Perception & Attention

Perception is effortless but its underlying mechanisms are incredibly sophisticated.

- Biology of the visual system
- Representations in primary visual cortex and Hebbian learning
- Object recognition
- Attention: Interactions between systems involved in object recognition and spatial processing

Some motivating questions:
1. Why does primary visual cortex encode oriented bars of light?
2. Why is visual system split into what/where pathways?
3. Why does parietal damage cause attention problems (neglect)?
4. How do we recognize objects (across locations, sizes, rotations with wildly different retinal images)?

Overview of the Visual System

Hierarchies of specialized visual pathways, starting in retina, to LGN (thalamus), to V1 & up:

LGN of the Thalamus

A "relay station", but so much more.

- Organizes different types of information into different layers.
- Performs dynamic processing: magnocellular motion processing cells, attentional processing.
- On- and off-center information from retina is preserved in LGN

The Retina

The retina is not a passive "camera"

- Key principle: contrast enhancement that emphasizes changes over a "pay station" but so much more.

Two Streams: Ventral "what" vs. Dorsal "where"

The Retina

The LGN

The Visual System

Recap: Attention and spatial processing
- Recall: Object recognition
- Supports interactions in primary visual cortex and attentional learning
- Should we design the visual system to do this?
- Perceptual competition: Is the underlying mechanism of light?
Primary Visual Cortex (V1): Edge Detectors

V1 combines LGN (thalamus) inputs into oriented edge detectors:

- On-center
- Off-center

Edges differ in orientation, size (spatial frequency), and position.

For coherent vision, need to detect varying degrees of all these.

Primary Visual Cortex (V1): Topography

Hypercolumns: Full set of coding for each position. Pinwheels can arise from learning and lateral connectivity: not hard-wired!

Rerouting of Visual Info to Auditory Cortex

Sharma, Angelucci & Sur (2000), Nature

Rerouted fibers from Retina → auditory thalamus (MGN) → A1

If visual properties are known, they should develop in A1.

Visual Behavior After Rerouting Right Visual Field

Visual Acuity After Rerouting

Learning is powerful, but so is evolution!

A Question

What makes visual cortex visual cortex? Why does it represent oriented bars of light?

Primary Visual Representations

Key idea: Oriented edge detectors can develop from Hebbian correlational learning based on natural visual scenes.

The Model: 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 are seen in layers 2/3 of V1 (Layer 4), but can be modulated by attention such as vision of object detection.

The Model: Simulating one Hypercolumn

- Hebbian learning only.
- KWTA inhibitory competition for specialization (see Ch 4).

[v1rf.proj]
Perception and Attention

1. Why does primary visual cortex encode oriented bars of light? Correlational learning based on natural visual scenes. Reflects reliable presence of edges in natural images, which vary in size, position, orientation and polarity. It models how documented V1 properties can result from interactions between learning and biophysical mechanisms.

Model Training on Faces

Red = on-center, Blue = off-center. How many babies are on the mountain tops? How many faces are in the mirror?

The Receptive Fields

Red = on-center, Blue = off-center. How many center faces? How many off-center faces?
Perception and Attention

1. Why does primary visual cortex encode oriented bars of light?
   Correlational learning based on natural visual scenes.

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

3. Why is visual system split into what/where pathways?

4. Why does parietal damage cause attention problems (neglect)?

The Object Recognition Problem

1. Why do paradigmatic shape transformation problems (rigid) not work?
2. How do we recognize objects (across location, size, rotations)?
   Correlational learning based on neural activity sources.
3. Why does visual system split into what/where pathways?
4. Why does premotor cortex encode objects at the level of shape?
The Model

LGN_On

LGN_Off

V1

V2

V4/IT

Output

V1 = oriented line (edge) detectors, hard-coded
V2 units encode conjunctions of V1 edges across a subset of space
Each V4 unit pays attention to all of V2

The Objects

Each object is presented at multiple locations, sizes.
Network's job is to activate the appropriate Output unit (0-19) for each object, regardless of location and size.

Activation-Based Receptive Fields
- How do we find receptive fields for V2?
- Do V2 units participate in processing multiple inputs simultaneously?
- Plot which output units are active when an object is present.
- Display a composite of all the input patterns that activate the unit.
- V1 units code for signal strength, not location.
- Every V4 unit is sensitive to all of V2.
- This should be a map that is initially unknown but becomes refined as input data is presented.
Generalization

• Can the network generalize to unseen views of studied objects?

• In other words: Does training the network to recognize a set of objects in a size/location invariant way help it recognize new objects in a size/location invariant way?

• Procedure:
  - Take a net trained on 10% of possible sizes/locations.
  - Approx. 75% correct on novel views following training on 10%.

  - Can the network generalize to unseen views of studied objects?

Generalization

• Can the network generalize to unseen views of studied objects?

• In other words: Does training the network to recognize a set of objects in a size/location invariant way help it recognize new objects in a size/location invariant way?

• Procedure:
  - Take a net trained on 10% of possible sizes/locations.
  - Approx. 75% correct on novel views following training on 10%.

  - Can the network generalize to unseen views of studied objects?
Yeah, but these objects are regularly shaped, straight lines...

What about real objects?
Benefit of Bidirectionality: Noise Robustness

Generalization (n=3), 4 splits

Error rate

Why?

Cluttered Backgrounds

Performance degrades significantly — in Y2

A Challenge

Generalization Results: 92.8%

Video Demo

Emer the robot recognizing objects.
State of the Art

Still missing...

Motion

- Neurons in area MT very sensitive to motion
- Lots of work on how downstream areas integrate motion signals across time to detect coherence (e.g. Shadlen, Newsome, etc)
- Thomas Serre has shown that motion signals very reliable for discriminating between particular actions (e.g. throwing a baseball)
- Should be able to solve problem via bidirectional influence of motion integration signals, object recognition, and spatial attention (next)....

Perception and Attention

1. Why do primary visual cortex encode oriented bars of light?
   - Correlational learning based on natural visual scenes.
2. How do we recognize objects (across locations, sizes, rotations with wildly different retinal images)?
   - Transformations: increasingly complex featural encodings, increasing levels of spatial invariance; Distributed representations.
3. Why is visual system split into what/where pathways?
4. Why does parietal damage cause attention problems (neglect)?

Spatial Attention: Unilateral Neglect

- Patient copying a scene
- Self portrait, copying, line bisection tasks:
  - In all cases, patients with parietal/temporal lesions seem to forget about 1/2 of space but they still see it!
Effects of Parietal Lesions on Posner Task

- Patients perform normally in the "neutral" 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.

Possible Models

- Alert
- Interrupt
- Localize
- Disengage
- Move
- Engage
- Inhibit

Attention emerges from bidirectional constraint satisfaction.

Simple Model

\[
\begin{align*}
\text{Input} &\quad \rightarrow \quad \text{V1} \\
\text{Spat1} &\quad \rightarrow \quad \text{Spat2} \\
\text{Obj1} &\quad \rightarrow \quad \text{Obj2} \\
\text{t a r g e} &\quad \rightarrow \quad \text{c u e} \\
\text{Output} &\quad \rightarrow \quad \text{Object 1 (Cue)} \quad \text{Object 2 (Target)}
\end{align*}
\]

Effects of Parietal Lesions on Posner Task

When about space, unilateral parietal lesions
- Hemispheric do not respond to target in the neglected space
- The opposite (contralateral) side of space
- Large unilateral parietal lesions result in neglect of

Effects of Parietal Lesions on Posner Task

Invalid cue

Invalid cue
Can explain in terms of accommodation (retinal lag).

Performance is better on valid vs invalid trials.

\[ \text{Valid cue} \]

Typically larger detection is easier on trials with valid cues.

### Posner Task Data

<table>
<thead>
<tr>
<th>Group</th>
<th>Normal</th>
<th>Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adult</td>
<td>350</td>
<td>640</td>
</tr>
<tr>
<td>Elderly</td>
<td>390</td>
<td>600</td>
</tr>
<tr>
<td>Elderly, normalized</td>
<td>40</td>
<td>760</td>
</tr>
<tr>
<td>Normalized</td>
<td>540</td>
<td>350</td>
</tr>
</tbody>
</table>

### More Posner Lesion Fun

- Returning to patient with left parietal lesion...
  - What happens if cues are presented in contralateral (affected) hemifield? ("Reverse Posner")

**Predictions:**
- Smaller benefit for valid cues
- Patients should be hurt less than controls by invalid cues.

### Inhibition of Return

- Typically, target detection is faster on trials with valid cues.

\[ \text{Valid cue} \]

- However, if the cue is presented for a longer time (eg. 500 ms),
  - Performance is better on invalid vs valid trials.

\[ \text{Invalid cue} \]

The model explains the basic finding that valid cues speed performance. However, if the cue is presented for a longer time (eg. 500 ms), performance is better on invalid vs valid trials.
Simple model: too simple?

- Has unique one-to-one mappings between low-level visual features and object representations (not realistic)
- Does not address issue of spatial attention when trying to perceive multiple objects simultaneously

"Complex" model combines more realistic model of object recognition (starting from LGN) with simple attention model

→ Can use spatial attention to restrict object processing pathway to one object at a time, enabling it to sequentially process multiple objects.

Lesions of entire spatial pathway cause simultanagnosia: inability to concurrently recognize two objects.

General Issues in Attention

- Attention:
  - Prioritizes processing
  - Coordinates processing across different areas
  - Solves binding problems via coordination

But attention should be much more flexible than just spatial bias!

Later: how to incorporate goals, reinforcement, probability into attention allocation