Neuro-physiological approach to visual perception

Three main categories of research are:

- **Direct substitution of known neural / physiological mechanisms**

- **Neural feature detectors**

- **Structures in the visual system can perform elaborate synthesis, as well as analysis, of incoming sensory data (did we hear of a visualization and perception lab. Somewhere ?).**

Neural region of the eye, the retina, is actually an outgrowth of the brain (not so with ear-drum, skin tissue, tongue, nasal cavity ?). It's a region of formidable complexity.

Human Visual Cognition is characterized by the flexibility and ease with which spatial relations and shape properties are computed. How is it possible for a system to be fast as well as being capable of computing a possibly infinite number of different spatial relations? What is the underlying computational structure?

Perceiving and acting upon the visual environment is something that humans are expert at. How does this amazing and seemingly effortless ability to represent, recognize, and reason about visual entities come about? Visual cognition is a diverse and interdisciplinary field of study that investigates the complex interplay of mental and brain function.

Topics in visual cognition: biological motion perception, face and object recognition, attention, and similarities between perception and action at the neural and functional levels.

The human visual system computes qualitatively different properties with great ease. For example, topological properties like closure and containment are quite different from metric properties like reflectional symmetries and pseudo symmetries. Yet, both are computed effortlessly by the human visual system in most typical situations.

There is a need for a computational framework that captures the intuition that there are qualitatively different, but interconnected levels in visual cognition.

Investigation of visual intelligence - the way in which the human visual system uses the light entering the eyes to create a variety of perceptual experiences.

It is necessary to explore the mechanisms that carry this out, and the ways in which this knowledge can help with the design of new computational models and devices.

Recognition -> Re – cognition

[Recognition:](#) An awareness that something being observed has been observed before;
Recognize: To know again; to identify as known or experienced before;

Cognition: Knowing in the widest sense, including sensation, perception etc.

Feature detectors in the visual system:

Shape perception and discrimination:

Contour plays an important role in delineating shape. Contour is the 2nd derivative of the intensity gradient.

Pattern Recognition or perception, gave rise to two broad categories of computational models:

template matching and feature detection.

We know drawbacks of template matching.

Train with *A*, and then recognize A.

Feature detection on the other hand, analyses shape by component parts of features.

There could be several low-level feature detectors, some of them are triggered, some of them respond better or produces larger output based on the input pattern.

A combination of the features detected by such demons will help in perceiving the shape.

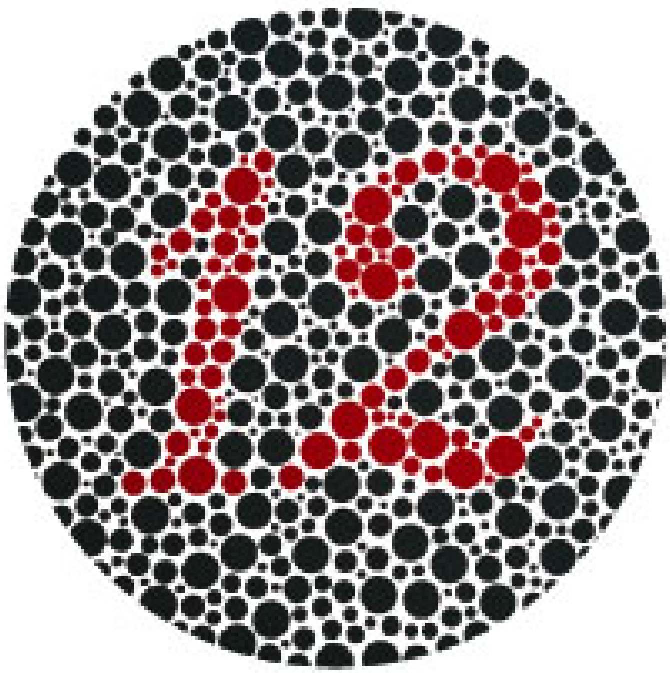
e.g. Table =
4 legs + 2 arm rests + 1 back rest + 1 flat platform

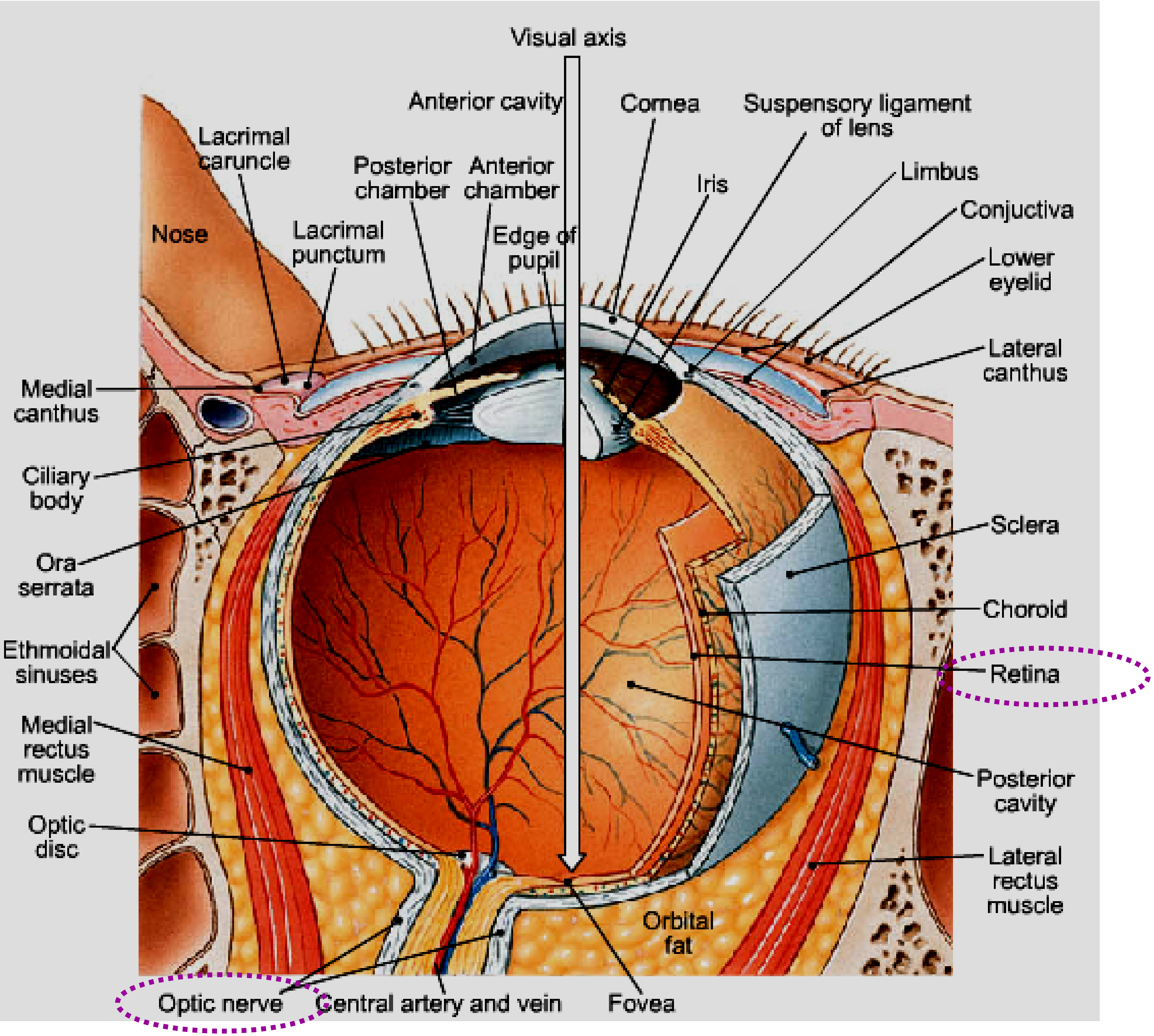
D. Marr – main job of vision was to derive a representation of shape.

Before we look into other issues in cognition, let us look into some physiological characteristics of the eye.

There are about 1.2-1.5 million retinal ganglion cells in the human retina. With about 105 million photoreceptors per retina, on average each retinal ganglion cell receives inputs from about 100 rods and cones.

Retinal ganglion cells spontaneously fire action potentials at a base rate while at rest. Excitation of retinal ganglion cells results in an increased firing rate while inhibition results in a depressed rate of firing.




Visual axis
Anterior cavity
Cornea
Suspensory ligament of lens
Lacrimal caruncle
Posterior chamber
Anterior chamber
Iris
Limbus
Conjuctiva
Nose
Lacrimal punctum
Edge of pupil
Lower eyelid
Lateral canthus
Medial canthus
Ciliary body
Ora serrata
Sclera
Chorold
Ethmoidal sinuses
Retina
Medial rectus muscle
Posterior cavity
Optic disc
Lateral rectus muscle
Orbital fat
Optic nerve
Central artery and vein
Fovea

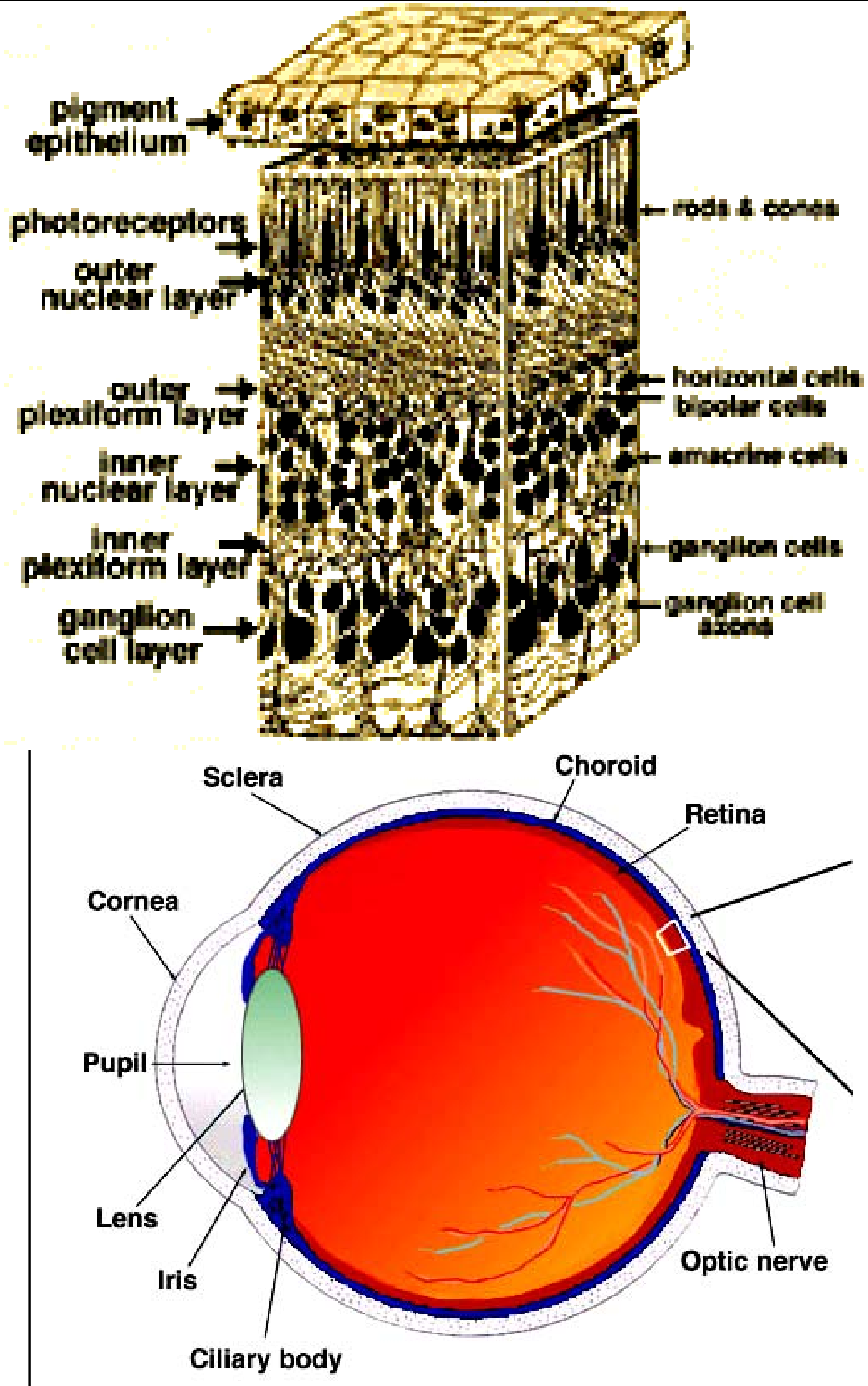


Fig. 1.1. A drawing of a section through the human eye with a schematic enlargement of the retina.

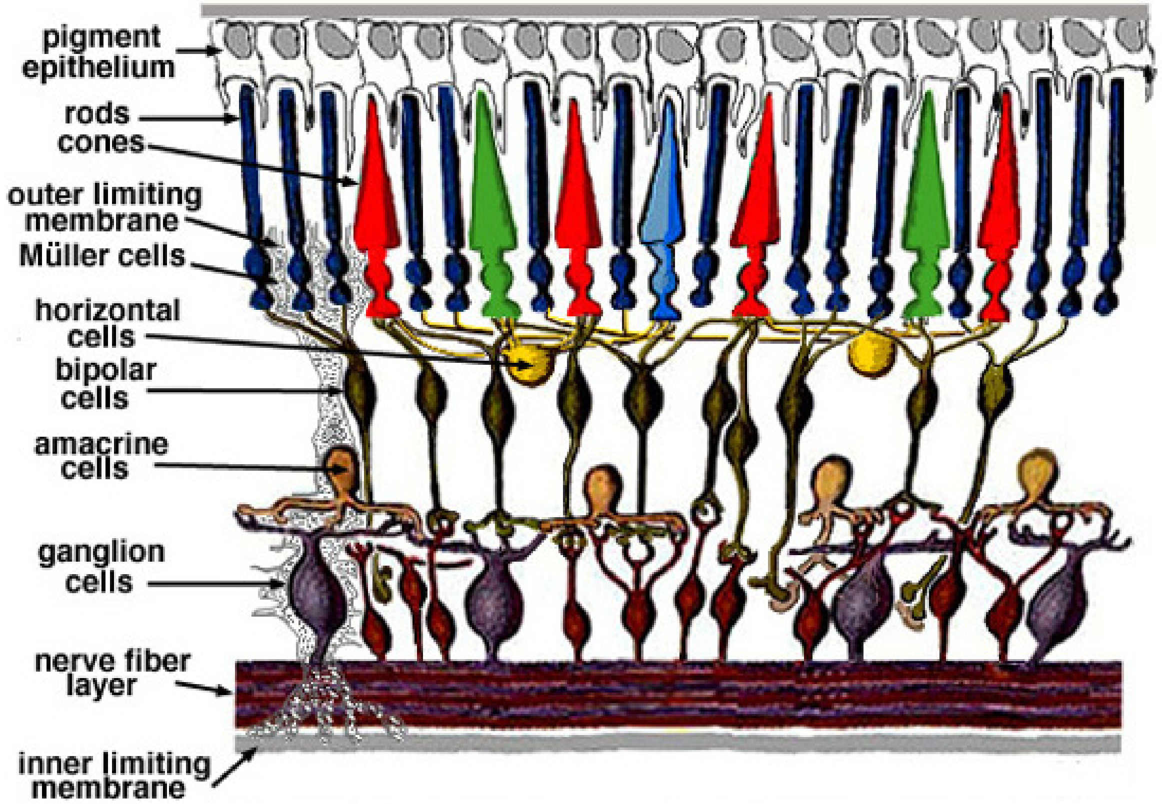
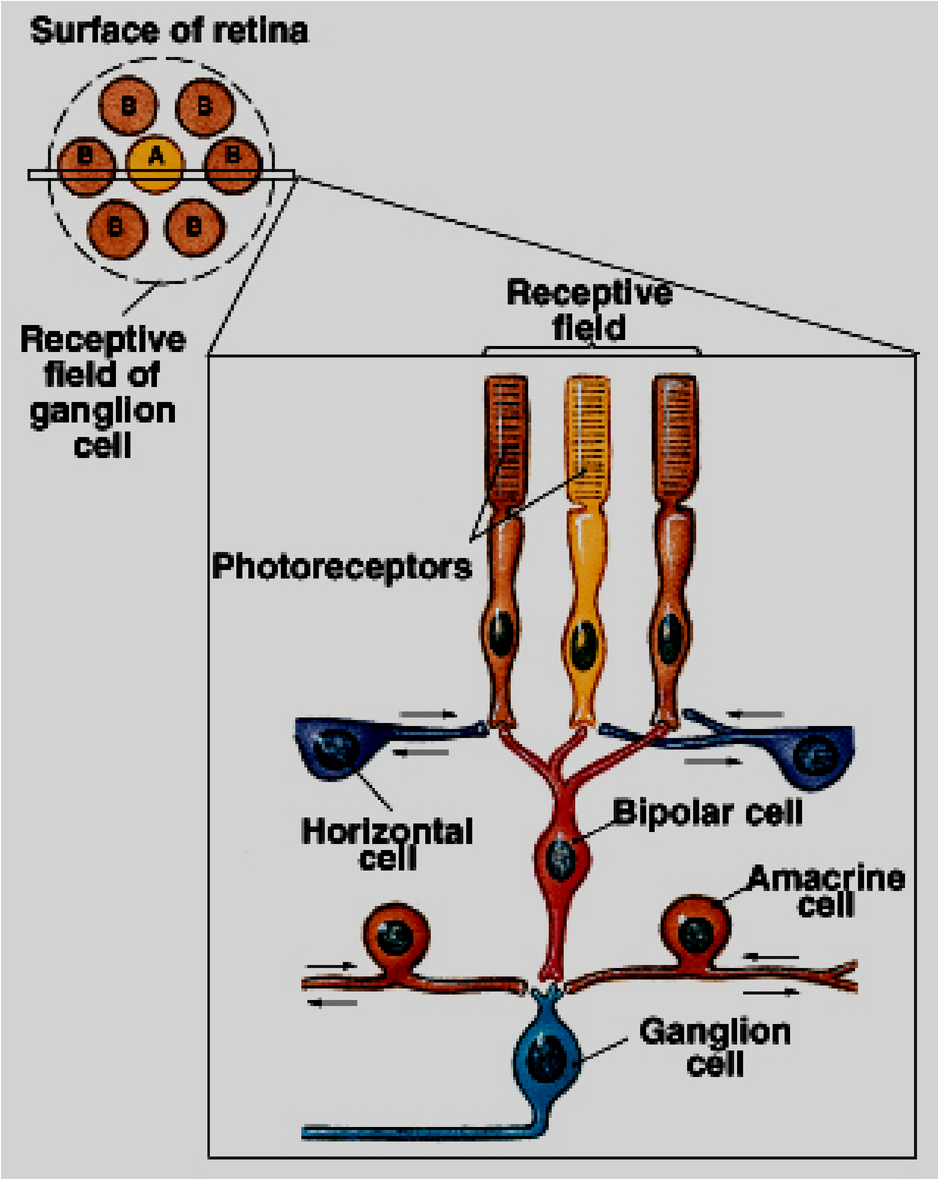
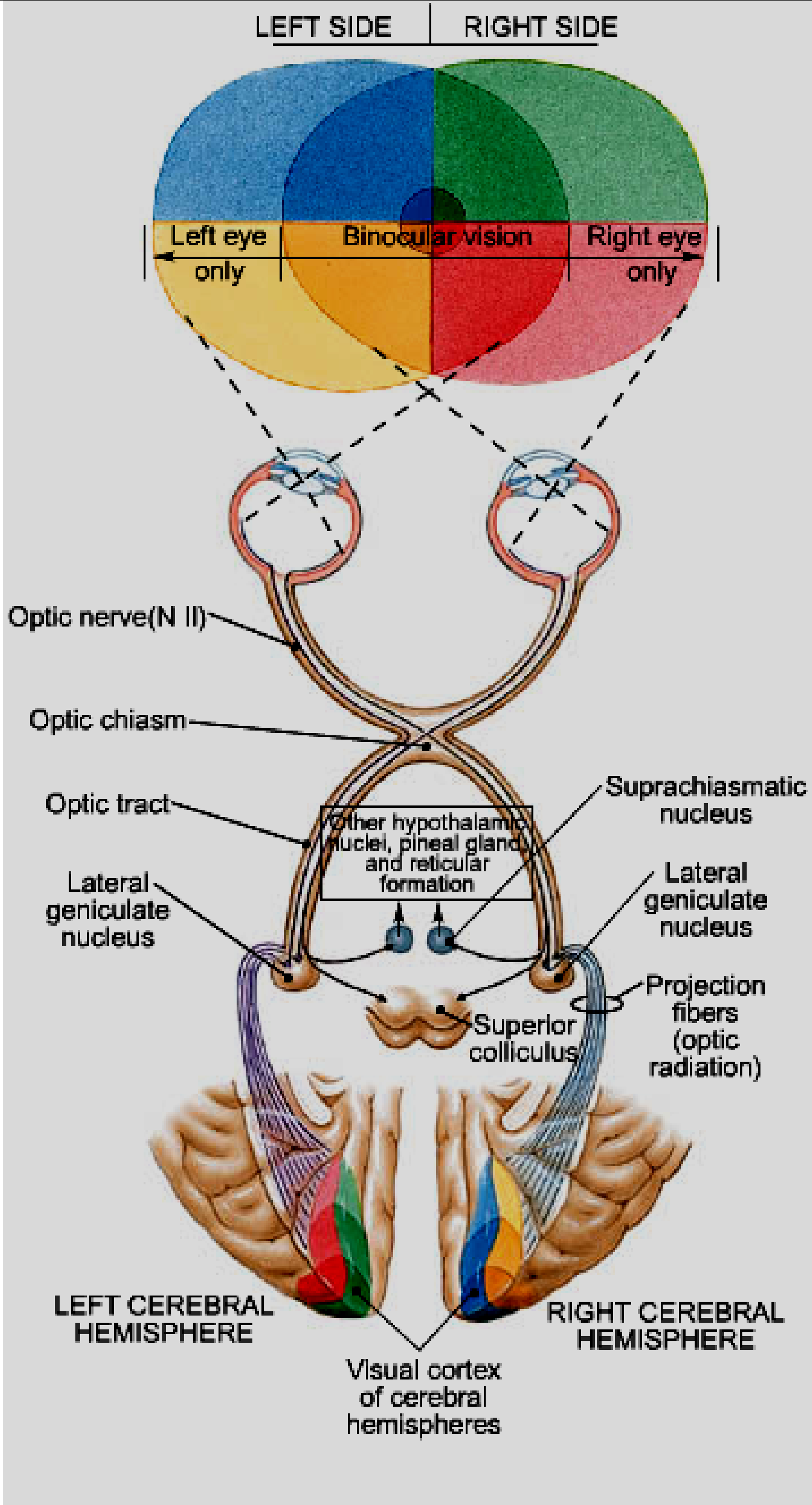


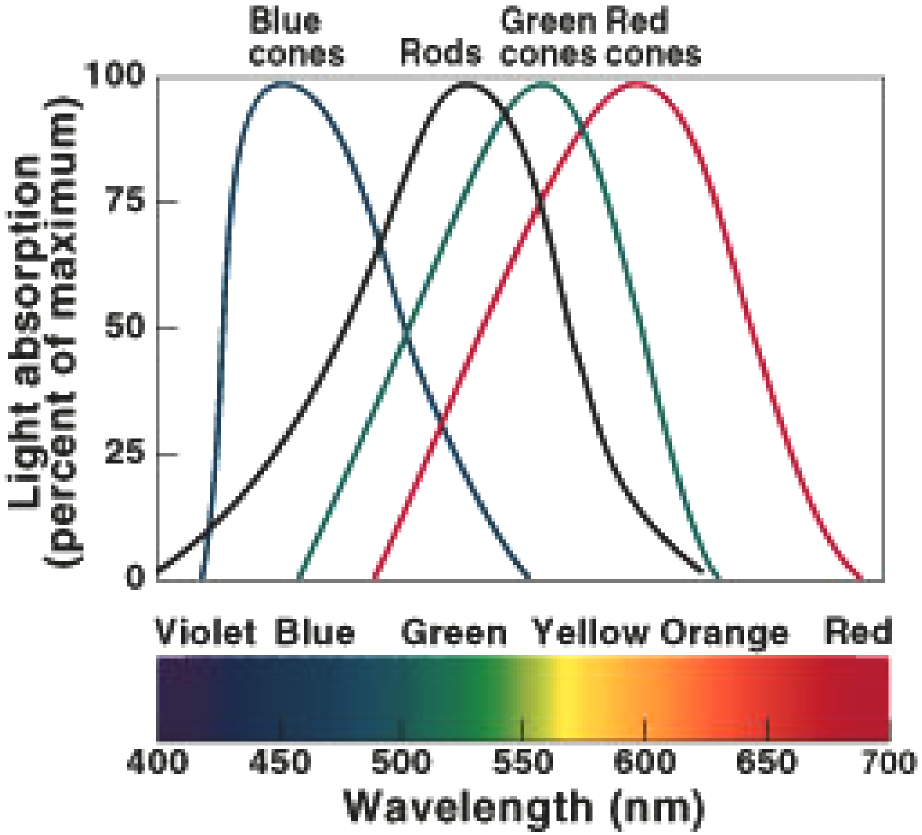
Fig. 2. Simple diagram of the organization of the retina.


Surface of retina
B
B
B
A
B
B
B
Receptive field of ganglion cell
Receptive field
Photoreceptors
Horizontal cell
Bipolar cell
Amacrine cell
Ganglion cell


LEFT SIDE
RIGHT SIDE
Left eye only
Binocular vision
Right eye only
Optic nerve(N II)
Optic chiasm
Optic tract
Other hypothalamic nuclei, pineal gland, and reticular formation
Suprachiasmatic nucleus
Lateral geniculate nucleus
Lateral geniculate nucleus
Projection fibers (optic radiation)
Superior colliculus
LEFT CEREBRAL HEMISPHERE
RIGHT CEREBRAL HEMISPHERE
Visual cortex of cerebral hemispheres

The rods and cones of the retina are called *photoreceptors* because they detect *photons,* basic units of visible light. Light is a form of *radiant energy.* This energy is radiated in waves that have a characteristic *wavelength.*

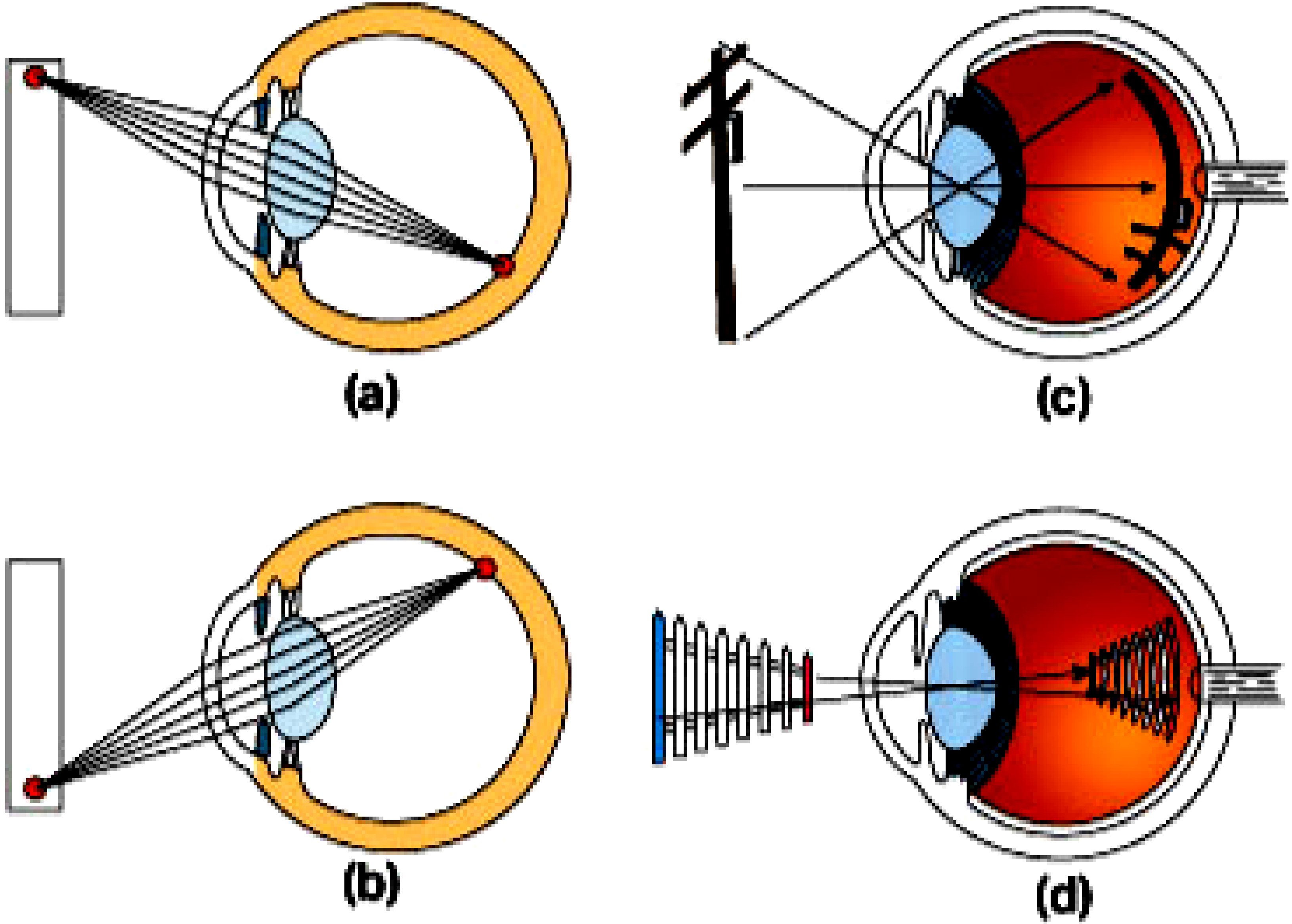
Rods provide the central nervous system with information about the presence or absence of photons, without regard to wavelength. Cones provide information about the wavelength of arriving photons, giving us a perception of color.



The three types of cones are blue cones, green cones, and red cones. Each type has a different composition and a sensitivity to a different range of wavelengths. Their stimulation in various combinations is the basis for color vision. In an individual with normal vision, the cone population consists of 16 percent blue cones, 10 percent green cones, and 74 percent red cones. Although their sensitivities overlap, each type is most sensitive to a specific portion of the visual spectrum.

Color discrimination occurs through the integration of information arriving from all three types of cones. For example, the perception of yellow results from a combination of inputs from green cones (highly stimulated), red cones (stimulated), and blue cones (relatively unaffected). If all three cone populations are stimulated, we perceive the color as white.

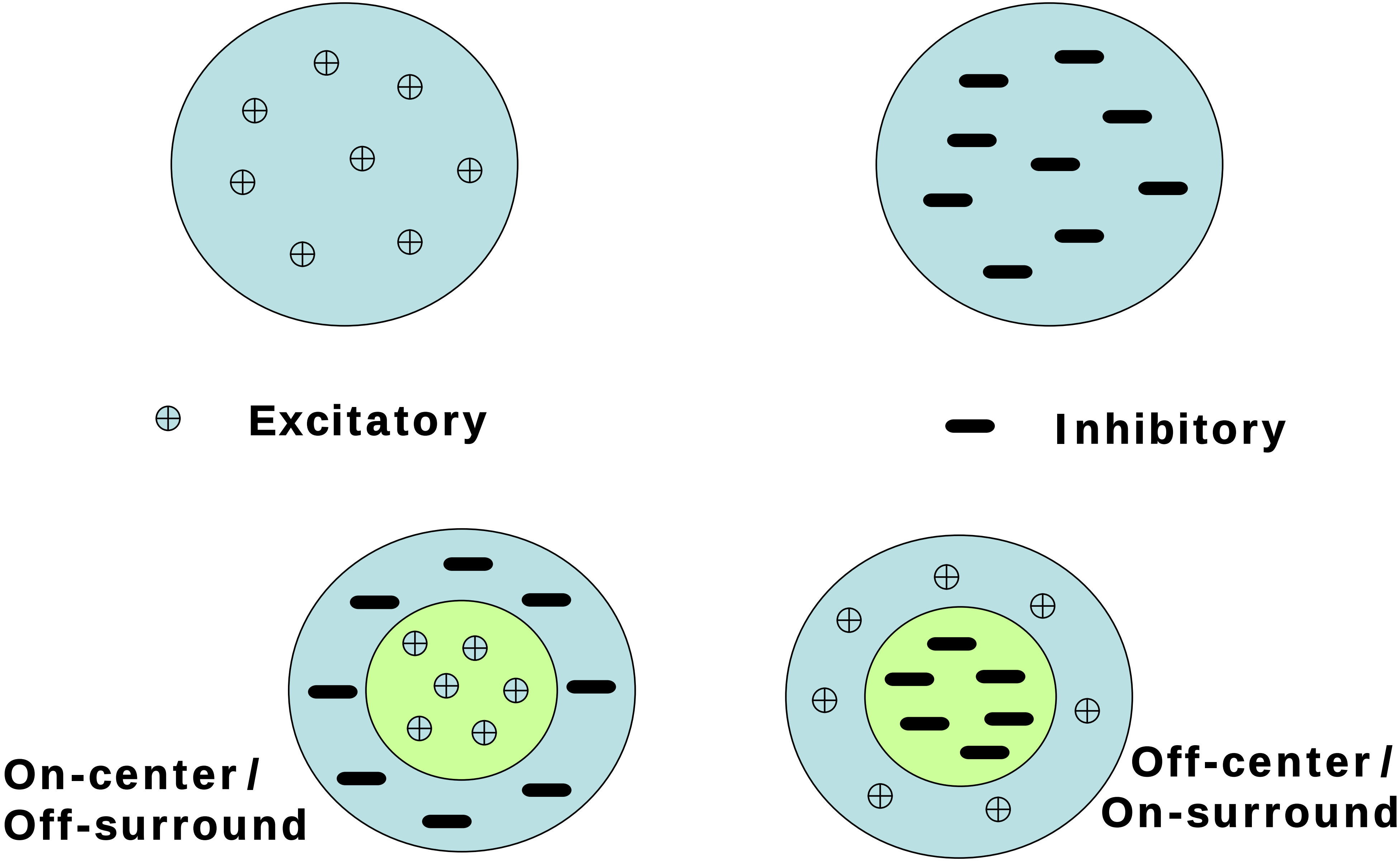
Because we also perceive white if rods, rather than cones, are stimulated, everything appears to be black and white when we enter dimly lit surroundings or walk by starlight.


(a)
(c)
(b)
(d)

The receptive field of a particular neuron in an animal's nervous system is a region of the information space around the animal in which stimulation there will alter the firing of that neuron. Receptive fields have been identified for neurons of the visual system, the auditory system, and the somatosensory system (touch, pressure, temperature, pain, itch, posture, movement and facial expression).

In the visual system, receptive fields are volumes in visual space. In the case of retinal ganglion cells for example, light from a particular visual direction might excite the cell, and light from an adjacent visual direction might inhibit the cell.

The receptive field is all the visual directions that alter the firing of the cell, that is a cone-shaped volume with its apex in the centre of the lens and its base essentially at infinity in visual space. It has been traditional to portray these receptive fields in two dimensions (i.e., as circles), but these are simply slices, the screen on which the researcher presented the light, of the volume of space to which this cell will respond.



Receptive field organization of the vertebrate visual system. Each diagram represents an area of the retina monitored by a retinal ganglion cell.
Hartline got Noble award in 1967 for this work.

Neurons in the visual cortex fire action potentials when visual stimuli appear within their receptive field. A receptive field is a small region within the entire visual field. Any given neuron only responds to a subset of stimuli within its receptive field. This property is called tuning.

In the earlier visual areas, neurons have simpler tuning. For example, a neuron in V1 may fire to any vertical stimulus in its receptive field. In the higher visual areas, neurons have complex tuning. For example, in the inferior temporal cortex (IT), a neuron may only fire when a certain face appears in its receptive field.

The primary visual cortex (V1) is the most well-studied visual area in the brain. It is the part of the cerebral cortex that is responsible for processing visual stimuli. It is the simplest, earliest cortical visual area. It is highly specialized for processing information about static and moving objects and is excellent in pattern recognition.

Current consensus seems to be that V1 consists of tiled sets of selective spatiotemporal filters. In the spatial domain, the functioning of V1 can be thought of as similar to many spatially local, complex Fourier transforms.

Individual V1 neurons tend to be selective for relatively narrow bands of spatial frequencies and orientations (similar to Gabor functions), in addition to being sensitive to certain directions and speeds (i.e. temporal frequencies) of motion of these spatial features.

Neurons in V1 are also selective for color and luminance, and are primarily driven by contrast. Thus, information relayed to subsequent visual areas is no longer coded in terms of spatial (or optical) imagery; rather, it is coded as increasingly non-local frequency/phase signals. However, it is important to remember that at this very early stage of cortical visual processing, spatial location of visual information is still preserved, and a given location in V1 does correspond rather precisely to a given location in the subjective visual field.

Visual area V2 is the second major area in the visual cortex, and first region within the visual association area. It receives strong feedforward connections from V1 and sends strong connections to V3, V4, and V5. It also sends strong feedback connections to the V1.

Functionally, V2 has many properties in common with V1. Cells are tuned to simple properties such as orientation, spatial frequency, and color.

Recent research has shown that V2 cells are tuned to the orientation of illusory contours, shows a small amount of attentional modulation (more than V1, less than V4), are tuned for moderately complex patterns, and may be driven by multiple orientations at different subregions within a single receptive field.

Understanding Vision: A Problem of Reverse Engineering.

Our attempt to understand the visual pathways is very much like approaching a machine about which we know nothing but its basic functions.

By way of analogy, we all know how to operate a car (more or less) and recognize its basic function, to get us from one place to another. Now let us assume that we knew absolutely nothing but this and, for some reason, decided we should learn every detail of the cars inner workings.

We might start by trying to figure out the parts of the car devoted to this central function and those that are not.

We remove the bumper, horn, AC, and windshields and find the car still runs splendidly. We remove small parts of the engine, piece by piece until the car no longer starts, analogous to the lesion studies in the brain.

Understanding Vision: A Problem of Reverse Engineering.

Slowly we begin to understand what parts of the car (brain) are involved in locomotion (vision).

In vision research we are very much at this stage of the game, still wanting to know what each part actually does, when it does it, and how all the individual parts act in concert.

It is the hope of many researchers, that a careful investigation into the structure and function of the visual pathways using chemical, electrophysiological, genetic, and behavioral approaches will culminate in a true understanding of how the brain provides us with this most crucial of sensory capabilities, vision.

Some examples:

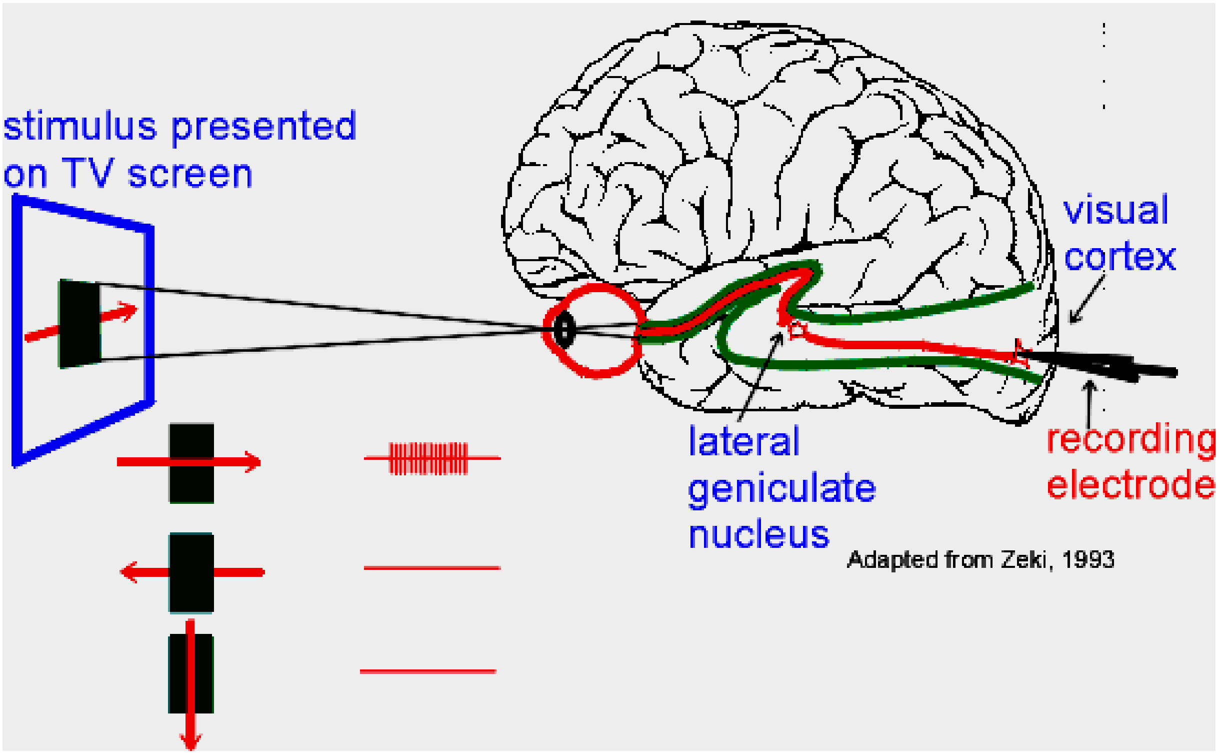
Frog's visual system appears to respond in a selective manner to stimulation at the retina. Some respond to edges, others respond to small dark objects moving across the visual field. Some respond to change in contrast, others when the visual field is darkened. The latter is necessary for their survival. Processing is peripheral rather than central – have a simple brain.

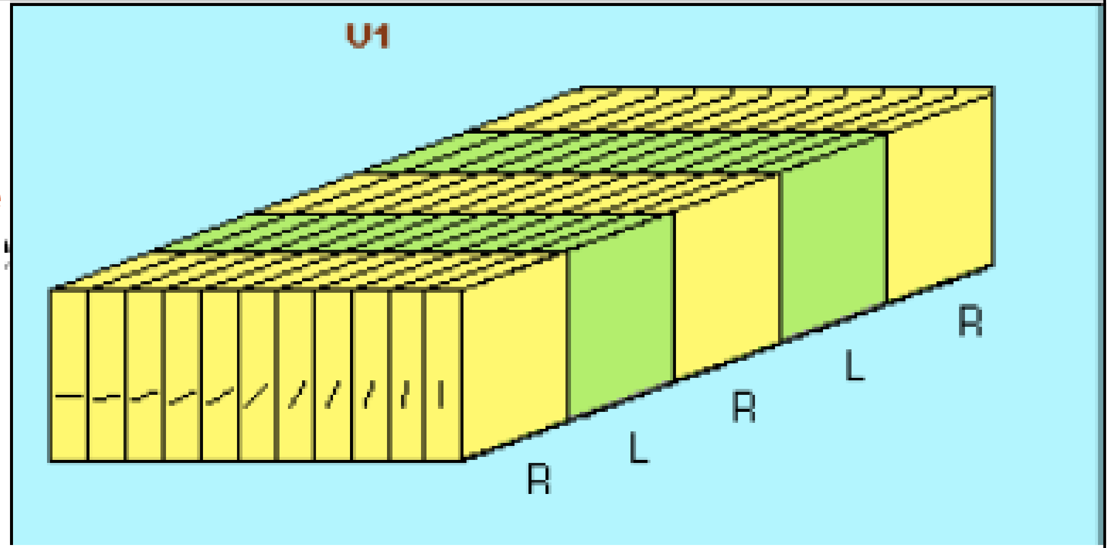
Observations of the visual cortex of vertebrates (cats and monkeys) were made using micro-electrodes. Animals kept alive under anesthesia and retina is stimulated in a controlled manner.

Cortical cells respond to the presence of moving edges or contours having a particular orientation. Some respond to vertical lines and their response falls off as the lines change away from their vertical. Others respond to horizontal or oblique lines and edges.

Nobel laureates (1981) David Hubel & Thorsten Wiesel discovered that there were cortical [receptive fields] that respond best when the stimulus was of a certain shape, had a given [orientation] and or moved in a given direction.

For example, one receptive field might respond best when a vertical rectangle moves to the right, but not when it moves in other directions.





Hubel and Wiesel were the first to discover that cells in V1 are arranged in a beautifully precise and orderly fashion.

Hubel and Wiesel found that as one advances deeper into the cortex through successive layers perpendicular to the surface, all cells that have orientation tuning prefer the same orientation.

On the other hand, moving across the surface of the cortex, orientation tuning mostly changes in an orderly fashion (as shown by the small lines in the picture). Hubel and Wiesel used the term "orientation columns" to describe this arrangement, but they are really slabs rather than columns.

Another major determinant of cell response is eye-of-origin. Most cortical cells can be driven by stimuli presented in either eye, but they generally prefer (ie. respond more to) one eye or the other - a property called 'ocular dominance'. Some cells prefer the right eye, and others prefer the left eye. Hubel and Wiesel discovered that ocular dominance is organised in a similar way to orientation preference - dominance is unchanging vertically but alternates as one moves horizontally across the cortex (the yellow and green 'ocular dominance columns' in the picture).

Cortical cells shows different types of responsiveness:

Some cells in the visual cortex exhibit a vertical, columnar organization. Deeper cells in the cortex produce subtly changing responses to stimuli. However all cells exhibit similar preference for the same orientation of the stimuli.

In some cases the receptive field has excitatory and inhibitory regions, whereas some have no inhibitory regions. Some are invariant to orientation and respond selectively to patterns of a particular height and width. Some are binocularly driven – either eye works.

Cells in a monkey's inferotemporal cortex (far from the visual cortex, responsible for recognition) respond selectively to the image of a hand.

Groups of neurons thus act in loose confederations – each group involved in specialized functions.

Receptive fields are present at birth, but early experience can modify the nature of fields.

Experiment with Kittens reared in artificial environments:

Kittens reared in vertical striped environment (surfaces) showed impaired acuity to horizontal stripes, and vice-versa.

Animals were not blind to unfamiliar stripes but their performance of perception was not as clear as those with familiar ones.

Cortical cell responses were also proportionate to that of the familiarity. The number of cells responding to familiar stripes were abnormally large.

Early experience hence has a major role to play. A critical period (3 weeks – 3 months after birth) experience is important for normal development or abnormal experiences can bias them.

Elements of perceiving are present at birth but not in a rigid and formidable form.

Interpretations of the discoveries of Hubel and Wiesel, and followers:

- Neurons in the brain are capable of responding selectively to certain aspects of stimuli.

- Features may be: lines in a particular orientation, combination of length-width of patterns, complex structures – hand shape etc.

- Basic building block of visual textures are dots, elongated blobs and termination of lines.

- Feature detectors play an important role in analysis, may not be in synthesis (visualization).

SPATIAL FREQUENCIES

Grating is a display comprising of alternate light and dark stripes. Four independent properties are:

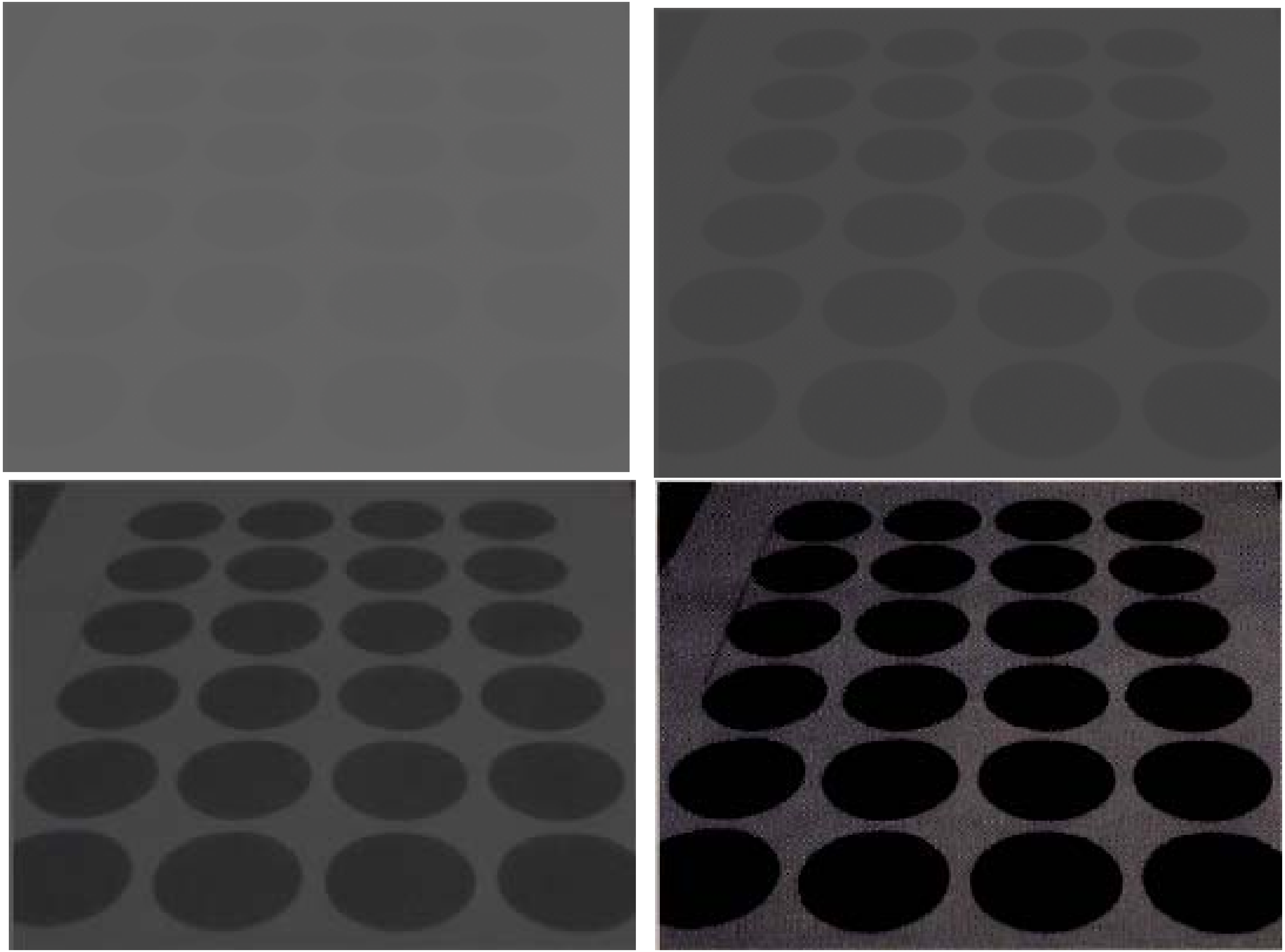
- **CONTRAST (light : dark intensity)**
- **Spatial frequency (change per degree of visual angle)**
- **orientation**
- **Phase**

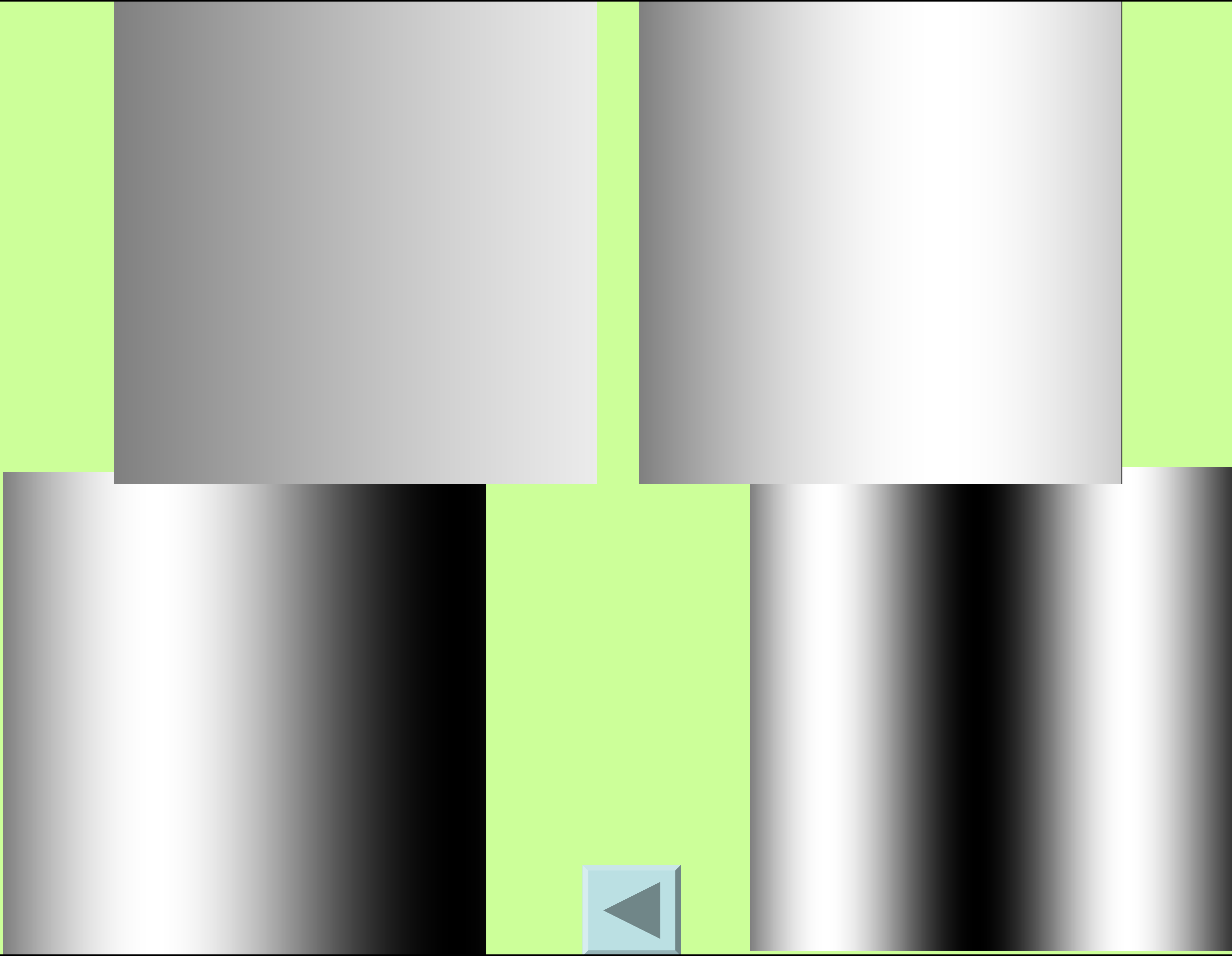
Too low a contrast or too high a frequency (fine grating) cannot be resolved. Hence there exists two thresholds associated with detection of gratings.

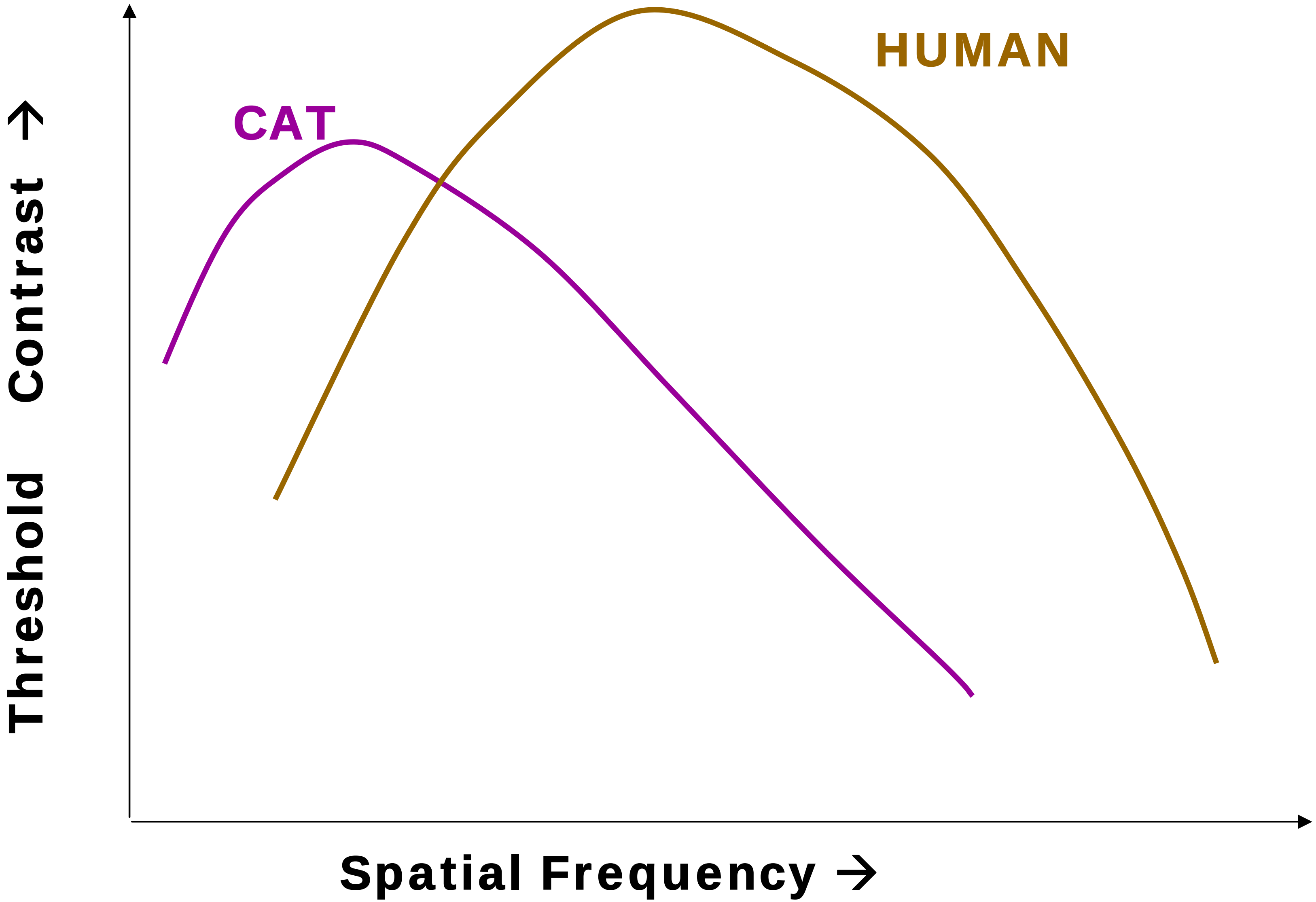
Campbell and Robson (1968) *measured* this. Select a particular frequency and set a very low contrast. Increase contrast till the stripes are barely visible.

Repeat the same with increasing frequency.

Get a threshold which depends on frequency.







HUMAN
CAT
Threshold Contrast →
Spatial Frequency →

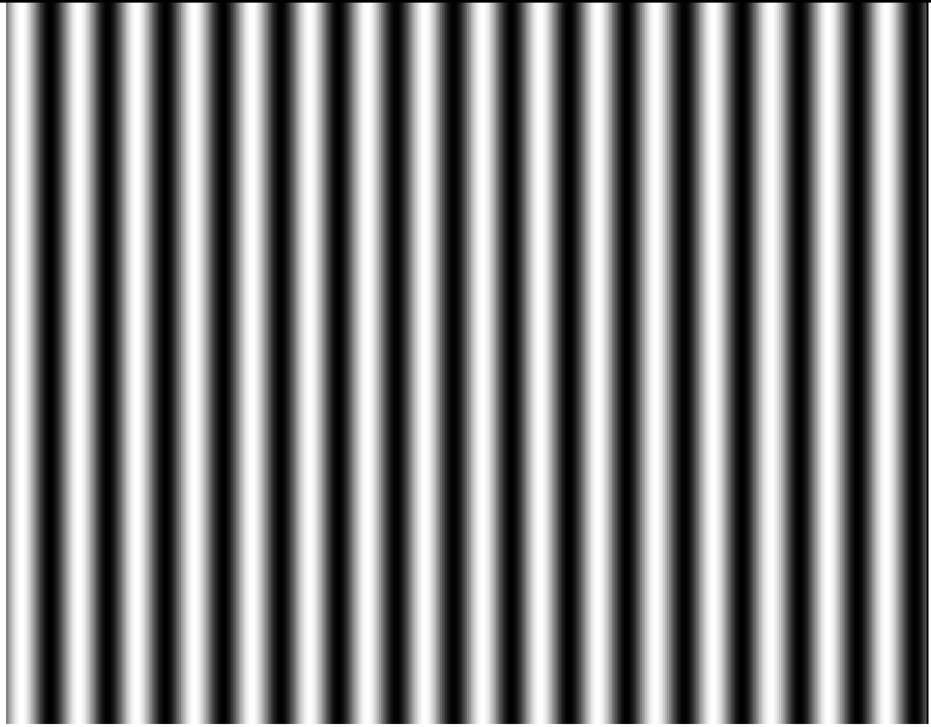
Cat's EYE: Sensitive to low spatial frequencies (say, shadows). Lots of stories of cats with super-natural phenomena.

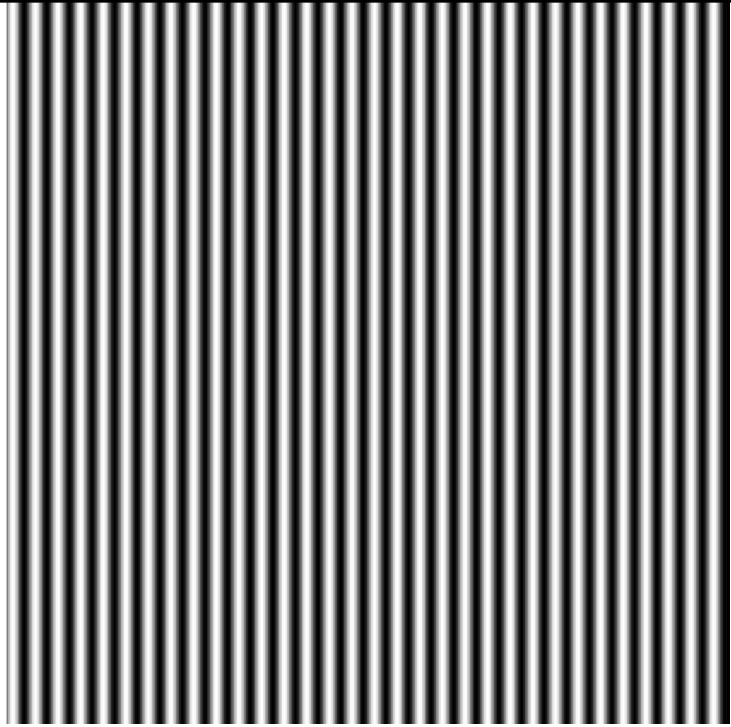
Hawks have better resolving power and nocturnal creatures have very high sensitivity. However, humans have the best all-round performance – largest area ("window of visibility") under the curve.

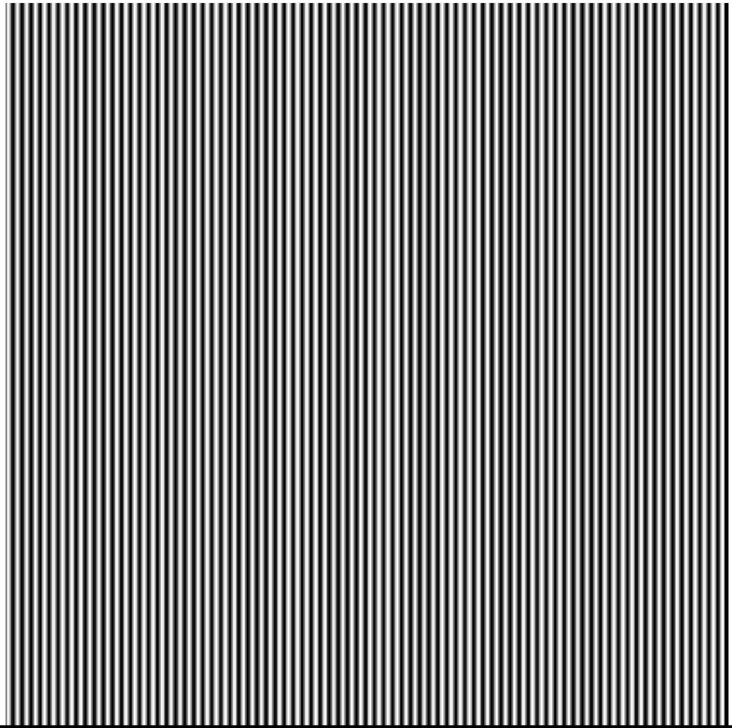
Studies on complex gratings revealed interesting discovery. Superimposition of thresholds was visible even with complex gratings (superimposition of sinusoids).

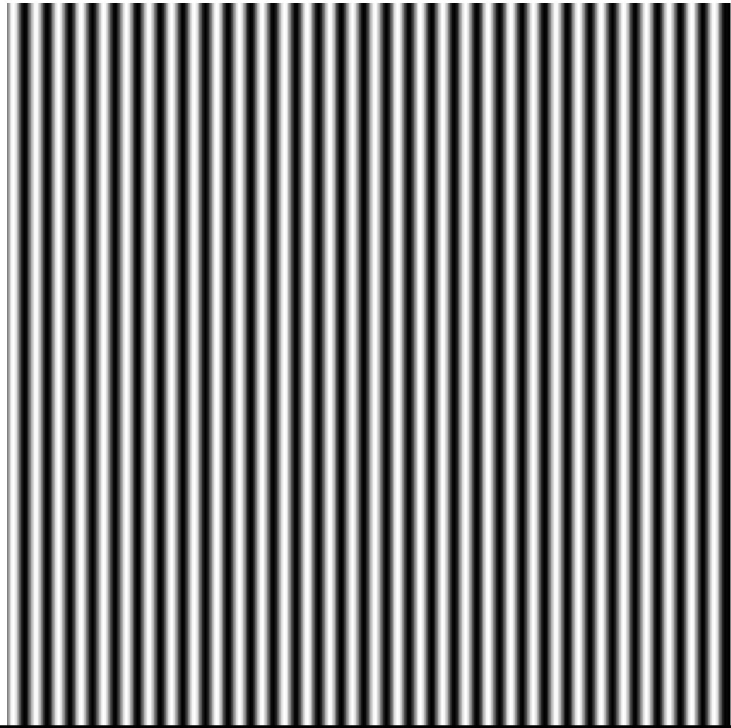
They also observed the adaptation phenomena in humans. This is also known as fatigue effect, where the drop in sensitivity is observed when exposed to certain gratings over longer periods of time.

Supra-threshold effect was also observed ??









Adaptation and fatigue are not applicable in general – Sensitivity is reduced in nearby frequencies, and that too in a particular direction in which the samples are arranged. It is orientation and frequency specific.

This gives rise to the idea:

THE VISUAL SYSTEM CONVEYS INFORMATION ABOUT SPATIAL FREQUENCIES IN TUNED CHANNELS.

Now think: If a visual system perceives signals (images) in different bands and orientations, it may have to fuse or combine this information or features somewhere to extract the knowledge of the scene contents (cognition).

e.g. computers can re-create back the picture from different (if not all) subbands (synthesis in IDWT).

Our system does this or (in other words) use some sort of synthesis for visualization ?? No one knows.

Lets observe one experiment with frequency vs. aperture.

Along with frequency and orientation sensitivity, we also have aperture size (σ for a Gabor filter) which is an important component in a spatial frequency detection model). We need a small aperture (visualize a light meter, as an example) to respond actively to finer (high-freq.) gratings and the vice-versa. Or else, what do we expect ?

What neural mechanism (not IRIS), is responsible for this function, analogous to an aperture – any guesses ?

- *Receptive Field*

Some important properties of receptive fields:

- Various forms of organization (On-Off / surround etc.)
- Orientation specific
- Various shapes and sizes
- Overlapping receptive fields will be stimulated by features within the visual stimuli, caused by moving lines, edges and also the eye itself
- Spatial-frequency analysis is hence performed, which gives the scope to recombine information to form percepts.

Rumelhart's PDP (connectionist network) properties: (for more refer to ANN course)

- A set of processing units – discrete, input, output, hidden
- State of Activation
- Pattern of connectivity
- Propagation Rule
- Current state of the network
- Modification by experience (new, old connections, changes weights).
- Represent the environment (labeled probabilities).

Emergent properties (Mcclelland, Rumelhart, Hinton, 1986):

- Spontaneous generalization
- Graceful degradation
- Default assignment (learning from similar but incomplete examples)

Recent techniques used for understanding brain physiology and also for medical diagnosis.

- PET (Position Emission Tomography
- FMRI

FMRI led to very interesting discoveries (on-going). One of them (2003):

Most active part of he brain during problem-solving is the left occipitotemporal region. This area was known to be involved in visual object recognition and visual imagery.

Kan et.al (2003) proposes the hypothesis:

Semantic knowledge may be grounded in perceptual regions of the brain.

Frisby – 1979:

It is dangerous to assume that a property of a neuron can be directly equated with its functions.

e.g. Line detectors are optimal to detect lines, but this does not mean their function is only to detect lines.

Mind vs Brain:

If the brain is organized according to some network models, which are not rule-dominated, how is the mind organized.

We know what is brain. What is mind ??

Reactions and responses of the brain functions.

Mind collectively refers to the aspects of intellect and consciousness manifested as combinations of thought, perception, memory, emotion, will, reasons and imagination.

Dennet (1991) –

Human consciousness is itself a huge complex of memes (or meme-effects in brains). This can be best understood as the operating of a 'von neumanesque' virtual machine implemented in the parallel architecture of a brain that was not designed for any such activities.

A meme consists of any idea or behavior that can pass from one person to another by learning or imitation. Examples include thoughts, ideas, theories, gestures, practices, fashions, habits, songs, and dances.

Brief Outline of concepts covered so far

- **Theories of visual perception vs. sensation**

- **Categories of theories in Visual perception**

- **Gestalt principles**

  - closure, Previous experience. figure-background distinction, grouping, wholes and parts, infant perception, symmetry, depth cues, size constancy etc.

- **Brunwick' Probabilistic functionalism**
  - **Proximal vs. distal**
  - **uncertain, probabilistic nature of cues**
  - **validity of cues**
  - **spatial proximity,**
  - **statistical relationship**
  - ***We are all smart intuitive statisticians*.**

Neurophisiology

- **Eye and brain**
- **Layers in Visual cortex**
- **Visual cognition**
- **Receptive field**
- **Studies of Cortical cell responses**

End of neuro-physiology

for this semester.