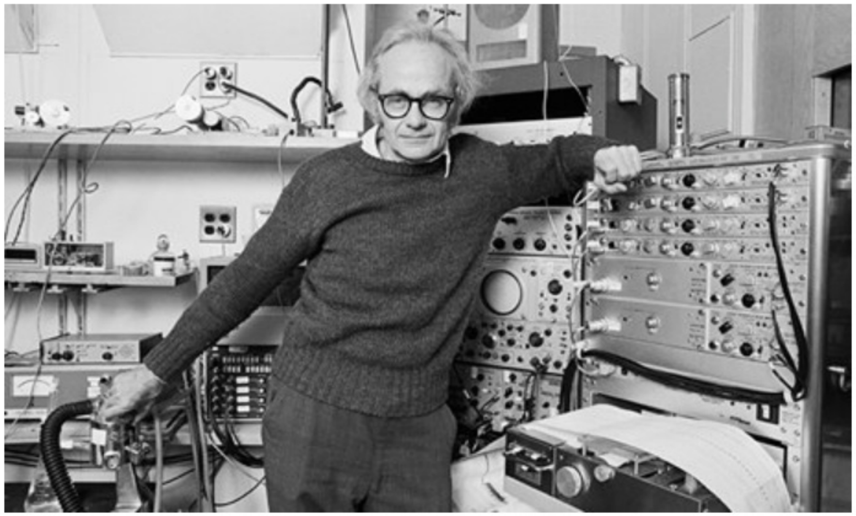
Bio-Inspired Artificial Intelligence

Lecture 5: Bio-Inspired Vision

http://www.informatik.uni-hamburg.de/WTM/



David Hubel in 1981. His work created new insights into the machinery of the brain.
Photograph: Ira Wyman/Sygma/Corbis

The knowledge we have now is really only the beginning of an effort to understand the physiological basis of perception, a story whose next stages are just coming into view; we can see major mountain ranges in the middle distance, but the end is nowhere in sight.

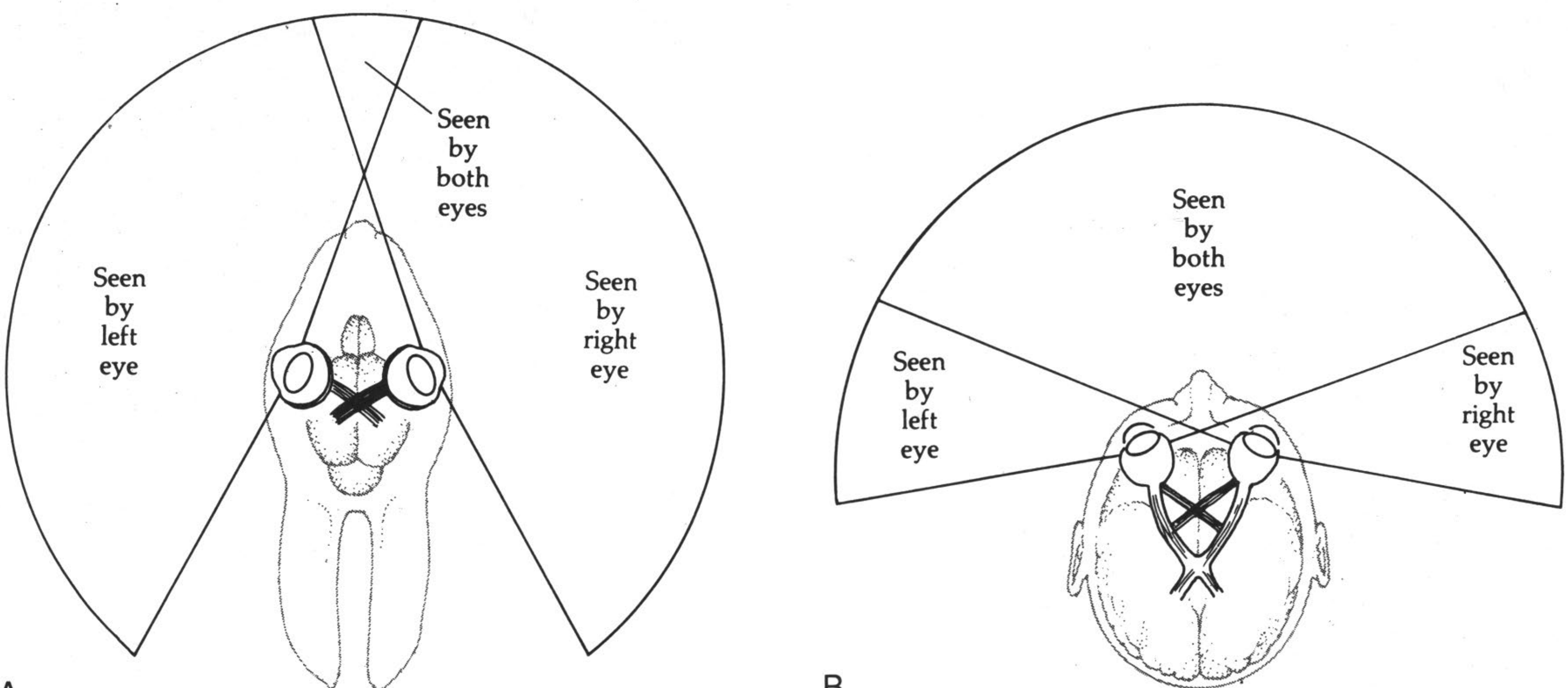
(Hubel 1926-Sept 2013)

Outline

- Biological motivation
- Different Cell Types
- Marr’s Edge Detection
- Computational Models and Simulation
- Summary

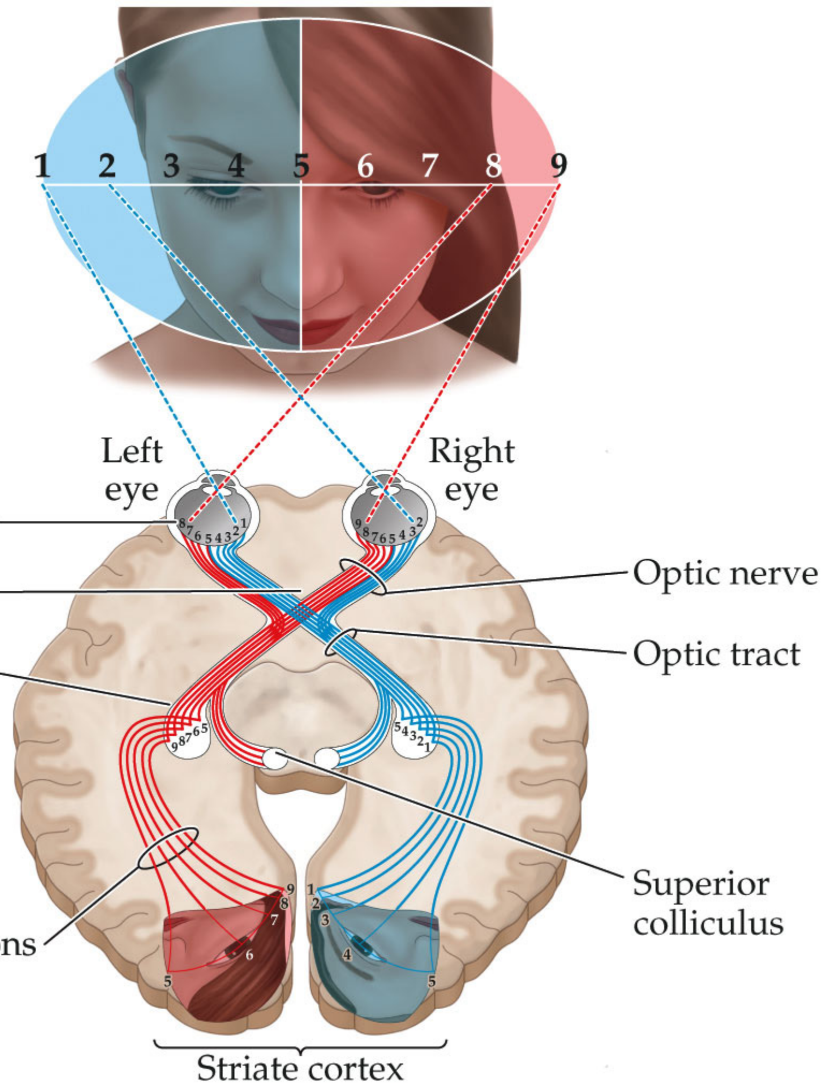
Bio-inspired Vision: Field of View

- Different visual fields of view in nature: why?
- Different from humans, rabbits´ visual system hardly integrates information from both eyes
- More complex neural processing in the human brain



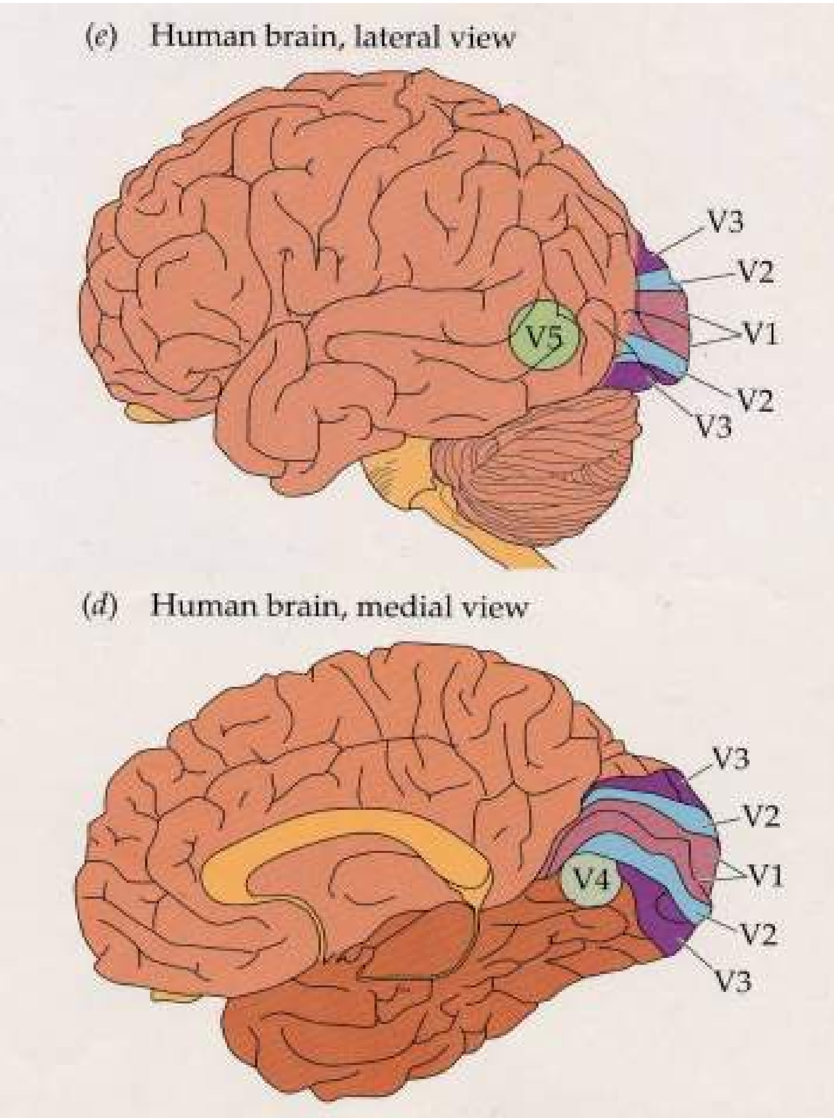
A B

Bioinspired Visual Information Processing



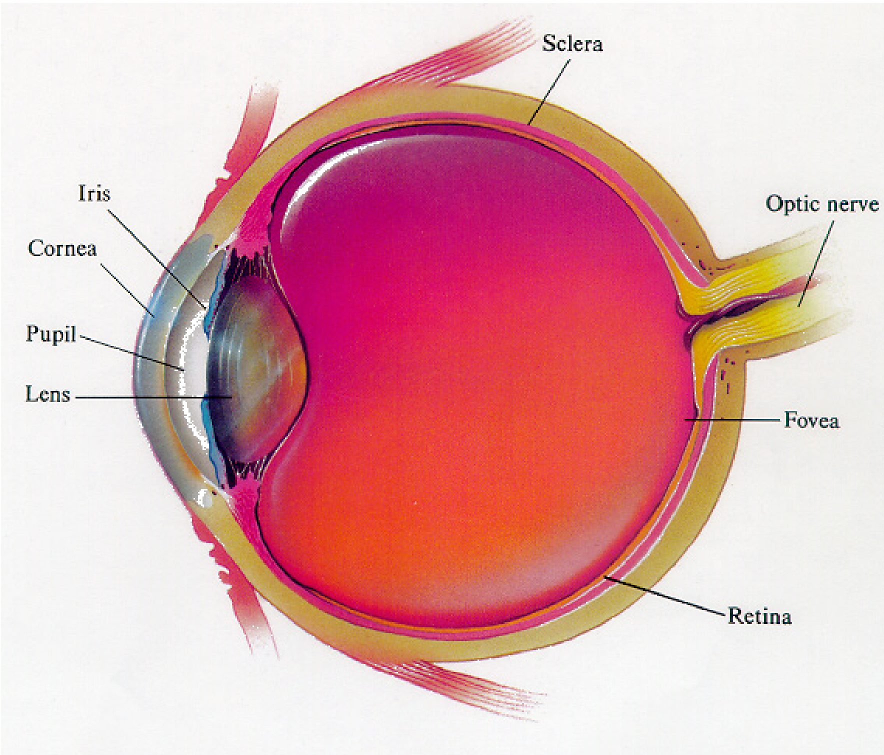
SENSATION & PERCEPTION 2e, Figure 3.14

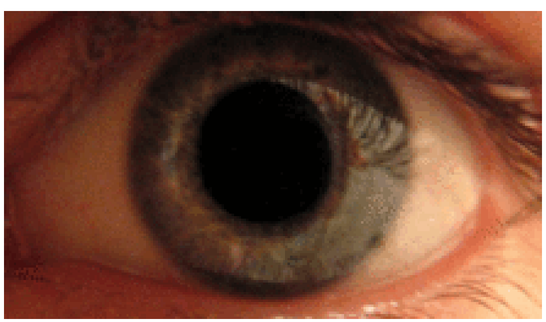
Human Visual Areas



- Visual information processing in occipital lobe (back of brain)
- Different areas for different visual information
- V1: Edge detection (also: striate cortex); start of dorsal / ventral stream
- V2: Responds to more complex patterns, e.g. illusionary contours, parts of figures; Object memory
- V3: Colour, motion
- (V4: Orientation and colour in macaque monkey brain, no human homologue known)
- V5: Motion, eye movements
- (V2-V5 also: extrastriate cortex)

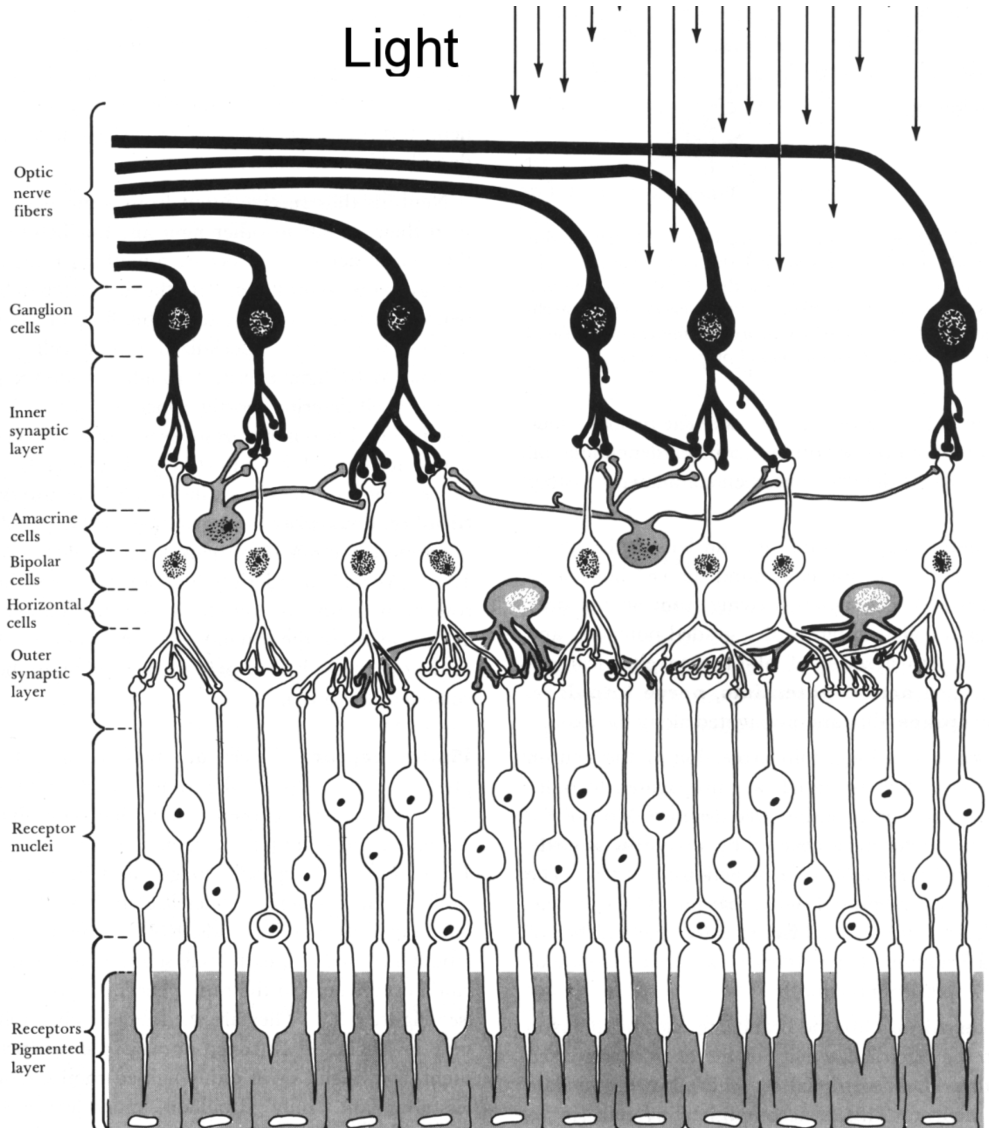
Retinal Processing





- Stimuli from retina project to four subcortical regions
- Lateral Geniculate Nucleus (LGN): provides information to V1
- Hypothalamus: controls circadian cycle (body clock)
- Pretectum: controls pupillary light reflex
- Superior Colliculus: controls eye movements

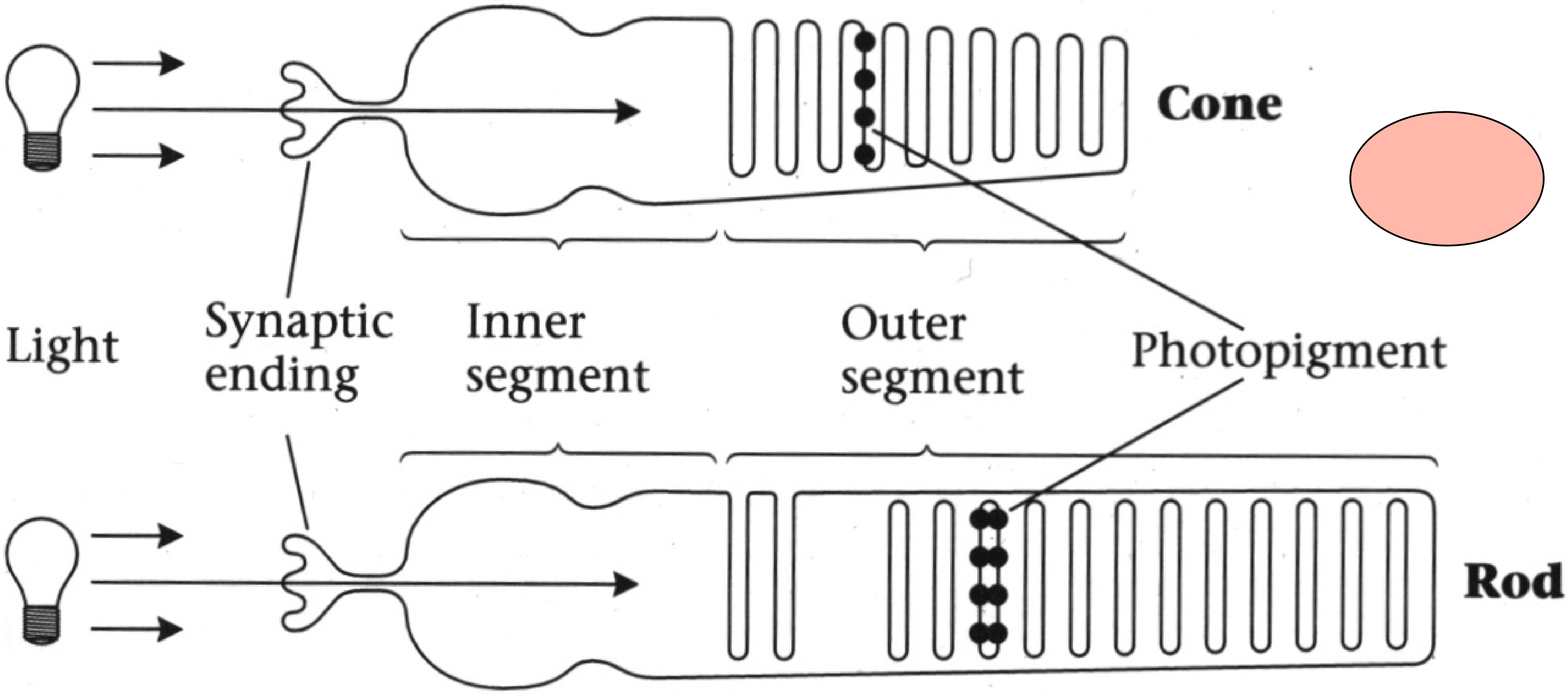
Retina



Light
Optic nerve fibers
Ganglion cells
Inner synaptic layer
Amacrine cells
Bipolar cells
Horizontal cells
Outer synaptic layer
Receptor nuclei
Receptors
Pigmented layer

Retina: Inside the rod and the cone cells

color vision; at fovea; they function best in relatively bright light

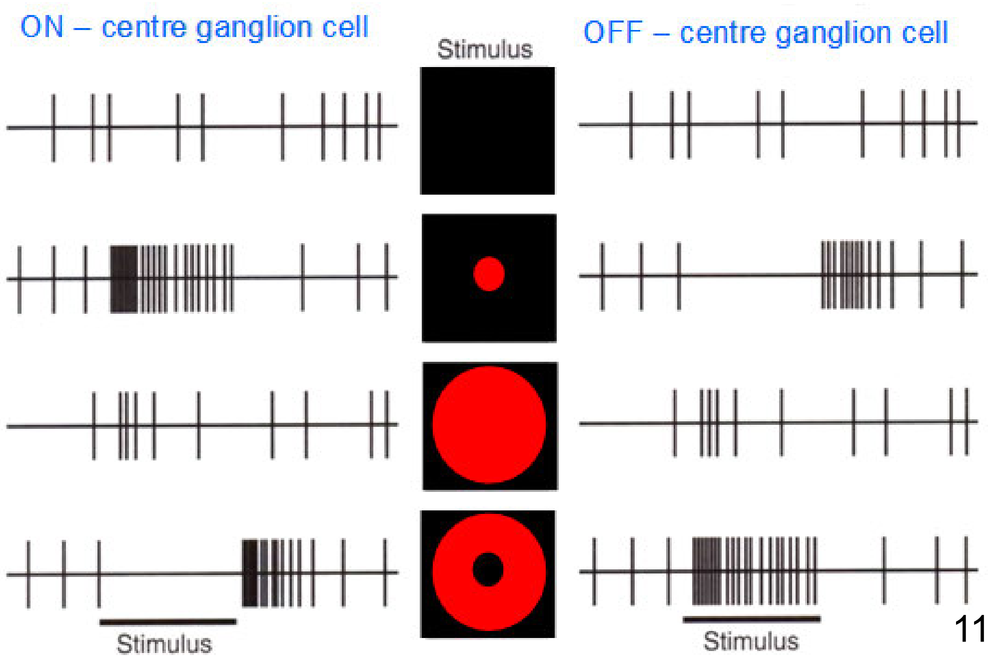


rod cells are almost entirely responsible for night vision; outer edge of retina

Inside the rod and the cone

Receptive Fields

- Region in which the presence of stimulus will alter the firing of that neuron
- Light stimulus evokes action potential (AP) in ON-ganglion cell
- Frequency increases with sensor strength
- Derivation of AP determines retinal area: receptive field
- ON: excitatory influence on stimulus, center of RF
- OFF: inhibitory influence on stimulus, periphery of RF



Visual Area 1 (V1)

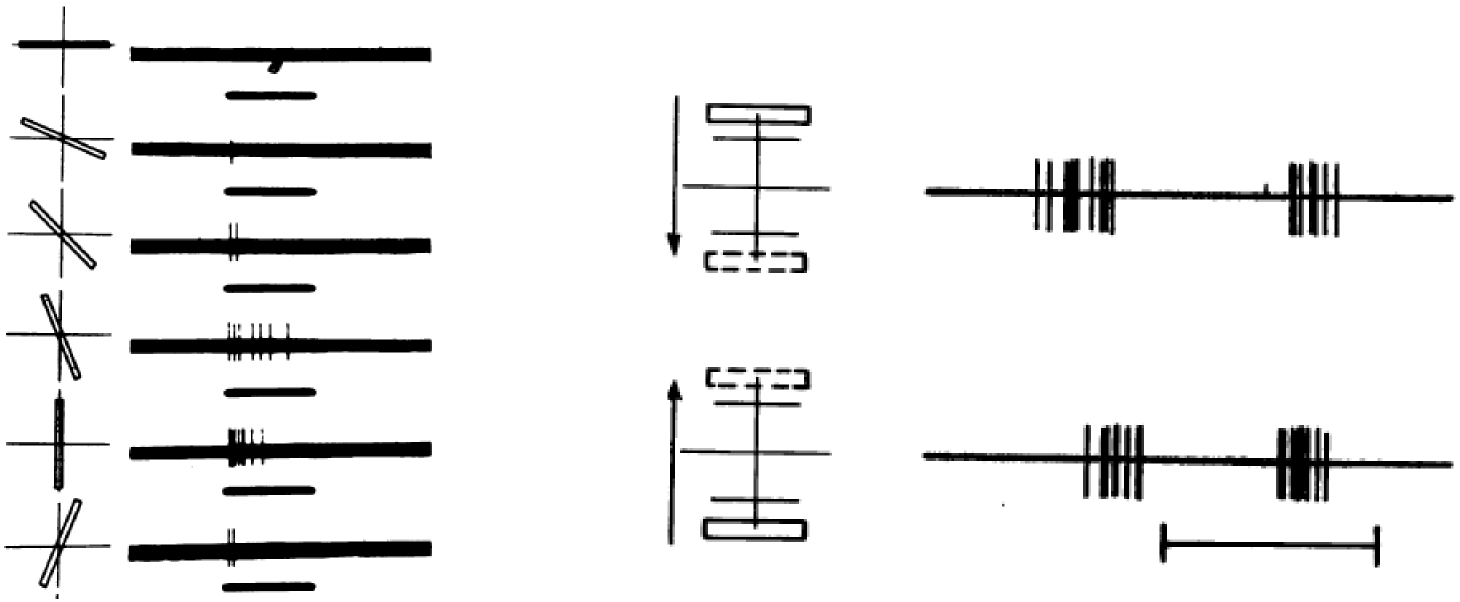
- Exhaustive research of V1 e.g. in cats and macaque monkeys
- Cell Types:
  - ***Simple Cells***: Line/Edge Detector, highly selective
  - ***Complex Cells***: Orientation Detector, implements invariant features e.g. translational invariance
  - Hypercomplex Cells: Angle/Length Detector; can be composed of several complex cells
- Later in this lecture: mechanisms of simple and complex cells built into computational model (e.g. ‘Pooling‘ over simple cells)

Hubel and Wiesel (Nobel Prize research)

- Cells in striate cortex respond best to bars of light rather than to spots of light
  - Some cells prefer bars of light, some prefer bars of dark (simple cells)
  - Some cells respond to both bars of light and dark (complex cells)
- Orientation tuning:
  - Tendency of neurons in striate cortex to respond more to bars of certain orientations and less to others
  - Response rate falls off with angular difference of bar from preferred orientation

Hubel and Wiesel

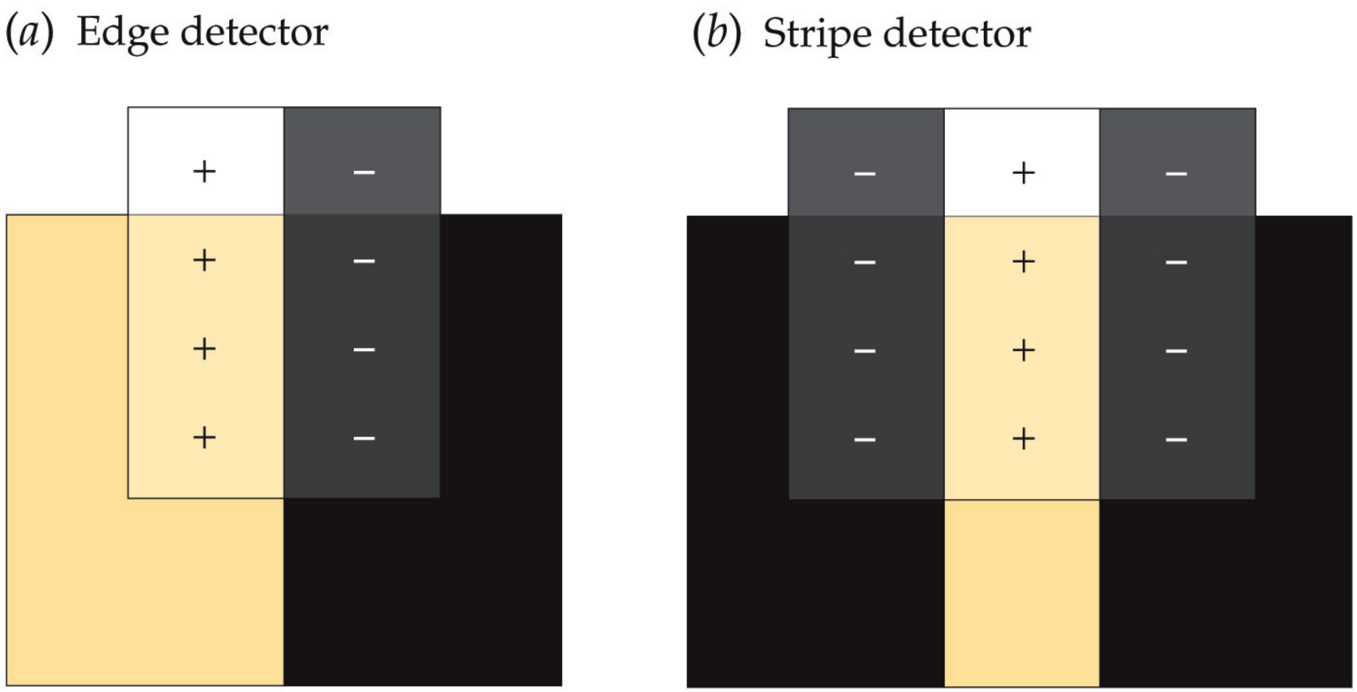
- 1959: Paper about cell tuning in cat striate cortex
- Experiments with bars in different positions
- The following video demonstrates the experiment



http://hubel.med.harvard.edu/papers/HubelWiesel1959JPhysiol.pdf

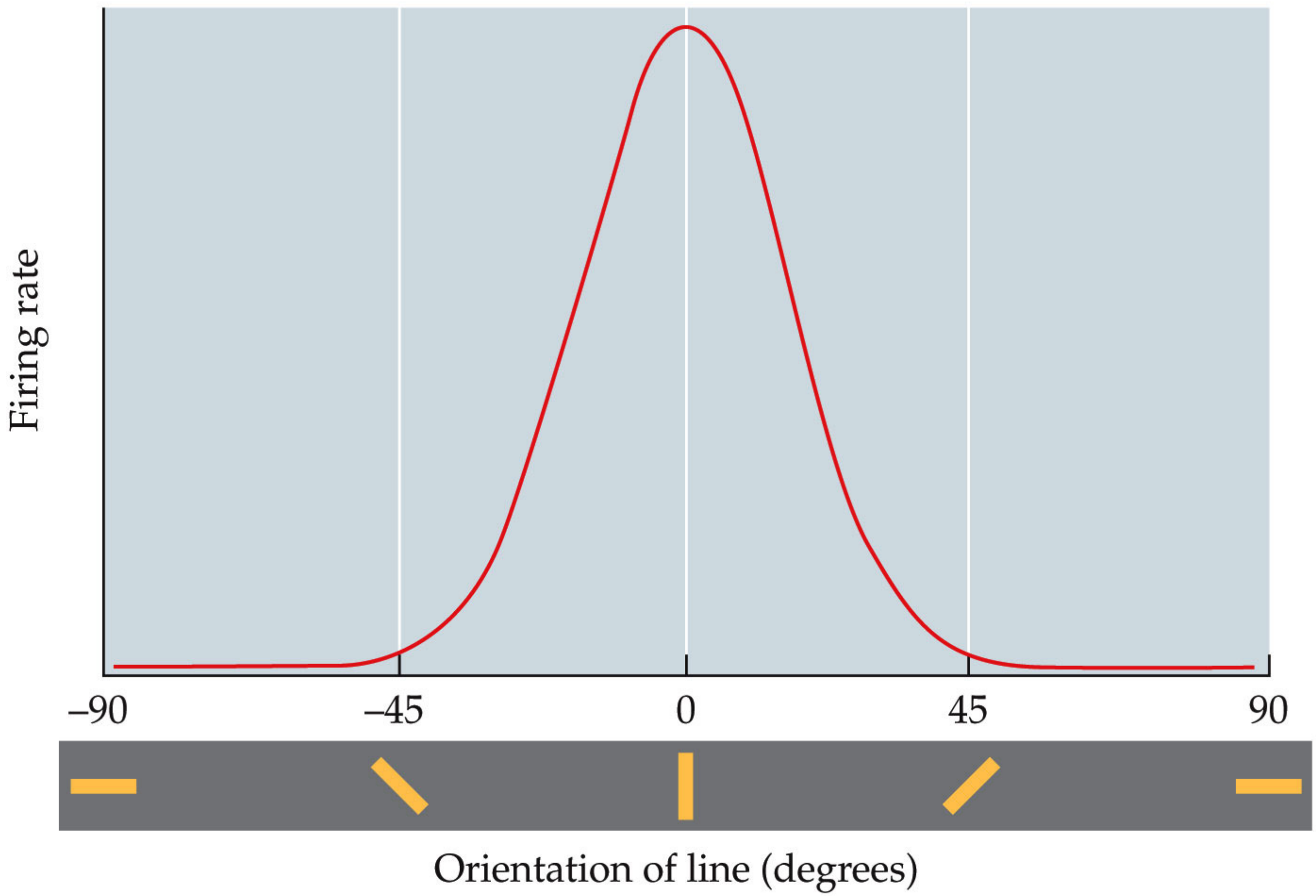
Hubel and Wiesel: Simple Cell Tuning

Simple Cell Processing



SENSATION & PERCEPTION 2e, Figure 3.19

Orientation Detection



SENSATION & PERCEPTION 2e, Figure 3.16

Computational Model by Marr

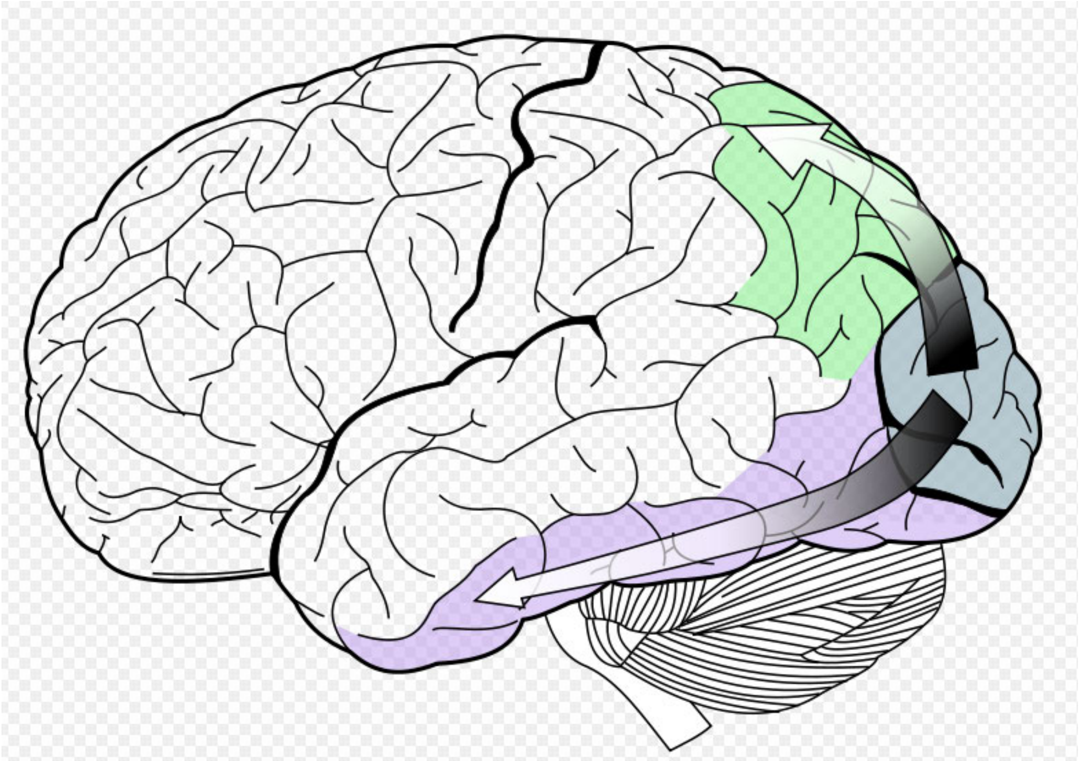
- *Marr 1982*. Vision: A computational investigation into the human representation and ***processing of visual information***
- The computational approach tries to formally define the visual processes necessary to ***represent*** the world
- Computational theorist: light in retina serves as input, visual processes are then formulated as ***algorithms***
- The computational approach emphasizes the importance of ***edge detection***, for early stages of visual perception.
- Zero Crossing (function - -> +) represents the change in intensity (*Mexican-Hat* filter)

Computational Model by Marr

- First 2-D sketch is the representation of the visible surfaces in the view
  - This is built up by the primal sketch along with information derived from motion and the differences of images from both retinas
- Secondly we transform this into an object-centered 3-D sketch
  - This sketch can be constructed mentally by combinations of simple 3-D shapes

Ventral and Dorsal Streams

- Dorsal in green: visual spatial location – Where
- Ventral in purple: object recognition -- What



Higher Visual Processing

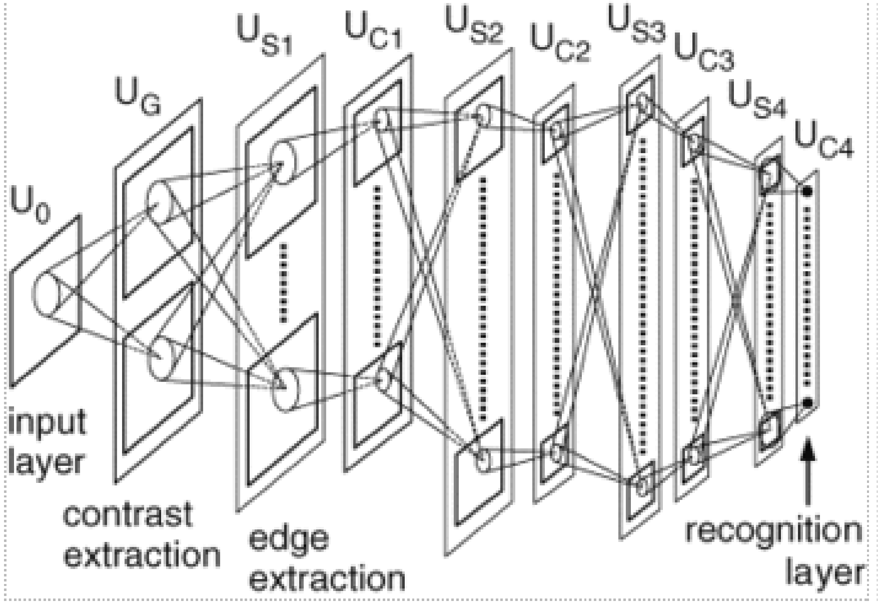
- Dichotomy of object processing in occipital lobe
- Found by Ungerleider and Mishkin in 1982
- Separation underpinned by lesion studies
- ***Ventral*** stream: projects to temporal cortex

  → Neural encoding of shape, colour, texture, … (***what)***
- ***Dorsal*** stream: projects to parietal cortex

  → Neural encoding of objects spatial information ***(where)***
- Sometimes referred to as ***how***-path as dorsal path responsible for sensory-motor transformation (see also affordances)

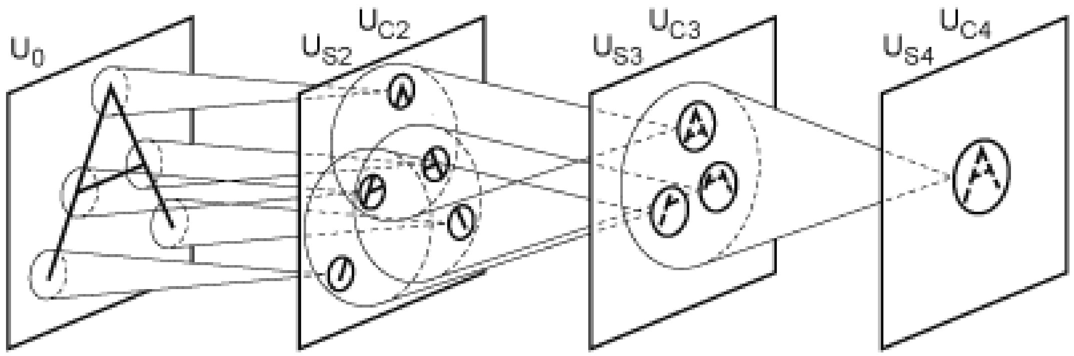
Building Models: Neocognitron (Fukushima)

- Hierarchical multilayer neural network for visual recognition
  - Alternate planes of simple ***S-cells*** (feature extraction) and complex ***C-cells*** (positional errors)
  - Feature extraction in different modules
  - Resembles processing stages in visual cortex



Neocognitron (Fukushima)

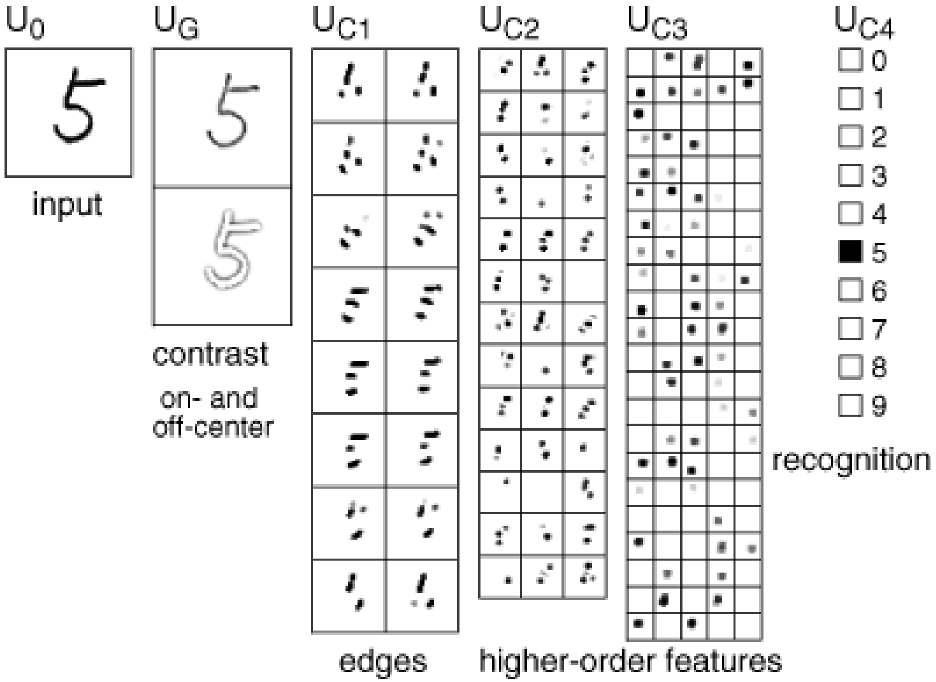
- Example for the recognition of letter ‘A‘
- S-cells trained to particular feature present in receptive field
- C-cells inserted to correct for positional errors: receive responses from S-cells coding for the same feature



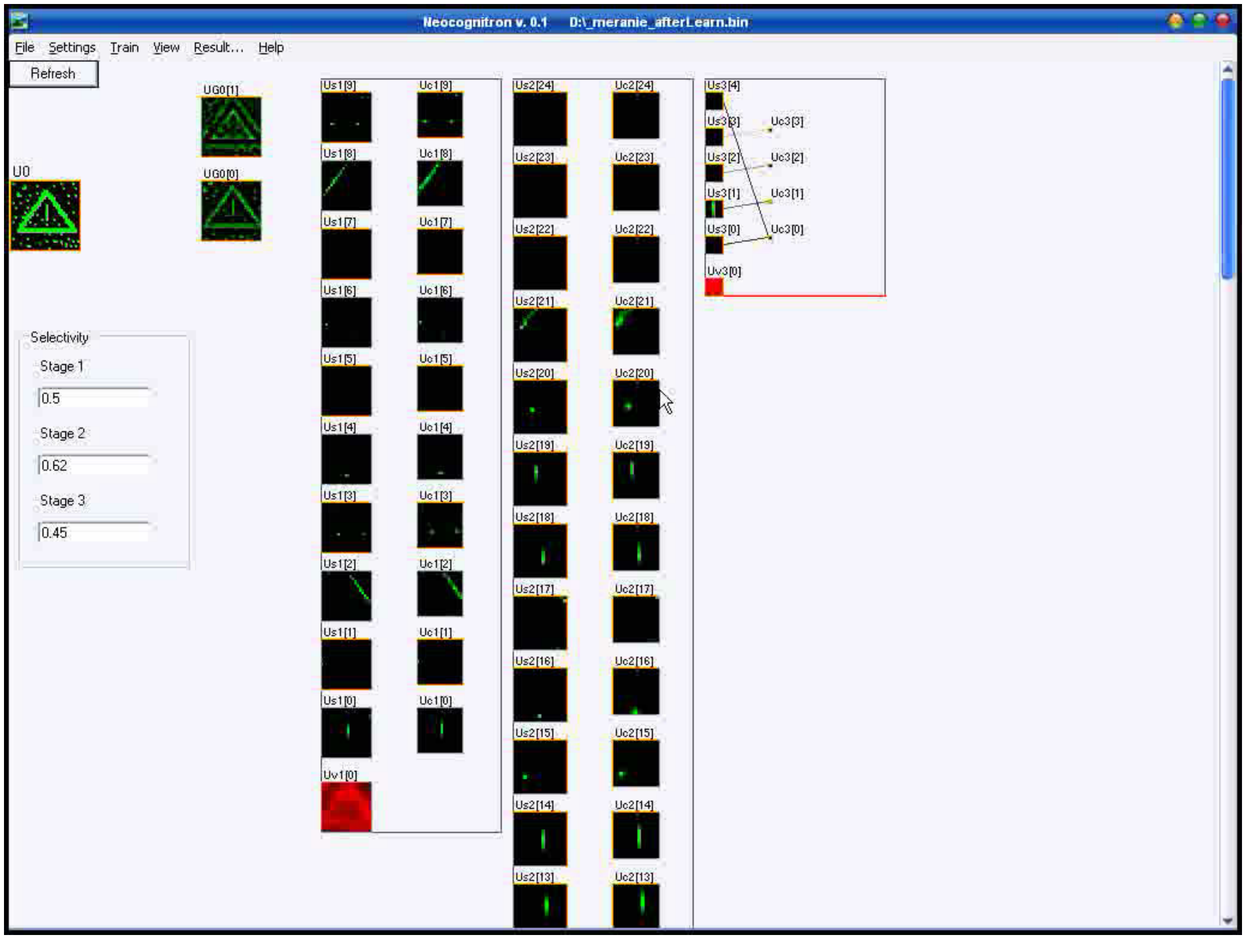
Kunihiko Fukushima, "A neural network for visual pattern recognition." IEEE Computer, March 1988

Neocognitron (Fukushima)

- Sample of recognition simulation of number ‘5‘ in trained neural network
- Training of S-cells with unsupervised or supervised methods; only S-cells have learning inputs

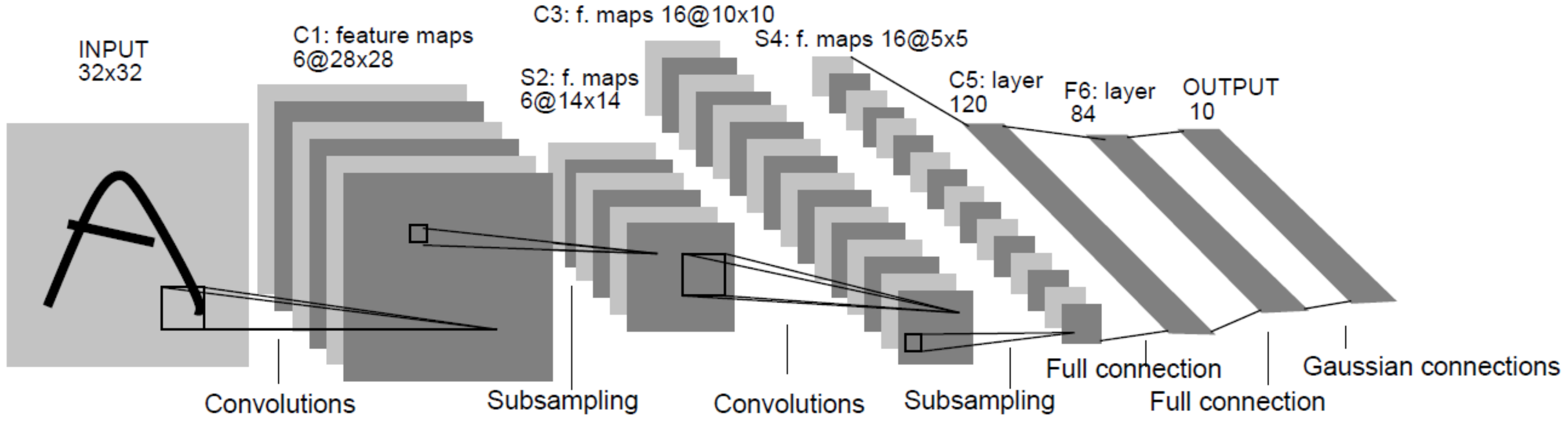


Neocognitron Simulator

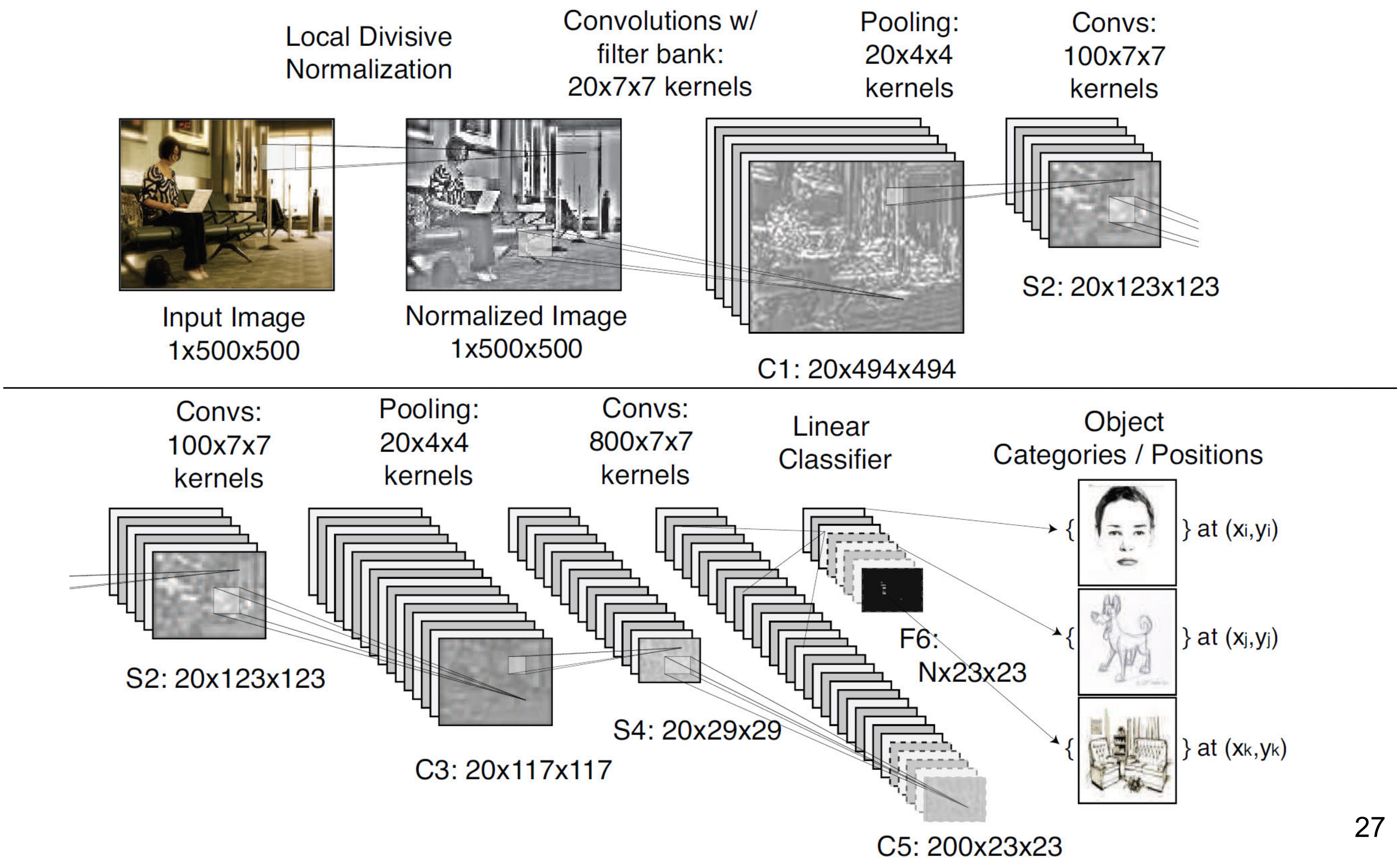


http://neocognitron.euweb.cz/index.html

Convolutional Neural Networks (LeCun)

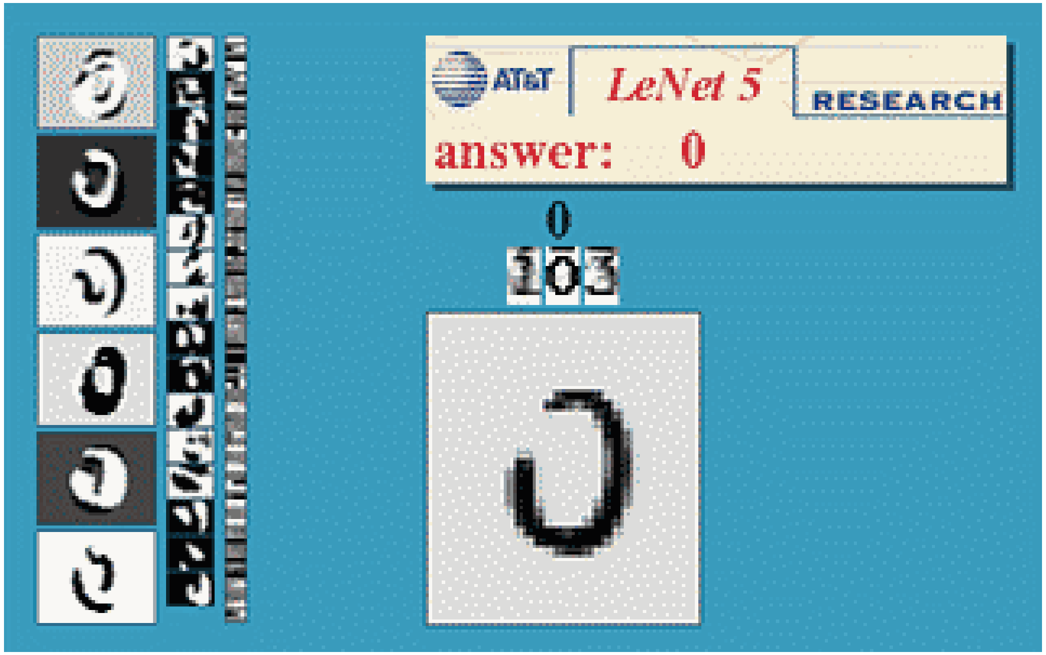


Convolutional Neural Networks (LeCun)



Convolutional Neural Networks (LeCun)

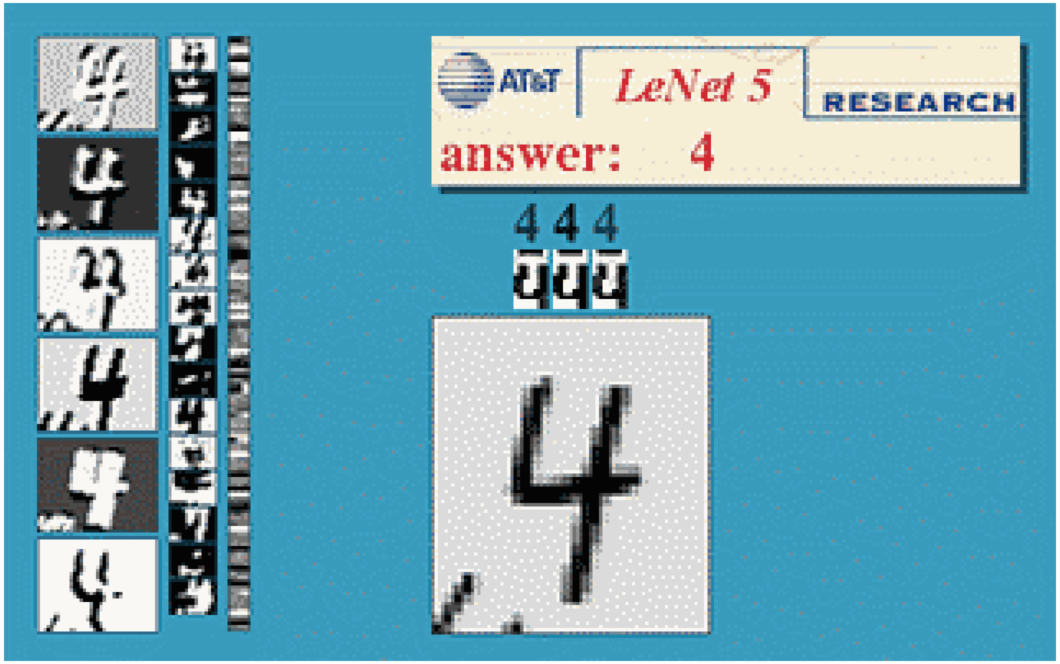
- Multi-Layer Neural Network + Backpropagation Training
- Similar to Neocognitron
- Pattern recognition directly from images with limited preprocessing



http://yann.lecun.com/exdb/lenet/

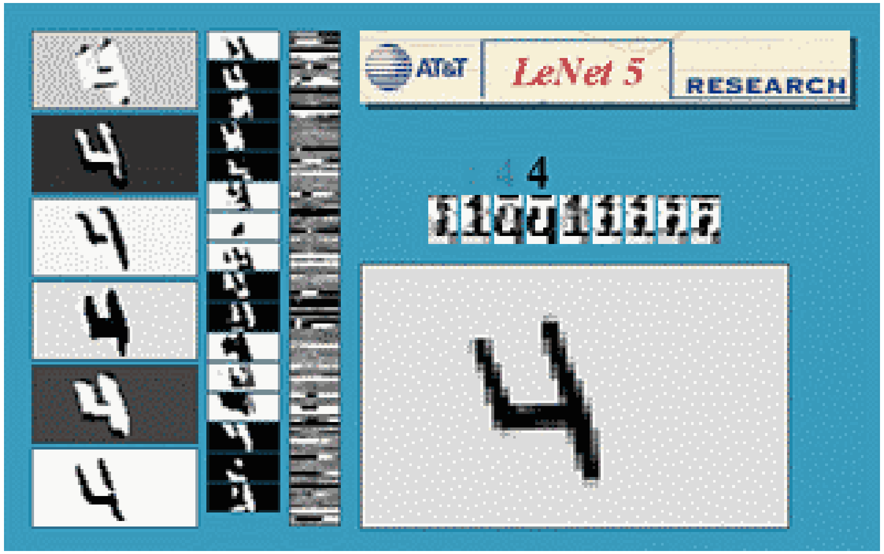
Convolutional Neural Networks (LeCun)

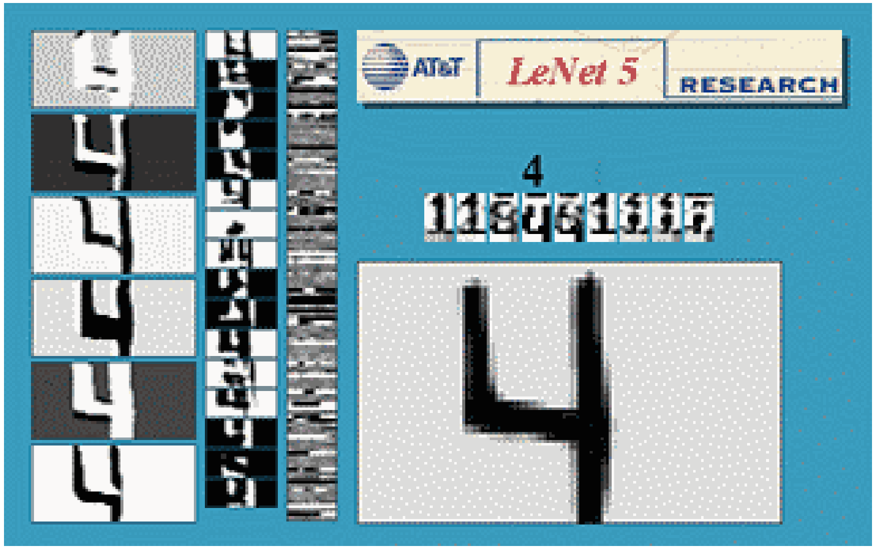
- Implementation robust to noise and image degradation
- Feature invariance in processing stages of visual cortex



Convolutional Neural Networks (LeCun)

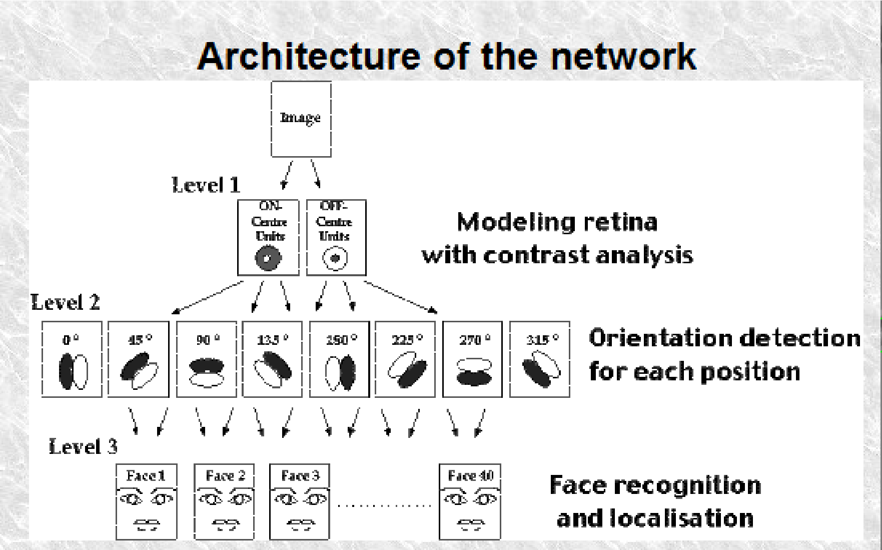
- Implementation and simulation of invariance to scaling and rotation of cipher '4'

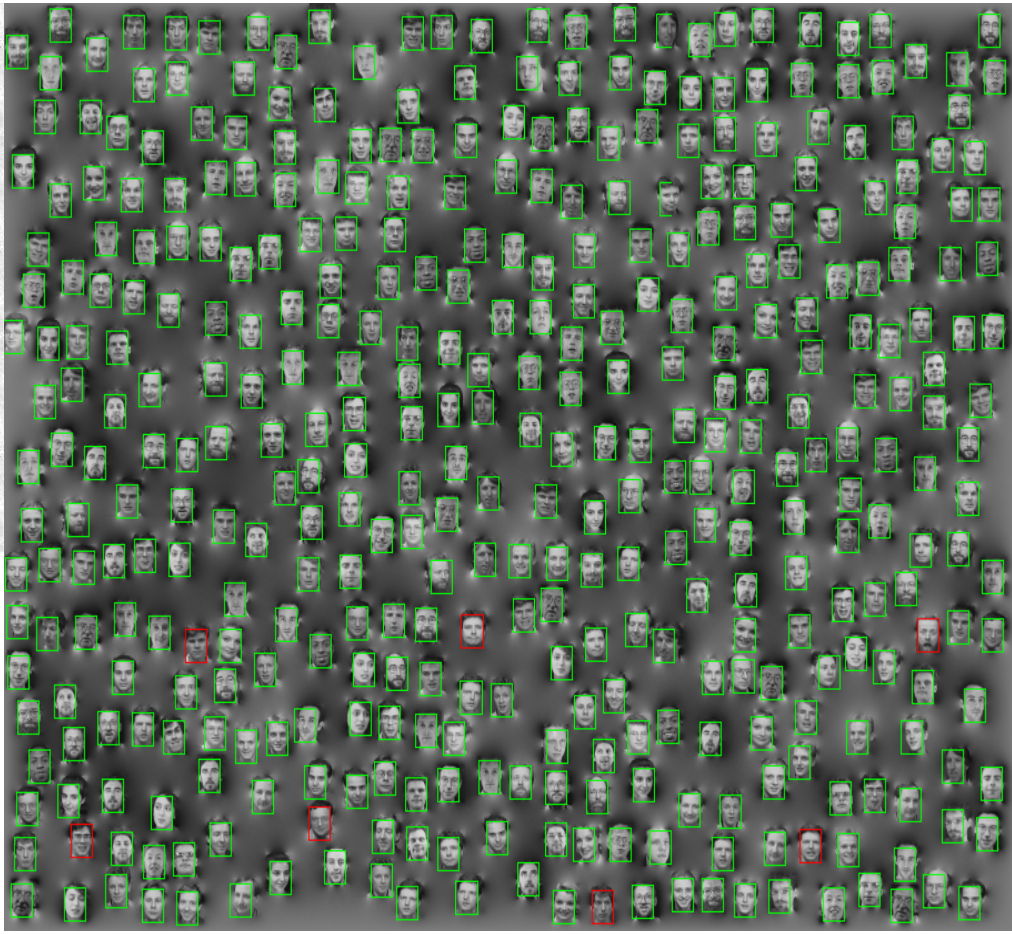




Simulation and Application: SpikeNET

- Simulator for large sets of asynchronous spiking neurons

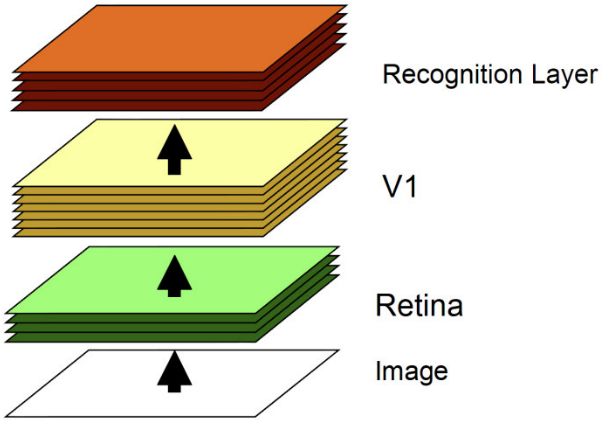




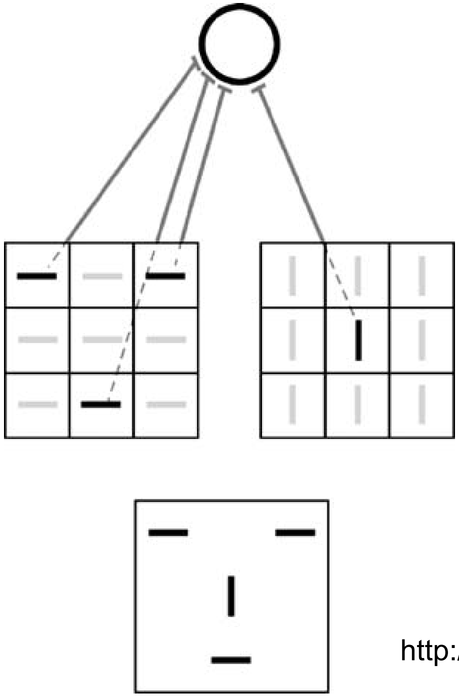
http://sccn.ucsd.edu/~arno/spikenet/

SpikeNet (Thorpe): Example

- Neuron receives different stimuli (bars) of two 3x3 arrays
- Simplification could represent a face
- Weights are concentrated on just four of the inputs: 3 horizontally tuned units, and one vertically
- Fix threshold for firing in output unit at 4 active inputs results in one spike

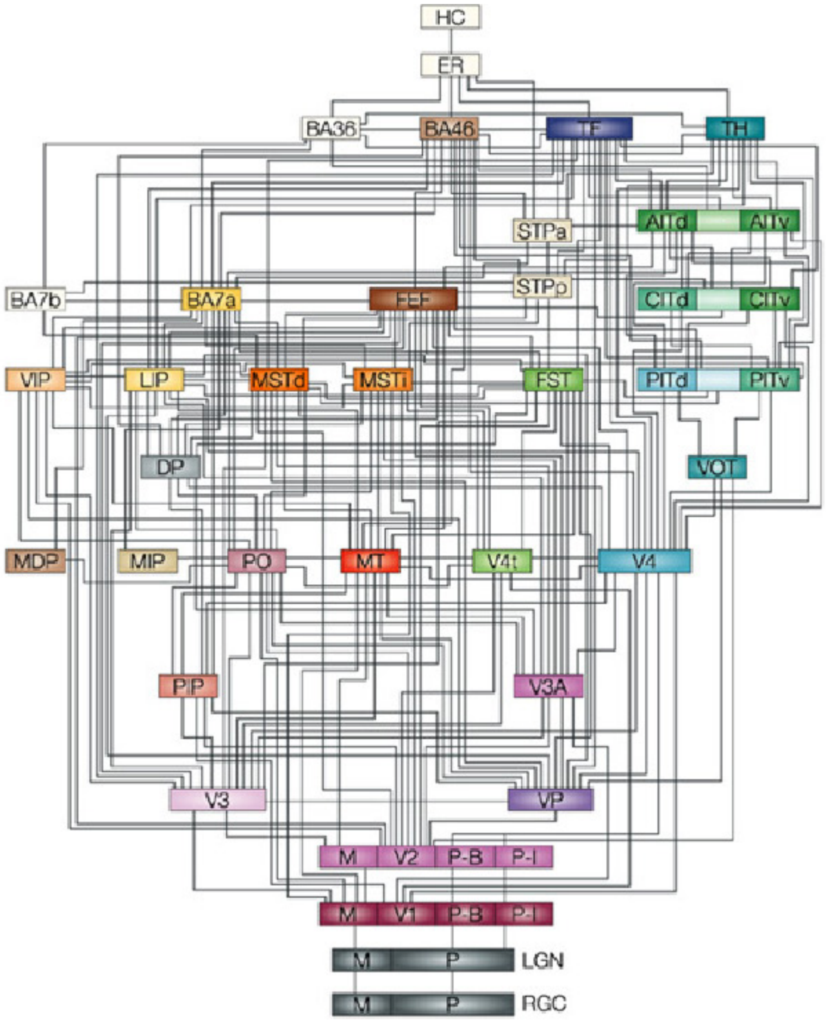


General architecture of SpikeNet



http://www.spikenet-technology.com/

Hierarchy and Connections in the Visual System



Rees Kreiman Koch. Neural correlates of consciousness in humans. Nature Reviews Neurosci, 2002

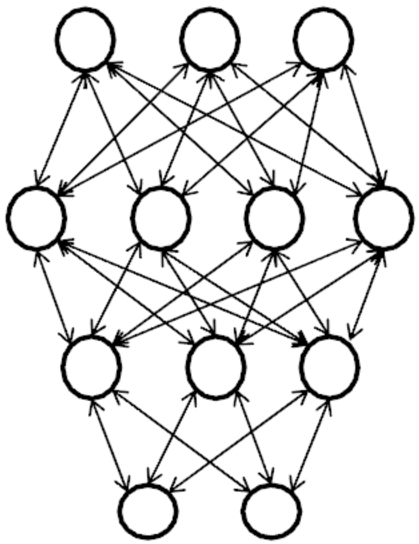
Directions in Deep Learning (1)

- Deep auto-encoders
- Learn one layer at a time

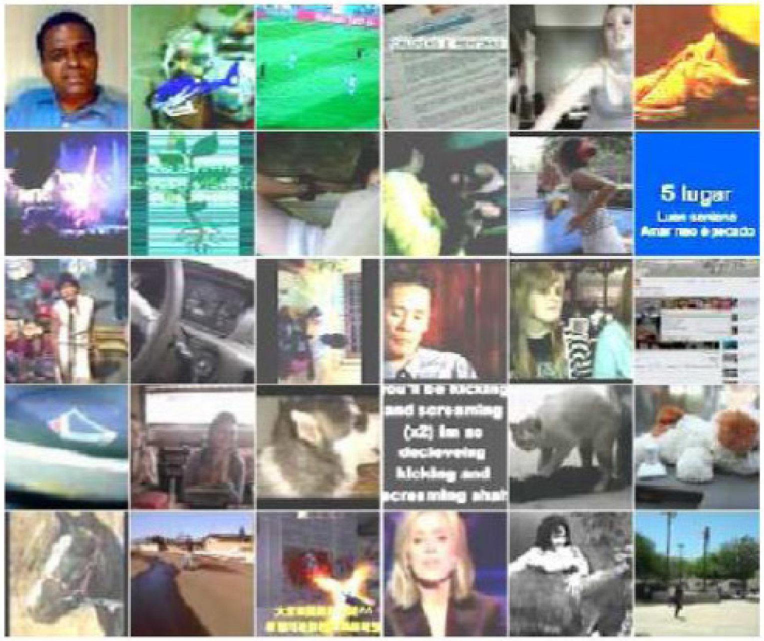
Stack auto-encoders on top of each other

Train layers one by one

Sparsity or other regularizations can be used



Directions in Deep Learning (2)



Trained on Youtube images

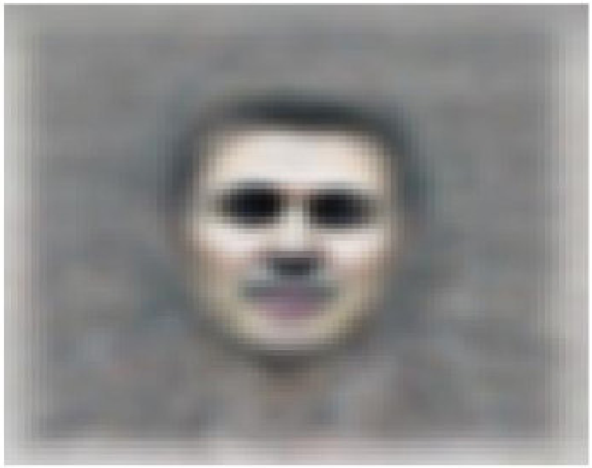
Tested on a mixture of Labeled Faces in The Wild and ImageNet

Le, et al., ICML 2012

Directions in Deep Learning (3)

"face neuron"

Images with strongest responses



Optimal stimulus

Directions in Deep Learning (4)

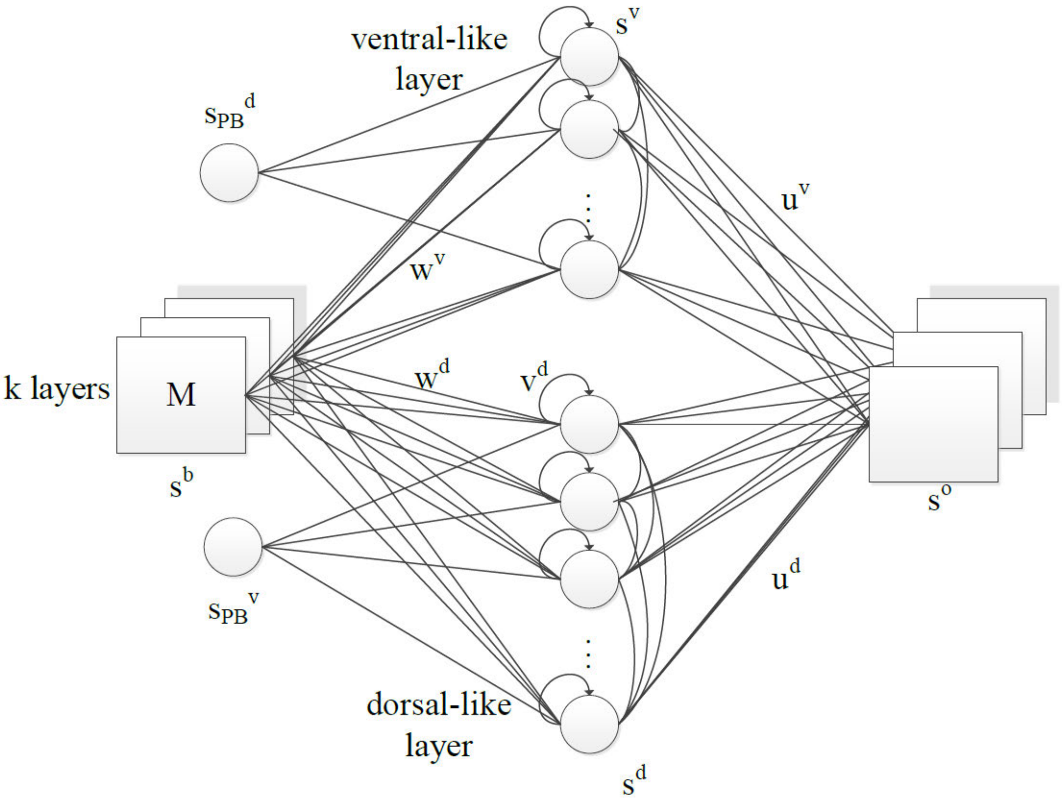
"cat neuron"

Images with strongest responses

Optimal stimulus

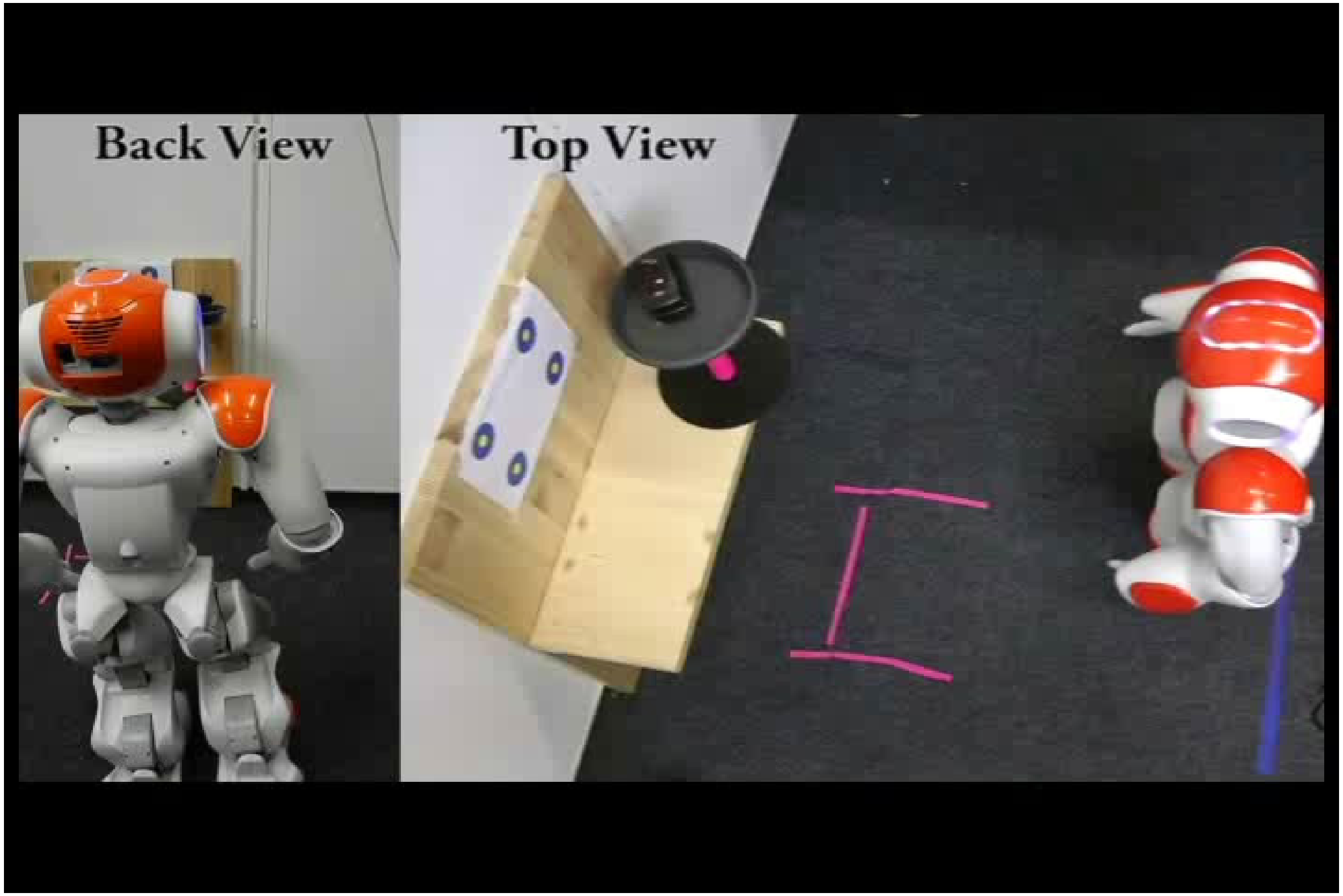
Recurrent connections for ventral and dorsal stream integration

- Questions:
  - How about storing multiple sensory sequences in one recurrent network?
  - Can the sensory information still be encoded in two streams?



J. Zhong, A. Cangelosi, and S. Wermter. Towards a self-organizing pre-symbol neural model representing sensorimotor primitives. Frontiers in Behavioral Neuroscience, 2013

Robust and Smooth Robot Docking Behaviour



Summary

- Visual information processed in different brain areas

- Images projected from retina to striate and extrastriate cortex, filtered according to different features (edges, orientation, colour, motion, …)

- Visual Processing inspired computational models for:
  - Object recogntion (LeCun, Fukushima)
  - Object detection and object recognition (Marr, Serre et al.)
  - Spiking model (Thorpe)

Further optional references

- Hubel, D.H. & Wiesel, T.N. Republication of The Journal of Physiology (1959) 148, 574-591: Receptive fields of single neurones in the cat's striate cortex. 1959. *The Journal of Physiology* 587, 2721-2732 (2009)
- Marr, D. Marr's Vision. *Vision A Computational Investigation into the Human Representation and Processing of Visual Information* (1982).
- Farabet, C., Martini, B., Akselrod, P.,Talay, S., (LeCun, Y., Culurciello, E.). Hardware Accelerated Convolutional Neural Networks for Synthetic Vision Systems. *Electrical Engineering,* 257–260 (2010)
- Thorpe, S., Guyonneau, R., Guilbaud, N., Allegraud, J. & Vanrullen, R. SpikeNet: real-time visual processing with one spike per neuron. *Neurocomputing* 58-60, 857-864 (2004).