6. Perception

A. Biology of Perception

Retinal Cells
- Retina is CNS tissue
- Nonuniform distribution of cells
  - rods in periphery specialized for sensitivity and motion
  - cones in macula/fovea specialized for color & form (acuity)
  - humans are foveating animals
- Information compression:
  - 126 million rods (low light)
    - large convergence of rods on bipolar ⇒ sensitivity
  - 6 million cones (color and high acuity)
  - 1 million RGC (retinal ganglion cells)
    - estimated to transmit 10^9 bits/sec
    - spontaneously active; information conveyed by change in rate

Outline
A. Biology of Perception
B. Primary Visual Cortex
C. Object Recognition and “What” Pathway
D. Attention and “How” Pathway

Optic Pathway
- temporal
- nasal chiasm
- LGN
- V1
- V2, V4
Key Organizing Principles

- Transduction of different information
  - wavelength (rods; blue, green, red cones)
  - spatial frequency (resolution)
  - motion
- Topographic organization
  - contrasting similar information
- Filtering to extract relevant information

Retinal Contrast Filtering

- On-center
- Off-center
- Contrast sensitive

LGN of the Thalamus

- A “relay station,” but also much more
- Organizes different types of information into different layers with aligned retinotopic maps
- Performs dynamic processing: magnocellular motion processing cells, attentional processing
- On- and off-center information from retina is preserved in LGN

Structure of LGN

- Cells have monocular input
- Six layers alternate input from two eyes (RGC)

“What” vs. “Where” Pathways

- “What” ignores differences in location, illumination, size, rotation
- “Where” emphasizes location, size, and ignores object identity
6. Perception

Hierarchy of Macaque Visual Areas

Principal Regions in “What” Pathway
- V1: Primary Visual Cortex
  - encodes image in terms of oriented edges
- V2: Secondary Visual Cortex
  - encodes in terms of intersections & junctions
- V4
  - more complex features over wider range of locations
- PIT: Posterior Inferotemporal (IT) Cortex
  - location & size invariant object recognition
  - includes FFA (fusiform face area)
- AIT: Anterior IT Cortex
  - abstract/semantic visual information

Hierarchy of Visual Detectors

B. Primary Visual Cortex
What is the origin of detectors for oriented bars of light?

Self-Organization of V1 Orientation Selective Neurons

Topographic Maps
- Map of orientations
- Hypercolumn: Full set of coding for each position
- Pinwheel can arise from learning and lateral connectivity: not hard-wired!
6. Perception

**What is Common?**

**V1Rf: Simulating One Hypercolumn**

- Natural visual scenes are preprocessed by passing them (separately) through layers of on-center and off-center inputs.
- Hidden layer: edge detectors seen in layers 2/3 of V1; Layer 4 (input) just represents unoriented on/off inputs like LGN (but can be modulated by attention).
- Circular neighborhood of lateral excitatory connectivity in Hidden layer.

**Self-Organized Topography**

Model shows how documented V1 properties can result from interactions between learning, architecture (connectivity), and structure of environment.

**Faces vs. Natural Scenes**

Some differences, but pinwheels still emerge.

**Rewiring Cortex**

- Experiments by Mriganka Sur & colleagues (MIT).
- What happens if retinal axons are redirected into auditory thalamus (MGN) instead of its usual inputs?
- **Answer:** Auditory cortex (A1) develops orientation columns and retinotopic maps similar to V1.
- Animals experience activity in A1 as visual perception.
6. Perception

Orientation Columns in A1

Orientation columns develop in A1 similar to those in V1

Are They Having Visual Experiences?

They Are Having Visual Experiences

Visual Acuity

• There is less visual acuity in the rewired pathway
• Suggests there may be intrinsic factors in organization of auditory cortex as well as extrinsic factors

C. Object Recognition and “What” Pathway

How do we recognize objects (across locations, sizes, rotations with wildly different retinal images)?

Invariant Object Recognition
6. Perception

Invariant Object Recognition

- Hierarchy of increasing:
  - Feature complexity
  - Spatial invariance
  - Increasing RF size:
    - Conjunction of features (to form more complex objects)
    - Collapsing over location information (“spatial invariance”)
  - Strong match to RF’s in corresponding brain areas

Biological Data: Increasing Complexity and Invariance

The Model: combining Fukushima with convolutional neural nets, bidirectional connectivity and learning!

V1 Receptive Fields

- 4x5 hypercolumns
- Two rows of simple cells at 4 orientations and two polarities
- Two rows of end-stop complex cells
- One row of length-sum complex cells
- 50% overlap with adjacent hypercolumns

Simple Textbook Test
Activation-Based Receptive Fields

- How do we plot receptive fields for V4?
- Receiving weights show which V1 units a V4 unit responds to, but they don’t show what thing in the world the unit responds to.
- **Solution**: Show the network lots of input patterns. Then, display a composite of all the input patterns that activate the unit (weighted by activity).

V4 Receptive Fields

- Some V4 units code for simple features in a location invariant way.
- This will show up as smeary parallel lines in Image input.

3D Object Recognition Test

3D models from Google SketchUp
100 categories
9–10 objects per category
2 objects left out for testing
-1–20° basic depth rotation
0–30° vertical depth rotation
14° 2D planar rotations
25% scaling
30% planar translations

http://grey.colorado.edu/CompCogNeuro/index.php/CU3D (slide < O'Reilly)
Generalization to Novel Category Exemplars — Better than 90%

emergent demonstration: Objrec

D. Attention and “How” Pathway

Why is visual system split into what/where pathways?
Why does parietal damage cause attention problems (neglect)?

Some Functions of Dorsal Pathway
- “Where” pathway (spatial relations)
  - visual attention (this chapter)
- But more broadly “how” pathway
  - maps perception to action (next chapter)
- Numerical and mathematical processing
- Representation of abstract relationships
- Modulation of episodic memory
- Aspects of executive control

Spatial Attention and Neglect

Hemispatial Neglect

Mainly from injuries to right parietal cortex
6. Perception

**Posner Task**
- Valid cues speed performance (relative to "no cue" condition)
- Invalid cues slow performance (relative to "no cue" condition)

- Patients perform normally in the "neutral" (no cue) condition, regardless of where the target is presented
- Patients benefit just as much as controls from valid cues
- Patients are hurt more than controls by invalid cues

**Models: Boxology vs. Biology**
- Model explains the basic finding that valid cues speed target processing, while invalid cues hurt
- Also explains finding that patients with small unilateral parietal lesions benefit normally from valid cues in ipsilateral field but are disproportionately hurt by invalid cues
- No need to posit "disengage" module
- Also explains finding of neglect of contralateral visual field after large, unilateral parietal lesions when some stimulus is present in ipsilateral field ("extinction")

**What’s Missing? Lacking Depth..**

**Gabor Uncertainty Principle and Gabor Elementary Functions**

MacLennan, B. J. Gabor Representations of Spatiotemporal Visual Imagery. University of Tennessee, Knoxville, Computer Science Department technical report CS-91-144, September 1991
Dennis Gabor

- Dennis Gabor (1900–79) is the father of holography (1947, 1971 Nobel Prize in Physics)
- “the future cannot be predicted, but futures can be invented”
- Developed a theory of information (1946) complementary to Shannon’s theory
- Gabor Uncertainty Principle based on same mathematics as derivation of Heisenberg Uncertainty Principle
- Nearly optimal Gabor representations are used in primary visual cortex

Time to Detect Difference in Frequency

\[ \Delta f \Delta t \geq 1 \]

A Little More Formally…

- **Nominal duration** \((\Delta t)\) = duration of rectangular pulse with same area as signal and height equal to amplitude at origin
- Hence, \(\Delta t |\varphi(0)| = \int_{-\infty}^{\infty} |\varphi(t)| dt\)
- Some details omitted

A Little More Formally (2)

- **Nominal bandwidth** \((\Delta f)\) of spectrum = width of rectangular pulse with height equal to spectrum’s amplitude at origin and same area as absolute value of spectrum
- Hence, \(\Delta f |\Phi(0)| = \int_{-\infty}^{\infty} |\Phi(f)| df\)

A Little More Formally (3)

- Computing the Fourier transform at origin, \(|\Phi(0)| \leq \Delta t |\varphi(0)|\)
- So \(\Delta t \geq |\varphi(0)|/|\varphi(0)|\)
- Computing the inverse Fourier transform at origin, \(|\varphi(0)| \leq \Delta f |\Phi(0)|\)
- So \(\Delta f \geq |\varphi(0)|/|\Phi(0)|\)
- Hence, \(\Delta f \Delta t \geq 1\)

1D Gabor Elementary Function

Gaussian modulated complex exponential
Orthogonal Components of 1D Gabor

\[ e^{if} = \cos ft + i \sin ft \]

Real Part of 2D Gabor Elem. Function

Maximum Logon Content of Signal

\[ \Delta t/\Delta f = 1 \]

Maximum Logon Content

- \( HT = M\Delta t \) is the duration and \( F = N\Delta f \) is the bandwidth
- The maximum number of logons \( MN \) is achieved when \( \Delta t\Delta f = 1 \) (i.e., Gabor elementary functions)
- In general, the area doesn’t have to be divided into rectangles of the same shape, so long as area is 1
- So the maximum logon content is \( TF \) (duration times bandwidth)
- Any such signal can be represented uniquely as a sum of \( TF \) Gabor elementary functions

Gabor Representations

- Any “finite energy” function \( \psi \) of finite duration \( D \) and finite bandwidth \( F \) is equal to a linear superposition of Gabor elementary functions:
  
  \[ \psi(t) = \sum_{j=0}^{M-1} \sum_{k=0}^{N-1} a_{jk} C_{jk}(t) + b_{jk} S_{jk}(t) \]

  where \( C_{jk}(t) = e^{-\pi (t-\Delta t)^2/a^2} \cos[2\pi k\Delta f (t - j\Delta t)] \)
  
  and \( S_{jk}(t) = e^{-\pi (t-\Delta t)^2/a^2} \sin[2\pi k\Delta f (t - j\Delta t)] \)

  - The same applies in higher dimensions.

Gabor Filters in Early Vision

- Measurements of receptive fields of simple cells in cat visual cortex have show them to be like Gaussian-modulated sinusoids (Jones & Palmer, 1987)
- Daugman (1984, 1985, 1993) showed 97% of them are statistically indistinguishable from the odd- or even-symmetric parts of a 2D Gabor elementary function
- Adjacent simple cells have grating patches that are 90\(^\circ\) out of phase, but matched in preferred orientation and frequency
- And more… (MacLennan, 1991)