Geometry and the visual brain

Jean Lorenceau *

UPR 640-LENA CNRS, 47 Bd de l’Hôpital, 75011 Paris, France

Abstract

Receptive fields structure of neurons in primary visual cortex suggests that they process visual stimuli in the frequency domain, in a way similar to the frequency analysis performed in the auditory system. As a consequence, both psychophysicists and electrophysiologists have long probed the visual system using extended sine wave gratings that are well localized in the frequency domain but poorly defined in visual space. Meanwhile, how the brain processes the geometrical properties and the spatial and temporal relationships between stimulus parts has received less attention. Recent progress in visual neuroscience that uncovered long-range horizontal connections between cortical neurons and revealed the complex architecture of primary visual cortex and feedback connectivity led to new insights concerned with the processing of geometrical properties of visual stimuli in V1. This paper presents a short historical perspective of the emergence of new issues related to the cortical architecture and its functional consequences on the processing of geometrical properties.

Keywords: Receptive field; Fourier analysis; Phase; Geometry; Long-range horizontal connections

1. Introduction

Early in the 20th century, Wertheimer [32] observed that two static flashes presented in succession with appropriate temporal and spatial separations elicit a perception of visual apparent motion. This seminal study, indicating that the visual system does more than simply registering external events, had influential theoretical consequences and launched the idea that the “whole is more than the sum of its parts”. From Wertheimer’s work emerged a scientific attempt to define general principles underlying perceptual organization, the gestalt psychology [19,20]. In vision, figure/ground segregation and perceptual grouping of individual “tokens” into a “whole” was found to strongly depend upon their spatial and temporal relationships, whatever the exact nature of these “tokens”. General principles, such as good continuity, proximity, closure, symmetry, common fate, synchronicity etc. appeared as strong determinants of perceptual organization. Importantly, the general model underlying the Gestaltist approach is a geometrical one, stressing the spatial relations between parts rather than concerned with the intrinsic processing of the parts themselves. However, the early attempt of the gestaltist school to offer a plausible neuronal account of their findings failed, as their idea of an isomorphism between the external Euclidean space and the internalization of geometrical rules whereby spatial relations between neurons would mimic the geometry of the stimulus was not confirmed experimentally.

In contrast with the Gestaltist approach, Fechner, at the end of the 19th century, aimed at establishing the quantitative relationships between physical stimulus inputs and perceptual outputs of sensory systems, in order to define and measure the elementary sensations processed by the human brain. Within this elementarist framework, less emphasis is put on perceptual organization and grouping of “tokens” and more on how the central nervous system processes the “tokens” themselves, measuring detection thresholds to probe the transfer function of sensory systems. The relationships between the physical inputs and perception thresholds were found to follow simple, although disputed, mathematical equations, among which the Bouguer–Weber’s law on discrimination thresholds or the Fechner’s law according to which sensation grows as a function of the logarithm of the physical input intensity. Incidentally, at about the same time, Ramon and Cajal discovered that neurons are independent autonomous brain elements,
invaldating the dominant reticularist view according to which neurons form a continuous single brain circuit. This finding, fitting well with the elementarist approach of Fechner, suggested a piecewise functioning of the brain whereby each neuron would independently perform some specific computation of its inputs.

Advance in electrophysiology soon allowed Hartline [16] to record the responses of frog’s optic nerve in response to visual stimulation. His discovery that retinal ganglion cells are selectively activated by localized distributions of luminance or chromatic contrast gave birth to the notion of “receptive field” defined as the small region of visual space that need be illuminated to elicit a response from the recorded cell [22]. Subsequent studies showed that this notion extends to primary visual cortex, where neurons have elongated receptive fields and are retinotopically organized, such that neighboring neurons analyze neighboring regions of the visual field [17]. These pioneering studies yielded the view that the primary visual cortex consists in a mosaic of independent filters acting in parallel. This advance in the understanding of the electrophysiological properties of neuronal processing led psychophysicists to look for the potential perceptual consequences of these new findings, as the spatial structure of receptive fields provided strong experimental constraints to define a model of the stimulus relevant to biological vision. Indeed, the center-surround spatial structure of retinal receptive fields implies that each ganglion cell is band pass in the frequency domain. As a matter of fact, Fourier’s theorem states that any image can be decomposed into a sum of sinusoidal waves of different frequency, amplitude and phase. This raised the question of whether the cortex analyses images in visual space or in its equivalent frequency space. As the Fourier analysis is a linear operation, determining whether the visual system itself behaves as a linear spatial frequency analyzer appeared of great interest as it could provide a powerful tool to study biological vision. In the frequency domain, the most elementary stimulus, represented by a single point, is a sinusoidal distribution of luminance contrast, consisting of a simple extended oriented grating. It was therefore relevant that, in their seminal study, Campbell and Robson [6] used oriented one-dimensional gratings to probe the relations between the center-surround structure of neurons’ receptive field and human contrast sensitivity. Their results brought support to the view that the visual system decomposes visual inputs into spatial frequency bands, each band being selectively processed by a sub-population of neurons composing a “spatial frequency channel”. Consistent with this view, electrophysiological recordings indicated that cortical visual neurons are selectively tuned to different spatial frequencies [9,10,23]. The idea that the visual system decomposes a spatial distribution of light into different spatial frequency bands selectively processed by a set of spatial frequency analyzers working independently and in parallel at different scales had a huge impact on subsequent research. Based on this filtering model of visual processing, studying the visual system required that the spatial frequency content of the stimuli was precisely controlled and specified.

Accordingly, psychophysicists and electrophysiologists used gratings at different spatial frequency, orientation and contrast to probe early vision, as they were mostly concerned with the effects of the energy spectrum on contrast sensitivity and neuronal selectivity. Numerous psychophysical studies used a variety of methods to assess the number and bandwidth of spatial frequency channels underlying visual perception, the degree of overlap and independence between channels, etc. [2,3,33]. Subsequent studies of motion or texture perception drew on what soon became the standard model to offer different schemes whereby the outputs of “filter neurons” were combined to yield units endowed with more complex capabilities [1,31].

More recently, the development of multiscale analysis and wavelet transform that use Gabor filters that are well localized both in space and in the frequency domain provided new tools to describe the morphological properties of images while modeling more accurately the response profile of cortical cells [8]. However, the extraction of the geometrical properties of images with the wavelet transform remains a difficult problem.

As mentioned above, the Fourier decomposition of a two-dimensional (2D) image result in both an energy spectrum, describing the distribution of amplitude in different spatial frequency bands at different orientations, and a phase spectrum, that contains information about the position of different spatial frequencies in visual space (Fig. 1a). One issue with this representation of images in the Fourier domain is that the position and geometrical relationships between different parts of an image, although they are somehow embedded in the phase spectrum, are difficult to visualize and analyze. This is mainly because the phase of each spatial frequency component is expressed relative to an arbitrary origin (absolute phase), but does not directly represent the phase between different spatial frequencies (relative phase) that gives information about their spatial relationships. This latter information is highly relevant to image processing as it characterizes both the different features in an image, such as edges and borders and the relative positions between different distant parts of the image, such as contour alignments, curvature, etc.

As a matter of fact, psychophysical studies showed that human observers rely heavily on the phase spectrum to recognize objects and scenes, and to a lesser extent on the energy spectrum. For instance, when blending, through image synthesis, an energy spectrum of an image A with the phase spectrum of an image B, the resulting image B’ is more easily recognized than
These experiments are strong evidence that phase information is more important for object recognition and the analysis of natural images than the amplitude spectrum, consistent with the observation that, although the amplitude spectrum may greatly vary, perceptual organization remains stable. Electrophysiological recordings in primary visual cortex showed that simple cells are sensitive to the phase of a grating with optimal spatial frequency and orientation as their responses depend on the position of a grating within their receptive field. However, complex cells proved to be phase insensitive, providing the same response independently of the position of an optimally oriented sine wave grating within their receptive field. Since these cells provide the main output for subsequent processing in other extracortical areas, this raises the question of whether the sensitivity of simple cells to spatial phase is used to process images [28].

At this point, the notion of phase, as represented in the Fourier domain, need to be clarified. The phases of different spatial frequencies may correspond to the distribution of luminance intensity in a single location of visual space and be processed by neurons with overlapping receptive fields; alternately, phase information may indicate that different spatial frequencies are present in different locations of the visual space and stimulate
distinct neuronal populations with non-overlapping receptive fields [24]. Although this distinction may be irrelevant in the standard linear filtering model that assumes that filters at an early stage work independently and in parallel—the integration of their outputs being left for neurons at a “higher” stage—it is not irrelevant if neurons already interact within primary visual cortex. In this latter case, the architecture of V1 and the spatial distribution of neurons within the cortical tissue—the intrinsic geometry of the visual brain—becomes of primarily importance.

As a matter of fact, the classical view that the distribution of neurons in primary visual cortex is purely retinotopic has recently been revised, as it was found the cortical surface presents fractures and singularities, called pinwheels, where neurons change rapidly their orientation selectivity as well as their positions in visual space [4]. These findings have important functional consequences in the light of the fact that the notion that neurons are immune to stimulation outside the spatially restricted region—the receptive field—has also been challenged by psychophysical and electrophysiological studies showing that the responses of V1 neurons to oriented stimuli presented within their receptive field can be markedly modulated by stimuli falling in surrounding regions, which by themselves fail to activate the cell. These influences, that can be mediated by feedback projections from higher cortical areas as well as by long-range horizontal connections within V1, call to a revision of the mere notion of receptive field (see Angelucci et al., Hayes and Field this issue). Indeed, anatomical studies indicate that horizontal connections link cortical regions over distances up to 6–8 mm, tend to connect cells with similar orientation preferences and more specifically cells whose receptive fields are topographically aligned along an axis of collinearity [12,14,30]. The non-trivial architecture of divergent feedback connections and experimental evidence also indicates that feedback exerts long-range modulation of neuronal activity at a lower processing level. Thus, this circuitry—feedback from other areas and horizontal connections within a single area—together with the highly complicated architecture of primary visual cortex provide a possible physiological substratum to compute various geometrical properties of the incoming image.

In support of a functional link between neurons through horizontal connections in primary visual cortex, a number of psychophysical studies uncovered strong contrast dependent center-surround interactions, either facilitatory or suppressive [7,26,27], that occur when one or several oriented test stimuli are presented together with surrounding oriented stimuli. For instance, contrast sensitivity is improved by similar flankers collinear and aligned with the test stimulus. Changing the relative distance, orientation, spatial frequency or contrast of the flankers modulates the change in sensitivity, allowing the analysis of the architecture of these spatial interactions [26,27]. In addition, the ability to detect the presence of a specific configuration of oriented bars immersed within a surrounding textures of randomly oriented elements with similar characteristics is better for configurations of collinear and aligned elements than for parallel configurations. Field et al. [11] proposed that perceptual “association fields” are involved in this contour integration process, and suggested that the architecture of horizontal long-range connections may underlie these effects. It is worth noting that the standard filtering model is unable to account for these findings (see Hess et al., this issue). Furthermore, the notion of association field is supported by studies showing that these long-range interactions are decreased or suppressed in amblyopic patients who suffer from a disorganization of horizontal connectivity. Electro-physiological recordings revealed that similar center-surround interactions are already present at the cellular level in primary visual cortex [18,21]. The additional finding that the activity elicited by a visual stimulus propagates slowly along horizontal connections [5,15] has important functional, perceptual and theoretical consequences [13,29], as the functional architecture of V1 cannot be viewed as a pure spatial circuitry but must incorporate the temporal dimension.

Overall, these studies are compatible with the view that long-range connections play a functional role in perceptual contour integration, and further suggest that they may constitute a physiological substrate that implement some of the gestalt rules at an early processing stage. In this respect, using Fourier analysis, and its amplitude spectrum in particular, as a tool and a model to specify the stimulus characteristics may not be fully adapted to embrace the capability of the striate cortex to process geometrical aspects of images, that carry the information relevant to figure/ground segregation, visual binding, motion analysis or object recognition. The articles of the present issue explore in details geometrical models of the functional architecture of primary visual cortex and its consequences on the processing of geometrical characteristics of images. They should be of great interest for researchers aiming at understanding biological vision.

References


