Image and Video Understanding

2VO  710.095  WS
Christoph Feichtenhofer, Axel Pinz

1KU  710.096  WS
Christoph Feichtenhofer, Axel Pinz
Vision and Brain (1) [J.V. Stone]

Figure 4.8
Combining the outputs from LGN cells with center-surround receptive fields to create a simple cell receptive field that responds best to a luminance step edge at one (horizontal) preferred orientation. The sign (+,-) indicates whether light has an excitatory or inhibitory effect on a cell within each region of its receptive field.
Why is it coded the way it is?

• Saliency $\rightarrow$ contrast $\rightarrow$ difference (between neighbouring features)

“The retina is a vast bank of spatial band-pass filters, which perform Fourier analysis on retinal images” [Stone]

• Change in contrast $\rightarrow$ difference of differences $\rightarrow$ zero crossing
• Opponency $\rightarrow$ linearity

Opponent pairs of cells also for color, motion, …
Vision and Brain (2) [J.V. Stone]

Figure 4.1 (plate 5)
Major divisions of the human brain and visual cortex. Reproduced with permission, copyright © 1999 Terese Winslow.
Vision and Brain (3) [J.V. Stone]

Figure 4.3
Lateral geniculate nuclei (LGNs). (a) Cross section of the lateral geniculate nucleus, with the magnocellular layers 1 and 2 at the base and the parvocellular layers 3–6 at the top. Reproduced with permission. (b) Schematic diagram showing the anatomical origins of each layer in the LGNs. From Frisby & Stone 2010.
Vision and Brain (4)  
[J.V. Stone]

Magnocellular layers → Motion-related

Parvocellular layers → Static form

Konio layers → Color

Figure 4.7
Unpacking the striate cortex. The outputs of different ganglion cells in each eye project to specific layers of the LGNs (C = contralateral, I = ipsilateral). These then project to striate cortex (V1), which project to discrete stripes of V2. These project to areas V3–V5, which are specialized for processing different types of information.
From Retina $\rightarrow$ Visual Cortex

photoreceptors $\rightarrow$ optic nerve $\rightarrow$ implicit $\rightarrow$ explicit

126 million $\rightarrow$ 1 million fibres $\rightarrow$ color, depth, motion, …

retina $\leftrightarrow$ visual pathway $\leftrightarrow$ visual cortex

“What the eye does not tell the brain, the brain doesn’t need to know” [Stone]
Vision and Brain [5] [J.V. Stone]

“packing problem”

“arises from the manifold richness of the representations constructed by the human brain” [Stone]

→ coding, pooling

Figure 4.17
The need to make implicit features explicit in every point in the retinal image may explain the massive increase in neurons in primary visual cortex compared to the number of neurons sending information from the retina.
The **only way** to model it ?! [Stone]

- Evolutonal evidence
- Very *different embodiments* along very different evolutonal paths (e.g. horseshoe crab, octopus, human)
- *Same processing* in terms of feature extraction encoding, pooling concepts!
- Contrast, opponency, linearity, layers, columns, …
Requirements (what you should already know) - 1

- The digital raster image and its main properties
  - Spatial resolution, spatial frequency (relationship image ↔ scene)
  - Radiometric resolution (bit/pixel)
  - Color (RGB additive and subtractive, IHS)
  - Neighborhood (4-, 8-, 6-, connectedness of contours and regions)

- Histogram
  - Modes (Normal distribution, bimodal, threshold selection)
Requirements (what you should already know) - 2

• Operations on dig. Rls: E(x,y) → A(x,y)
  – Point-op.
  – Local op. (convolution, rank-order (median, morphology))
  – Global op. (2-dim discrete Fourier-Transformation)

• Segmentation
  – From pixels → image objects
  – Formal definition of segmentation
  – Basic concepts
    • Region-based
    • Contour-based