CISC 3250
Systems Neuroscience

Perception

Professor Daniel Leeds
dleeds@fordham.edu
JMH 332

Pathways to perception in 3 (or fewer) synaptic steps

0 Input through sensory organ/tissue

1 Synapse onto neurons in spinal cord/brain stem

2 Synapse onto neurons in thalamus

3 Synapse onto cortical neurons in "primary ____ cortex"

4⁺ Further cortical processing

Types of percepts in this lecture:
- Tactile (touch)
- Audition (sound)
- Vision (sight)

Touch/"Tactile"

Skin

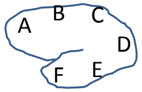
Spinal cord

Dorsal horn

Spinal cord

Thalamus

Ventral postero-lateral (VPL) nucleus

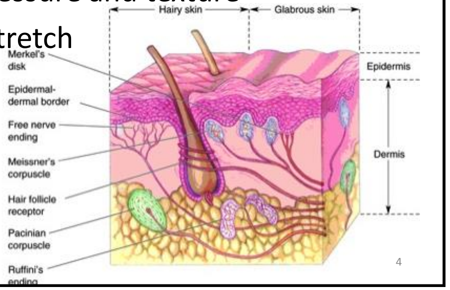


Cortex

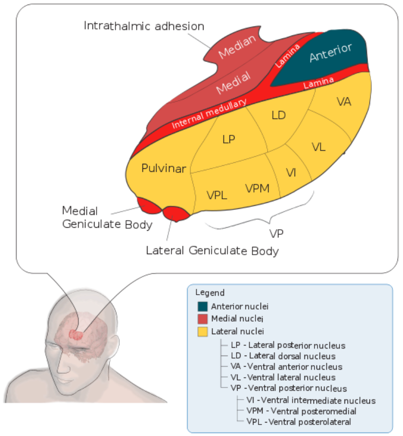
Post-central gyrus (SI)

Touch: Inputs

Mechanoreceptors in skin
- Pacinian corpuscles – vibrations
- Meissner's corpuscles – light touch
- Merkel's discs – pressure and texture
- Ruffini endings – stretch



Thalamus – the "relay" station



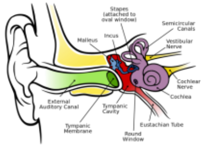
Region names largely based on location

VPL for somatosensation

VPL =
Ventral (bottom)
Posterior (back)
Lateral (side) Nucleus

http://en.wikipedia.org/wiki/File:Thalamus-schematic.svg

Hearing/"Auditory"



Cochlea

Cochlear nerve

Cochlear nucleus (-> Superior olive) -> Inferior colliculus

Brain stem

Recall: in cochlea have tonotopy
Neurons selective for specific frequencies

Thalamus

Medial geniculate nucleus (MGN)

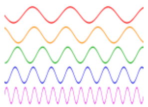
Geniculate nuclei at most posterior ventral spots in thalamus

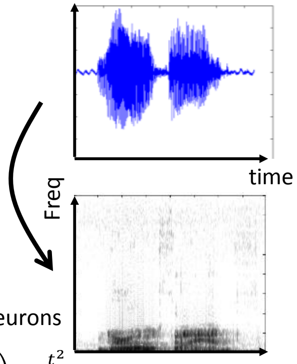
Cortex

Primary auditory cortex (AI)

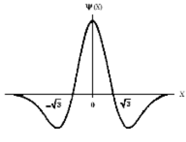
Hearing and frequency decomposition

Sound consists of times and frequencies





Time-bound wavelets:



"Mexican hat"

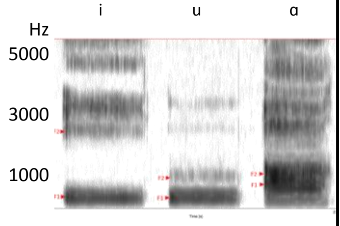
Similar to cochlear neurons

$$\frac{2}{\sqrt{3\sigma}\pi^{1/4}}\left(1-\left(\frac{t}{\sigma}\right)^2\right)e^{-\frac{t^2}{2\sigma^2}}$$

Spectrogram

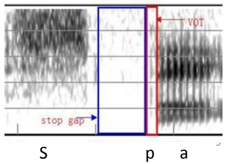
Common patterns in speech

- Vowels (a,e,i,o,u) correspond to steady frequency combinations



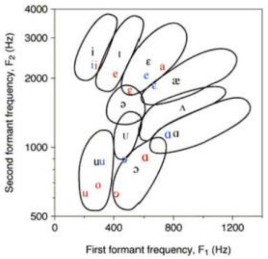
Top 2 freqs: i 300, 2500; u 300, 1000; ɑ 500, 1000

- Consonants may be broad-range frequencies, or sweeps

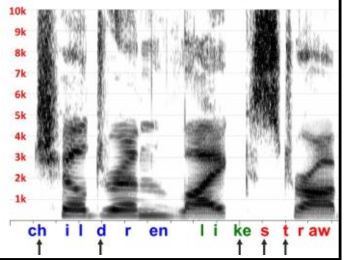


More speech pattern

- Speech **formant** ranges by frequency



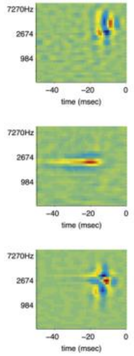
- ch, s – long high freq
- d, k, t – broad freq burst
- l, r, n, m – freq slide



Spectro-temporal receptive fields

Nagel 2008 Neuron
Zebra Finch (field L)

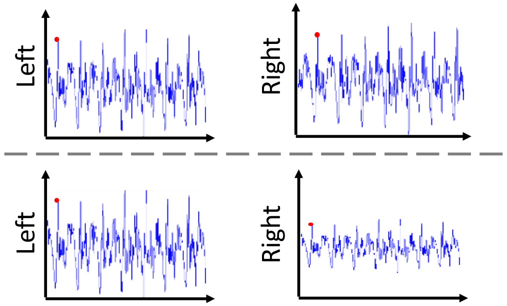
AI (primary auditory cortex) neurons selective for patterns in space and time



Binaural hearing

Comparing sounds from left and right

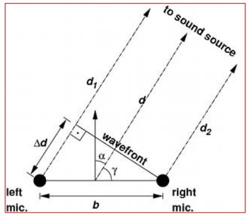
- Time shift and/or Volume Change



Applications:

- Localize sound source
- Distinguish sounds from multiple sources

Math of sound localization

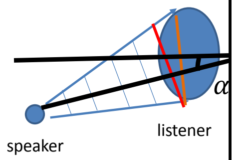


Speed of sound

c=343 m/s

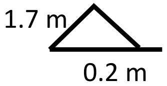
Human head

b=0.2m



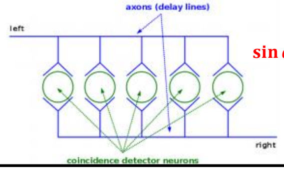
If Δt =5ms

Δd=343x0.005 = 1.7m



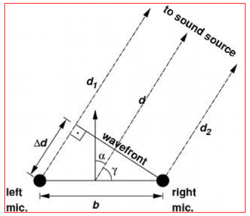
$$\alpha = \sin^{-1}\frac{c\Delta t}{b}$$

Jeffres ’48
sound delay/
axon delay model



$$\sin\alpha = \frac{opposite}{hypotenuse} = \frac{\Delta d}{b}$$

Math of sound localization

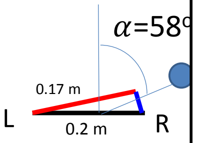


Speed of sound

c=343 m/s

Human head

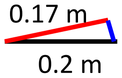
b=0.2m



If Δt =**0.5ms**

Δd=343x0.0005 = 0.17m

$$\alpha = \sin^{-1}\frac{c\Delta t}{b}$$



$$\sin\alpha = \frac{0.17}{0.2} = 0.85$$
$$\alpha = \sin^{-1}0.85$$
$$\alpha = 58^{o}$$

90° = 1.57 rad ($\frac{\pi}{2}$ rad)

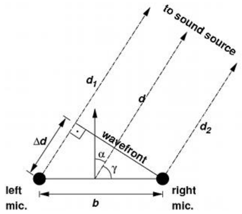
x radians:

$$\frac{90x}{1.57} \approx \frac{360x}{2\pi}\text{ degrees}$$

Sound gets R ear @ 1345.**2** ms

to L ear @ 1345.**7** ms

Math of sound localization



Speed of sound

c=343 m/s

Human head

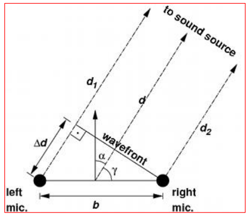
b=0.2m

$$\alpha = \sin^{-1}\frac{c\Delta t}{b}$$

Pick direction for comparison

$$\Delta t = \begin{cases} > 0 \ \text{rightSound earlier} \\ < 0 \ \ \text{leftSound earlier} \end{cases}$$

Math of sound localization



Speed of sound

c=343 m/s

Human head

b=0.2m

x radians:
$$\frac{90x}{1.57} \approx \frac{360x}{2\pi}\text{ degrees}$$

$$\alpha = \sin^{-1}\frac{c\Delta t}{b}$$

Sound gets R ear @

~~258.5 ms~~ **285.5ms**

Get to L ear @

285.3 ms

What's my α?

Closer to L ear

Which ear is sound closest to?

The ear that sound arrives at first

Ear with smaller time of arrival

$$\Delta d = c\Delta t = 343 \times (0.2855 - 0.2853)$$
$$\sin\alpha = \frac{\Delta d}{b} = \frac{343 \times 0.0002}{0.2} = \frac{.0686}{0.2} = 0.35$$
$$\alpha = asin(0.35) = 20^{o}$$

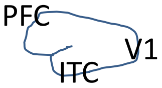
Seeing/"Visual"

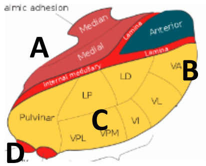
Retina

Optic nerve

Thalamus

Lateral geniculate nucleus (LGN)



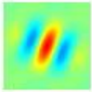


Cortex

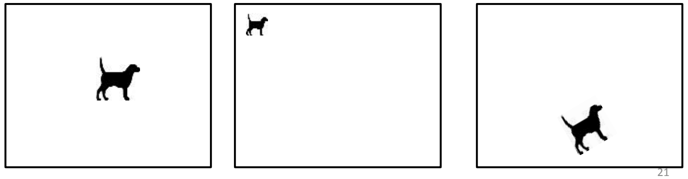
Primary visual cortex (VI)

Sensitivity to perceptual variations

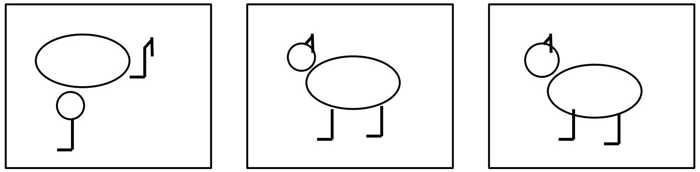
- V1: Surround-suppression for shifted edges



- ITC/PFC: Same object detected at diverse locations and scales

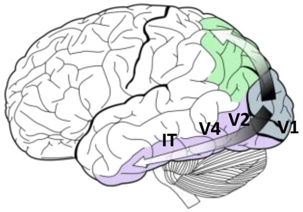


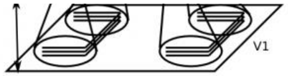
Selectivity to perceptual variations



- More complex percepts invariant to greater spatial transformations
  - But if transformation are TOO large, invariance breaks down

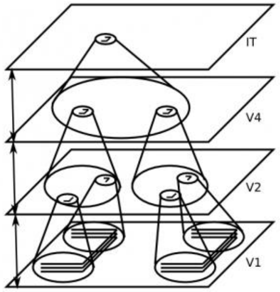
HMAX – model of hierarchical vision



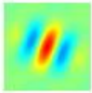


- Higher cortical levels cover larger visual spans
- Object recognition invariant to changes in location and orientation

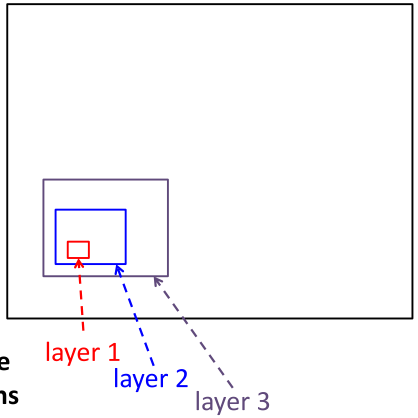
HMAX – model of hierarchical vision



1. Gabor "filters" (edge detectors)
2. Perform "Max pooling" (semi-invariance over space)
3. Weighted combination of space-invariant edges
4. Further max pooling



Higher HMAX layers cover more space



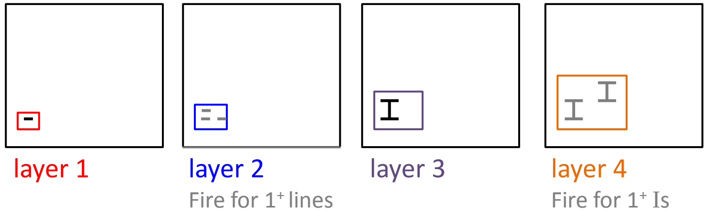
Example coverage for layer *x* neurons

Functions of HMAX layers

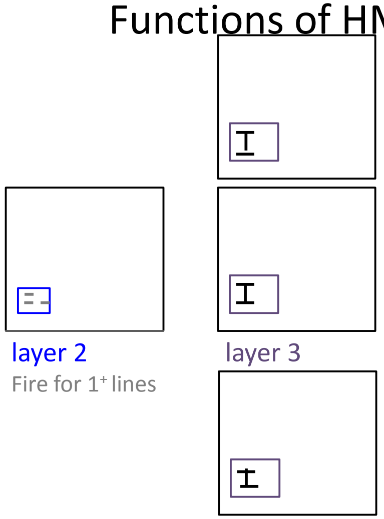
- Odd layers (layer 1, 3, 5, …) look for specific combinations of lower-level features

Neuron in layer L combines info from neurons in layer L-1

- Even layers (layer 2, 4, 6, …) provide invariance to some feature changes (e.g., shift in position)



Functions of HMAX layers



One neuron in layer 3

Looking for I shape

Combo of 3 lines in layer 2

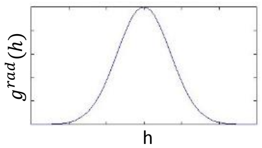
Layer 2 allows for slight position variation in each line

Functions of HMAX layers

- Odd layers (layer 1, 3, 5, …) look for specific combinations of lower-level features

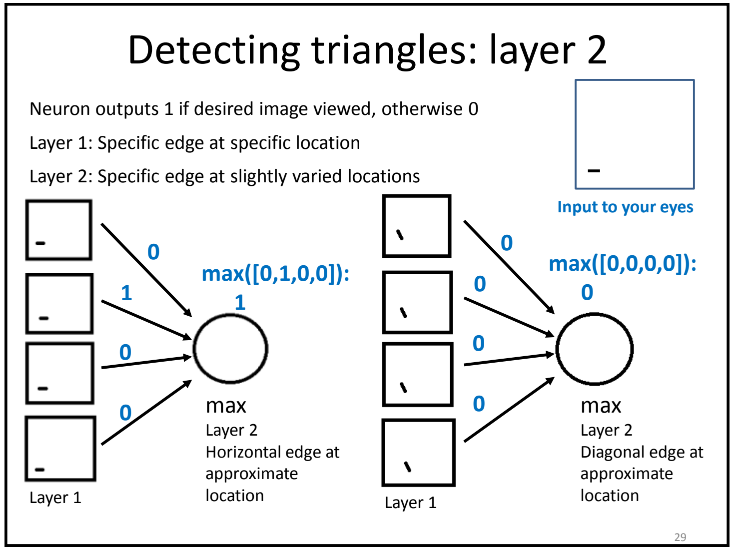
$h=\sum_j w_j r_j^{in}$ $\qquad r^{out} = g^{rad}(h)$

Radial basis function



- Even layers (layer 2, 4, 6, …) provide invariance to some feature changes (e.g., shift in position)

$$r^{out} = \max([r_1^{in} \quad r_2^{in} \quad \cdots \quad r_J^{in}])$$

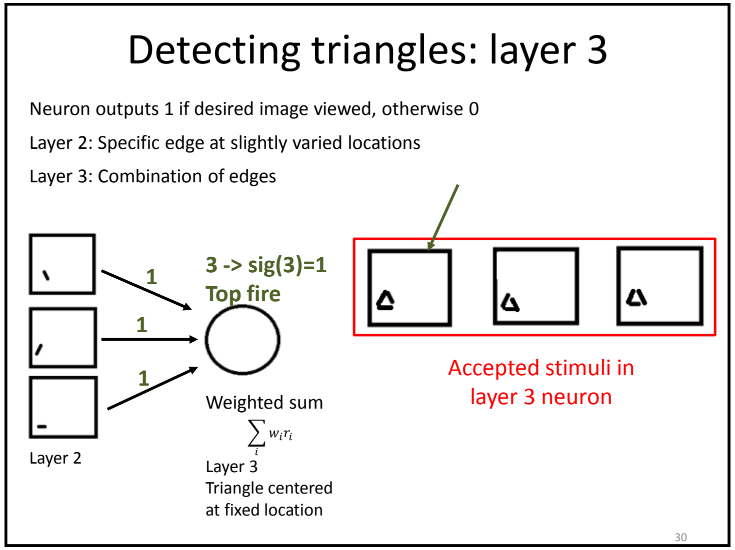


Detecting triangles: layer 2

Neuron outputs 1 if desired image viewed, otherwise 0

Layer 1: Specific edge at specific location

Layer 2: Specific edge at slightly varied locations

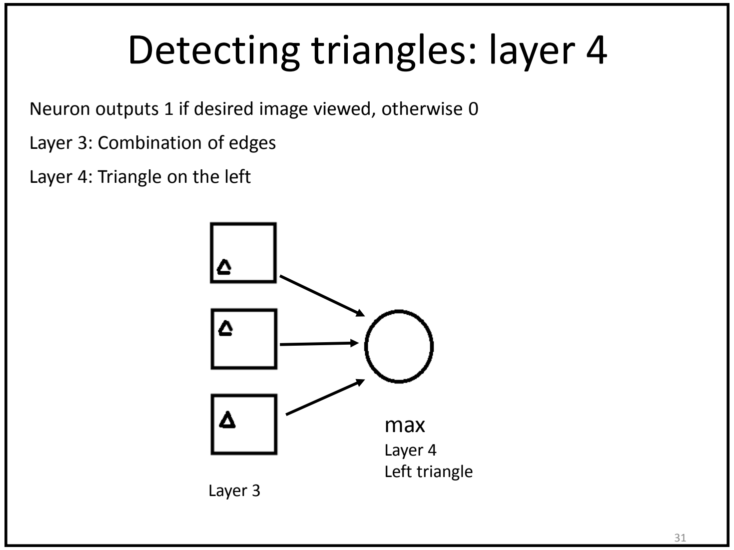


Detecting triangles: layer 3

Neuron outputs 1 if desired image viewed, otherwise 0

Layer 2: Specific edge at slightly varied locations

Layer 3: Combination of edges

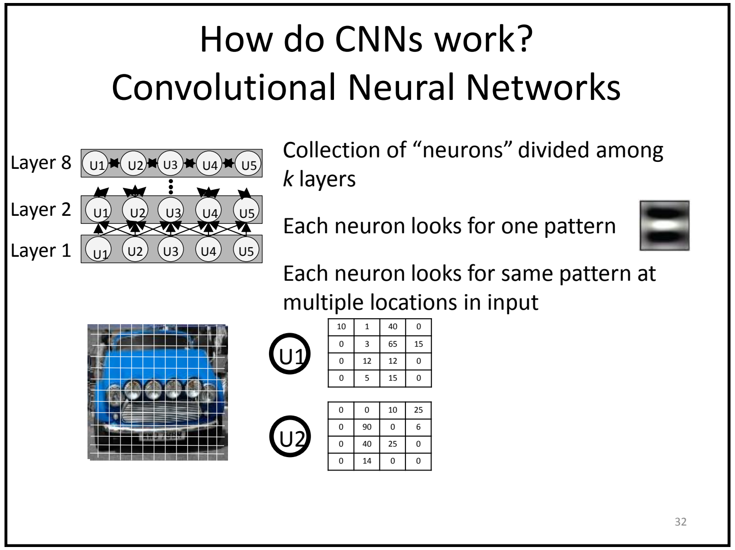


Detecting triangles: layer 4

Neuron outputs 1 if desired image viewed, otherwise 0

Layer 3: Combination of edges

Layer 4: Triangle on the left



How do CNNs work?
Convolutional Neural Networks

Collection of “neurons” divided among *k* layers

Each neuron looks for one pattern

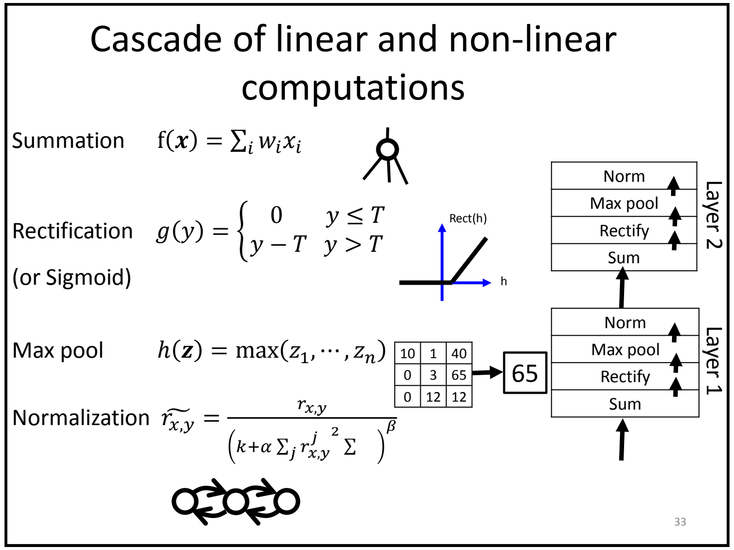
Each neuron looks for same pattern at multiple locations in input

U1

10	1	40	0
0	3	65	15
0	12	12	0
0	5	15	0

U2

0	0	10	25
0	90	0	6
0	40	25	0
0	14	0	0



Cascade of linear and non-linear computations

Summation $f(\boldsymbol{x}) = \sum_i w_i x_i$

Rectification $g(y) = \begin{cases} 0 & y \leq T \\ y - T & y > T \end{cases}$

(or Sigmoid)

Max pool $h(\boldsymbol{z}) = \max(z_1, \cdots, z_n)$

10	1	40
0	3	65
0	12	12

65

Normalization $\widetilde{r_{x,y}} = \frac{r_{x,y}}{\left(k + \alpha \sum_j {r_{x,y}^j}^2 \Sigma \quad \right)^{\beta}}$

33