A gap in the theory?

from Hubel, 1995
Eye anatomy

- Photoreceptors: rods (night vision) and cones (day vision)
- Other layers: processing to enhance SNR and maximize information transmission
- All visual information conveyed to brain through ~1 million optic nerve fibers

(from Hubel, 1995)
Retinal cross section: the real thing

from Wandell, 1995

Ganglion cell layer
Inner plexiform layer
Inner nuclear layer
Outer plexiform layer
Outer nuclear layer
Photoreceptors
Distribution of rods and cones

- Maximum accuracy is in the fovea:
  - about 50,000 cones, each about 30 seconds of visual angle
  - no sharp border, but about 250 cones across
  - about 2° of visual angle
Photoreceptor mosaic in macaque monkey


from Wandell, 1995
Light intensity spans an enormous range

<table>
<thead>
<tr>
<th>Luminance (cd/m²)</th>
<th>Catch Rate (photons/s)</th>
<th>Environment</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>$10^{-6}$</td>
<td>1/5100 s</td>
<td>overcast night sky</td>
<td>rod threshold</td>
</tr>
<tr>
<td>$10^{-3}$</td>
<td>1/5 s</td>
<td>starlight</td>
<td>cone threshold</td>
</tr>
<tr>
<td></td>
<td>3/s</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$10^{-2}$</td>
<td>10/s</td>
<td>full moon</td>
<td></td>
</tr>
<tr>
<td>$10^{-1}$</td>
<td>100/s</td>
<td>sunrise</td>
<td>colors visible</td>
</tr>
<tr>
<td>$10^2$</td>
<td></td>
<td>indoor lighting</td>
<td></td>
</tr>
<tr>
<td>$10^3$</td>
<td></td>
<td></td>
<td>best accuracy</td>
</tr>
<tr>
<td>$10^4$</td>
<td></td>
<td>clear sky at noon</td>
<td></td>
</tr>
<tr>
<td>$10^5$</td>
<td>300,000/s</td>
<td>patch of snow at noon</td>
<td>maximal cone rate</td>
</tr>
<tr>
<td></td>
<td>1,000,000/s</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Scotopic Vision: rods and night vision

from Rodeick, 1998

Photons on retina: sky-tree border in overcast night sky
Rods sensitivity

- Rods respond reliably to single photons.
  - Do we really need this degree of sensitivity?

- Suppose instead of 1 photon, a rod needed 2 photons (within 0.1 sec).
  - How much more light would we need to get same output?

- Relevant facts:
  - photon catch rate is \( \sim 1/5100 \) secs or once every 85 hours
  - but, there are \( \sim 10^8 \) rods
  - photons arrive at random times ⇒ Poisson distribution

- What is the rate of seeing two photon within 0.1 sec?
  - Once every 16 years.
  - Need 100,000 times intensity to achieve same discriminability.
  - In short: two photon sensitivity ⇒ no night vision.
Dynamic range of rods

- Retinal circuitry enhances sensitivity by summing the output across many rods
- Rods are also insensitive to the direction of light
- What about cones?
Dynamic range of rods and cones

Cones do not saturate

from Rodeick, 1998
Cones are adapted for daylight vision

Cones are sensitive to the direction of photon arrival

 photons arrive at photoreceptors from different directions from Rodeick, 1998

Cones are aligned to optimize focus

alignment of cones toward nodal point (cones greatly enlarged)
Color vision: Cone spectral sensitivity curves

- Three types of spectral sensitivity curves: long, medium, and short wavelength.
- Response of an individual cone is the integral of the product of the curve and the spectrum at that point.

(from Wandell, 1995)
Sampling of color space

- These two spectra are indistinguishable.
- The perception of color starts from here, but gets far more complicated.

from Wandell, 1995
There are far more M&L cones than S cones: Why?


from Wandell, 1995
Chromatic aberration

from Wandell, 1995
Focus is matched to M & L cones

from Rodeick, 1998
• 3 mm pupil diameter. Eye in focus at 580 nm and diffraction limited.
• No significant contrast beyond 4 cpd at short wavelengths.
Cone spatial mosaic measured in living human retina

- Roorda and Williams (1999). False color: blue, green, red represent S, M, and L cones. Subject JW temporal (a) and nasal (b) retina, Subject AN nasal retina (c)

<table>
<thead>
<tr>
<th>Subject</th>
<th>% L cones</th>
<th>% M cones</th>
<th>% S cones</th>
<th>L:M ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>JW</td>
<td>75.8</td>
<td>20.0</td>
<td>4.2</td>
<td>3.8 : 1</td>
</tr>
<tr>
<td>AN</td>
<td>50.6</td>
<td>44.2</td>
<td>5.2</td>
<td>1.2 : 1</td>
</tr>
</tbody>
</table>
Back to computation
Redundancy reduction for noisy channels (Atick, 1992)

\[ y = x + \nu \]

\[ y = Ax + \delta \]

Mutual information

\[ I(x, s) = \sum_{s,x} P(x, s) \log_2 \left[ \frac{P(x, s)}{P(s)P(x)} \right] \]

\[ I(x, s) = 0 \text{ iff } P(x, s) = P(x)P(s), \text{ i.e. } x \text{ and } s \text{ are independent.} \]
Profiles of optimal filters

- high SNR
  - reduce redundancy
  - center-surround

- low SNR
  - average
  - low-pass filter

- matches behavior of retinal ganglion cells
An observation: Contrast sensitivity of ganglion cells

Luminance level decreases one log unit each time we go to lower curve.
Natural images have a $1/f$ amplitude spectrum

Field, 1987

\[ \log_{10} \text{amplitude} \]

\[ \log_{10} \text{spatial frequency (cycles/picture)} \]
Components of predicted filters

from Atick, 1992
Predicted contrast sensitivity functions match neural data

from Atick, 1992
Robust coding of natural images

- Theory refined:
  - image is noisy and blurred
  - neural population size changes
  - neurons are noisy
## Problem 1: Real neurons are “noisy”

<table>
<thead>
<tr>
<th>system (area)</th>
<th>stimulus</th>
<th>bits / sec</th>
<th>bits / spike</th>
</tr>
</thead>
<tbody>
<tr>
<td>fly visual (H1)</td>
<td>motion</td>
<td>64</td>
<td>~1</td>
</tr>
<tr>
<td>monkey visual (MT)</td>
<td>motion</td>
<td>5.5 - 12</td>
<td>0.6 - 1.5</td>
</tr>
<tr>
<td>frog auditory (auditory nerve)</td>
<td>noise &amp; call</td>
<td>46 &amp; 133</td>
<td>1.4 &amp; 7.8</td>
</tr>
<tr>
<td>Salamander visual (ganglion cells)</td>
<td>rand. spots</td>
<td>3.2</td>
<td>1.6</td>
</tr>
<tr>
<td>cricket cercal (sensory afferent)</td>
<td>mech. motion</td>
<td>294</td>
<td>3.2</td>
</tr>
<tr>
<td>cricket cercal (sensory afferent)</td>
<td>wind noise</td>
<td>75 - 220</td>
<td>0.6 - 3.1</td>
</tr>
<tr>
<td>cricket cercal (10-2 and 10-3)</td>
<td>wind noise</td>
<td>8 - 80</td>
<td>avg. = 1</td>
</tr>
<tr>
<td>Electric fish (P-afferent)</td>
<td>amp. modulation</td>
<td>0 - 200</td>
<td>0 - 1.2</td>
</tr>
</tbody>
</table>

Limited capacity $\Rightarrow$ neural codes need to be **robust**.
Traditional codes are *not* robust

encoding neurons

sensory input

Original
Traditional codes are not robust

- Encoding neurons
- Sensory input

Add noise equivalent to 1 bit precision

**1x** efficient coding

Original

1 bit precision

Reconstruction (34% error)
Redundancy plays an important role in neural coding

Response of salamander retinal ganglion cells to natural movies

From Puchalla et al, 2005
Robust coding (Doi and Lewicki, 2005, 2006)

Model limited precision using additive channel noise:

Encoder

\[ \mathbf{x} \rightarrow \mathbf{W} \rightarrow \mathbf{u} \rightarrow \mathbf{A} \rightarrow \hat{\mathbf{x}} \]

Data

Noiseless Representation

Reconstruction

Channel Noise

\[ \mathbf{n} \]

Encoder

\[ \mathbf{x} \rightarrow \mathbf{W} \rightarrow \mathbf{u} \rightarrow \mathbf{r} \rightarrow \mathbf{A} \rightarrow \hat{\mathbf{x}} \]

Data

Noiseless Representation

Noisy Representation

Reconstruction

\[ \mathbf{r} = \mathbf{Wx} + \mathbf{n} = \mathbf{u} + \mathbf{n}. \]

\[ \hat{\mathbf{x}} = \mathbf{Ar} = \mathbf{AWx} + \mathbf{An}. \]

Caveat: linear, but can evaluate for different noise levels.
Generalizing the model: sensory noise and optical blur

(a) Undistorted image  (b) Fovea  retinal image  (c) 40 degrees eccentricity

Can also add sparseness and resource constraints
How do we learn robust codes?

**Objective:**

Find $W$ and $A$ that minimize reconstruction error.

- Channel capacity of the $i^{th}$ neuron:
  \[
  C_i = \frac{1}{2} \ln(\text{SNR}_i + 1)
  \]

- To limit capacity, fix the coefficient signal to noise ratio:
  \[
  \text{SNR}_i = \frac{\langle u_i^2 \rangle}{\sigma_n^2}
  \]

Now robust coding is formulated as a *constrained optimization problem.*
Robust coding of natural images

encoding neurons

sensory input

Add noise equivalent to 1 bit precision

Original

1x efficient coding

1 bit precision

reconstruction (34% error)
Robust coding of natural images

encoding neurons

sensory input

Weights adapted for optimal robustness

Original

1x robust coding

1 bit precision

reconstruction (3.8% error)
Reconstruction improves by adding neurons

Weights adapted for optimal robustness

encoding neurons

sensory input

Original

8x robust coding

1 bit precision

reconstruction error: 0.6%
Adding precision to 1x robust coding

original images  \hspace{1cm} 1 bit: 12.5% error  \hspace{1cm} 2 bit: 3.1% error
Can derive minimum theoretical average error bound

\[ \mathcal{E} = \frac{1}{M} \cdot \text{SNR} + 1 \cdot \frac{1}{N} \left[ \sum_{i=1}^{N} \sqrt{\lambda_i} \right]^2 \quad \text{if } \text{SNR} \geq \text{SNR}_c \]

\( \lambda_i \) - ith eigenvalue of the data covariance

\( N \) - input dimensionality

\( M \) - # of coding units (neurons)

Algorithm achieves theoretical lower bound

<table>
<thead>
<tr>
<th>Results</th>
<th>Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5x</td>
<td>19.9%</td>
</tr>
<tr>
<td>1x</td>
<td>12.4%</td>
</tr>
<tr>
<td>8x</td>
<td>2.0%</td>
</tr>
</tbody>
</table>

Balcan, Doi, and Lewicki, 2007; Balcan and Lewicki, 2007
Sparseness localizes the vectors and increases coding efficiency

Encoding vectors ($W$) vs. decoding vectors ($A$)

Robust coding:
- Average error is unchanged: $12.5\% \rightarrow 12.5\%$

Robust sparse coding:

Normalized histogram of coeff. kurtosis

$\text{avg.} = 4.9$

$\text{avg.} = 7.3$

Normalized histogram of coeff. kurtosis

$\text{avg.} = 4.9$

$\text{avg.} = 7.3$
Non-zero resource constraints localize weights

average error is also unchanged

from Doi and Lewicki, 2007