Simulations (Part1): Human Visual System

Earl Wong

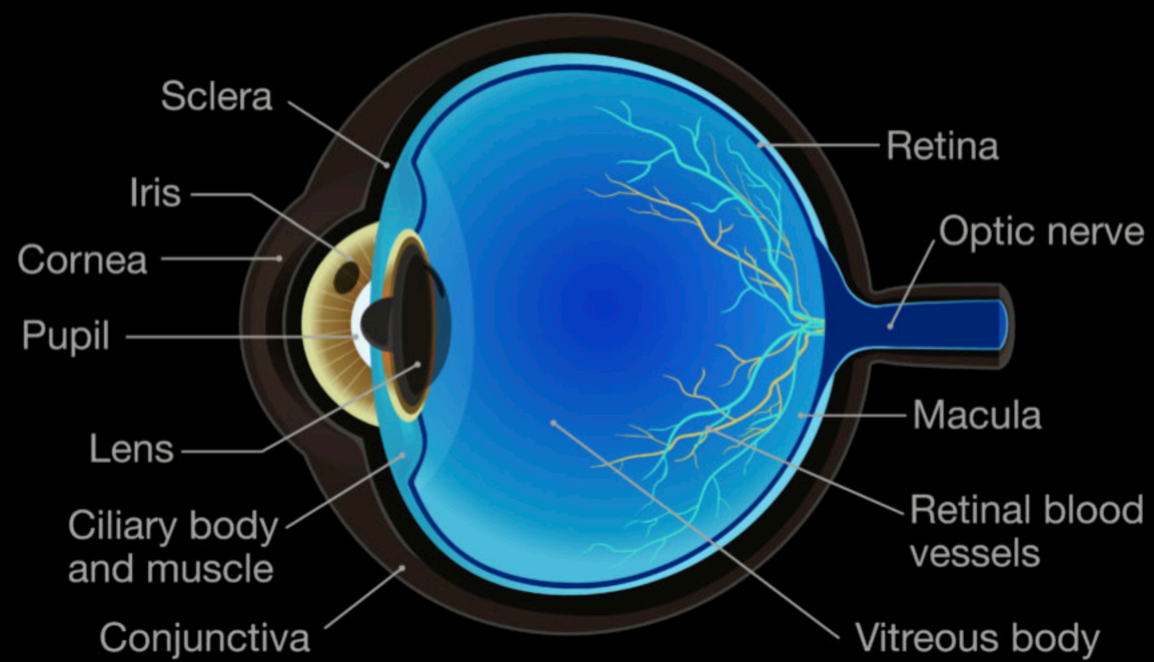
Overview

- Part I discusses the foundations of the human visual system and vision.
- Armed with this understanding, the role of simulations in RL - specifically, it’s applications and shortcoming - can be better understood.
- Part II discusses different simulation techniques used in vision based reinforcement learning.

Overview

- The brain consists of approximately 200 structures for processing five senses: sight, sound, smell, taste, touch
- In order to understand why simulations for vision related tasks may or may not be effective, it is important to understand the visual pathway / how visual information is processed.
- There are several main processing units in the visual pathway: the eye, the LGN and the primary visual cortex
- Each unit is composed of different entities and cells with different functionalities.
- Eventually, signals emerge from the visual cortex (via motor neurons) that control muscle contractions.

The Anatomy of the Eye



We begin with the eye.

Light enters through the cornea, the clear front window of the eye. The cornea transmits and focuses light into the eye.

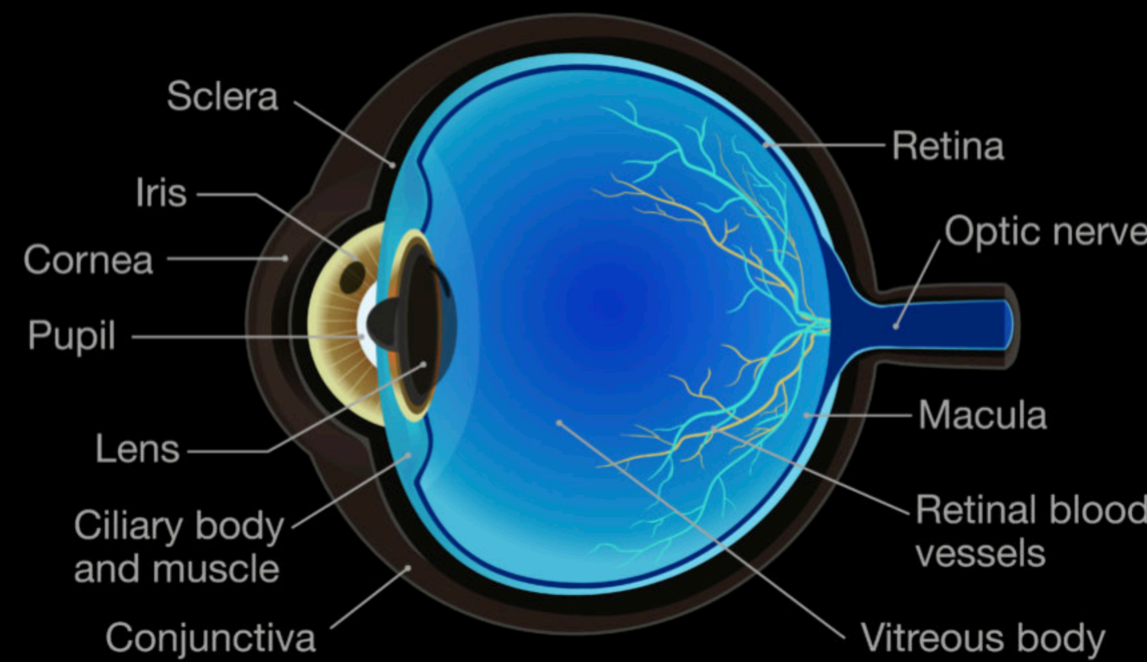
(Corrective laser surgery is often used to reshape the cornea, thereby changing the focus.)

The pupil controls the amount of light entering the eye, and functions much like the aperture of a camera. In bright scenes, the pupil contracts, reducing the amount of light entering the eye. In dark scenes, the pupil expands, increasing the amount of light entering the eye.

The iris (=the colored part of the eye) controls the pupil size.

The outer white part of the eye that surrounds the iris is called the sclera.

The Anatomy of the Eye



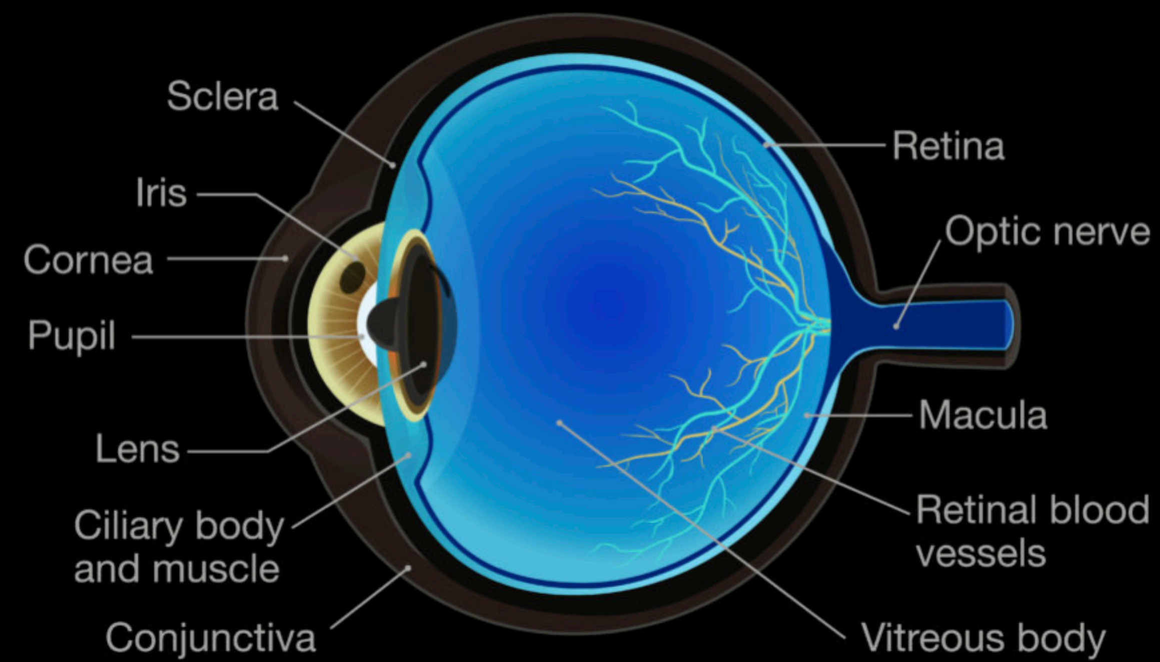
The lens is a transparent structure that focuses incoming light rays onto the retina.

(As we age, proteins in the eye break down. Eventually, a cataract develops, making the lens foggy. This results in blurry vision. In cataract surgery, the blurry lens is replaced with an artificial lens implant.)

The ciliary body and muscle control the focus of the lens.

The vitreous body is the central cavity of the eye. It is filled with a gelatinous substance.

The Anatomy of the Eye



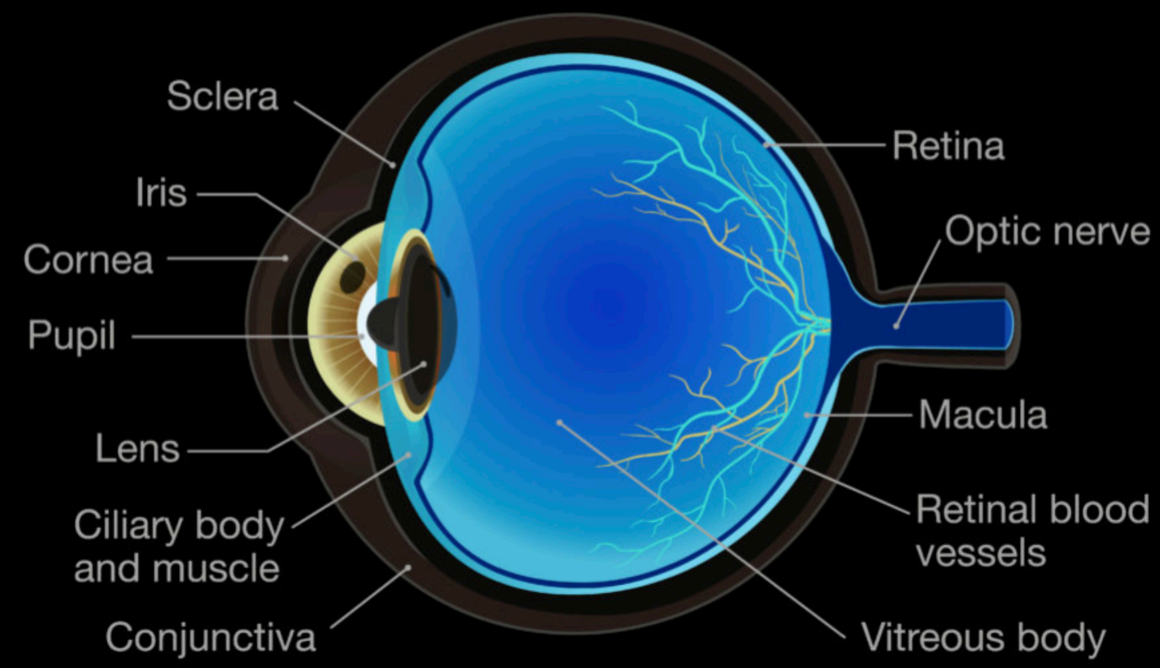
The retina is a nerve layer that lines the back of the eye. It is ~1/4 mm thick. It senses light and sends electrical impulses to the brain via the optic nerve.

The macula is the area of the retina that contains light sensitive cells.

The fovea is the “center part” of the macula, and provides sharp vision.

The retinal blood vessels supply oxygen nutrients to the retina, keeping the retina alive.

The Functionality of the Eye



Human vision depends on the conversion of light to electrochemical signals.

Light flows through the cornea and the lens.

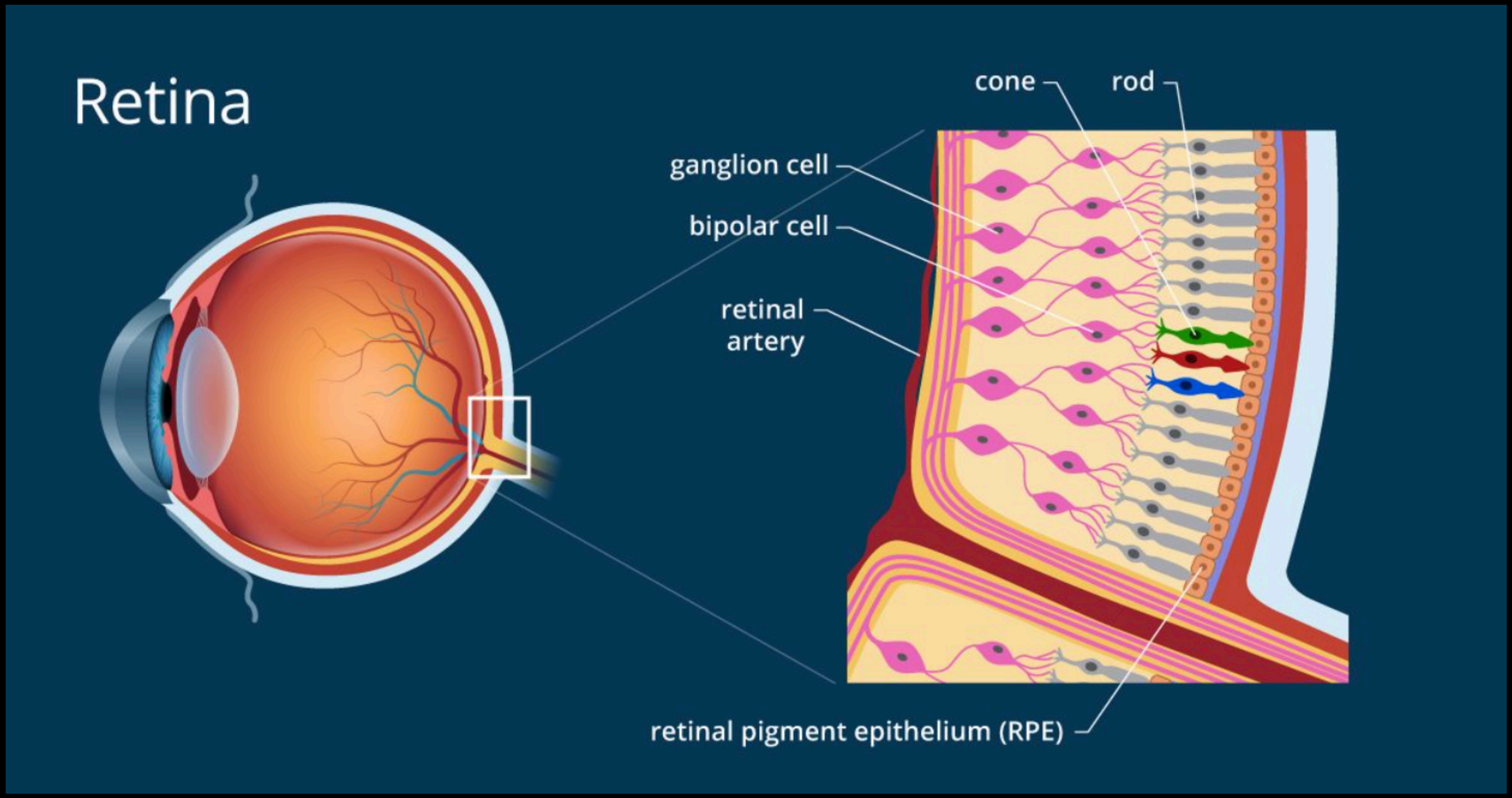
Both the cornea and the lens focus the incoming light onto the retina.

The cells in the retina then convert the light to electrochemical impulses.

These impulses are sent to the LGN via the optic nerve.

The pupil and the iris regulate the amount of light that reaches the retina, much like the aperture of a camera.

The Retina



The retina is a sensory membrane found at the back of the eye.

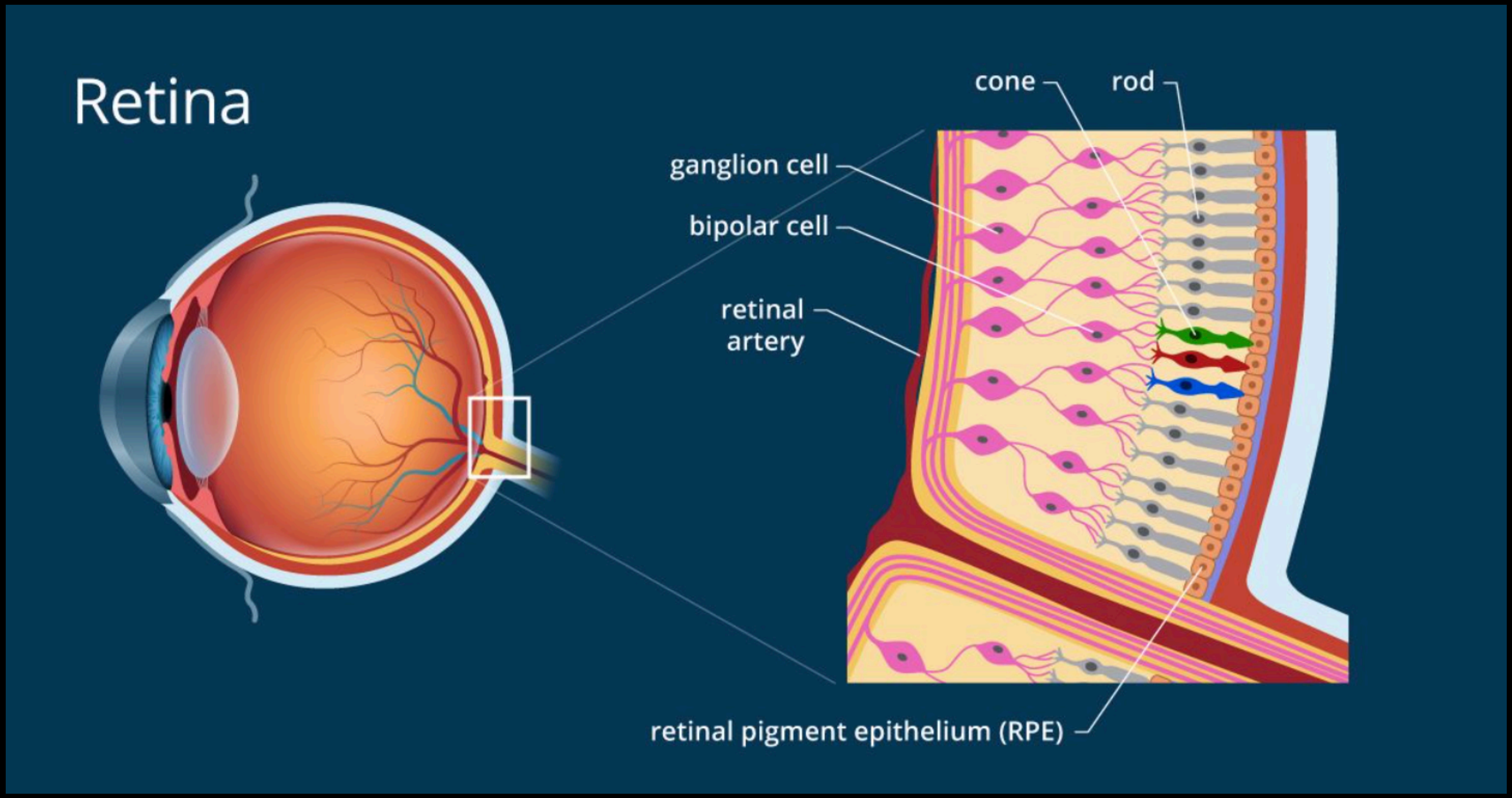
The retina consist of several cell layers: photoreceptors, bipolar cells and ganglion cells

The special photo (light) receptors in the retina consist of rods and cones.

Cones provide color vision. (R, G and B cones have spectral sensitivities closely tied to the the red, green and blue wavelengths.) Cones function well in medium to bright light.

Rods detect scene motion and provide black and white vision. Rods functions well in low light.

The Retina



Rods are located throughout the retina.

Cones are concentrated in the macula.

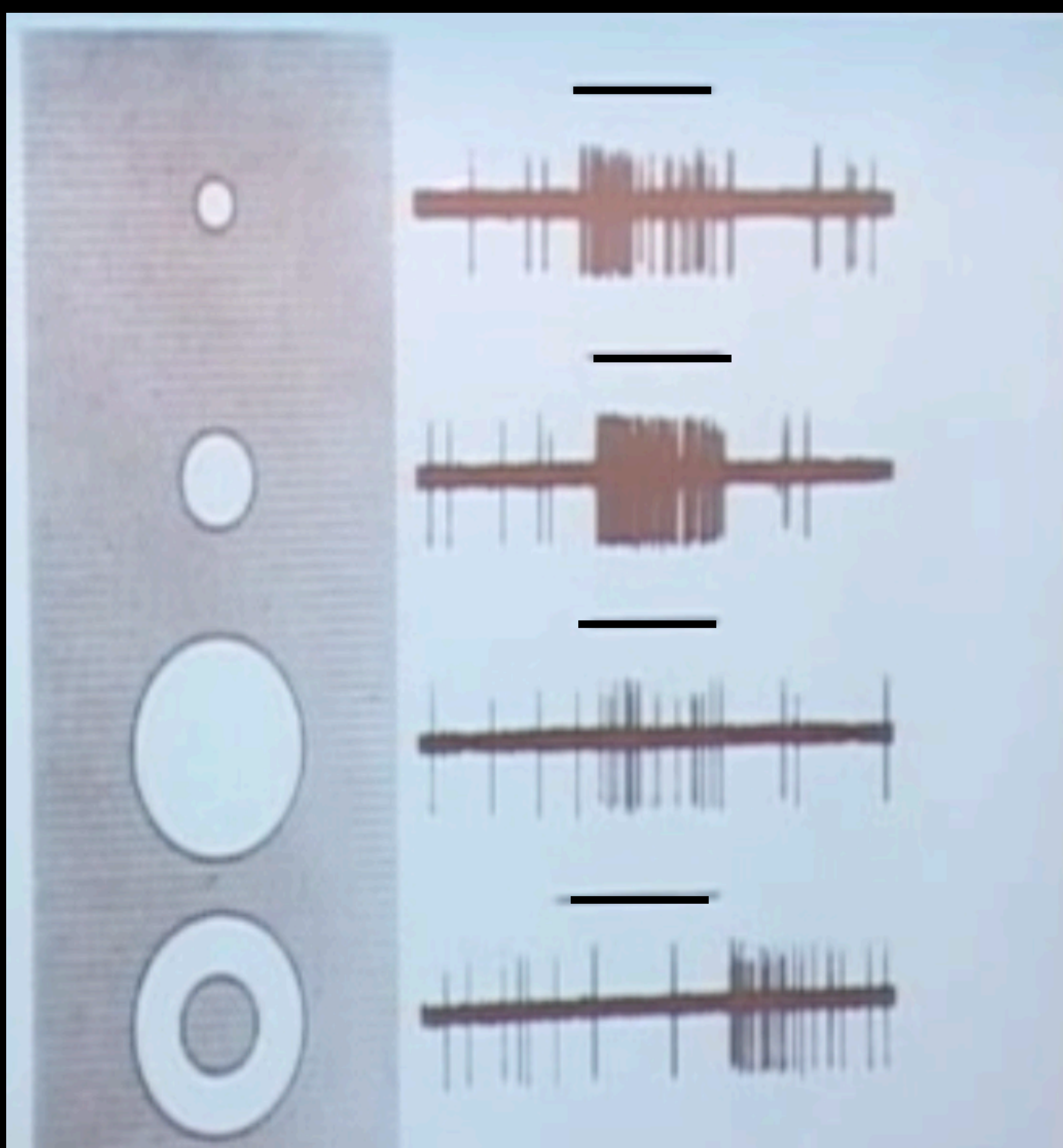
At the center of the macula is the fovea. Maximum visual acuity occurs at the fovea. Only cones are found at the fovea.

The signals from the photoreceptors are passed through the bipolar cells to the retinal ganglion cells.

Retinal Ganglion Cells

- The field of psychophysics studies the relationship between physical stimuli and their response.
- In 1953, Kuffler conducted psychophysical experiments to determine the functionality of the retinal ganglion cells.
- By placing an electrode through the sclera (white portion of eye), Kuffler was able to probe the activity of the retinal ganglion cells.
- His findings are shown on the next slide.

Retinal Ganglion Cells



By varying the types of visual stimuli applied, different output signals were produced by the retinal ganglion cells.

The black bar indicated the time interval that the stimuli was displayed on a screen.

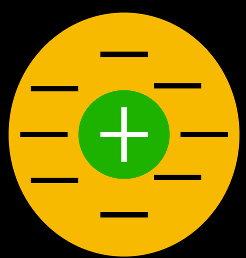
Without the application of any stimuli, random firings were recorded. By introducing a disk shaped visual stimuli and increasing it's size, the firings increased - up to a point.

As the disk grew even larger, the firings began to decrease, indicating that there was a "right sized" disk stimuli that maximized the firing rate.

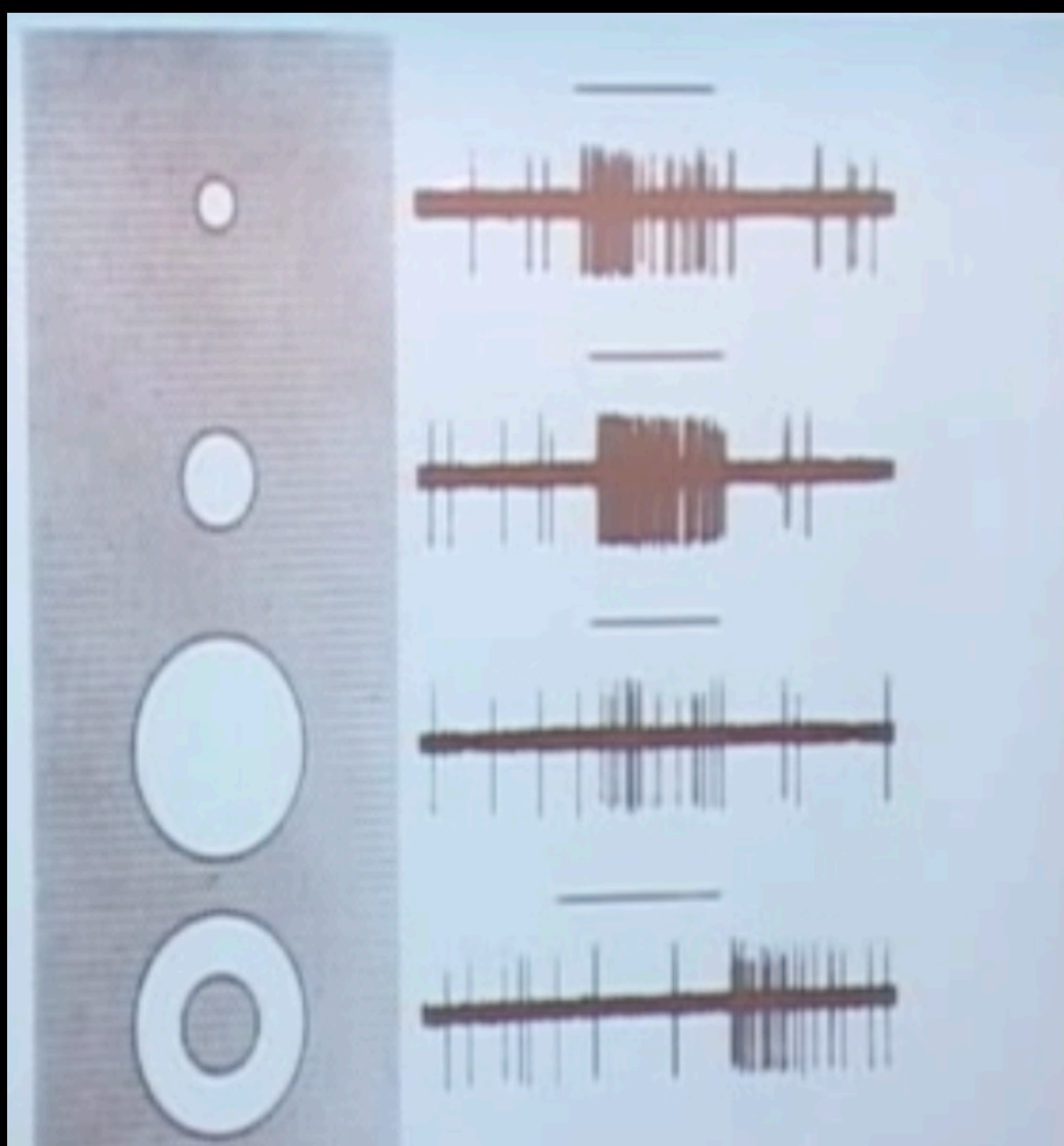
By blocking out the center of the stimuli (donut stimuli), the firings decreased.

These experiments demonstrated the presence of "surround" spatial regions in the receptive field that were responsible for decreasing the firing rates of the retinal ganglion cells.

Retinal Ganglion Cells

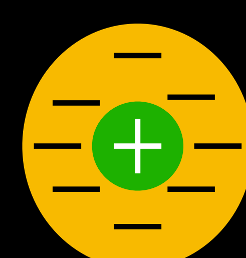


Receptive Field for an On Center Ganglion Cell
+ region encourages firing
- region discourages firing



When the light stimuli fell on only the + region, increased firings occurred.

When the light stimuli matched the + region, maximal firing occurs.



When the light stimuli overlapped the - region, decreased firings occurred.

Retinal Ganglion Cells

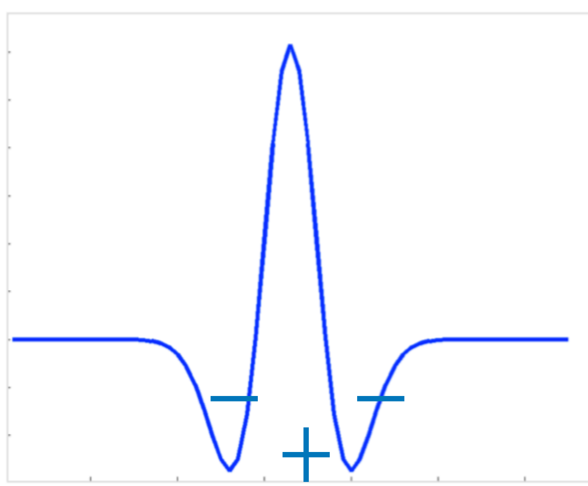
- These experiments identified both on center and off center retinal ganglion cells.
- Off center ganglion cells behaved in an opposite fashion.
- i.e. Off center ganglion cells experienced increased firings with the donut stimuli.
- i.e. Off center ganglion cells fired when the light in the receptive field was turned off.
- It was also determined that retinal ganglion cells closer to fovea had smaller receptive fields than those further away.

Retinal Ganglion Cells

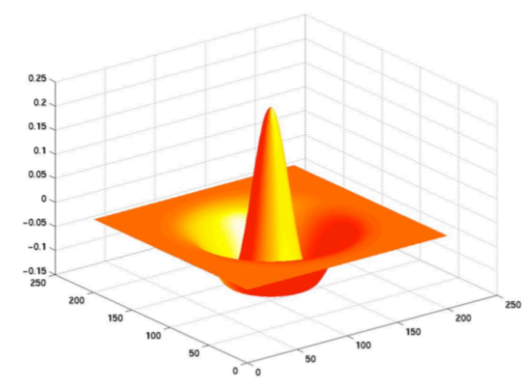
Robert Collins
CSE486

Second Derivative of a Gaussian

$$g''(x) = \left(\frac{x^2}{\sigma^4} - \frac{1}{\sigma^2}\right) e^{-\frac{x^2}{2\sigma^2}}$$



2D analog



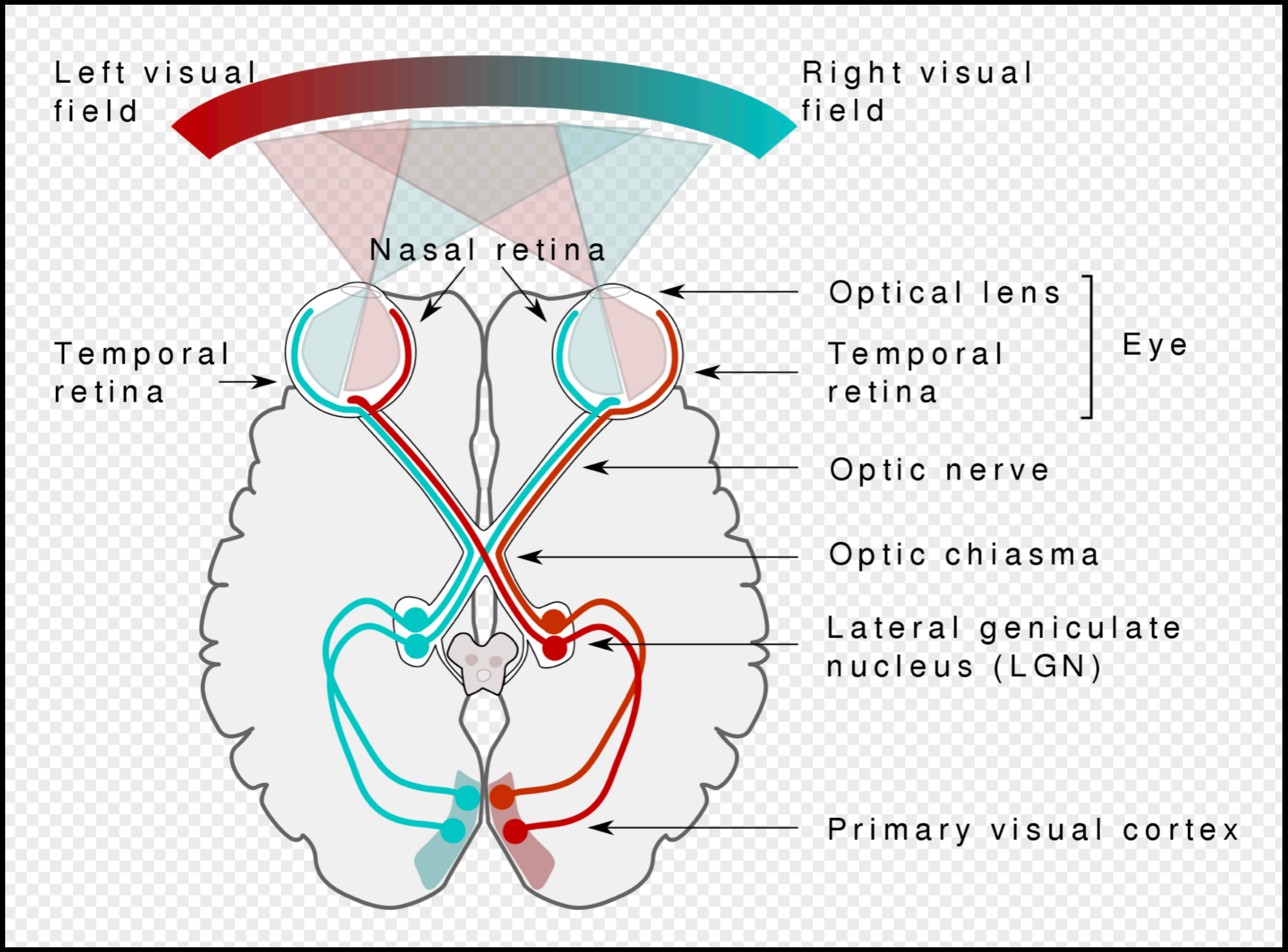
LoG "Mexican Hat"

On center retinal ganglion cells possess characteristics similar to a circularly symmetric laplacian of gaussian filter ~ difference of Gaussian filter.

Retinal Ganglion Cells

- Based on psychophysical experiments, it became clear that diffuse light was a bad stimulus for the human visual system.
- Specifically, it was observed that the firings became nearly non-existent, when a diffuse disk occupied the entire screen.
- i.e. Retinal ganglion cells did not care about the brightness of the light, as much as they cared about the differences in brightness between the center and the surround (=contrast).
- This result provided further evidence that retinal ganglion cells possessed negative inhibitory regions, encouraging contrast detection over brightness detection.
- In general, retinal ganglion cells enable humans to a read newspaper (printed in black ink on white paper) in both bright sunlight, and, a very low lit room.

LGN



The optic nerve was attached to the retinal ganglion cells, and passed information from both eyes into the left and right LGN structures.

Each LGN structure had 6 distinct cell layers.

Each cell layer was associated with information from a specific eye and side of the retina. (i.e. The left LGN received input from the left side of both retinas.)

The cell layers in the LGN were comprised of cells similar to the retinal ganglion cells.

Only now, some cells were more / less sensitive to color and some cells were more / less sensitive to movement.

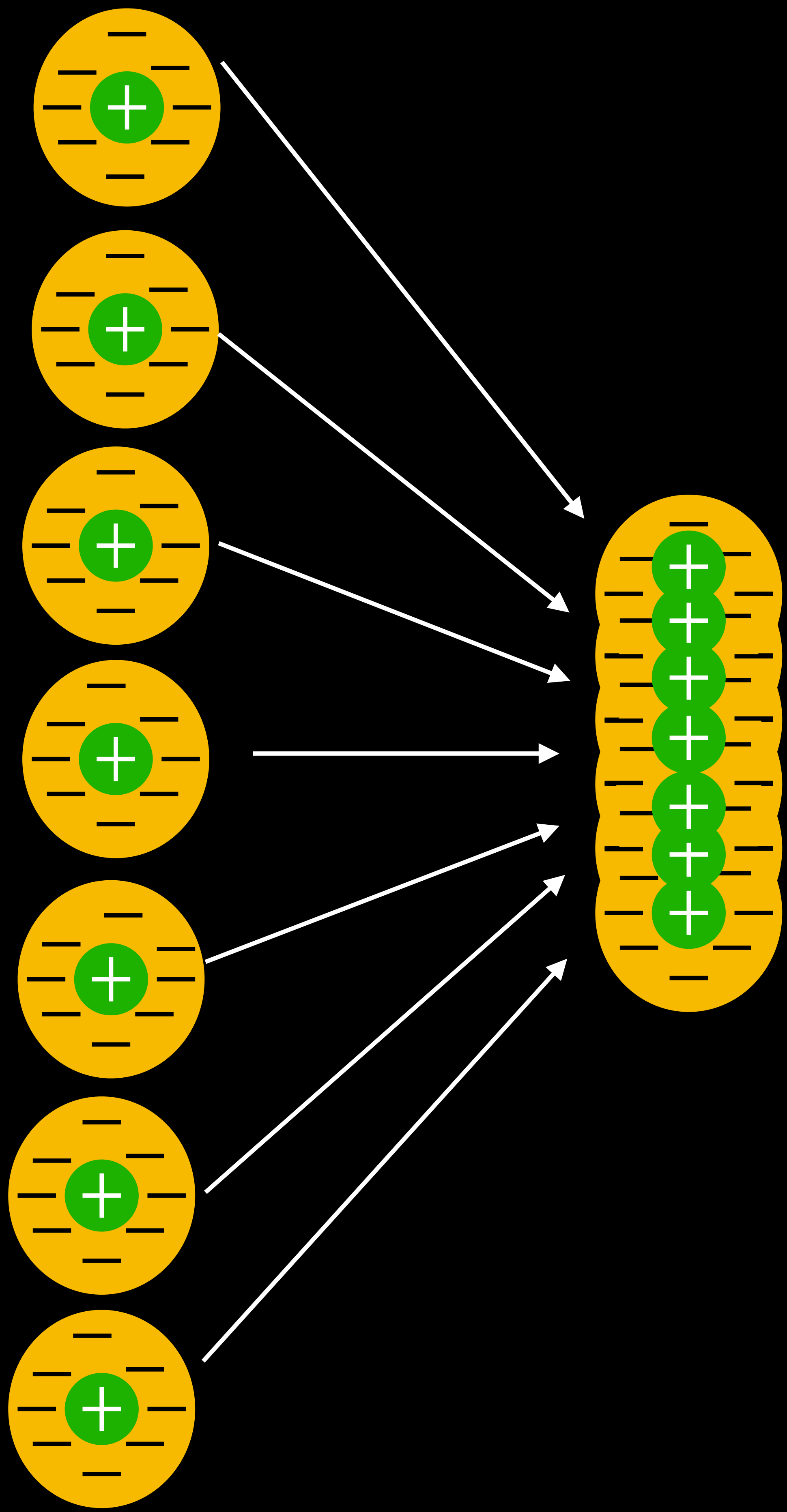
Primary Visual Cortex (V1)

- The primary visual cortex was the first area in the human visual system where the cells received inputs from both eyes.
- Hubel and Wiesel conducted psychophysic experiments to determine the functionality of V1.
- Initially, they applied stimulus similar to the stimulus used in the Kuffler experiments.
- However, no output signal was generated.
- By pure accident, Hubel and Wiesel then discovered that the cells in V1 were orientation sensitive cells.

Primary Visual Cortex (V1)

Overall, V1 consisted of the following cells:

- Oriented simple cells
- Oriented complex cells
- Oriented end stopped cells

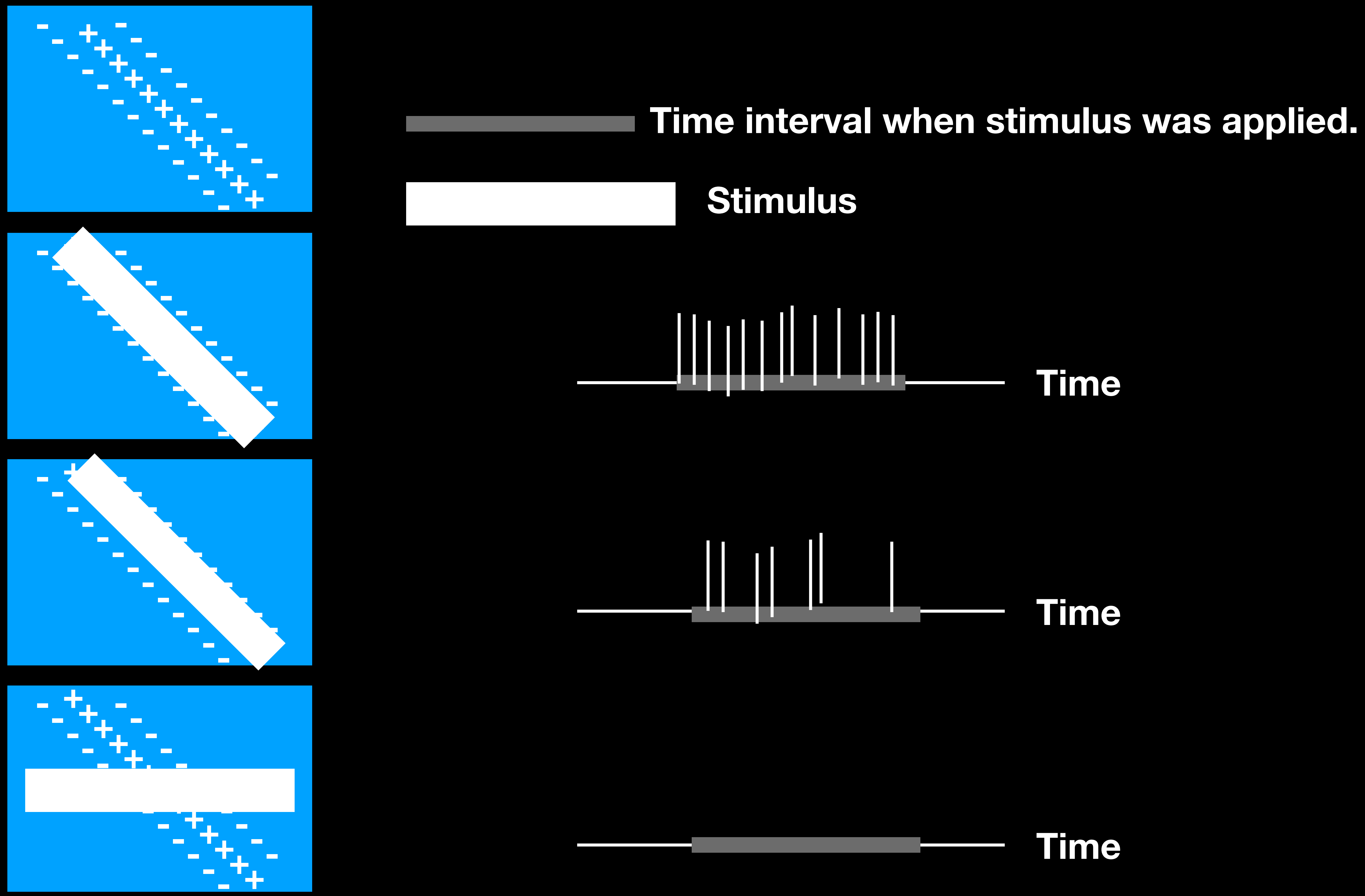


Primary Visual Cortex (V1)

Oriented simple cells were able to detect oriented lines.

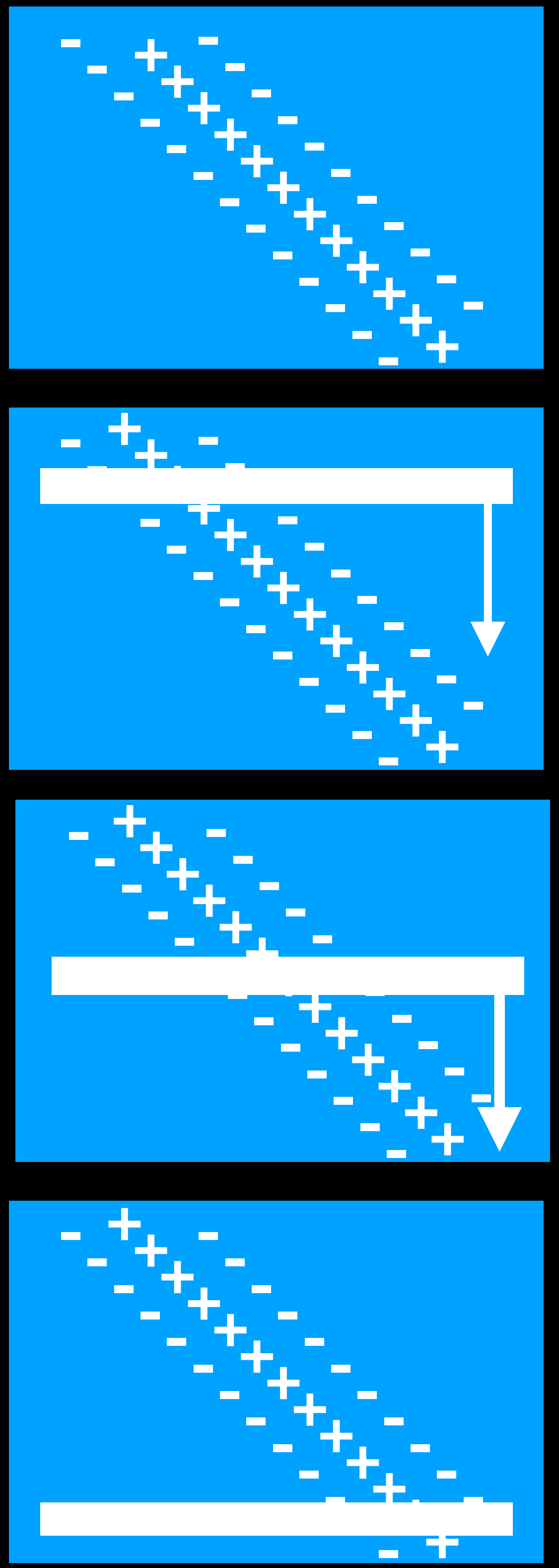
The receptive field for a simple cell attributed to a vertical line detector is shown to the left.

Example (Simple Cell)



The firing rate for various stimulus applied to an oriented cell is shown above.

Example (Complex Cell)



Time interval when stimulus was applied.

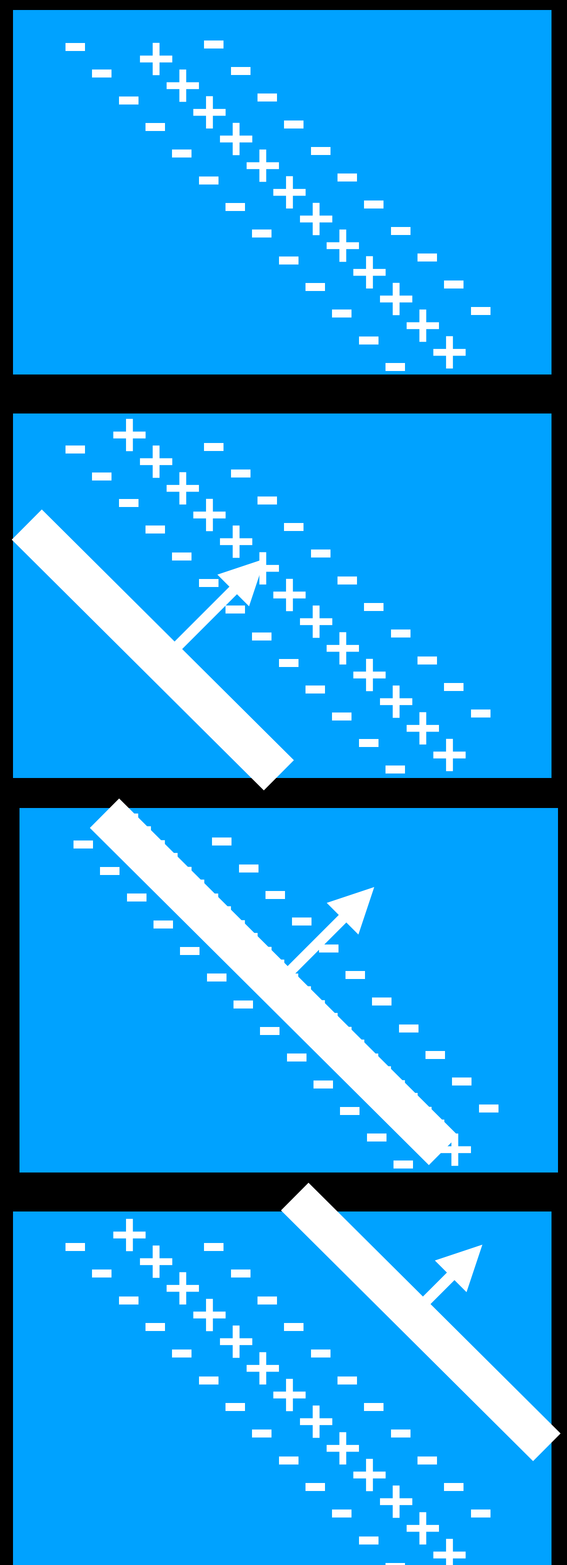
t0

t1

t2

For complex cells, the firing rate only increased when the stimulus moved across the receptive field in the appropriate orientation.

Example (Complex Cell)

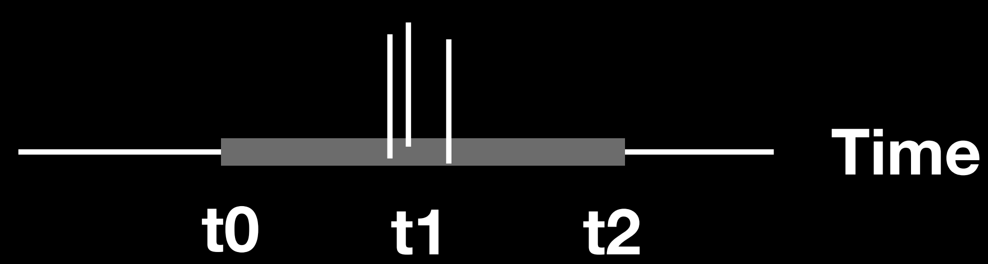


Time interval when stimulus was applied.

t0

t1

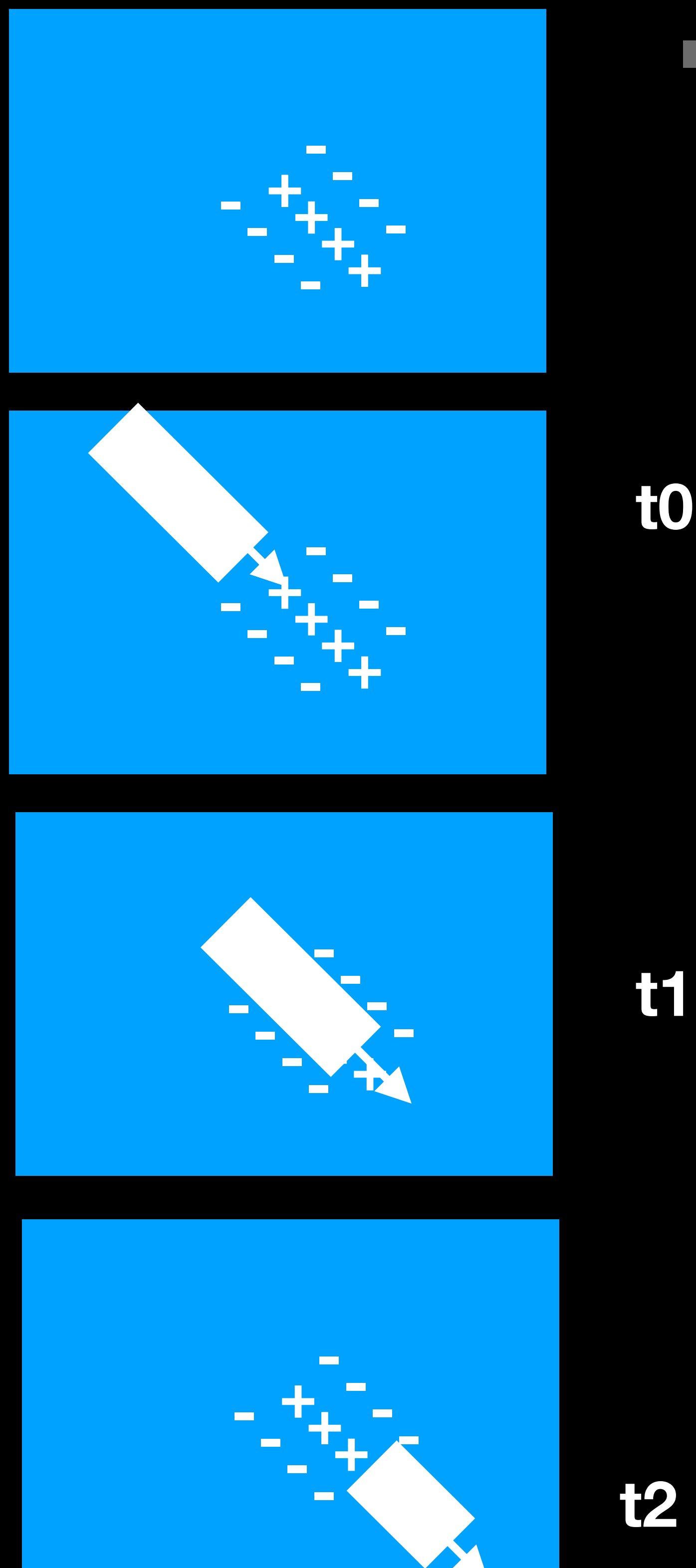
t2

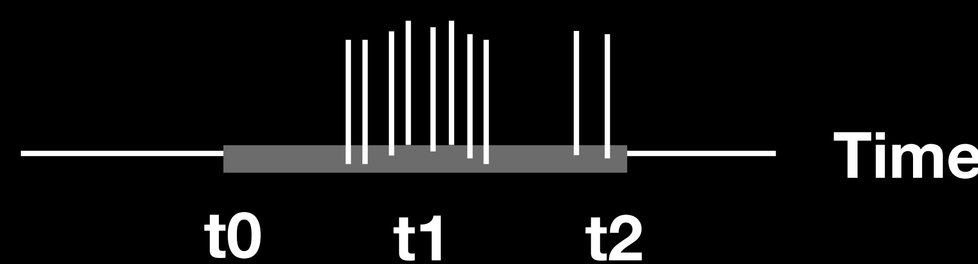


For complex cells, the firing rate only increased when the stimulus moved across the receptive field in the appropriate orientation.

Example (End Stopped Cell)

Time interval when stimulus was applied.



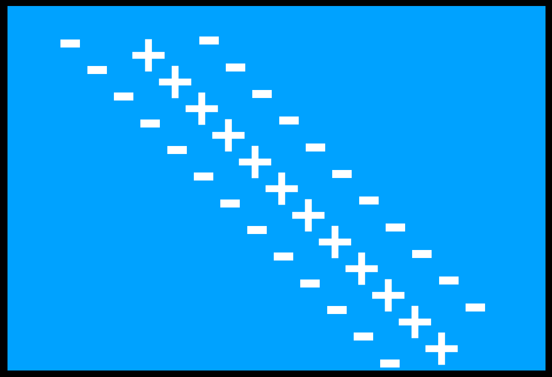


End stopped cells responded best to moving lines of a specific length.

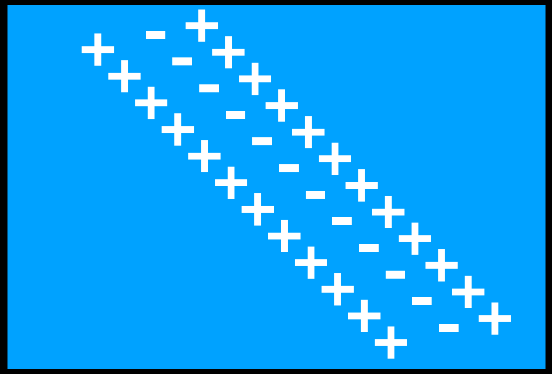
Primary Visual Cortex (V1)

- Simple cells - fires without stimulus motion
- Complex cells - fires best, with stimulus motion
- End stopped cells - fires best, with stimulus motion

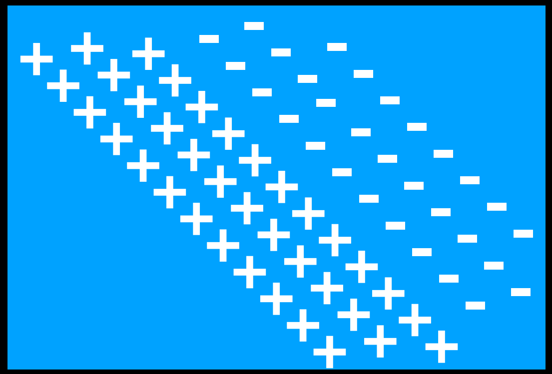
Initial V1 Discoveries



Oriented On Center Line Detector



Oriented Off Center Line Detector



Oriented Edge Detector

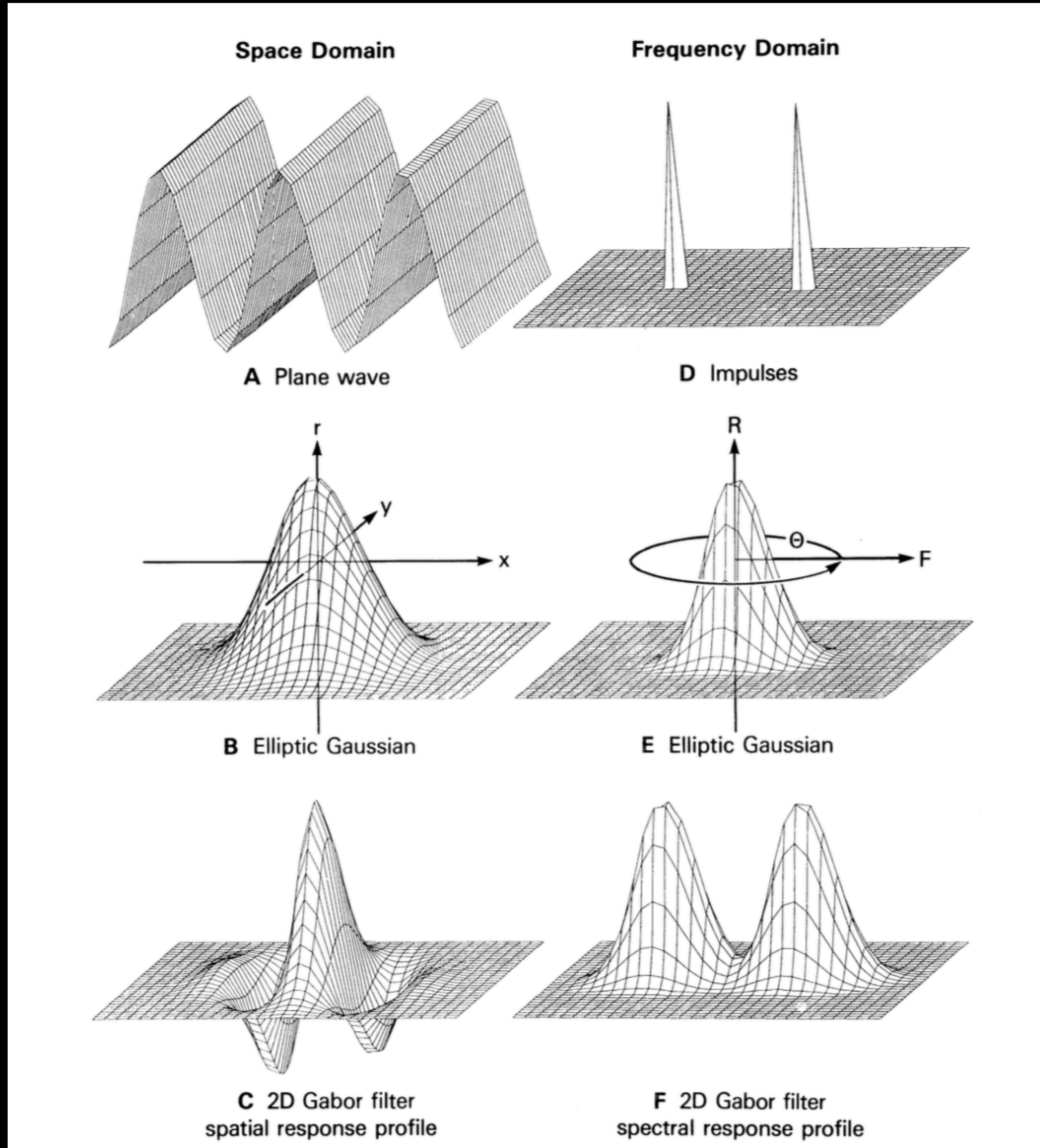
Gabor Filters

- It was hypothesized that 2D Gabor filters were a good model for the cells in the V1 cortex.
- In 1987, this hypothesis was tested and verified.
- Mathematically, Gabor filters consisted of oriented sinusoids modulated with gaussian envelopes.

$$g(x,\ y) = s(x,\ y)w(x,\ y)$$

$$s(x,\ y) = \mathrm{e}^{j(2\pi(u_0x+v_0y)+\phi)}$$

$$w(x,\ y) = K\mathrm{e}^{-\pi(a^2(x-x_0)^2+b^2(y-y_0)^2)}$$



Taken from the Jones paper in 1987, the above figure shows the sinusoidal signal (shown as "plane wave") and the modulating gaussian signal (shown as "elliptic gaussian"). When multiplied together, the 2D gabor filter is produced.

Summary:
Two Filters For Vision

- The retinal ganglion cell were shown to be modeled by circularly symmetric difference of Gaussian filters.

- These cells excelled at detecting contrast.

- The simple and complex cells (found in V1) were shown to be modeled by 2D Gabor filters.

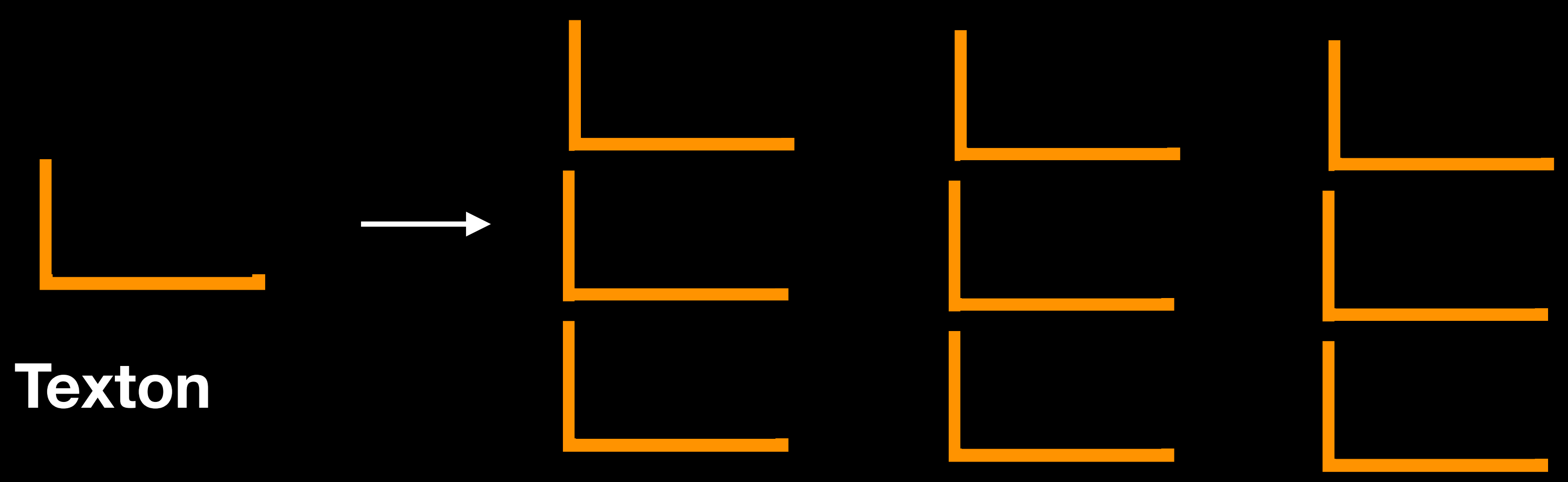
- These cells excelled at detecting oriented structures.

Texture

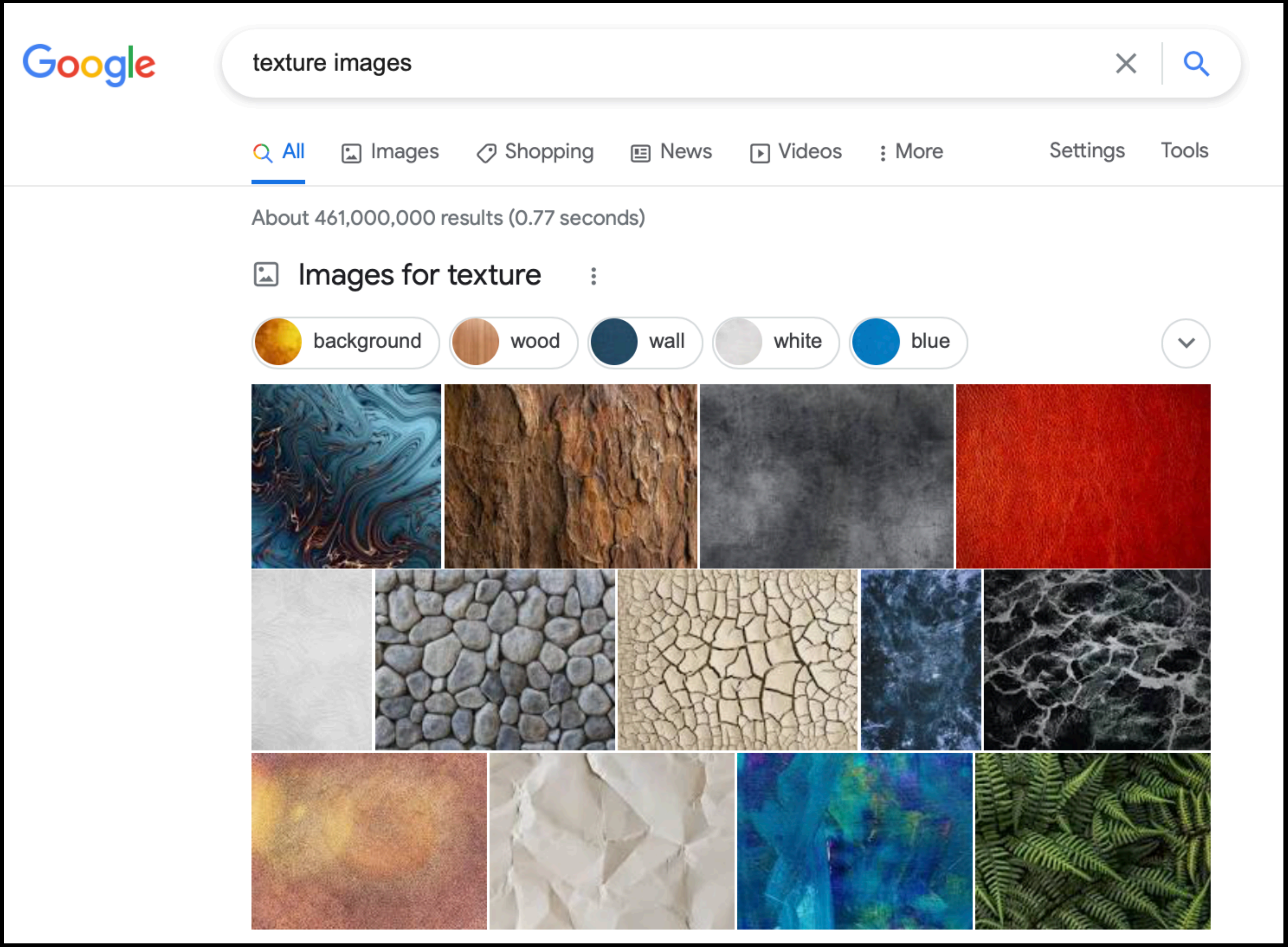
- Texture is another important aspect of vision.
- Using psychophysical experiments, it was determined that texture detection occurred during pre-attentive vision.
- Note: The human visual system operates in 2 distinct modes: pre-attentive vision and attentive vision. Pre-attention vision occurs in the first 100-200 msec of vision processing, and is associated with “pop out” vision.
- Important contributions to the theory of human visual perception for texture, were made by Julesz.
- Julesz viewed textures as the histograms that resulted from images processed by filter banks (= histogram of filtered response images).

Texture

- It was believed that novel textures could be synthesized using only these gradient based histogram distributions.
- Julesz was also responsible for introducing the idea of textons - “the units of pre-attentive texture perception”
- Textons were fundamental texture elements consisting of elongated blobs, terminators (ends of line segments), corners and crossing line segments.



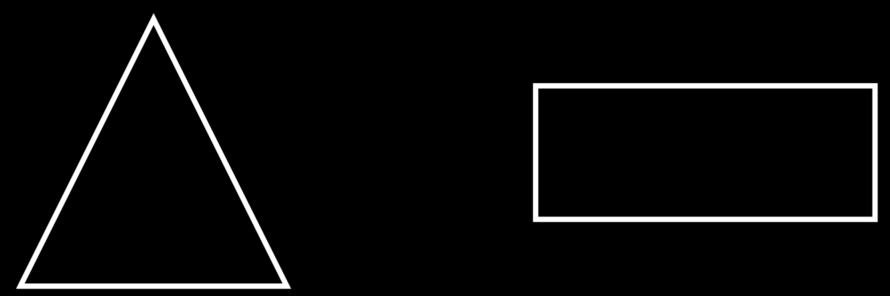
Texture



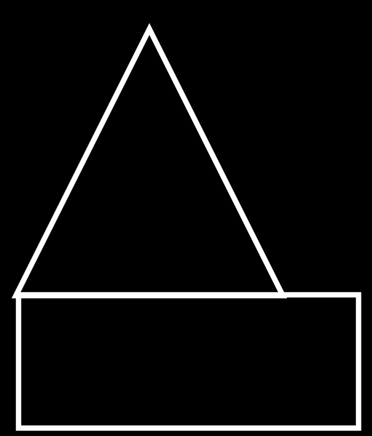
In general, textures can be best understood visually, by seeing examples.

Polygons

Points and lines can be used to construct simple polygons, such as those shown below.

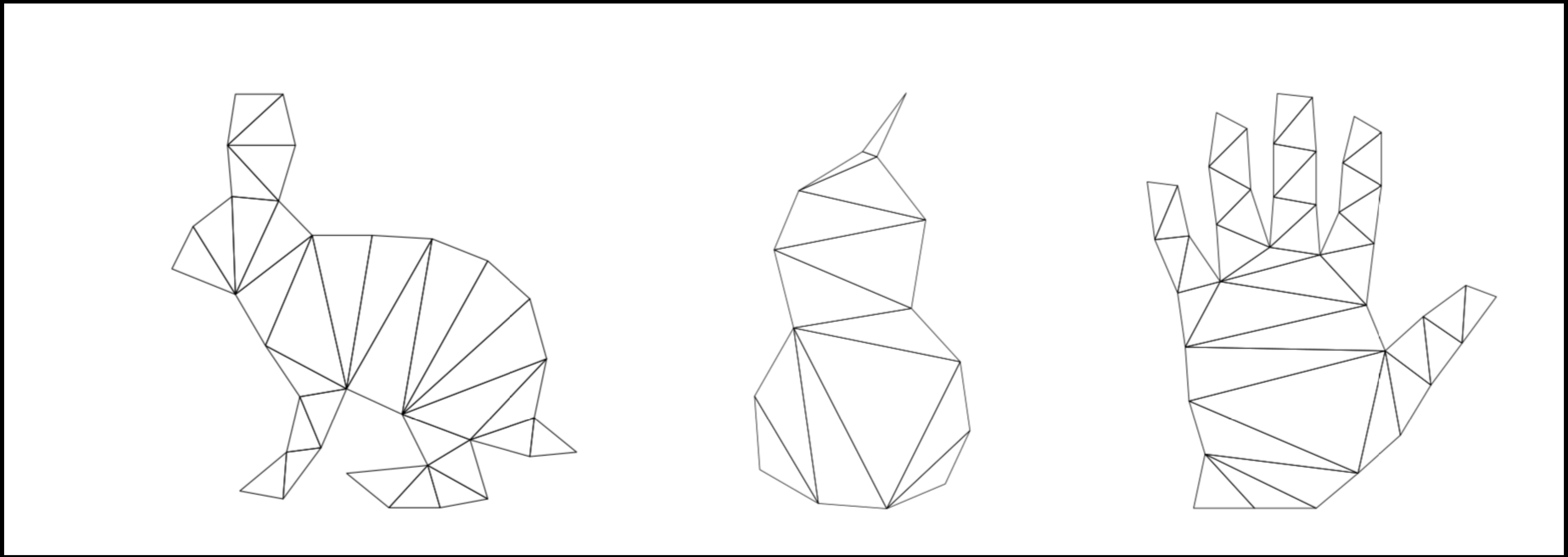


By combining these polygons, simple shapes associated with different structures can be produced.



Structures

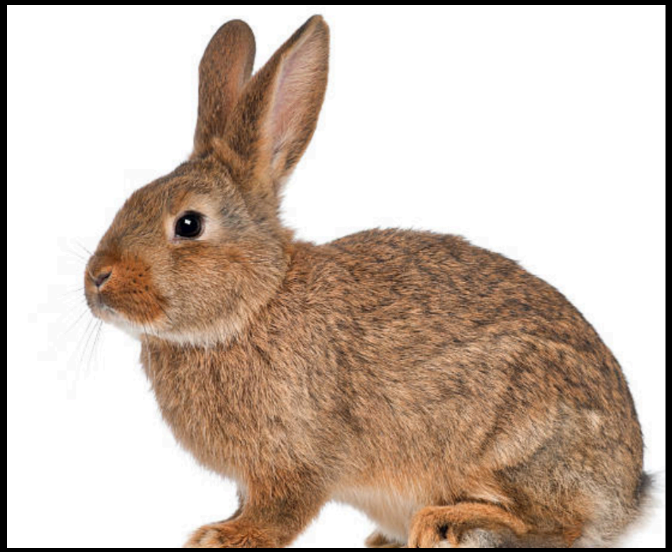
In the field of computer graphics, numerous triangular polygons are combined, to produce structures.

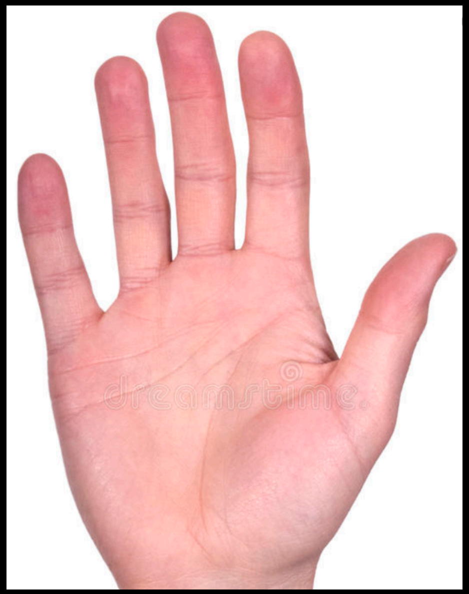


Rabbit **Pear** **Hand**

Image taken from P. Felzenszwalb's PhD dissertation.

With Texture Added





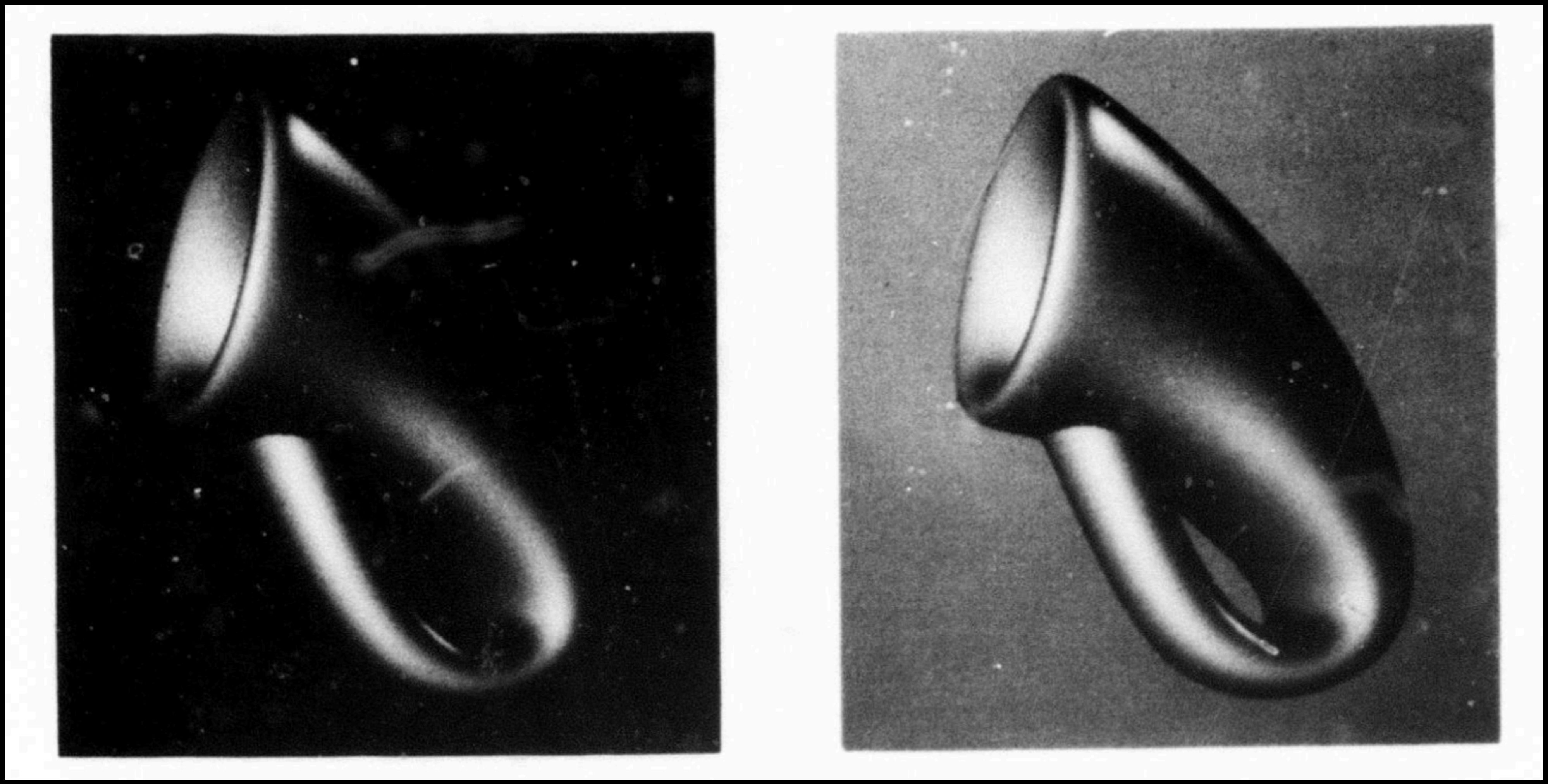
By using smaller and smaller triangular polygons and overlaying texture onto the polygons, 2D images can be created.

What’s Missing?

- Images are projections from the 3D world, onto the 2D plane.
- Although the image plane is 2D, we perceive depth.
- This is because our two eyes function as a binocular depth detecting, visual system.
- In practice, one of the ways that depth is conveyed in still images, is through shadows.

Hard Shadows Cast By the Objects

Shadows On the Objects Themselves



Some of the seminal work in the latter, initially involved the creation of shadows on simple, curved surfaces.

Image taken from Ed Catumull’s PhD dissertation.

Summary

- For any simulated image, we need to consider all of the previously discussed items, that comprise the human vision system.

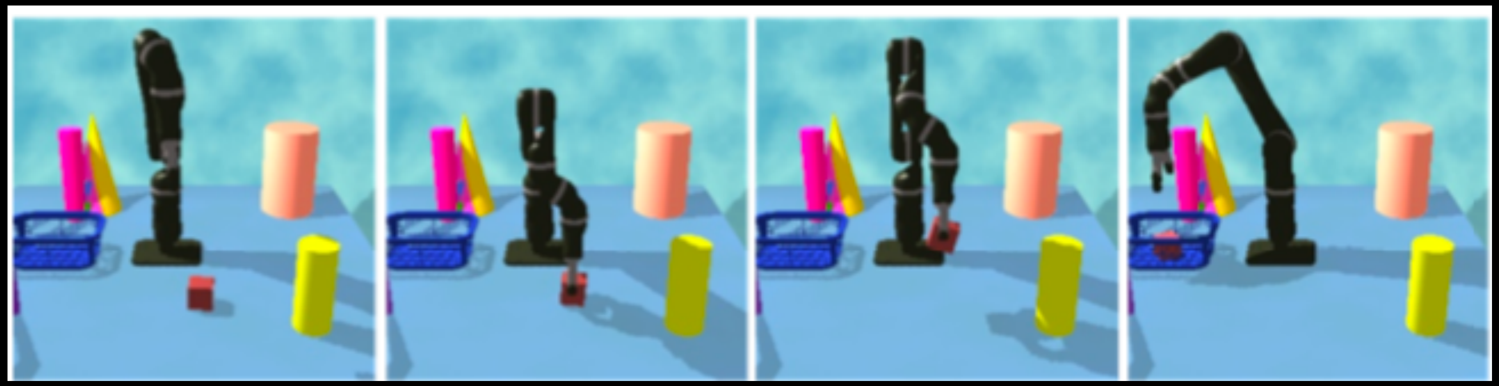
- Exposure (iris and pupil), Focus (cornea and lens), Light receptors / color (rods and cones), Contrast (ganglion cells), Edges (V1), Texture (pre-attentive vision), Depth (stereo vision) and shadows.

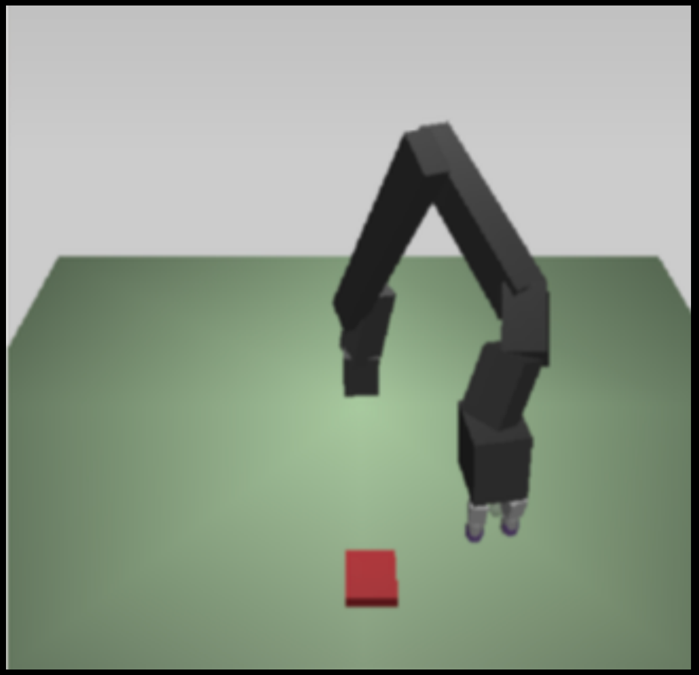
- All of the above mentioned items will impact vision based policy learning, during the training phase.

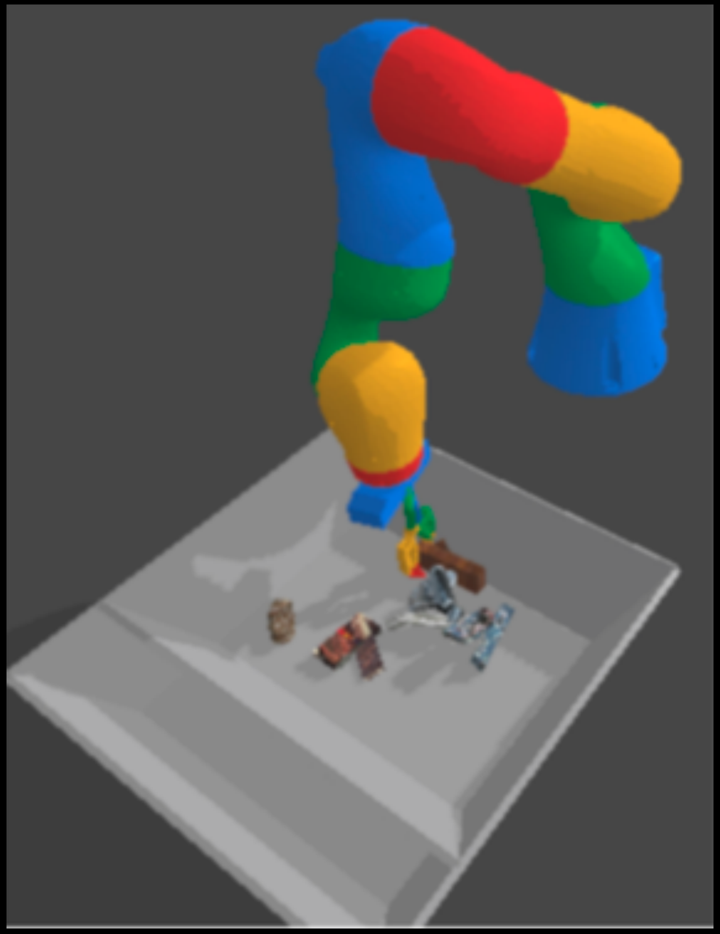
- In addition, if the training data results from real images captured by physical cameras, how will the variations in color, exposure, etc. from different cameras, affect the performance of the deployed system?

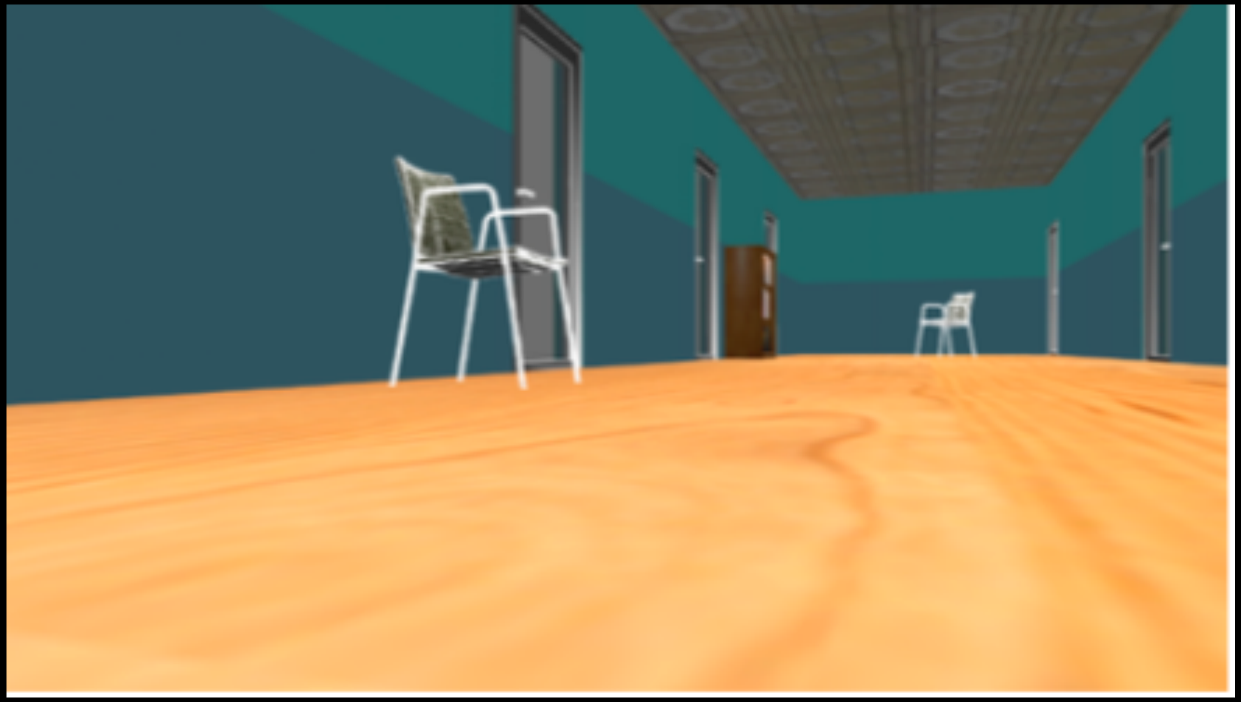
Prelude to Part2

- Having proceeded through the vision / perception journey, we are now ready to examine the application of simulations to reinforcement learning.
- As a teaser, the next slide shows some of the simulated images currently used by different RL algorithms for training and test.









References

- Recounting the Impact of Hubel and Wiesel, Wurtz, Journal of Physiology, June 2009
- Discharge Patterns and Functional Organization of Mammalian Retina, Kuffler, Journal of Neurophysiology, 1953
- Receptive Fields of Single Neurons in the Cat's Striate Cortex, Hubel and Wiesel, Journal of Physiology, October 1959
- An Evaluation of the Two Dimensional Gabor Filter Model of Simple Receptive Fields in the Cat Striate Cortex, Jones and Palmer, Journal of Neurophysiology, 1987
- Textons, the Elements of Texture Perception, and Their Interactions, Julesz, Nature 1981
- Representation and Detection of Shapes in Images, Felzenszwalb, PhD Dissertation 2003
- A Subdivision Algorithm for the Computer Display of Curved Surfaces, Catmull, PhD Dissertation, 1974