Sensory Encoding

Eero Simoncelli
Stimulus

Neural response

“Encoding”
Transform and represent sensory information
Stimulus

Neural response

Behavior

“Encoding”
Transform and represent sensory information

“Decoding”
Extract encoded information

psychometric function
- simple to measure
- very incomplete:
  - limited stimuli
  - limited response model (no noise, spikes, dynamics)
a black box
Some big numbers:
Some big numbers:

seconds since big bang: $10^{17}$
Some big numbers:

seconds since big bang: \(10^{17}\)

atoms in the visible universe: \(10^{80}\)
Some big numbers:

seconds since big bang: $10^{17}$

atoms in the visible universe: $10^{80}$

10-bit images, 1000x1000: $10^{3,000,000}$
Some big numbers:

seconds since big bang:  $10^{17}$

atoms in the visible universe:  $10^{80}$

10-bit images, 1000x1000:  $10^{3,000,000}$

we can’t enumerate them => we need models
Problems with lookup tables...

- Curse of dimensionality
- Mimics, but doesn’t explain how or why
- Doesn’t make predictions

=> Need a model
Model constraints

- Implementation (anatomy, biophysics, etc)
- Functional physiological properties
- Fundamental computational principles
An algorithm is likely to be understood more readily by understanding the nature of the problem being solved than by examining the mechanism in which it is embodied.

- Marr, 1982
Model constraints

- Implementation (anatomy, biophysics, etc)
- Functional physiological properties
- Fundamental computational principles
Model constraints

- Implementation (anatomy, biophysics, etc)
- Functional physiological properties
- Fundamental computational principles
  - Environmental
Model constraints

- Implementation (anatomy, biophysics, etc)
- Functional physiological properties
- Fundamental computational principles
  - Environmental
  - Task/goal
Model constraints

- Implementation (anatomy, biophysics, etc)
- Functional physiological properties
- Fundamental computational principles
  - Environmental
  - Task/goal
- Occam’s Razor
[Maunsell, after Lennie, after Felleman and Van Essen]
[Adelson & Bergen, 1990]
Stimulus  

Neural response  

Behavior  

“Encoding”  

“Decoding”  

psychometric function
Stimulus

Neural response

Behavior

“Encoding”

“Decoding”

Bayesian estimation/decision

psychometric function
Stimulus Neural response Behavior

“Encoding”
Efficient coding hypothesis [Barlow ’61]

“Decoding”
Bayesian estimation/decision

psychometric function

p(v |v )

P(v 1 > v 2 )
“Encoding”
Efficient coding hypothesis [Barlow ’61]

“Decoding”
Bayesian estimation/decision

Both theories rely on statistical models of environment

Wednesday, 3 February 2010
Common frameworks

• Science: optimality principles for neurobiology (evolution, development, learning, adaptation)

• Engineering: compression, denoising, restoration, enhancement/modification, synthesis, manipulation
Efficient Coding

[Attneave ’54; Barlow ’61; Laughlin ’81; Atick ’90; Bialek etal ‘91]
Efficient Coding

Maximize information about stimulus in response, subject to constraints (e.g. metabolic)
Efficient Coding

[Attneave ’54; Barlow ’61; Laughlin ’81; Atick ’90; Bialek et al ‘91]

Maximize information about stimulus in response, subject to constraints (e.g. metabolic)

\[ I(r, s) = H(r) - H(r|s) \]
Efficiency in single neurons

• Utilize full response range
• Subject to constraints
• Subject to noise
Efficiency in single neurons

- Input noise: $p(n|s)$
- Response cost: $C(r)$
- Output noise: $p(nr)$
- Prior: $p(s)$
- Input noise: $p(nls)$
Efficiency in single neurons

\[ \text{response cost } C(r) \hspace{2cm} \text{output noise } p(n|r) \]

\[ \text{input noise } p(nls) \hspace{2cm} \text{prior } p(s) \]
Efficiency in single neurons

input noise $p(n|s)$

response cost $C(r)$

output noise $p(nr)$
Efficiency in single neurons

Response cost $C(r)$

Output noise $p(nr)$

Input noise $p(nls)$

Prior $p(s)$

Wednesday, 3 February 2010
Efficiency in single neurons

- Prior: \( p(s) \)
- Input: \( p(n|s) \)
- Response: \( C(r) \)
- Output noise: \( p(nr) \)
- Input noise: \( p(nls) \)
- Prior: \( p(s) \)
Efficiency in single neurons

Prior \( p(s) \)

Input noise \( p(n|s) \)

Response cost \( C(r) \)

Output noise \( p(n|r) \)

Prior \( p(s) \)
Efficiency in single neurons

- Response cost $C(r)$
- Output noise $p(nr)$
- Prior $p(s)$
- Input noise $p(nls)$
Efficiency in single neurons

response cost $C(r)$
output noise $p(nr)$

input noise $p(nls)$
prior $p(s)$
- Measure contrast in natural scenes
- Construct cumulative PDF
- Compare to neural response (fly, large monopolar neuron)

[Laughlin, 1981]
The Cost of Cortical Computation

Peter Lennie*
Center for Neural Science
New York University
4 Washington Place
New York, New York 10003

Summary

Electrophysiological recordings show that individual neurons in cortex are strongly activated when engaged in appropriate tasks, but they tell us little about how many neurons might be engaged by a task, which is important to know if we are to understand how cortex encodes information. For human cortex, I estimate the cost of individual spikes, then, from the known energy consumption of cortex, I establish how many neurons can be active concurrently. The cost of a single spike is high, and this severely limits, possibly to fewer than 1%, the number of neurons that can be substantially active concurrently. The high cost of spikes requires the brain not only to use representational codes that rely on very few active neurons, but also to allocate its energy resources flexibly among cortical regions according to task demand. The latter constraint explains the investment in local control of hemodynamics, exploited by functional magnetic resonance imaging, and the need for mechanisms of selective attention.

Introduction

rat neocortex. Neurons in human neocortex are larger than those in rat and receive and make more synapses, but they are not otherwise known to differ in their basic structure or organization [5]. Thus, with appropriate scaling of parameters for the larger neurons, Attwell and Laughlin’s analysis can be used to estimate the energy consumed by a pyramidal neuron in human neocortex.

In different mammals, the number of neurons under a unit area of cortical surface is relatively constant (~100,000/mm²), except in primate striate cortex, where it may be twice as high [6]. Increasing brain size brings an increase in cortical thickness and a proportionately lower density of neurons [5, 6] without an increase in cell body size, which remains approximately constant at 15 μm diameter [7]. The volume of axons and dendrites increases with cortical thickness. This reflects an increase in the lengths of dendrites and axons without an increase in diameter [5]. Table 1 summarizes relevant statistics for human cortex.

Postsynaptic Potentials

Individual synapses are assumed to be the same in rat and human neurons, so the energy costs associated with transmitter uptake and release will be the same, as will the current flow through receptor channels. Given (from Table 1) $7 \times 10^8$ synapses per mm³ of cortex, and 40,000 neurons/mm³, the average neuron will make 17,500 synaptic contacts. If we use this number, and assume a 50% failure rate [8, 9], the cost of EPSPs arising from a single spike will be $1.2 \times 10^9$ ATP molecules [4].
Distributions appear exponential, optimally efficient under a constraint on the mean firing rate.

Caveats/concerns:

- Firing rate constraint not always the right one. [Laughlin ’81]
- Why are responses to white noise also somewhat exponential?
- Noise probably matters here
Efficiency in multiple neurons

• Utilize joint response range
• Subject to constraints
• Subject to noise
Efficiency in multiple neurons

- Utilize joint response range $\Rightarrow$ independent
- Subject to constraints
- Subject to noise
Efficiency in multiple neurons

- Utilize joint response range $\Rightarrow$ independent
- Subject to constraints
- Subject to noise

Simple approximation (low noise): responses should be statistically independent
Pixel correlation
Pixel correlation

I(x,y)

Spatial separation (pixels)

Correlation

l(x+1,y)
l(x+2,y)
l(x+4,y)

l(x,y)

l(x,y)

l(x,y)

l(x,y)

0
1

0 10 20 30 40

Wednesday, 3 February 2010
Principal Components Analysis (PCA)
Principal Components Analysis (PCA)

Rotate
(orthogonal matrix)
Principal Components Analysis (PCA)

- Rotate (orthogonal matrix)
- Squeeze/stretch (diagonal matrix)
Principal Components Analysis (PCA)

For Gaussian sources: guaranteed independence
observations

• spectral power

Field, 1994

figure from Simoncelli 05

Spectral power

Structural:

\[ F(s\omega) = s^p F(\omega) \]

\[ F(\omega) \propto 1/\omega^p \]

[Ritterman 52; DeRiugin 56; Field 87; Tolhurst 92; Ruderman/Bialek 94; ...]

Assume scale-invariance: then:

Empirical:

Natural image amplitude spectra
Efficient coding in the retina?

[Srinivasan et. al. 82; Atick & Redlich 90; van Hateren 92]
Retinal ganglion cell receptive fields

“ON”

“OFF”

[Kuffler, 1953]
Efficient coding in the retina?

[Srinivasan et. al. 82; Atick & Redlich 90; van Hateren 92]

\[ S \xrightarrow{+} \text{filter} \xrightarrow{+} r \]
Components of predicted filters

- Low-pass filter
- Whitening filter
- Optimal filter

[Atick & Redlich '92]
compared with contrast sensitivity functions (human)
Optimal filters depend on SNR

(a) \( S/N=10 \)

(b) \( S/N=2 \)

(c) \( S/N=0.1 \)
Limitations

• Signal model: Gaussian, 1/f
• Noise model(s): Gaussian, uncorrelated
• Photoreceptors assumed regularly spaced
• Transformation assumed linear
• Transformation assumed convolutional
• Cone:RGC ratio is not 1:1 outside fovea
• Cone:RGC ratio is not 1:1 outside fovea
• Cone lattice is not regular
• Cone:RGC ratio is not 1:1 outside fovea
• Cone lattice is not regular

Curcio et al 91
• Cone:RGC ratio is not 1:1 outside fovea
• Cone lattice is not regular
• RGC receptive fields are irregular

Curcio et al 91
• Cone:RGC ratio is not 1:1 outside fovea
• Cone lattice is not regular
• RGC receptive fields are irregular

Curcio et al 91

[Gauthier et al, 09]
Efficient Coding in Retina

\[ \text{image} \quad \text{blur} \quad \text{cone} \quad \text{retinal} \quad \text{response} \quad \text{noise} \]

\[ \text{noise} \]

[Doi et. al., SfN 2008]
Information Maximization

\[ I(s; r) = \frac{1}{2} \log \frac{|WH\Sigma_s H^T W + \sigma^2_\nu WW^T + \sigma^2_\delta I_M|}{|\sigma^2_\nu WW^T + \sigma^2_\delta I_M|} \]

(total power of responses is constrained)

Solution: \( W = P\Omega Q^T \)

[Doi et al, SfN 08]
Information Maximization

\[ I(s; r) = \frac{1}{2} \log \frac{|WHP\Sigma_s H'^T W^T + \sigma^2 \nu W W'^T + \sigma^2 I_M|}{|\sigma^2 \nu W W^T + \sigma^2 I_M|} \]

(total power of responses is constrained)

Solution: \( W = P \Omega Q^T \)

\( P \) orthonormal, unconstrained

[Doi et al, SfN 08]
Information Maximization

\[ I(s; r) = \frac{1}{2} \log \frac{|WH\Sigma_sH^T W^T + \sigma^2 \nu W W^T + \sigma^2_\delta I_M|}{|\sigma^2_\nu W W^T + \sigma^2_\delta I_M|} \]

(total power of responses is constrained)

Solution: \( W = P\Omega Q^T \)

\( Q \) orthonormal, evecs of blurred signal

[Doi et al, SfN 08]
Information Maximization

\[ I(s; r) = \frac{1}{2} \log \frac{|WH\Sigma_sH^TW^T + \sigma^2_\nu WW^T + \sigma^2_\delta I_M|}{|\sigma^2_\nu WW^T + \sigma^2_\delta I_M|} \]

(total power of responses is constrained)

Solution: \[ W = P\Omega Q^T \]

\[ \Omega_{\text{eff}} \text{ with different photoreceptor SNR} \]

\[ \Omega_{\text{eff}} \text{ with different neural SNR} \]

\[ \Omega_{\text{eff}} \text{ with different population size} \]
• Macaque retina, 27 deg eccentricity
• full mosaics: ON/OFF Midget/Parasol
• 145 RGCs, 665 cones [ratio = 4.6]
• receptive fields, as weights on cones, determined by STA
• W contains receptive fields

[Gauthier et. al., 09]
Theoretical W

Optimal $Q, \Omega$ based on:

- 62 natural images [Doi et al. 2003]
- Human optical blur at 30 deg ecc [Navarro et al. 93]
- 10 dB photoreceptor SNR
- 10 dB ganglion cell SNR [Borst & Theunissen 99]

[Doi et. al., SfN 2008]
Comparison to data

Measured W

Efficient manifold

[Doi et. al., SfN 2008]
Comparison to data

65.7% variance explained (34.3% error)

Efficient manifold

Measured $W$

[Doi et. al., SfN 2008]
Comparison to data

65.7% variance explained (34.3% error)

[Doi et. al., SfN 2008]
Comparison to data

65.7% variance explained (34.3% error)

16.6% variance (+/- 0.1%) explained (83.4% error)

[Doi et. al., SfN 2008]
the diagonal matrix whose entries can be found using the KKT condition

Parasol

ON

OFF

Midget

ON

OFF

data

N=5

N=6

N=26

N=32

theory

[Doi et. al., SfN 2008]
Testing the Independence Hypothesis

- Experimental: Examine dependency between neuronal responses under natural stimulation conditions [eg: Vinje/Gallant ‘00; Nirenberg/Latham ’01]
Testing the Independence Hypothesis

- Experimental: Examine dependency between neuronal responses under natural stimulation conditions [eg: Vinje/Gallant ‘00; Nirenberg/Latham ’01]
Testing the Independence Hypothesis

- Experimental: Examine dependency between neuronal responses under natural stimulation conditions [eg: Vinje/Gallant ’00; Nirenberg/Latham ’01]
Testing the Independence Hypothesis

• Experimental: Examine dependency between neuronal responses under natural stimulation conditions [eg: Vinje/Gallant ‘00; Nirenberg/Latham ‘01]

• Theoretical: Derive a model that can produce independent responses; Compare with physiology
Toward Independent Image Components...

A methodology for constructing image models.

Extensive history in engineering and theoretical neuroscience.
Independent Component Analysis (ICA)

Seek linear transform that maximizes statistical independence of transform coefficients.
A variety of algorithms exist.
ICA on Images

Linear operators with maximally independent (or maximally non-Gaussian) responses are oriented bandpass filters

[Bell/Sejnowski '97; Olshausen/Field '96]
Algorithm selects best of many possible sensor codes.

Lewicki & Olshausen, 1999
Learned kernels share features of auditory nerve filters.

Optimized kernels

Scale bar = 1 msec

Auditory nerve filters
from Carney, McDuffy, and Shekhter, 1999

- Smith & Lewicki, 2006
For most filters, there’s a matching auditory nerve fiber!
Trouble in paradise
Trouble in paradise

- Biology: Visual system uses a nonlinear cascade
  - Where’s the retina? The LGN?
  - What about V1 nonlinearities?
  - What happens after V1?
Trouble in paradise

• Biology: Visual system uses a nonlinear cascade
  - Where’s the retina? The LGN?
  - What about V1 nonlinearities?
  - What happens after V1?

• Statistics: Images don’t obey ICA source model
  - The responses of ICA filters are highly dependent
    [Wegmann & Zetzsche ‘90, Simoncelli ‘97]
  - Sparsity seems unlikely to be a goal for the brain
Linear responses are not independent, even for optimized filters!
Different Filter Pairs:

Different Images:

Baboon

Flowers

White noise
Modeling the Dependency

One filter:
\[ \text{var}(L_1|L_2) = w L_2^2 + \sigma^2 \]

Generalized neighborhood:
\[ \text{var}(L_1|\{L_n\}) = \sum_n w_n L_n^2 + \sigma^2 \]
Local Linear Variance Estimator

- State-of-the-art image compression [Buccigrossi&Simoncelli ’99]
- State-of-the-art image denoising [Simoncelli ’99, Portilla etal ’03]
- State-of-the-art parametric texture modeling/synthesis [Portilla&Simoncelli ’00]
- Related models: turbulent fluid flow, financial time series (ARCH), speech (SIRP).
Reducing Dependency

\[ R_n = \frac{L_n^2}{\sum_k w_{nk} L_k^2 + \sigma^2} \]
Functional Physiological Model

- Divisive suppression by *weighted* sum of neighbors increases independence
- Optimal weights determined from image statistics
Divisive Normalization: Physiological Evidence

Steady-state neural responses = linear projection, rectification, and division by the summed responses of other neurons [Heeger '92; Carandini/Heeger/Movshon '97]

Such models can account for some nonlinear striate cortical behaviors. Examples [Carandini et al. 1997]:

- Tuning curves independent of contrast
- Contrast saturation level depends on stimulus parameters
- Cross-orientation suppression
- Increasing phase lag at lower contrast
Methods

• Statistically-determined model:
  1. Linear basis: multi-scale, oriented 3rd derivative operators
  2. “Neuron”: vertical, optimal spatial frequency 0.125 cycle/pixel
  3. Neighborhood: 2 scales, 4 orientations, $3 \times 3$ array
  4. Weights: optimized (ML) for statistics of 10 images (faces, landscapes, and animals).

• Physiological simulations:
  1. Compute linear responses of full neighborhood
  2. Square
  3. Divide chosen neuron response by weighted combination of squared neighbor responses.

[Schwartz & Simoncelli ’01]
Parameter Optimization

Assume a Gaussian form for the conditional distribution:

$$
P \left( L_n \mid \{ L_k \} \right) \sim \mathcal{N} \left( 0; \sum_k w_{nk} \left| L_k \right|^2 + \sigma^2 \right)$$

Maximize the likelihood over the image data:

$$\hat{w}_{nk}, \hat{\sigma} = \arg \max_{w_{nk}, \sigma} \prod_i \frac{1}{\sqrt{2\pi \Sigma_k}} \frac{1}{w_{nk}} \left| L_k \right|^2 + \sigma^2 \exp \left[ \frac{-L^2_n}{2 \sum_k w_{nk} \left| L_k \right|^2 + \sigma^2} \right]$$
Cross-orientation Suppression

Cell
(Bonds, 1989)

Model

Mean firing rate

Signal Contrast

0.1
1

Signal Contrast

0.1
1

Mask contrast:

- 0.05
- 0.1
- 0.2
Tuning Curves Independent of Contrast

Cell
(Skottun et al., 1987)

Model

Contrast:
- 0.8
- 0.2
- 0.05
Surround Suppression

(Cavanaugh et al., 2000)

**Cell**

**Model**

**Mean firing rate**

**Signal contrast**

- 0.03
- 0.3
- 1

**Mask contrast**:

- 0.13
- 0.5
Surround Spatial Frequency

Cell

Model

[Data: Müller, Krauskopf, & Lennie.]
Stimulus Diameter

Cell
(Cavanaugh et al., 2000)

Model

Contrast:
- 0.25
- 0.06
Natural Sounds

Different Sounds:

- Cat
- Speech
- White noise

Different Filter Pairs:
Auditory Nerve Fiber (cat)

Cell
(Javel et al., 1978)

Model

a

![Graph showing mean firing rate vs. signal intensity with and without mask intensity.]  
- **Signal**
- **Mask**

Mask intensity:
- No mask
- 80 dB

b

![Graph showing mean firing rate vs. signal intensity with and without mask intensity.]  
- **Signal**
- **Mask**

Mask intensity:
- No mask
- 80 dB

(CSH-02)
Auditory Nerve Fiber (cat)

Cell
(Rose et al., 1971)

Model

<table>
<thead>
<tr>
<th>Decibels</th>
<th>90</th>
<th>40</th>
</tr>
</thead>
</table>

Mean firing rate

Rel. Freq.
Summary

Divisive gain control can get us closer to a statistically independent decomposition for images (or sounds)

• Ecological justification for physiological models
• Quantitative predictions of sensory nonlinearities:
  – suppression by non-optimal stimuli
  – intensity-independence of some tuning curves (eg orientation)
  – intensity-independence of other tuning curves (eg diameter)
• Canonical computation?
• Numerous applications in signal/image processing
General Summary

Efficiency hypothesis provides a principle linking environmental properties to response properties of sensory neurons. There are many ways to study/test this. How far into the system can we go?
Stimulus

Neural response

Behavior

“Encoding”
Efficient coding hypothesis [Barlow ’61]

“Decoding”
Bayesian estimation/decision

"v1 seen faster"
psychometric function

Wednesday, 3 February 2010
Canonical computation
Canonical computation

• Encoding: efficient use of resources to represent/transmit/store information
Canonical computation

- Encoding: efficient use of resources to represent/transmit/store information
- Decoding: Optimal estimation/decision
Canonical computation

- Encoding: efficient use of resources to represent/transmit/store information
- Decoding: Optimal estimation/decision
- Self-calibrating/learning/adapting
Canonical computation

- Encoding: efficient use of resources to represent/transmit/store information
- Decoding: Optimal estimation/decision
- Self-calibrating/learning/adapting
- Cascadable
Canonical computation

- Encoding: efficient use of resources to represent/transmit/store information
- Decoding: Optimal estimation/decision
- Self-calibrating/learning/adapting
- Cascadable
- Constraints: functional “toolbox”, noise, metabolism, wiring
Credits

• Efficient Coding: Eizaburo Doi, Deep Ganguli

• Physiology:
  Jonathon Shlens, EJ Chichilnisky (Salk Institute)