Abstracting the Computational Principles That Give Rise to Sensory Experience

Joshua Bowren

Where does experience come from

- **The brain?**
  - Specific location?
  - Everywhere?
- **Is experience equivalent to consciousness?**
  - Some thought the soul resides in the heart
  - But “heart” has several meanings
  - Starts to get philosophical

- **Studying conscious experience directly is hard**

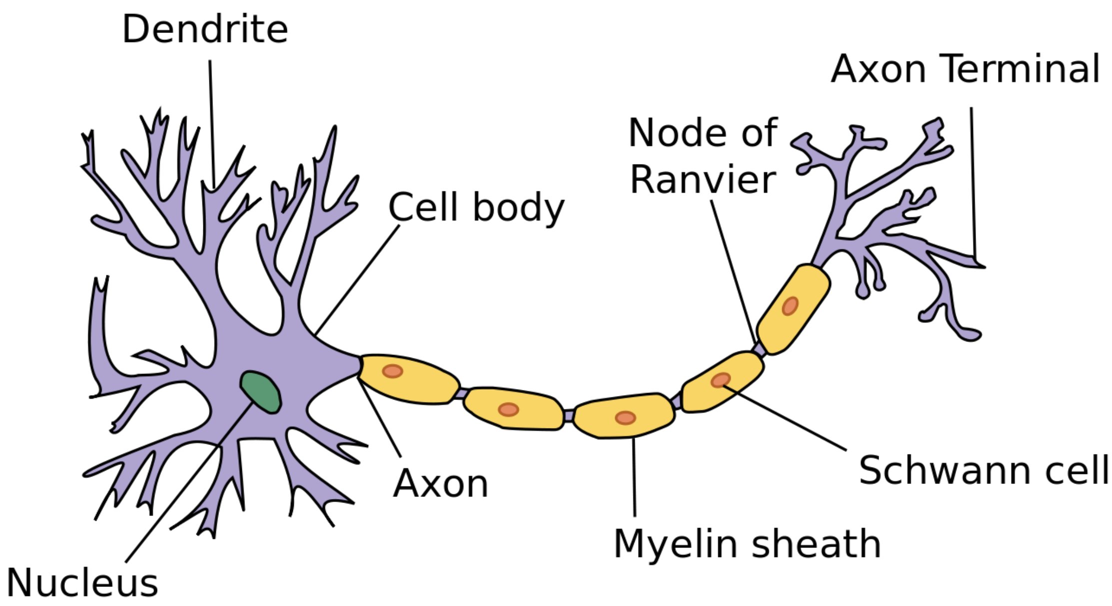
- **Perhaps we can study the processes that lead to sensory experience**
  - Study sensory tissue in the brain

Approaches

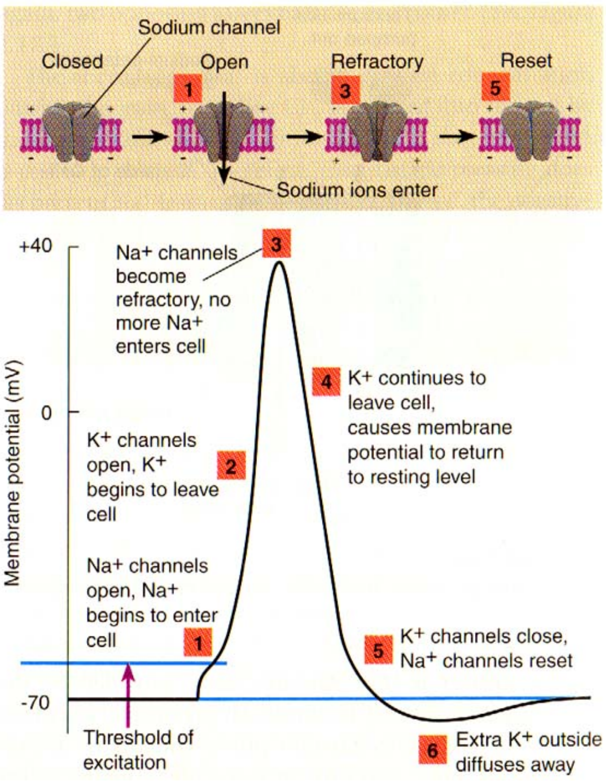
- **Neuroanatomy**
  - Studying morphology
- **Neurobiology**
  - Recording neuron spiking activity
  - Noninvasive methods (e.g. fMRI, EEG)
- **Psychophysics**
  - Measuring human cognitive abilities
- **Computational modeling**
  - Programs that implement mathematical models

What do we mean by sensory experience

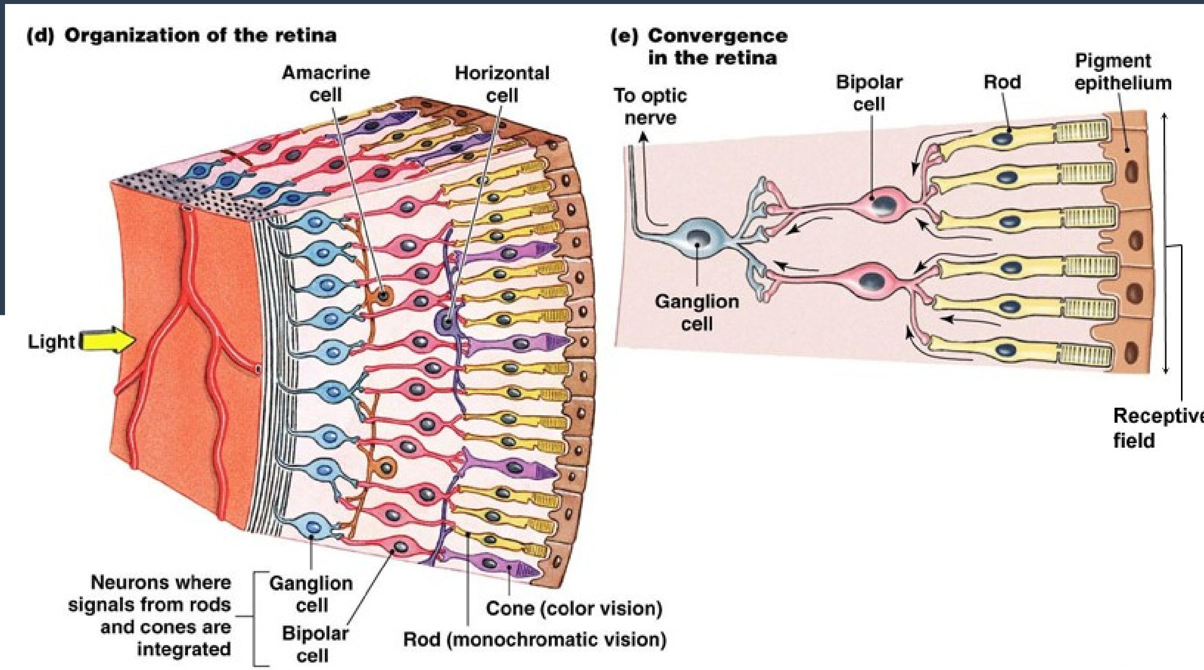
- **Vision**
- **Audition (Hearing)**
- **Somatosensation (Touch)**
- **Olfaction (Smell)**
- **Gustation (Taste)**



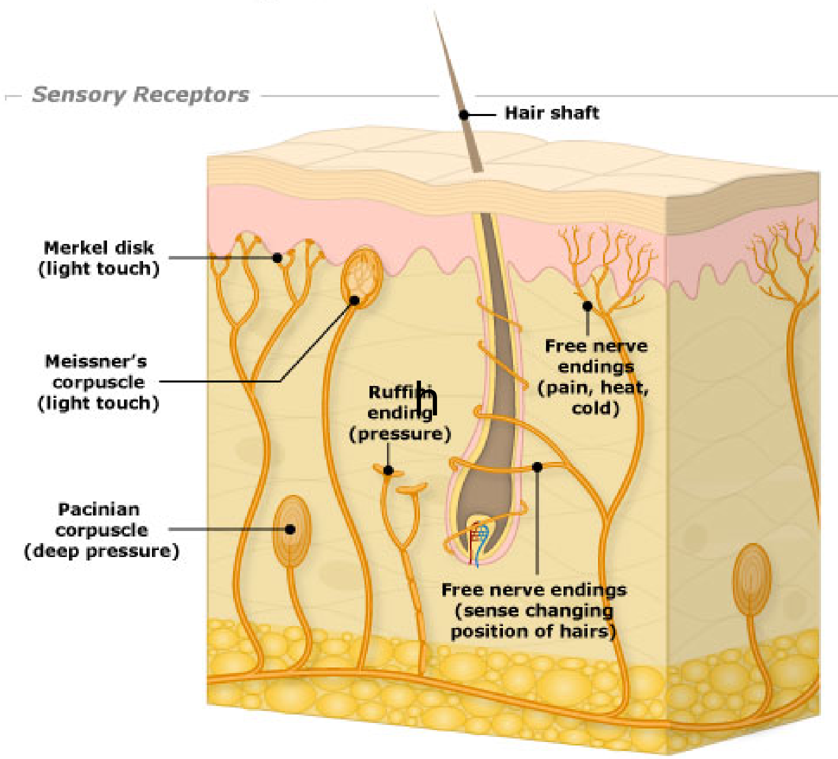
https://upload.wikimedia.org/wikipedia/commons/thumb/b/b5/Neuron.svg/1280px-Neuron.svg.png

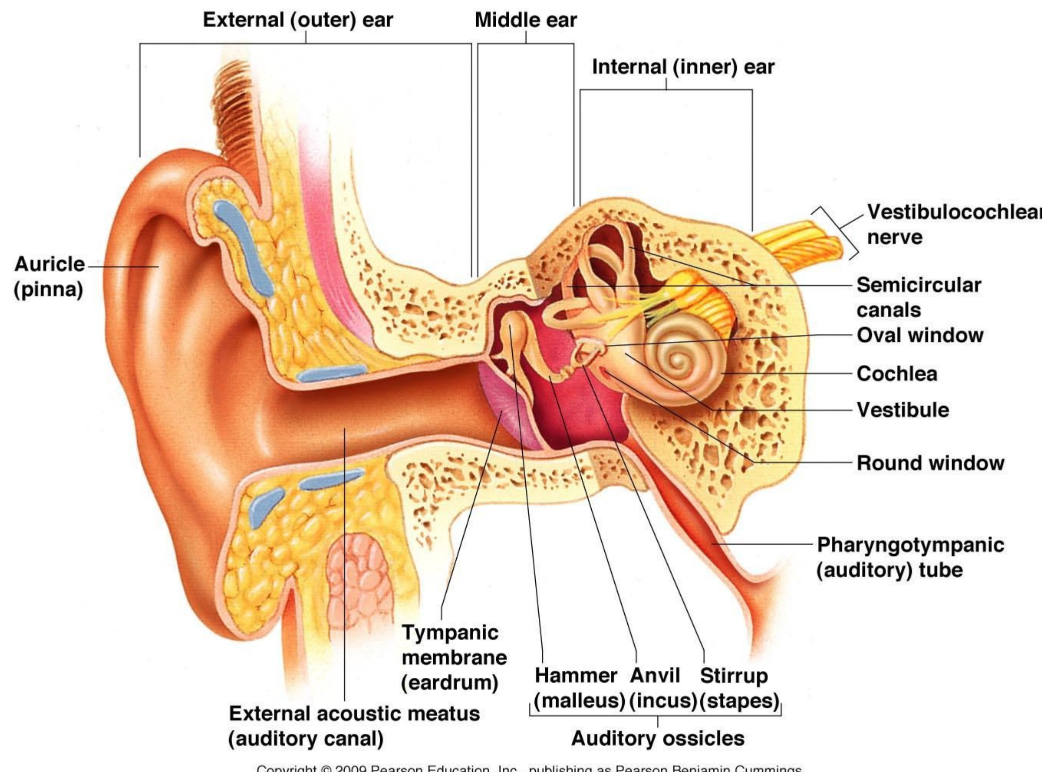


https://2.bp.blogspot.com/-mGTtnGanaP0/Txy_ShfDfgI/AAAAAAAAASg/kUegSPnKJ_k/s1600/action_potential.jpg



https://upload.wikimedia.org/wikipedia/commons/thumb/b/b5/Neuron.svg/1280px-Neuron.svg.png





https://online.science.psu.edu/sites/default/files/bisc004/content/sensory_receptors.jpg

https://daydreamanatomy.com/wp-content/uploads/2017/12/diagram-of-inner-ear-unlabeled-diagram-of-human-inner-ear-ear-anatomy-diagram-blank.jpg

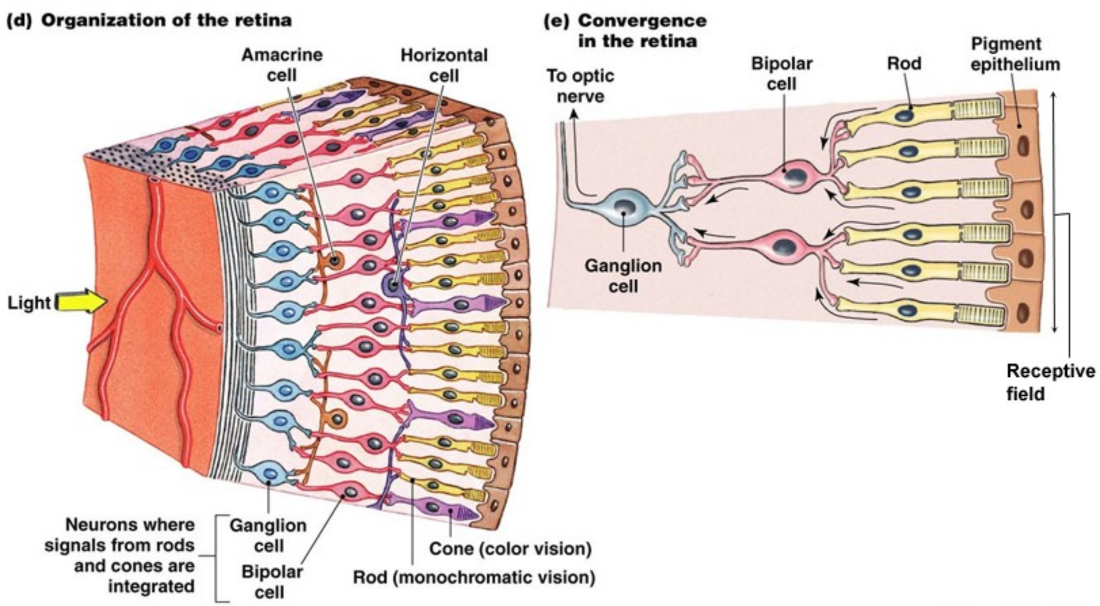
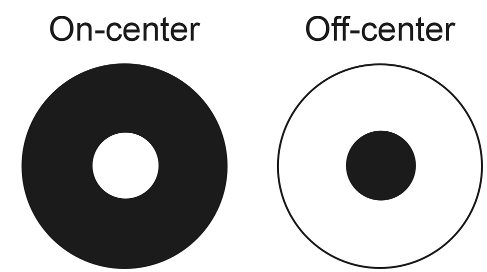
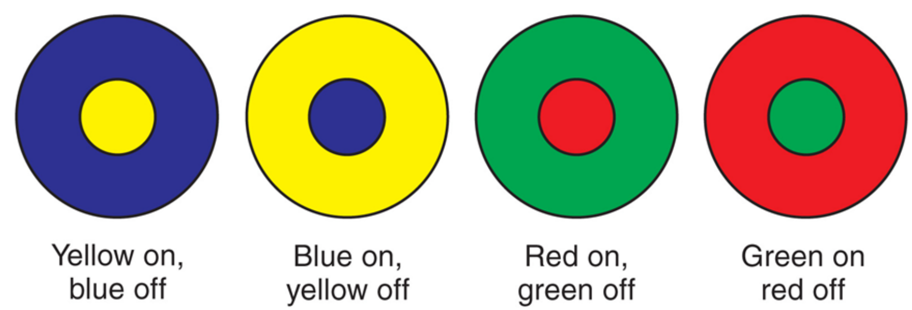


Fig. 10-35

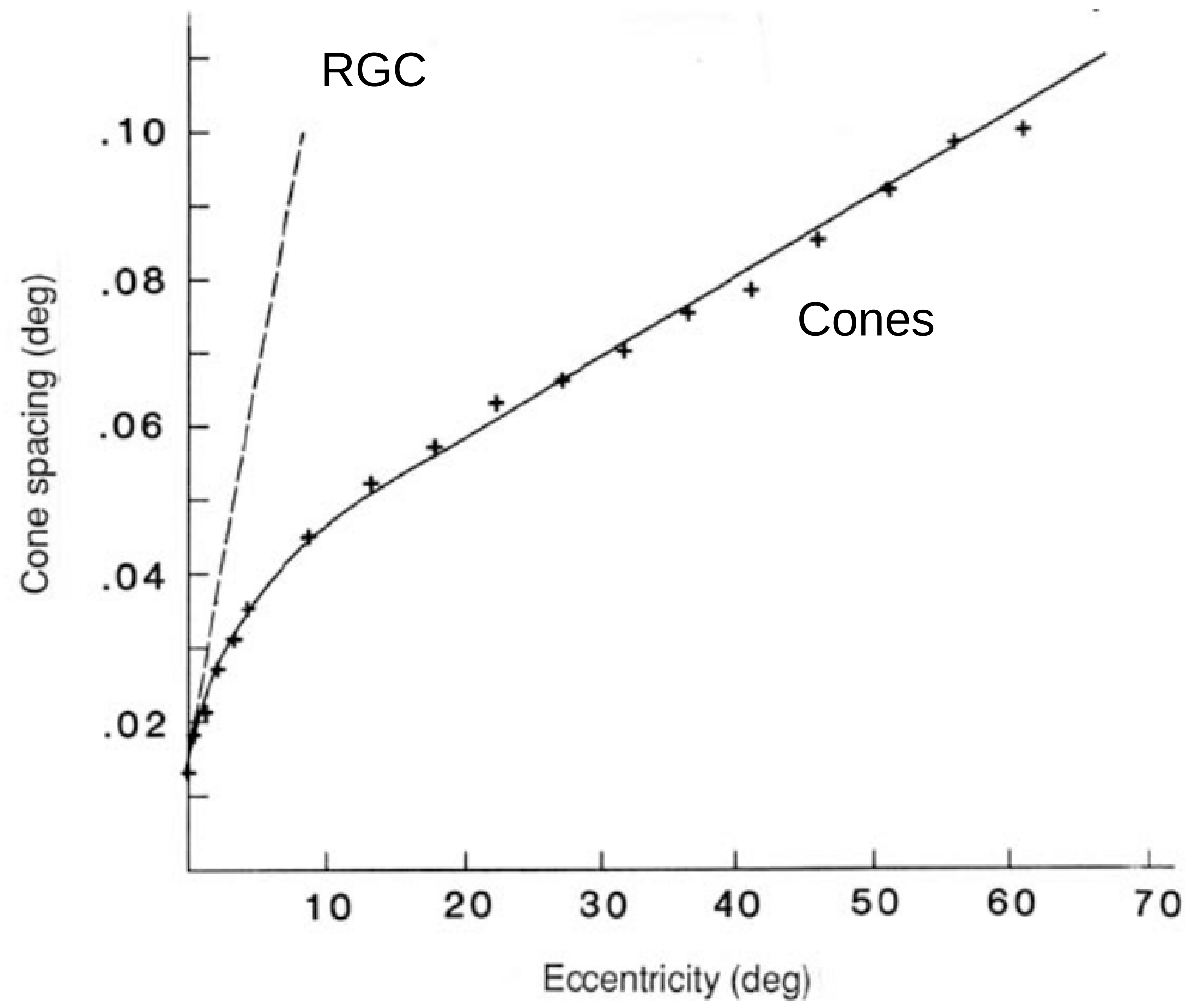
https://upload.wikimedia.org/wikipedia/commons/thumb/b/b5/Neuron.svg/1280px-Neuron.svg.png

http://www.scholarpedia.org/w/images/4/4b/Receptive_field_figure2.jpg

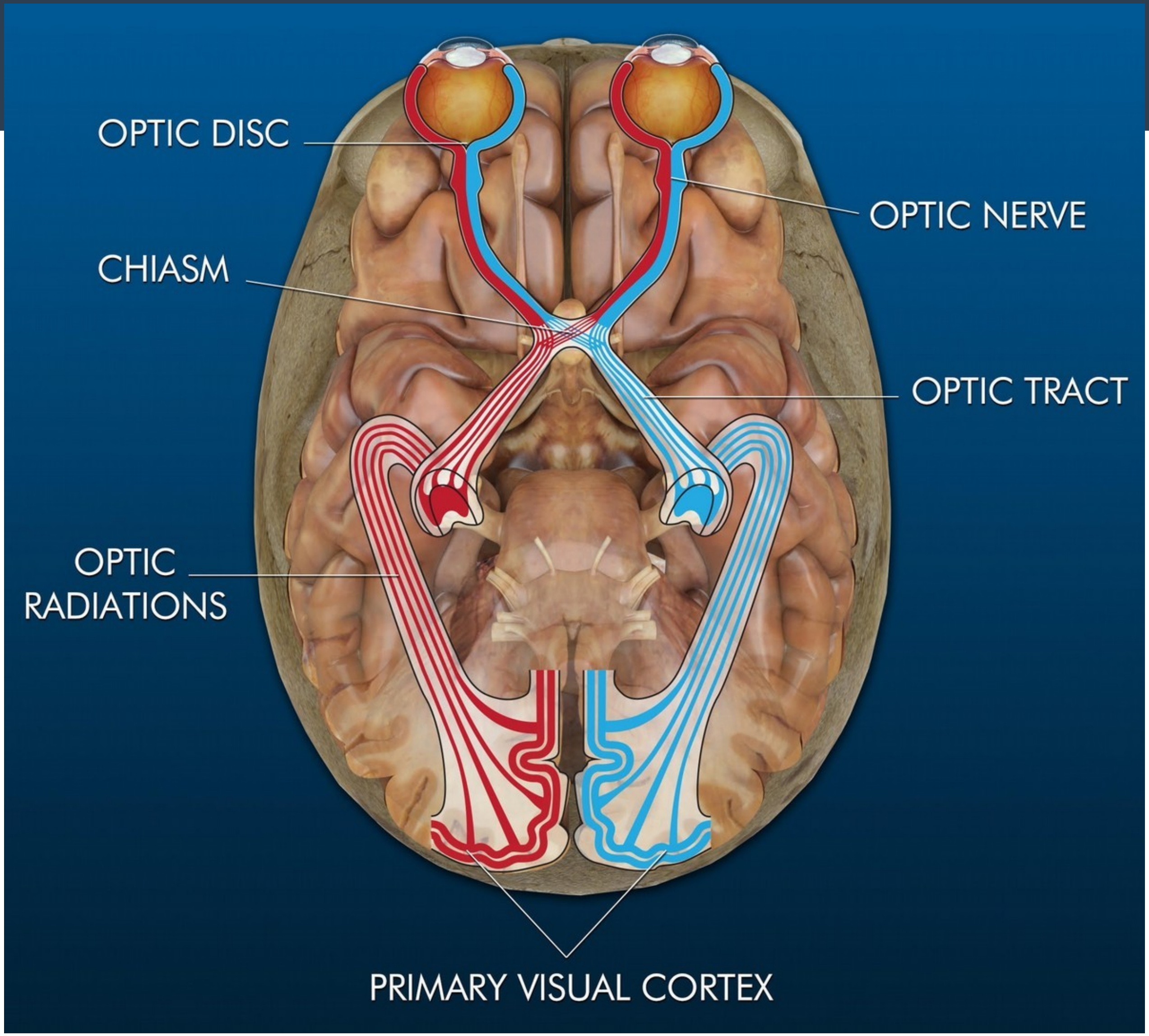




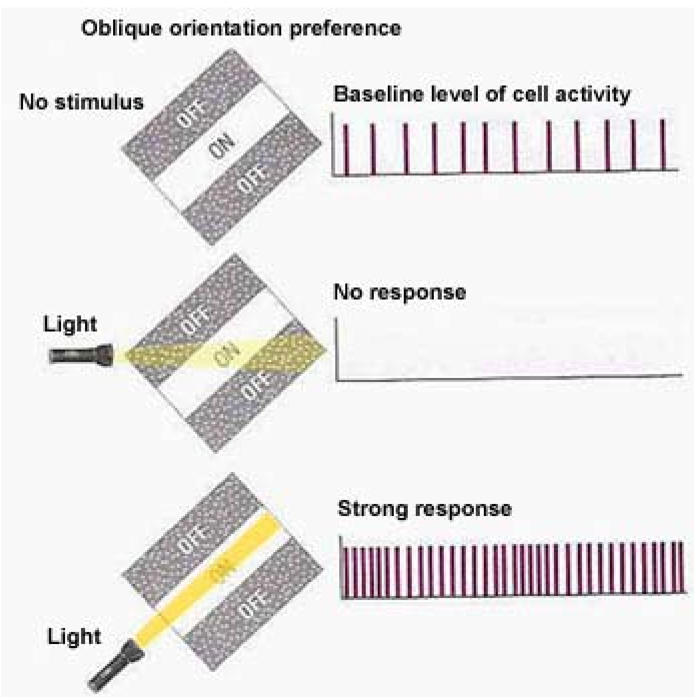
https://classconnection.s3.amazonaws.com/764/flashcards/151764/png/receptive_field_for_color_cells1352764540211.png

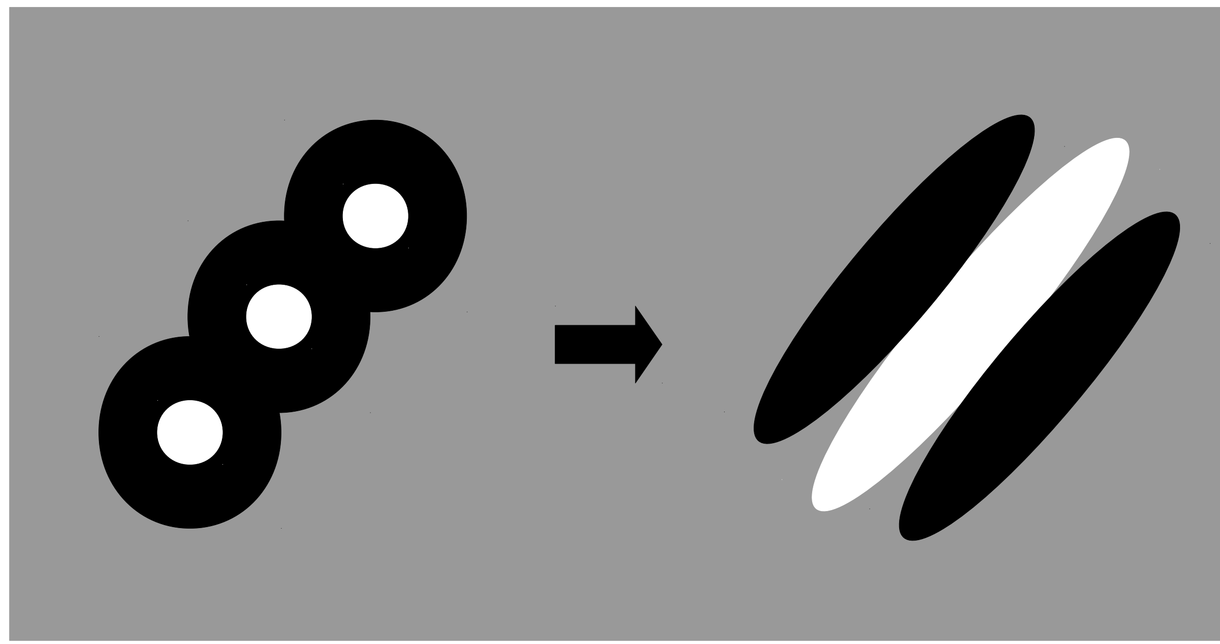


Andersen & Van Essen (1995)

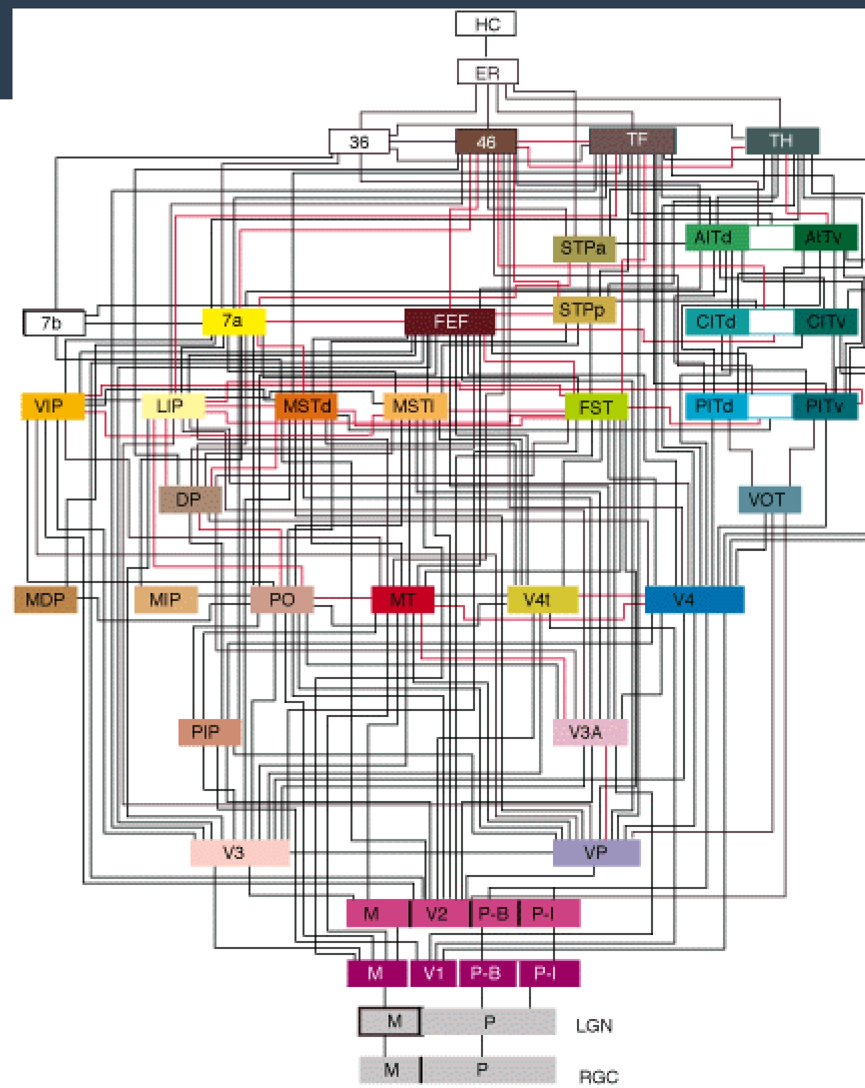


http://www.ophthalmologytraining.com/images/Stills/Visual%20Pathway.jpg





http://toritris.weebly.com/uploads/1/4/1/3/14134854/7597815_orig.jpg



http://www.cse.yorku.ca/~billk/images/VisHierarchy.gif

Artifacts of neural computation

https://www.gla.ac.uk/schools/humanities/research/philosophyresearch/cspe/illusions/#/figuresforproducingafter-images

Artifacts of neural computation

https://www.gla.ac.uk/schools/humanities/research/philosophyresearch/cspe/illusions/#/figuresforproducingafter-images

Artifacts of neural computation

http://nivea.psycho.univ-paris5.fr/ASSChtml/kayakflick.gif

Artifacts of Neural Computation

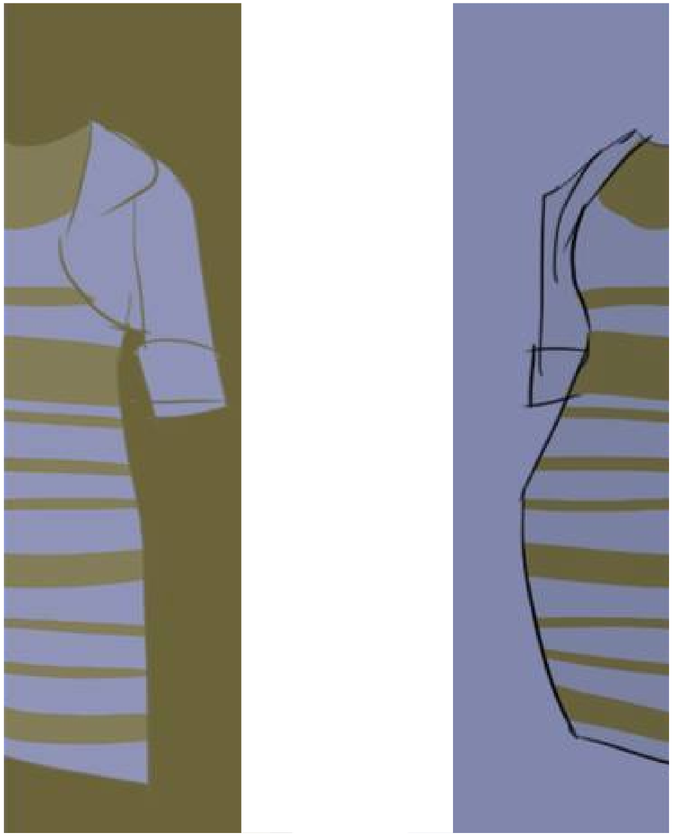
Artifacts of neural computation

http://nivea.psycho.univ-paris5.fr/ASSChtml/kayakflick.gif

Artifacts of neural computation

https://petapixel.com/assets/uploads/2015/02/Ios1eOY.jpg

Artifacts of neural computation



https://petapixel.com/assets/uploads/2015/02/Ios1eOY.jpg

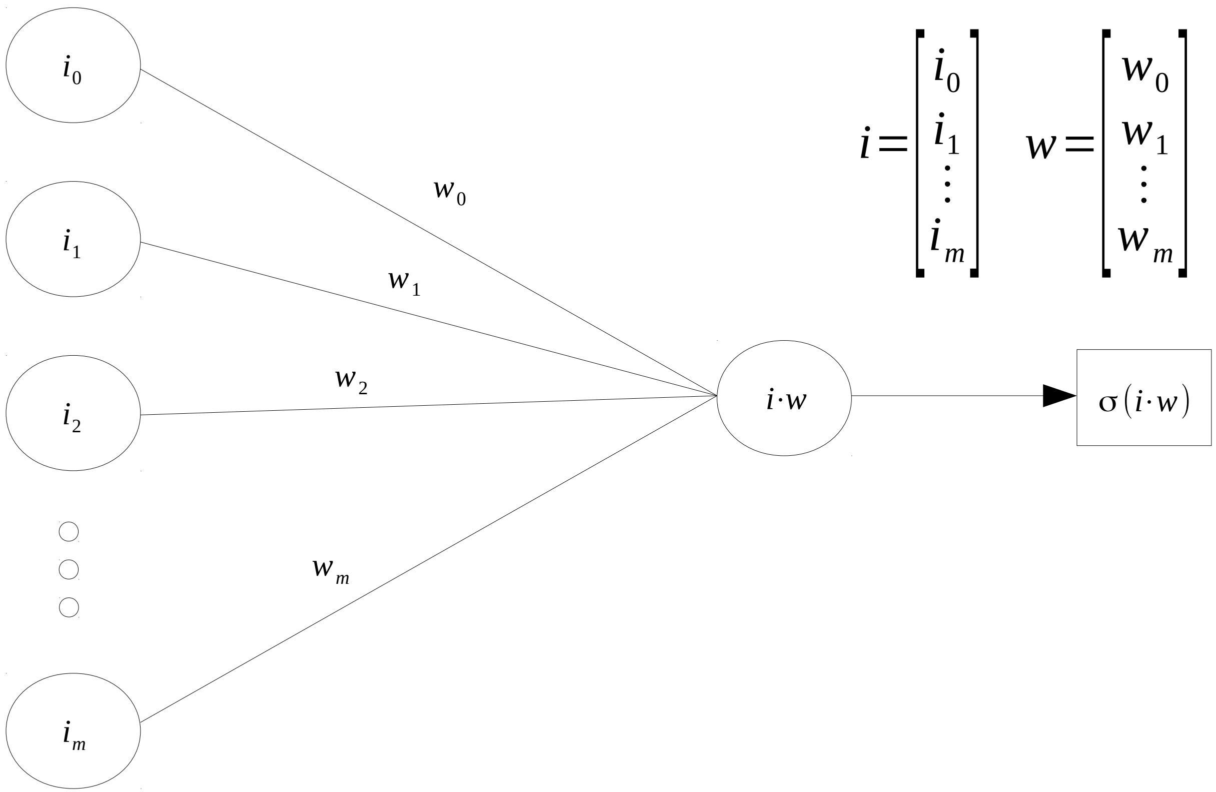
Artifacts of neural computation

https://www.gla.ac.uk/schools/humanities/research/philosophyresearch/cspe/illusions/#/change-blindness,colourcontrasteffects

Models of Neural Systems

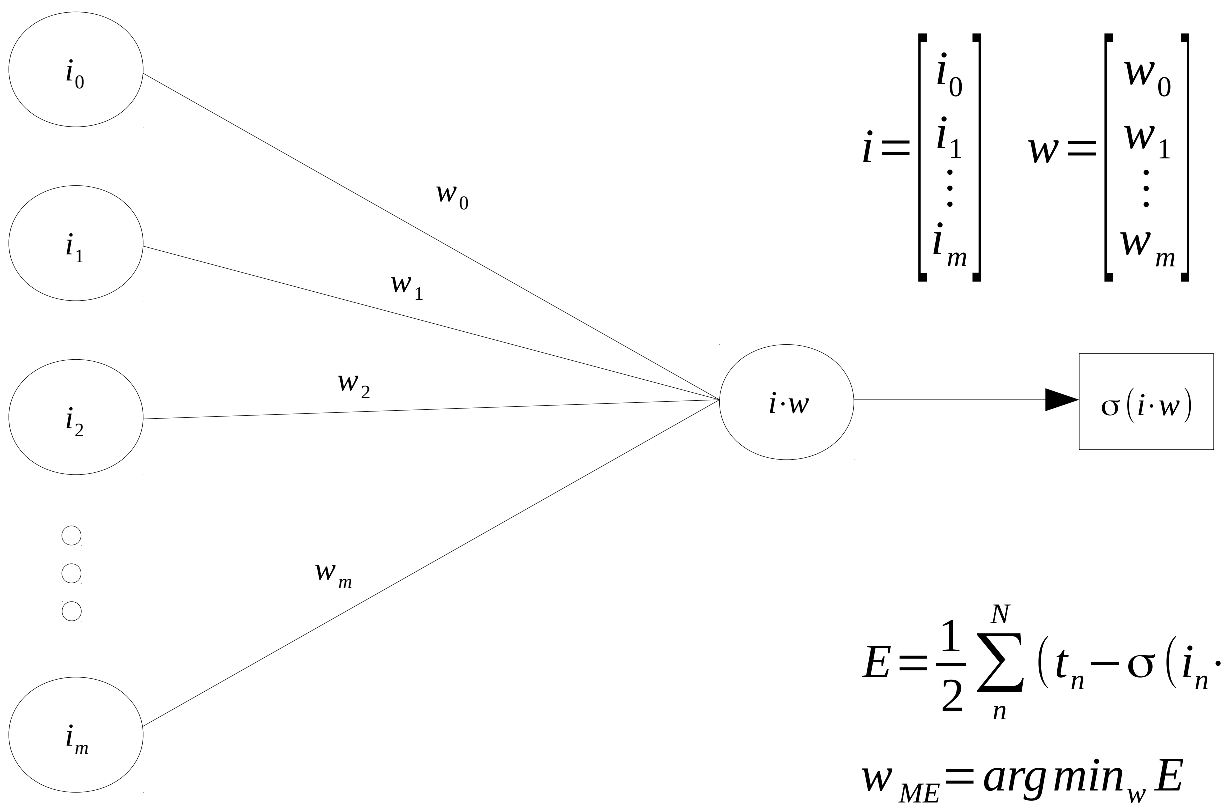
- **Artificial neural networks**
- **Computational models based on theories of neural computation**

Artificial neural networks



$$i = \begin{bmatrix} i_0 \\ i_1 \\ \vdots \\ i_m \end{bmatrix} \quad w = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix}$$

Artificial neural networks



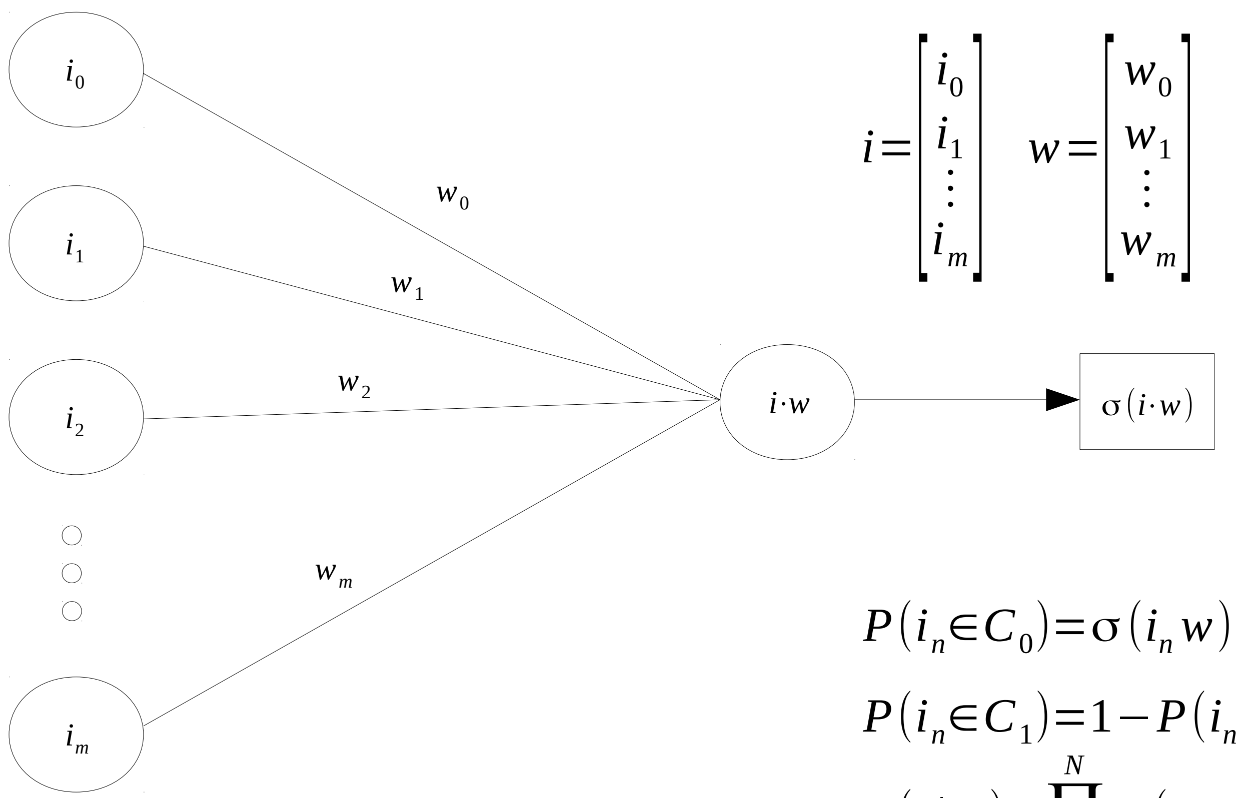
$$i = \begin{bmatrix} i_0 \\ i_1 \\ \vdots \\ i_m \end{bmatrix} \quad w = \begin{bmatrix} w_0 \\ w_1 \\ \vdots \\ w_m \end{bmatrix} \quad I = \begin{bmatrix} i_{00} & i_{01} & & i_{0m} \\ i_{10} & i_{11} & \dots & i_{1m} \\ \vdots & \vdots & & \vdots \\ i_{n0} & i_{n1} & & i_{nm} \end{bmatrix} = \begin{bmatrix} i_0 \\ i_1 \\ \vdots \\ i_n \end{bmatrix}$$

$$E = \frac{1}{2} \sum_{n}^{N} \left(t_n - \sigma\left(i_n \cdot w\right)\right)^2 = \frac{1}{2} \left\| t - \sigma\left(Iw\right) \right\|^2$$

$$w_{ME} = arg\,min_w E$$

$$w = w - \alpha \nabla_w E$$

Artificial neural networks



$$i=\begin{bmatrix} i_0\\ i_1\\ \vdots\\ i_m\end{bmatrix} \quad w=\begin{bmatrix} w_0\\ w_1\\ \vdots \\ w_m\end{bmatrix} \quad I=\begin{bmatrix} i_{00} & i_{01} & & i_{0m}\\ i_{10} & i_{11} & \dots & i_{1m}\\ \vdots & \vdots & & \vdots\\ i_{n0} & i_{n1} & & i_{nm}\end{bmatrix}=\begin{bmatrix} i_0\\ i_1\\ \vdots\\ i_n\end{bmatrix}$$

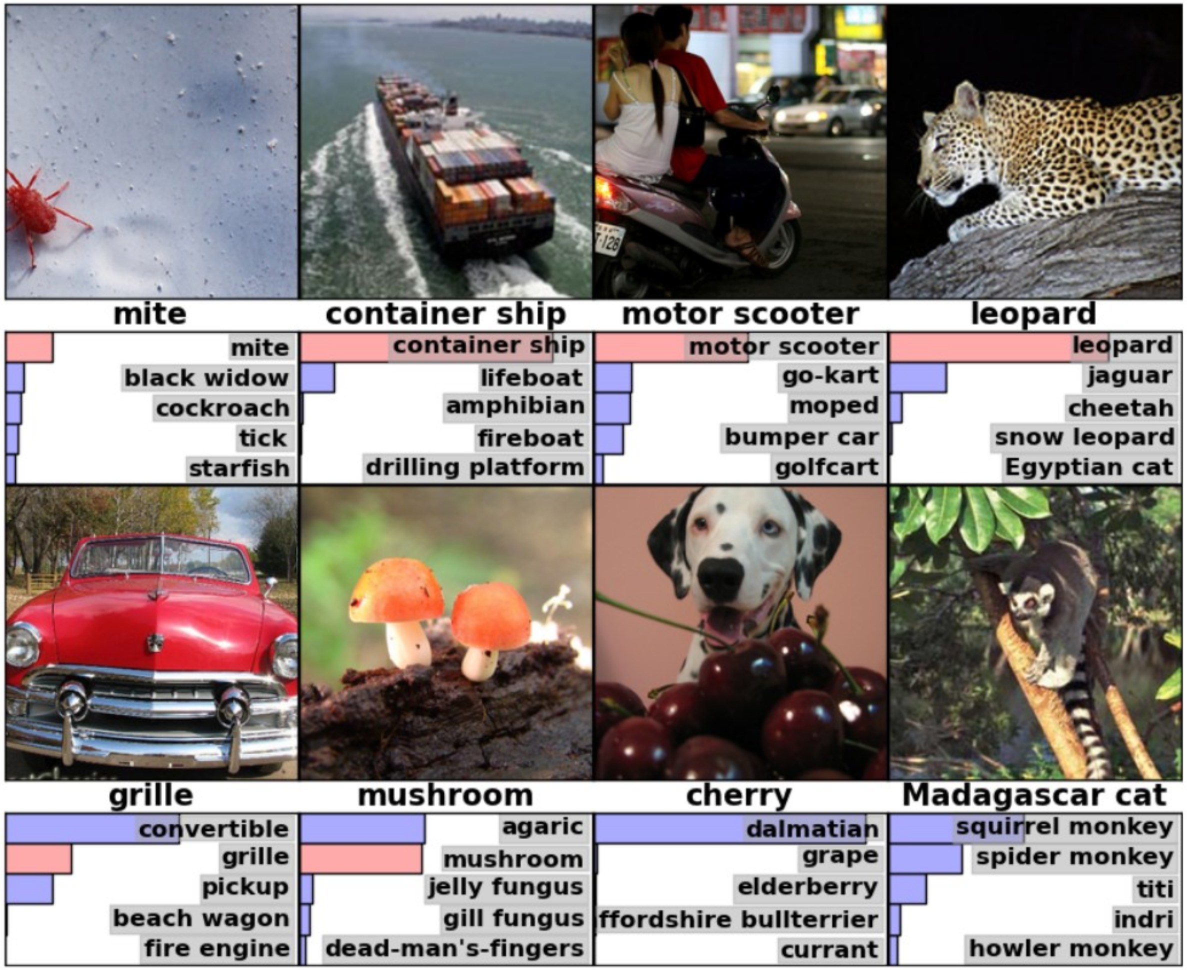
$$P(i_n \in C_0)=\sigma(i_n w)$$

$$P(i_n \in C_1)=1-P(i_n \in C_0)=1-\sigma(i_n w)$$

$$P(t|w)=\prod_n^N P(i_n \in C_0)^{t_n}(1-P(i_n \in C_0))^{1-t_n}$$

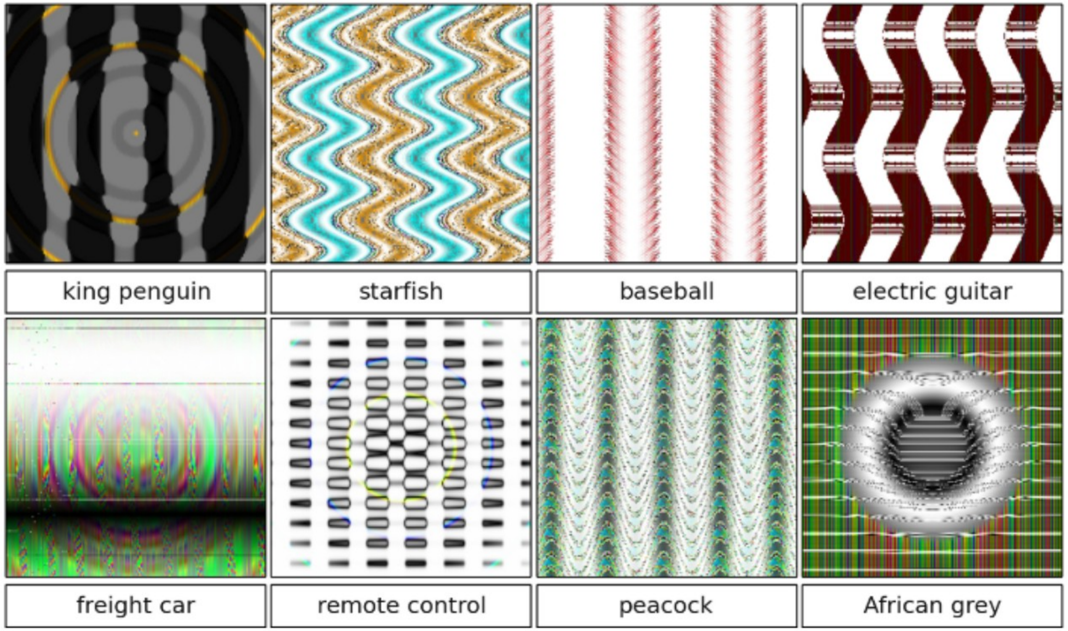
$$w_{ML}=arg\,max_w\, p(t|w)$$

$$w=w+\alpha \nabla_w p(t|w)$$

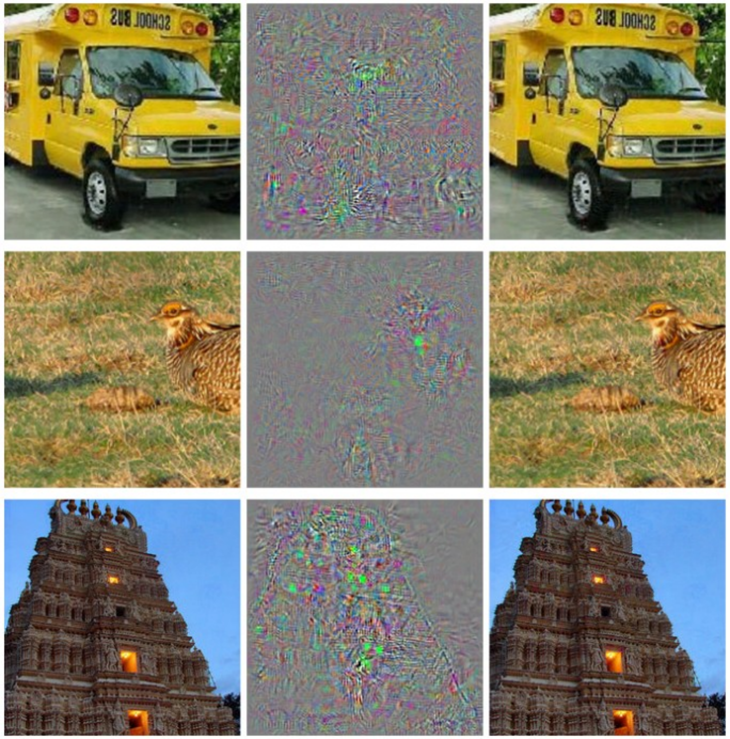


Krizhevsky, Sutskever, & Hinton (2012)

Neural Network Fooling



Nguyen, Yosinski, & Clune (2015)



Szegedy et al. (2014)

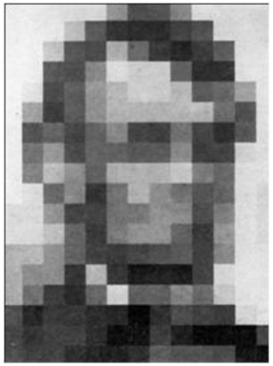
Computation and Theory Approaches

- **Efficient coding – Barlow (1961)**

- **Redundancy reduction and whitening – Atick, Li, & Redlich (1992)**

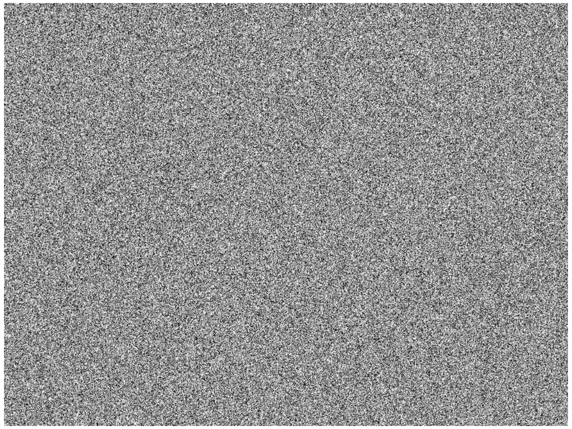
The retina as a low pass whitening filter

- **Spatial frequency**
  - Low spatial frequencies are important



Low SF

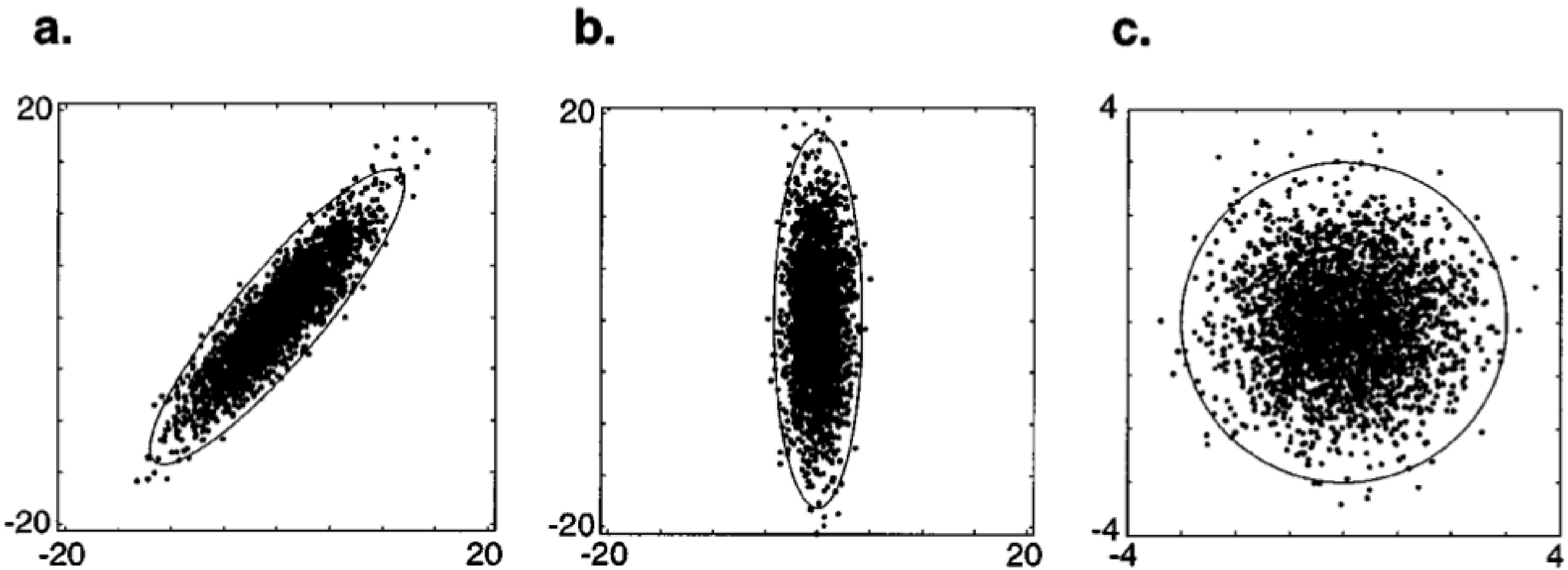
https://blogs.scientificamerican.com/illusion-chasers/files/2014/06/Lincoln.jpg



High SF

https://zuriest.files.wordpress.com/2011/11/white-noise2.jpg

- **Fitting in the light of the optic nerve bottleneck**



Simoncelli and Olshausen (2001)

In the cortex

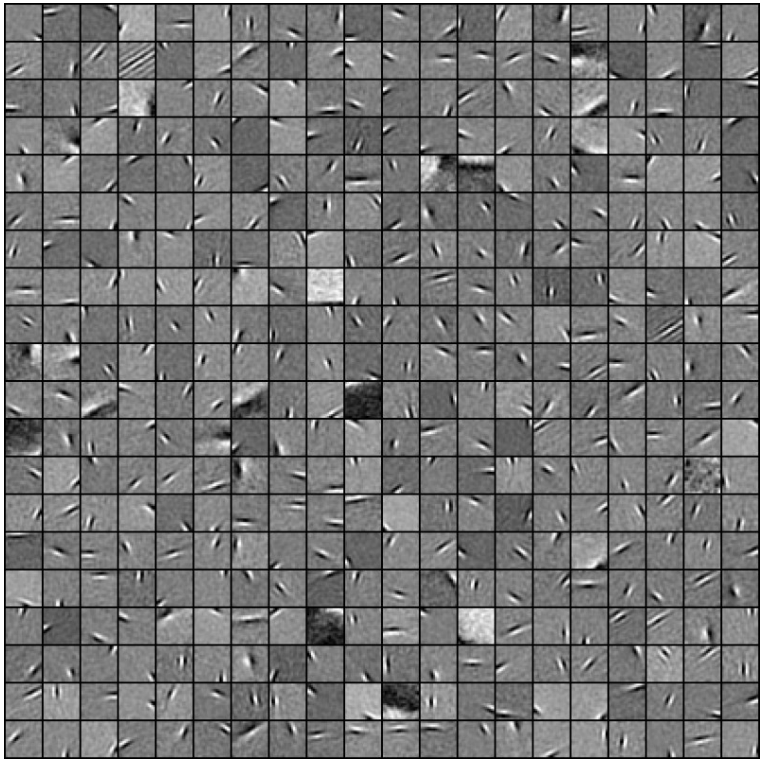
- **Sparse coding (Olshausen & Field, 1996)**
  - Model of V1

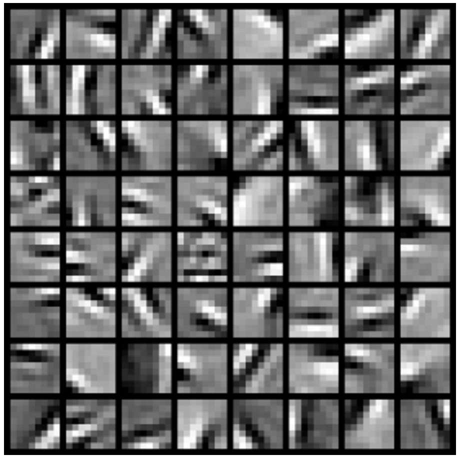
$$I = \sum_{n}^{N} a_n \Phi_n = \Phi a$$

$$a \sim Laplace(\lambda) \qquad a_{ML} = arg\,max_a P(a|I)$$

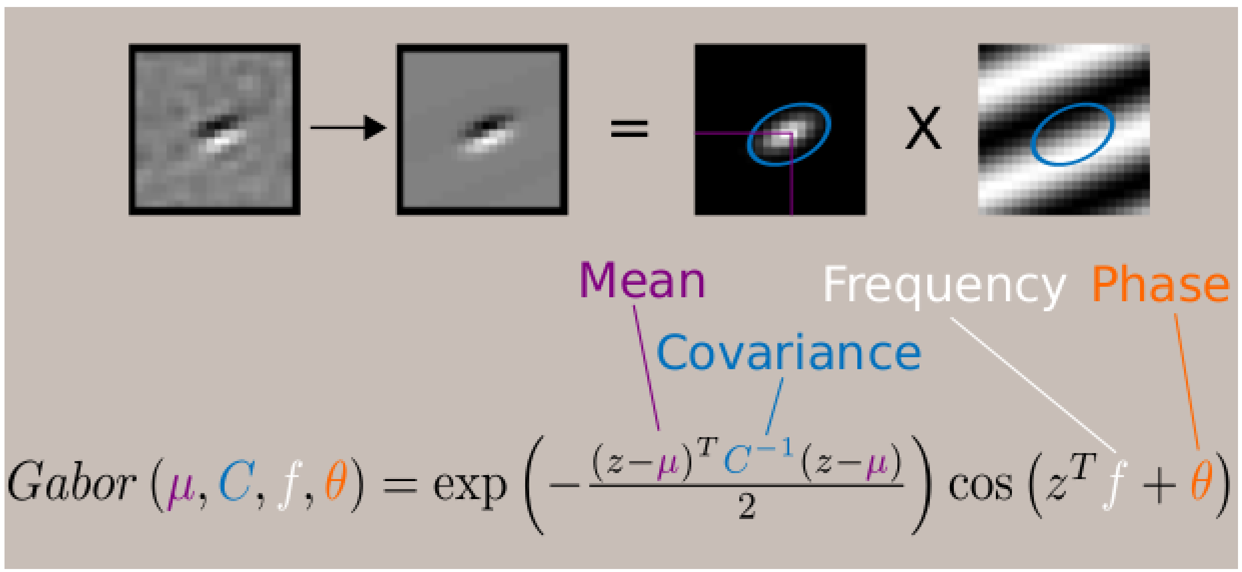
$$I \sim N(\Phi a, \sigma^2) \qquad \Phi_{ML} = arg\,max_{\Phi} P(I|a)$$

Sparse coding





V1 simple cells as Gabor filters



$$Gabor\left(\mu, C, f, \theta\right) = \exp\left(-\frac{(z-\mu)^T C^{-1} (z-\mu)}{2}\right) \cos\left(z^T f + \theta\right)$$

Implications for technology

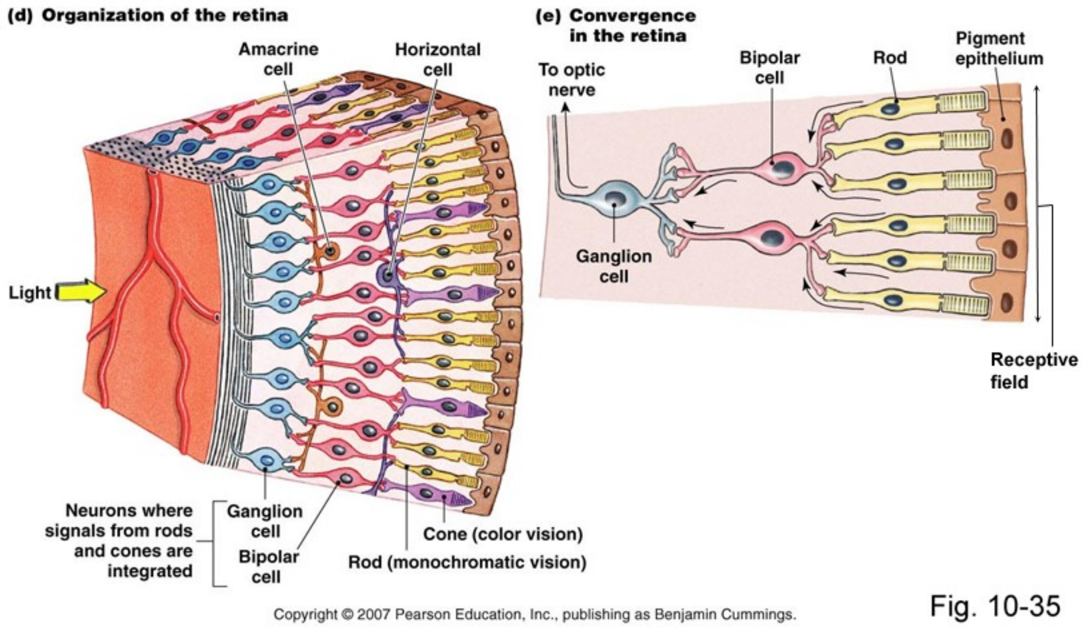
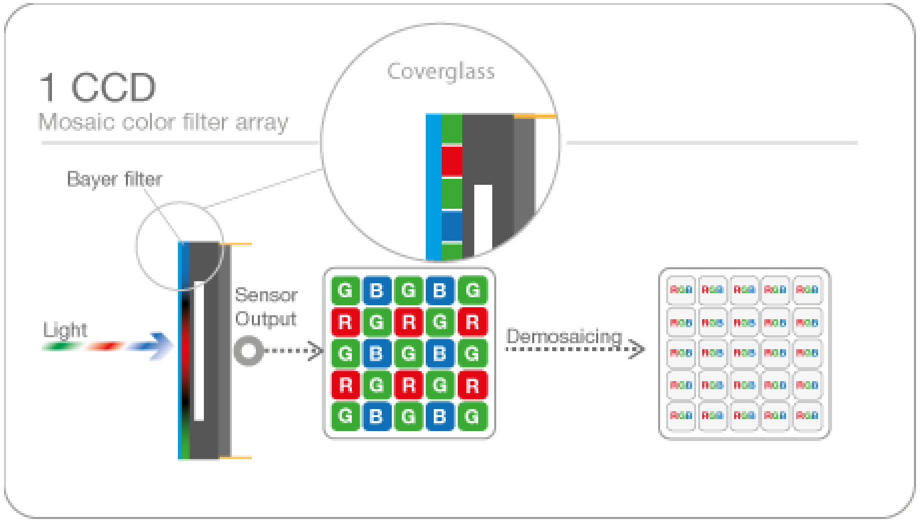


Fig. 10-35



https://upload.wikimedia.org/wikipedia/commons/thumb/b/b5/Neuron.svg/1280px-Neuron.svg.png

https://info.adimec.com/hs-fs/hub/20421/file-39331860-png/images/1-ccd-and-bayer-color.png?t=1515527438180

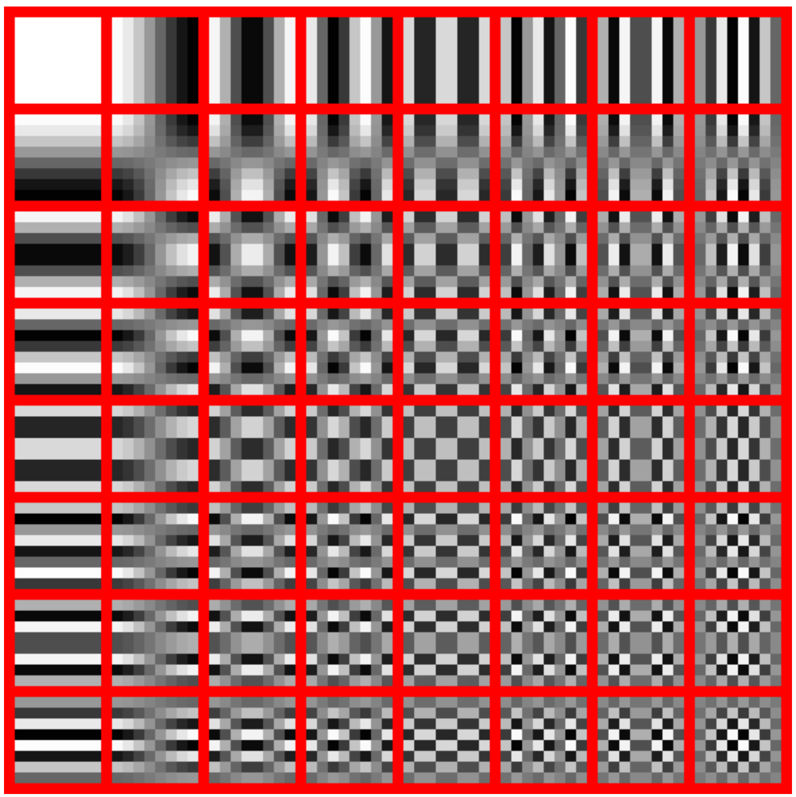
Implications for technology

- **Image compression**

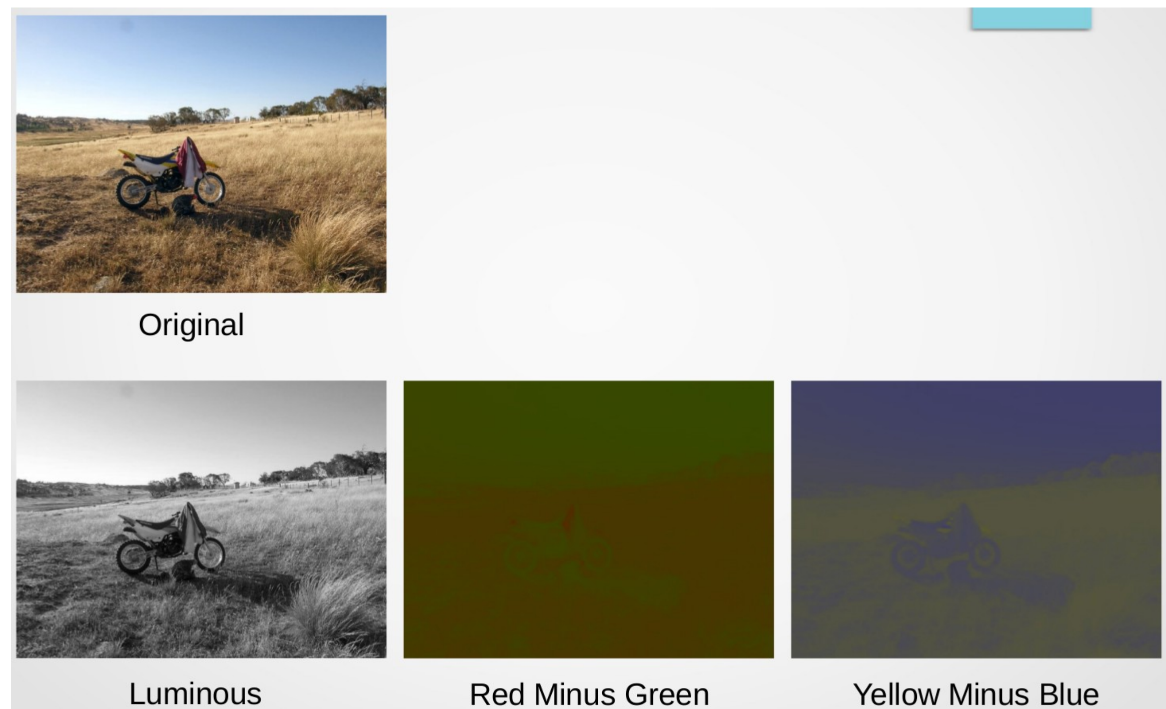
Based on an illustration from
Brian Wandell's book
Foundations of Vision

- **JPEG**
  - Discrete Cosine Transform

Implications for technology



https://upload.wikimedia.org/wikipedia/commons/2/23/Dctjpeg.png



Implications for technology

- Brain Machine Interfaces (BMIs)
  https://youtu.be/YJMckMlaPrY?t=159

Takeaways

- Abstracting neural computation for understanding the brain is promising
- Need to understand models to address potential problems
- If we never understand the brain in its entirety, the understanding we gather will still have valuable lessons

PIs in the field

- **Bruno Olshausen**
  - Redwood Center for Theoretical Neuroscience (UC Berkeley)
- **Eero Simoncelli**
  - Center for Neural Science (New York University)
- **Zhaoping Li**
  - Gatsby Computational Neuroscience Unit (University College London)
- **Jonathan Pillow**
  - Princeton Neuroscience Institute (Princeton University)
- **Christopher Rozell**
  - Georgia Institute of Technology
- **Odelia Schwartz**
  - University of Miami