Neurogeometry of vision. Lecture II. Main principles of organisation. Neurons. Eye. Retina.

Dmitri Alekseevsky

Institute for Information Transmission Problems, Moscow,Russia and
Faculty of Science University of Hradec Kralove,
Rokitanskeho 62, Hradec Kralove, 50003, Czech Republic
Spring School in Topology and Geometry, Gradec Karlove,
May 2018

Gradec Kralove, May 2018

I. Main principles of organisation of visual system.

II. Visual neurons

1, Visual neurons as filters.

2. Gauss, Marr and Gabor filters. Definition of simple and complex neurons by M. Hansard and R. Horaud.

3. Set of Gauss filters as homogeneous convex cone and Information Geometry.

4. Multiscale approximation of differential geometry by Koenderink.

III. Eye

1. Eye as an optical device. Central projection to retina.

2. Helholtz definition of straight line.

3. Eye as a rotation rigid body. Fixation eye movements.

4. Stochastic model of eye movements and diffusion geometry.

5. Problem of stability.

IV. Retina

1. Architecture of retina. Structure of retina images, metric.

2. Retina as information processing system. Problem of input function.

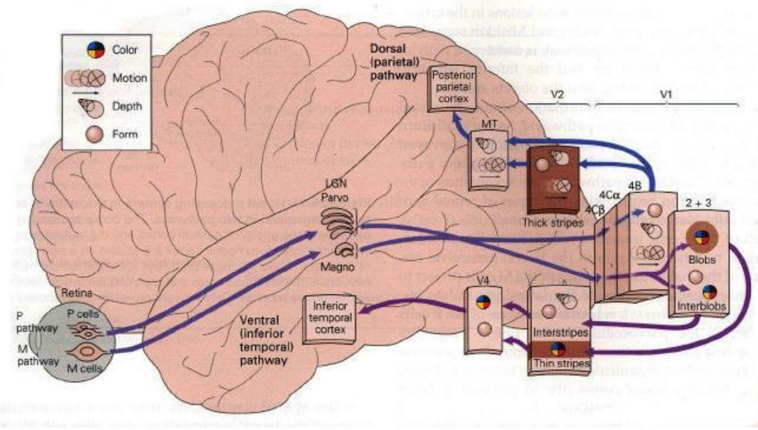
3. Aim of information processing in retina .

Main principles of organisation of visual system

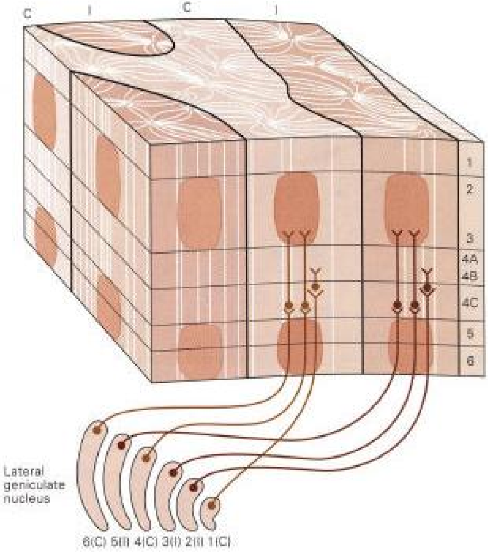
Aim of the vision. Similarity between vision and DG.
What visual information comes to retina and is detected by receptors?

Principle 0. Light, coming through eye to retina, is the only source of information for visual system.

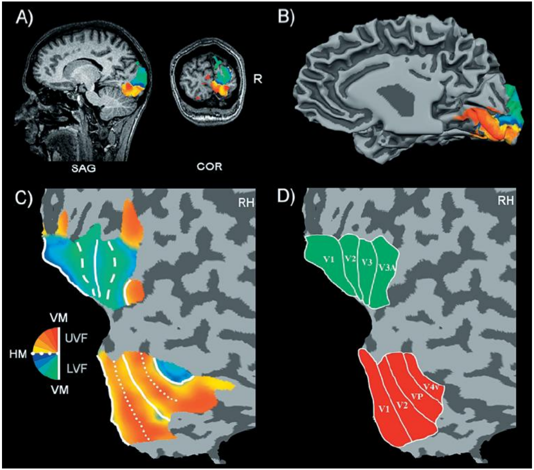
Hierachical organization of Visual Systems



Retinotopic Map to VI



Location of visual cortex



Light in approximation of geometric optics

Light is described in terms of the symmetric pseudo-Kähler space $L(E^3) = TS^2 = T^*S^2$ of straight lines.
Light travels along a line $\ell \in L(E^3)$ with energy density $I(\ell)$ (the average value of the square norm of electric filed).
We ignore the wave length (color) and polarization of light. Then all information for eye is coded in the energy function
$I : L(E^3) \to \mathbb{R}$ of light.
Maxwell electrodynamics : light is a superposition of plane waves. Plane waves is associated with a shear-free congruence of isotropic lines in the Minkowski space and with a complex surface in the Penrose twistor space $\mathbb{C}P^3$ (Kerr).
QED (quantum electrodynamics).

Energy function associated with a surface $\Sigma \subset E^3$

A point $A \in \Sigma$ is a source of (reflected) light which travels along rays in all directions outside of the surface.
In most cases . the energy density $I(AX)$ of a ray from a point $A \in \Sigma$ depends only on the source $A \in \Sigma$(**diffuse reflection**) and is constant in time. Hence the (density of) energy of light eminated from Σ is described by a function $I : \Sigma \to \mathbb{R}$ (energy function).

Part I. Main principles of organization of visual system

a) All visual information is extracted from the light, coming to retina, more precisely, from the density of energy of light falling to retina.
b)Hierarchical structure of VS : Eyes-retina-LGN-cortex V1-V2-V3-V4-V5-...
c)Strong feedback.
d)Visual system do not record the value of the input function , but its change.
e) Contours are the main objects of recognition in early vision.
e) Locality of information processing. Visual neurons as generalized functions (filters).
d)The aim of information processing in retina is a regularization and contourization of the input function.
e) There is a retinotopic conformal map from retina to LGN and cortex VI.

Part II. Visual neurons Visual neurons as filters (generalized functions)

Receptive field (RF) and receptive profile (RP). Gauss, Marr and Gabor filters. Principle of locality. Idea of fibre bundle (Hubel and Wiezel) and problem of local parameters. Orientation and spatial frequency.

Local information about image is coded in visual neurons.

Excitation of a visual neuron depends on the values of the input function I in the receptive field (RF) $D \subset R$ of the neuron (a small domain of the retina) .

A linear neuron acts as a linear functional (filter) of the form

$$T_F : I \mapsto \int_D F(z)I(z)dxdy, \ z = (x, y)$$

which calculate the "mean value "of the energy function I in D with some weight $F(z)$ called the receptive profile (RP).

Gauss filter

Gauss filter is a filter, whose RP is the Gauss function

$$G = G_{z_0}^{\sigma}(z) = \frac{1}{\sqrt{2\pi\sigma}} \exp(-\frac{|z - z_0|^2}{2\sigma^2}).$$

When the standard deviation $\sigma \to 0$, the Gauss functional tends to the Dirac delta function, i.e. the functional $\delta(I) = I(z_0)$.
So the Gauss filter can be considered as a σ-approximation of the delta function , which calculate the value of a function in a point.

Manifold of (multivariate) Gauss distributions in $\mathbb{R}^n$

The standard Gauss distribution with mean value $\mu = 0$ and standard deviation 1 is defined by the volume form

$$\gamma_0 dx = (2\pi)^{n/2} \exp(-1/2|x|^2) dx.$$

The manifold $\mathcal{N}_n$ of Gauss distribution is the orbit $\mathcal{N}_n = Aff_n^+(\gamma_0 dx) = Aff_n^+ / SO_n \cdot \mathbb{R}^n$ of this point under the natural action of the unimodular affine group $Aff_n^+ = SL_n \cdot \mathbb{R}^n$, given by

$$T_{A,a}\gamma_0 dx = (2\pi)^{n/2} (\det A)^{-1} \exp(-1/2|A^{-1}x - a|^2) dx.$$

This action is extended to the natural action of the group SL_{n+1} in the cone $\mathcal{P}_{n+1}$ of positively defined matrices under the identification $\mathcal{N}_n = \mathcal{P}_{n+1}$, given by

$$(T^{-1}_{(A,a)})^* \gamma_0 dx = (\det A)^{\frac{-2}{n+1}} \begin{pmatrix} AA^t + aa^t & a \\ a^t & 1 \end{pmatrix}.$$

Derived filters

Let X be a vector field. Then the functional $T_{X \cdot G}$ with RP $X \cdot G$ is called the derivative of the Gauss filter in direction X.
Integration by part shows that the limit for $\sigma \to 0$ of $T_{X \cdot G}$ is the functional
$X_{z_0} : I \to -(X \cdot I)(z_0)$ (i.e. the tangent vector $-X_0$ at z_0).
So $T_{X \cdot G}$ is a sigma approximation of the tangent vector $-X_{z_0}$.
Similarly, the functional with RP $Y \cdot (X \cdot G)$ is a sigma approximation of the second order operator $Y \cdot (X \cdot I)(z_0)$ at the point z_0.

Simple visual cells of order k (M. Hansard and R. Horaud)

According to M. Hansard and R. Horaud , a simple visual cell of order k is defined functionally as the filter associated with k-th directional derivatives

$$X_1 \cdot X_2 \cdot \dots \cdot X_k \cdot G$$

of the Gauss function (or their linear combinations). It is sigma approximation of the differential operator $(-1)^k X_1 \circ \dots \circ X_k|_{z_0}$. M. Hansard and R. Horaud give also a general definition of complex visual cells as a composition of simple cells, following an idea by D. Hubel and T. Wiesel.

Gabor filter and simple cells of VI cortex

Gabor filter is defined by RP which is "Gauss function modulated by cos or sin ". More precisely,the standard (even and odd) Gabor filter has RP

$$Gab_{+}^{\sigma}(z) = G_0^{\sigma}(z)\cos y, \;\; Gab_{-}^{\sigma}(z) = G_0^{\sigma}(z)\sin y.$$

Arbitrary Gabor profile is defined by application to $Gab_{\pm}^{\sigma}$ a transformation from the special isometry group $SE(2) = SO_2 \cdot \mathbb{R}^2$. So the set of (even or odd) Gabor filters on the plane is parametrized by the isometry group $E(2) = SO(2) \cdot \mathbb{R}^2$ and the parameter σ (the standard deviation).
Appropriate simple filter of M. Hansard and R. Horaud of degree two and one have RP which are close to the RP of (even and odd) Gabor filter.
25% of cells in VI cortex acts (in the first approximation) as Gabor filters. Due to anisotropy of Gab they fire when the the tangent direction of the contour which intersects the RF is (approximately) orthogonal to the distinguished direction y.

Inner parameters ("engrafted variables"by Hubel) and idea of fibre bundle

Hubel remarked that the activity of a visual neuron depends not only on the coordinates $z = (x, y)$ of the RF, but on many other internal parameters, which may be considered as coordinates of total space of some fibre bundle. The complete set of inner parameters, which are relevant for visual system, is not known. N.V.Swindale ("How many maps are in visual cortex 2000.) estimates this number as 6-7 or 9-10. It is closed to the number of inner parameters of the physical model of our Universe. (The dimension of space-time in Supergravity is $D \leq 11$ and the space of inner parameters has dimension $D - 4 \leq 7$.)
In neurophysiology, the most important internal parameters are orientation and spatial frequency. Other relevant parameters are 3 parameters of the color space, curvature, temporal frequency, ocular dominance, disparity etc.

What is the input function on the retina : the energy function I, 1-form dI or 1-distribution $D = [dI] = \mathrm{ker} dI$? Basic global objects of early vision

Basic global objects of early vision are
contours = curves on the retina $R \subset S^2$ which are level sets of the input function with "big"gradient Thet are the images of edge (boundary of objectz of external world.)

Basic infinitesimal objects of early vision

First order infinitesimal approximation of a non parametrized curve (contour) is a tangent line. The space of such not oriented infinitesimal contours is the contact bundle

$$PTS^2 = PT^*S^2 = \{(x, y, p = \frac{dy}{dx})\}$$

with the contact structure $\ker(dy - pdx)$.
The space of oriented infinitesimal contours is the sphere bundle $S^1(S^2)$. Sections of these two bundles are 1-dimensional non oriented and , respectively, oriented distributions.
An infinitesimal contour of order k is a k-jet of a contour. The space of such objects can be identified with the sp;ace of k-jets

$$J^k(\mathbb{R}, \mathbb{R}) = \{x, y, p = y', \cdots, y^{(k)}\}.$$

A k-th order infinitesimal part of an input function $I \in \mathcal{F}(S^2)$ is the k-jet $j^k_z(I) \in J^k(S^2, \mathbb{R})$. For $k = 1$, $J^1(S^2, \mathbb{R}) = \mathbb{R} \times T^*S^2$.

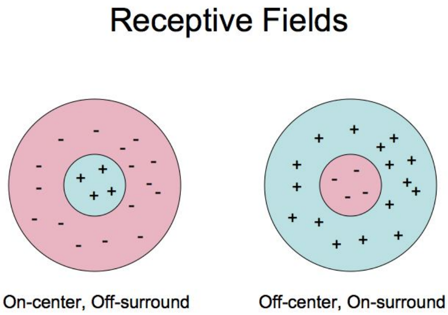
Koenderink's Multiscale Geometry of Image Processing

From algebraic point of view (used in Algebraic Geometry and Quantum Physics), DG can be considered as geometry associated with algebra of functions $\mathcal{F}(M)$.

Jan Koenderink defined Multiscale Geometry as geometry which considers simultaneously sigma approximations to DG for all resolution parameter σ . In his seminal paper (Structure of Images, 1984), he showed that any image is embedded in a one-parameter family of derived images, parametrized by resolution and the structure of family is governed by the heat equation.

Marr filters and isotropic cells by Kuffler

Steven Kuffler (1952) was the first who detected a response of a visual neuron in retina on a stimulus and describe the (isotropic) RF of cells and RP. David Marr (1975) showed that the Laplacian of Gauss ΔG (which is a second order simple filter with isotropic profile) gives a good approximation to RP of Kuffler cells. He explained that system of such filters produces a regularization and contourization of the input function I. This is the aim of data processing in retina.



On and Off cells

Action of Marr filters

Eye as an optical device.Central projection to retina.

The eye is a transparent ball B^3 together with a lens L with center F of the boundary sphere which focuses light rays to the retina $R \subset S^2$.
The lens is formed by the cornea and the eye crystal.
The light emitted by a surface Σ and coming to retina defines a map

$$\pi : \Sigma \to R, \ A \mapsto \bar{A} = \ell_{AF} \cap R$$

from Σ to retina R which depends on the position of the eye ball $B(OF)$. Here $\bar{A}$ means the second point of the intersection of the line ℓ_{AF} with retina.
Denote by $D_A = \{Y \in S^2(A), \ell_{AY} \cap L \neq \emptyset\}$ the domain on the unite sphere $S^2(A)$ with center at $A \in \Sigma$ which is the intersection with $S^2(A)$ the cone of beams from $A \in \Sigma$ which come to the eye lens L.

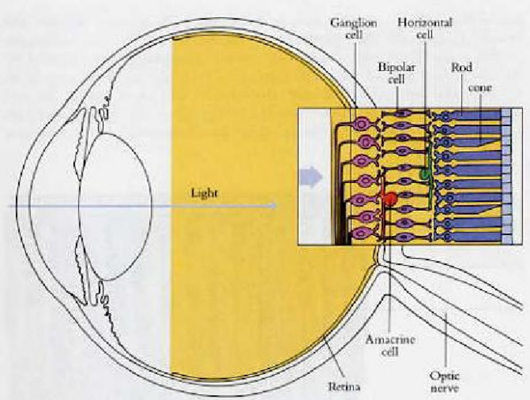
Then the energy of light coming from a point $A \in \Sigma$ to its image $\bar{A} \in R$ on retina is equal $I(A) \cdot area(D_A)$. where $I(A)$ is the density of energy of a beam emitted from A.
So the function $I : \Sigma \to \mathbb{R}$ of density of energy of light , emitted from the surface determines the function

$$I_R : R \to \mathbb{R}. \bar{A} \mapsto I_R(\bar{A}) = I(A) \cdot area(D_A)$$

of density of energy of lights emitted from the surface Σ and coming to retina.
The function $I = I_R$ on the retina is called the energy function or the input function on the retina.

Eye as an optical device

Eye



Central projection of a surface $M \subset E^3$ to a sphere S^2

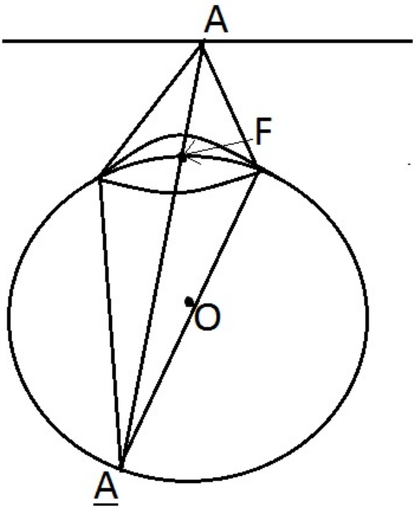
The central projection of a surface M to a sphere S^2 ()"retina"with center $F \in S^2$ is defined by

$$\varphi : M \ni A \to \bar{A} = \ell_{AF} \cap S^2 \subset S^2$$

where $\ell_{AF} \cap S^2$ is the second point of intersection of the line ℓ_{AF} which goes through the center F ("the center of eye lens") with the sphere ("retina").

Central projection of a surface to retina

Central projection



Definition of straight line by Helmholtz

The image of a straight line ℓ on retina $R \subset S^2$ is a conformal circle, i.e. the intersection of the eye sphere with a plane.
If the gaze (the line of sight (OF) where O is the center of eye ball) moved along a line ℓ , the image $\varphi(\ell)$ does not change.
(This is the physiological definition of a straight line by Helmholtz).
If $M = \Pi$ is the frontal plane (orthogonal to the line of sight), then the central projection is a conformal map.

EYE AS A ROTATING RIGID BODY.Donder's and Listing's laws. Eye movements:tremor, drift and saccades. Problem of stability.

Eye is a rigid ball B^3_O which can rotate around the center O w.r.t. three mutually orthogonal axes i, j, k.
The center $F_0 \in B^3_O$ of the eye lens belongs to the eye sphere $S^2_O = \partial B^3_O$ with center at point O and the retina $R \subset S^2_O$ occupies a big part of S^2_O. For a fixed position of head, there is a privilege initial position $B(OF_0)$ of the eye ball corresponding to the standard (frontal) direction (OF_0) of the gaze.

Donder's and Listing's laws

Donder's law (1846)(No twist). If the head is fixed, the result of a movement from position $B(OF_0)$ to a new position $B(OF)$ is uniquely defined by the gaze OF and do not depend on previous movement.

Mathematically, it defines a section $s : S^2 \to SO_3$ of the frame bundle $SO_3 \to S^2 = SO_3/SO_2$ such that a curve $\gamma(t)$ in S^2 has lift $\tilde{\gamma}(t)$ to the group of rotations SO_3.

Due to this law, a movement of the eye is determined by a curve on the eye sphere.

Listing's law (1845) The movement from $B(OF_0)$ to $B(OF)$ is obtained by rotation with respect to the axe $\vec{OF_0} \times \vec{OF}$.

The curve in SO_3 is the parallel lift of the initial frame along the arc $F_0F \subset S^2$.

Eyes movements. Tremor, drift, microsaccades and macrosaccades

Eyes participate in different involuntary types of movements. We consider only two types of movements:
fixation eye movements when the gaze is "fixed"and
macrosaccades, very rapid (up to $900^\circ/sec$ in humans) rotation of the eyes with big amplitude.

Fixation eye movements

Involuntary fixation eye movements include tremor, drifts and microsaccades.

Tremor is an aperiodic, wave-like motion of the eyes of high frequency but very small amplitude.

Drifts occur simultaneously with tremor and are slow motions of eyes, in which the image of the fixation point for each eye remains within the fovea!

Drifts occurs between the fast, jerk-like, linear microsaccades.

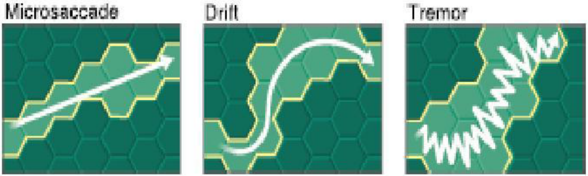
Characteristics of fixation eye movements

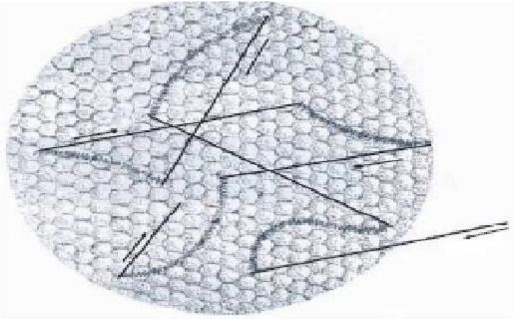
	Amplitude	Duration	Frequency	Speed
Tremor	20-40 sec	-	30-100 Hz	Max 20 min/s
Drift	1-9 min	0.2-0.8 s	95-97% of time	1-30 min/s
Micsac	1-50 min	0.01-0.02 s	0.1-5 Hz	$10-50^{\circ}/s$

Per 1 s tremor moves on 1-1.5 diameters of the fovea cone
drift moves on 10-15 diameters
microsaccads moves on 15-300 diameters.

Under tremor the axis of eye draws a cone for 0.1 s.

In 2-3 sec after compensation of fixation eye movement, a human lost ability to see an immobile object.(Yarbus)


Microsaccade
Drift
Tremor



Stochastic model of eyes movements by R.Engbert, K. Mergenthaler, P. Sinn, A. Pikovsky "Self-avoiding Random Walk"

The authors describe involuntary eye movement as a self-avoiding random walk on the square lattice $\mathbb{Z}^2$ with quadratic potential ("Random walk in a swamp on a paraboloid").
Physiological aim of such fixation eyes movement (when gaze fix a point A):
the images $(\bar{A})(t)$ of A on retina for some interval of time must be homogeneously distributed between all receptors of the fovea.

Diffusion distance between retina images

Corollary. Image of a point on retina is a trajectory of a Brownian motion.

Conjecture. The brain calculates the distance between retina images of two points A, B of external world as diffusion distance between Brownian trajectories $x_t(A).x_t(B)$, which correspond to these points in the retina. Diffusion distance was defined by R.R. Coifman and S. Lefon (2004).in the framework of Diffusion Geometry. Diffusion Geometry is an effective way to construct dimensional reduction , i.e. to construct a low dimensional manifold , which comprise data in a high-dimensional vector space. It is more effective method for dimensional reduction, then Principal Component Analysis (PCA) of K.Pearson (1901) and Multi-Dimensional Scaling (MDS).

Why we need fixation eye movements

1.(Geometry) Provided that the gaze OF is fixed, the retina gets information only from the 2-dimensional Lagrangian submanifold $L(F) = \mathbb{R}P^2 \subset L(E^3)$ of lines going through the point F.
When eye moves in a neighborhood of a fixed point OF, it gets information from a neighborhood of $L(F)$ in 4-manifold $L(E^3)$.
2. To see immobile objects (Yarbus)
3. To determine direction of moving external objects (Roords et al., 2013)
4(Neuroscience) For better identification of contours in V1 cortex.

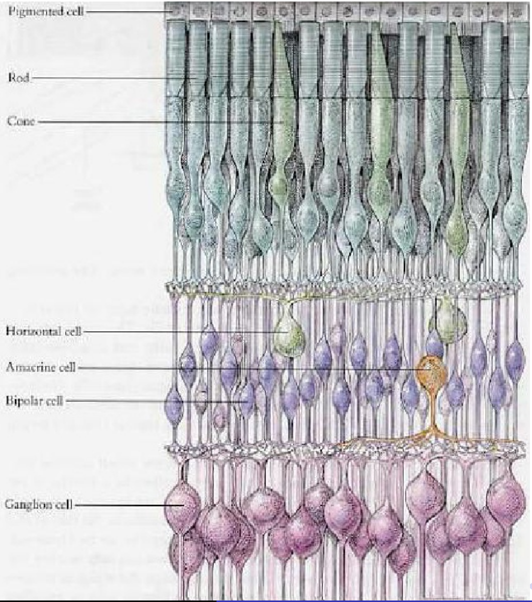
Physiological structure of retina

Retina consists of 5 layers. In rabbit's retina, there are 55 different types of cells and in human approx. 80.
The bottom layer consists of receptors, photoelements which transform light energy into electric signals. They measure the energy function

$$I_R : R \to \mathbb{R}^{\geq 0}$$

and send information to ganglion cells.
In fovea one cone is connected with 1 ganglion. In peripherP,,, one rode is connected with $10^2 - 10^3$ ganglions.
There are 1 million of ganglions and $125 - 150$ millions of receptors.
RF of ganglion cells are rotationally invariant and contain central disc and surround ring.

Retina

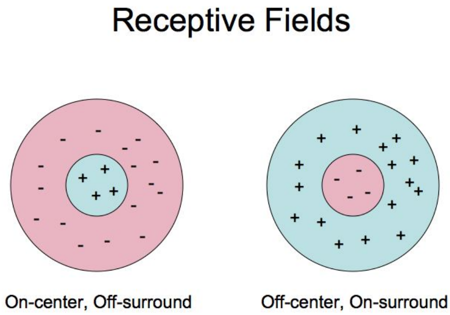


Pigmented cell
Rod
Cone
Horizontal cell
Amacrine cell
Bipolar cell
Ganglion cell

Contourization in retina

Action of Marr filter

On and Off Marr cells



Structure of retina. Fovea

The fovea (from Latin pit or pitfall) was discovered by Leonardo da Vinci. It is a small pit in retina which contains mostly cones. Diameter of fovea is $0.35mm \sim 1^o$ it occupies 1% of retina, but is projected onto 50% of area of the visual cortex.
When we fix gaze on a point A, the image $\bar{A}$ of the points on retina moved due to fixation eye movements (FEM), but always remains inside fovea.

Metric in retina

The physical metric in retina as a (part of sphere) is standard metric of sphere. The distance in retina is described in mm or in degrees. $1mm = 3.5^o \sim 6cm$ on distance $1.5m$,
$1^o \sim 0.3mm \sim 2,5cm$ on distance 135 cm.
diameter Moon and Sun is $0.5^o = 0,15mm = 150\mu$.
RF of neurons inside fovea has diameter
$0.25^0 - 0.5^o = 0.07 \times 0.15mm$, area $0.12mm^2$.
RF in periphery has diameter up to 8^o, in average in 30 times larger then in fovea and RF contains thousands of rods.

Magnification = distance between two points in primary visual cortex VI which corresponds to 1^o distance in retina.
Magnification in fovea $1^o = 0.3$ mm corresponds to $6mm$ in V1 (18 times)
Maginification in periphery is $1^o = 0.3$ mm corresponds to $0.15mm$ (0.5 times).
Shift in 2 mm at any point of primary cortex VI corresponds to shift on diameter of corresponding RF in retina (Hubel).
The natural metric in cortex , proportional to the physical metric in retina, is given by the number of neurons along a given curve.

Conjecture

Under fixation movements , the image of a point at periphery remains inside the RF of a neuron and such neurons do not see the movements and form a stable coordinate system.

Diameter of RF at periphery has several degrees , amplitude of tremor $0,5'$, drift $1' - 9'$, microsaccade: average $15'$, maximum $50'$.

Area of RF of 10-12 sequentially located VI cells is 2-4 time larger then RF of one point. The collection of columns, which has any possible orientation $0^o - 180^o$ is called a hypercolumn (HC) or module. Roughly speaking is it a square with side 12 cells (or 1 mm) in central domain which corresponds to fovea. Union of RF of a HC is called the composed RF of a point (which corresponds to HC). More generally, a MC associated with a local parameter , which is measured by a visual cell is a domain in minimal domain in cortex ("fundamental domain") which contains columns, measured any value of this parameter.

Information precessing in retine. Two pathways from receptors to ganglions

Direct path: receptor - bipolar - ganglion activates the center of ganglion cells, which work as a linear filter.
Antagonistic surround is activated by (linear) negative feed back from horizontal cells via indirect path : receptor-horizontal cell-(amacril)- bipolar - ganglion.
A nonlinear rectifying mechanism (associated with contrast gain control) is related with amacril cells.
For sufficiently small contrast, ganglion P-cells is working as linear Marr filter. M-cells, responsible for perception of moving objects, are working as essentially non-linear filters. Response depends on stimulus contrast and temporal frequency. (E. Kaplan, E. Benardette, Dynamics of ganglion cells, 2001).