

Geometric Analysis of the Conformal Camera for Intermediate-Level Vision and Perisaccadic Perception

Jacek Turski

University of Houston-Downtown

Department of Computer and Mathematical Sciences

March 2009

Abstract

A binocular system developed by the author in terms of projective Fourier transform (PFT) of the conformal camera, which numerically integrates the head, eyes, and visual cortex, is used to process visual information during saccadic eye movements. Although we make three saccades per second at the eyeball's maximum speed of 700 deg/sec, our visual system accounts for these incisive eye movements to produce a stable percept of the world. This visual constancy is maintained by neuronal receptive field shifts in various retinotopically organized cortical areas prior to saccade onset, giving the brain access to visual information from the saccade's target before the eyes' arrival. It integrates visual information acquisition across saccades. Our modeling utilizes basic properties of PFT. First, PFT is computable by FFT in complex logarithmic coordinates that approximate the retinotopy. Second, a translation in retinotopic (logarithmic) coordinates, modeled by the shift property of the Fourier transform, remaps the presaccadic scene into a postsaccadic reference frame. It also accounts for the perisaccadic mislocalization observed by human subjects in laboratory experiments. Because our modeling involves cross-disciplinary areas of conformal geometry, abstract and computational harmonic analysis, computational vision, and visual neuroscience, we include the corresponding background material and elucidate how these different areas interweave in our modeling of primate perception. In particular, we present the physiological and behavioral facts underlying the neural processes related to our modeling. We also emphasize the conformal camera's geometry and discuss how it is uniquely useful in the intermediate-level vision computational aspects of natural scene understanding.

Keywords: The conformal camera, projective Fourier transform, complex projective geometry, intermediate-level vision, retinotopy, binocular vision, saccades, efference copy, predictive remapping, perisaccadic mislocalization

1 Introduction

In the last few years, we have developed projective Fourier analysis for computational vision in the framework of the representation theory of the semisimple Lie group $\mathbf{SL}(2, \mathbb{C})$ [58, 59, 60, 61, 62]. It was done by restricting the group representations to the image plane of the conformal camera—the camera with image projective transformations given by the action of $\mathbf{SL}(2, \mathbb{C})$. This analysis provides an efficient image representation and processing that are not only well adapted to the projective transformations of retinal images, but are also to the retinotopic mappings of the brain’s oculomotor and visual pathways. This latter assertion stems from the fact that the projective Fourier transform (PFT) is computable by a fast Fourier transform algorithm (FFT) in coordinates given by a complex logarithm that transforms PFT into the standard Fourier integral and at the same time approximates the retinotopic mappings [54].

However, the conformal camera is somewhat abstract and noticeably different than any other camera model used in computer vision. Nevertheless, its remarkable advantages are revealed to us every time we model specific physiological processes involved in visual perception. For instance, one could reasonably expect that a stationary camera and a moving object is similar to a moving camera and a stationary because the relative position of the camera and the object could be the same in both cases. Remarkably, it fails in primate vision systems. In fact, when the image of a fast-moving object sweeps across a static retina, though we are normally aware of its motion, we fail to detect the comparable motion of images as they sweep across the retina during fast eye movements. Computational modeling presented in this article demonstrates that the conformal camera naturally supports this asymmetry.

Recently, building on projective Fourier analysis of the conformal camera, a mathematical model integrating the head, eyes, and visual cortex into a single computational binocular system was introduced in [63] with particular focus on stereopsis. Here it is demonstrated that this integrated system may efficiently process visual information during fast scanning eye movements called saccades, employed to build up understanding of a scene despite the highest acuity only present in the central foveal region of a 2 deg visual angle. We make about three saccades per second at the eyeball’s maximum speed of 700 deg/sec. Visual sensitivity is markedly reduced during saccades as we do not see moving retinal images. These fragmented pieces of visual information are sent to the cortical areas, with a minor part going to subcortical areas where they are integrated into a stable coherent percept of a 3D world despite of the persistence of incisive eyes movements. This constancy of vision is maintained by a widespread neural network with multiple mechanisms receiving inputs from several sources. Not surprisingly, in spite of a significant recent progress, how this problem is solved by the brain has been the topic of many theories, see [69] for a recent review.

The modeling presented in this article, first proposed in [64], utilizes basic properties of PFT to capture some of the very first computational aspects of the neural processes during the saccadic eye movements. First, because the PFT of an image can be efficiently computed by FFT in complex logarithmic coor-

dinates that also approximate the retinotopy, the output from the inverse PFT resembles the cortical representation of the image. Second, a simple translation in retinotopic (logarithmic) coordinates that is efficiently modeled here by the standard shift property of the inverse PFT when expressed in these coordinates, remaps the presaccadic scene in the reference frame centered on the fovea into a postsaccadic reference frame centered on the impending saccade target. Equivalently, it uniformly shifts images around the target in cortical periphery to the cortical foveal location. Moreover, this shift that takes place in retinotopic (logarithmic) coordinates accounts for perceptual space compression seen around the time of saccadic eye movements by human subjects in psychophysical laboratory experiments [34, 53].

The idea of remapping is supported by the fact that the neural correlates of a copy of the oculomotor command to move eyes, known as efference copy or corollary discharge [24], have been found in the form of a neuronal receptive field shift about 50 ms before a saccade onset in various retinotopically organized visual cortical areas [18, 37]. This shift points to the possibility that prior to the eyes arriving at the target, the brain has access to visual information from that peripheral region. In fact, in the recent experiment [31], when human subjects shifted fixation to the clock, their reported time was earlier than the actual time on the clock by about 40 ms. It may integrate visual information from an object across saccades, and therefore, eliminate the need for starting visual information processing anew three times per second at each fixation and speed up a costly process of visual information acquisition [66]. It may also build up perceptual continuity across fixations [45].

The conformal camera was initially constructed for the purpose of developing projectively adapted image representation in the framework of the only well understood ‘projective’ Fourier analysis formulated as a direction in the representation theory of semisimple Lie groups, a great achievement of the 20th-century mathematics [29]. In the case of the conformal camera, it is the representation theory of the group $\mathbf{SL}(2, \mathbb{C})$, the group generating image projective transformations in a conformal geometry setting; see [61] where a brief introduction to the group representations is also given. When writing this article, it became apparent that we should carefully set a stage for our modeling that involves conformal projective geometry, abstract and computational harmonic analysis, image processing, and computational vision including visual neuroscience and machine vision. Thus, the overarching aim of this article is to elucidate how these cross-disciplinary areas interweave in our modeling of primate perception.

To this end, the paper is organized as follows. In the next section, we include, in some detail, physiological and behavioral facts that underlie the neural processes of human vision related to our computational modeling. In the following three sections, we lay down the background that explains the mathematical tools we use in modeling human vision processes. In Section 3, we introduce the conformal camera and discuss the image projective transformations. We end this section with the construction of the group of image projective transformations in the conformal camera. In Section 4, we review the geometry underlying the conformal camera and demonstrate that the fundamental properties of this

geometry should be uniquely useful in the early- and intermediate-level vision computational aspects of natural scene understanding. In the last of these three sections, Section 5, we show that the conformal camera possesses its own harmonic analysis—projective Fourier analysis—which gives efficient image representation well adapted to both the retinal image projective transformations and the retinotopy of the brain’s visual pathways. Finally, in Section 6, we discuss some implementation issues when working with the discrete PFT. In particular, the binocular system with head, eyes, and visual cortex numerically integrated by PFT is discussed. Further, using this integrated binocular system, we model the perisaccadic perception, including the perisaccadic mislocalizations observed in psychophysical laboratory experiments. This perisaccadic mislocalization, in the form of perceptual space compression around the saccade target, is simulated in the model by the standard shift property of Fourier transform. Also, the future direction in advancing our modeling and its implementation are discussed. The paper is summarized in the last section.

The research program presented here advances our mathematical modeling intended for computational vision, including visual neuroscience and machine vision systems. It is guided by a strategy important in the contemporary neurocomputing research: linking known anatomical and physiological details with efficient computational modeling and engineering designs should be vital not only to the emerging field of neural engineering but also to interpreting relevant neurophysiological data.

2 Visual Neuroscience Background

2.1 Visual Perception is a Creative Process

When light reflected from objects in the 3D world is impinged upon the retina, it activates the neuronal pathways, beginning with phototransduction by about 125 million photoreceptors. Next, the visual information passes through a multi-layered circuitry of the retina where substantial processing takes place.

The only recently emerging picture [20] of the retinal processing tells us that more than a dozen of distinct visual recordings of the retinal image are extracted. For example, one recording emphasizes the boundaries between objects while another carries information about movement in specific directions. The result is that more than a dozen of the most essential features of the original retinal image are extracted in parallel and sent to the brain as a train of spikes along about 1.5 million axons of ganglion cells to more than 30 association cortex areas containing about 30 billion neurons where the details: depth, texture, color, form, motion, etc., are added and integrated into a coherent view of the 3D world. This integration is entirely dependent upon visual experience; almost all higher order features of vision are influenced by expectations based on past experience. Although such influences occasionally allow the brain to be fooled into misperception, as is the case with the optical illusion in Fig.1, they also give us the ability to see and respond to the visual world quickly.

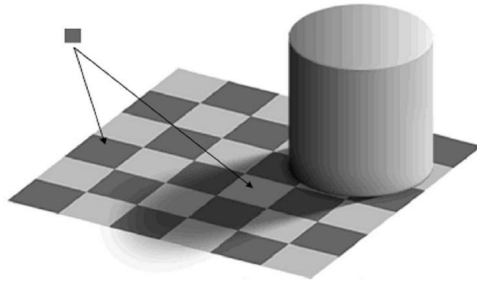


Figure 1. This illusion created by Adelson illustrates how perception may reflect the complex properties of the environment.

We see from this very brief description that visual perception is a creative process and, for this one reason alone, its quantitative modeling must be extremely difficult. Therefore, we try to develop a model that captures only some of the very first computational aspects of visual perception that takes place the first seconds following the opening of our eyes in daylight. Even with this limited goal, we find that those aspects are controlled by extremely sophisticated neural processes that involve nearly every level of the brain.

2.2 Early Visual Pathways

When humans open eyes in daylight and direct their gazes to attend a scene, they only see with the highest clarity, the central part of about a visual angle of 2 deg. This region is projected onto the central fovea where its image is sampled by the hexagonal mosaic of photoreceptors consisting of mainly cone cells that are color-selective type of photoreceptors for a sharp daylight vision. The visual acuity decreases rapidly away from the fovea because the distance between cones increases with eccentricity as they are outnumbered by rod cells, photoreceptors for a low acuity black-and-white night vision. Moreover, there is a gradual loss of hexagonal regularity of the photoreceptor mosaic. For example, at 2.5 deg radius, which corresponds to the most visually useful region of the retina, acuity drops 50%.

The distribution of axons in the optic nerve, which carries the retinal processing output to the brain, is precisely organized, but varies along the visual pathways. One aspect of this organization, or the retinotopy, is that axons corresponding to neighboring places in the retina are positioned closely in the nerve bundle, with notable exception along the vertical meridian. This exception stems from the fact that the output of each eye splits along the retinal vertical meridian when the axons originating from the nasal half of the retina cross at the optic chiasm to the contralateral brain's hemisphere and join the temporal half, which remains on the same side of its eye-of-origin. This splitting and crossing re-organizes the retina outputs so that the left hemisphere destinations receive information from the right visual field, and the right hemisphere destinations

receive information from the left visual field. According to the split theory [39, 43], which provides a greater understanding of vision cognitive processes than the bilateral theory of overlapping projections, there is a sharp foveal split along the vertical meridian of hemispherical cortical projections. Although it is crucial for synthesizing 3D representation from the binocular disparities in the pair of 2D retinal images, it presents a challenge in modeling retino-cortical image processing across visual hemifields.

2.3 Beyond Early Visual Pathways: Visuo-Saccadic Perception

One of the most important functions of any nervous system is sensing the external environment and responding in a way that maximizes immediate survival chances. For this reason, the perception and action have evolved in mammals by supporting each other's functions. This functional link between visual perception and oculomotor action is well demonstrated in primates when they execute the eye-scanning movements (saccades) in order to overcome the eye's acuity limitation in building up the scene understanding (see Fig. 2).

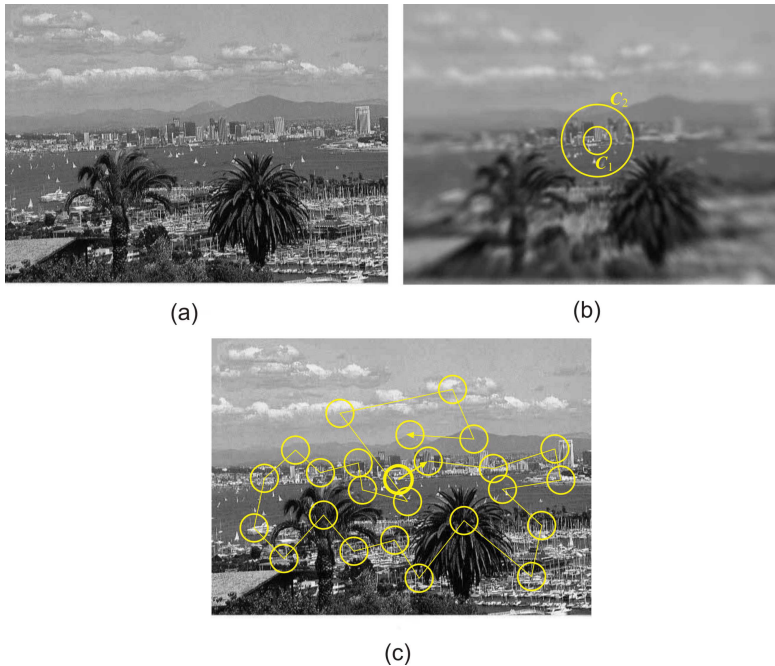


Figure 2. (a) San Diego skyline and harbor. (b) Progressively blurred image from (a) simulating the progressive loss of retinal acuity with eccentricity. The circle C_1 encloses the part of the scene projected onto the high acuity fovea of a 2 deg diameter. The circle C_2 encloses the part projected onto the visually useful faveal region of a 5 deg diameter. (c) A scanning path the eyes may take to build the scene understanding. Adapted from [2].

The saccadic eye movement is the most common bodily movement since we make about three saccades per second at the eyeball's maximum speed of 700 deg/sec. The eyes remain relatively still (while undergoing tremors, drifts and microsaccades—a miniature, random eye movement important for proper functioning of eyes [44]) between consecutive saccades for about 180-320 ms, depending on the task performed. During this time period, the image is processed by the retinal circuitry and sent mainly to the visual cortex (starting with the primary visual cortex, or V1, and reaching higher cortical areas, including cognitive areas) with a minor part going to oculomotor midbrain areas.

The sequence of saccades, fixations, and, often, also smooth-pursuit eye movements for tracking a slowly moving small object in the scene, is called the scanpath, first studied in [71]. In Fig. 2, (b) shows a progressively blurred image from (a), simulating the progressive loss of acuity with eccentricity. In Fig. 2 (c) we depict the scanpath that eyes might actually take to build up understanding of the scene.

Although they are the simplest of bodily movements, the eyes' saccades are controlled by widespread neural network that involves nearly every level of the brain. Most prominently, it includes the superior colliculus (SC) of the midbrain for representing possible saccade targets, the parietal eye field (PEF) and frontal eye field (FEF) in the parietal and frontal lobes of the neocortex (which obtain inputs from many visual cortical areas) for assisting the SC in the control of the involuntary (PEF) and voluntary (FEF) saccades. They also project to the simple neural circuits in the brainstem reticular formation in the midbrain that ensure the saccade's outstanding speed and precision.

Remarkably, many of the neural processes involved in saccade generation and control are amenable to precise quantitative studies such that even questions regarding the operation of the whole structure can be addressed by building on the existing models [21]. This not only carries immense clinical significance [15], but also forms an essential preliminary stage in building our understanding of human vision, the knowledge that will eventually be transferred to the emerging field of neural engineering.

Nevertheless, some neural processes of the visuo-saccadic system remain virtually unknown. Visual sensitivity is markedly reduced during saccadic movements as we do not see moving images on the retinas. This barely understood neural process is known as saccadic suppression. There is accumulating evidence that viewers integrate information poorly across fixations during tasks such as reading, visual search, and scene perception [50]. It means that, three times per second, there are instant large changes in the retinal images without almost any information consciously carried between images. Furthermore, because the next saccade target selection for the voluntary saccades takes place in the higher cortical areas involving cognitive processes [22], the time needed for the oculomotor system to plan and execute the saccadic eye movement could take as long as 150 ms. Therefore, it is critical that visual information is efficiently acquired during each fixation period of about 300 ms without repeating much of the whole process at each fixation since it would require too much

computational resources. However, visual constancy, the fact that we are not aware of any discontinuity in the scene perception when executing the scanpath, is not perfect. About 50 ms before the onset of the saccade, during saccadic movement (~ 30 ms) and about 50 ms after the saccade, perceptual space is transiently compressed around the saccade target [34, 53], a phenomenon called perisaccadic mislocalization. We continue this discussion in Section 6.5 where we present our modeling of the perisaccadic perception based on projective Fourier transform of the conformal camera.

3 The Conformal Camera

We model the human eyes' imaging functions with the conformal camera, the name of which will be explained later. The camera has many remarkable properties, the first following directly from its construction: the group of image projective transformations in the conformal camera is generated internally and has the 'minimal' property as explained in Fig. 3. In the remaining pages of this article, the other properties will be carefully examined in their relation to many computational aspects of visual perception.

In the conformal camera, the retina is represented by the image plane $x_2 = 1$ with complex coordinates $x_3 + ix_1$, on which a 3D scene is projected under the mapping

$$j(x_1, x_2, x_3) = (x_3 + ix_1) / x_2. \quad (1)$$

The implicit assumption $x_2 \neq 0$ will be removed later. Next, we give the precise form of the 'k' and 'h' image transformations introduced in Fig. 3.

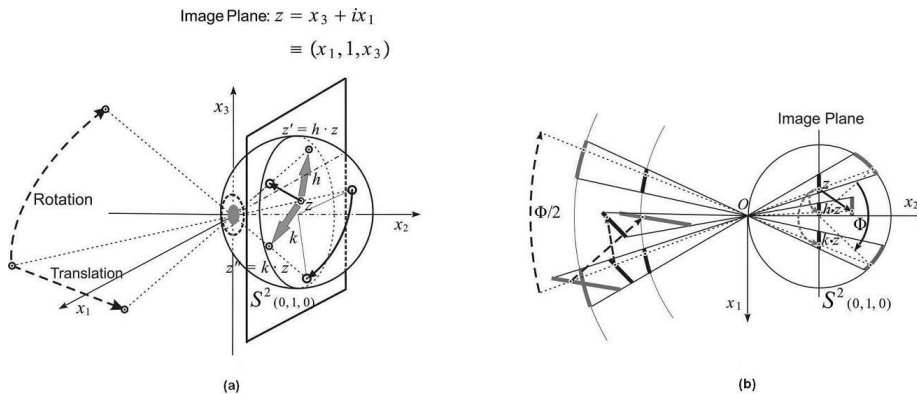


Figure 3. (a) Image projective transformations are generated by iterations of transformations covering translations 'h' and rotations 'k' of **planar** objects in the scene. (b) The 2D section of the conformal camera further explains how image projective transformations are generated and how the projective degrees of freedom are reduced in the camera; one image projective transformation in the conformal camera corresponds to different planar objects translations and rotations in the 3D world.

3.1 Basic Image Transformations

The image projective transformations in the conformal camera are generated by the following two transformations: (1) an image is projected by $(j|_{\mathbf{S}_{(0,1,0)}^2})^{-1}$ into the unit sphere $S_{(0,1,0)}^2$ centered at $(0, 1, 0)$, then the sphere is rotated and the (rotated) image is projected by j back to the image plane, (2) the image is translated out of the image plane then projected by j back to the image plane. The (1) and (2) transformations result in the ‘ k ’ and ‘ h ’ mappings in Fig. 3, respectively. They are explicitly given as follows:

1. **k transformations:** $\mathbf{SU}(2) = \left\{ \begin{pmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{pmatrix} \right\}$ is the maximal compact subgroup in $\mathbf{SL}(2, \mathbb{C})$, the group of 2×2 complex matrices of determinant 1. We let the group $\mathbf{SO}(3)$ of three dimensional rotations act on the sphere $S_{(0,1,0)}^2$ by rotating it about $(0, 1, 0)$. Furthermore, we parametrize $\mathbf{SO}(3)$ by the Euler angles (ψ, ϕ, ψ') , where ψ is the rotation about the x_2 -axis, followed by the rotation ϕ about the y_3 -axis, which is parallel to the x_3 -axis and passes through $(0, 1, 0)$, and finally by the rotation ψ' about the rotated x_2 -axis. Then, to each $R(\psi, \phi, \psi')$ in $\mathbf{SO}(3)$ there correspond two elements in $\mathbf{SU}(2)$,

$$k(\psi, \phi, \psi') = \pm \begin{pmatrix} e^{i(\psi+\psi')/2} \cos \frac{\phi}{2} & ie^{i(\psi-\psi')/2} \sin \frac{\phi}{2} \\ ie^{-i(\psi-\psi')/2} \sin \frac{\phi}{2} & e^{-i(\psi+\psi')/2} \cos \frac{\phi}{2} \end{pmatrix}, \quad (2)$$

such that $j \circ R(\psi, \phi, \psi') \circ (j|_{\mathbf{S}_{(0,1,0)}^2})^{-1}(z) = k \cdot z$ are given by the following linear fractional mappings

$$k(\psi, \phi, \psi') \cdot z = \frac{(e^{-i(\psi+\psi')/2} \cos \frac{\phi}{2}) z + ie^{i(\psi-\psi')/2} \sin \frac{\phi}{2}}{(ie^{i(\psi-\psi')/2} \sin \frac{\phi}{2}) z + e^{i(\psi+\psi')/2} \cos \frac{\phi}{2}}. \quad (3)$$

2. **h transformations:** Similarly, for each translation vector $\vec{b} = (b_1, b_2, b_3)$ where $b_2 \neq -1$ acting on the image plane $T_{\vec{b}}(x) = x + \vec{b}$, there are two elements $\mathbf{SL}(2, \mathbb{C})$,

$$h(b_1, b_2, b_3) = \pm \begin{pmatrix} (1+b_2)^{1/2} & 0 \\ (b_3+ib_1)(1+b_2)^{-1/2} & (1+b_2)^{-1/2} \end{pmatrix} \quad (4)$$

such that $j \circ T_{\vec{b}} \circ (j|_{x_2=1})^{-1}(z) = h \cdot z$ are given by the corresponding linear fractional mappings by the same action as before,

$$h(b_1, b_2, b_3) \cdot z = \frac{(1+b_2)^{-1/2} z + (b_3+ib_1)(1+b_2)^{-1/2}}{(1+b_2)^{1/2}}. \quad (5)$$

Now, if $f(z)$ is an image intensity function and g is either k or h mapping, the corresponding image transformation is the following: $f(g^{-1} \cdot z)$.

Both $k \cdot z$ and $h \cdot z$ mappings have forms of special linear-fractional transformations

$$g \cdot z = \frac{\alpha z + \beta}{\gamma z + \delta}; \quad \alpha\delta - \gamma\beta = 1.$$

These mappings are conformal, that is, they preserve the oriented angles of two tangent vectors $z'_k(t_0)$ to any two curves $z_k(t)$ ($k = 1, 2$) intersecting at the point $q = z(t_0)$. In fact,

$$\frac{d}{dt} \left(\frac{\alpha z_k(t) + \beta}{\gamma z_k(t) + \delta} \right)_{t=t_0} = \frac{z'_k(t_0)}{(\gamma q + \delta)^2} = \frac{e^{i\chi(q)} z'_k(t_0)}{|\gamma q + \delta|^2}, \quad k = 1, 2, \quad (6)$$

and both vectors $z'_1(t_0)$ and $z'_2(t_0)$ are rotated by the same angle $\chi(q)$.

3.2 The Group of Image Projective Transformations

3.2.1 The $\mathbf{PSL}(2, \mathbb{C})$ Group

The group of image transformations in the conformal camera is generated by all finite iterations of k and h mappings. To derive this group, we recall that $k \in \mathbf{SU}(2)$ and note that $h \in \mathbf{AN} \subset \mathbf{SL}(2, \mathbb{C})$ if $1 + b_2 > 0$ and $h = \varepsilon \mathbf{AN} \subset \mathbf{SL}(2, \mathbb{C})$ if $1 + b_2 < 0$, where

$$\mathbf{A} = \left\{ \begin{pmatrix} \rho & 0 \\ 0 & \rho^{-1} \end{pmatrix} \right\}, \quad \mathbf{N} = \left\{ \begin{pmatrix} 1 & 0 \\ \xi & 1 \end{pmatrix} \right\}, \quad \varepsilon = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix}. \quad (7)$$

Now, it follows from the polar decomposition $\mathbf{SL}(2, \mathbb{C}) = \mathbf{SU}(2)\mathbf{ASU}(2)$, that all these finite iterations result in the group $\mathbf{SL}(2, \mathbb{C})$ acting by linear-fractional mappings

$$\mathbf{SL}(2, \mathbb{C}) \ni \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot z = \frac{dz + c}{bz + a}; \quad z = x_3 + ix_1 \equiv (x_1, 1, x_3). \quad (8)$$

Because $\pm \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ have the same action, we need to identify matrices in $\mathbf{SL}(2, \mathbb{C})$ that differ in sign. The result is the quotient group $\mathbf{PSL}(2, \mathbb{C}) = \mathbf{SL}(2, \mathbb{C})/\{\pm Id\}$, where Id is the identity matrix, and the action (8) establishes a group isomorphism between linear-fractional mappings and $\mathbf{PSL}(2, \mathbb{C})$. Thus,

$$\mathbf{PSL}(2, \mathbb{C}) \ni g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto f(g^{-1} \cdot z) = f\left(\frac{az - c}{-bz + d}\right) \quad (9)$$

gives the image projective transformations of the intensity function $f(z)$.

3.2.2 Conformality

As we showed in (6), the mappings in (8) are conformal. Because of this property, the camera is called ‘conformal’. Although, the conformal part of an image projective transformation can be removed with almost no computational cost,

leaving only a perspective transformation of the image (see [61, 62]); the conformality provides an advantage in imaging because the conformal mappings rotate and dilate the image infinitesimal neighborhoods, and, therefore, locally preserve the image ‘pixels’.

To complete the description of the conformal camera, we need to address some implicit assumptions, such as the restriction $-bz + d \neq 0$ in (9) we have frequently made in this section.

4 Geometry of the Conformal Camera

In the homogeneous coordinate framework of projective geometry [6], the conformal camera is embedded into the complex plane

$$\mathbb{C}^2 = \left\{ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \mid z_1 = x_2 + iy, z_2 = x_3 + ix_1 \right\}.$$

In this embedding, the ‘slopes’ ξ of the complex lines $z_2 = \xi z_1$ are numerically identified with the points on the extended image plane $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ where ∞ corresponds to the line $z_1 = 0$. We note that if $x_2 \neq 0$ and $y = 0$, the slope ξ corresponds to the point $x_3 + ix_1$ at which the ray (line) in \mathbb{R}^3 that passes through the origin is intersecting the image plane of the conformal camera.

Now, the standard action of the group $\mathbf{SL}(2, \mathbb{C})$ on nonzero column vectors \mathbb{C}^2 ,

$$\begin{pmatrix} z'_1 \\ z'_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} az_1 + bz_2 \\ cz_1 + dz_2 \end{pmatrix}$$

implies that the slope $\xi = \frac{z_2}{z_1}$ is mapped to the slope

$$\xi' = \frac{z'_2}{z'_1} = \frac{cz_1 + dz_2}{az_1 + bz_2} = \frac{c + d\xi}{a + b\xi}$$

agreeing with the linear fractional mappings in (8).

However, the action must be extended to include the line $z_1 = 0$ of ‘slope’ ∞ as follows:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \infty = d/b, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot (-a/b) = \infty.$$

The stereographic projection $\sigma = j|_{S^2_{(0,1,0)}}$ (with j in (1)) maps $S^2_{(0,1,0)}$ bijectively onto $\widehat{\mathbb{C}}$ and $\sigma(0, 0, 0) = \infty$ gives a concrete meaning to the point ∞ such that it can be treated as any other point of $\widehat{\mathbb{C}}$. Thus, geometry of the image plane $\widehat{\mathbb{C}}$ of the conformal camera with the image projective transformations given by the group $\mathbf{PSL}(2, \mathbb{C})$ acting by linear-fractional transformations can be dually described as follows:

1. $\widehat{\mathbb{C}}$ is the complex projective line, i.e., $\widehat{\mathbb{C}} \cong P^1(\mathbb{C})$ where

$$P^1(\mathbb{C}) = \{ \text{complex lines in } \mathbb{C}^2 \text{ through the origin} \}$$

with the group of projective transformations $\mathbf{PSL}(2, \mathbb{C})$. Thus, the image projective transformations acting on the points of the extended image plane (or simply, the image plane) of the conformal camera can be identified with projective geometry (containing Euclidean geometry as a sub-geometry) of the one-dimensional complex line [6].

2. $\widehat{\mathbb{C}}$ is the Riemann sphere since under stereographic projection $\sigma = j|_{S^2_{(0,1,0)}}$ we have the isomorphism $\widehat{\mathbb{C}} \cong S^2_{(0,1,0)}$. The group $\mathbf{PSL}(2, \mathbb{C})$ acting on $\widehat{\mathbb{C}}$ consists of the bijective meromorphic mappings of $\widehat{\mathbb{C}}$ [32]. Thus, it is the group of holomorphic automorphisms of the Riemann sphere that preserve the intrinsic geometry imposed by complex structure, known as Möbius geometry [27] or inversive geometry [13].

What we have just described shows the following fundamental property: projective geometry underlying the conformal camera, also called Möbius or inversive geometry, and holomorphic complex structure that provides the framework for the development of complex numerical analysis, are in fact two faces—one ‘geometric’ and the other ‘numerical’—of the same coin. We stress that the real projective geometry underlying the pinhole camera and usually employed in computer vision [49, 55] does not possess this fundamental property which sets apart our modeling of primate visual perception from other approaches.

4.1 The Conformal Camera and Visual Perception

The image plane of the conformal camera does not admit a distance that is invariant under image projective (that is, linear-fractional) transformations. Therefore, geometry of the conformal camera does not possess a Riemannian metric; for instance, there is no curvature measure. As customary in complex projective (Möbius or inversive) geometry, we consider a line as a circle passing through the point ∞ . Then, the fundamental property of this geometry can be expressed as follows: linear-fractional mappings take circles to circles. Thus, circles can play the role of geodesics. Moreover, each circle carries a signature of curvature—the inverse of the radius. We showed before that linear-fractional mappings are conformal; we add here for completeness that stereographic projection $\sigma = j|_{S^2_{(0,1,0)}}$ is also conformal and maps circles in the sphere $S^2_{(0,1,0)}$ onto circles in $\widehat{\mathbb{C}}$. In conclusion, circles play a crucial role in the conformal camera geometry and it should be reflected in psychological and computational aspects of natural scene understanding if this camera is relevant to modeling primate visual perception.

Neurophysiological experiments demonstrate that the retina performs filtering of impinged images that extract local contrast spatially and temporally. For instance, center surround cells at the retinal processing stage are triggered by local spatial changes in intensity referred to as edges or contours. This filtering is enhanced in the primary visual cortex, the first cortical area receiving, via LGN, the retinal output, which itself is a case study in dense packing of

overlapping visual submodalities: motion, orientation, frequency (color), and oculomotor dominance (depth). In psychological tests, humans easily detect a significant change in spatial intensity (low-level vision), and effortlessly and unambiguously group this usually fragmented visual information (contours of occluded objects, for example), into coherent, global shapes (intermediate-level vision). Considering its computational complexity, it is one of the most difficult problems that primate visual system has to solve [65].

The Gestalt phenomenology and quantitative psychological measurements established the rules, summarized in the ideas of good continuation [35, 68] and association field [19], that determine interactions between fragmented edges such that they extend along continuous contours joining them in the way they will normally be grouped together to faithfully represent a scene. Evidence accumulated in psychological and physiological studies suggests that the human visual system utilizes a local grouping process (association field) with two very simple rules: collinearity and co-circularity with underlying scale invariant statistics for both geometric arrangements in natural scenes. These rules were confirmed in [56, 16] by statistical analysis of natural scenes. Two basic intermediate-level descriptors that the brain employs in grouping elements into global objects are the medial axis transformation [10], or symmetry structure [40, 41], and the curvature extrema [3, 30]. In fact, the medial axis, which visual system extracts as a skeletal (intermediate-level) representation of objects [36], can be defined as the set of the centers of maximal circles inscribed inside the contour. The curvatures at the corresponding points of a contour are given by the inverse radii of the circles.

From the above discussion we see that, on one hand, co-circularity and scale invariance emerge as the most basic concepts used by intermediate-level vision in solving the difficult problems of grouping local elements into individual objects of natural scenes. On the other hand, the non-metric projective geometry of the conformal camera that models eye imaging functions can be entirely constructed from circles such that co-circularity is preserved by projective transformations. Thus, it seems that the conformal camera would be very useful in modeling eye's imaging functions related to the lower and intermediate-level natural vision.

Other characteristics of the conformal camera that are uniquely useful in modeling primate visual perception are discussed in the remaining part of this article. Next, we briefly review the unity of geometry and numerical methods by showing that the conformal camera has its own projective Fourier transform (PFT).

5 Projective Fourier Analysis

The projective Fourier analysis has been constructed by restricting geometric Fourier analysis of $\mathbf{SL}(2, \mathbb{C})$ —a direction in the representation theory of the semisimple Lie groups [33]—to the image plane of the conformal camera (see Section 5.1 in [62]). The resulting projective Fourier transform (PFT) of a given

image intensity function $f(z) \in L^2(\mathbb{C})$ is the following

$$\widehat{f}(s, k) = \frac{i}{2} \int f(z) |z|^{-is-1} \left(\frac{z}{|z|} \right)^{-k} dz d\bar{z} \quad (10)$$

where $(s, k) \in \mathbb{R} \times \mathbb{Z}$ and if $z = x_3 + ix_1$, then $\frac{i}{2} dz d\bar{z} = dx_3 dx_1$ is the Euclidean measure on the image plane. In this work, we consider only the noncompact picture of PFT and, for a complete mathematical account, that includes also the compact picture we refer to [62]. The noncompact and compact pictures in the case of Euclidean group correspond to the classical and spherical Fourier analyses, respectively (Section 3 in [61]). In the next remark we justify the name ‘projective Fourier transform’, and, for comprehensive account, we refer to [62].

Remark 1 *The functions $\Pi_{s,k}(z) = |z|^{is} \left(\frac{z}{|z|} \right)^k$; $s \in \mathbb{R}$, $k \in \mathbb{Z}$ are all one dimensional unitary representations of the Borel subgroup $\mathbf{B} = \mathbf{MAN}$ of $\mathbf{SL}(2, \mathbb{C})$, and they play in (10) the role complex exponentials play in the classical Fourier transform. These one dimensional representations are all finite unitary representations of the Borel subgroup \mathbf{B} , as opposed to the fact that all nontrivial unitary representations of $\mathbf{SL}(2, \mathbb{C})$ are infinite. Furthermore, the group \mathbf{B} ‘exhausts’ the projective group $\mathbf{SL}(2, \mathbb{C})$ by Gauss decomposition $\mathbf{SL}(2, \mathbb{C}) \doteq \overline{\mathbf{N}}\mathbf{B}$, where ‘ \doteq ’ means that the equality holds up to lower dimensional subset, that is, almost everywhere, and $\overline{\mathbf{N}}$ in (7) represents Euclidean translations.*

In log-polar coordinates (u, θ) given by $\ln r e^{i\theta} = \ln r + i\theta = u + i\theta$, $\widehat{f}(k, s)$ has the form of the standard Fourier integral

$$\widehat{f}(s, k) = \int \int f(e^{u+i\theta}) e^u e^{-i(us+\theta k)} du d\theta, \quad (11)$$

where we used $\frac{i}{2} dz d\bar{z} = e^{2u} du d\theta$. We see that a function f that is integrable on $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, has finite PFT,

$$\left| \widehat{f}(s, k) \right| \leq \int_0^{2\pi} \int_{-\infty}^{u_1} f(e^{u+i\theta}) e^u du d\theta = \int_0^{2\pi} \int_0^{r_1} f(re^{i\theta}) dr d\theta < \infty. \quad (12)$$

Therefore, this f can be extended to \mathbb{C} by $f(0) = 0$. Thus, in spite of the logarithmic singularity of log-polar coordinates, the projective Fourier transform of integrable functions on \mathbb{C} is finite. This observation will be crucial when we discretize the PFT in the next section.

Inverting (11), which is done in the (u, θ) -space, we get

$$e^u f(u, \theta) = \frac{1}{(2\pi)^2} \sum_{k=-\infty}^{\infty} \int \widehat{f}(s, k) e^{i(us+\theta k)} ds, \quad (13)$$

where $f(u, \theta) = f(e^{u+i\theta})$. We stress that although $f(e^{u+i\theta})$ and $f(u, \theta)$ are numerically equal, they are given on different spaces; $f(e^{u+i\theta})$ is on the image

plane in polar coordinates and $f(u, \theta)$ is on the space defined by rectangular (u, θ) -coordinates.

Finally, by expressing (13) in the z -variable, we obtain the inverse projective Fourier transform

$$f(z) = \frac{1}{(2\pi)^2} \sum_{k=-\infty}^{\infty} \int \widehat{f}(s, k) |z|^{is-1} \left(\frac{z}{|z|} \right)^k ds. \quad (14)$$

5.1 Discrete Projective Fourier Transform

To discretize the PFT we use the fact that $\widehat{f}(s, k)$ is finite for an integrable function f , see (12). By removing a disk $|z| \leq r_a$, we can assume that the support of $f(u, \theta)$ is contained within $(\ln r_a, \ln r_b) \times [0, 2\pi)$. We approximate the integral in (11) by a double Riemann sum

$$\widehat{f}(2\pi m/T, n) \approx \frac{2\pi T}{NM} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} e^{u_k} f(e^{u_k} e^{i\theta_l}) e^{-2\pi i(mk/M + nl/N)}$$

with $M \times N$ partition points

$$(u_k, \theta_l) = (\ln r_a + kT/M, 2\pi l/N); 0 \leq k \leq M-1, 0 \leq l \leq N-1, T = \ln(r_b/r_a). \quad (15)$$

Then, introducing

$$f_{k,l} = (2\pi T/MN) f(e^{u_k} e^{i\theta_l}) \text{ and } \mathbf{f}_{k,l} = (2\pi T/MN) f(u_k, \theta_l) \quad (16)$$

and defining $\widehat{f}_{m,n}$ by

$$\widehat{f}_{m,n} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f_{k,l} e^{u_k} e^{-i2\pi mk/M} e^{-i2\pi nl/N}, \quad (17)$$

we obtain

$$\mathbf{f}_{k,l} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \widehat{f}_{m,n} e^{-u_k} e^{i2\pi mk/M} e^{i2\pi nl/N}. \quad (18)$$

We note that $\widehat{f}_{m,n} \approx \widehat{f}(2\pi m/T, n)$ and refer to [28] for a discussion of numerical aspects on the approximation. Both expressions (17) and (18) can be computed efficiently by FFT algorithms since the exponents are taken at equidistant points. See simulation for a bar pattern in Fig. 4.

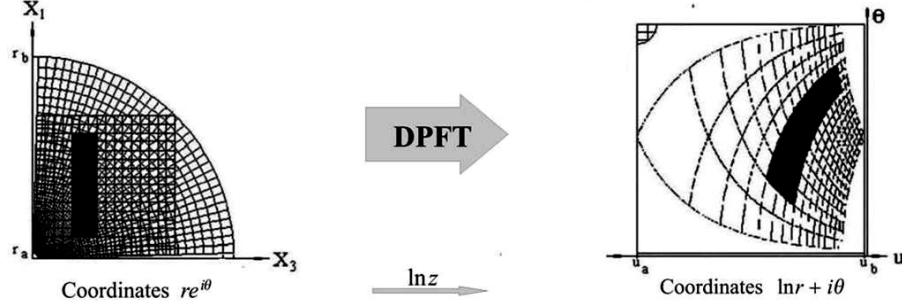


Figure 4. Exp-polar sampling (the distance between circles partially displayed in the first quadrant changes exponentially) of a bar pattern on the retina is shown on the left. The bar pattern in the cortical space rendered by the inverse DPFT computed with FFT is shown on the right. The cortical uniform sampling grid, which is obtained by applying complex logarithm to the exp-polar grid in (a), is shown only in the upper left corner.

Finally, on introducing $z_{k,l} = e^{u_k + i\theta_l}$ into (17) and (18), we arrive at the (M, N) -point discrete projective Fourier transform (DPFT) and its inverse:

$$\hat{f}_{m,n} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f_{k,l} \left(\frac{z_{k,l}}{|z_{k,l}|} \right)^{-n} |z_{k,l}|^{-i2\pi m/T+1} \quad (19)$$

and

$$f_{k,l} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{f}_{m,n} \left(\frac{z_{k,l}}{|z_{k,l}|} \right)^n |z_{k,l}|^{i2\pi m/T-1}, \quad (20)$$

now with $f_{k,l} = (2\pi T/MN)f(z_{k,l})$. The projectively adapted characteristics of the discrete projective Fourier analysis can be expressed as follows:

$$f'_{k,l} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{f}_{m,n} \left(\frac{z'_{k,l}}{|z'_{k,l}|} \right)^n |z'_{k,l}|^{i2\pi m/T-1}, \quad (21)$$

where $z'_{k,l} = g^{-1} \cdot z_{k,l}$, $g \in \mathbf{SL}(2, \mathbb{C})$ and $f'_{k,l} = (2\pi T/MN)f(z'_{k,l})$.

Although projective characteristics must be derived in z -coordinates, in practical image processing, (21) should be expressed in log-polar coordinates to be fast computable by FFT. To this end, let $(u'_{m,n}, \theta'_{m,n})$ denote log-polar coordinates of $z'_{m,n} = e^{u'_{m,n}} e^{i\theta'_{m,n}}$. In these coordinates, (21) is given by the following expression (see [61, 62] for details)

$$f'_{m,n} = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \hat{f}_{k,l} e^{-u'_{m,n}} e^{i2\pi u'_{m,n} k/T} e^{i\theta'_{m,n} lL}.$$

Thus, we can render image projective transformations in terms of projective Fourier transform of the original image only.

6 DPFT in Computational Vision

We discussed before the relevance of the conformal camera to the intermediate-level vision task of grouping image elements into individual objects in natural scenes. Here we want to discuss the relevance of the data model of image representation based on projective Fourier analysis to image processing in computational vision, including visual neuroscience and biologically motivated machine vision systems.

6.1 Modeling the Retinotopy

The mappings $w = \ln(z \pm a) - \ln a$, with $a > 0$ and $\pm a$ indicating, for different signs, the left or right brain hemisphere, are accepted approximations of the topographic structure of primate primary visual cortex (V1) [54], where the parameter a removes the singularity of the logarithm. However, the discrete projective Fourier transform (DPFT) that provides the data model for retinal image representation, can be efficiently computed by FFT only in log-polar coordinates given by the complex logarithm $w = \ln z$, the mapping with distinctive rotational and zoom symmetries:

$$\ln(e^{i\theta} z) = \ln z + i\theta, \quad \ln(\rho z) = \ln z + \ln \rho.$$

Thus, we see that the Schwartz model of the retina comes with drastic consequences; it destroys rotation and zoom symmetries. We also recall that PFT in log-polar coordinates does not have a singularity at the origin, see (12).

The following facts support our modeling with DPFT. First, for small $|z| \ll a$, $\ln(z \pm a) - \ln a$ is approximately linear while, for large $|z| \gg a$, it is dominated by $\ln z$. Secondly, to construct discrete sampling for DPFT, the image was regularized by removing a disc representing the fovea (see previous section). Thirdly, there is accumulated evidence pointing to the fact that the fovea and periphery have different functional roles in vision [51, 52, 70] and likely involve different image processes. Finally, by the split theory of hemispherical image representation, which we mentioned before, the foveal region has a discontinuity along the vertical meridian, with each half processed in a different brain hemisphere [39]. We note that the two hemispheres are connected by a massive bridge of 500 million neuronal axons called the corpus callosum.

We conclude this discussion with the following remarks: both models our and Schwartz' model in [54] (see Fig. 5), as well as all other similar models, are, in fact, fovea-less models [67].

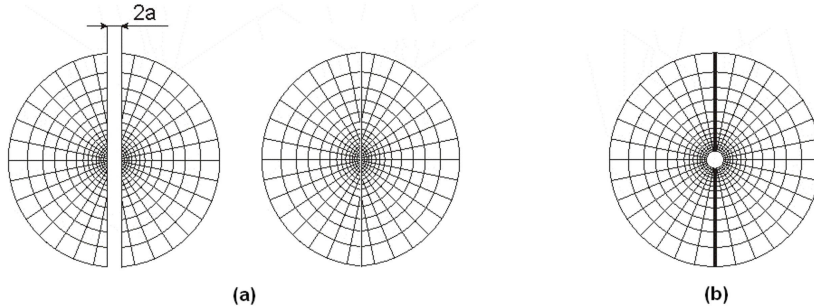


Figure 5. (a) Schwartz model of the retina: the strip of width $2a$ is removed and two half-maps of $\ln z$ are shifted to meet along the vertical meridian. (b) Our model: the fovea is removed and the retina is split along the vertical meridian, conforming to the split theory of the retino-cortical projection.

Furthermore, since the fovea is explicitly removed in our modeling, we expect to extend the present model to include foveal representation in the next stage of this modeling. In fact, the lack of the fovea in our modeling is one of the challenges that is stalling implementation of the model. We continue this discussion in Section 6.5.1.

6.2 On Numerical Implementation of DPFT

The DPFT approximation was obtained using the rectangular sampling grid (u_k, θ_l) in (15), corresponding, under the mapping,

$$w_{k,l} = u_k + i\theta_l \mapsto z_{k,l} = e^{u_k + i\theta_l} = r_k e^{i\theta_l}$$

to nonuniform sampling grid with equal sectors

$$\alpha = \theta_{l+1} - \theta_l = \frac{2\pi}{N}, \quad l = 0, 1, \dots, N-1 \quad (22)$$

and with ring radii increasing exponentially

$$\rho_k = r_{k+1} - r_k = e^{u_{k+1}} - e^{u_k} = e^{u_k}(e^\delta - 1) = r_k(e^\delta - 1), \quad k = 0, 1, \dots, M-1, \quad (23)$$

where $\delta = u_{k+1} - u_k$. The radii $r_k = r_0 e^{k\delta}$ are given in terms of the spacing $\delta = \frac{T}{M}$ and $r_0 = r_a$, where r_a is the radius of the disc that has been removed to regularize logarithmic singularity, see (15).

Lets assume that we have been given a picture of the size $A \times B$ in pixel units, which is displayed with K dots per unit length (dpl). Then, the physical dimensions, in the chosen unit of length, of the pixel and the picture are $1/K \times 1/K$ and $A/K \times B/K$, respectively. Also, we assume that the retinal coordinates' origin (fixation) is the picture's center.

The central disc of radius r_0 represents the fovea with a uniformly distributed of grid points and the number of the foveal pixels N_f given by $\pi r_0^2 = N_f/K^2$. This means that the fovea cannot increase the resolution, which is related to

the distance of the picture from the eye. The number of sectors is obtained from the condition $2\pi(r_0 + r_1)/2 \approx N(1/K)$, where $N = \lceil 2\pi r_0 K + \pi \rceil$. Here $\lceil a \rceil$ is the closest integer to a . To get the number of rings M , we assume that $\rho_0 = r_0(e^\delta - 1) = 1/K$ and $r_b = r_M = r_0 e^{M\delta}$. We can take either $r_b = (1/K) \min(A, B)/2$ or $r_b = (1/K) \sqrt{A^2 + B^2}/2$. Thus, $\delta = \ln[(1 + 1/r_0 K)]$ and $M = (1/\delta) \ln(r_b/r_0)$.

Example 2 We let $A \times B = 512 \times 512$ and $K = 4$ per mm, so the physical dimensions in mm are 128×128 and $r_b = 128\sqrt{2}/2 = 90.5$. Furthermore, we let $N_f = 296$, so $r_0 = 2.427$ and $N = 64$. Finally, $\delta = \ln(10.7084/9.7084) \approx 0.09804$ and $(1/0.09804) \ln(90.5/2.427) \approx M = 37$. The sampling grid consists of points in polar coordinates: $(r_k + \rho_{k+1}/2, \theta_l + \pi/64) = (2.552e^{k \cdot 0.09804}, (2l + 1)\pi/64)$ $k = 0, 1, \dots, 36, l = 0, 1, \dots, 63$.

In this example, the number of pixels in the original image is 262,144, whereas the foveal (uniform sampling) and peripheral (log-polar sampling) representation of the image contain only 2,664 pixels.

We stress again that $f_{k,l}$ and $\hat{f}_{k,l}$ are discretizations of the same image in different planes; $f_{k,l}$ are the image samples in the image plane sampled on a nonuniform grid $(e^{u_k} e^{i\theta_l})$, while the inverse DPFT output (18) gives the image samples $\hat{f}_{k,l}$ on the uniform grid (u_k, θ_l) , where $u_k = \ln r_k$.

In summary, a simple description of the imaging model based on DPFT is as follows: an image (analog or digital) of a scene impinged on the retina is sampled on a nonuniform exp-polar grid, $\{r_k e^{i\theta_l}\}_{M \times N}$, that approximates the density distribution of retinal ganglion cells, giving the set of pixels $\{f_{k,l}\}_{N \times M}$. In this grid, the radial spacing changes exponentially: $r_k = r_a e^{\delta k}$, $k = 1, 2, \dots, M$, and the angular spacing is constant: $\theta_l = \alpha l$, $l = 1, 2, \dots, N$. As it was shown in Example 2, this sampling results in about 100 times less pixels than in the original image. To render $\{f_{k,l}\}_{N \times M}$, the DPFT is formed and computed by FFT in log-polar coordinates (u_k, θ_l) obtained by applying a complex logarithm as follows: $\ln(r_a e^{\delta k} e^{i\alpha l}) = \ln r_a + \delta k + i\alpha l = u_k + i\theta_l$, resulting in the set $\{\hat{f}_{k,l}\}_{M \times N}$. Next, the IPFT is assembled and computed again by FFT, this time giving the image samples $f_{k,l} = \hat{f}_{k,l}$ rendered in cortical (log-polar) coordinates (u_k, θ_l) .

6.3 Relation to Other Numerical Approaches

From the numerical approaches to foveate (or space-variant) vision, involving, for example, Fourier-Mellin transform or log-polar Hough transform, the most closely related to our work are results reported by Schwartz' group at Boston University. We note that the approximation of the retinotopy by a complex logarithm (see Section 6.1) was first proposed by Eric Schwartz in 1977. This group introduced the fast exponential chirp transform (FECT) [11] in their attempt to develop numerical algorithms for space-variant image processing. Basically, both FECT and its inverse were obtained by the change of variables in both the spatial and frequency domains in the standard Fourier integrals.

The discrete FECT was introduced somehow ad hoc, without references to numerical aspects of the approximation. Moreover, some basic components of Fourier analysis, such as underlying geometry or Plancherel measure was not considered. In comparison, projective Fourier transform (PFT) provides an efficient image representation well adapted to projective transformations produced in the conformal camera by the group $\mathbf{SL}(2, \mathbb{C})$ acting on the image plane by linear-fractional mappings. Significantly, PFT can be obtained by restricting geometric Fourier analysis of the Lie group $\mathbf{SL}(2, \mathbb{C})$ to the image plane of the conformal camera. Thus, the conformal camera comes with its own harmonic analysis. Moreover, PFT is computable by FFT in log-polar coordinates given by a complex logarithm that approximates the retinotopy. It implies that PFT can integrate the head, eyes, and visual cortex into a single computational system. This aspect is discussed, with special attention to perisaccadic perception, in the remaining part of the paper. Another advantage of PFT is the complex (conformal) geometric analysis underlying the conformal camera. We discussed, in Section 4.1, the relation of this geometry to the intermediate-level vision problem of grouping local contours into individual objects and the background of natural scenes.

The other approaches to space-variant vision use the geometric transformations, mainly based on a complex logarithmic function between the nonuniform (retinal) sampling grid and the uniform (cortical) grid for the purpose of developing computer programs. These approaches can be classified into two different groups. The first group of problems deal with visualizing and classifying large information data sets. We give two examples for the first group. The first deals with the problem of mapping information space to the image space for navigation through complex two-dimensional data sets when viewing small details and at the same time the general overview [12]. The second gives the model based image processing in mathematical morphology for qualifying/segmenting/quantifying spots topology in genomic microarray-based data [1]. The second group of problems is related to robotic vision. We give only a few examples of such problems, which include tracking [7], navigation [5], detection salient regions [57], and disparity estimation [42]. However, it seems that they share one common problem: high computational costs in the geometric transformation process.

In the next figure, we show a simulation applied to Fig. 2 (a) with the software available over internet [8]. In Fig. 6, the San Diego skyline and harbor shown in (a) is sampled in retinal exp-polar coordinates (with the vertical meridian deleted according to the split theory discussed before) and mapped by a complex logarithm transformation to rectangular log-polar coordinates (b). The inverse geometric transformation shown in (c) results in the retinal image that simulates the sampling by the ganglion cells density as a function of eccentricity.

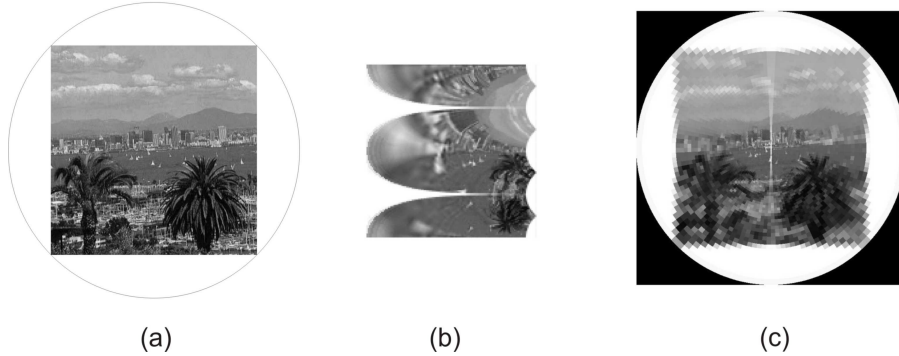


Figure 6. (a) San Diego skyline and harbor. (b) Its log-polar image, the vertical meridian deleted, obtained by the geometric transformation of both the polar samples with the radial partition changing exponentially and a constant angular partition, into regular samples in log-polar rectangular plane (c)

We note that the image processing presented here (see the last paragraph in the previous section) differs from the above simulation by one crucial aspect: we use projective Fourier analysis framework for image representation that provides low computational cost of the retino-cortical (logarithmic) transformation.

6.4 DPFT and Binocular Vision

In order to carry out numerical experiments with the discrete PFT, the conformal camera should work in the following setup: we get a set of samples $f_{k,l} = f(e^{u_k} e^{i\theta_l})$ of an image f from a camera with anthropomorphic visual sensors [9] or an ‘exp-polar’ scanner with the sampling geometry similar to the distribution density of the retinal ganglion cells. Next, we form DPFT $\hat{f}_{k,l}$ according to (17) and compute it with FFT. Then, we compute IDPFT of $\hat{f}_{k,l}$ given in (18), again with FFT. However this output from IDPFT renders the retinotopic image $f_{k,l}$ of the retinal samples in cortical log-polar coordinates. This setup provides an efficient model that integrates the head, eyes, and the cortex into a single computational system, which is introduced next.

We discuss this integrated system by assuming that a 3D scene consists of a gray square with a red bar located in front of it (see Fig. 7).

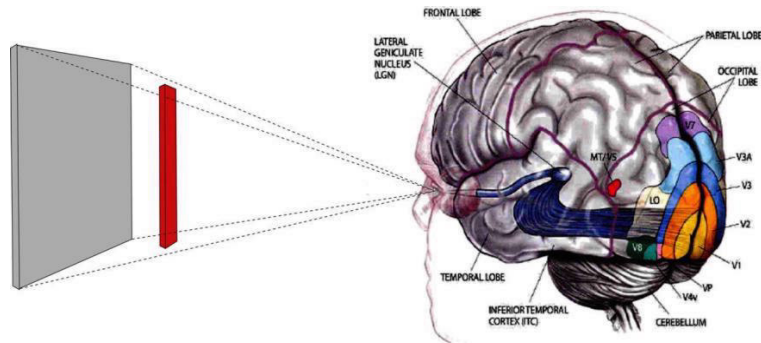


Figure 7. The scene consisting of a gray square with a red bar in front of it is seen by an observer. The visual pathway with the major cortical areas is shown.

The integrated binocular system with eyes modeled by the conformal cameras and this scene as seen from above is shown in Fig. 8.

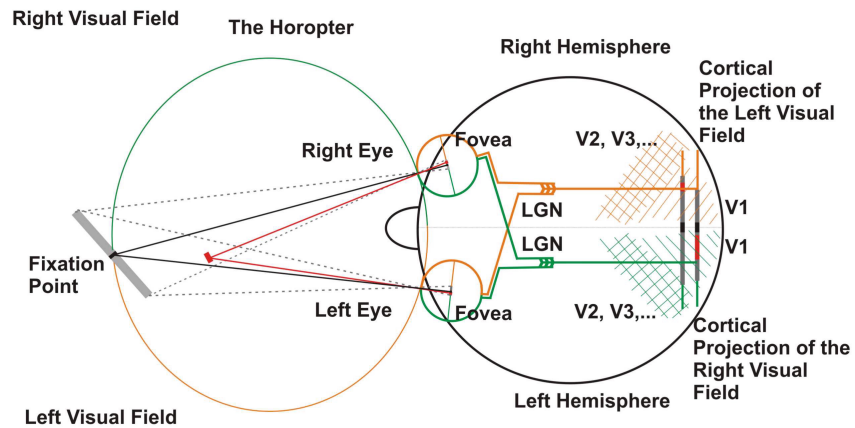


Figure 8. The head-eyes-visual cortex integrated system. Following from the fact that eyes are modeled by the conformal camera, theoretical horopters are conics that resemble empirical horopters.

A simulation of the integrated binocular system with the grey square-red bar scene can be seen in Fig. 9. Each eye sees the scene from a different vantage point ((a) and (c) in Fig. 9), as the eyes are separated laterally. The retinal projections are sampled on the exp-polar grid with the meridian line removed as implied by the split theory.

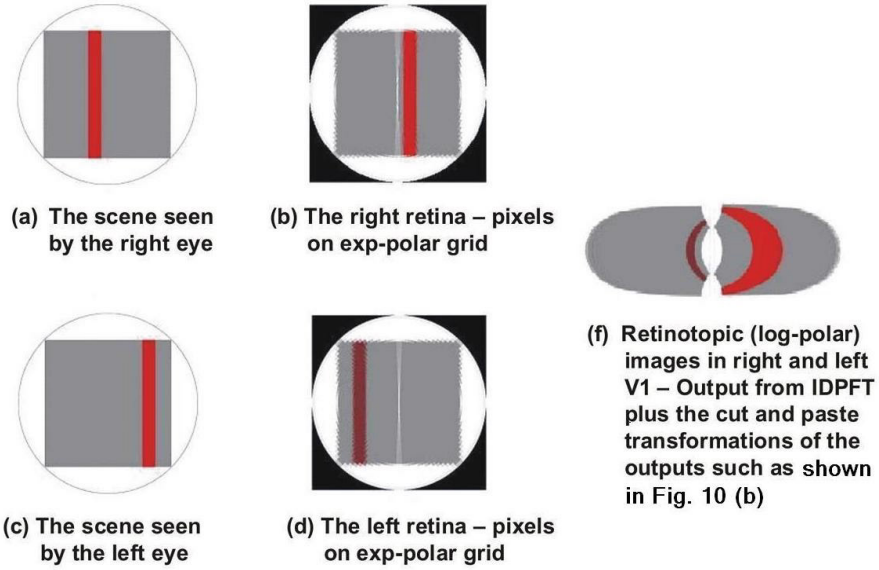


Figure 9. In (a) and (c), the 3D scene from Fig. 8 is seen from a different vantage point by each eye (i.e., the conformal camera) due to eyes lateral displacement. The Matlab-simulated right and left retinal projections and the retinotopic image can be seen in (b), (d) and (f), respectively.

The retinotopic images are simulated in Matlab using the program from [8], and the cut-and-paste transformations are used to account for the global retinotopy topology. For example, the output from FFT computing the inverse DPFT of the scene projected on the right eye and sampled by ganglion cells is shown in Fig 10 (b).

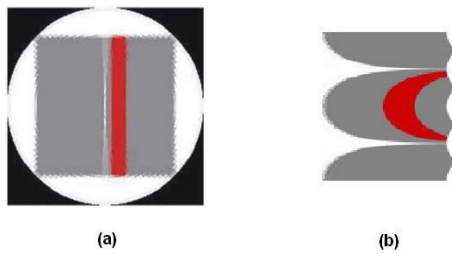


Figure 10. (a) The simulation of the right eye's projected scene sampled by ganglion cells. (b) The retinotopic image of the sampled projection shown in (a); the vertical size corresponds to the length of the angular interval $[-\pi, \pi]$

The cut-and-paste operation is applied to the output in Fig. 10 (b) and to the corresponding DPFT output of the left eye to obtain (f) in Fig. 9.

6.5 Modeling Perisaccadic Perception with DPFT

Because of acuity limitations of foveate vision, a sequence of fast eye rotations is necessary for processing the details of the scene by fixating eyes consecutively on the targets of interest. The sequence of fixations, saccades and smooth pursuits, called the scanpath, is the most basic feature of foveate vision (cf., Fig 2). The fact that we do not see moving images on the retinas points to a poor integration of visual information across fixations during tasks such as reading, visual searching, or looking at a scene. Given the limited computational resources, it is critical that visual information is not only efficiently acquired during each fixation, but also that it is done without starting anew much of this acquisition process at each fixation.

Although we are not aware of discontinuities in a scene perception when executing a scanpath, this visual constancy is not perfect. In psychophysical laboratory experiments, the phenomenon of perisaccadic compression is observed: before the onset of the saccade, brief flashes are perceived by human subjects to be compressed around the impending saccade target [34, 53], see Fig. 11. However, perisaccadic perception experiments have revealed a multitude of mislocalization phenomena, pointing to the involvement of many different neural processes. Accordingly, many different theories have been proposed, see [26].

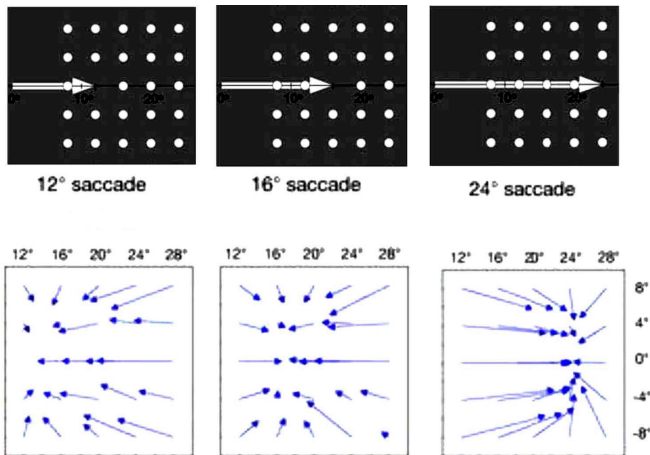


Figure 11. The spatial pattern of perisaccadic compression. It shows experimental data of the absolute mislocalization (lower row), reference to the true position of flashed dot randomly chosen from an array of 24 dots and four different saccade amplitude (upper row). Adapted from [38].

Two computational theories of perisaccadic vision that have been proposed in visual neuroscience are related to our modeling. The first theory, suggested in [66], states that an efference copy generated by SC, a copy of an oculomotor command to rotate eyes in order to execute the saccade, is used to uniformly shift cortical neural activity representing spatial locations of the saccade target

area toward foveal representation. It was proposed that this shift is reflected on the neuronal level by a transient spatial remapping of the receptive fields in numerous retinotopically organized cortical areas ([18, 37]), including the superior colliculus (SC), parietal eye field (PEF), and frontal eye field (FEF). It can explain the perceived increase in spatial resolution around the saccade target as more foveal neurons are available there to process the details of objects. Furthermore, because the shift occurs in logarithmic coordinates that approximate retinotopy, the model can also explain perceived perisaccadic compression.

The second theory, [38], explains the perisaccadic compression by directing spatial attention to the target of a planned saccade. The proposed computational model assumes that the initial stimulus neuronal activity in the visual cortical area is distorted by the feedback of the retinotopically organized activity hill of the saccade target in the oculomotor SC layer, what pushes the population response of the flashed stimulus in retinotopically organized cortical areas (including PEF and FEF) towards the saccade target. This boost of performance at the target location of the saccade occurs immediately before the saccade onset increases spatial discrimination. The shift of the neuronal activity in logarithmic coordinates, and hence perisaccadic compression, is a direct consequences of it.

Because circuitry underlying receptive field remapping is widespread and not well understood, it cannot be easily decided whether saccadic remapping is the cause or consequence of saccadic compression. For example, only recently it was reported in [46] that a phenomenon very similar to the remapping occurs in extrastriate (V4, and, though progressively weaker, V3, V2 and V1) cortical areas in humans. Remarkably, remapping in extrastriate cortex could be functionally related to the integration of visual information from a constant object across saccades [23].

In this section, we model perisaccadic perception using the integrated binocular system, addressing the process of presaccadic activity consisting shifts of neurons current receptive fields to their future postsaccadic locations, that is thought to underlie the scene remapping based on anticipated saccadic eye movement (efference copy) with the accompanied perisaccadic perceptual space compression. The postsaccadic activity during which actual integration of visual features takes place, will be considered in the next stage of our modeling. Although, our modeling directly conforms to the theory in [66], it may also be useful, on the image processing level, in representing the resulting receptive field shifts from ‘attentional multiplicative gain field interaction’ [38], especially since the efficiency of the whole modeling, which must be repeated three times per second, was not addressed by the authors.

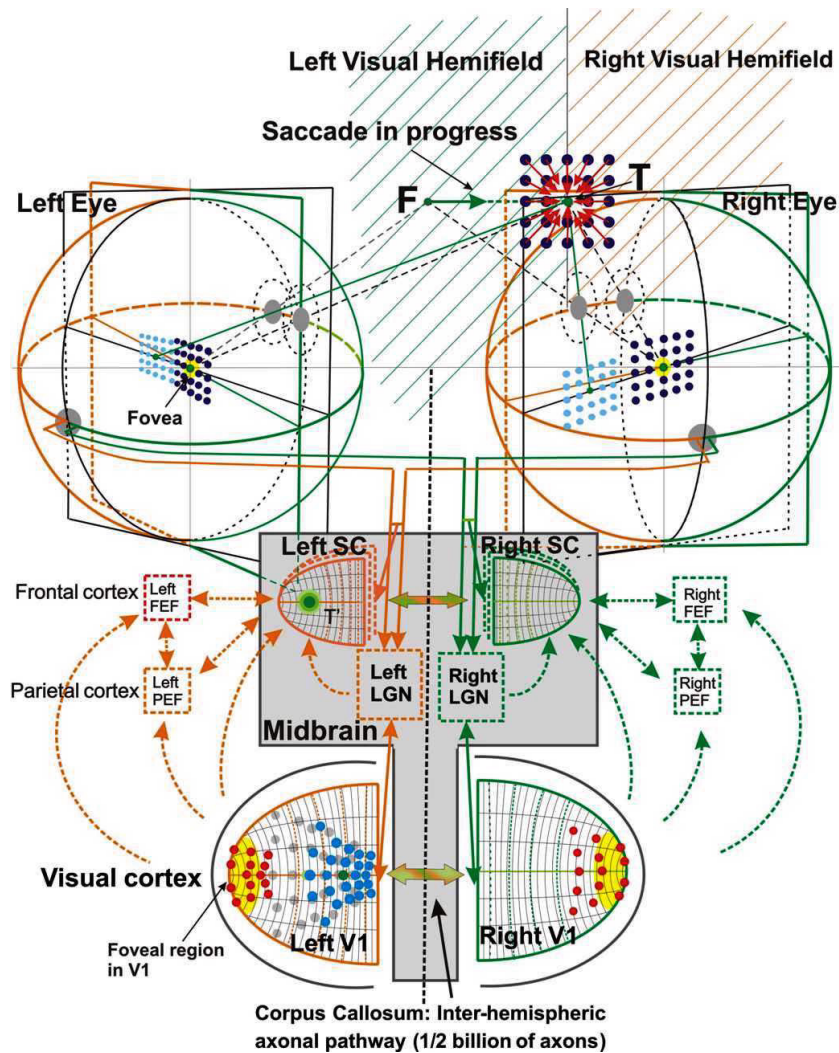


Figure 12. The description is given in the text of the article.

We start here by supplementing the integrated binocular system presented in Section 6.4 with the most important subcortical and cortical pathways of the visuo-saccadic neural processes. These pathways depicted by arrowhead lines in Fig. 12, include the SC of the midbrain, which contains retinotopically organized visual and oculomotor layers, the PEF, and the FEF in the parietal and frontal lobes of the neocortex (which themselves obtain inputs from many visual cortical areas) for assisting the SC in the control of the involuntary (PEF) and voluntary (FEF) saccades. We also include the interhemispheric pathways, the corpus callosum (about 500 million of neuronal axons connecting cerebral cortical hemispheres), and the intercollicular commissure, because the coordi-

nated movement of two eyes is a bihemispheric event. The motor commands that originate from the brain's major hemisphere (the left hemisphere for most right-handed people) travel across the corpus callosum to the minor hemisphere then down to brainstem, where part of it again crosses to the other side of the brain before both eyes are finally moved in coordination [17]. We believe that building on the existing models [21] and accelerating advances in visual neuroscience will soon allow the inclusion of these pathways such that the operation of a more complete system of perisaccadic perception can be addressed in numerical modeling in a way that could be useful in neural engineering designs.

The course of events taking place during perisaccadic perception, shown in Fig. 12, is as follows: the eyes are fixated at \mathbf{F} and the new stimulus appears at \mathbf{T} . The SC population \mathbf{T}' at the retinotopic image of \mathbf{T} (green spot in the left SC) calculates the position of the target \mathbf{T} of an impending saccade. The SC also codes the motor command for the execution of the saccade.

About 50 ms before the onset of the saccade, during the saccade (about 30 ms), and about 50 ms after the saccade, the visual sensitivity is reduced and flashes (dark blue dots) around \mathbf{T} are not perceived in veridical locations. Instead, a copy of the motor command (efference copy) is sent to translate the cortical image (light blue dots in V1) of flashes to remap it into a target-centered frame (red dots in V1).

This internal remapping results in the illusory compression of flashes, shown by red arrows. The compression is perceived around the incoming target \mathbf{T} even though the eyes fixation is moving from \mathbf{F} to \mathbf{T} . The location of the cortical area of neural correlates of remapping is uncertain; it is required that this area is retinotopically organized. Although it could be PEF/FEF, here, for simplicity, this area is represented by V1.

During the fixation of eyes at \mathbf{F} , lasting on average about 300 ms, the image is sampled by ganglion cells $f_{k,l} = f(e^{u_k} e^{i\theta_l})$ and its DPFT $\hat{f}_{k,l}$ is computed by FFT in log-polar coordinates (u_k, θ_l) where $u_k = \ln r_k$. The inverse DPFT, computed again by FFT, gives a cortical image representation

$$f_{k,l} = f(u_k, \theta_l) = f(e^{u_k} e^{i\theta_l})$$

where disparity-sensitive cells contribute to the building 3D understanding of the scene. In the same fixation period, the next saccade's target \mathbf{T} is selected (PEF/FEF) and its position in respect to the fovea is calculated and converted into the motor command to move the eyes (SC). During that time interval of about 130 ms, the visual sensitivity is reduced, neural processes, using a copy of the eyes motor command (efference copy), transiently shift the cortical image. In our modeling, this shift is generated using the shift property of Fourier transform as follows

$$f(u_k + j\delta, \theta_l) = f_{k+j,l} = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} e^{i2\pi mj/M} \hat{f}_{k,l} e^{-(u_k+j\delta)} e^{i2\pi mk/M} e^{i2\pi nl/N},$$

where δ is the corresponding spacing. It brings the presaccadic scene at \mathbf{F} in

fovea-centered coordinates into postsaccadic scene \mathbf{T} in target-centered coordinates. However,

$$f(u_k + j\delta, \theta_l) = f(e^{u_k + j\delta} e^{i\theta_l}) = f(e^{j\delta} r_k e^{i\theta_l})$$

compresses perceptual space.

6.5.1 Challenges with Implementation

There are some problems that must be addressed before we can implement our modeling of primate perception. One problem is due to the fact that the model of retinotopy is fovea-less. The other is related to the global topology of retinotopy, and in particular, to the vertical meridian split in retinas (and hence in the visual field) of the brain’s hemispheric projections.

In order to address the first problem, we need to develop a model of retinotopy that will include both foveal and peripheral regions. Hence, the projective Fourier transform that gives extrafoveal image representation must be complemented with a transform for the foveal image representation. Two different transforms, foveal and extrafoveal, could conform to the accumulated evidence indicating that the fovea and periphery have different functional roles in vision and may have visual processing differences [51, 52, 70]. Maybe the simplest way to construct the foveal image transform is restricting the group $\mathbf{SL}(2, \mathbb{C})$ action (which gives both image projective transformations and Möbius geometry), to Euclidean or affine subgroups. We refer to Section 3.2 in [61] where Euclidean Fourier transform is introduced in the framework of representation theory to motivate the construction of PFT and can be seen as its ‘restriction’ to the Euclidean subgroup of $\mathbf{SL}(2, \mathbb{C})$. The affine subgroup could bring the wavelet transform to supplement PFT.

The second problem involves two facts that are not compatible with each other: the computation of DPFT of an image in log-polar coordinates by FFT and the foveal split along the vertical meridian and partial crossing that reorganizes the retina outputs so that the left hemisphere destinations receive information from the right visual field, and the right hemisphere destinations receive information from the left visual field. The retina (that is, the image plane of the conformal camera) with the foveal disc removed has the visual field represented by an annulus, which under the complex logarithm $w = \ln z$ is mapped into a rectangle. In order to discretize PFT, this rectangle must be extended periodically, which forces a quasiperiodic extension of the annulus, see Eq. 21 in [58]. In our numerical experiments with the image translation by the corresponding shift property of DPFT, the image ‘disappeared’ into the foveal region of the cortical area (the foveal region of the retina) to reappear from the opposite side of the rectangle (opposite circular boundary of the annulus). Also, we need to modify the FFT to account for the global retinotopy simulated in Fig. 9 (f) by the cut-and-paste transformations. Clearly, the two problems are interdependent.

7 Conclusions

In this article we presented a comprehensive account of our approach to computational vision developed over the last decade. It was done by bringing in one place physiological and behavioral aspects of primate visual perception and the conformal camera's computational harmonic analysis with the underlying geometry. This allowed us to discuss remarkable advantages that the conformal camera possesses over other cameras used in computational vision. First, the conformal camera geometry fully accounts for the basic concepts of co-circularity and scale invariance employed by human vision system in solving the difficult intermediate-level vision problems of grouping local elements into individual objects of natural scenes. Second, the conformal camera has its own harmonic analysis—projective Fourier analysis—for image representation and processing that is well adapted to image projective transformations and the retinotopic mapping of the brain visual and oculomotor pathways. Projective Fourier analysis integrates the binocular model consisting of the head, eyes (conformal cameras), and the visual cortex into a single computational system. Based on this binocular system, we proposed a model of the perisaccadic perception, including perisaccadic mislocalizations observed in laboratory experiments. More precisely, we modeled the presaccadic activity, which, through shifts of neurons current receptive fields to their future postsaccadic locations, is thought to underlie remapping based on anticipated saccadic eye movement (efference copy). The postsaccadic activity, during which the actual integration of visual features takes place, will be considered in the next stage of our modeling.

Finally, we presented numerous challenges with the implementation of our modeling. First, the fovea-less model of the retina, based on the discrete projective Fourier transform (DPFT) of an image, must be supplemented with the foveal image transform. Second, the computations of the DPFT with a fast Fourier transform algorithm (FFT) has to be modified in order to account for the global retinotopy of the brain visual pathway.

It was observed that saccades cause, not only a compression of space, but also of time [47]. In order to preserve visual stability during the saccadic scan-path, receptive fields undergo a fast remapping at the time of saccades. When the speed of this remapping approaches the physical limit of neural information transfer, relativistic-like effects are psychophysically observed and may cause space-time compression [14, 48]. Curiously, this suggestion can also be accounted for in our model based on projective Fourier analysis since the group of image projective transformations in the conformal camera is the double cover of the group of Lorentz transformations of Einstein's special relativity.

References

- [1] J. Angulo, Polar Modelling and Segmentation of Genomic Microarray Spots Using Mathematical Morphology, Image Analysis and Stereology, 27 (2008), pp. 107-124.

- [2] S. Anstis, Picturing peripheral acuity, *Perception*, 27 (1998), pp. 817-125.
- [3] F. Attneave, Some informational aspects of visual perception, *Psychological Review*, 61 (1954), pp. 183-193.
- [4] H. Awater and M. Lappe, Mislocalization of perceived saccade target position induced by perisaccadic visual stimulation, *Journal of Neuroscience*, 26 (2006), pp. 12-20.
- [5] G. Baratoff, C. Toepfer, and H. Neumann, Combined Space-Variant Maps for Optical Flow Based Navigation, *Biological Cybernetics: Special Issue on Navigation in Biological and Artificial Systems*, 83 (2000), pp. 199-209.
- [6] M. Berger, *Geometry I*, Springer-Verlag, New York, 1987.
- [7] A. Bernardino, J. Santos-Victor, and G. Sandini, Foveated active tracking with redundant 2D motion parameters, *Robotics and Autonomous Systems*, 39 (2002), pp. 205-221.
- [8] A. Bernardino, *LogPolar Wrapper and LogPolar Mapper*, Matlab codes available over the Internet, 2002.
- [9] F. Berton, G. Sandini, and G. Metta, Anthropomorphic Visual Sensors, in *Encyclopedia of Sensors*, Vol. X, C.A. Grimes, E.C. Dickey, and M.V. Pishko, ed., American Scientific Publishers, 2006, pp. 1-16.
- [10] H. Blum, Biological shape and visual science, *Journal of Theoretical Biology*, 38 (1973), pp. 205-287.
- [11] G. Bonmassar and E. L. Schwartz, Space-Variant Fourier Analysis: The Exponential Chirp Transform, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19 (1997), pp. 1080-1089.
- [12] J. Böttger, M. Balzer, and O. Deussen, Complex Logarithmic Views for Small Details in Large Context, *IEEE Transactions on Visualization and Graphics*, 12 (2006), pp. 845-852.
- [13] D.A. Brannan, M.F. Esplen, and J.J. Gray, *Geometry*, Cambridge University Press, Cambridge, 1998.
- [14] D. Burr and C. Morrone, Time Perception: Space-Time in the Brain, *Current Biology*, 16 (2006), pp. R171-R173.
- [15] R.H.C. Carpenter, The Saccadic System: A Neurological Microcosm, *Advances in Clinical neuroscience and Rehabilitation*, 4 (2004), pp. 6-7.
- [16] C.C. Chow, D.Z. Jin, and A. Treves, Is the world full of circles?, *Journal of Vision*, 2 (2002), pp. 571-526.
- [17] I. Derakhshan, How do eyes move together? New understandings help explain eye deviations in patients with stroke, *Canadian Medical Association Journal: Neuroanatomy*, 18 (2005), pp. 171-173.

- [18] J.R. Duhamel, C.L. Colby, and M.E. Goldberg, The updating of the representation of visual space in parietal cortex by intended eye movements, *Science*, 255 (1992), pp. 90-92.
- [19] D.J. Field, A. Hayes, and R.F. Hess, Contour Integration by the Human Visual System: Evidence for a Local ‘Association Field’, *Vision Research*, 33 (1993), pp. 173-193.
- [20] G.D. Field and E.J. Chichilnisky, Information Processing in the primate Retina: Circuitry and Coding, *Annual Review Neuroscience*, 30 (2007), pp. 1-30.
- [21] B. Girard and A. Berthoz, From brainstem to cortex: Computational models of saccade generation circuitry, *Progress in Neurobiology*, 77 (2005), pp. 215-251.
- [22] P.W. Glimcher, Making choices: the neurophysiology of visual-saccadic decision making, *TRENDS in Neuroscience, Rev.*, 24 (2001), pp. 654-659.
- [23] J. Gottlieb, From a Different Point of View: Extrastriate Cortex Integrates Information Across Saccades. Focus on "Remapping in Human Visual Cortex", *Journal of Neurophysiology*, 97 (2007), pp. 961-962.
- [24] O.J. Grüsser, On the history of the ideas of efference copy and reafference, *Clio. Med.* 33 (1995), pp. 35–55.
- [25] F.H. Hamker, M. Zirnsak, D. Calow, and M. Lappe, The Peri-Saccadic Perception of Objects and Space, *PLOS Computational Biology*, 4 (2008), pp. 1-15.
- [26] F.H. Hamker, M. Zirnsak, and M. Lappe, About the influence of post-saccadic mechanisms for visual stability on peri-saccadic compression of object location, *Journal of Vision*, 8 (2008), pp. 1-13.
- [27] M. Henle, *Modern Geometries. The Analytical Approach*, Prentice Hall, Upper Saddle River, NJ, 1997.
- [28] P. Henrichi, *Applied and Computational Complex Analysis, Vol. 3*, Wiley, New York, 1986.
- [29] R.A. Herb, Harish-Chandra and his work, *Bulletin of the AMS*, 25 (1991), pp. 1-17.
- [30] D.D. Hoffman and W.A. Richards, Parts of recognition, *Cognition*, 18 (1984), pp. 65-96.
- [31] A. Hunt and P. Cavanagh, Clocking saccadic remapping, *Journal of Vision*, 8 (2008), pp. 818.
- [32] G. Jones and D. Singerman, *Complex Functions*, Cambridge University Press, Cambridge, 1987.

- [33] A.W. Knapp, Representation Theory of Semisimple Groups: An Overview Based on Examples, Princeton University Press, Princeton, NJ, 1986.
- [34] M. Kaiser and M. Lappe, Perisaccadic mislocalization orthogonal to saccade direction, *Neuron*, 41 (2004), pp. 293-300.
- [35] K. Kofka, Principles of Gestalt Psychology, Harcourt & Brace, New York, 1935.
- [36] I. Kovacs and B. Julesz, Perceptual sensitivity maps within globally defined visual shapes, *Nature*, 370 (1994), pp. 644-646.
- [37] B. Krekelberg, M. Kubischik, K.P. Hoffmann, and F. Gremmer, Neural correlates of visual localization and perisaccadic mislocalization, *Neuron*, 37 (2003), pp. 537-545.
- [38] M. Lappe, H. Awater, and B. Krekelberg, Postsaccadic visual references generate presaccadic compression of space, *Nature*, 403 (2000), pp. 892-895.
- [39] M. Lavidor and V. Walsh, The nature of foveal representation, *Nature Reviews: Neuroscience*, 5 (2004), pp. 729-735.
- [40] M. Leyton, A theory of information structure I: General principles, *Journal of Mathematical Psychology*, 30 (1986), pp. 103-160.
- [41] M. Leyton, A theory of information structure II: A theory of perceptual organization, *Journal of Mathematical Psychology*, 30 (1986), pp. 257-305.
- [42] R. Manzotti, A. Gasteratos, G. Metta, and G. Sandini, Disparity Estimation on Log-Polar Images and Vergence Control, *Computer Vision and Image Understanding*, 83 (2001), pp. 97-117.
- [43] C.D. Martina, T. Guillaume, J-F. Démonet, M. Roberts, and T. Nazira, ERP evidence for the split fovea theory, *Brain Research*, 1185 (2007), pp. 212-220.
- [44] S. Martinez-Conde, S.L. Macknik, and D. Hubel, The Role of Fixational Eye Movements in Visual Perception, *Nature Reviews: Neuroscience*, 5 (2004), pp. 229-240.
- [45] D. Melcher, Predictive remapping of visual features precedes saccadic eye movements, *Nature neuroscience*, 10 (2007), pp. 903-907.
- [46] E.P. Merriam, C.R. Genovese, and C.L. Colby, Remapping in human visual cortex, *Journal of Neurophysiology*, 97 (2007), pp.1738-1755.
- [47] M.C. Morrone, J. Ross, and D. Burr, Saccadic eye movements cause compression of time as well as space, *Nature Neuroscience*, 8 (2005), pp. 950-954.

- [48] M.C. Morrone, J. Ross, and D. Burr, Keeping vision stable: Rapid updating of spatiotopic receptive fields may cause relativistic-like effect, In Problems of Space and Time in Perception, R. Nijhawan, ed., 2008, Cambridge.
- [49] J.L. Mundy and A. Zisserman, eds., Geometric Invariance in Computer Vision, MIT Press, 1992.
- [50] J. Najemnik and W.S. Geisler, Optimal eye movement strategies in visual search, *Nature*, 434 (2005), pp. 387-391.
- [51] Y. Petrov, M. Carandini, and S. McKee, Two distinct Mechanisms of Suppression in Human Vision, *Journal of Neuroscience*, 25 (2005), pp. 8704-8707.
- [52] J. Prado, S. Clavagnier, H. Otzenberger, C. Scheiber, H. Kennedy, and M-T. Perenin, Two Cortical Systems for Reaching in Central and Peripheral Vision, *Neuron*, 48 (2005), pp. 849-858.
- [53] J. Ross, M.C. Marrone, and D.C. Burr, Compression of visual space before saccades. *Nature*, 386 (1997), pp. 698-601.
- [54] E. L. Schwartz, Computational anatomy and functional architecture of striate cortex, *Vision Research*, 20 (1980), pp. 645-669.
- [55] L.S. Shapiro, *Affine Analysis of Image Sequences*, Cambridge University Press, 1995.
- [56] M. Sigman, G.A. Cecchi, C.D. Gilbert, and M.O. Magnasco, On a common circle: Natural scenes and Gestalt rules, *Proceedings of the National Academy of Sciences of the United States of America*, 98 (2001), pp. 1935-194.
- [57] N. Tamayo and V.J. Traver, Entropy-Based Saliency Computation in Log-Polar Images, *Proceedings of the International Conference on Computer Vision Theory and Applications*, 2008, pp. 501-506.
- [58] J. Turski, Projective Fourier analysis in computer vision: Theory and computer simulations, in *Vision Geometry VI*, SPIE Vol. 3168, R.A. Malter, A.Y. Wu, and L.J. Latecki eds, 1997, pp. 124-135.
- [59] J. Turski, Harmonic analysis on $\mathbf{SL}(2, \mathbb{C})$ and projectively adapted pattern representation, *Journal of Fourier Analysis and Applications*, 4 (1998), pp. 67-91.
- [60] J. Turski, Projective Fourier analysis for patterns, *Pattern Recognition*, 33 (2000), pp. 2033-2043.
- [61] J. Turski, Geometric Fourier Analysis of the Conformal Camera for Active Vision, *SIAM Review*, 46 (2004), pp. 230-255.
- [62] J. Turski, Geometric Fourier Analysis for Computational Vision, *Journal of Fourier Analysis and Applications*, 11 (2005), pp. 1-23.

- [63] J. Turski, Computational Harmonic Analysis for Human and Robotic Vision Systems, *Neurocomputing*, 69 (2006), pp. 1277-1280.
- [64] J. Turski, Harmonic Analysis for Cognitive Vision: Perisaccadic Perception, in *Human Vision and Electronic Imaging XIV*, SPIE Vol. 7240, B.E. Rogowitz and T.N. Pappas, eds, 2009, pp. 72401A-1 - 72401A-5.
- [65] S. Ullman, *Higher-Level Vision: Object Recognition and Visual Cognition*, MIT Press, 1996.
- [66] R. VanRullen, A simple translation in cortical log-coordinates may account for the pattern of saccadic localization errors. *Biological Cybernetics*, 91 (2004), pp. 131-137.
- [67] C.F.R. Weiman, *Log-polar Binocular Vision System*, Transition Research Corporation: NASA Phase II SBIRFinal Report, 1994.
- [68] M. Wertheimer, *Laws of Organization in Perceptual Forms*, Harcourt, Brace & Jovanovitch, London, 1938.
- [69] R.H. Wurtz, Neuronal mechanisms of visual stability, *Vision Research*, 48 (2008), pp. 2070-2089.
- [70] J. Xing and D.J. Heeger, Center-surround interaction in foveal and peripheral vision, *Vision Research*, 40 (2000), pp. 3065-3072.
- [71] Y. Yarbus, *Eye Movements and Vision*, Plenum Press, New York, 1967.