Percept Decoding with Sparse Latent Variable Models

Marcel van Gerven
Donders Centre for Cognition
Radboud University Nijmegen

Machine Learning and Neuroimaging Workshop

woensdag 9 november 2011
Percept decoding

encoding

luminance
gist
contrast
motion
color
semantics
Percept decoding

- encoding
- brain activity
- luminance
gist
contrast
motion
color
semantics
Percept decoding

brain activity

encoding

decoding

luminance
gist
contrast
motion
color
semantics

woensdag 9 november 2011
Percept decoding

encoding

match?

decoding

brain activity

match?

luminance
gist
contrast
motion
color
semantics

luminance
gist
contrast
motion
color
semantics

Donders Institute
for Brain, Cognition and Behaviour

woensdag 9 november 2011
Motivation

Understand how percepts are encoded in the brain by exploiting multivariate analysis methods
Decoding subjective experience

- Imagery
- Memory

Match?

Brain activity

Decoding

- Luminance
- Contrast
- Motion
- Color
- Semantics

Gist
Decoding low-level stimulus properties

Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).
Decoding low-level stimulus properties

Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).

generative approach:
Decoding low-level stimulus properties

Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).

generative approach:
Decoding low-level stimulus properties

Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).

generative approach:  

---

discriminative approach:  

---
Decoding low-level stimulus properties

Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).

generative approach: sparse latent variable models: interpretable, stable
discriminative approach:
Decoding low-level stimulus properties

Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).

generative approach:

discriminative approach:

sparse latent variable models: interpretable, stable

Outlook:

- generative sparse model
- generative latent variable model
- discriminative sparse latent variable model
- decoding high-level stimulus properties

woensdag 9 november 2011
Generative approach: sparse encoding model

Generative approach: sparse encoding model


For each voxel $k$: $p(r \mid s) = \mathcal{N}(r; \alpha_k + \beta_k^T s, \sigma_k)$
Generative approach: sparse encoding model


For each voxel $k$: 
$$p(r \mid s) = \mathcal{N}(r; \alpha_k + \beta_k^T s, \sigma_k)$$

Choose 
$$- \log p(\alpha_k, \beta_k, \sigma_k^2) \propto R_{\lambda, \tau}(\beta_k)$$  
with elastic net regularizer

$$R_{\lambda, \tau}(\beta_k) = \lambda \sum_{k=1}^{K} \{(1 - \tau) \frac{1}{2} \| \beta_k \|_2^2 + \tau \| \beta_k \|_1 \}$$
Generative approach: sparse encoding model


For each voxel $k$: $p(r \mid s) = \mathcal{N}(r; \alpha_k + \beta_k^T s, \sigma_k)$

Choose $- \log p(\alpha_k, \beta_k, \sigma_k^2) \propto R_{\lambda, \tau}(\beta_k)$ with elastic net regularizer

$$R_{\lambda, \tau}(\beta_k) = \lambda \sum_{k=1}^{K} \{(1 - \tau) \frac{1}{2} \| \beta_k \|_2^2 + \tau \| \beta_k \|_1 \}$$

Solve $k$ independent elastic net problems:

$$\hat{\theta}_k = \arg \min_{\alpha_k, \beta_k, \sigma_k^2} \left\{ - \log p(\alpha_k, \beta_k, \sigma_k^2) - \sum_n \log \mathcal{N}(r_k^n; \alpha_k + \beta_k^T s_n, \sigma_k^2) \right\}$$
It can be shown that

\[ p(r | s) = \frac{1}{Z} \prod_i \psi_i(s_i) \prod_{i \sim j} \psi_{i,j}(s_i, s_j) \]

where

\[ \psi_i(s_i) = \exp \left( s_i \sum_k \frac{\beta_{ki}}{\sigma_k^2} (r_k - \alpha_k - \frac{1}{2} \beta_{ki}) \right) \]

\[ \psi_{i,j}(s_i, s_j) = \exp \left( -s_i s_j \sum_k \frac{\beta_{ki}}{\sigma_k^2} \beta_{k,j} \right) \]
It can be shown that

\[ p(r|s) = \frac{1}{Z} \prod_i \psi_i(s_i) \prod_{i \sim j} \psi_{i,j}(s_i, s_j) \]

where

\[ \psi_i(s_i) = \exp \left( s_i \sum_k \beta_{ki} (r_k - \alpha_k - \frac{1}{2} \beta_{ki}) \right) \]

\[ \psi_{i,j}(s_i, s_j) = \exp \left( -s_i s_j \sum_k \frac{\beta_{ki}}{\sigma_k^2} \beta_{kj} \right) \]
Define appropriate MRF prior:

\[ p(s) = \frac{1}{Z} \prod_i \phi_i(s_i) \prod_{i \sim j} \phi_{i,j}(s_i, s_j) \]
Generative approach: decoding

Define appropriate MRF prior:

\[
p(s) = \frac{1}{Z} \prod_i \phi_i(s_i) \prod_{i \sim j} \phi_{i,j}(s_i, s_j)
\]

Estimate the mode of the following MRF:

\[
p(s|r) = \frac{1}{Z} \prod_i (\phi_i(s_i)\psi_i(s_i)) \prod_{i \sim j} (\phi_{i,j}(s_i, s_j)\psi_{i,j}(s_i, s_j))
\]
Results

- Miyawaki et al., Neuron, 2008
- 10x10 images (random/geometric)
- BOLD response measured in
  1017 voxels in primary visual cortex
Results

- Miyawaki et al., Neuron, 2008
- 10x10 images (random/geometric)
- BOLD response measured in 1017 voxels in primary visual cortex

encoding

\[
\begin{array}{cc}
V1 & R^2 = -0.01 \\
\uparrow & R^2 = 0.49 \\
V2 & R^2 = -0.04 \\
\downarrow & R^2 = 0.42 \\
V3 & R^2 = 0.47 \\
\end{array}
\]
Results

- Miyawaki et al., Neuron, 2008
- 10x10 images (random/geometric)
- BOLD response measured in 1017 voxels in primary visual cortex

encoding

decoding

\[ R^2 = -0.01 \]

\[ R^2 = 0.49 \]

\[ R^2 = -0.01 \]

\[ R^2 = 0.42 \]

\[ R^2 = -0.04 \]

\[ R^2 = 0.47 \]

random

geometric

\[ \text{Manhattan distance} \]

\[ \text{included voxels} \]
Generative approach: latent variables

Generative approach: latent variables

Deep belief network

hidden layer

hidden layer

hidden layer

input layer

Generative approach: latent variables

Deep belief network

hidden layer

hidden layer

hidden layer

input layer

restricted Boltzmann machine

Generative approach: latent variables

Deep belief network

hidden layer

hidden layer

hidden layer

input layer

restricted Boltzmann machine


woensdag 9 november 2011
Induce a coupling with conditional restricted Boltzmann machines

\[ E(v, h \mid z) = -h^T W v - z^T C v - z^T B h \]
Using a hierarchical generative model

Reconstruction phase
Experimental results

12.5 s

12500 ms

12.5 s

12500 ms

12.5 s

12500 ms

woensdag 9 november 2011
Experimental results

12500 ms
12.5 s
12500 ms
12.5 s
12500 ms
12.5 s

Donders Institute
for Brain, Cognition and Behaviour
Radboud University Nijmegen

woensdag 9 november 2011
Experimental results

12.5 s
12500 ms

12.5 s
12500 ms

12.5 s
12500 ms

Donders Institute
for Brain, Cognition and Behaviour

Radboud University Nijmegen

woensdag 9 november 2011
Experimental results

12.5 s
12.5 s
12.5 s

woensdag 9 november 2011
• learn deep belief network
Putting things together

**BOLD response**

- Learn deep belief network
- Learn responses using elastic net
Putting things together

- learn deep belief network
- learn responses using elastic net
- represent as a Markov random field
Putting things together

- learn deep belief network
- learn responses using elastic net
- represent as a Markov random field
- decode by estimating the mode of the MRF
Discriminative approach: sparse partial least squares

Predict image from a restricted set of responses using a small number of latent variables

Discriminative approach: sparse partial least squares

Predict image from a restricted set of responses using a small number of latent variables

Key features:

Discriminative approach: sparse partial least squares

Predict image from a restricted set of responses using a small number of latent variables

Key features:
- Linear: not enough data to (consistently) find strong nonlinear effects, stable, fast.

Discriminative approach: sparse partial least squares

Predict image from a restricted set of responses using a small number of latent variables

Key features:

- Linear: not enough data to (consistently) find strong nonlinear effects, stable, fast.
- Dimension reduction: gives a smooth image-like output, helps prevent overfitting.

Discriminative approach: sparse partial least squares

Predict image from a restricted set of responses using a small number of latent variables

Key features:

▶ Linear: not enough data to (consistently) find strong nonlinear effects, stable, fast.
▶ Dimension reduction: gives a smooth image-like output, helps prevent overfitting.
▶ Sparsity: small number of relevant voxels makes the model interpretable.

Partial least squares

Linear heteroencoder:

- Unique optimal solution (no local minima).
- reduces to principal component analysis in case \( x=y \);
  rows of \( Q \) correspond to principal components of \( y \).

\[
\begin{align*}
  z &= P^\top x \\
  y &= Qz
\end{align*}
\]
Sparse partial least squares

\[ y \]

\[ P \]

\[ Q \]

\[ x \]

\[ z \]

\[ y \]
Sparse partial least squares

Objective:

\[
(\hat{P}, \hat{Q}) = \arg \min_{P, Q} \left[ \frac{1}{2N} \sum_{n=1}^{N} \left\| y^{(n)} - QP^T x^{(n)} \right\|_2^2 + R_{\nu, \Lambda}(P) \right]
\]

with \( R_{\nu, \Lambda}(P) = \nu \sum_{i=1}^{k} \|P_i\|_1 + \frac{1}{2} \sum_{j=1}^{k} P_j^T \Lambda P_j \)
Sparse partial least squares

\[ (\hat{P}, \hat{Q}) = \arg\min_{P, Q} \left[ \frac{1}{2N} \sum_{n=1}^{N} \| y^{(n)} - QP^T x^{(n)} \|^2 + R_{\nu, \Lambda}(P) \right] \]

with \[ R_{\nu, \Lambda}(P) = \nu \sum_{i=1}^{k} \| P_i \|_1 + \frac{1}{2} \sum_{j=1}^{k} P_j^T \Lambda P_j \]

\( \triangleright \) reduces to sparse PCA in case \( x = y \) \( \text{(Zou et al., J Comput Graph Stat, 2006)} \)
Iterative sparse PLS algorithm

Fix $Q$, reconstruct $Z = Q^T Y$, and solve

$$\hat{P} = \arg\min_{P} \left[ \frac{1}{2N} \sum_{n=1}^{N} \|z^{(n)} - P^T x^{(n)}\|_2^2 + R_{\nu, \Lambda}(P) \right]$$
Iterative sparse PLS algorithm

Fix $Q$, reconstruct $Z = Q^T Y$, and solve

$$\hat{P} = \arg\min_P \left[ \frac{1}{2N} \sum_{n=1}^{N} \|z^{(n)} - P^T x^{(n)} \|_2^2 + R_{\nu,\Lambda}(P) \right]$$

- set of standard elastic net problems

Iterative sparse PLS algorithm

Fix $P$, reconstruct $Z = P^T X$, and solve

$$
\hat{Q} = \arg \min_P \left[ \frac{1}{2N} \sum_{n=1}^{N} \| y^{(n)} - Q^T z^{(n)} \|_2^2 \right]
$$

subject to $Q^T Q = I_k$
Iterative sparse PLS algorithm

Fix $P$, reconstruct $Z = P^T X$, and solve

$$\hat{Q} = \arg \min_P \left[ \frac{1}{2N} \sum_{n=1}^{N} \|y^{(n)} - Q^T z^{(n)}\|_2^2 \right]$$

subject to $Q^T Q = I_k$

$$\hat{Q} = \Sigma_{yz} \left( \Sigma_{yz}^T \Sigma_{yz} \right)^{-1/2} \quad \text{with} \quad \Sigma_{yz} \equiv \frac{1}{N} \sum_{n=1}^{N} y^{(n)} (z^{(n)})^T$$
Experiment

- Miyawaki et al., Neuron, 2008
- 10x10 images (geometric)
- BOLD response measured in 1017 voxels in primary visual cortex
- 10 latent variables, $\nu = 0.01$

Learned features

Learned features (rows of the matrix $\mathbf{Q}$) are similar to principal components of the original images but change as a function of $\nu$
Reconstructions on hold-out data
Sparseness

selected voxels for the first latent variable out of all voxels in primary visual cortex

For $\nu = 0.01$, 80% of the parameters are set to zero
Decoding high-level stimulus properties?

Classification of handwritten characters classes (6 vs 9) from BOLD response:

![Handwritten characters (6 vs 9)]

![Brain image with regions labeled 6/9]
Adding prior knowledge

We require:

\[ p(\theta \mid D) \propto p(D \mid \theta)p(\theta) \]

likelihood: logistic regression model

prior: sparse and smooth solutions

\[ p(\theta) = \int duv \left( \prod_k N(\theta_k; 0, u_k^2 + v_k^2) \right) N(u; 0, Q)N(v; 0, Q) \]

• preference for small regression coefficients
• large magnitude in one coefficient reduces regularization of coupled coefficients
Bayesian logistic regression with the multivariate Laplace prior

- Posteriors are computed with expectation propagation
- Scales well with the number of variables
- Computation time depends on the amount of coupling


Conclusions and future work
Conclusions and future work

- Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance
Conclusions and future work

- Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

- Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:
Conclusions and future work

- Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

- Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:
- other regularizers (e.g., total variation)
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:

• other regularizers (e.g., total variation)
• multitask formulation of SPLS
Conclusions and future work

- Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

- Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:

- other regularizers (e.g., total variation)
- multitask formulation of SPLS
- Bayesian formulation of SPLS
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:

• other regularizers (e.g., total variation)
• multitask formulation of SPLS
• Bayesian formulation of SPLS
• combine low-level and high-level information
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:

• other regularizers (e.g., total variation)
• multitask formulation of SPLS
• Bayesian formulation of SPLS
• combine low-level and high-level information
• focus on reactivation during imagery and memory retrieval
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:

• other regularizers (e.g., total variation)
• multitask formulation of SPLS
• Bayesian formulation of SPLS
• combine low-level and high-level information
• focus on reactivation during imagery and memory retrieval
• use intracranial EEG and high-field fMRI
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:

• other regularizers (e.g., total variation)
• multitask formulation of SPLS
• Bayesian formulation of SPLS
• combine low-level and high-level information
• focus on reactivation during imagery and memory retrieval
• use intracranial EEG and high-field fMRI

Acknowledgements:

Tom Heskes, Floris de Lange, Eric Maris
Conclusions and future work

• Generative approach via Bayesian inversion of sparse forward model, possibly on abstract stimulus features (also see Thirion, 2006, Gallant, 2009) - easier to interpret

• Discriminative approach via sparse partial least squares (also see Kamitani, 2008) - better decoding performance

Extensions:
• other regularizers (e.g., total variation)
• multitask formulation of SPLS
• Bayesian formulation of SPLS
• combine low-level and high-level information
• focus on reactivation during imagery and memory retrieval
• use intracranial EEG and high-field fMRI

Acknowledgements:
Tom Heskes, Floris de Lange, Eric Maris

methods part of FieldTrip multivariate module (http://fieldtrip.fcdonders.nl)
Face decoding
Spatial statistics

woensdag 9 november 2011
Decoding spatial statistics
Pilot results

spatial statistics

semantics

decoding of gender 80% correct
Face identification
Face identification

laughing
Face identification

laughing

young
Face identification

laughing

young

woman
Imagined face identification