Pattern-information fMRI and representational similarity analysis

Mathematics in Brain Imaging
IPAM, UCLA, 25 July 2008, Los Angeles

Nikolaus Kriegeskorte
Section on Functional Imaging Methods, Laboratory of Brain and Cognition
National Institute of Mental Health
response channels

functional region

response amplitude

faces

objects

stimuli
Activation analysis

functional region

response channels

response amplitude

stimuli

faces

objects
Pattern-information analysis

functional region

response channels

response amplitude

stimuli

objects

faces
Goal: honor all these distinctions
Talk overview

Specific neuroscientific experiments

• inferior temporal object representations in human and monkey

General methodology

• every stimulus is a condition
• condition-rich fMRI design
• representational similarity analysis
Related literature

• multidimensional scaling

• second-order isomorphism
  Shepard & Chipman (1970)

• first application of multidimensional scaling to fMRI data
  Edelman et al. (1998)

• some more recent studies with similarity analyses
Collaborators

Bethesda, MD, USA
• Marieke Mur
• Douglas Ruff
• Jerzy Bodurka
• Peter Bandettini

Seattle, WA, USA
• Roozbeh Kiani

Tehran, Iran
• Hossein Esteky

Wako, Saitama, Japan
• Keiji Tanaka
Overview of experiment and analysis

• present images of real-world objects to subjects (human, monkey)

• measure the brain-activity pattern during perception of each particular image (fMRI, cell recording)

• study similarity structure of object representations focusing on inferior temporal (IT) cortex

• compare representations in human and monkey IT by relating representational similarity matrices
Core concept
representational similarity matrix
dissimilarity
dissimilarity matrix
dissimilarity-graph icon

compute dissimilarity
(1-correlation across space)

activity patterns

brain

experimental conditions

p< 1%
5%
Transformation of representational similarity

Parahippocampal place area (PPA)

Fusiform face area (FFA)

Anterior inferotemporal face-exemplar region

Early visual areas

Kriegeskorte et al. 2007
rubberband graph
- stretched effects: thin
- compressed effects: thick
- effect [bits] $\propto$ length $\cdot$ tickness

Anterior inferotemporal face-exemplar region

Kriegeskorte et al. 2007
96-stimulus experiment
Brain-activity measurements

**Human**
- fMRI in four subjects (>12 runs per subject)
- task: fixate, discriminate fixation-point color changes
- occipitotemporal measurement slab (5-cm thick)
- voxels: 1.95×1.95×2mm³
- 3T magnet, 16-channel coil (SENSE, acc. fac. 2)
- stimulus duration: 300ms
- quick event-related (SOA=4s)

**Monkey** *(Kiani et al. 2007)*
- single-cell recordings in two monkeys
- task: fixation
- electrodes in anterior IT (left in monkey 1, right in monkey 2)
- 674 cells total
- stimulus duration: 105ms
- rapid serial presentation
- 140-ms window spike count starting 71ms after stimulus onset
Single-image hemodynamic response predictors
Design matrix

- 96 experimental events
- run 1, run 2, run 3, run 4
- TR = 2s
- predictor
- confound means
- head-motion parameters
- trend models
Stimulus quartets
(multidimensional scaling, metric stress)

(1) multidimensional scaling (1-r, metric stress)
(2) rigid alignment for easier comparison (Procrustes alignment)
Stimulus quartets
(multidimensional scaling, metric stress)

human
monkey
Stimulus quartets
(multidimensional scaling, metric stress)

human

monkey
Stimulus quartets
(multidimensional scaling, metric stress)

human

monkey
Stimulus quartets
(multidimensional scaling, metric stress)

human

monkey
Stimulus quartets
(multidimensional scaling, metric stress)

human

monkey
Stimulus quartets
(multidimensional scaling, metric stress)

human

monkey
Stimulus quartets
(multidimensional scaling, metric stress)

human
monkey
compute dissimilarity (1-correlation across space)
Human IT

(1000 most visually responsive voxels within anatomical IT mask)

<table>
<thead>
<tr>
<th>human</th>
<th>animate</th>
<th>inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>body</td>
<td>not ~</td>
<td>natural</td>
</tr>
<tr>
<td>face</td>
<td>body</td>
<td>artificial</td>
</tr>
</tbody>
</table>

- human face 1
- human face 2

<table>
<thead>
<tr>
<th>animate</th>
<th>human</th>
<th>inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>not ~</td>
<td>body</td>
<td>natural</td>
</tr>
<tr>
<td>body</td>
<td>face</td>
<td>artificial</td>
</tr>
</tbody>
</table>

- animate
- inanimate

matrices of 4 subjects averaged
To visualize the representation...

Let's arrange the images in a figure reflecting response-pattern similarity

- images close $\Rightarrow$ similar response patterns
- images far $\Rightarrow$ dissimilar response patterns

Multidimensional scaling

first application to fMRI: Edelman et al. (1998)
Human IT
(1000 visually most responsive voxels)
Human IT
(1000 visually most responsive voxels)

animate  inanim
Human IT
(1000 visually most responsive voxels)

animate

inanim
Human IT
(1000 visually most responsive voxels)
Human IT
(1000 visually most responsive voxels)
**Human IT (larger)**

(3162 visually most responsive voxels)

<table>
<thead>
<tr>
<th>human body</th>
<th>face</th>
<th>animate</th>
<th>not ~ body</th>
<th>face</th>
<th>inanimate</th>
<th>natural</th>
<th>artificial</th>
</tr>
</thead>
</table>

![Heatmap and scatter plot showing the distribution of visually most responsive voxels. The heatmap is color-coded with red and blue indicating different activity levels, and the scatter plot on the right distinguishes between animate and inanimate categories.](image-url)
Human IT (smaller)
(316 visually most responsive voxels)
Human early visual cortex
(1057 visually most responsive voxels)
Human IT
(1000 visually most responsive voxels)
similarity matrices

activity patterns

human brain

monkey brain

experimental events
Monkey IT
(674 cells in left and right aIT, windowed spike count, fixation task, 2 monkeys)

data from Roozbeh Kiani et al. (2007), Journal of Neurophysiology
average of 4 subjects
fixation-color task
316 voxels

human

monkey

average of 2 monkeys
fixation task
674 cells

Kriegeskorte et al. (submitted)
Stimulus arrangements reflecting response-pattern similarity

(1) multidimensional scaling (1-r, metric stress)
(2) rigid alignment for easier comparison (Procrustes alignment)
monkey IT dissimilarity

human IT dissimilarity (1-r)

monkey IT dissimilarity (1-r)
average of 4 subjects
fixation-color task
316 voxels

average of 2 monkeys
fixation task
>600 cells

Kriegeskorte et al. (under revision)
Man-to-monkey correlation of dissimilarities

within all images: $r=0.49$, $p<0.0001^{***}$
Man-to-monkey correlation of dissimilarities

within all images: $r=0.49$, $p<0.0001^{***}$
Man-to-monkey correlation of dissimilarities

within all images: $r=0.49$, $p<0.0001^{***}$

within animates: $r=0.51$, $p<0.0001^{***}$
Man-to-monkey correlation of dissimilarities

within all images: $r=0.49$, $p<0.0001^{***}$
within animates: $r=0.51$, $p<0.0001^{***}$
within inanimates: $r=0.20$, $p<0.0001^{***}$
Man-to-monkey correlation of dissimilarites

significant also...
• within faces
• within bodies
• within humans
• within nonhuman animals

not significant...
• within human faces
• within animal faces
Representational connectivity

EVC $\rightarrow$ LO

- All pairs: $r=0.43$
- Within-animate: $r=0.48$
- Within-inanimate: $r=0.59$
- Between-animate & -animate: $r=0.43$
(1) multidimensional scaling (1-r, metric stress)

(2) rigid alignment for easier comparison (Procrustes method)

bilateral EVC

animate inanimate

right LO
Summary so far...

• **Natural categorical structure in primate IT**
  – top-level distinction: animate | inanimate
  – faces form a separate tight cluster

• **Man-monkey match of IT representational similarity**
  – matching major categorical clusters
  – matching within-category similarity structure

**Interpretation**

Human and monkey IT may host a common code for continuous and categorical object information.
But questions remain...

• Can the apparent categorical structure be explained by low-level features?

• More generally, what kind of computational model can explain our findings?
From description to explanation...

• use a range of models to predict representational similarity

• models can be computational, conceptual, or based on behavioral data
similarity matrices

activity patterns

experimental events

brain

model

compare
Model dissimilarity matrices
Early visual cortex

(dissimilarity SNR=0.61)
Human IT

(dissimilarity SNR=1.29)

brain region
naive model
neuro-model
category model
Simultaneously relating multiple representational similarity matrices

multidimensional scaling (metric stress, 1-r(Spearman))

brain region
naive model
neuro-model
category model
Summary of model-based analyses

• We compared a range of low- and intermediate-level model representations to the IT representation.
• None of them could explain the categorical clustering we saw in human and monkey IT.

Interpretation

An IT model probably needs to include category-specific visual knowledge – as might be acquired by supervised learning.
Methodological concepts
Searchlight mapping

Kriegeskorte et al. 2006
Searchlight mapping

\[ (p_2 - p_1)^T \Sigma^{-1}(p_2 - p_1) \]

Kriegeskorte et al. 2006
Representational-similarity searchlight mapping

searchlight

local fMRI similarity matrix

model similarity matrix

compare
computational models
- symbolic models
- connectionist models
- biological neural models

brain-activity data
- cell recordings
- fMRI
- EEG, MEG

behavioral data
- reaction time
- errors
- explicit judgements
behavioral data
• reaction time
• errors
• explicit judgements

brain-activity data
• cell recordings
• fMRI
• EEG, MEG

computational models
• symbolic models
• connectionist models
• biological neural models

representational similarity matrix

behavioral data
• reaction time
• errors
• explicit judgements