Neural mechanisms of vision

Giuseppe di Pellegrino
Department of Psychology, University of Bologna

g.dipellegrino@unibo.it

Cognition and Neuroscience
Second cycle Degree in Artificial Intelligence – 2022/23

0

The anatomical and functional organization of vision

Low-level visual processing: the retina and primary visual cortex

1

1

What is vision?

What does it mean, to see?
Vision is the process of discovering what is present in the world, and where it is.
(Marr, Vision, 1982)

Vision is a process that produces from images of the external world a description that is useful to the viewer and not cluttered with irrelevant information (Marr and Nishihara, 1978).

2

2

Vision dominates our perceptions and memories of the world and appears even to frame the way we think.

Vision is used not only for object recognition but also for guiding our movements.
These separate functions are mediated by at least two parallel and interacting pathways.

Vision, and more generally the brain, is a system that analyzes information (information processing device): receives inputs and transforms them into outputs.

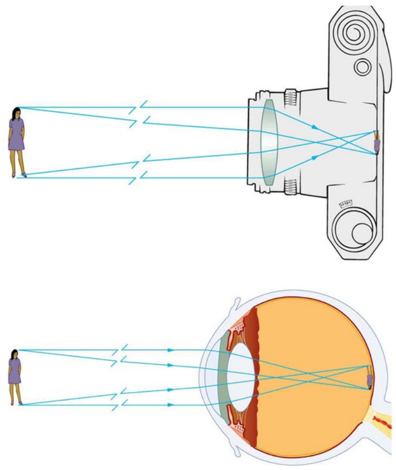
3

3

Vision is often incorrectly compared to the operation of a camera.

A camera simply reproduces point-by-point the light intensities in one plane of the visual field.

The visual system, in contrast, does something fundamentally different. It interprets the scene and parses it into distinct components, separating foreground from background.

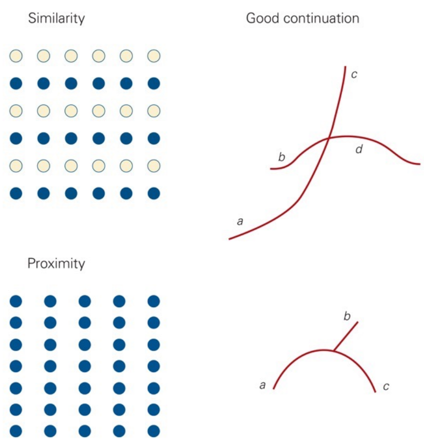


4

4

Vision is an active and bidirectional process

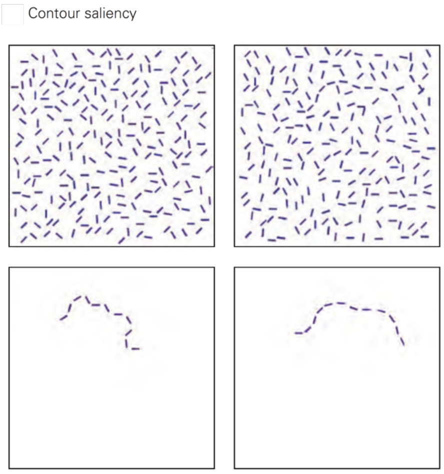
Vision is a generative process that involves more than just the information provided to the retina. The brain has a way of looking at the world, a set of expectations about the structure of the world that derives in part from experience and in part from built-in neuronal wiring neural wiring.



To link the elements of a visual scene into unified percepts, the visual system relies on organizational rules such as similarity, proximity, and good continuation.

5

5



Visual priming
Higher-order representations of shape (in memory) guide lower-order processes of surface segmentation.

The principle of good continuation is also seen in contour saliency. On the right, a smooth contour of line elements pops out from the background, whereas the jagged contour on the left is lost in the background.

6

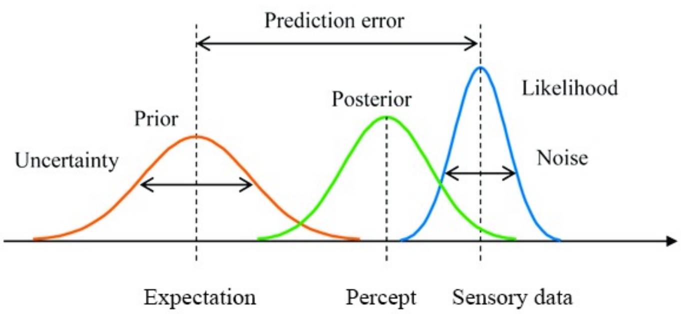
6

Bayesian theories treat the visual system as an ideal observer that uses prior knowledge about visual scenes and information in the image to infer the most probable interpretation of the image.

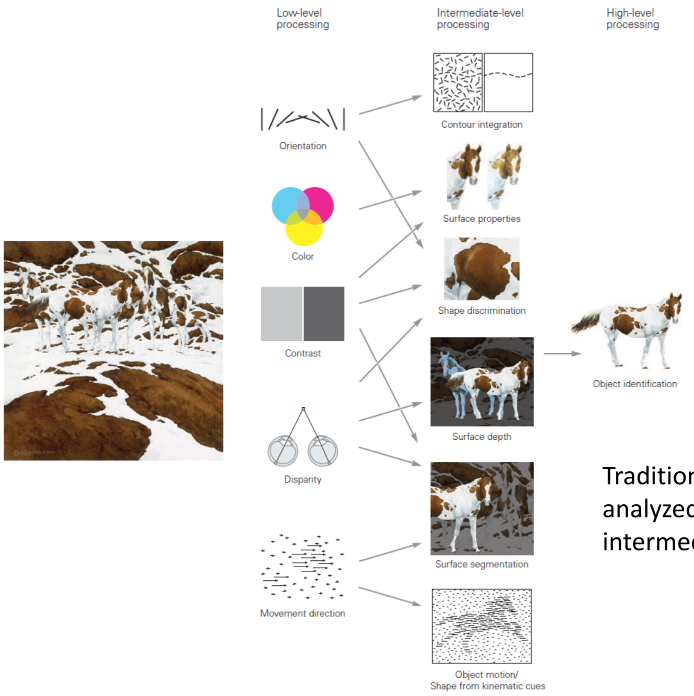
The posterior probability of a possible real–world stimulus S (i.e., percept) is proportional to the product of the prior probability of S (that is, the probability of S before receiving the stimulus I, e.g., expectation) and the likelihood (the probability of I given S, i.e., sensory data).

Is it reasonable to assume that the visual system knows the probability calculus and operates according to it?

$$p(S|I) = \frac{p(I|S)p(S)}{p(I)}$$



7



Traditionally, a visual scene is analyzed at three levels: low, intermediate and high.

8

At the lowest level, visual attributes such as local contrast, orientation, color, depth and motion are processed.

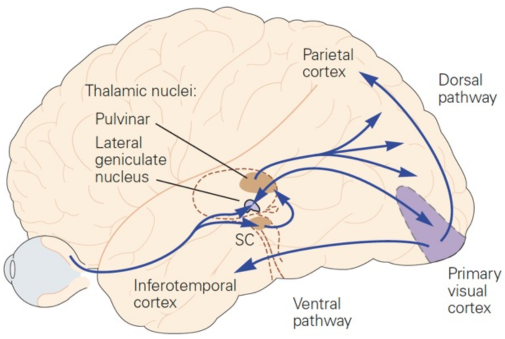
Intermediate-level processing: low-level features are used to parse the visual scene.
Local orientation is integrated into global contours (contour integration); local visual features are assembled into surfaces, objects are segregated from background (surface segmentation), surface shape is identified from depth, shading and kinematic cues.

The highest level concerns object recognition.

Once a scene has been analyzed by the brain and the objects have been recognized, the objects can be associated with memories of shapes and their meanings.

9

Visual processing is mediated by the retino-geniculo-striate pathway

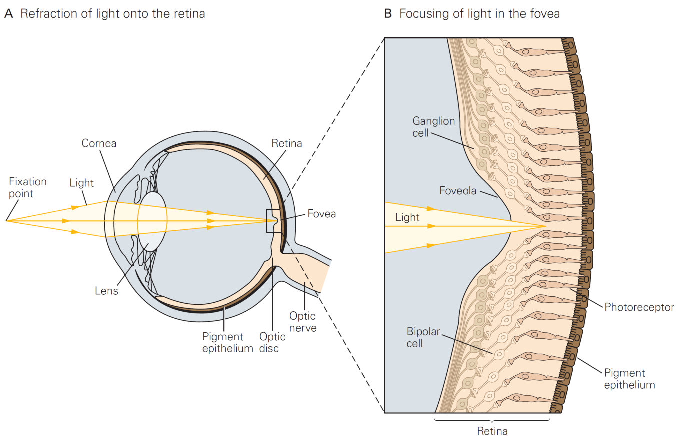


This pathway includes:

- Retina;
- Lateral geniculate nucleus (LGN) of the thalamus;
- Primary visual cortex (V1) or striate cortex;
- Extrastriate visual areas

10

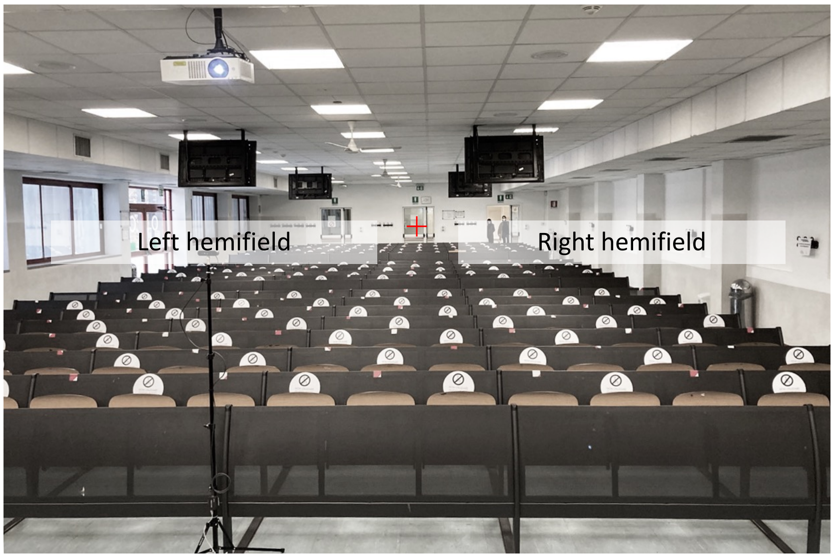
10



The brain's analysis of visual scenes begins in the two retinas, which transform the visual input into neural signals, a process known as phototransduction.

11

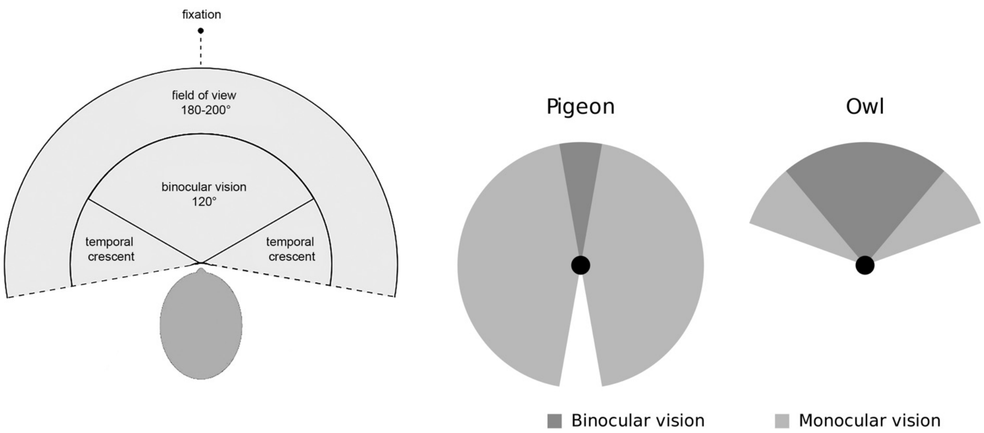
11



Visual field refers to the area that can be seen as you fixate your eyes on a central point (cross).

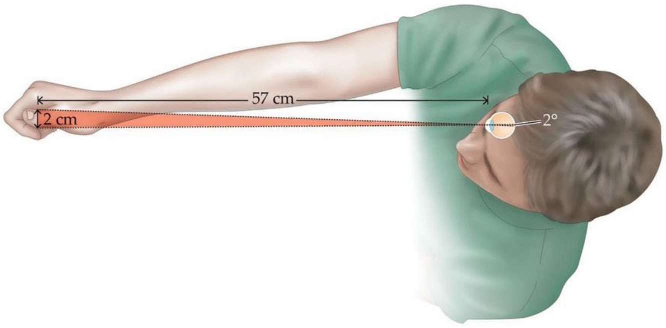
12

12



13

13



fovea

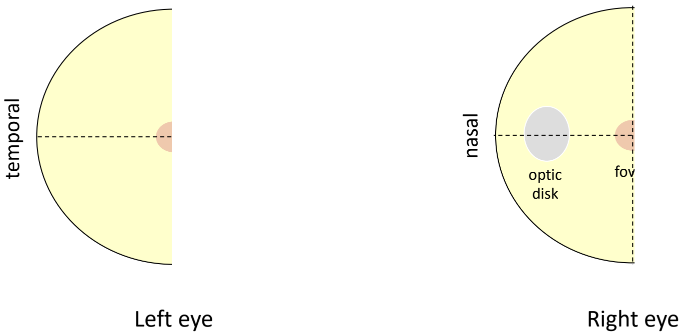
Visual angle of the width of the thumb held at arm's length is about 2 deg.

14

14

The retina of each eye can be divided into

Nasal hemiretina, medial to the fovea
Temporal hemiretina, lateral to the fovea

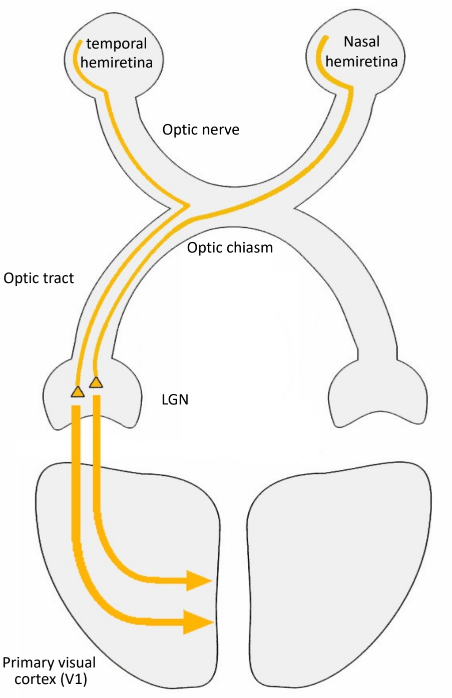


The left hemiretinas (temporal of the left eye, and nasal of the right eye) see the right (opposite) visual hemifield

The right hemiretinas (nasal of the left eye, and temporal of the right eye) see the left (opposite) visual hemifield

15

15

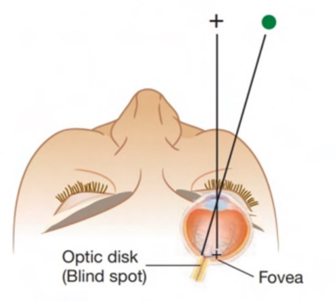


Partial decussation

Because the temporal hemiretina of one eye sees the same visual hemifield as the nasal hemiretina of the other eye, partial decussation of the optic nerve fibers at the chiasm ensures that all information related to each hemifield is processed in the visual cortex of the contralateral hemisphere.

16

16



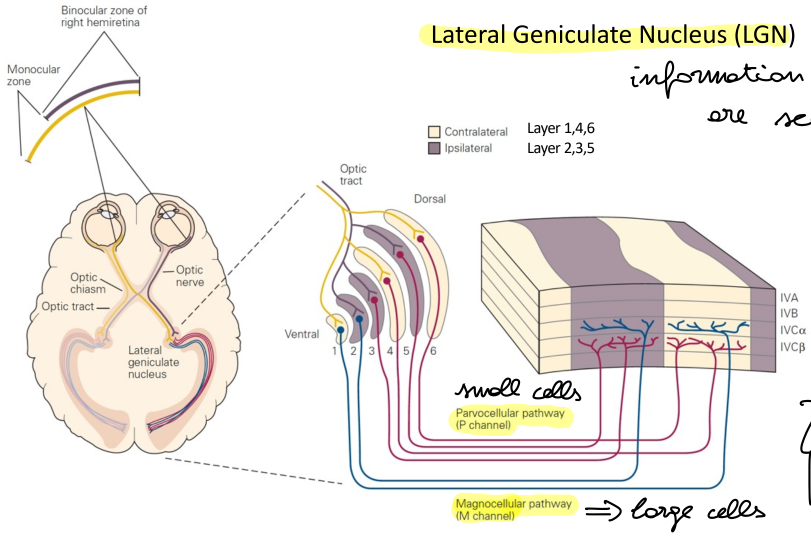
Retina blind spot

The optic disc is where the axons of retinal ganglion cells leave the eye to reach the LGN of the thalamus.
This region of the retina does not contain photoreceptors and, because it is insensitive to light, produces the perceptual phenomenon known as a blind spot.

17

Lateral Geniculate Nucleus (LGN)

information from left and right are separated

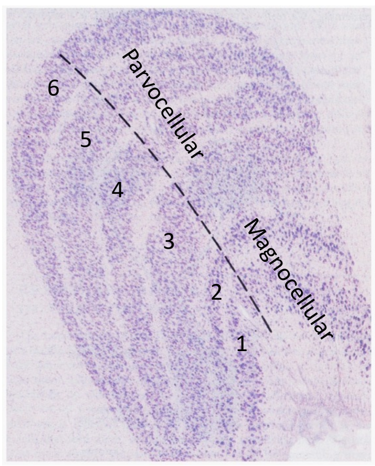


within it some layers are for one eye, others for the other eye

Beyond the optic chiasm, the axons from nasal and temporal hemiretinas carrying input from one hemifield join in the optic tract, which extends to the LGN of the thalamus.

18

18



Interspersed between the Magno and Parvo layers, are thin but dense layers called Koniocellular which receive from the K retinal ganglion cells.

Lateral Geniculate Nucleus (LGN)

In primates, LGN is a layered structure consisting of six layers, of which two Magnocellular layers (layers 1 and 2), and four Parvocellular layers (layers 3 to 6).
Each layer receives input from either the ipsilateral eye (temporal hemiretina, layers 2, 3, 5) or the contralateral eye (nasal hemiretina, layers 1, 4, 6).

Since each layer contains a map of the contralateral hemifield, the six maps are stacked on top of each other and in spatial register.

The magnocellular layers project to the IVCα layer while the Parvocellular layers project to the IVCβ layer of V1. The Koniocellular intercalated layers project to the blobs (layers 1-2) of V1.

19

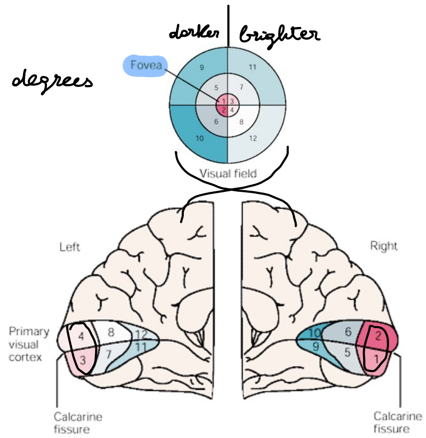
19

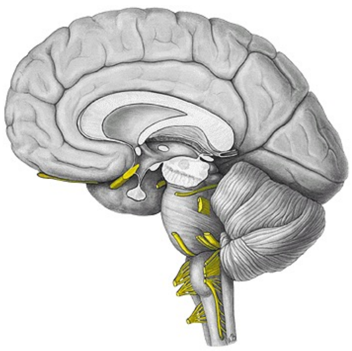
Primary visual cortex (V1)

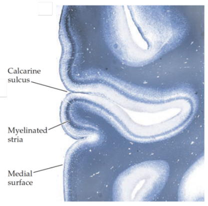
In humans, V1 (BA17) is located in the occipital portion of the brain along the calcarine fissure of the brain.
V1 constitutes the first level of cortical information processing.

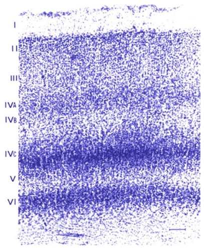
fovea: very few degrees
⇒ thumb

darker | brighter









Nissl stain of V1 reveals the different layers quite clearly.

topographic map → closer in the receptive field, closer in the brain

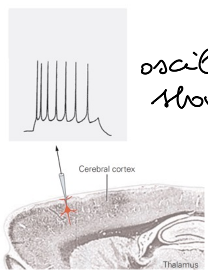
20

Single-cell recording → single neuron

This technique allows recording signals (firing rate) from single neurons.

A fine-tipped, usually metal (platinum), electrode is inserted in the animal brain to record extracellularly change in electrical activity called action potential (AP, 1ms duration) or spike. Collected signals are appropriately amplified, filtered, viewed through an oscilloscope, and saved to a computer for offline analysis.
Since spikes are all-or-none highly stereotyped signals, most information is encoded in the brain as neuron firing rate, i.e., the number of AP in 1s.
The primary goal of single-cell recording experiments is to determine what experimental manipulations produce a consistent change in the firing rate of an isolated neuron.

Disadvantages
– invasive
Advantages
– high spatial and temporal resolution
– differentiation between excitation and inhibition



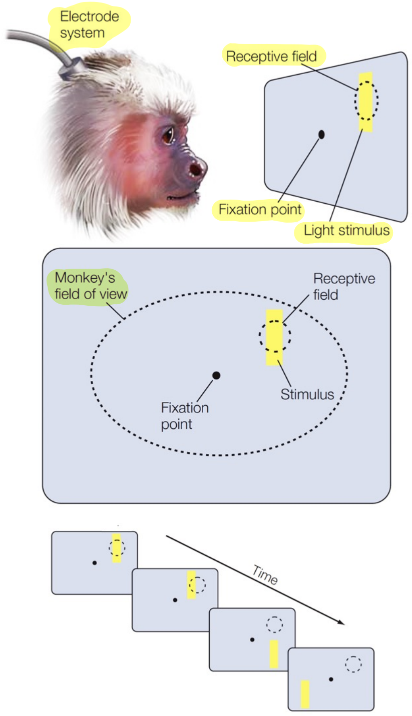
oscilloscope shows the activity

21

21

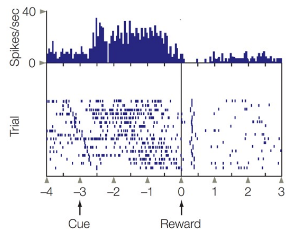
receptive field: portion of the visual field where, if you put a stimulus, something activates
⇒ it is perceived

5/3/23



Receptive field

Visual neurons respond to stimuli in only a limited region of space.
Th is region of space is referred to as that cell's receptive field (RF).
Monkey is required to maintain fixation, and stimuli are presented at various positions in the field of view.
The neuron below fires vigorously only when the stimulus is presented in the upper right quadrant, thus defining the upper right region as the RF for this cell.

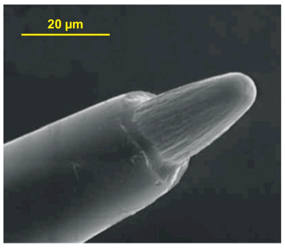


activation potential has all-or-none nature

spike gives the importance
⇒ 1 per ms ~ 1000 per ms

22

22

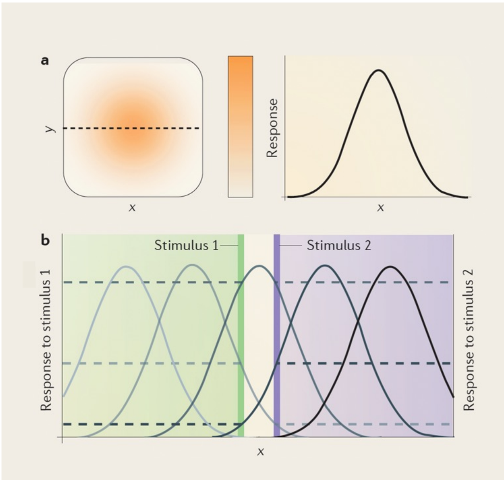


Recording of the activity of a V1 neuron. In the receptive field a moving stimulus is presented, directed downward or upward. Note that neuron firing rate is higher when the stimulus moves up.

23

23

smaller RF, higher resolution



The receptive field (RF; Charles Sherrington, 1906) of a neuron is defined as the part of stimulus space within which a stimulus elicits a response from the neuron.

In the visual system, the neuron responds to a stimulus presented in a region of space in the visual field (that is, its RF) but not to the same stimulus when it is presented outside this region.

A given stimulus typically elicits the strongest response from the centre of the RF, with the response gradually declining as the stimulus is presented further away from the centre of the RF.

Thus, the RF can be well described by a two-dimensional Gaussian distribution.

24

24

Organization of receptive fields in the visual system

The receptive field size varies based on:

- position of the neuron along the visual pathway
- eccentricity, i.e., RF position relative to the fovea

At each eccentricity, receptive fields are relatively small at the first levels of visual processing and become progressively larger at later levels.

advantages and disadvantages of high resolution

25

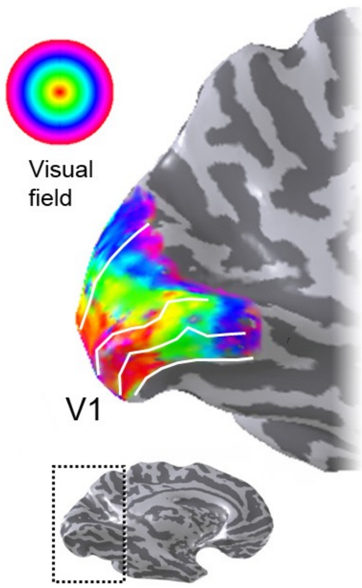
25

Retinotopy

In early visual areas (e.g., V1 to V5), neuron RFs reveal an ordered organization, termed a retinotopic or visuotopic map.

This refers to the existence of a non-random relationship between the position of neurons in the visual areas.

Neuron RFs form a 2D map of the visual field, such that neighbouring regions in the visual image (and therefore on the retinal surface) are represented by adjacent regions of the visual cortical area (i.e., orderly mapping of RF positions in retinotopic coordinates)

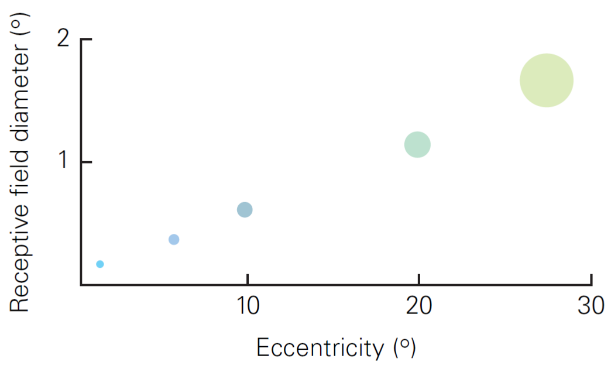


26

26

Eccentricity

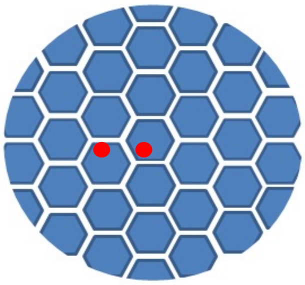
The receptive fields of the retinal ganglion cells that monitor portions of the fovea subtend about 0.1 ° (equal to 6 min of arc), while those in the visual periphery reach up to 1 ° of visual angle or more.



1 Arc min = 1/60 degree

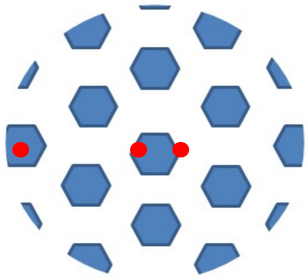
27

27



High density of photoreceptors in the center of the retina (fovea)

High spatial resolution

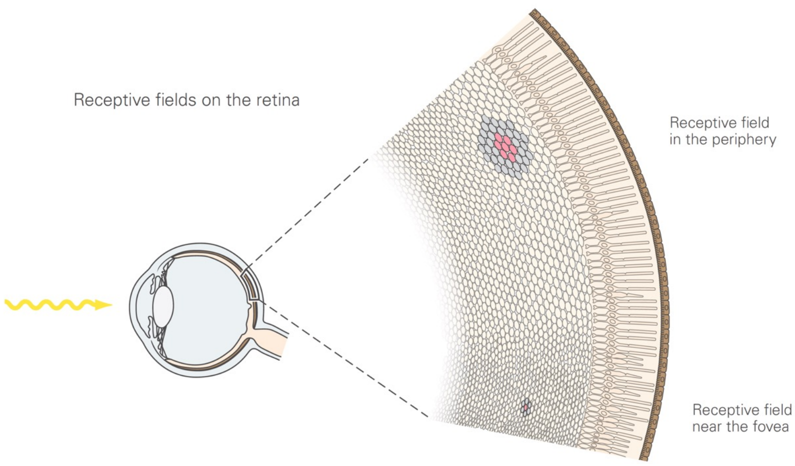


Low density of photoreceptors in the periphery of the retina

Low spatial resolution

28

28



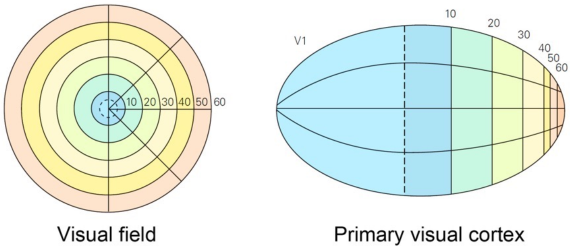
A retinal ganglion cell near the fovea receives input from fewer receptors covering a smaller area (less converging inputs), while a cell further away from the fovea receives inputs from many more receptors covering a larger area (high convergence inputs).

29

29

Cortical magnification

The amount of cortical area devoted to each degree of the visual field, known as the magnification factor, varies with eccentricity (i.e., the neural maps of the visual field are not isometric).
In fact, the central part of the visual field controls the largest area of the cortex.
For example, in V1 more cortex is dedicated to the central 10 ° of the visual space than to everything else.



30

30

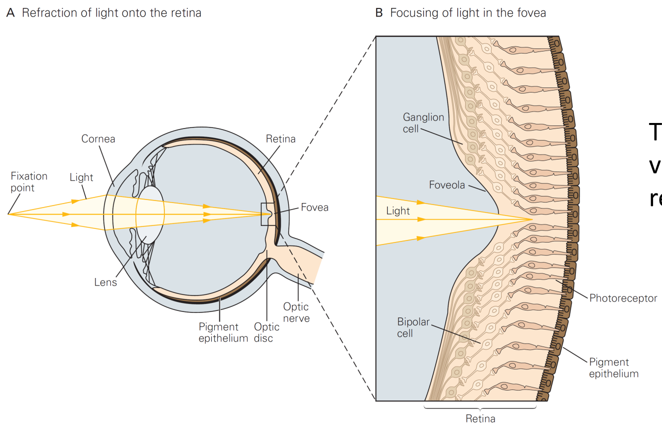
Low-level visual processing involves:

retinal circuits
LGN
some (but not all) of the neurons of V1

This level allows you to extract some elementary characteristics from the visual image.

31

31



The eye projects the visual scene onto the retina's photoreceptors.

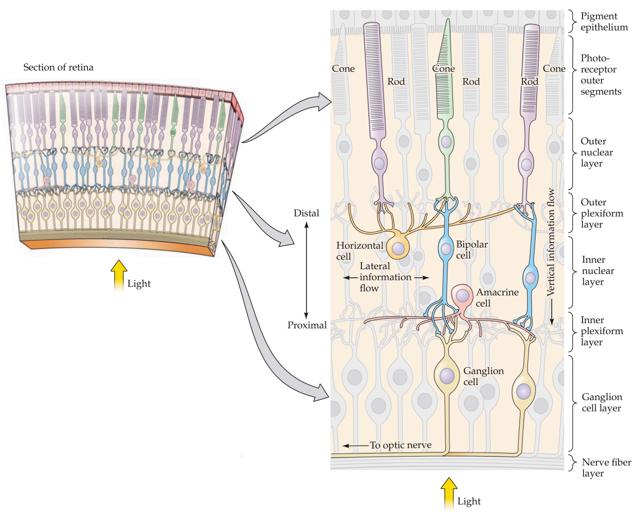
The retina is a thin (0.5mm) sheet of neurons at the back of the eye, made up of five main types of cells arranged in three layers.

The fovea corresponds to the gaze center that we direct towards the objects we pay attention to.

The optic disc is an area of the retina without a receptor and is therefore a blind area.

32

32



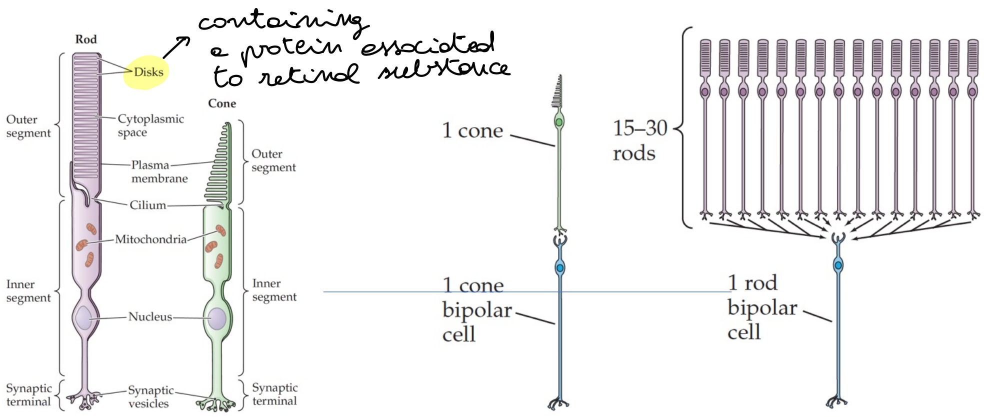
Basic circuitry of the retina.
A three-neuron chain—photoreceptor, bipolar cell, and retinal ganglion cell—provides the most direct route for transmitting visual information to the brain.

33

33

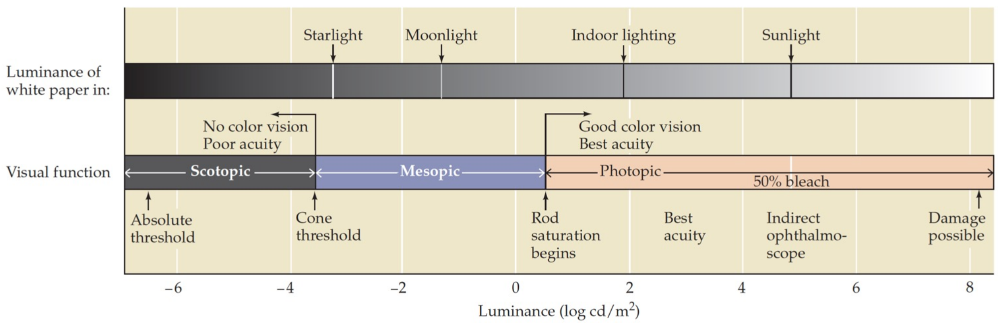
Humans, like most vertebrates, possess two types of photoreceptors, rods and cones, differing in shape, function, connectivity and distribution in the retina.

containing a protein associated to retinal substance



34

34



Rods and cones differ in their sensitivity to light

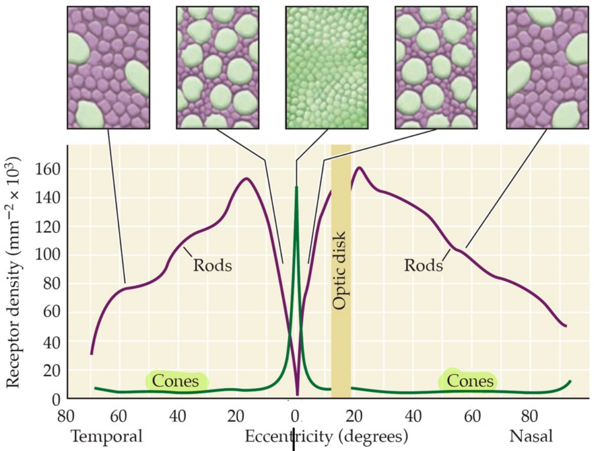
Rods (100 million) are mainly for night (or scotopic) vision. They can signal the absorption of 1 single photon; their response saturated in daylight, and they cease to respond to variation of intensity;

Cones (5 million) are only for daytime (or photopic) vision.
They are much less sensitive to light, make no contribution to night vision, but have higher spatial and temporal resolution (their response is considerably faster than that of rods);

35

35

Distribution of photoreceptors in the human retina



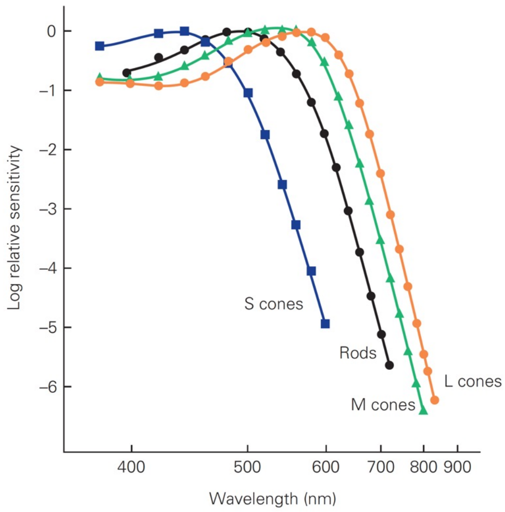
fovea region → outside no cones

Cones are present at a low density throughout the retina, with a sharp peak in the center of the fovea (the foveola).

Conversely, rods are present at high density throughout most of the retina, with a sharp decline in the fovea; rods are absent in the foveola.

36

36



Spectral sensitivity curves of retinal photoreceptors

For each wavelength, the greater the sensitivity the lower the intensity of light required to evoke the response in the photoreceptor.

a single photoreceptor sees just in black and white

Primates have only one type of rod and three types of cones, distinguishable on the basis of the different photopigments (S, M and L) and, consequently, of the range of wavelengths to which they respond: L (long waves), M (medium waves) and S (short waves).

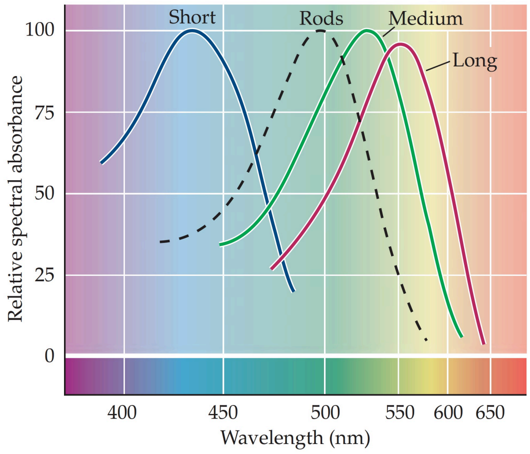
37

37

without a cone, we have a bichromatic vision

Because a single photoreceptor cannot distinguish between a change in the wavelength of light and a change in its intensity, the analysis of colour requires the comparison of signals from different types of cones

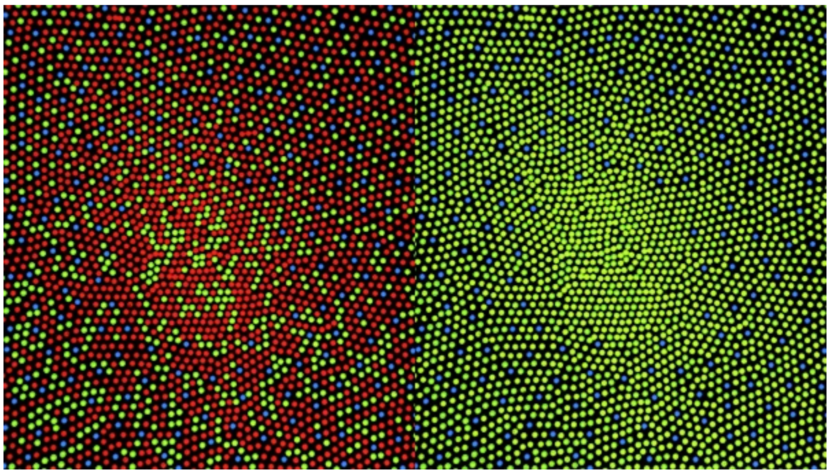
In night vision, when only rods are active, a green light consequently has exactly the same effect on the visual system as a red light of a greater intensity.



38

38

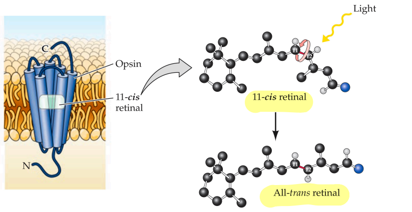
No long wavelengths

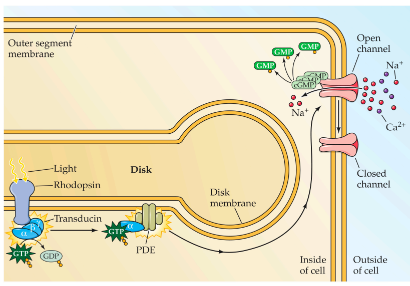


Distribution of cone cells in the fovea of an individual with normal color vision (left), and a color blind (protanopic) retina.
Note that the center of the fovea holds very few blue-sensitive S cones.

39

39





Phototransduction in rod photoreceptors.

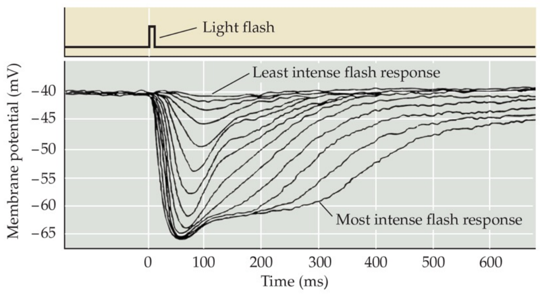
Rhodopsin resides in the disk membrane of the photoreceptor. The opsin molecule enclose the light-sensitive retinal molecule.

Absorption of a photon of light by retinal leads to a change in configuration from the 11-cis to the all-trans isomer.

This in turn activates a phosphodiesterase (PDE) which then hydrolyzes cGMP, reducing its concentration and leading to the closure of channels in the outer segment membrane.

40

40



In the retina, photoreceptors do not exhibit action potentials; rather, light activation causes a graded change in membrane potential and a corresponding change in the rate of transmitter release.

In the dark, the receptor is in a state of continuous depolarization, with a membrane potential of approximately –40mV (current in the dark).

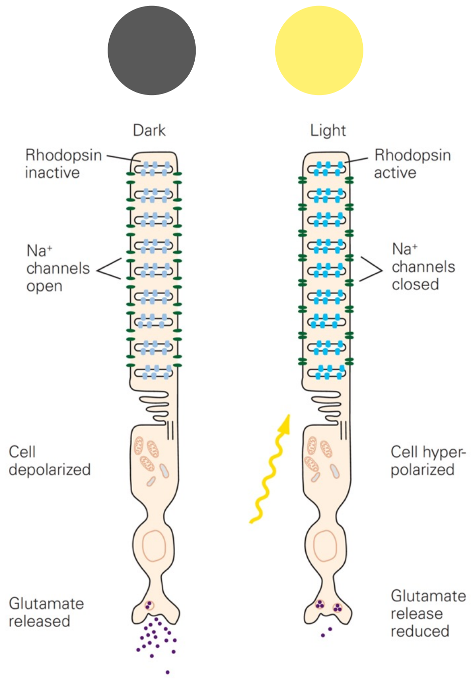
The presentation of a light flash of increasing intensity produces an ever greater (i.e., more negative) hyperpolarization that becomes saturated when the membrane reaches about -65mV.

41

41

Photoreceptors produce a relatively simple neural representation of the visual scene:

Neurons in the bright regions are hyperpolarized, while those in the dark regions are depolarized.



42

response on the amount of photons
↓
no action potential

42

they connect the retinal cells with the brain

RGCs respond to light with action potentials, unlike photoreceptors and bipolar cells.

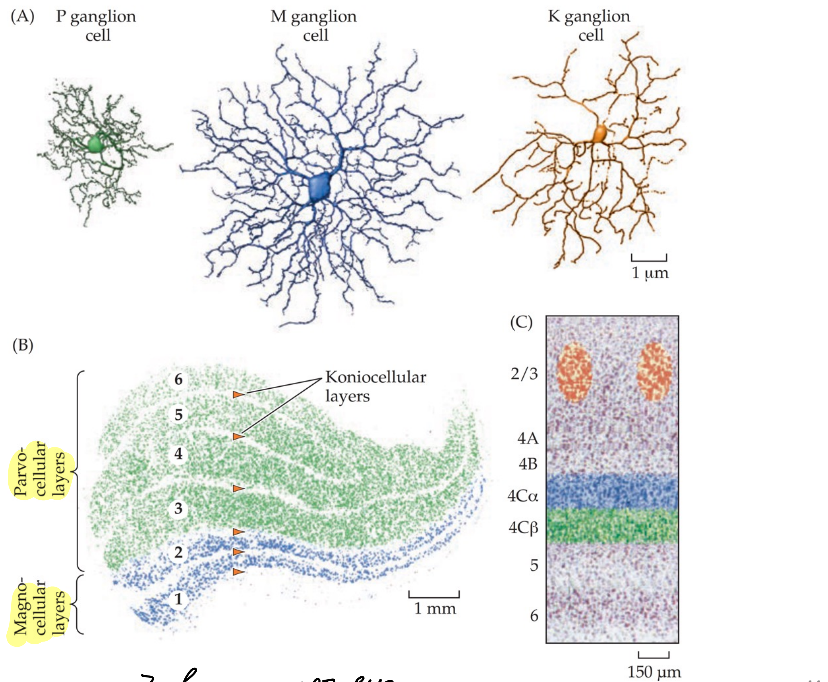
There are several types of RGCs that differ in shape, connectivity and function:

Midget or P (parvo) cells are the most numerous (70%), have smaller cell bodies and dendritic fields, and supply the parvocellular layers of LGN;

Parasol or M (magno) cells, about 10%, have large-diameter cell bodies and large dendritic fields supply the magnocellular layers of LGN;

K cells, about 10%, have small cell bodies and intermediate-sized dendritic fields. They supply the koniocellular layers of LGN;

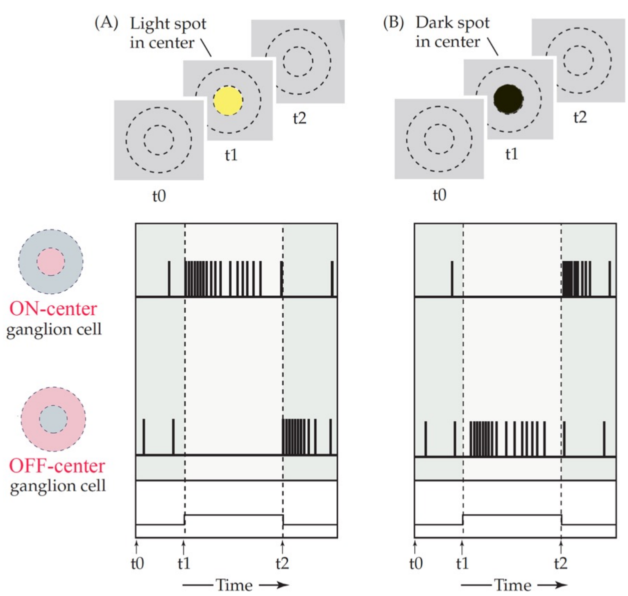
43



3 layers per eye

44

44



RGCs have concentric circular RFs and fall into one of two categories:

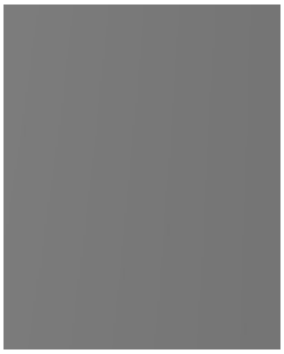
ON-center light
OFF-center dark

The ON-center cells discharge in response to the increase in brightness in the center of the receptive field, while the OFF-center cells discharge in response to the reduction of brightness in the center of the receptive field.

45

45

retina can display brigther and darker images



The ON- and OFF-center cells serve to communicate rapidly and independently both brightening and dimming (relative to mean illumination, left pcvciture) in the visual scene.

If the retina had only ON cells, a dark object would be encoded by a decrease in firing rate.

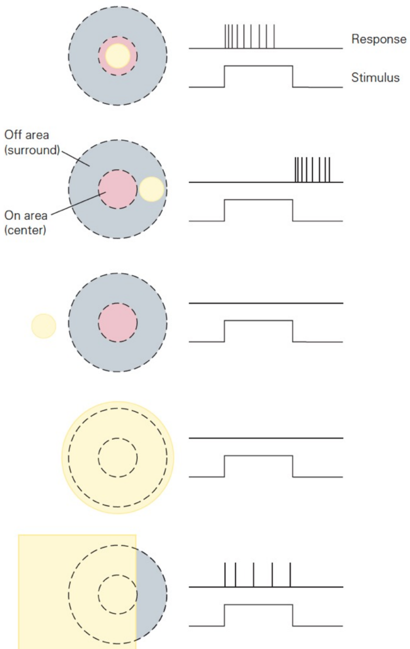
46

46

Center and surround reveal opposite response (lateral inhibition) mutually antagonistic

A uniform stimulus that activates the center and surround simultaneously causes a weak or no response.

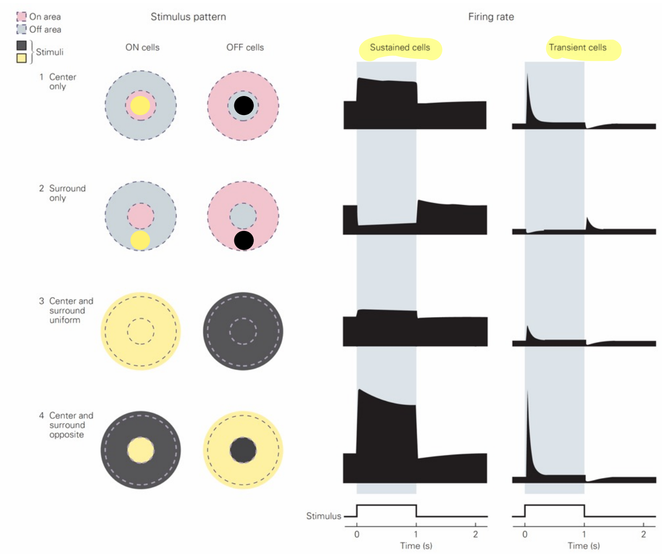
Note that ganglion cells are not selective for the orientation of lines or edges.



borders are meaningful

47
47

47



tonic response

Stimulation of the center combined with the opposite stimulus in the boundary produces the strongest response.

Therefore, the output produced by a population of retinal ganglion cells emphasizes luminance contrast regions in the input, such as a border between two areas of different intensity, and gives less emphasis to the regions of homogeneous illumination.

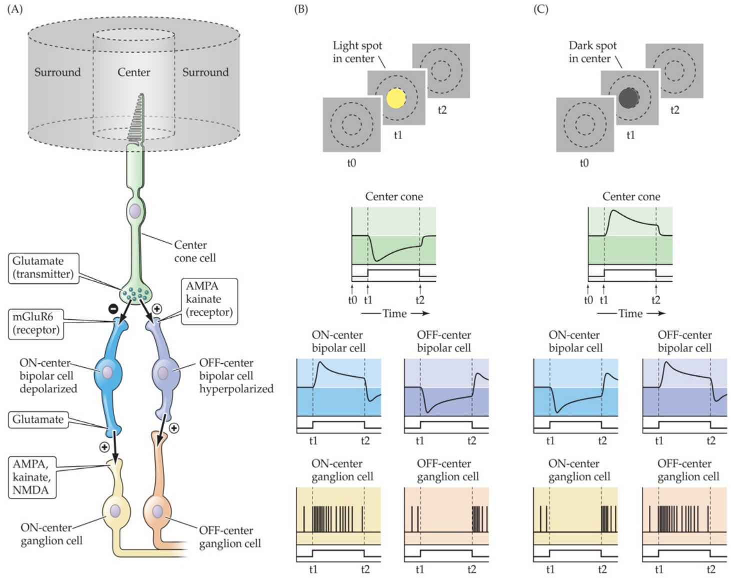
48

48

- RGCs respond only weakly to uniform stimulation.
- RGC mainly emphasize the contrasts of brightness present in the visual scene, and not the absolute intensity of the lighting itself.
- Most of the useful information in a visual scene lies in the distribution of brightness contrasts (edges).
- In fact, the absolute light values reflected by the different objects give very little information as they depend on the intensity of the lighting source.

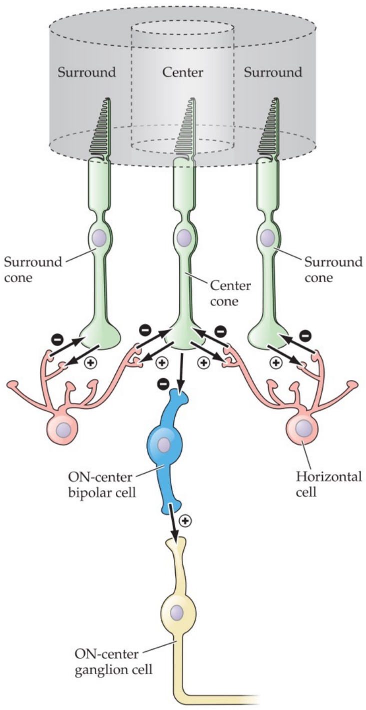
49

The mechanism responsible for generating ON-center and OFF-center responses of retinal ganglion cells depends on bipolar cells.

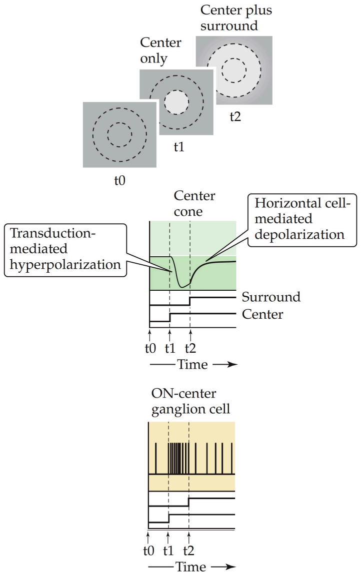


50

50



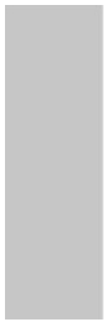
The mechanism responsible for the lateral inhibition of retinal ganglion cells depends on the horizontal cells.



51

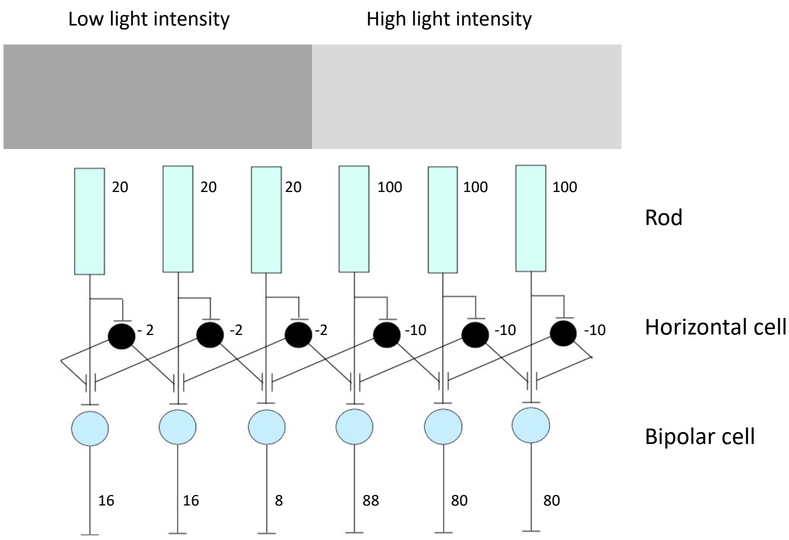
51

The lateral inhibition mechanism mediated by the horizontal cells enhances the contrast at each border, thus increasing the ability to see objects that contrast only weakly with the background, as in the image below.



Chevreul-Mach bands

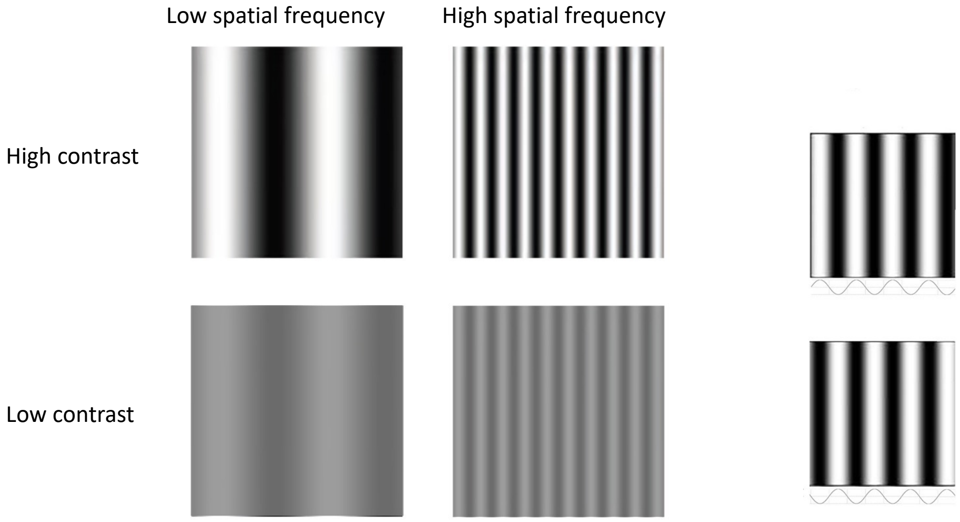
52



Chevreul-Mach bands

53

53

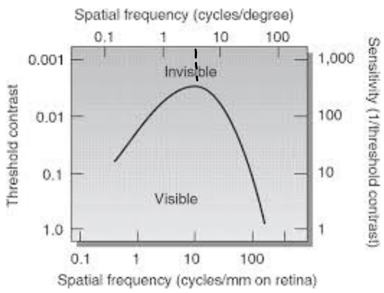


Sinusoidal gratings used for the study of contrast sensitivity in humans

54

54

55



medium spatial frequency

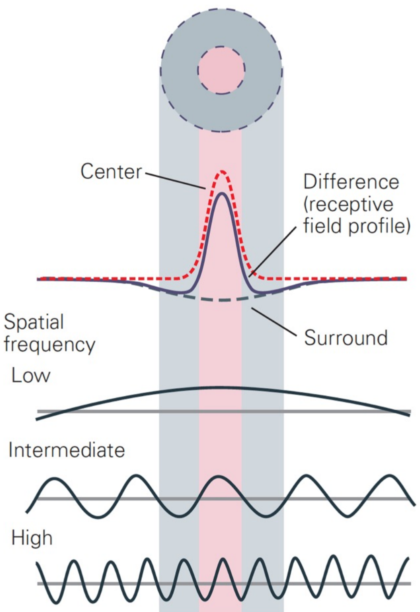
The contrast sensitivity (CSF) function describes an observer's sensitivity to sinusoidal gratings as a function of their spatial frequency.

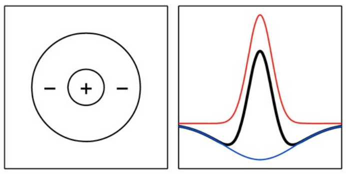
This is measured using a contrast detection experiment in which the minimum (threshold) contrast required to detect sinusoidal gratings of various spatial frequencies is determined.

Sensitivity is defined as 1 / (threshold contrast) (so if the threshold is low, the sensitivity is high).

56

56





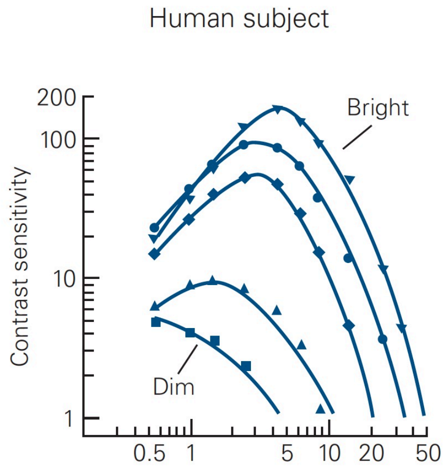
The response of the center-surround concentric RF is modeled as two Gaussian curves, a positive with a narrow base corresponding to the center of the receptive field, and a negative with a wide base, corresponding to the center and surround.

The neuron's response is given by the difference of the two Gaussian curves.

In daylight, contrast sensitivity declines both a high and low spatial frequencies (band-pass behavior)

57

57



In dim light, the visual system's contrast sensitivity declines, but more so at high than at low spatial frequencies.

Thus, the peak sensitivity shifts to lower spatial frequencies.

In this state, the visual system has so-called low-pass behavior, for it preferentially encodes stimuli of low spatial frequency.

The fact that in dim light the receptive fields of ganglion cells lose their antagonistic surrounds explains the transition from band-pass to low-pass spatial filtering

58

58

P cells
dense in central retina (fovea)
narrow dendritic trees and thin axons
cone inputs
small receptive fields
slow and sustained responses
color opponency respond differently to different colours
low contrast sensitivity
respond to high spatial frequencies from small areas of the visual field

M cells
more common in peripheral retina
large dendritic trees and thick axon
rod input
wide receptive fields
fast and transitory responses
no color opponency
high contrast sensitivity
respond to low spatial frequencies from large areas of the visual field

59

Magnocellular	Parvocellular
Large RF	Smail RF
ON- and OFF-center	ON- and OFF-center
High contrast sensitivity	Low contrast sensitivity
Color blind	Color sensitive
Transient	Sustained
Low spatial resolution	High spatial resolution
High temporal resolution	Low temporal resolution
Global aspect	Local aspect
Fast and raw	Slow and detailed

60

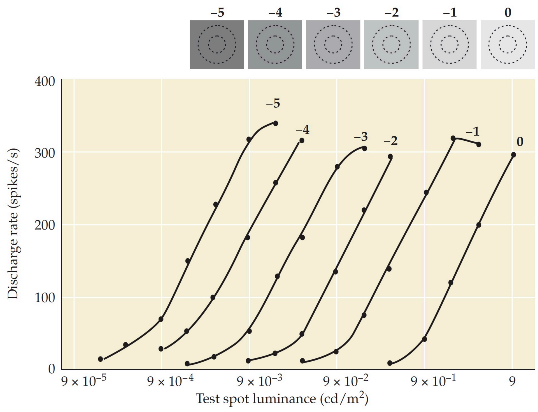
60

Selective P cell lesion
complete loss of color perception
deficits high spatial frequencies discrimination (reduced visual acuity)

Selective M cell lesion
reduced perception of movement
Deficits high temporal frequencies discrimination;

61

Retinal ganglion cells have adjustable operating range



Sigmoid curves illustrating the discharge rate of a single ON-center ganglion cell to the onset of a small test spot of light in the center of its receptive field.

Each curve represents the discharge rate evoked by spots of varying intensity at a constant background level.

The response rate is proportional to stimulus intensity over a range of 1 log unit, but the operating range shifts to the right as the background level of illumination increases.

Sakmann and Creutzfeldt, Pflugers Arch, 1969

62

62

The primary visual cortex (V1) has several functional categories of cells:

simple cells RF becomes elongated in one dim → orientation sensitivity

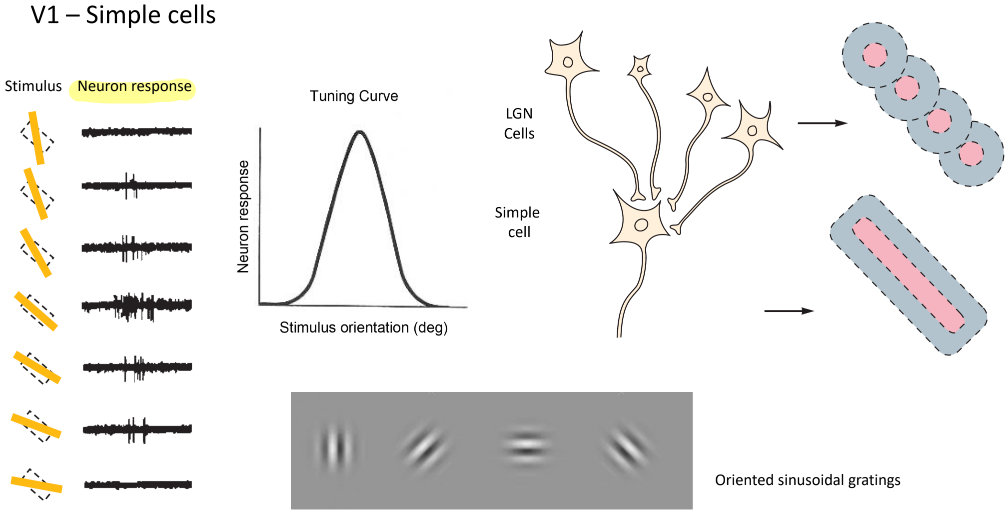
complex cells

end-stopped cells

V1 responds to contrast too

63

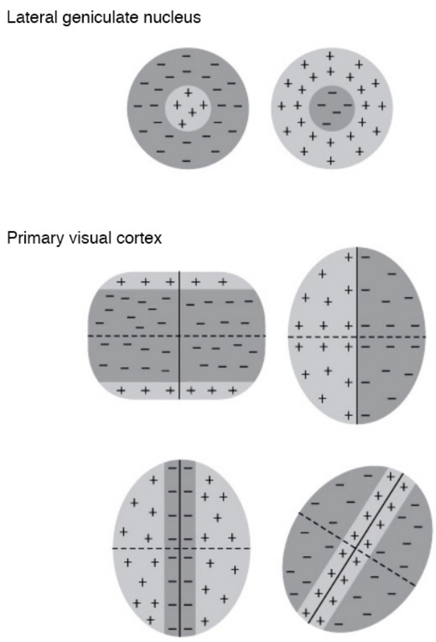
V1 – Simple cells



elungated center → light
periphery → dark

In V1, neurons selectively respond to oriented bars or gratings.
In simple cells, the receptive fields have separate ON and OFF areas.

64

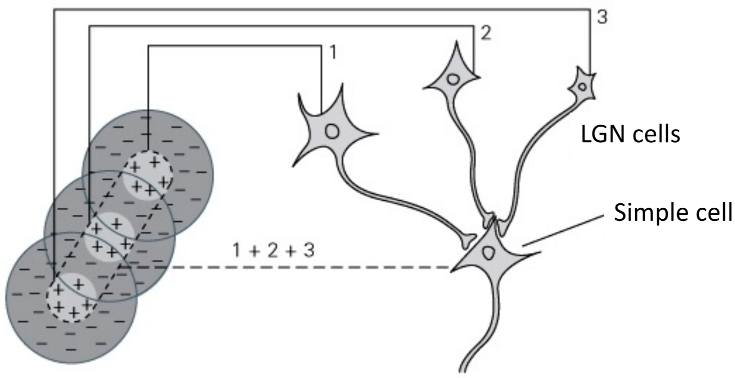


The receptive fields of the simple cells of the primary visual cortex are different and less homogeneous than those of the ganglion cells of the retina and the LGN

orientation matters

there can be also special frequencies or polarities

65



The linear receptive fields of simple cells of the primary visual cortex arise from the convergence of multiple cells of the lateral geniculate nucleus on a single simple cell.

66

Complex cells

Have rectangular receptive fields, larger than those of simple cells;

Respond to linear stimuli with specific orientation;

The position of the stimulus within the receptive field is not critical as the demarcation between on and off zones is not so clear;

Movement of the stimulus in the receptive field is particularly effective in activating the cells;

Complex cells selectively respond to stimuli that move in particular directions;

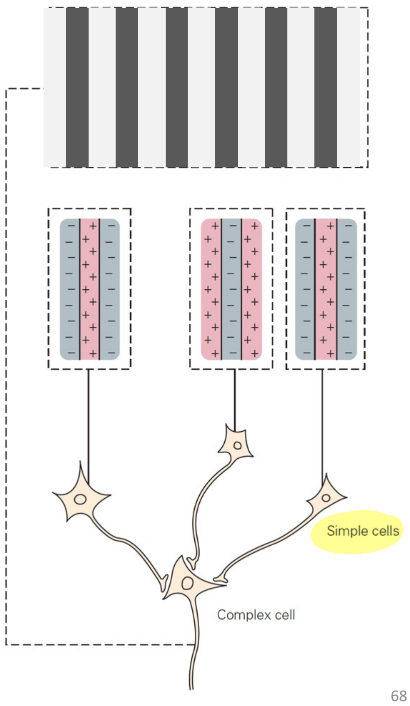
67

V1 – Complex cells

In complex cells, the ON and OFF regions are superimposed, i.e. each position in the receptive field responds to both white and black bars, and the cells respond when a line or edge crosses the receptive field along an axis perpendicular to the orientation of the receptive field.

This constancy in the response to variations of stimulus location in the RF is commonly called position invariance.

important for object recognition



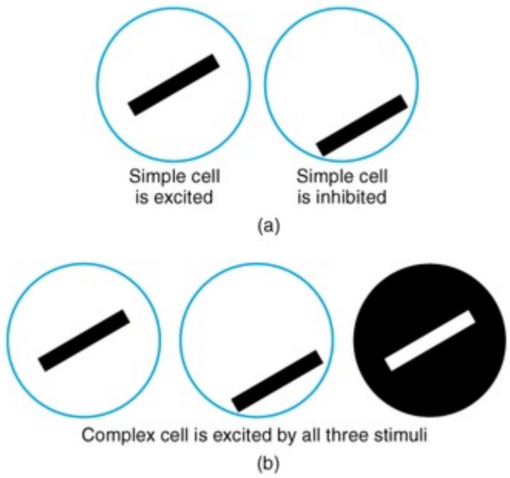
3 simple

converge to 1 complex

68

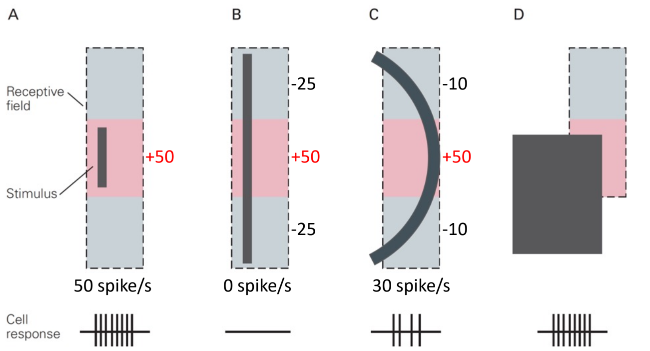
68

- **Response Characteristics of Neurons to Orientation in the Primary Visual Cortex**



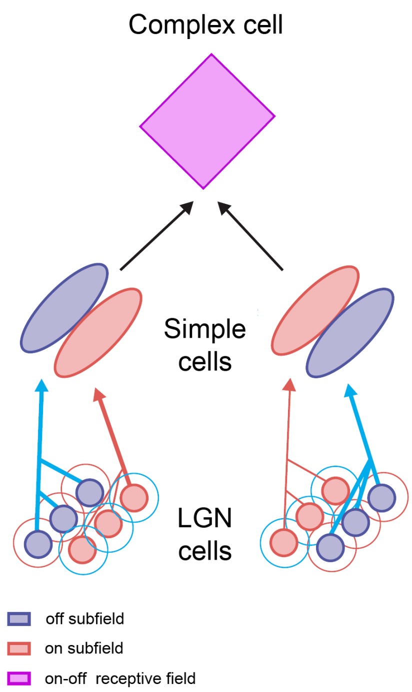
69

End-stopped cells



The receptive fields of some cells have a central excitatory region flanked by inhibitory regions that have the same preferential orientation.
A short linear segment (A) or a long curved line (C) will be effective in activating the neuron, but not a long straight line (B).
A neuron with a receptive field that has only one inhibitory region in addition to the excitatory region can signal the presence of angles (D).

70



Hierarchical model of the receptive field

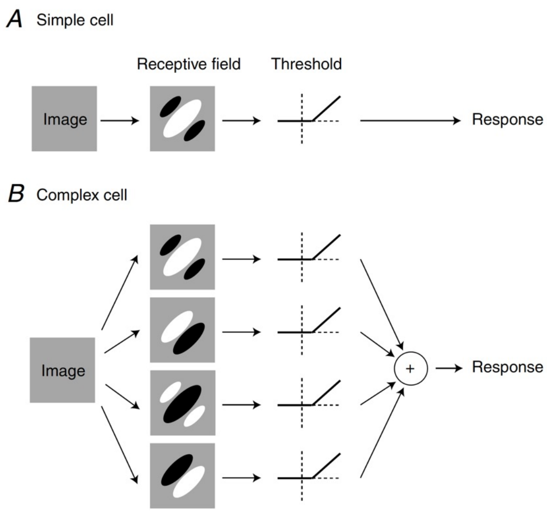
The processing of some characteristics of the visual images is performed through a progressive convergence of information within the visual system.

Each RF at one level of the visual processing hierarchy emerges from the convergence of inputs from many neurons of the previous level.

The size of the RF increases along the visual hierarchy.

Hubel, Wiesel, J. Neurophysiol. 1962

71



The models of simple and complex cells proposed by Movshon, Thompson and Tolhurst (Movshon et al. 1978)

A, simple cells. The first stage is linear filtering, i. e. a weighted sum of the image intensities, with weights given by the receptive field. The second stage is rectification: only the part of the responses that is larger than a threshold is seen in the firing rate response.

B, complex cells. The first stage is linear filtering by a number of receptive fields such as those of simple cells (here we show four of them with spatial phases offset by 90 deg). The subsequent stages involve rectification, and then summation.

72

72

In addition to stimulus position, V1 neurons are selective for a number of attributes:

Orientation. This selectivity must arise from computations that take place within cortex, because LGN responses are not selective for orientation.

Spatial frequency. V1 neurons are typically sharply selective for the spatial frequency of a stimulus. Spatial frequency is best defined for a grating pattern, where frequency is the inverse of the distance between bars. This selectivity arises naturally from the shape of the receptive fields that have multiple ON and OFF regions.

Direction. Cells in area V1 are commonly selective for direction of stimulus motion.

Temporal frequency. Is the inverse of the period between temporal oscillations between dark and light. V1 neurons typically prefer lower temporal frequencies than those that can drive LGN neurons.

Disparity. In animals with front-facing eyes (such as carnivores and primates), much of the visual field is covered jointly by both eyes. This poses a challenge as signals need to be integrated, but also an opportunity for computing binocular depth (stereoscopy). The signals from corresponding regions in the two eyes are kept separate in the LGN, and are combined in V1.

Color. Retinal ganglion cells respond along one of three “cardinal directions”, known informally as red-green, blue-yellow, and black-white. V1 neurons is also organized along "cardinal directions".

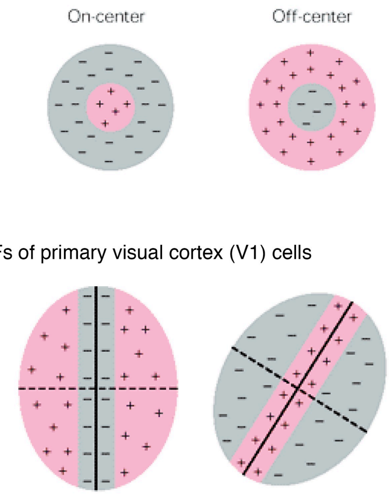
73

Receptive field properties

Properties change from relay to relay along a visual pathway.

By determining these properties, one can test the function of each relay nucleus and how visual information is progressively analysed.

Whereas retinal ganglion cells and neurons in the LGN have concentric center-surround receptive fields, those in the primary visual cortex, although equally sensitive to contrast, analyse oriented contours (orientation selectivity).



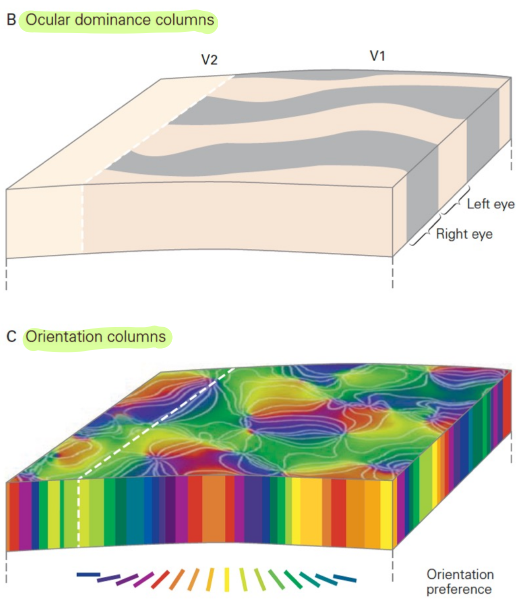
74

74

Cortical columns

In V1, neurons with similar functional properties are found close together in columns or functional modules (about 50-75μm in diameter) that extend from the cortical surface to the white matter (about 2mm high).

V1 includes specific columns for stimulus orientation, and ocular dominance.



75

75

Orientation columns

Neurons with the same orientation preference (i.e., vertical) are grouped together into orientation columns. Each column contains a few hundred cells and is 50-75µm wide;

Moving from one column to the adjacent one, orientation preference changes systematically by 10-15 °, both clockwise and counterclockwise, completing a 180 ° cycle (12 steps) every 750-1000µm.

The set of columns corresponding to a complete sequence of orientations (a period) is called a hypercolumn.

There are approximately 3-4 thousand hypercolumns, each monitoring a position of the visual field, in accordance with the retinotopic topology.

76

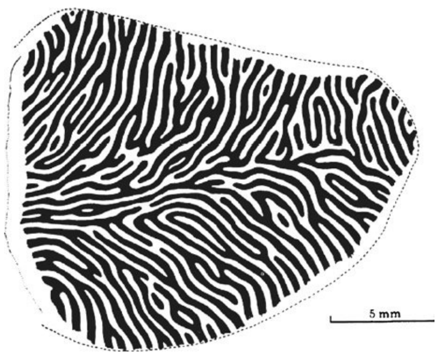
76

Ocular dominance columns

The ocular dominance columns group neurons that respond more vigorously to stimuli presented to one of the two eyes. They are stripes with an average width of approximately 750 µm, running tangentially for various mm.

The ocular dominance columns reflect the segregation of inputs from different layers of the LGN, which receive inputs from retinal ganglion cells located in the ipsilateral or contralateral retina.

In tangential penetrations of V1, the dominance columns of the left and right eye have been found to alternate regularly with a periodicity of 750 to 1,000µm.



Ocular dominance columns in primary visual cortex (V1) of macaque monkey shown in tangential section. Regions receiving input from one eye are shaded black and regions receiving input from the other eye are unshaded. The dashed line signifies the border between areas V1 and V2 (taken from Hubel and Wiesel, 1977).

77

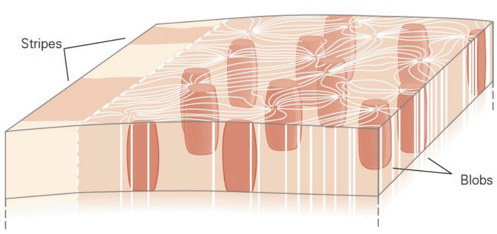
77

Blob e interblob

The columns of orientation and ocular dominance include groups of neurons that are poorly selective for orientation (they have circular receptors) but with strong preferences for the color of the stimulus.

These cell groups are located in the superficial layers (II and III) of V1. They are detectable by a specific marker for the cytochrome oxidase (CO) enzyme, which distributes in a regular pattern of regions defined as blobs (CO rich and color responsive) separated by interblob areas (CO poor and orientation responsive) .

D Blobs, interblobs (V1), and stripes (V2)



78

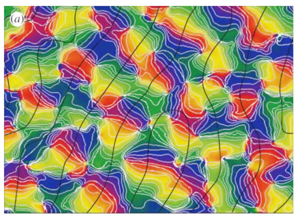
78

Pinwheels

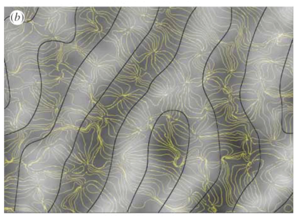
More recently, optical imaging technique has enabled to visualize a surface representation of the orientation and ocular dominance columns in living animals.

The cycles of orientation columns form various structures, from parallel stripes to pinwheel-like shapes.

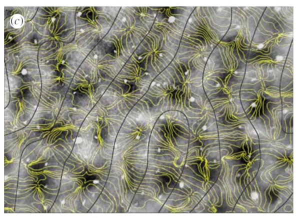
Sharp jumps or singularities (discontinuities) in orientation preference occur at the pinwheel centers and cause "fractures" in the orientation map.



Orientation columns



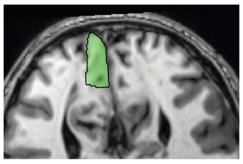
Ocular dominance columns

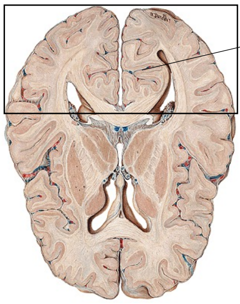


Blobs

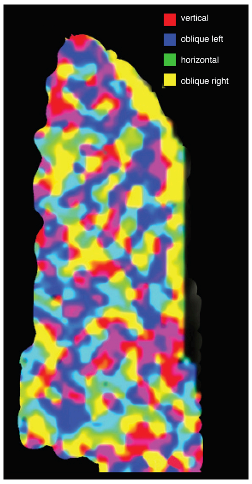
79

79

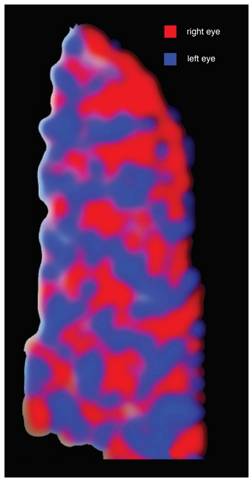




Calcarine fissure



Orientation colums



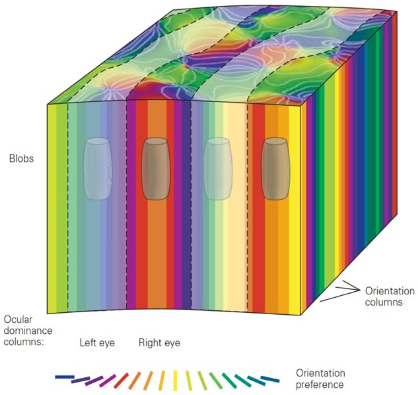
Ocular dominance columns

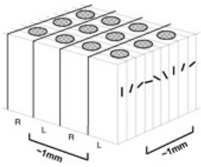
Yacub et al. (2008) showed, by using high-resolution (7-T) fMRI, the existence and spatial features of orientation selective columns in humans.

80

80

Ice cube model (Hubel e Wiesel, 1977)





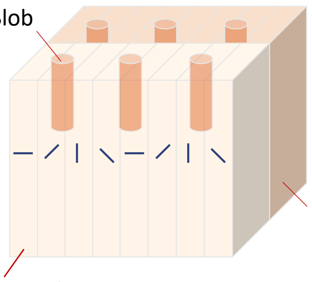
A region of cortical tissue of about 1mm contains two orientation hypercolumns (a complete cycle of selective vertical columns for orientation), one for the left eye and one for the right that alternate regularly, blob and interblob. This computational module contains all the anatomical-functional types of V1 neurons, and would be repeated thousands of times to cover the entire surface of the visual field.

81

81

The functional organization of the primary visual cortex is therefore based on two systems running orthogonally to each other:

orientation system
ocular dominance system



1mm size

similar columns closed

82

Area V2

In area V2, thick and thin dark stripes separated by pale stripes are evident with cytochrome oxidase labeling.

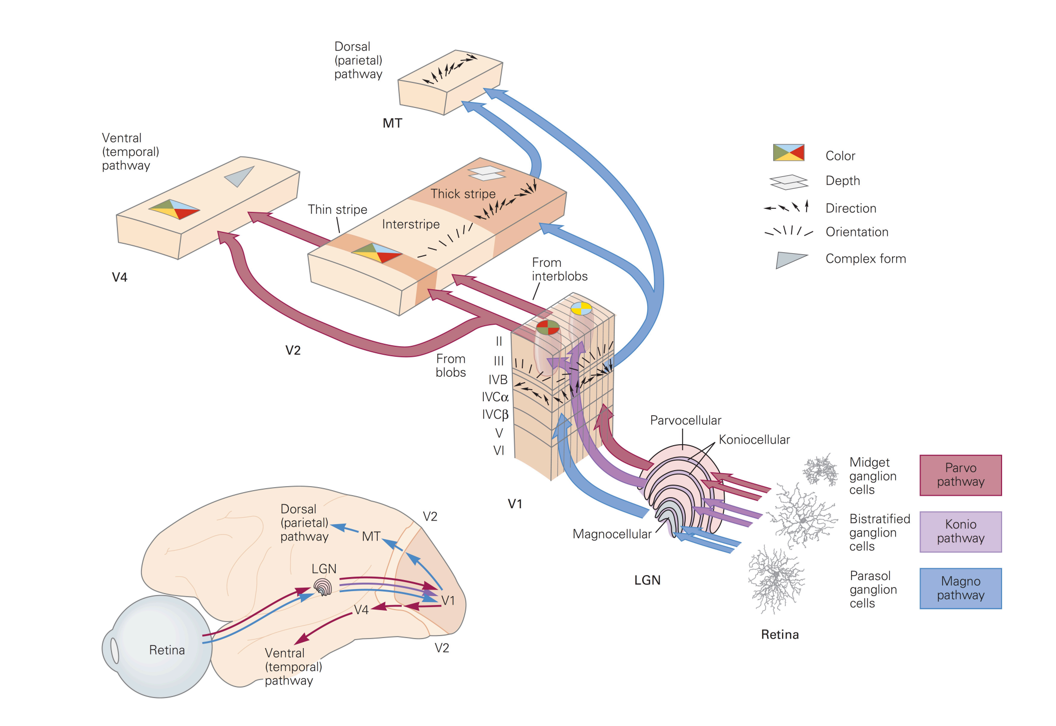
- The thick stripes contain neurons selective for direction of movement and for binocular disparity;
- The thin stripes contain cells specialized for color.
- The pale stripes contain orientation-selective neurons.

For every visual attribute to be analyzed at each position in the visual field, there must be adequate tiling, or coverage, of neurons with different functional properties.

Any given position in the visual field can therefore be analysed adequately in terms of the orientation of contours, the color and direction of movement of objects, and stereoscopic depth by a single computational module.

83

83



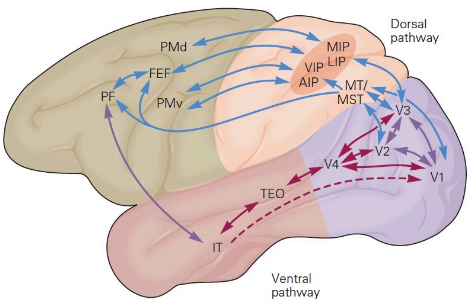
84

Beyond V1 are the extrastriate visual areas (more than 30 areas in macaques), a set of higher-order visual areas organized as neural maps of the visual field.

Visual areas are organized in two hierarchical pathways, a ventral pathway involved in object recognition and a dorsal pathway dedicated to the use of visual information for guiding movements.

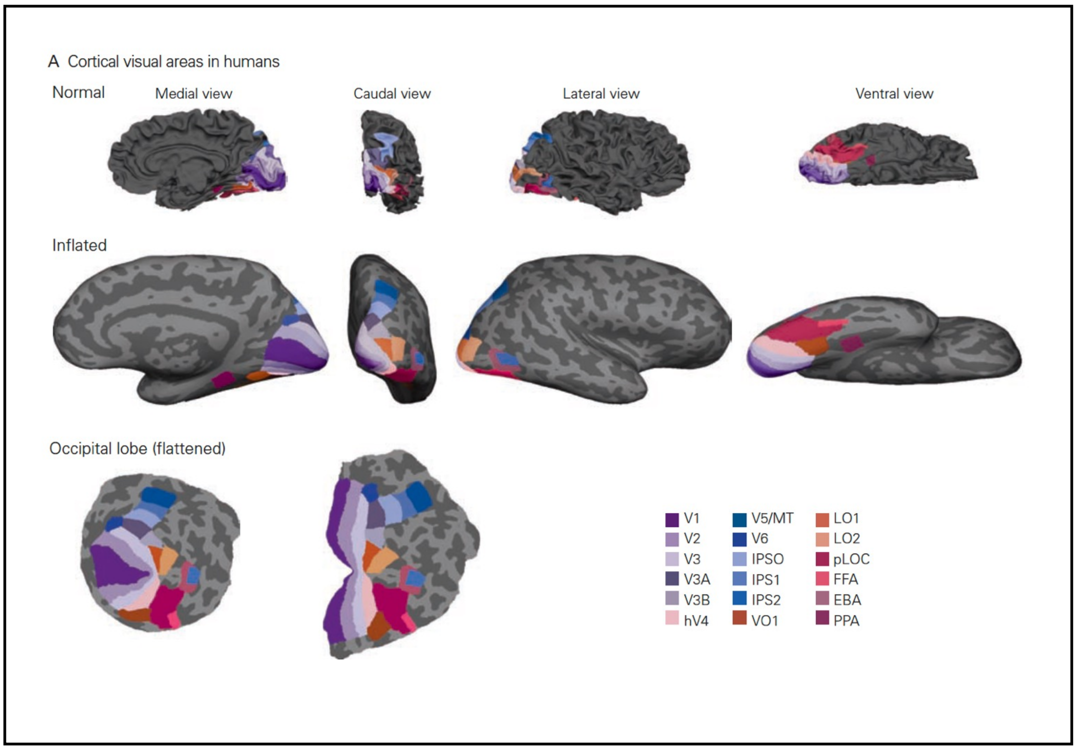
efficiency

The ventral or object recognition pathway extends from V1 to the temporal lobe
The dorsal or movement-guidance pathway connects V1 with the parietal lobe and then with the frontal lobes.



Ungerleider & Mishkin, Two cortical visual systems. 1982

85


A Cortical visual areas in humans
Normal
Medial view
Caudal view
Lateral view
Ventral view
Inflated
Occipital lobe (flattened)
V1
V2
V3
V3A
V3B
hV4
V5/MT
V6
IPSO
IPS1
IPS2
VO1
LO1
LO2
pLOC
FFA
EBA
PPA

86

87

3D vision

Psychophysical studies indicate that 3D vision is based on:

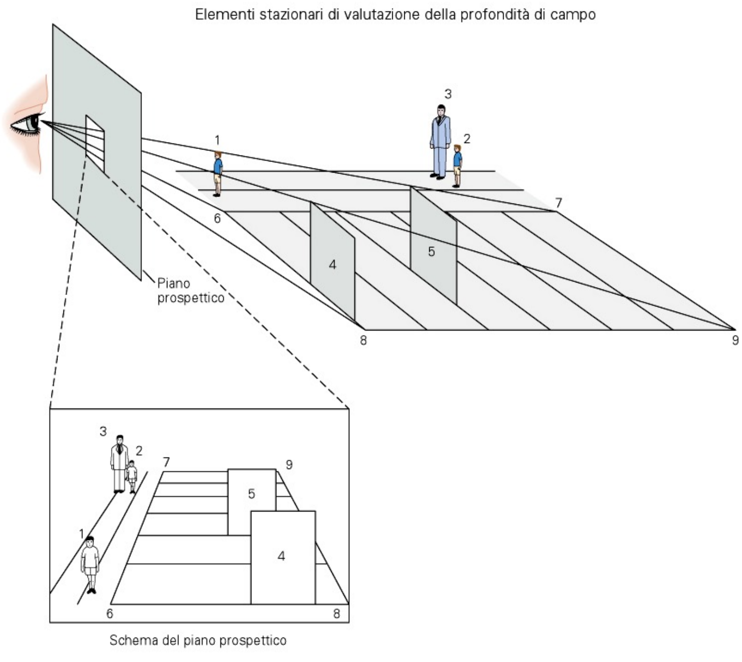
monocular elements
stereoscopic elements

Monocular elements are able to create the sense of depth for distances greater than 30 meters:

- familiarity with the object
- interposition
- linear perspective
- size of objects
- distribution of shadows and lighting
- parallax movement

parallax movement

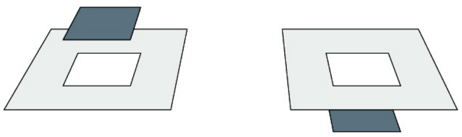
88



89

1	0	1	0	1	0	0	1	0	1
1	0	0	1	0	1	0	1	0	0
0	0	1	1	0	1	1	0	1	0
0	1	0	W	A	A	B	B	0	1
1	1	1	Z	B	A	B	A	0	1
0	0	1	Z	A	A	B	A	1	0
1	1	1	W	B	B	A	B	0	1
1	0	0	1	1	0	1	1	0	1
1	1	0	0	1	1	0	1	1	1
0	1	0	0	0	1	1	1	1	0

1	0	1	0	1	0	0	1	0	1
1	0	0	1	0	1	0	1	0	0
0	0	1	1	0	1	1	0	1	0
0	1	0	A	A	B	B	Z	0	1
1	1	1	B	A	B	A	W	0	1
0	0	1	A	A	B	A	W	1	0
1	1	1	B	B	A	B	Z	0	1
1	0	0	1	1	0	1	1	0	1
1	1	0	0	1	1	0	1	1	1
0	1	0	0	0	1	1	1	1	0

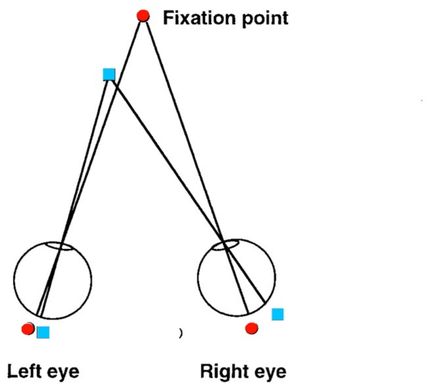


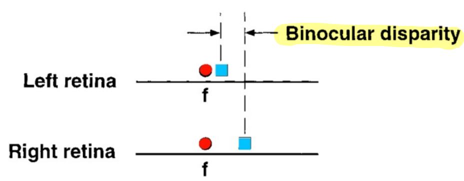
The experiments carried out by Bela Julesz with random dot stereograms show that stereopsis is a visual faculty separate from the perception of forms.

not completely random → little shift

90

When we fix our eyes on a point, the convergence of the eyes causes the fixed point to fall on identical areas of both retinas. Points lying outside the fixation plane stimulate slightly different parts of each eye's retina creating binocular disparity.

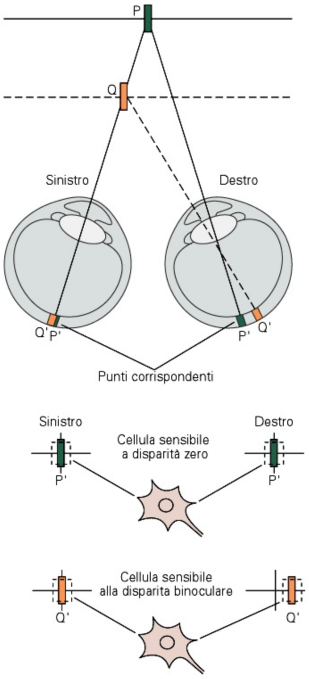




red is fovea

91

neurons catch this disparity

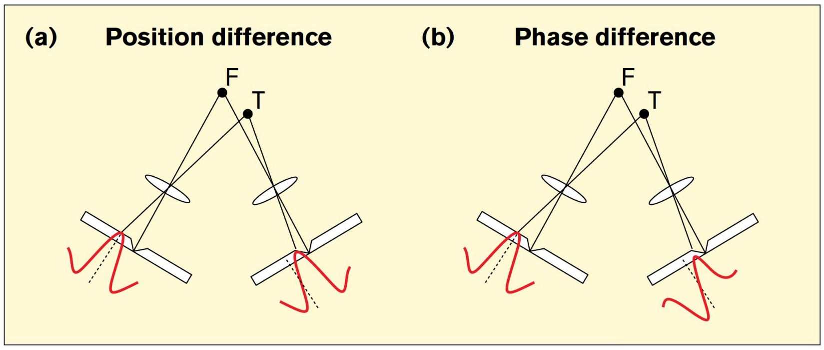


disparity is computed by V1

The primary visual cortex is the first site of the central nervous system where information from the two eyes converges on a single neuron.

Some neurons of V1 (simple and complex cells) are specifically selective for the horizontal disparity of the retinal images, that is, they respond better to stimuli localized to particular depths of the visual field.

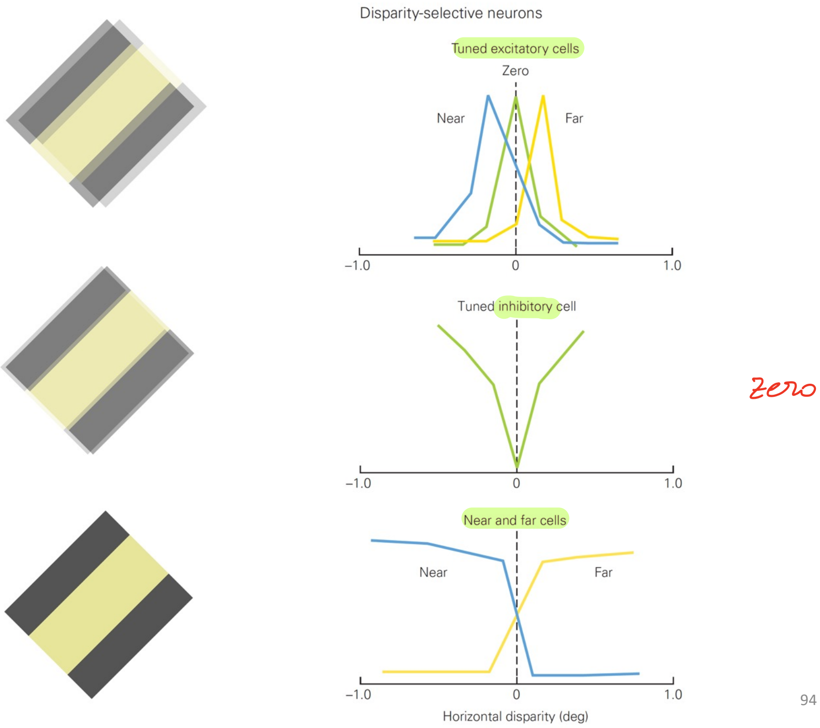
92



The dotted line marks the equivalent retinal location in the two eyes.
Neurons with receptive fields in both eyes could detect this disparity in two ways.
(a) Position difference: the right eye receptive field is an exact copy of the left eye receptive field, but in a different retinal location.
(b) Phase difference: the envelope enclosing the right receptive field profile sits in the same position as for the left receptive field, but within the envelope, the right receptive field has a different structure, responding best to white light on the right hand side.

93

93

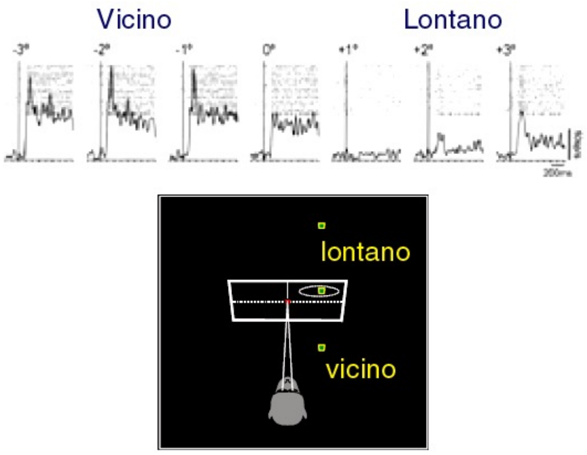


zero disparity

94

In addition to V1, neurons sensitive to retinal disparity are also observed in V2, V3 and particularly in MT and MST.

MST neurons appear sensitive to stimuli that move in particular direction at specific depths of field.



95

V5 is particularly important for stereopsis

Visual motion

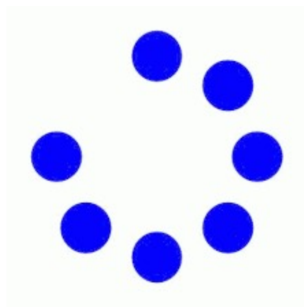
- Contributes to the recognition of objects (background figure segregation);
- It is used to establish the depth or distance of an object (parallax);
- Helps in navigation and interaction with the outside world;
- It serves to direct attention.

96

The movement of visual images is essentially analyzed from the dorsal visual pathway.

The (illusory) perception of apparent movement constitutes proof of the existence of specific mechanisms for the analysis of visual movement.

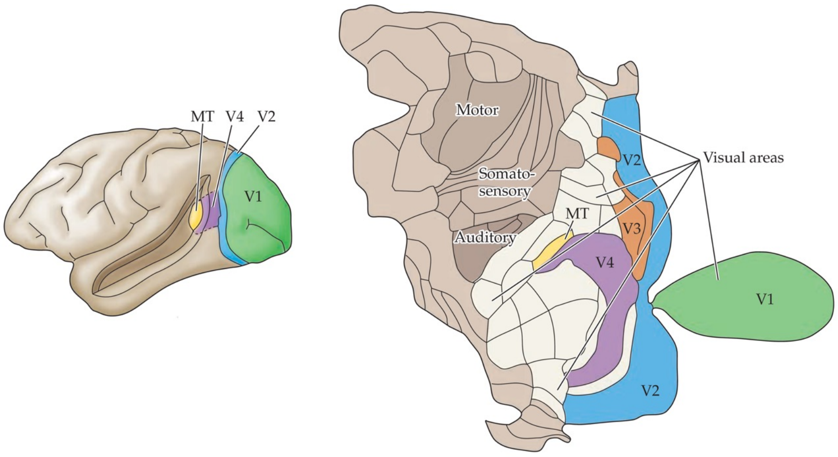
The neurons of the IVB layer of V1 are particularly sensitive to the direction of movement of a stimulus within the receptive field;



phi phenomenon

97

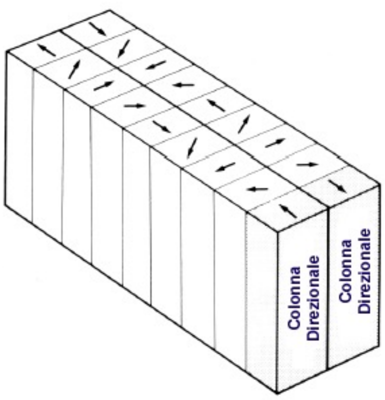
A key role in the perception of movement is played by the MT area, (Middle Temporal) a small region of the dorsal path, located in the posterior part of the superior temporal sulcus, and first described in 1971 independently by J. Kaas and S. Zeki .



98

MT shows a columnar cortical organization (analogous to V1):

the neurons present in each column selectively respond to stimuli in the same region of the visual field (same RF) and with the same direction of movement. The preferred direction changes gradually from one column to the contiguous one.

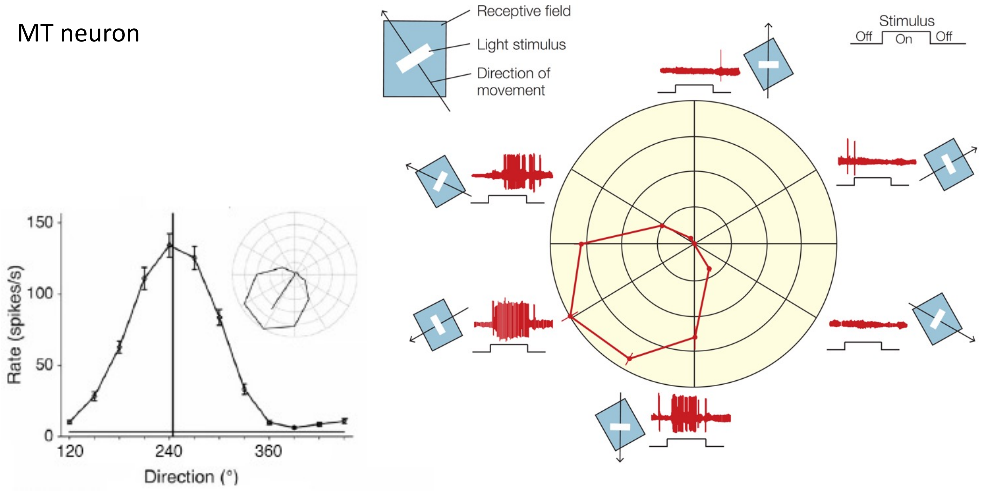


In addition to the direction of movement, MT cells are selective for the speed of the stimulus. Conversely, they do not selectively respond to shape and colors.

99

Example of MT neuron responding maximally to downward and leftward directions of movement (225 ° = preferred direction). Note that the neuron also responds, but less, to directions contiguous to the preferred one.

MT neuron

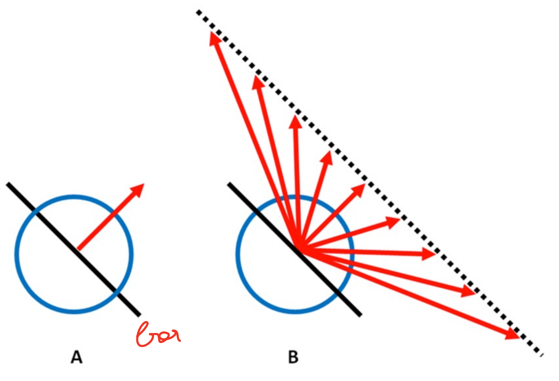


100

If a moving contour is seen through a small aperture, only the velocity component perpendicular to the contour can be detected through the opening (Wallach 1935).

The velocity component parallel to the orientation of the contour can only be observed at the extremes.

Since neurons in the primary visual cortex (V1) have small receptive fields, this ”aperture problem" (Marr, 1982) must be overcome if the visual system is to function properly.



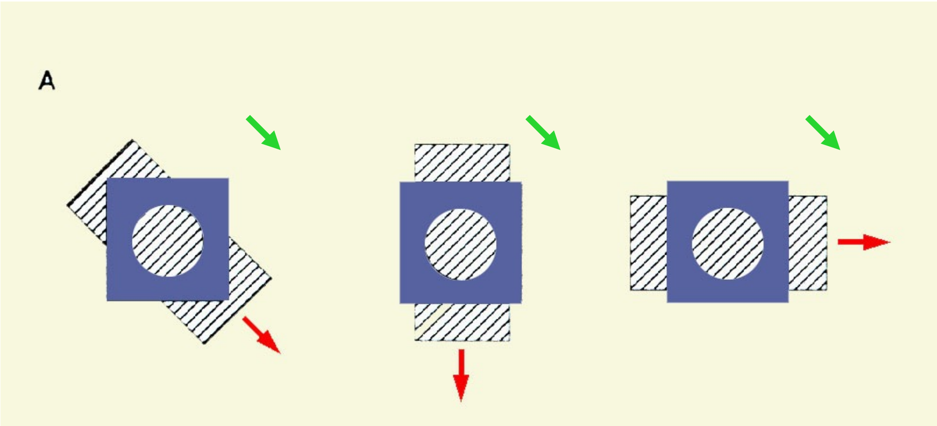
we see just ⊥ movements

// we don't perceive anything

101

101

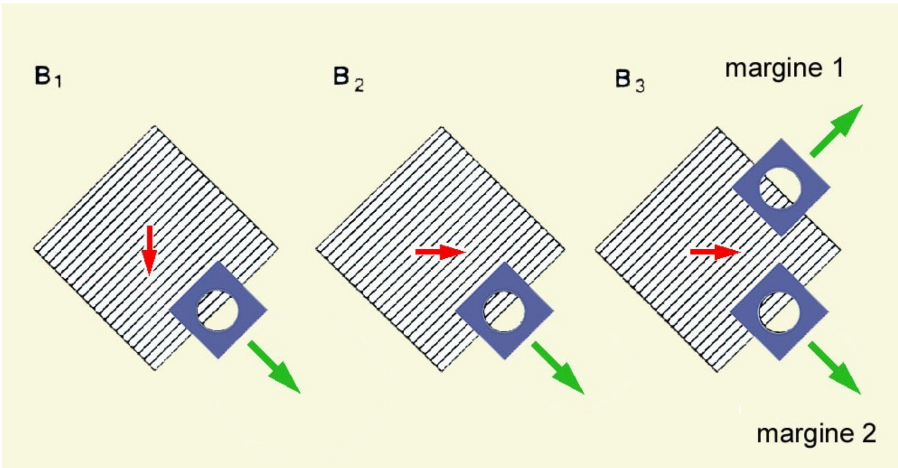
The aperture problem



Due to the aperture problem, movements of a stimulus in three different directions (red arrows) are always interpreted by the neuron as movement in a single direction (green arrows), perpendicular to the orientation of the contour.

102

The aperture problem



if we read 2 RF we can catch something

The responses of the neurons in B1 and B2 are highly ambiguous, and fail to detect the true direction of the figure's movement.

In B3, the aperture problem is solved by integrating the response of two or more neurons whose receptive fields are placed in different points of the figure.

103

Visual motion processing occurs in two stages

To solve the aperture problem, it has been hypothesized by some researchers that the visual processing of moving stimuli occurs through two successive cortical stages:

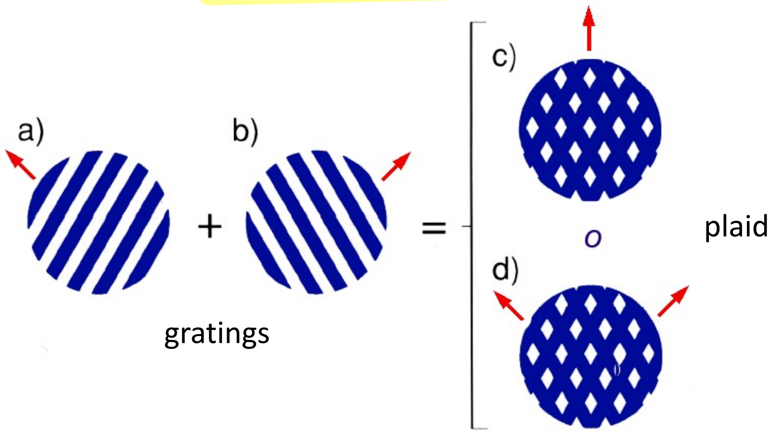
1) In the first stage, the individual components of the stimuli are analyzed, through the responses of neurons capable of signaling only the movement of the local components, perpendicular to the movement of the contour.

2) In the second stage, higher order neurons would integrate the different local components of the stimulus, analyzed by the neurons of the previous stage.

The integrated signal produced by higher-order neurons corresponds to the real direction of motion of the object and to the perception that an observer has of it.

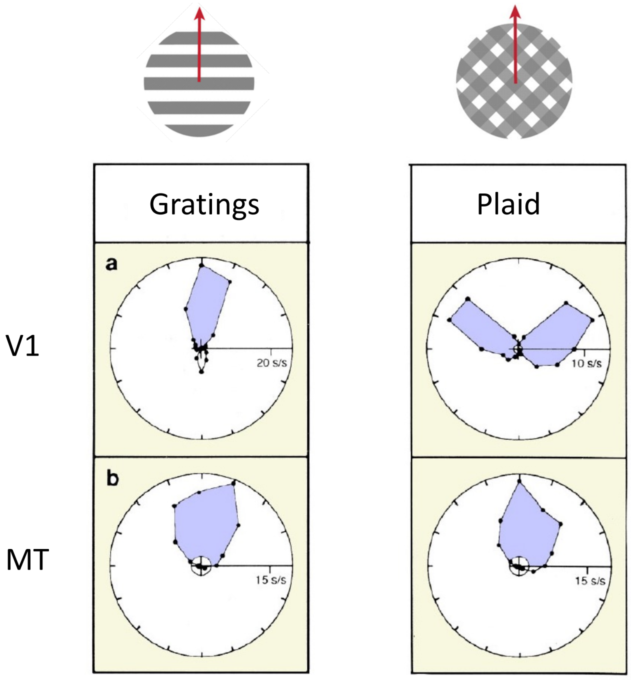
104

This hypothesis was evaluated by Movshon and colleagues (1982) in the MT area of a monkey, through the use of complex stimuli ('PLAID') obtained from the superimposition of sinusoidal gratings directed in opposite directions.



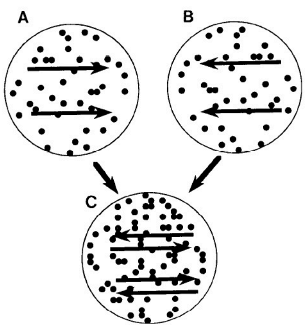
By superimposing two gratings (a and b) that move in the opposite direction (top-right and top-left) we obtain a gratings or PLAID that appears to move clearly in only one direction (top, resulting from the sum of the directions of the two gratings, as in c).

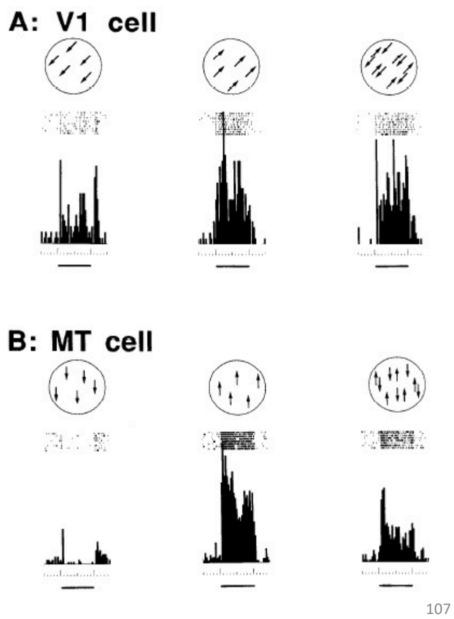
105



106

The response of area MT and VI neurons to transparent motion

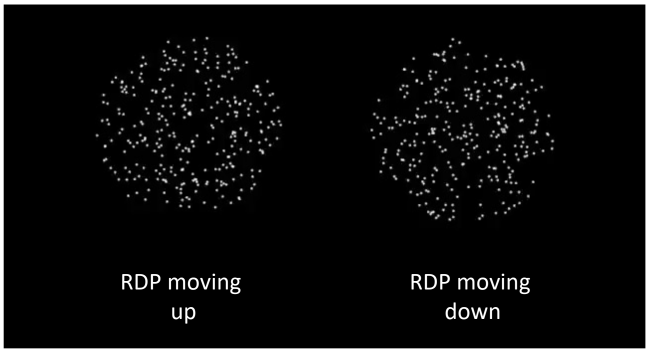




107

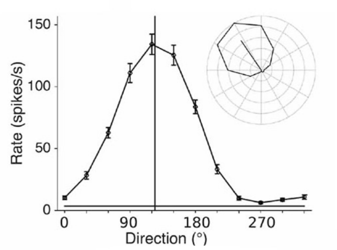
Selective visual attention can be directed toward a specific region of space (i.e., space-based attention),

Middle temporal (MT) area is crucial for the perception of visual motion.
MT neurons show direction tuning curves (bell-shaped response profiles) depending on the direction of motion of the stimulus (random dot pattern, RDP).



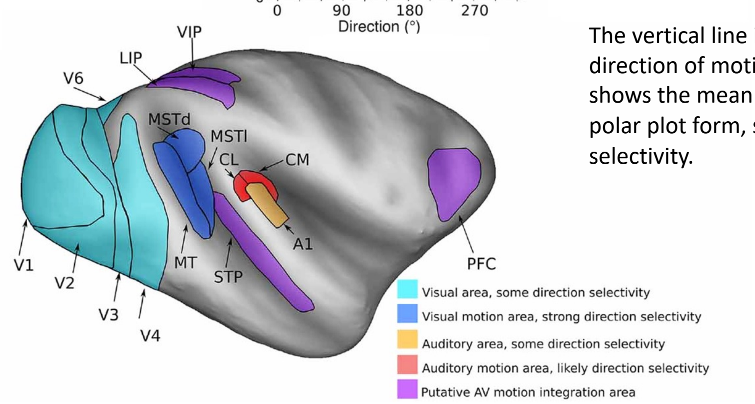
Treue & Martınez-Trujillo, Nature, 1999

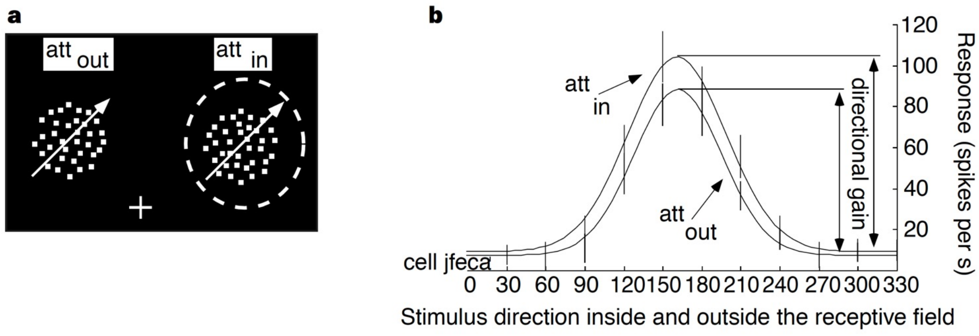
Spatial attention modulates the responses of neurons selective to direction of motion in area MT and MST.



A typical visual direction tuning curve from a neuron in the monkey visual cortex (area MT) in response to a moving dot stimulus.

The vertical line indicates the preferred direction of motion, and the inset shows the mean spiking responses in polar plot form, showing clear direction selectivity.





Experiment 1: Effect of spatial attention

a, One RDP was presented inside the RF (dashed circle) while the other was presented in the opposite hemifield. In a given trial, both RDPs moved in the same of 12 possible directions.

b, The upper curve shows the response when the monkey was attending to the stimulus inside the receptive field (marked att in), and the lower curve plots the responses when the monkey was attending to the stimulus outside the receptive field (marked att out).
These tuning curves show an increase in directional gain in the attended condition.

Treue & Martınez-Trujillo, Nature, 1999

110

110

Attending to a given direction enhances the responses of neurons whose preferred direction aligns with the attended direction and reduces the responses of those neurons preferring the opposite direction.

These findings shows that spatial and feature-based attention represent separate (and summable) processes that have a multiplicative effect on the responses of neurons.

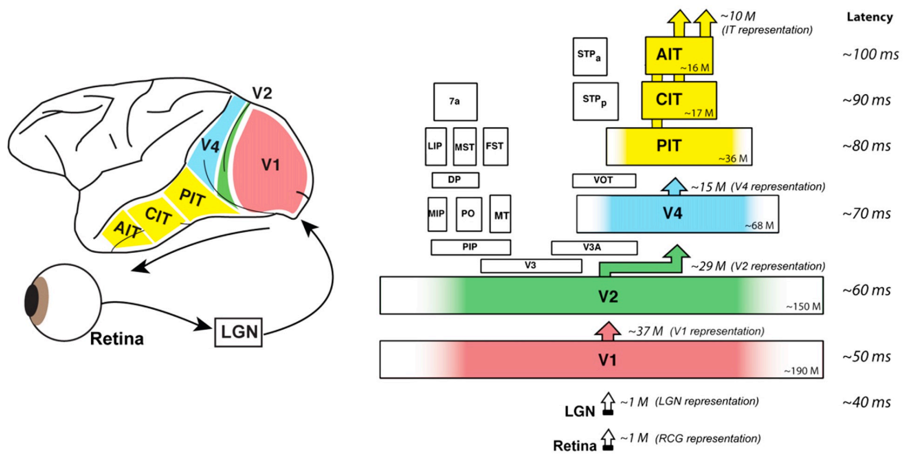
Treue & Martınez-Trujillo, Nature, 1999

111

111

112

112



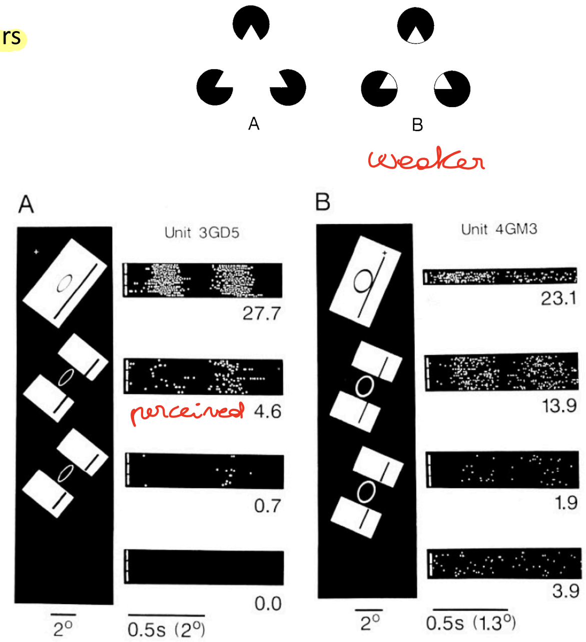
113

113

Neural responses to illusory contours

Figures in which humans perceive illusory contours evoke responses in cells of area 18 (V2) in the monkey visual cortex.
Modifications that weakens the perception of contours also reduces the neuronal responses.

In contrast, cells in V1 (area 17) were apparently unable to see these contours.



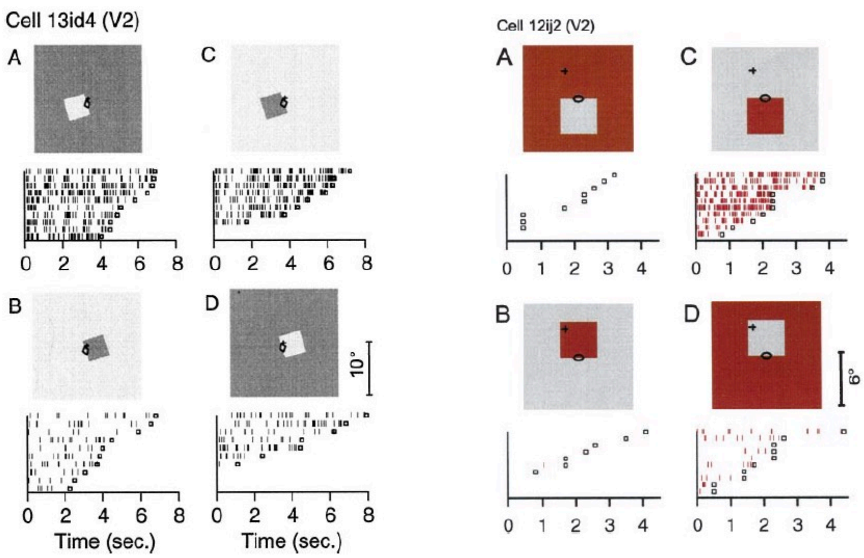
V2 RF

Von der Heydt et al., Science, 1984

114

114

Coding of border ownership in monkey area V2



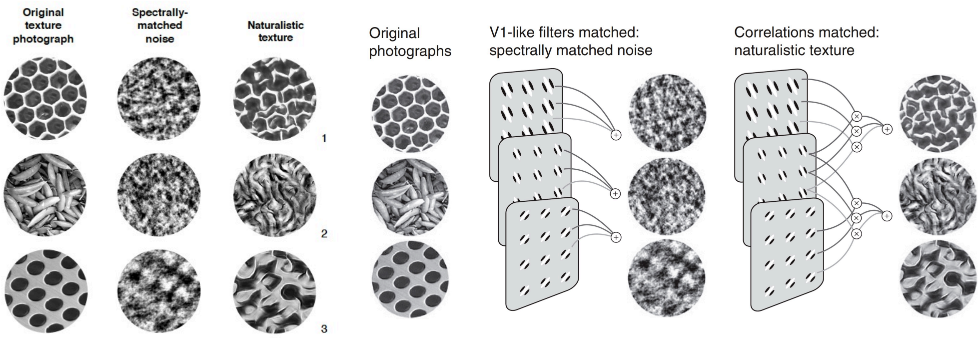
Zhou et al., J Neurosci, 2000

115

115

There is no generally accepted account of the function of V2, partly because no simple response properties robustly distinguish V2 neurons from those in V1.

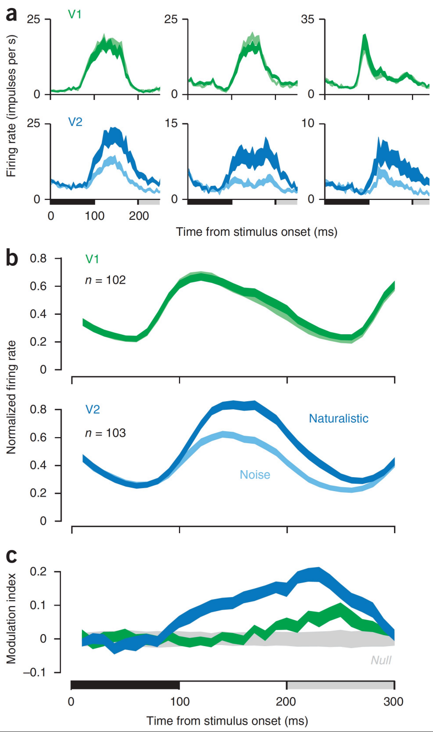
Stimuli replicating the higher-order statistical dependencies found in natural texture images were used to stimulate macaque V1 and V2 neurons.



Freedman et al., Nat. Neurosci., 2013

116

116



Neuronal responses to naturalistic textures differentiate V2 from V1 in macaques.

a) Time course of firing rate for three single units in V1 (green) and V2 (blue) to images of naturalistic texture (dark) and spectrally matched noise (light). Black bar indicates the presentation of the stimulus; gray bar indicates the presentation of the subsequent stimulus.

b) Time course of firing rate averaged across neurons in V1 and V2. Each neuron's firing rate was normalized by its maximum before averaging.

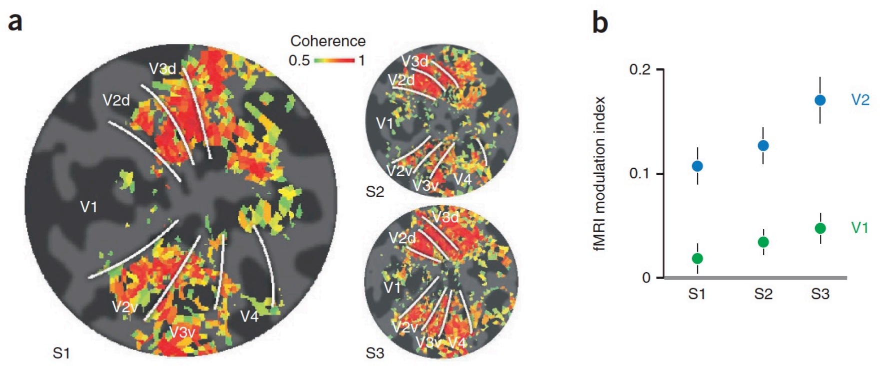
c) Modulation index, computed as the difference between the response to naturalistic and the response to noise, divided by their sum. Modulation was computed separately for each neuron and texture family, then averaged across all neurons and families.

Freedman et al., Nat. Neurosci., 2013

117

117

Naturalistic texture stimuli modulate the responses of neurons in area V2, while having only a minimal effect on neurons in area V1.
These modulations were similar and substantial in both macaques and humans.

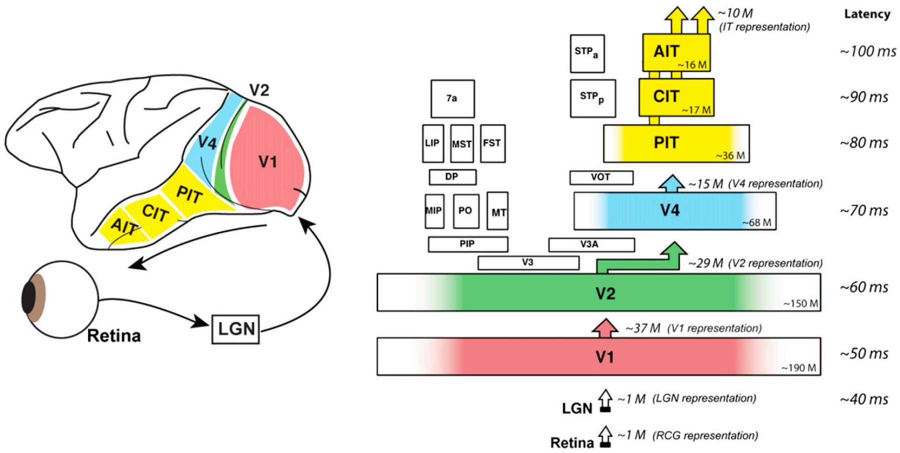


fMRI responses to naturalistic textures differentiate V2 from V1 in humans.

Freedman et al., Nat. Neurosci., 2013

118

118

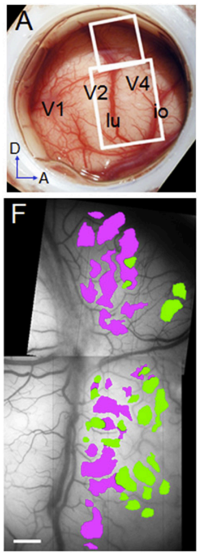


Visual area V4 is an intermediate cortical area in the ventral visual pathway.

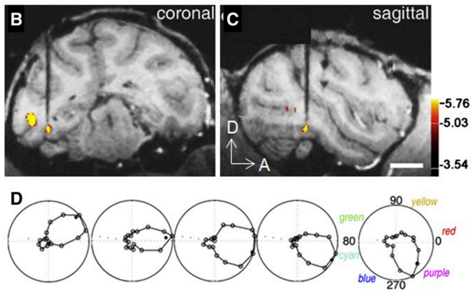
It considered to be crucial for visual object recognition and visual attention.
A primary role of V4 is to facilitate figure-ground segmentation of the visual scene, thus enabling both bottom-up and top-down visual processes.

119

119



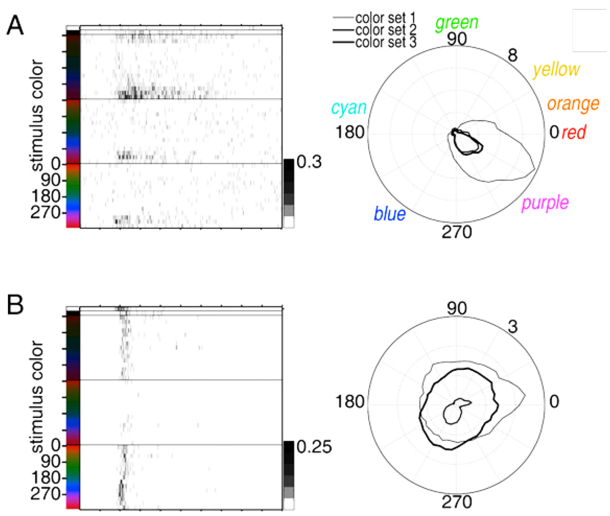
Recent advances in fMRI and optical imaging methods have revealed that V4 is not a homogenous area and may comprise segregated functional domains, dubbed globs (in analogy with the color blobs of V1), specialized for color processing, and interglobs, for orientation and shape processing



(F) illustrates segregation of color/ Lum and orientation preference bands in area V4

120

Figure shows color-response measurements for a typical cell in a glob (A), and a cell in the interglob (B) region. Left panels show time histograms to an optimally shaped bar of various colors; right panels show the color tuning in polar coordinates.



Most glob cells were excited by a specific hue, as shown by hue tuning in the polar plots.

This hue selectivity was luminance invariant. That is, hue tuning does not change with change in luminance between stimulus and background.

Conway et al., Neuron, 2007

121

Color constancy

It refers o the effect whereby the perceived or apparent color of a surface remains constant despite changes in the intensity and spectral composition of the illumination.

The following shows a striking example in which the same physical patch appears yellow when the scene appears to be illuminated by blue light but appears blue when the scene appears to be illuminated by yellow light.

Surrounded by white, the patch appears gray.

This example illustrates the potency of color contrast in generating color perception.

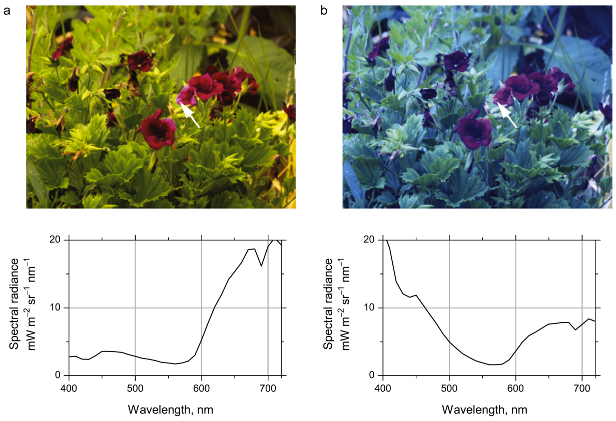
122

122

Color constancy illustration: the patches indicated by the asterisks in the images are the same

123

123



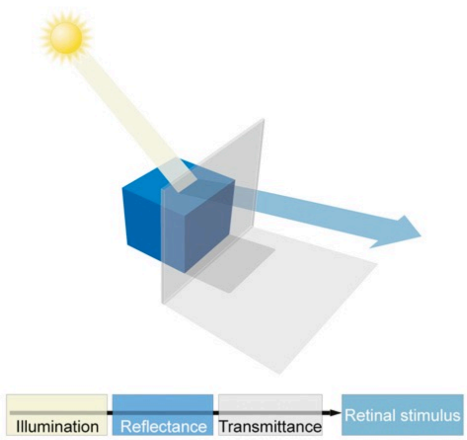
Images of a pelargonium under sunlight and skylight with respective correlated color temperatures (a) 4000 K and (b) 25,000 K and the corresponding radiance spectra reflected from the arrowed region of a petal.

Constancy is achieved because the brain determines an object's color from a local comparison of the light reflected from the object with the light reflected from adjacent regions of the scene.

124

124

Retinal luminance in visual stimuli (the inverse optics problem)



Retinal luminance, the spectral light reflected from an object is a product of two variables, Illumination (the spectral content of the illuminant, i.e., the light source) and reflectance (the absorptive properties of the object, as well as a variety of other factors (e.g., atmospheric transmittance).

These physical determinants of retinal luminance values are conflated in visual stimuli and cannot be disentangled by any algorithmic process.

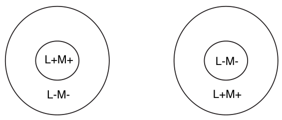
125

125

Retina and LGN: Color Opponency

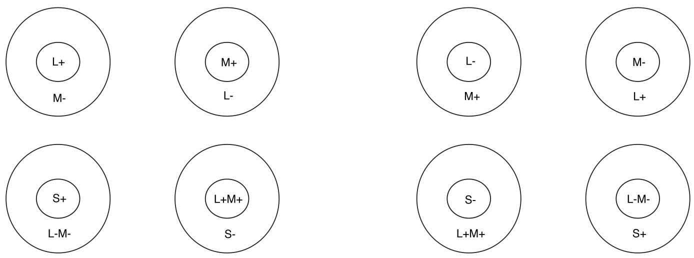
The principle dimensions in the space of cone excitations produced by natural objects are:
1) a luminance axis where the L- and M-cone signals are added;
2) a red-green opponent axis where the difference of L- and M-cone signals is taken;
3) a blue-yellow color opponent axis where the S-cone signal is differenced with the sum of the L- and M-cone signals;

M-cells Magnocellular (luminance cells)



Receptive fields of single-opponent cells in the retina and LGN are not capable of color contrast.

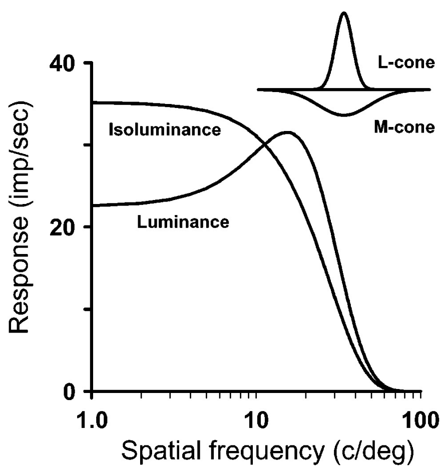
P-cells Parvocellular (color-luminance cells)



126

126

Spatial frequency tuning curves for a typical P-/parvocellular color-luminance cell



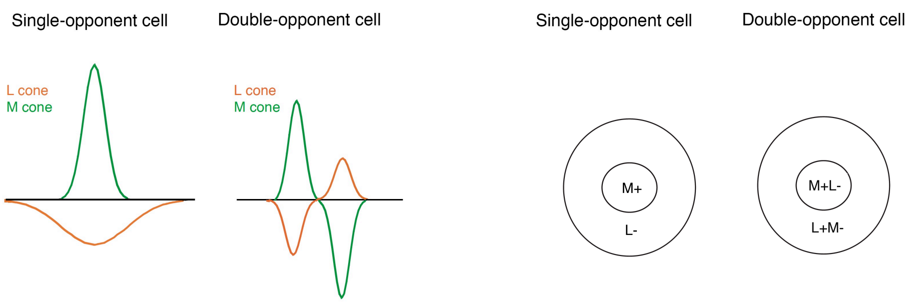
These cell perform “double duty,” carrying information for both color vision and achromatic spatial vision, which are “de-multiplexed” by circuitry at higher levels of processing in the geniculostriate pathway.

These cells show color opponency to large spots and luminance sensitivity to small spots.

127

127

Primary visual cortex: calculating color contrast



Single-opponent cells cannot resolve color contrast.
Double-opponent cells may represent the neural basis for color contrast and color constancy.

The optimal stimulus for a Green-on double-opponent cell is a green spot on a red background. Because of their specialized receptive-field structure, double-opponent cells are candidates for the neural basis for color contrast and color constancy.

128

128

VISION → perception and motion

Object identification and categorization

The visual experience of the world is fundamentally centered on objects.

By visual object we mean a set of visual characteristics (e.g., visual features) grouped or joined perceptually in discrete units on the basis of the organizational principles of the Gestalt, such as proximity, similarity, closure, good continuation, good form, connection, etc.

By visual recognition we mean the ability to assign a verbal label (e.g., a name) to objects in the visual scene.

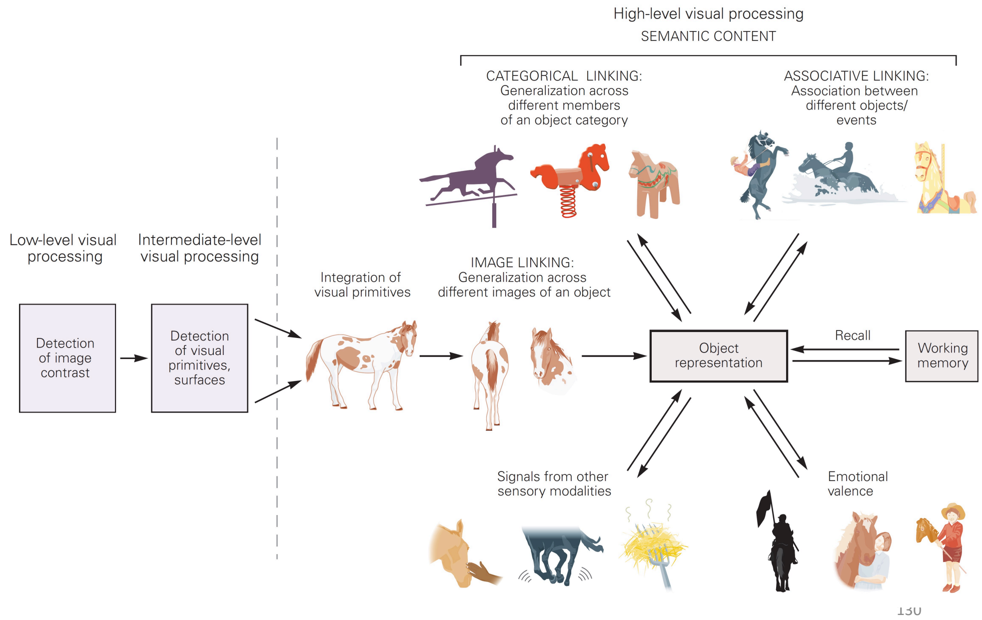
There are at least two possible object recognition tasks, distinguished by level of specificity: identification and categorization.

An object can be recognized at an individual level (e.g., a Siamese cat), or at a more general categorical level, as an object belonging to a given class (a cat, a mammal, an animal, and so on).

129

129

High-level visual processing is concerned with object recognition



130

Selectivity and object constancy (or invariance)

A computational difficulty of object recognition is that it requires both:

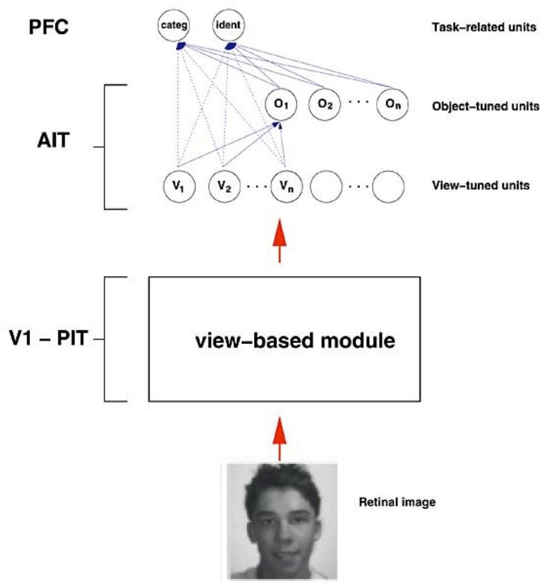
selectivity (different responses to distinct objects, such as one face with respect to another face);
and invariance with respect to image transformations (similar responses to, for example, rotations or translations of the same face);

In fact, we are able to recognize the same object even when the image it projects on the retina varies considerably.

131

131

Object recognition



(view-invariant units)

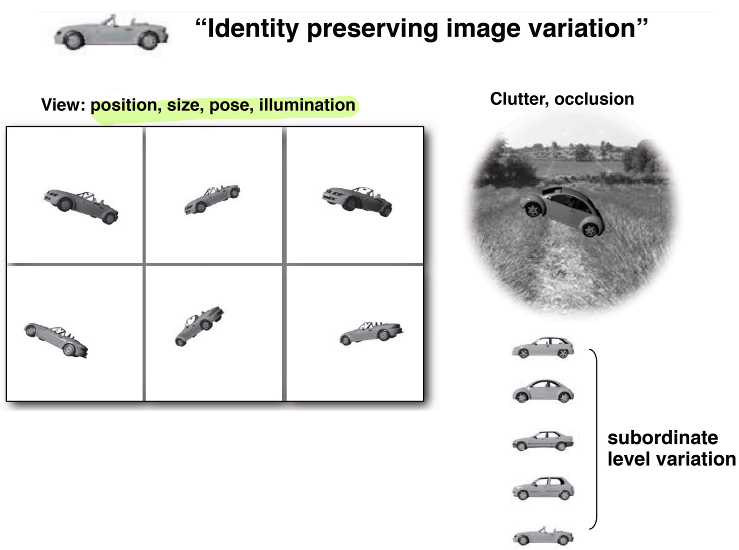
(view-dependent units)

PFC = Prefrontal cortex
AIT = Anterior IT
PIT = Posterior IT

Riesenhuber & Poggio, Nat. Neurosci. 2000

132

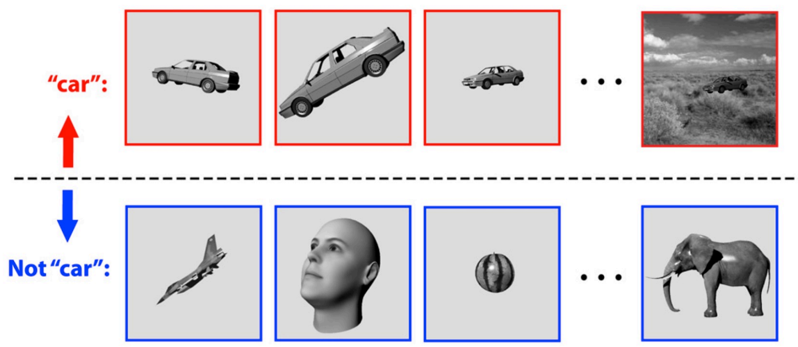
The phenomenon of perceptual constancy, or "object invariance problem", is the crucial point of object recognition, and also the main obstacle for computer vision recognition systems.



core object recognition
⇒ we are very fast to recognize objects invariant to lots of factors

133

133

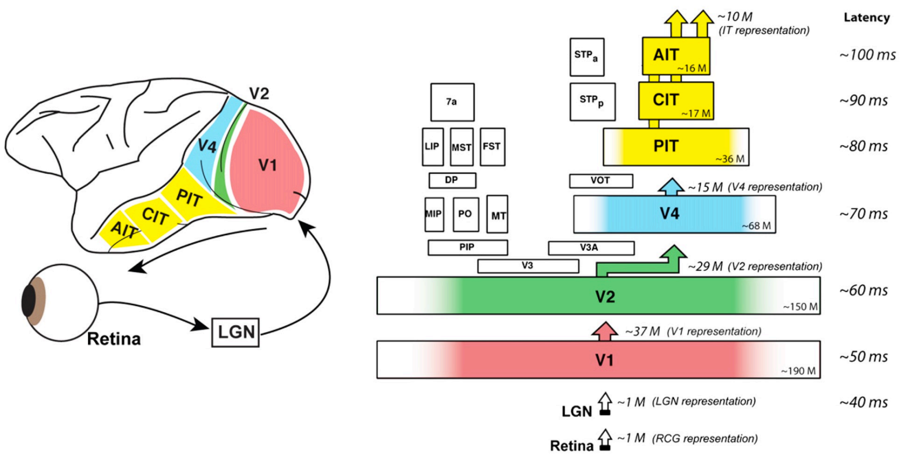


Core object recognition is the ability to rapidly (<200 ms viewing duration) discriminate a given visual object (e.g., a car, top row) from all other possible visual objects (e.g., bottom row).

Primates perform this task remarkably well, even in the face of identity-preserving transformations (e.g., changes in object position, size, viewpoint, and visual context).

134

134

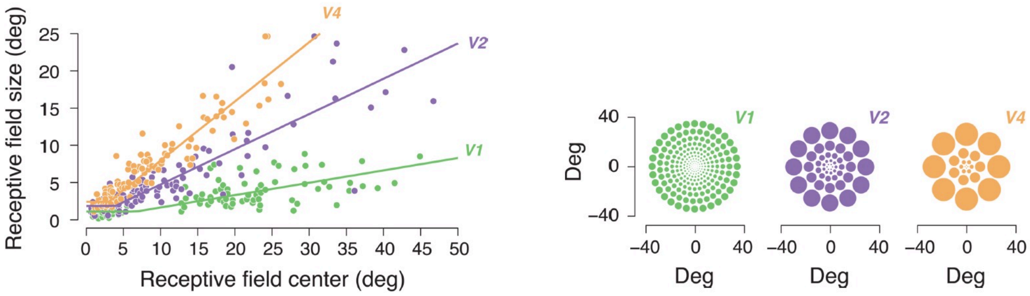


Studies in animals (e.g., primates) primarily implicates the inferior temporal cortex in object perception and recognition.

The inferior temporal cortex is a large region of the brain that includes at least two major functional subdivisions - the posterior area, or temporo-occipital cortex (TEO area), and the anterior area, or inferotemporal cortex (IT area).

135

135

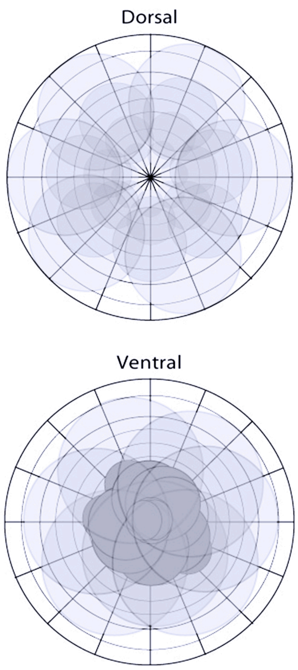


In contrast to V1, V2, V4 and partly PIT, CIT and AIT show no clear retinotopy.

Large receptive fields (from 2 to 30 degree). Most of the receptive fields include the fovea. Foveal stimuli evoke stronger responses.

Freeman and Simoncelli, Nat. Neurosci., 2013

136



Representation of the visual field in the dorsal and ventral visual pathway

The dorsal visual pathway is characterized by a more complete sampling of the visual field with wide receptive fields, which include the visual periphery, and, in some cases, sparing of foveal regions.

In contrast, the areas of the ventral visual pathway show greater focalization, which always involve coverage of the foveal regions.

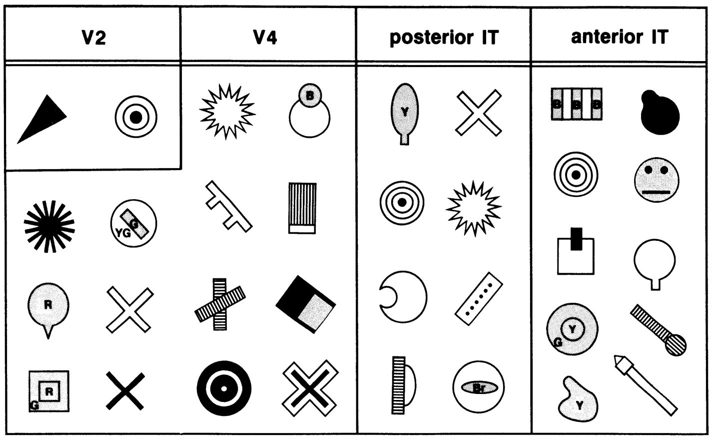
This difference would indicate a crucial role of the dorsal pathway in the exploration of salient stimuli of the visual field, while the ventral pathway would be involved in object recognition and use (exploration vs. exploitation).

Sheth and Young, Front. Integrat. Neurosci, 2016

137

137

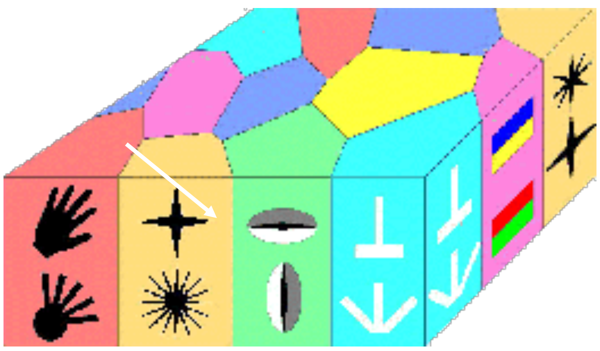
Increased complexity of effective stimuli along the ventral visual path



Kobatake and Tanaka, J. Neurophysiol., 1994

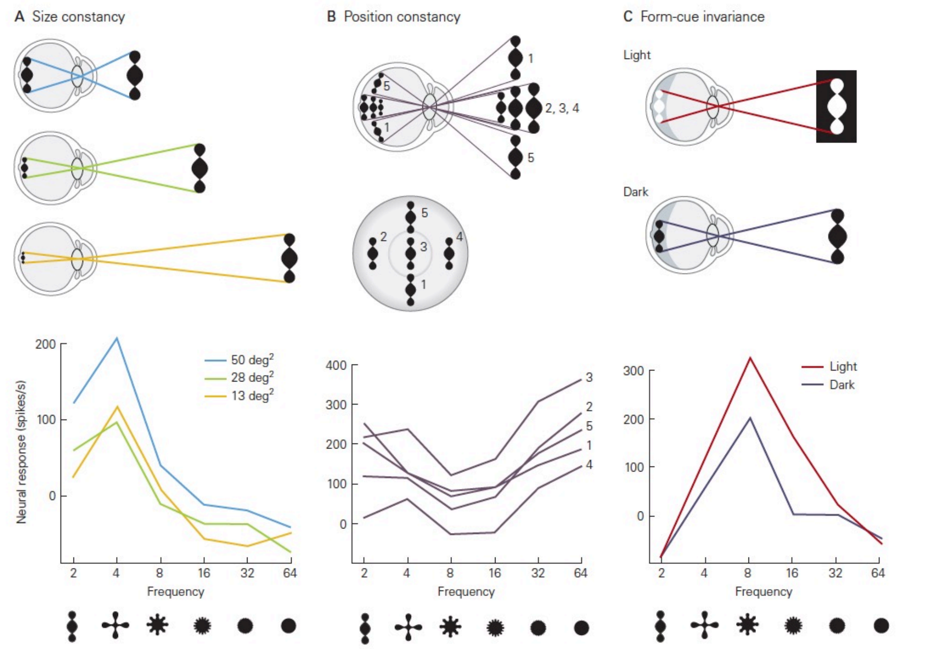
138

Tanaka (1996) suggested that responses to effective stimuli are organized in topographical order, so that two adjacent neurons are much more likely to respond similarly to a set of stimuli than neurons that are more distant.



Functional columns in IT

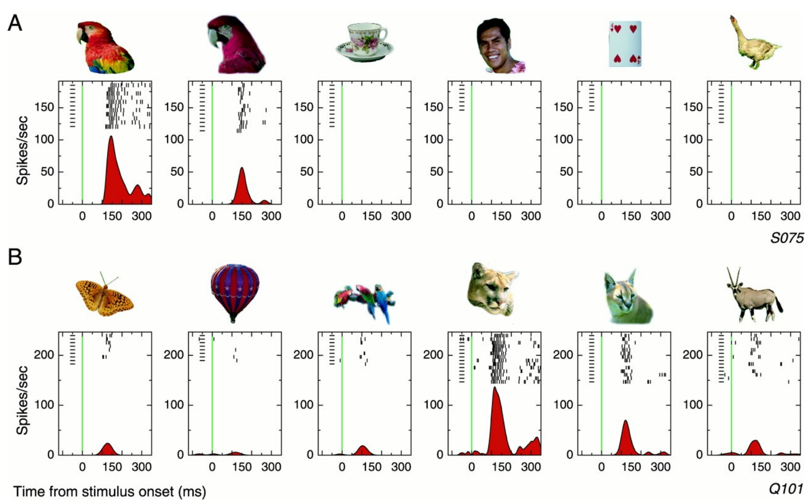
139



preference regardless to the size

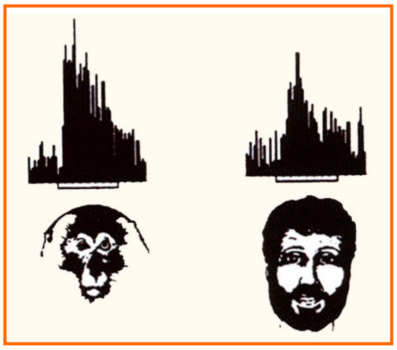
140

Neurons in the IT respond to relatively complex stimuli, often to biologically relevant objects such as human and other animal faces, hands and other parts of the body.



141

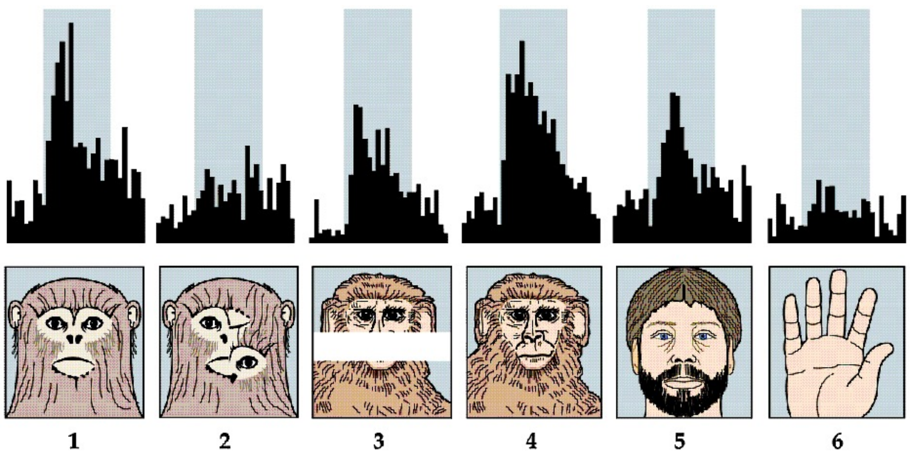
In the early 1980s, several researchers (Bruce et al., Perrett et al.) Identified in the monkey a group of IT neurons that responded selectively to faces.



Question: Are there in IT, as for faces, selective cells for the different types of objects that can be encountered in the outside world (neurons for chairs, for flowers, for cars, etc ...)?

142

Neuron that responds to faces: The neuron responds to faces of different species (1, 4, 5). The discharge is reduced if the elements of the face are mixed (2) or occluded (3). The neuron does not respond to other biologically relevant stimuli (5).



Desimone, Albright, Gross and Bruce, J. Neurosci, 1984

143

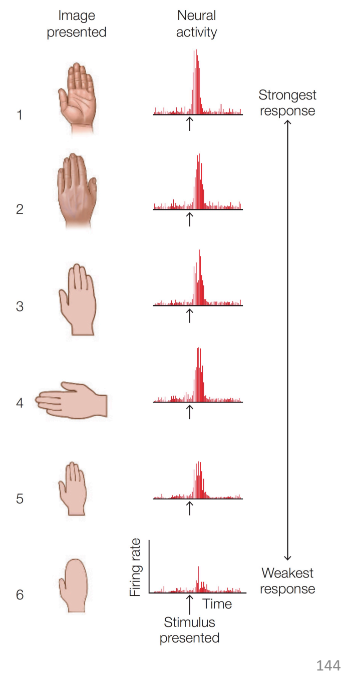
Recording of a single neuron from monkey IT cortex (Desimone et al., 1984)

This cell is activated by the vision of the human hand.

The first five images in the figure show the cell's response to various perspectives of a hand.

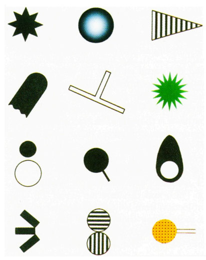
Activity is high regardless of hand orientation and only decreases slightly when the hand is noticeably smaller.

The sixth image shows that the response decreases if the stimulus has the same shape, but does not have well-defined fingers.



144

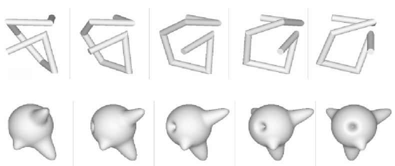
Hung et al., Science, 2005

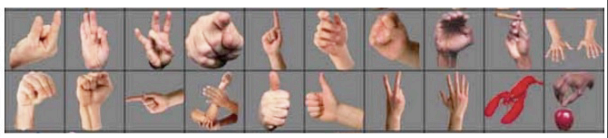


Tanaka, Science, 1993

neurons respond to parts as we have a dictionary for each part

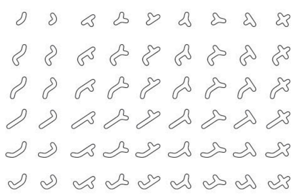
Logothetis et al., Curr, Biol., 1995





Kiani et al., J. Neurophysiol., 2007

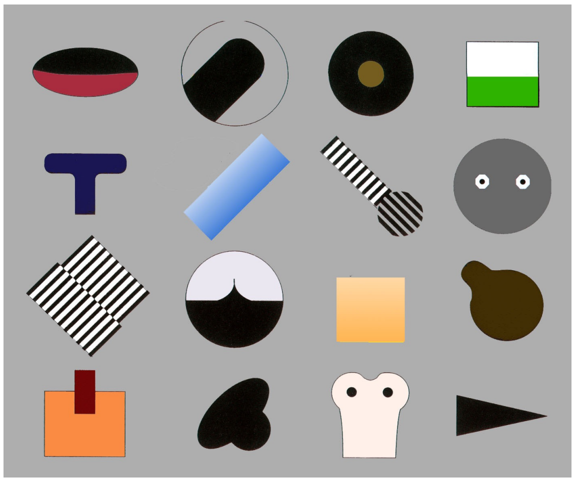
Freiwald and Tsao, Science, 2010



Brincat and Connor, Nat. Neurossci, 2007

145

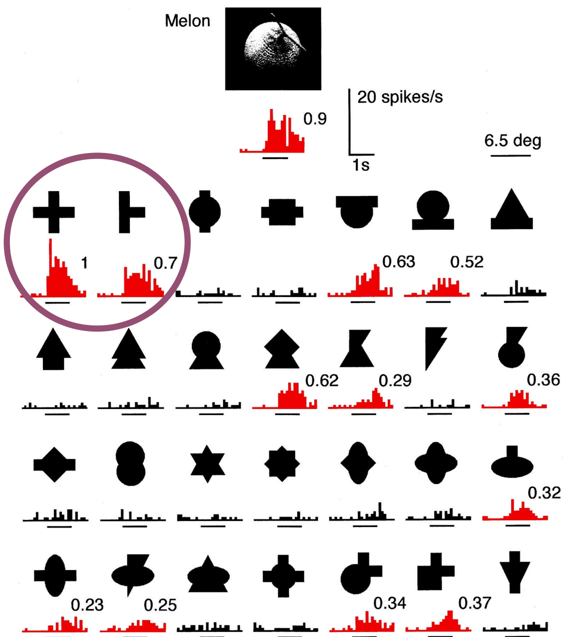
biological stimuli are important



Tanaka hypothesized that the representation of each object occurs not through the activation of a single neuron but of a population of IT neurons (distributed representation), each of which encodes only one element or a particular (which can be present in several objects) and not the whole object.

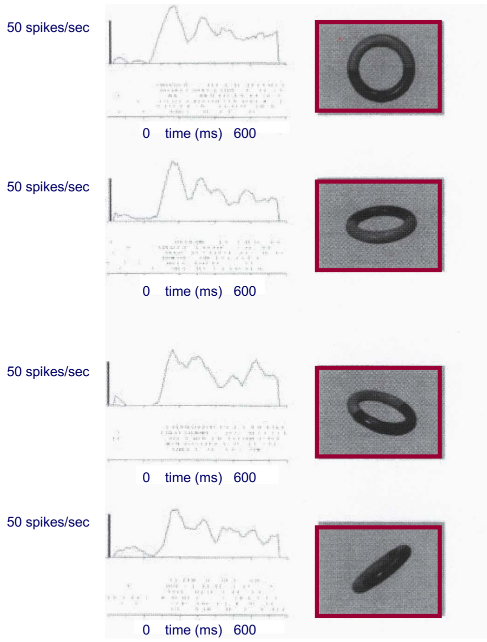
Schematic representation of the different elements to which IT neurons respond (from Tanaka, 1996).

146



For example, the neuron shown in the figure seems to respond to the complex image of a fruit (melon). However, a detailed study (method of progressive reduction) shows that the neuron also responds to simpler stimuli that represent the visual elements of an object to which the neuron is sensitive.

147



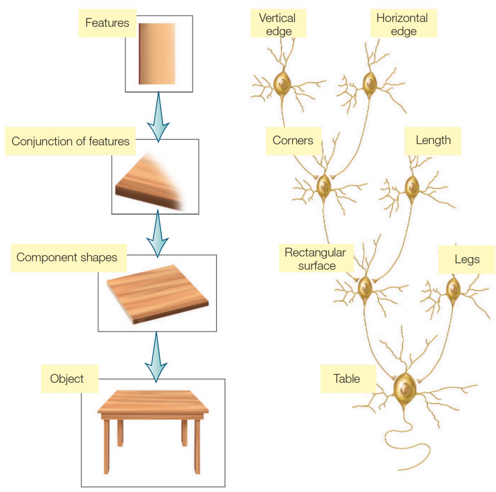
The majority of IT neurons respond to a stimulus only when it is presented from specific points of view (view-dependent responses).

Some neurons (10%) selectively respond to familiar stimuli regardless of their position with respect to the observer (view-independent responses).

These response, although rare, indicate that IT is capable of forming a (relatively abstract) representation of the object, rather than responding to one of the different forms that the object can take when its position with respect to the observer changes.

148

Hierarchical model of the object recognition



The finding that IT cells selectively respond to more complex stimuli than V1, V2 and V4 is consistent with a hierarchical model of object perception.

According to this model, each subsequent level encodes more complex combinations from the inputs of the previous level.

The type of neuron that can recognize a complex object has been called the gnostic unit, referring to the idea that the cell (or cells) signals the presence of a complex, highly specific, and significant stimulus: that is, a known object, place or animal that has been encountered in the past.

149

Local or distributed coding?

It is tempting to conclude that the cell represented by the activity of IT cells signal the presence of an object (a hand or face), independent of the point of view.

In this regard, the researchers coined the term 'grandmother cell' to convey the idea that people's brains may have a gnostic unity that is activated only when the grandmother comes into view.

Other Gnostic units would specialize in recognizing, for example, a blue Volkswagen or the Golden Gate Bridge.

150

150

Distributed code hypothesis

An alternative to the Grandmother cell hypothesis is that object recognition is the result of a distributed activation pattern on the population of IT neurons.

According to this hypothesis, recognition is due not to one unit but to the collective activation of many units.

Distributed code theories easily explain why we can recognize similarities between objects (say, a tiger and a lion) and make mistakes between visually similar objects - both objects activate many of the same neurons.

Losing some units may degrade our ability to recognize an object, but the remaining units may be enough.

Distributed code theories also explain our ability to recognize new objects. New objects have a resemblance to familiar things, and our perceptions result from activating units that represent their characteristics.

151

151

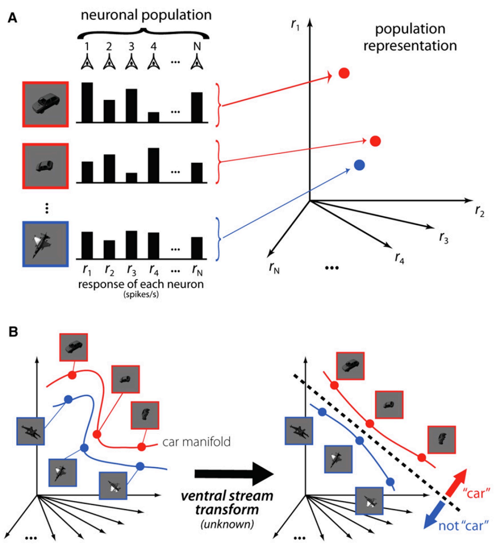
- IT neurons respond only to visual stimuli.
- The receptive fields always include the fovea, that is the part of the retina most involved in the fine recognition of a visual stimulus.
- The receptive fields tend to be large, providing the opportunity to generalize the stimulus within the receptive field, and often extend along the midline in both visual hemifields, thus joining the two halves of the space for the first time. This property depends on the interhemispheric connections through the splenium of the corpus callosum and the anterior commissure.
- IT neurons encode complex characteristics of the stimulus (not simple features, such as color, form orientation, depth).

152

- A small percentage of IT units are selective for faces. Some are sensitive to emotional expression and others to the direction of gaze. Hand-selective cells are also found.
- Faces and probably other shapes appear to be encoded by a pattern of activity distributed over a set of cells, rather than by gnostic units (grandmother's cell,) that is, a cell that responds to complex and highly specific visual stimuli such as one's grandmother.
- The selectivity of IT cells is usually invariant with respect to changes in stimulus size, contrast, color, and exact location on the retina.
- There appears to be a vertical organization for the selectivity of the stimulus of IT neurons.
- The activity of IT neurons can be modulated by the animal's attention.
- IT cells can exhibit both short- and long-term memory effects for visual stimuli, and their selectivity can be changed by experience.

153

Ventral visual pathway gradually "untangles" information about object identity



Response of a population of neurons to a particular view of one object can be represented by a response vector in a space whose dimensionality is defined by the number of neurons in the population.

When an object undergoes an identity-preserving transformation, it produces a different pattern of population activity, which corresponds to a different response vector.

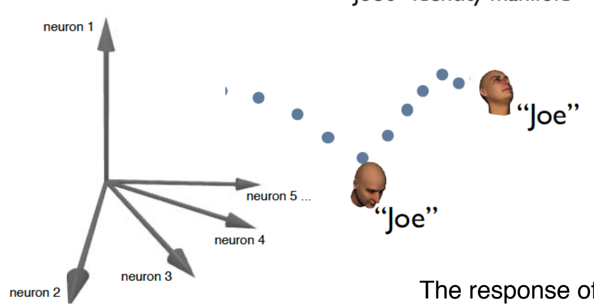
Together, the response vectors corresponding to all possible identity preserving transformations define a low-dimensional surface in this high-dimensional space—an object identity manifold.

DiCarlo et al., Neuron, 2012

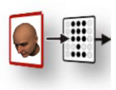
154

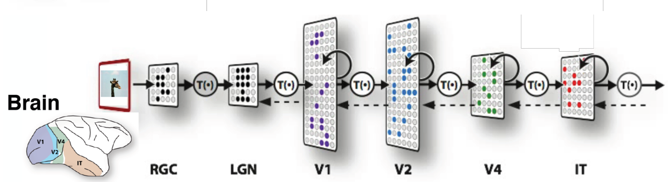
154

"Joe's" identity manifold



The response of a population of neurons to a particular view of one object can be considered as a response vector in a space whose dimensionality is defined by the number of neurons in the population.

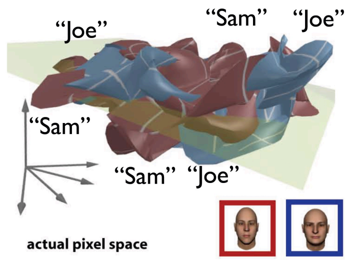


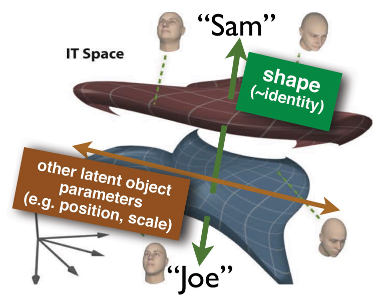


DiCarlo et al., Neuron, 2012

155

155





Object recognition is the ability to separate representation that contain one particular object from representation that do not.

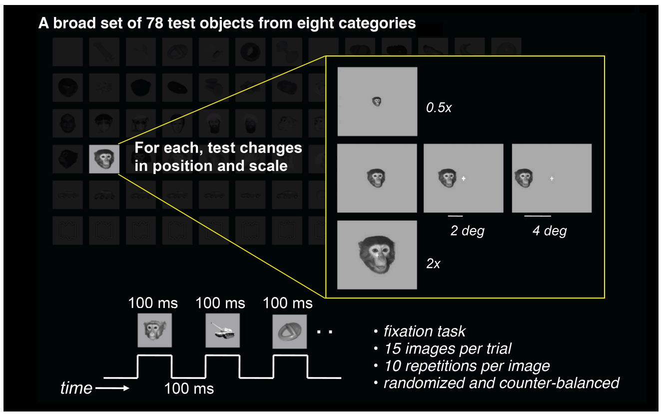
Thus, object manifolds are thought to be gradually untangled through nonlinear selectivity and invariance computations applied at each stage of the ventral pathway.

At higher stages of visual processing, neurons tend to maintain their selectivity for objects across changes in view; this translates to manifolds that are more flat and separated (more ''untangled'').

DiCarlo et al., Neuron, 2012

156

156



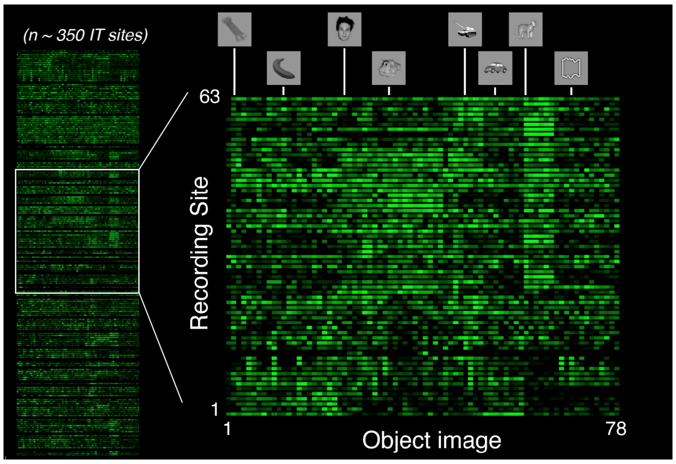
By using a classifier-based readout technique, Hung et al (2005) investigated the neural coding of selectivity and invariance at the IT population level.

They showed that the activity of small neuronal populations (~300 units) over very short time intervals (as small as 12.5 milliseconds) contain accurate and robust information about both object ''identity'' and ''category.''

Hung et al., Science, 2005

157

157



The readout technique consists of training a regularization classifier to learn the map from neuronal responses (from the independently recorded neurons) to each object label.

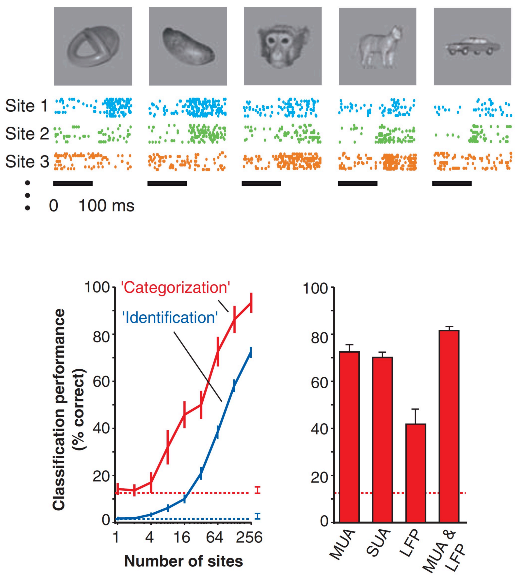
The input consists of the neuronal responses

.

The activity of ~ 250 randomly selected multi- and single-unit activity in response to 78 images of different objects were recorded at 350 IT sites in two monkeys

Hung et al., Science, 2005

158



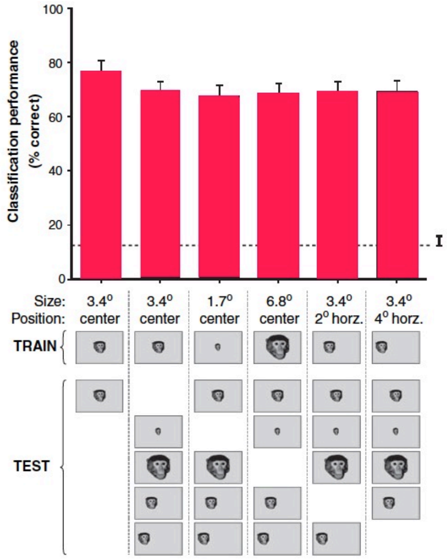
The spiking activity of 256 randomly selected multi-unit activity sites was sufficient to categorize the objects with 94% accuracy.

we have lots of neurons but they do the same things

Classifier performance increased approximately linearly with the logarithm of the number of sites, which is indicative of a distributed representation in contrast to a grandmother-like representation.

Hung et al., Science, 2005

159



After training, the classifier was used to decode the responses to novel stimuli.

The classifier performance is maintained high over a range of object positions and scales, even for novel objects.

A one-versus-all approach was used whereby for each class of stimuli (8 classes for categorization, 77 classes for identification, 3 classes for scale and position readout), one binary classifier was trained.

Recognition performance is what downstream neurons (for instance PFC neurons) could, in theory, perform by simply computing a weighted sum of IT spikes over a short time interval (100- to 300-ms interval divided into bins of 50 ms in this case).

This is notable considering the high trial-to-trial variability of cortical neurons.

Hung et al., Science, 2005

160

160

161

161