A big data approach to functional characterization of the human brain

Jack Gallant
University of California at Berkeley
CCC-NSF Meeting Dec. 2014
How complicated is the brain?

How much data will we need to model it?

How can we model the system?

How do we deal with non-stationarity?

What about individual differences?

Can we do brain decoding?

Can we get better data?
Macaque vision is a typical brain subsystem

- Dozens of distinct anatomical and functional areas arranged in a hierarchical, parallel network with substantial recurrent feedback.
- Transformations between areas are nonlinear.
- Each area represents a different high-dimensional feature space projected onto the cortical surface.
- Each area is differentially affected by bottom-up and top-down information.

Felleman and Van Essen, *Cerebral Cortex*, 1992
The brain is a big place

**Macaque brain:**
- 2 billion cortical neurons
- 1-10 thousand connections/neuron
- 383 total areas and nuclei
- 6602 bi-directional, inter-areal connections

**Human brain:**
- 18 billion cortical neurons
- 1-10 thousand connections/neuron
- 5 million cortical columns?
- 500 areas and nuclei?
- 12000 inter-areal connections?
Information is organized at multiple scales

Retinotopic Mapping
Stimulus

2DG map on flattened Macaque cortex

Orientation
Spatial frequency

Ocular dom.

Structural and functional measurements are complimentary.
How complicated is the brain?

How much data will we need to model it?

How can we model the system?

How do we deal with non-stationarity?

What about individual differences?

Can we do brain decoding?

Can we get better data?
Predictions are limited by data size

Approximate size of MT data sets for single neurons

1.38 Hours of video

Oliver & Gallant, *in preparation*
Methods are time and/or space limited

- Blue: Less activity
- Green: Average activity
- Red: More activity

Nishimoto et al., *Current Biology*, 2011; Huth et al., *Neuron*, 2012;
Data visualization by James Gao’s pyCortex ([http://pycortex.org](http://pycortex.org))
Online visualization available at [http://gallantlab.org](http://gallantlab.org)
How complicated is the brain?

How much data will we need to model it?

How can we model the system?

How do we deal with non-stationarity?

What about individual differences?

Can we do brain decoding?

Can we get better data?
Neuroscience as system identification

The amount of data required to fit the model is proportional to the degree of nonlinearity.

State variables (e.g. Attention) multiply the dimensionality of the problem.
The explicit feature space approach

Nonlinear transformation

Linear transformation

Linear or Nonlinear transformation

Hypothesis

Linear regression

Experimental measurement
Modeling multiple, explicit feature spaces

1. Collect Data
2. Calculate Stimulus Features
   - Gabor Wavelets
   - Object Contours
   - Scene Contours
   - Semantic Categories
3. Estimate Model Weights
4. Evaluate Predictions
5. Interpret Results
   - Voxel-wise receptive field
   - Population analysis in feature space
   - Projection onto flattened cortex
Explicit modeling of MT neurons

Nishimoto & Gallant, *Neuron*, 2011
A motion-energy model for low-level vision

A stimulus-response encoding model for one voxel

Output (predicted BOLD signals)

A motion energy filter

Converting color space

Space-time Gabor filtering

Motion-energy computation

Compressive nonlinearity

Temporal downsampling

A filtered output

Time

Nishimoto, Vu, Naselaris, Benjamini, Yu and Gallant, *Current Biology*, 2011
A semantic model for high-level vision

Huth, Nishimoto, Vu & Gallant, *Neuron*, 2012
The machine learning approach

Nonlinear transformation
Learning algorithm or regression algorithm

Linear or Nonlinear transformation
Experimental measurement
A deep network for V2 neurons

\[ a(x) = \max(0, x) \]
\[ a(x) = \text{Pois}(a(x)) \]
\[ a(x) = \mathcal{N}(a(x), a(x)) \]
\[ a(x) = a(x) + \sqrt{a(x)} \mathcal{N}(0, 1) \]

\[ r(x) = \frac{1}{k} \log(1 + e^{k \cdot x}) \]

\[ LL = \sum_t \left( R_{\text{obs}}(t) \log(r(t)) - r(t) \right) \]

\[ \frac{\partial LL}{\partial r} = \sum_t \left( \frac{R_{\text{obs}}(t)}{k \log(1 + e^{k \cdot x})} - 1 \right) \left( \frac{e^{k \cdot x}}{1 + e^{k \cdot x}} \right) \]

Deep time-delay network trained with DropOut and Poisson loss
A convolutional network for fMRI data

Note: The CNN used here contained 5 convolutional layers and 2 classifier layers.

Images predicted to increase BOLD

Images predicted to decrease BOLD

Agrawal, Cheung, Stansbury, Malik & Gallant, in preparation
How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?
Attention alters tuning in single neurons

Olshausen, Anderson & Van Essen, *J. Neurosci.*, 1993

Connor, Gallant, Preddie & Van Essen, *J. Neurophys.*, 1996

David, Mazer, Hayden & Gallant, *Neuron*, 2008
Attention changes cortical semantic maps

Cukur, Nishimoto, Huth & Gallant, *Nature Neuroscience*, 2013
How complicated is the brain?

How much data will we need to model it?

How can we model the system?

How do we deal with non-stationarity?

What about individual differences?

Can we do brain decoding?

Can we get better data?
Narrative language semantic maps for 5 subjects

Huth, de Heer, Griffiths, Theunissen & Gallant, *in review*
**PrAGMATiC: Probabilistic And Generative Model of Areas Tiling Cortex**

**Arrangement Model**
- Area centroid
- Functional anchor
- Spring connector

**Emission Model**

**Voronoï tessellation of centroids**

**PrAGMATiC Parameters**
- Shared across subjects:
  - \( L \) - ideal spring lengths
  - \( K \) - spring constants
  - \( M \) - area functional means
- Unique to each subject:
  - \( H \) - exact centroid locations
  - \( V \) - functional values on cortex

**Total probability**
\[
P(V; H; L, K, M) = P(H; L, K) \cdot P(V|H; M)
\]

**Maximum Likelihood Estimation of Parameters**

1. Resample centroid locations
   
   \( H_j \sim P(H; L, K) \)

2. Compute likelihood of observed functional values
   
   \( P(V_{obs}|H_j; M) \)

3. Compute gradients for \( K, L, \) and \( M \)

4. Update parameters
   
   \[
   \begin{align*}
   K^{t+1} &= K^t - \epsilon \frac{\partial \mathcal{L}(K; V_{obs})}{\partial K} \\
   L^{t+1} &= L^t - \epsilon \frac{\partial \mathcal{L}(L; V_{obs})}{\partial L} \\
   M^{t+1} &= M^t - \epsilon \frac{\partial \mathcal{L}(M; V_{obs})}{\partial M}
   \end{align*}
   \]

**Huth, de Heer, Griffiths, Theunissen & Gallant, in review**
Likelihood tests for PrAGMATiC

Huth, de Heer, Griffiths, Theunissen & Gallant, in review
How complicated is the brain?

How much data will we need to model it?

How can we model the system?

How do we deal with non-stationarity?

What about individual differences?

Can we do brain decoding?

Can we get better data?
Using encoding models to decode

\[
P(f(S)|R) \propto P(R|f(S))P(f(S))
\]
Decoding semantics from higher visual areas

Movie

Likely Objects and Actions

Huth, Lee, Nishimoto & Gallant, *in review*
You can decode anything that is represented in measured brain activity
The current and future state of brain decoding

- Our ability to measure brain activity.
- Our ability to model measured activity.
- Computer power.
How complicated is the brain?
How much data will we need to model it?
How can we model the system?
How do we deal with non-stationarity?
What about individual differences?
Can we do brain decoding?
Can we get better data?
Improved methods for recording brain data

Optical imaging

Functional MRI

Electrocorticography
Trying to measure orientation columns

Yacoub, Harel & Ugurbil, PNAS, 2008
The brain is very, very complicated structure.
Current methods for measuring the brain have limited spatial, temporal and conditional sampling.
To make progress we need to push the boundaries of spatial, temporal and conditional sampling.
If we do this then we can use a variety of computational methods to build robust computational models that predict accurately and generalize well.
When the resultant models perform poorly (which is likely), then we can focus on improving computational theory rather than mere neuroscience data collection.
We are entering the age of “Big Data” in neuroscience, and there is a pressing need for data processing algorithms that can operate on these meso-scale data. The required algorithms fall into several domains: matrix (tensor) libraries, image processing algorithms, data mining tools, database tools and visualization tools.

The brain is a nonlinear dynamical system with feedback, but these are the least well understood sorts of systems from a theoretical point of view.

Therefore, even if we had a complete functional (or structural) map of the brain, we wouldn't know what to do with the data or how to model it!

This would be a good problem to have, and would likely drive theoretical developments in computer science and mathematics.