



## OPEN ACCESS

## EDITED BY

James Elder,  
York University, Canada

## REVIEWED BY

Carlos Vazquez,  
École de technologie supérieure (ÉTS), Canada  
Johannes Burge,  
University of Pennsylvania, United States

## \*CORRESPONDENCE

Maria Virginia Bolelli  
✉ maria.bolelli2@unibo.it

RECEIVED 11 January 2023  
ACCEPTED 06 December 2023  
PUBLISHED 08 January 2024

## CITATION

Bolelli MV, Citti G, Sarti A and Zucker SW (2024)  
Good continuation in 3D: the neurogeometry  
of stereo vision. *Front. Comput. Sci.* 5:1142621.  
doi: 10.3389/fcomp.2023.1142621

## COPYRIGHT

© 2024 Bolelli, Citti, Sarti and Zucker. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](#). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Good continuation in 3D: the neurogeometry of stereo vision

Maria Virginia Bolelli<sup>1,2\*</sup>, Giovanna Citti<sup>1</sup>, Alessandro Sarti<sup>2</sup> and Steven W. Zucker<sup>3</sup>

<sup>1</sup>Department of Mathematics, University of Bologna, Bologna, Italy, <sup>2</sup>Centre d'analyse et de mathématique sociales (CAMS), CNRS, École des hautes études en sciences sociales (EHESS), Paris, France, <sup>3</sup>Departments of Computer Science and Biomedical Engineering, Yale University, New Haven, CT, United States

Classical good continuation for image curves is based on 2D position and orientation. It is supported by the columnar organization of cortex, by psychophysical experiments, and by rich models of (differential) geometry. Here, we extend good continuation to stereo by introducing a neurogeometric model to abstract cortical organization. Our model clarifies which aspects of the projected scene geometry are relevant to neural connections. The model utilizes parameterizations that integrate spatial and orientation disparities, and provides insight into the psychophysics of stereo by yielding a well-defined 3D association field. In sum, the model illustrates how good continuation in the (3D) world generalizes good continuation in the (2D) plane.

## KEYWORDS

stereo vision, sub-Riemannian geometry, 3D space of position-orientation, 3D association field, neurogeometry

## 1 Introduction

Binocular vision is the ability of the visual system to provide information about the three-dimensional environment starting from two-dimensional retinal images. Disparities are among the main cues for depth perception and stereo vision but, in order to extract them, the brain needs to determine which features coming from the right eye correspond to those from the left eye, and which do not. This generates a coupling problem, which is usually referred to as the *stereo correspondence problem*. Viewed in the large, stereo correspondence must be consistent with stereo perception more generally, and knowing the relevant features is key for both issues. In this paper we develop an approach to stereo based on the functional organization of the visual cortex, and we identify the geometric features extracted by the binocular cells. This model will be able to extend the notion of good continuation for planar curves to that for 3D spatial curves. A simple example demonstrates their application in computing stereo correspondence.

Good continuation in the plane (retinotopic coordinates) is one of the foundational principles of Gestalt perceptual organization. It enjoys an extensive history (Wagemans et al., 2012). It is supported by psychophysical investigations (e.g., Field et al., 1993; Geisler et al., 2001; Elder and Goldberg, 2002; Hess et al., 2003; Lawlor and Zucker, 2013), which reveal connections to contour statistics; it is supported by physiology (orientation selectivity), which reveals the role for long-range horizontal connections (Bosking et al., 1997); and it is supported by computational modeling (Ben-Shahar and Zucker, 2004; Sarti et al., 2007), which reveals a key role for geometry. The notion of orientation underlies all three of these aspects: neurons in visual cortex are selective for orientations, pairs of dots in grouping experiments indicate an orientation, and edge elements in natural images are oriented and related to image statistics. Orientation in space involves two angles, which we shall exploit. Nevertheless, good

continuation in depth is much less well-developed than good continuation in the plane, despite having comparable historical origins. (Koffka, 1963, p. 161-162):

...a perspective drawing, even when viewed monocularly, does not give the same vivid impression of depth as the same drawing if viewed through a stereoscope with binocular parallax... for in the stereoscope the tri-dimensional force of the parallax co-operates with the other tri-dimensional forces of organization; instead of conflict between forces, stereoscopic vision introduces mutual reinforcement.

Our specific goal in this paper is to develop a neurogeometrical model of stereo vision, based on the functionality of binocular cells. The main application will be a good continuation model in three dimensions that is analogous to the models of contour organization in two dimensions. We will develop *ad hoc* mathematical instruments, supported by a number of neural and psychophysical investigations (Malach et al., 1993; Uttal, 2013; Deas and Wilcox, 2014, 2015; Khuu et al., 2016; Scholl et al., 2022).

Although only one dimension higher than contours in the plane, contours extending in depth raise subtle new issues; this is why a geometric model can be instructive. First among the issues is the choice of coordinates which, of course, requires a mathematical framework for specifying them. In the plane, position and orientation are natural; smoothness is captured by curvature or the relationship between nearby orientations along a contour. For stereo, there is monocular structure in the left eye and in the right. Spatial disparity is a standard variable relating them, and it is well-known that primate visual systems represent this variable directly (Poggio, 1995). Spatial disparity is clearly a potential coordinate. However, other physiological aspects are less clear. The columnar architecture so powerful for contour organization in the plane is not only monocular: the presence of columns for spatial disparity of binocular cells has been experimentally described in V2 (Ts'o et al., 2009). However, orientation disparity does not seem to be coded in the cortex (see next section). Nevertheless, orientation-selective cells provide the input for stereo so, at a minimum, both position disparity and orientation – one orientation for the right eye and (possibly) another for the left – should be involved. While it is traditional to assume only “like” orientations are matched (Hubel and Wiesel, 1962; Nelson et al., 1977; Marr and Poggio, 1979; Bridge and Cumming, 2001; Chang et al., 2020), our sensitivity to orientation disparity questions this, making orientation disparity another putative variable. We shall show that orientations do play a deep role in stereo, but that it is not necessarily efficient to represent them as a disparity. Furthermore, there is a debate in stereo psychophysics about orientation: since its physiological realization could be confounded with disparity gradients (Mitchison and McKee, 1990; Cagenello and Rogers, 1993), orientation may be redundant. This is not the case, since it is the orientation of the “gradient” that matters. Thus we provide a representation of the geometry of spatial disparity and orientation in support of using good continuation in a manner that both incorporates the biological “givens” and provides a rigorous foundation for the correspondence problem. As has been the case with curve organization, we further believe that our modeling will illuminate the underlying functional architecture for stereo.

Hubel and Wiesel reported disparity-tuned neurons in early, classic work (Hubel and Wiesel, 1970). They observed that single units could be driven from both eyes and that it was possible to plot separate receptive fields (RF) for each eye. We emphasize that these monocular receptive fields are tuned to orientation (Cumming and DeAngelis, 2001; Parker et al., 2016), and a review of neural models can be found in Read (2015).

The classical model for expressing the left/right-eye receptive field combination is the *binocular energy model (BEM)*, first introduced in Anzai et al. (1999b). It encodes disparities through the receptive profiles of simple cells, raising the possibility of both position and phase disparities (Jaeger and Ranu, 2015). However, Read and Cumming (2007), building upon (Anzai et al., 1999a), showed that phase disparity neurons tend to be strongly activated by false correspondence pairs. Other approaches are based on the statistics of natural images (Burge and Geisler, 2014; Jains and Burge, 2017; Burge, 2020) utilized in an optimal fashion; these lead to more refined receptive field models. Nevertheless, the orientation differences between the two eyes (Nelson et al., 1977), or orientation disparity, should not be neglected. Although there were attempts to incorporate it (Bridge et al., 2001) in energy models, they are limited. The geometrical model we will present incorporates orientation differences directly.

Many other mathematical models for stereo vision based on neural models have been developed. Some claim (e.g., Marr and Poggio, 1979) that orientations should match between the two eyes, although small differences are allowed. This, of course, assumes the structure is frontal-parallel. Subsequently, Jones and Malik (1991) used a set of linear filters tuned to different orientations (and scales) but their algorithm was not built on a neurophysiological basis. Alibhai and Zucker (2000), Li and Zucker (2003), and Zucker (2014) built a more biologically-inspired model that addressed the connections between neurons. Their differential-geometry model employed position, orientations and curvatures in 2D retinal planes, modeling binocular neurons with orientations given by tangent vectors of Frenet geometry. Our results here are related, although the geometry is deeper (We develop this below.). A more recent work, based on differential and Riemannian geometry, is developed in Neilson et al. (2018). Before specifying these results, however, we introduce the specific type of geometry that we shall be using. It follows directly from the columnar organization often seen in predators and primates.

## 1.1 Columnar architectures and sub-Riemannian geometry

We propose a sub-Riemannian model for the cortical-inspired geometry underlying stereo vision based on the encoding of positional disparities and orientation differences in the information coming from the two eyes. We build on neuromathematical models, starting from the work of Koenderink and van Doorn (1987) and Hoffman (1989), with particular emphasis on the neurogeometry of monocular simple cells (Petitot and Tondut, 1999; Citti and Sarti, 2006; Sarti et al., 2007; Petitot, 2008; Sanguinetti et al., 2010; Sarti and Citti, 2015; Baspinar et al., 2020).

To motivate our mathematical approach, it is instructive to build on an abstraction of visual cortex. We start with monocular information, segregated into ocular dominance bands (LeVay et al., 1975) in layer 4; these neurons have processes that extend into the superficial layers. We cartoon this in Figure 1, which shows an array of orientation hypercolumns arranged over retinotopic position. It is colored by dominant eye inputs; the binocularly-driven cells tend to be closer to the ocular dominance boundaries, while the monocular cells are toward the centers. A zoom emphasizes the orientation distribution along a few of the columns near each position; horizontal connections (not shown) effect the interactions between these units. This raises the basic question in this paper: *what is the nature of the interaction among groups of cells representing different orientations at nearby positions and innervated by inputs from the left and right eyes?* The physiology suggests (Figure 1C) the answer lies in the interactions among both monocular and binocular cells; our model specifies this interaction, starting from the monocular ones and building analogously into a columnar organization.

## 1.2 Informal setup and overview

Since much of the paper is technical, we here specify, informally, the main ingredients of the model and the results. We first list several of the key points, then illustrate them directly.

- Stereo geometry enjoys a mathematical structure that is a formal extension of plane curve geometry. In the plane, points belonging to a curve are described by an orientation at a position, and these are naturally represented as elements (orientation, position) of columns. In our model, these become abstract fibers. The collection of fibers across position is a fiber bundle. Elements of the (monocular) fiber can be thought of as neurons.
- Our geometrical model is based on tangents and curvatures. Tangents naturally relate to orientation selectivity, and are commonly identified with “edge” elements in the world. We shall occasionally invoke this relationship, for intuition and convenience, but some caution is required. While edge elements comprising, e.g., a smooth bounding contour are tangents, the converse is not necessarily true (e.g., elongated attached highlights or hair textures). Instead, our model should be viewed as specifying the constraints relevant to understanding neural circuitry; see Section 1.3.
- To elaborate the previous point: the tangents in our model need not be edges in the world; they are neural responses. The constraints in our model can be used to determine whether these responses should be considered as “edges.” This is why the model is built from the geometry of idealized space curves: to support such inferences.
- For stereo, we shall need fibers that are a “product” of the left and right-eye monocular columns. This is the reason why we choose position, positional disparity and orientations from the left and right eyes respectively, as the natural variables that describe the stereo fiber over each position.

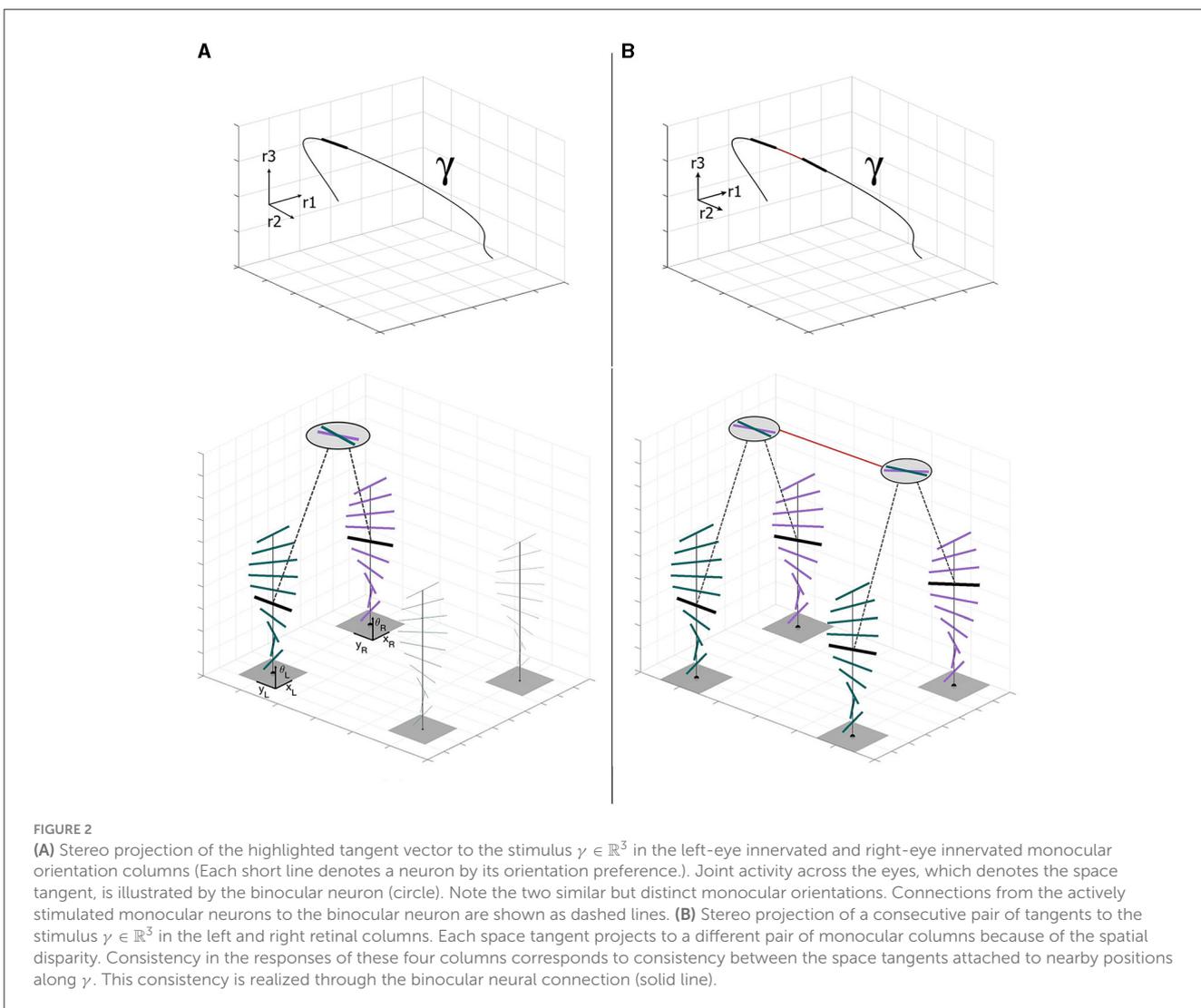
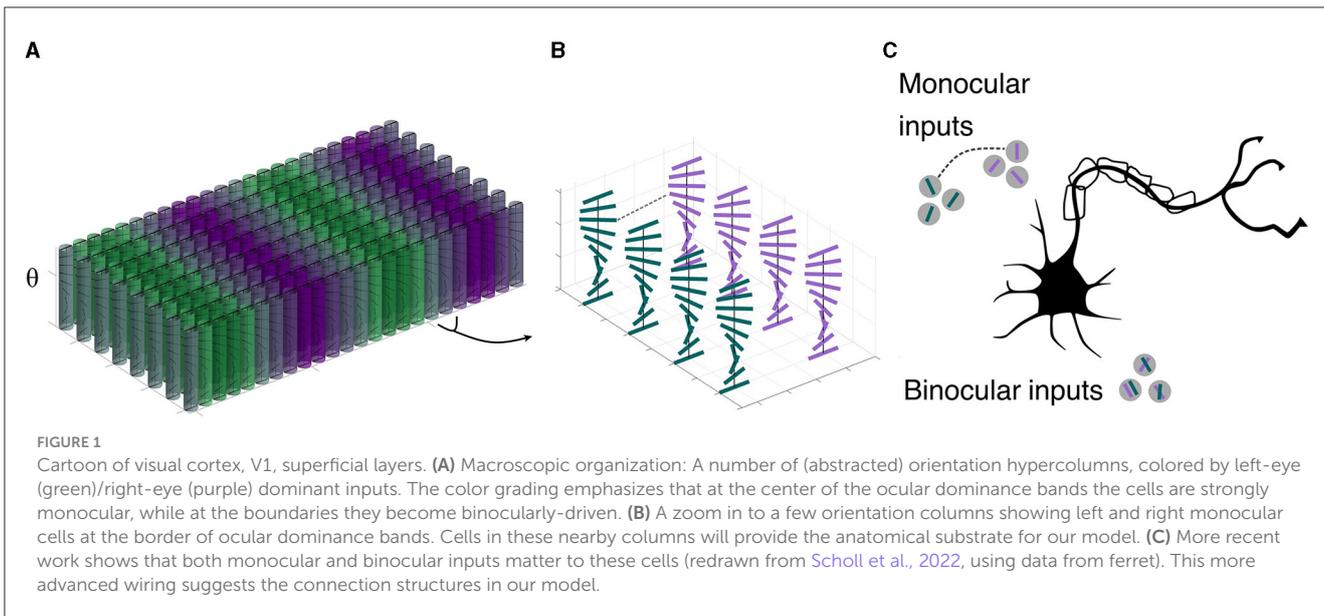
We stress that these fibers are not necessarily explicit in the cortical architecture.

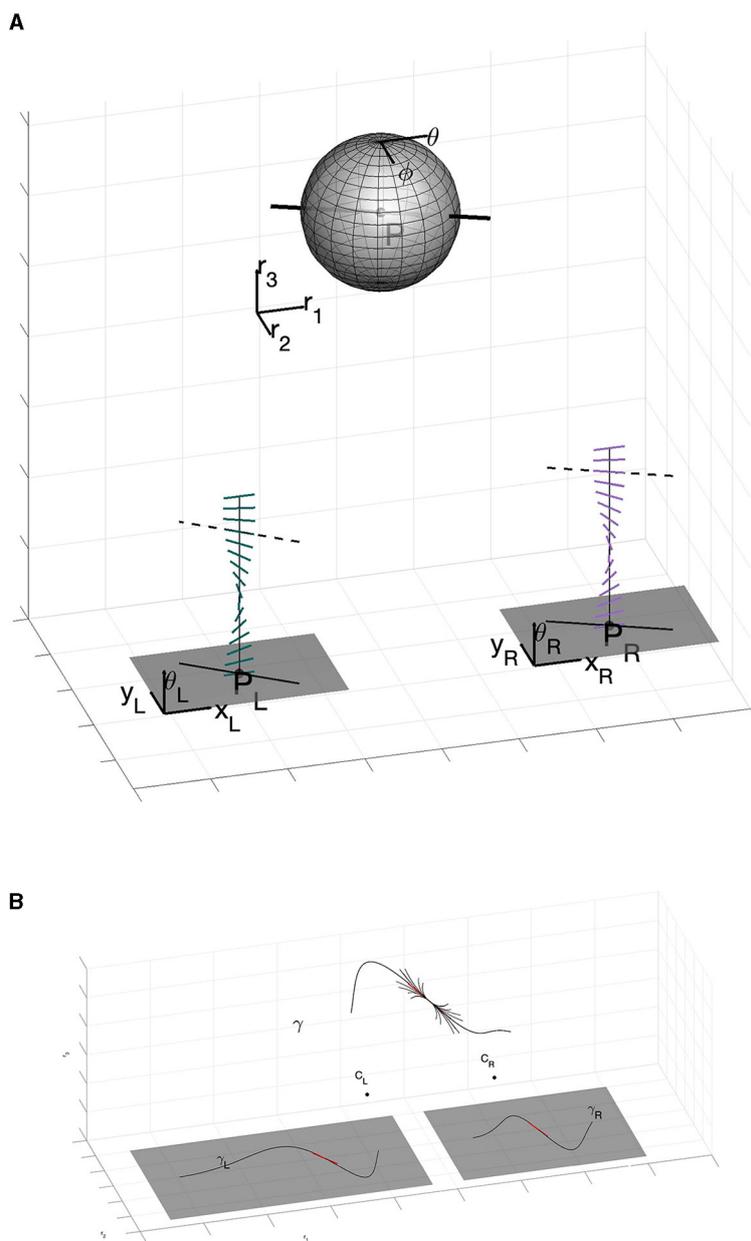
- Curvature provides a kind of “glue” to enable transitions from points on fibers to nearby points on nearby fibers. These transitions specify “integral curves” through the stereo fiber bundle.
- The integral curve viewpoint provides a direction of information flow (information diffuses through the bundle) thereby suggesting underlying circuits.
- The integral curves formalize *association field* models. Their parameters describe the spray of curves that is well in accordance with 3D curves as studied in psychophysical experiments in Hess and Field (1995), Hess et al. (1997), and Khuu et al. (2016).
- Our formal theory addresses several conjectures in the literature. The first is the *identity hypothesis* (Kellman et al., 2005a,b) and the organization of units for curve interpolation (cf. Anderson et al., 2002); we show how tangents are natural “units” and how they can be organized. The second concerns the nature of the organization (Li and Zucker, 2006), where we resolve a conjecture regarding the interpolating object (see Proposition 3.2 below).
- Our formal theory provides a new framework for specifying the correspondence problem, by illustrating how good continuation in the 3-D world generalizes good continuation in the 2-D plane. This is the point where consistent binocular-binocular interactions are most important.
- Our formal theory has direct implications for understanding torsional eye movements. It suggests, in particular, that the rotational component is not simply a consequence of development, but that it helps to undo inappropriate orientation disparity changes induced by eye movements. This role for Listing’s Law will be treated in a companion paper (in preparation); see also the excellent paper (Schreiber et al., 2008).

We now illustrate these ideas (Figure 2). Consider a three-dimensional stimulus as a space curve  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^3$ , with a unit-length tangent at the point of fixation. Since the tangent is the derivative of a curve, the binocular cells naturally encode the unitary tangent direction  $\dot{\gamma}$  to the spatial 3D stimulus  $\gamma$ . This space tangent projects to a tangent orientation in the left eye<sup>1</sup>, and perhaps the same or a different orientation in the right eye. A nearby space tangent projects to another pair of monocular tangents, illustrated as activity in neighboring columns. Note how connections between the binocular neurons support consistency along the space curve. It is this consistency relationship that we capture with our model of the stereo association field.

Since space curves live in 3D, two angles are required to specify its space tangent at a point. In other words, monocular tangent angles span a circle in the plane; space tangent angles span a 2-sphere in 3D. In terms of the projections into the left-eye and the right-eye, the space tangent can be described by the

<sup>1</sup> We are here being loose with language. By a tangent orientation in the left eye, we mean the orientation of a left-eye innervated column in V1.





**FIGURE 3**  
**(A)** The full geometry of stereo. Note how the stereo correspondence problem allows to establish the relationship between the 3D tangent point  $(P, \theta, \phi)$  and the projections  $p_L$  and  $p_R$ , the disparity and the orientations  $\theta_L$  and  $\theta_R$ . **(B)** Main result of the paper. The three-dimensional space curve  $\gamma$  is enveloped by the 3D association field centered at a point. Formally, this association field is a fan of integral curves in the sub-Riemannian geometry computed entirely within the columnar architecture (It is specifically described by Equation (36) with varying  $c_1$  and  $c_2$  in  $\mathbb{R}$ , but that will take some work to develop.).

parameters  $n = (\theta, \phi)$  of  $\mathbb{S}^2$  (Figure 3A). Thus, we can suitably describe the space of stereo cells – the full set of space tangents at any position in the 3D world – as the manifold of positions and orientations  $\mathbb{R}^3 \times \mathbb{S}^2$ . Moving from one position in space to another, and changing the tangent orientation to the one at the new position, amounts to what is called a *group action* on the appropriate manifold. We informally introduce these notions in the next subsection; a more extensive introduction to these ideas is in Appendix A (Supplementary material).

### 1.2.1 Sub-Riemannian geometry

We live in a 3D world in which distances are familiar; that is, a space of points with a Euclidean distance function defined between any pair of them. Apart from practical considerations we can move in any direction we would like. Cars, however, have much more restricted movement capabilities. They can move forward or backward, but not sideways. To move in a different direction, cars must turn their wheels. Here is the basic analogy: in cortical space information can move to a new retinotopic position in a tangent

direction, or it can move up or down a column (orientation fiber) to change direction. Moving in this fashion, from an orientation at a position to another orientation at a nearby position, is clearly more limited than arbitrary movements in Euclidean space. Euclidean geometry, as above, is an example of a Riemannian geometry; the limitations involved in moving through a cortical columnar space specify a sub-Riemannian geometry (Citti and Sarti, 2014; Citti et al., 2015; Sarti et al., 2019). Just as cars can move along roads that are mostly smooth, excitatory neurons mainly connect to similarly “like” (in orientation) excitatory neurons. This chain of neurons indicates a path through sub-Riemannian space (Agrachev et al., 2019); the fan of such paths is the cortical connectivity which can be considered the neural correlate of association fields. Again, for more information please consult Appendix A (Supplementary material).

Moving now out to the world, we must be able to move between all points. Repeating the above metaphor more technically, we equip  $\mathbb{R}^3 \times \mathbb{S}^2$  with a group action of the three-dimensional Euclidean group of rigid motions  $SE(3)$ . Notice, importantly, that this group is now acting on the product space of positions and orientations. A bit more is required, though, since the geometry of stereo vision is not solved only with these punctual and directional arguments. As we showed in Figure 2 there is the need to take into account the relationships between nearby tangents; in geometric language this involves a suitable type of connections. It is therefore natural to look at integral curves of the sub-Riemannian structure, which encode in their coefficients the fundamental concept of 3D curvature and torsion. An example of this is shown in Figure 3B. Notice how the 3D association field envelopes a space curve, in the same way that a 2D association field envelopes a planar curve. This figure illustrates, in a basic way, the fundamental result in this paper.

### 1.3 On the neuro-geometric approach

There are many different ways to approach mathematical modeling in vision. One could, for example, ask what is the best an ideal observer could do for the stereo problem working directly on image data (Burge and Geisler, 2014; Burge, 2020). This requires specifying the task, e.g., disparity at a point; a database of images on which the estimation is to be carried out; and a specification of the output. The approach is fundamentally statistical, and has been successful at predicting discrimination thresholds and optimal receptive field designs for patches of natural images. We seek to go the next step – to specify the relationship between receptive fields; i.e., between neurons. Note that the complexity multiplies enormously. At the behavioral level this raises the question of grouping, or determining the combinations of disparities, or Gabor patch samples, that belong together. The complexity arises because this must be evaluated over all possible arrangements of patches, be they along curves, or surfaces, or combinations thereof. In effect, the output specification is pushed toward co-occurrence phenomena, and these toward neural connections.

Our working hypothesis is that there is a deep functional relationship between structure in the brain and structure in the world, and that geometry is the right language with which to capture this relationship, especially as regards connectivity between

neurons and their functionality. The *neuro-geometric approach* is precisely this; an attempt to capture how the structure of cortical connectivity (and other functional properties) are reflected in the phenomena of visual perception.

At first blush this might seem completely unrelated to the statistics of natural images, and how these could be informative of neural connections, but we believe that there is a fundamental relationship. Consider, to start, the distribution of oriented edge elements in a small patch. Pairwise edge statistics are well-studied (August and Zucker, 2000; Geisler et al., 2001; Elder and Goldberg, 2002; Sanguinetti et al., 2010), and indicate how orientation changes are distributed over (spatially) nearby edge elements. Co-linear and co-circular patterns emerge from these studies, as well as in third-order statistics of edges (Lawlor and Zucker, 2013). Interestingly, in Singh and Fulvio (2007) and Geisler and Perry (2009) deviation from co-circular behavior emerges.<sup>2</sup> In particular, Geisler and Perry (2009) proposes a parabolic model to explain these statistical evidences, that is consistent with previous results if we consider a composition of the joint action of cocircularity and parallelism cues (as found to factor for example in Elder and Goldberg, 2002). To elaborate, it begins by following either a co-circular or linear term, followed by the composition with another circular or linear term. The outcome of this process is described as a spline-like behavior that can approximate a parabola. In Sanguinetti et al. (2010), it has been shown that the histogram of the co-occurrence of edges in a natural image provides the same probability kernel we could find with geometric analysis instruments. As a result, statistical measurements are integrated into the geometric approach.

The geometric analysis that we shall use is continuous mathematics, and is essentially differential (Tu, 2011). This has important implications. First, the relationships that matter are those over small neighborhoods, not over “long” distances. Thus at a point there is an orientation (tangent) and a curvature. These barely change as one moves a tiny distance from the point. Thus we are not considering (in this paper) what happens behind (relatively) large occluders, when longer distances devoid of intermediate structure separate structure (Singh and Fulvio, 2005; Fulvio et al., 2008). Such problems are important but are outside the scope of this paper. Second, because the mathematics is continuous, we shall not consider sampling issues (Warren et al., 2002). To the extent that it matters, we shall assume discrete entities are sufficiently densely distributed that they function as if they were continuous (Zucker and Davis, 1988). In this sense our analysis is restricted to early vision. It does not necessarily account for the full range of cognitive tasks, which may well invoke higher-order computations over longer distances and even richer abstractions.

It has been observed that edge statistics for curves in the world depart from co-circularity. To quote (Geisler and Perry, 2009): “Except for a direction of zero, where the orientation difference is consistent with a collinear relationship, the highest-likelihood orientation differences are less than those predicted by a co-circular

<sup>2</sup> Of course we need to take into account the difficulty of measurements of coupled position-orientation variables for small difference of angle and position. This is due to the well-known intrinsic uncertainty of measurement in the non-commutative group of position and orientation (Barbieri et al., 2012).

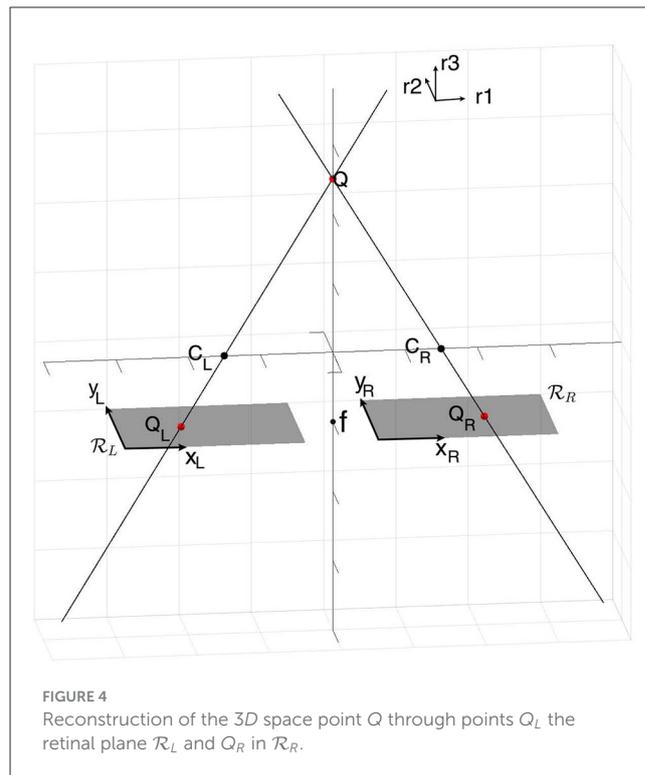
relationship.” We believe this has to do with the notion of curvature used: whether it is purely local, or an estimate over distance. In summary of the geometric approach, we explain this as follows. Our constraints can be used in two rather different ways. First, from a computational perspective, one can “integrate” the local constraints into more global objects. This is the approach used in the example section of this paper, and could give rise to an “average” curvature over some distance. The second approach is more distributed, and may be closer to a neurobiological implementation (Ben-Shahar and Zucker, 2004). In this second approach (not developed in this paper, but see Li and Zucker (2006) and Figure 6), the local computations overlap to enforce consistency (The scale of such computations would be a small factor larger than that indicated in Figure 2B). This scale corresponds to the extent of biological “long range horizontal connections” but is smaller than many of the occluders used in psychophysical experiments. In other words, to emphasize this distinction, in the former case the use of integral curves may be closer to the parabolic relations observed in scene statistics (Geisler and Perry, 2009).<sup>3</sup> Our use of the term “co-circularity” is in the latter sense.

## 1.4 Overview of paper

The paper is organized as follows: in Section 2, we describe the geometrical and neuro-mathematical background underlying the problem of stereo vision. In particular, we review the standard stereo triangulation technique to relate the coordinate system of one retina with the other, and put them together in order to reconstruct the three-dimensional space. Then, we briefly review the classical neurogeometry of monocular simple cells selective for orientation and the underlying connections. The generalization of approximate co-circularity for stereo is also introduced. In Section 3, starting from binocular receptive profiles, we introduce the neuro-mathematical model for binocular cells. First we present the cortical fiber bundle of binocular cells. It follows the differential interpretation of the binocular profiles in terms of the neurogeometry of the simple cells, and we show how this is well in accordance with the results of the stereo triangulation. Then, we give a mathematical definition of the manifold  $\mathbb{R}^3 \times \mathbb{S}^2$  with the sub-Riemannian structure. Finally, we study the integral curves and the suitable change of variables that allow us to switch our analysis from cortical to external space. In Section 4 we proceed to the validation of our geometry with respect to psychophysical experiments. We combine information about the psychophysics of 3D perception and formal conjectures; it is here that we formulate a 3D association field analogous to the 2D association field. At the end, we show an example of a representation of a stimulus (from image planes to the full 3D and orientation geometry) and how our integral curves properly connect corresponding points. This illustrates the use of our model as a basis for solving the correspondence problem.<sup>4</sup>

<sup>3</sup> The crucial point is that the curves demonstrate locally quadratic (not linear) behavior.

<sup>4</sup> Portions of this material were presented at Bolelli et al. (2023a).



## 2 Stereo vision and neuro-mathematical background

### 2.1 Stereo geometry

In this subsection, we briefly recall the geometrical configuration underlying 3D vision, to define the variables that we use in the rest of the paper, mainly referring to (Faugeras, 1993, Ch. 6). For a complete historical background see Howard (2012); Howard and Rogers (1995).

#### 2.1.1 Stereo variables

We consider the global reference system  $(O, i, j, k)$  in  $\mathbb{R}^3$ , with  $O = (0, 0, 0)$ , and coordinates  $(r_1, r_2, r_3)$ . We introduce the optical centers  $C_L = (-c, 0, 0)$  and  $C_R = (c, 0, 0)$ , with  $c$  real positive element, and we define two reference systems:  $(C_L, i_L, j_L)$ ,  $(C_R, i_R, j_R)$ , the reference systems of the retinal planes  $\mathcal{R}_L$  and  $\mathcal{R}_R$  with coordinates respectively  $(x_L, y)$ ,  $(x_R, y)$ . In the global system we suppose the retinal planes to be parallel and to have equation  $r_3 = f$ , with  $f$  denoting the focal length. This geometrical set-up is shown in Figure 4.

**Remark 2.1.** If we know the coordinate of a point  $Q = (r_1, r_2, r_3)^T$  in  $\mathbb{R}^3$ , then it is easy to project it in the two planes via perspective projection, having  $c$  the coordinate of the optical centers and  $f$  focal length. This computation defines two projective maps  $\Pi_L$  and  $\Pi_R$ , respectively, for the left and right retinal

planes:

$$\begin{aligned} \Pi_L : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 & \Pi_R : \mathbb{R}^3 &\longrightarrow \mathbb{R}^2 \\ \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} &\mapsto \begin{pmatrix} f(r_1+c) \\ \frac{fr_2}{r_3} \\ \frac{fr_3}{r_3} \end{pmatrix}, & \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix} &\mapsto \begin{pmatrix} f(r_1-c) \\ \frac{fr_2}{r_3} \\ \frac{fr_3}{r_3} \end{pmatrix}. \end{aligned} \tag{1}$$

*Proof.* A point on the left retinal plane of local coordinates  $(x_L, y)^T$  has global coordinates  $Q_L = (-c + x_L, y, f)^T$ , and it corresponds to a point  $Q = (r_1, r_2, r_3)^T$  in the Euclidean  $\mathbb{R}^3$  such that  $C_L$ ,  $Q_L$  and  $Q$  are aligned. This means that the vectors  $Q_L - C_L = (x_L, y, f)^T$  and  $Q - C_L = (r_1 + c, r_2, r_3)^T$  are parallel, obtaining the following relationships:

$$x_L = f \frac{r_1 + c}{r_3}, \quad y = f \frac{r_2}{r_3}. \tag{2}$$

Analogously, considering  $Q_R$  and  $C_R$ , we get:

$$x_R = f \frac{r_1 - c}{r_3}, \quad y = f \frac{r_2}{r_3}. \tag{3}$$

□

In a standard way, the *horizontal disparity* is defined as the differences between retinal coordinates

$$d := \frac{x_L - x_R}{2}, \tag{4}$$

up to a scalar factor. Moreover, it is also possible to define the coordinate  $x$  as the average of the two retinal coordinates  $x := \frac{x_L + x_R}{2}$ , leading to the following change of variables:

$$\begin{cases} x = \frac{fr_1}{r_3} \\ y = \frac{fr_2}{r_3} \\ d = \frac{fc}{r_3} \end{cases} \longleftrightarrow \begin{cases} r_1 = \frac{xc}{d} \\ r_2 = \frac{yc}{d} \\ r_3 = \frac{fc}{d} \end{cases}, \tag{5}$$

where the set of coordinates  $(x, y, d)$  is known as *cyclopean coordinates* (Julesz, 1971).

### 2.1.2 Tangent estimation

Corresponding points in the retinal planes allow to project back into  $\mathbb{R}^3$ . An analogous reasoning can be done for the tangent structure: if we have tangent vectors of corresponding curves in the retinal planes, it is possible to project back and recover an estimate of the 3D tangent vector. Let us recall here this result; a detailed explanation can be found in Faugeras (1993).

**Remark 2.2.** Let  $\gamma_L$  and  $\gamma_R$  be corresponding left and right retinal curves; i.e., perspective projections of a curve  $\gamma \in \mathbb{R}^3$  through optical centers  $C_L$  and  $C_R$  with focal length  $f$ . Knowing the left and right retinal tangent structures, it is possible to recover the direction of the tangent vector  $\dot{\gamma}$ .

*Proof.* Starting from a curve  $\gamma \in \mathbb{R}^3$ , we project it in the two retinal planes obtaining  $\gamma_L = \Pi_L(\gamma)$  and  $\gamma_R = \Pi_R(\gamma)$  from Equation (1). The retinal tangent vectors are obtained through the Jacobian matrix<sup>5</sup> of the left and right retinal projections  $\dot{\gamma}_{L,R}(t) = (J_{\Pi_{L,R}})_{\gamma(t)} \dot{\gamma}(t)$ :

$$\dot{\gamma}_R(t) = \begin{pmatrix} f(\gamma_3 \dot{\gamma}_1 + (c - \gamma_1) \dot{\gamma}_3) \\ \frac{\gamma_3(t)^2}{f(\gamma_3 \dot{\gamma}_2 - \gamma_2 \dot{\gamma}_3)} \\ \frac{\gamma_3(t)^2}{\gamma_3^2} \end{pmatrix}, \quad \dot{\gamma}_L(t) = \begin{pmatrix} f(\gamma_3 \dot{\gamma}_1 - (c + \gamma_1) \dot{\gamma}_3) \\ \frac{\gamma_3(t)^2}{f(\gamma_3 \dot{\gamma}_2 - \gamma_2 \dot{\gamma}_3)} \\ \frac{\gamma_3(t)^2}{\gamma_3^2} \end{pmatrix}. \tag{6}$$

Extending the tangent vectors and the points into  $\mathbb{R}^3$ , we get  $\tilde{t}_L = (\dot{\gamma}_{L1}, \dot{\gamma}_{L2}, 0)^T$ , and  $\tilde{m}_L = (\gamma_{L1} - c, \gamma_{L2}, f)^T$ , and  $U_{t_L} = (P_L)^{-1} \tilde{m}_L \times (P_L^{-1}) \tilde{t}_L$ , with the projection matrix  $P_L = \begin{pmatrix} 1 & 0 & -c/f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . The same

reasoning holds for the right structure, with projection matrix  $P_R = \begin{pmatrix} 1 & 0 & c/f \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ .

Then  $U_{t_R} \times U_{t_L}$  is a vector parallel to the tangent vector  $\dot{\gamma}$ :

$$\begin{aligned} U_{t_R} \times U_{t_L} &= \begin{pmatrix} \frac{f^4 2c(\dot{\gamma}_2 \gamma_3 - \dot{\gamma}_3 \gamma_2)}{\gamma_3^4} \dot{\gamma}_1, \\ \lambda(t) \frac{f^4 2c(\dot{\gamma}_2 \gamma_3 - \dot{\gamma}_3 \gamma_2)}{\gamma_3^4} \dot{\gamma}_2, \frac{f^4 2c(\dot{\gamma}_2 \gamma_3 - \dot{\gamma}_3 \gamma_2)}{\gamma_3^4} \dot{\gamma}_3 \end{pmatrix}^T \\ &= \lambda(t) (\dot{\gamma}_1(t), \dot{\gamma}_2(t), \dot{\gamma}_3(t))^T \\ &= \lambda(t) \dot{\gamma}(t). \end{aligned} \tag{7}$$

□

Although this section has been based on the geometry of space curves and their projections, we observe that related geometric approaches have been developed for planar patches and surfaces; see, e.g., Li and Zucker, 2008; Oluk et al., 2022 and references therein.

## 2.2 Elements of neurogeometry

We now provide background on the geometric modeling of the monocular system, and good continuation in the plane. Our goal is to illustrate the role of sub-Riemannian geometry in the monocular system, which will serve as the basis for generalization to the stereo system.

### 2.2.1 Classical neurogeometry of simple cells

We model the activation map of a cortical neuron's receptive field (RF) by its receptive profile (RP)  $\varphi$ . A classical example is the receptive profiles of simple cells in V1, centered at position

<sup>5</sup> The Jacobian matrix  $(J_\Pi)_p$  evaluated at point  $p$  represents how to project displacement vectors (in the sense of derivatives or velocities or directions). In details, if  $\dot{\gamma}(t)$  is the displacement vector in  $\mathbb{R}^3$ , then the matrix product  $(J_\Pi)_{\gamma(t)} \dot{\gamma}(t)$  is another displacement vector, but in  $\mathbb{R}^2$ . In other words, the Jacobian matrix is the differential of  $\Pi$  at every point where  $\Pi$  is differentiable; common notation includes  $J_\Pi$  or  $D\Pi$ .

$(x, y)$  and orientation  $\theta$ , modeled (e.g., in [Daugman, 1985](#); [Jones and Palmer, 1987](#); [Barbieri et al., 2014b](#)) as a bank of Gabor filters  $\varphi_{\{x,y,\theta\}}$ . RPs are mathematical models of receptive fields; they are operators which act on a visual stimulus.

Formally, it is possible to abstract the primary visual cortex as  $\mathbb{R}^2 \times \mathbb{S}^1$ , or position-orientation space, thereby naturally encoding the Hubel/Wiesel hypercolumnar structure ([Hubel and Wiesel, 1962](#)). An example of this structure is displayed in [Figure 5D](#) from [Ben-Shahar and Zucker \(2004\)](#).

Following the model of [Citti and Sarti \(2006\)](#), the set of simple cells' RPs can be obtained via translations along a vector  $(x, y)^T$  and rotation around angle  $\theta$  from a unique "mother" profile  $\varphi_0(\xi, \eta)$ :

$$\varphi_0(\xi, \eta) = \exp\left(\frac{2\pi i\xi}{\lambda}\right) \exp\left(-\frac{\xi^2 + \eta^2}{2\sigma^2}\right), \quad (8)$$

This RP is a Gabor function with even real part and odd imaginary part ([Figure 7A](#)). Translations and rotations can be expressed as:

$$T_{(x,y,\theta)}(\xi, \eta) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \begin{pmatrix} x \\ y \end{pmatrix}, \quad (9)$$

where  $T_{(x,y,\theta)}$  denotes the action of the group of rotations and translations  $SE(2)$  on  $\mathbb{R}^2$ . This group operation associates to every point  $(\xi, \eta)$  a new point  $(\tilde{x}, \tilde{y})$ , according to the law  $(\tilde{x}, \tilde{y}) = T_{(x,y,\theta)}(\xi, \eta)$ . Hence, a general RP can be expressed as

$$\varphi_{(x,y,\theta)}(\xi, \eta) = \varphi_0(T_{(x,y,\theta)}^{-1}(\xi, \eta)), \quad (10)$$

and this represents the action of the group  $SE(2)$  on the set of receptive profiles.

The retinal plane  $\mathcal{R}$  is identified with the  $\mathbb{R}^2$  plane, whose coordinates are  $(x, y)$ . When a visual stimulus  $I: \mathcal{R} \rightarrow \mathbb{R}^+$  of intensity  $I(x, y)$  activates the retinal layer, the neurons centered at every point  $(x, y)$  produce an output  $O(x, y, \theta)$ , modeled as the integral of the signal  $I$  with the set of Gabor filters:

$$O(x, y, \theta) = \int_{\mathcal{R}} \varphi_{\{x,y,\theta\}}(\xi, \eta) I(\xi, \eta) d\xi d\eta, \quad (11)$$

where the function  $I$  represents the retinal image.

For  $(x, y)$  fixed, we will denote  $\bar{\theta}$  the point of maximal response:

$$\max_{\theta} |O(x, y, \theta)| = |O(x, y, \bar{\theta})|. \quad (12)$$

We will then say that the point  $(x, y)$  is lifted to the point  $(x, y, \bar{\theta})$ . This is extremely important conceptually to understand our geometry: it illustrates how an image point, evaluated against a simple cell RP, is lifted to a "cortical" point by introducing the orientation explicitly. If all the points of the image are lifted in the same way, the level lines of the 2D image  $I$  are lifted to new curves in the 3D cortical space  $(x, y, \bar{\theta})$ .

We shall now recall a model of the long range connectivity which allows propagation of the visual signal from one cell in a column to another cell in a nearby column. This is formalized as a set of directions for moving in the cortical space  $(x, y, \bar{\theta})$ , in the sense of vector fields. This is important because it will be necessary to move within this space, across both positions and orientations.

To begin, in the right hand side of the Equation (11) the integral of the signal with the real and imaginary part of the Gabor filter

is expressed. The two families of cells have different shapes, hence they detect (or play a role in detecting) different features. Since the odd-symmetry cells suggest boundary detection, we concentrate on them, but this is a mathematical simplification. The output of a simple cell can then be locally approximated as  $O(x, y, \theta) = -X_{3,p}(I_{\sigma})(x, y)$ , where  $p = (x, y, \theta) \in SE(2)$ ,  $I_{\sigma}$  is a smoothed version of  $I$ , obtained by convolving it with a Gaussian kernel, and

$$X_{3,p} = -\sin\theta\partial_x + \cos\theta\partial_y, \quad (13)$$

is the directional derivative in the direction  $\vec{X}_{3,p} = (-\sin\theta, \cos\theta, 0)^T$ . From now on, we will denote (by a slight abuse of notation)  $\omega^* := \vec{X}_{3,p}$  to remind the reader familiar with the language of 1-forms the correspondence of these quantities, and the relation with the Hodge star operator.<sup>6</sup>

Now, think of vector fields as defining a coordinate system at each point in cortical space. Then, in addition to above, the vector fields orthogonal to  $X_{3,p}$  are:

$$X_{1,p} = \cos\theta\partial_x + \sin\theta\partial_y, \quad X_{2,p} = \partial_{\theta} \quad (14)$$

and they define a 2-dimensional admissible tangent bundle<sup>7</sup> to  $\mathbb{R}^2 \times \mathbb{S}^1$ . One can define a scalar product on this space by imposing the orthonormality of  $X_{1,p}$  and  $X_{2,p}$ : this determines a sub-Riemannian structure on  $\mathbb{R}^2 \times \mathbb{S}^1$ .

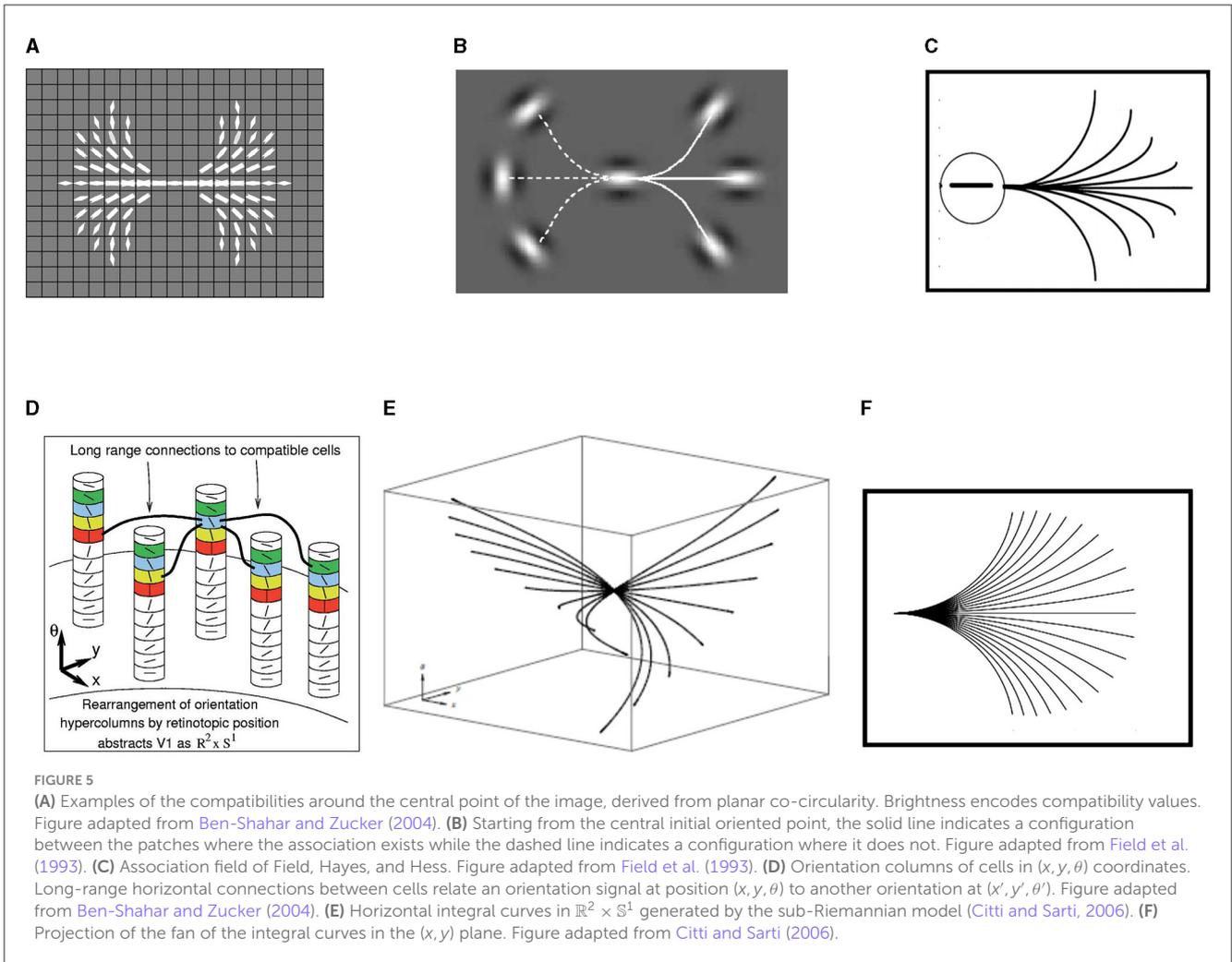
The visual signal propagates, in an anisotropic way, along cortical connectivity and connects more strongly cells with comparable orientations. This propagation has been expressed by the geometry just developed and 2-dimensional contour integration. This is the neural explanation of the Gestalt law of good continuation ([Koffka, 1963](#); [Kohler, 1967](#)). It can be directly expressed as co-circularity in the plane ([Parent and Zucker, 1989](#)), to describe the consistency and the compatibility of neighboring oriented points, in accordance with specific values of curvature. An example of these compatibilities can be found in [Figure 5A](#). It is complemented by psychophysical experiments, e.g., [Uttal, 1983](#); [Smits and Vos, 1987](#); [Ivry et al., 1989](#). In particular, [Field et al. \(1993\)](#) describe the association rules for 2-dimensional contour integration, introducing the concept of *association fields*. A representation of these connections can be found in [Figures 5B, C](#). Note that this is equivalent to the union (over curvature) in [Parent and Zucker \(1989\)](#). Neurophysiological studies ([Blasdel, 1992](#); [Malach et al., 1993](#); [Bosking et al., 1997](#); [Schmidt et al., 1997](#); [Hess et al., 2014](#)) suggest that the cortical correlate of the association field is the long-range horizontal connectivity among cells of similar (but not necessarily identical) orientation preference.

Based on these findings, [Citti and Sarti \(2006\)](#) modeled cortical propagation as propagation along integral curves of the vector fields  $X_1$  and  $X_2$ , namely curves  $\gamma: [0, T] \subset \mathbb{R} \rightarrow \mathbb{R}^2 \times \mathbb{S}^1$  described by the following differential equation:

$$\dot{\gamma}(t) = \vec{X}_{1,\gamma(t)} + k\vec{X}_{2,\gamma(t)}, \quad t \in [0, T], \quad (15)$$

<sup>6</sup> The purpose of introducing this notation is also to motivate an implication of the mathematical model in [Citti and Sarti \(2006\)](#); see Appendix B.2.1 ([Supplementary material](#)) for explanation.

<sup>7</sup> as defined in Appendix A3 ([Supplementary material](#)).



obtained by varying the parameter  $k \in \mathbb{R}$ . ( $k$  acts analogously as curvature.) An example of these curves is in Figure 5E. Their 2D projection is a close approximation of the association fields (Figure 5F).

A related model has been proposed by Duits et al. (2013). They study the geodesics of the sub-Riemannian structure to take into account all appropriate end-conditions of association fields.

### 2.2.2 Generalizing co-circularity for stereo

The concept of co-circularity in  $\mathbb{R}^2$  has been developed by observing that a bidimensional curve  $\gamma$  can be locally approximated at 0 via the osculating circle.<sup>8</sup> Alibhai and Zucker (2000), Li and

Zucker (2003), and Li and Zucker (2006) generalize this concept with the Frenet differential geometry of a three dimensional curve.

While in the two-dimensional case the approximation of the curve using the Frenet 2D basis causes the curvature to appear in the coefficient of the Taylor series development (1st order), in the three-dimensional case the coefficients involve both the curvature and torsion. So, in Alibhai and Zucker (2000) the authors propose heuristically to generalize the osculating circle for space curves with an osculating helix, with a preference for  $r_3$ -helices to improve stability in terms of camera calibration. In this way the orientation disparity is encoded in the behavior of the helix in the 3D space: there is no difference in orientation in the retinal planes if the helix is confined to be in the fronto-parallel plane (the helix becomes a circle); otherwise moving along the 3D curves the retinal projections have different orientations.

In Li and Zucker (2003, 2006) they observe that, by introducing the curvature variable as a feature in the two monocular structures, and assuming correspondence, it is possible to reconstruct the 3D Frenet geometry of the curve, starting from the two-dimensional Frenet geometry, up to the torsion parameter. In particular, they prove:

**Proposition 2.1.** Given two perspective views of a 3D space curve with full calibration, the normal  $N$  and curvature  $k$  at a curve

<sup>8</sup> Locally, a curve can be approximated by its osculating circle and, at a slightly larger scale, by the integral (parabolic) curve through the first two Taylor terms. The first approximation is co-circularity; the second is a parabolic curve. The second is an accurate model over large distances; see discussion in Section 1.3. However, since in this paper we are working over small distances and with cortical sampling (Figure 2), there is essentially no difference between them; see Figure 22.4 in Zucker (2006) and Sanguinetti et al. (2010) for a direct comparison.

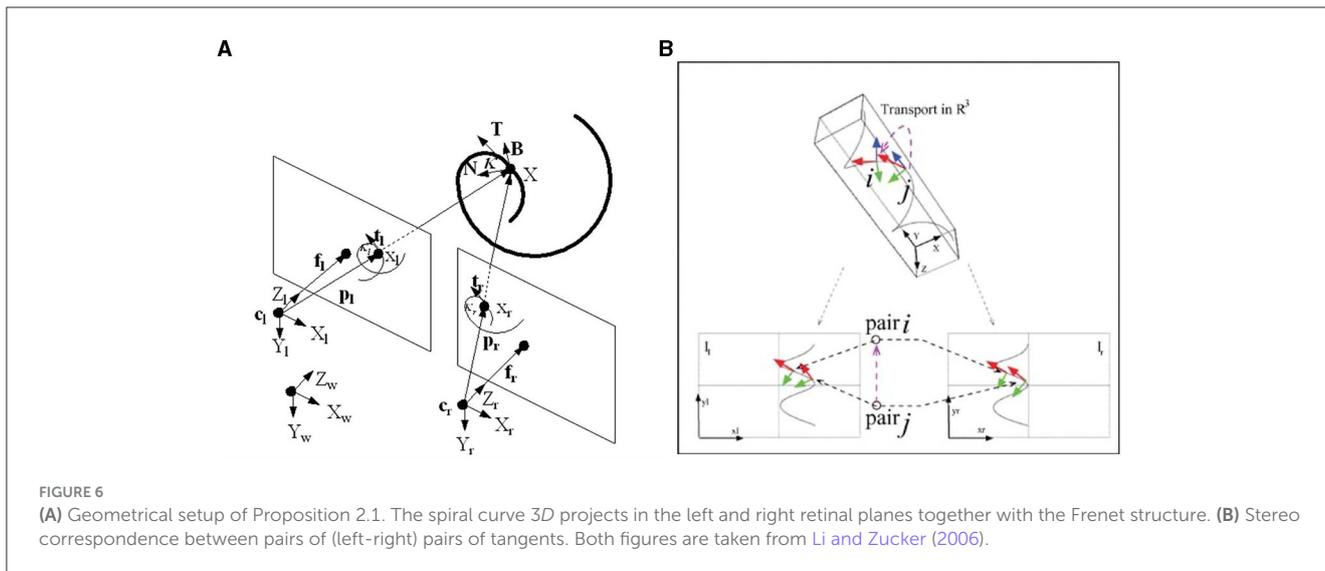


FIGURE 6 (A) Geometrical setup of Proposition 2.1. The spiral curve 3D projects in the left and right retinal planes together with the Frenet structure. (B) Stereo correspondence between pairs of (left-right) pairs of tangents. Both figures are taken from Li and Zucker (2006).

space point are uniquely determined from the positions, tangents, and curvatures of its projections in two images. Thus the Frenet frame  $\{T, N, B\}$  and curvature  $k$  at the space point can be uniquely determined.

Hence, using the knowledge of the Frenet basis together with the fundamental addition of the curvature variable, Zucker et al. applied the concept of *transport*. This allowed moving the 3D Frenet frame in a consistent way with the corresponding 2D Frenet structures of the left and right retinal planes, to establish stereo correspondence between pairs of (left and right) pairs of tangents (see Figure 6B).

Remark 2.3. The model that we propose in this paper is related to, but differs from, what has just been stated. In particular, to remain directly compatible with the previous neuro-geometric model, we will work only with the monocular variables of position and orientation. Rather than using curvature directly, we shall assume that these variables are encoded within the connections; mathematically they appear as parameters. A theoretical result of our model is that the heuristic assumption regarding the  $r_3$ -helix can now be established rigorously.

Let us also mention the paper (Abbasi-Sureshjani et al., 2017), where the curvature was considered as independent variable and helices have been obtained in the 2D space.

### 3 The neuromathematical model for stereo vision

Here, we do not want to directly impose a co-circularity property: our scope is to model the behavior of binocular cells, and deduce properties of propagation, which will ultimately induce a geometry of 3D good continuation laws.

### 3.1 Binocular profiles

Binocular neurons receive inputs from both the left and right eyes. To facilitate calculations, we assume these inputs are first combined in simple cells in the primary visual cortex, a widely studied approach (Anzai et al., 1999b; Cumming and DeAngelis, 2001; Menz and Freeman, 2004; Kato et al., 2016). It provides a first approximation in which binocular RPs are described as the product of monocular RPs; see Figure 7. This model is clearly an oversimplification, in several senses. First, it leaves out the more refined receptive fields discussed in Section 1.3. Second, it leaves out the role for complex cells (Sasaki et al., 2010). Third, it leaves out different ways to get the position and orientation information, such as eye fixations (Intoy et al., 2021). And fourth, it avoids the delicate question of whether the max operation over a column (Equation 12) truly captures a tangent element. Nevertheless, since our focus is geometric, it does capture all of the necessary ingredients and simplifies computations.

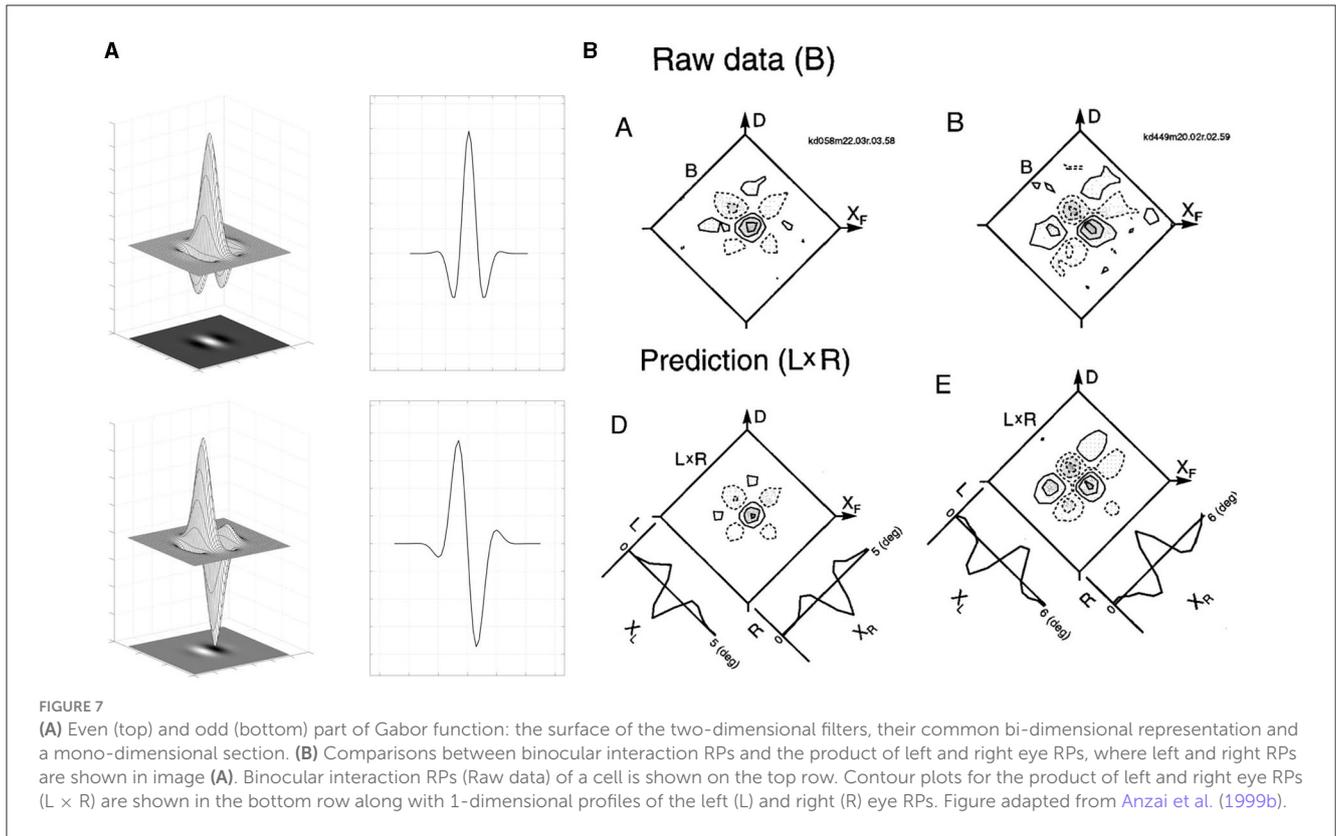
This binocular model allows us to define disparity and frontoparallel coordinates as

$$\begin{cases} d = \frac{x_L - x_R}{2} \\ x = \frac{x_R + x_L}{2}, \end{cases} \tag{16}$$

perfectly in accordance with the introduction of cyclopean coordinates in (4). In this way  $(x, y, d)$  correspond to the neural correlate of  $(r_1, r_2, r_3)$ , via the change of variables (5).

### 3.2 The cortical fiber bundle of binocular cells

The hypercolumnar structure of monocular simple cells (orientation selective) has been described as a jet fiber bundle in the works of Petitot and Tondut (1999), among many others. We concentrate on the fiber bundle  $\mathbb{R}^2 \times \mathbb{S}^1$ , with fiber  $\mathbb{S}^1$ ; see, e.g., Ben-Shahar and Zucker, 2004 among many others.



In our setting, the binocular structure is based on monocular ones; recall the example illustrations from the Introduction. In particular, for each cell on the left eye there is an entire fiber of cells on the right, and vice versa, for each cell on the right there is an entire fiber of cells on the left. This implies that the binocular space is equipped with a symmetry that involves the left and right structures, allowing us to use the cyclopean coordinates  $(x, y, d)$  defined in (16).

Hence, we define the cyclopean retina  $\mathcal{R}$ , identified with  $\mathbb{R}^2$ , endowed with coordinates  $(x, y)$ . The structure of the fiber is  $\mathcal{F} = \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$ , with coordinates  $(d, \theta_L, \theta_R) \in \mathcal{F}$ . The total space is defined in a trivial way,  $\mathcal{E} = \mathcal{R} \times \mathcal{F} = \mathbb{R}^2 \times \mathbb{R} \times \mathbb{S}^1 \times \mathbb{S}^1$ , and the projection  $\pi : \mathcal{E} \rightarrow \mathcal{R}$  is the trivial projection  $\pi(x, y, d, \theta_L, \theta_R) = (x, y)$ . The preimage of the projection  $\mathcal{E}_{(x,y)} := \pi^{-1}(\{(x, y)\})$ , for every  $(x, y) \in \mathcal{R}$ , is isomorphic to the fiber  $\mathcal{F}$ , and the local trivialization property is naturally satisfied.

A schematic representation can be found in Figure 8. The base has been depicted as 1-dimensional, considering the restriction  $\mathcal{R}|_x$  of the cyclopean retina  $\mathcal{R}$  on the coordinate  $x$ . The left image displays only the disparity component of the fiber  $\mathcal{F}$ , encoding the relationships between left and right retinal coordinates. The right image shows the presence of the left and right monodimensional orientational fibers.

### 3.3 Binocular energy model

To simplify calculations, as stated in the Introduction, we follow the classical binocular energy model (Anzai et al., 1999b)

for binocular RPs. The basic idea is a binocular neuron receives input from each eye; if the sum  $O_L + O_R$  of the inputs from the left and right eye is positive, the firing rate of the binocular neuron is proportional to the square of the sum, and it vanishes, if the sum of the inputs is negative:

$$O_B = (Pos(O_L + O_R))^2, \tag{17}$$

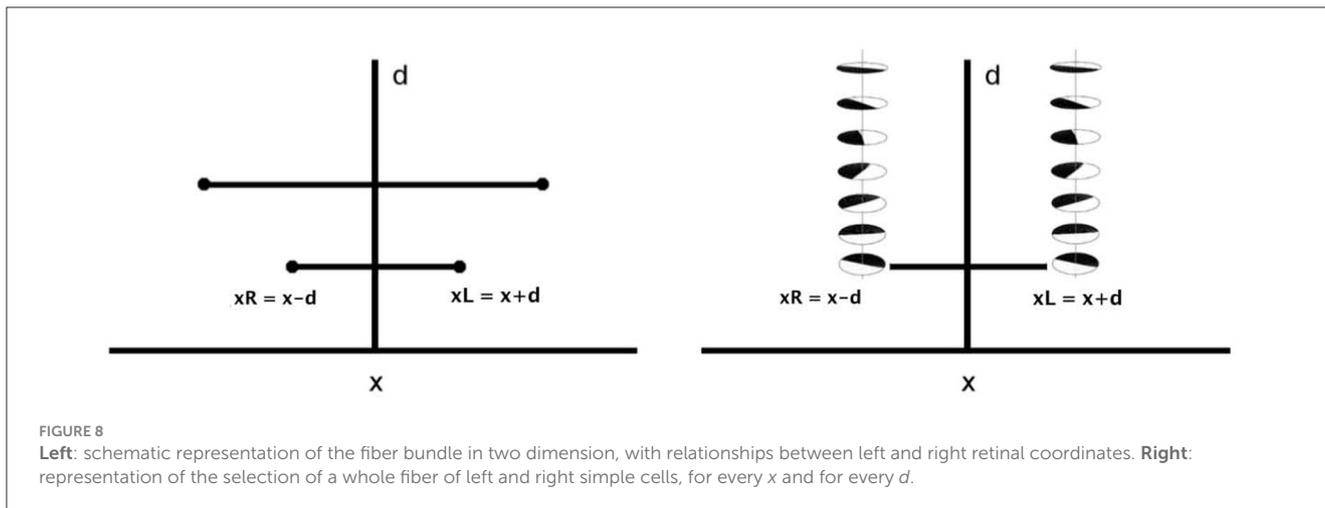
with  $Pos(x) = \max\{x, 0\}$ ,  $O_B$  the binocular output.

If  $O_L + O_R > 0$ , then the output of the binocular simple cell can be explicitly written as  $O_B = O_L^2 + O_R^2 + 2O_L O_R$ . The first two terms represent responses due to monocular stimulation while the third term  $2O_L O_R$  can be interpreted as the binocular interaction term.

The activity of a cell is then measured from the output and will be strongest at points that have a higher probability of matching each other. The maximum value over  $d$  of this quantity is the extracted disparity.

It is worth noting that neurophysiological computations of binocular profiles displayed in Figure 7B assume the monodimensionality of the monocular receptive profile, ignoring information about orientation of monocular simple cells. However, this information will be needed to encode different types of orientation disparity.

**Remark 3.1 (Orientation matters).** In 2001, the authors of Bridge et al. (2001) conducted investigations on the response of binocular neurons to orientation disparity, by extending the energy model of Anzai, Ohzawa, and Freeman to incorporate binocular differences in receptive-field orientation. More recently, the difference between



orientations in the receptive fields of the eyes has been confirmed (Sasaki et al., 2010).

The binocular energy model is a type of minimal model. It serves as a starting point, allowing the combination of monocular inputs. But is not sufficient to solve the stereo-matching problem.

**Remark 3.2 (Connections).** It is argued in Samonds et al. (2013) and Parker et al. (2016) that, in addition to the neural mechanisms that couple characteristics (such as signals, stimuli, or particular features) relating the left and right monocular structures, there must be a system of connections between binocular cells, which characterizes the processing mechanism of stereo vision; see also Samonds et al. (2013) in particular.

### 3.4 Differential interpretation of binocular RPs

It is possible to write the interaction term  $O_L O_R$  coming from (17), in terms of the left and right receptive profiles:

$$\begin{aligned}
 O_L O_R &= \int \varphi_{\theta_L, x_L, y}(\tilde{x}_L, \tilde{y}_L) I_L(\tilde{x}_L, \tilde{y}_L) d\tilde{x}_L d\tilde{y}_L \\
 &\quad \int \varphi_{\theta_R, x_R, y}(\tilde{x}_R, \tilde{y}_R) I_R(\tilde{x}_R, \tilde{y}_R) d\tilde{x}_R d\tilde{y}_R \\
 &= \int \int \varphi_{\theta_L, x_L, y}(\tilde{x}_L, \tilde{y}_L) \varphi_{\theta_R, x_R, y}(\tilde{x}_R, \tilde{y}_R) I_L(\tilde{x}_L, \tilde{y}_L) \\
 &\quad I_R(\tilde{x}_R, \tilde{y}_R) d\tilde{x}_R d\tilde{y}_R d\tilde{x}_L d\tilde{y}_L.
 \end{aligned}
 \tag{18}$$

If we fix  $(\tilde{x}_R, \tilde{y}_R, \tilde{x}_L, \tilde{y}_L)$ , we derive the expression of the binocular profiles  $\varphi_{L,R} = \varphi_{\theta_R, x_R, y} \varphi_{\theta_L, x_L, y}$  as the product of monocular left and right profiles. This is in accordance with the measured profiles of Figure 7B).

**Proposition 3.1.** The binocular interaction term can be associated with the cross product of the left and right directions defined through (13), namely  $\omega_L^*$  and  $\omega_R^*$  of monocular simple cells:

$$O_L O_R = \omega_L^* \times \omega_R^*. \tag{19}$$

*Proof.* The idea is that the binocular output is the combined result of the left and right actions of monocular cells, thus identifying a direction in the space of cyclopean coordinates. The detailed proof of this proposition can be found in Appendix B (Supplementary material).  $\square$

To better understand the geometrical idea behind Proposition 3.1, we recall that the retinal coordinates can be expressed in terms of cyclopean coordinates (4) as  $x_R = x - d$  and  $x_L = x + d$ , and so we can write  $\omega_L^*$  and  $\omega_R^*$  in the 3D space of coordinates  $(x, y, d)$  as:

$$\begin{aligned}
 \omega_R^* &= (-\sin \theta_R, \cos \theta_R, \sin \theta_R)^T \\
 \omega_L^* &= (-\sin \theta_L, \cos \theta_L, -\sin \theta_L)^T.
 \end{aligned}
 \tag{20}$$

We define  $\omega_{bin} := \omega_L^* \times \omega_R^*$  as the natural direction characterizing the binocular structure:

$$\omega_{bin} = \begin{pmatrix} \sin(\theta_R + \theta_L) \\ 2 \sin \theta_R \sin \theta_L \\ \sin(\theta_R - \theta_L) \end{pmatrix}. \tag{21}$$

**Remark 3.3.** The vector  $\omega_{bin}$  of Equation (21) can be interpreted as the intersection of the orthogonal spaces defined with respect to  $\omega_R^*$  and  $\omega_L^*$  when expressed in cyclopean coordinates  $(x, y, d)$ . More precisely, if

$$\begin{aligned}
 (\omega_L^*)^\perp &= \text{span} \left\{ \begin{pmatrix} \cos \theta_L \\ \sin \theta_L \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}, \\
 (\omega_R^*)^\perp &= \text{span} \left\{ \begin{pmatrix} \cos \theta_R \\ \sin \theta_R \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}
 \end{aligned}
 \tag{22}$$

then

$$\omega_{bin} = (\omega_L^*)^\perp \cap (\omega_R^*)^\perp. \tag{23}$$

The result of the intersection of these monocular structures identifies a direction, as shown in Figure 9A.

We earlier showed that the result of the action of a monocular odd simple cell is to select directions for the propagation of information. We now combine these, for the two eyes, to show that in the three-dimensional case the binocular neural mechanisms also lead to a direction. We will see in the next sections that this direction is the direction of the tangent vector to the 3D stimulus, provided points are corresponding.

### 3.5 Compatibility with stereo geometry

We consider the direction characterizing the binocular structure  $\omega_{bin}$  defined in (21) and we show that it can be associated with the 3D tangent vector to the 3D curve. The idea is that this tangent vector is orthogonal both to  $\omega_R^*$  and to  $\omega_L^*$ , and therefore it has the direction of the vector product  $\omega_L^* \times \omega_R^*$ .

Precisely, we consider the normalized tangent vector  $t_L$  and  $t_R$  on retinal planes

$$t_R = (\cos \theta_R, \sin \theta_R)^T \quad t_L = (\cos \theta_L, \sin \theta_L)^T, \quad (24)$$

to the points  $(x_R, y)$  and  $(x_L, y)$  respectively. Taking into account that  $f$  is the focal coordinate of the retinal planes in  $\mathbb{R}^3$ , then we associate to these points the correspondents in  $\mathbb{R}^3$ , namely  $\tilde{m}_L = (x_L - c, y, f)^T$ ,  $\tilde{m}_R = (x_R + c, y, f)^T$ . Applying Equation (7), it is possible to derive the tangent vector of the three dimensional contour:

$$\begin{aligned} U_{t_L} &= P_L^{-1} \tilde{m}_L \times P_L^{-1} \tilde{t}_L &&= \begin{pmatrix} x_L \\ y_L \\ f \end{pmatrix} \times \begin{pmatrix} \cos \theta_L \\ \sin \theta_L \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -f \sin \theta_L \\ f \cos \theta_L \\ x_L \sin \theta_L - y_L \cos \theta_L \end{pmatrix}, \\ U_{t_R} &= P_R^{-1} \tilde{m}_R \times P_R^{-1} \tilde{t}_R &&= \begin{pmatrix} x_R \\ y_R \\ f \end{pmatrix} \times \begin{pmatrix} \cos \theta_R \\ \sin \theta_R \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} -f \sin \theta_R \\ f \cos \theta_R \\ x_R \sin \theta_R - y_R \cos \theta_R \end{pmatrix}, \end{aligned} \quad (25)$$

and the tangent direction is recovered by

$$\begin{aligned} U_{t_L} \times U_{t_R} &= \\ f \begin{pmatrix} \frac{x_L + x_R}{2} \sin(\theta_R - \theta_L) - \frac{x_R - x_L}{2} \sin(\theta_L + \theta_R) \\ y \sin(\theta_R - \theta_L) - (x_R - x_L)(\cos(\theta_R - \theta_L) - \cos(\theta_L + \theta_R)) \\ f \sin(\theta_R - \theta_L) \end{pmatrix} \end{aligned} \quad (26)$$

If we define

$$\tilde{\omega}_L^* := \frac{d}{fc} U_{t_L}, \quad \tilde{\omega}_R^* := \frac{d}{fc} U_{t_R} \quad (27)$$

and the corresponding 2 form  $\omega_{\mathbb{R}^3} := \tilde{\omega}_L^* \times \tilde{\omega}_R^*$ , using the change of variables (16) we observe that:

$$\tilde{\omega}_L^* = \omega_L^*, \quad \tilde{\omega}_R^* = \omega_R^*, \quad \omega_{\mathbb{R}^3} = \omega_{bin}, \quad (28)$$

up to a scalar factor. See Appendix C (Supplementary material) for explicit computation.

In this way, the disparity binocular cells couple in a natural way positions, identified with points in  $\mathbb{R}^3$ , and orientations in  $\mathbb{S}^2$ , identified with three-dimensional unitary tangent vectors. As already observed in Remark 3.2, the geometry of the stereo vision is not solved only with these punctual and directional arguments, but there is the need to take into accounts suitable type of connections. In Alibhai and Zucker (2000); Li and Zucker (2003, 2006), Zucker et al. proposed a model that considered the curvature of monocular structures as an additional variable. Instead, we propose to consider simple monocular cells selective for orientation, and to insert the notion of curvature directly into the definition of connection. It is therefore natural to introduce the perceptual space via the manifold  $\mathbb{R}^3 \times \mathbb{S}^2$ , in line with the theoretical toolbox proposed in Miolane and Pennec (2016) to generalize 2D neurogeometry to 3D images, and adapt this framework to our problem, looking for appropriate curves.

### 3.6 A perceptual model in the space of 3D position-orientation

We now derive the objects in Figure 3A. We have clarified (end of Section 3.5) that binocular cells are parameterized by points in  $\mathbb{R}^3$ , and orientations in  $\mathbb{S}^2$ . An element  $\xi$  of the space  $\mathbb{R}^3 \times \mathbb{S}^2$  it is defined by a point  $p = (p_1, p_2, p_3)$  in  $\mathbb{R}^3$  and an unitary vector  $n \in \mathbb{S}^2$ . Since the topological dimension of this geometric object is 2, we introduce the classical spherical coordinates  $(\theta, \varphi)$  such that  $n = (n_1, n_2, n_3) \in \mathbb{S}^2$  can be parameterized as:

$$\begin{aligned} n_1 &= \cos \theta \sin \varphi \\ n_2 &= \sin \theta \sin \varphi \\ n_3 &= \cos \varphi \end{aligned} \quad (29)$$

with  $\theta \in [0, 2\pi]$  and  $\varphi \in (0, \pi)$ . The ambiguity that arises using local coordinate chart is overcome by the introduction of a second chart, covering the singular points.

Translations and rotations are expressed using the group law of the three-dimensional special Euclidean group  $SE(3)$ , defining the group action

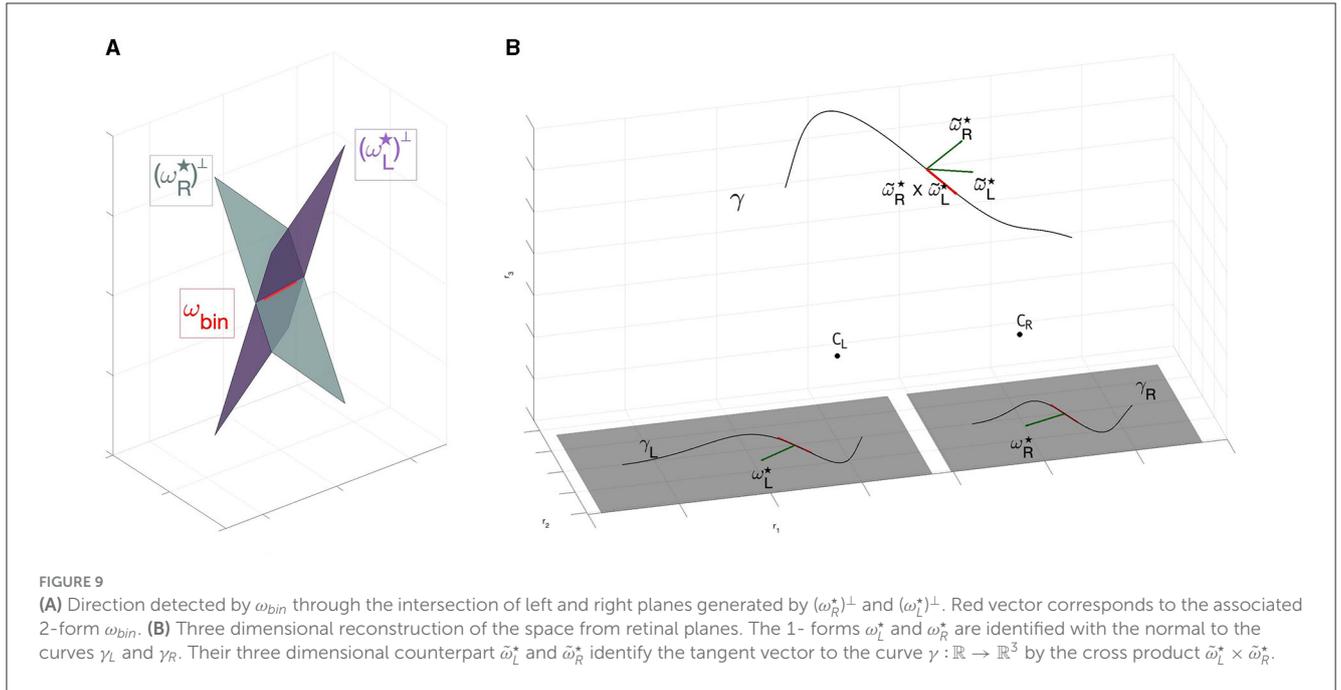
$$\sigma : \mathbb{R}^3 \times \mathbb{S}^2 \times SE(3) \longrightarrow \mathbb{R}^3 \times \mathbb{S}^2 \text{ s.t. } \sigma((p, n), (q, R)) = (Rp+q, Rn), \quad (30)$$

with  $(p, n) \in \mathbb{R}^3 \times \mathbb{S}^2$ ,  $(q, R) \in SE(3)$ , namely  $R \in SO(3)$  tridimensional rotation, and  $q \in \mathbb{R}^3$ .

#### 3.6.1 Stereo sub-Riemannian geometry

The emergence of a privileged direction in  $\mathbb{R}^3$  (associated with the tangent vector to the stimulus) is the reason why we endow  $\mathbb{R}^3 \times \mathbb{S}^2$  with a sub-Riemannian structure that favors the direction in 3D identified by  $\omega_{bin}$ .

Formally, we consider admissible movements in  $\mathbb{R}^3 \times \mathbb{S}^2$  described by vector fields:



**FIGURE 9** (A) Direction detected by  $\omega_{bin}$  through the intersection of left and right planes generated by  $(\omega_R^*)^\perp$  and  $(\omega_L^*)^\perp$ . Red vector corresponds to the associated 2-form  $\omega_{bin}$ . (B) Three dimensional reconstruction of the space from retinal planes. The 1-forms  $\omega_L^*$  and  $\omega_R^*$  are identified with the normal to the curves  $\gamma_L$  and  $\gamma_R$ . Their three dimensional counterpart  $\tilde{\omega}_L^*$  and  $\tilde{\omega}_R^*$  identify the tangent vector to the curve  $\gamma: \mathbb{R} \rightarrow \mathbb{R}^3$  by the cross product  $\tilde{\omega}_L^* \times \tilde{\omega}_R^*$ .

$$\begin{aligned}
 Y_{\mathbb{R}^3, \xi} &= \sin \varphi \cos \theta \partial_1 + \sin \varphi \sin \theta \partial_2 + \cos \varphi \partial_3 \\
 Y_{\theta, \xi} &= -\frac{1}{\sin \varphi} \partial_\theta \\
 Y_{\varphi, \xi} &= \partial_\varphi
 \end{aligned} \tag{31}$$

with  $\xi \in \mathbb{R}^3 \times \mathbb{S}^2$  for  $\varphi \neq 0, \varphi \neq \pi$ . The admissible tangent space<sup>9</sup> at a point  $\xi$

$$\mathcal{A}_\xi := \text{span}\{Y_{\mathbb{R}^3, \xi}, Y_{\theta, \xi}, Y_{\varphi, \xi}\} \tag{32}$$

encodes the coupling between position and orientations, as remarked by [Duits and Franken \(2011\)](#). In particular, the vector field  $Y_{\mathbb{R}^3}$  identifies the privileged direction in  $\mathbb{R}^3$ , while  $Y_\theta$  and  $Y_\varphi$  allow changing this direction, involving just orientation variables of  $\mathbb{S}^2$ . The vector fields  $\{Y_{\mathbb{R}^3}, Y_\theta, Y_\varphi\}$  and their commutators generate the tangent space of  $\mathbb{R}^3 \times \mathbb{S}^2$  in a point, allowing to connect every point of the manifold using privileged directions (*Hörmander condition*). Furthermore, it is possible to define a sub-Riemannian structure by choosing a scalar product on the admissible tangent bundle  $\mathcal{A}$ : the simplest choice is to declare the vector fields  $\{Y_{\mathbb{R}^3}, Y_\theta, Y_\varphi\}$  orthonormal, considering on  $\mathbb{S}^2$  the distance inherited from the immersion in  $\mathbb{R}^3$  with the Euclidean metric.

### 3.6.2 Change of variables

We have already expressed the change of variable in the variables  $(x, y, d)$  to  $(r_1, r_2, r_3)$  in Equation (5). However, the cortical coordinates also contain the angular variables  $\theta_R$  and  $\theta_L$  which involve the introduction of the spherical coordinates  $\theta, \varphi$ .

<sup>9</sup> see Appendix A ([Supplementary material](#)) for the definition of admissible tangent space.

To identify a change of variable among these variables, we first introduce the function

$$(r_1, r_2, r_3, \theta, \varphi) \xrightarrow{F} (x, y, d, \theta_L, \theta_R):$$

$$F: \mathbb{R}^3 \times \mathbb{S}^2 \longrightarrow \mathbb{R}^3 \times \mathbb{S}^2$$

$$\begin{pmatrix} r_1 \\ r_2 \\ r_3 \\ \theta \\ \varphi \end{pmatrix} \mapsto \begin{pmatrix} \frac{fr_1}{r_3} \\ \frac{fr_2}{r_3} \\ \frac{cf}{r_3} \\ \tan^{-1}\left(\frac{r_3 \sin \theta \cos \varphi - r_2 \cos \varphi}{r_3 \cos \theta \sin \varphi - (c+r_1) \cos \varphi}\right) \\ \tan^{-1}\left(\frac{r_3 \sin \theta \cos \varphi - r_2 \cos \varphi}{r_3 \cos \theta \sin \varphi - (c-r_1) \cos \varphi}\right) \end{pmatrix}, \tag{33}$$

where the retinal right angle  $\theta_R = \tan^{-1}\left(\frac{r_3 \sin \theta \cos \varphi - r_2 \cos \varphi}{r_3 \cos \theta \sin \varphi - (c+r_1) \cos \varphi}\right)$  and the retinal left angle  $\theta_L = \tan^{-1}\left(\frac{r_3 \sin \theta \cos \varphi - r_2 \cos \varphi}{r_3 \cos \theta \sin \varphi - (c-r_1) \cos \varphi}\right)$  are obtained considering Equation (6).

Analogously, it is possible to define the change of variable  $(x, y, d, \theta_L, \theta_R) \xrightarrow{G} (r_1, r_2, r_3, \theta, \varphi)$ :

$$G: \mathbb{R}^3 \times \mathbb{S}^2 \longrightarrow \mathbb{R}^3 \times \mathbb{S}^2$$

$$\begin{pmatrix} x \\ y \\ d \\ \theta_R \\ \theta_L \end{pmatrix} \mapsto \begin{pmatrix} \frac{cx}{d} \\ \frac{cy}{d} \\ \frac{d}{d} \\ \frac{cf}{d} \\ \tan^{-1}\left(\frac{2 \sin \theta_R \sin \theta_L}{\sin(\theta_R + \theta_L)}\right) \\ \tan^{-1}\left(\frac{\sqrt{\sin^2(\theta_R + \theta_L) + 4 \sin^2 \theta_R \sin^2 \theta_L}}{\sin(\theta_R - \theta_L)}\right) \end{pmatrix}, \tag{34}$$

where the angles  $\theta = \tan^{-1}\left(\frac{2 \sin \theta_R \sin \theta_L}{\sin(\theta_R + \theta_L)}\right)$  and  $\varphi = \tan^{-1}\left(\frac{\sqrt{\sin^2(\theta_R + \theta_L) + 4 \sin^2 \theta_R \sin^2 \theta_L}}{\sin(\theta_R - \theta_L)}\right)$  are obtained considering that  $\tan \theta = \frac{(\tilde{Y}_{\mathbb{R}^3})_2}{(\tilde{Y}_{\mathbb{R}^3})_1}$  and  $\tan \varphi = \frac{\sqrt{(\tilde{Y}_{\mathbb{R}^3})_1^2 + (\tilde{Y}_{\mathbb{R}^3})_2^2}}{(\tilde{Y}_{\mathbb{R}^3})_3}$ .

### 3.6.3 Integral curves

The connectivity of the space is described by admissible curves of the vector fields spanning  $\mathcal{A}$ . In particular, a curve  $\Gamma : [0, T] \rightarrow \mathbb{R}^3 \times \mathbb{S}^2$  is said to be *admissible*<sup>10</sup> if:

$$\dot{\Gamma}(t) \in \mathcal{A}_{\Gamma(t)}, \leftrightarrow \dot{\Gamma}(t) = a(t)\vec{Y}_{\mathbb{R}^3, \Gamma(t)} + b(t)\vec{Y}_{\theta, \Gamma(t)} + c(t)\vec{Y}_{\varphi, \Gamma(t)}, \tag{35}$$

where  $a, b, c$  are sufficiently smooth function on  $[0, T]$ . We will consider a particular case of these admissible curves, namely constant coefficient integral curves with  $a(t) = 1$ , since the vector field  $Y_{\mathbb{R}^3}$  represents the tangent direction of the 3D stimulus (and so it never vanishes):

$$\dot{\Gamma}(t) = \vec{Y}_{\mathbb{R}^3, \Gamma(t)} + c_1 \vec{Y}_{\theta, \Gamma(t)} + c_2 \vec{Y}_{\varphi, \Gamma(t)}, \tag{36}$$

with  $c_1$  and  $c_2$  varying in  $\mathbb{R}$ .

These curves can be thought of in terms of trajectories in  $\mathbb{R}^3$  describing a movement in the  $\vec{Y}_{\mathbb{R}^3}$  direction, which can eventually change according to  $\vec{Y}_{\theta}$  and  $\vec{Y}_{\varphi}$ . An example of the fan of integral curves was shown in the Introduction in Figure 3B.

It is worth noting that in the case described by coefficients  $c_1$  and  $c_2$  equal to zero, the 3D trajectories would be straight lines in  $\mathbb{R}^3$ ; by varying the coefficients  $c_1$  and  $c_2$  in  $\mathbb{R}$ , we allow the integral curves to follow curved trajectories, twisting and bending in all space directions.

Formally, the amount of “twisting and bending” in space is measured by introducing the notions of curvature and torsion. We then investigate how these measurements are encoded in the parameters of the family of integral curves, and what constraints have to be imposed to obtain different typologies of curves.

Remark 3.4. The 3D projection of the integral curves (36) will be denoted  $\gamma$  and satisfy  $\dot{\gamma}(t) = (\cos \theta(t) \sin \varphi(t), \sin \theta(t) \sin \varphi(t), \cos \varphi(t))^T$ . Classical instruments of differential geometry let us compute the curvature and the torsion of the curve  $\gamma(t)$ :

$$k = \sqrt{(\dot{\varphi})^2 + \sin^2 \theta (\dot{\theta})^2},$$

$$\tau = \frac{1}{k^2} (-\cos \varphi \sin^2 \varphi (\dot{\theta})^3 - \sin \varphi \dot{\varphi} \ddot{\theta} + \dot{\theta} (-2 \cos \varphi (\dot{\varphi})^2 + \sin \varphi \ddot{\varphi})). \tag{37}$$

Using the explicit expression of the vector fields  $Y_{\theta}$  and  $Y_{\varphi}$  in Equation (36), we get

$$\dot{\theta} = -\frac{c_1}{\sin \varphi}, \quad \dot{\varphi} = c_2, \tag{38}$$

from which it follows that:

$$k = \sqrt{c_1^2 + c_2^2}$$

$$\tau = c_1 \cotan \varphi. \tag{39}$$

Proposition 3.2. By varying the parameters  $c_1$  and  $c_2$  in (39) where we explicitly find solutions of (36), we have:

1. If  $\varphi = \frac{\pi}{2}$  then  $k = \sqrt{c_1^2}$ ,  $\tau = 0$ , and so the family of curves (36) are circles of radius  $1/c_1^2$  on the fronto-parallel plane  $r_3 = \text{cost}$ .

2. If  $\varphi = \varphi_0$ , with  $\varphi_0 \neq \pi/2$ , then  $k = \sqrt{c_1^2}$  and  $\tau = c_1 \cotan \varphi_0$ , and so the family of curves (36) are  $r_3$ -helices.
3. If  $\theta = \theta_0$  then  $k = \sqrt{c_2^2}$ ,  $\tau = 0$ , and so the family of curves (36) are circles of radius  $1/c_2^2$  in the osculating planes.

*Proof.* The computation follows immediately from the computed curvature and torsion of (39) and classical results of differential geometry.  $\square$

Remark 3.5. If we know the value of the curvature  $k$ , and we have one free parameter,  $c_2$ , in the definition of the integral curves (36), then we are in the setting of Proposition 2.1. In fact, the coefficient  $c_1$  is obtained by imposing  $c_1 = \pm \sqrt{k^2 - c_2^2}$ , and in particular the component that remains to be determined is the torsion.

Examples of particular cases of the integral curves (36) according to Proposition 3.2 and Remark 3.5 are visualized in Figure 10.

## 4 Comparison with experimental data

Our sub-Riemannian model enjoys some consistency with the biological and psychophysical literature. We here describe some initial connections.

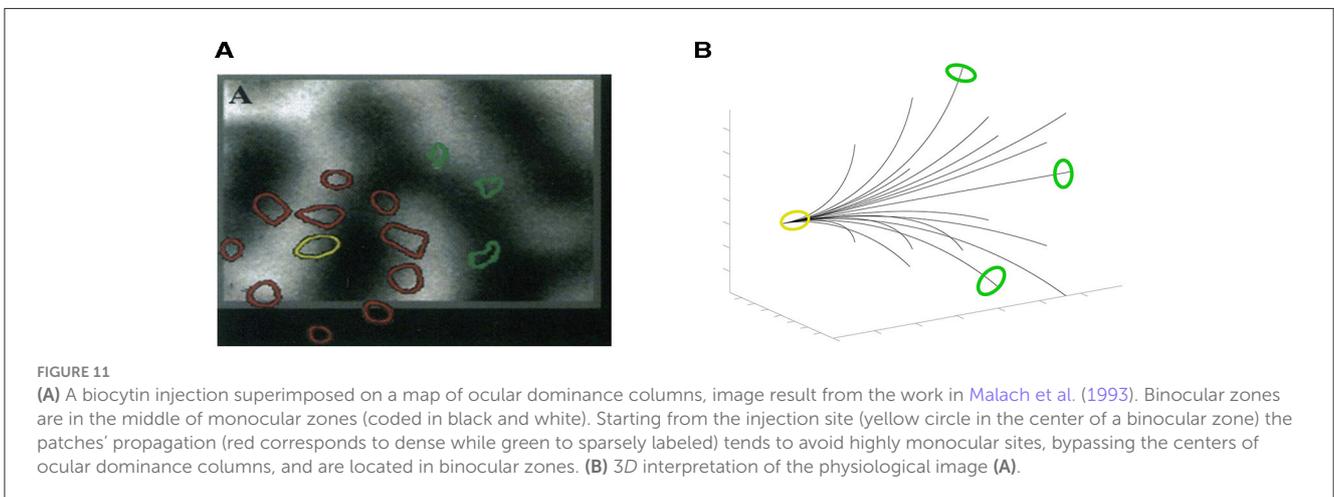
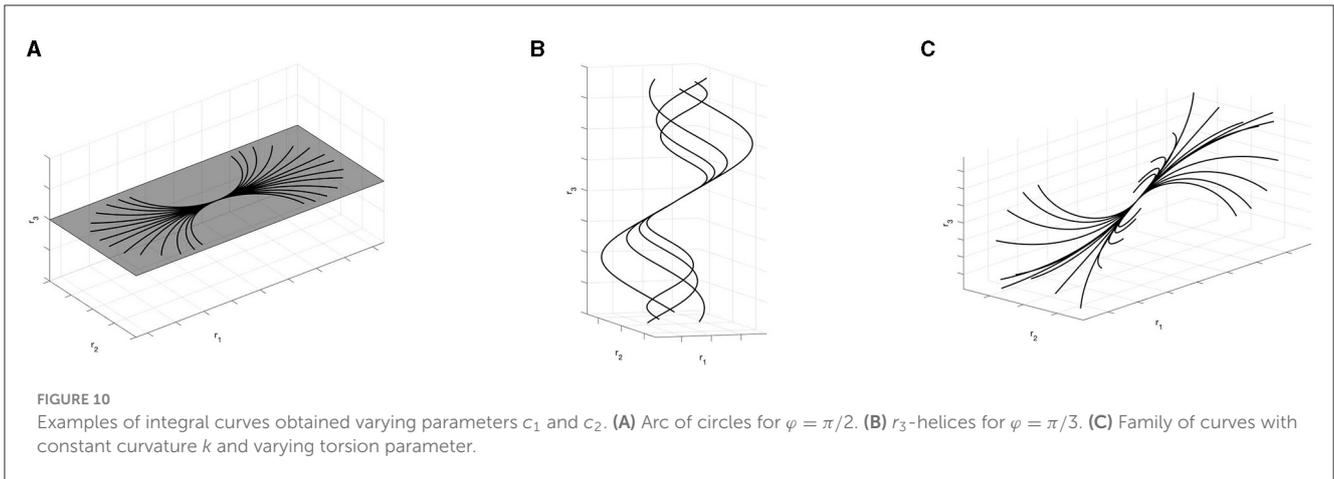
### 4.1 Biological connections

The foundation for building our sub-Riemannian model of stereo was a model of curve continuation, motivated by the orientation columns at each position. The connections between cells in nearby columns were, in turn, a geometric model of long-range horizontal connections in visual cortex (Bosking et al., 1997). In the Introduction we cartooned aspects of the cortical architecture that support binocular processing. Although the inputs are organized into ocular dominance bands, there is no direct evidence for “stereo columns” in V1 analogous to the monocular orientation columns. But such columns are not strictly necessary for our model. Rather, what is central is how information propagates. We showed in Figure 1C that there is evidence of long-range connections between binocular cells, and our model informs, abstractly, what information could propagate along these connections. Although less extensive than in the monocular case, some measurements are beginning to emerge that are informative.

The Grinwald group first established the presence of long-range connections between binocular cells (Malach et al., 1993) (see also Figure 11A), using biocytin. This is a molecule that is taken up by neurons, propagates directly along neuronal processes and is deposited at excitatory synapses. These results were refined, more recently, by the Fitzpatrick group (Scholl et al., 2022), using *in vivo* calcium imaging. As shown in Figure 1C the authors demonstrated both the monocular and the binocular inputs for stereo, and (not shown) the dependence on orientations.

More precisely, Malach et al. (1993) showed selective *anisotropic* connectivity among binocular regions: the biocytin tracer does not spread uniformly, but rather is highly directional with

<sup>10</sup> sometimes the term *horizontal* is preferred.



distance from the injection point. (This was the case with monocular biocytin injections as well.) Putting this together with Scholl et al. (2022), we interpret the anisotropy as being related to (binocular) orientation (Scholl et al., 2022), which is what the integral curves of our vector fields suggest. Our 3D association fields are strongly directional, and information propagates preferentially in the direction of (the starting point of) the curve. An example can be seen in Figure 11B, where the fan of integral curves (36) is represented, superimposed with colored patches, following the experiment proposed in Malach et al. (1993). We look forward to more detailed experiments along these lines.

## 4.2 Psychophysics and association fields

In this section, we show that the connections described by the integral curves in our model can be related to the geometric relationships from psychophysical experiments on perceptual organization of oriented elements in  $\mathbb{R}^3$ ; in other words, that our connections serve as a generalization of the concept of an association field in 3D.

### 4.2.1 Toward a notion of association field for 3D contours

The perception of continuity between two elements of position-orientation in  $\mathbb{R}^3$  has been studied experimentally. To start, Kellman, Garrigan, and Shipley (Kellman et al., 2005a,b) introduce 3D reliability, as a way to extend to 3D the experiments of Field, Hayes and Hess (Field et al., 1993) in 2D.

Particularly, in a system of 3D Cartesian coordinates, it is possible to introduce oriented edges  $E$  at the application point  $(r_1, r_2, r_3)^T$  and with an orientation identified with the angles  $\theta$  and  $\varphi$ . This orientation can be read, in our case, through the direction expressed by  $(\cos \theta \sin \varphi, \sin \theta \sin \varphi, \cos \varphi)^T$ . For an initial edge  $E_0$ , with application point on the origin of the coordinate system  $(0, 0, 0)^T$  and orientation lying on the  $r_1$ -axis, described by  $\theta = 0, \varphi = \pi/2$ , the range of possible orientations  $(\theta, \varphi)^{11}$  for 3D-relatable edges with  $E_0$  is given by:

11 The angle  $\varphi$  here has been modified to be compatible with our set of coordinates. The relationship between the angle  $\tilde{\varphi}$  in works (Kellman et al., 2005a,b) can be expressed as :  $\tilde{\varphi} = \text{acos}(\sin \varphi) + \pi$ .

$$\tan^{-1}\left(\frac{r_2}{r_1}\right) \leq \theta \leq \frac{\pi}{2} \quad \text{and} \quad \frac{\pi}{2} \leq \frac{3\pi}{2} - \varphi \leq \tan^{-1}\left(\frac{r_3}{r_1}\right). \quad (40)$$

The bound on these equations identified with the quantity  $\frac{\pi}{2}$  incorporates the 90 degree constraint in three dimensions, while the bounds defined by the inverse of the tangent express the absolute orientation difference between the reference edge  $E_0$  and an edge positioned at the arbitrary oriented point  $E_{(r_1, r_2, r_3)}$  so that its linear extension intersects  $E_0$ ; see [Kellman et al. \(2005a,b\)](#) for further details.

Numerical simulations allow us to visually represent an example of the 3D positions and orientations that meet the 3D reliability criteria. Starting from an initial edge  $E_0$  with endpoints in  $(p_{01}, p_{02}, p_{03})^T$  and orientation on the  $e_1$ -axis, we represent for an arbitrary point  $(p_1, p_2, p_3)^T$  the limit of the reliable orientation  $(\theta, \varphi)$ . Results are shown in [Figure 12A](#).

**Remark 4.1.** By projecting on the retinal planes of the 3D fan of reliable points, it is possible to notice that these projections are in accordance with the notion of 3D compatibility field in [Alibhai and Zucker \(2000\)](#) (see [Figures 12B, C](#)).

Psychophysical studies, see [Hess and Field \(1995\)](#); [Hess et al. \(1997\)](#); [Deas and Wilcox \(2015\)](#), have investigated the properties of the curves that are suitable for connecting these reliable points. These curves are well-described as smooth and monotonic. In particular, using non-oriented contour elements for contours, [Hess et al. \(1997\)](#) indicate that contour elements can be effectively grouped based primarily on the good continuation of contour elements in depth. This statement is confirmed by the more recent work of [Deas and Wilcox \(2015\)](#) who, in addition, observe that detection of contours defined by regular depth continuity is faster than detection of discontinuous contours. All these results support the existence of depth grouping operations, arguing for the extension of Gestalt principles of continuity and smoothness in three dimensional space. Finally, on the relationship of the three-dimensional curves to 2-dimensional association fields, see [Kellman et al. \(2005b\)](#); [Khuu et al. \(2016\)](#). These authors have assumed that the strength of the reliable edges in the co-planar planes of  $E_0$  must meet the relations of the bi-dimensional association fields of [Field et al. \(1993\)](#).

#### 4.2.2 Compatibility with the sub-Riemannian model

To model the associations underlying the 3D perceptual organization discussed in the previous paragraph, we consider again the constant coefficient family of integral curves studied in (36):

$$\dot{\Gamma}(t) = \bar{Y}_{\mathbb{R}^3, \Gamma(t)} + c_1 \bar{Y}_{\theta, \Gamma(t)} + c_2 \bar{Y}_{\varphi, \Gamma(t)}, \quad \text{with } c_1, c_2 \in \mathbb{R}. \quad (41)$$

Importantly, these curves locally connect the association fan generated by the geometry of 3D reliability. In particular, [Figure 13B](#) shows the family of the horizontal curves connecting the

initial point  $E_0$  with 3D reliable edges ([Figure 13A](#)). These curves are computed using Matlab solver function `ode45`.

In analogy with the experiment of [Field, Hayes, and Hess in Field et al. \(1993\)](#), we choose to represent non-reliable edges to the left of the starting point  $E_0$ , while on the right are 3D reliable edges. So, filled lines of the integral curves indicate the correlation between the central horizontal element  $E_0$  and the ones on its right, while dotted lines connect the starting point  $E_0$  with elements not correlated with it, as represented on the left part of the image.

Restricting the curves on the neighborhood of co-planar planes with an arbitrary edge  $E$ , we have different cases. First, on the  $r_1$ - $r_2$  plane (fronto-parallel) and the  $r_1$ - $r_3$  plane we have arcs of circle, as proved with [Proposition 3.2](#). Furthermore, for an arbitrary plane in  $\mathbb{R}^3$  containing an edge  $E$ , we observe that the curves generating with fixed angle  $\varphi$  are helices, and locally they satisfy the bidimensional constraint in the plane. Examples can be found in [Figures 13C–E](#). In particular, the curves displayed in [Figures 13C, D](#) are well in accordance with the curves of the Citti-Sarti model, depicted in [Figure 5](#).

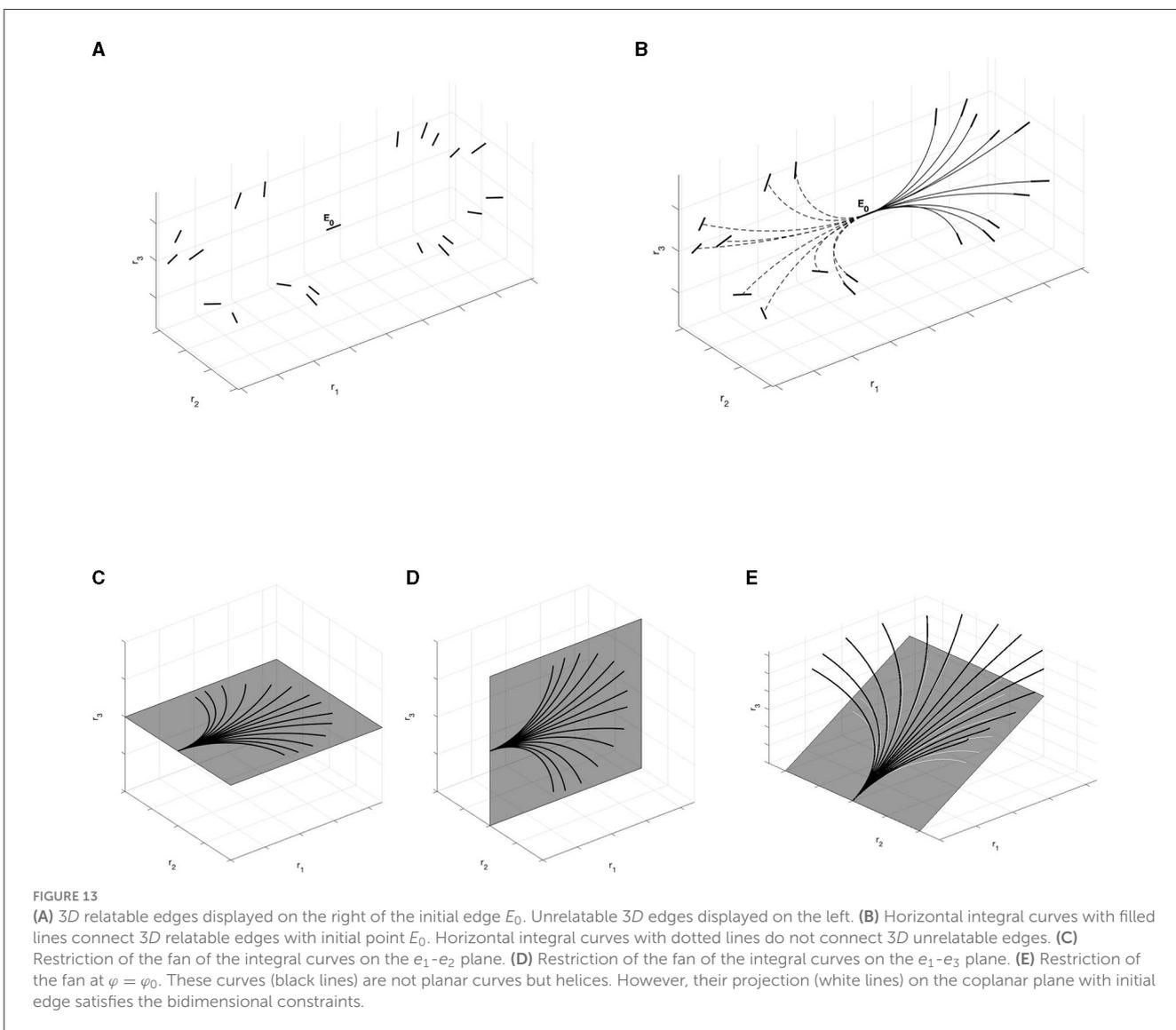
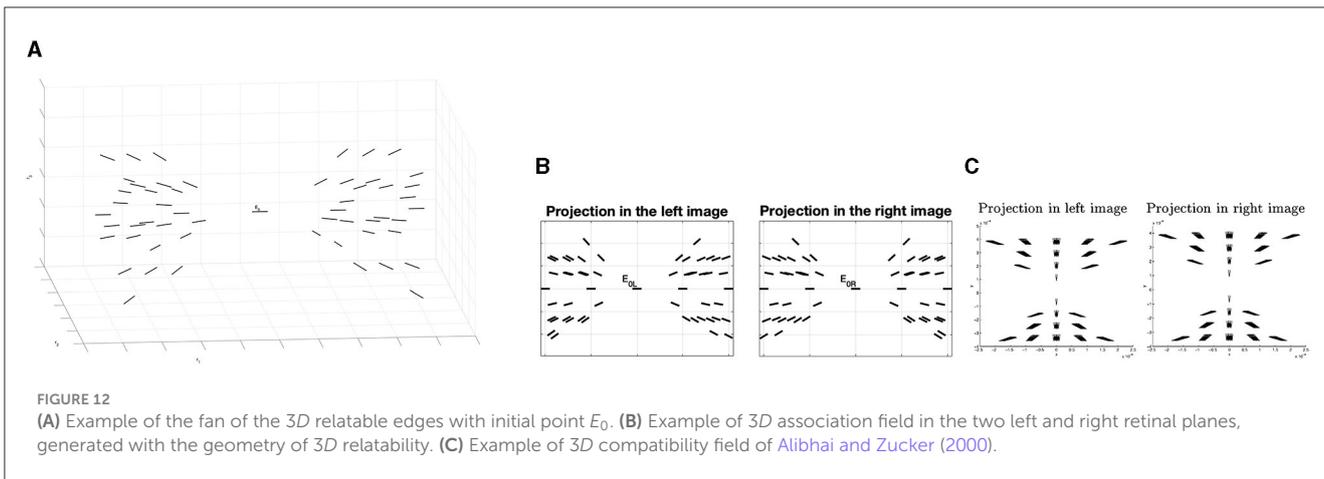
One final connection with the psychophysical literature concerns how depth discrimination thresholds increase exponentially with distance ([Burge, 2020](#) and references therein). This is related to how the fan of integral curves “spreads out” with distance ([Figures 11, 12](#)), which is also exponential. These notions are developed more fully in [Bolelli et al. \(2023a\)](#).

### 4.3 Integration of contours and stereo correspondence problem

Although the goal of this paper is not to solve the stereo correspondence problem, we can show how our geometry is helpful in understanding how to match left and right points and features. These ideas are developed more fully in [Bolelli \(2023\)](#).

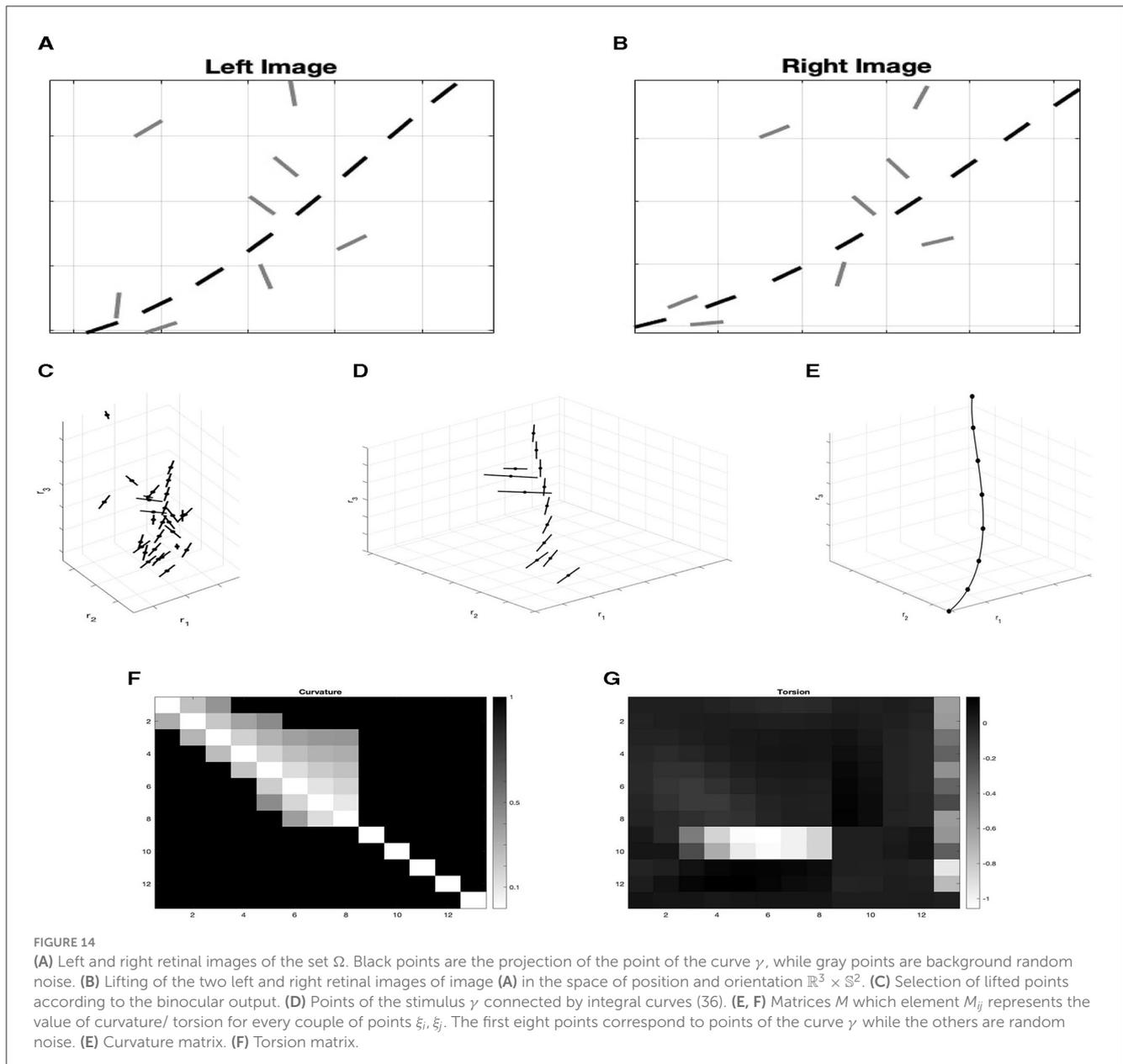
Inspired by [Hess and Field \(1995\)](#), we consider a path stimulus  $\gamma$  interpreted as a contour, embedded in a background of randomly oriented elements: left and right retinal visual stimuli are depicted in [Figure 14A](#). We perform an initial, simplified lift of the retinal images to a set  $\Omega \subset \mathbb{R}^3 \times \mathbb{S}^2$ . This set contains all the possible corresponding points, obtained by coupling left and right points which share the same  $y$  retinal coordinate, see [Figure 14B](#). The set  $\Omega$  contains false matches, namely points that do not belong to the original stimulus. It is the task of correspondence to eliminate these false matches.

We compute for every lifted point the binocular output  $O_B$  of Equation (17). This output can be seen as a probability measure that gives information on the correspondence of the pair of left and right points. We then simply evaluate which are the points with the highest probability of being in correspondence, applying a process of suppression of the non-maximal pairs over the fiber of disparity. In this way, noise points are removed ([Figure 14C](#)). We now directly exploit the Gestalt rule of good continuation by filtering out any couple of elements with high curvature. This qualitative rule could be quantitatively modeled by considering the statistics of distribution of curvature and torsion in natural 3D images ([Geisler and Perry, 2009](#)). The remaining noise elements are orthogonal to the directions of the elements of the curve that we



would like to reconstruct. Calculating numerically the coefficients  $c_1$  and  $c_2$  of integral curves (36) that connect all the remaining pairs of points, we can obtain for every pair the value of curvature and torsion using (39).

Figures 14E, F show matrices  $M$  representing the values of curvature or torsion for every pair of points  $\xi_i, \xi_j$  in the element  $M_{ij}$ . In particular, we observe that random points are characterized by very high curvature and deviating torsion. So, by discarding



these high values, we select only the three-dimensional points of the curve  $\gamma$ , which are well-connected by the integral curves, as shown in Figure 14D. This is in accordance with the idea developed in Alibhai and Zucker (2000); Li and Zucker (2003, 2006), where curvature and torsion provide constraints for reconstruction in 3D.

In this artificial example we assumed that local edge elements have already been detected. Our goal was simply proof-of-concept. To apply this approach to realistic images, of course, stages of edge detection would have to be adopted, for which there is a huge literature well outside the scope of our theoretical study.

## 5 Summary and conclusions

Understanding good continuation in depth, like good continuation for planar contours, can benefit from basic

physiological constraints; from psychophysical performance measures, and from mathematical modeling. In particular, good continuation in the plane is supported by orientation selectivity and cortical architecture (orientation columns), by association field grouping performance, and by geometric modeling. We showed that the same should be true for good continuation in depth. However, while the psychophysical data may be comparable, the physiological data are weaker and the geometry of continuation is not as well-understood. In this paper, we introduced the neuro-geometry of stereo vision to fill this gap. It is strongly motivated by an analogical extension to 3D of 2D geometry, while respecting the psychophysics. In the end, it allowed us to be precise about the type of geometry that is relevant for understanding stereo abstractly, and concretely was highly informative toward the physiology. Although a “stereo columnar architecture” is not obvious from the anatomy, it is well-formed computationally.

Technically, we proposed a sub-Riemannian model on the space of position and orientation  $\mathbb{R}^3 \times \mathbb{S}^2$  for the description of the perceptual space of the neural cells involved. This geometrical structure favors the tangent direction of a 3D curve stimulus. The integral curves of the sub-Riemannian structure encode the notions of curvature and torsion within their coefficients, and are introduced to describe the connections between elements. This model can be seen as an extension in the three-dimensional scene of the 2-dimensional association field. In particular, the integral curves of the sub-Riemannian structure of the 3D space of position-orientation are exactly those that locally correspond to psychophysical association fields.

Although the goal of this paper is not to solve the stereo correspondence problem, we have seen how the geometry we propose is a good starting point to understand how to match left and right points and features. We used binocular receptive fields to prioritize orientation preferences and orientation differences under the assumption that neuronal circuitry has developed to facilitate the interpolation of contours in 3D space. On the other side, the neurogeometrical method has been coupled with a probabilistic methods for example in Sanguinetti et al. (2010) and Sarti and Citti (2015). Here, the authors studied an analogous problem for generation of perceptual units in monocular vision: they introduced stochastic differential equations, analogous to the integral curves of vector fields, and used its probability density as a kernel able to generate monocular perceptual units. In Montobbio et al. (2019), the probability kernel is built in a direction starting from the receptive fields. A future development of the model will consist in adapting the technique of Sarti and Citti (2015) to find the probability of the co-occurrence between two elements, and individuate percepts in 3D space. Individuation of percepts through harmonic analysis on the sub-Riemannian structure has been proposed in the past, both for 2D spatial stimuli (Sarti and Citti, 2015) and in 2D + time spatio-temporal stimuli (Barbieri et al., 2014a). It would be interesting to develop a similar analysis and extend it to stereo vision.

## Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## References

- Abbasi-Sureshjani, S., Favali, M., Citti, G., Sarti, A., and ter Haar Romeny, B. M. (2017). Curvature integration in a 5D kernel for extracting vessel connections in retinal images. *IEEE Trans. Image Process.* 27, 606–621. doi: 10.1109/TIP.2017.2761543
- Agrachev, A., Barilari, D., and Boscain, U. (2019). *A Comprehensive Introduction to Sub-Riemannian Geometry*. Vol. 181. Cambridge: Cambridge University Press. doi: 10.1017/9781108677325
- Alibhai, S., and Zucker, S. W. (2000). "Contour-based correspondence for stereo," in *Computer Vision - ECCV 2000* (Berlin; Heidelberg: Springer), 314–330. doi: 10.1007/3-540-45054-8\_21
- Anderson, B. L., Singh, M., and Fleming, R. W. (2002). The interpolation of object and surface structure. *Cogn. Psychol.* 44, 148–190. doi: 10.1006/cogp.2001.0765
- Anzai, A., Ohzawa, I., and Freeman, R. (1999a). Neural mechanisms for encoding binocular disparity: receptive field position versus phase. *J. Neurophysiol.* 82, 874–890. doi: 10.1152/jn.1999.82.2.874
- Anzai, A., Ohzawa, I., and Freeman, R. (1999b). Neural mechanisms for processing binocular information I. Simple cells. *J. Neurophysiol.* 82, 891–908. doi: 10.1152/jn.1999.82.2.891

## Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. MB, GC, and AS were supported by EU Project, GHAI, Geometric and Harmonic Analysis with Interdisciplinary Applications, H2020-MSCA-RISE-2017. SZ was supported in part by US NIH EY031059 and by US NSF CRCNS 1822598.

## Acknowledgments

The current article is part of the first named author's Ph.D. thesis (Bolelli, 2023) and a preprint version is available at Bolelli et al. (2023b).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fcomp.2023.1142621/full#supplementary-material>

- August, J., and Zucker, S. W. (2000). "The curve indicator random field: curve organization via edge correlation," in *The Kluwer International Series in Engineering and Computer Science* (Boston, MA: Springer US), 265–288. doi: 10.1007/978-1-4615-4413-5\_15
- Barbieri, D., Citti, G., Cocci, G., and Sarti, A. (2014a). A cortical-inspired geometry for contour perception and motion integration. *J. Math. Imaging Vision* 49, 511–529. doi: 10.1007/s10851-013-0482-z
- Barbieri, D., Citti, G., Sanguinetti, G., and Sarti, A. (2012). An uncertainty principle underlying the functional architecture of V1. *J. Physiol.* 106, 183–193. doi: 10.1016/j.jphysparis.2012.03.001
- Barbieri, D., Citti, G., and Sarti, A. (2014b). How uncertainty bounds the shape index of simple cells. *J. Math. Neurosci.* 4, 5. doi: 10.1186/2190-8567-4-5
- Baspinar, E., Sarti, A., and Citti, G. (2020). A sub-riemannian model of the visual cortex with frequency and phase. *J. Math. Neurosci.* 10, 1–31. doi: 10.1186/s13408-020-00089-6
- Ben-Shahar, O., and Zucker, S. W. (2004). Geometrical computations explain projection patterns of long-range horizontal connections in visual cortex. *Neural Comput.* 16, 445–476. doi: 10.1162/089976604772744866
- Blasdel, G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex. *J. Neurosci.* 12, 3139–3161. doi: 10.1523/JNEUROSCI.12-08-03139.1992
- Bolelli, M. V. (2023). *Neurogeometry of stereo vision* (theses). Sorbonne Université; Università degli studi, Bologna, Italy.
- Bolelli, M. V., Citti, G., Sarti, A., and Zucker, S. (2023a). "A neurogeometric stereo model for individuation of 3D perceptual units," in *International Conference on Geometric Science of Information* (Springer), 53–62. doi: 10.1007/978-3-031-38271-0\_6
- Bolelli, M. V., Citti, G., Sarti, A., and Zucker, S. W. (2023b). Good continuation in 3D: the neurogeometry of stereo vision. *arXiv preprint arXiv:2301.04542*.
- Bosking, W. H., Zhang, Y., Schofield, B., and Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *J. Neurosci.* 17, 2112–2127. doi: 10.1523/JNEUROSCI.17-06-02112.1997
- Bridge, H., and Cumming, B. G. (2001). Responses of macaque V1 neurons to binocular orientation differences. *J. Neurosci.* 21, 7293–7302. doi: 10.1523/JNEUROSCI.21-18-07293.2001
- Bridge, H., Cumming, B. G., and Parker, A. J. (2001). Modeling V1 neuronal responses to orientation disparity. *Visual Neurosci.* 18, 879–891. doi: 10.1017/S0952523801186049
- Burge, J. (2020). Image-computable ideal observers for tasks with natural stimuli. *Annu. Rev. Vision Sci.* 6, 491–517. doi: 10.1146/annurev-vision-030320-041134
- Burge, J., and Geisler, W. S. (2014). Optimal disparity estimation in natural stereo images. *J. Vision* 14, 1. doi: 10.1167/14.2.1
- Cagenello, R., and Rogers, B. J. (1993). Anisotropies in the perception of stereoscopic surfaces: the role of orientation disparity. *Vision Res.* 33, 2189–2201. doi: 10.1016/0042-6989(93)90099-1
- Chang, J. T., Whitney, D., and Fitzpatrick, D. (2020). Experience-dependent reorganization drives development of a binocularly unified cortical representation of orientation. *Neuron* 107, 338–350. doi: 10.1016/j.neuron.2020.04.022
- Citti, G., Grafakos, L., Pérez, C., Sarti, A., and Zhong, X. (2015). Harmonic and geometric analysis. *Springer*. doi: 10.1007/978-3-0348-0408-0
- Citti, G., and Sarti, A. (2006). A cortical based model of perceptual completion in the roto-translation space. *J. Math. Imaging Vision* 24, 307–326. doi: 10.1007/s10851-005-3630-2
- Citti, G., and Sarti, A. (2014). *Neuromathematics of Vision*. Vol. 32. Berlin; Heidelberg: Springer. doi: 10.1007/978-3-642-34444-2
- Cumming, B. G., and DeAngelis, G. C. (2001). The physiology of stereopsis. *Annu. Rev. Neurosci.* 24, 203–238. doi: 10.1146/annurev.neuro.24.1.203
- Daugman, J. G. (1985). Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A* 2, 1160. doi: 10.1364/JOSAA.2.001160
- Deas, L. M., and Wilcox, L. M. (2014). Gestalt grouping via closure degrades suprathreshold depth percepts. *J. Vision* 14, 14. doi: 10.1167/14.9.14
- Deas, L. M., and Wilcox, L. M. (2015). Perceptual grouping via binocular disparity: the impact of stereoscopic good continuation. *J. Vision* 15, 11. doi: 10.1167/15.11.11
- Duits, R., Boscain, U., Rossi, F., and Sachkov, Y. (2013). Association fields via cusplless sub-riemannian geodesics in SE(2). *J. Math. Imaging Vision* 49, 384–417. doi: 10.1007/s10851-013-0475-y
- Duits, R., and Franken, E. (2011). Left-invariant diffusions on the space of positions and orientations and their application to crossing-preserving smoothing of hard images. *Int. J. Comput. Vision*. doi: 10.1007/s11263-010-0332-z
- Elder, J. H., and Goldberg, R. M. (2002). Ecological statistics of gestalt laws for the perceptual organization of contours. *J. Vision* 2, 5–5. doi: 10.1167/2.4.5
- Faugeras, O. (1993). *Three-Dimensional Computer Vision: A Geometric Viewpoint*. Cambridge, MA; London: MIT Press.
- Field, D. J., Hayes, A., and Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Res.* 33, 173–193. doi: 10.1016/0042-6989(93)90156-Q
- Fulvio, J. M., Singh, M., and Maloney, L. T. (2008). Precision and consistency of contour interpolation. *Vision Res.* 48, 831–849. doi: 10.1016/j.visres.2007.12.018
- Geisler, W. S., and Perry, J. S. (2009). Contour statistics in natural images: grouping across occlusions. *Visual Neurosci.* 26, 109–121. doi: 10.1017/S0952523808080875
- Geisler, W. S., Perry, J. S., Super, B. J., and Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Res.* 41, 711–724. doi: 10.1016/S0042-6989(00)00277-7
- Hess, R. F., and Field, D. J. (1995). Contour integration across depth. *Vision Res.* 35, 1699–1711. doi: 10.1016/0042-6989(94)00261-J
- Hess, R. F., Hayes, A., and Field, D. J. (2003). Contour integration and cortical processing. *J. Physiol.* 97, 105–119. doi: 10.1016/j.jphysparis.2003.09.013
- Hess, R. F., Hayes, A., and Kingdom, F. A. A. (1997). Integrating contours within and through depth. *Vision Res.* 37, 691–696. doi: 10.1016/S0042-6989(96)00215-5
- Hess, R. F., May, K. A., and Dumoulin, S. O. (2014). "Contour integration: psychophysical, neurophysiological, and computational perspectives," in *The Oxford Handbook of Perceptual Organization*, ed J. Wagemans (Oxford Academic). doi: 10.1093/oxfordhb/9780199686858.013.013
- Hoffman, W. C. (1989). The visual cortex is a contact bundle. *Appl. Math. Comput.* 32, 137–167. doi: 10.1016/0096-3003(89)90091-X
- Howard, I. P. (2012). *Perceiving in Depth, Volume 1: Basic Mechanisms*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780199764143.001.0001
- Howard, I. P., and Rogers, B. J. (1995). *Binocular Vision and Stereopsis*. New York, NY: Oxford University Press. doi: 10.1093/acprof:oso/9780195084764.001.0001
- Hubel, D. H., and Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J. Physiol.* 160, 106–154. doi: 10.1113/jphysiol.1962.sp006837
- Hubel, D. H., and Wiesel, T. N. (1970). Stereoscopic vision in macaque monkey: cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225, 41–42. doi: 10.1038/225041a0
- Intoy, J., Cox, M. A., Alicic, E., Victor, J. D., Banks, M. S., and Rucci, M. (2021). Fixational eye movements contribute to stereopsis. *J. Vision* 21, 2112–2122. doi: 10.1167/jov.21.9.2112
- Ivry, R., Beck, J., and Rosenfeld, A. (1989). Line segregation. *Spat. Vision* 4, 75–101. doi: 10.1163/156856889X00068
- Jaeger, D., and Ranu, J. (2015). *Encyclopedia of Computational Neuroscience*. New York, NY: Springer. doi: 10.1007/978-1-4614-6675-8
- Jaini, P., and Burge, J. (2017). Linking normative models of natural tasks to descriptive models of neural response. *J. Vision* 17, 16–16. doi: 10.1167/17.12.16
- Jones, D. G., and Malik, J. (1991). *Determining Three-Dimensional Shape from Orientation and Spatial Frequency Disparities I-Using Corresponding Line Elements*. Technical Report UCB/CSD-91-656, EECS Department, University of California, Berkeley, CA.
- Jones, J. P., and Palmer, L. A. (1987). An evaluation of the two-dimensional gabor filter model of simple receptive fields in cat striate cortex. *J. Neurophysiol.* 58, 1233–1258. doi: 10.1152/jn.1987.58.6.1233
- Julesz, B. (1971). *Foundations of Cyclopean Perception*. Chicago, IL: The University of Chicago Press.
- Kato, D., Baba, M., Sasaki, K. S., and Ohzawa, I. (2016). Effects of generalized pooling on binocular disparity selectivity of neurons in the early visual cortex. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150266. doi: 10.1098/rstb.2015.0266
- Kellman, P. J., Garrigan, P., and Shipley, T. F. (2005a). Object interpolation in three dimensions. *Psychol. Rev.* 112, 586–609. doi: 10.1037/0033-295X.112.3.586
- Kellman, P. J., Garrigan, P., Shipley, T. F., Yin, C., and Machado, L. (2005b). 3-D interpolation in object perception: evidence from an objective performance paradigm. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 558–583. doi: 10.1037/0096-1523.31.3.558
- Khuu, S. K., Honson, V., and Kim, J. (2016). The perception of three-dimensional contours and the effect of luminance polarity and color change on their detection. *J. Vision* 16, 31. doi: 10.1167/16.3.31
- Koenderink, J. J., and van Doorn, A. J. (1987). Representation of local geometry in the visual system. *Biol. Cybernet.* 55, 367–375. doi: 10.1007/BF00318371
- Koffka, K. (1963). *Principles of Gestalt Psychology*. New York, NY: A Harbinger Book.
- Kohler, W. (1967). Gestalt psychology. *Psychol. Forschung* 31, 18–30. doi: 10.1007/BF00422382
- Lawlor, M., and Zucker, S. W. (2013). "Third-order edge statistics: contour continuation, curvature, and cortical connections," in *Advances in Neural Information Processing Systems* 26.
- LeVay, S., Hubel, D. H., and Wiesel, T. N. (1975). The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. *J. Comp. Neurol.* 159, 559–575. doi: 10.1002/cne.901590408

- Li, G., and Zucker, S. W. (2003). "A differential geometrical model for contour-based stereo correspondence," in *Proc. of IEEE Workshop on Variational, Geometric and Level set Methods in Computer Vision* (Nice).
- Li, G., and Zucker, S. W. (2006). Contextual inference in contour-based stereo correspondence. *Int. J. Comput. Vision* 69, 59–75. doi: 10.1007/s11263-006-6853-9
- Li, G., and Zucker, S. W. (2008). Differential geometric inference in surface stereo. *IEEE Trans. Pattern Anal. Mach. Intell.* 32, 72–86. doi: 10.1109/TPAMI.2008.270
- Malach, R., Amir, Y., Harel, M., and Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and *in vivo* targeted biocytin injections in primate striate cortex. *Proc. Natl. Acad. Sci. U.S.A.* 90, 10469–10473. doi: 10.1073/pnas.90.22.10469
- Marr, D., and Poggio, T. (1979). A computational theory of human stereo vision. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* 204, 301–328. doi: 10.1098/rspb.1979.0029
- Menz, M. D., and Freeman, R. D. (2004). Functional connectivity of disparity-tuned neurons in the visual cortex. *J. Neurophysiol.* 91, 1794–1807. doi: 10.1152/jn.00574.2003
- Miolane, N., and Pennec, X. (2016). "A survey of mathematical structures for extending 2D neurogeometry to 3D image processing," in *Medical Computer Vision: Algorithms for Big Data: International Workshop, MCV 2015* (Munich: Springer), 155–167. doi: 10.1007/978-3-319-42016-5\_15
- Mitchison, G. J., and McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Res.* 30, 1781–1791. doi: 10.1016/0042-6989(90)90159-1
- Montobbio, N., Citti, G., and Sarti, A. (2019). From receptive profiles to a metric model of V1. *J. Comput. Neurosci.* 46, 257–277. doi: 10.1007/s10827-019-00716-6
- Neilson, P., Neilson, M., and Bye, R. (2018). A riemannian geometry theory of three-dimensional binocular visual perception. *Vision* 2, 43. doi: 10.3390/vision2040043
- Nelson, J., Kato, H., and Bishop, P. O. (1977). Discrimination of orientation and position disparities by binocularly activated neurons in cat striate cortex. *J. Neurophysiol.* 40, 260–283. doi: 10.1152/jn.1977.40.2.260
- Oluk, C., Bonnen, K., Burge, J., Cormack, L. K., and Geisler, W. S. (2022). Stereo slant discrimination of planar 3d surfaces: Frontoparallel versus planar matching. *J. Vision* 22, 6–6. doi: 10.1167/jov.22.5.6
- Parent, P., and Zucker, S. W. (1989). Trace inference, curvature consistency, and curve detection. *IEEE Trans. Pattern Anal. Mach. Intell.* 11, 823–839. doi: 10.1109/34.31445
- Parker, A. J., Smith, J. E. T., and Krug, K. (2016). Neural architectures for stereo vision. *Philos. Trans. R. Soc. B Biol. Sci.* 371, 20150261. doi: 10.1098/rstb.2015.0261
- Petitot, J. (2008). *Neurogéométrie de la vision: modèles mathématiques et physiques des architectures fonctionnelles*. Palaiseau: Editions Ecole Polytechnique.
- Petitot, J., and Tondut, Y. (1999). Vers une neurogéométrie. Fibrations corticales, structures de contact et contours subjectifs modaux. *Mathématiques et Sciences humaines* 145, 5–101. doi: 10.4000/msh.2809
- Poggio, G. F. (1995). Mechanisms of stereopsis in monkey visual cortex. *Cereb. Cortex* 5, 193–204. doi: 10.1093/cercor/5.3.193
- Read, J. C. (2015). "Stereo vision, models of," in *Encyclopedia of Computational Neuroscience*, eds D. Jaeger and R. Jung (New York, NY: Springer), 2873–2881. doi: 10.1007/978-1-4614-6675-8\_560
- Read, J. C., and Cumming, B. G. (2007). Sensors for impossible stimuli may solve the stereo correspondence problem. *Nat. Neurosci.* 10, 1322–1328. doi: 10.1038/nn1951
- Samonds, J. M., Potetz, B. R., Tyler, C. W., and Lee, T. S. (2013). Recurrent connectivity can account for the dynamics of disparity processing in V1. *J. Neurosci.* 33, 2934–2946. doi: 10.1523/JNEUROSCI.2952-12.2013
- Sanguinetti, G., Citti, G., and Sarti, A. (2010). A model of natural image edge co-occurrence in the roto-translation group. *J. Vision* 10, 37. doi: 10.1167/10.14.37
- Sarti, A., and Citti, G. (2015). The constitution of visual perceptual units in the functional architecture of V1. *J. Comput. Neurosci.* 38, 285–300. doi: 10.1007/s10827-014-0540-6
- Sarti, A., Citti, G., and Petitot, J. (2007). The symplectic structure of the primary visual cortex. *Biol. Cybernet.* 98, 33–48. doi: 10.1007/s00422-007-0194-9
- Sarti, A., Citti, G., and Piotrowski, D. (2019). Differential heterogenesis and the emergence of semiotic function. *Semiotica* 2019, 1–34. doi: 10.1515/sem-2018-0109
- Sasaki, K. S., Tabuchi, Y., and Ohzawa, I. (2010). Complex cells in the cat striate cortex have multiple disparity detectors in the three-dimensional binocular receptive fields. *J. Neurosci.* 30, 13826–13837. doi: 10.1523/JNEUROSCI.1135-10.2010
- Schmidt, K. E., Goebel, R., Löwel, S., and Singer, W. (1997). The perceptual grouping criterion of colinearity is reflected by anisotropies of connections in the primary visual cortex. *Eur. J. Neurosci.* 9, 1083–1089. doi: 10.1111/j.1460-9568.1997.tb01459.x
- Scholl, B., Tepohl, C., Ryan, M. A., Thomas, C. I., Kamasawa, N., and Fitzpatrick, D. (2022). A binocular synaptic network supports interocular response alignment in visual cortical neurons. *Neuron* 110, 1573–1584. doi: 10.1016/j.neuron.2022.01.023
- Schreiber, K. M., Hillis, J. M., Filippini, H. R., Schor, C. M., and Banks, M. S. (2008). The surface of the empirical horopter. *J. Vision* 8, 7–7. doi: 10.1167/8.3.7
- Singh, M., and Fulvio, J. M. (2005). Visual extrapolation of contour geometry. *Proc. Natl. Acad. Sci. U.S.A.* 102, 939–944. doi: 10.1073/pnas.0408444102
- Singh, M., and Fulvio, J. M. (2007). Bayesian contour extrapolation: geometric determinants of good continuation. *Vision Res.* 47, 783–798. doi: 10.1016/j.visres.2006.11.022
- Smits, J. T. S., and Vos, P. G. (1987). The perception of continuous curves in dot stimuli. *Perception* 16, 121–131. doi: 10.1068/p160121
- Ts'o, D. Y., Zarella, M., and Burkitt, G. (2009). Whither the hypercolumn? *J. Physiol.* 587, 2791–2805. doi: 10.1113/jphysiol.2009.171082
- Tu, L. W. (2011). *An Introduction to Manifolds*. New York, NY: Springer. doi: 10.1007/978-1-4419-7400-6
- Uttal, W. R. (1983). *Visual Form Detection in 3-Dimensional Space*. Hillsdale, NJ: L. Erlbaum Associates.
- Uttal, W. R. (2013). *Visual Form Detection in Three-Dimensional Space*. New Jersey, NJ: Psychology Press. doi: 10.4324/9780203781166
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., et al. (2012). A century of gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychol. Bull.* 138, 1172. doi: 10.1037/a0029333
- Warren, P. A., Maloney, L. T., and Landy, M. S. (2002). Interpolating sampled contours in 3-D: analyses of variability and bias. *Vision Res.* 42, 2431–2446. doi: 10.1016/S0042-6989(02)00266-3
- Zucker, S. (2006). "Differential geometry from the frenet point of view: boundary detection, stereo, texture and color," in *Handbook of Mathematical Models in Computer Vision* (Boston, MA: Springer), 357–373. doi: 10.1007/0-387-28831-7\_22
- Zucker, S. W. (2014). Stereo, shading, and surfaces: curvature constraints couple neural computations. *Proc. IEEE* 102, 812–829. doi: 10.1109/JPROC.2014.2314723
- Zucker, S. W., and Davis, S. (1988). Points and endpoints: a size/spacing constraint for dot grouping. *Perception* 17, 229–247. doi: 10.1068/p170229