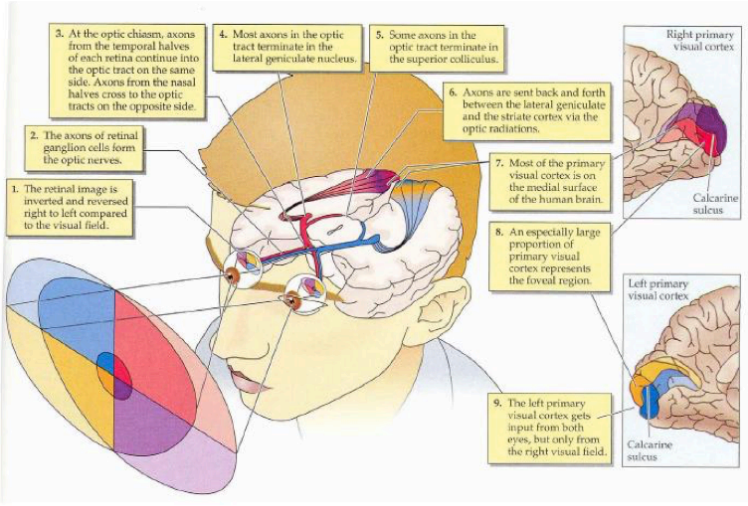
# Computational Vision – Daniel Kiper – 2022/2023

## Lecture 1 – Computation in Neural Systems: Biological and Computational Vision

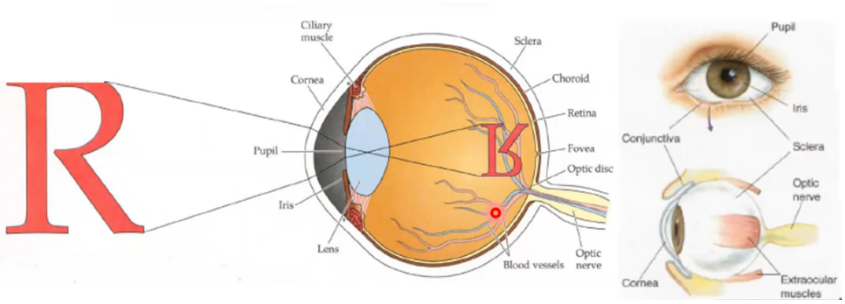
### The Visual Pathways

There are many structures in the visual system and the majority of them does not give rise to visual perception. We will discuss the conscious vision system.




The retinocortical pathways give rise to conscious vision:

1. **Visual field** (approx. 180°). The same point is projected onto the nasal and the temporal part of the retina. **The retinal image is inverted and reversed right to left compared to the visual field**. The less saturated parts (colors) of the visual field are those areas only seen by one eye.
2. The **axons of retinal ganglion cells form the optic nerves**.
3. At the **optic chiasm**, axons from the temporal halves of each retina continue into the optic tract on the same side. Axons from the nasal halves cross to the optic tracts on the opposite side. The left hemisphere is analyzed in the right brain hemisphere and viceversa. While the fovea region is represented in both hemispheres.
4. Most axons in the optic tract terminate in the **lateral geniculate nucleus**.
5. Some axons in the optic tract terminate in the **superior colliculus**.
6. Axons are sent back and forth between the lateral geniculate and the striate cortex via the **optic radiations**.
7. Most of the primary visual cortex is on the medial surface of the human brain. It is buried into the **calcarine sulcus**. And everything below the calcarine sulcus processes the lower quadrant of the visual field.
8. An especially large proportion of primary visual cortex represents the foveal region.
9. The left primary visual cortex gets input from both eyes, but only from the right visual field.

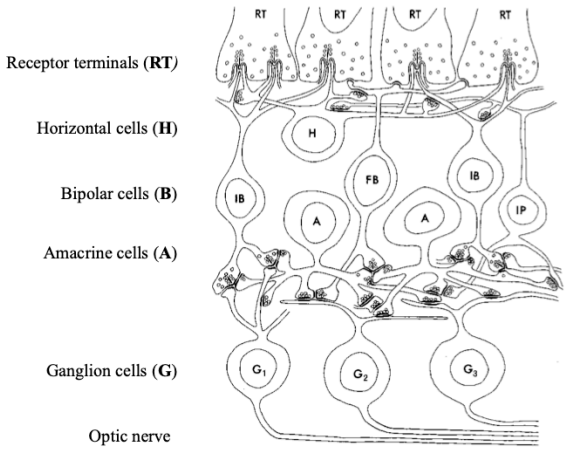



### The Human Eye

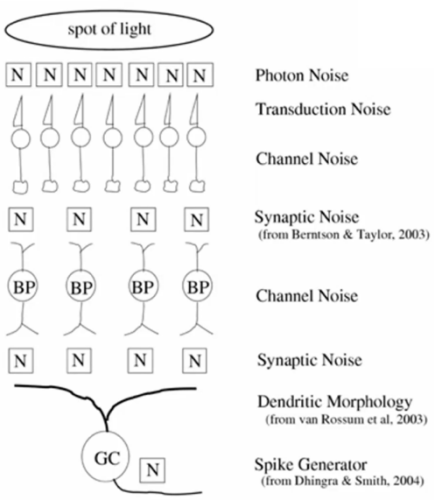
The **Optic Nerve** represents a "blind spot" as there are no photoreceptors in its location. However, the brain compensates by "filling" the missing information. The **fovea** is involved in focused vision and allows only a few degrees of visual angle.

### Basic Retinal Circuitry

The retina is the part of the eye that transforms light energy into electrical signals interpretable by the brain (AP). The light arrives from the bottom and passes through all these other cells



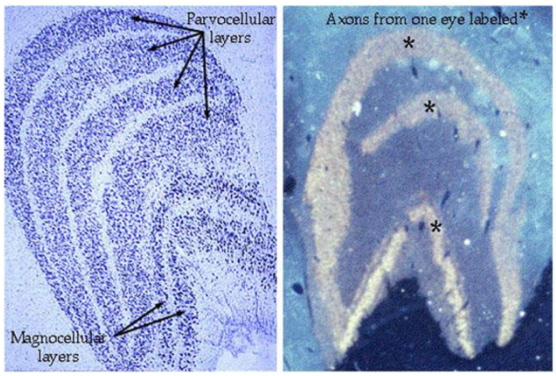
before reaching the photoreceptors (receptor terminals), so all the circuitry before the photoreceptors has to be transparent.

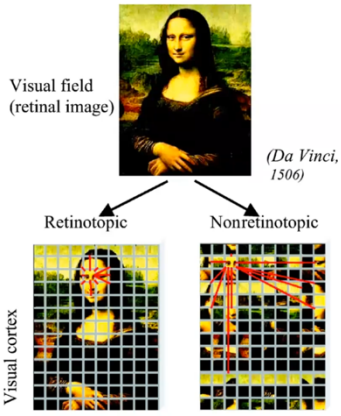



Light is a noisy source of information (photon noise), but also the retina is a noisy system. Indeed, at each level of processing some noise is added. The brain is capable (via compensation) to process reliably visual information coming from unreliable components featuring noise. The action potentials generated by the retina are then projected through ganglionic cells axons to the lateral geniculate nucleus (LGN) of the thalamus.

**The Lateral Geniculate Nucleus (LGN)**

From this cross-section frame of LGN in a primate, it is possible to notice that it has a layer structure, where each layer receives input from one eye (three layers per eye), hence each layer is "monocular". The LGN presents Magnocellular (bottom two layers, bigger cells), Parvocellular (top four layers) and Koniocellular (white layers, smallest cells) layers of cells. Magnocellular layers do not care about the color, but rather about light intensity differences. They are extremely contrast sensitive. Parvocellular layers are less contrast sensitive, but they are sensitive to color and are more capable of discriminating fine details than their Magnocellular counterparts. Finally, Koniocellular layers deal with short wavelength information. Hence, there is an anatomical segregation according to the information and the eye from which it is coming from. Notice that there exist two LGN, one for each hemisphere. (Computation done by ganglionic cells is very similar to the computation performed by LGN cells).



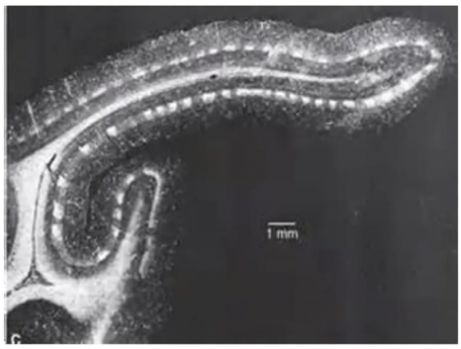



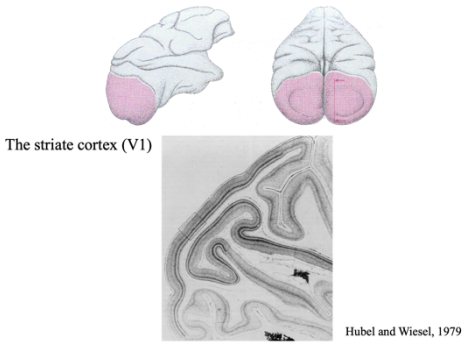

**Retinotopic Organization of the Brain**

The retinotopic organization implies that two points close in the image are represented closely also in the neural representation (Retina, LGN and V1). This peculiarity is common to most animals. Outside of V1, this retinotopical organization is not respected.

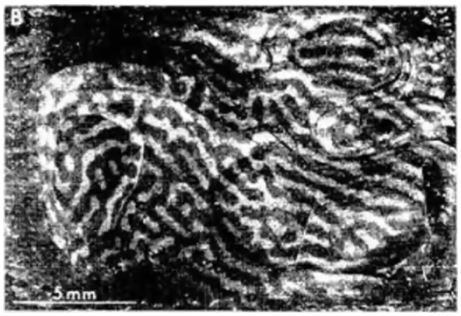
**The Striate Cortex (V1) – Hubel Wiesel**

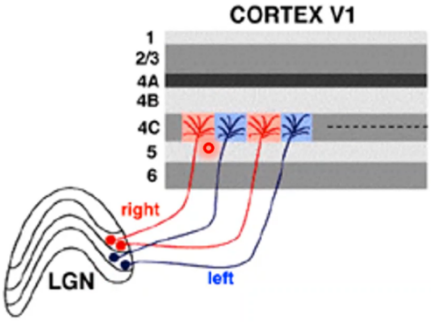
The Striate Cortex presents a thick black line, which represents the layer 4 of V1, i.e., the layer where all the connections from LGN are received. In the dark picture: Cat V1 that has been labelled with radioactive label injected in one eye of the cat that has been transported all the way to V1. We can observe that there is an alternation between cells in L4 that receive from one eye and from the other, via the LGN. However, the vast majority of cells in V1 are binocular, even if in L4 the separation of ocular dominated cells is more evident.



You seem the same thing on the surface of L4 (**monocular columns**):



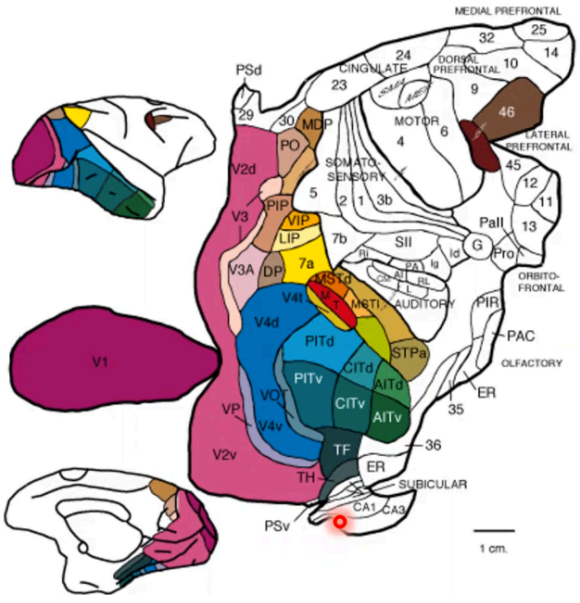



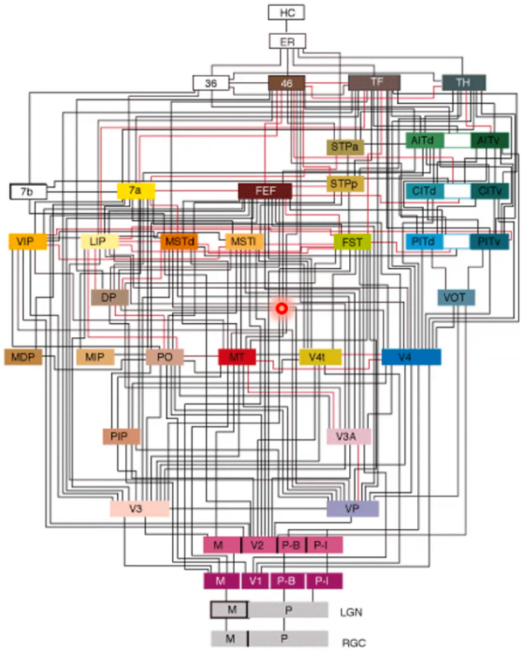


However, as soon as you get out of L4 (input layer), the cells are not monocular anymore, but rather binocular. Although receiving inputs from both eyes, most of these cells present ocular dominance, i.e., one eye induces an higher excitation of the cell. These cells are organized in **ocular dominance columns**.

These binocular property of V1 cells outside L4 is important to compare signals from both eyes, which in turns allows **depth perception**.

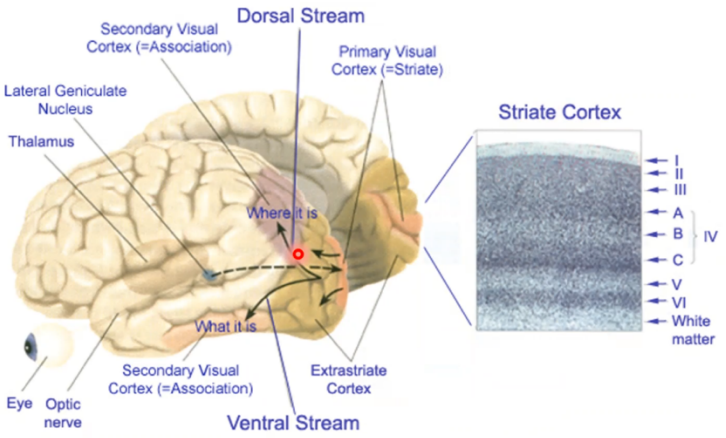
About half of the area of the cortex is dedicated to vision (because we are visual animals), however not all animals are visual animals, for example mice have more developed olfactory areas compared to visual ones. Some of these areas are only involved in visual perception and processing, while other are multi-sensory areas that combine signals from different sensory systems.




Each of these areas have a representation of the visual field, some a full one while others only a partial one. For example, V2 has a representation of the visual field which is almost retinotopic. These regions are organized in a hierarchy (right picture), which shows the connections between the areas. The connections between ganglionic cells and LGN are unidirectional, while the connections from LGN onwards are bidirectional (which have been distinguished in feedforward and feedback connections). V1 is strongly connected to V2 and V4, V2 is strongly connected to V4 and V3. The

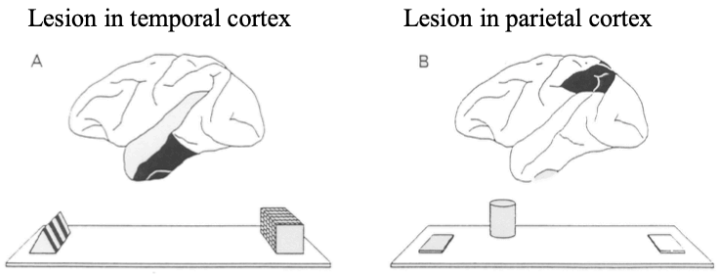
striate cortex of monkeys is very similar to the one of human. All the regions in the **left** of the hierarchy are part of the **parietal lobe**, while the **right** side is part of the **temporal lobe**. This two part represents respectively the **What** and **Where Pathways**.

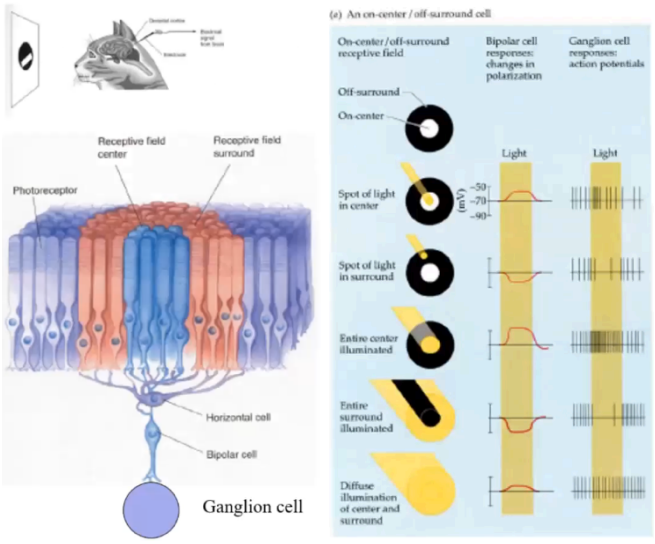



**Two Cortical Functional Streams – Ungerleider and Mishkin, 1982**

These two pathways are doing different things and it has been found through lesion studies.

Monkeys with parietal cortex lesions were not able to perform tasks associated with object localization, but they were capable of object identification. On the contrary, monkeys with temporal cortex lesions were still capable of object localization but were defecting of object recognition skills. Hence, the conclusion that temporal and parietal cortices are both involved in processing visual information, but they have distinct roles. The **temporal cortex** is primarily involved in **object discrimination**, while the **parietal cortex** is involved in **landmark discrimination**. **Visual Agnosia**: Inability to recognize objects, often associated with face recognition problems (Prosopagnosia). In the case of parietal cortex lesions, it is often associated with **neglect** symptoms, i.e., patients tend to ignore a complete visual hemisphere.



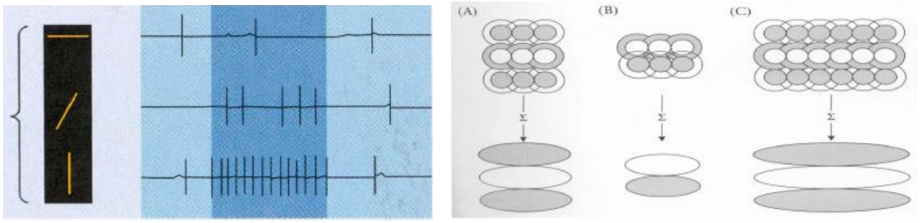




**The Receptive Field (RF)**

The Receptive Field of a cell is the region of visual space in which light can affect (increase or decrease) the cell's firing frequency. The further we go away from the retina, the wider the receptive fields of the cells increase. Ganglion cells do not receive light directly, they receive signals from the photoreceptors. The photoreceptors projecting to a single ganglion cell are a round area of the retina. Half of these receptive fields present an **On-center surrounded by an Off-region**, while the other half presents an opposite. The On-center fires to increase in light intensity, while Off-surround regions fire to decrease in light intensity.

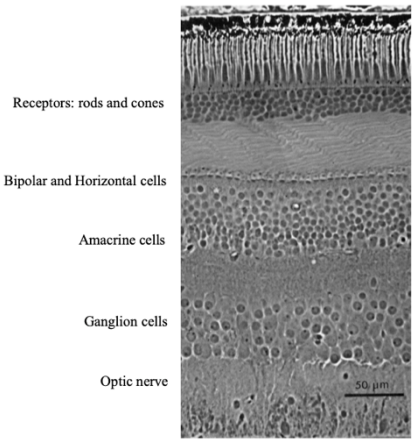
**Cortical Receptive Fields**

Orientation-sensitive cortical cell. This cell responds strongly only when the stimulus is a vertical stripe. Due to the shape of cortical receptive fields, the stimuli that strongly excite the cells are bars rather than spots of lights. They are **edge-detectors**.



# Lecture 3 – Retina & Linear Model

## A Section Through the Human Retina

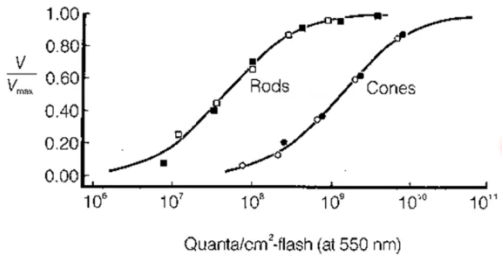



Horizontal cells modulate the output of photoreceptors and play many roles in early visual processing contributing to contrast enhancement, color opponency, and the generation of center-surround receptive fields in cones. In the human retina there are approx. 1 million ganglion cells (per eye), which collect information about the visual world from bipolar and amacrine cells (retinal interneurons). There are many types of amacrine cells that are distinguished based on the way they respond to visual inputs. Then, the layer of photoreceptors responds to changes in membrane potential. Indeed, when the membrane changes potential, photoreceptors release neurotransmitters (All known NTs are present in the retina). Since they need nutrients, they are located closely to blood vessels, i.e., at the back of the eye in humans.

## Phototransduction in Rods and Cones

The retina has a very broad dynamic range, it functions under very varied illumination conditions.

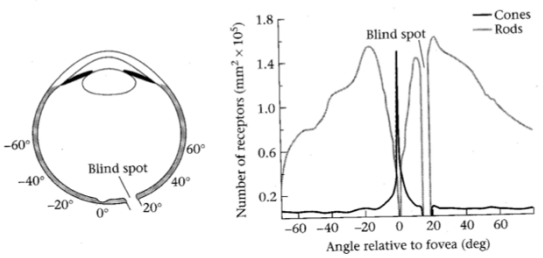
- Rods: Vision in low light (e.g., night)
- Cones: Vision in stronger light (e.g., day).



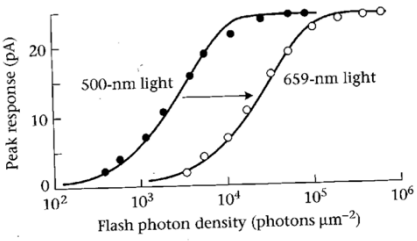

The graph to the right represents how changes the membrane potential to changes in light intensity shed on the photoreceptors.

## Distribution of Rods and Cones: A View from the Side

A different distribution of cones and rods is present on the retina. The **cones** work in day light and are capable of color vision, while the **rods** are saturated during day-time, but they provide uncolored night vision.



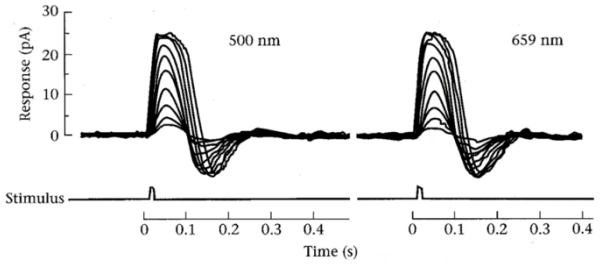

The figure to the right above shows the distribution of rods and cones in the retina. We can notice that the fovea is rod-free and has a very high density of cones. The density of cones falls rapidly to a constant level at about 10-15 degrees from the fovea. Notice the blind spot which has no receptors. At about 15-20 degrees from the fovea, the density of the rods reaches a maximum. The color vision and resolution become quite poor in the periphery.

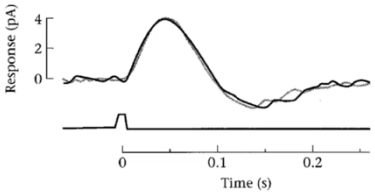



## Response of a Cone to Light of Two Different Wavelengths

In the picture to the right, light is flashed through one photoreceptor and changes in membrane potential are observed. What we can observe is that this cone responds differently to different wavelengths of photon light. Hence, the response of cones is wavelength-dependent. Indeed, there exist **3 types of cones**:

- **short**-wavelength sensitive **(blue** cones**)**
- **middle**-wavelength sensitive **(green** cones**)**
- **long**-wavelength sensitive **(red** cones**)**



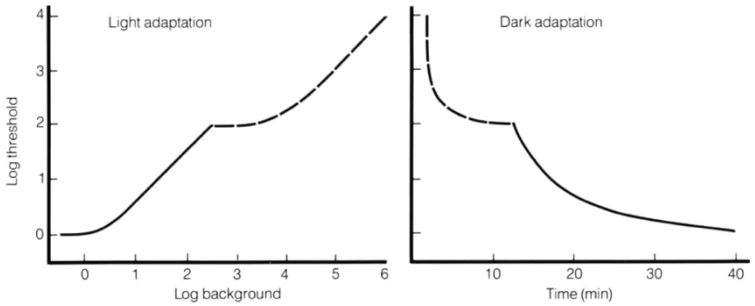



and only 1 type of rods. These different cones allow us to process different wavelengths and interpret them as colors. This concept is shown in the following figure, where the **Principle of Univariance** is explained.

**Principle of Univariance**
This principle states that one and the same visual receptor cell can be excited by different combinations of wavelength and intensity, so that the brain cannot know the color of a certain point of the retinal image. One individual photoreceptor type can therefore not differentiate between a change in wavelength and a change in intensity. Thus, the wavelength information can be extracted only by comparing the responses across different types of receptors. E.g., a rod could be equally excited by a low-intensity, well-tuned wavelength stimuli and by a high-intensity, badly-tuned wavelength stimuli. This also explains why rods (associated with vision in dark situation) tend to make us see everything in grey (but more or less bright) because cones are not active in such a situation, which instead allow us to disambiguate between the wavelength of the light stimuli.
There are cases of **monochromats** humans that have only one type of cones, which implies the loss of color perception. In such cases, humans can only perceive changes in light intensity. On the other side, there are women that are **tetrachromats**, which means they have 4 types of cones that allow the perception of "more" colors than a normal **trichromats** person.

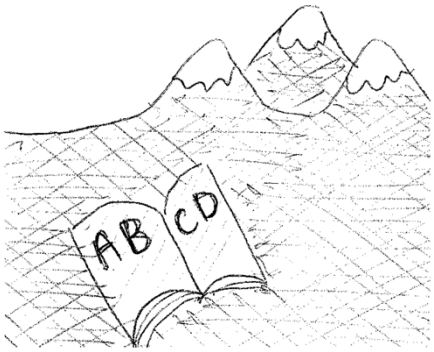
**Light Adaptation**
- One of the main functions of the retina is light adaptation.
- "Cells get used to the light stimuli".
- The photoreceptors adapt very quickly to the average level of stimulation.
- It adjusts its sensitivity to the overall level of illumination.

**Human Light and Dark Adaptation**




The dark adaptation curve on the right shows the threshold (how much light) do you need to perceive something, i.e., how strong the light stimulus has to be in order for us to perceive it as a function of time when you walk in a dark environment. With the time passing we become more and more sensitive, which lowers the (log) threshold of the amount of light needed for us to see "something". The first dotted-line is the curve describing the adaptation of cones, followed by the adaptation of rods, which induce a significant decrease in threshold levels. The result is that the retina as a whole changes its sensitivity.
On the left we have the curve describing the minimal amount of light that is needed to see something. However, in this graph the horizontal axis defines the light intensity of the background instead of time. Thus, we can see that with a low-intensity of background light the threshold for perceiving a light is much lower than when the background has an "overwhelmingly" high-intensity level. Again, the dotted line describes the adaptation of cones, while the full-line describes the adaptation of rods.

**The Jungfrau Viewed from Wengen**




We care for surface reflectance, not light intensity.
Contrast is proportional to reflectance.
The **intensity** of light in different parts of the image is dependent on: **reflectance*illuminance**. So, assuming that the illuminance is always the same in an image, the intensity is then determined by the reflectance of the elements. The relevant parameter is the **contrast**, which tells us how much each part of the image differ in terms of intensity. In particular, the contrast remains constant to changes in illumination.

| | Reflectance | Intensity *I* at noon (1000000 W) | Intensity *I* at dusk (1000 W) | Local contrast *c* at noon (1000000 W) | Local contrast *c* at dusk (1000 W) |
|---|---|---|---|---|---|
| Snow | 90% | 900000 W | 900W | 1.25 | 1.25 |
| Grass | 40% | 400000 W | 400 W | 0 | 0 |
| Paper | 80% | 800000 W | 800 W | 1 | 1 |
| Ink | 10% | 100000 W | 100 W | -0.75 | -0.75 |
| **Mean** | 40% | 400000 W | 400 W | 0 | 0 |

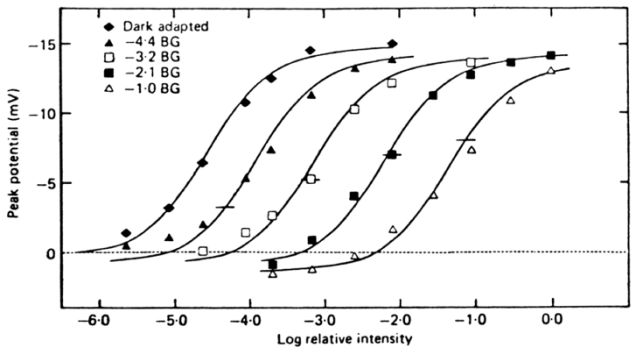
*Intensity I is reflectance*illuminance.*

*Local contrast is c = (I-Imean)/Imean.*

**Cones Responses Adapt to Background Illumination**

This graph shows how individual cones respond to flashes of different light intensity presented in different background light conditions. It shows that cones become less and less sensitive the higher the light level in the background, which shows that the cone has adapted.

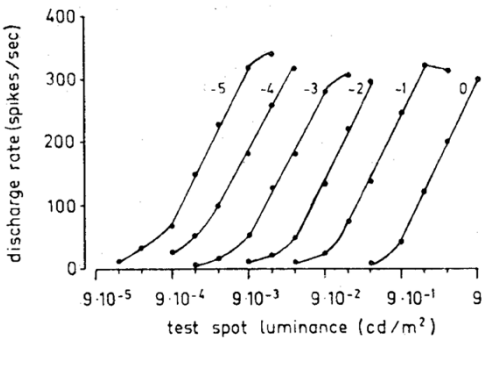
Different parts of the retina can be adapted to different light intensities. Light adaptation is somewhat local in space.




The above effect can be explained by the **local contrast**, which changes across space.

**Ganglion Cells Adapt to the Mean Light Intensity**

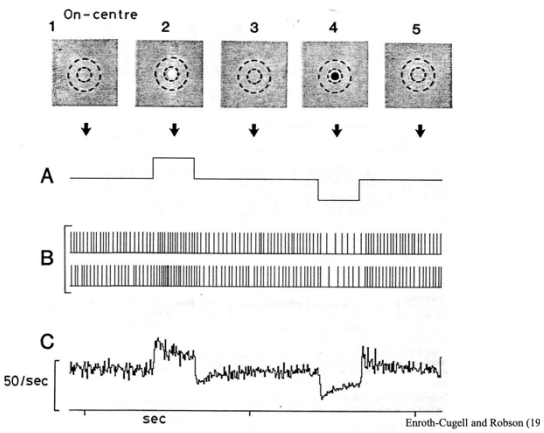
In this graph, the response of an individual ganglion cell measured in terms of how many AP/s the cell fires under different background light conditions. When the background is dark the cell is quite responsive for low intensity stimuli and then adapts and reduces the firing rate for low-intensity stimuli.

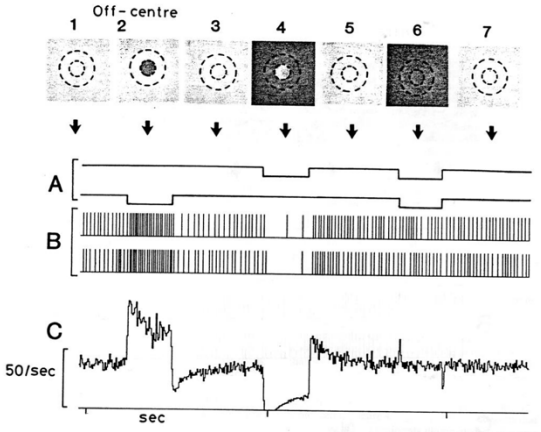


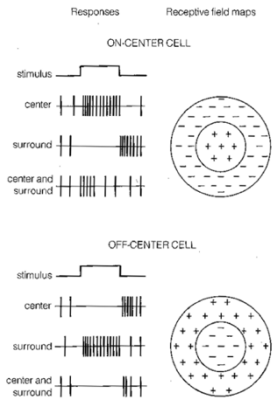
**Ganglion Cells have Center-Surround Receptive Fields**

An Off-region strongly responds to reductions in light-intensity.

The On-region responds to the opposite. In Center-Surround regions we can see that the two responses do not completely cancel each other.



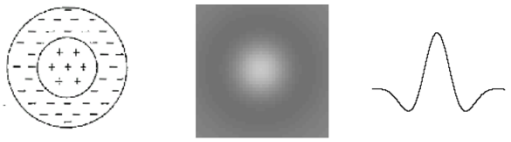



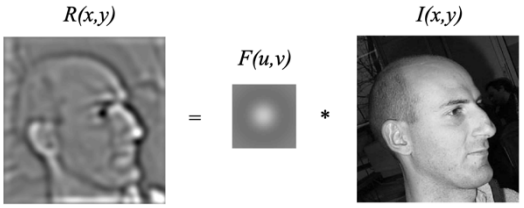

**Center-Surround Receptive Fields Enhance Edges**

Ganglion cells can be considered like **filters**, they do some spatial averaging of what is happening in the receptive field. The filter enhances the spatial locations associated with high contrast. Ganglion cells are good local contrast detectors. By changing the size of the receptive field (for example in the periphery RF are larger, while in the fovea are smaller), for example by reducing it as in the example to the right, the resolution is increased.

**The Linear Model**

**A Model of the Ganglion Cell Receptive Field**




$$R(x,y) = \iint F(u,v)\, I(x+u, y+v)\, du\, dv$$

**Assumptions Implicit in the last 3 Slides:**

- Receptive fields are difference of gaussians.
- Responses are a weighted average of the stimulus intensity, where the map of the weights is the receptive field.

Are these assumptions reasonable?

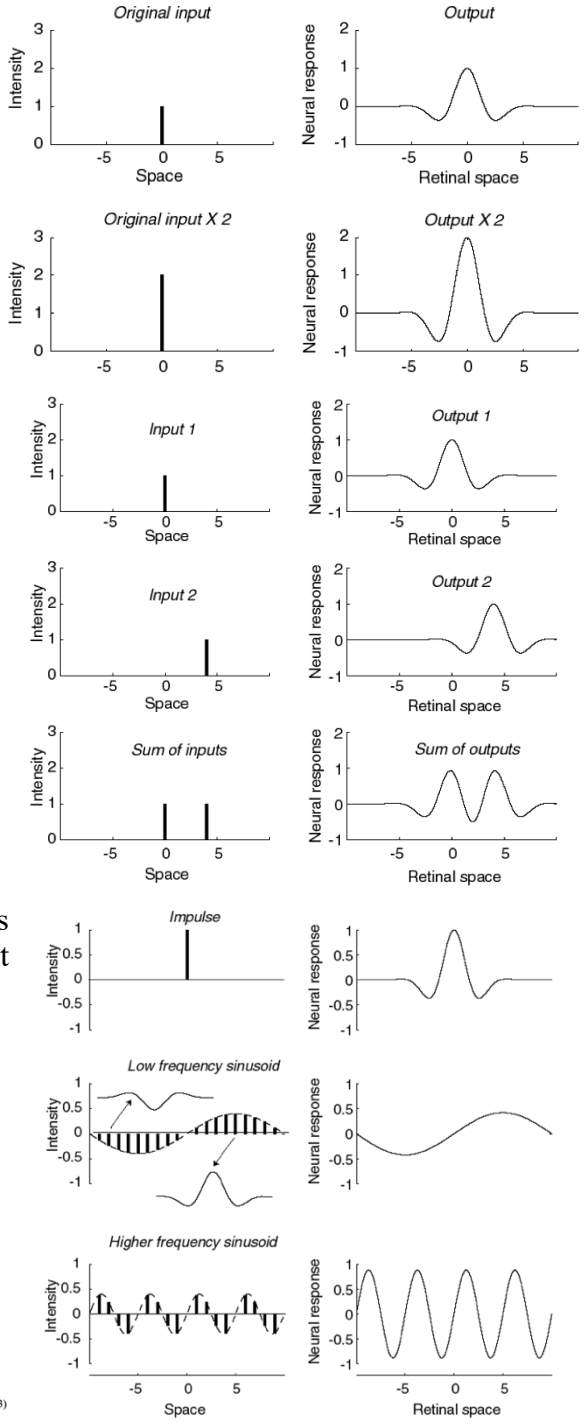
The second assumption is true if and only if the cell is a linear system.

Linear systems L(x) obey:

- Homogeneity: L(a*x) = a L(x)
- Superposition (Additivity): L(x+y) = L(x) + L(y)

**Homogeneity**

It means that if you have an input to the system x and you measure the response of the system to the input L(x), then this response satisfies: L(a*x) = a*L(x). The picture to the right presents this property by showing a plot of the retinal surface against the neural response, which evidences how the excitation elicited by an input scales linearly.




**Superposition**

If you measure the response of the system to an input x: L(x) and consecutively to an input y: L(y). Then, if the inputs are presented together, the response of the system satisfies: L(x + y) = L(x) + L(y). The same discussion performed for the figure above applies here, which indeed shows the behavior that the retinal surface should follow to satisfy the superposition property.

Linearity is often checked by using sinusoidal stimuli, because for a linear system:

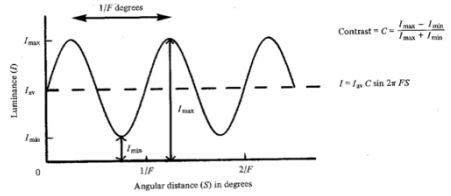
- The responses to sinusoids are sinusoids.
- The dependence of response on stimulus frequency can be predicted from the shape of the receptive field.

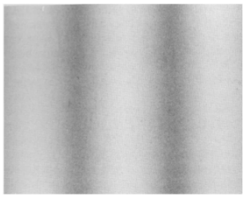
So, if any of these two are false, the system is not linear.

**Responses of a Linear System to Sinusoids**

Sinusoidal stimuli are a sequence of black/white bars that follows a sinusoidal distribution presented synchronously with different intensities.

**A Sinusoid in 2-D: A Sinusoidal Grating**

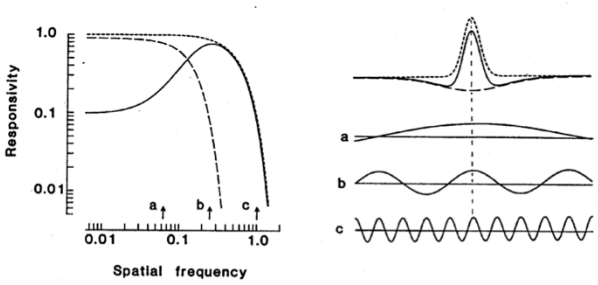


Barlow and Mollon, 1982 (Fig 1.3)

**Predictions of the Linear Model with a "Difference of Gaussians" Receptive Field**

If you have a linear system, you can predict the particular response of each ganglion cell receptive field (center/surround/different of center and surround) to a specific spatial frequency, with the knowledge of its difference of gaussians

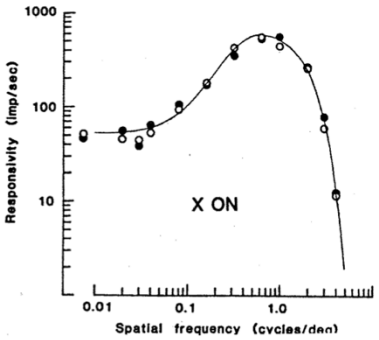



Enroth-Cugell and Robson (1984)

**Spatial frequency**: within one degree of visual angle, how many sinusoidal cycles do we have?

**Fitting the Model to the Data**

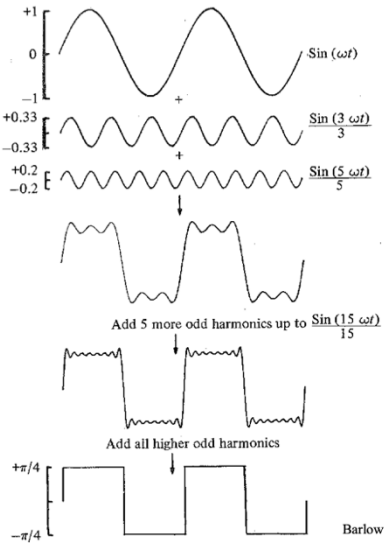
In the figure we have measures of ganglion cells receptive fields responses to changes in spatial frequency, which were made on an on-center cell. They presented sinusoidal stimuli of different spatial frequency, from which we see that by increasing the spatial frequency we obtain a frequency which is optimal for the on-center and then drops again.



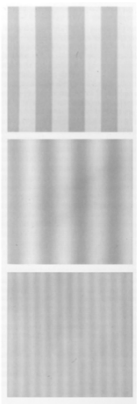

The fits are good: the responses to sinusoids are predictable by a linear model with a "difference of gaussians" receptive field. Let's try another test of linearity. If it succeeds as well, we'll be happy with the model.

**Making a Square Wave with Sinusoids**

Fourier theorem tells us that the function at the bottom of the image (luminance as a function of space) can be approximated by summation of sinusoids.




Barlow and Mollon, 1982 (Fig 1.2)

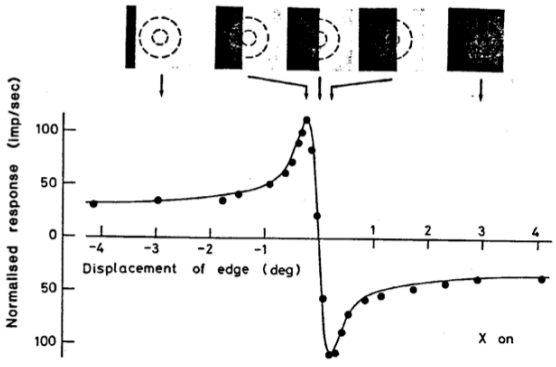


**Square Waves in 2-D**

By adding sine waves components we can obtain a square waves.

**Responses of a Ganglion Cell to Edges**

By making the assumption that the system is linear and that the receptive field is approximated by a "Mexican hat" we can see estimate the response of ganglion cells to edges. We can image an edge that is moving over the receptive field, we notice that the maximum response is obtained when the edge is bordering the on-center such that it covers the off-center and leaves the on-center fully illuminated. As soon as the edge covers the on-center, the response drops.

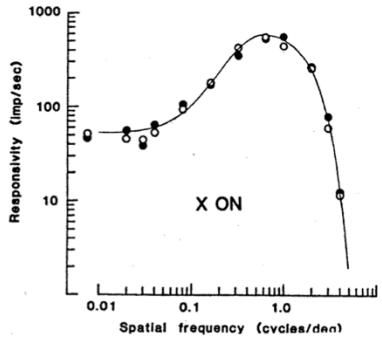



Enroth-Cugell and Robson (1984)

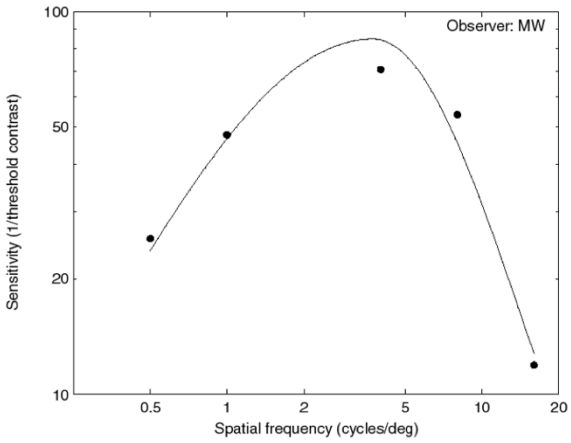
**Chevreuil Illusion – Mach Bands**

If you look at the edges of the image to the right, you should notice that the contour to the edge appears brighter on the left than on the right. These slight increases/decreases are due to the shape (center/surround) of the ganglion cells receptive fields.

**Sensitivity for Different Spatial Frequencies**
**Spatial Frequency Tuning of a Ganglion Cell**
This represents again the responses to different spatial frequencies of the single ganglion cells receptive fields.

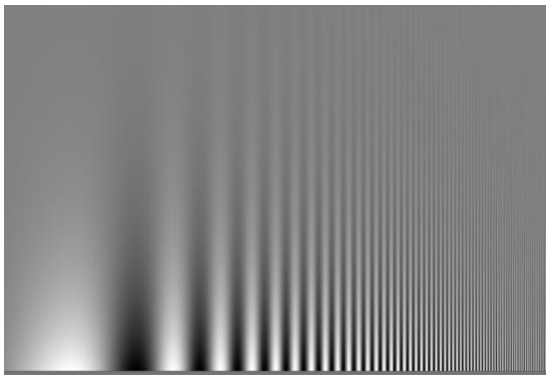



**Spatial Frequency Sensitivity Curve of a Whole Brain**
Same kind of slide as before, but rather than being for a single cell, represents the overall spatial frequency sensitivity curve for a whole observer (person). To measure this kind of curve, you present on a screen a sinusoidal wave, then you make it flatter and flatter until it becomes undetectable (no more contrast respect to the background). However, if the frequency is increased over a certain frequency, the brain is incapable of detecting them and you will end up seeing it as a continuous stream rather than single sinusoidal waves.

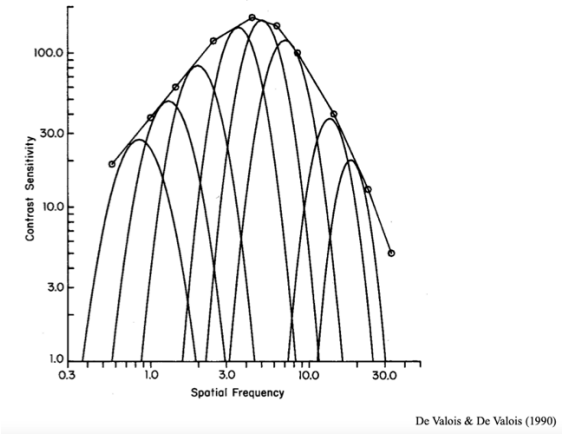



**Contrast Sensitivity Varies with Spatial Frequency**
On the x-axis the frequency of the sinusoidal waves increases, while on the y-axis the contrast decreases.

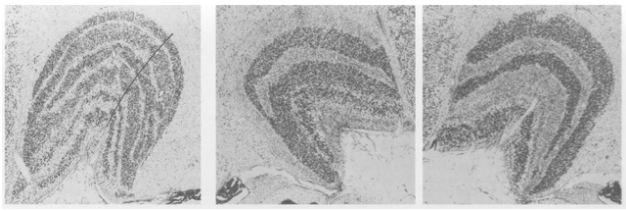
The contrast sensitivity function represents what is in principle visible for an observer.



**One Interpretation of the Contrast Sensitivity Curve**
This picture shows spatial frequency against contrast sensitivity in a whole observer (Macaque monkey in this case). The curves without datapoints are the contrast sensitivity curves of a cortical cell/neuron, and what can be seen is that the cells are highly selective for spatial frequency. So, each individual cell sees only a small range of spatial frequency. Hence, the overall contrast sensitivity of an observer derives from the sum of many individual cells contrast sensitivities.



# Lecture 5 – LGN & V1



## Lateral Geniculate Nucleus (LGN): A Few Facts

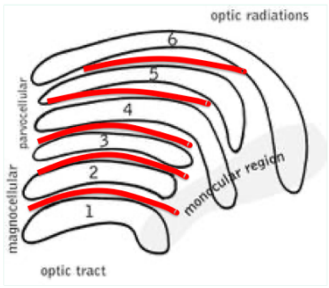
The LGN presents 6 layers as it can be seen in the Macaque section shown to the left. Also, in the picture, the figures in the middle and to the right show that both LGN (left and right of the brain) receive inputs from both eyes, but these inputs are not mixed and remain segregated in the various layers.

Layers 1 and 2 are **Magnocellular** and are bigger cells with particular high contrast sensitivity but not color selective. While layers 3, 4, 5 and 6 are **Parvocellular** which are dimensionally smaller than Magnocellular. Although not particularly color sensitive either, they do show some sensitivity for red/green color opponency. **Koniocellular** layers are the smallest cells and show sensitivity for short-wave lengths, i.e., blue/yellow color opponency. The cells in the LGN share the center-surround structure of ganglion cells.

«Parvo» Cells/layers

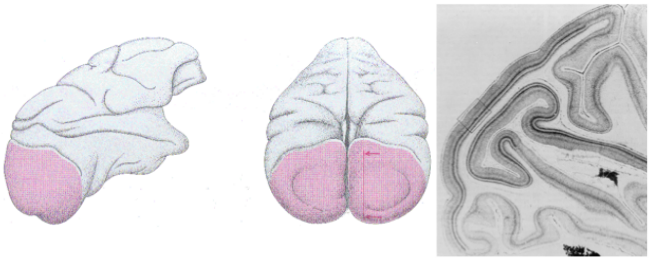
«Konio» Cells/Layers

«Magno» Cells/Layers.




LGN plays a major role as a relay station to project the visual information towards other areas of the cortex. However, it doesn't seem to play any additional role in the processing of the information, compared to the retina.

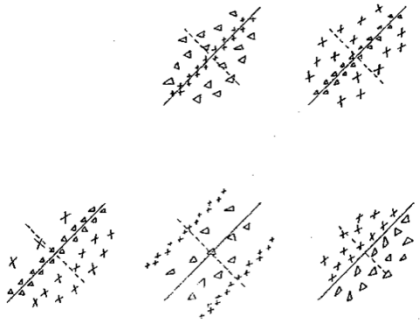
**Primary Visual Cortex** (V1 = Areal 17 = Striate Cortex)

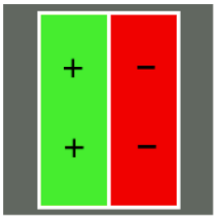


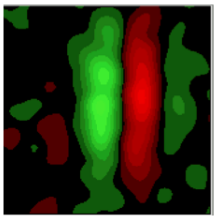
## Cortical Processing

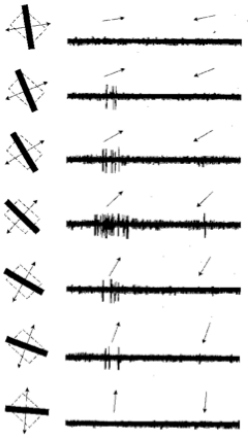
### V1 "Simple Cells" & Selectivity for Stimulus Orientation and Direction

**Orientation** refers to the "orientation" of the bar, while **Direction** refers to the "direction of motion" of the bar.



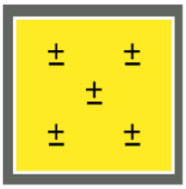


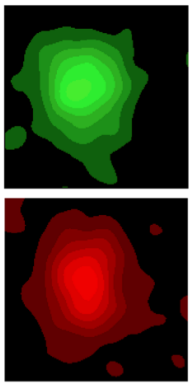





### Receptive Field of a Complex Cell

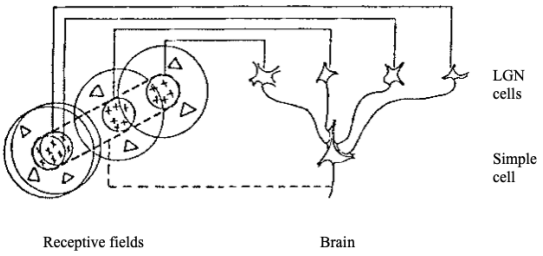
Complex cells show superimposed on/off subregions. Also, the receptive fields of these cells do not have clear orientations. Nonetheless, they are extremely orientation selective, even though it is not easy to identify their preferred orientation preference. Complex cells represent the majority of cells in the cortex.



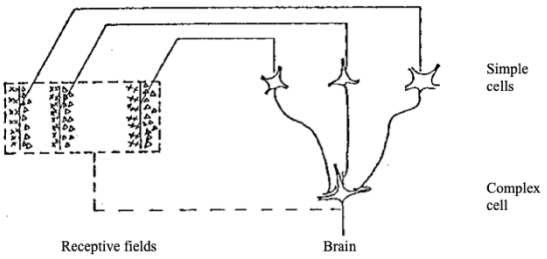
**Hubel & Wiesel's Feedforward Model of Simple Cells**

In this particularly successful model, we have 4 LGN neurons projecting on a particular simple cell in the cortex. Now, on the left we can see the receptive fields of the LGN cells, which happen to be aligned in space. Then, the receptive field of the simple cell will be maximally activated when the on-centers of the LGN cells are stimulated, which occurs when a properly oriented bar is presented. This model is accurate qualitatively, but not quantitatively. Indeed, it has been proven that orientation selectivity is enhanced also by cortical connections.



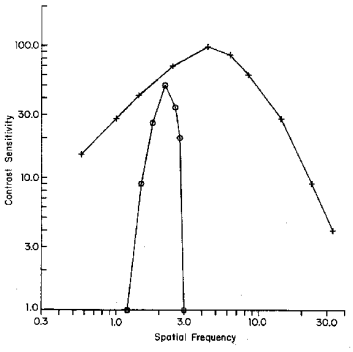

**Hubel & Wiesel's Feedforward Model of Complex Cells**

In this model, the complex cell receives inputs from 3 simple cells. We should imagine the simple cells receptive fields as superimposed rather than spread out like this, which makes it untruthfully easy to determine the preferred orientation of the complex cells.



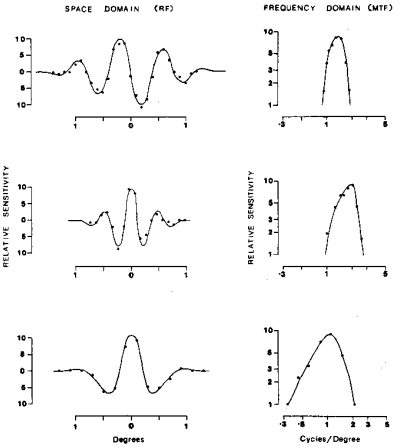

**Perceptual and Neural Sensitivity: Data from a Monkey**

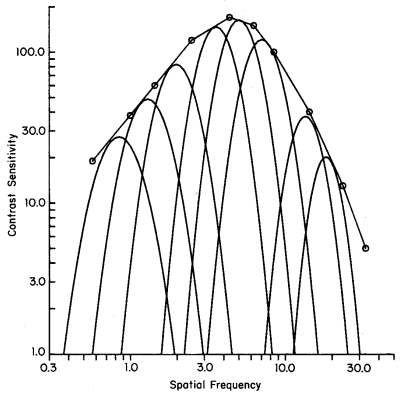
The plot to the right shows the relation between spatial frequency and contrast sensitivity, indicating the level of contrast and frequency at which an individual (x-dotted line) and a single cell (o-dotted line) can detect the stimulus. From the graph we can notice that a single cell is particularly tuned to a narrow range of spatial frequencies.




**For Simple Cells, Knowing the Receptive Field is Knowing Spatial Frequency Tuning**

As we have seen for the ganglion cells, knowing the shape of the receptive field allows us to infer/predict the spatial frequency tuning of a cell. If you make a map of the receptive fields of these simple cells and you know how many on/off regions they have, their size and sensitivity, you can predict the spatial frequency tuning with an almost perfect match. We can notice that the more complex a cell is in terms of on/off regions the narrower the spatial frequency tuning will be. This shows that not all the cells have the same spatial frequency selectivity and that they "peak" at different frequencies.

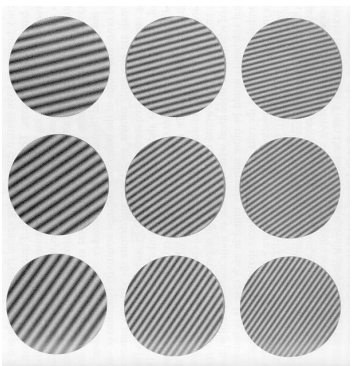



This, in turn, explains the "whole-individual" perceptual sensitivity to spatial frequencies against contrast sensitivity, that we have seen before.

**Selectivity in V1 is Extremely Sharp**

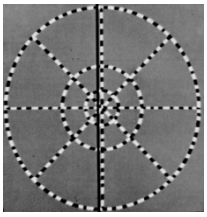
The grating at the center of the gratings' matrix here shows the optimal grating for a particular cell (specificity). All the gratings around it represent gratings to which the cell does not respond anymore. And we can see that just by changing the orientation a little bit (vertical axis), keeping the spatial frequency constant, the cell doesn't respond anymore. And the same can be observed in the opposite case (horizontal axis), where we change the frequency keeping the orientation constant.

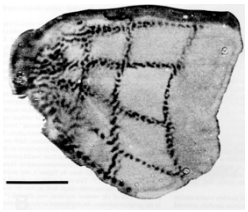


**Retinotopy**

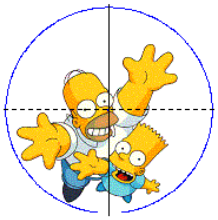
**Cortical Representation Measured with 2-Deoxy-Glucose**

Retinotopy means that you can recognize the image as projected in the retina if we inspect the pattern of cortical activity in V1. There is an orderly representation (tonotopy) of neighboring points in the retina among neighboring neurons in V1. As you can see on the right figure, the right hemisphere is tonotopically organized in monkey's V1.





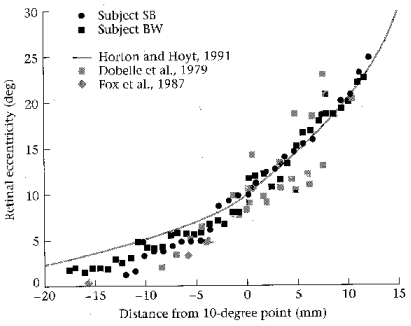
The Simpsons' image shows how the image would be represented in the cortex, i.e., upside-down, reversed and magnified in the foveal region.




**Cortical Magnification**

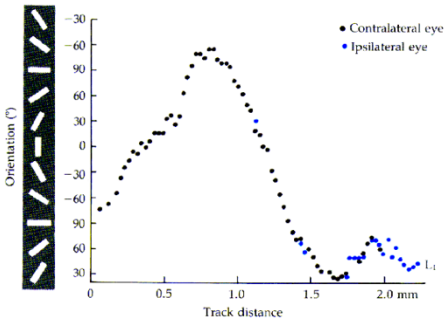
The foveal region of the retina is "magnified" in V1 and takes most of the cortical area (as it can be seen in the top right figure).

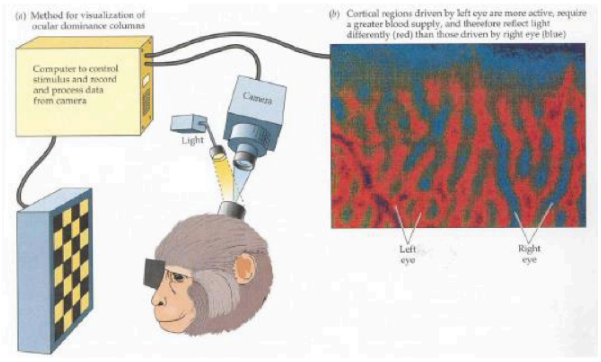
In the figure to the right we can see the relation between the distance of cortical representation and the degree of eccentricity in the retina. Thus, showing that the foveal region is magnified and the relation is not linear, but rather exponential.




**Ocular Dominance**

The leftmost figure shows how the preferred orientation of the various cells in V1 varies gradually (oblique electrode insertions), thus signifying a "columnar" organization of cells in **orientation preference columns**. With the same type of measuring technique, it was realized that certain eyes where dominated by stimuli received from one particular eye, thus indicating a "columnar" organization of cells in **ocular dominance columns**. The middle picture shows another technique to assess ocular dominance regions in the monkeys cortex (red: left eye, blue: right eye).
V1 is the first area where the information of right and left eyes are combined.

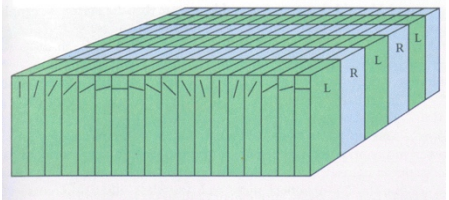




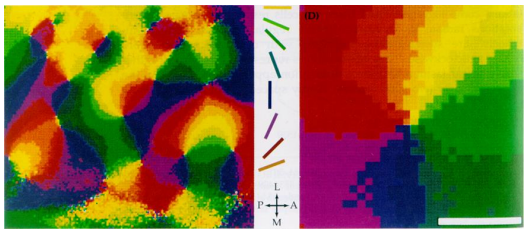
**The "Ice-Cube" Model of Hubel and Wiesel**

This model explains how visual information would be analyzed by the brain. Each "module" (composed of a green and light blue set of cells) analyzes a specific region of the visual field. Where the green and light blue represent the column of ocular dominance, and the other subdivisions represent the orientation preferences of cells in these columns. The subsequent module would analyze the neighboring part of the visual field, thus allowing retinotopy.



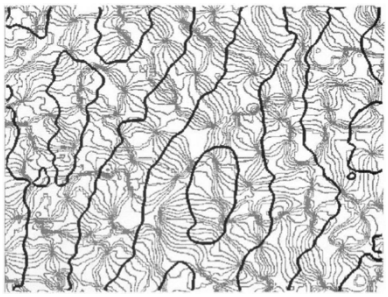
**Orientation Columns Measured with Optical Imaging**

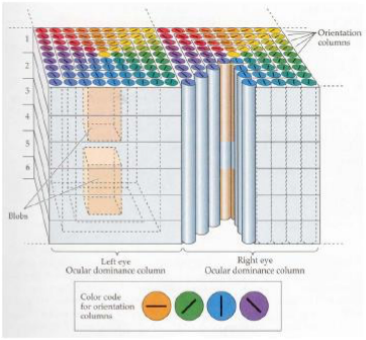
As it can be expected, the model proposed by Hubel & Wiesel is not perfectly accurate, as in biology we rarely see such well defined columns. What it has been measured and represented to the right is the organization of cells in orientation columns.



**Orientation and Ocular Dominance Columns**

While in this figure to the right, we have a representation of the relation between ocular dominance (thick dark lines) and orientation preference columns (grey lines). It is peculiar to notice that most of the intersections between the two types of columns tends to be perpendicular.






**Columnar Organization**

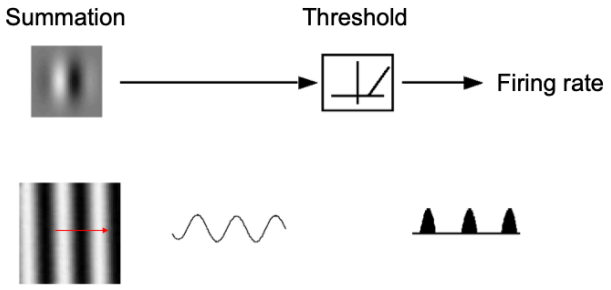
The column is the elementary unit of calculation in the visual cortex. The signals from both eyes for all orientations are represented in one column.

**Linear Model of V1 Simple Cells**

Responses are a weighted average of the stimulus intensity, where the receptive field is the map of the weights.

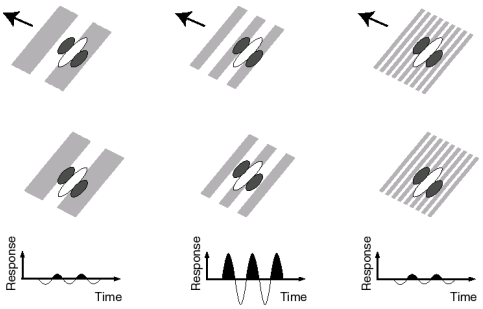
**The Linear Model**

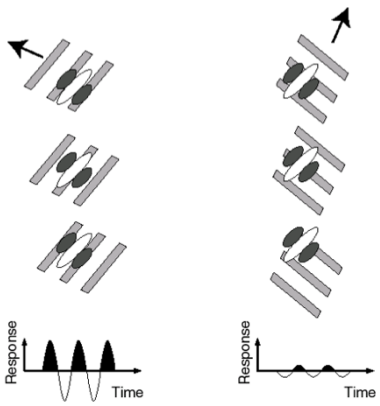
At the bottom of the figure we can see a light stimulus, represented by a sine wave, followed by the response of the cell. The modelling idea is that we have a receptive field as shown in the top left of the figure (vertically oriented in this case) and the "product" of light stimulus and receptive field is passed through a trivial non-linearity, i.e., a threshold. We obtain a "rectified sinewave", where only the positive part "makes it through" the filter.




**For a Linear Cell, Knowing the Receptive Field is Knowing Everything**

Through this linear model, you can predict the responses of neurons, which are dependent on where the light is shown on the receptive field.




**Dependence of Responses on Orientation**

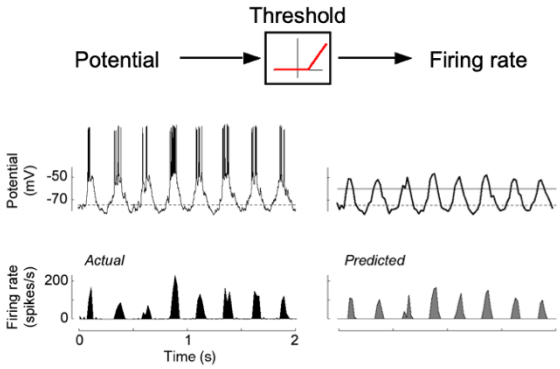
The same thing is true for orientation, you can multiply (convolve) the receptive field by stimuli of different orientations and the model will return the predicted neuronal response.

**Nonlinearities in V1 Responses**

Over the study of V1 cells, it was noticed that different types of non-linearities take place in V1 cells. We will further discuss, which are the situations that make the linear model insufficient and violate the "laws" governing the linear systems.

Linear systems L(x) obey:

- Homogeneity: L(a*x) = a L(x)
- Superposition (Additivity): L(x+y) = L(x) + L(y)

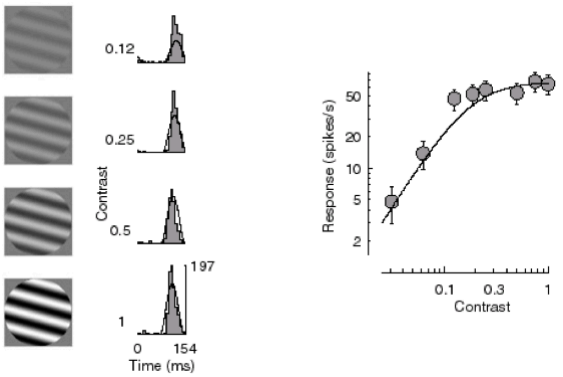



**A Basic Nonlinearity: Thresholding**

In the right figure, we see the membrane potential recorded via intracellular electrode. It analyses how the membrane potential varies in the presence of light stimulus. We can see that it varies sinusoidally and when it reaches a threshold it generates a peak, i.e., an Action Potential. In the lower part of the figure, we can see the measurement made via an extracellular electrode, thus indicating the AP. On the right part of the figure, we see the predicted curves via the linear model, which resemble quite accurately the measured ones.

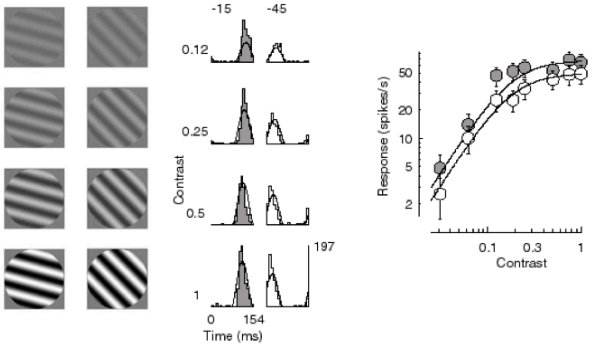
**A Violation of Homogeneity: Saturation**

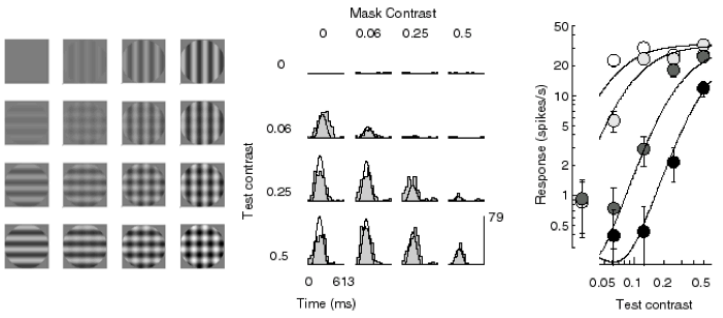
The figure to the right shows that by increasing the contrast in the image, the neuronal response increases as well. However, it does not increase linearly, indeed the response for a 100% contrast, is just a slightly greater than the response for 50% contrast. The graph to the right shows this relationship, which is linear for low contrasts, but it is not linear for higher contrasts (saturation). Thus, violating homogeneity of linear systems.




**Saturation Depends on Contrast**

The figure to the rights explains that saturation does not always occur at the same contrast. Indeed, with different orientations of light stimuli, saturation occurs at different levels. Thus, showing that saturation is not a biophysical limit of neurons, but due to something else.



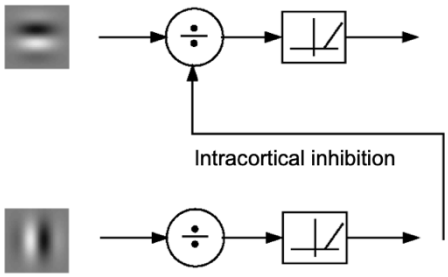




**A Violation of Superposition: Masking**

In the figure to the left we have a cell whose preferred orientation is horizontal and shows practically no response to vertical stimuli. In this experiment, the concept of masking was explained. It refers to the phenomenon where the response of a V1 neuron to a visual stimulus is suppressed by the presence of a nearby stimulus, referred to as the "mask". Specifically, the authors proposed that the response of a V1 neuron to a stimulus is suppressed by the activity of other neurons whose receptive fields overlap with the neuron's receptive field, but whose preferred orientations are different.

## A Nonlinear Model of V1 Simple Cells

To include the concept of masking in the model of V1 neurons, the model in the right figure has been proposed. It includes a divisive inhibition term in the model, which represent the inhibition of a neuron's response by the activity of other neurons. This divisive inhibition term is sensitive to the similarity between the preferred orientations of the neuron and its neighboring neurons. When the preferred orientations are similar, the inhibition is stronger, leading to greater masking.




Overall, the masking effect helps to sharpen the tuning of V1 neurons for specific visual features by reducing their responses to stimuli that are similar but not identical to their preferred features. This tuning is important for extracting relevant visual information from a complex visual scene.

## Adaptation

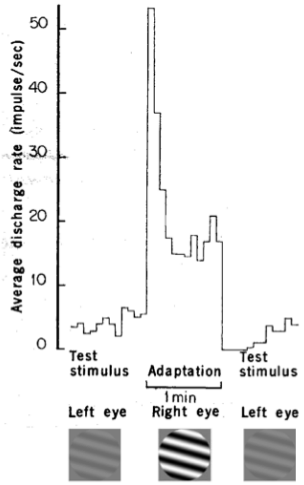
Motion-after effect due to adaptation. It is a perceptual phenomenon in which a stationary stimulus appears to move in the opposite direction after prolonged exposure to a moving stimulus. This effect occurs because the visual system adapts to the prolonged motion in one direction, causing a temporary imbalance in the neural response to motion.

The motion aftereffect can be explained by the process of adaptation, which is a fundamental property of the visual system. Adaptation refers to the ability of neurons in the visual system to adjust their sensitivity to a particular stimulus feature based on prior exposure. In the case of the motion aftereffect, prolonged exposure to a moving stimulus causes neurons in the visual system to adapt to the specific direction and speed of motion. This adaptation results in a decreased response to the adapted motion and an increased response to the opposite motion, which leads to the perception of motion in the opposite direction.

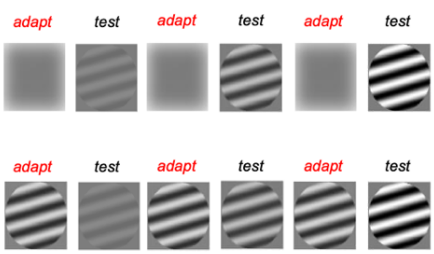
Adaptation is an important mechanism for the visual system to adjust to changes in the environment and maintain sensitivity to relevant stimuli. It allows the visual system to filter out unchanging or irrelevant features of a visual scene and focus on relevant information. In addition to the motion aftereffect, adaptation is responsible for a wide range of perceptual phenomena, including color aftereffects, orientation aftereffects, and size aftereffects.

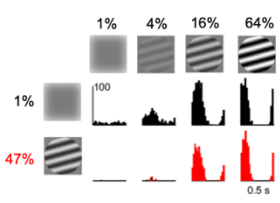
## Adaptation in a V1 Neuron (Violates Linearity)

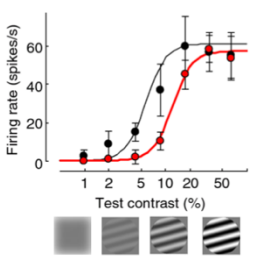
The figure to the right describes one of the first experiments that showed contrast adaptation in the cortex. They presented a low-contrast stimulus to one eye. Then in the other eye they presented a high-contrast stimulus. Finally, they present again the low-contrast stimulus to the left eye and they measured that the response is lower than the response elicited during the first stimulus presentation. In addition, we can see that the adaptation occurs also during the presentation of the high-contrast stimulus as the neuronal response decreases over time.




## Contrast Adaptation Controls V1 Neuron Sensitivity



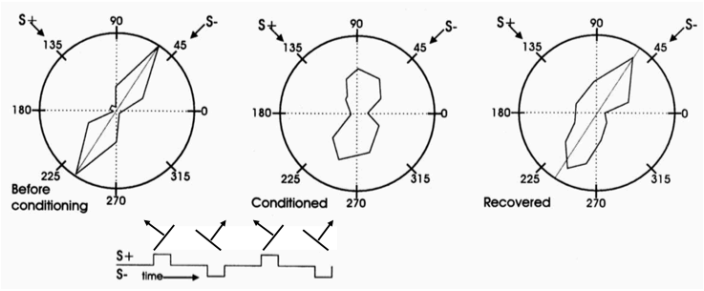



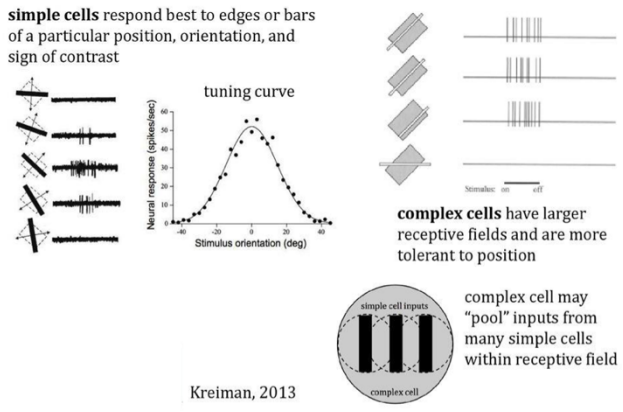
**Learning – Hebbian Learning**


Learning represents another type of nonlinearity. By learning we mean the change in neuronal response based on the experience that the neuron has had. In neuroscience, we often discuss about Hebbian Learning. This concept postulates that the synapses between two neurons are plastic, dynamic. "Neurons that fire together, wire together". The fact that we have dynamic synapses represents a non-linearity, as the response to the same stimulus could potentially vary in time. Is Hebbian Learning happening in V1 simple cells?

**Evidence for Hebbian Learning in V1**

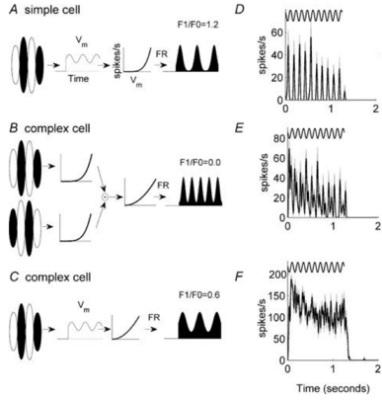



The figure to the right shows that the orientation tuning of simple cells in cats V1. In the experiment, the visual response of a single cell to a light bar was drove to a "high" level (via electrode stimulation) when presented with an initially nonpreferred orientation (S+), and alternately reducing it to a "low" level when presented with the preferred orientation (S). The goal was to test the possible role of neuronal coactivity in controlling the plasticity of orientation selectivity. Approx. 40% of the tested cells showed significant long-lasting changes in their relative orientation preference. The results support the hypothesis that covariance levels between pre- and post-synaptic activity determine the sign and the amplitude of the modification of efficacy or cortical synapsesThus, meaning that V1 simple cells feature Hebbian learning nonlinearities.




**V1 Simple & Complex Cells**

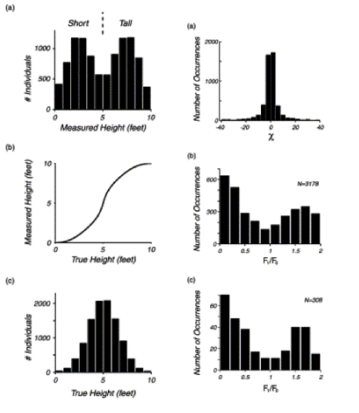
Both simple and complex cells are orientation selective. In the case of the simple cells, presentation of the bars in different areas of the receptive fields have differential effects due to on/off regions, while this doesn't occur in complex cells that show a constant response across the whole receptive field.




**Simple vs Complex Cells Responses**

The figure to the right is interesting to compare the responses of simple cells and complex cells to sinusoidal stimuli. In the case of a simple cell, we obtain a rectified sinewave as we have also seen before. In the case of the complex cell, that receives inputs from two simple cells with opposite receptive fields, we have a doubled frequency of the sinewave, **frequency doubling**.

**Could Simple and Complex Cells be Extremes of a Continuum?**




Simple cells in V1 have segregated On/Off subregions in their receptive fields, while complex cells have overlapping On/Off subregions. These two cell types form the extremes at each end of a continuum of receptive field types. Hubel & Wiesel suggested a hierarchical scheme of processing whereby spatially offset simple cells drive complex cells. In their pioneering studies of V1, Hubel & Wiesel described the existence of two classes of cells, which they termed "simple" and "complex". The original classification scheme was based on a number of partly subjective tests of linear spatial summation. Later, investigators adopted an objective classification method based on the ration between the amplitude of the

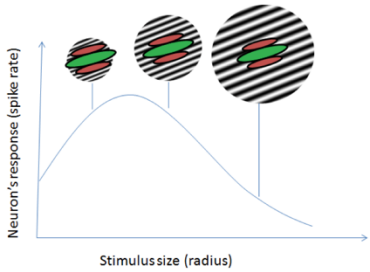
first harmonic of the response and the mean spike rate (or the F1/F0 ratio) when the neuron is stimulated with drifting sinusoidal gratings. This measure is bimodally distributed over the population and divides neurons into two classes that correspond closely to the classical definition by Hubel & Wiesel. Here we show that a simple rectification model can predict the observed bimodal distribution of F1/F0 in V1 when the distributions of the intracellular response modulation and mean are unimodal. Thus, contrary to common belief, the bimodality of F1/F0 does not necessarily imply the existence of two discrete cell classes.

**End-stopping and Hypercomplex Cells**

Hubel & Wiesel found out that certain cells, called hypercomplex cells, showed a decrease in response when the stimulus exceeds the border of the cell receptive field. This phenomenon takes the name of **surround suppression**.

**Surround Suppression Example**

Decreasing firing rate with increasing stimulus size outside the classical receptive field.



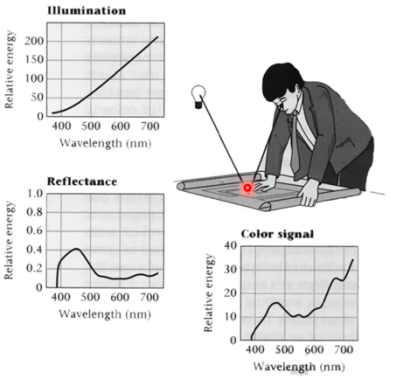

**Surround and Saliency**

In the figure, we have the same sinewave grating presented 3 times with different surrounds. It can be noticed that changing the surround gives us the sense of an increased contrast. When the surround stimulus orientation is different, the stimulus becomes more salient and "pops out" from the background.

# Lecture 8 – Color Vision

**Real Life Color Perception**

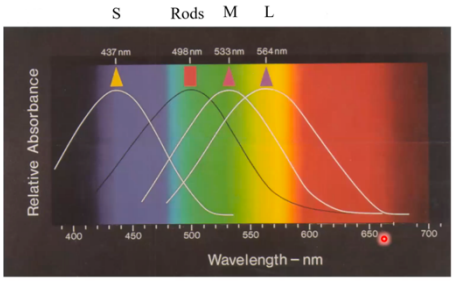
The **spectral composition** of **illumination** represents the set of wavelengths present in light source. Each surface has a particular **reflectance spectrum**, not all surfaces reflect light in the same way. The product of illumination and reflectance gives us the **color signal**, i.e., the color that arrives at the eye, which is a combination of which wavelengths are generated by the source of illumination and which wavelengths are reflected by the surface. The color signal that reaches the eye is then filtered by the three populations of cones receptors present in the retina.




Hence, the signal representing colors that is propagated as information in the brain can be summarized by three numbers, i.e., how much activity the signal generated in the **Long (Red), Medium (Green) /Short (Blue) wavelength sensitive cones**. Keep in mind, that the wavelength composition of the color signal is not sufficient to assess with certainty the color appearance in the eye. Indeed, adaptation, context, background and various factors affect the color perception.

**Photoreceptors Sensitivity Profiles**

The picture to the right shows the same concept as before. The curves show the relative absorbance of wavelengths depending on the cones/rods type. It is important that also the rods are wavelength selective, even though they are mostly used in low-lighting settings. Since we only have three cones type, our vision is defined **trichromatic**. Side note: computer screens feature only three types of pixels matching the three wavelengths preferences of cones photoreceptors.



About 10% of human population suffers from color blindness. People with color blindness have a restricted spectrum of colors. Color blindness occurs when one or more types of cone cells in the retina are absent, non-functioning, or have an altered response to light. In a person with normal color vision, each cone type is most sensitive to a specific range of wavelengths. When light enters the eye and stimulates these cones, the brain integrates the signals from each cone type to perceive color. In individuals with color blindness, one or more of the cone types are either absent or have an altered response to light. This can result in an inability to distinguish certain colors or a reduced ability to see certain color combinations. Color blindness can be inherited or acquired. Inherited color blindness is usually due to a genetic mutation that affects the development or function of the cone cells. Acquired color blindness can occur due to certain diseases or conditions that damage the retina or optic nerve, or due to exposure to certain chemicals or medications that can affect the function of the cone cells.



## Cone Mosaic

The cone mosaic refers to the pattern of cones across the retina, which varies across the retina and between individuals. In the fovea, which is the central region of the retina responsible for high-resolution vision, the cone cells are densely packed and arranged in a regular pattern. This regular arrangement of cone cells is important for high visual acuity, as it allows for precise and detailed information to be transmitted to the brain. In this area, as it can be seen from the figure, we have close-to-none blue cones and an higher density of red cones over green ones. In the peripheral retina, the cone cells are more widely spaced and arranged in a less regular pattern. This arrangement is not as critical for high-resolution vision, as the peripheral retina is more specialized for detecting motion and changes in brightness. The second image shows the inter-individual variability in the cones distribution among human subjects with normal color vision.

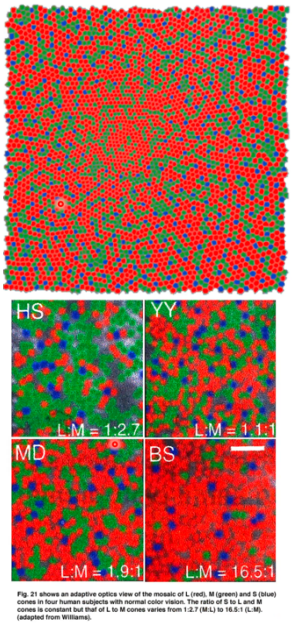
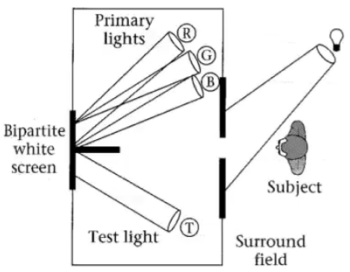
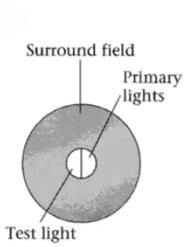


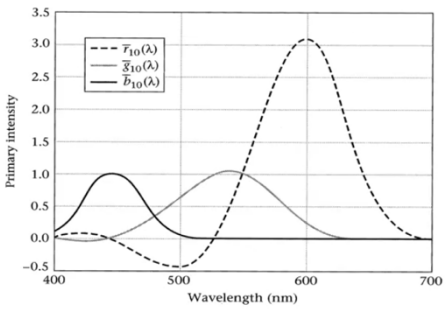

Fig. 21 shows an adaptive optics view of the mosaic of L (red), M (green) and S (blue) cones in four human subjects with normal color vision. The ratio of S to L and M cones is constant but that of L to M cones varies from 1:2.7 (M:L) to 16.5:1 (L:M). (adapted from Williams).

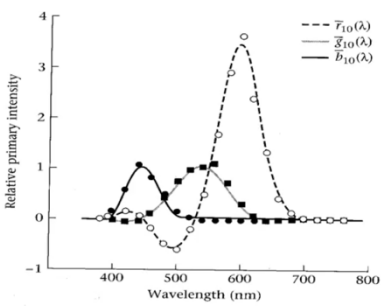
## Color Matching Experiments

The color matching experiment (Maxwell) is a common technique used to study color vision and color perception. In this experiment, a subject is presented with a test color and is asked to adjust the intensity of three primary lights (usually red, green and blue) to match the color of the test stimulus. The primary lights are mixed together in varying intensities until the subject perceives a color that matches the test stimulus. The process is repeated for several test stimuli, each of which has a different wavelength. The resulting data can be plotted as a color matching function, which shows how the intensities of the primary lights change to match different test stimuli. This experiment provides information about the sensitivity and response properties of the three types of cone cells in the retina. By analyzing the data, researchers can estimate the relative sensitivity of each type of cone cell to different wavelengths of light. They can also identify any anomalies or deficiencies in color vision, such as color blindness or abnormal color discrimination.



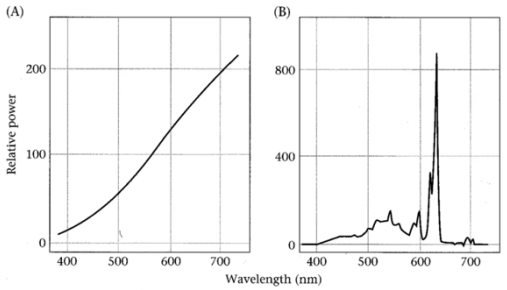



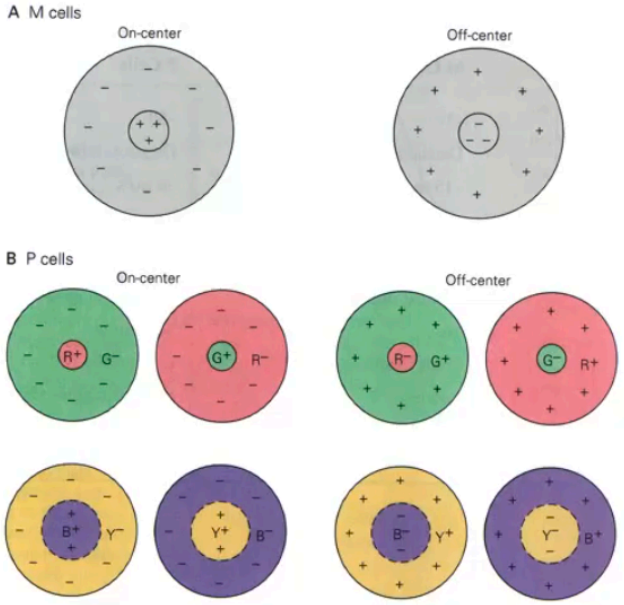
**Two Metameric Spectral Distributions**

Color matching experiments reveal two metameric colors, i.e., colors that are different from a physical point of view, but appear identical to us. Referring to the figure to the right, these two spectral compositions give rise to the same color perception when filtered through the three cones types.




**Receptive Fields of LGN Neurons**

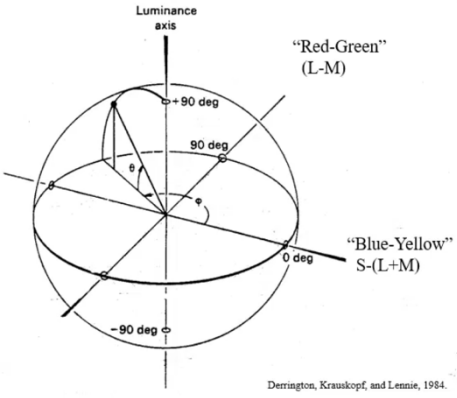
M-cells and Ganglion cells show very similar receptive fields. Approximately 50% of these cells have an on-center, where the cell strongly responds to an increase of light intensity. While the other 50% shows off-center, where the cell strongly responds to a decrease in light intensity. This type of receptive field is great to detect contrast. P-cells also have round receptive fields, but rather than discriminating light intensity, they discriminate between middle and long wavelengths. Furthermore, the size of the centers of the receptive fields changes depending on the location of the cell in the retina. Finally, K-cells show again round receptive fields which distinguish among short wavelengths.




**The "DKL" Color Space**

It is a color space defined on the base of cone pigments and sensitivity of cones in macaque monkeys. They represent the colors in a tridimensional space based on cones excitation. This space has three axis:

- **Luminance (L + M)**, represents the overall brightness of a color, which is determined by the combined responses of the long-wavelength (L) and medium-wavelength (M) sensitive cones in the retina.
- **Chromaticity "Red – Green" (L – M**), represents the difference in the responses of the long-wavelength (L) and medium-wavelength (M) sensitive cones, which corresponds to the red-green axis of color perception.
- **Chromaticity "Blue – Yellow" (S – (L + M))**, represents the difference in the responses of the short-wavelength (S) sensitive cones and the combined responses of the long-wavelength (L) and medium-wavelength (M) sensitive cones, which corresponds to the blue-yellow axis of color perception.

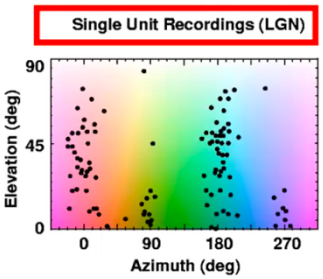



Any color perceived by humans can be represented as a point in this system. In this coordinates system, **azimuth** and **elevation** are defined as follows:

- **Azimuth** (also called **hue angle**, is the angle on the horizonal plane) represents the hue of the color in the chromatic plane. It is the angle formed between the red-green axis (L – M) and the color's chromaticity vector projected onto the chromatic plane.
- **Elevation** (also called **saturation angle**, is the angle on the vertical plane) represents the saturation or chroma of the color. It is the angle formed between the luminance axis (L + M) and the color's chromaticity vector.

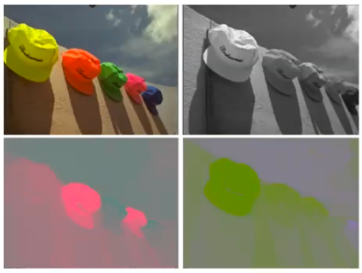
**Preferred Color of pLGN neurons**

If you plot the preferred colors of LGN neurons based on this coordinates system, you obtain a figure like the one to the right. We can notice that the points distribution is condensed in four main "columns" approximately 90 degrees from one another. The 0 and 180 deg populations are the red and green cells, while the 90 and 270 deg are the blue and yellow cells. This demonstrates that LGN cells come in the two families we have discussed except few outliers (which are very unusual). In addition, we can draw some analyses based on the cell position in the plot, e.g., a cell positioned at high-levels of elevation indicates a cell that strongly response to high-levels of light. On the other side, cells positioned at low-levels of elevation indicates cells that strongly respond to colors regardless of illumination.




**The World Seen through the LGN**

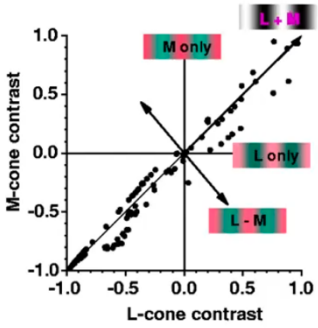
The image to the right gives us an idea of how color vision at the early stages of visual processing might look like. Indeed, superimposing the three subfigures to the right and to the bottom, we obtain the top-left subfigure. Thus, we can notice one channel indicating the red-green information, another the blue-yellow information and one channel describing the light variation information.



There are several studies suggesting that this represents an efficient way of encoding color information, thus explaining why the brain evolved and developed this technique.

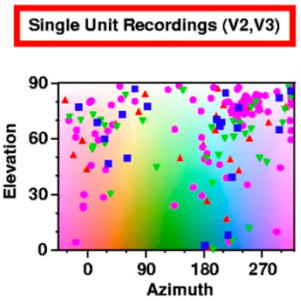
**Color Contrasts of Natural Objects**

In the figure to the right, we have excitation in the long-wavelength sensitive cones as horizontal axis, against the excitation in medium-wavelength sensitive cones. The points represent the excitation of these two types of cones to the pixels of a "natural" image. The most efficient way to transmit this information in the brain is by taking these two orthogonal diagonals, which represent (L + M) and (L – M).



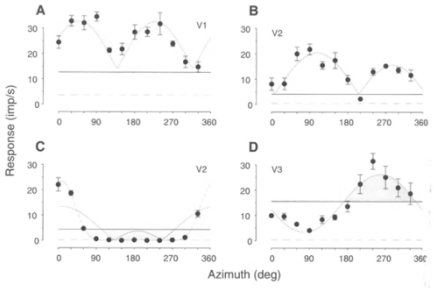

**Preferred Color of Cortical Cells**

The same exact experiment discussed above is performed in V2 and V3. Here, we see major differences. We do not have any more a red-green and a blue-yellow channel, but we have an almost-equal distribution across the azimuth axis. There is a much more widespread distribution over color preferences. In addition, we notice that there is more density in high-levels of elevation, thus indicating that cells are more "interested" in luminance rather than color. Overall, the cells become more specialized with intermediate color preferences and stronger responses to higher levels of illumination.



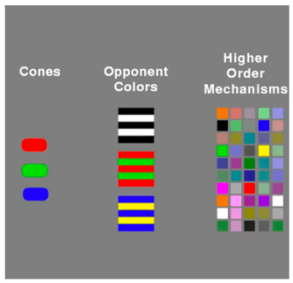

**Examples of Cortical Cells' Responses to Color Variations**

At earlier levels, by varying the azimuth of the stimulus (i.e., changing the color) you will observe plots like the ones to the right. Indeed, the cell will respond differently depending on its color preference. While ganglion cells and LGN cells are pretty much the same, they linearly combine cones' inputs in the same way (red-green cells and blue-yellow cells), when you reach the cortex level of information processing things get more complicated, all the color spectrum is considered, and cells can become selective to extremely specific colors.

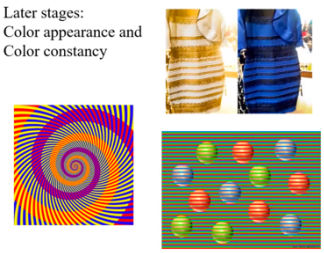


**Three Stages of Color Processing**

What we discussed in color vision can be summarized by the figure to the right, where we have the cones tuned to the three wavelengths associated with red, blue and green. Then, at the level of LGN, we have colors opponents with red-green, blue-yellow and luminance opponency. Finally, in higher order areas (e.g., V1, V2, V3) we have a wide plethora of possibilities, with cells tuned to specific colors.

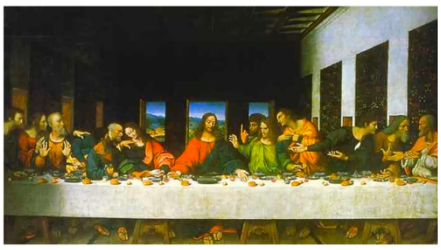


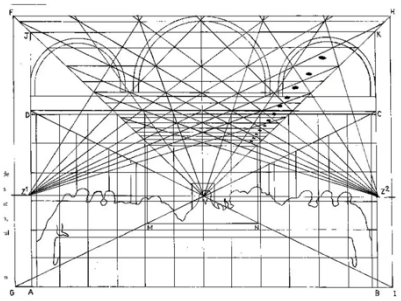

Other major areas of research are **color constancy** and **color appearance**. Color constancy is the ability of humans of maintaining unchanged color appearance of objects even when varying illumination or context. Our brain is capable of discerning the contribution of illumination and reflectance in the color signal perceived, such that we are capable of interpreting different color signals as equal when varying conditions. This level of information processing takes place in extrastriate areas.




# Lecture 9 – Depth Perception

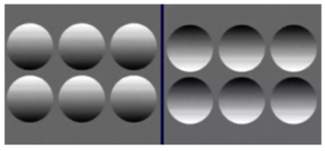
There are many cues that give a 2D image the sense of depth. For example, objects in the front cover objects behind them, or the perspective used by Leonardo Da Vinci in the famous "Ultima Cena".

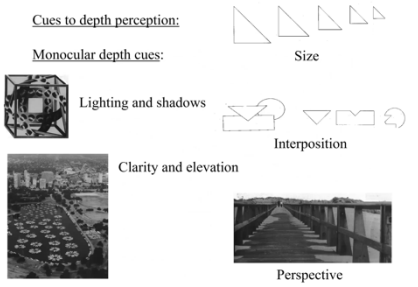


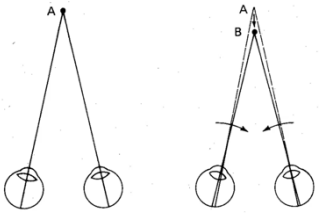


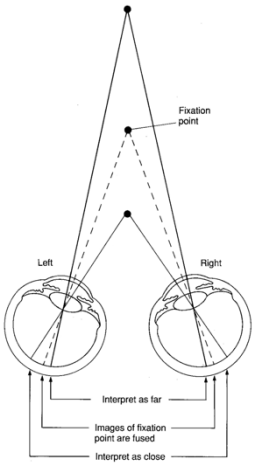
**Cues to Depth Perception**

- **Monocular Depth Cues**:
  - **Size**
  - **Lighting** and **Shadows**: the image to the right shows six dots whose coloring is just rotated by 180 degrees, we perceive the left set as being a concave and the second set as being convex because we are biased to imagine the source of light as coming from above.
  - **Interposition**
  - **Clarity** and **Elevation**
  - **Perspective**
- **Binocular Depth Cues**:
  - **Convergence**: if we are observing point A and we want to observe point B, we will have to perform a convergent movement of the eyes. Such movement allows the brain to understand that point B is closer to us that point A. Potentially, via the convergence angle we could estimate the absolute distance of an objects. However, our brain has not evolved to do that.
  - **Stereopsis** (**Binocular Disparity**): the fact that the two images that we have in the two retinas are not exactly the same allows the brain to estimate the depth of objects in the scene. Referring to the image, if we fix a point, that would be represented in the foveal point of the retina. However, farther points would be represented to the left of the fovea in the right eye and opposite for the other eye. These represent non-corresponding points on the two eyes, which are exploited by the brain to estimate depth. Contrarily to the possibility of evaluating the absolute





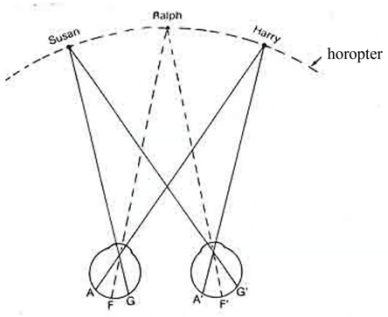



distance of an object thanks to the convergence of eyes, binocular disparity allows the evaluation of relative distance of an object from the fixation point.

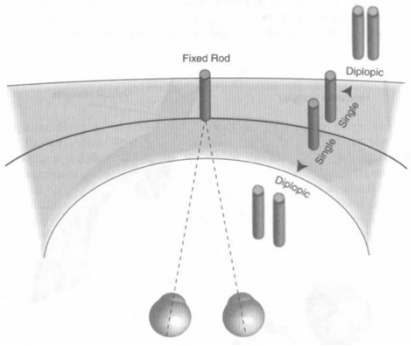
**The Horopter**

The horopter refers to the set of points in space that are perceived as being at the same depth or distance as the fixation point when both eyes are focused on that point. The horopter is an important concept in binocular vision and stereopsis, as it helps explain how the brain uses the disparity in the images captured by the two eyes to perceive depth and create a 3D representation of the environment. When both eyes are focused on a single point in space, the visual system aligns the corresponding retinal points of the two eyes. Points on the horopter fall on these corresponding retinal points in both eyes, meaning they have zero disparity and are perceived as being at the same depth as the fixation point. The horopter is not a fixed geometric shape but changes depending on the position of the fixation point and the angle of convergence of the eyes.




**The Horopter and Panum's Fusion Area**

When both eyes are focused on a point in space, the visual system aligns the corresponding retinal points of the two eyes. Points on the horopter fall on these corresponding retinal points and are perceived as being at the same depth as the fixation point. However, the visual system can tolerate small disparities and still fuse the images from the two eyes into a single perception of depth. This tolerance is described by Panum's fusion area, which is the region around the horopter where the disparities between the images are small enough for the brain to combine the images and perceive them as a single 3D image.




Outside of Panum's fusion area, the disparities between the images from the two eyes become too large for the visual system to fuse them effectively, leading to a phenomenon called diplopia, or double vision. This occurs because the visual system cannot reconcile the differences between the two images and perceive them as a single, coherent 3D image. Panum's fusion area is important in understanding how the brain processes depth information from the two eyes and forms a coherent, single perception of the environment. It highlights the visual system's ability to tolerate and fuse small disparities, which is crucial for achieving binocular depth perception and stereopsis.

**The Wheatstone Stereoscope**

It is an optical instrument that allows to demonstrate the perception of depth in binocular vision. It uses two separate images, each representing the view of a scene from the perspective of one eye, to create the illusion of a single, 3D image when viewed with both eyes.

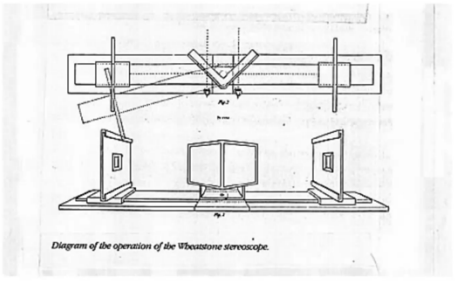
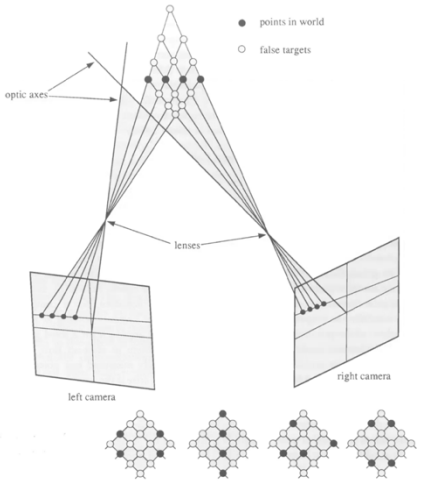


Diagram of the operation of the Wheatstone stereoscope.

- Two images are created by capturing or drawing the same scene from two slightly different viewpoints, mimicking the horizontal separation of the left and right eyes. These images, known as stereopairs or stereograms, contain slightly different information, just as the images captured by our eyes do.
- The stereoscope consists of a frame with two mirrors placed at a 90-degree angle to each other, with one mirror in front of each eye. The stereopair images are positioned on either side of the mirrors, facing them.
- When the viewer looks into the mirrors, each eye sees the reflection of one of the images from the stereopair. The mirrors ensure that the images are aligned correctly, with the left eye viewing the left image and the right eye viewing the right image.

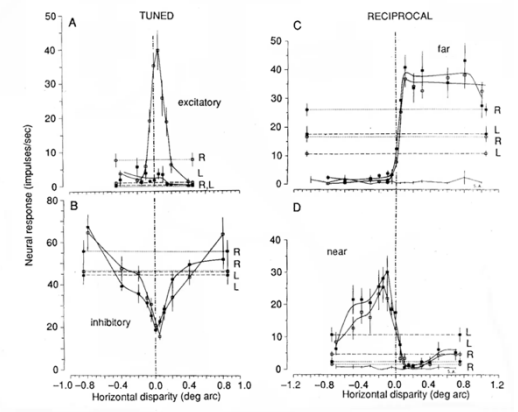
- The brain processes the images from each eye separately and then combines them into a single, coherent image. Due to the differences between the images, the brain perceives depth in the scene, creating the illusion of a 3D image.

**The Correspondence Problem in Stereopsis**

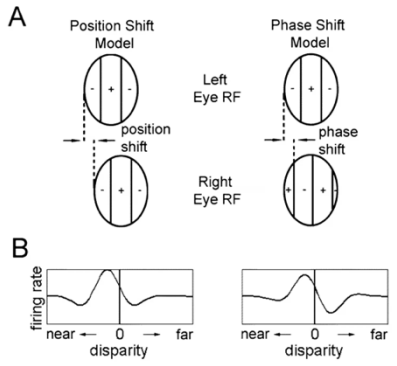



Referring to the figure on the right, it is difficult to establish the relative position of the objects on the two retinas. In order to be sure that a certain point on the left retina matches another specific point on the right retina, we have to assume a distribution of the points. What the image shows at the bottom are all valid distributions for the 4 points based on their representation on the retina. In order to do so, the brain exploits a heuristic approach. The correspondence problem arises because the visual system must search for matching points or features in two different images that contain many similar or repetitive elements. This can be particularly challenging in complex or textured scenes, where many elements might appear similar and finding the correct correspondence becomes more difficult.

**Neurons Tuned for Binocular Disparity (monkey V1)**



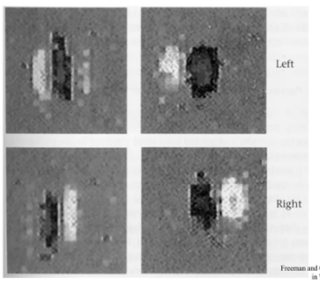

The graphs to the right show individual V1 cells responses measured in monkeys during individual stimulation of both eyes varying the amount of binocular disparity. They labelled positive disparity for objects farther away from the fixation point and negative disparity for objects closer than the fixation point. They find out that there is a number of different cells profiles based on how they respond to different disparity stimuli. For example, the graph on top-right corner shows a cell that strongly response to a positive disparity in the stimulus (i.e., a cell that encodes objects far away from the fixation point), while the bottom-left figure shows a cell that behaves oppositely, i.e., increases firing in the presence of negative disparity. The graph on top-right shows a cell that is perfectly tuned to a specific amount of binocular disparity and strongly responds in such case. The bottom-left graph shows a cell that is again tuned to a specific amount of binocular disparity, but shows an inhibitory effect rather than excitatory.




How are these cells in the cortex tuned for non-zero disparity?
Two models have been proposed:

- **Position Shift Model**: according to this model, the brain features some "disparity-tuned neurons" that are shifted in position such that when the stimulus with binocular disparity arrives to V1 it maximally excites both neurons**.**
- **Phase Shift Model**: according to this model, the receptive fields of the neurons match in terms of position but differ in terms of internal organization such that they are maximally excited when the stimulus has a binocular disparity.

**Binocular Receptive Fields of Disparity Tuned Neurons**

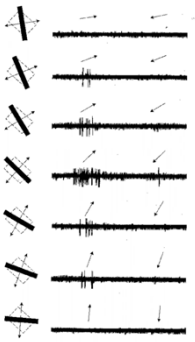



In the image to the right, the receptive fields of disparity tuned neurons have been mapped in a cat cortex. It has been observed that both phase shift and position shift models seem to occur and also a mix between them. The first column of images shows a phase shift of the receptive fields between left and right eye. While in the second column we have both a phase shift and a position shift of the

receptive fields between the left and the right eye. (white region = on-region, black region = off-region).

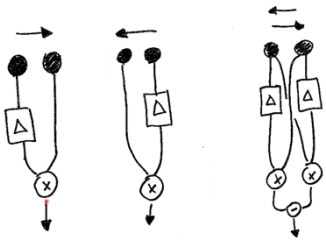
# Lecture 10 – Motion Perception

## Selectivity for Stimulus Orientation and Direction



Hubel and Wiesel performed single neurons recording of cats V1 neurons. On top of what we already discussed, i.e., orientation selectivity, this image shows us that certain neurons show selectivity for the direction of motion of the stimulus. In primates this selectivity arises in the cortex, but certain species feature this selectivity also in the retina.

## Reichardt Detector



This is a computational model or neural circuit designed to identify motion or movement in a visual scene by computing the relative change in intensity across space and time. It compares the output of two spatially separated photoreceptor neurons, which are sensitive to light intensity changes. These neurons have their outputs delayed and multiplied, and the final output is a measure of correlation between the intensity changes at the two spatial locations. If the change an intensity at one location closely follows the change in intensity at the other location, with a small delay, the detector signals motion in a specific direction.

## Visual Motion: 2D

## The "Aperture Problem"

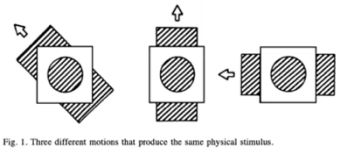
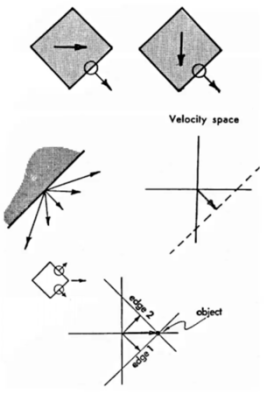


Fig. 1. Three different motions that produce the same physical stimulus.

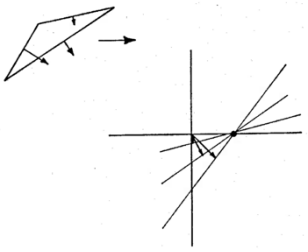
This is a fundamental issue in the perception of motion that arises due to the limited spatial extent of the receptive fields of visual neurons, such as those in the primary visual cortex (V1). When a neuron's receptive field is limited or "apertured", it can only "see" a small portion of the visual scene, making it challenging to accurately determine the true motion direction of an object or pattern. For example, in the barber pole illusion we have diagonal stripes on a vertically rotating cylinder that appear to move vertically rather than diagonally, which is their true motion direction. The illusion occurs due to the limited spatial extent of V1 neurons' receptive fields and their sensitivity to specific orientations. When observing the barber pole, V1 neurons with receptive fields that match the orientation of the diagonal stripes are activated. However, because of their small receptive fields, they only "see" a small portion of the stripes and cannot determine the true motion direction. Instead, they detect motion along the stripes' orientation, which is perpendicular to the stripes. At the same time, the edges of the cylinder constrain the visible motion of the stripes to the vertical direction. This vertical motion is consistent with the motion detected by V1 neurons along the stripes' orientation, and it becomes the dominance percept.

## The "Aperture Problem" in the Visual System

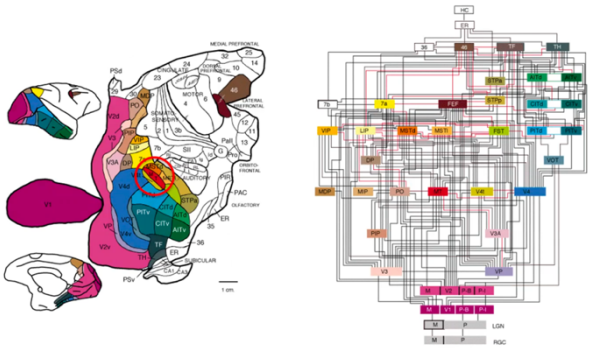



This is problem is relevant because our visual system is "seeing" the world through a bunch of small apertures. In the experiment to the right, it was investigated how the visual system resolves the aperture problem by studying the responses of direction-selective MT neurons in the macaque monkey. Their findings demonstrated that MT neurons can integrate local motion signals from V1 neurons to determine the true global motion direction and that they are highly sensitive to coherent motion, even in noisy visual stimuli. Referring to the picture, we can notice that by using only the receptive field of a single V1 cell we are not sure regarding the velocity of movement of the object and multiple velocities are possible, however integrating over two V1 receptive fields allows us to disambiguate and identify the single velocity vector representing the

global motion. Referring to the picture on the right: the left-hand diamond moves to the right; the right-hand diamond moves down. Note that in both cases, in the local region indicated on each diamond by the small circle, the border moves downward and to the right. The moving edge (under the "diamonds") which could represent a magnified view of the circled regions of the diamonds' borders can be generated by any of the motions shown by the arrows. Motion parallel to the edge is not visible, so all motions that have the same component of motion normal to the edge are possible candidates for the "true" motion giving rise to the observed motion of the edge. We may map this set of possible motions as a locus in "velocity space".

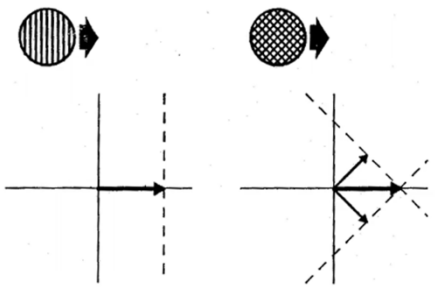


**Intersection of Constraints**: This concept summarizes what we have been discussing: each neuron sensitive to motion provides a constraint on the possible motion direction of an object or pattern in the visual scene, based on the neuron's preferred orientation. However, due to the aperture problem, the motion information provided by a single neuron is ambiguous. By combining the constraints from multiple neurons with different preferred orientations, the visual system can determine the true motion direction where the constraints intersect.



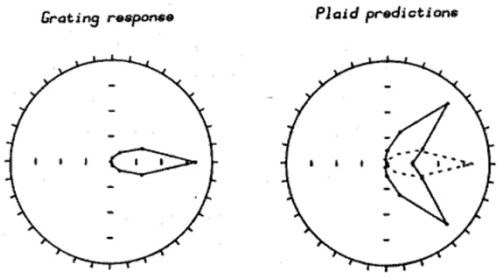
**Area MT – Responses to Moving Stimuli**

**Perceived Direction of Gratings and Plaids**



We can construct a stimulus that incorporates the aperture problem. For example, gratings are regular patterns of alternating light and dark bars, while plaids are formed by superimposing two or more gratings of different orientations. In the context of the aperture problem, gratings and plaids can be used to demonstrate the limitation of local motion processing. For instance, a grating moving behind a circular aperture will produce an ambiguous motion signal, as V1 neurons with small receptive fields will only detect motion along the grating's orientation. In the case of Plaids, there is no ambiguity about the motion of the whole pattern, since the two families of possible velocities (shown by the dotted lines) intersect at a single point.

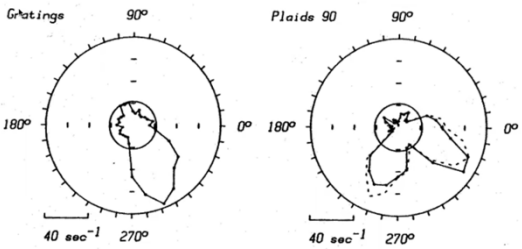
**Component and Pattern Direction Selectivity**




The figure to the right illustrates the response of a hypothetical direction selective neuron. In each plot the direction of motion of the stimulus is given by the angle, and the response of the cell to that direction is given by the distance of the point from the origin. The left-hand plot reveals that this "neuron" responded best to gratings moving directly rightward and did not respond to leftward motion. The direction tuning curve for a single grating therefore has a single peak corresponding to the best direction of motion. When one component of a 90 degree plaid (one whose components are oriented at 90 degrees to one another) is within the direction bandwidth of the neuron, the other component will be outside the acceptable range. If the neuron is component direction selective, the predicted direction tuning curve to a plaid then, is the

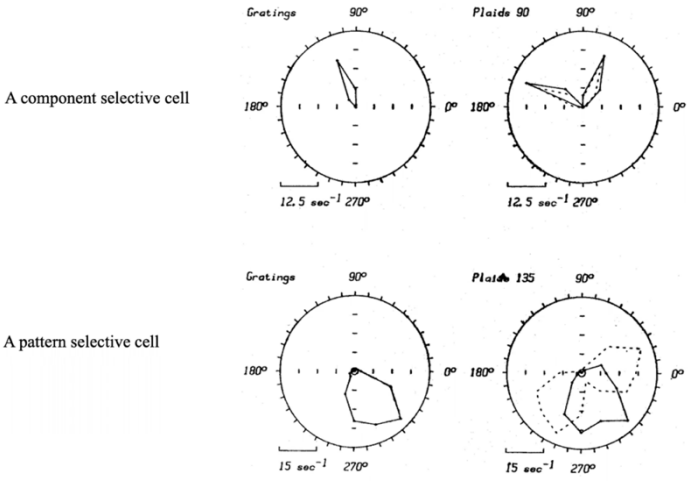
sum of the responses to the two components presented separately. Before the responses are added, however, any spontaneous firing rate (here zero) is subtracted from each. After the two responses are added, the spontaneous rate is added back in. In the right-hand plot, responses are plotted as a function of the direction of motion of a plaid. When the plaid is moving in the optimal direction (as determined with a single grating), the components will be oriented 45 degrees to either side of the optimum. Thus the response peaks are also shifted to either side by 45 degrees, and the predicted tuning curve for the plaid is a bi-lobed curve whose peaks straddle the single peak derived from the single grating experiment. This prediction is shown by solid lines in the right-hand plot. The prediction for pattern direction selectivity is even simpler: the neuron's tuning curves for the two stimuli should be similar since their directions of motion are the same. The predicted tuning curve is thus simply the curve derived from the single grating experiment, and is shown by dashed lines in the right-hand plot. The basis of this test is to dissociate the oriented components of a pattern from the direction in which they move: a single grating always moves at right angles to its orientation, but the plaids move at a different angle to their oriented components (+5 deg in the figure). The two predictions for the different types of direction selectivity are radically different and one may simply see whether the neuron's response depends on the overall direction of motion, or on the orientation of the moving components.

**Responses of a V1 Cell**



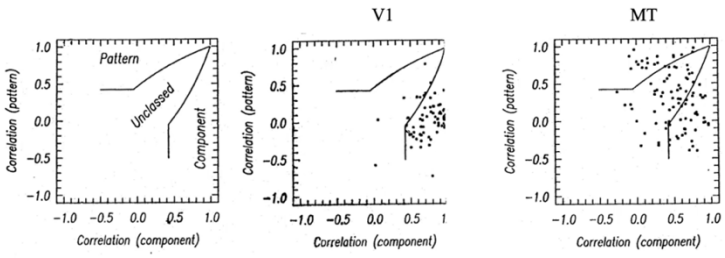

In the figure: directional selectivity of a special complex cell recorded in area 17 of a cat. On the left is shown the neuron's tuning for the direction of motion of single gratings, and on the right is shown the neuron's response to moving 90 deg plaids. The dashed curve on the right shows the expected response of a component direction selective neuron. The inner circles in each plot show the neuron's maintained discharge level.

**Responses of two MT Cells**




This figure shows data, in a format similar to the figure before, for two neurons recorded from MT. The neuron in the first "row" preferred upward movement of single gratings, like its component direction selective counterpart in V1 (previous image), this preference was translated into a dual preference for two directions 45 deg apart when it was tested with 90 deg plaids. As comparison of the data with the dashed lines in the right-hand plot of the first line reveals, the component direction selective prediction provided a very good description of this behavior. About 40% of the cells we studied in MT were clearly component direction selective. The second line of the figure shows data from a neuron in MT whose behavior was rather different. This neuron preferred downward and rightward movement of grating stimuli, and maintained this preference when tested with 135 deg plaids. The actual response to plaids differed very dramatically from the component direction selective prediction.

**Population Analysis**




In order to examine the distribution of behavior of neurons in different areas, this figure shows scatter diagrams in which the values of the pattern and component correlation coefficients were plotted against one another. The first subplot illustrates the significance of various regions of these plots. The region marked "component" is a zone in which the component correlation coefficient significantly exceeds either zero or the pattern correlation coefficient, whichever is larger. The region marked "pattern" similarly marks neurons that were unambiguously pattern direction selective. The region marked "unclassed" represents cases in which both pattern and component correlations significantly exceeded zero, but did not differ significantly from one another, or cases in which neither correlation coefficient differed significantly from zero. The middle subplot shows a scatter plot of data in this space for 69 neurons recorded from cat and monkey V1. It is clear that these cluster around a component correlation value of 1 and a pattern correlation value of zero. While a few neurons lie in the two indeterminate regions of the plot, no clearly pattern direction selective cases exist. The left subplot shows a scatter diagram of the directional selectivity of 108 neurons in MT, tested with 135 deg plaids. Most neurons in MT are rather more broadly tuned for direction than their counterparts in V1, and in consequence the distinction between the component and pattern predictions cannot often be made very clearly with 90 deg plaids.