A M/EEG-fMRI fusion primer:

Resolving Human Brain Responses in Space and Time

Aude Oliva

CSAIL, MIT-IBM Watson AI Lab, MIT Quest for Intelligence
MIT Schwarzman College of Computing
Massachusetts Institute of Technology
Brain: A **Dynamic** Deep Learning Network

- **100 billion neurons**
  (80 billion in cerebellum)
- **100 trillion connections**
  (Each neuron connects to 5000 to 200,000 other neurons)
- **10,000 different types of neurons**
- **1000 > new neurons per day, all your life**

- **Consumes 20 watts**
Spatio-temporal maps of human brain processes

- Basic sensory processing
- Recognition & Memory
- Learning & Plasticity
- Multi-sensory integration
- Working memory, Attention
- Clinical research

Brain-Machine Interfaces, Deep Neural Network Architectures, Comparison across Species
MEG: Time
Every millisecond

fMRI: Space
Each millimeter

Sensors (time)
Evoked activity
every msec

Voxels (space)

Relative distances

Shepard et al., 1980; Kruskal and Wish., 1978; Edelman et al. 1998; Kriegeskorte et al., 2008; Mur et al., 2009; Liu et al., 2013
Representational Dissimilarity Matrices (RDMs)

Nikolaus Kriegeskorte (2008): "RDMs as a hub to relate different representations across sensors and models"
Spatio-temporal maps of brain responses

A spatially unbiased view of the relations in similarity structure between MEG and fMRI

MEG

306 x 1 vector

t

MEG, 170ms

SVM classification

Representational Dynamics: MEG Dissimilarity Matrices
One matrix per millisecond

Cichy, Pantazis, Oliva (2014). Nature Neuroscience
Spatio-temporal maps of brain responses

A spatially unbiased view of the relations in similarity structure between MEG and fMRI

MEG

306 x 1 vector

fMRI

Voxels within searchlight

SVM classification

MEG, 170ms

Spearman R assign to voxel

(dissimilarity 1 - R)

fMRI voxel

Pattern comparison (Pearson R)


Kriegeskorte et al., 2008
Spatio-temporal maps of brain responses
Visual Perception

• A hierarchy of responses from occipital to ventral and parietal regions during object recognition

• The spatio-temporal maps time stamps align with previous neurophysiology results

• Both ventral and dorsal responses peaks occurring around 170 msec (IPS, LO)

• **Lesson for models:** Different high-level representations of visual events after an initial common processing


<table>
<thead>
<tr>
<th>Region of Interest</th>
<th>Peak latency (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PAC (TE1.0 &amp; TE1.1)</td>
<td>115 (100, 191)</td>
</tr>
<tr>
<td>TE1.2</td>
<td>117 (111, 207)</td>
</tr>
<tr>
<td>Planum Temporale (PT)</td>
<td>115 (103, 197)</td>
</tr>
<tr>
<td>Planum Polare (PP)</td>
<td>117 (110, 207)</td>
</tr>
<tr>
<td>Temporal Voice Area (TVAx)</td>
<td>199 (188, 251)</td>
</tr>
<tr>
<td>Left Inferior Frontal Gyrus (LiFG)</td>
<td>196 (188, 380)</td>
</tr>
<tr>
<td>Fusiform Face Area (FFA)</td>
<td>298 (200, 368)</td>
</tr>
<tr>
<td>Parahippocampal Place Area (PPA)</td>
<td>293 (219, 380)</td>
</tr>
<tr>
<td>Medial Place Area (MPA)</td>
<td>199 (196, 400)</td>
</tr>
<tr>
<td>Lateral Occipital Complex (LOC)</td>
<td>377 (169, 410)</td>
</tr>
</tbody>
</table>
Auditory Perception

• Timing of auditory cortex, pre-frontal and higher-level regions responses in human

• Specificity of human voices in auditory, pre-frontal and fusiform gyrus with time stamps

• A late response (~ 300 msec) for auditory stimuli in regions classically associated with vision

• Lesson for models: Human voices have specialized modules, being separated early on; audition is mixed with visual features at a later stage in the hierarchy of processing
Memorability

Visual memorability is a consequence of the optimizations required for visual processing.
Early Visual Cortex, V5/MT+, Posterior Temporal Lobe

With Y. Mohsenzadeh, B. Lahner, C. Mullin
Memorability

• Late responses for memorable images in higher cortical regions associated with object recognition

• Memorable images have a stronger signals during visual processing

• Visual memorability is a consequence of the optimizations required for visual processing

• **Lesson for models:** Memorability as a measure of the *utility* of information
Human & Artificial Cognition

- Characterizing the bandwidth of human perception and cognition is critical.

- A new field of investigation: Cognitive / Clinical / Social / Perceptual Computational Experimentalist / Synthetic Neuroscientist.

- Studying the implementation that works best for performing specific tasks.

- Exploring the alternatives that have not been taken by biological systems.
Datasets and Models

Memento10k: Video Memorability Dataset (to be released soon)
Memento10k is the largest-in-the-wild video memorability dataset to date, with more than 10,000 videos and close to 1 million human annotations. Videos represent varied everyday events, captured in a homemade fashion. Each video was annotated through our Memento Memory Game and possesses 90 human annotations on average. We also release action labels, as well as 5 detailed captions for each video.

Moments in Time and Multi-Moments in Time
Moments in Time is a research project aiming to build a very large-scale dataset to help AI systems recognize events in videos. The first release includes one million 3 second videos each with one activity label (covering >500 action categories). The second release Multi-Moments in Time (M-MIT) includes over 2 million labels. The third version Spoken Moments will be released by ECCV 2020.

GANalyze: Generate Memorable and Forgettable Images
A framework that uses Generative Adversarial Networks (GANs) to study cognitive properties like memorability, aesthetics, and emotional valence. GANs allow us to generate a manifold of natural-looking images with fine-grained differences in their visual attributes. By navigating this manifold in directions that increase memorability, we can visualize what it looks like for a particular generated image to become more or less memorable.

The Algonauts Project: 2019 Edition
The Algonauts Project brings biological and artificial intelligence researchers together on a common platform to exchange ideas and advance both fields. Our first challenge Explaining the Human Visual Brain, focused on building computer vision models that are able to efficiently recognize objects. The released dataset includes multiple image sets with fMRI and MEG human brain data. The second Algonauts challenge will be announced during Summer 2020.

MEG and fMRI Data of Images
We recorded magnetoencephalography (MEG) and functional magnetic resonance imaging (fMRI) data while 15 participants viewed a set of 156 natural images. These images can be subdivided into five categories (faces, bodies, animals, objects, scenes) or two twinsets of 78 images each.