

to as motion integration. In particular, we know from many psychophysiological experiments that the chances of detecting a straight or curvilinear trajectory in noise increase with stimulus duration and with direction coherence [127, 118].

To do so, I created multiple instances from the stimulus paradigm described above, and shown in Fig. 2.11, assigning different values to the parameters ΔT and $\Delta \theta$. It is worth noting that I am using an elongated object whose axis of motion is always orthogonal to its major axis of eccentricity. The choice of this particular stimulus is naturally explained by the geometry, as the points $\in \mathcal{M}$ that are connected by its continuation property implicitly define a local orientation and a direction of motion that is orthogonal to that orientation. The Gabor filtering of a different kind of stimulus (for example a moving dot) would indeed measure high responses also for those velocity projections that are coherent with the real axis of motion, due to the motion streak effect [38], requiring a different neural

model to detect unambiguously the direction of motion. Thus, even if the majority of the psychophysiological experiments use moving dots as stimuli, the validity of the following experiment is not influenced, as exploring the exact physiological implementation of motion integration is out of the scope of this chapter.

The image sequences to process are 51×51 pixels wide and are composed of 102 frames. The value of eccentricity of the moving ellipsoidal object is set to 2, and its velocity is set to 0.5 pixels per frame. Each stimulus instance is uniquely identified by the couple $(\Delta T, \Delta \theta)$. I convolve each sequence with a set of Gabor filters with $v_m = 1$ and p set to match width of the object's minor axis, and again, I threshold the output according to (2.29) using parameters $\mu = 20$, $\beta = 0.7$, obtaining the set of measurements F_T. In the previous simulation I have taken a temporal slice of this output, and then I have propagated the activity using the connectivity kernel $\Gamma_{\mathfrak{g}^v}$. Similarly, here I will propagate the activity present in F_T, but without discarding time. The corresponding facilitation pattern $P(\eta_i)$ is then obtained convolving it with the 5-dimensional kernel $\Gamma_{\mathfrak{g}_0}$ defined on the same discrete domain $\mathcal{M}$ of the lifted stimulus using the same approach that led to (2.35)

$$P(\eta_i) = \sum_{\eta_i' \in \mathcal{M}} \Gamma_{\mathfrak{g}_0}(\eta_i | \eta_i') F_T(\eta_i')$$

and the total population activity $F_0(\eta_i)$ is computed following (2.36) with $c_f = 20$. Since this is a discretization of (2.32), again I am considering a convolution-type operator.

The results obtained by processing two stimulus configurations are shown in Fig. 2.12, where only the central portion of the full temporal domain is plotted, to highlight the different effects that activity propagation has in the two cases. When the change in direction of motion of the object is under a certain threshold, the trajectory is completed smoothly, yielding a strong activation in the temporal interval when the object is not present in the scene. If the value of $\Delta \theta$ increases,
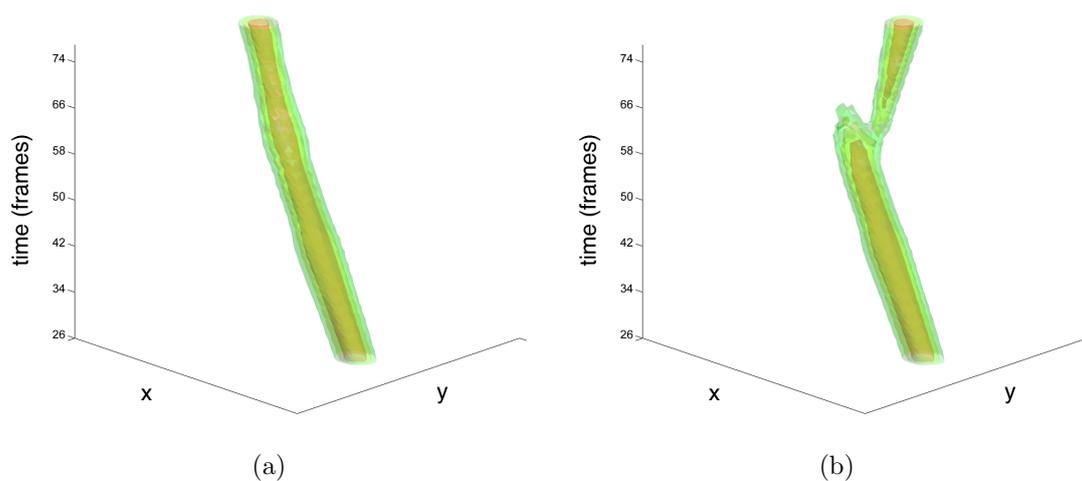
Figure 2.12: A comparison between the spatio-temporal continuation capabilities of the model, applied to stimulus instances with different values of $\Delta\theta$. The plot shows three isosurfaces of $\max_{(\theta,v)} F_0$ (isovalues 0.9, 0.5, 0.1 in red, yellow an green): a) $\Delta T = 12$, $\Delta\theta = \frac{\pi}{6}$, b) $\Delta T = 12$, $\Delta\theta = \frac{5\pi}{12}$. The probabilistic continuation given by the geometry in the configuration with the smallest change in direction of motion allows a smooth trajectory completion. The same is not true for the stimulus with an higher value of $\Delta\theta$: even if a weak response connecting and interpolating the two parts of the stimulus is still present, the strongest facilitation component do not deviate from the early straight path of the object.

(a)

(b)

Figure 2.13: A direct comparison between the electrophysiological recordings made in [138] and the results obtained with the proposed model for the same stimulus parameters: in both cases $\Delta T = 0$, while $\Delta\theta$ is equal to $\frac{\pi}{4}$ (a) and $\frac{\pi}{2}$ (b). To compare responses, I integrate over x and y and depict three isosurfaces with isovalues taken as 0.9, 0.5, 0.1 (in red, yellow and green respectively) times the maximum value among the outputs shown. Note that, disregarding the weaker physiological responses to opposite directions, the non-linear facilitating behavior is very well reproduced by the model.

however, the spatio-temporal curvature of the optimal subjective trajectory progressively becomes too high to be completed by the connectivity, even if a weak response linking the two parts of the stimulus could yet be present.

It is worth noting that due to the thresholding stage, this subjective interpolation effect is strongly non-linear, as it cannot be explained by the sum of the responses to the first $(t_i < t_1)$ and second $(t_i > t_2)$ part of the stimulus alone (see Fig. 2.13). This is coherent with the findings in the work of Wu et al, where electrophysiological experiments recorded a similar non-linear behavior of trajectory interpolation for small values of $\Delta\theta$ [138]. In that work, the cause of this effect is left unexplained. Even if the stimulus paradigms are slightly different (they use a field of moving dots with a null value of ΔT), the result of the simulations allow us to make some speculations.

A good qualitative description of the effects that the modeled excitatory connectivity has on stimulus response can be found in Fig. 2.14, where for every stimulus configuration I plot the difference

$$F_{fac} = F_0(S_3) - F_0(S_1) - F_0(S_2) \tag{2.37}$$

where S_3 is the full stimulus and S_1 and S_2 represent the first $(S1(t_i > t_1) = 0)$ and the second $(S2(t_i < t_2) = 0)$ part of the stimulus. The visualization of the output F_{fac} highlights the role that a trajectory-specialized cortical connectivity could have in performing tasks of motion integration. The spatio-temporal extension of F_{fac} over $\mathcal{M}$ gets smaller for higher values of ΔT and $\Delta\theta$. The decay of the facilitation effect is coherent with the observations made in [125] (where $\Delta\theta = 0$) and [138] (where $\Delta T = 0$), even if the experiments are methodologically different. Moreover, even if, coherently with the results, some psychophysiological experiments showed that a broken trajectory in noise is easily detectable [126, 109], as far as I know little has been done to explore the effect of changing the duration ΔT of the temporal occluding gap.

2.5.4   Discussion

Regarding the implementation of the model proposed in this chapter, I would like to highlight functional meaning of the parameter couple (κ, α), driving the diffusion along the fiber variables (θ, v) when calculating the kernels $\Gamma_{\mathfrak{g}^v}$ and $\Gamma_{\mathfrak{g}_0}$. In particular, the parameter κ seems to be strongly related to the maximum perceived curvature of illusory contours and trajectories, while α sets the maximum rate of change of local velocity along admissible subjective contours (trajectories). It would be interesting to try to fine-tune the parameters of the model in order to reproduce quantitatively as precisely as possible the psychophysiological findings found in literature. I aim to carry out this analysis in a future work.

It is worth noting that the physiological correlates of the first simulation are not well documented and, as far as I know, no one has yet studied the neural activity in cortical regions responding to subjective contours in motion. I suppose that it would be a difficult issue to address, as multiple cortical layers, as V1, V4 and MT/V5, may be involved [61]. Recently, some significative results have been obtained with static illusory contours, using promising electrophysiological techniques [84]. Those kinds of experiments, targeted to the detection of illusory motion contours, could give additional clues about the neural computation that governs the influence of spatio-temporal features in the detection of moving shapes and boundaries.

As discussed in the beginning of Section 2.4, a possible physiological implementation of the geometry used in the second numerical simulation, could be a trajectory-specialized cortico-cortical connectivity between neurons in the higher visual areas. The advantages of having such an excitatory connectivity implemented in the visual cortex could be very important in performing complex cognitive tasks. For example, I assume that the brain could use the improvements in contour detection with respect to the only $SE(2)$ related functional architecture,
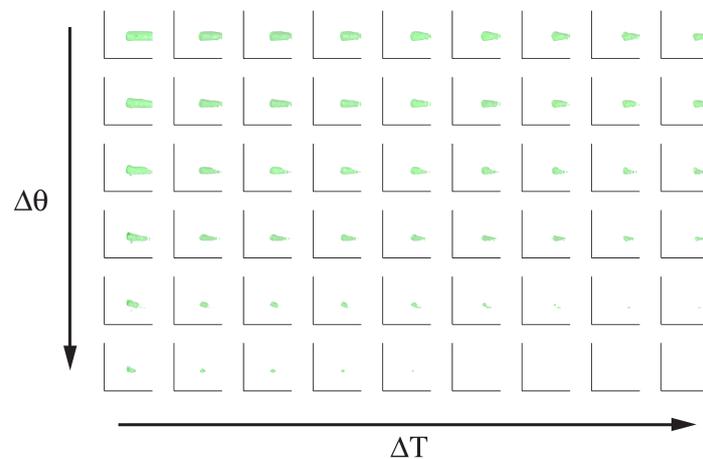
Figure 2.14: The non-linear effect of the facilitation by varying the stimulus parameters ΔT and $\Delta \theta$. Each subplot shows the isosurface of F_{fac} relative to different stimulus instances (isolevels calculated as in Fig. 2.13). The non-linear effect of the facilitating kernel $\Gamma_{\mathfrak{g}_0}$ remains appreciable just for small values of ΔT and $\Delta \theta$, rapidly decaying as the value of the characteristic stimulus parameters increases.

due to an additional information, that of velocity, that indeed in practical situation is a feature that is coherent on objects. The influence that the connectivity arising from the proposed geometrical model has in the processing of visual tasks such as spatio-temporal grouping or segmentation will be the subject of a future work.

Chapter 3

Spatio-temporal grouping properties of the connectivity of the visual cortex

3.1 Introduction

The importance of key global properties of the visual stimulus, such as neighboring or good continuation, for the execution of visual cognitive tasks like image segmentation and grouping has been well understood since the beginning of the century, when the phenomenological concepts of the Gestalt began to be defined [124].

The study of these phenomena through psycho-physiological experiments lead to the definition of *association fields*, entities that describe which properties the stimuli near a given sample should have to be recognized as belonging to the same perceptual unit – in the classical case, the property of local orientation [34]. Further phenomenological experiments demonstrated that also movement direction and velocity are important features in order to perceive global shapes [92], and

81

that similarly to what happens for the integration of spatial visual information, the brain is also capable to easily predict stimulus trajectories, and to group together elements having similar motion or apparent motion paths [118, 119, 125].

Several physiological experiments showed how these principles seem to be implemented in the V1 of mammals, where long-range connections between two cells is stronger when they respond to stimuli having similar position and orientation [13]. Interpreting cortical columns as directional differential operators, Citti and Sarti showed how this specialized functional organization of V1 for oriented stimuli, implemented by long-range horizontal connections, naturally leads to the definition of Lie algebras operating within a sub-Riemannian space. In particular, they modeled the position/orientation association fields with a family of integral curves, obtained from the generators of the Lie algebra associated to a particular direction-selective visual neuron, embedded in the $R^2 \times S^1$ domain [19].

The spatio-temporal properties and organization of cortical visual neurons though, together with the phenomenological findings on spatial and motion integration, suggest extensions of the model to include local stimulus velocity, and give indications for the different geometries to be chosen in order to include the elaboration of trajectories. Indeed, the mathematical instruments given by group theory form a flexible framework in which the properties of different geometries can be inherited and embedded in a layered fashion, coherently integrating more their associated mechanisms.

While in the cited papers the geometry has been used to perform spatio-temporal image completion, in this chapter I will focus on its grouping capabilities. Indeed, while the grouping properties of the first kind of connections have been previosly studied [88, 39, 60], little has been done to follow this path in the spatio-temporal case. The aim of this chapter is to show that a connectivity pattern

based also on local stimulus velocity can enhance the spatial (stationary) grouping capabilities of a visual system.

In Section 2 I will describe the geometry arising from the study of the spatio-temporal functional architecture in the visual cortex. The geometrical tools that I will use trhoughout this chapter have been introduced in chapter 2. The notation and the theoretical framework exposed in this section will also be used in Section 3.4, in order to propose a new version of the previously described connectivity models.

In Section 3 I will recall the concept of spectral clustering and graph partitioning by means of locality-preserving embeddings of high-dimensional data sets in the low-dimensional Euclidean eigenspace of a suitable affinity matrix, and then I will use the connectivity properties of the proposed cortical-inspired geometry to represent a visual stimulus living in the feature space of position/orientation to build a *cortical* affinity matrix. I will also propose a simple clustering algorithm that is able to perform grouping of elements and background/object separation depending on the kernel used to build its input affinity matrix.

In Section 4 I will use the previously introduced spectral clustering algorithm to automatically extract perceptual information from artificial stimuli living in the extended cortical spaces described in Section 2, so to group together the points in the lifted feature space that are coherent with the proposed cortical-inspired geometry. I will show the effect of changing kernel parameters on the results of the algorithm, and finally I will propose two different connectivity geometries that the visual cortex could have implemented in order to perform spatio-temporal segmentation of moving contours and shapes.

3.2   The geometry of V1

In this section I resume the work presented in the previous chapters, where it is shown how examining the functional architecture of the visual cortex naturally leads to the definition of a contact structure on a sub-Riemannian manifold, whose connectivity properties are able to explain some physiological and phenomenological findings on spatio-temporal visual perception.

3.2.1   The cortical feature space $\mathcal{M}$

It is well known since the fundamental studies of Hubel and Wiesel [51, 52] that primary visual cortex (V1) is one of the first physiological layers along the visual pathway to carry out geometrical measurements on the visual stimulus, decomposing it in a series of local feature components. In particular, the development of suitable electrophysiological techniques [96] has made possible to reconstruct the linear filtering behavior of V1 simple and complex cells, i.e. their spatio-temporal receptive profiles (RPs).

Spatio-temporal RPs of V1 simple cells can be modeled by 3-dimensional Gabor functions of the form

$$g_{q,p}^{\sigma}(x,y,t) = e^{2\pi i(p_x(x-q_x)+p_y(y-q_y)+p_t(t-q_t))}e^{-\left(\frac{(x-q_x)^2}{2\sigma_x^2}+\frac{(y-q_y)^2}{2\sigma_y^2}+\frac{(t-q_t)^2}{2\sigma_t^2}\right)}, \qquad (3.1)$$

where $q = (q_x, q_y, q_t)$ is the spatio-temporal center of the Gabor filter, $p = (p_x, p_y, p_t)$ is the spatio-temporal frequency and $\sigma = (\sigma_x, \sigma_y, \sigma_t)$ is the spatio-temporal width. It is worth noting that this model strictly captures the features of so-called inseparable RPs, tuned for position, orientation and motion detection, depicted in Fig. 3.1), while separable RPs have been treated elsewhere, see e.g. [20]. One of the crucial features of (3.1) is minimization of the uncertainty of simultaneous measurements in space-time and frequency [24].
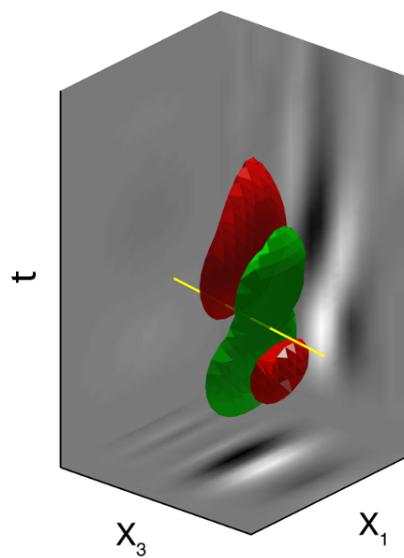
Figure 3.1: Isolevel surfaces of an inseparable receptive profile defined in space-time, as the ones usually found in V1. Green and red surfaces enclose, respectively, excitatory and inhibitory regions. The yellow line indicates the direction of the local vector $\vec{X}_5$ (see text for details).

Further analyses have also shown that the Gabor parameter distribution found in cortical cells cover only a subset of the whole Gabor family. Such subsets are optimized for the detection of the local features of orientation θ and speed v [20, 8], defined as

$$
\begin{aligned}
\theta &= \arctan\left(\tfrac{p_y}{p_x}\right) \\
v &= \tfrac{p_t}{\sqrt{p_x^2 + p_y^2}}
\end{aligned}
$$

that can be interpreted as fundamental features of the visual stimulus. For this reason I will not deal with the dependence on the spatial frequency $\kappa = \sqrt{p_x^2 + p_y^2}$ or on the scale σ (see also Chapter 2), and consider RPs of the form

$$
g_{q,\theta}^{\sigma,\kappa}(x,y,t) = e^{-2\pi i \kappa((x-q_x)\cos\theta + (y-q_y)\sin\theta + v(t-q_t))} e^{-\left(\frac{(x-q_x)^2}{2\sigma_x^2} + \frac{(y-q_y)^2}{2\sigma_y^2} + \frac{(t-q_t)^2}{2\sigma_t^2}\right)}
$$

where κ and σ are considered as fixed.

This corresponds to a neural processing stage where the visual stimulus is lifted from the spatio-temporal image space $\mathbb{R}^2 \times \mathbb{R}^+$ to the extended 5-dimensional feature space

$$
\mathcal{M} = \mathbb{R}^2 \times \mathbb{R}^+ \times S^1 \times \mathbb{R}^+ = \left\{ \eta = (q_x, q_y, q_t, \theta, v) \right\},
$$

where every point $\eta \in \mathcal{M}$ corresponds to a Gabor filter (3.1), up to the spatial frequency and scale factors (see also Fig.3.2). The activity of V1 simple cells is indeed modeled by the map

$$
f \mapsto F^{\sigma,\kappa}(q,\theta,v) \doteq \langle g_{q,\theta}^{\sigma,\kappa}, f \rangle_{L^2(\mathbb{R}^3)}. \tag{3.2}
$$

3.2.2   Connectivity on $\mathcal{M}$ as a differential constraint

The functional behavior of V1 simple cells modeled by (3.2) can be interpreted as a finite scale spatio-temporal directional derivative of the stimulus around position q, performed along the direction

$$
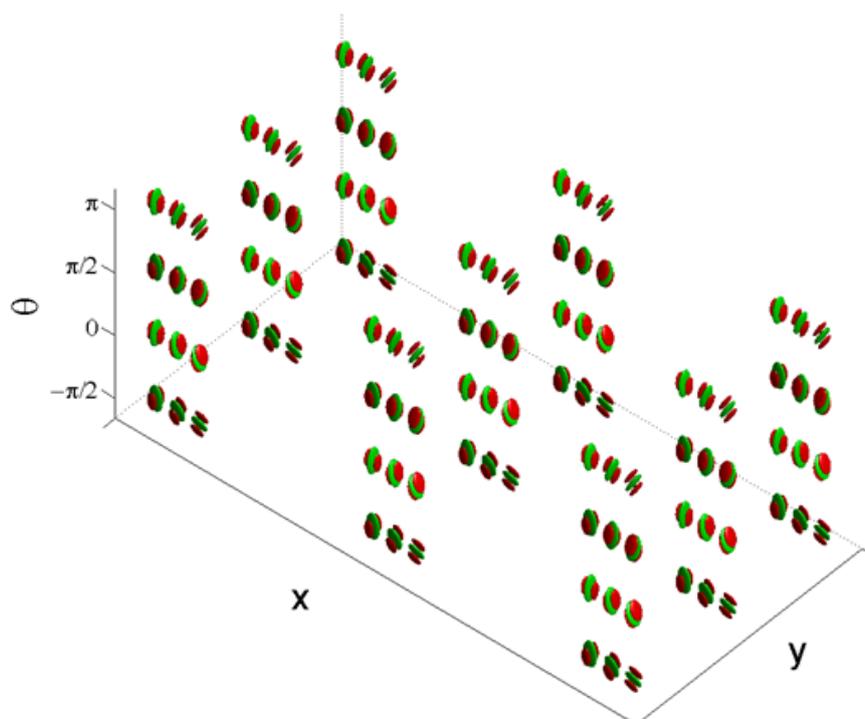\vec{X}_{\theta,v} = (\cos\theta, \sin\theta, -v)
$$

Figure 3.2: Schematization of the feature-wise organization of the primary visual cortex. For each spatio-temporal point (x, y, t) of the image hyperplane there is a two dimensional fiber of representing local orientation θ and local velocity v.

expressed in the coordinates $\{\hat{e}_x, \hat{e}_y, \hat{e}_t\}$. Accordingly, as discussed in Chapter 2, this derivation is maximal along the direction of the gradient of f. This implies that the lifting to $\mathcal{M}$ of any smooth level set of f is always orthogonal to the vector field

$$\vec{X} = (\cos\theta, \sin\theta, -v, 0, 0) \in \mathcal{T}\mathcal{M}$$

expressed in the coordinates $\{\hat{e}_{q_x}, \hat{e}_{q_y}, \hat{e}_{q_t}, \hat{e}_\theta, \hat{e}_v\}$.

Hence the present problem induces to consider as admissible surfaces on $\mathcal{M}$ those whose tangent space at any point is spanned by the vector fields

$$\vec{X}_1 = (-\sin\theta, \cos\theta, 0, 0, 0) \, , \; \vec{X}_2 = (0, 0, 0, 1, 0)$$
$$\vec{X}_4 = (0, 0, 0, 0, 1) \, , \; \vec{X}_5 = (v\cos\theta, v\sin\theta, 1, 0, 0) \tag{3.3}$$

defining the orthogonal complement to $\vec{X}$ in $\mathcal{TM}$. The four dimensional hyperplanes generated by $\{X_1, X_2, X_4, X_5\}$ is called contact plane, and the whole structure is named a contact structure. Contact structures have been used for modeling the functional architecture of the visual cortex in several works, see e.g. [89, 19].

Due to this contact structure, the connectivity among V1 cells on $\mathcal{M}$ was modeled geometrically in Chapter 2 in terms of advection-diffusion processes along the directions of the vector fields (3.3). Two corresponding stochastic processes were introduced in order to provide concrete realizations of the mechanisms of propagation of information along connections.

A first mechanism, aimed to model connectivity along lifted contours of a spatial image at a fixed time, consists of a propagation along the direction $\vec{X}_1$ forced by a diffusion over $\vec{X}_2$ and $\vec{X}_4$. This will be used for a single-frame segmentation out of a spatio-temporal streaming. It lives on a codimension 1 submanifold of $\mathcal{M}$ at fixed time t_0,

$$\mathcal{M}_0 = \mathbb{R}^2 \times S^1 \times \mathbb{R}^+ = \{\xi = (q_x, q_y, q_t, \theta, v) : q_t = t_0\}$$

and can be formally described by the following system of stochastic differential equations

$$
\begin{cases}
dq_x(s) &= -\sin\theta(s)\,ds \\[2mm]
dq_y(s) &= \cos\theta(s)\,ds \\[2mm]
d\theta(s) &= \kappa\,dW_1(s) \\[2mm]
dv(s) &= \alpha\,dW_2(s) \\[2mm]
\xi(0) &= \xi_0
\end{cases}
\qquad (3.4)
$$

where $W = (W_1, W_2)$ is a two dimensional Brownian motion and (κ, α) are the corresponding diffusion constants.

A second mechanism, aimed to model connectivity among moving contours of a spatio-temporal stimulus, will be used for spatio-temporal segmentation of apparent point trajectories. It consists of a propagation along $\vec{X}_5$ again forced by a diffusion over $\vec{X}_2$ and $\vec{X}_4$, and is described by

$$
\begin{cases}
dq_x(s) &= v\cos\theta(s)\,ds \\[2mm]
dq_y(s) &= v\sin\theta(s)\,ds \\[2mm]
dq_t(s) &= ds \\[2mm]
d\theta(s) &= \kappa\,dW_1(s) \\[2mm]
dv(s) &= \alpha\,dW_2(s) \\[2mm]
\eta(s=0) &= \eta_0.
\end{cases}
\qquad (3.5)
$$

The construction of these processes takes into account the different role of the spatio-temporal variables q, where is defined the stimulus, and that of the engrafted variables (θ, v). Particularly, in the q variables we have advective processes while in the (θ, v) variables we have diffusion. It is worth noting that this naturally extend the process proposed by Mumford in [73] for the case of static images.

I can define for the processes (3.4) and (3.5) two probability densities that a point on $\mathcal{M}_0$ or on $\mathcal{M}$ is reached by a random path. They can be computed

as fundamental solutions of partial differential differential equations, derived in Chapter 2 from the associated Fokker Planck equations. In particular, for the process (3.4) I can obtain the corresponding connectivity kernel $\Gamma_0(\xi|\xi_0)$ on $\mathcal{M}_0$ from

$$(X_1 - \kappa^2 X_2^2 - \alpha^2 X_4^2)\Gamma_0 = \delta(\xi - \xi_0) \qquad (3.6)$$

where as customary I have denoted with X_i the directional derivative along the vector field $\vec{X}_i$. Analogously, the process (3.5) defines connectivity kernels $\Gamma(\eta|\eta_0)$ on $\mathcal{M}$ as solutions to

$$(X_5 - \kappa^2 X_2^2 - \alpha^2 X_4^2)\Gamma = \delta(\eta - \eta_0). \qquad (3.7)$$

A crucial feature of the differential operators in equations (3.6) and (3.7) is that they are hypoelliptic (see [50] and the discussion in Chapter 2). This implies in particular that the resulting kernels Γ_0 and Γ are smooth outside the pole and nonzero everywhere on the manifolds $\mathcal{M}_0$ and $\mathcal{M}$.

Hypoellipticity is due to the fact that the vector fields $\{\vec{X}_1, \vec{X}_2, \vec{X}_4\}$ together with their commutators span the whole tangent space of $\mathcal{M}_0$, while the vector fields $\{\vec{X}_5, \vec{X}_2, \vec{X}_4\}$ together with their commutators span the whole tangent space of $\mathcal{M}$. Such span condition allows to connect any couple of points on the reference manifolds with of integral curves of these reduced sets of tangent vectors, hence defining an anisotropic geodesic distance [42] whose balls are squeezed in the directions of the commutators [75]. Moreover, the obtained connectivity kernels decay to zero at infinity according to such a distance [100].

3.2.3   Discrete connectivity kernels

In order to use the connectivity properties of Γ_0 and Γ to model the visual grouping capabilities of the visual cortex, I must define a numerical method for their

calculation that is both accurate and efficient. While in [31] several representations of the exact solution to a problem analogous to (3.6) were presented, in this work I will implement the fundamental solutions with standard Markov Chain Monte Carlo methods (MCMC) [101], a technique that is suitable to treat also the more involved case of equation (3.7). As written in Chapter 2, the comparison of the resulting kernel numerical estimates with [31] confirms the accuracy of the stochastic method. A deeper study of numerical accuracy, that goes out of the scope of this thesis, will be published in a future work.

To get a stochastic estimate of the kernel Γ_0, I must implement the calculation of a random path originating from a point $\xi_0 = (x_0, y_0, \theta_0, v_0)$ of the discretized contact manifold $\mathcal{M}_0$. This is done by discretizing (3.4) over the parameter s, and without loss of generality I will use as step of the discretization $\Delta h = 1$, so to have, for unitary displacements over the image plane $\mathbb{R}^2$:

$$\begin{cases} x_{h+1} - x_h &= -\sin\theta \\ y_{h+1} - y_h &= +\cos\theta \\ \theta_{h+1} - \theta_h &= \delta(\kappa, 0) \\ v_{h+1} - v_h &= \delta(\alpha, 0) \end{cases} , \quad h \in \{0, \dots, H\} \tag{3.8}$$

where H is the number of steps performed by the random path and $\delta(\sigma, \mu)$ is a generator of numbers taken from a normal distribution with mean μ and variance σ^2. Solving this finite difference n times will give n different realizations of the stochastic path: the kernel estimate will then be obtained by assigning to each region of the discretized contact domain a value proportional to the probability that a path passes through it at step h, $\rho^{\kappa,\alpha}(\xi_i, h | \xi_0, 0)$, and then integrating over h:

$$\hat{\Gamma}_0^{H,\kappa,\alpha}(\xi | \xi_0) = \sum_{h=0}^{H} \rho^{\kappa,\alpha}(\xi_i, h | \xi_0, 0). \tag{3.9}$$

Note that with this notation I explicitly underline the role of the diffusion constants κ and α and the path length limit H as fundamental parameters of the estimate kernel $\hat{\Gamma}_0$, as later I will vary these parameters to build the analyses. In Fig. 3.4 I show the effect of the variation of two of these parameters on the shape and level of stochastic approximation of $\hat{\Gamma}_0$.

Similarly, the estimate of the kernel for particle motion integration can be calculated by discretizing the stochastic system (3.5), so that, for unitary displacements over the image plane $\mathbb{R}^2$ in one time unit and starting point $\eta_0 = (x_0, y_0, t_0, \theta_0, v_0)$, we have the random paths:

$$
\begin{cases}
x_{h+1} - x_h & = & +v\cos\theta \\
y_{h+1} - y_h & = & +v\sin\theta \\
t_{h+1} - t_h & = & 1 \qquad\qquad , \quad h \in \{0, \ldots, H\} \\
\theta_{h+1} - \theta_h & = & \delta(\kappa, 0) \\
v_{h+1} - v_h & = & \delta(\alpha, 0)
\end{cases}
\tag{3.10}
$$

As before, integrating over h and counting for each region in the discrete domain $\mathcal{M}$ how many paths passed through it, I can build the kernel estimate $\hat{\Gamma}^{H,\kappa,\alpha}(\eta|\eta_0)$.

In Fig. 3.3 I show the isosurfaces of two projections of $\hat{\Gamma}_0$ and $\hat{\Gamma}$. As can be seen, these kernels reach their maximum values in the proximity of the horizontal curves that originate from the same starting point. Thus, assuming that the fan composed of integral curves with constant coefficients on θ and v can be considered good models of spatio-temporal association fields, the kernel estimates $\hat{\Gamma}_0$ and $\hat{\Gamma}$ can be interpreted as connectivity distributions on $\mathcal{M}_0$ and $\mathcal{M}$, since they provide stronger connections between points that may belong to the same kind of lifted level set.

In the next sections I will make use also of the simpler connectivity kernel described in [19], that is, the position-orientation probability distribution Γ_q defined
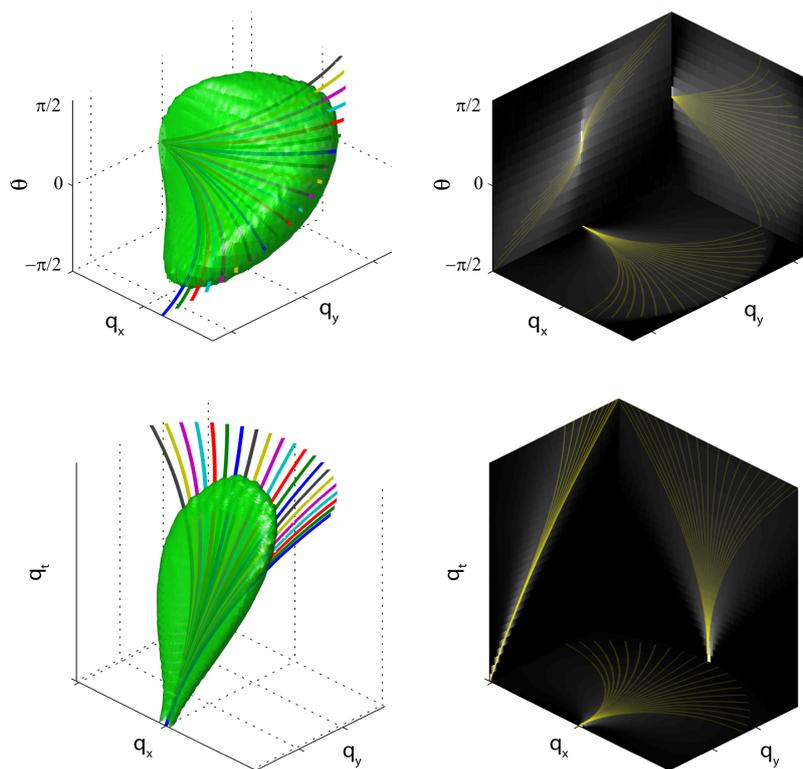
Figure 3.3: Horizontal curves calculated as in [19] and in Chapter 2 compared to the stochastic estimates $\hat{\Gamma}_0$ and $\hat{\Gamma}$. Left: isosurface plot of the kernels (isovalues: $0.002 \max \hat{\Gamma}_0$ and $0.001 \max \hat{\Gamma}$). Right: kernel projections relative to the variables (x, y, θ) for $\hat{\Gamma}_0$ and b) (x, y, t) for $\hat{\Gamma}$ are plotted in gray under the projections of the horizontal curves (in yellow). This kernels were obtained by calculating the evolution of 10^6 stochastic paths..

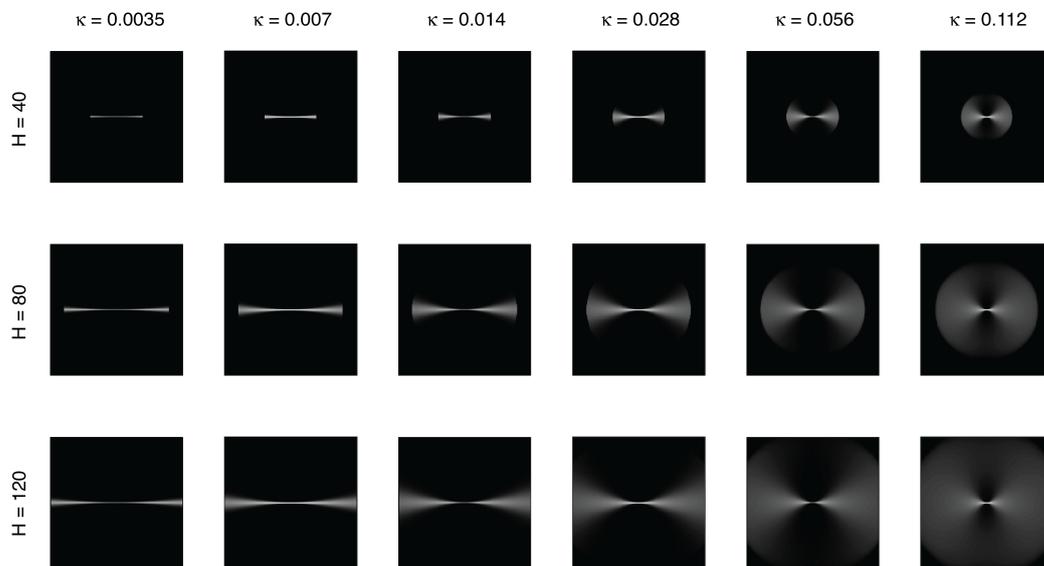Figure 3.4: Marginal distributions over the $x-y$ plane of the stochastic kernel estimates $\hat{\Gamma}_0^{H,\kappa,\alpha}$ for different values of parameters k and H. For visualization purposes, the values were exponentiated by $\frac{1}{4}$

on the cortical contact manifold $\mathcal{M}_q = \mathbb{R}^2 \times S^1$. The equations for its stochastic estimate calculation are exactly the ones of Eq. (3.8), without the additional engrafted variable of local velocity:

$$\begin{cases} x_{h+1} - x_h &=& -\sin\theta \\ y_{h+1} - y_h &=& +\cos\theta \quad , \quad h \in \{0, \dots, H\} \\ \theta_{h+1} - \theta_h &=& \delta(\kappa, 0) \end{cases} \qquad (3.11)$$

As before, the integration over the steps of the random walks and over the stochastic paths lead to the definition:

$$\hat{\Gamma}_q^{H,\kappa}(\xi|\xi_0) = \sum_{h=0}^{H} \rho^{\kappa}(\xi_i, h|\xi_0, 0). \qquad (3.12)$$

3.3 Spectral analysis

Two examples of visual grouping are portrayed in the first column of Fig.3.5. In the first row, three dense Gaussian distributions of 2-dimensional points are embedded within a sparser set of random points uniformly scattered throughout the domain. The human eye normally associates the points of the Gaussian point clouds to three separate objects lying on a noisy environment. In the second row, two dashed continuous lines are embedded in a field of segments having random position and orientation. In this case, stimulus collinearity clearly gives rise to a pop-out effect that makes the two lines easily distinguishable from the background. This powerful effect was quantified with psycho-physiological experiments in [34], and was firstly simulated using spectral techniques in [88, 63].

Though these examples show two very different grouping effects, the underlying mechanism is the same and can be generalized and formalized as follows: given a set of points $S = \{x | x \in \mathbb{R}^d\}$ living in an arbitrary d-dimensional feature space, the task of grouping together the points that are more similar to each other
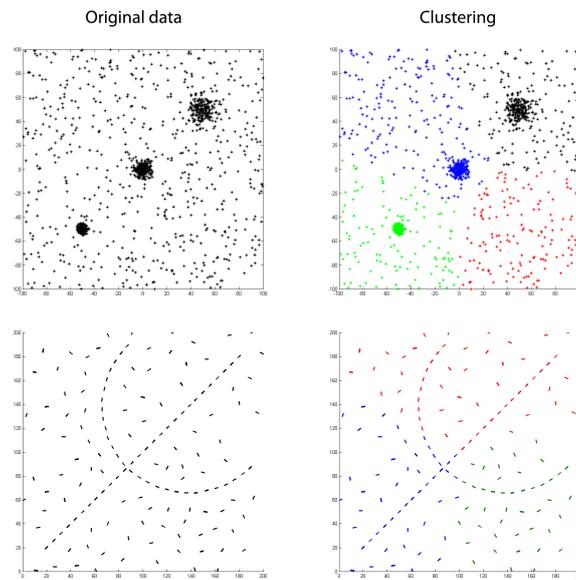
Figure 3.5: Left column: two sample data set of perceptual units embedded in a random point field. Right column: result of k-means algorithm. The parameter k, determining the number of clusters to find, was set to, respectively, 4 and 3, as I know the number of perceptual units (including the noise/background) that I want to find.

(so that their ensemble can be considered as being an *object*) is carried out by discriminating these groups from the rest. More formally, I want to search for K disjoint subsets $S_i \subset S$, with $i \in \{1, \ldots, K\}$, where we have $K - 1$ relevant groups of points and an additional group containing noise, that is, the ensemble of data that cannot be associated to any other point.

The cognitive task of spatial or spatio-temporal visual grouping can be interpreted as a form of clustering. A vast literature is available on the theoretical and practical aspects of several clustering algorithms: a survey of the main and most known techniques can be found in [123].

It is widely known that this problem is not easily resolvable with popular and efficient algorithms like *k-means*. In the second column of Fig.3.5 I show how its direct implementation can lead to big clustering failures in both cases: for the point clouds data set, I used the standard Euclidean distance in the image plane $\mathbb{R}^2$, while for the segment data set I used the Euclidean distance in the extended feature space $\mathbb{R}^2 \times S^1$. Another downside of many clustering algorithms, including k-means, is that the number of clusters have to be given as input, chosen by means of a priori knowledge or some arbitrary heuristics.

Trying to solve these issues, a major branch of research on this subject has been recently developed by studying the mechanisms of methods that analyze the spectral properties of positive, semi-definite affinity matrices constructed from a set of input data. The ensemble of these techniques can generally be subdivided into two classes [58]: methods for locality-preserving embeddings of large data sets, that project the data points onto the eigenspaces of the affintiy matrices [22, 11, 102], and methods for data segregation and partitioning, that basically perform an additional clustering step taking as input the projected data set [88, 131, 111, 70, 76].

Here, I want to formalize and describe a spectral clustering method to perform visual grouping, that is dependent on as few parameters as possible, including the number of clusters into which subdivide the input data set. In section 4, I will adopt this method to test and analyze the grouping properties of the geometry described in section 2.

3.3.1   Spectral clustering for visual grouping

As previously mentioned, dimensionality reduction and clustering can be obtained by inspectioning the spectral properties of a $n \times n$ affinity matrix A (for a data set containing n points) whose elements a_{ij} have values proportional to the similarity of point p_i to point p_j. The formalization of the concept of similarity strictly depends on the domain in which the points are defined: in continuity with the previous examples, without loss of generality we can think of n points $p_j{}_{j=1}^n$ living in a Euclidean space $\mathbb{R}^d$, for which an affinity matrix can be built as:

$$A = a_{ij} = \exp\left(\frac{\|p_i - p_j\|^2}{2\sigma^2}\right), \tag{3.13}$$

where σ is a scale parameter that has to be chosen upon the characteristics of the data set to be clustered.

It has been originally shown by Perona that the first eigenvector of A can serve as indicator vector for basic grouping purposes [88]. In fact, it can be shown that simply by thresholding the first eigenvectors of the affinity matrix it is possible to obtain a basic background/foreground contour separation, with the foreground layer represented by the most salient eigenvalue (these results will be published in a future paper). Basing upon this property, Perona formalized a partitioning algorithm that recursively separate the foreground information from the data set. While this algorithm's implementation is straightforward and efficient, it was

demontrated that it can easily lead to clustering errors due to noise, non-linear distributions or outliers [111, 123].

To overcome these issues, the affinity matrix can be normalized using the diagonal degree matrix D having elements

$$d_i = \sum_{j=1}^{n} a_{ij}. \tag{3.14}$$

Though many authors have proposed different kinds of normalization, it has been found that the one giving the best results is the one proposed by Meila and Shi [70, 123], where the affinity matrix A is transformed into a Markov matrix P via the row-wise normalization

$$P = D^{-1}A, \tag{3.15}$$

Thus this matrix won't in general be symmetric, it can be shown that its eigenvalues $\{\lambda_j\}_{j=1}^{n}$ are real and satisfy $0 \leq \lambda_j \leq 1$, and its eigenvectors $\{u_j\}_{j=1}^{n}$ can accordingly be chosen with real components [22, 58].

This matrix offers very nice interpretations of the clustering properties of its eigenvectors. Suppose we have a graph $G(V, E)$ - V being the set of vertices and E the sets of edges connecting them - with n nodes and c connected components $\{G_i\}_{i=1}^{c}$, where elements of a component have the same edge weight connecting them. The resulting normalized affinity matrix P would be a block diagonal matrix, and thus will have c non-null eigenvalues $\{\lambda_i\}_{i=1}^{c} = 1$ and $n-c$ null eigenvalues $\{\lambda_i\}_{i=c+1}^{n}$, with piece-wise constant indicator eigenvectors $\{u_i\}_{i=1}^{c}$ [111, 123].

In real applications, the affinity matrices are perturbed versions of the block diagonal ones representing a clear partitioning, and thus we won't generally have ideally binary spectra. Anyway, if their connectivity is well defined, it is sufficient for points in a cluster to be weakly connected mainly to their cluster neighbours, to get a good approximation of the ideal case. Indeed, considering the affinity values a_{ij} as the edges of a weighted undirected graph with n nodes, several authors

demonstrated that the first egenvectors $\{u_i\}_{i=1}^c$ of the normalized affinity matrix P, corresponding to the ordered eigenvalues $(\lambda_1 > \lambda_2 > \ldots > \lambda_c)$ solve the relaxed optimization problem of Normalized Cuts [111], thus giving a nice probabilistic interpretation of the clustering problem.

Thus, since we know that the first eigenvectors of P possess clustering information, we have to decide which ones we have to take into consideration. While many authors proposed different solutions, like looking for the maximum eigengap or trying to minimize a particular cost function [123, 140], the simplest and most efficient way to go is to take all those u_i whose $\lambda_i > 1 - \epsilon$, where ϵ is a suitable thresholding parameter. The more P is far from being similar to a block diagonal matrix, the more its spectrum will be far from be dichotomous, with ordered λs decreasing more smoothly (see Fig. 3.6), in which case the sensitivity to small changes on the values of ϵ becomes very high. To solve this issue, I will use another property of P.

Due to the Markovian normalization of Eq. (3.15), the rows of the exponentiated matrix P^τ all sum to 1, so that row p_i^τ can be interpreted as the transitional probability of a random walk to go from point x_j to point x_i in τ steps. If the points in a cluster are mainly connected, even weakly, to their cluster neighbors (that is, the connectivity is geometrically well-defined), a random walk on the graph associated to the data set starting from one of those point will end inside the cluster with a higher probability than outside of it. This property also permitted Coifman to underline the relationship of P with heat diffusion with the definition of a diffusion distance between the points that is dependent on the parameter τ[22, 58]. For the scope of this thesis, it is sufficient to note that the values of P^τ can be calculated as simply as $\{\lambda_i^\tau\}_{i=1}^n$, and by giving a sufficiently high value to the parameter τ the spectrum of the exponentiated affinity matrix will tend to dichotomy, so to avoid criticities on the choice of the thresholding parameter ϵ.

Once selected the number q of eigenvectors to use, I have to chose how to extract the clustering information. Because of the piece-wise constant structure of the first eigenvectors of P^τ, and the relationship of the Euclidean distance between the points projected onto the eigenspace spanned by $\{u_i\}_{i=1}^q$ to a more general probabilistic distance [22, 58], many authors encourage the use of a simple algorithm like k-means carried out on the projected data set. In this chapter, thus I will use a variation of a much simpler and efficient clustering technique based on the eigenvectors $\{u_i\}_{i=1}^q$ proposed in [55], tht is, for a data set of n points $\{x_i\}_{i=1}^n \in \mathbb{R}^d$, given the clustering parameters (ϵ, τ, M):

- Build the affinity matrix A upon an appropriate connectivity measure — for example with Eq.(3.13).

- Calculate the normalized affinity matrix $P = D^{-1}A$.

- Solve the eigenproblem $PU = \lambda U$, where U is the matirx formed by the column eigenvectors $\{u_i\}_{i=1}^n$.

- Find the q eigenvalues $\{\lambda_i^\tau\}_{i=1}^q > 1 - \epsilon$.

- Take the eigenvectors $\{u_i\}_{i=1}^q$ to form the $n \times q$ matrix $\hat{U} = [u_1 \ldots u_q]$.

- For $j \in 1 \ldots n$, assign the data set point x_i to the cluster labeled by the position $arg\,max\,\{u_i(j)\}_{i=1}^q$.

- Join together the clusters with less than M elements.

The results of the application of this algorithm for the two previously described data sets are presented in Fig. 3.6. In both cases I used the gaussian kernel of Eq.(3.13) to build the affinity matrices, considering positions in $\mathbb{R}^2$ and using $\sigma = 0.75$ for the point data sets and $\sigma = 5$ for the segment data set. The parameters used for clustering were $\epsilon = 0.1$, $\tau = 100$ and $M = 10$. In the first case,
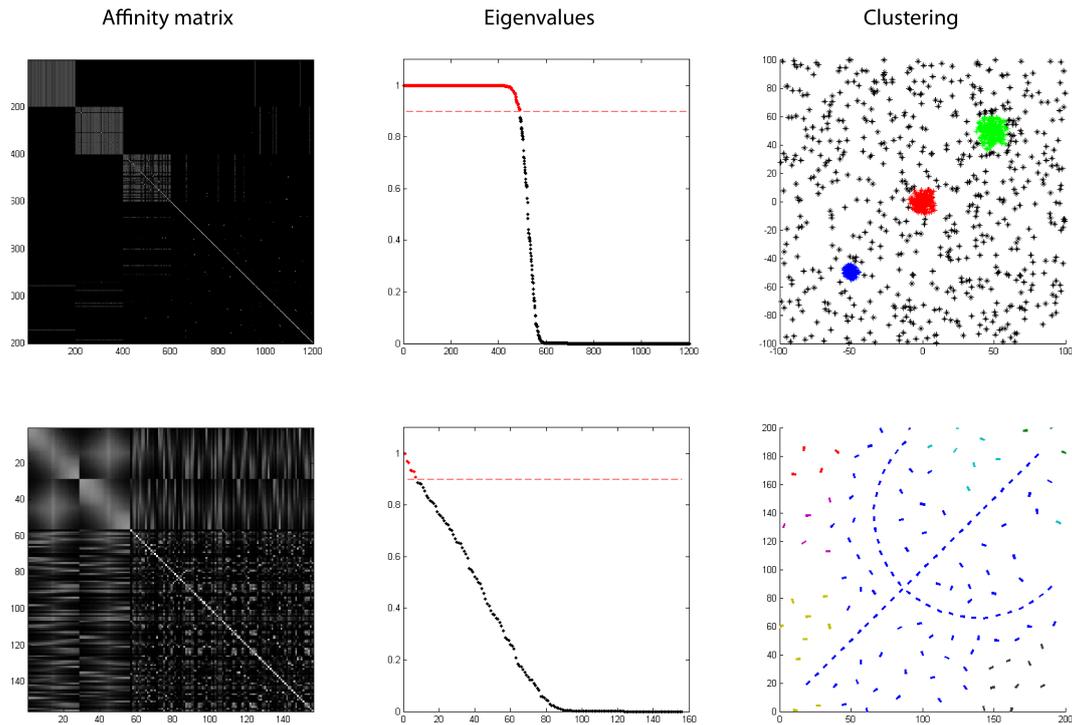
Figure 3.6: Result of the proposed algorithm for the two example data sets. The picture shows how a Gaussian kernel happens to be an optimal choice to cluster groups of points living in the image plane $\mathbb{R}^2$, but results inadequate when trying to separate boundaries or contours, naturally defined on the contact manifold $\mathcal{M}_q = \mathbb{R}^2 \times S^1$.

the algorithm performs the clustering process correctly, finding automatically the number of the main perceptual units and assigning the remaining elements to the noise/background cluster. It is worth noting that the spectrum of the normalized affinity matrix P^τ counts many eigenvalues that are close to 1, each only representing a single perceptual unit, that are mostly composed of few elements, as it is indicated by the associated eigenvectors. The segment data set, on the other hand, was not clustered correctly: this was predictable, as we have seen that these kind of stimuli lives on the position-orientation cortical domain $\mathbb{R}^2 \times S^1$, where to correctly assign a similarity measure between points I have to use the sub-Riemannian geometry described in section 2.

3.3.2 The cortical affinity matrix

As I have described in Section 3.2 and in Chapter 2, the connectivity patterns arising from many electrophysiological and phenomenological findings in V1 allow to naturally define a cortical-inspired distance with a related notion of affinity between points of a contact manifold $\mathcal{M}$, that in the proposed model represent cells in the visual cortex. As we have seen, this particular affinity measure can be modeled by the probability distributions Γ_0, and Γ, and I showed how to efficiently implement their estimates $\hat{\Gamma}_0$ and $\hat{\Gamma}$ by the integration of several stochastic paths generated by the processes (3.4) and (3.5). For convenience, in this section I will refer to these stochastic estimates using the original notation Γ_0 and Γ.

As we have seen in the previous section, the notion of distance given by the standard Euclidean geometry in the image plane lead us to good clustering results when applied to Gaussian distribution of points, but fails in finding the $\mathbb{R}^2 \times S^1$ lifted level set of boundaries and contours. Thus, it seems reasonable for the construction of the affinity matrix A to substitute the classical Gaussian kernel with the stochastic kernels previously described. Note that in doing that, I assume

a direct relationship between the concept of affinity and connectivity: the higher is the probability that two points $(\xi_i, \xi_j) \in \mathcal{M}_q$ are connected, the higher is their affinity or similarity.

I will proceed, thus, with the construction of the affinity matrix in the case of the segment sample data set of Fig. 3.6 using the kernel $\Gamma_q^{H,\kappa}$ as defined in Section 3. In general, for each stimulus like the one presented in the example (where to every point in $\mathbb{R}^2 \times S^1$ is assigned an activity equal to either 1 or 0) we will have n active points within the whole contact manifold $\mathcal{M}_q$, permitting us to define the $n \times n$ similarity matrix A_q as

$$A_q = a_q(i,j) = \Gamma_q^{H,\kappa}(\xi_i | \xi_j) = \frac{\sum_{h=0}^{H} \rho^\kappa(\xi_i, h | \xi_j, 0)}{\max \Gamma_q^{H,\kappa}}., \qquad (3.16)$$

where $(i,j) \in \{0 \ldots n\}$. The normalization obtained by dividing for the maximum kernel value allow us to formally define the relationship between the concepts of connectivity and affinity. This way, the point couple of maximum connection probability (in this model, the points with themselves) will have a unique maximum affinity value, that I arbitrarily set to 1. The definition of affinity matrices for stimuli living in the extended contact manifolds $\mathcal{M}_0$ and $\mathcal{M}$, given Eq. (3.16), is straightforward.

As the kernels described in Section 3 are a stochastic estimation of the theoretical solutions of the associated Fokker-Planck operators, in general we will not have a symmetric affinity matrix, though the kernel methods of spectral clustering and dimensionality reduction assume that the underlying affinity matrices are positive-definite real symmetric. To make the algorithm work, we thus obtain a symmetric matrix $\hat{A}$ by transposition:

$$\hat{A}_q = \hat{a}_q(ij) = \frac{a_q(i,j) + a_q(j,i)}{2}. \qquad (3.17)$$

Note that in doing this, I am just adjusting the stochastic errors to replicate the simmetry of the theoretical solutions. By using symmetrical kernels, derived from

the geometrical properties of the contact structure of the model, I am assuming that the modeled long-range horizontal connection between cells in the visual cortex are reciprocal. This is actually what happens in reality, as repeatedly found by different studies [13, 56].

The results of the application of the spectral clustering algorithm on the cortical affinity matrix are displayed in Fig. 3.7, and they show us that a good (geometrically and semantically correct) definition of distance or affinity between points of a complex domain can influence positively on the grouping capabilities of the method. In particular, I show 3 results obtained by using different sets of kernel parameters (H, κ). The clustering parameters used here and in the examples of the next sections were $\epsilon = 0.05$, $\tau = 150$ and $T = 3$. I maintained those parameters for all the tests, so to correctly analyze the differences in the grouping properties stricly relative to the kernels.

In the first case (top plots) I used $H = 40$ steps and an orientation diffusion coefficient $\kappa = 0.0140$ (the same value of the curvature of the semi-circular object). The algorithm clearly succeeds in distinguishing the two perceptual units from each other, and correctly assigns the remaining elements to the same background/noise partition, as I expect that the human eye would actually do. Changing the two parameters, though, greatly influences the results.

In the second case for example (middle plots), I reduced the value of the diffusion constant to $\kappa = 0.0035$: while the algorithm correctly retrieves the straight contour and distinguish the units from the background, the semi-circle gets over-partitioned. Note that, in fact, as in the previous case the affinity matrix is close to being a block diagonal matrix, but in this case the represented connected components of the sub-graph relative to the two objects are more than two.

Setting again the diffusion coefficient to the original value, in the third case I increased the step value to $H = 100$. Note that as the stimulus $\mathbb{R}^2$ domain
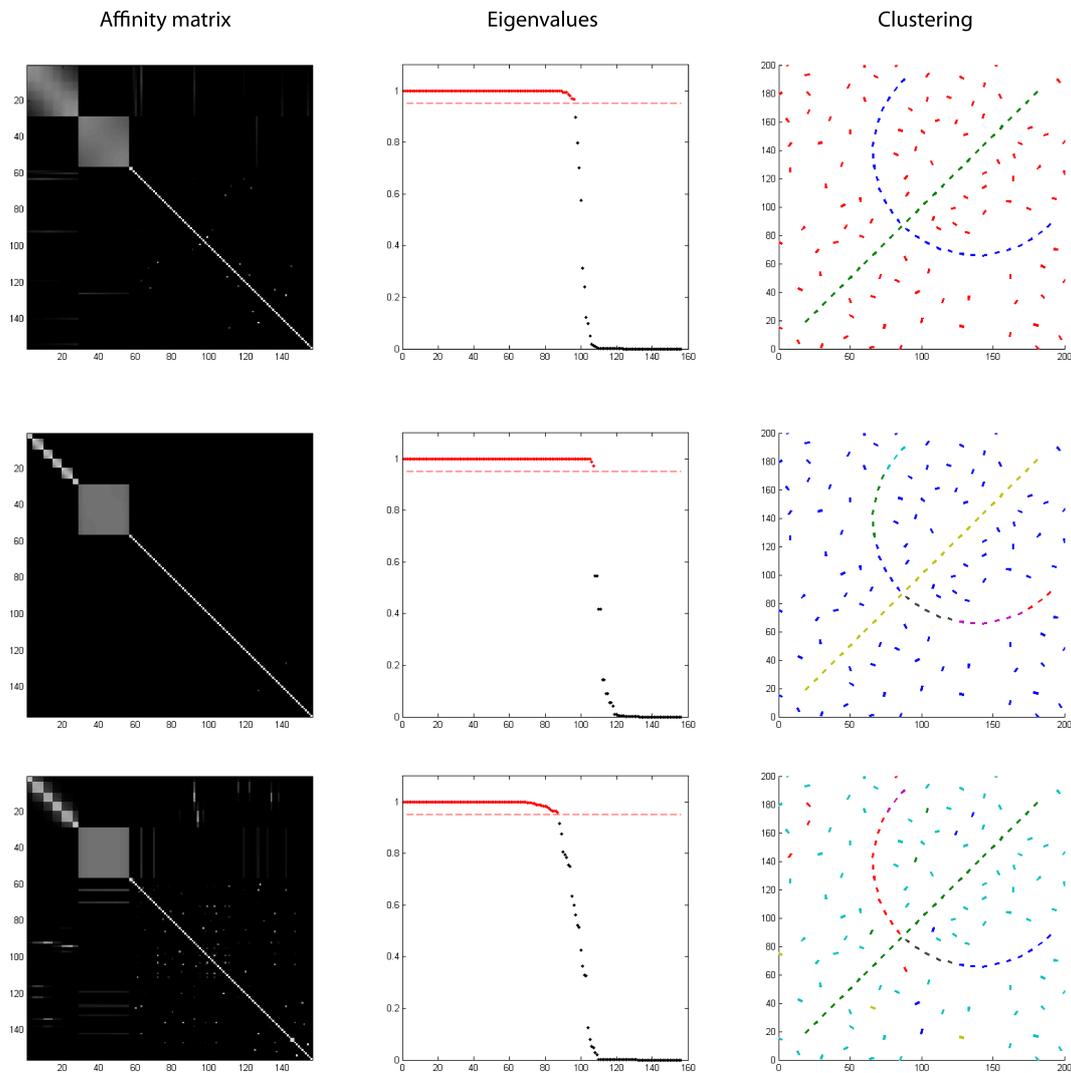
Figure 3.7: Result of the proposed algorithm for the second example data set and different parameters for the kernel Γ_q. Note how different kernel parameters modify the look of the affinity matrix, of its spectrum, and of the resulting data set partitioning.

is 200×200 pixel wide, this means that every segment in the example could potentially have a non-null connectivity value with almost an half of the other segments, if the co-circularity conditions are satisfied. Indeed, by observing A_q, I notice high affinity values between the objects and the random noise and in between the background elements. Again, even if the straight line is correctly retrieved as a single object, two random elements, approximately co-circular with the beginning of the line, are uncorrectly interpeted as being part of the object. Furthermore, the semi-circular contour gets again over-segmented, and many of the randomly collinear points, very far from each other, are interpreted as being a perceptual unit.

Since the stochastic kernel are good models of long-range horizontal cortical connectivity, it's interesting to better understand the origin of these differences in their grouping capabilities, as they could let us explore and better define their role in the cognitive process of spatio-temporal visual grouping. In the next section, I will carry out a simple parametric analysis of the stochastic kernels Γ_q, Γ_0 and Γ using the same spectral clustering algorithm described for the examples above.

3.4 Grouping properties of Γ_q, Γ_0 and Γ

In this section I try to simulate some of the cited grouping capabilities of the brain by assuming a central role played by horizontal connections in the cortical areas V1 and V5/MT of the visual cortex.

3.4.1 Spatial grouping using Γ_q or Γ_0

The perceptual bias towards collinear stimuli has classically been associated to the long-range horizontal connections linking cells in V1 having similar preferences in stimulus orientation. This specialized form of intra-striate connectivity pattern is

found across many species, including cats [56], tree shrews [15], and primates [4], the main difference being the specificity and the spatial extent of the connections. Furthermore, axons seem to follow the retinotopic cortical map anisotropically, with the axis of anisotropy being related to the orientation tuning of the originating cell [13]. Assuming that spatial integration, grouping and shape perception are fundamentally modulated by the position and the orientation of the elements in the visual space, here I use the clustering algorithm by building the affinity matrix A_q as previosly described using the kernel Γ_q.

But other prominent features of the visual stimulus, namely temporal dynamics, seem to play an important role in the spatial integration of oriented elements. The receptive fields (RFs) of orientation-selective cells in V1 have classically been modeled with two-dimensional Gabor functions [53, 54], which basically compute a local approximation of the directional derivative of the visual stimulus, minimizing the uncertainty between localization in position and spatial frequency [24]. In addition to stimulus orientation, cells in the striate areas are also selective for the direction of motion orthogonal to the cell's preferred orientation. These kind od preference is directly analyzable if one reconstructs the temporal dynamics of a neuron spatio-temporal RF, representing the motion pattern that would most excite the neuron [26]. These preferences are also structurally mapped in the cortical surface, with nearby neurons being tuned for similar motion direction [132], and it has been shown that excitatory horizontal connections in the V1 of the ferret are strictly iso-direction-tuned [99]. Modeling the grouping effect of these connectivities by building the affinity matrix A_0 using the extended kernel Γ_0, I want to analyze the influence of local velocuty on the detection of contours.

A set of points defined on the cortical manifolds $\mathcal{M}_q = \mathbb{R}^2 \times S^1$ and $\mathcal{M}_0 = \mathbb{R}^2 \times S^1 \times \mathbb{R}^+$ can be represented on the image plane by segments having orientation θ and moving with local velocity v in the corresponding $\vec{X}_3$ direction. In
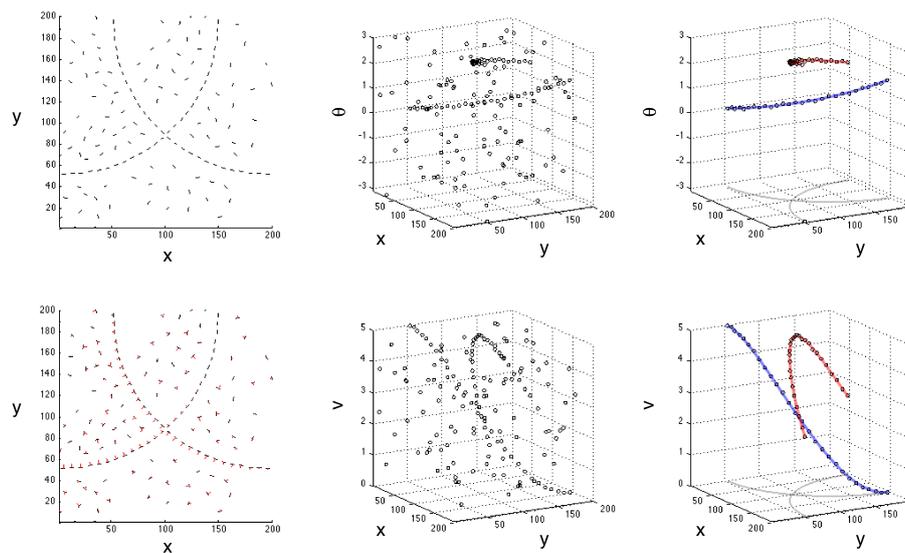
Figure 3.8: An example of the stimuli living in $\mathbb{R}^2 \times S^1 \times \mathbb{R}^+$ used in the tests for spatial grouping. The top row represents the stimulus with respect to the feature of orientation, while the bottom row represents the stimulus with respect to the feature of velocity. The perceptual units in this case have curvature $k = 0.023$ and the local velocity of istantaneous deformation changes sinusoidally along the contours.

particular, a set of points $C = \{\xi_i\}_{i=1}^n \in \mathcal{M}_0$ whose $\theta \in S^1$ values are mutually cocircular and whose $v \in \mathbb{R}^+$ values have low differences for neighboring points can be semantically regarded as a deforming contour or boundary. For the sake of clarity, I focused the first analyses on stimuli composed of two sets of this kind, together with r points placed at random coordinates within $\mathcal{M}_q$ and $\mathcal{M}_0$. The two perceptual units are always circles or circle arcs, so that both contours are characterized by a constant curvature value k, like the ones pictured in Fig. 3.8. I will refer to these stimuli with the notation S_r^k.

The analysis consists in the application of the method described in the previous section, with the same tuning parameters used for the three examples (i.e., $\epsilon = 0.05$, $\tau = 150$ and $T = 3$), on a set of different stimuli S_r^k, using various kernel parameter couples (H, κ). For each iteration I then evaluate the goodness of the grouping results by calculating a percentage error measure

$$E_P = \frac{E_1 + E_2 + E_3}{n}, \tag{3.18}$$

where n is the total number of points in the stimulus, and the numerator is obtained by summing together the number of points that were incorrectly assigned to the noise/background set (E_1), the number of random points that were incorrectly recognized as part of a perceptual unit (E_2), and the points pertaining to an over- or under-partitioned contour (E_3). In order to correctly compare results obtained by partitioning stimuli with a different number of random points, I calculated the mean percentage error measure $\hat{E}_P$ by averaging E_P over 100 repetitions, where for each repetition I changed the random part of the stimulus and calculated new kernel stochastic estimates Γ_q and Γ_0.

In Fig. 3.9a I show the results of the analysis of $S_{200}^{0.056} \in \mathbb{R}^2 \times S^1$, that is the stimulus in the space-orientation manifold composed of the perceptual units with the highest curvature and the highest number of random elements that I used in the analysis. The plot axes correspond to the kernel parameters (κ, H) with
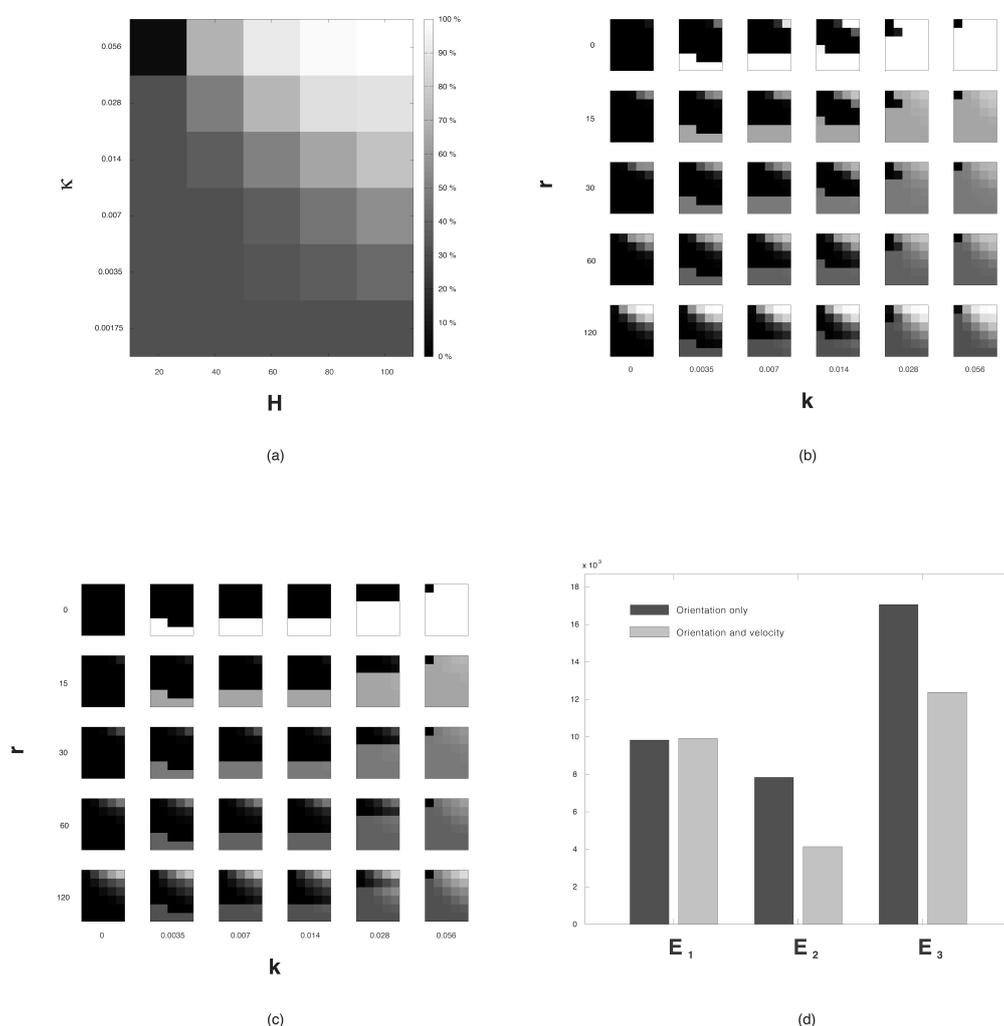
Figure 3.9: Parametric analysis for visual grouping in $\mathbb{R}^2 \times S^1$ and $\mathbb{R}^2 \times S^1 \times \mathbb{R}^+$. a) Grouping results of the algorithm applied to the stimulus $S_{120}^{0.056}$, every region of the image indicating one couple of kernel parameters (κ, H). The color intensity of each region is proportional to the mean percentage of misinterpreted points, calculated over 100 repeats. While taking local contour velocity greatly reduce grouping errors due to noise, the main effect of the added variable v on kernel parameters, compared to the position-orientation case, is to give low error indices also with a high number of kernel steps h.

which I have run the clustering algorithm, while color intensities represent the mean percentage error measure $\hat{E}_P$. From the color map we can visually notice some significative patterns: first of all, it is clear that the only set of kernel parameters capable to give a low grouping error value is $\kappa = 0.056, H = 20$, that is, the same curvature value of the contours in the stimulus and the shortest stochastic path length. Maintaining the kernel parameter H set to its minimum value and decreasing the kernel diffusion coefficient κ we have a constant error value $\hat{E}_P$, whose dominating component is given by E_1 and E_3. This means that, a change in the width of the fan of stochastic curves generating Γ_q (in this case, a reduction) impairs the connectivity between high curvature contour elements, so that the algorithm would see them as separate units and assign part or all of them to the background/noise set. It is worth noting that, regardless of the kernel diffusion coefficient, increasing the parameter H negatively impacts the goodness of the grouping, this time mainly because of the error component E_2. This was predictable, as longest stochastic path lengths can generate affinities between elements very distant from each other, a condition that with a high number of random elements can induce the algorithm in recognizing them as part of a distant contour. This gives its worse effects when both κ and H have high values, because with the associated kernel a point in $\mathcal{M}_q$ can potentially be connected to distant elements having a very different orientation. In such cases, also the error component E_3 gives its contribute, as the two perceptual units can be under-partitioned and interpreted as one unique object, having each one reciprocally affine contour elements.

Fig. 3.9b resumes the same kind of grouping analysis carried out with 30 different stimuli $S_r^k \in \mathcal{M}_q$. The first thing we can notice is the correlation between the stimulus contour curvature k and the smallest possible value of the kernel diffusion coefficient κ that does not present influences by the error component E_1.

Furthermore, while values smallest than these minimums lead to over-partitioned or unrecognized contours, increasing κ does not impairs the contour grouping capability of the algorithm, even if at wider κs the error component E_2 will grow together with the number of random elements in the space r and the length of the stochastic paths H. These correlations generally remain true also when significantly changing the parameters of the algorithm (data not shown). From this observation I can infer that the spread of the connectivity kernel Γ_q, governed by the diffusion coefficient κ, set a first functional constraint on its grouping capabilities. In general thus, I can say that object contours cannot be correctly segmented or recognized by the algorithm if the connectivity used to assign the affinity between their elements has a diffusion coefficient significantly below the contour's curvature.

In a second experiment, I repeated the same analysis with the same parameters with stimuli S_r^k defined in the extended space $\mathcal{M}_0 = \mathbb{R}^2 \times S^1 \times \mathbb{R}^+$, thus using the kernel stochastic estimate Γ_0 to build the affinity matrix between points. Each point composing these stimuli has a non-negative component of local velocity taken from the set $\{v | v \in \mathbb{R}, 0 \leq v \leq V_m\}$. I assigned to the elements of the perceptual units velocity values coherently with shape deformation. In particular, I let the values change sinusoidally, with a semi-period of L, i.e. the length of the contour, so that, parametrizing the curve of the stimulus by using the arc parameter s:

$$v(s) = \frac{V_m}{2}\sin(\frac{\pi s}{L}) + \frac{V_m}{2}, \tag{3.19}$$

For the sake of clarity, as we have seen in the previous experiment the grouping constraint that exist for the diffusion coefficient of the fiber variable θ, assuming that the same type of constraint would act also on the fiber variable v I set the local velocity diffusion coefficient α as

$$\alpha = 2\max\left|\frac{dv(s)}{ds}\right| = \frac{\pi}{L}V_m, \tag{3.20}$$

so that grouping capabilities will be preserved even at the maximum local speed change rate along the contour. For the analysis, I set $V_m = 5$ and let the noise/background elements of the stimuli have random velocity values within the range $(0, V_m)$.

In Fig. 3.9c I show the grouping results obtained by analyzing the stimuli $S_r^k \in \mathcal{M}_0$. Note that the stimuli differ from the ones defined on $\mathbb{R}^2 \times S^1$ just for the addition of the fiber variable v to each point in the space. From a first visual inspection, we can see that, while the correlation constraint between the contour curvature k and the orientation diffusion coefficient κ is still present, the influence of the error component E_2 at the highest values of κ and H is significantly reduced, if not almost completely eliminated for the stimuli having fewer random elements. This is mainly due to the fact that when considering just the orientation of a segment, two aligned samples p_1 and p_2 moving in opposite directions ($\theta_2 = \theta_1 + \pi$) result to be very affine to each other, while they should be considered very different objects. When velocity information is added, on the other hand, this ambiguity ceases to exist, as the directions of moving segments cover the whole S^1 domain without repetitions.

The beneficial effect of the additional fiber variable v on the grouping capabilities of the algorithm is also evidenced in Fig. 3.9d, where, in order to obtain an error measure to globally and quantitatively compare the results obtained with Γ_q and Γ_0, I integrated the three error components E_1, E_2 and E_3 over all the stimuli, the kernel parameters and the repetitions. From this plot we can clearly see that while both kernels tend to confuse the perceptual units with noise approximately in the same way (E_1), information on local velocity greatly improves the behavior of the algorithm in the presence of noise, correctly assigning the random elements to the background, as I have previously commented (E_2). Furthermore, the third group of bars shows us that by using Γ_0 in the grouping algorithm helps to avoid
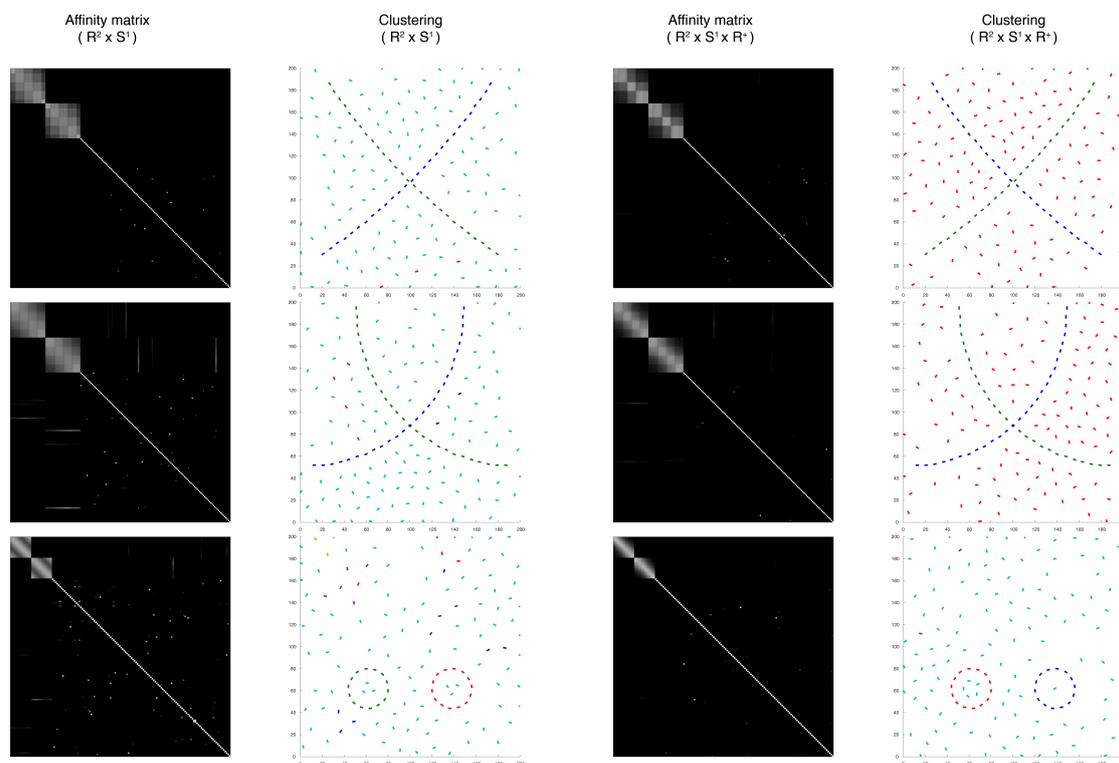
Figure 3.10: Results of the proposed algorithm for data set with perceptual units with different curvatures. Left column: results obtained by using the position-orientation connectivity given by Γ_q. Right column: results obtained by using the position-orientation-velocity connectivity given by Γ_0. It is worth noting that the affinity matrices built with Γ_0 are more clean than those built with Γ_q: they generally avoid spurious affinities between perceptual units and noise, while maintaining the approximate object block diagonal structure that is essential for a correct clustering.

the under-partitioning of the stimulus when the parameters (κ, H) of the kernel make it a long range, widespread connectivity tool.

Some punctual examples of the mentioned improvement in the grouping capabilities are pictured in Fig. 3.10: in all three cases, the use of Γ_0 instead of Γ_q concurs in reducing different kinds of grouping errors of the algorithm. The consistency and generality of these results show the strong effects of augmenting the dimensionality of the visual space on visual perception and object recognition, thus suggesting an additional functionality, that of noise/surround suppression, for the presence of specialized neurons and connectivities in in the first visual cortical areas of mammals, such as neurons optimized for the measurement of local velocities (see Chapter 1) and long-range horizontal connections between movement-sensitive cells [56, 132] in V1.

3.4.2   Spatio-temporal grouping using Γ_0 and Γ

At the physiological level, the estimation of stimulus motion is an elaboration process that has classically been associated with neurons in visual area MT/V5, for their high selectivity in direction of movement and their extended sensitivities to a wide range of stimulus velocities [68, 98]. Extra-striate areas are retinotopically organized, and the horizontal connectivity pattern is highly structured, with anisotropic and asymmetric connectivity bundles reaching columns of cells tuned for similar orientation and direction preference [65]. This highly organized functional architecture can be related to the different spatial symmetries of the spatial surround suppression found by Xiao et al, who found cells with inhibitory regions placed on one or both sides of the axis of motion direction [139].

Though the importance of this area in complex motion and depth perception, grouping and segmentation is widely recognized, decades of neurophysiology have undiscovered just a little part of the mechanisms therein implemented. Even the
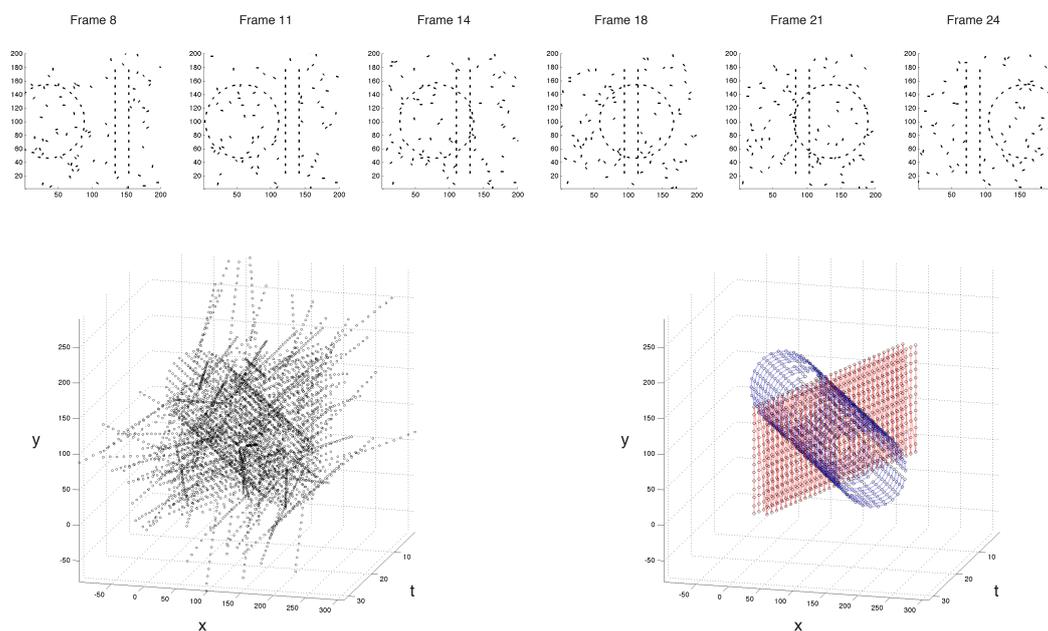
Figure 3.11: An instantiation of the stimulus living in $\mathbb{R}^2 \times \mathbb{R}^+ \times S^1 \times \mathbb{R}^+$ used in the tests for spatio-temporal grouping of moving shapes. The circle has curvature $k = 0.02$ and its velocity is $7,5$ units/frame, that is twice that of the bars $(3,75$ units/frame).

functional difference for the estimation of motion with the primary visual cortex, which serves as its major input and to which projects strong feedback connections, has not yet been clearly outlined [16]. More generally, striate and extra-striate cortical areas seem to cooperate, and surround modulation in V1 can be given by the connectivity patterns implemented in both areas by means of fast feedforward and feedback inter-areal projections [5].

The stimulus that I use in this simulation is the one pictured in Fig. 3.11: I created a set o points $\eta_i = (x_i, y_i, t_i, \theta_i, v_i) \in \mathcal{M} = \mathbb{R}^2 \times \mathbb{R}^+ \times S^1 \times \mathbb{R}^+$, represented in the figures by segments moving in time for a total of $n_t = 32$ frames, forming a circular shape of curvature $k = 0.02$ and two bars translating in opposite directions.

Note again that the fiber cohordinate of local velocities v_i of each point represents just the projection onto $\vec{X}_3$ of the real velocity vector, driving the rigid movement of the circle and of the bars. Similarly to what I did in the previous example, the total number n of points in the stimulus if obtained by adding a variable number r of noise/background elements, each one having a rigid motion path in their $\vec{X}_3$ direction that remains consistent over all the stimulus frames. In this case, the aim of the grouping algorithm will be to carry out a segmentation of the complete spatio-temporal surfaces representing the moving objects.

In Section 2, I have described two different connectivity kernels, one modeling the interactions between points of a motion contour (Γ_0), and one specifically modeled for motion integration of point trajectories (Γ). As the domain on which the first kernel is defined, $\mathcal{M}_0$, doesn't include a temporal dimension, it is straightforward to conclude that to carry out the detection of contours and shapes over time, i.e. to perform the grouping of the spatio-temporal surfaces depicted in Fig. 3.11, I need to combine the influence of both definitions of affinity between points.

At the psychophysiological level, Ledgeway et al confirmed the presence of more than one grouping law governing the detection of contours, with different underlying implementing structures, and that the composition of them is compatible with the probability summation hypothesis [60]. Thus, to account for the interaction between striate (V1) and extra-striate (V5/MT) connectivities, I decided to sum together the connection probability between points of the manifold by building two distinct $n \times n$ affinity matrices. For the first I use the kernel for motion contours, Γ_0, to assign affinities between the points at the same temporal position:

$$A_0 = a_v(i,j) = \frac{\Gamma_0(\xi_i|\xi_j)}{\max \Gamma_0} \quad , \quad \xi_i, \xi_j \in \mathcal{M}_0 = \mathbb{R}^2 \times S^1 \times \mathbb{R}^+, \qquad (3.21)$$

so that we have $a_v(i,j) = 0$ for all the points η_i and η_j with temporal coordinates $t_i \neq t_j$. A second affinity matrix is then calculated by assigning affinities between
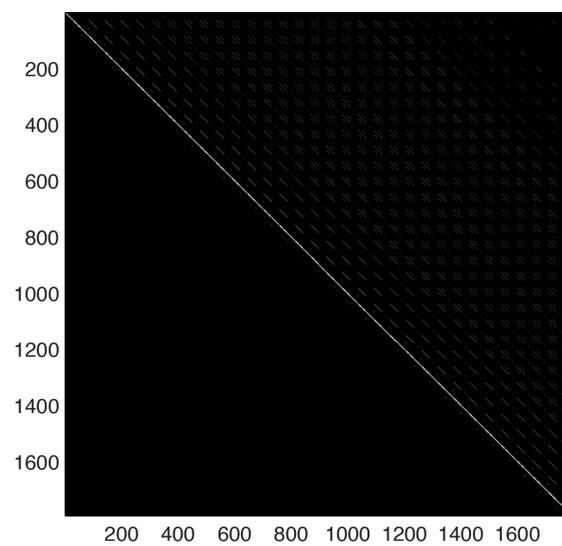
Figure 3.12: Affinity matrix A_{xt} of the spatio-temporal stimulus of Fig.3.11 without background/noise elements. Due to the intrinsic causality of Γ, the matrix is real positive but not symmetric. This will have no consequences for the clustering algorithm, as described in the text.

points using Γ on the entire 5-dimensional manifold $\mathcal{M}$:

$$A = a_t(i,j) = \frac{\Gamma(\eta_i|\eta_j)}{\max \Gamma} \quad , \quad \eta_i, \eta_j \in \mathcal{M} = \mathbb{R}^2 \times \mathbb{R}^+ \times S^1 \times \mathbb{R}^+ \qquad (3.22)$$

Sticking to the assumption of probability summation of [60], thus, I create a global $n \times n$ affinity matrix $A_{xt} = A_0 + A$. It is worth noting that, due to the characteristics of the model, I am assuming an instantaneous effect of the connectivity defined with Γ_0, meaning high horizontal transmission velocities at the physiological level. On the other hand, this kind of connectivity could be thought as an approximation of the steady state of an unmodeled feedback/feedforward interaction system between movement-sensitive cells of the visual cortex.

Another important thing to note is that the kernel Γ intrinsically implements causality, so that $a_t(i,j) = 0$ if $t_i > t_j$ and A, A_{xt} are strongly asymmetric matrices (see Fig. 3.12). At first thought, this could potentially affect the efficiency of the grouping algorithm, as the previously cited dimensionality reduction methods have always been defined for real positive symmetric affinity matrices. Indeed, the first eigenvectors of the transition probability matrix $P_G = D_G^{-1} A_{xt}$ will in general be complex. Even if the operation of symmetrizing the matrix by using its transpose, by generating a reversible Markov chain on an undirected graph, might seem a reasonable and practical solution, it has been shown that it leads to a strong loss of information, thus preventing any algorithm from a correct grouping or partitioning of the data set to cluster [71]. Following these issues, here I use the method proposed in [87] by changing the algorithm accordingly, so that the clustering is not carried out by taking the first q eigenvectors of P_G, $\{u_i\}_{i=1}^q$, but the composition of their real and imaginary parts:

$$\{\bar{u}_i\}_{i=1}^q = \Re\{\{u_i\}_{i=1}^q\} + \Im\{\{u_i\}_{i=1}^q\}. \qquad (3.23)$$

The results obtained by using the modified algorithm for various instances of the circle/bars stimulus, each time by varying the number of background/noise
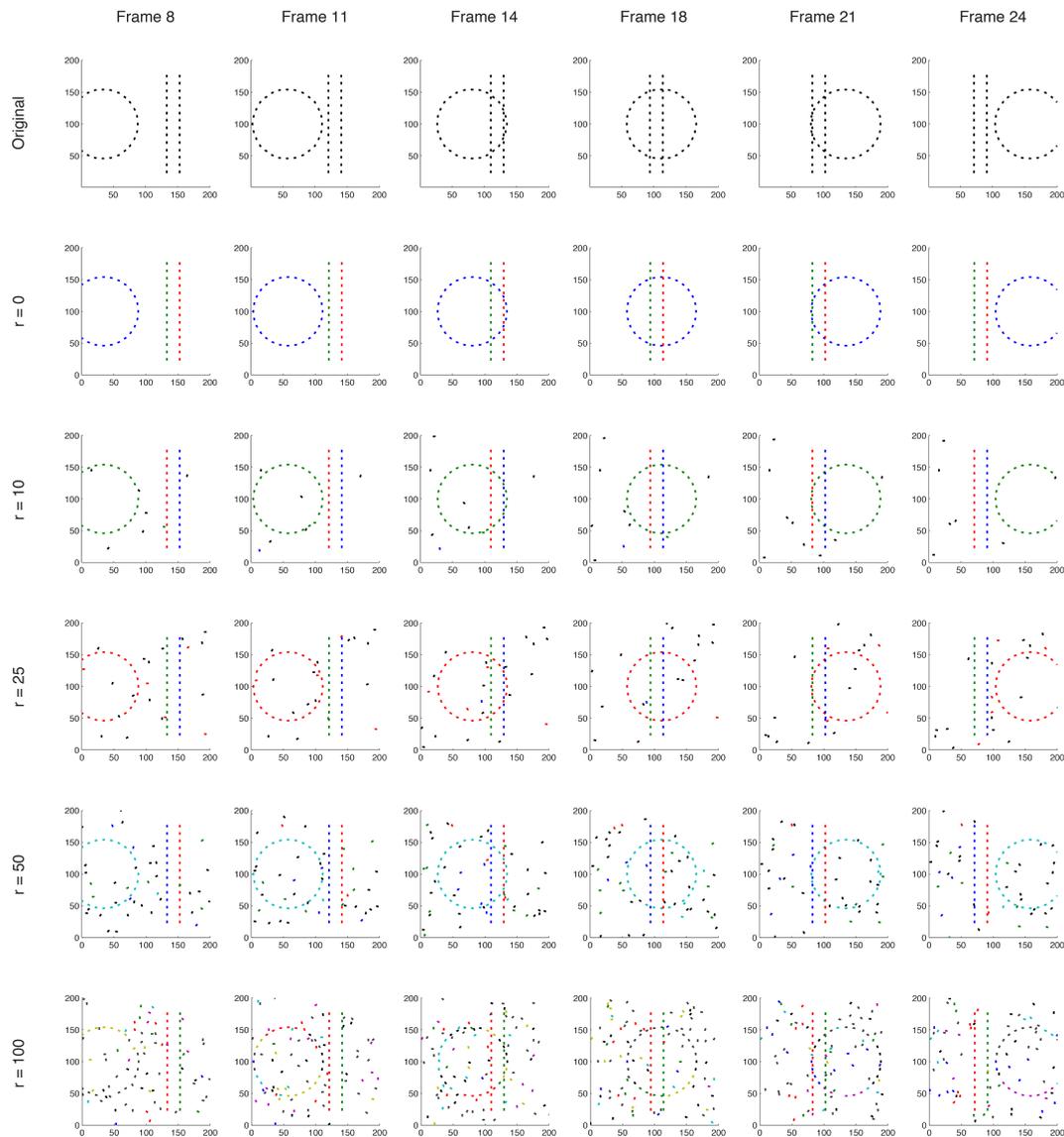
Figure 3.13: Results obtained by using Γ_0 and Γ. The grouping is succesful at all noise/background conditions apart from $r = 100$, where the algorithm fails by over-partitioning the countours of the moving circle.

elements r, are displayed in Fig.3.13. The parameters chosen to run the algorithm were $\epsilon = 0.01$, $\tau = 150$ and $T = 3n_t$, and the number of stochastic paths was set to $H = 40$ for both Γ_0 and Γ. While the diffusion coefficients of Γ_0 were set accordingly to the logic used in the previous analysis, for Γ I used a larger α parameter, to take into account the considerations made in [46], where it is specified that changes in the local velocity of a motion contour impairs the perception of the visual unit significantly less if the velocity is tangential to the contour, thus defining a partial trajectory. In particular all the results shown in the next figures were obtained by setting $\kappa_v = \kappa_t = 0.7$ for both kernels, and a diffusion coefficient over velocity of $\alpha_v = 0.5$ and $\alpha_t = 1$.

From the figure it is possible to see how the composition of the kernels with which I built the affinity matrix A_{xt} allows to correctly recognize the spatio-temporal surfaces relative to the moving circle and the bars, and to separate clearly their boundaries from the background (always plotted in black for clarity) and between themselves. When the number r of random elements was lower than 25, I always obtained the correct clustering for various stimulus instances. For higher noise values, though, the algorithm began to give poor grouping results, as pictured in the bottom row of Fig.3.13 in the case of $r = 100$ random segments. The example with $r = 50$ was the best grouping result I obtained among many clustering failures.

It is worth noting that even in the case with the highest noise values, the bars are always correctly retrieved, so in that case the algorithm is failing to detect only contours with higher curvature confusing them with the background segments, thus leading to an over-partitioning of the perceptual units. In general, though, I show that the connectivity kernels defined by the proposed cortical-inspired geometrical model applied to a simple spectral clustering algorithm are able to carry out a non-trivial grouping task. To better understand the powerful mechanics involved

in the calculations, let's consider for example that the only segments of the circle that present a positive affinity value with their corresponding points at future temporal positions are the ones having an orientation value near $\pm\frac{\pi}{2}$, as only for them the vector field $\vec{X}_5$ of connectivity propagation has the same direction of the global movement of the shape.

In fact, while the ability and the reliability of visual neurons in areas V1 and MT/V5 in measuring local stimulus orientation and speed have been studied extensively, the majority of cells in those areas respond solely to the local characteristics they are tuned for. In doing so, the measurements available in the first stages of the visual cortex are subject to the well-known aperture problem: with no information other than the local direction of movement, it cannot be said much about the real direction and speed of the object to which that local measurement refers. For a continuously moving contour, for example, classical orientation- and direction-selective cortical cells measure, for each position along the contour, just the velocity component that is orthogonal to the contour tangent direction at that point (see Fig. 3.14, left plot). In the framework presented throughout the thesis, this is modeled so that the fiber variable of local velocity refers to movements in the $\vec{X}_3$ direction.

Even if this component alone is ambiguous and generally not sufficient for making global assumptions, in this section we have seen that spatio-temporal grouping is still possible. To analyze the influence of the locality of the velocity/orientation measurement on grouping, in the following section I tested the algorithm by first solving the problem of aperture, and then using the recovered global information to partition the visual space.

Solving the aperture problem using Γ_0

The ability of the visual cortex in resolving the aperture problem has been widely investigated, both from the phenomenological and the physiological point of view. In [62, 112], it was observed that the velocity of a moving line is often underestimated when its orientation diverges from being orthogonal to the direction of movement, and that this bias is dependent on line orientation, speed, contrast and length. The study of the temporal dynamics of neurons in visual area MT showed that after an orientation-selective transient response, the cells begin to being activated by stimuli moving in the preferred direction, irrespective to their orientation [82]. As demonstrated in [83], this kind of neurons is also found in V1, and their capability to decode real movement directions strongly depends on the information presented in their receptive field surround. In that experiment, the cells managed to resolve the aperture problem more easily if the line used as stimulus didn't extend too far beyond their classical receptive field. In [44, 64] this notion has been extended by the finding that a consistent subpopulation of cells in V1 is capable to measure global directions even if the oriented stimulus length was well beyond receptive field surround.

The cortical mechanisms underlying the capability of resolving the aperture problem by integrating motion information are still unclear. All of the cited studies, together with some of the models proposed in literature (see [78, 113, 103]) stress the importance to rely on unambiguous measurements, such as the output of motion end-stopping cells, in order to recover real motion. While this is true in the case of a line moving continuously in space, we don't always need a discontinuity along a contour to estimate its axis of motion when its tangent orientations vary smoothly, think for example to the translating circle of Fig. 3.11. The literature is rich of models that are capable to solve the geometrical problem also known as intersection of constraints (IOC) using multiple neural layers linked by
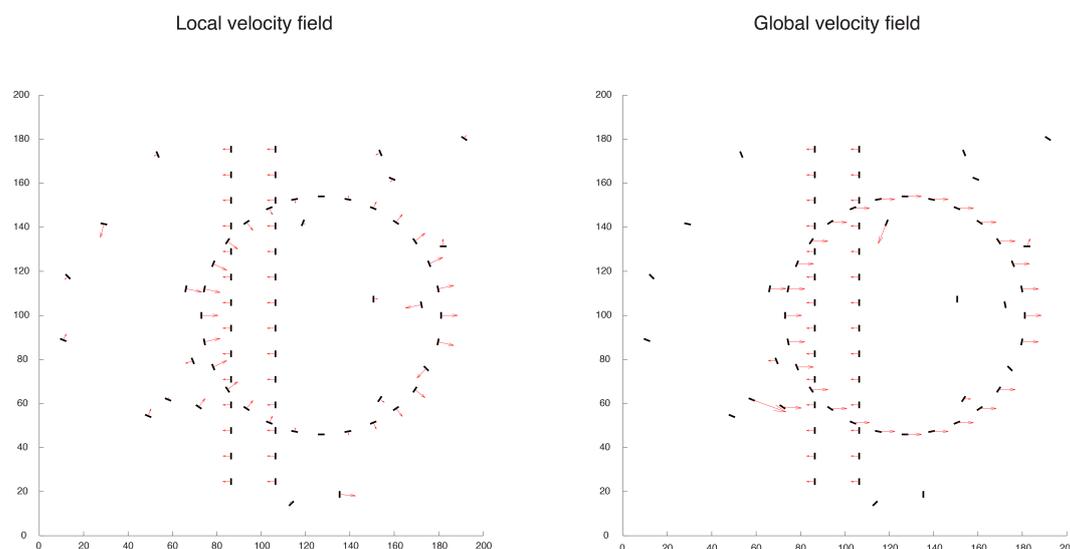
Local velocity field

Global velocity field

Figure 3.14: Local versus global velocity fields plotted for the points belonging to the stimulus at a given temporal position. In the global case, the elements pertaining to the circles all share the object's direction ϕ and velocity ω.

feed-forward, feed-back and horizontal connections. In particular, the most widely accepted model of V5/MT neuron integrates the responses of the subset of local movement-selective V1 cell whose orientation and velocity preferences are coherent with a specific global direction and speed [64]. Since the model was originally defined in the spatio-temporal Fourier space, the pooling region over space of the input was initially modeled with an iso-tropic Gaussian kernel. Recently, though, it has been shown that the V5/MT cells can pool input signals from non-trivial neighborhood regions that are strongly anisotropic, with a possible preference for geometrical patterns [103].

Following those observations, here I want to simulate a part of the cortical elaboration of global spatio-temporal features by assuming that the connectivity in V1 modeled by Γ_0 strongly influences the integration of local measurements. First of all, I solve the IOC system associated to each couple of points in the

stimulus, so to build two matrices of estimated global velocity components

$$
\begin{aligned}
V_x &= v_x(i,j) = \frac{v_i \cos\theta_j - v_j \cos\theta_i}{\sin(\theta_j - \theta_i)} \\
V_y &= v_y(i,j) = \frac{v_i \sin\theta_j - v_j \sin\theta_i}{\sin(\theta_j - \theta_i)},
\end{aligned}
\tag{3.24}
$$

where for point couples having $\theta_i = \theta_j$ the velocity components of the two segments were simply averaged. Note that in this case I am addressing the problem of aperture with an implementation that is the reverse of the supposed physiological one: for each point ξ_i, we have $n-1$ global estimates that reflect the feature preferences of MT/V5 cells. Among those possibilities, I want to chose the one that would most probably generate a response from the extra-striate neurons, by choosing a method to weigh, or pool, the inputs from the each point's surround. To do so I average all the global estimates, weighing them with the horizontal connectivity model Γ_0 by first modifying and normalizing A_0, so that self-similarity is excluded from the weighted mean

$$
P_s = D_s^{-1} A_s,
\tag{3.25}
$$

where $A_s = A_0 - I$ and D_s is the diagonal degree matrix of A_s. Since in this way the rows of P_s all sum to 1, obtaining the wighted average is as simple as multiplying it element-wise with V_x and V_y

$$
\begin{aligned}
v_x(i) &= \sum_{j=1}^{n} v_x(i,j) p_v(i,j) \\
v_y(i) &= \sum_{j=1}^{n} v_y(i,j) p_v(i,j),
\end{aligned}
\tag{3.26}
$$

so that I can assign two new fiber variables to each point of the stimulus, indicating its real (global) direction and velocity of movement:

$$
\begin{aligned}
\phi_i &= \operatorname{atan2}(v_x(i), v_y(i)) \\
\omega_i &= \sqrt{v_x(i)^2 + v_y(i)^2}.
\end{aligned}
\tag{3.27}
$$

In Fig.3.14 I show the resulting velocity vector field compared to the one of local velocities. While for the elements of the bars $\theta_i = \phi_i$ and $v_i = \omega_i$, the segments composing the circle and the background have different associated directions and velocities. In particular we see that this method assigns to all the elements of the circle the values (ϕ, ω) of the entire shape, irrespective of their local orientation θ. Note that the object's direction has been assigned also to a background element (on the left), as its local features matched significantly the connectivity conditions imposed by the circle.

A second model for spatio-temporal grouping

Once I have calculated the global velocity of each point of the stimulus, I can define a new manifold, $\mathcal{M}_g$, composed of the three spatio-temporal base variables (x, y, t) and the fiber variables (ϕ, ω). Similarly to what I described in the first section, in this space we can think about a propagation along the direction

$$\vec{X}_5^g \;\; = \;\; (\omega \cos \phi, \omega \sin \phi, 1, 0, 0), \tag{3.28}$$

forced by a diffusion over the new fiber variables. The propagation will be described by an analogous stochastic differential system, but in this case the diffusion coefficients κ and α will drive the two dimensional Brownian motion over, respectively, ϕ and ω. The stochastic process associated to the system leads as before to the definition of a connectivity kernel Γ_g on $\mathcal{M}_g$, giving us the probability that two points of the manifold are connected by a random path generated by one of them. Using the new kernel, I can build a modified affinity matrix for the points of the stimulus defined on $\mathcal{M}_g$,

$$A_t^g = a_t^g(i, j) = \tfrac{\Gamma_g(\eta_i | \eta_j)}{\max \Gamma_g} \;\; , \;\; \eta_i, \eta_j \in \mathcal{M}_g = \mathbb{R}^2 \times \mathbb{R}^+ \times S^1 \times \mathbb{R}^+ \tag{3.29}$$

to add to the previously calculated A_0 in order to proceed with the clustering algorithm.
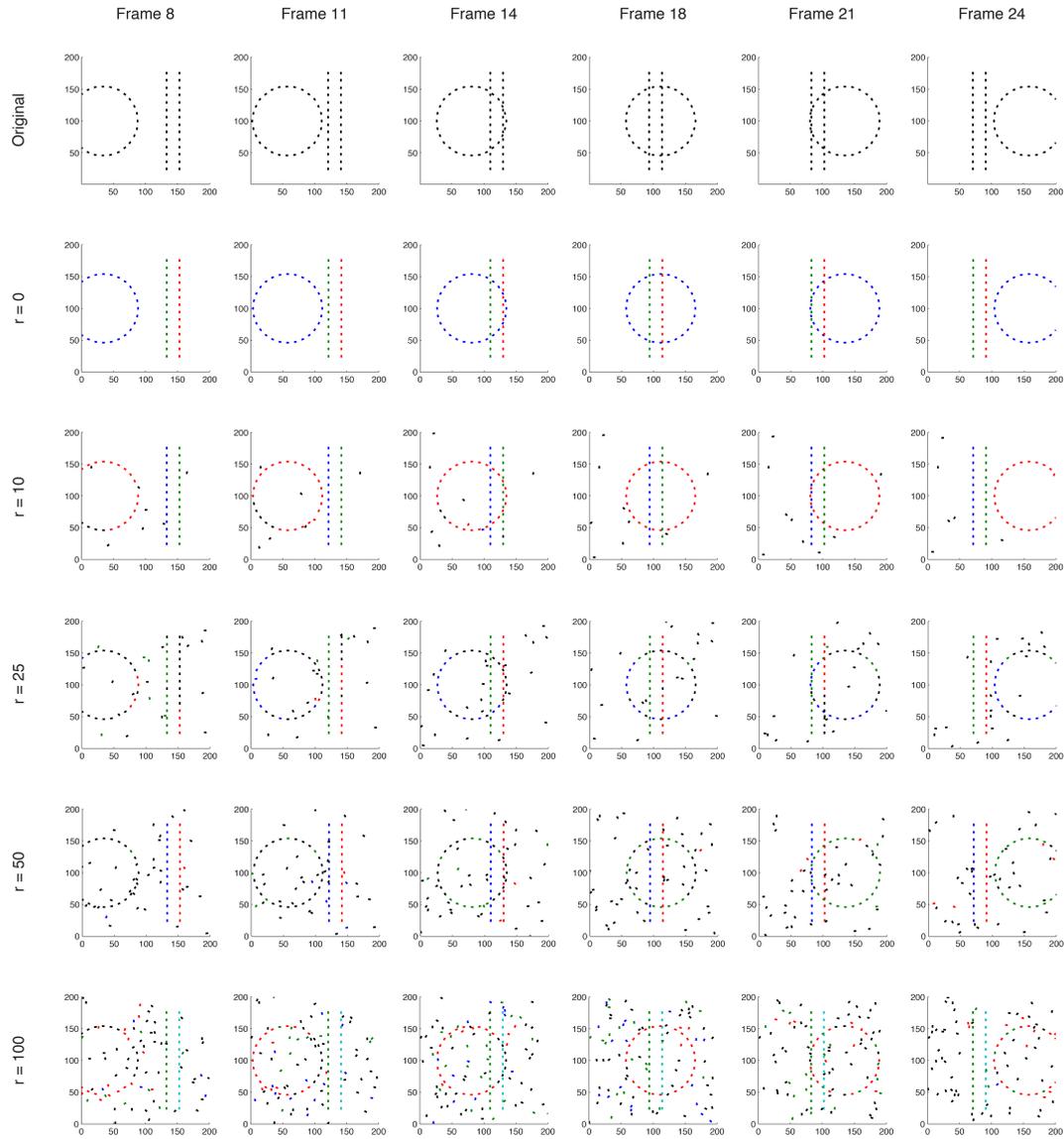
Figure 3.15: Results obtained by using Γ_0 and Γ_g. Grouping capabilities are maintained even at the highest number of background random elements, even if the circle seem to be recognized only in the late part of the stimulus.

The results of this analysis carried out on the same stimuli of the previous experiment are displayed in Fig.3.15. The diffusion coefficients κ and α were not changed. The plots clearly show that also by using the kernel Γ_g, the spectral clustering algorithm is capable of grouping the elements of the circle and the bars in space-time. Furthermore, this time the objects are being correctly recognized also at the highest random/background noise condition, with $r = 100$, even if some of the random elements have been assigned to a perceptual unit.

Thus, while the underlying physiological mechanism of motion integration is still an open subject of research, assuming that the algorithm for spatio-temporal visual grouping is a valid methodology to test the characteristics of different connectivity kernels, I showed two possible geometries that are both capable of solving the segmentation problem of recognizing moving contours and shapes from background/noise. My preliminary results seem to suggest that by delevolping a geometry based on estimated global velocity, we get a reduced sensitivity to noise. Future experiments, though, should concentrate on analyzing the grouping effect of the kernel parameters H, κ and α, by carrying out a statistical analysis similar to the one proposed in this chapter for the connectivity in $\mathcal{M}_0$.

It is also worth noting the tendency of the algorithm to correctly recognize all the elements of the circle as being part of the same object mainly after a certain temporal position. Even if this behavior was detected in some instances also when using Γ, the consinstency with which I obtain this effect with Γ_g suggests to explore more deeply the relationships between the algorithm parameter τ, used by Coifman to explain non-linear diffusion on the graphs represented by real positive affinity matrices, and the temporal base bariable t. A possible extension of the geometrical model proposed in Chapter 2, that adds the temporal variable to Γ_0 to model real cortical propagation, and that analyze the effect of that modification on visual grouping capabilities, could be the object of a future work.

Conclusions

The field of cognitive neuroscience is a very fascinating area of research, containing a wide range of subjects from many disciplines, such as physiology, psychology, mathematical modeling, computer science, etc. In particular, a full and complete understanding of the mechanics underlying visual cognition could have a strong impact in a huge variety of sectors, like medical and biomedical science or more generally, image analysis.

In this work I tried to explore many aspects of cognitive visual science, each one based on different academic fields, proposing mathematical models capable to reproduce both neurophysiological and phenomenological results that were described in the recent literature. The structure of my thesis is mainly composed of three chapters, corresponding to the three main areas of research on which I focused my work. The results of each work put the basis for the following, and their ensemble form an homogeneous and large-scale survey on the spatio-temporal properties of the architecture of the visual cortex of mammals.

Starting from the very basis of neurophysiology of the primary visual cortex, in Chapter 1 I explored the spatio-temporal characteristics of the direction-selective cells in V1, by reconstructing a data set of spatio-temporal receptive profiles belonging to both simple and complex cells. The available raw data were cat recordings recollected with three different experimental procedures (oriented bars, natural stimuli, 1D Gaussian noise). The RF profiles were then fitted using a 3D Gabor

model, which has the property of minimizing uncertainty on stimulus localization simultaneously over both space and time.

It has been established that RF size has an upper and an inferior limit. The scale factor is indeed one of the main physiological constraints limiting the subset of Gabor functions used in the primary visual cortex to process visual information, and the range of sizes effectively present in the cortex is much narrower than the one defined by the spatial limit boundaries. Spatial and temporal frequencies measured by a single neuron are also important parameters linked to the spatial and temporal sizes, and somewhat constrained by them, as I have previously seen.

Nevertheless, the limits in the parameters that are found within V1 are much stronger than the ones dictated by physical reality or logical assumptions. It appears that only a relatively small portion of the entire 3D Gabor model family is used by the cortex and some of the parameters show very strong linear correlations or non-linear relationships between each other. Some of these relationships have already been described in the literature, but few studies have been carried out on the time-space simultaneous cross-correlations between properties of cells belonging to V1 as a whole, describing a general spatio-temporally organized architecture.

These results are also consistent with the anatomical studies that show V1 to be one of the main inputs to the MT (Middle-Temporal) area, also known as V5, whose neurons are highly selective for stimulus velocities. RF profiles of cells belonging to this movement-dedicated area of the cortex are represented within the spatio-temporal Fourier plane by skewed elongated ellipsoids that follow iso-velocity lines with relative precision [91]. Indeed, the relationship between spatial and temporal parameters, found in this work by modeling RF profiles of simple and complex cortical cells with three-dimensional Gabor functions, seem to form an ideal basic architecture for building elements that have similar responses to the ones present in the MT area.

Possible future developments could include the confirmation of these findings through specifically targeted experimental procedures; the construction of a mathematical framework able to better describe the functional spatio-temporal architecture of V1, introducing for example a formal uncertainty principle for the velocity variable; the design of a numerical simulation of the visual information processing labor done within the cortical stage; and finally, the construction of solid algorithms for the compression and the processing of movies or image sequences, based on the efficiency rules followed by the brain throughout the visual pathway.

Then in Chapter 2, I proposed a model of cortical functional architecture for the processing of spatio-temporal visual information. Motion features are detected first by simple and complex cells with RP modeled by 2D+time Gabor filters. Linear filtering with such a profiles lifts the visual stimulus from the base space $\mathbb{R}^3$ to the phase space $\mathbb{R}^6$ comprising spatial and temporal frequencies, in which a Liouville form can be defined.

I defined a 5D phase space with fixed frequency by taking a reduction of the previous differential form. Then, exploiting the commutation properties of the horizontal basis, I regarded the tangent space of the contact manifold $\mathcal{M}$ as a possible constraint acting on the connectivity between points, giving the definition of admissible integral, or horizontal, curves. Possible linear combinations of the horizontal basis that are compatible with the definition of admissible curve have been studied, in order to model the possible lifting of the visual stimuli as association fields, in the sense of [34]. I considered the corresponding deterministic integral curves for two modeling limit cases: contours in motion and trajectories of a point in motion.

Then I considered horizontal stochastic paths, i.e. trajectories of points that always move along the tangent space of $\mathcal{M}$ and are allowed to change the value associated to the fiber variables (θ, v) in a random, equidistributed way. I have

seen that the resulting integrations over the evolution parameter of the paths coincide with the kernels of the Fokker Planck operators defined on the geometries. Both deterministic curves and stochastic kernels inherit the symmetries of the non associative reduced Galilean group as described in Appendix.

After a discussion about the compatibility of the presented theoretical framework with various phenomenological and psycho-physiological findings on visual perception and cognition reported in the literature, I introduced the stochastic kernels as facilitation inducers in a neural population activity model.

Suitable numerical simulations have been carried out, by processing pre-determined artificial stimuli with the neural population activity model previously described, showing that the Fokker Planck stochastic kernels endow the model with the capability of completion and continuation of contours in motion and trajectories of points, coherently with the phenomenological experiments of Rainville [92] and Wu et al [138]. In conclusion, results have shown that the proposed functional geometry is compatible with existing psychophysical and physiological experiments, even if a complete knowledge about the effective neural implementation needs supplementary empirical data.

In Chapter 3, I described and implemented a clustering algorithm in order to test the visual grouping properties of the connectivity kernels that were previously introduced. In particular, I used one of the most recent dimensionality reduction methods [58, 22] to perform spectral clustering on the spatio-temporal cortical feature space of position, orientation and velocity, by using the stochastic cortical kernels Γ, Γ_0 and Γ_t instead of the classical Gaussian kernel to assign a similarity measure to each couple of points. The algorithm is capable of grouping together elements belonging to a single contour or shape moving in time, forming a spatio-temporal surface, distinguishing them from a noisy background composed of randomly placed elements.

The first analysis that I carried out was based on phenomenological findings that abrupt changes in local velocity can drastically impair the perception of a contour, even if its elements are collinear to each other [92]. The grouping results of the algorithm confirmed that if the affinity between points in the cortical feature space is assigned not only following their position in $\mathbb{R}^2 \times S^1$, but also taking local velocity into account, we get a strongly reduced negative influence of random elements and a significantly lower percentage of grouping errors due to over- or under-partitioning, thus allowing the horizontal connectivity to be spatially extended without suffering noise, as it happens in the visual cortex [13].

A second analysis extended the algorithm in order to use also the connectivity kernel for motion integration, to test if its combination with Γ_0, that is defined as an instantaneous connectivity, could be capable to segment entire spatio-temporal surfaces drawn by contours and shapes moving in time. The connectivity model was also modified, defining a new affinity kernel Γ_g, so to take into account global direction and velocity information at each point, and to compare the algorithm outcome with its local counterpart. The results showed that the integration of local informations about movement, with the use of Γ_0 and Γ_t, manage to give a good segmentation at global scale, even if by integrating global information in the first place by using Γ_g noise seems to be less effective.

The neurophysiological counterparts of the proposed connectivities have been discovered and confirmed only in part, while it is not yet completely clear if they may have multiple functional roles. The models proposed in this thesis may be regarded more as perceptual and phenomenological modeling than strictly physiological models of cortical connectivity, even if in Chapter 2 their use for the simulation of neural population response lead to the reproduction of physiological non-linearities that were previously unexplained. At the physiological level, the mechanisms involved in the perception of complex shapes and motion patterns are

likely to be implemented differently (see for example the proposed method to estimate global velocity at the end of Chapter 3, against the functionality of V5/MT neurons described and modeled in [103, 113]).

Finally, the flexibility and the generality of the geometrical framework where I defined the mathematical tools used in this work, make possible to extend and to define many new differential constraints on the cortical feature space, leading to new stochastic connectivity kernels, that are a powerful tool to explore, compare and parametrize the actual horizontal connections in the visual cortex. For example, it would be interesting to introduce the scale or color variables in the geometry, defining their implications on the connectivity between points, for it has been suggested in the literature that different stimulus scales are probably processed by different connectivity networks [60], or to slightly modify the model to try to reproduce additional neurophysiological findings. Furthermore, the results of my work could be used as a suggestion for future phenomenological or electro-physiological experiments, trying to tune the model's parameters in order to fit real visual perception and cognition behaviors.

Bibliography

[1] Adelson, E.H., Bergen, J.R.: Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A 2(2), 284-299 (1985)

[2] Ahmed, B., Cordery, P.M., McLelland, D., Bair, W., Krug, K.: Long-range clustered connections within extrastriate visual area V5/MT of the rhesus macaque. Cereb. Cortex 22(1), 60-73 (2012)

[3] Ambrosio, L., Masnou, S.: A direct variational approach to a problem arising in image recostruction. Interfaces and Free Boundaries 5 (1), 63-81 (2003)

[4] Angelucci, A., Levitt, J.B., Walton, E.J.S., Hupe, J., Bullier, J., Lund, J.S.: Circuits for local and global signal integration in primary visual cortex. J. Neurosci. 22(19), 8633-46 (2002)

[5] Angelucci, A., Bressloff, P.C.: Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Visual Perception 154(A), 93-120 (2006)

[6] August, J., Zucker, S. W.: The curve indicator random field: curve organization via edge correlation. In: Boyer, K., Sarkar, S. (eds.) Perceptual organization for artificial vision systems, pp. 265-288. Kluwer (2000)

[7] August, J., Zucker, S. W.: Sketches with curvature: the curve indicator random field and Markov processes. IEEE Trans. Pattern Analysis and Machine Intelligence 25(4), 387-400 (2003)

[8] Barbieri, D., Citti, G., Sarti, A.: How Uncertainty bounds the shape index of simple cells. To appear on The Journal of Mathematical Neuroscience.

[9] Barbieri, D., Citti, G., Cocci, G., Sarti, A.: A cortical-inspired geometry for contour perception and motion integration. J. Math. Imag. Vis. Jan (2014)

[10] Boscain, U., Duits, R., Rossi, F., Sachkov, Y: Curve cuspless reconstruction via sub-Riemannian geometry. arXiv:1203.3089 (2012)

[11] Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. Neur. Comp. 15(6), 1373-1396 (2003)

[12] Blanche, T.J., Spacek, M.A., Hetke, J.F., Swindale, N.V.: Polytrodes: high density silicon electrode arrays for large scale multiunit recording. J. Neurophysiol. 93, 2987-3000 (2005)

[13] Bosking, W.H., Zhang, Y., Schofield, B., Fitzpatrick, D.: Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. J. Neurosci. 17(6), 2112-27 (1997)

[14] Chirikjian G. S.: Stochastic Models, Information Theory, and Lie Groups, Volume 2. Birkhäuser 2012.

[15] Chisum, H.J., Mooser, F., Fitzpatrick, D.: Emergent properties of layer 2/3 neurons reflect the collinear arrangement of horizontal connections in tree shrew visual cortex. J. Neurosci. 23(7), 2947-60 (2003)

[16] Curchland, M.M., Priebe, N.J., Lisberger, S.G.: Comparison of the Spatial Limits on Direction Selectivity in Visual Areas MT and V1. J. Neurophysiol. 93(3), 1235-1245 (2005)

[17] Cleland, B.G., Dubin, M.W., Levick, W.R.: Sustained and transient neurones in the cats retina and lateral geniculate nucleus. J. Physiol. 217, 473-496 (1971)

[18] Citti, G., Manfredini, M., Sarti, A.: Neuronal oscillations in the visual cortex: convergence to the Riemannian Mumford-Shah functional. SIAM Journal of Mathematical Analysis 35 (6), 1394-1419 (2004).

[19] Citti, G., Sarti, A.: A cortical based model of perceptual completion in the roto-translation space. J. Math. Imag. Vis. 24(3), 307-326 (2006)

[20] Cocci, G., Barbieri, D., Sarti, A., Spatio-temporal receptive fields of cells in V1 are optimally shaped for stimulus velocity estimation. J. Opt. Soc. Am. A 29(1), 130-138 (2012)

[21] Cocci, G., Barbieri, D., Citti, G., Sarti, A.: Spatio-temporal grouping properties of cortical-inspired connectivity models. (submitted article)

[22] Coifman, R.R., Lafon, S.: Diffusion maps. Appl. Comp. Harm. Anal. 21, 530 (2006)

[23] CRCNS Data Sharing website. *http://crcns.org*

[24] Daugman, J., Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J. Opt. Soc. Am. A 2(7), 1160-1169 (1985)

[25] DeAngelis, G.C., Ohzawa, I., Freeman, R.D.: Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. J. Neurophysiol. 69(4), 1091-1117 (1993)

[26] DeAngelis, G.C., Ohzawa, I., Freeman, R.D.: Receptive-field dynamics in the central visual pathways. Trends Neurosc. 18, 451-458 (1995)

[27] DeAngelis, G.C., Ghose, G.M., Ohzawa, I., Freeman, R.D.: Functional micro-organization of primary visual cortex: receptive field analysis of nearby neurons. J. Neurosc. 19, 4046-4064 (1999)

[28] Duits, R., Franken, E.M.: Left invariant parabolic evolution equations on SE(2) and contour enhancement via invertible orientation scores, part I: Linear left-invariant diffusion equations on SE(2). Q. Appl. Math. 68, 255-292 (2010)

[29] Duits, R., Franken, E.M.: Left invariant parabolic evolution equations on SE(2) and contour enhancement via invertible orientation scores, part II: Nonlinear left-invariant diffusion equations on invertible orientation scores. Q. Appl. Math. 68, 293-331 (2010)

[30] Duits, R., Führ, H., Janssen, B., Bruurmijn, M., Florack, L., Van Assen, H.: Evolution equations on Gabor transforms and their applications. arXiv:1110.6087 (2011)

[31] Duits, R., van Almsick, M.: The explicit solutions of linear left-invariant second order stochastic evolution equations on the 2D Euclidean motion group. Quart. Appl. Math. 66, 27-67 (2008)

[32] Ermentrout, G.B., Cowan, J.D.: Large scale spatially organized activity in neural nets. SIAM J. Appl. Math. 38(1), 1-21 (1980)

[33] Fermuller, C., Ji, H., Kitaoka, J.: Illusory motion due to causal time filtering. Vision Research 50(3), 315-329 (2009)

[34] Field, D.J., Hayes, A., Hess, R.F.: Contour integration by the human visual system: evidence for a local association field. Vision Res. 33, 173-193 (1993)

[35] Folland, G.B.: Harmonic analysis on phase space. Princeton University Press (1989)

[36] Gabor, D.: Theory of communication. J. IEE 93, 429-59 (1946)

[37] Geiges, H.: An introduction to contact topology. Cambridge University Press (2008)

[38] Geisler, W.S.: Motion streaks provide a spatial code for motion direction. Nature 400(6739), 65-69 (1999)

[39] Geisler, W.S., Perry, J.S., Super, B.J., Gallogly, D.P.: Edge co-occurrence in natural images predicts contour grouping performance. Vision Res. 41, 711-724 (2001)

[40] Geisler, W.S., Albrecht, D.G., Crane, A.M., Stern, L.: Motion direction signals in the primary visual cortex of cat and monkey. Visual Neurosc.. 18(4), 501-516 (2001)

[41] Graham, N.V.: Beyond multiple pattern analyzers modeled as linear filters (as classical V1 simple cells): Useful additions of the last 25 years. Vision Research 51, 1397-1430 (2011).

[42] Gromov, M.: Carnot-Carathéodory spaces seen from within. Progr. Math. 144, 79-323 (1996).

[43] Grzywacz, N.M., Watamaniuk, S.N.J., McKee, S.P.: Temporal coherence theory for the detection and measurement of visual motion. Vision Res. 35(22), 3183-203 (1995)

[44] Guo, K., Robertson, R., Nevado, A., Pulgarin, M., Mahmoodi, S., Young, M.P.: Primary visual cortex neurons that contribute to resolve the aperture problem. Neuroscience 138(4), 1397-1406 (2006)

[45] Hartline, H.K.: The response of single optic nerve

bers of the vertebrate eye to illumination of the retina. J. Physiol. 121, 400-415 (1938)

[46] Hess, R.F., Ledgeway, T.: The detection of direction-defined and speed-defined spatial contours: One mechanism or two? Vision Research 43, 597606 (2003)

[47] Hladky, R.K., Pauls, S.D.: Minimal Surfaces in the Roto-Translation Group with Applications to a Neuro-Biological Image Completion Model. J. Math Imaging Vis 36, pp. 1-27 (2010)

[48] Hoffman, W.C.: The Lie Algebra of Visual Perception. Journal of Mathematical Psychology 3, 65-98 (1966)

[49] Hoffman, W.C.: The visual cortex is a contact bundle. Appl. Math. Comput. 32, 137-167 (1989)

[50] Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. 119(1), 147-171 (1967).

[51] Hubel, D.H.: Eye, Brain and Vision. Scientific American Library, New York (1988)

[52] Hubel, D.H., Wiesel, T.N.: Ferrier lecture: Functional architecture of macaque monkey visual cortex. Proceedings of the Royal Society of London Series B 198(1130), 1-59 (1977)

[53] Jones, J.P., Palmer, L.A.: The two-dimensional spatial structure of simple receptive fields in cat striate cortex. J. Neurophysiol. 58(6), 1187-1210 (1987)

[54] Jones, J.P., Palmer, L.A.: An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. J. Neurophysiol. 58(6), 1233-1258 (1987)

[55] Kannan, R., Vempala, S., Vetta, A.: On clusterings: Good, bad and spectral. Journal of the ACM 51(3), 497515 (2004)

[56] Kisvarday, Z.F., Toth, E., Rausch, M., Eysel, U.T.: Orientation-specific relationship between populations of excitatory and inhibitory lateral connections in the visual cortex of the cat. Cereb. Cortex 7(7), 605-18 (1997)

[57] Koenderink, J.J., van Doorn, A.J: Representation of local geometry in the visual system. Biol. Cybernet. 55,367-375 (1987)

[58] Lafon, S., Lee, A. B.: Diffusion maps and coarse-graining: A unified framework for dimensionality reduction, graph partitioning, and data set parameterization. IEEE Trans. Pattern Anal. Mach. Intell. 28(9), 1393-403 (2006)

[59] Ledgeway, T., Hess, R.F.: Rules for combining the outputs of local motion detectors to define simple contours. Vision Research 42, 653659 (2002)

[60] Ledgeway, T., Hess, R.F., Geisler, W.S.: Grouping local orientation and direction signals to extract spatial contours: empirical tests of "association field" models of contour integration. Vision Res. 45(19), 2511-22 (2005)

[61] Lee, T.S., Nguyen, M.: Dynamics of subjective contour formation in the early visual cortex. Proc. Natl. Acad. Sci. USA 98(4), 1907-11 (2001)

[62] Lorenceau, J., Alais, D.: Form constraints in motion binding. Nat. Neurosci. 4(7), 745-51 (2001)

[63] S. Mahamud, L. R. Williams, K. K. Thornber, K. Xu: Segmentation of multiple salient closed contours from real images. IEEE Transactions on pattern analysis and machine intelligence 25, PP-PP (2003)

[64] Majaj, N.J., Carandini, M., Movshon, J.A.: Motion integration by neurons in macaque MT is local, not global. J. Neurosci. 27(2), 366-370 (2007)

[65] Malach, R., Schirman, T.D., Harel, M., Tootell, R.B., Malonek, D.: Organization of intrinsic connections in owl monkey area MT. Cereb. Cortex 7(4), 386-93 (1997)

[66] Marcelja S.: Mathematical description of the responses of simple cortical cells. J. Opt. Soc. Am. 70 (1980)

[67] Maunsell, J.H.R., Gibson, J.R.: Visual response latencies in striate cortex of the macaque monkey J. Neurophysiol. 68, 1332-1344 (1992)

[68] Maunsell, J.H.R., Van Essen, D.C.: Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. J. Neurophysiol. 49(5), 1127-1147 (1983)

[69] Mechler, F., Ringach, D.L.: On the classification of simple and complex cells Vision Research 42, 1017-1033 (2002)

[70] Meila, M., Shi, J: A random walks view of spectral segmentation. 8th International Workshop on Articial Intelligence and Statistics (2001)

[71] Meila, M., Pentney, W.: Clustering by weighted cuts in directed graphs. Proceedings of the 2007 SIAM International Conference on Data Mining (2007)

[72] Montgomery R., A tour of subriemannian geometries, their geodesics and applications. AMS 2002.

[73] Mumford, D.: Elastica and computer vision. In: Bajaj, C.L. (eds.) Algebraic Geometry and its Applications, pp. 491-506. Springer-Verlag (1994)

[74] Mumford, D., Shah, J.: Optimal approximation by piecewise smooth functions and associated variational problems. Comm. Pure Appl. Math. 17, 577-685 (1989)

[75] Nagel, A., Stein, E.M., Wainger, S.: Balls and metrics defined by vector fields I: Basic properties. Acta Math. 155, 103-147 (1985)

[76] Ng, A.Y., Jordan, M.I., Weiss, Y.: On spectral clustering: analysis and an algorithm. Cambridge, MA. Advances in Neural Information Processing Systems 14. MIT Press (2002)

[77] Nitzberg, M., Mumford, D., Shiota, T.: Filtering, segmentation and depth. Lecture Notes in Computer Science 662, Springer 1993.

[78] Nowlan, S.J., Sejnowski, T.J.: A selection model for motion processing in area MT of primates. Journal of Neuroscience, 15, 1195-1214 (1995)

[79] Ohzawa, I., DeAngelis, G.C., Freeman, R.D.: Encoding of binocular disparity by simple cells in the cats visual cortex. J. Neurophysiol. 75, 1779-1805 (1996)

[80] Øksendal, B.: Stochastic Differential Equations: an Introduction with Applications. Springer (2010)

[81] Olson, I.R., Gatenby, J.C., Leung, H.C., Skudlarski, P., Gore, J.C.: Neuronal representation of occluded objects in the human brain. Neuropsychologia 42(1), 95-104 (2004)

[82] Pack, C., Berezovskii, V.K., Born, R.T.: Dynamic properties of neurons in cortical area MT in alert and anaesthetized macaque monkeys. Nature 414, 905-908 (2001)

[83] Pack, C., Livingstone, M.S., Duffy, K.R., Born, R.T.: End-stopping and the aperture problem: two-dimensional motion signals in macaque V1. Neuron 39(4), 671-680 (2003)

[84] Pan, Y., Chen, M., Yin, J., An, X., Zhang, X., Lu, Y., Gong, H., Li, W., Wang, W.: Equivalent representation of real and illusory contours in macaque V4. J. Neurosci. 32(20), 6760-70 (2012)

[85] Pauls, S.D., Hladky, R.K.: Minimal surfaces in the roto-translation group with applications to a neuro-biological image completion model. J. Math. Imaging Vis. 36(1), 563-594 (2010)

[86] Pavan, A., Casco, C., Mather, G., Bellacosa, R. M., Cuturi, L. F., Campana, G.: The effect of spatial orientation on detecting motion trajectories in noise. Vis. Res. 51, 2077-2084 (2011)

[87] Pentney, W., Meila, M.: Proceedings of the 20th national conference on Artificial intelligence (AAAI) 2, 845-850 (2005)

[88] Perona, P., Freeman, W.T.: A factorization approach to grouping. Proc. ECCV, 655-670 (1998)

[89] Petitot, J., Tondut, Y.: Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux. Math. Sci. Hum. 145, 5-101 (1999)

[90] Petkov, N., Subramanian, E.: Motion detection, noise reduction, texture suppression and contour enhancement by spatiotemporal Gabor filters with surround inhibition. Biol. Cybern. 97(5-6), 423-439 (2007)

[91] Priebe, N.J., Casanello, C.R., Lisberger, S.G.: The neural representation of speed in macaque area MT/V5. J. Neurosc. 23, 5650-5661 (2003)

[92] Rainville, S.J.M., Wilson, H.R.: Global shape coding for motion-defined radial-frequency contours. Vision Res. 45(25-26), 3189-201 (2005)

[93] Reid, R.C., Soodak, R.E., Shapley, R.M.: Directional selectivity and spatiotemporal structure of receptive fields of simple cells in cat striate cortex. J. Neurophysiol. 66, 505-529 (1991)

[94] Ringach, D.L.: Spatial structure and symmetry of simple-cell receptive fields in macaque primary visual cortex. J. Neurophysiol. 88(1), 455-463 (2002)

[95] Ringach, D.L., Hawken, M.J., Shapley, R.: Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. J. Vision 2, 12-24 (2002)

[96] Ringach, D., Shapley, R.: Reverse correlation in neurophysiology. Cognitive Science 28(2), 147-166 (2004)

[97] Rodieck, R.W.: Quantitative analysis of cat retinal ganglion cell response to visual stimuli. Vision Research 5, 583-601 (1965)

[98] Rodman, H.R., Albright, T.D.: Coding of visual stimulus velocity in area MT of the macaque. Vis. Res. 27(12), 20352048 (1987)

[99] Roerig, B., Kao, J.P.: Organization of intracortical circuits in relation to direction preference maps in ferret visual cortex. J Neurosci. 19(24) (1999)

[100] Rothschild, L.P., Stein, E.M.: Hypoelliptic differential operators and nilpotent Lie groups. Acta Math. 137, 247320 (1977)

[101] Robert, C.P., Casella, G.: Monte Carlo statistical methods. Springer, 2nd Edition (2004)

[102] Roweis, S., Saul, L.: Nonlinear dimensionality reduction by locally linear embedding. Science, 290(5500), 2323-2326 (2000)

[103] Rust, N.C., Mante, V., Simoncelli, E.P., Movshon, J.A.: How MT cells analyze the motion of visual patterns. Nature Neurosc. 9(11), 1421-1431 (2006)

[104] Sabinin, L.V.: Smooth quasigroups and loops. Kluwer (1999)

[105] Sanguinetti, G., Citti, G., Sarti, A.: Image Completion Using a Diffusion Driven Mean Curvature Flowin A Sub-Riemannian Space. VISAPP (2)'08, 46-53 (2008).

[106] Sanguinetti, G., Citti, G., Sarti, A.: A model of natural image edge co-occurrence in the rototranslation group. J. Vis. 10(14) (2010)

[107] Sarti, A., Citti, G., Petitot, J.: The symplectic structure of the primary visual cortex. Biol. Cybern. 98, 33-48 (2008)

[108] Sasaki, S.K., Ohzawa, I.: Internal spatial organization of receptive fields of complex cells in the early visual cortex. J. Neurophysiol. 98, 1194-1212 (2007)

[109] Scholl, B.J., Pylyshyn, Z.W.: Tracking multiple items through occlusion: clues to visual objecthood. Cogn. Psychol. 38(2), 259-290 (1999)

[110] Sherrington, C.S.: The integrative action of the nervous system. C. Scribner and Sons, New York (1906)

[111] Shi, J., Malik, J.: Normalied cuts and image segmentation. Proc. IEEE Conf. Comp. Vis. and Pattern Recogn, 731-737 (1997)

[112] Series, P., Georges, S., Lorenceau, J., Fregnac, Y.: Orientation dependent modulation of apparent speed: a model based on the dynamics of feed-forward and horizontal connectivity in V1 cortex. Vis. Res. 42(25), 2781-2797 (2002)

[113] Simoncelli, E.P., Heeger, D.J.: A model of neuronal responses in visual area MT. Vis. Res. 38(5), 743-761 (1998)

[114] Sorba, P.: The Galilei group and its connected subgroups. J. Math. Phys 17(6), 941-953 (1976)

[115] Tan, Z., Yao, H.: The spatiotemporal frequency tuning of LGN receptive field facilitates neural discrimination of natural stimuli. J. Neurosc. 29, 11409-11416 (2009)

[116] Theunissen, F.E., David, S.V., Singh, N.C., Hsu, A., Vinje, W.E., Gallant, J.L.: Estimating spatio-temporal receptive fields of auditory and visual neurons from their responses to natural stimuli. Network 12, 289-316 (2001)

[117] Touryan, J., Lau, B., Dan, Y.: Isolation of relevant visual features from random stimuli for cortical complex cells. J. Neurosc. 22 10811-10818 (2002)

[118] Verghese, P., Watamaniuk, S.N.J., McKee, S.P., Grzywacz, N.M.: Local motion detectors cannot account for the detectability of an extended trajectory in noise. Vision Res. 39(1), 19-30 (1999)

[119] Verghese, P., McKee, S.P., Grzywacz, N.M.: Stimulus configuration determines the detectability of motion signals in noise. J. Opt. Soc. Am. A 17(9), 1525-34 (2000)

[120] Verghese, P., McKee, S.P.: Predicting future motion. J. Vis. 2(5) (2002)

[121] Verghese, P., McKee, S.P.: Motion grouping impairs speed discrimination. Vis. Res. 46(8-9), 1540-1546 (2006)

[122] Visiome Network website. *http://visiome.neuroinf.jp.*

[123] Von Luxburg, U.: A tutorial on spectral clustering. Statistics and Computing 17(4), 395-416 (2007)

[124] Wagemans, J., Elder, H. J., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., von der Heydt, R.: A Century of Gestalt Psychology in Visual Perception: I. Perceptual grouping and Figure-Ground Organization. Psychol Bull. 138(6), 1172-217 (2012).

[125] Watamaniuk, S.N.J.: The predictive power of trajectory motion. Vision Res. 45(24), 2993-3003 (2005)

[126] Watamaniuk, S.N.J., MkKee, S.P.: Seeing motion behind occluders. Nature 377(6551), 729-30 (1995)

[127] Watamaniuk, S.N.J., McKee, S.P., Grzywacz, N.M.: Detecting a trajectory embedded in random-direction motion noise. Vision Res. 35(1), 65-77 (1995)

[128] Watson, A.B., Ahumada, A.J. Jr: A look at motion in the frequency domain. J. K. Tsotsos (Ed.). Motion: Perception and representation, 1-10 (1983)

[129] Watson, A.B., Ahumada, A.J. Jr: Model of human visual-motion sensing. J. Opt. Soc. Am. A 2, 322-342 (1985)

[130] Watson, A.B., Ahumada, A.J. Jr, Farrell, J.A.: Window of visibility: a psychophysical theory of fidelity in time-sampled visual motion displays. J. Opt. Soc. Am. A 3, 300-307 (1986)

[131] Weiss, Y.: Segmentation using eigenvectors: a unifying view. Proc. IEEE Int. Conf. on Comp. Vis, 975-982 (1999)

[132] Weliky, M., Bosking, W.H., Fitzpatrck, D.: A systematic map of direction preference in primary visual cortex. Nature 379(6567), 725-728 (1996)

[133] Weng, C., Yeh, C., Stoelzel, C.R., Alonso, J.M.: Receptive field size and response latency are correlated within the cat visual thalamus. J. Neurophysiol. 93, 3537-3547 (2005)

[134] Wertheimer, M.: Laws of organization in perceptual forms. Ellis, Willis D. (Ed). A source book of Gestalt psychology, 71-88 (1938)

[135] Whitaker, D., Levi, D.M., Kennedy, G.J.: Integration across time determines path deviation discrimination for moving objects. PLoS One 3(4) (2008)

[136] Williams, L.R., Jacobs, D.W.: Stochastic completion fields. ICCV Proceedings (1995)

[137] Wilson, H.R., Cowan, J.D.: Excitatory and inhibitory interactions in localized populations of model neurons. Biophys. J. 12, 1-24 (1972)

[138] Wu, W., Tiesinga, P.H., Tucker, T.R., Mitroff, S.R., Fitzpatrick, D.: Dynamics of population response to changes of motion direction in primary visual cortex. J. Neurosci. 31(36), 12767-77 (2011)

[139] Xiao, D.K., Raiguel, S., Marcar, V., Orban, G.A.: The spatial distribution of the antagonistic surround of MT/V5 neurons. Cereb. Cortex 7(7), 662-677 (1997)

[140] Zelnik-manor, L., Perona, P.: Self-tuning spectral clustering. Advances in Neural Information Processing Systems 17, 1601-1608 (2004)

[141] Zucker, S.W.: Differential geometry from the Frenet point of view: boundary detection, stereo, texture and color. In: Paragios, N., Chen, Y., Faugeras, O. (eds.) Handbook of Mathematical Models in Computer Vision, pp. 357-373. Springer, US (2006)

[142] Zweck, J.W., Williams, L.R.: Euclidean group invariant computation of stochastic completion fields using shiftable-twistable functions. Journal of Mathematical Imaging and Vision 21 (2), 135-154 (2004)