

# CISC 3250 Systems Neuroscience

## Perception



Professor Daniel Leeds  
dleeds@fordham.edu  
JMH 332

## Pathways to perception in 3 (or fewer) synaptic steps

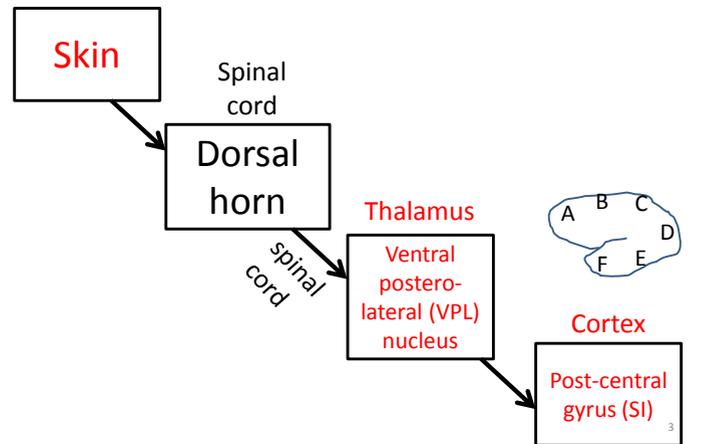
- 0 Input through sensory organ/tissue
- 1 Synapse onto neurons in spinal cord/brain stem
- 2 Synapse onto neurons in thalamus
- 3 Synapse onto cortical neurons in "primary \_\_\_\_ cortex"
- 4+ Further cortical processing

Types of percepts  
in this lecture:

- Tactile (touch)
- Audition (sound)
- Vision (sight)



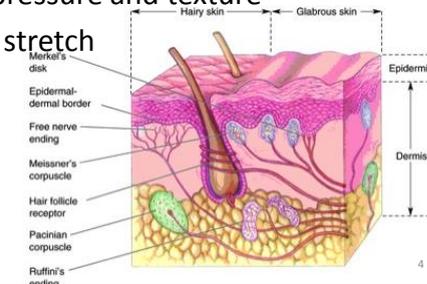
## Touch/"Tactile"



## 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

Thalamus

Medial geniculate nucleus (MGN)

Cortex

Primary auditory cortex (AI)

Geniculate nuclei at most posterior ventral spots in thalamus

Recall: in cochlea have tonotopy  
Neurons selective for specific frequencies

## Hearing and frequency decomposition

Sound consists of times and frequencies

Time-bound wavelets:

Similar to cochlear neurons

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

“Mexican hat”

Spectrogram

## Common patterns in speech

- Vowels (a,e,i,o,u) correspond to steady frequency combinations
- Consonants may be broad-range frequencies, or sweeps

i      u      a

Hz

5000

3000

1000

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

s      p      a

### 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

The top plot shows vowel space with axes for First formant frequency, F<sub>1</sub> (Hz) and Second formant frequency, F<sub>2</sub> (Hz). The bottom plot is a spectrogram of the sentence 'children like straw' with a frequency scale from 1k to 10k Hz. Arrows point to specific phonemes: 'ch' (high frequency), 'l' (broad burst), 'r' (frequency slide), 'en' (broad burst), 'l' (broad burst), 'i' (high frequency), 'k' (broad burst), 'e' (high frequency), 's' (high frequency), 't' (broad burst), 'r' (frequency slide), 'a' (broad burst), and 'w' (broad burst).

### Spectro-temporal receptive fields

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

Nagel 2008 Neuron Zebra Finch (field L)

The three plots show spectro-temporal receptive fields for a Zebra Finch neuron. Each plot has frequency (984, 2674, 7270 Hz) on the y-axis and time (-40 to 0 msec) on the x-axis. The plots show selective responses to specific spectro-temporal patterns.

### Binaural hearing

Comparing sounds from left and right

- Time shift and/or Volume Change

The diagram shows sound waves from a speaker reaching a listener's ears. Below it, two sets of waveforms illustrate 'Time shift' (where the right ear receives the sound earlier) and 'Volume Change' (where the right ear receives a louder signal).

Applications:

- Localize sound source
- Distinguish sounds from multiple sources

### Math of sound localization

Speed of sound  $c=343$  m/s

Human head  $b=0.2$  m

The diagram shows a speaker, a listener, and a sound source. The distance from the speaker to the listener is  $b$ . The angle of the sound source is  $\alpha$ . The path difference between the two ears is  $\Delta d$ .

If  $\Delta t = 5$  ms  
 $\Delta d = 343 \times 0.005 = 1.7$  m

1.7 m

0.2 m

Jeffres '48 sound delay/axon delay model

The diagram shows a coincidence detector neuron with axons from both ears. The axons have different lengths (delay lines) to compensate for the time delay of sound reaching the ears.

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

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

### Math of sound localization

Speed of sound  
 $c=343 \text{ m/s}$

Human head  
 $b=0.2\text{m}$

If  $\Delta t = 0.5\text{ms}$   
 $\Delta d = 343 \times 0.0005 = 0.17\text{m}$

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

$\sin \alpha = \frac{0.17}{0.2} = 0.85$   
 $\alpha = \sin^{-1} 0.85$  x radians:  
 $\alpha = 58^\circ$

Sound gets R ear @ 1345.2 ms  
 to L ear @ 1345.7 ms

$90^\circ = 1.57 \text{ rad } (\frac{\pi}{2} \text{ rad})$   
 $\frac{90x}{1.57} \approx \frac{360x}{2\pi} \text{ degrees}$

### Math of sound localization

Speed of sound  
 $c=343 \text{ m/s}$

Human head  
 $b=0.2\text{m}$

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

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  
 $c=343 \text{ m/s}$

Human head  
 $b=0.2\text{m}$

Sound gets R ear @ 258.5 ms  
 285.5 ms

Get to L ear @ 285.3 ms

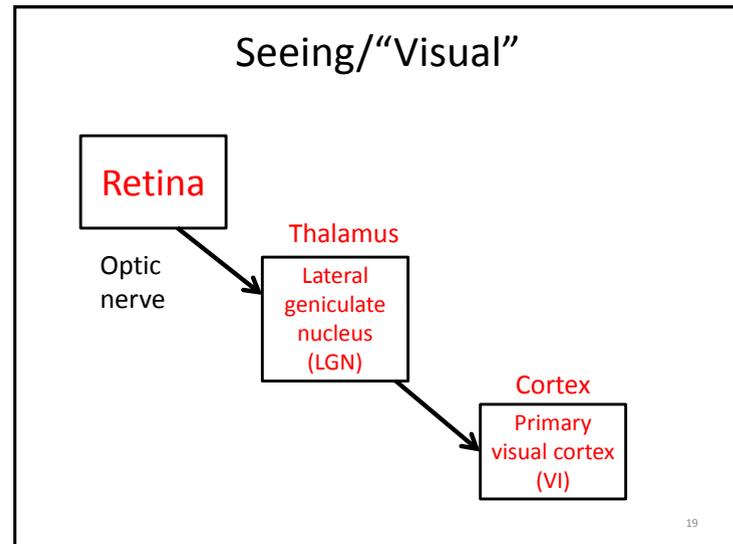
What's my  $\alpha$ ?  
 Closer to L ear

Which ear is sound closest to?  
 The ear that sound arrives at first  
 Ear with smaller time of arrival

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

$\Delta d = c\Delta t = 343 \times (0.2855 - 0.2853)$   
 $\sin \alpha = \frac{\Delta d}{b} = \frac{343 \times 0.0002}{0.2} = \frac{0.0686}{0.2} = 0.35$   
 $\alpha = \text{asin}(0.35) = 20^\circ$

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



## Sensitivity to perceptual variations

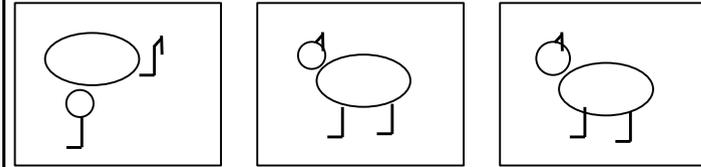
- V1: Surround-suppression for shifted edges



- PFC: Same object detected at diverse locations and scales



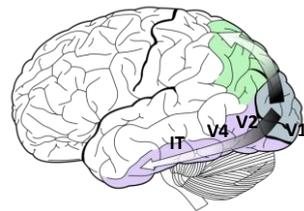
## Selectivity to perceptual variations



- More complex percepts invariant to greater spatial transformations

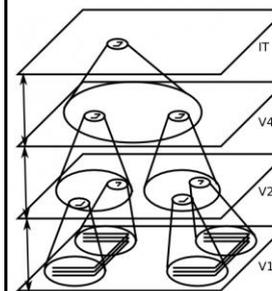
22

## 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

24

### Higher HMAX layers cover more space

Example coverage for layer  $x$  neurons

layer 1  
layer 2  
layer 3

25

### Functions of HMAX layers

- Odd layers (layer 1, 3, 5, ...) look for specific combinations of lower-level features
- Even layers (layer 2, 4, 6, ...) provide invariance to some feature changes (e.g., shift in position)

layer 1  
layer 2  
Fire for 1+ lines  
layer 3  
layer 4  
Fire for 1+ Is

26

### Functions of HMAX layers

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

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

Radial basis function

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

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

27

### 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

Layer 1  
max  
Layer 2  
Horizontal edge at approximate location  
Layer 1  
max  
Layer 2  
Diagonal edge at approximate location

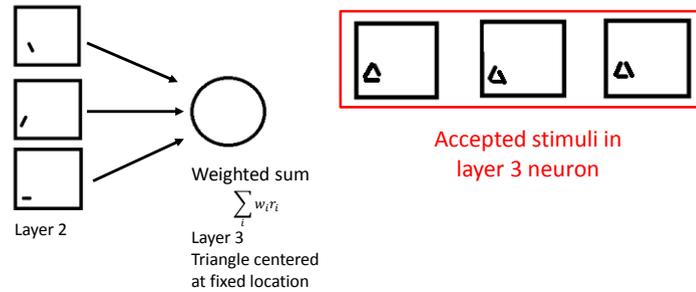
28

## 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



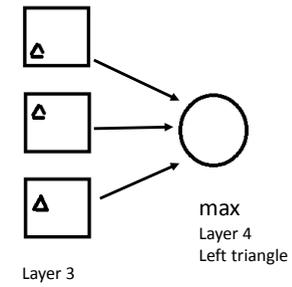
29

## Detecting triangles: layer 4

Neuron outputs 1 if desired image viewed, otherwise 0

Layer 3: Combination of edges

Layer 4: Triangle on the left



30