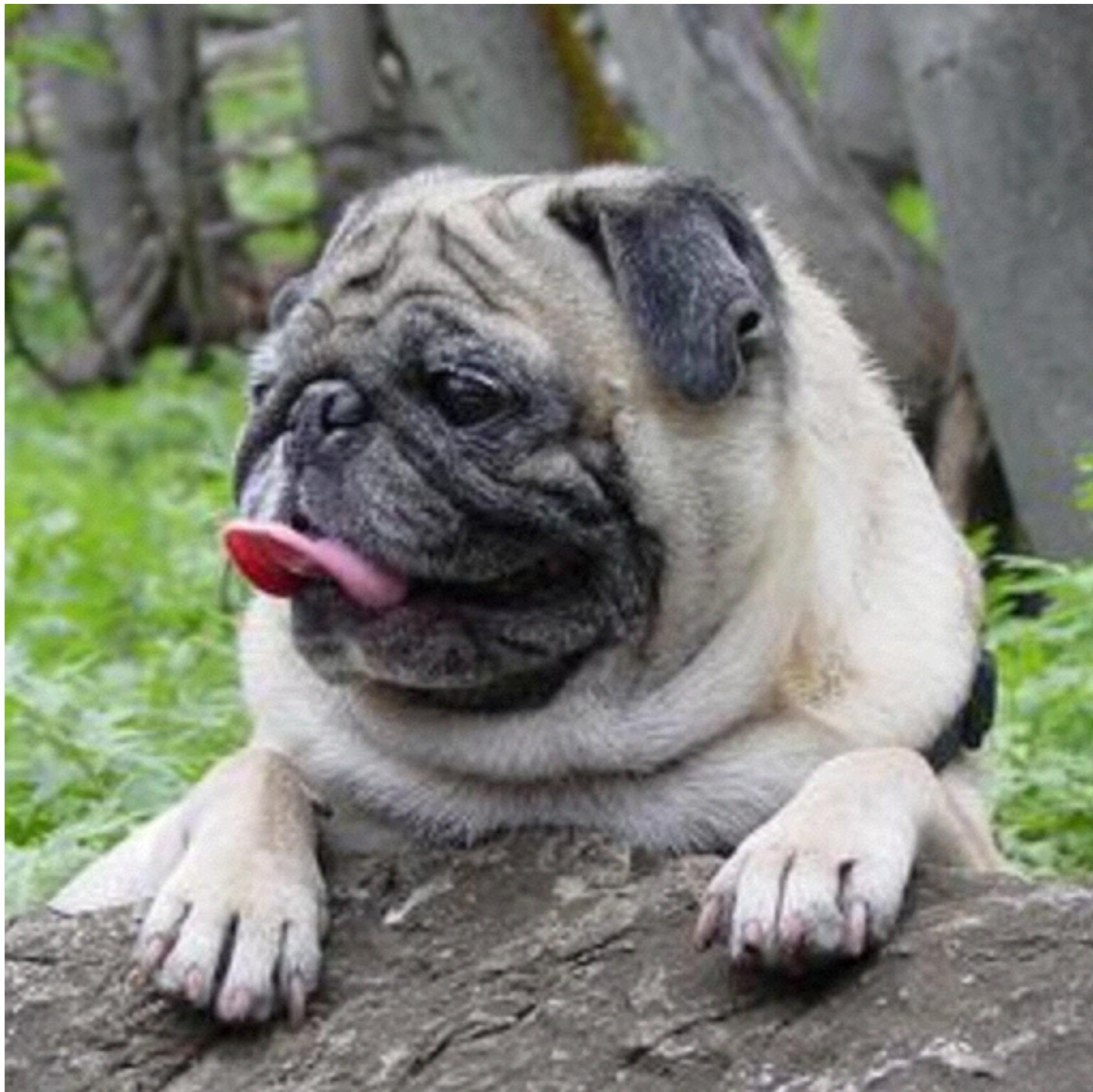
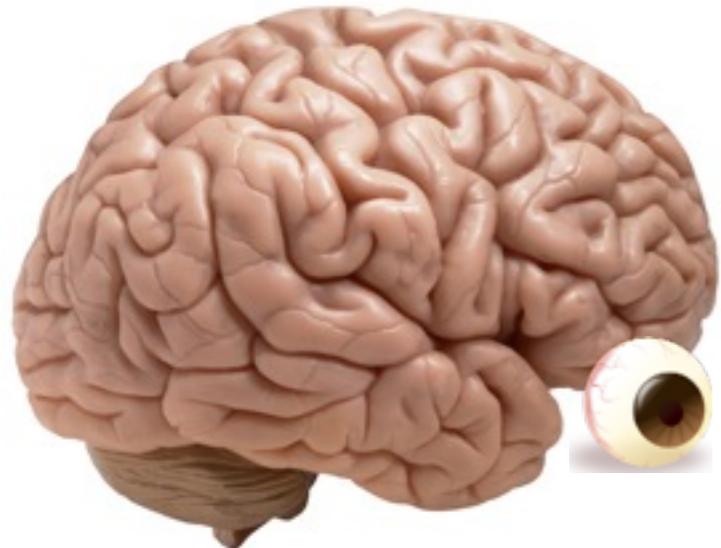


# A Primer on Human Vision: Insights and Inspiration for Computer Vision



Guest Lecture: Marius Cătălin Iordan  
CS 131 - Computer Vision: Foundations and Applications  
27 October 2014





detection  
recognition  
segmentation

...

**visual  
understanding**



# 1. The Mammalian Visual System

- anatomy and processing pathways
- from processing to perception

# **1. The Mammalian Visual System**

- anatomy and processing pathways
- from processing to perception

# **2. Early Computation**

- first stages of information processing
- edges: a basis for representing the visual world

# **1. The Mammalian Visual System**

- anatomy and processing pathways
- from processing to perception

# **2. Early Computation**

- first stages of information processing
- edges: a basis for representing the visual world

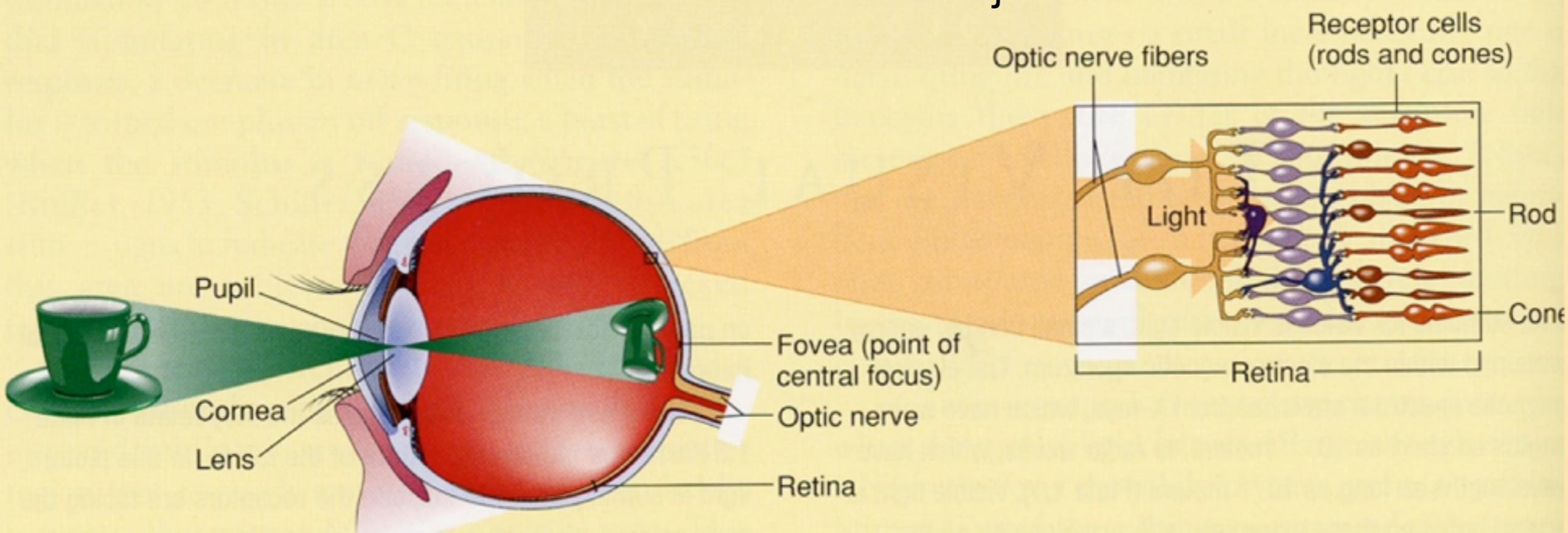
# **3. Object Recognition in the Human Visual System**

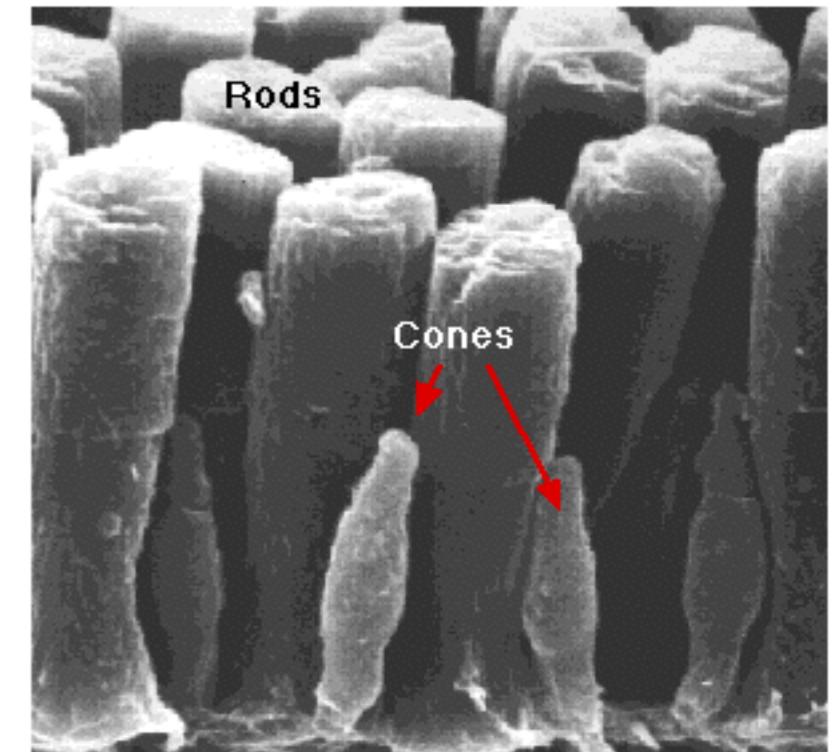
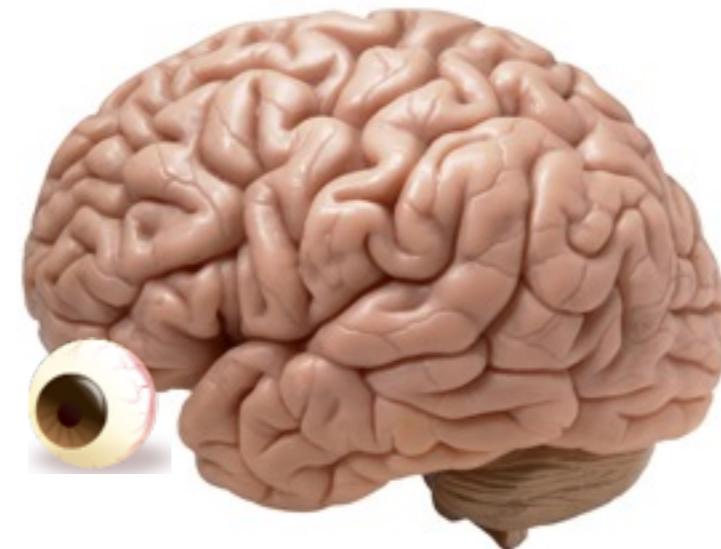
- building invariance: pooling and transformation
- untangling object representations



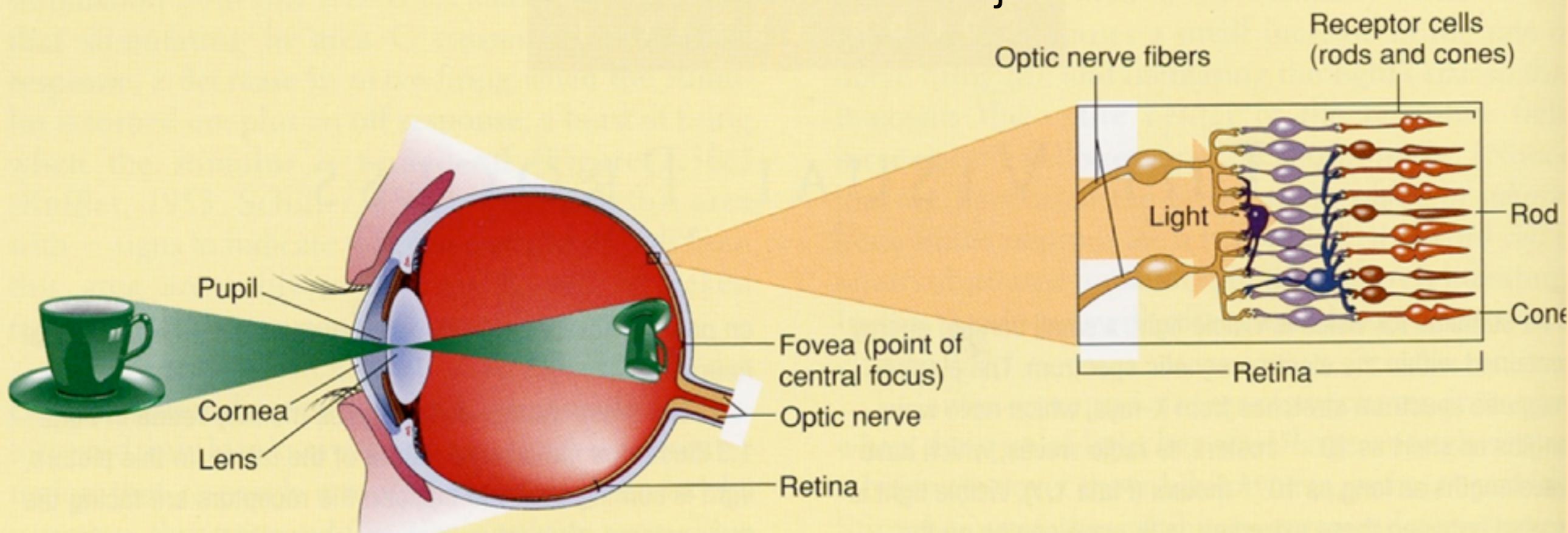


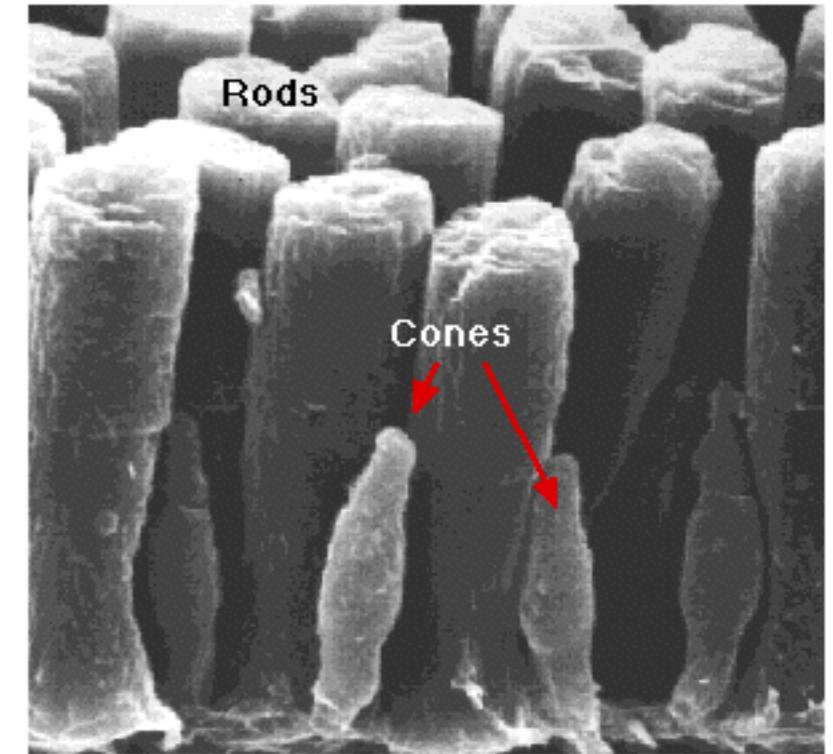
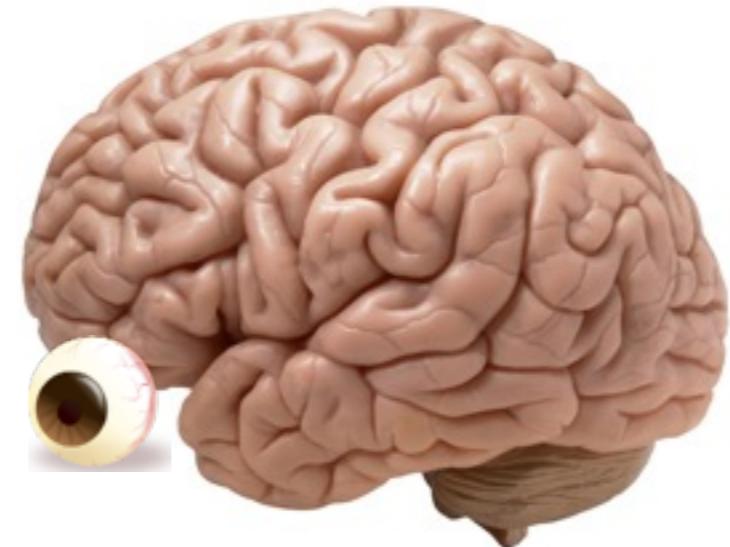
## The “Camera”: Retinal Projection



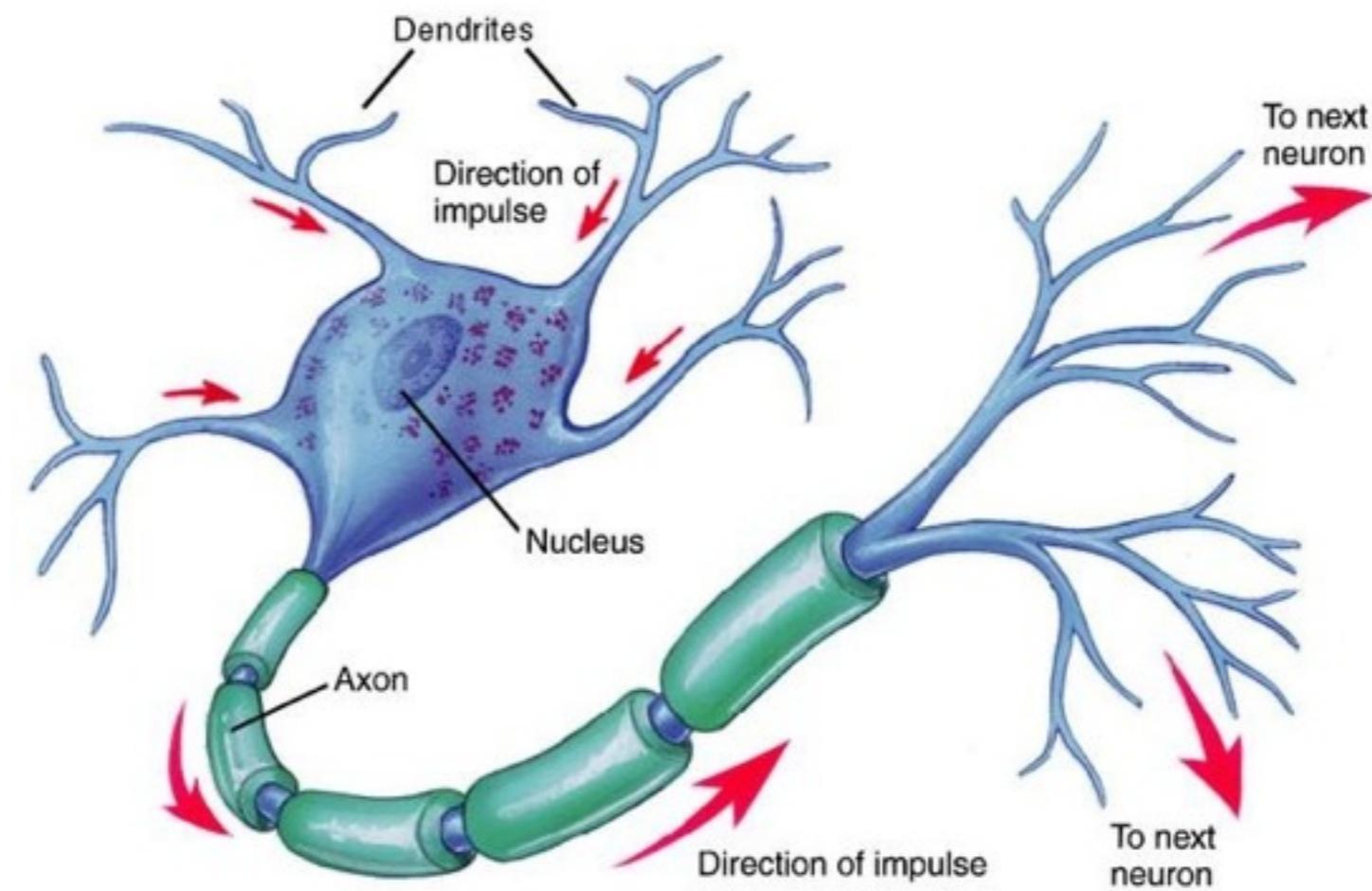


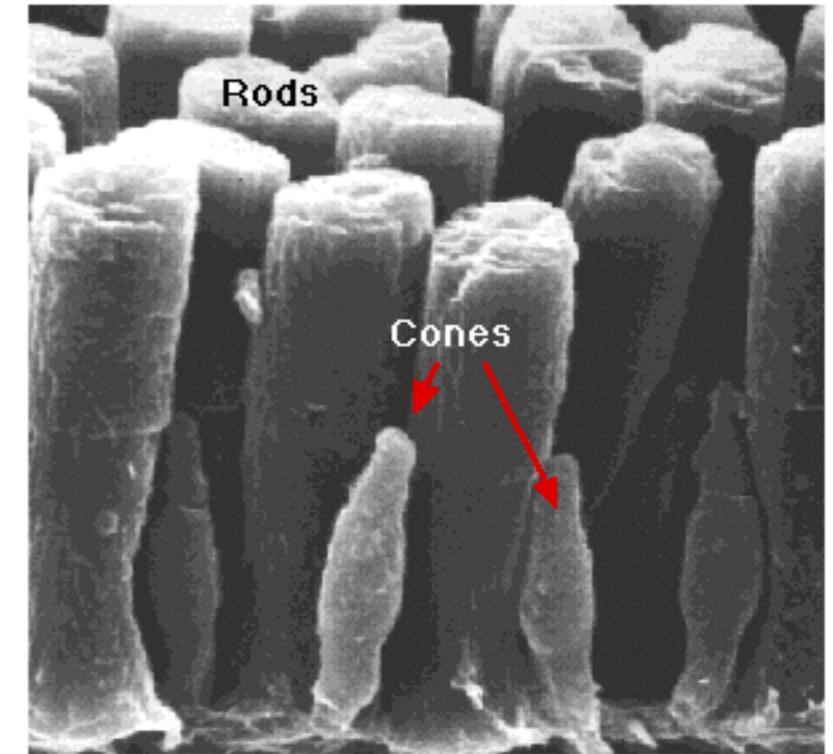
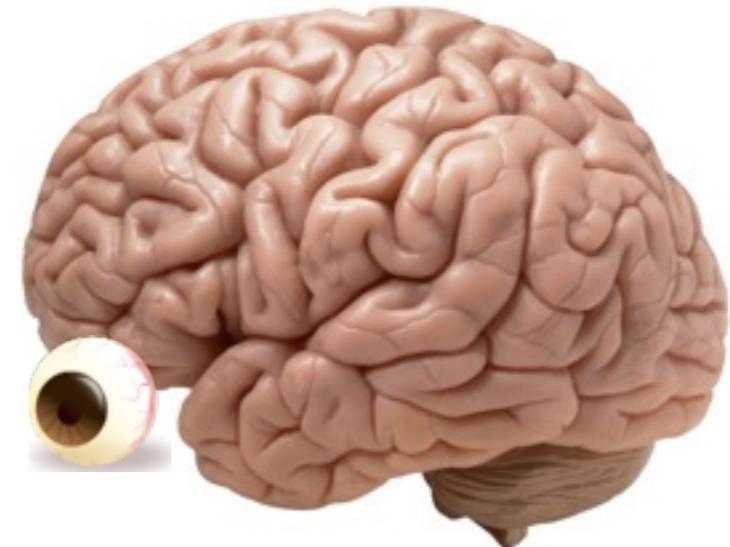
## The “Camera”: Retinal Projection



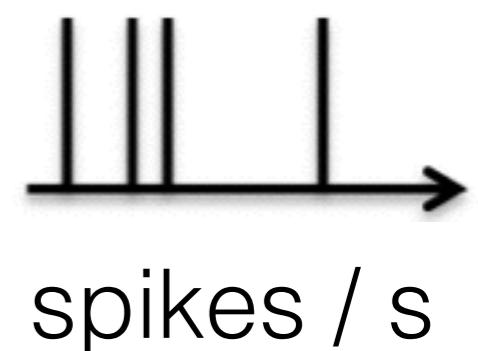
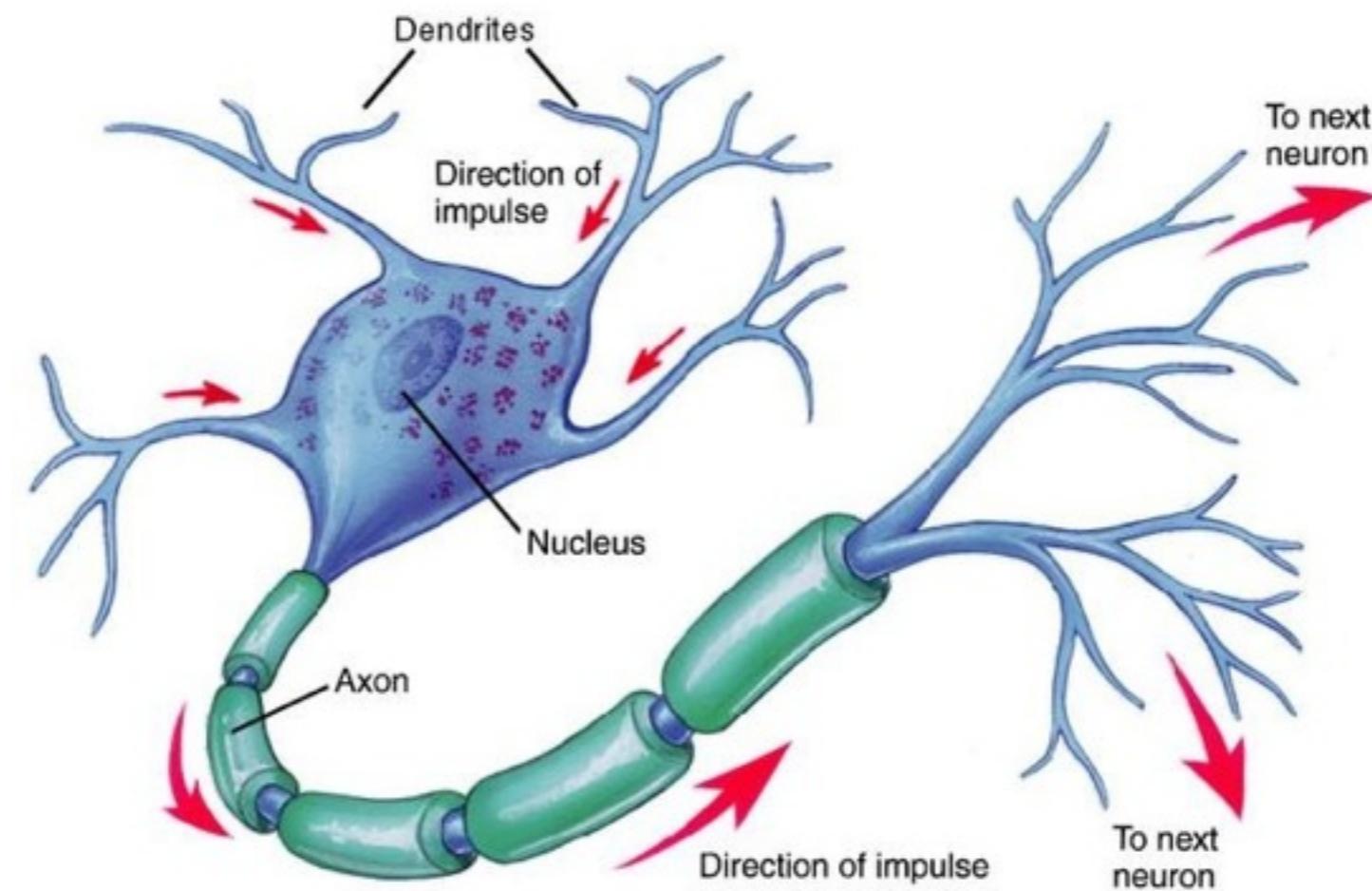


## Neuron





Neuron



spikes / s

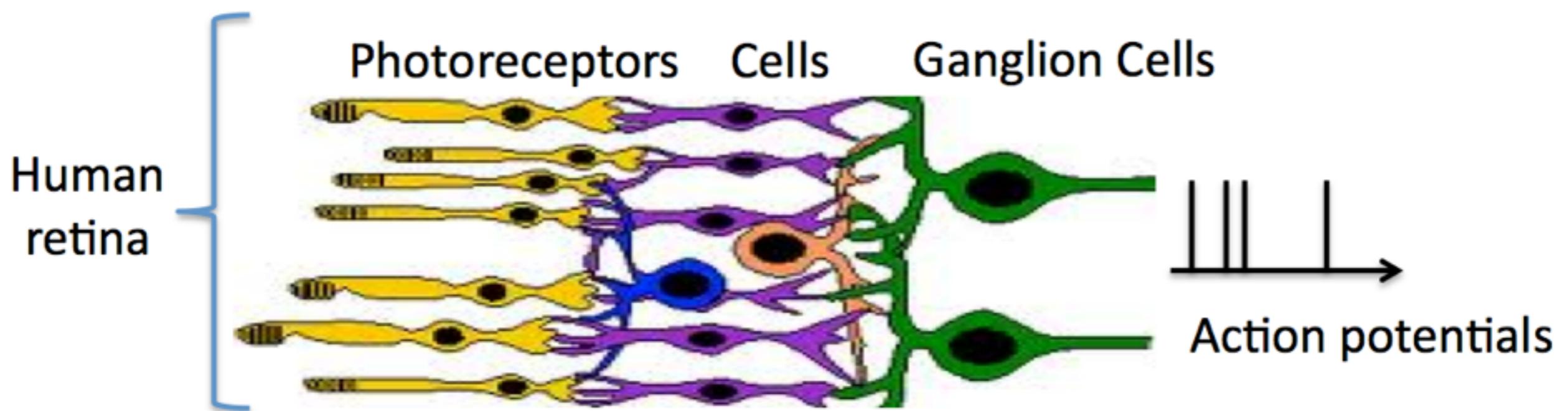
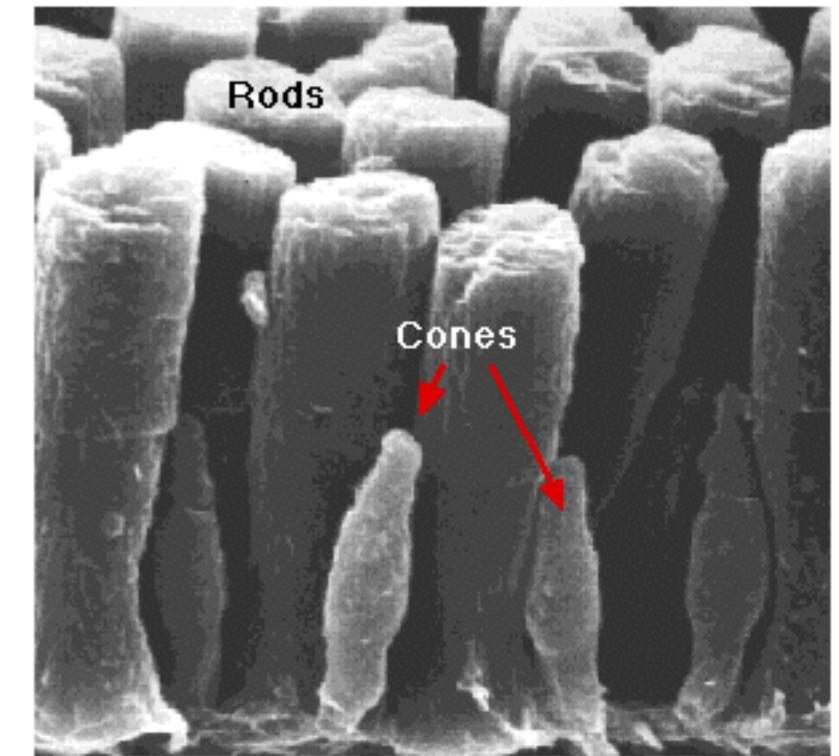
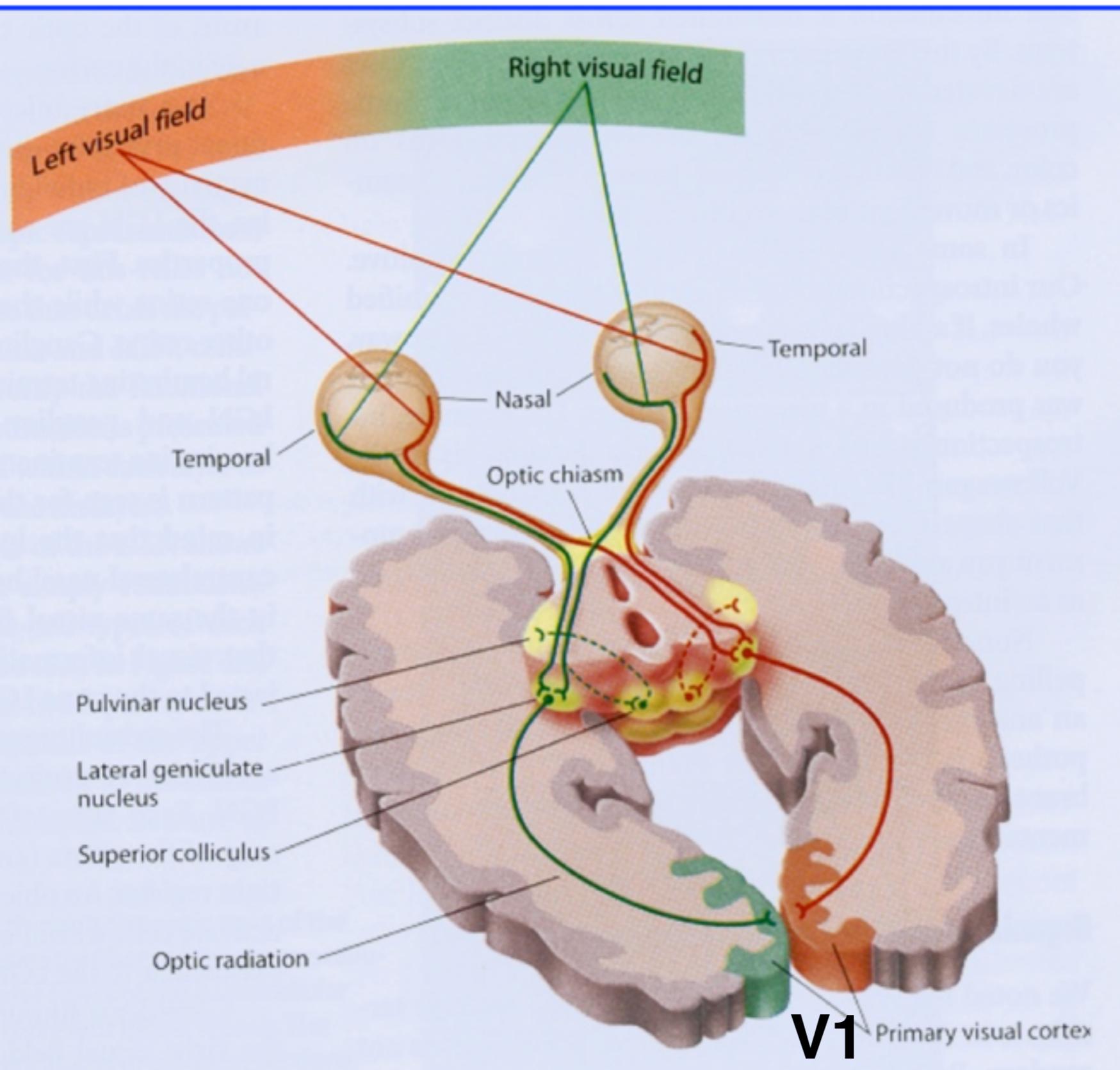


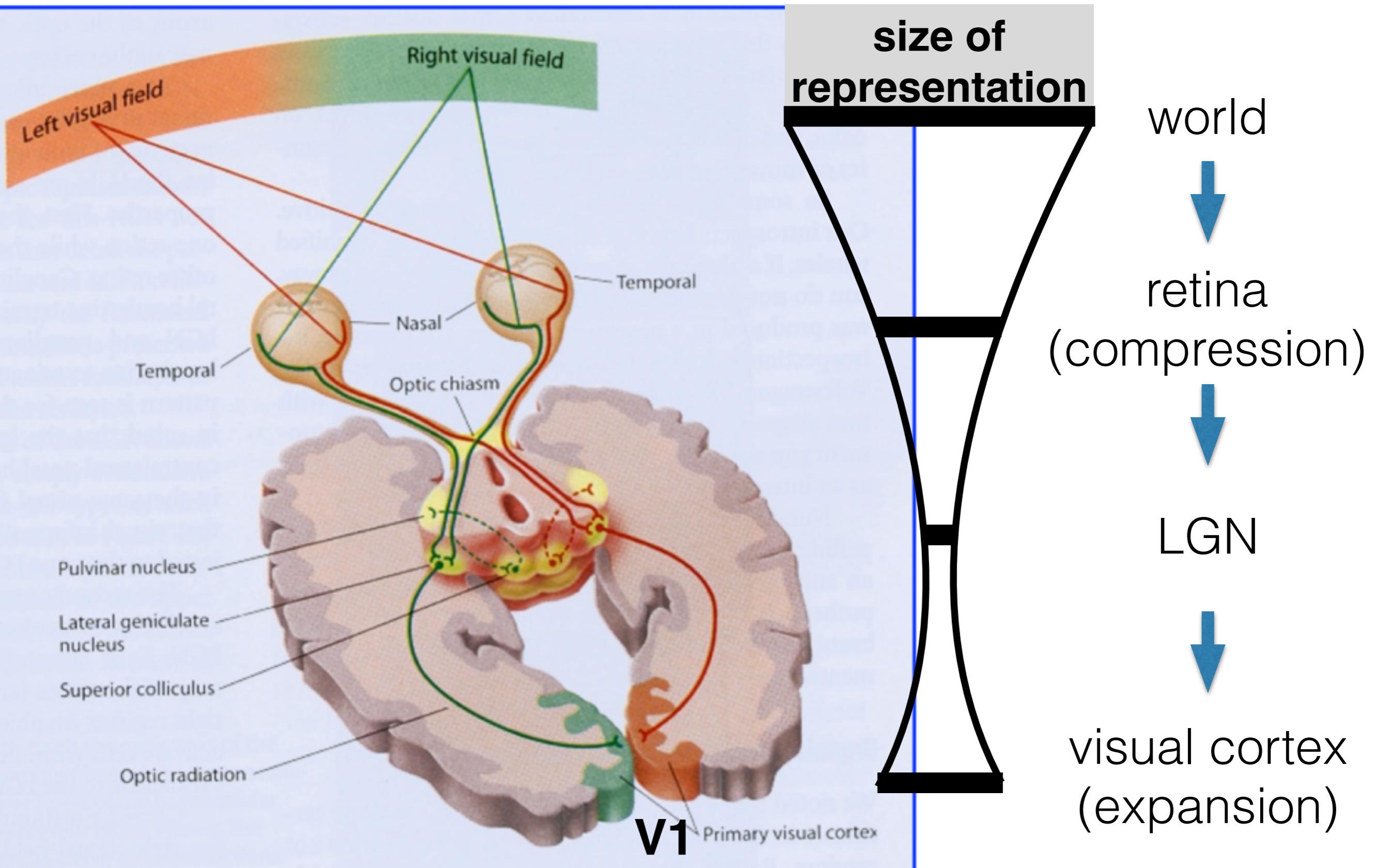
figure courtesy of A. Alahi

# From Retina to Cortex

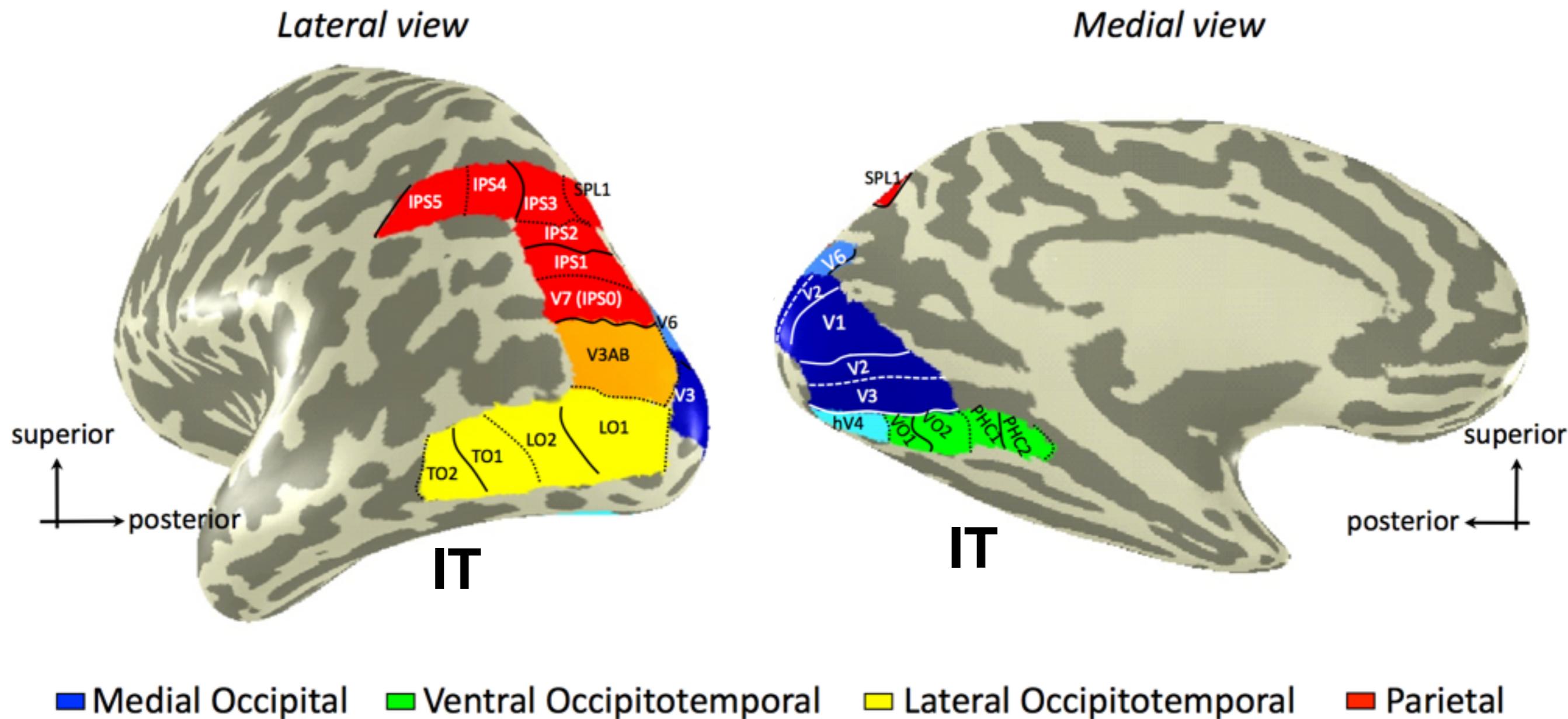


world  
↓  
retina  
(compression)  
↓  
LGN  
↓  
visual cortex  
(expansion)

# From Retina to Cortex



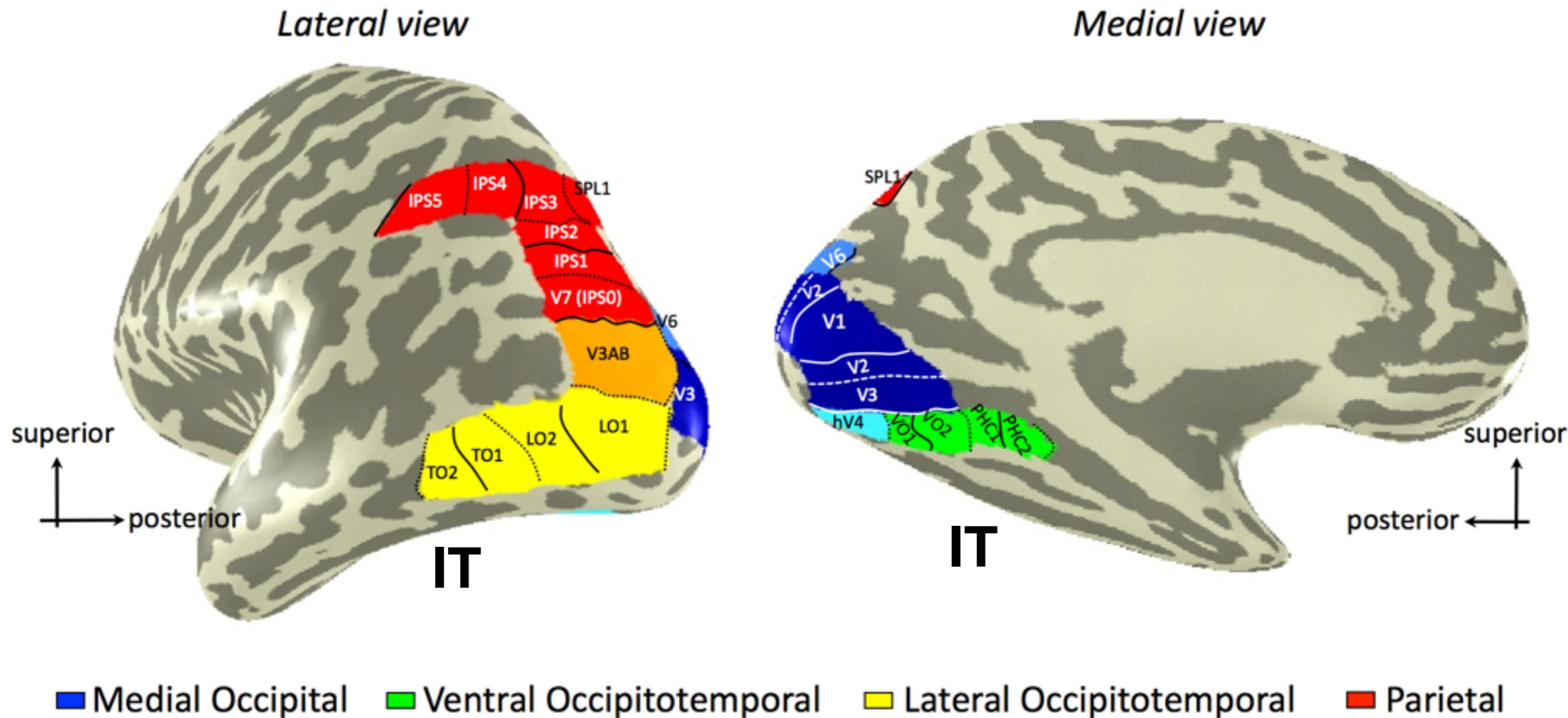
# Cortex: At Least Two Dozen Visual Areas



Weiner & Grill-Spector (2012)

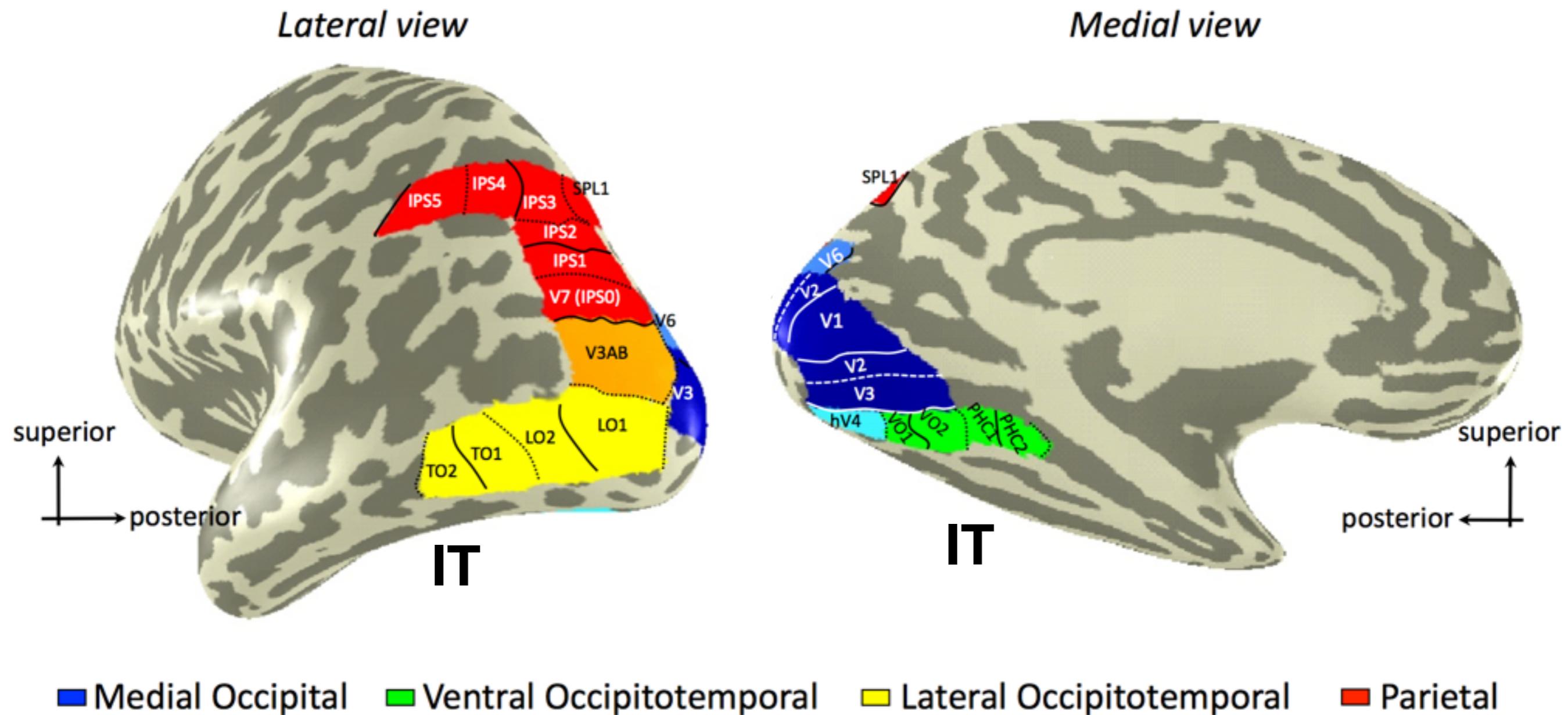
# Cortex: At Least Two Dozen Visual Areas

plus more: motion areas, “functional” regions, etc.



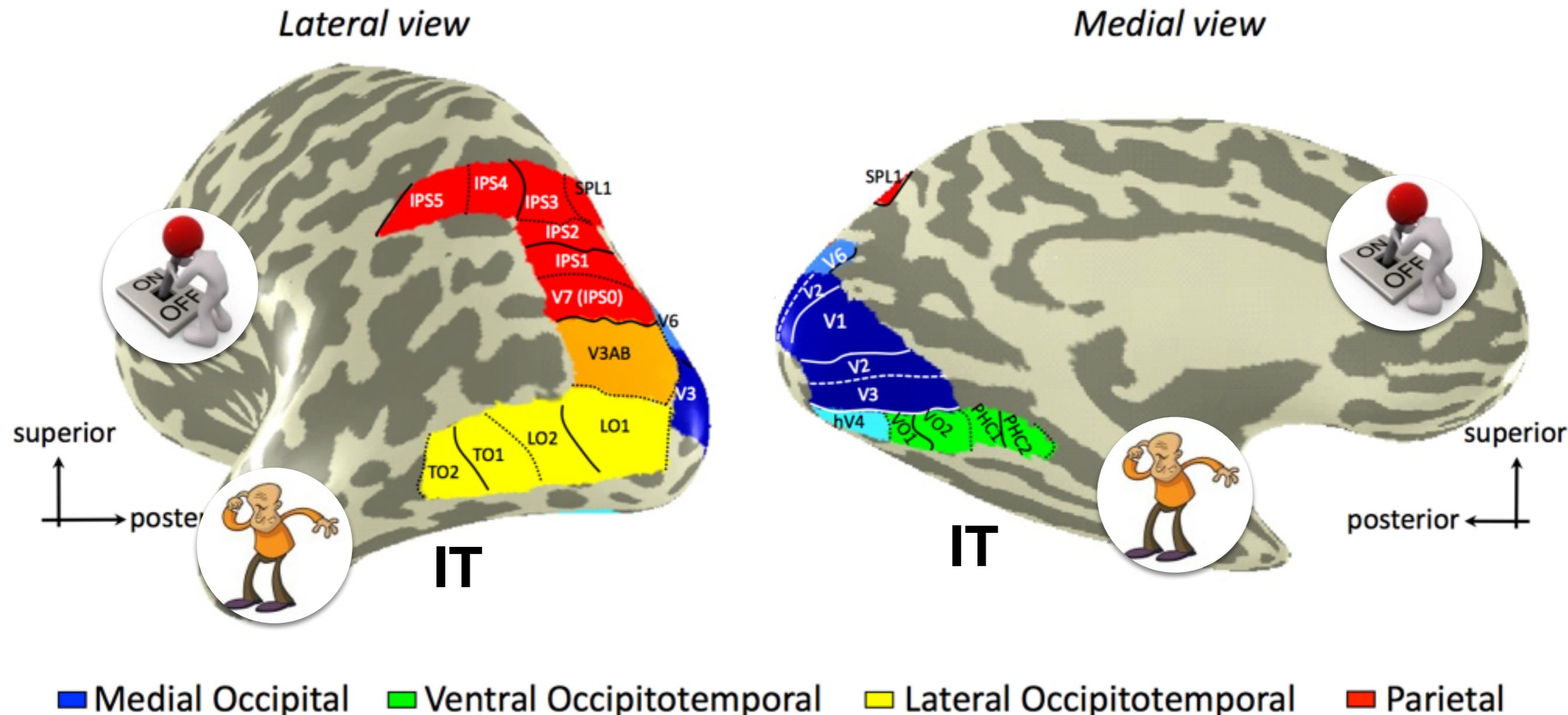
Weiner & Grill-Spector (2012)

# Visual Processing vs. Perception



Weiner & Grill-Spector (2012)

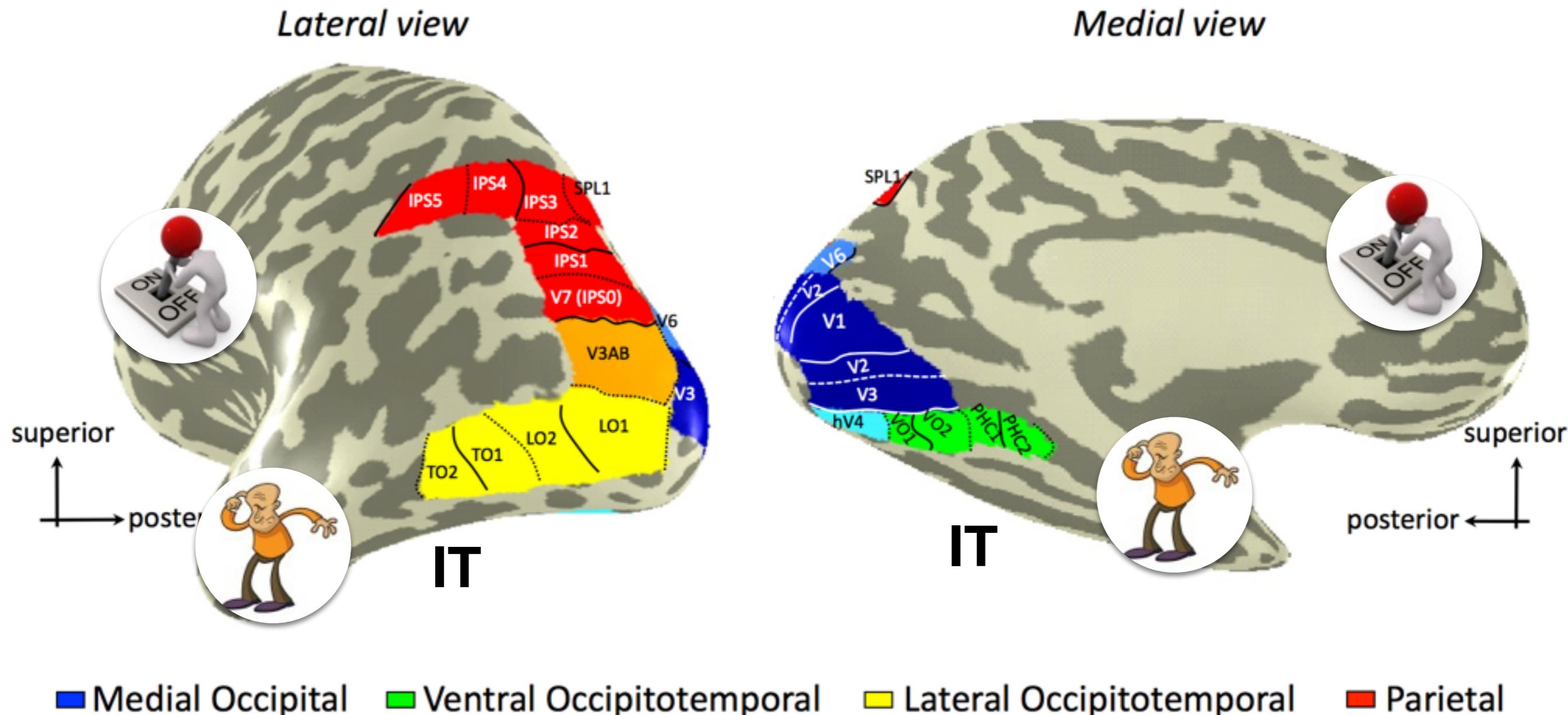
# Visual Processing vs. Perception



Weiner & Grill-Spector (2012)

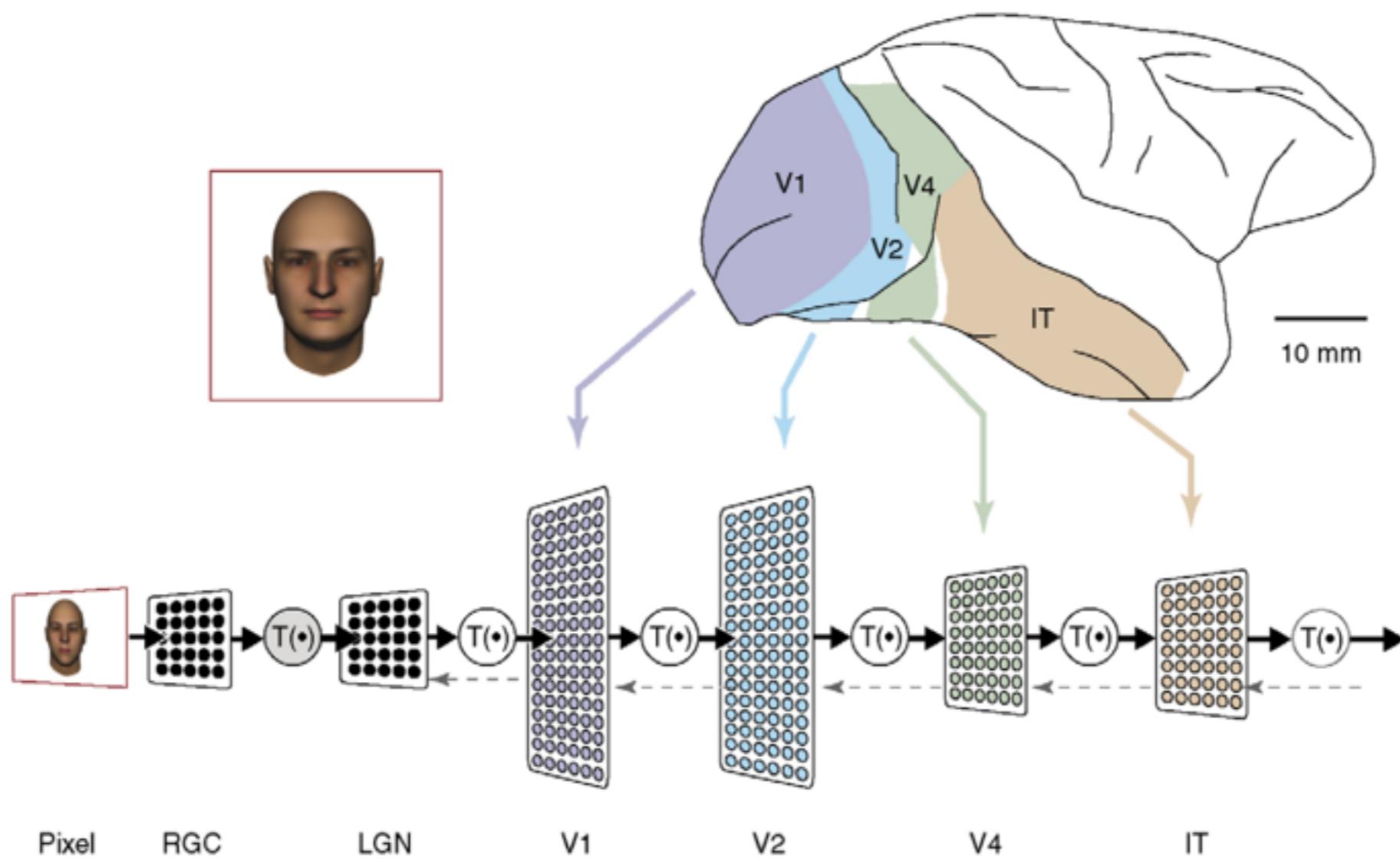
# Visual Processing vs. Perception

perception involves integration and higher functions



Weiner & Grill-Spector (2012)

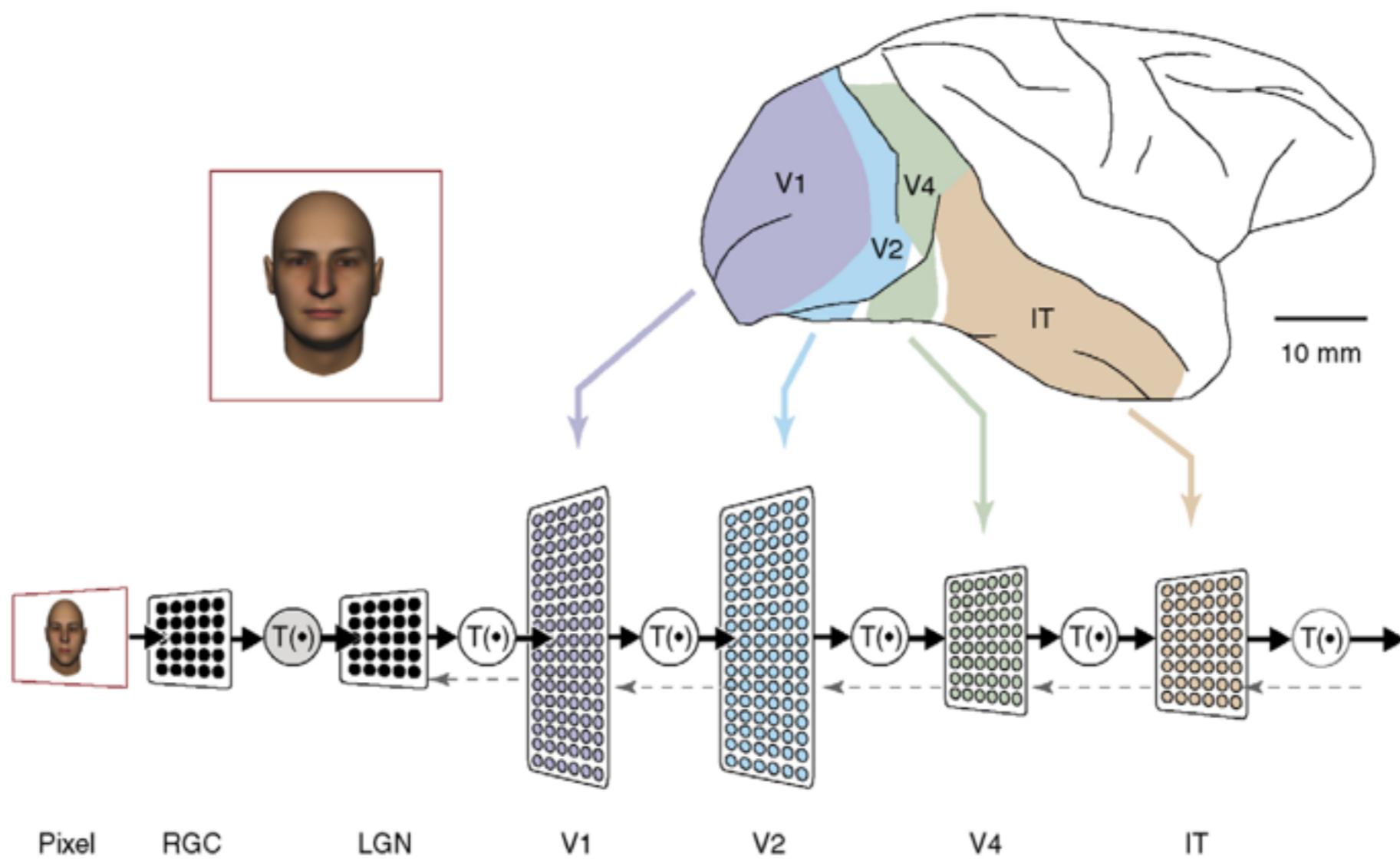
# Mammalian Visual System



DiCarlo & Cox (2007)

# Mammalian Visual System

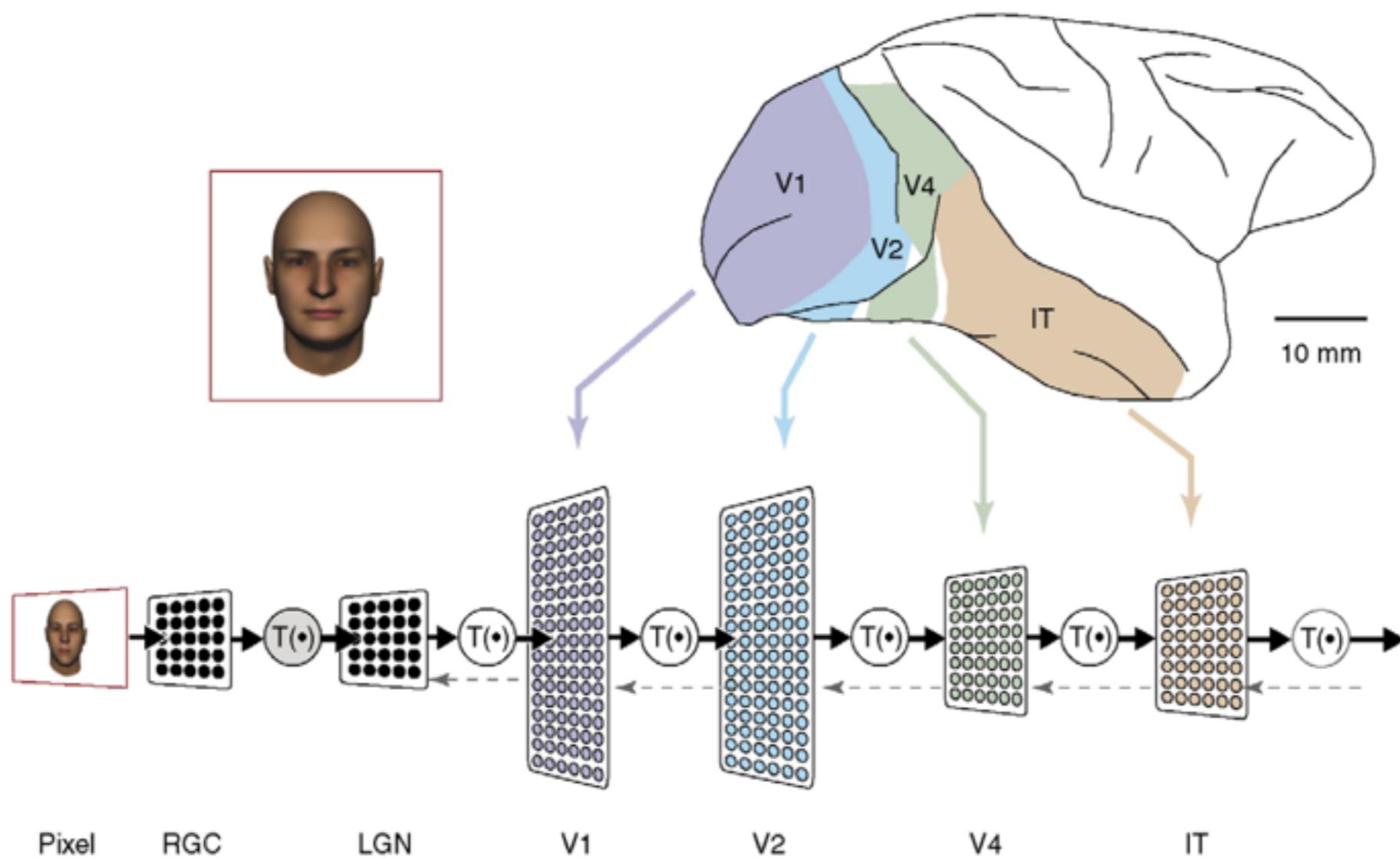
visual processing is done in stages



DiCarlo & Cox (2007)

# Mammalian Visual System

visual processing is done in stages  
but it is not synonymous with perception



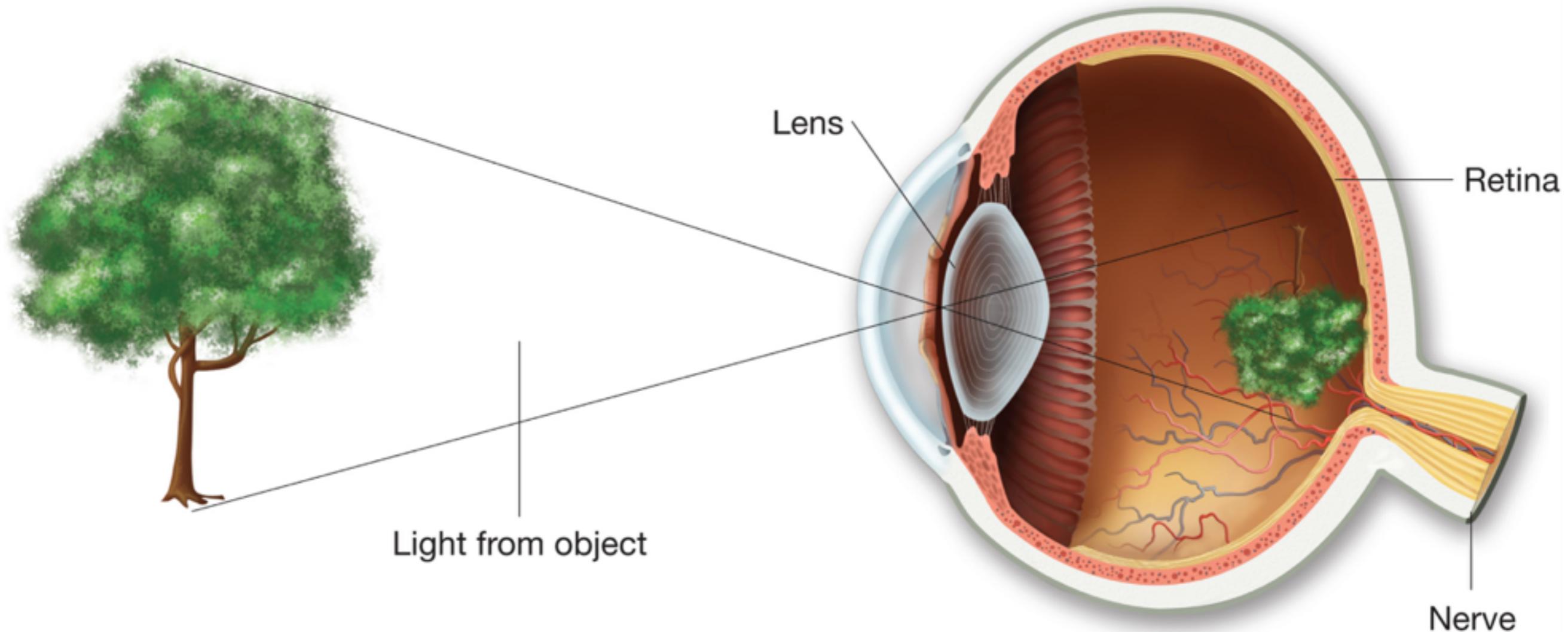
DiCarlo & Cox (2007)

## **2. Early Computation**

local information processing

# Retinal Projection

picture is inverted, but spatial relationships are preserved



# Receptive Fields

extract similar features at each position in the visual field

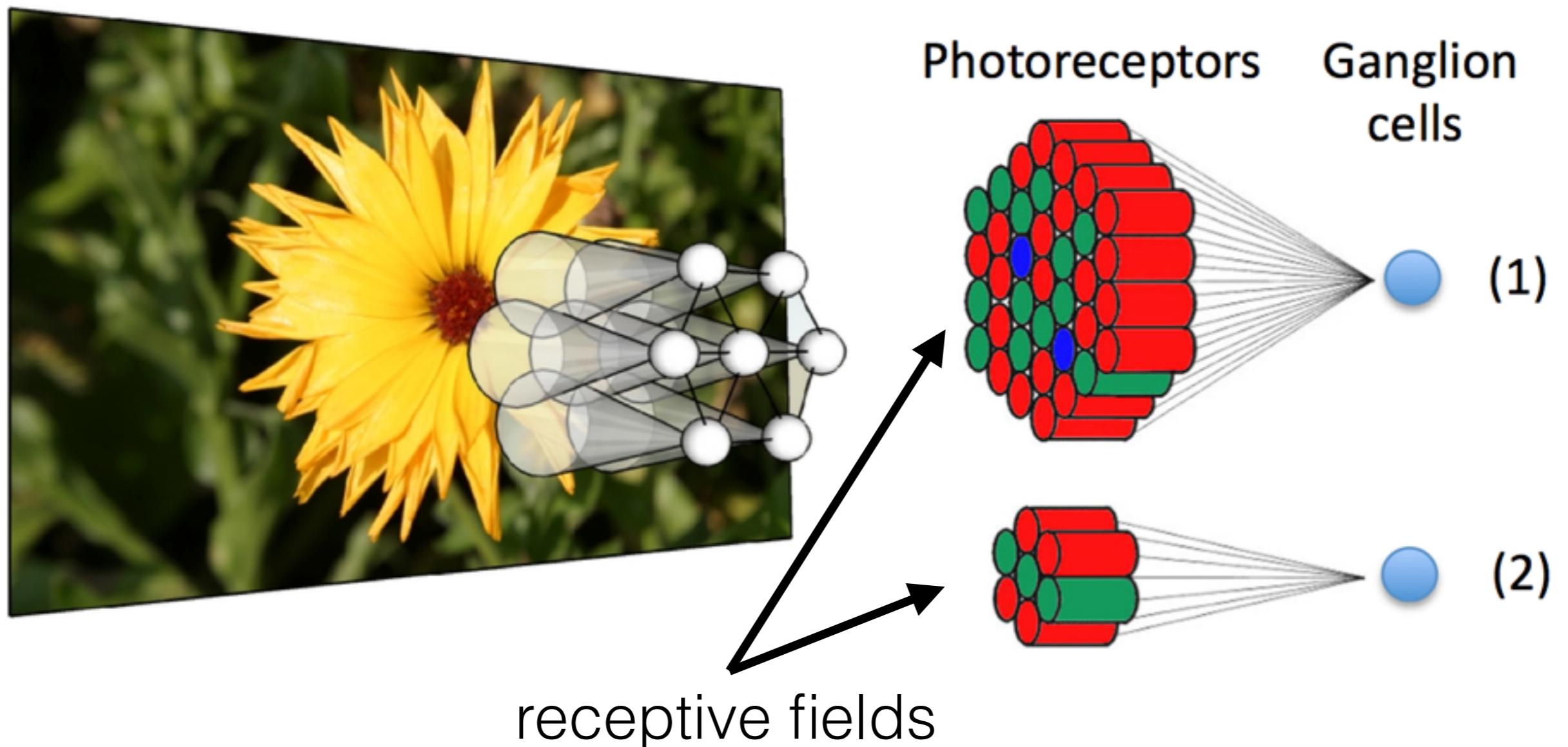
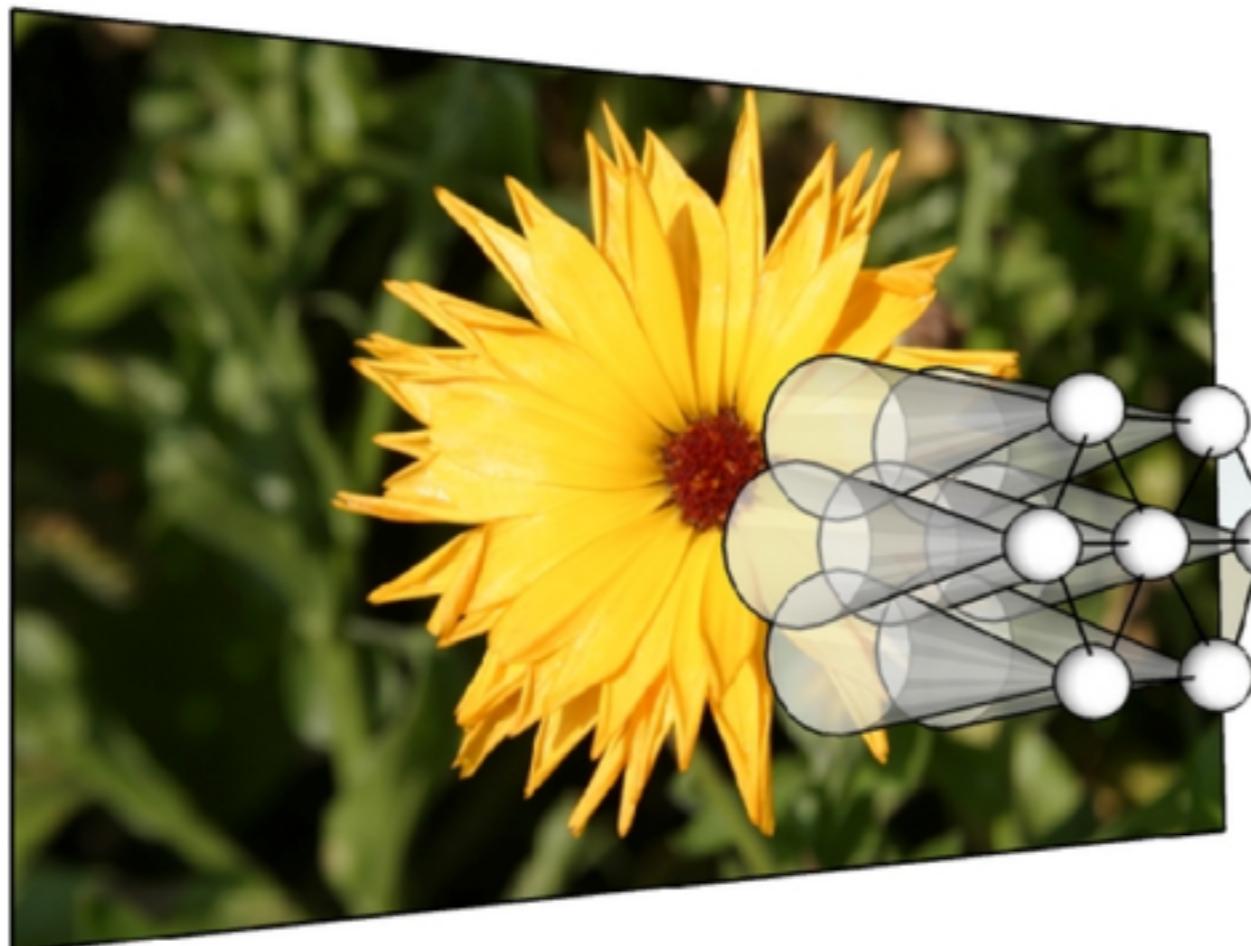


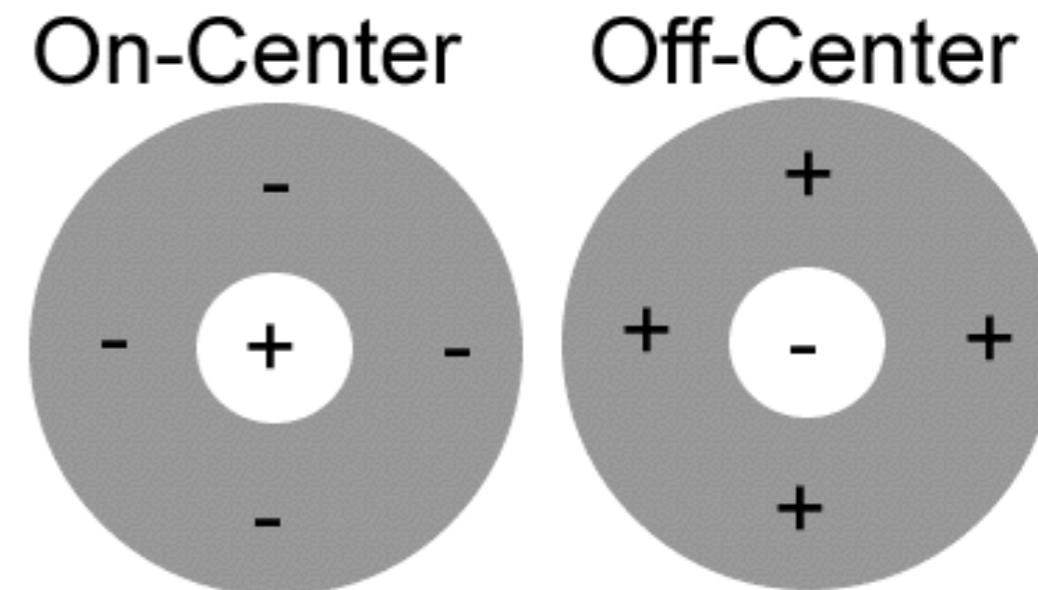
figure adapted from Ebner & Hameroff (2011)

# Receptive Fields

extract similar features at each position in the visual field



ganglion cells



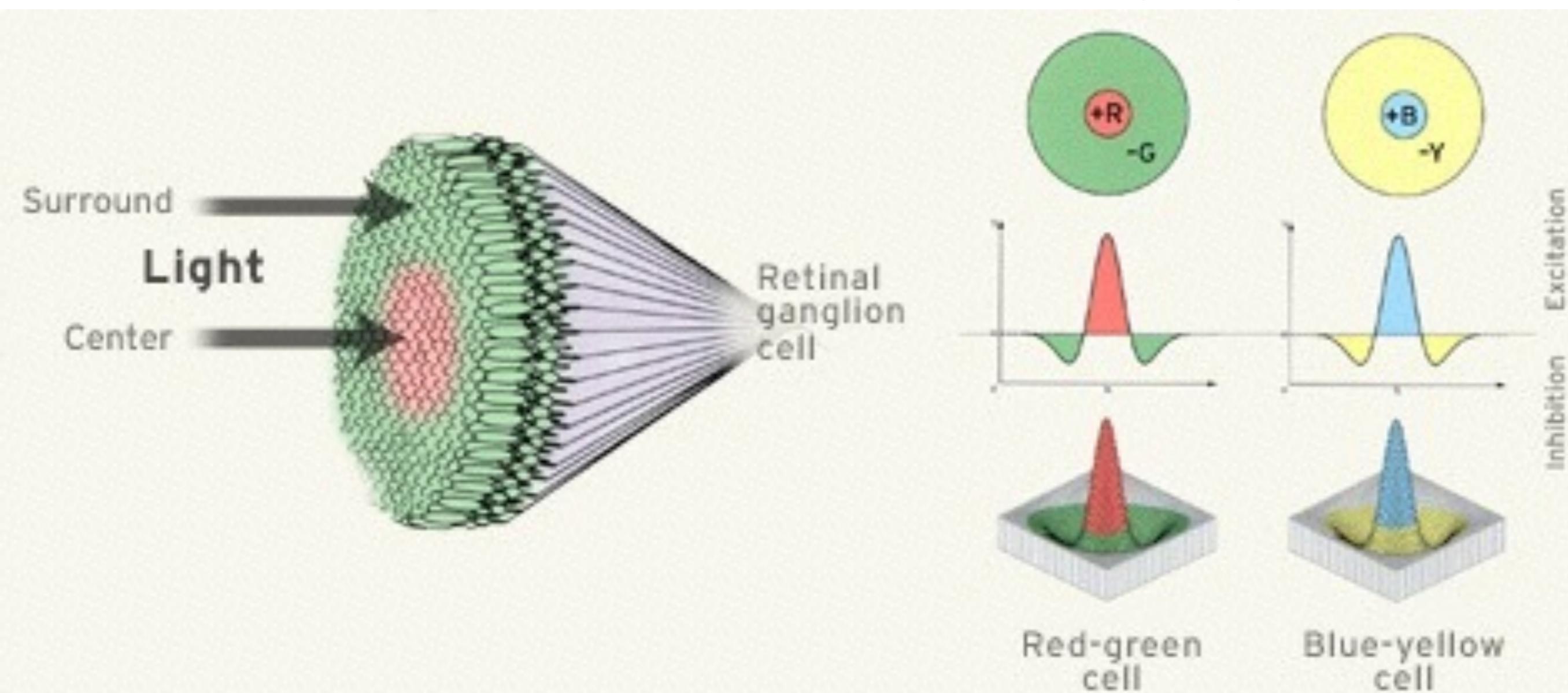
center-surround receptive fields

figure adapted from Ebner & Hameroff (2011)

# Receptive Fields

extract similar features at each position in the visual field

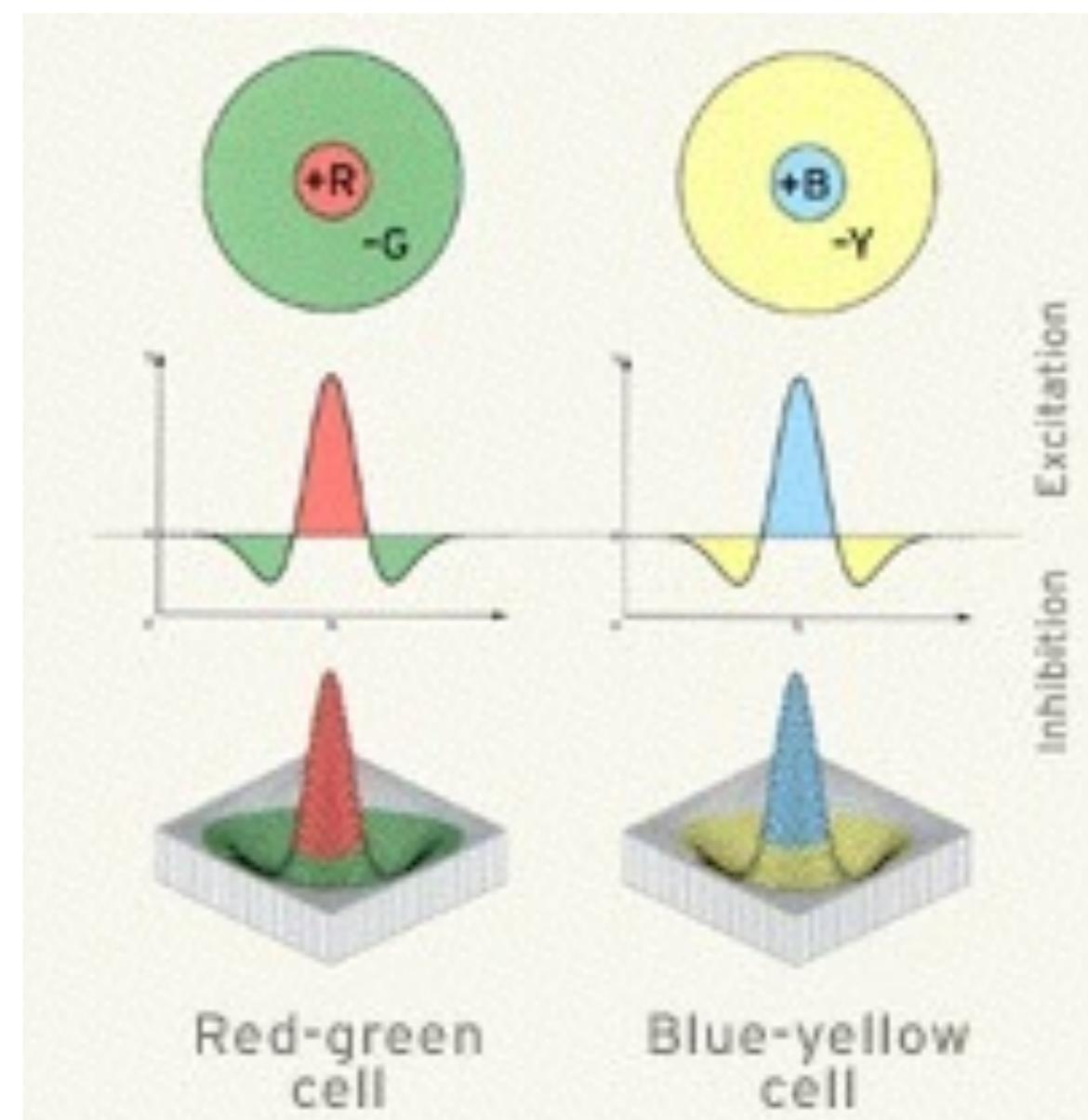
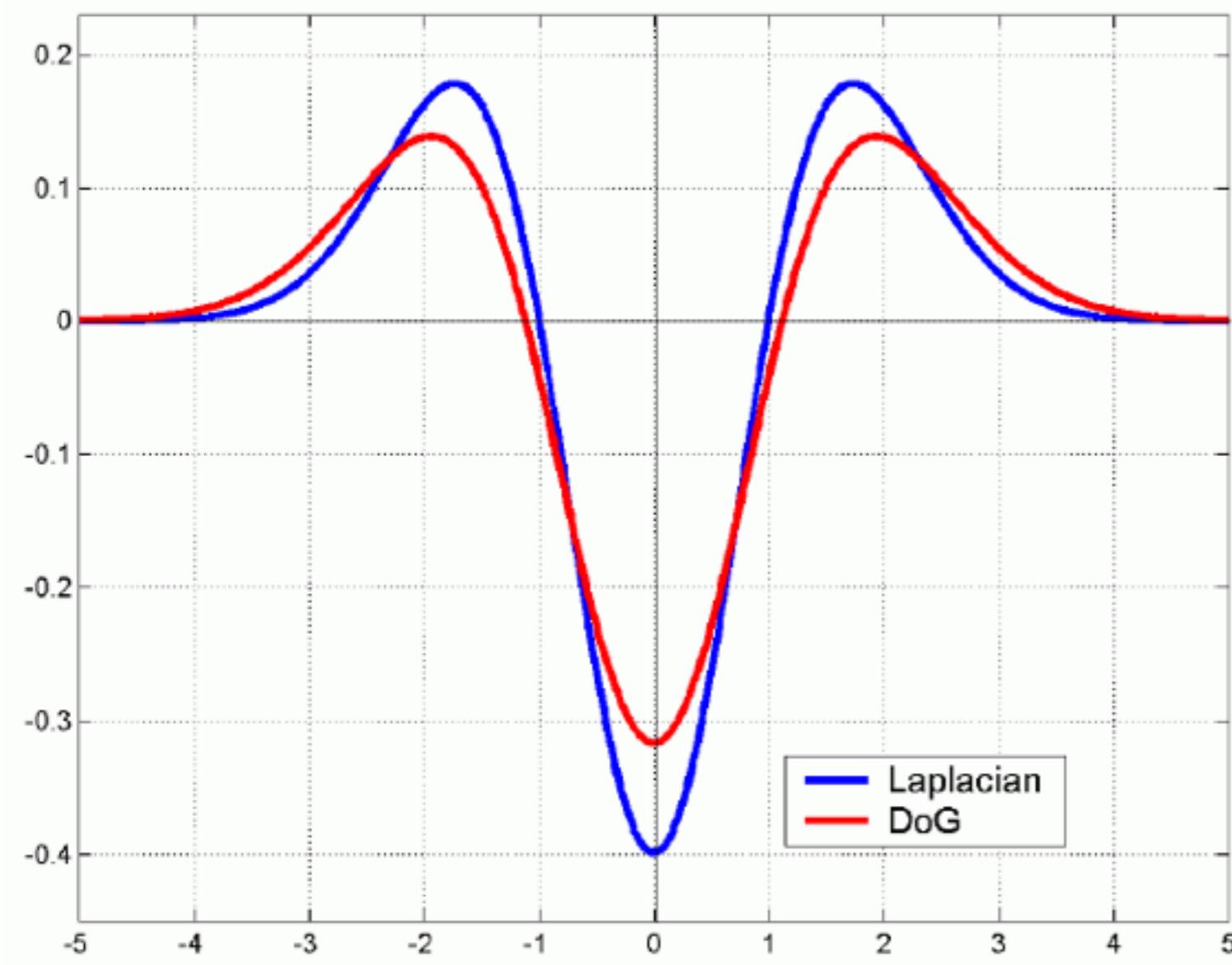
ON ganglion cells



# Receptive Fields

extract similar features at each position in the visual field

ON ganglion cells



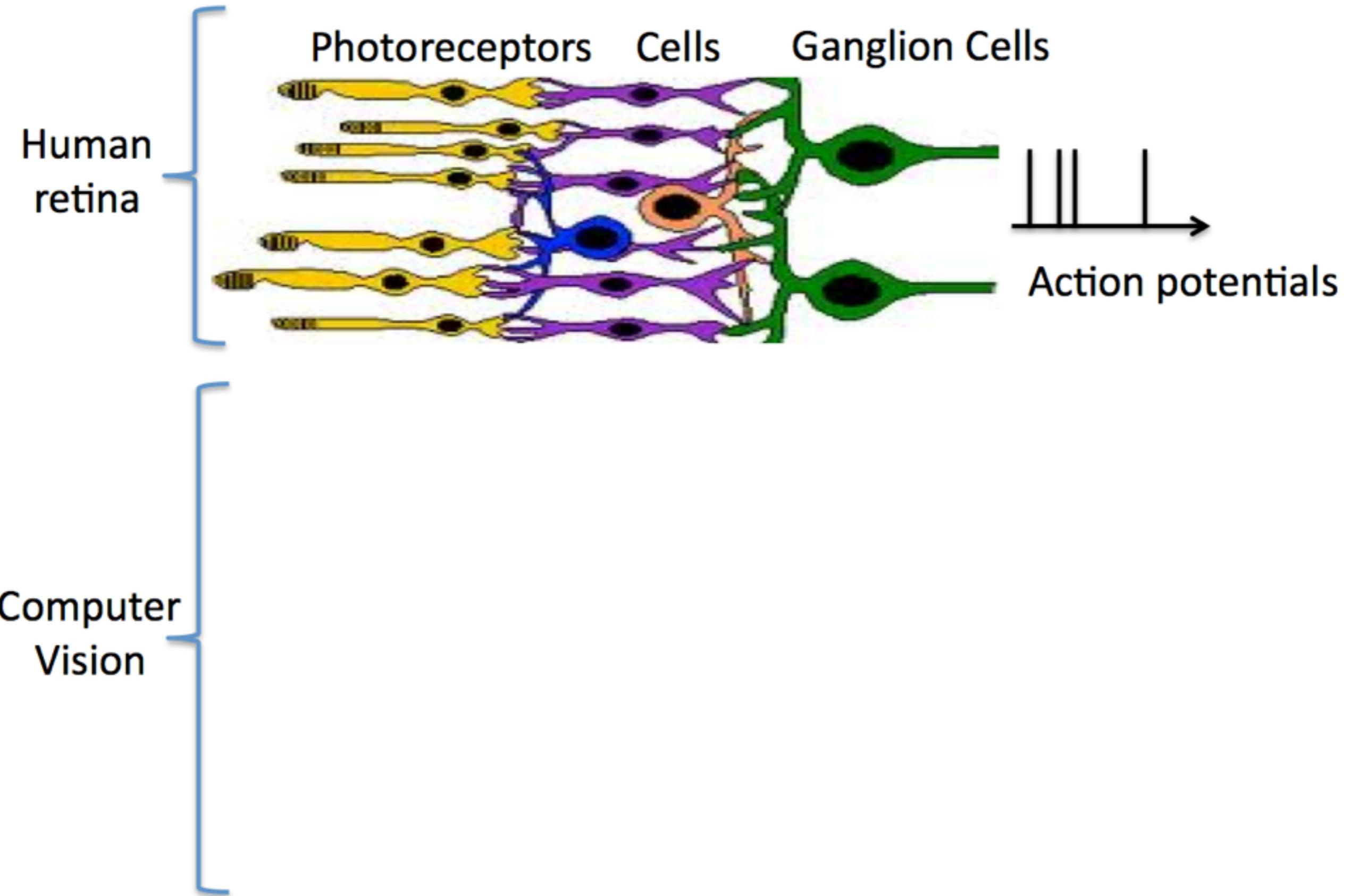


figure courtesy of A. Alahi

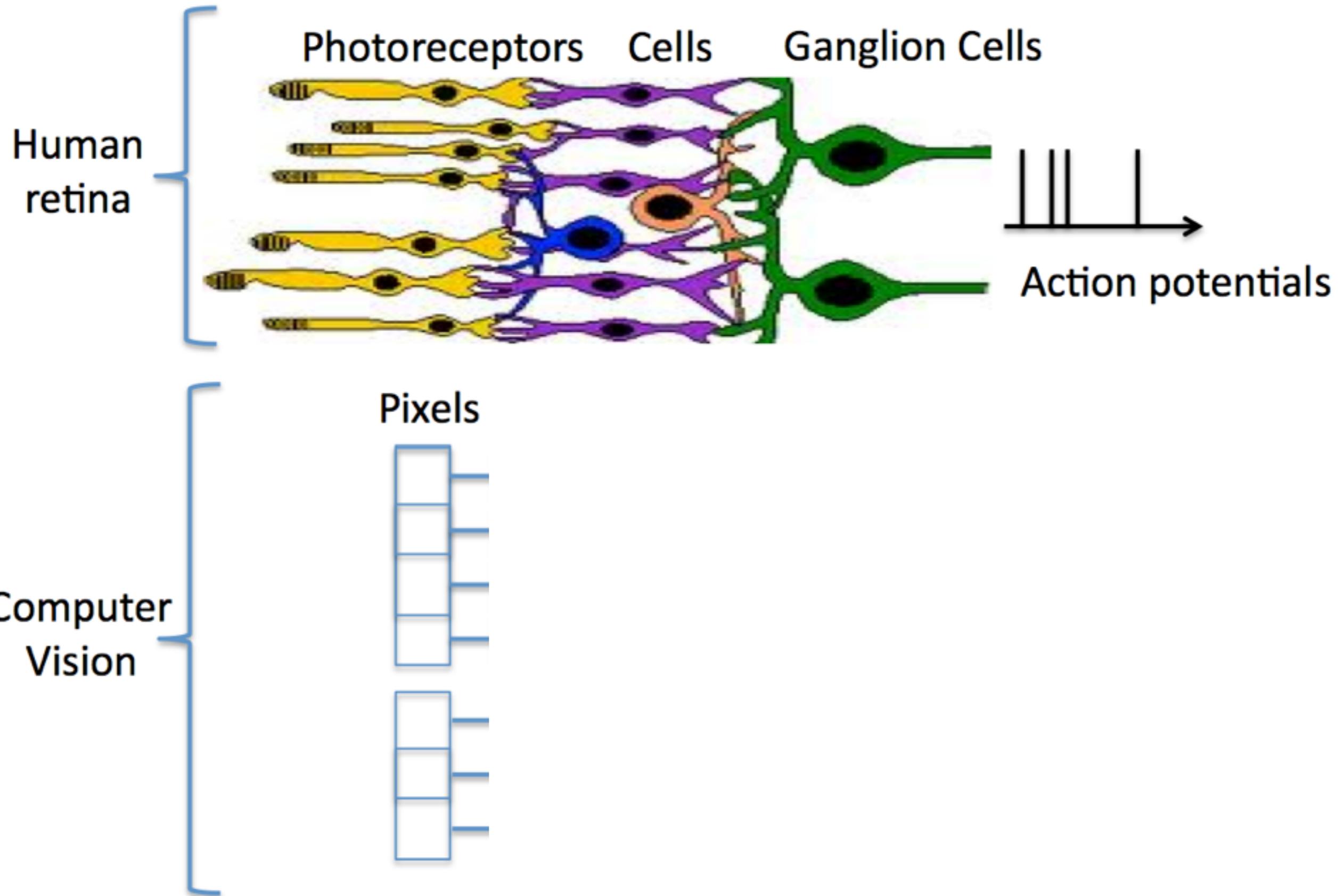


figure courtesy of A. Alahi

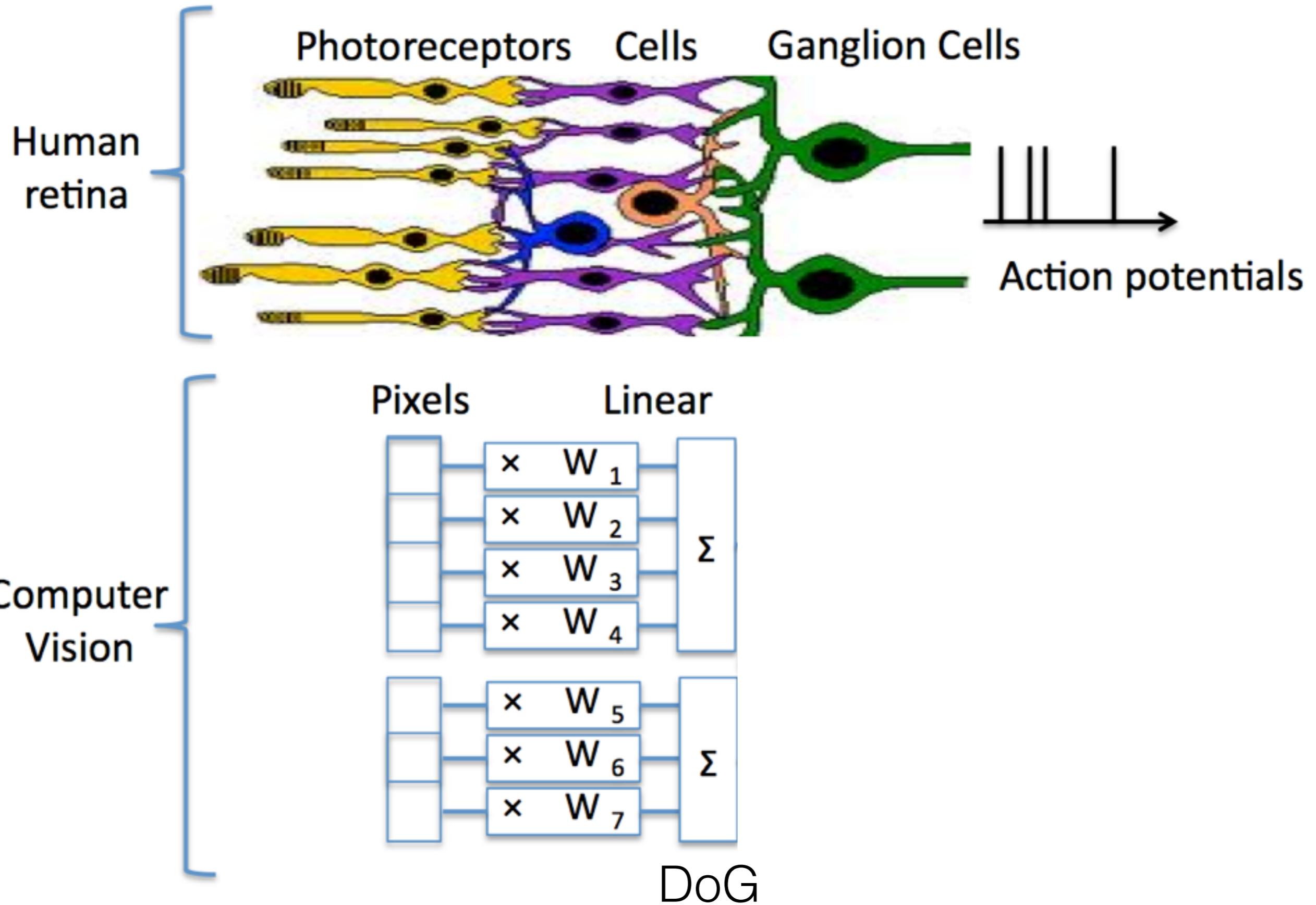


figure courtesy of A. Alahi

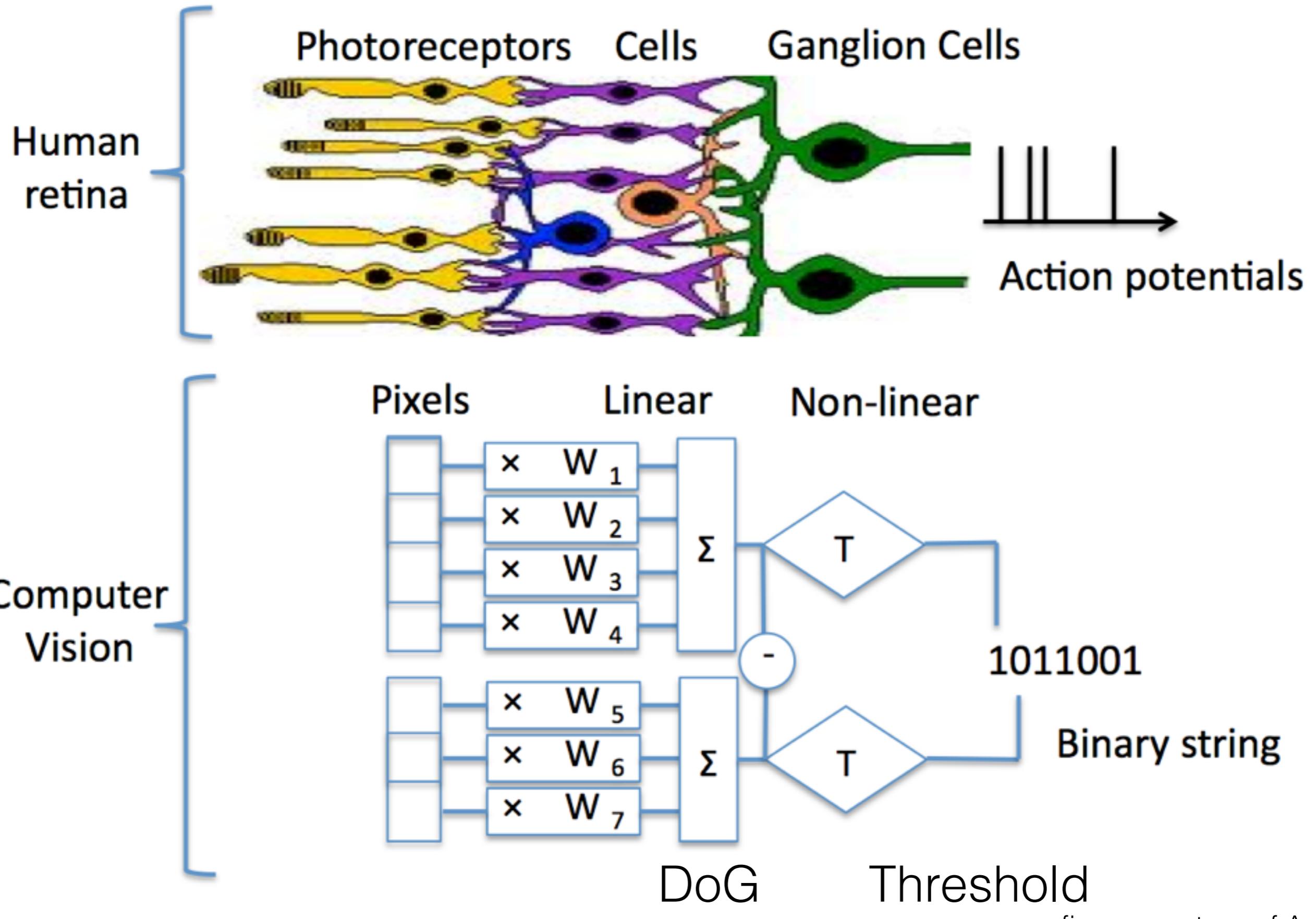
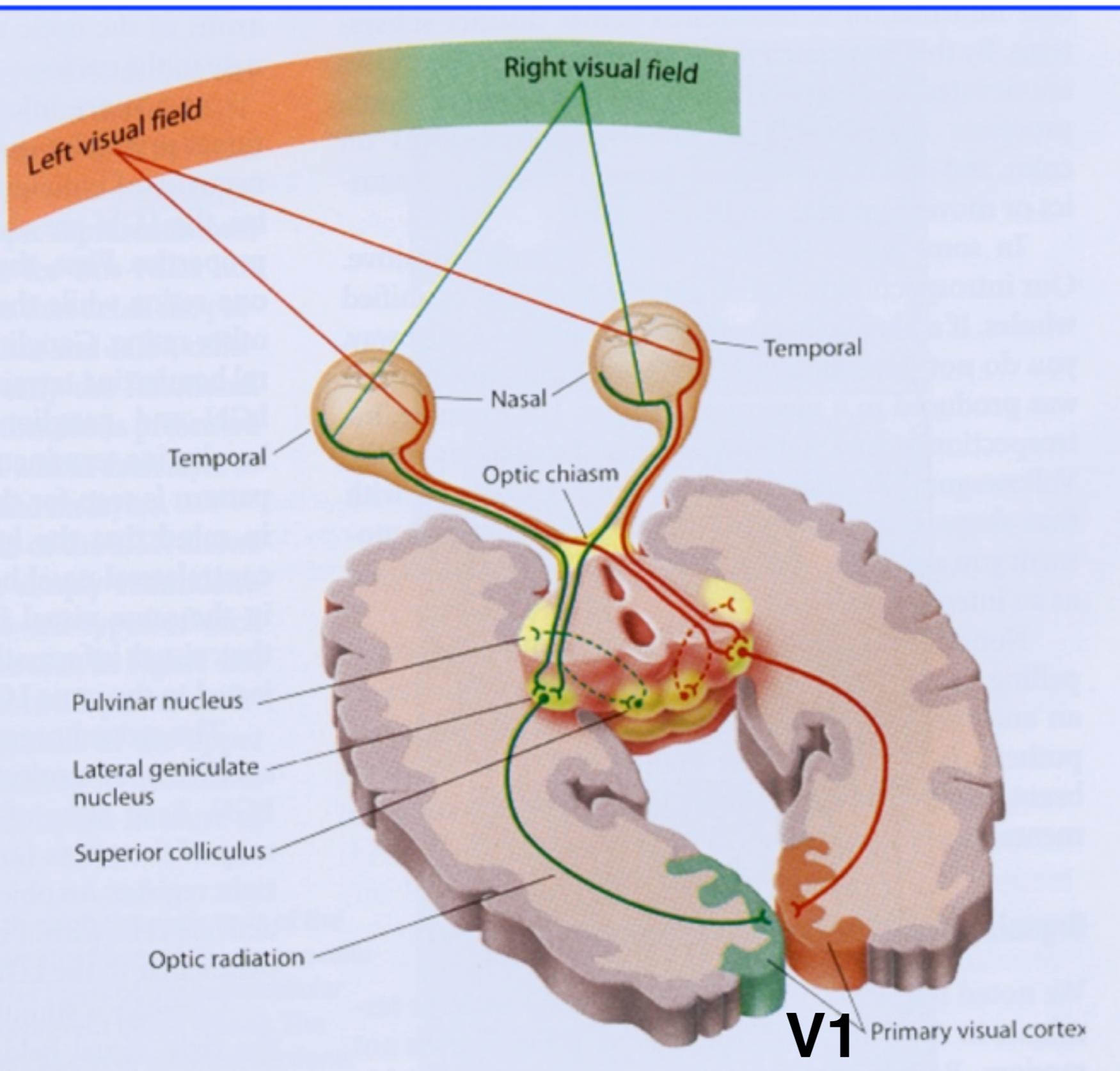


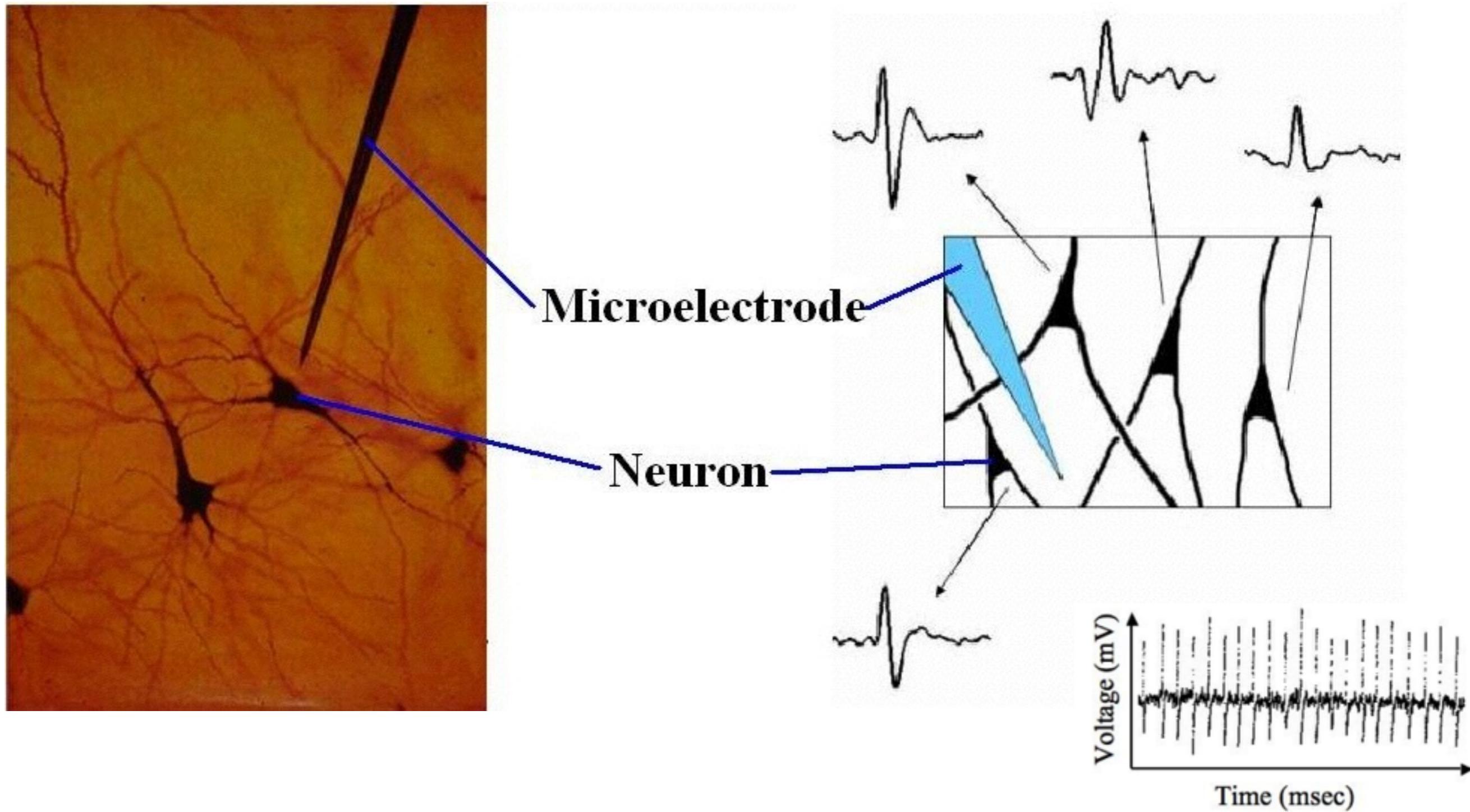
figure courtesy of A. Alahi

# From Retina to Cortex



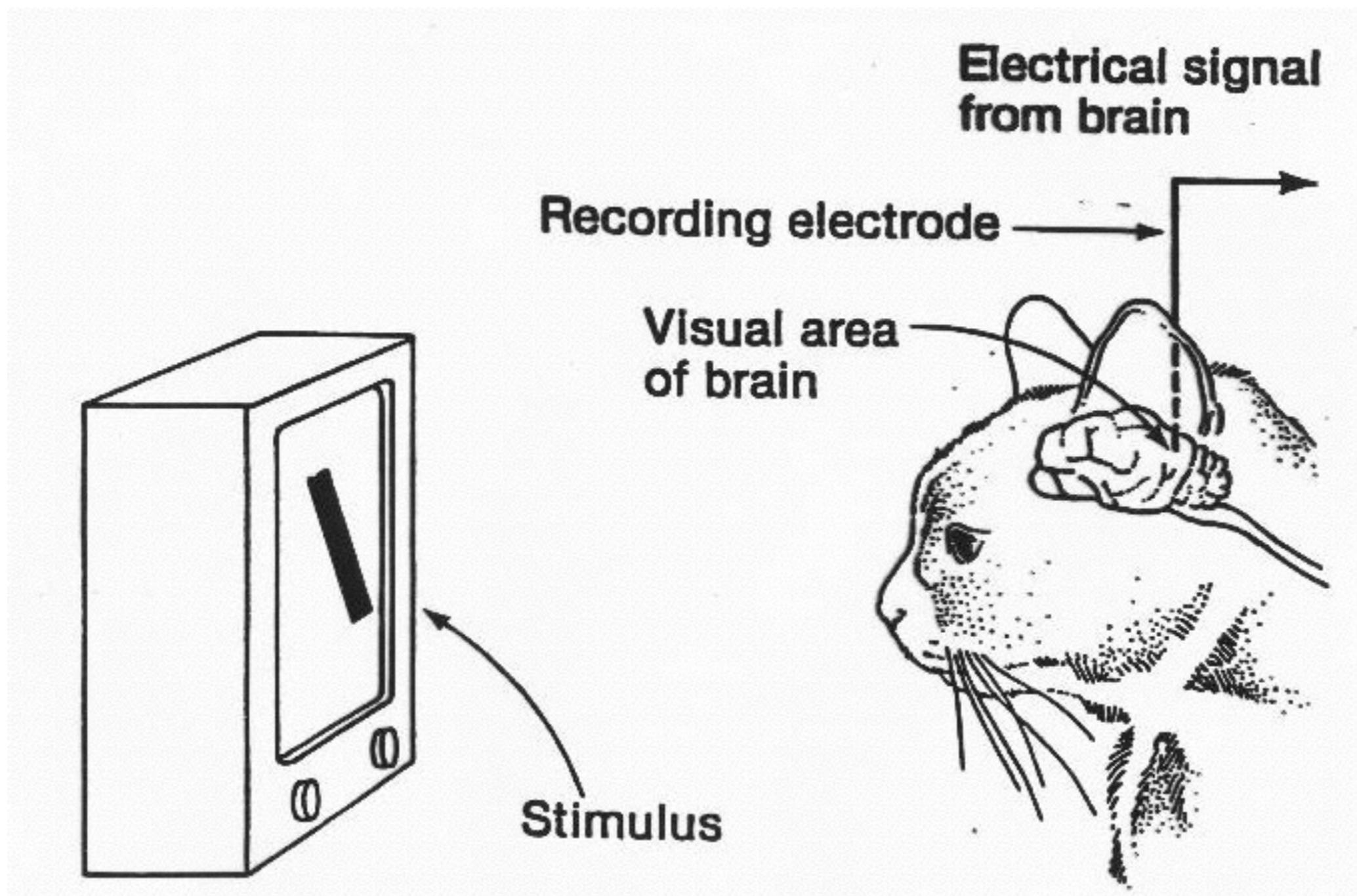
world  
↓  
retina  
(compression)  
↓  
LGN  
↓  
visual cortex  
(expansion)

# Single Electrode Recording



# Single Electrode Recording

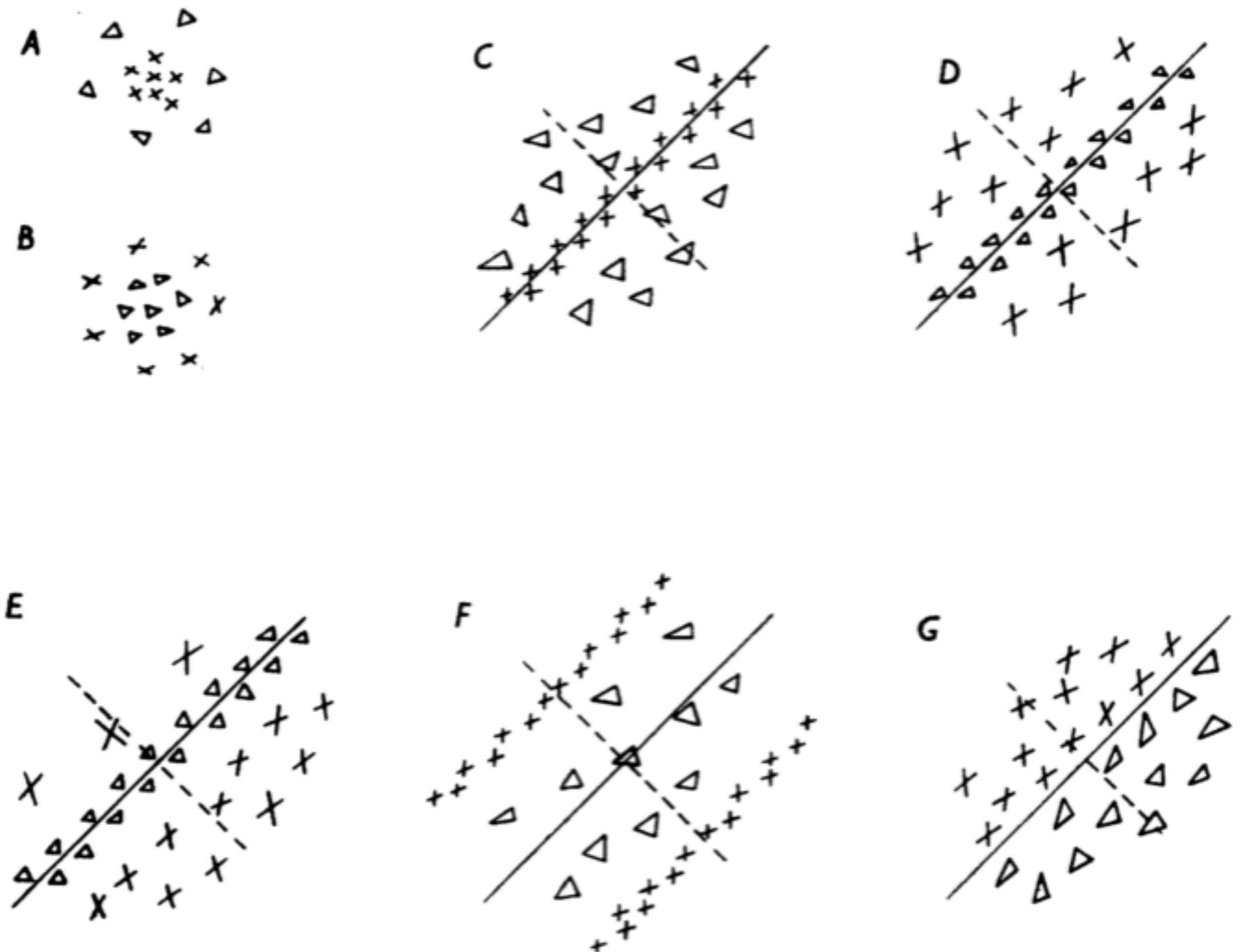
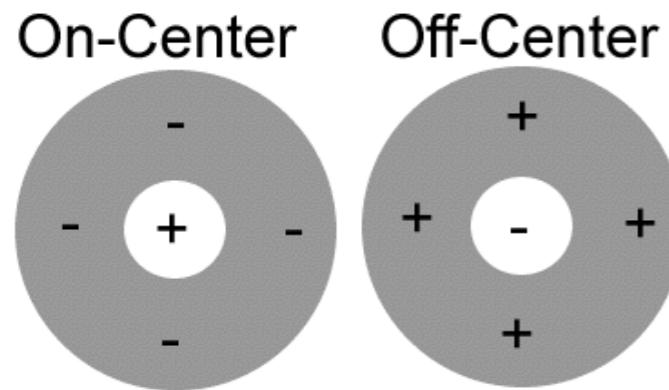
<https://www.youtube.com/watch?v=8VdFf3egwfg>



Hubel & Wiesel (1962)

# Types of V1 cells

simple cells

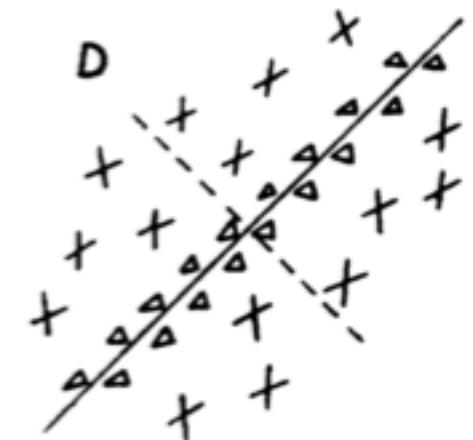
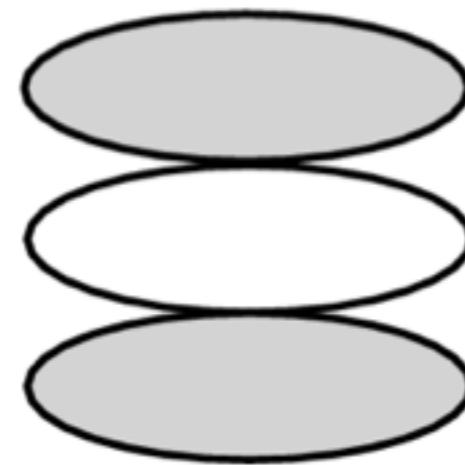
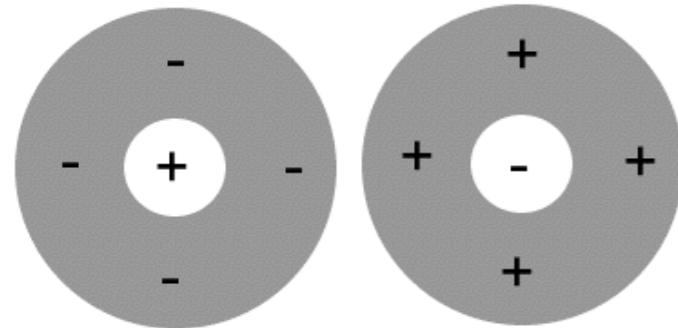


Hubel & Wiesel (1962)

# Types of V1 cells

simple cells

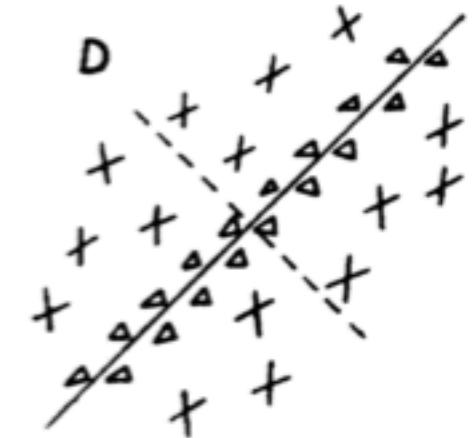
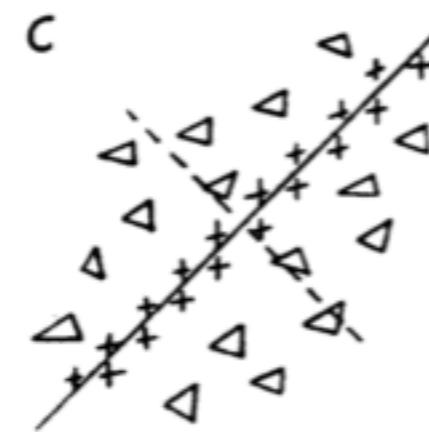
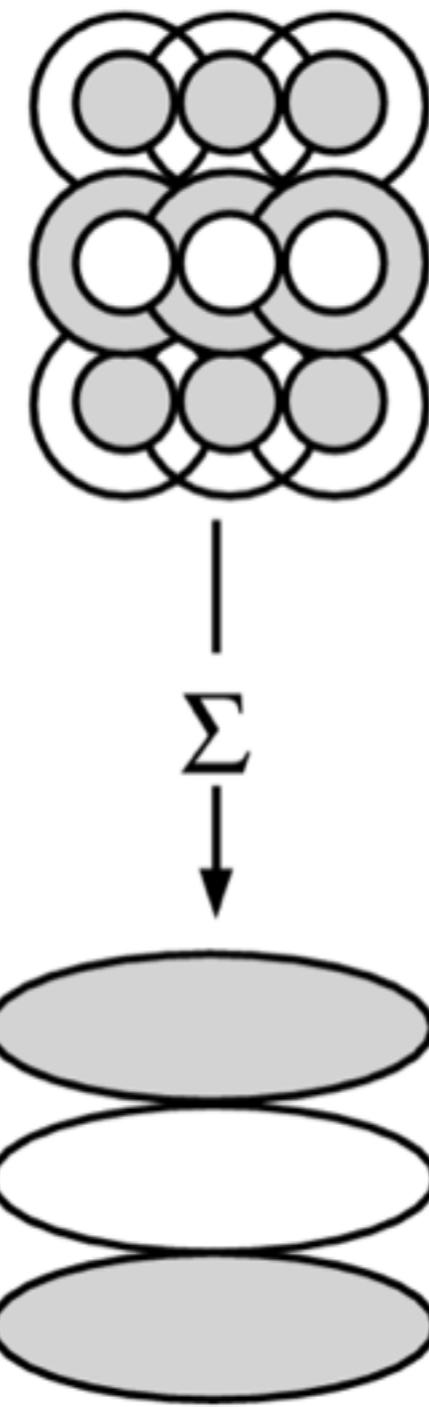
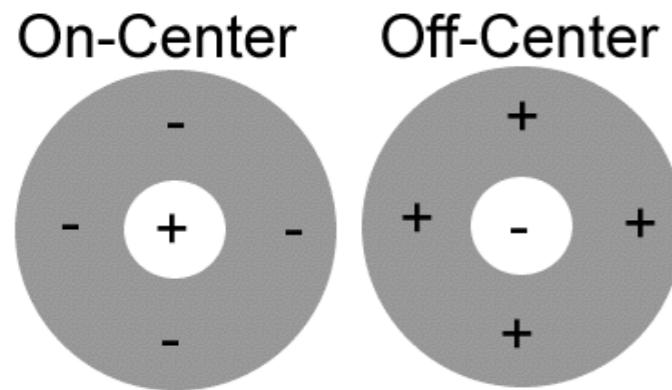
On-Center      Off-Center



Hubel & Wiesel (1962)

# Types of V1 cells

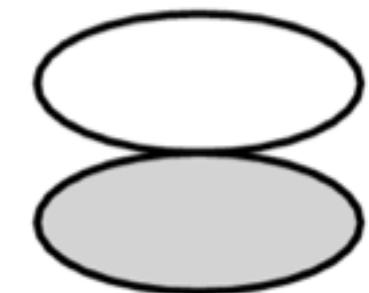
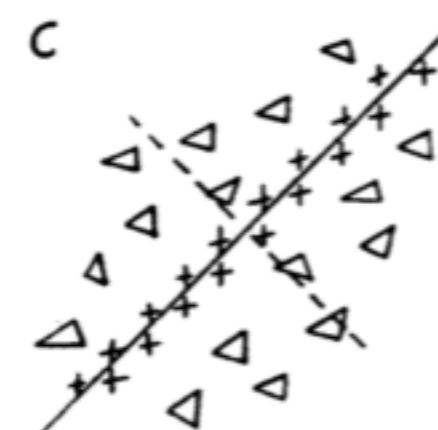
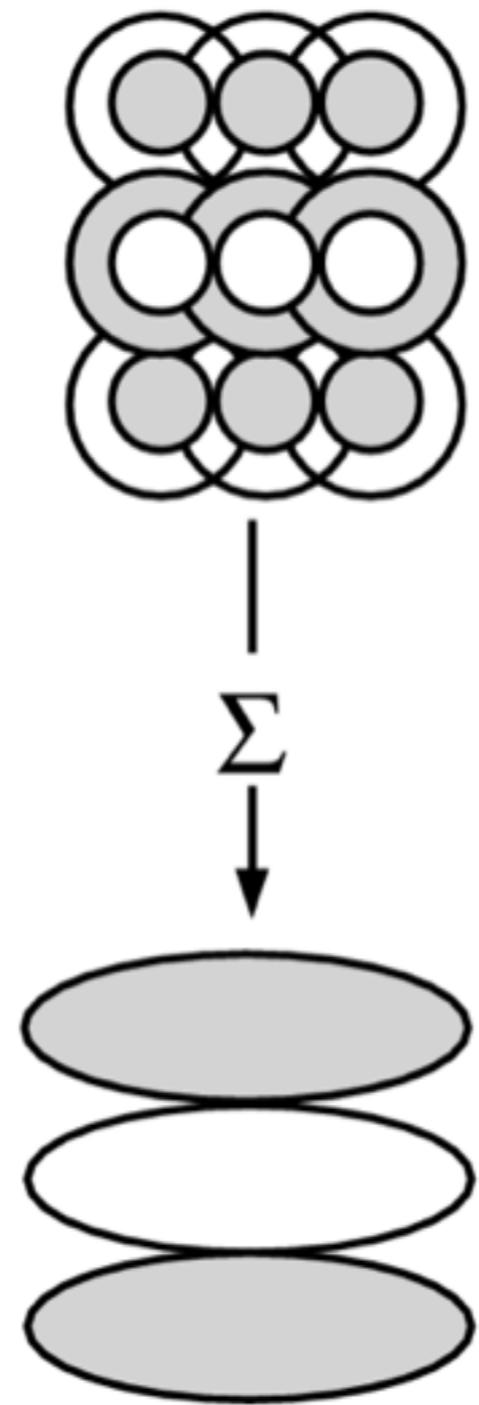
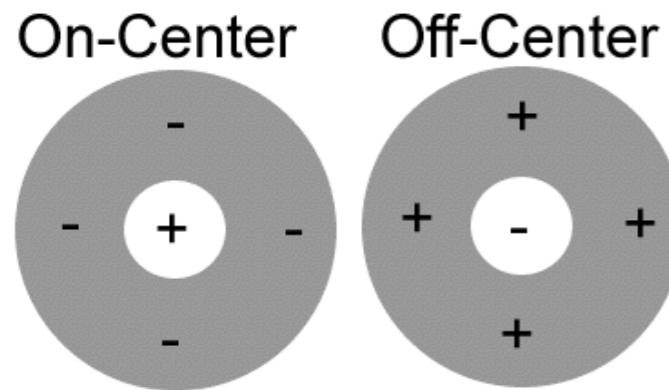
simple cells



Hubel & Wiesel (1962)

# Types of V1 cells

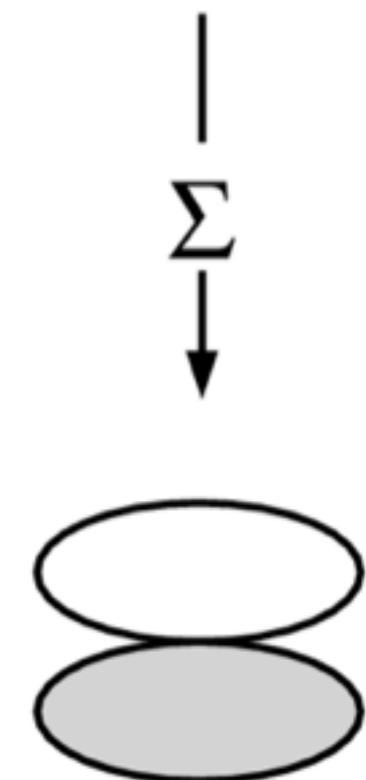
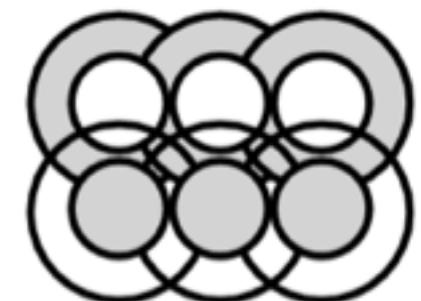
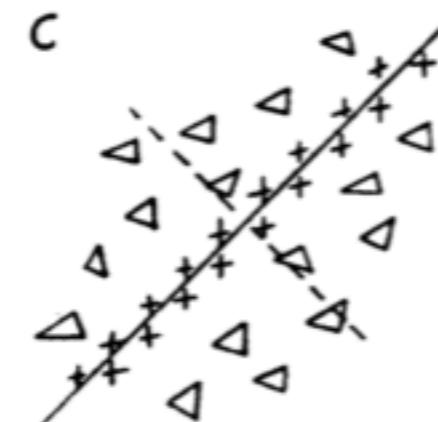
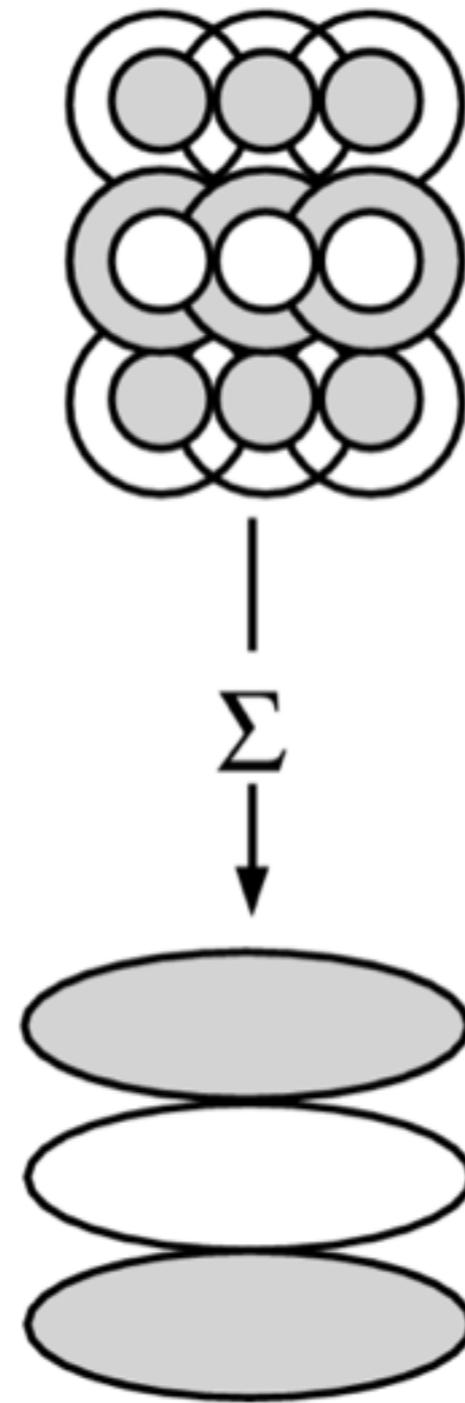
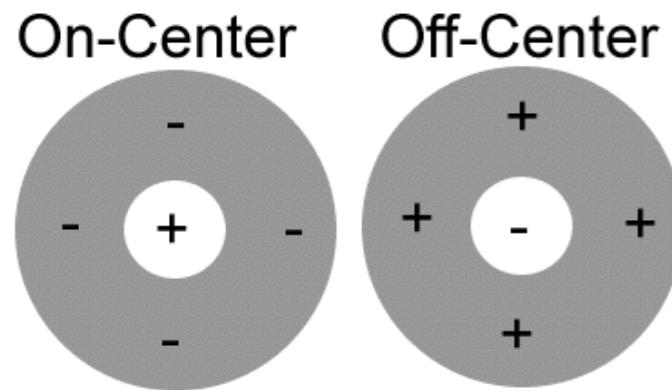
simple cells



Hubel & Wiesel (1962)

# Types of V1 cells

simple cells

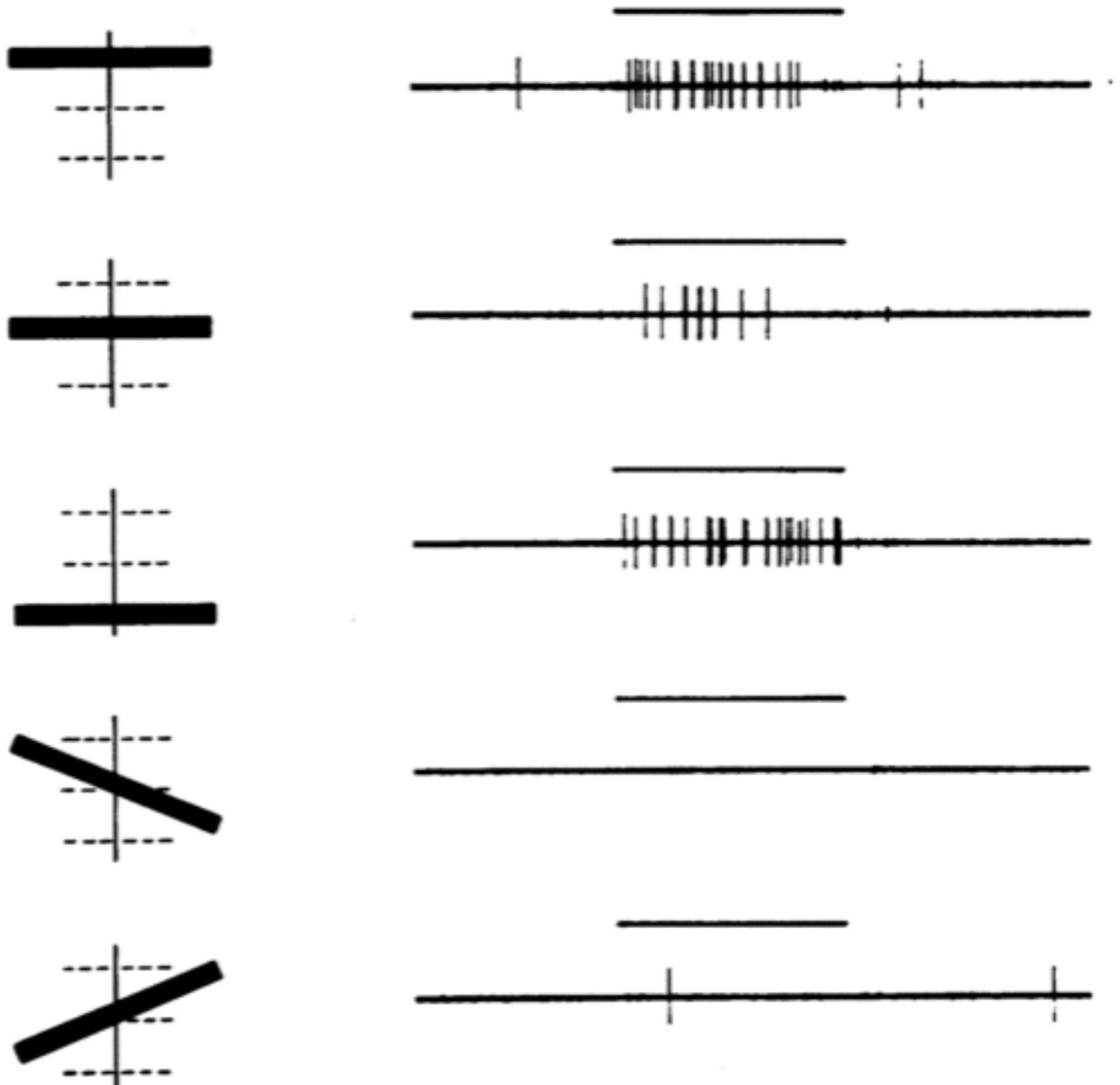
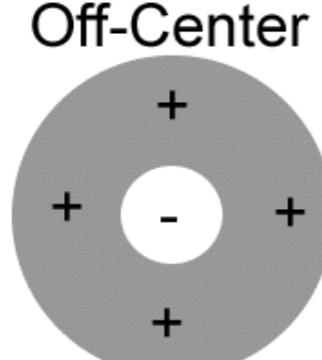
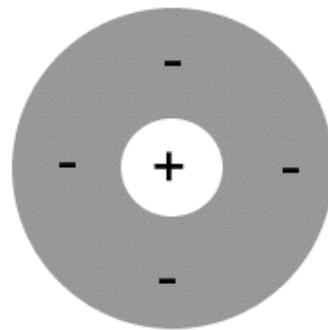


Hubel & Wiesel (1962)

# Types of V1 cells

complex cells

On-Center      Off-Center



Hubel & Wiesel (1962)

# Types of V1 cells

simple and complex cells are sensitive to:



center-surround (difference of gaussians!)



edges (symmetrical and asymmetrical)



rectangles of various elongation, visual half fields

Hubel & Wiesel (1962)

# Types of V1 cells

simple and complex cells are sensitive to:



center-surround (difference of gaussians!)

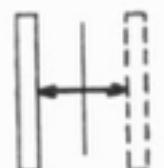


edges (symmetrical and asymmetrical)



rectangles of various elongation, visual half fields

V1 also has cells that are sensitive to:



motion (in fact, it's a separate processing stream!)



color (groups of cells called “color blobs”)

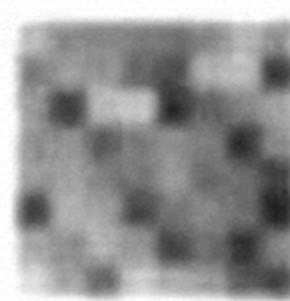


other stuff

Hubel & Wiesel (1962)

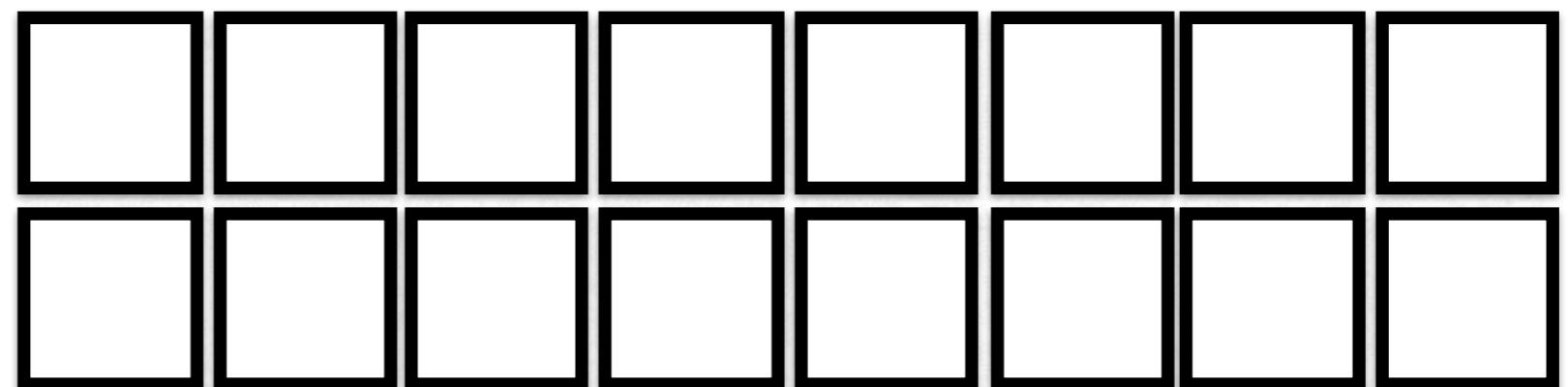
# Why Are Edges Special?

image



=

neurons



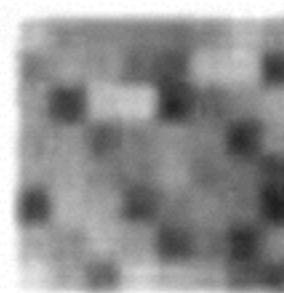
weights

?	?	?	?	?	?	?	?
?	?	?	?	?	?	?	?

Olshausen & Field (1996)

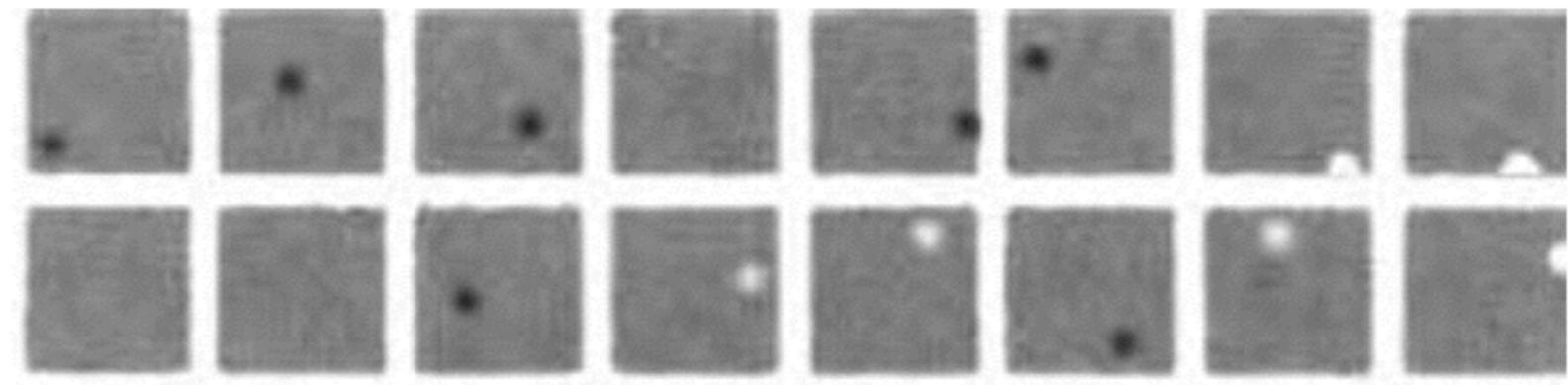
# Why Are Edges Special?

image



=

neurons



weights

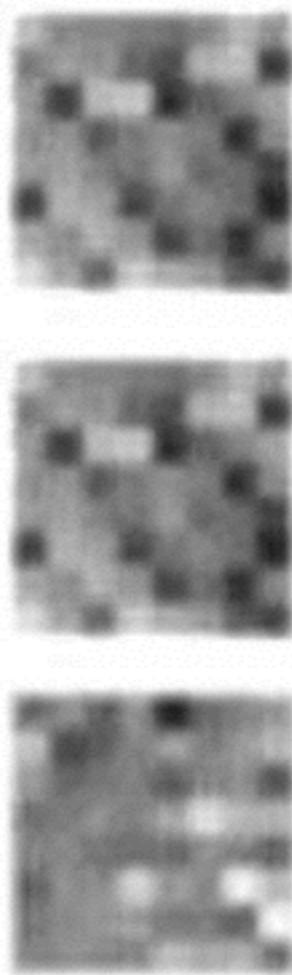
0	0	0	1	0	0	0	0
0	0	0	1	0	2	0	0

Olshausen & Field (1996)

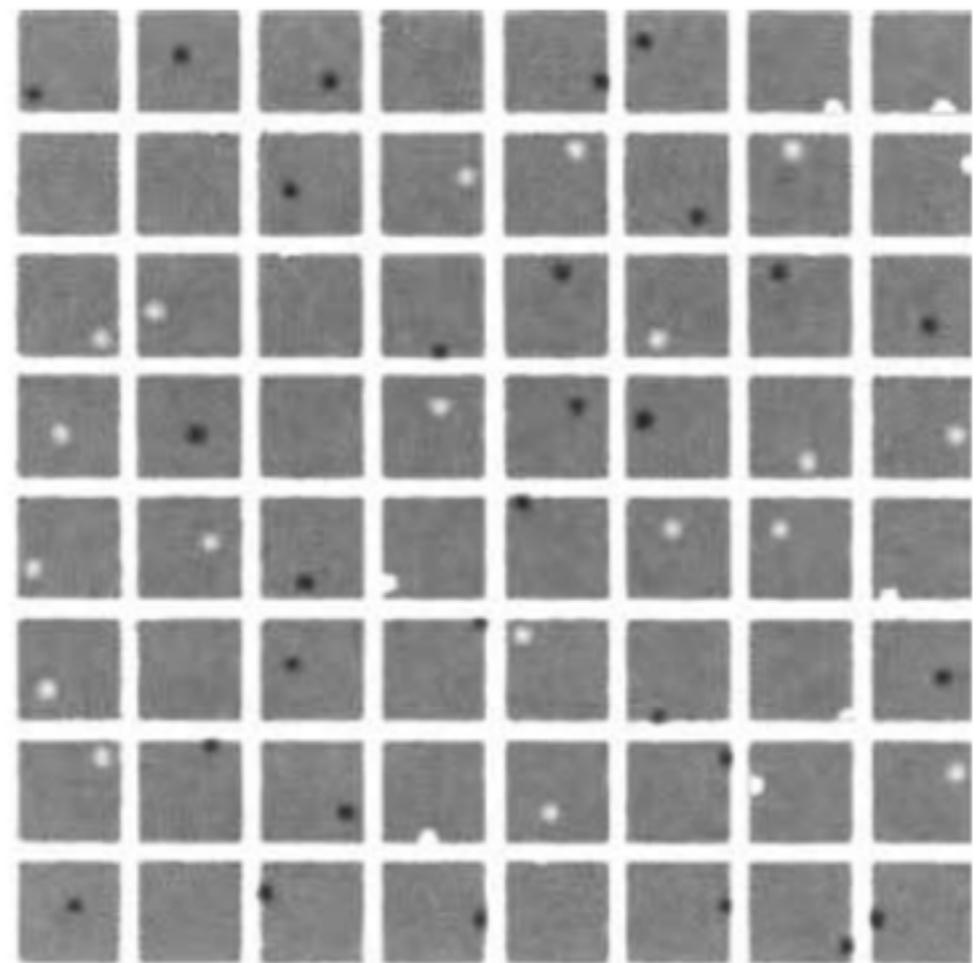
# Why Are Edges Special?

each pixel activated independently

INPUT



BASIS FOUND

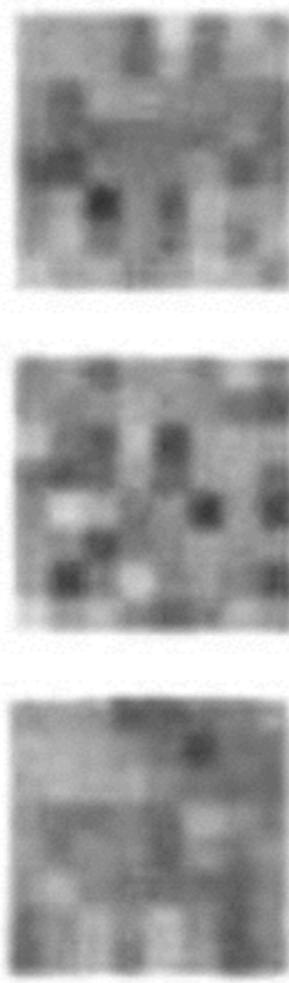


Olshausen & Field (1996)

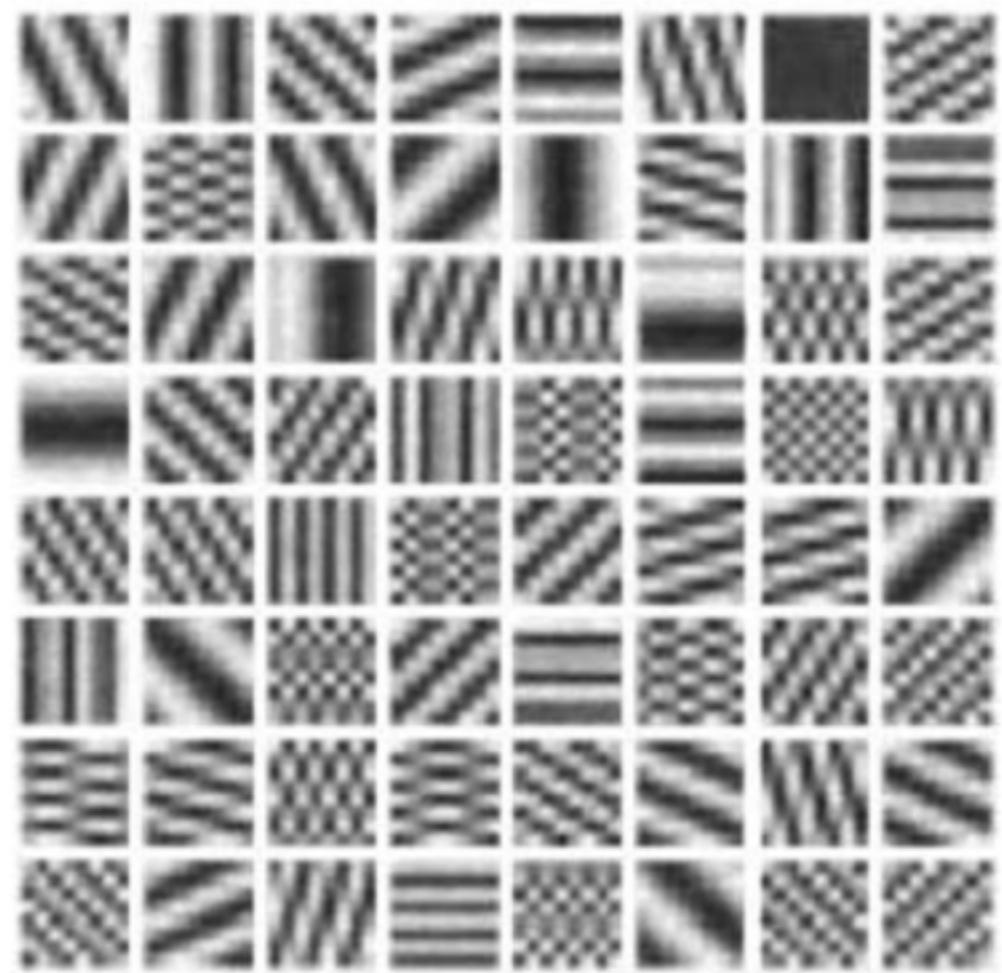
# Why Are Edges Special?

images are sums of independent gratings

INPUT



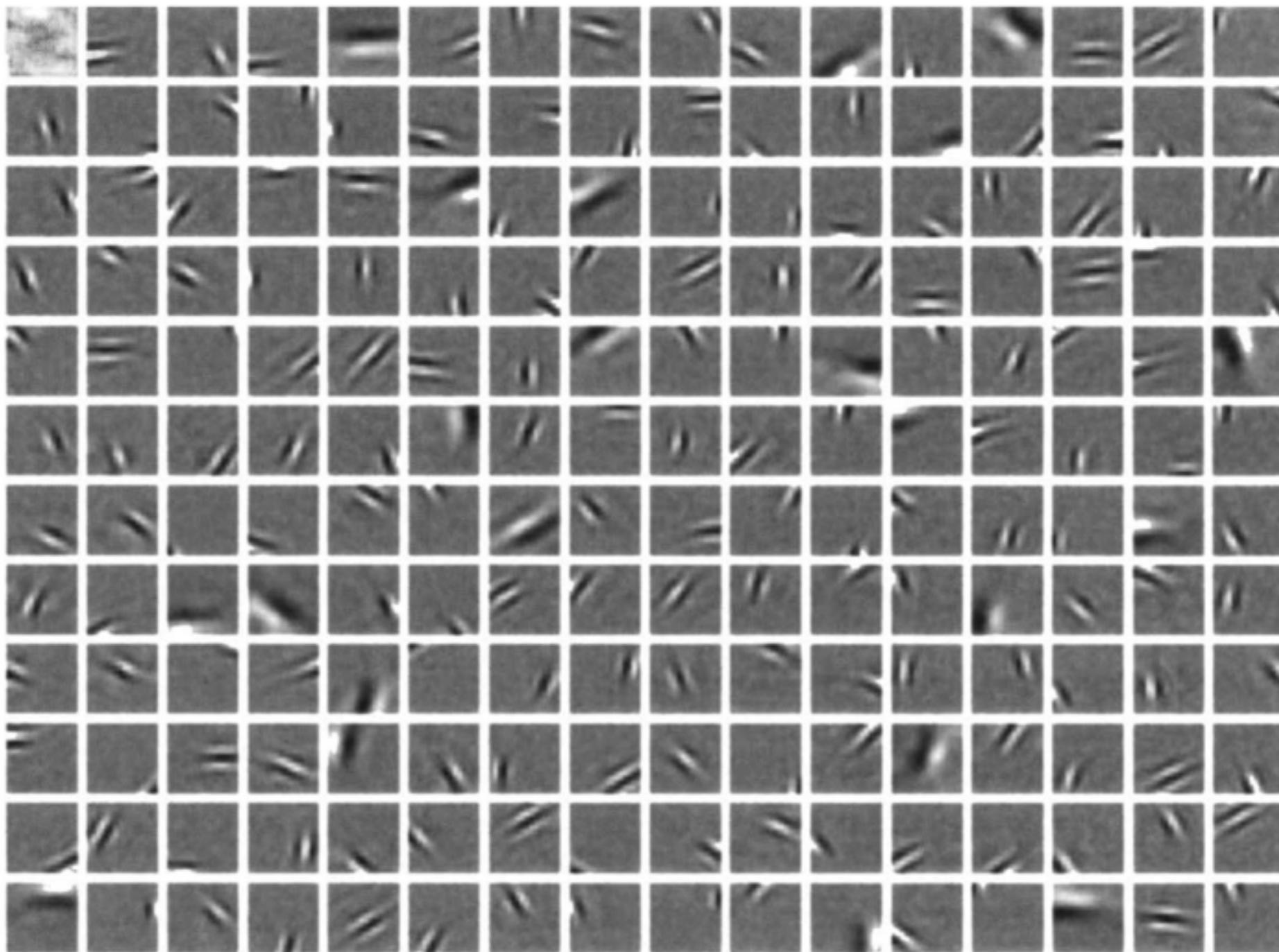
BASIS FOUND



Olshausen & Field (1996)

# Why Are Edges Special?

patches of real-world images



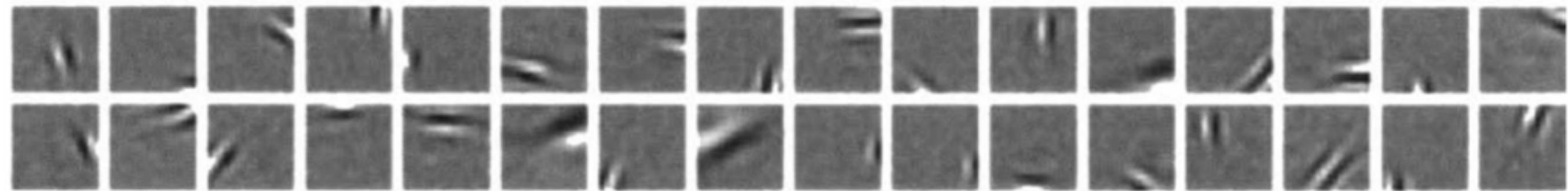
Olshausen & Field (1996)

# Early Computation

V1 cells encode edge orientation and position across visual field: simple and complex receptive fields



edges are a sufficient basis for real-world images!



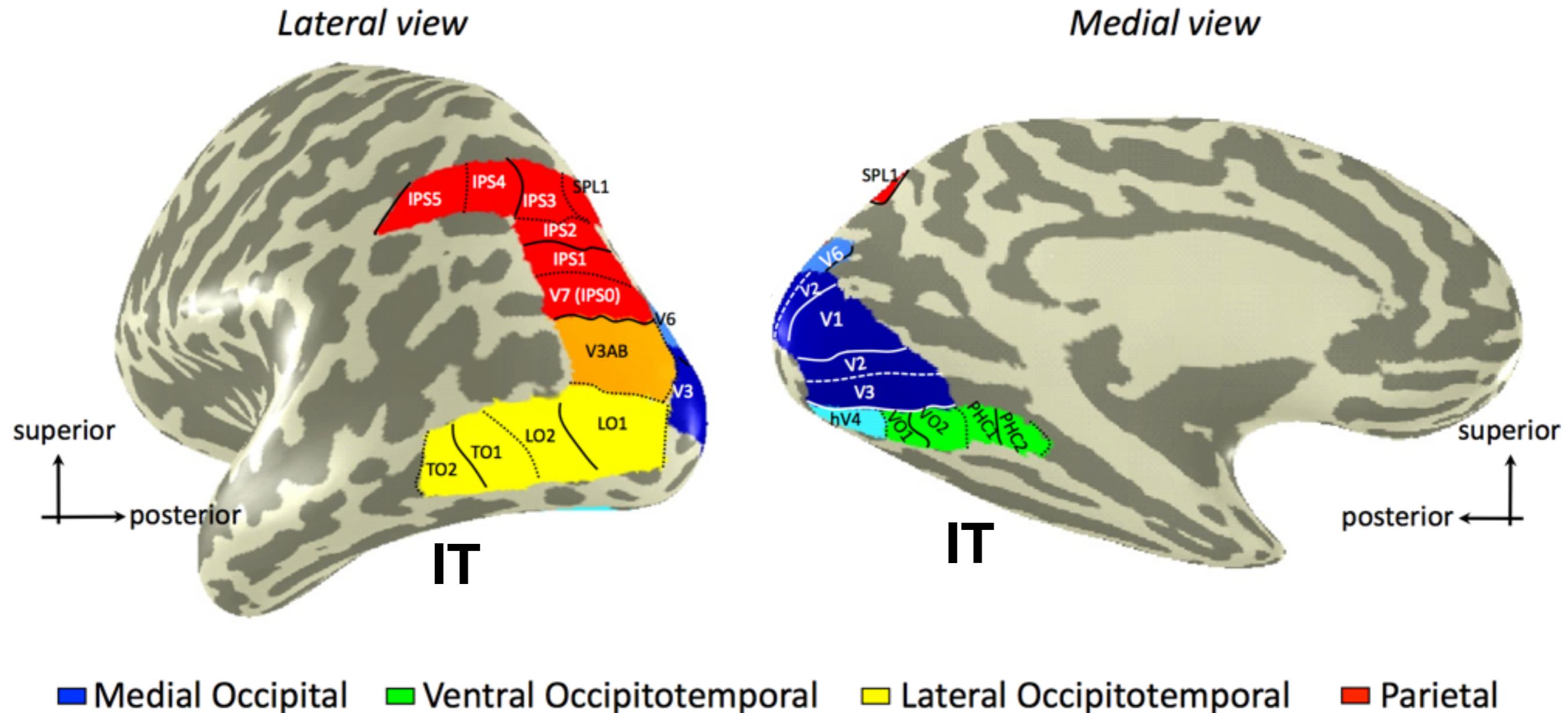
Hubel & Wiesel (1962), Olshausen & Field (1996)

### **3. Object Recognition in the Human Visual System**

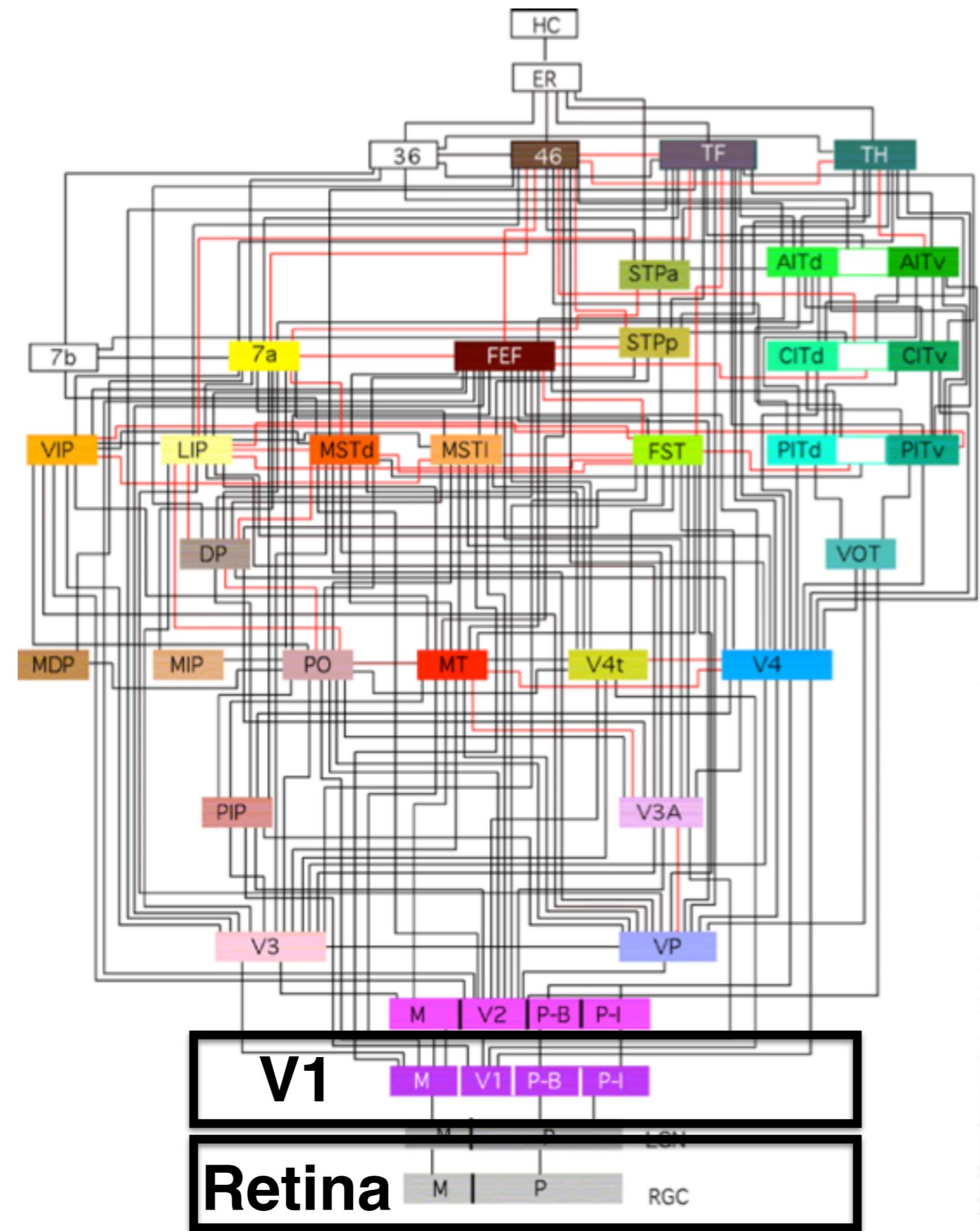
sequential transformations



# The Flow of Information



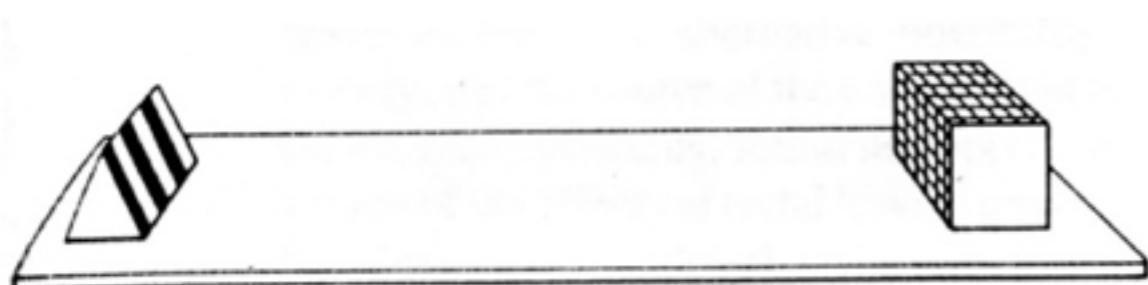
Weiner & Grill-Spector (2012)



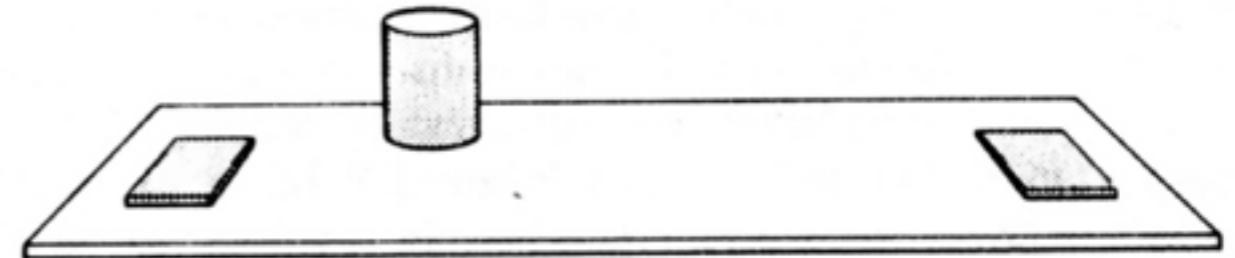
Van Essen (1991)

# Specialization: “What” and “Where” Pathways

monkey lesion studies



“what”



“where”

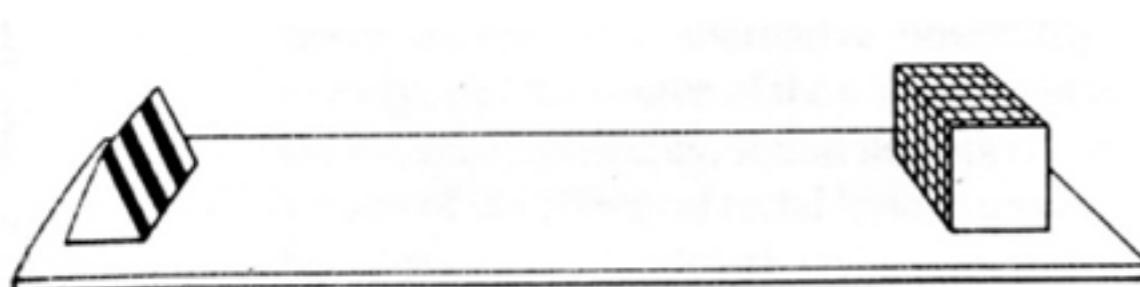
Mishkin & Ungerleider 1982

# Specialization: “What” and “Where” Pathways

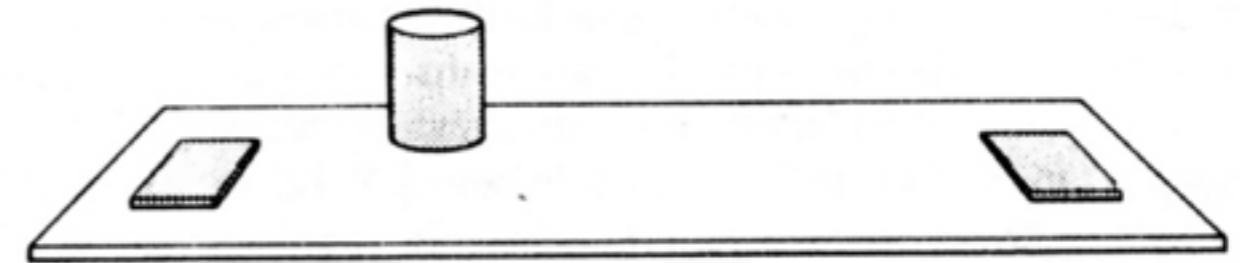
monkey lesion studies

lesion “where” pathway: difficulty in spatial reasoning

lesion “what” pathway: difficulty in object recognition



“what”



“where”

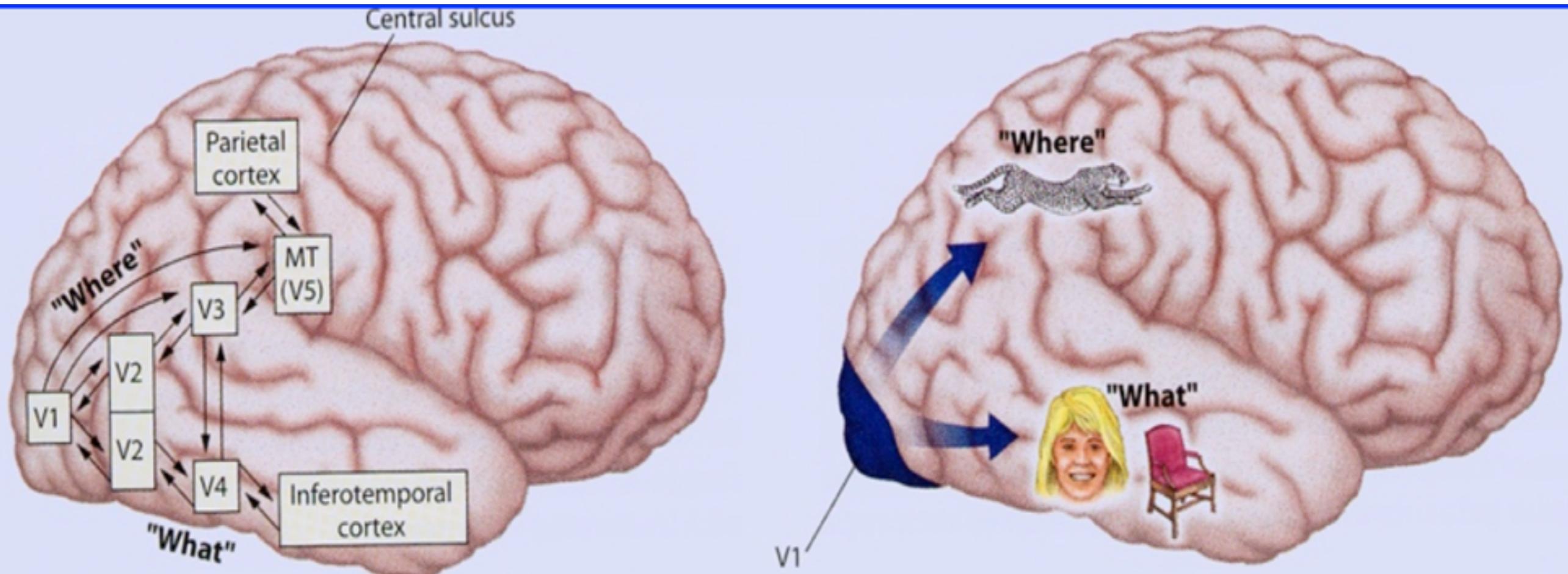
Mishkin & Ungerleider 1982

# Specialization: “What” and “Where” Pathways

monkey lesion studies

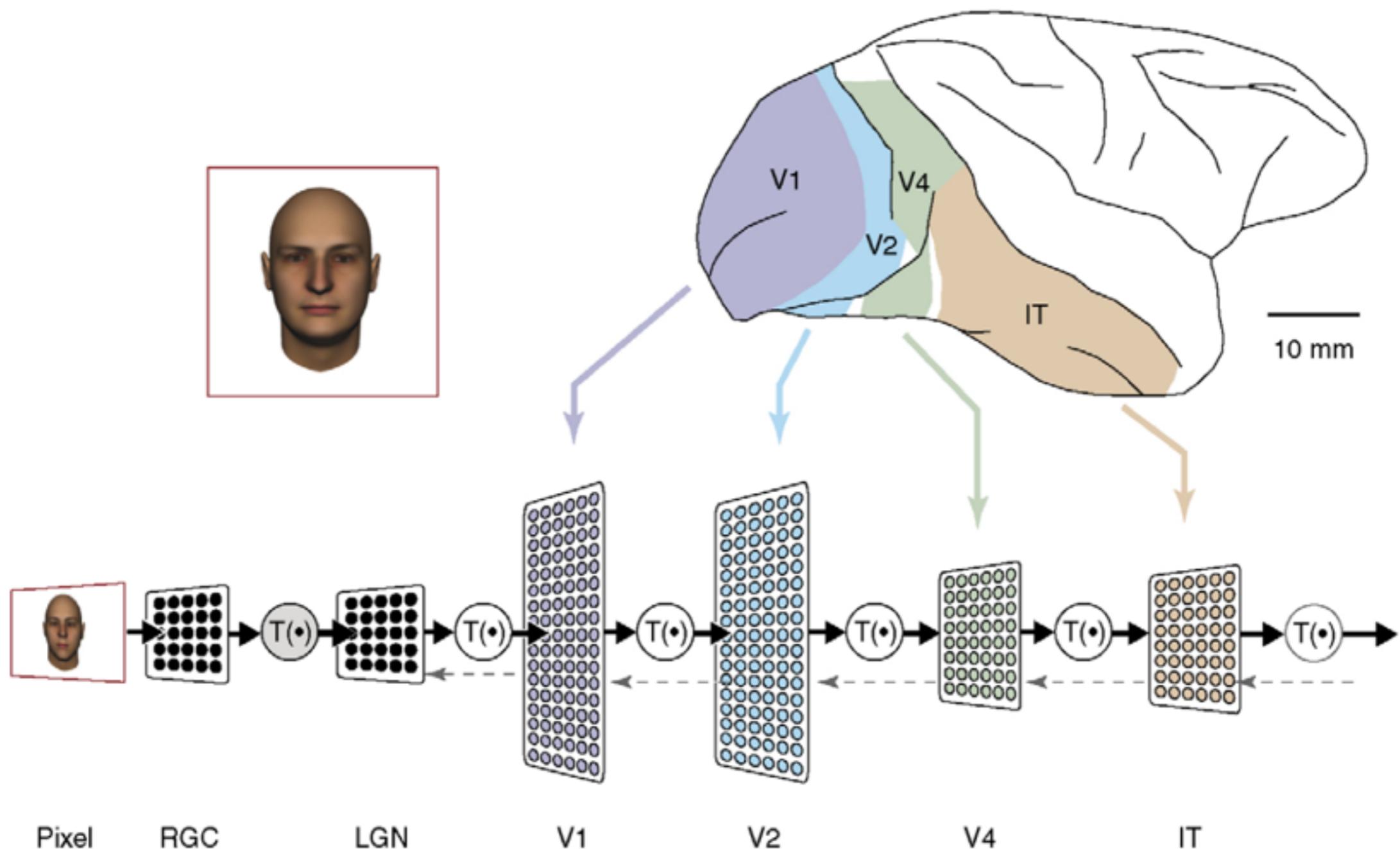
lesion “where” pathway: difficulty in spatial reasoning

lesion “what” pathway: difficulty in object recognition



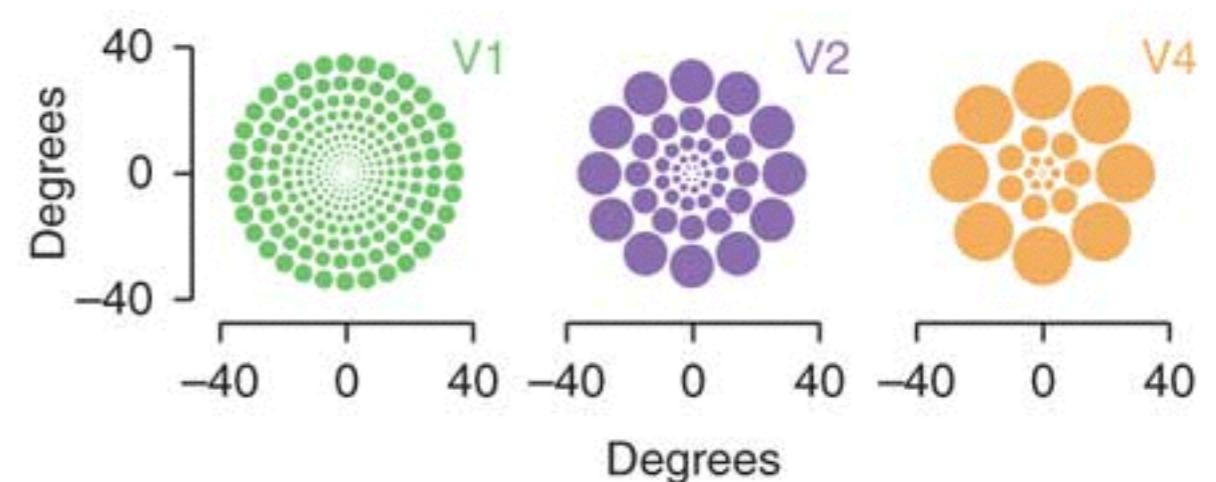
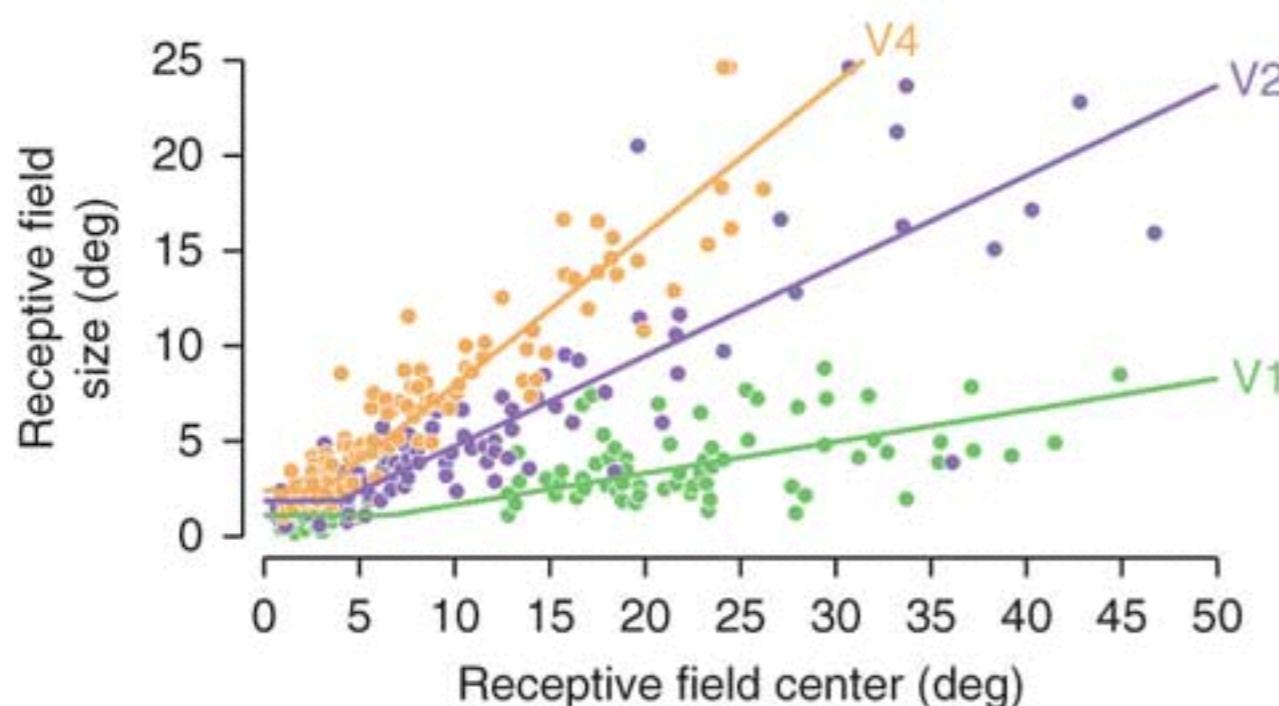
Mishkin & Ungerleider 1982

# Object Recognition: The “What” Pathway



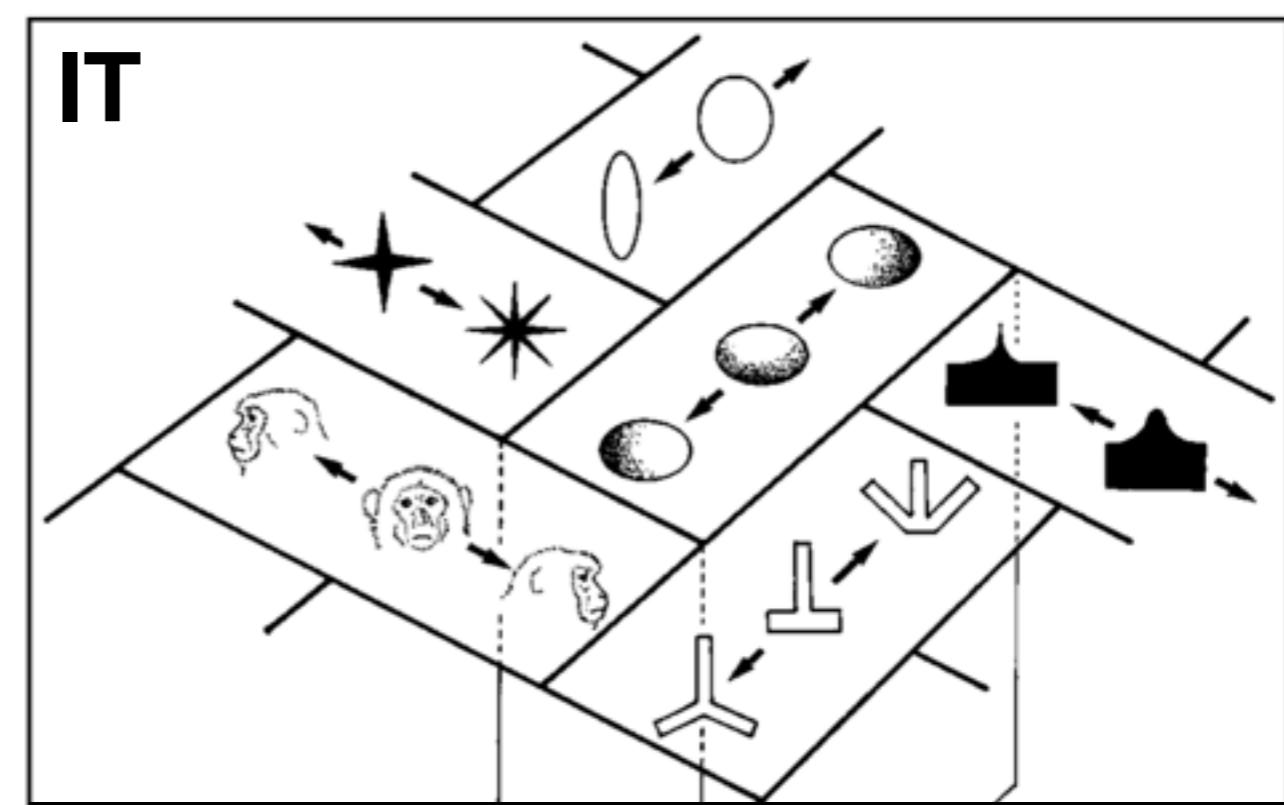
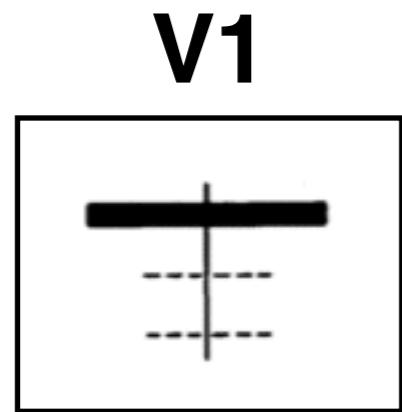
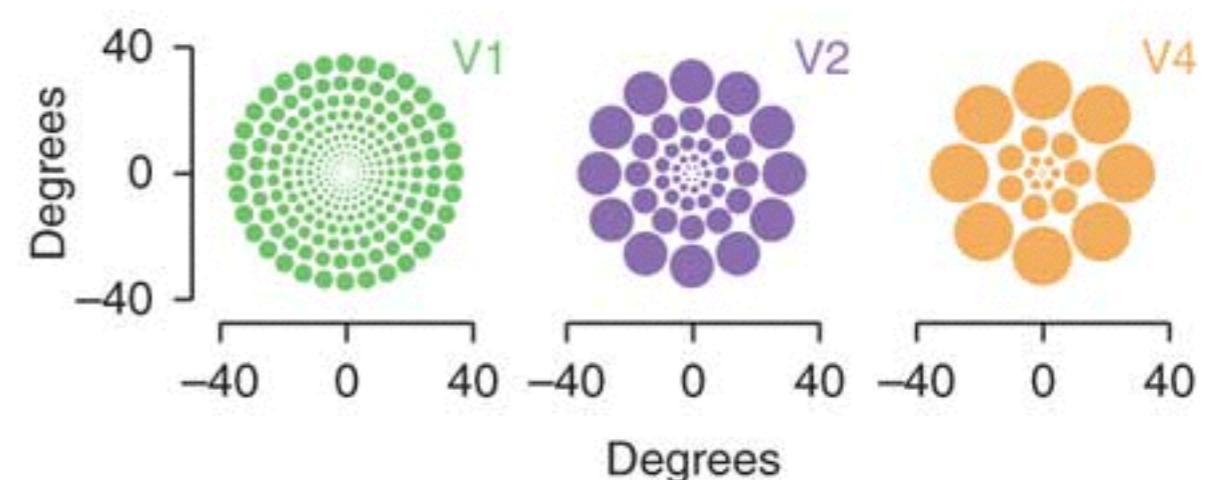
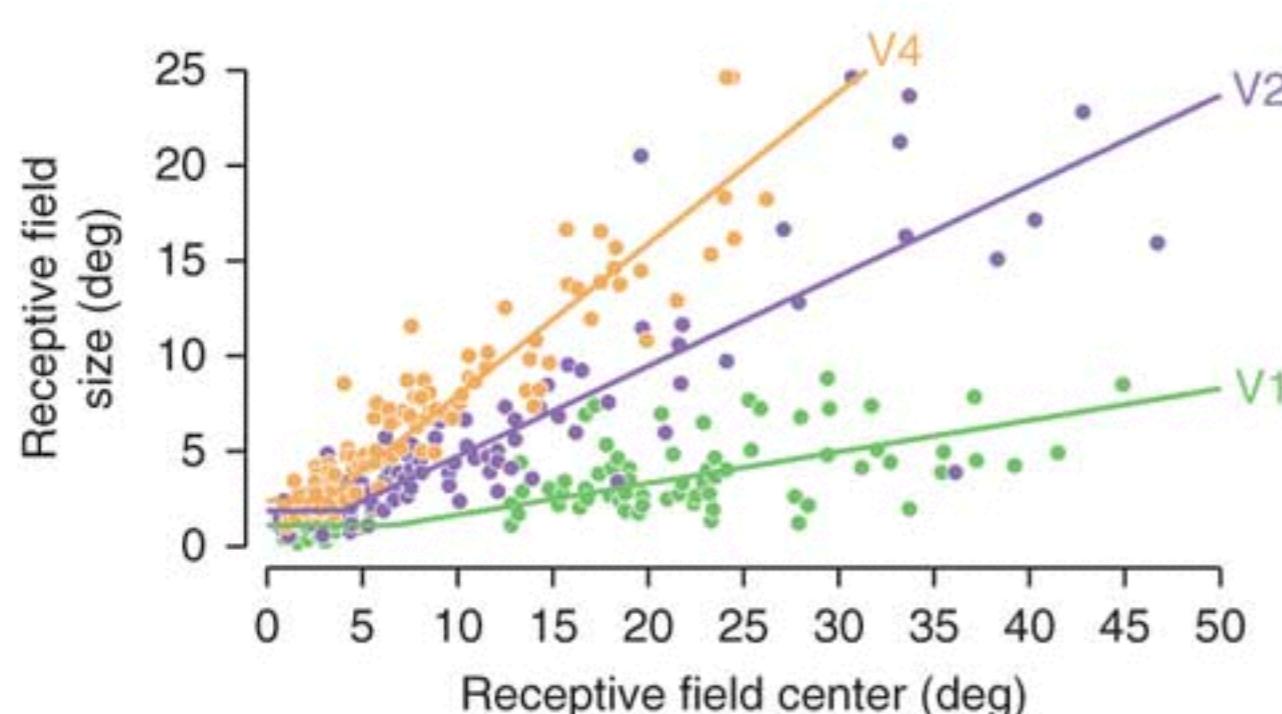
DiCarlo & Cox (2007)

# Object Recognition: The “What” Pathway



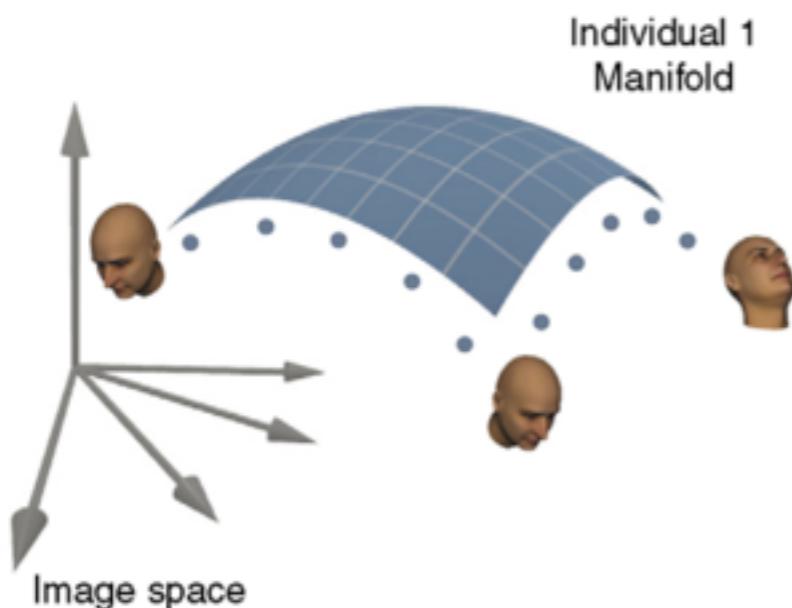
Freeman & Simoncelli (2011), Tanaka (1997)

# Object Recognition: The “What” Pathway



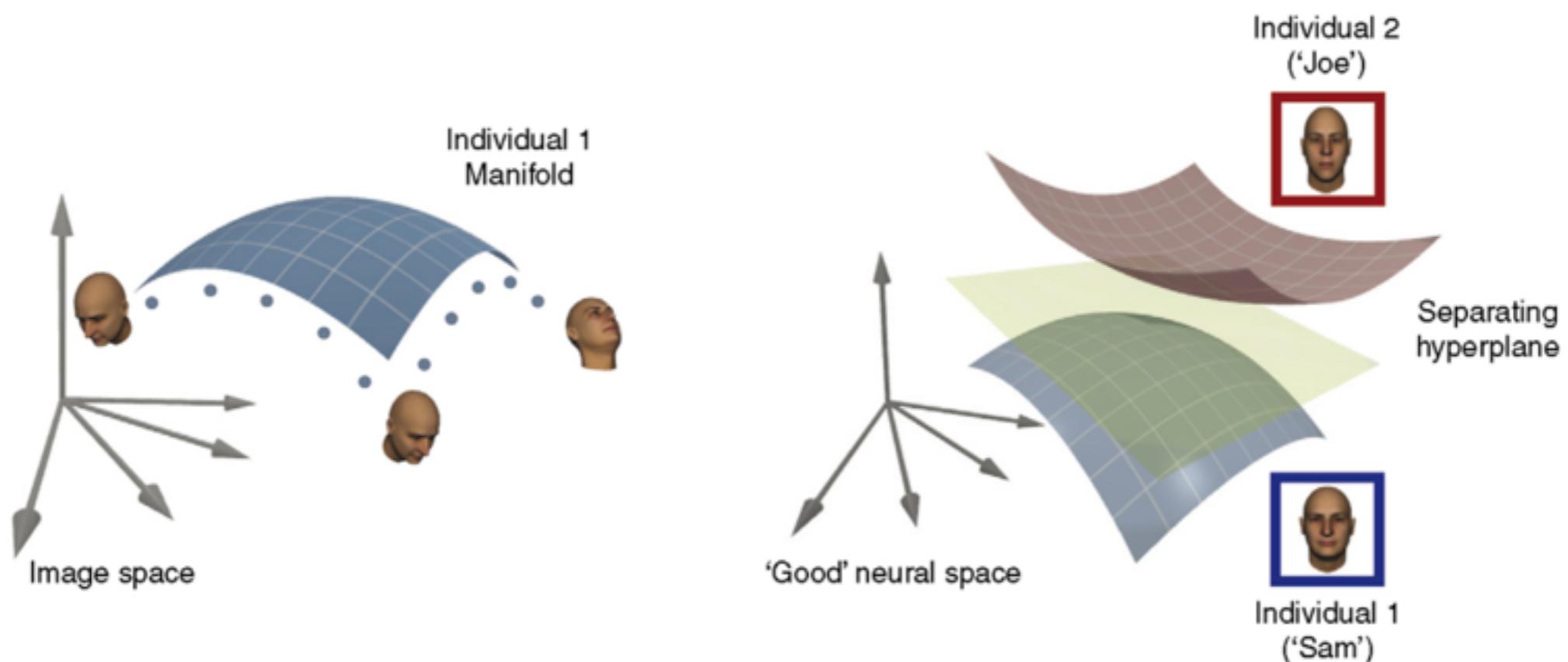
Freeman & Simoncelli (2011), Tanaka (1997)

# Object as Manifolds in High Dimensional Space



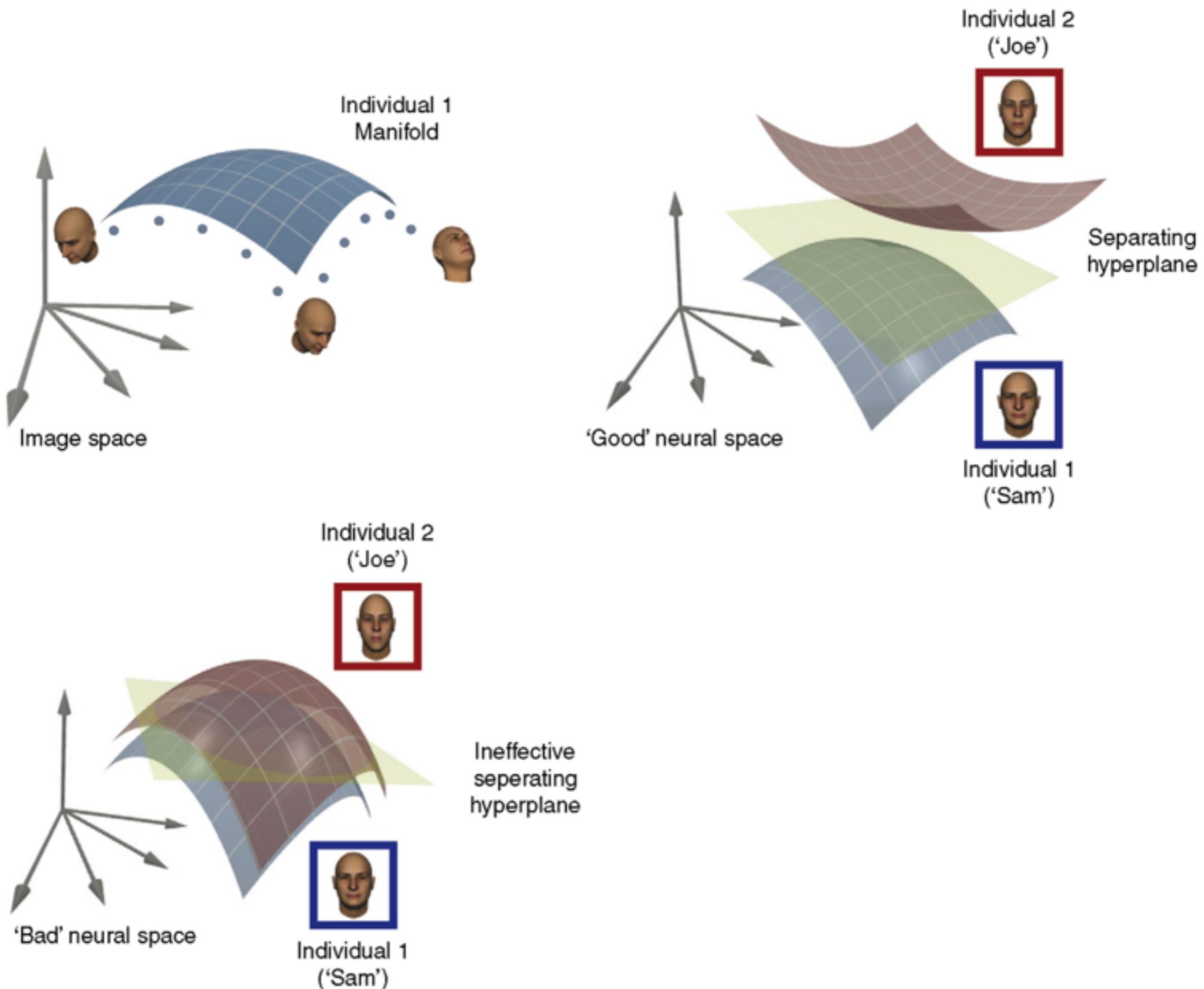
DiCarlo & Cox (2007)

# Object as Manifolds in High Dimensional Space



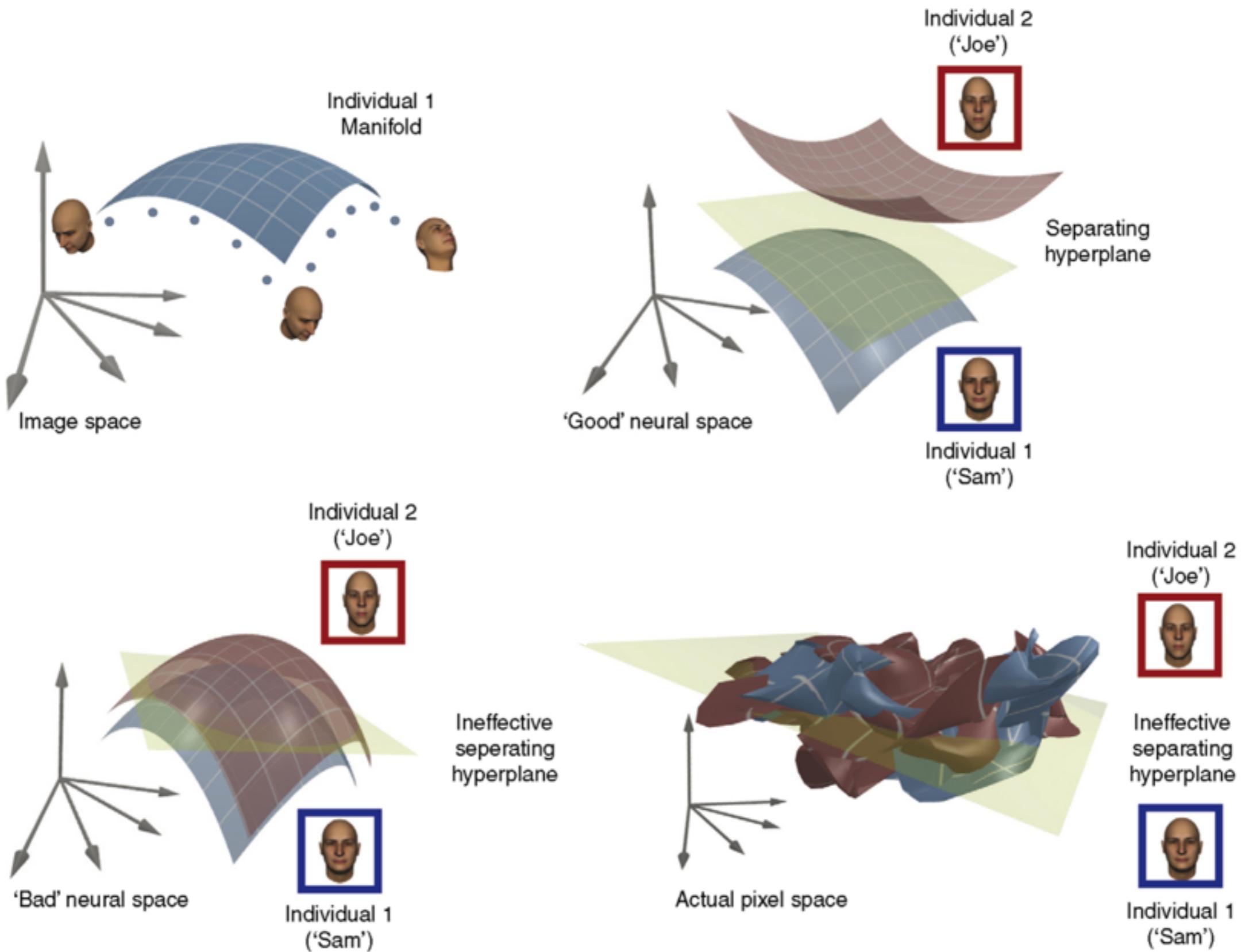
DiCarlo & Cox (2007)

# Object as Manifolds in High Dimensional Space



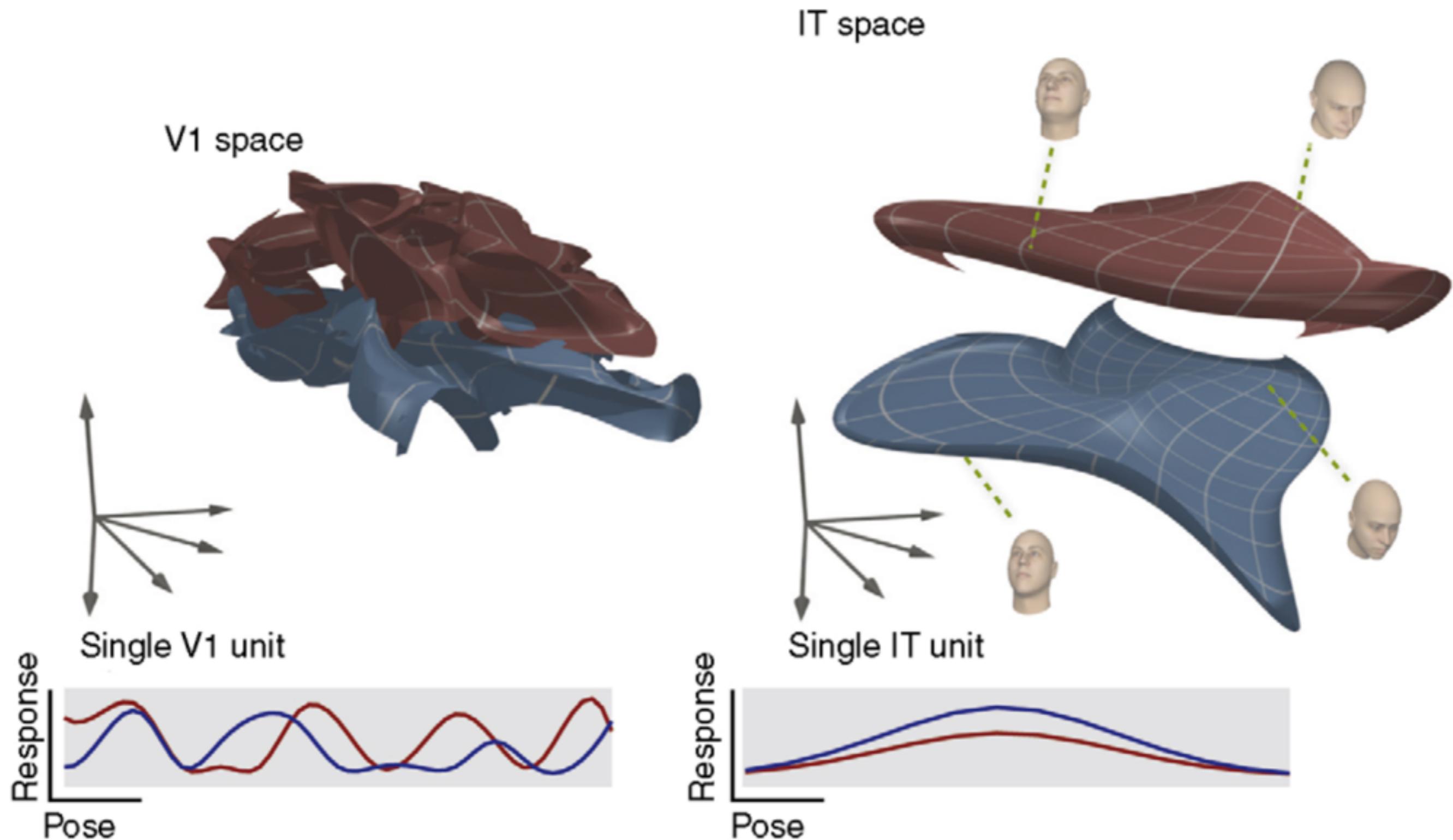
DiCarlo & Cox (2007)

# Object as Manifolds in High Dimensional Space



DiCarlo & Cox (2007)

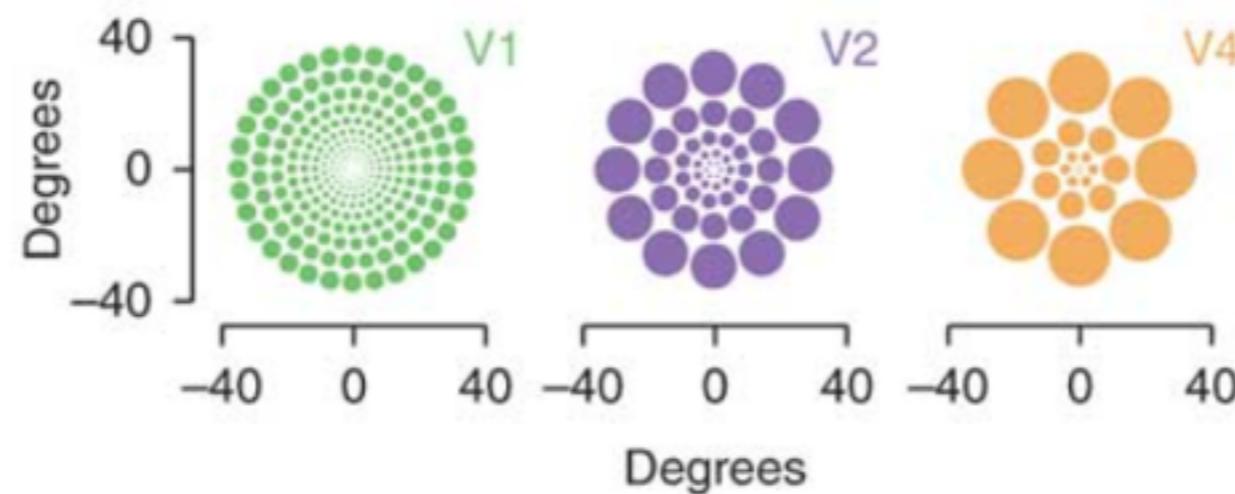
# Untangling Object Manifolds



DiCarlo & Cox (2007)

# Object Recognition: A Step Towards Visual Understanding

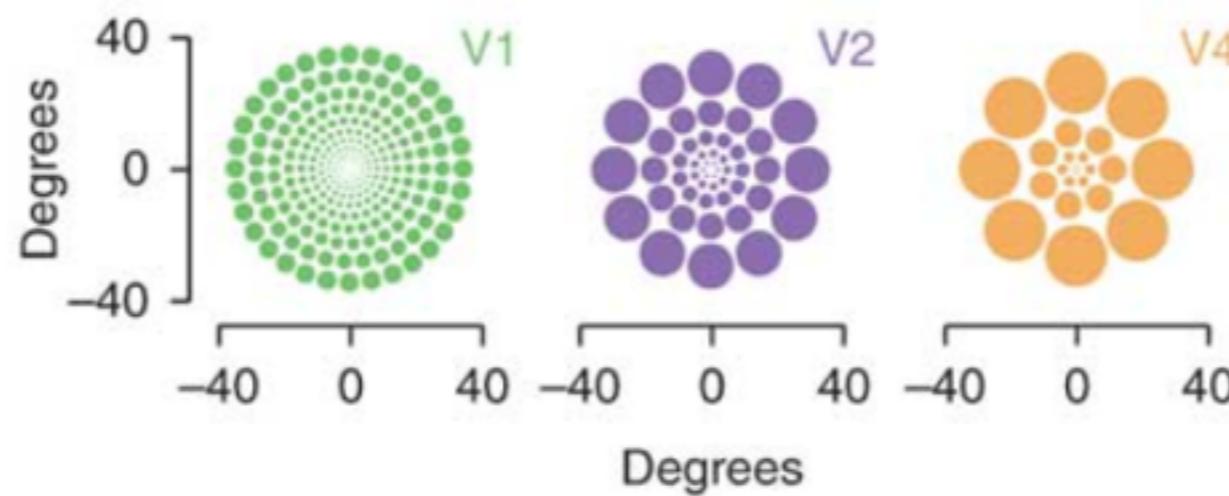
in the human visual system, invariance is built gradually across many successive transformations



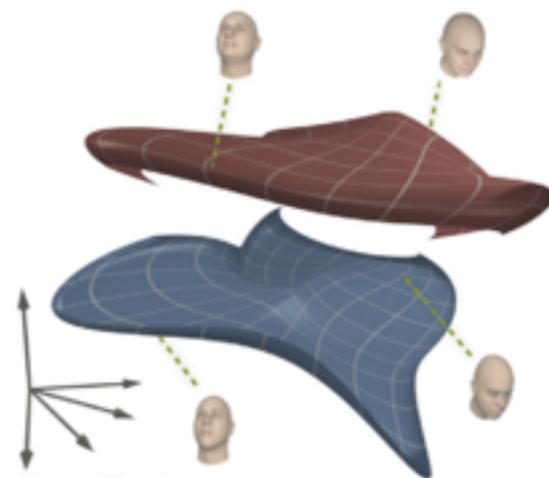
Freeman & Simoncelli (2011), DiCarlo & Cox (2007)

# Object Recognition: A Step Towards Visual Understanding

in the human visual system, invariance is built gradually across many successive transformations



human and computer vision both strive to achieve invariant representations



Freeman & Simoncelli (2011), DiCarlo & Cox (2007)

# **Discussion**

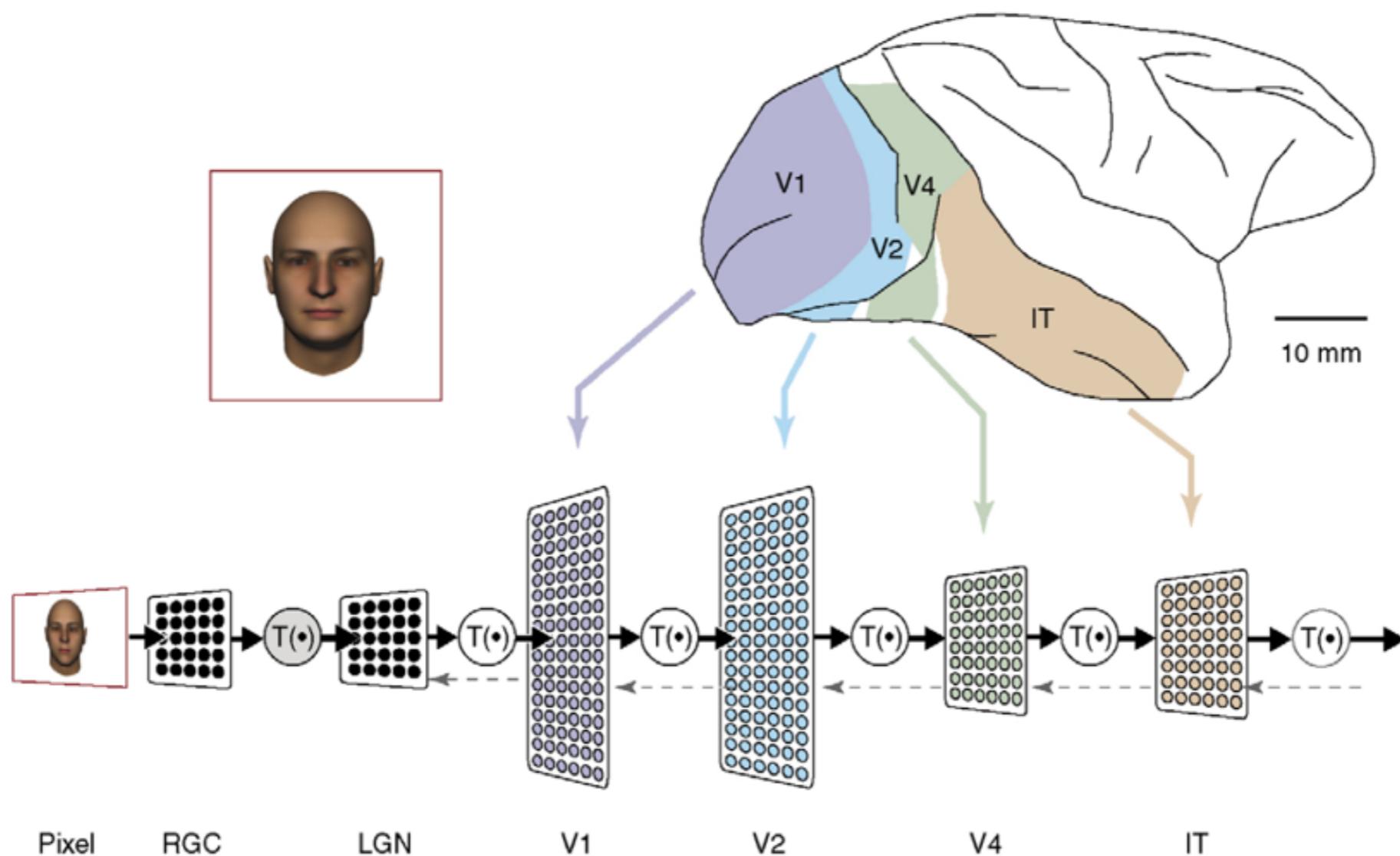
# Discussion

human visual invariance vs. CV features

incremental invariance vs. all-at-once

should we build detectors invariant to “everything”?

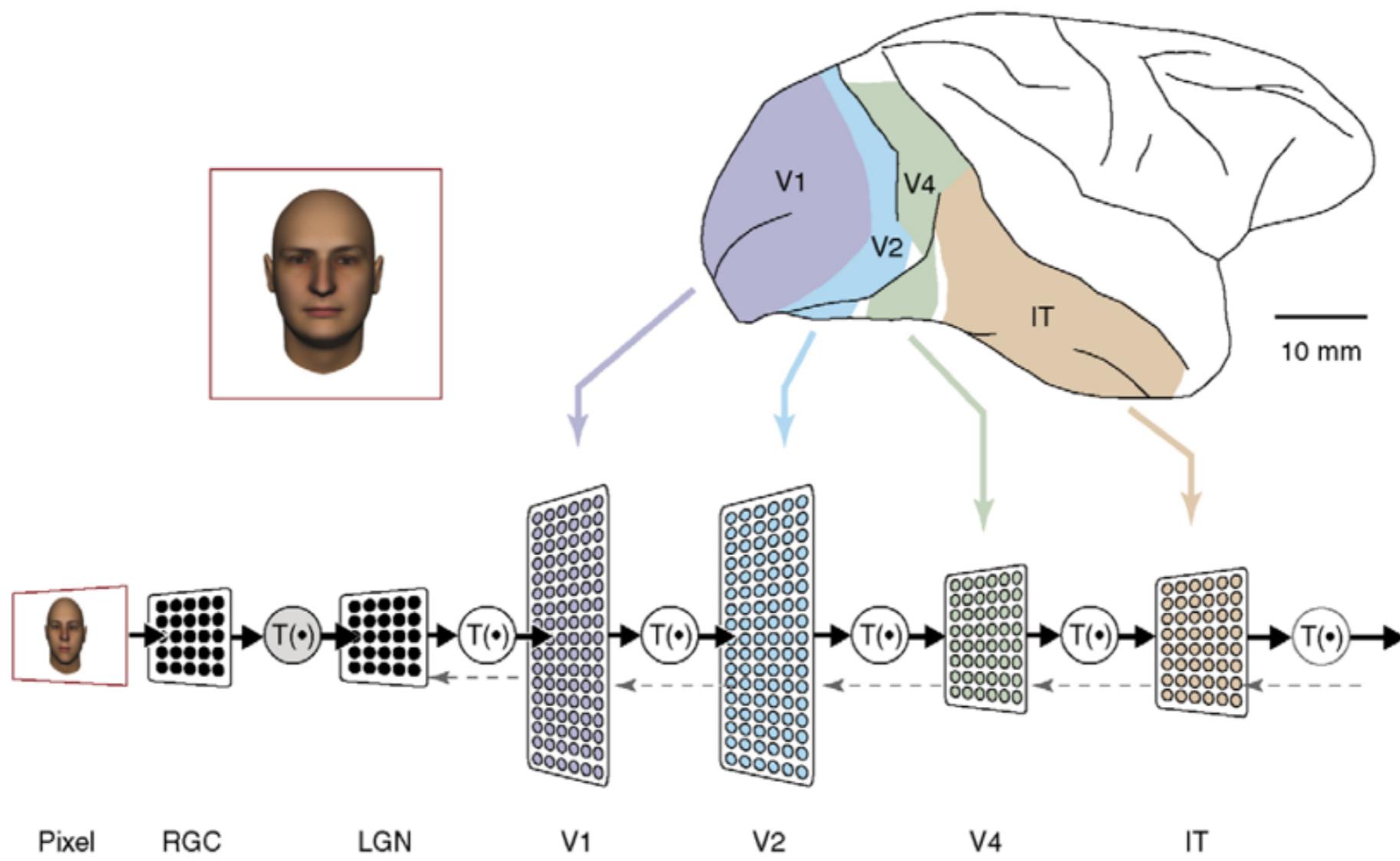
# Mammalian Visual System



DiCarlo & Cox (2007)

# Mammalian Visual System

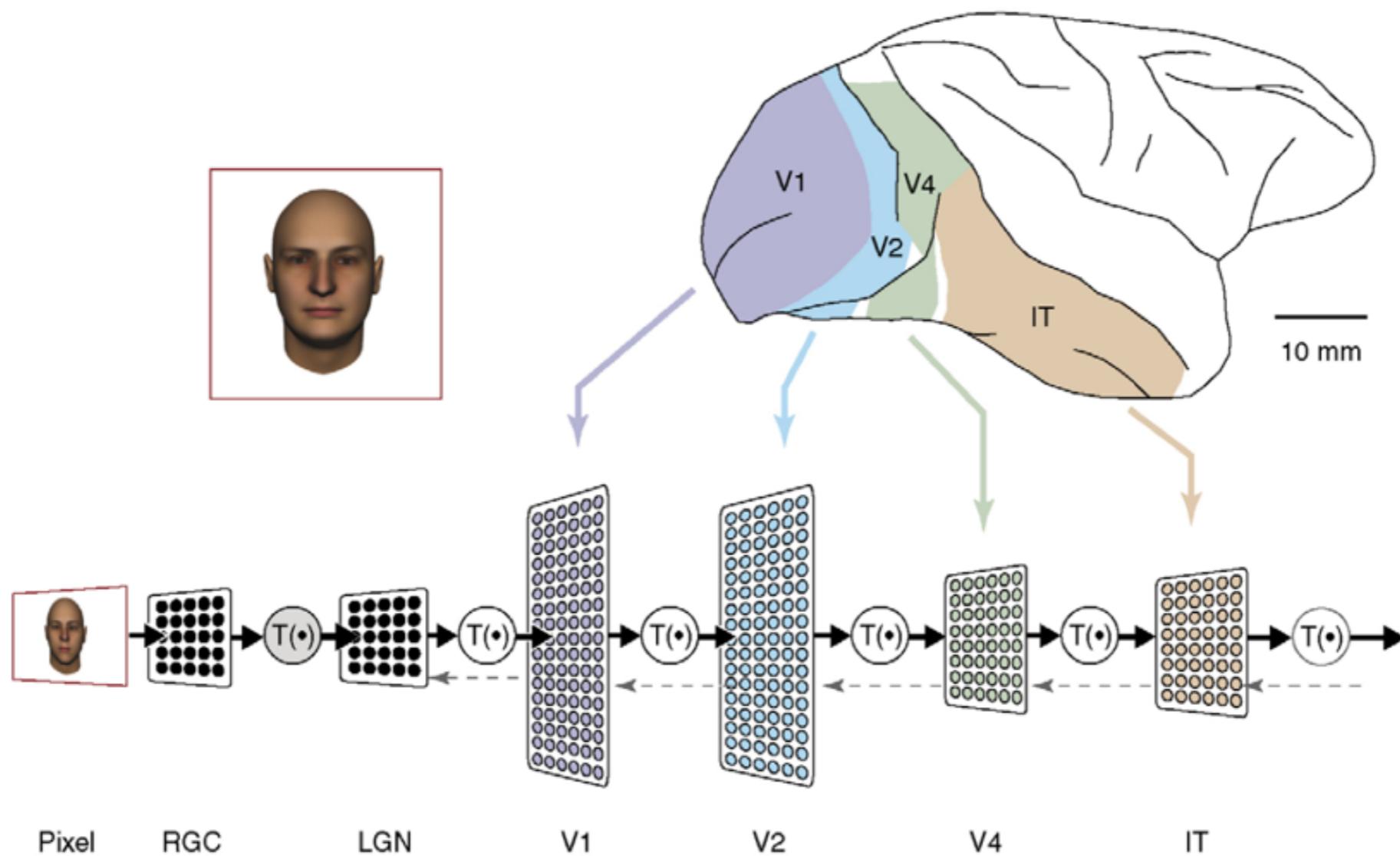
visual processing is done in stages



DiCarlo & Cox (2007)

# Mammalian Visual System

visual processing is done in stages  
but it is not synonymous with perception



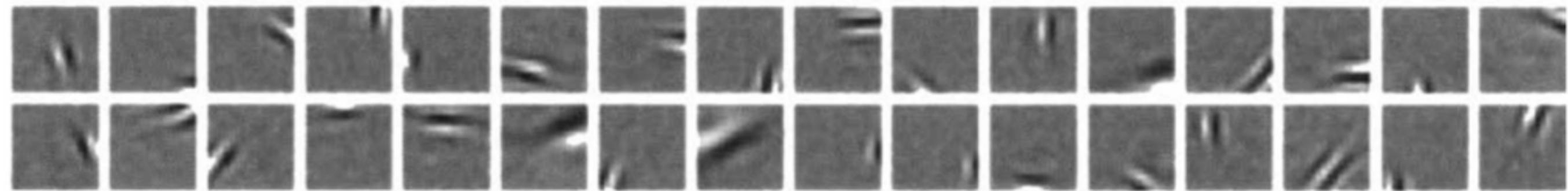
DiCarlo & Cox (2007)

# Early Computation

V1 cells encode edge orientation and position across visual field: simple and complex receptive fields



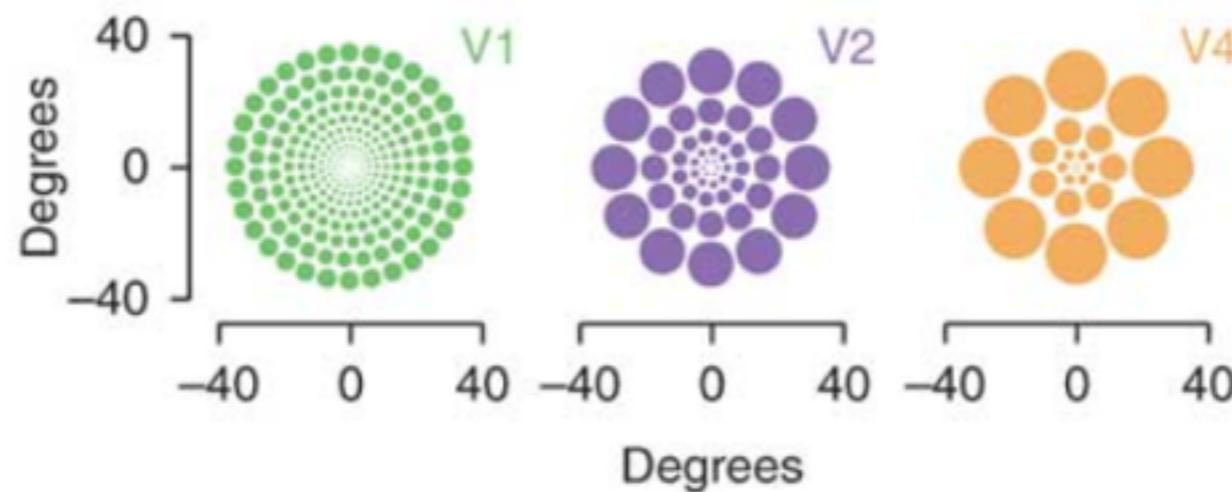
edges are a sufficient basis for real-world images!



Hubel & Wiesel (1962), Olshausen & Field (1996)

# Object Recognition: A Step Towards Visual Understanding

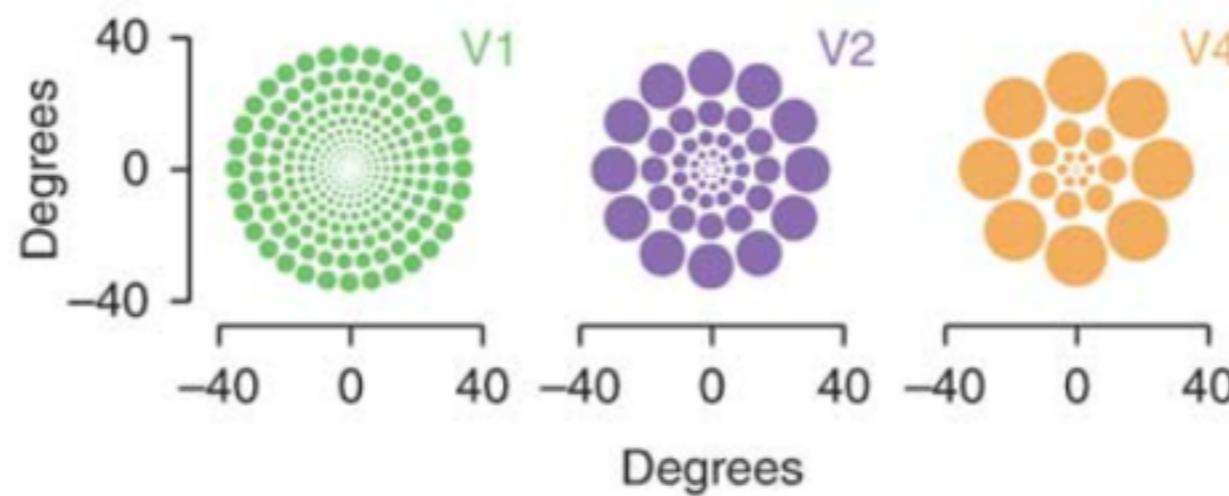
in the human visual system, invariance is built gradually across many successive transformations



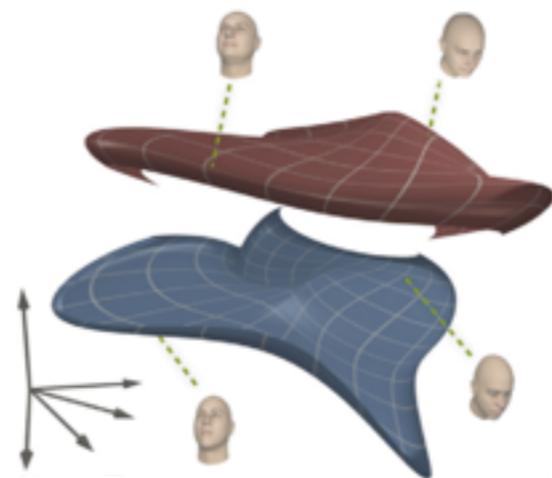
Freeman & Simoncelli (2011), DiCarlo & Cox (2007)

# Object Recognition: A Step Towards Visual Understanding

in the human visual system, invariance is built gradually across many successive transformations



human and computer vision both strive to achieve invariant representations



Freeman & Simoncelli (2011), DiCarlo & Cox (2007)

# **Mid-Quarter Feedback Forms**