Understanding Visual Processing with fMRI

Eli Merriam

Computational Neuroimaging and Perception (CNaP)
Laboratory of Brian and Cognition
What fMRI has taught us about human vision
Susan M Courtney* and Leslie G Ungerleider†

The recent application of functional magnetic resonance imaging (fMRI) to visual studies has begun to elucidate how the human visual system is anatomically and functionally organized. Bottom-up hierarchical processing among visual cortical areas has been revealed in experiments that have correlated brain activations with human perceptual experience. Top-down modulation of activity within visual cortical areas has been demonstrated through studies of higher cognitive processes such as attention and memory.

Addresses
Laboratory of Brain and Cognition, Building 10, Room 4C104, 10 Center Drive, Bethesda, Maryland 20892-1366, USA
* e-mail: susan_courtney@nih.gov
† e-mail: lgu@nlm.nih.gov

Current Opinion in Neurobiology 1997, 7:554–561
http://biomednet.com/elecref/0959438800706554
© Current Biology Ltd ISSN 0950-4388

Abbreviations
fMRI functional magnetic resonance imaging
LO lateral occipital (area)
MT middle temporal (area)
PET positron emission tomography
V1 primary visual cortex

results in a localized increase in the fMRI signal in the brain [3–5]. Therefore, fMRI signal intensity is correlated with localized changes in neural activity (typically averaged over 2–6 s and over 1–27 mm3 of cortex). At least within the primary visual cortex (V1), the fMRI signal increases monotonically with stimulus contrast [6*].

Organization of visual cortical areas in monkeys
Vision is the most richly represented sensory modality in primates. Visual information is processed in over 30 functional cortical areas. In Old World monkeys [7]—our seemingly closest evolutionary ancestors, aside from apes—these cortical areas cover about one-half of the total cortex. Visual cortical areas are organized into two processing pathways, or ‘streams’, both of which originate in area V1 [8]. The ventral stream, projecting from area V1 through areas V2 and V4 to the inferior temporal cortex, processes the physical attributes of stimuli that are important for object identification, such as color, shape, and pattern. The dorsal stream, projecting from V1 through areas V2 and V3 to the middle temporal area (MT) and thence to additional areas in superior temporal and parietal cortex, processes attributes of stimuli important for localizing objects in space and for the visual guidance of movement towards them, such as the direction and velocity of stimulus motion [9].

Curr Opin Neurobiol, 1997, 554-61

20 years later…what have we learned? what questions remain?
overview

• why study human vision?
• retinotopic mapping
• functional specialization
  
  orientation
  motion
  color
  objects, faces, and letters

• bottom-up hierarchical processing within visual cortical areas
• top-down influences on visual cortical area
Abstract—Can the human brain itself serve as a model for a systems neuroscience approach to understanding the human brain? After all, how the brain is able to create the richness and complexity of human behavior is still largely mysterious. What better choice to study that complexity than to study it in humans? However, measurements of brain activity typically need to be made non-invasively which puts severe constraints on what can be learned about the internal workings of the brain. Our approach has been to use a combination of psychophysics in which we can use human behavioral flexibility to make quantitative measurements of behavior and link those through computational models to measurements of cortical activity through magnetic resonance imaging. In particular, we have tested various computational hypotheses about what neural mechanisms could account for behavioral enhancement with spatial attention (Pestilli et al., 2011). Resting both on quantitative measurements and considerations of what is known through animal models, we concluded that weighting of sensory signals by the magnitude of their response is a neural mechanism for efficient selection of sensory signals and consequent improvements in behavioral performance with attention. While animal models have many technical advantages over studying the brain in humans, we believe that human systems neuroscience should endeavor to validate, replicate and extend basic knowledge learned from animal model systems and thus form a bridge to understanding how the brain creates the complex and rich cognitive capacities of humans.

INTRODUCTION

A peculiar phenomenon had taken hold of the elevators of the Meyer building when I first arrived as a post-doc at NYU’s Center for Neural Science. Students, post-docs and professors all seemed to have a different algorithm for hitting the buttons on the elevator. Some would simply hit the button for their floor and wait. Others, though, would use different cryptic combinations of buttons, stretching their fingers wide to simultaneously press the floor they wanted and the current floor. For some, the order was apparently crucial – hitting first their floor before reaching for the current floor. Others used the exact opposite order. After inquiring around about this curious behavior, I was earnestly informed that these combinations of button presses were required to make the doors of the elevator close more quickly – a matter of great importance to impatient occupants of the building. But, what could explain the diversity of different techniques I had witnessed? After some time in the department, I developed my own (incompletely tested) theory – that the elevator had a time-out of a few seconds, after which any button press, or combination thereof, would trigger the doors to close. Thus, the occupants of the building had all learned various completely different behaviors, all of which produced the same reward of a swift start to the elevator ride.
overview

• why study human vision?
• retinotopic mapping
• functional specialization

  orientation
  motion
  color
  objects, faces, and letters

• bottom-up hierarchical processing within visual cortical areas
• top-down influences on visual cortical areas
what is a visual area?

PhACT

• Physiology
• Architecture
• Connections
• Topography
Topography (human V1)

Left visual cortex

Right visual field

Radial component

Angular component

Retinotopic mapping stimuli


Retinotopic map, timing of activity
retinotopic mapping stimuli

Radial component

Angular component


Retinotopic map, timing of activity
retinotopic mapping

0 sec

12 sec

0º

180º

Polar angle

DeYoe et al., 1996
Engel et al., 1997
retinotopic map, timing of activity

Time

fMRI signal
computational flattening
fMRI response (% signal change)

Time (s)
-2 0 2

Fovea
V1
V2d
V2v

0 sec
24 sec

0º
180º

Polar angle

Right hemisphere

fMRI response (% signal change)

Time (s)

0 24 48 72 96 120 144 168

0
2

a

φ

Right hemisphere
visual area boundaries

Sereno et al., Science, 1995
retinotopic mapping

progress and challenges…
**additional areas discovered**

Left hemisphere

Right hemisphere

Larsson & Heeger, 2006
multiple maps in IPS

Swisher et al., J. Neurosci, 2006
face topography in IPS

Sereno et al., Nat Neurosci (2006)
face topography in IPS

Right face air puff location
Left face air puff location

LH
RH

Face air puff maps

Central sulcus
Postcentral
Intraparietal
LIP+

Alignment index hist.

1 cm

90°

10^{-3}

Alignment index hist.
many (>25) visual areas!

retinotopic map, timing of activity
better mapping methods (pRFs)

- allows for more flexible stimulus sets
- more full characterization of visual responses (e.g., pRF size)
- richer set of models (surround suppression, non-Gaussian receptive fields)

https://dbirman.github.io/learn/prf/prf.html

after Doumolin & Wandell, 2008
pRFs are elliptical and radially oriented

Silson et al., J. Neurosci, 2018
retinotopic mapping

progress and challenges...
but we need better methods

A. Individual subjects’ retinotopic maps as predicted via Data alone

Training Data can be used as a prediction of Validation Data

Subject S1204, Training Dataset 1

B. Individual subjects’ retinotopic maps as predicted via the Prior alone

(1) Apply Prior Warping:
\[ \dot{x}_p = x_w \] (Fig. 1)

\[ \dot{x}_0 \rightarrow \dot{x}_p \]

(2) Apply Model to \( \dot{x}_p \)

\[ \dot{x}_p \rightarrow \dot{x}_w \]

(3) Unwarp Prediction

\[ \dot{x}_0 \rightarrow \dot{x}_w \rightarrow \dot{x}_0 \]

C. Individual subjects’ retinotopic maps as predicted via Bayesian inference (Data + Prior)

(1) Apply Prior Warping:
\[ \dot{x}_p = x_w \]

(2) Further Warp using Data:
\[ \dot{x}_w = \text{argmin}_x F(x; \dot{x}_p, \dot{E}, \dot{\Theta}, \dot{P}, \Phi) \]

(3) Apply Model to \( \dot{x}_w \)

(4) Unwarp Prediction

\[ \dot{x}_0 \rightarrow \dot{x}_p \rightarrow \dot{x}_w \rightarrow \dot{x}_w \rightarrow \dot{x}_0 \]

Benson & Winawer, bioRxiv, 2018
Winawer et al., 2010
better mapping methods (pRFs)

- allows for more flexible stimulus sets
- more full characterization of visual responses (e.g., pRF size)
- richer set of models (surround suppression, non-Gaussian receptive fields)

after Doumolin & Wandell, 2008
overview

- *why study human vision?*
- *retinotopic mapping*
overview

- why study human vision?
- retinotopic mapping
- functional specialization
  - orientation
  - motion
  - color
  - objects, faces, and letters
test case: **orientation selectivity**
can we measure it with fMRI?

hubel and wiesel, 1962
ringach, hawken, and shapley, 1997
blasdel and salama, 1986
doesn't look good...
solution 1: high resolution fMRI
human ocular dominance and orientation columns

yacoub, harel, ugurbil, PNAS (2008)
solution 2: fMRI decoding (classifiers)

small biases in fMRI response because voxels sample an irregular underlying columnar architecture

fMRI decoding: an imager’s microelectrode?

kamitani & tong, 2005
haynes & rees, 2005
boynton, 2005
**fMRI decoding with multi-voxel pattern analysis (MVPA)**

1) train classifier on subset of data
2) test classifier on left-out data
3) get a single number (% correct)
4) compare to ‘chance’

Haynes & Rees, Nature Neurosci, 2006
Figure 1. Orientation and polar-angle topographic maps for a single subject shown on a flattened representation of the occipital lobe.

(a) Responses to phase-encoded oriented grating (shown in inset). The stimulus cycled through 16 steps of orientation, ranging from 0° to 180° every 24 s. The map is thresholded at a coherence of 0.3.

(b) Stimulus (shown in inset). The stimulus cycled through 16 steps of polar angle, ranging from 0° to 180° every 24 s. The map is thresholded at a coherence of 0.68 to account for differences in signal-to-noise ratio between the two experiments (see Materials and Methods). Color indicates phase of best-fitting sinusoid; white lines indicate the V1/V2 boundaries.
Figure 1. Orientation and polar-angle topographic maps for a single subject shown on a flattened representation of the occipital lobe.

(a) Responses to phase-encoded oriented grating (shown in inset). The stimulus cycled through 16 steps of orientation, ranging from 0° to 180° every 24 s. The map is thresholded at a coherence of 0.3.

(b) Stimulus (shown in inset). The stimulus cycled through 16 steps of polar angle, ranging from 0° to 180° every 24 s. The map is thresholded at a coherence of 0.68 to account for differences in signal-to-noise ratio between the two experiments (see Materials and Methods). Color indicates phase of best-fitting sinusoid; white lines indicate the V1/V2 boundaries.

?
Figure 1. Orientation and polar-angle topographic maps for a single subject shown on a flattened representation of the occipital lobe.

(a) Responses to phase-encoded oriented grating (shown in inset). The stimulus cycled through 16 steps of orientation, ranging from 0° to 180° every 24 s. The map is thresholded at a coherence of 0.3.

(b) stimulus (shown in inset). The stimulus cycled through 16 steps of polar angle, ranging from 0° to 180° every 24 s. The map is thresholded at a coherence of 0.68 to account for differences in signal-to-noise ratio between the two experiments (see Materials and Methods). Color indicates phase of best-fitting sinusoid; white lines indicate the V1/V2 boundaries.
radial bias
map is **sufficient** for decoding
map is *sufficient* for decoding
classify averaged responses
map is **sufficient** for decoding

classify averaged responses

average

orientation

polar angle
map is **sufficient** for decoding

classify averaged responses

![Graph showing decoding accuracy for polar angle averaging]

**Decoding accuracy**

<table>
<thead>
<tr>
<th>Polar-angle bin width (°)</th>
<th>Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
</tr>
<tr>
<td>100</td>
<td></td>
</tr>
</tbody>
</table>
map is **sufficient** for decoding

classify averaged responses

![Diagram showing decoding accuracy for different polar-angle bin widths](image)
map is \textit{necessary} for decoding

classify residuals after removing response component predicted by polar angle
map is **necessary** for decoding

classify residuals after removing response component predicted by polar angle
overview

• why study human vision?
• retinotopic mapping
• functional specialization
  orientation
  motion
  color
  objects, faces, and letters
what (ventral) and where (dorsal) pathways
human MT

Huk, Dougherty, & Heeger (2002)

Amano, Dumoulin, & Wandell, J Neuophysiol (2009)
topography in human MT

Huk, Dougherty, & Heeger (2002)
topography in human MT

TO-1 = MT
TO-2 = MST

Amano, Dumoulin, & Wandell, J Neuophysiol (2009)
Direction-selective adaptation in human MT

Adaptation improves speed discrimination thresholds

<table>
<thead>
<tr>
<th>Subject</th>
<th>Adapted</th>
<th>Opposite</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACH</td>
<td>4.3</td>
<td>6.9</td>
</tr>
<tr>
<td>ARW</td>
<td>7.3</td>
<td>9.1</td>
</tr>
<tr>
<td>DJH</td>
<td>6.6</td>
<td>9.3</td>
</tr>
</tbody>
</table>

(\% speed increment)

Huk, Ress, & Heeger (2001)
Direction selectivity across visual areas

Huk, Ress, & Heeger (2001)

Motion adaptation index (adaptation response / baseline response)

Visual area

Huk, Ress, & Heeger (2001)
direction selective columns?

columnar architecture for motion direction in MT
motion direction bias
motion direction bias

Time

Fovea

V1

V2v

V3v

V2d

V3d

MT

MST

V1

V2v

V3v

V2d

V3d

MT

MST

Fovea

left hemisphere

right hemisphere

O1

Decoding accuracy (%)

Number of voxels

SNR ≥ 0.4

SNR ≥ 0.3

V1 V2 V3 MT+ GM

left hemisphere

right hemisphere
direction bias depends on aperture
direction bias depends on aperture
direction bias depends on aperture

V1

Visual field location (deg)

n = 4
suppression along path of motion

fMRI response (% change img intensity) vs. Cortical distance (mm)

Fovea-inward

Fovea-outward
overview

- why study human vision?
- retinotopic mapping
- functional specialization
  - orientation
  - motion
  - color
  - objects, faces, and letters

fMRI good for measuring coarse-scale patterns of activity
overview

• why study human vision?
• retinotopic mapping
• functional specialization
  orientation
  motion
  color
  objects, faces, and letters

• **bottom-up hierarchical processing within visual cortical areas**
• **top-down influences on visual cortical area**
hierarchical processing in visual cortex

Figure 4. Responses of human visual cortex to natural images, naturalistic textures, and spectrally matched noise controls. Flattened maps of the posterior pole of the right hemisphere of a representative observer showing modulation of BOLD fMRI responses to alternating sequences of stimuli in a block design whose time course is schematically diagrammed at the bottom: briefly presented repeated stimuli (200 msec on, 200 msec off) were presented in 9-sec blocks, alternating between the conditions being compared. Area boundaries are derived from a separate topographic mapping experiment. In each panel, the pseudocolor scale represents the coherence of the BOLD response component synchronized with the time course of the stimulus exchange.

(A) Alternation between naturalistic textures and spectrally matched noise controls. (B) Alternation between natural images and spectrally matched noise controls. (C) Alternation between natural images and naturalistic textures. (From Freeman et al. 2013b.)
overview

• why study human vision?
• retinotopic mapping
• functional specialization
  - orientation
  - motion
  - color
  - objects, faces, and letters
• bottom-up hierarchical processing within visual cortical areas
• top-down influences on visual cortical areas
source of top-down modulation
source of top-down modulation

Dugue, Merriam, et al., 2018
overview

• why study human vision?
• retinotopic mapping
• functional specialization

  orientation
  motion
  color
  objects, faces, and letters

• bottom-up hierarchical processing within visual cortical areas
• top-down influences on visual cortical areas

The future is layer fMRI!!
thank you!