Geometric Fourier Analysis in Computational Vision

Biological and Mathematical Background

Jacek Turski

University of Houston-DTN

9th International Conference on Cognitive and Neural Systems

Boston University, May 18-21, 2005

Outline

I. The Brain Visual pathway

1. Biological Facts
2. Local V1 Topography
3. On Global Topography
4. Cyclopean Symmetry

II. GFA for Computational Vision

5. Mathematical Background
6. The Conformal Camera
7. Projective Fourier Analysis

III. On Binocular Image Processing

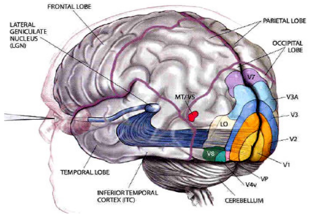
8. Cortical Image Processing
9. Binocular Vision
   - Cortical Images
   - Horopter

References

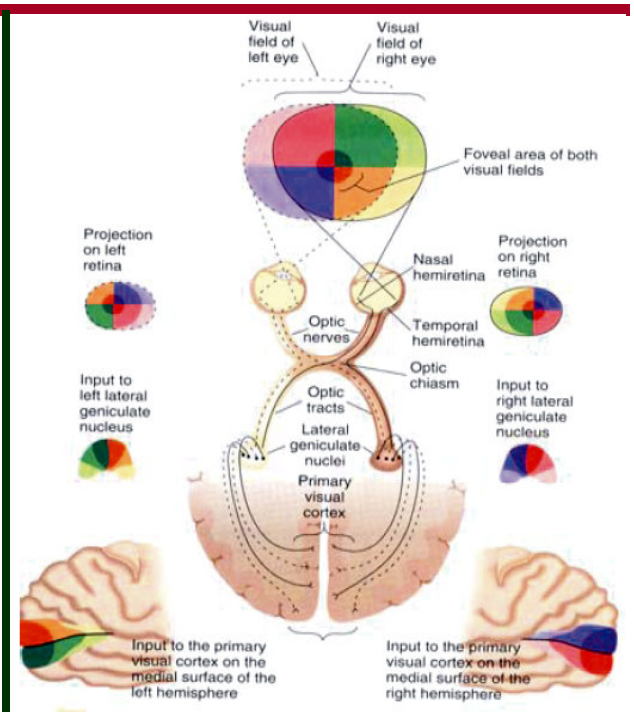
[1] M. Balasubraminian, J. Polimeni and E.L. Schwartz, The V1-V2-V3 complex: Quasiconformal dipole maps in primate striate and extra-striate cortex, Neural Networks, 15, 1157-1163, 2002.

[2] J. Turski, Geometric Fourier Analysis for Computational Vision, *JFAA*, 11, 1-23, 2005.

***Acknowledgement*: Some pictures have been borrowed from Eric Schwartz web page http://www.cns.bu.edu/~eric**



1. Biological Facts

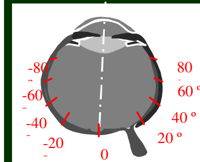


RETINOTOPY

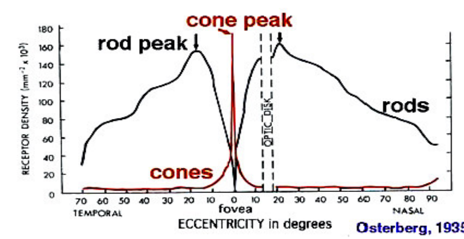
Image is transmitted from the retina to the visual cortex along the visual pathway in a precise, although changing, retinotopic arrangement, resulting in many maps of features (topography, orientation, eye dominance, motion direction, etc) superimposed in visual cortex.

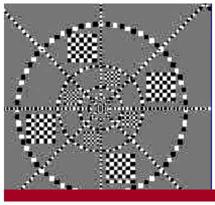
TOPOGRAPHIC MAGNIFICATION

50% of the primary visual cortex (V1) is used to process input from the foveal region of 4% of the retinal area

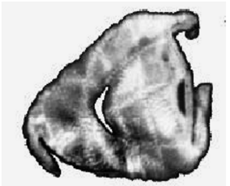


Cones are concentrated in the fovea.
Rods occur only in the periphery





2. Local V1 Topography



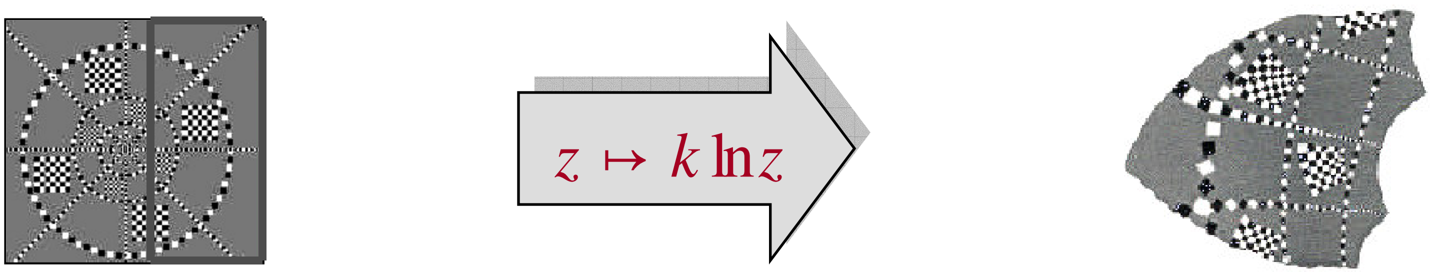
With the ganglion density $\rho(r) = \frac{\mu}{r^2}$ (away from the central part of the fovea) and a uniform V1 packing call density c ,

$$\rho(r) r dr d\theta = c du d\theta$$

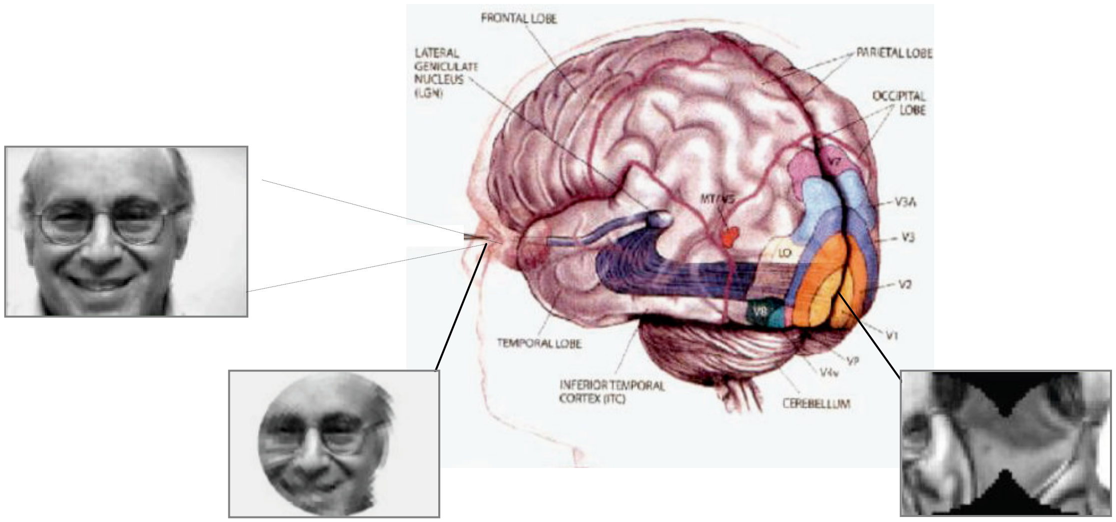
gives the coordinate mapping

$$(k \ln r, \theta) \equiv k \ln z$$

which defines the local V1 topography



3. On Global Topography



After the work of Eric Schwartz and his group:

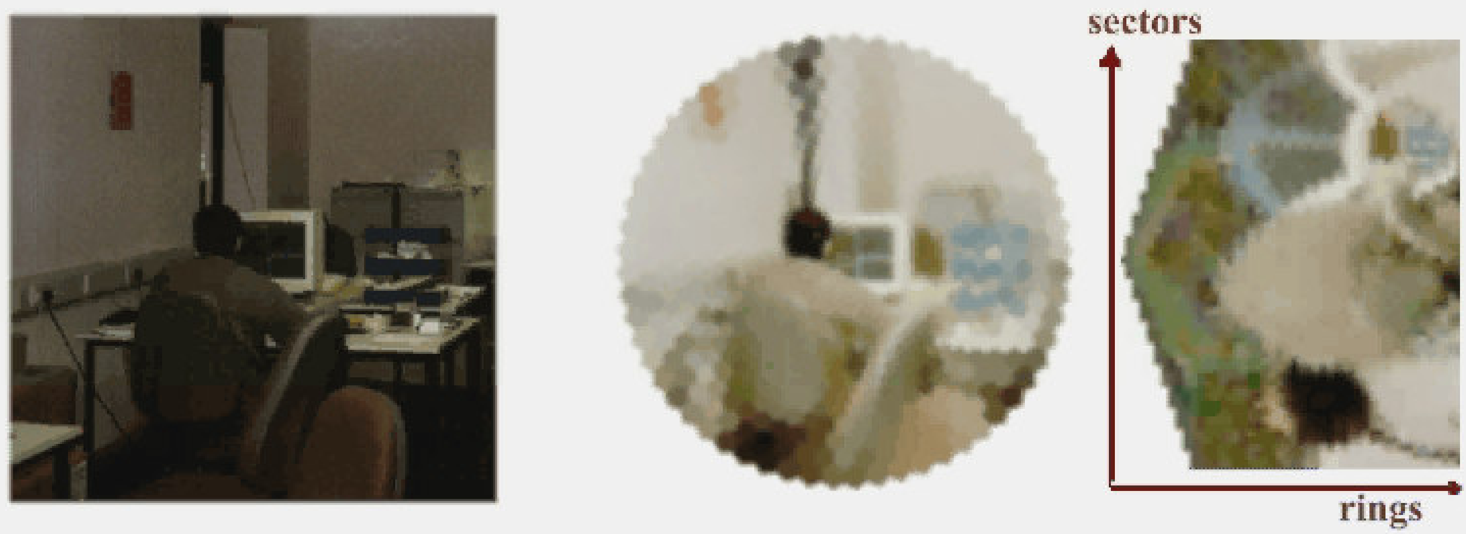
(1) **The mapping** $w = k\ln(z+a)$ **is an accepted approximation of the V1 topographic structure.**

(2) **A better model represents a full topography in terms of the mapping**

$$w = k\ln\frac{z+a}{z+b}$$

4. Cyclopean Symmetry

$$\ln re^{i\theta} = \ln r + i\theta = (\ln r, \theta)$$
$$\ln\left(r\varrho e^{i(\theta+\phi)}\right) = \left(\ln r + \ln \varrho, \theta + \phi\right)$$



Input Image — Retinal Image — Log-polar Image

Author: Herman Gomes

5. Mathematical Background

Geometric Fourier Analysis (GFA) – Harmonic Analysis associated with the **Unitary Group Representations**

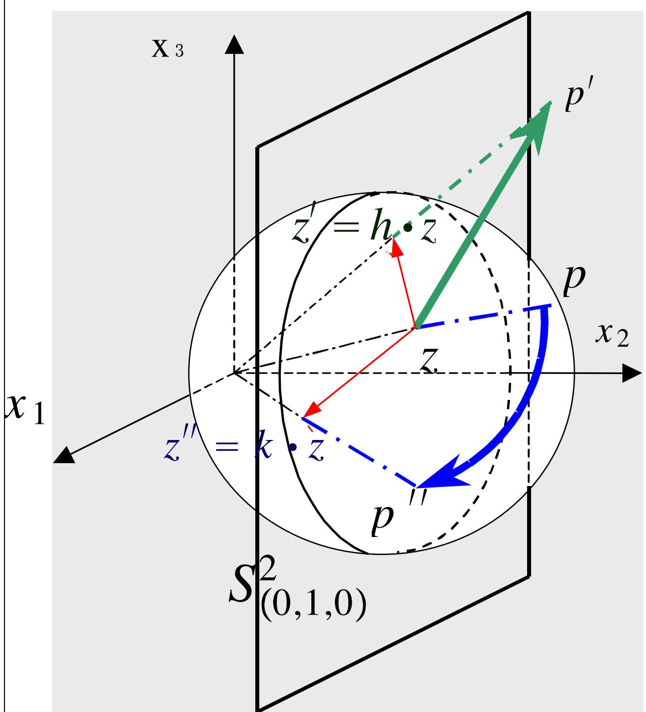
Note: The Classical Fourier Transform is associated with the unitary representations of the group of translations.

We constructed Computational Harmonic Analysis of the group SL(2,C) -- the group that provides image projective transformations in the Conformal Camera. Remarkably, the resulting image representation is also *well adapted to retinotopic mapping of the brain's visual pathway*.

6. The Conformal Camera

The Camera

Image plane: $(x_1, 1, x_3) \equiv x_3 + ix_1$



Geometry of the Image Plane

Complex Projective transformations

$$z \mapsto g \cdot z = \frac{dz+c}{bz+a}; \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbf{SL}(2,\mathbb{C})$$

are finite iterations of $h \cdot z = \frac{\delta^{-1} z + \gamma}{\delta}$ and $k \cdot z = \frac{\alpha z + \beta}{-\overline{\beta} z + \overline{\alpha}}$

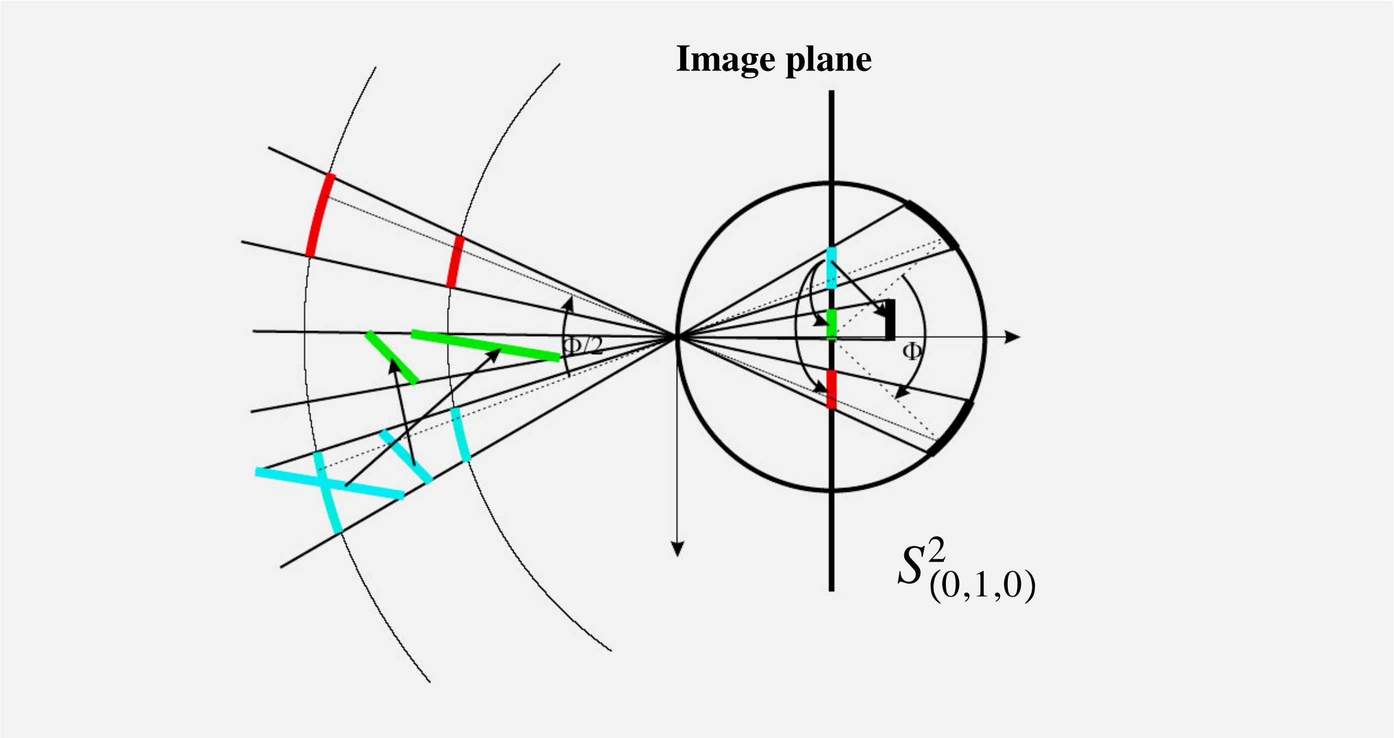
Complex Projective Geometry

The embedding: $\begin{pmatrix} x_2 + iy \\ x_3 + ix_1 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \in \mathbb{C}^2$

$g\begin{pmatrix} z_1 \\ z_2 \end{pmatrix}$ induces $g \cdot z$ on slopes of lines $z_2 = z\, z_1$

Image conformal – projective transformations: $f(z) \mapsto f(g^{-1} \cdot z)$

6. The Conformal Camera



The conformal camera reduces the projective degrees of freedom to a minimal set of image projective transformations.

We will show later that this set of projective transformations is relevant to the way the human vision system acquires the understanding of 3D scenes.

7. Projective Fourier Analysis

Geometric Fourier Analysis of the Conformal Camera

Given the action of SL$(2,\mathbb{C})$: $\widehat{\mathbb{C}} \ni z \mapsto g\cdot z \in \widehat{\mathbb{C}}$, **decompose a pattern** $f \in L^2(\mathbb{C})$ **in terms of the irreducible unitary representations of SL**$(2,\mathbb{C})$ **present in** $L^2(\mathbb{C})$.

The Projective Fourier Transform (PFT)

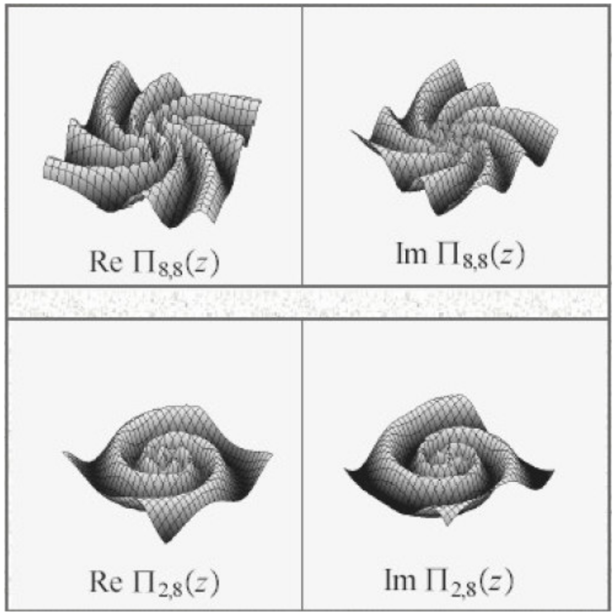
$$\widehat{f}(s,k) = \frac{i}{2}\int f(z)|z|^{-is-1}\left(\frac{z}{|z|}\right)^{-k} dz\,d\overline{z}$$

The Inverse PFT

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

7. Projective Fourier Analysis

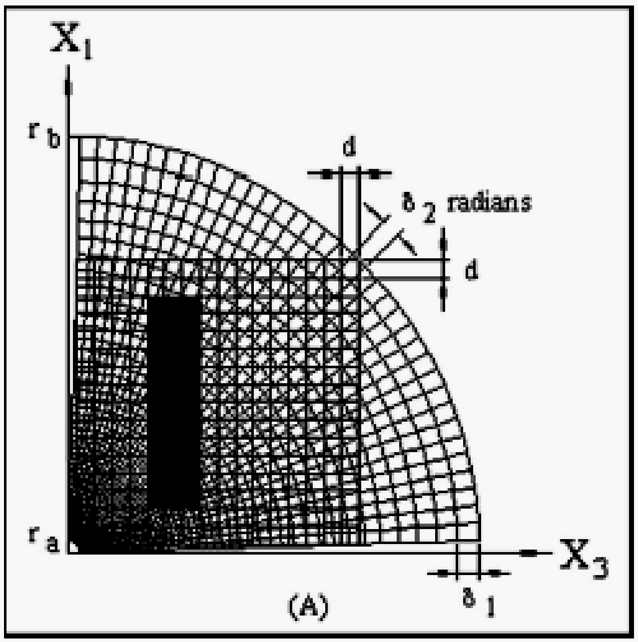
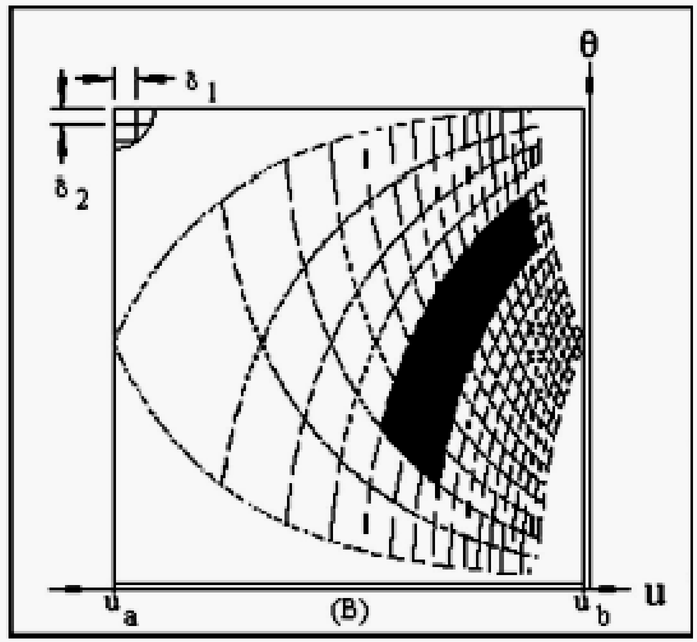
The Borel Characters $\Pi_{k,s}(B) = |z|^{-is}(|z/|z|)^{-k}$



Gauss Decomposition: $\mathbf{SL}(2,\mathbb{C}) \doteq \widetilde{\mathbf{N}}\mathbf{B}$, where $\widetilde{\mathbf{N}} = \begin{pmatrix} 1 & 0 \\ \gamma & 1 \end{pmatrix}$ and $\mathbf{B} = \begin{pmatrix} z & \beta \\ 0 & z^{-1} \end{pmatrix}$, implies that the Borel subgroup $\mathbf{B}$ exhausts the projective part of $\mathbf{SL}(2,\mathbb{C})$.

8. Cortical Image Processing

- **PFT** is the standard **FT** in log-polar (u, θ) where $u = \ln r$
- We discretize **PFT** and compute it by FFT in log-polar

	
Nonuniform Log-polar Sampling **– Retinal Image**	**Inverse DPFT in Log-polar Coordinates** **– Cortical Image**

8. Cortical Image Processing

The Discrete PFT and Its Inverse (Log-polar)

$$\hat{f}_{m,n} = \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} f_{k,l} e^{u_k} e^{-i2\pi u_m k/T} e^{-i2\pi \theta_n l L}$$

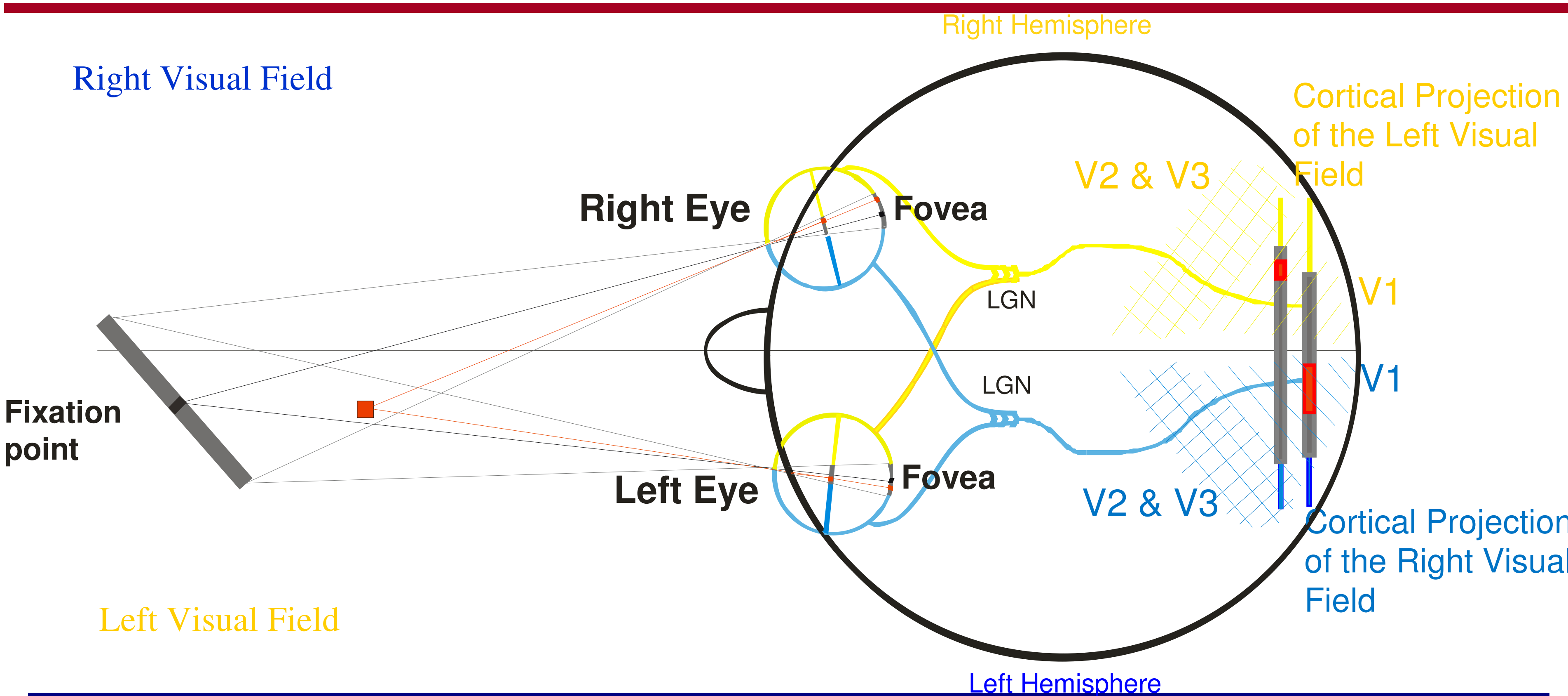
$$f_{k,l} = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \hat{f}_{m,n} e^{-u_k} e^{i2\pi u_m k/T} e^{i2\pi \theta_n l L}$$

Projective "covariance" of PFT

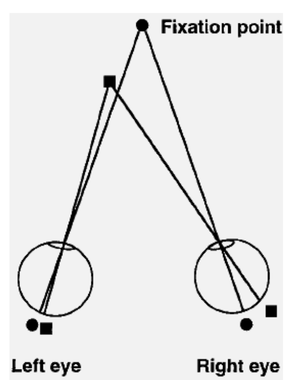
$$f'_{m,n} = \frac{1}{MN} \sum_{k=0}^{M-1} \sum_{l=0}^{N-1} \hat{f}_{k,l} e^{-u'_{m,n}} e^{i2\pi u'_{m,n} k/T} e^{i\theta'_{m,n} l L}$$

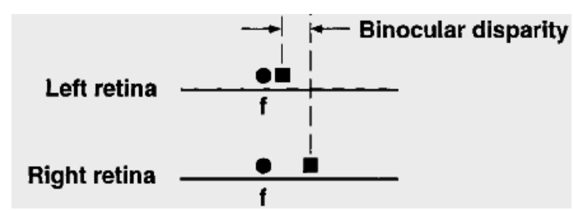
$$u'_{m,n} + i\theta'_{m,n} = \ln z'_{m,n} = \ln(g^{-1} \cdot z_{m,n})$$

9. Binocular Vision



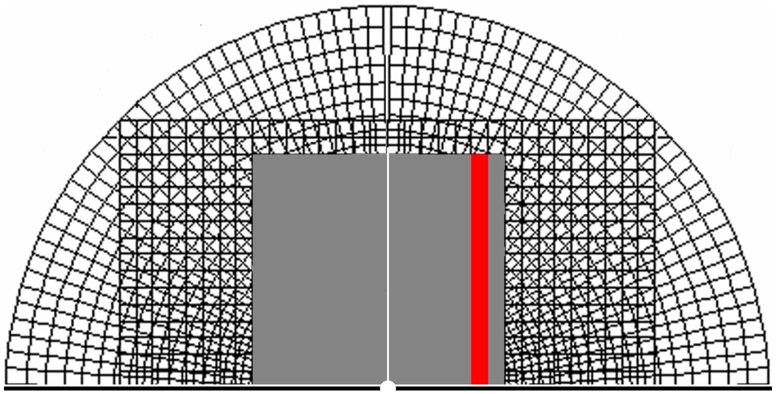
Binocular Disparity



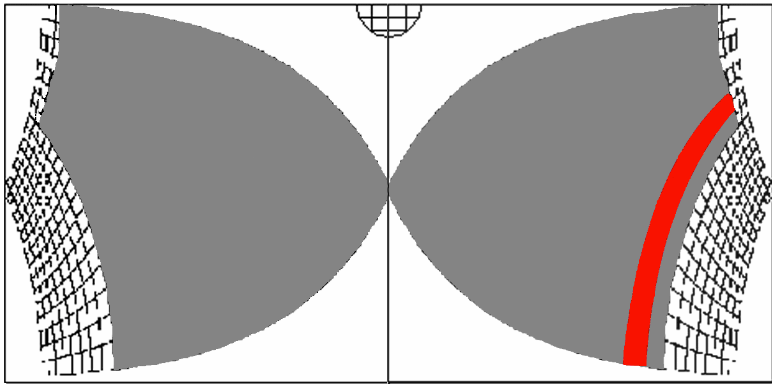


- Cortical Images

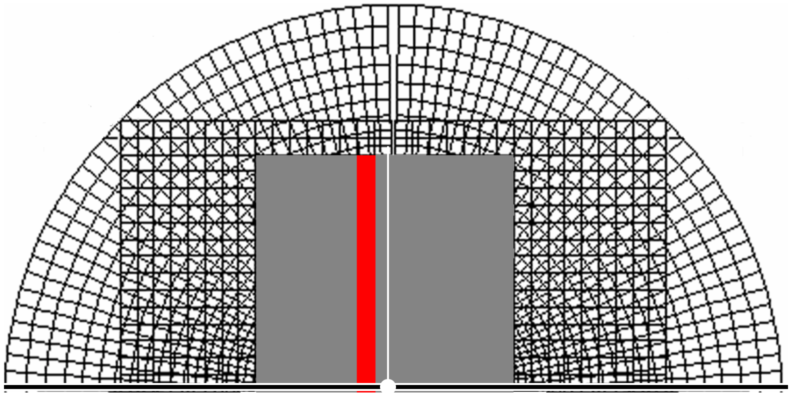
The Projection onto the Right Eye Retina



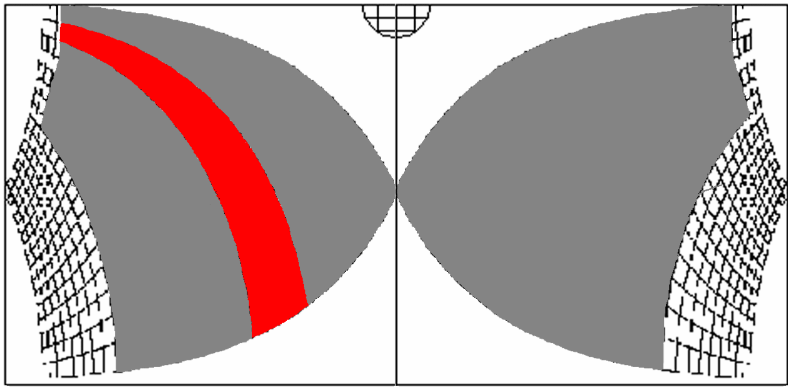
Cortical Image from The Right Eye



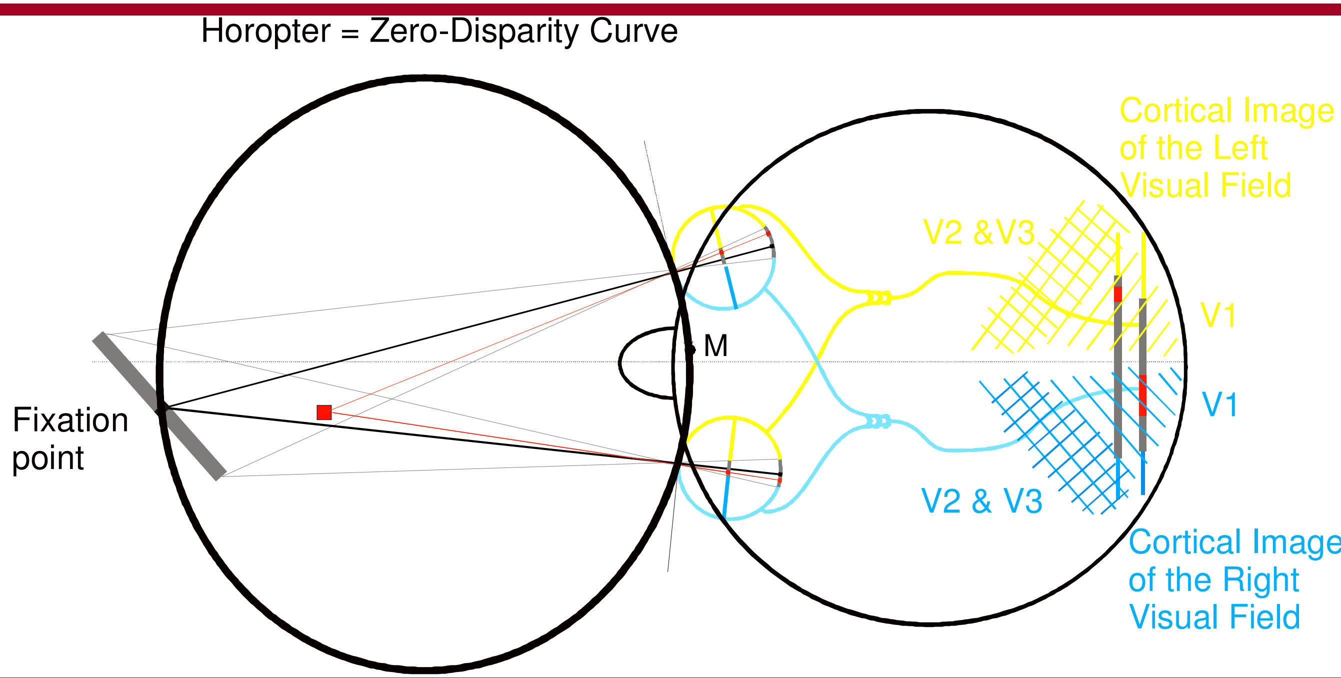
The Projection onto the Left Eye Retina



Cortical Image from The Left Eye

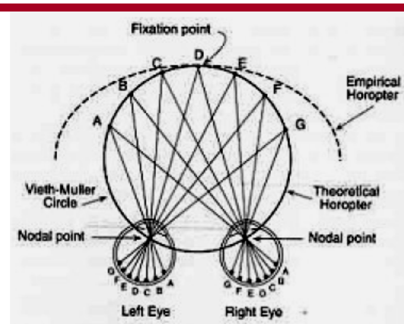


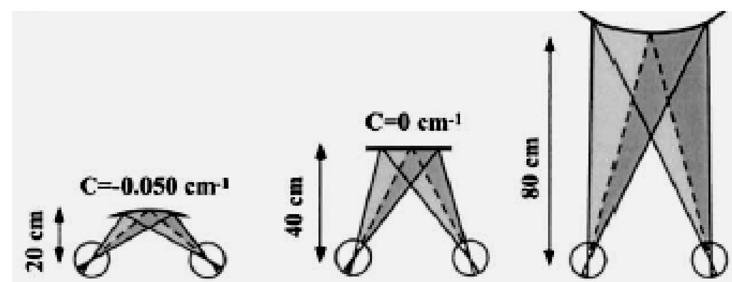
- Horopter



Theorem. *For binocular system with the conformal cameras, the horopters in the visual plane are conics, closely matching the empirical horopters.*

With retinas acting as if they were spheres, the geometric horopters are circles, known as the Vieth-Muller circles.





Conclusions

Because:

(1) the camera with silicon retina sensors produces the image similar to the topographic image in V1,

(2) the line singularity most likely exists in the fovea,

the head–eye–visual cortex integrated system makes possible an efficient, biologically realistic computational approach to binocular vision for robotic systems

The End

APPENDIX

Starting in [SM] with the space S of 2D shapes given by the set of all closed smooth curves in the plane, it is shown that a shape can be uniquely represented by a diffeomorphism of the unit circle S^1 modulo a Möbius transformation preserving the circle, called the "fingerprint" of the shape.

Then, using "welding" problem of "sewing" together conformally the interior of the circle with its exterior along the given "fingerprint", the classification the shapes up to an arbitrary Möbius transformation in $\mathrm{PSL}(2,\mathbb{C})$ is obtained.

[SM] *Sharon, E. and Mumford, D. 2D-Shape Analysis using Conformal mapping, IEEE on Conference CVPR Vol. 2, 2004.*

CONCLUSION. Since the conformal camera generates image projective transformations by the action of the Möbius group $\mathrm{PSL}(2,\mathbb{C})$, our work implies that the classifications of 2D shapes in [SM] is both *projectively* and *retinotopically invariant.*