Recognizing objects: Part I

Ames room – optical illusion

Briefly revisiting MVPA and RSA

Multi-voxel pattern analysis

- Allows for the identification of *representative patterns* for categories, visual objects, cognitions, etc. within a brain region.
- This information can be used in different ways:
  - ***Classification Analyses***

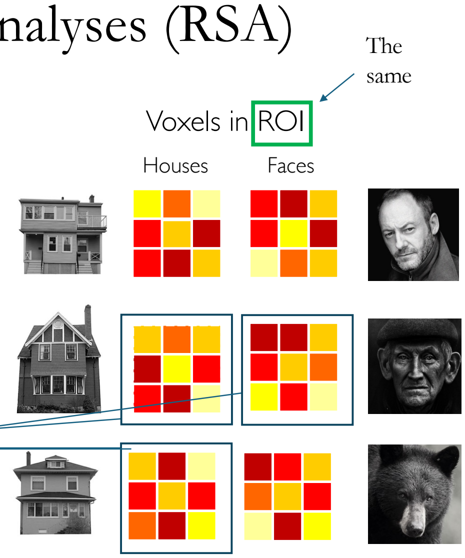
  -> Can the brain (region) dissociate this information?
    - Searchlights, ROIs, RFE
  - ***Representational Similarity Analyses***

  -> What does the brain group together?

  OR What shares a representational space?

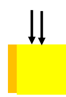
Representational similarity analyses (RSA)

- What is a similar representation?
  - When the voxel pattern evoked by two stimuli is correlated



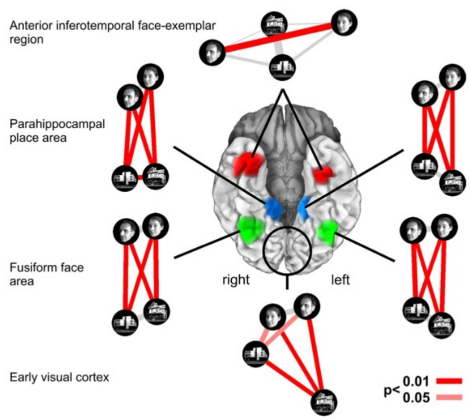
Correlation of these two matrices: $r = 0.8$

Correlation of the two matrices: $r = -0.2$



Representational similarity analyses (RSA)

- A comparison of patterns of voxels at different time points in the task
- Evaluate how much the representations correlate with each other



Digby learns a lesson

1/21/25 PSYC 365: Class 5 5

Q: So what happened there? Algae is not grass! That was an error in perception. To Digby's peril.
Seeing things in the world: We take it for granted. Our intuition is that the world is there, our eyes are like cameras, and we record the world accurately. But it's not that simple. And the more you think about how hard it is to actually recognize objects the way we do, it's actually surprising our visual systems don't deceive us us all the time. And these challenges in effectively recognizing objects in the world and acting appropriately are not particular to dogs and people. They're also relevant to AIs, and particularly relevant as companies try to develop

self-driving cars that can recognize objects well enough not to kill people.

Self-driving cars learn a lesson!

https://www.carscoops.com/2023/09/over-a-dozen-cruise-robotaxis-get-stuck-in-dystopian-traffic-jam-in-austin/

Q: So what happened there? Algae is not grass! That was an error in perception. To Digby's peril.
Seeing things in the world: We take it for granted. Our intuition is that the world is there, our eyes are like cameras, and we record the world accurately. But it's not that simple. And the more you think about how hard it is to actually recognize objects the way we do, it's actually surprising our visual systems don't deceive us us all the time. And these challenges in effectively recognizing objects in the world and acting appropriately are not particular to dogs and people. They're also relevant to AIs, and particularly relevant as companies try to develop

self-driving cars that can recognize objects well enough not to kill people.

Learning objectives (2 lectures)

- Appreciate our amazing capacity for object recognition
- Describe the roles of specific visual cortex systems in object recognition
- Discuss the contribution of classic and more recent fMRI research to our understanding of object processing stages in humans
- Evaluate challenges posed to computer vision

Object Recognition: Part 1

1. Challenges to object recognition as seen through the lens of machine learning
2. How does mother nature do it?
    1. Basics of human visual systems

Why has it been hard to teach computers to see?

Stop around 13 min, after maps. Since this was made CNNs have gotten faster & more efficient, better at recognizing moving objects in entire scenes in real time. A lot of course has been driven by work on self-driving cars.

How does Mother Nature do it?

Ok – when it comes to object recognition neural networks are getting there, and getting better all the time . But how do our brains handle object recognition? Do they face some of the same challenges as computers? To get start to answer that question let's look in a bit more detail at how the visual system works. As a reference, Figure 5 at the beginning of Chapter 2 in the Passingham textbook gives a general outline of how information travels in the visual system once it reaches the cortex. But first it has to get from the eyes to the cortex via the retina-geniculate-striate system .

Goals of Video

- Describe key structures in retina-geniculate-striate system
- Explain how neural signals arrive in both ipsilateral and contralateral cortex
- Describe receptive field properties of lateral geniculate nucleus (LGN) and V1 neurons
  - Receptive field: Region of sensory space (e.g., body surface or visual field) where a stimulus will influence a neuron's firing pattern

Pathway from eye to cortex

1/21/25 PSYC 365: Class 5 12

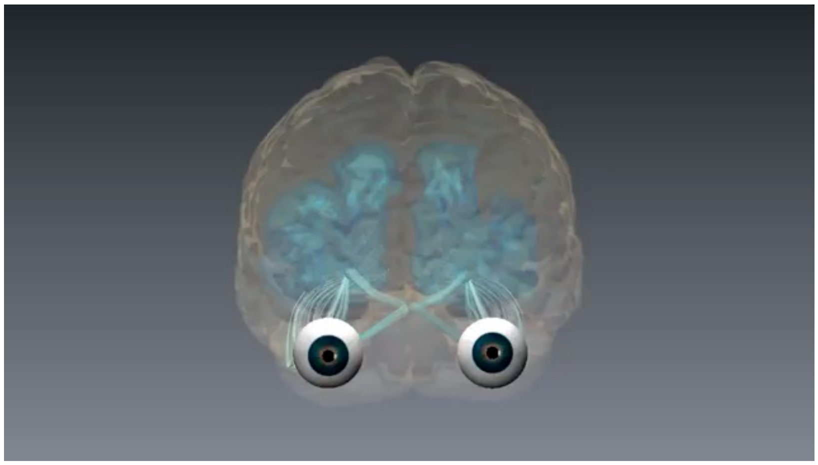
Cut it off after cutting through optic chiasm

Summary

- Information leaves retina via optic nerve
- Information from the nose half of each retina crosses to the other side of the brain via the optic chiasm
- Ipsilateral = same side
- Contralateral = opposite side
- Information from the outer half of each retina travels ipsilaterally
- It is then relayed via the lateral geniculate nucleus of the thalamus...
- ...To the primary visual cortex (V1), which is where visual information first enters the cortex

Here are the points I want you to take home from that video. 1, 2 *Ipsi* means self or same in latin. *Contra* means opposite, against, contrasting. Thalamus, as you'll recall functions as a relay station among many other functions. Coming up is a visual to help your review the pathway from eye to cortex. [before v1 cells]And then one piece of additional information you need to know is that ... We'll talk more about that later.

Review: Signal path from eye to cortex



1/21/25 PSYC 365: Class 5 14

Cut it off after cutting through optic chiasm

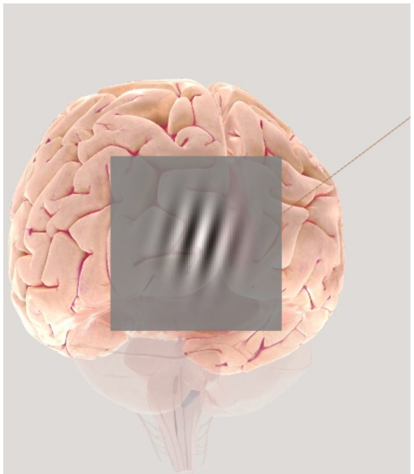
V1 mapping

Image on retina

V1 mapping

Your Passingham text describes how in 1604 Johannes Keppler discovered that the image on your retina is upside down. This because because the process of light refraction through a convex lens causes the image to be flipped. Each eye receives a 2-d image map on the retina, as you see on the left, and, as you just saw, passes signal on via the lateral geniculate nucleus of the thalamus to the primary visual cortex, or V1. Primary visual cortex is retinotopic. That means it functions as a topographic map of what's on the retina. But in V1 the world is not only upside down upside down, as it is on the retina, but also distorted as you see on the right. In V1, more real estate is given to the centre than the periphery because there are more light receptors in the fovea or centre of the eye, and we take in more information when we fixate our gaze on things. This is like the distortions we saw in the somatosensory and motor homunculi in the neuroanatomy module, where parts we use a lot, such as the hand is huge, and parts we use less such as the inner elbow not so much.

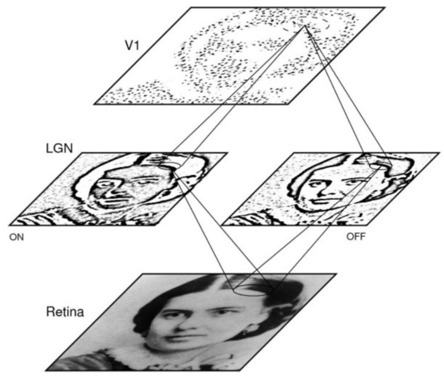
Primary visual cortex (V1) maps



- Where visual information enters cortex
- Retinotopically organized
- Neurons have small receptive fields
- V1 cells respond to specific types of information that allow us to detect contrast, edges, motion direction

V1 retinotopic maps

- Each neuron sees only one point!

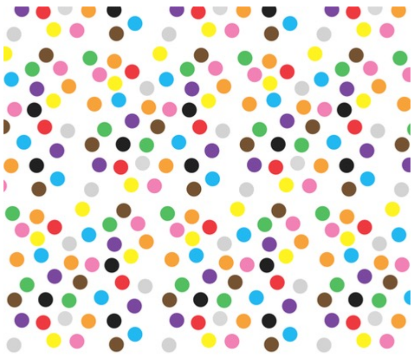


What V1 sees

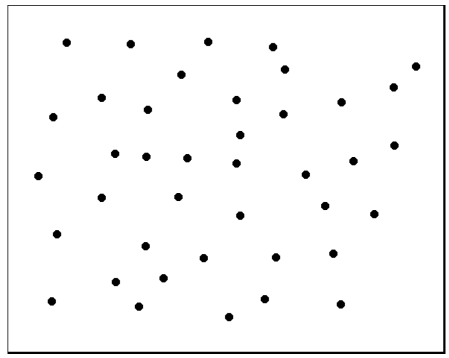
Mind you this SHOULD also be upside down. This is the whole visual field so each neuron is just seeing a small piece of this. ? Then what happens? From V1 Information is passed along further down in primary visual cortex, where you have more areas called V2, V3, V4, MT, which independently code for things like motion & colour.

V2-V4 & MT

V4

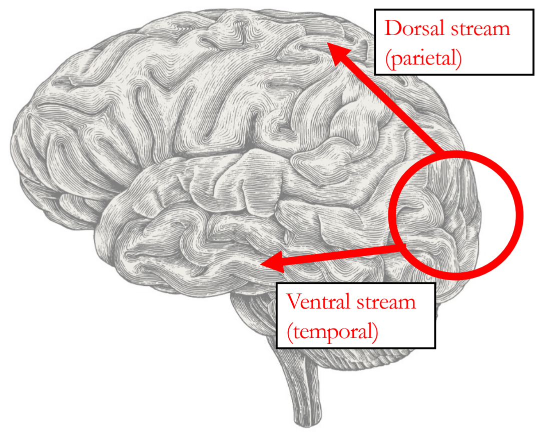


MT



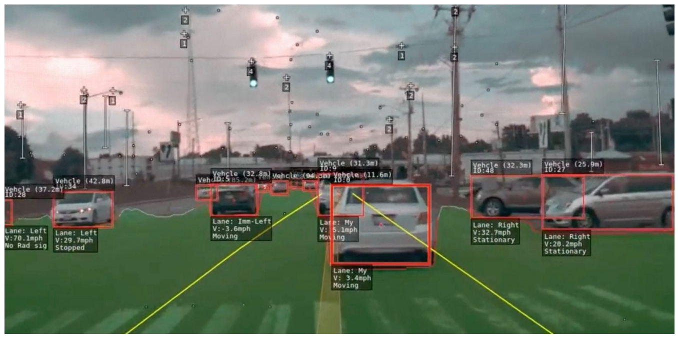
Feedforward sweep of processing

- Information travels from occipital cortex to parietal and temporal cortices
- Early stages process elements
- Later stages integrate them



So here’s the thing:

- If the 2D retinotopic map in V1 has no more information than a photograph, then the problem the brain must solve is the same as the problem a computer vision algorithm must solve.



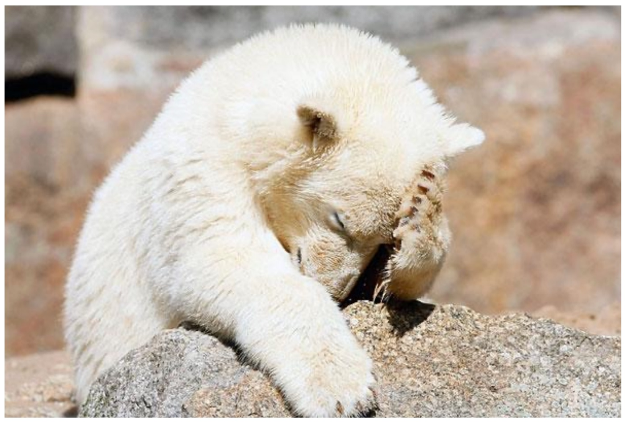
https://ai.plainenglish.io/teslas-next-autonomous-driving-release-will-be-a-quantum-leap-bce415a72c85

Learning Objectives (object processing continued)

- Appreciate our amazing capacity for object recognition
- **Describe the roles of two visual cortex systems in object recognition**
- Discuss the contribution of classic and more recent fMRI research to our understanding of object processing stages in humans
- Compare challenges posed to human and computer vision

To revisit our learning objectives for this and the next lecture, I'm hoping you're starting to get a grasp on the first objective, and are beginning to understand how challenging it is for our visual systems to recognize objects in 3D under all sorts of circumstances given the impoverished 2D maps in V1. We're now going to look at what we know about the roles of the dorsal and ventral visual streams in object recognition. In our methods modules we talked about how some of our most foundational knowledge of how the human brain works came from studying lesion patients. So to start we'll look at some of with what we've learned about visual pathways from studying lesion patients.

1a. What have we learned about object perception from lesion patients?

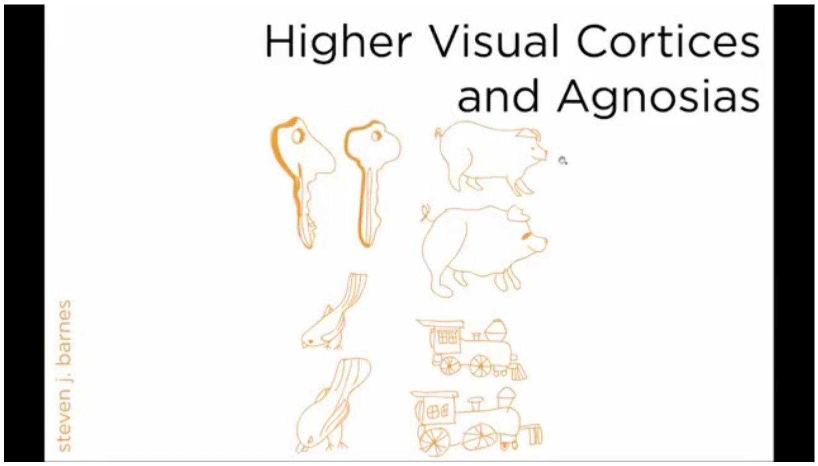


We're going to start with another video by Professor Barnes.

Goals of Video

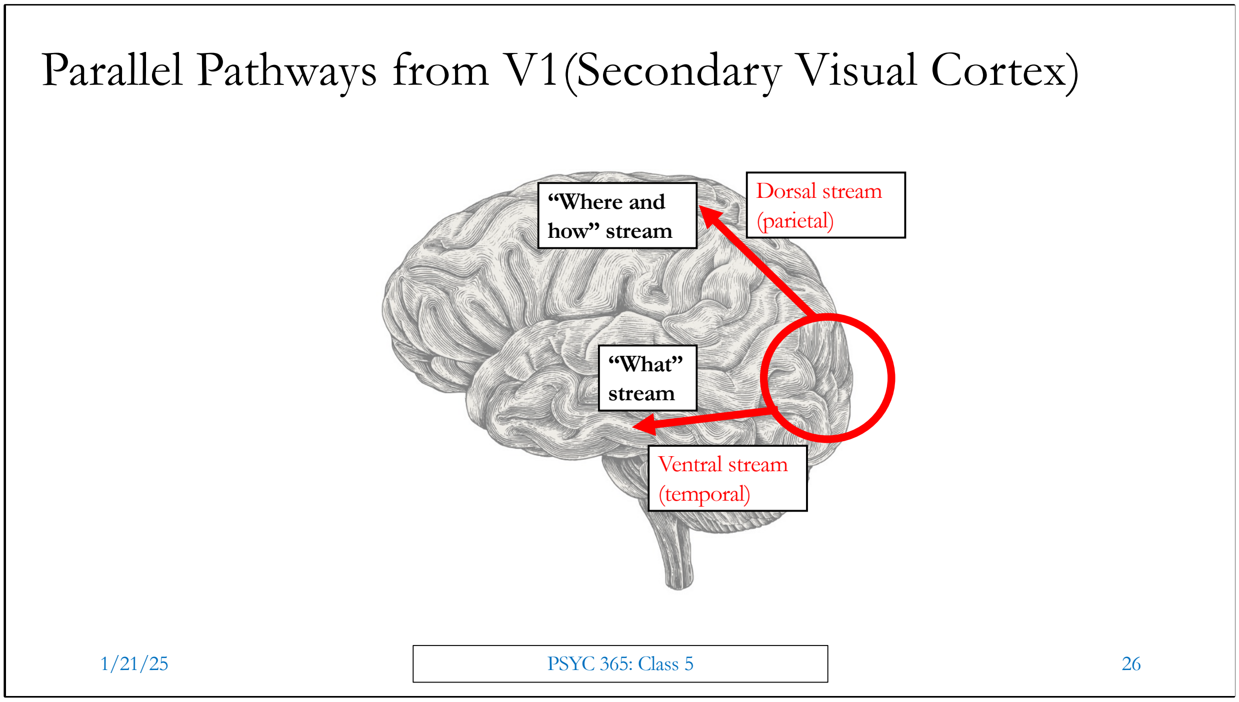
- Review of the difference between primary visual cortex, secondary visual cortex, and association cortex
- Introduction to theories of functions of dorsal and ventral streams
- Describe the difference between apperceptive and associative agnosias

Higher Level Vision



1/21/25 PSYC 365: Class 5 25

7 min. Apperceptive agnosia is kind of like always having the problem a CNN has with the cat before it's been trained to recognize all kinds of different angles and shapes of cats. It's very literal and small scale, like it's stuck in the smaller receptive fields of earlier areas of visual cortex, where the information to higher level regions of the ventral stream, where objects can be recognized from different viewpoints and when theye are obscured, happens.

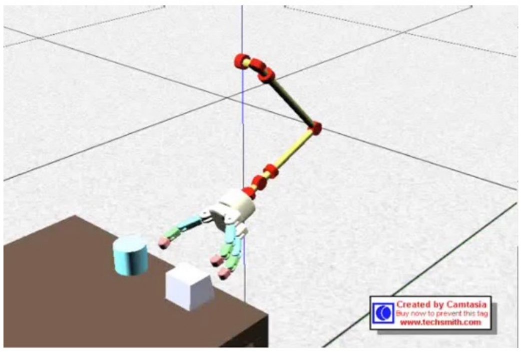


The dorsal and ventral streams are two parallel visual pathways branching out from primary (orange) to cortex into the temporal and parietal lobes. Goodale & Milner identified may of the functions of these streams by studying their patient Dee. Dee had brain damage to occipital cortex from carbon monoxide poisoning (from a faulty water heater) while taking a shower. She couldn't describe things like the orientation of a pencil or the width of a block. But she could USE a pencil and grab a block. Her dorsal stream regions were intact and were activated when Dee has to reach for and grasp an object in the fMRI scanner.They concluded that activity in the dorsal stream allows you to use the shape and orientation of an object to guide the motion of your hand when reaching for it. The take home was that there are parallel pathways for visual perception in the brain, one specialized for perception of shape and one specialized to use shape to guide action. Passingham suggests that each stream's unique function is a result of their connections. The dorsal stream has direct connections to frontal regions that control movement. So its set up to use vision for guided action. Passingham tells us the parallel architecture of the brain is ideal for speedy processing.

Parallel Pathway Functions

Dorsal Stream

- Planned, object-directed action



Ventral Stream

• LOC uses shape (and other) information to identify an object

Dorsal stream. It's very specific. It's not about ALL action. Guided action directed at objects. Again, uses shape and orientation to guide the action of reaching and grasping a block. In contrast, the ventral stream allows us to piece the bits that V1 perceives into recognizable objects that can be named.

Visual agnosia: Case study (5 min. video)

Do you think this is a case of apperceptive or associative agnosia, or neither, or both? Why?

Summary from Lesion Studies

- The visual system is divided into two parallel pathways
- Each has a distinct function
  - Dorsal: Vision for directed action
  - Ventral: Vision for object identification
- Damage to different regions of the ventral stream lead to specific agnosias
  - Apperceptive
  - Associative

Two Ventral Stream Agnosias

Associative

- Intact perceptual representations, inability to recall associations
- Can't recognize a bear when it's right in front of you
- Can't recognize the trees in the forest
- Can't tell a hawk from a handsaw (that's Shakespeare btw)

Problem with knowing

Apperceptive

- Failures in perceptual processing, for objects, words or faces:
- Can't pass the reCAPTCHA tests where you type in a code to prove you're not a bot
- Can't recognize a chair from below
- Can't recognize the forest for the trees

Problem with seeing

Both are ventral. Associative. You can perceive objects but you don't know what they mean. Apperceptive – had trouble with unusual views, degraded images, global processing (better at detail than gist) Goal for next bit of lecture: build enough knowledge of what the ventral stream does to make a good educated guess about where the damage would be for each.

Reading question

Neuropsychology research by Goodale and Milner revealed a crucial role for the ventral visual stream in:

a. Using object shape and orientation for directed action

b. Using shape to identify objects

c. Feeling the sensation of touch

d. All of the above

Ans b. Q:What stream carries out the role described in A?

Questions?

See you next class!

b. Going beyond comparing 2 experimental conditions and “reading” the brain to see what a person is looking at. O