Donders Institute
for Brain, Cognition and Behaviour

Machine Learning and Neuroimaging Workshop

Percept Decoding with Sparse Latent Variable Models

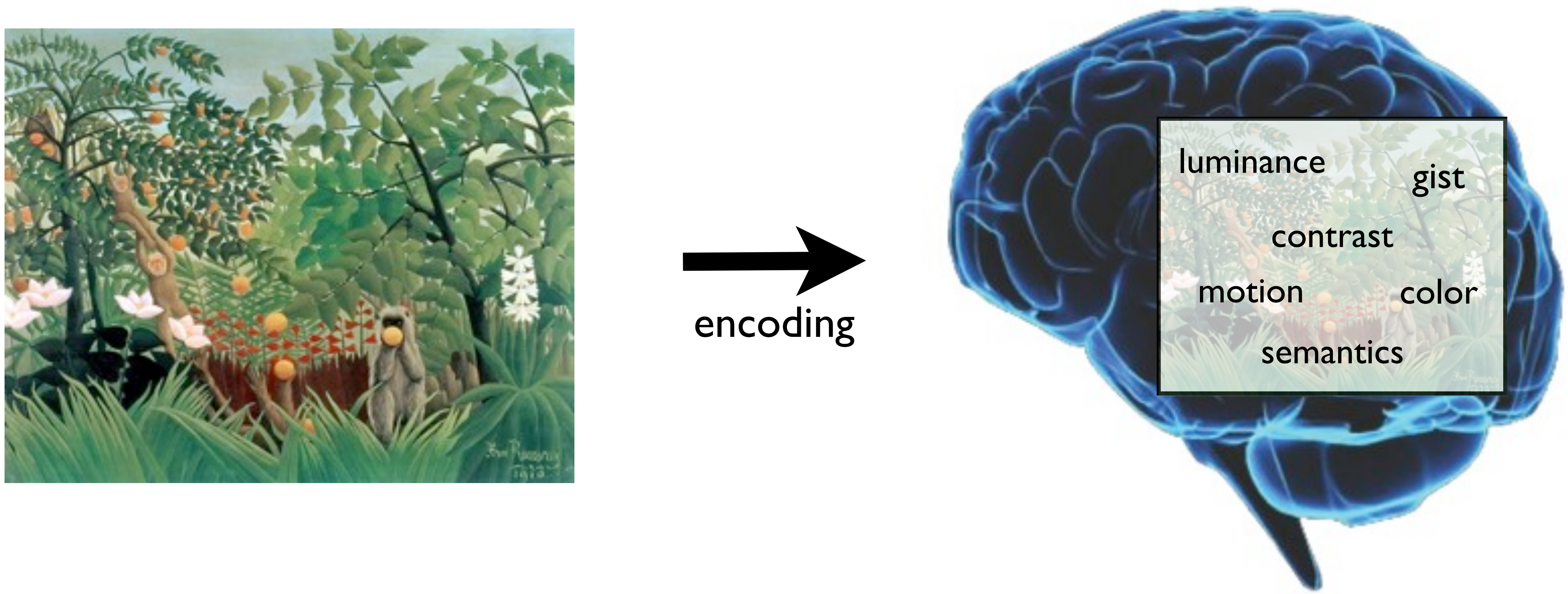
Marcel van Gerven

Donders Centre for Cognition

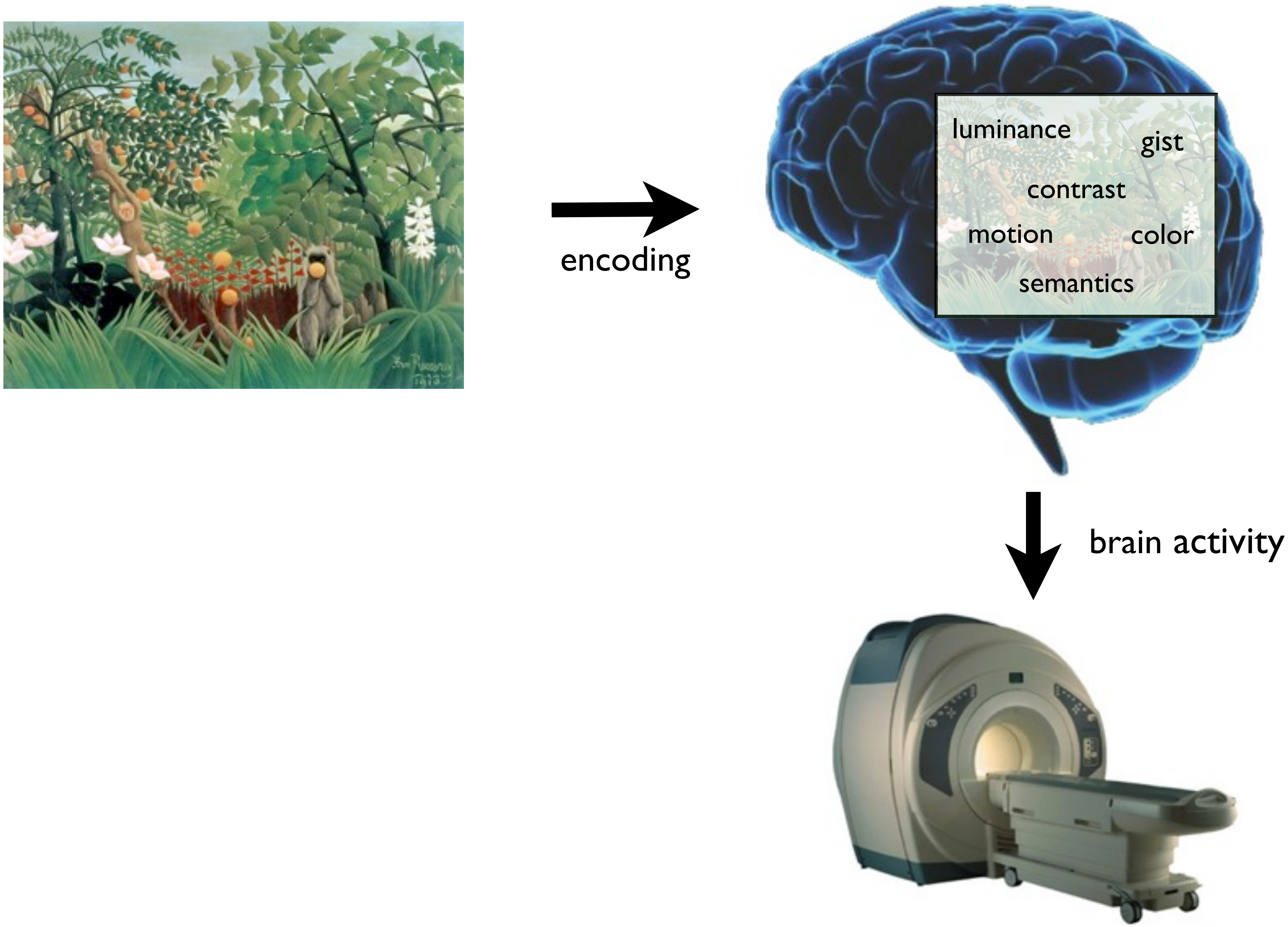
Radboud University Nijmegen

Percept decoding

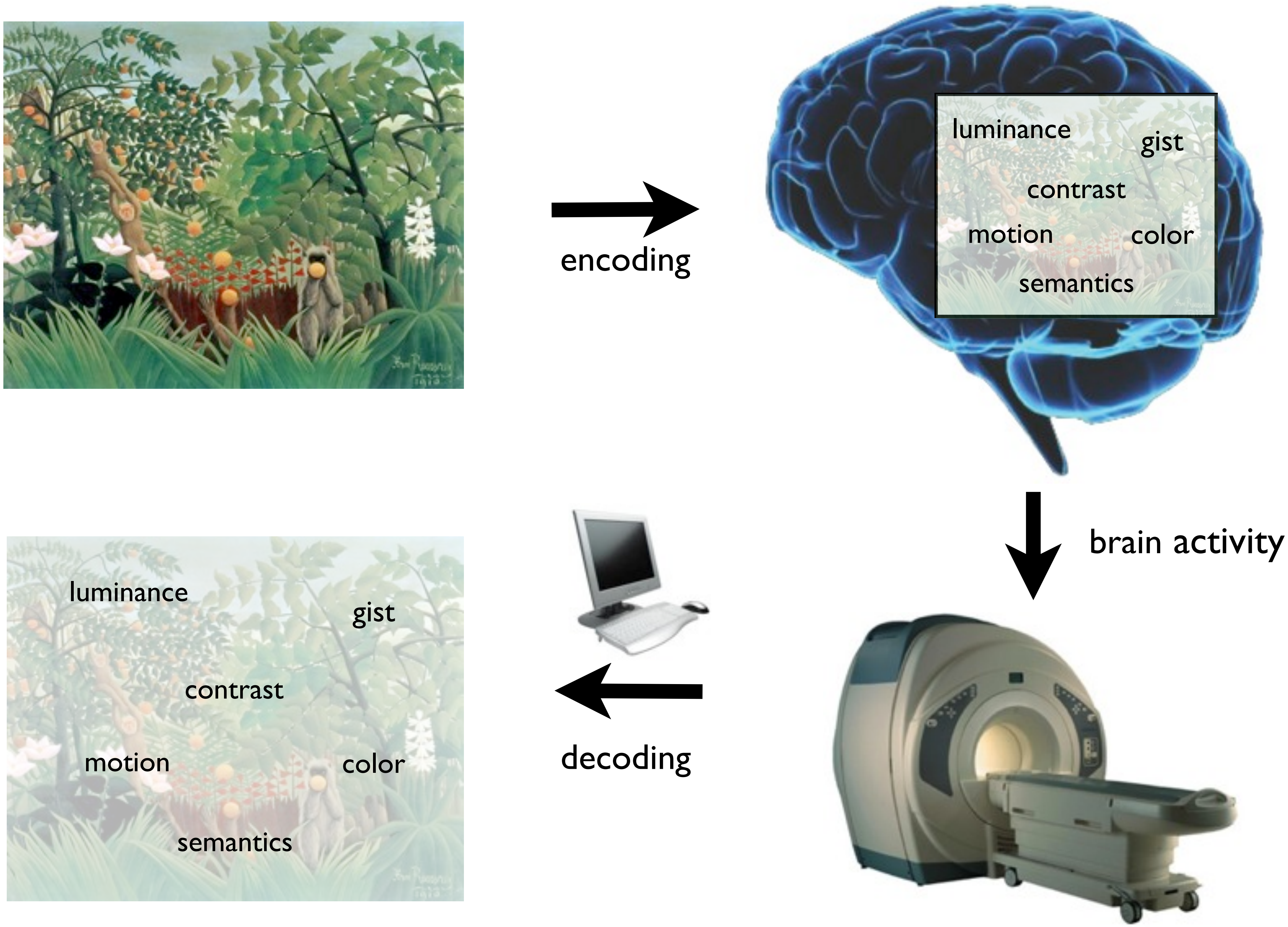
Percept decoding



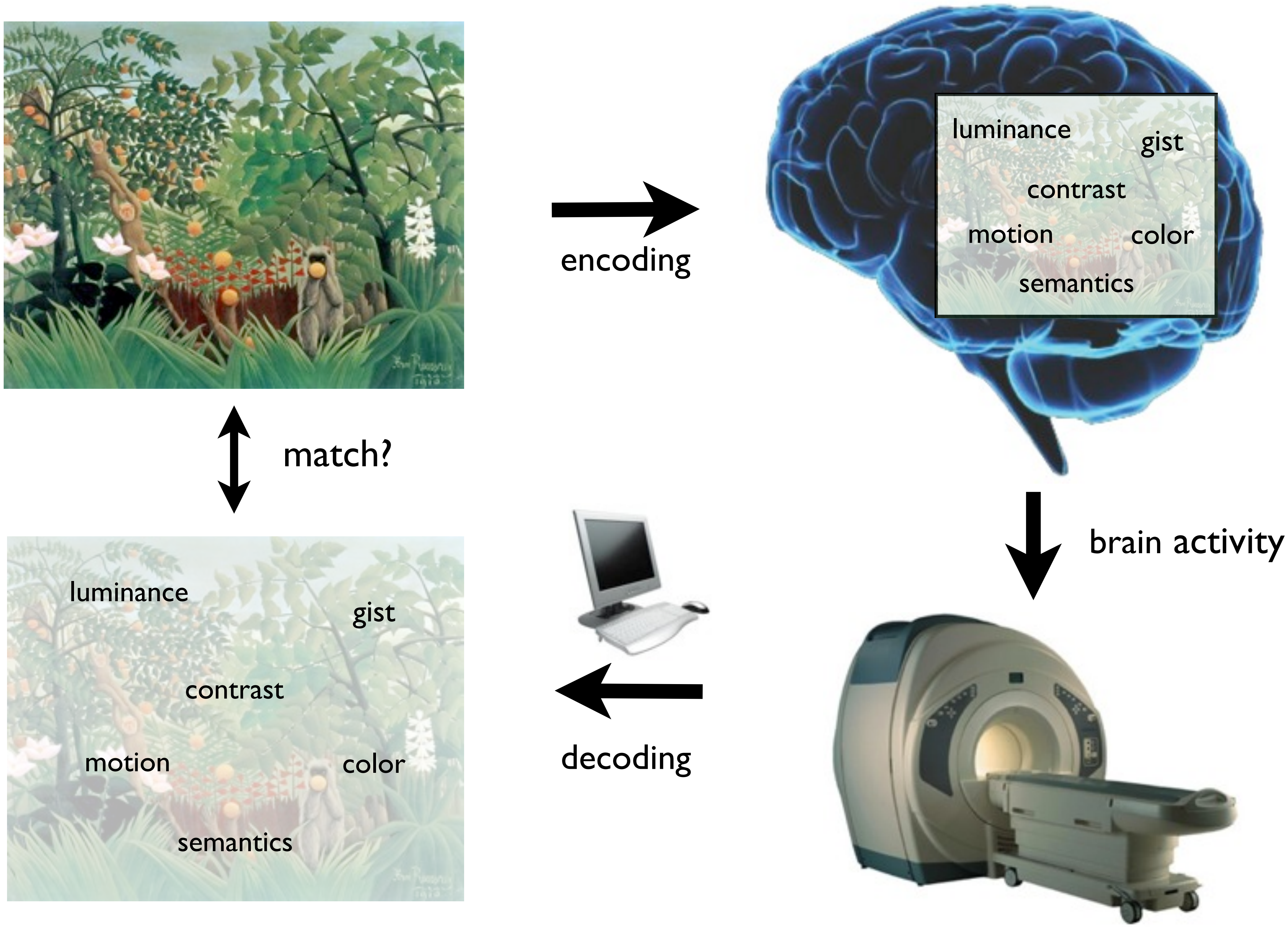
Percept decoding



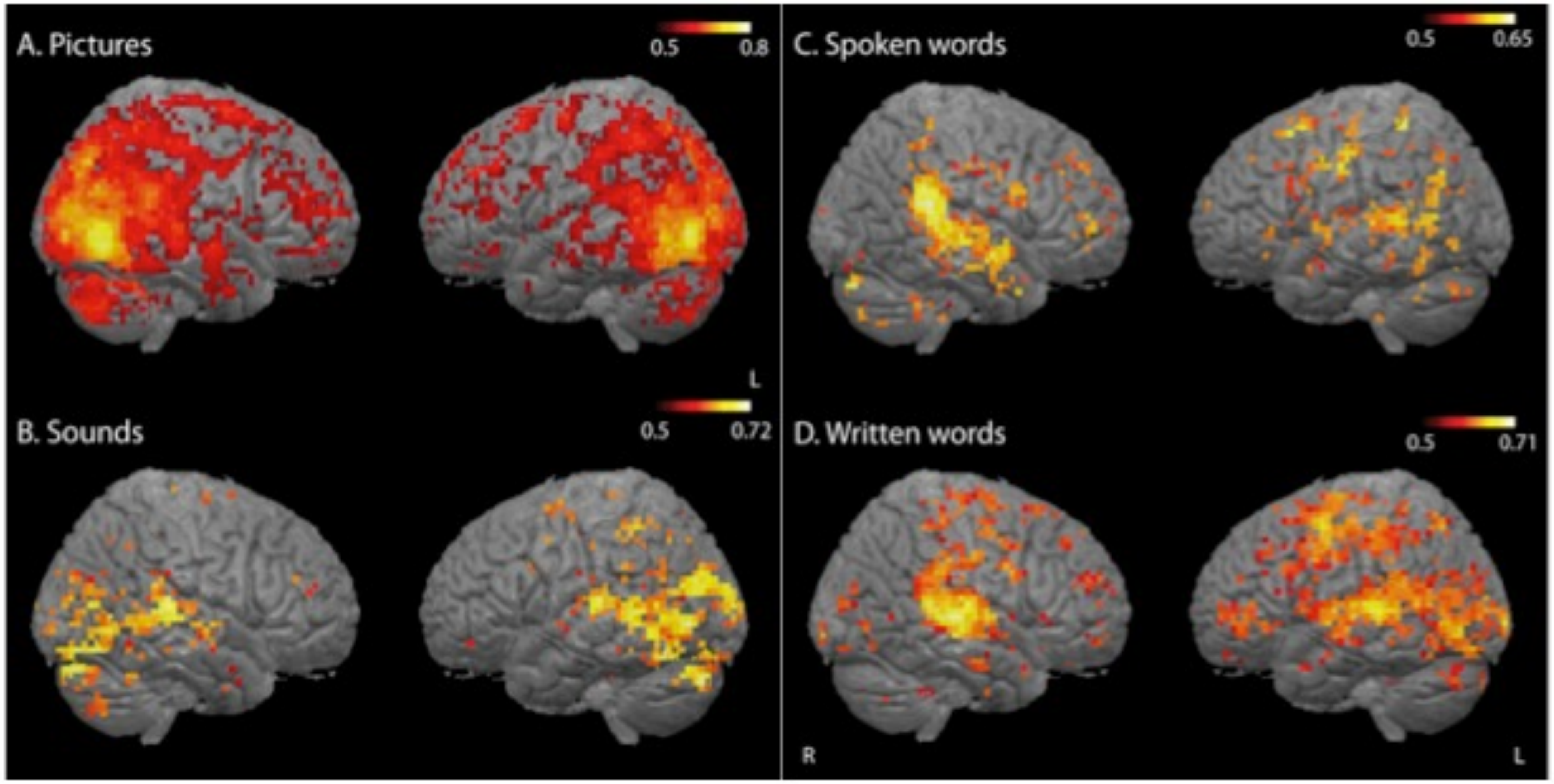
Percept decoding



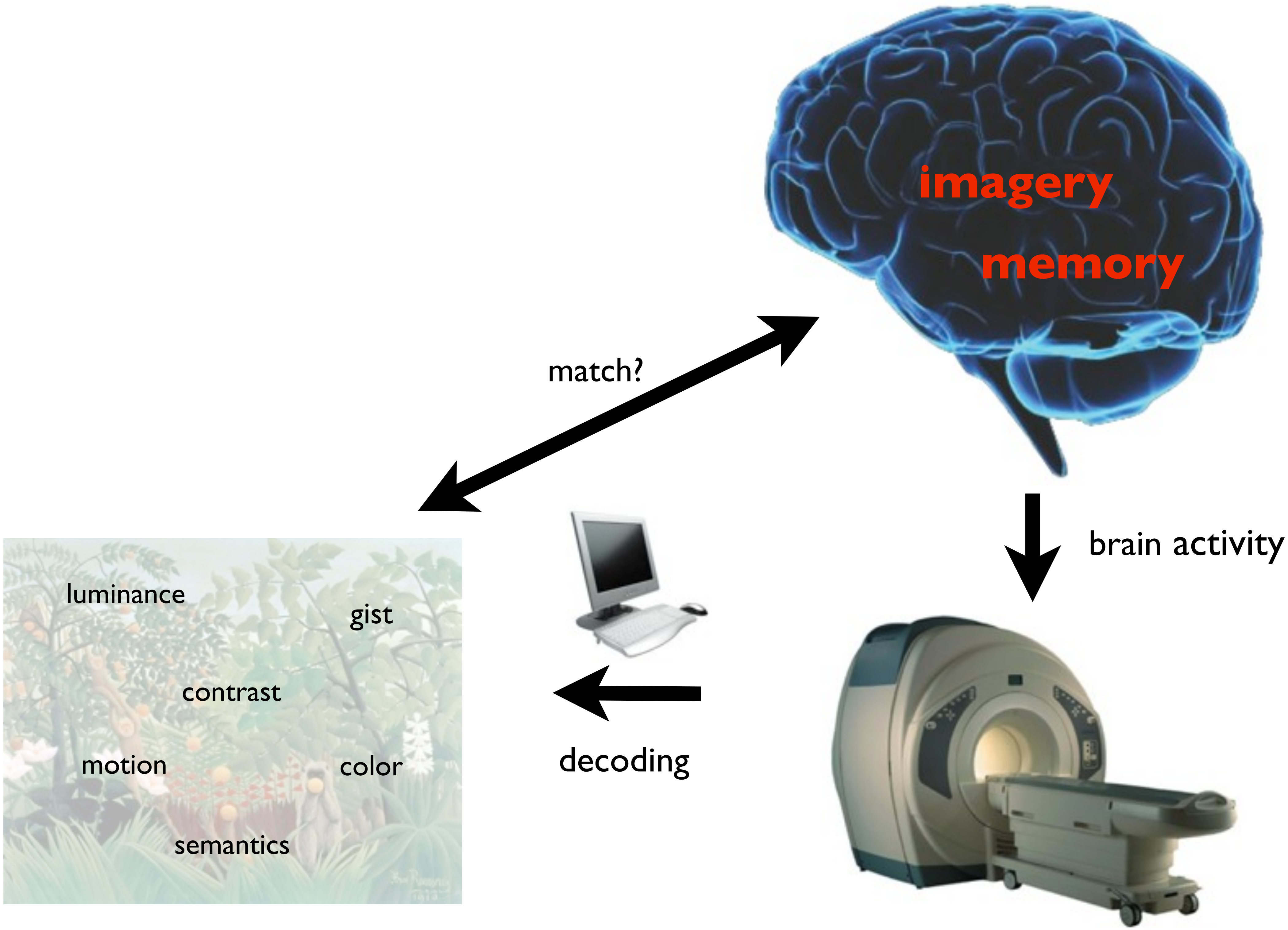
Percept decoding



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



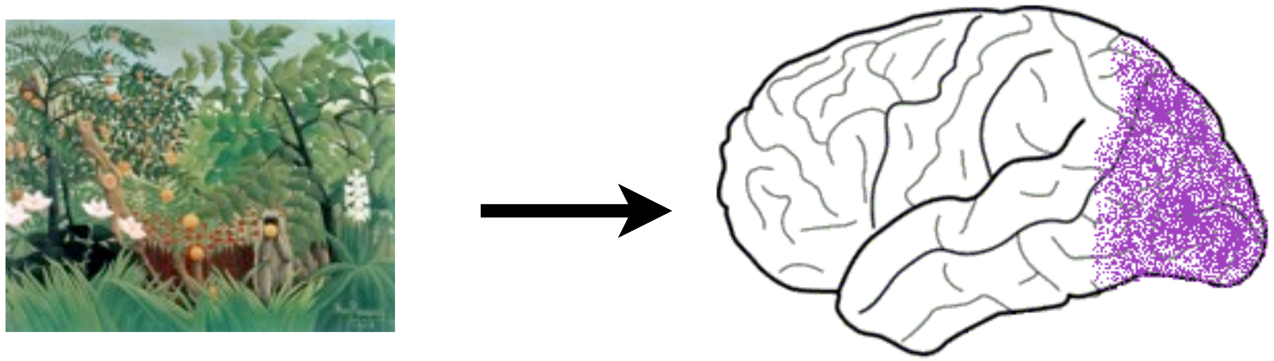
Decoding subjective experience



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

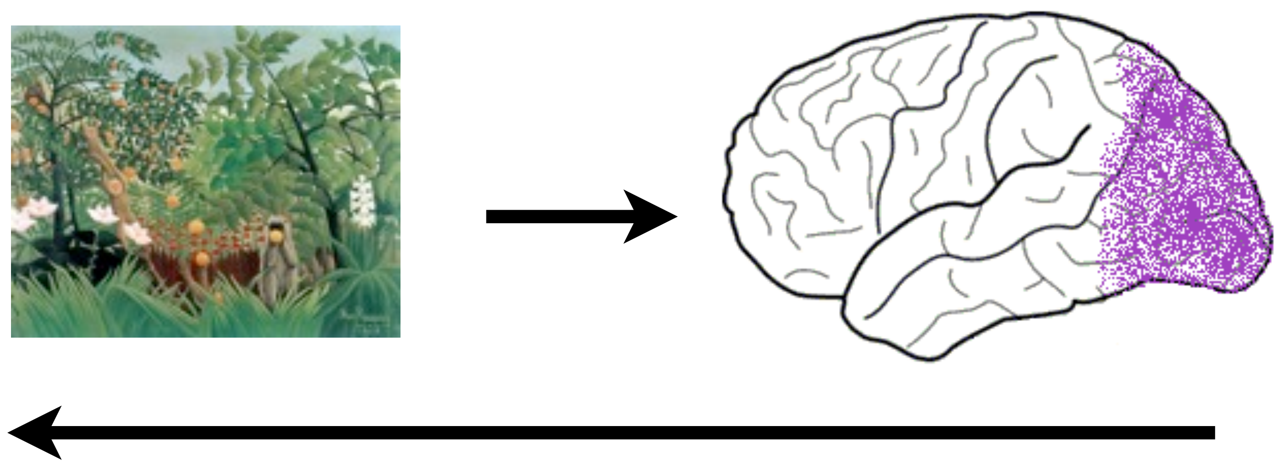
Predict from a very high-dimensional input (fMRI voxels) to a (possibly) very high-dimensional output (image pixels).

generative approach:



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

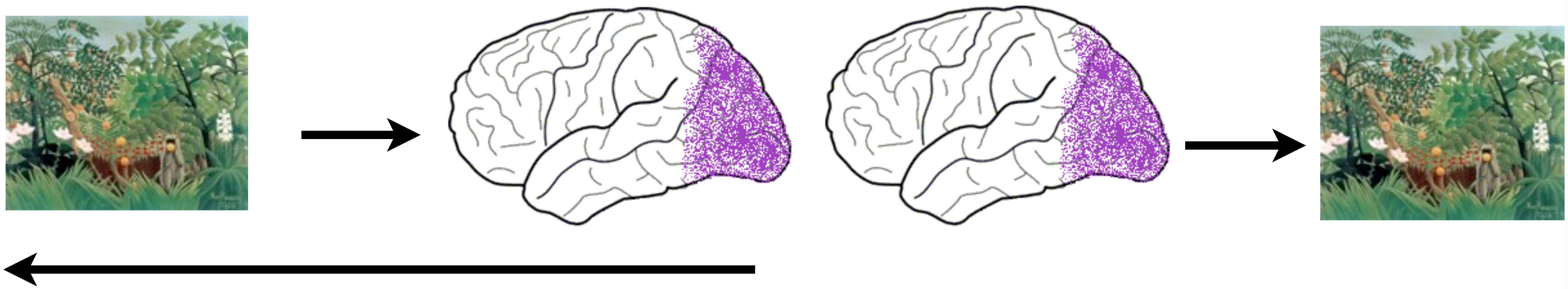
generative approach:



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

generative approach:

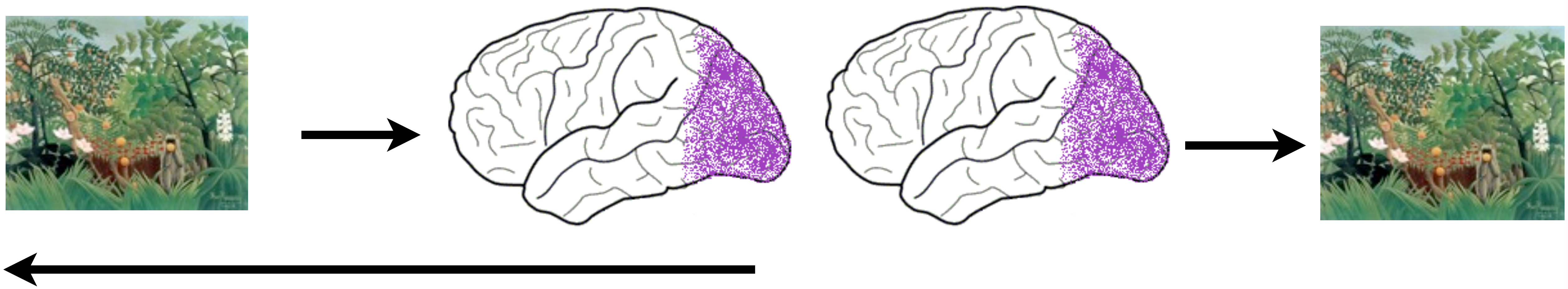
discriminative approach:



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

generative approach:

discriminative approach:

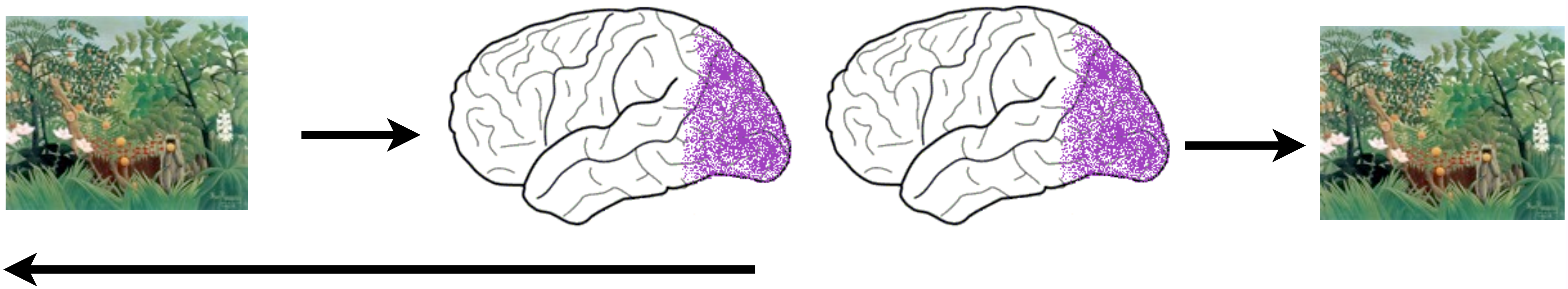


sparse latent variable models: interpretable, stable

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

generative approach:

discriminative approach:



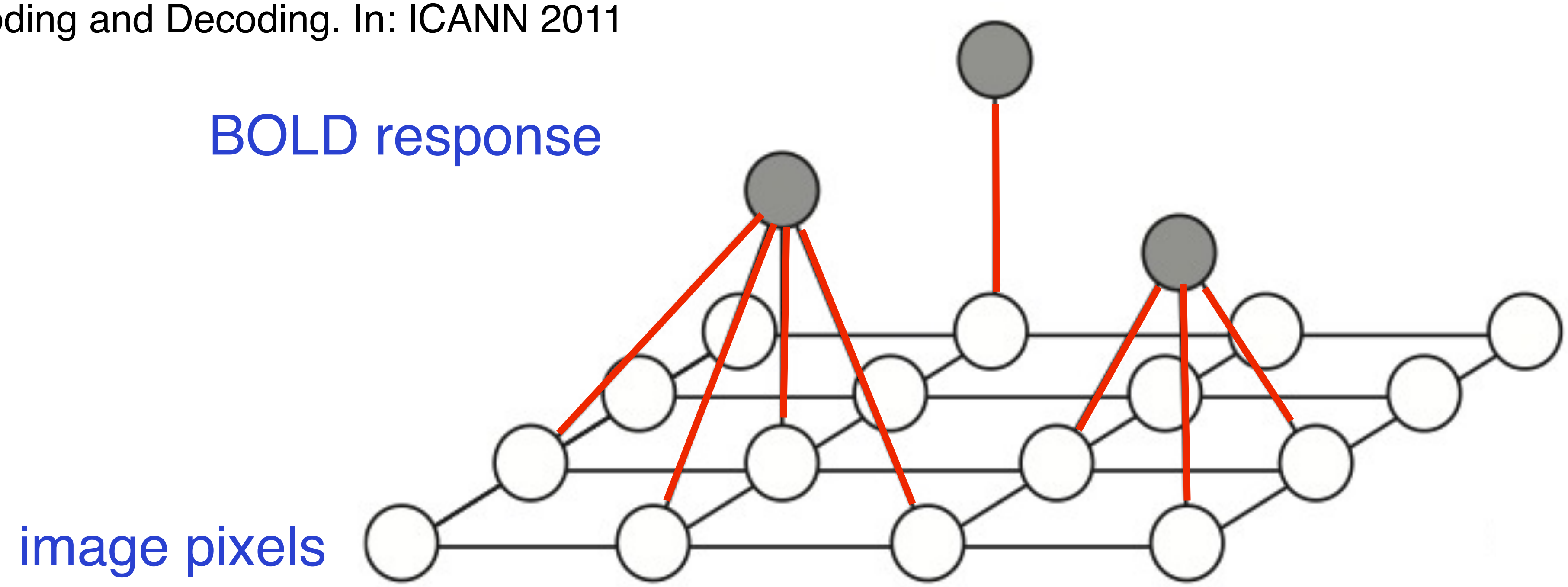
sparse latent variable models: interpretable, stable

Outlook:

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

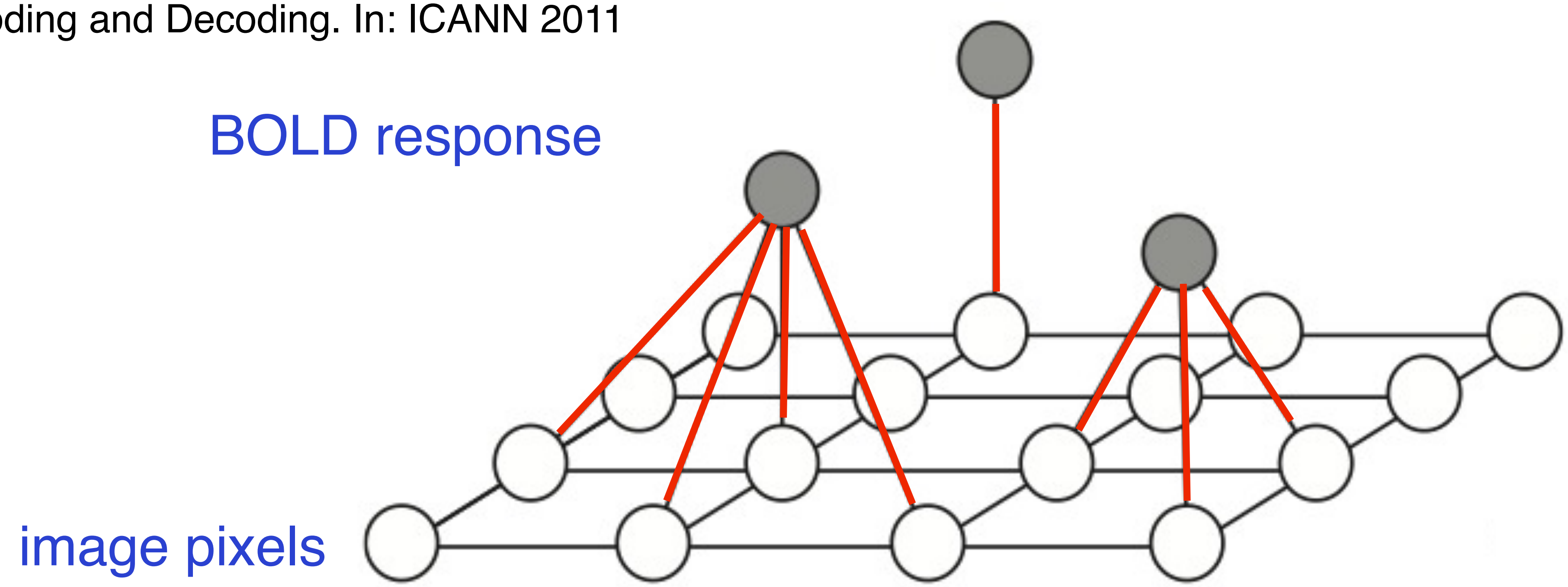
Generative approach: sparse encoding model

van Gerven et al. A Markov Random Field Approach to
Neural Encoding and Decoding. In: ICANN 2011



Generative approach: sparse encoding model

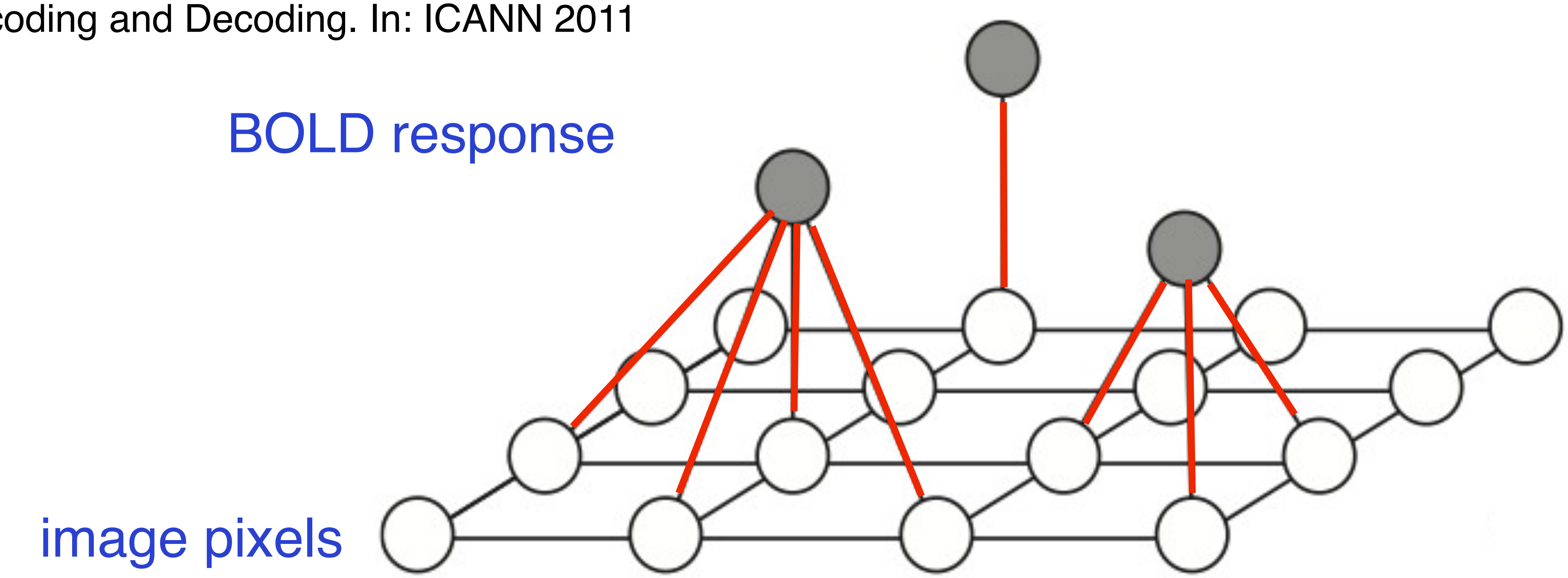
van Gerven et al. A Markov Random Field Approach to
Neural Encoding and Decoding. In: ICANN 2011



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

Generative approach: sparse encoding model

van Gerven et al. A Markov Random Field Approach to Neural Encoding and Decoding. In: ICANN 2011



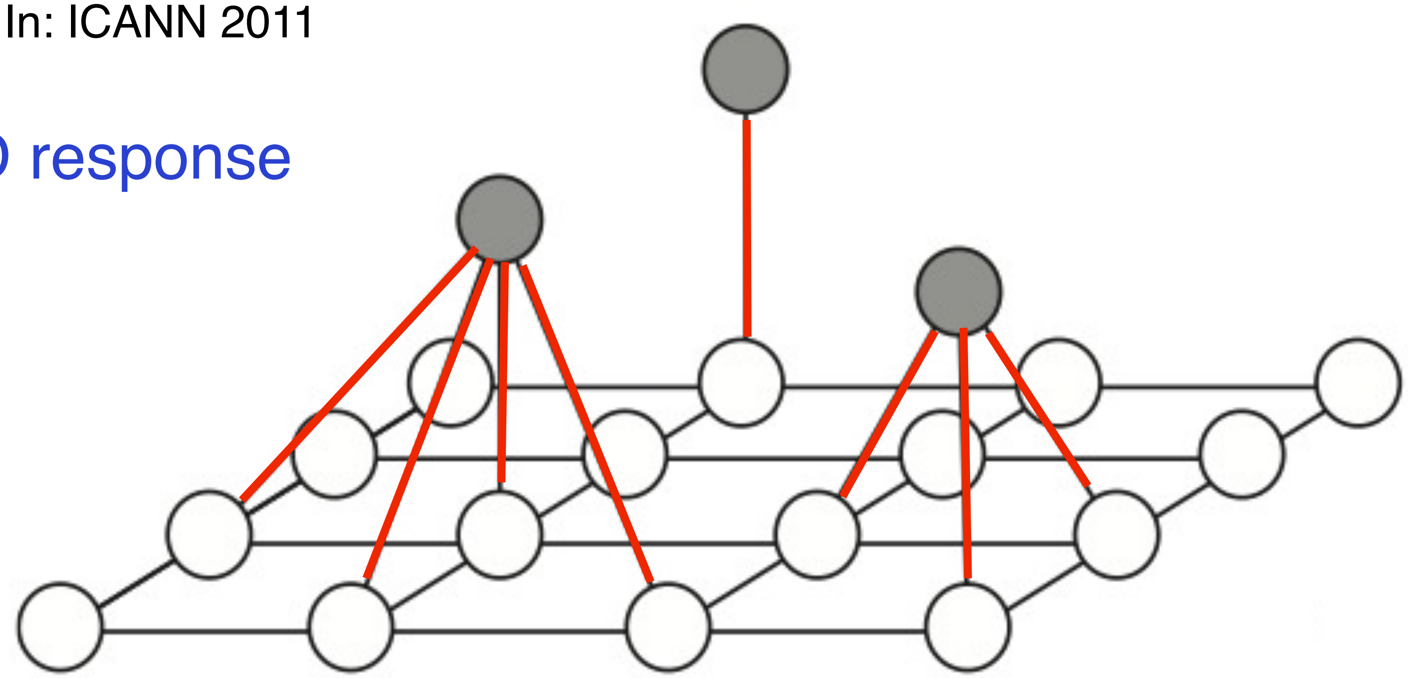
For each voxel *k*: $p(r \mid s) = \mathcal{N}(r; \alpha_k + \beta_k^\top s, \sigma_k)$

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

van Gerven et al. A Markov Random Field Approach to Neural Encoding and Decoding. In: ICANN 2011



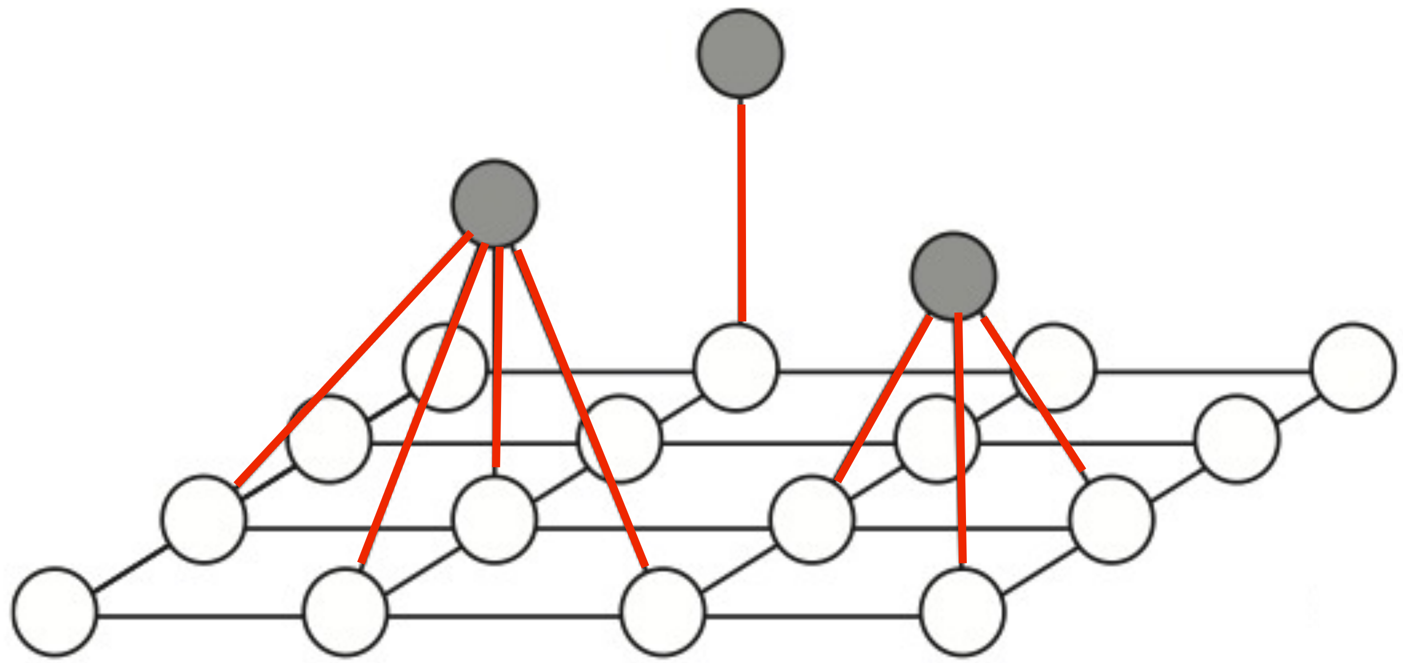
For each voxel *k*: $p(r \mid s) = \mathcal{N}(r; \alpha_k + \beta_k^\top s, \sigma_k)$

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:

$$\hat{\theta}_k = \arg \min_{\alpha_k, \beta_k, \sigma_k^2} \left\{ -\log p(\alpha_k, \beta_k, \sigma_k^2) - \sum_n \log \mathcal{N}\left(r_k^n; \alpha_k + \beta_k^\top s^n, \sigma_k^2\right) \right\}$$

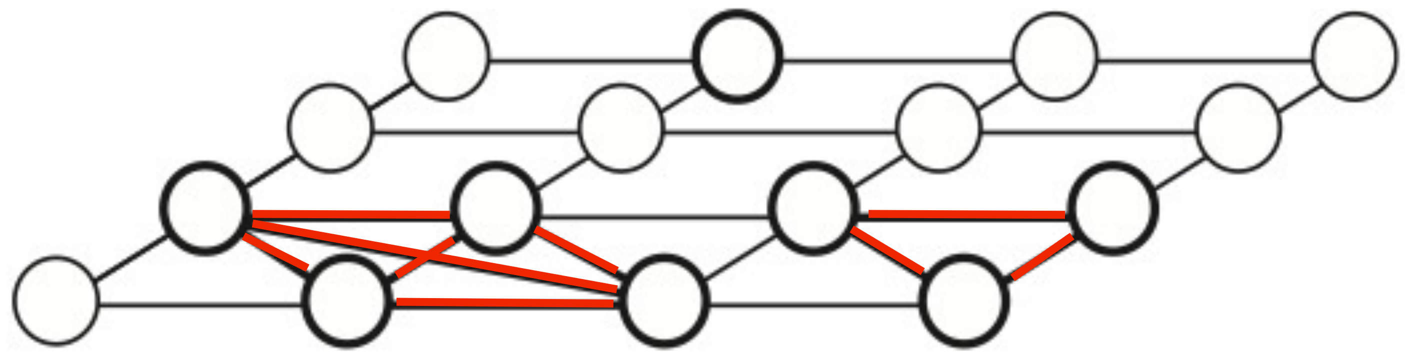


It can be shown that $p(r|s) = \frac{1}{Z} \prod_i \psi_i(s_i) \prod_{i\sim j} \psi_{i,j}(s_i, s_j)$

where

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

$$\psi_{i,j}(s_i, s_j) = \exp\left(-s_i s_j \sum_k \frac{\beta_{ki}}{\sigma_k^2}\beta_{kj}\right)$$

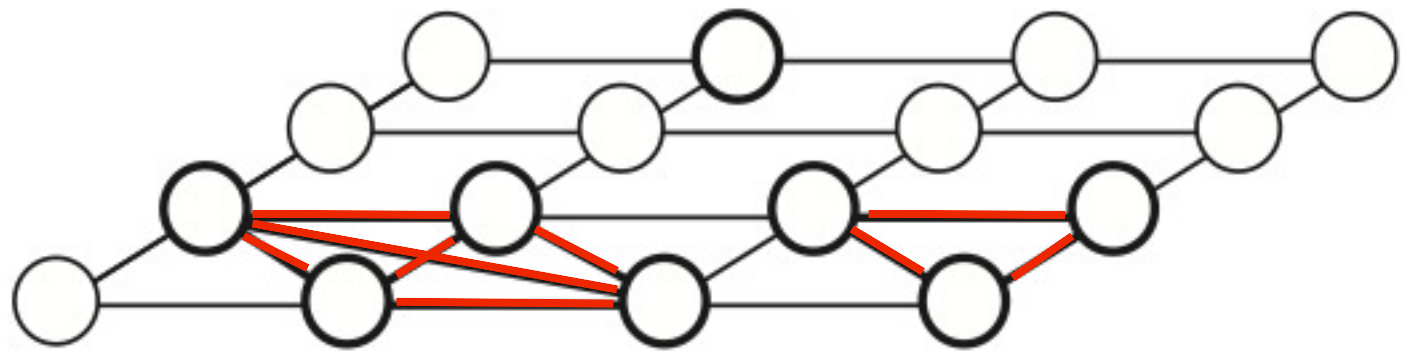


It can be shown that $p(r|s) = \frac{1}{Z} \prod_i \psi_i(s_i) \prod_{i \sim j} \psi_{i,j}(s_i, s_j)$

where

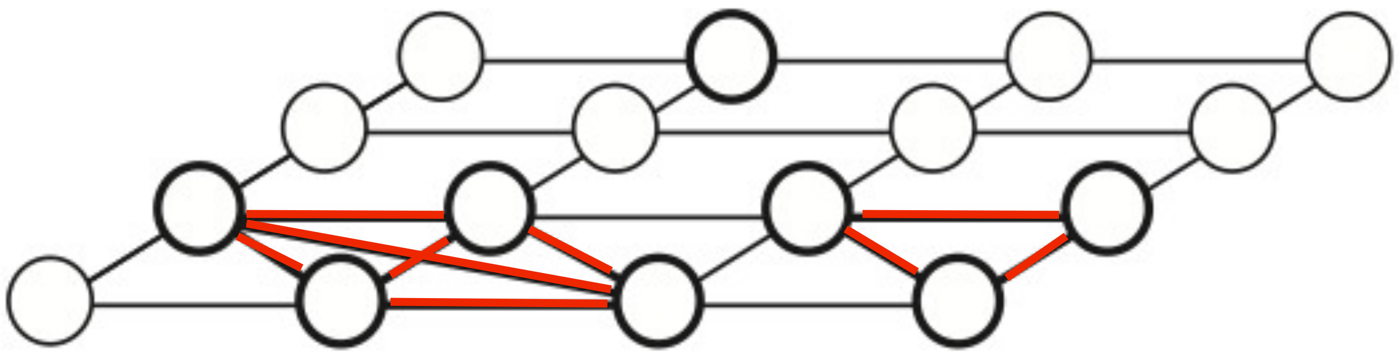
$$\psi_i(s_i) = \exp\left(s_i \sum_k \frac{\beta_{ki}}{\sigma_k^2}(r_k - \alpha_k - \frac{1}{2}\beta_{ki})\right)$$

$$\psi_{i,j}(s_i, s_j) = \exp\left(-s_i s_j \sum_k \frac{\beta_{ki}}{\sigma_k^2}\beta_{kj}\right)$$



Define appropriate MRF prior:

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



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:

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

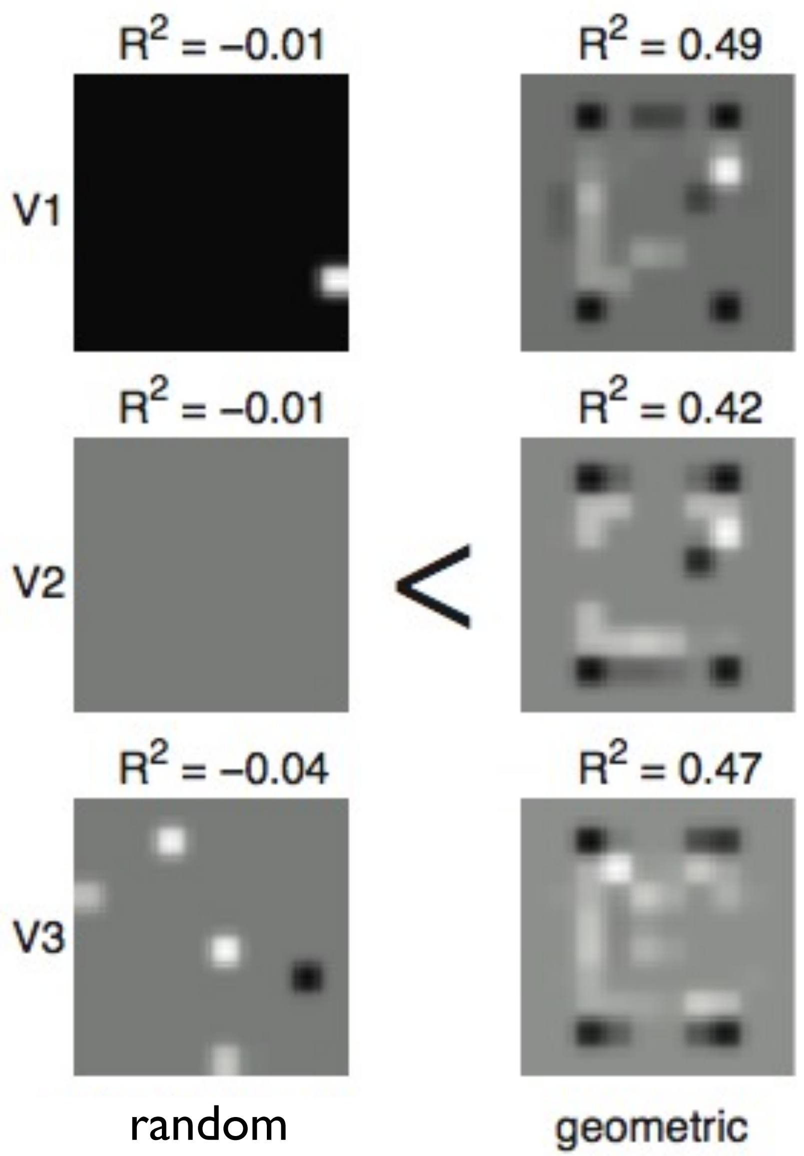
Results

- 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

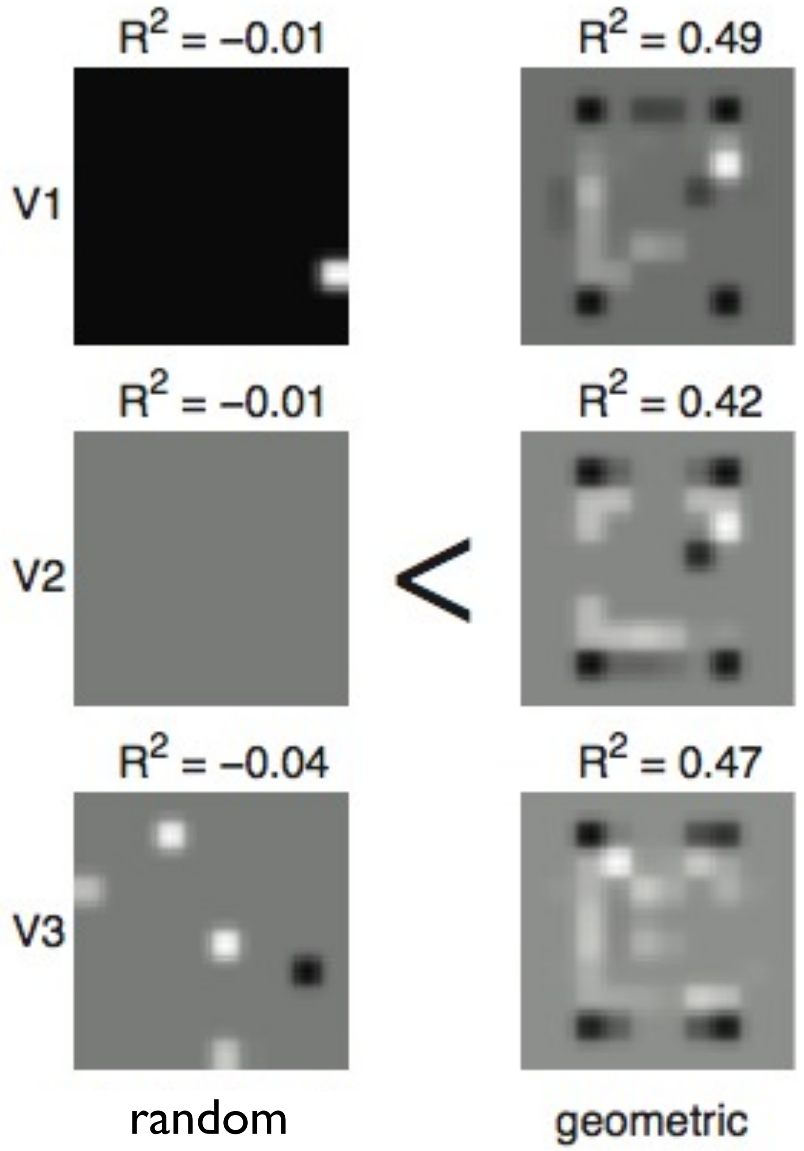
encoding



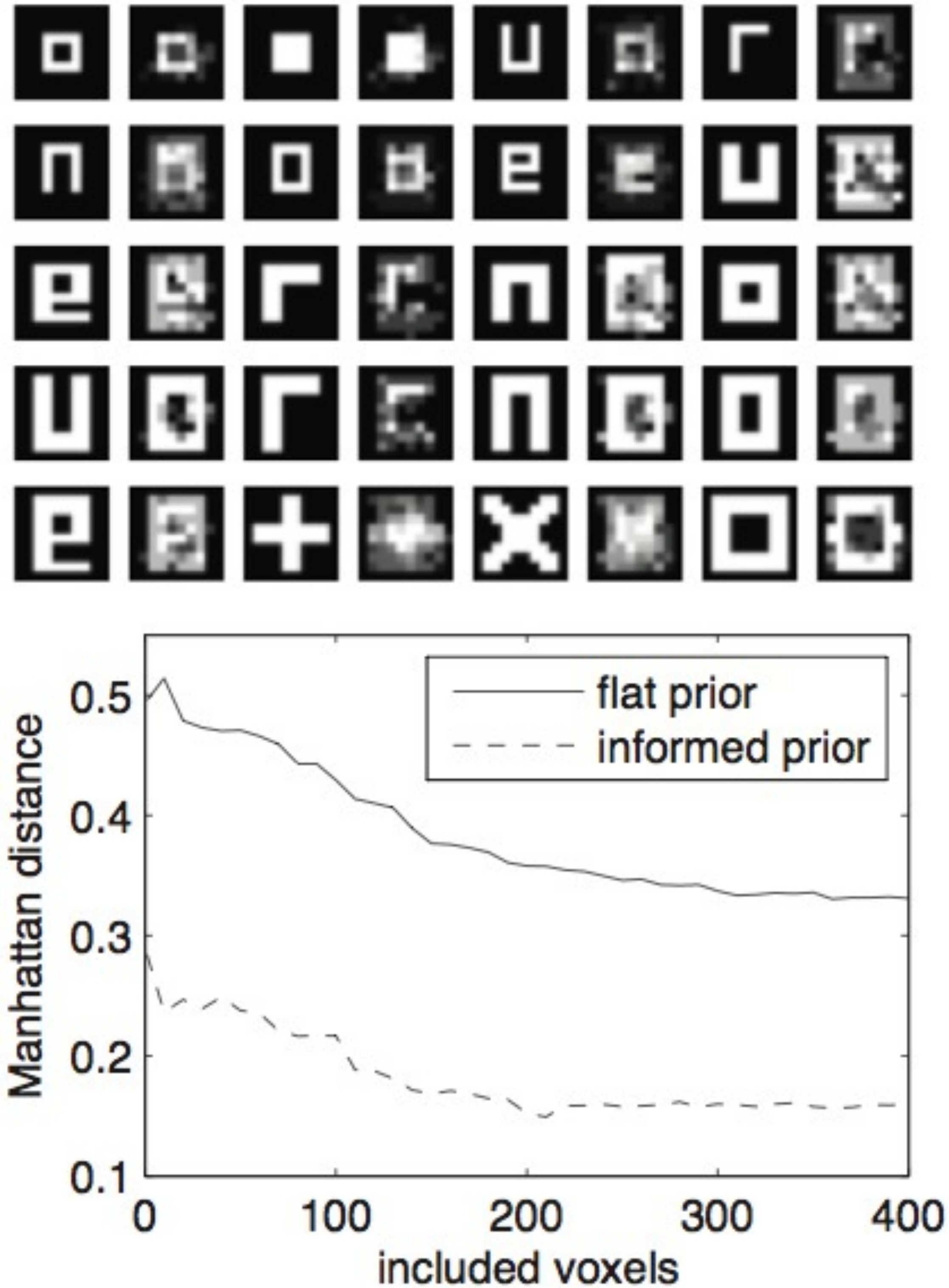
Results

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

encoding

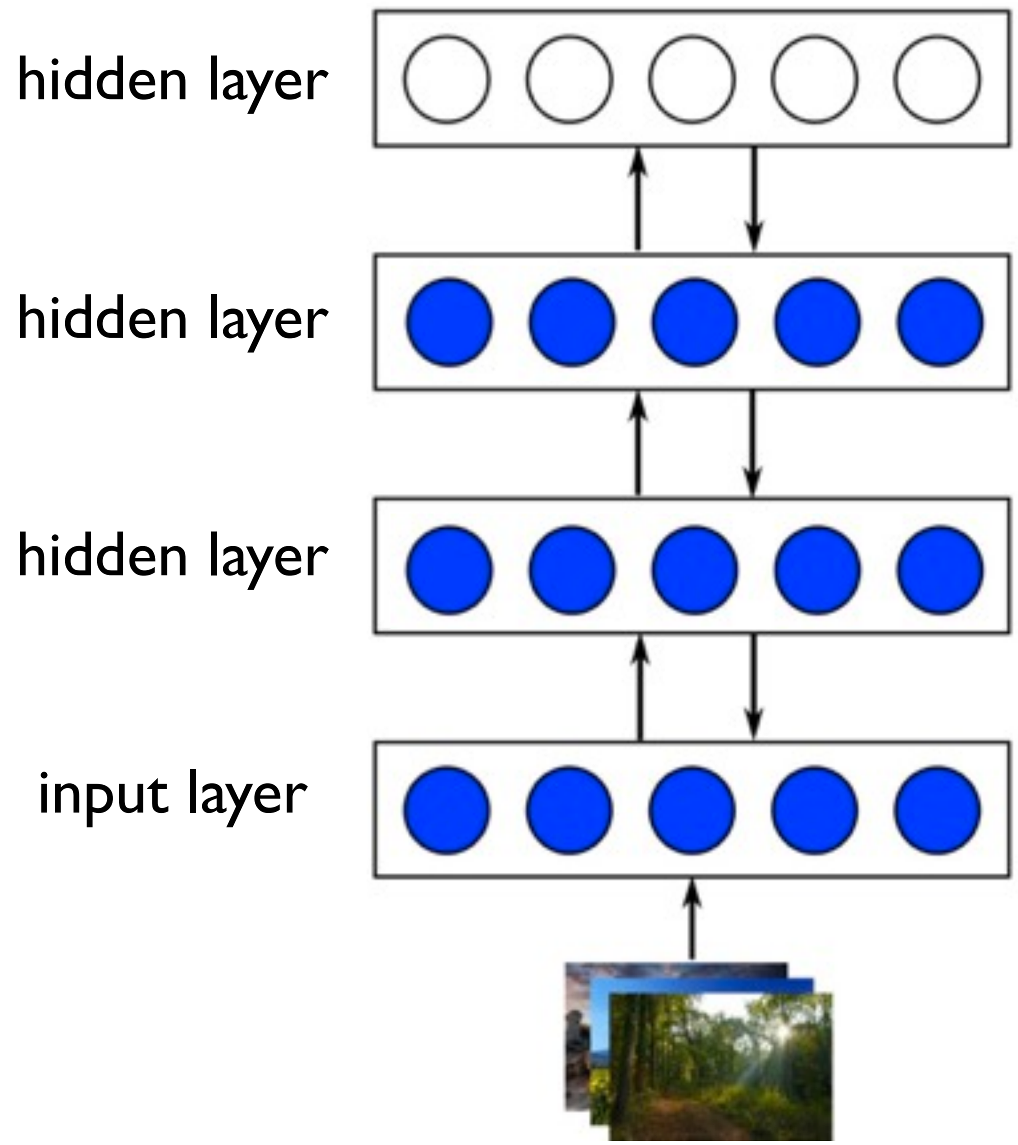


decoding



Generative approach: latent variables

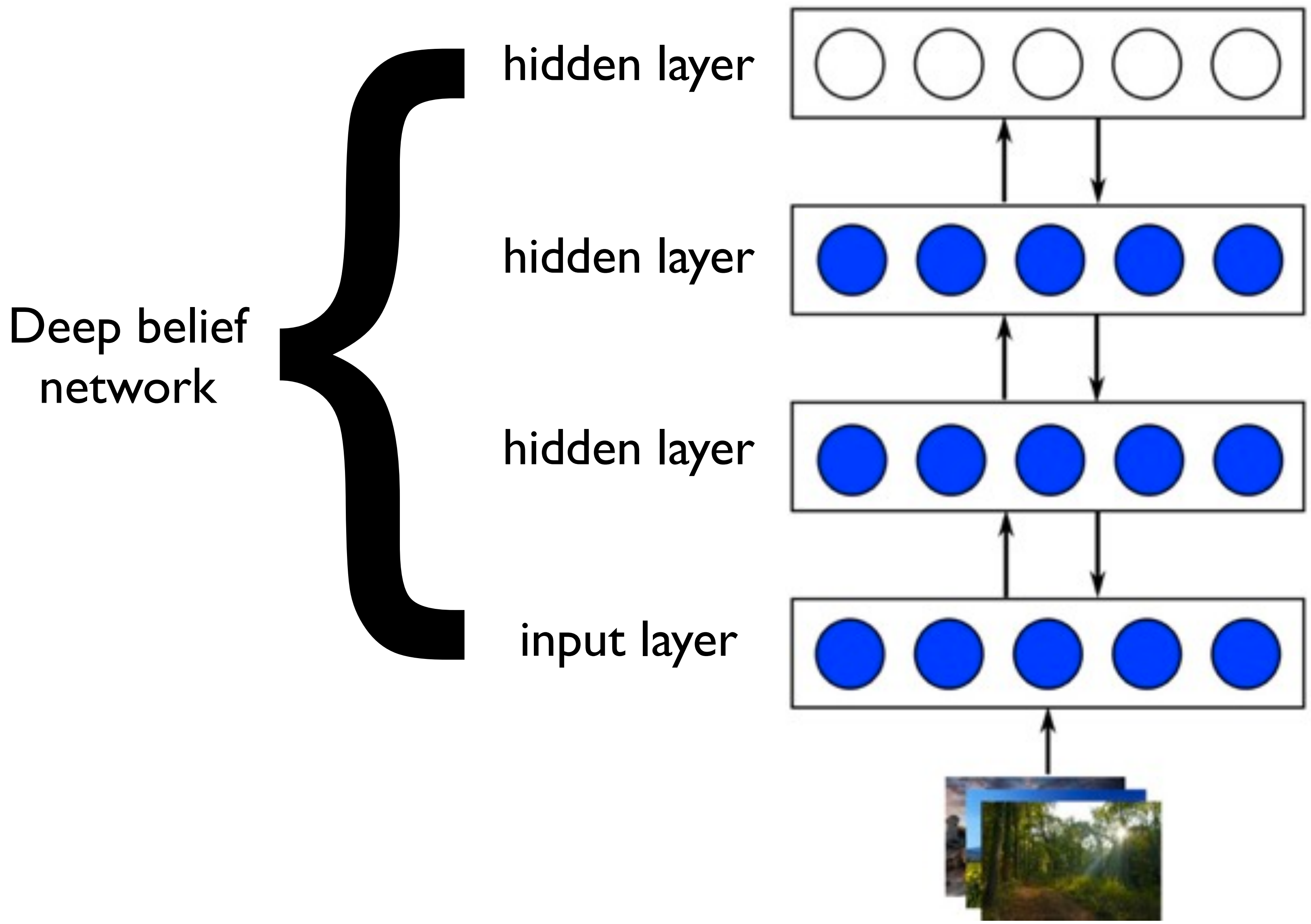
Unsupervised phase



van Gerven et al. Neural decoding with hierarchical generative models. *Neural Computation*, 2010.

Generative approach: latent variables

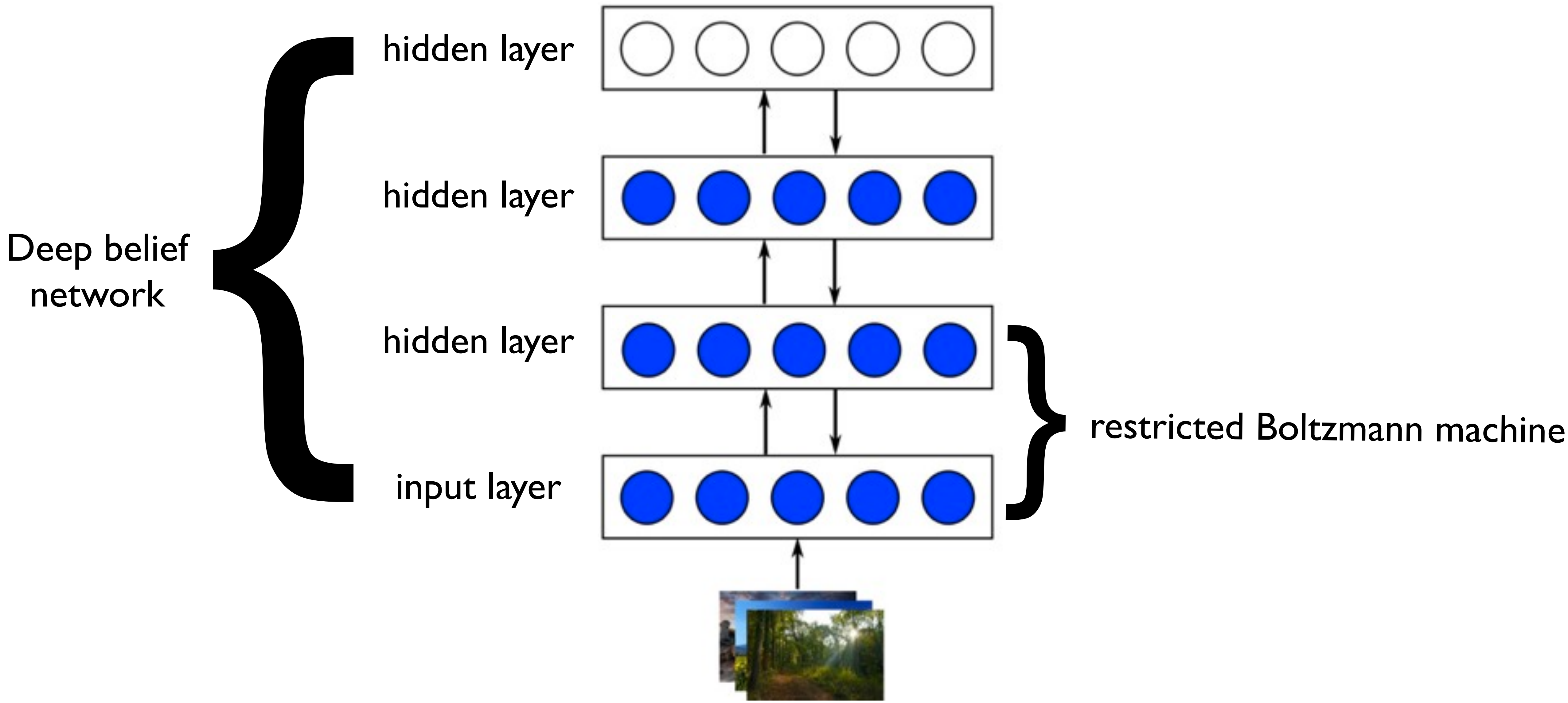
Unsupervised phase



van Gerven et al. Neural decoding with hierarchical generative models. *Neural Computation*, 2010.

Generative approach: latent variables

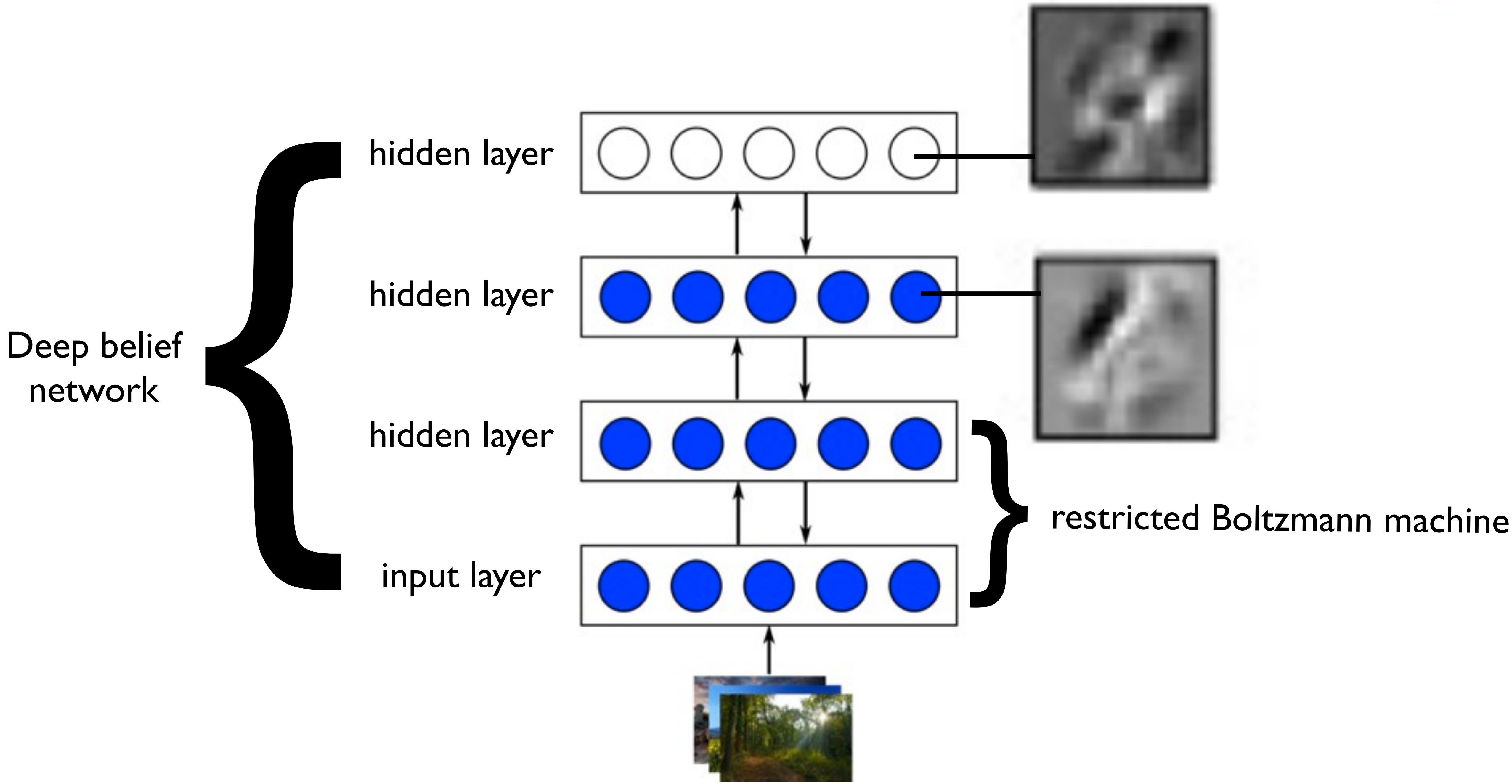
Unsupervised phase



van Gerven et al. Neural decoding with hierarchical generative models. *Neural Computation*, 2010.

Generative approach: latent variables

Unsupervised phase

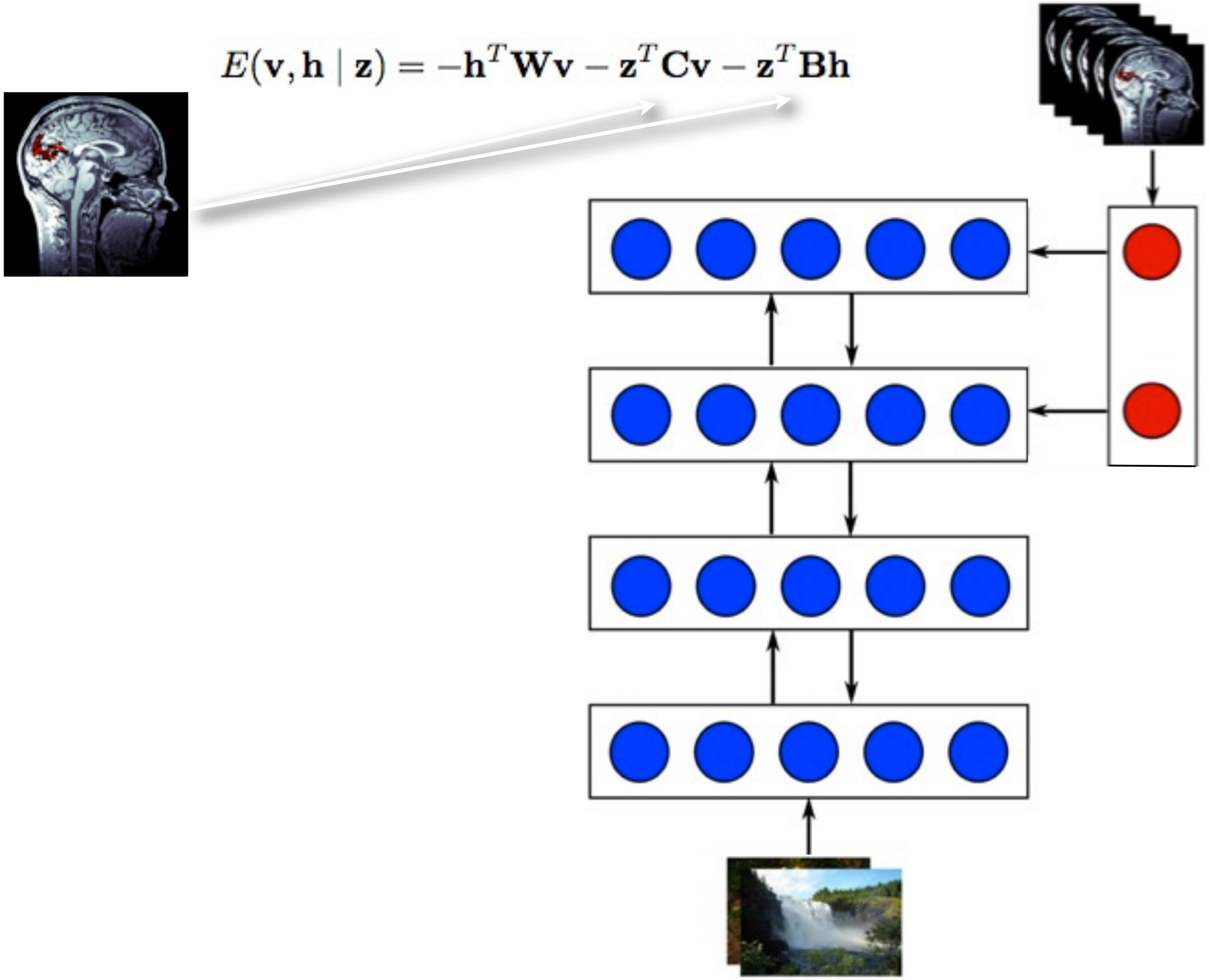


van Gerven et al. Neural decoding with hierarchical generative models. *Neural Computation*, 2010.

Induce a coupling with conditional restricted Boltzmann machines

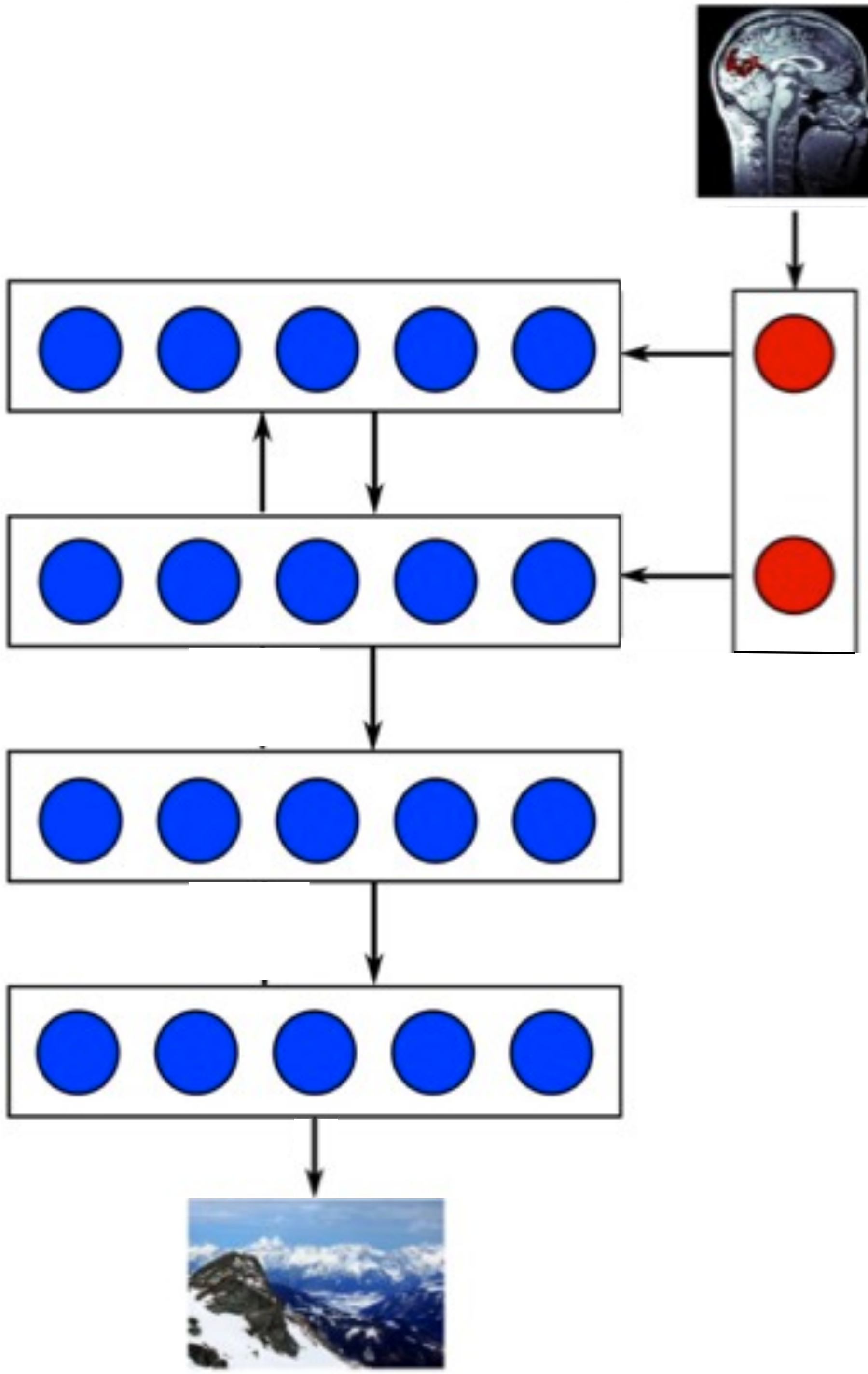
Supervised phase

$$E(\mathbf{v},\mathbf{h} \mid \mathbf{z}) = -\mathbf{h}^T\mathbf{W}\mathbf{v} - \mathbf{z}^T\mathbf{C}\mathbf{v} - \mathbf{z}^T\mathbf{B}\mathbf{h}$$

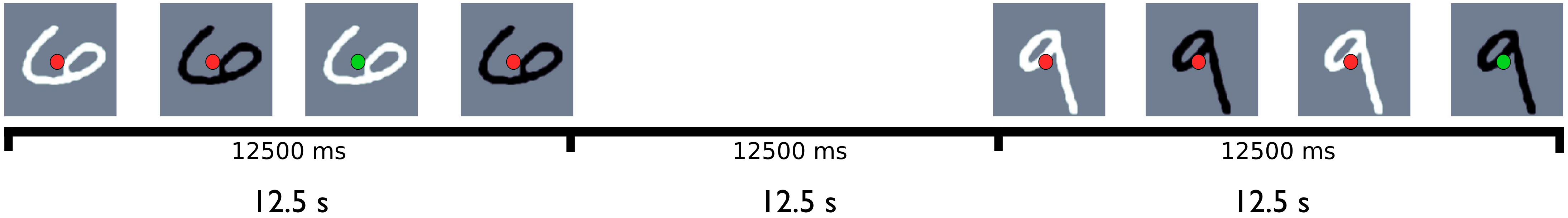


Using a hierarchical generative model

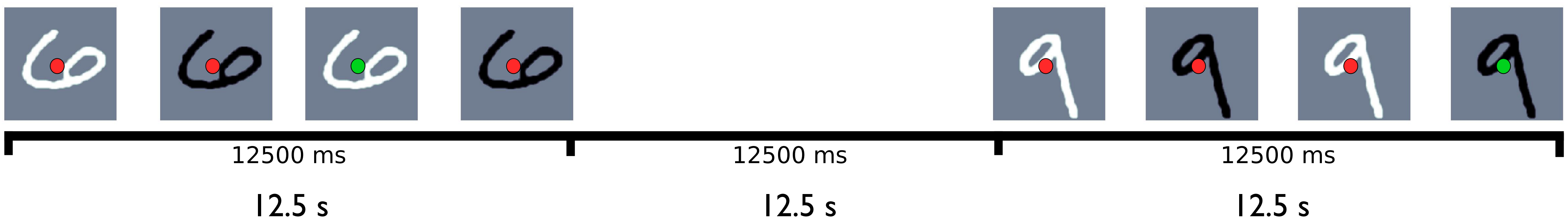
Reconstruction phase

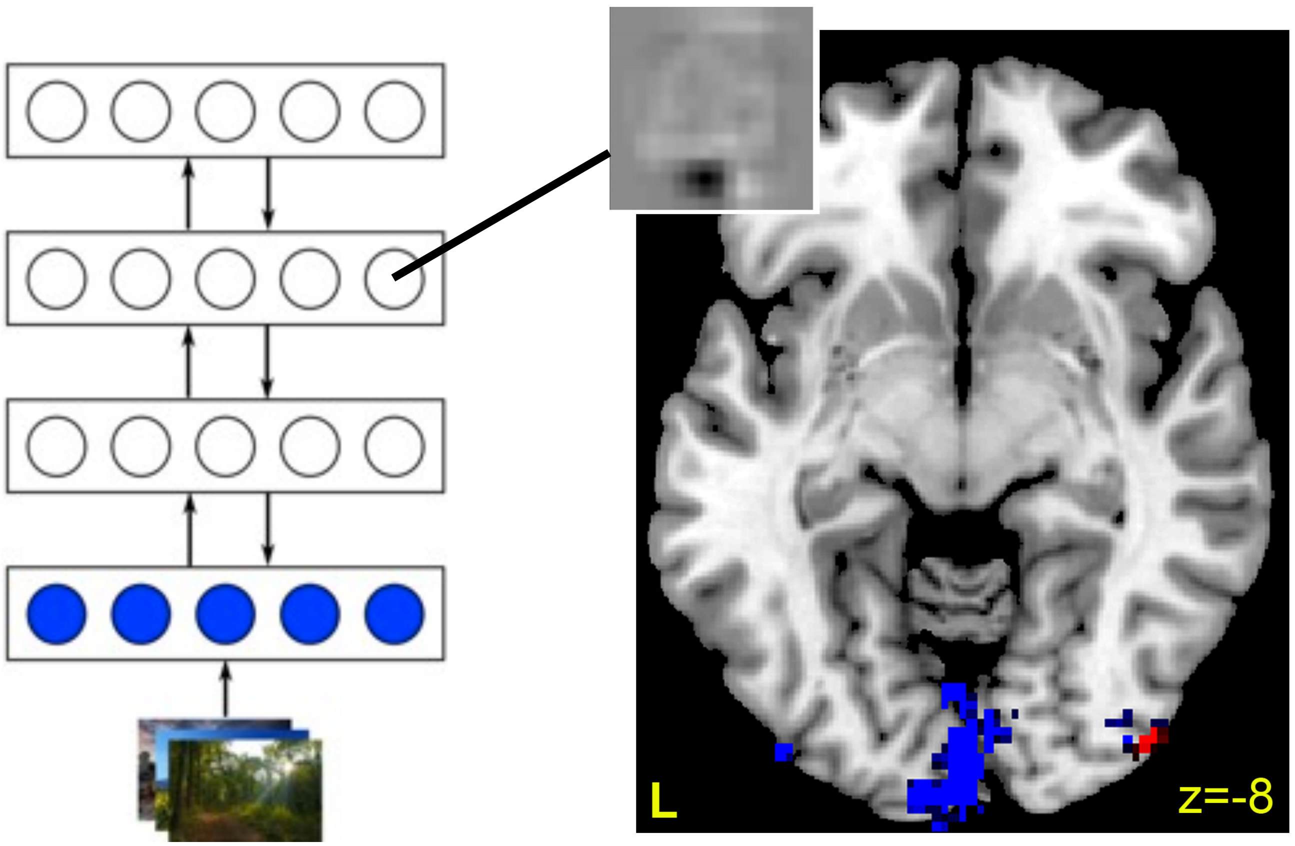


Experimental results

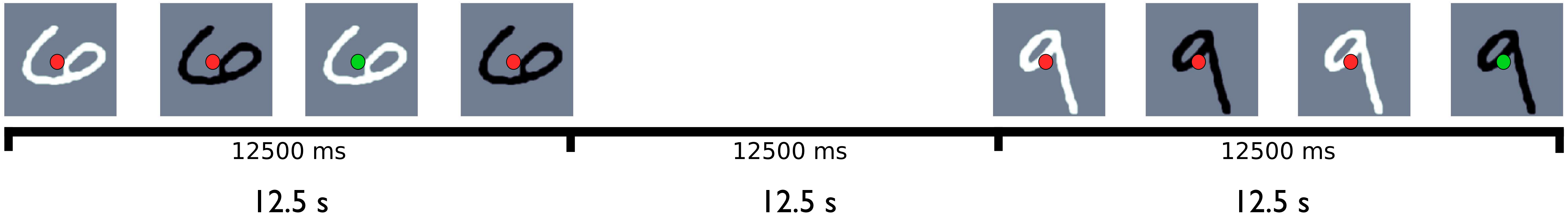


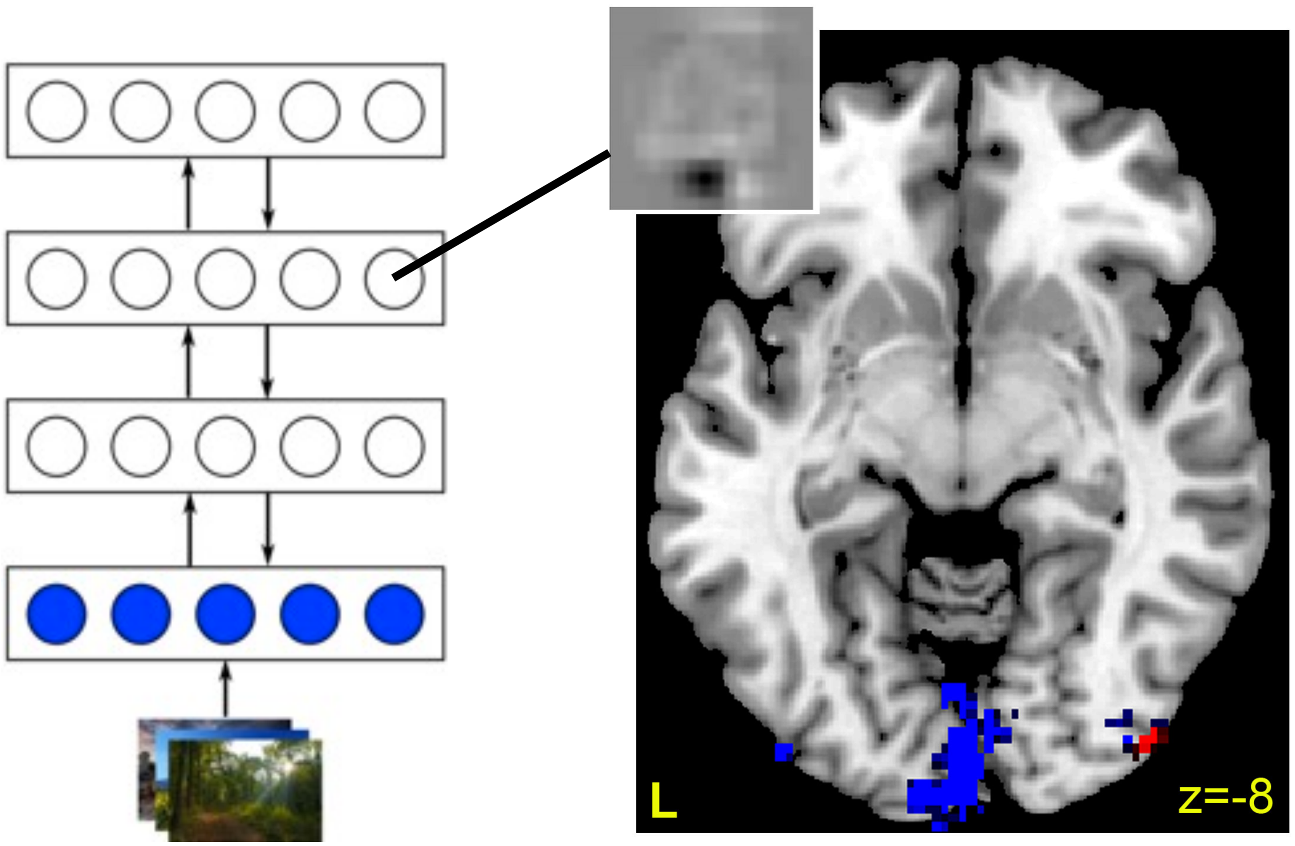
Experimental results



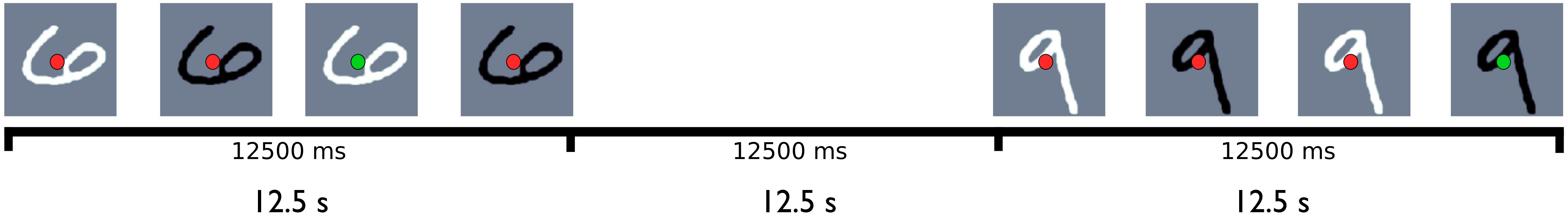


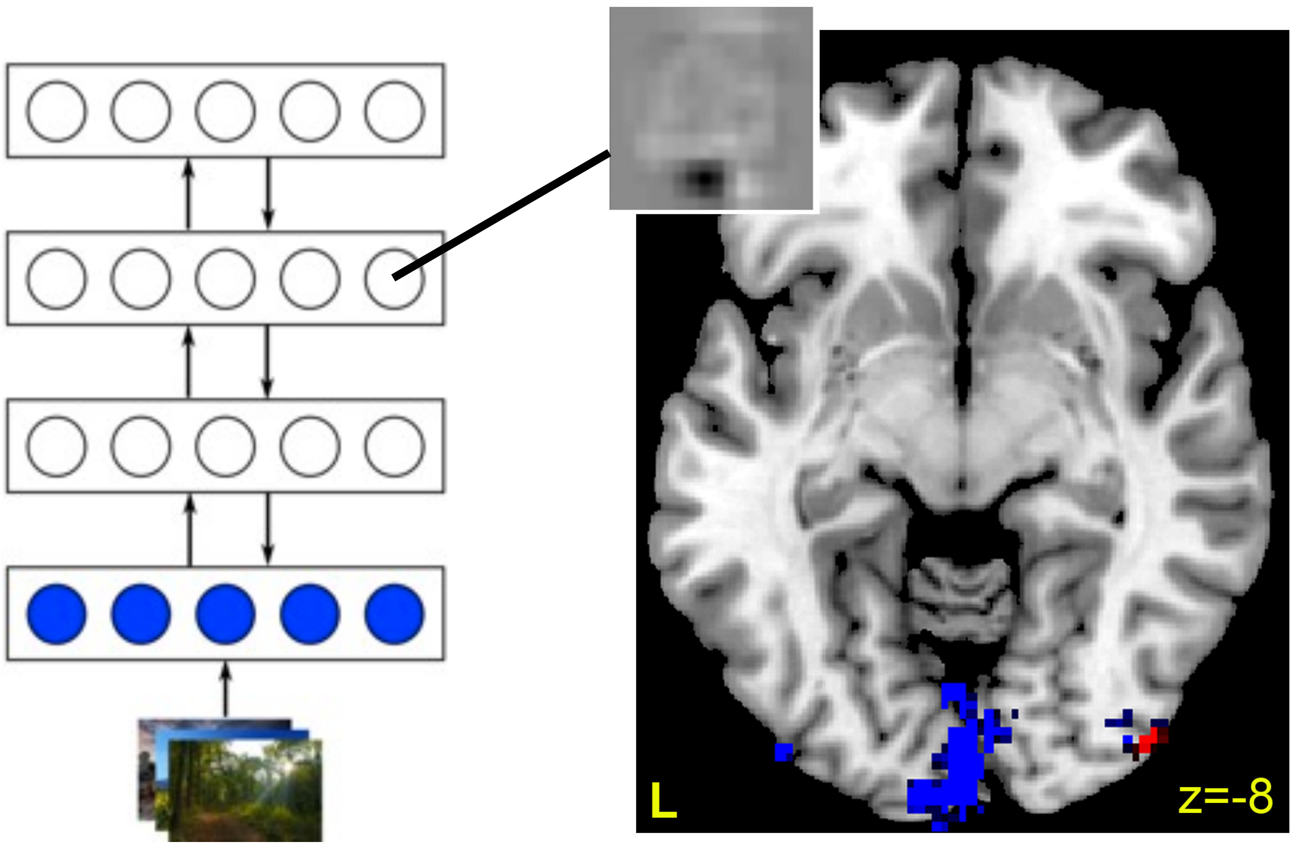
Experimental results

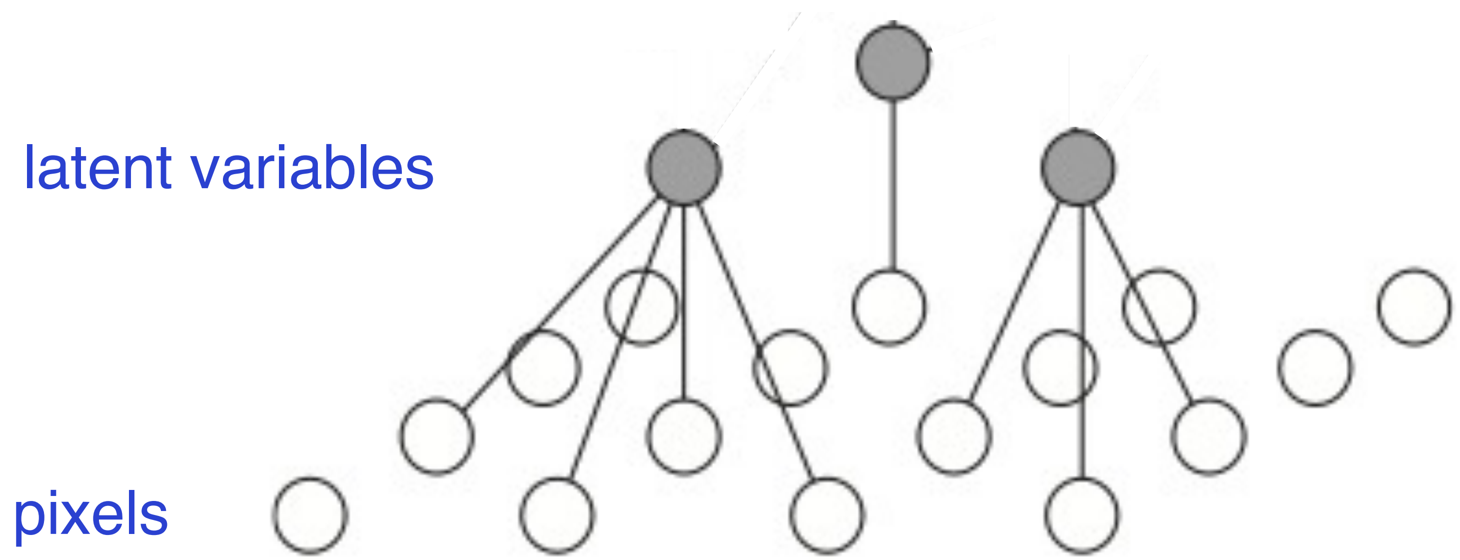




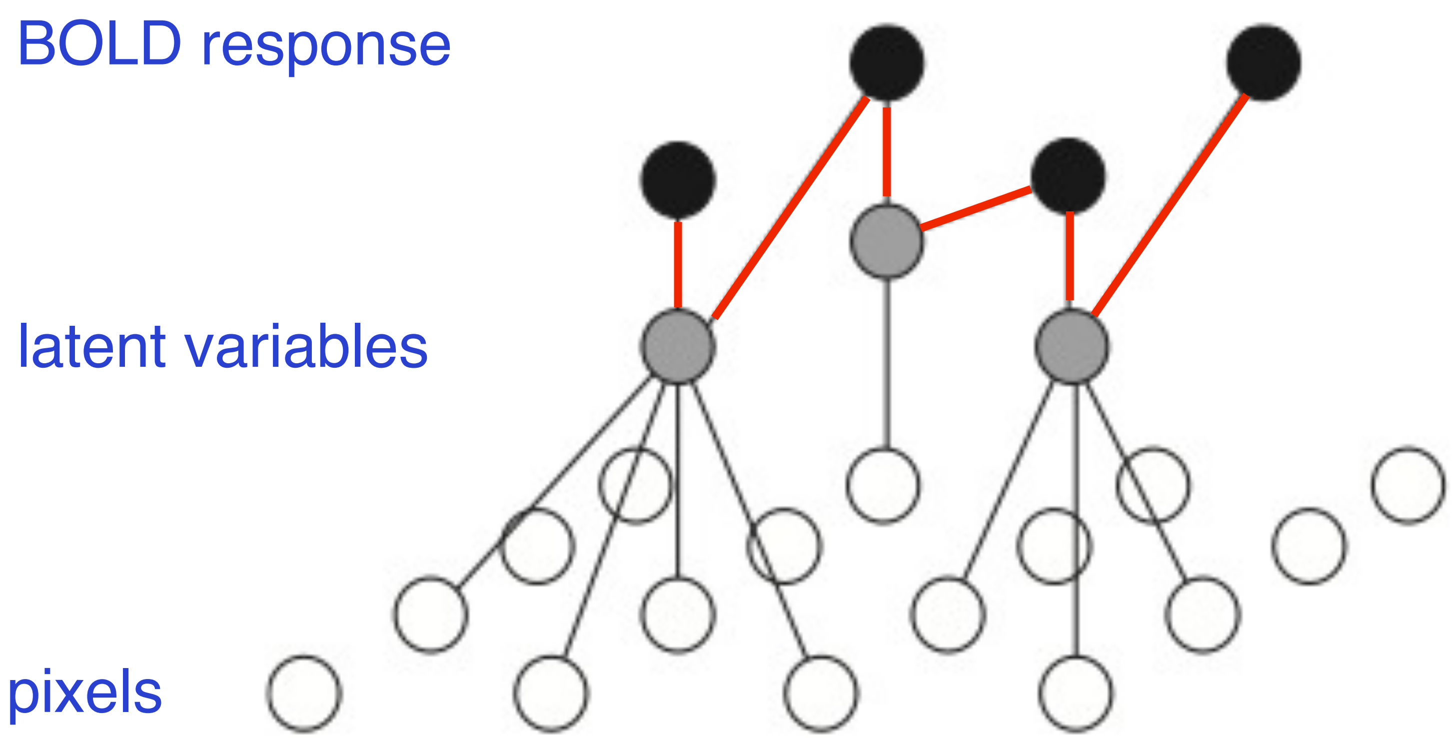
Experimental results



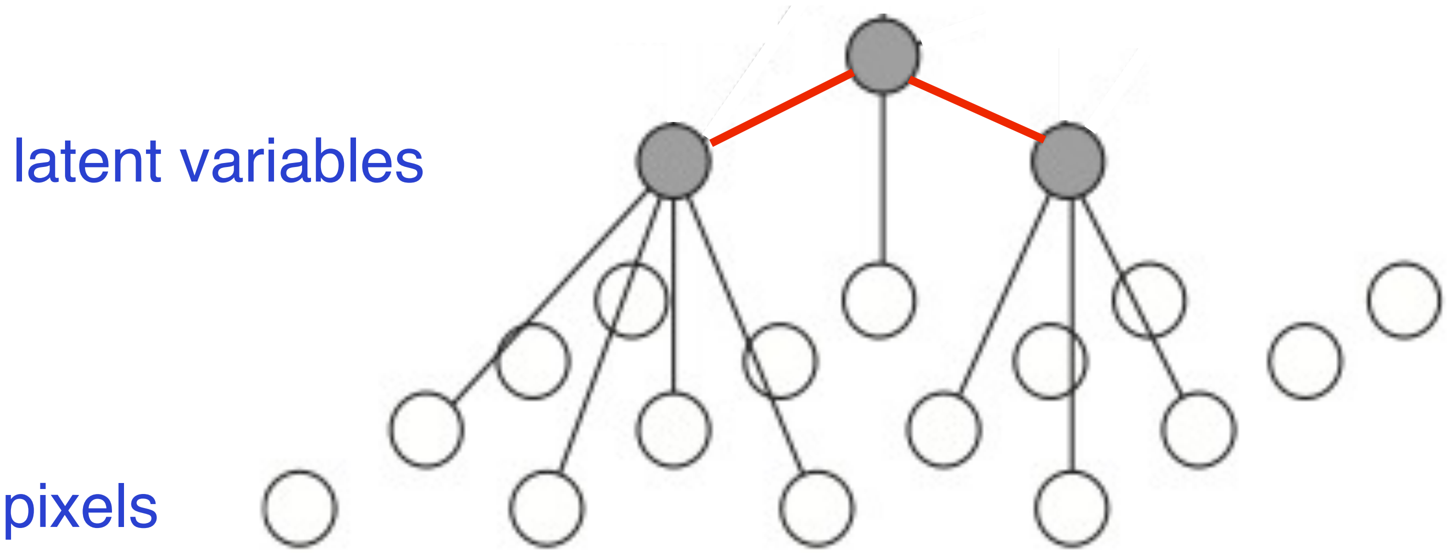




- learn deep belief network

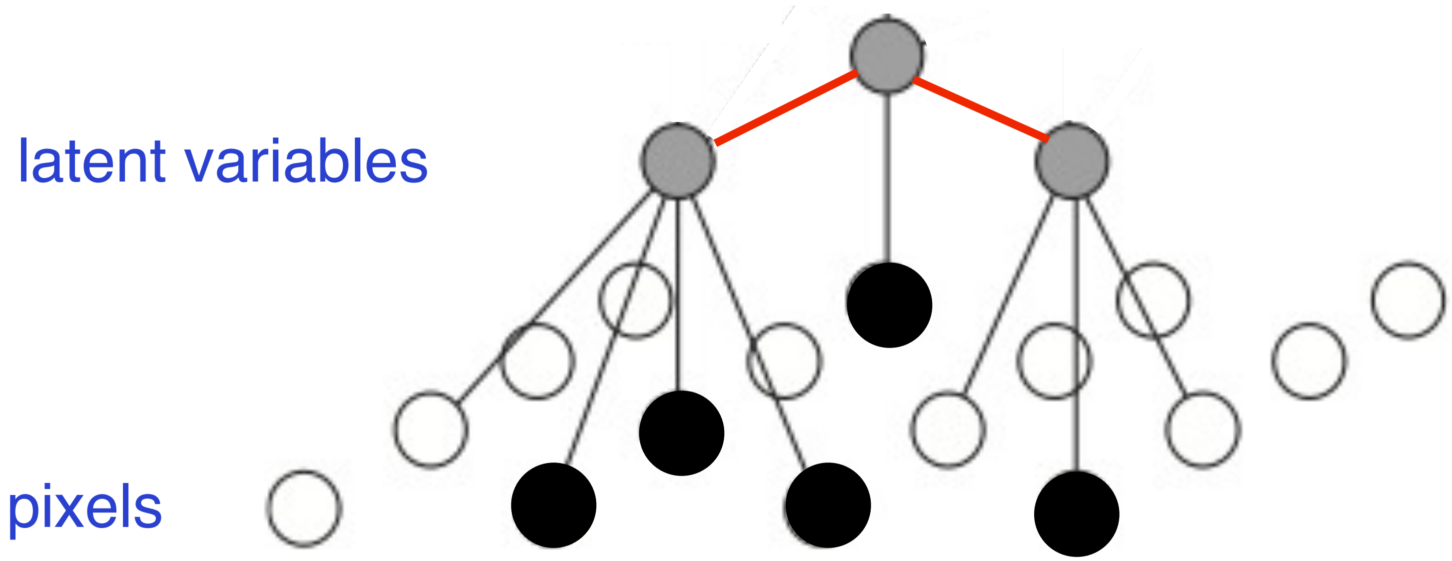


- learn deep belief network
- learn responses using elastic net



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

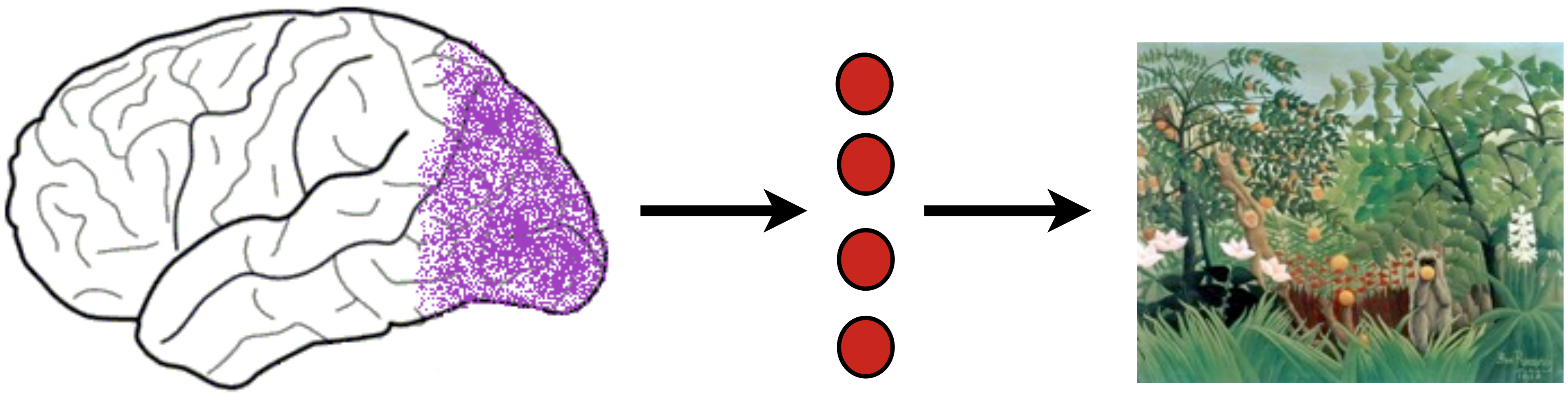
Putting things together



- learn deep belief network
- learn responses using elastic net
- represent as a Markov random field
- decode by estimating the mode of the MRF

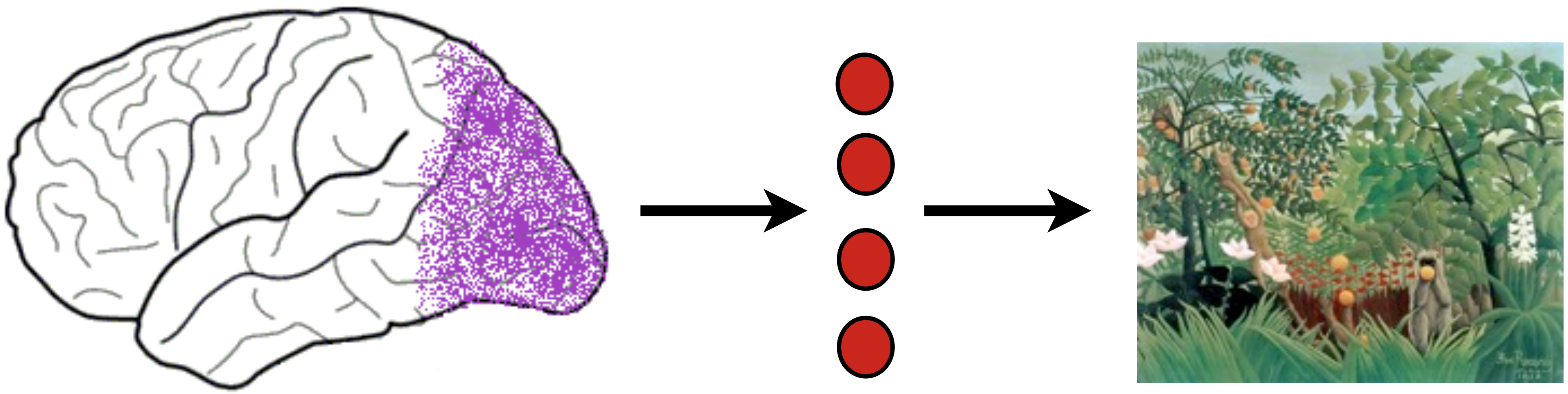
Discriminative approach: sparse partial least squares

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



van Gerven and Heskes. Sparse Orthonormalized Partial Least Squares. In: BNAIC. 2010.

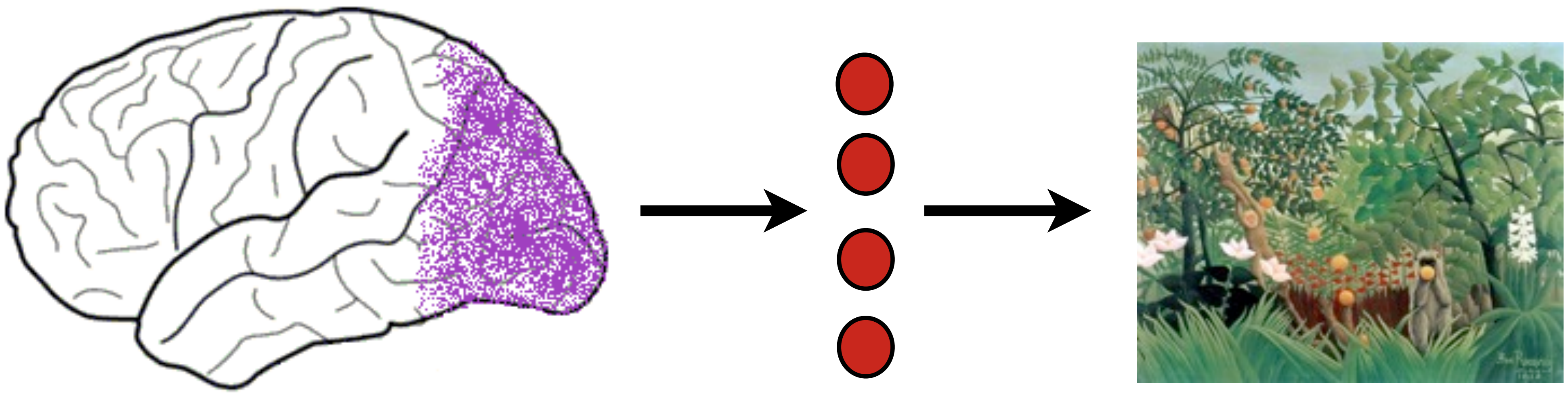
Predict image from a restricted set of responses using a small number of latent variables



Key features:

van Gerven and Heskes. Sparse Orthonormalized Partial Least Squares. In: BNAIC. 2010.

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

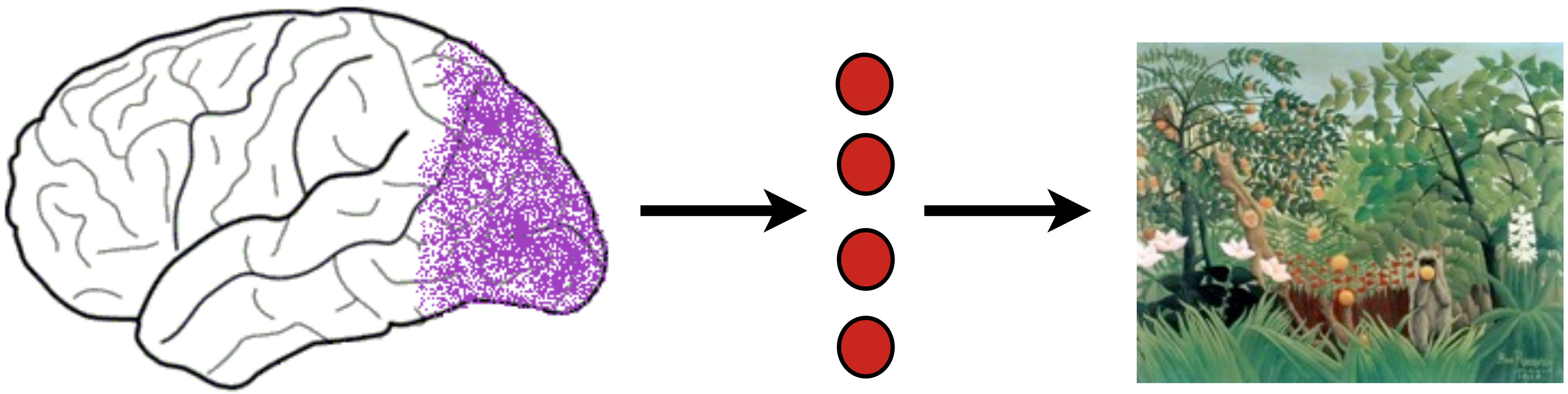


Key features:

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

van Gerven and Heskes. Sparse Orthonormalized Partial Least Squares. In: BNAIC. 2010.

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

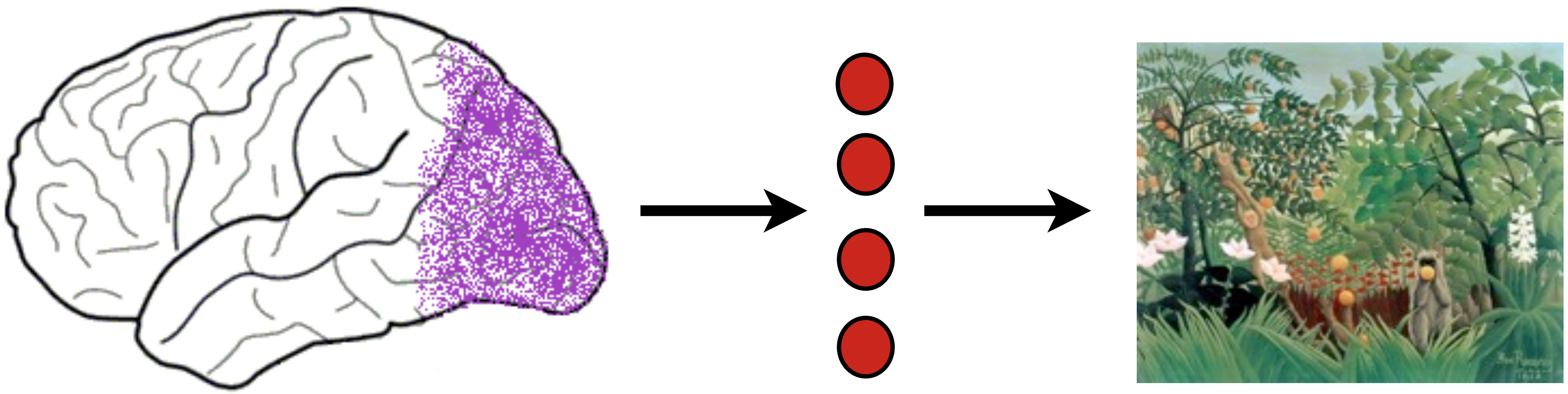


Key features:

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

van Gerven and Heskes. Sparse Orthonormalized Partial Least Squares. In: BNAIC. 2010.

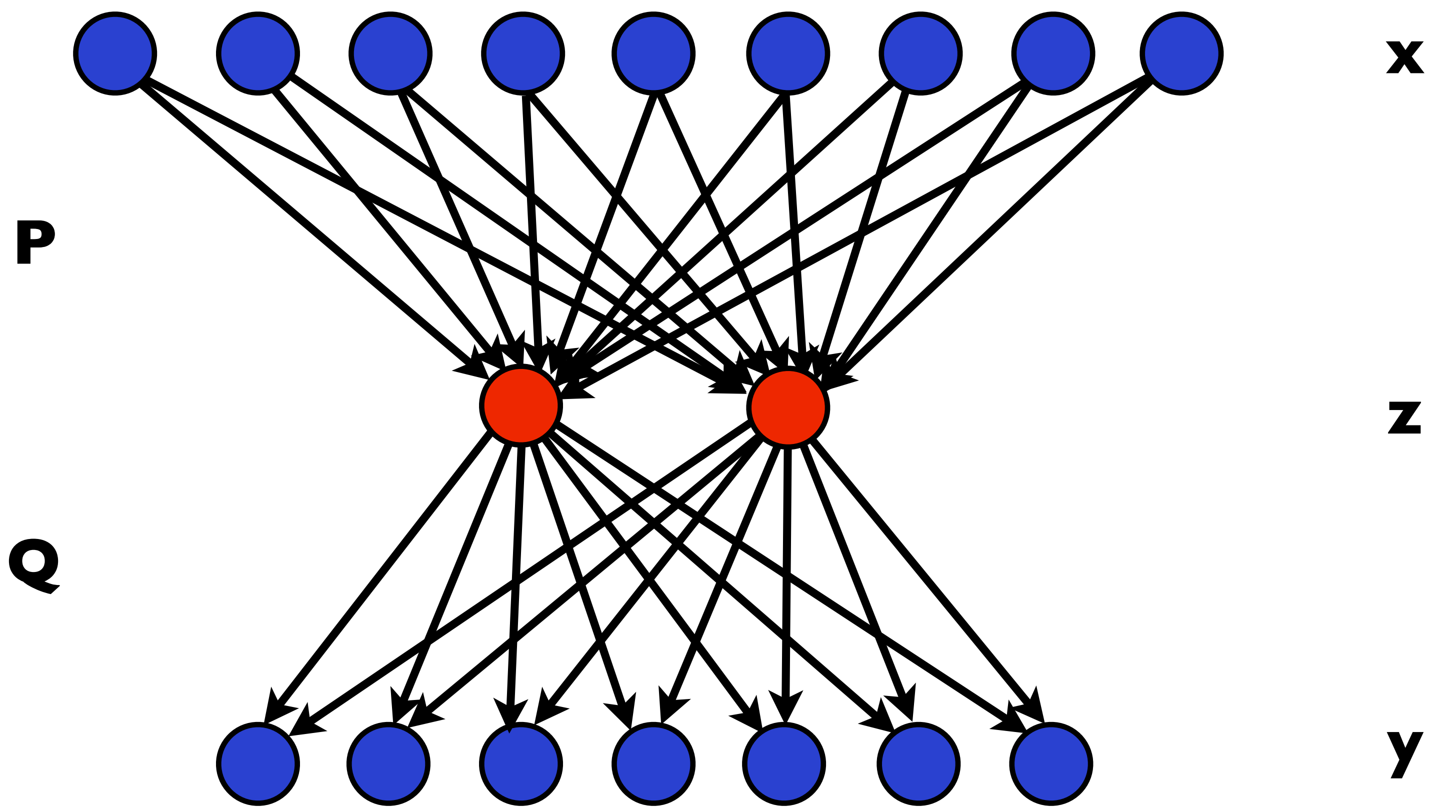
Predict image from a restricted set of responses
using a small number of latent variables



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.

van Gerven and Heskes. Sparse Orthonormalized Partial Least Squares. In: BNAIC. 2010.



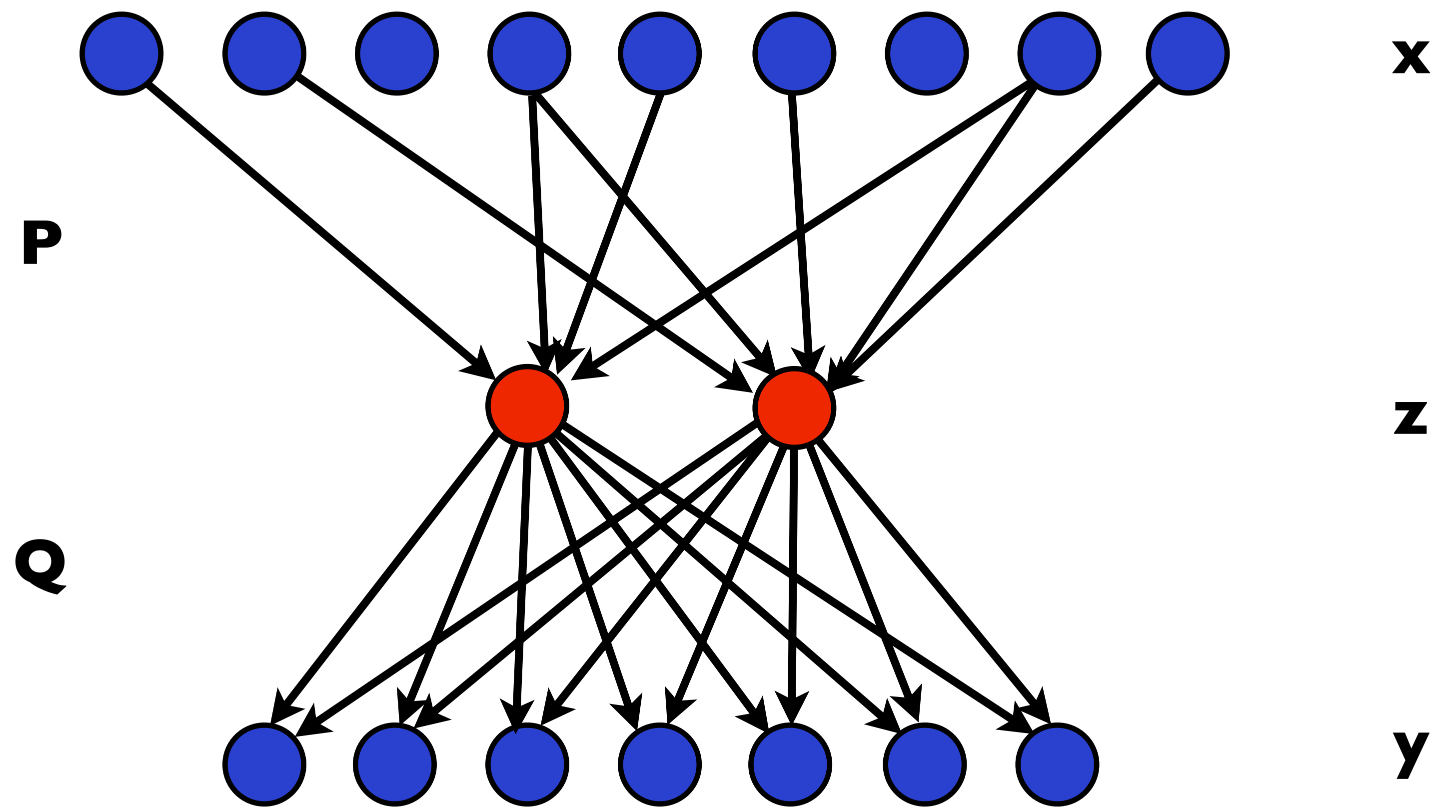
$$
\begin{aligned}
\mathbf{z} &= \mathbf{P}^\top \mathbf{x} \\
\mathbf{y} &= \mathbf{Q}\mathbf{z}
\end{aligned}
$$

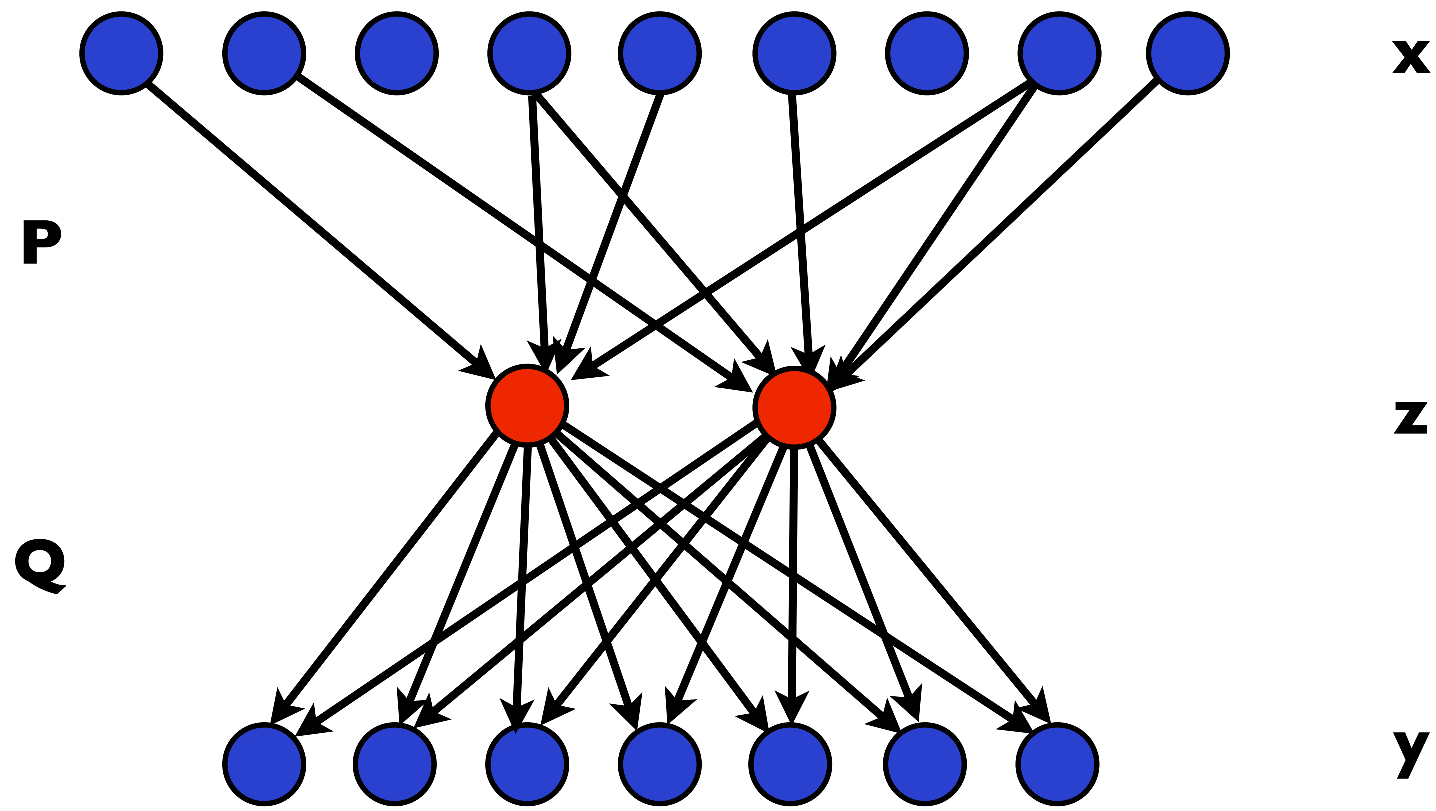
Linear heteroencoder:

- Unique optimal solution (no local minima).
- reduces to principal component analysis in case x=y;

  rows of **Q** correspond to principal components of y.

Sparse partial least squares

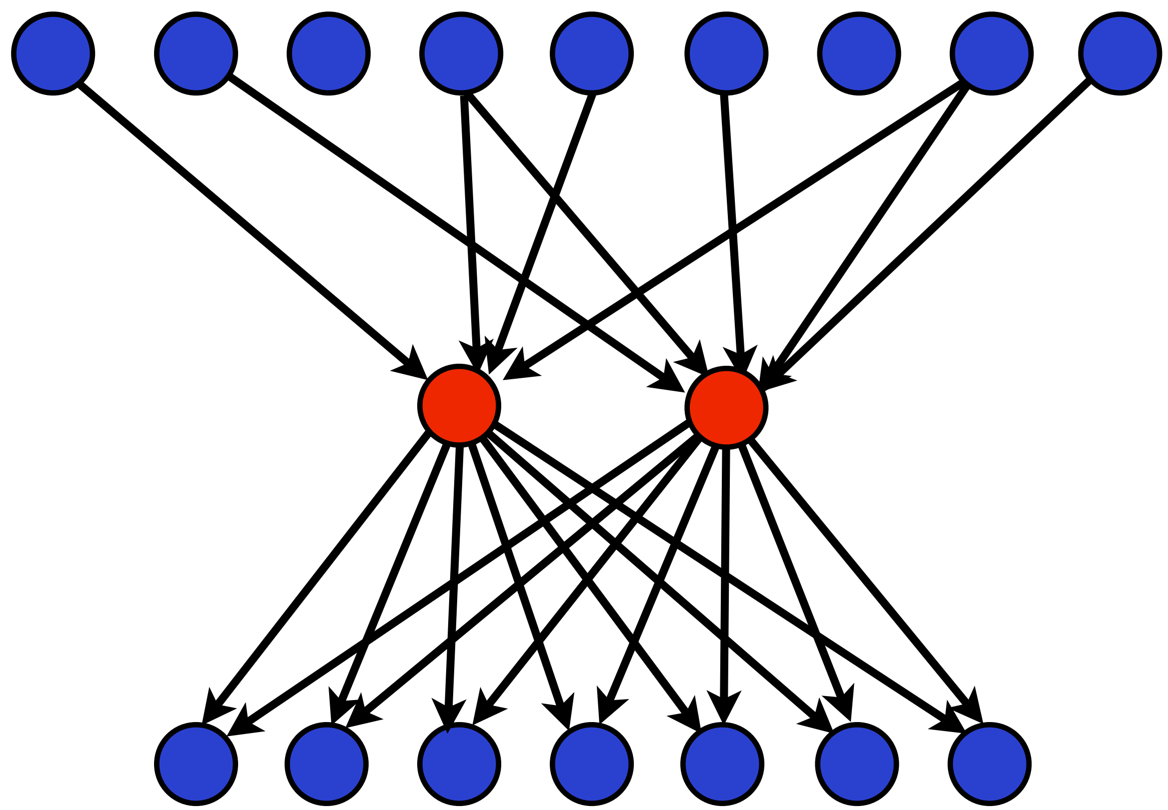




Objective:

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

with $R_{\nu,\Lambda}(\mathbf{P}) = \nu \sum_{i=1}^{k} ||\mathbf{P}_i||_1 + \frac{1}{2} \sum_{j=1}^{k} \mathbf{P}_j^\top \Lambda \mathbf{P}_j$



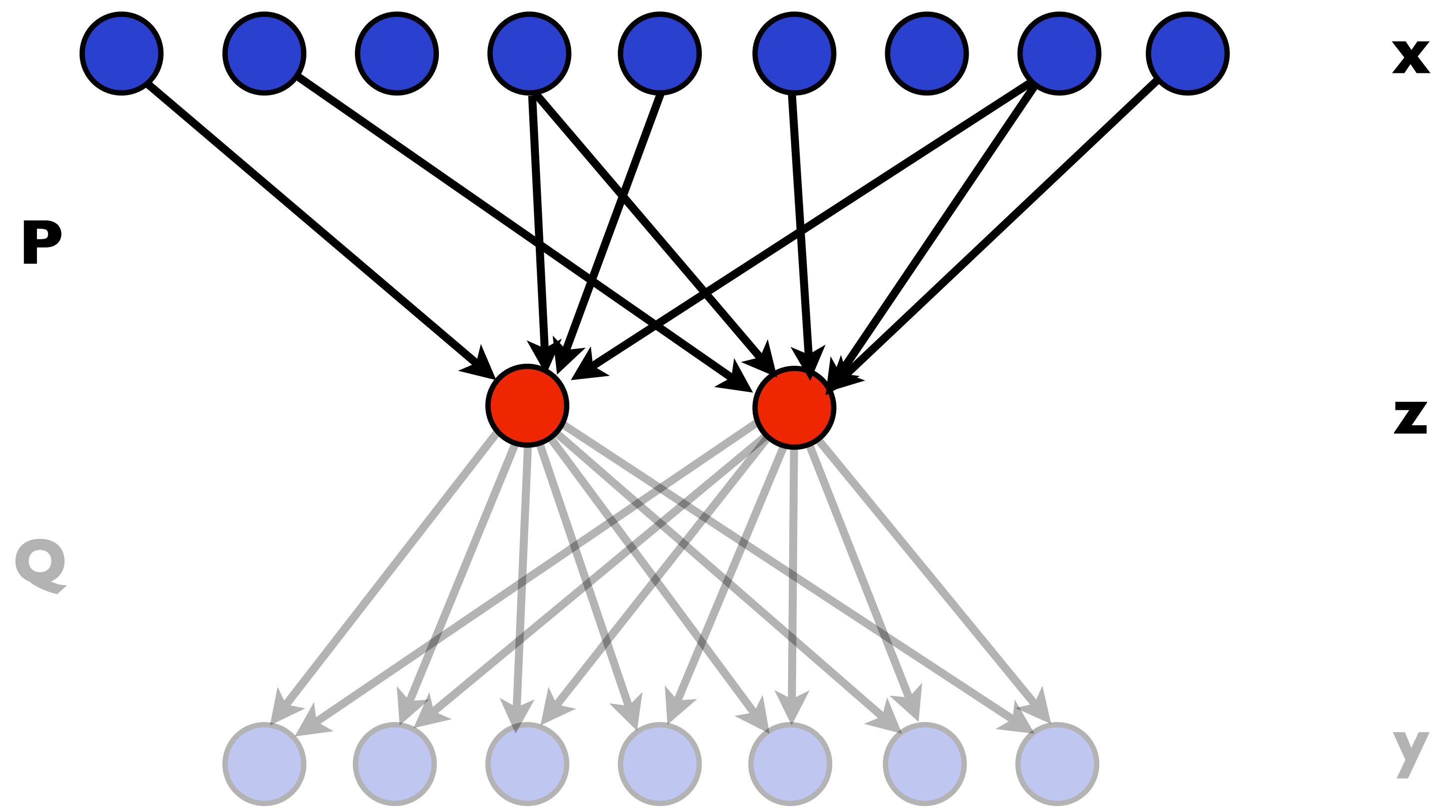
Objective:

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

with $R_{\nu,\Lambda}(\mathbf{P}) = \nu \sum_{i=1}^{k} ||\mathbf{P}_i||_1 + \frac{1}{2} \sum_{j=1}^{k} \mathbf{P}_j^\top \Lambda \mathbf{P}_j$

- reduces to sparse PCA in case x=y (Zou et al., J Comput Graph Stat, 2006)

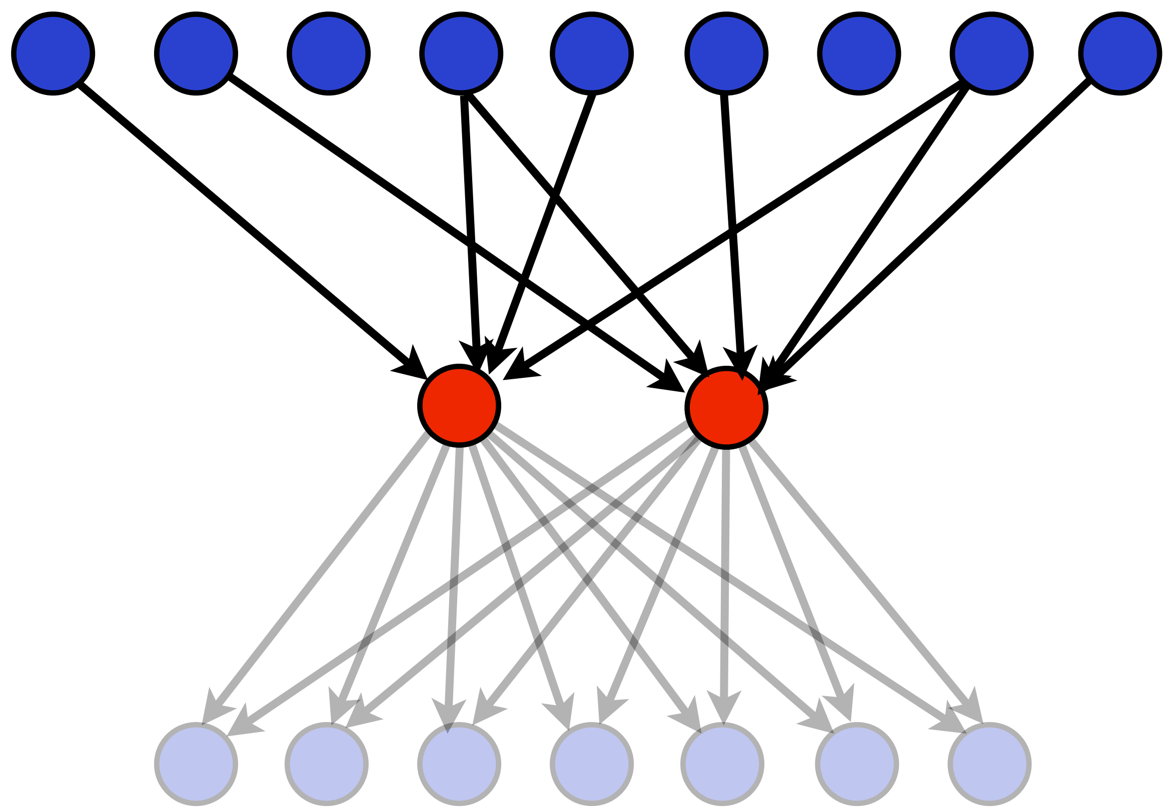
Iterative sparse PLS algorithm



Fix $\mathbf{Q}$, reconstruct $\mathbf{Z} = \mathbf{Q}^T\mathbf{Y}$, and solve

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

Iterative sparse PLS algorithm



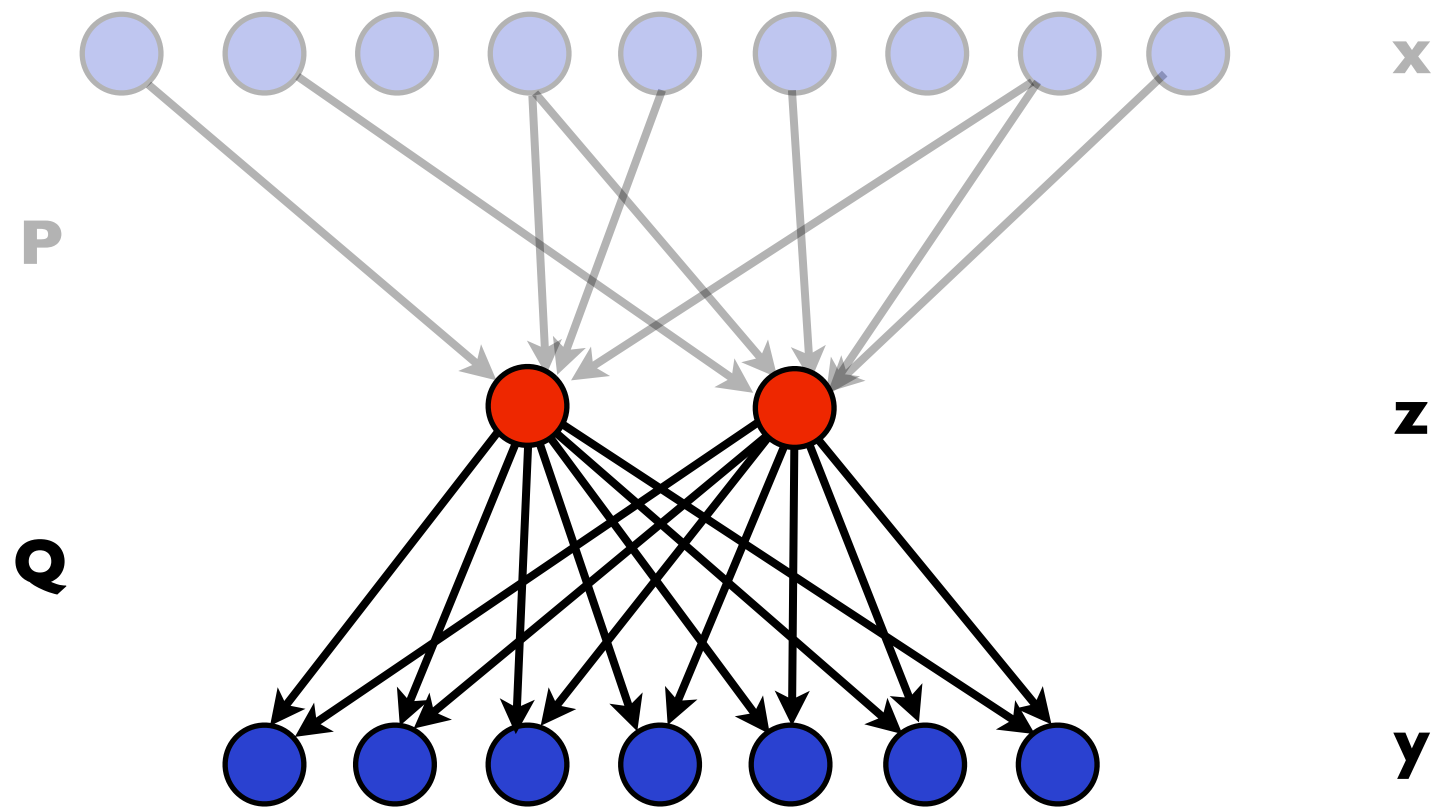
Fix $\mathbf{Q}$, reconstruct $\mathbf{Z} = \mathbf{Q}^T\mathbf{Y}$, and solve

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

- set of standard elastic net problems

Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. J. Stat. Softw. 2010;33(1):1–22.

Iterative sparse PLS algorithm

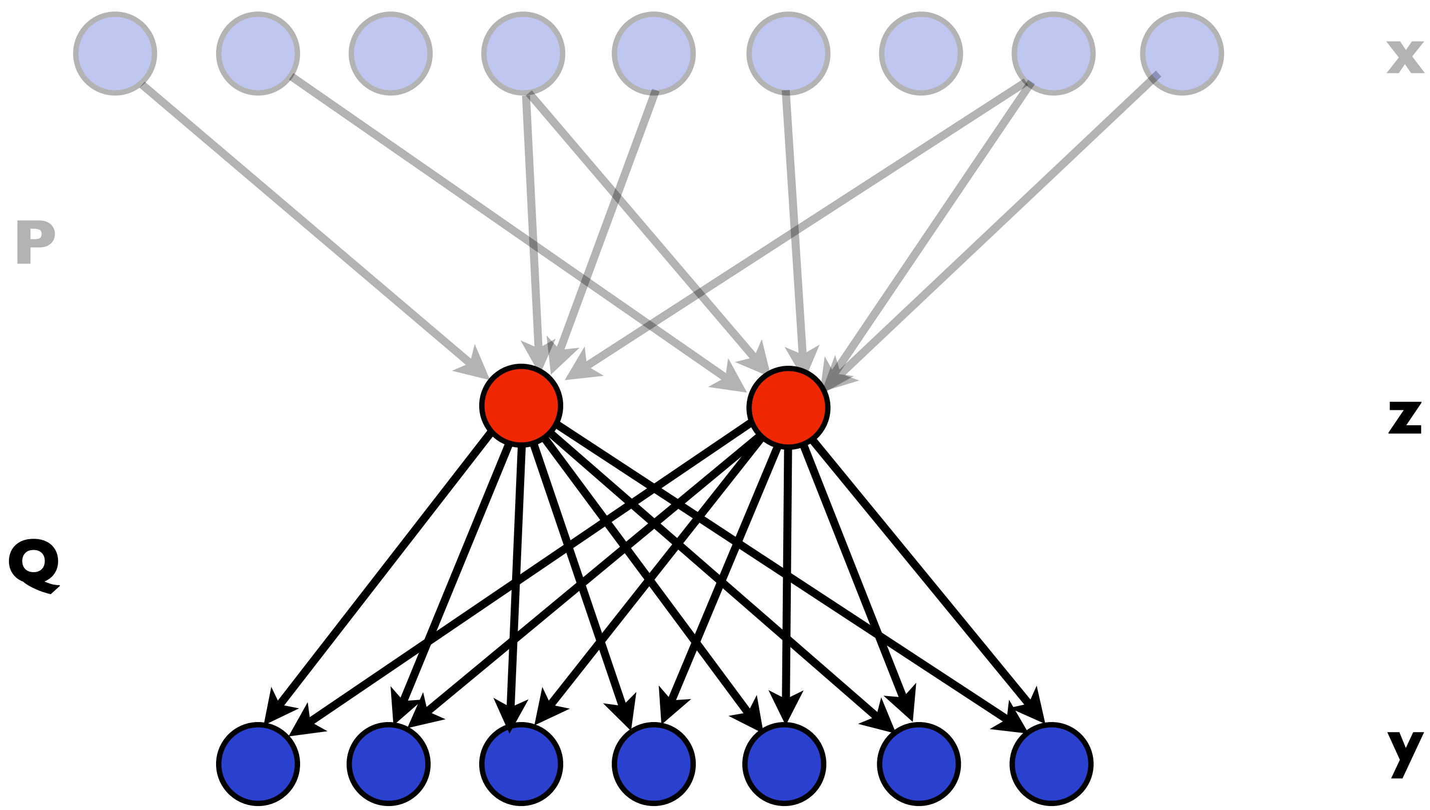


Fix $\mathbf{P}$, reconstruct $\mathbf{Z} = \mathbf{P}^T\mathbf{X}$, and solve

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

subject to $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_k$

Iterative sparse PLS algorithm



Fix $\mathbf{P}$, reconstruct $\mathbf{Z} = \mathbf{P}^T\mathbf{X}$, and solve

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

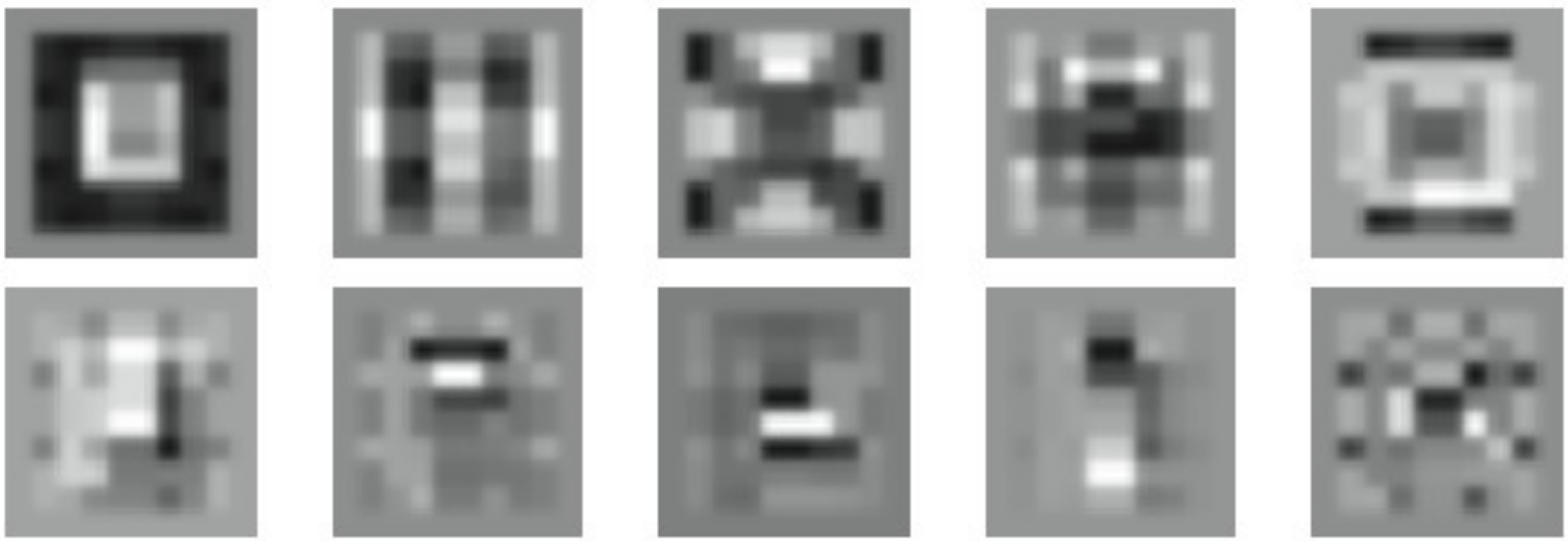
subject to $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}_k$

- $\hat{\mathbf{Q}} = \boldsymbol{\Sigma}_{yz} \left(\boldsymbol{\Sigma}_{yz}^T \boldsymbol{\Sigma}_{yz} \right)^{-1/2}$ with $\boldsymbol{\Sigma}_{yz} \equiv \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}^{(n)} \left(\mathbf{z}^{(n)} \right)^T$

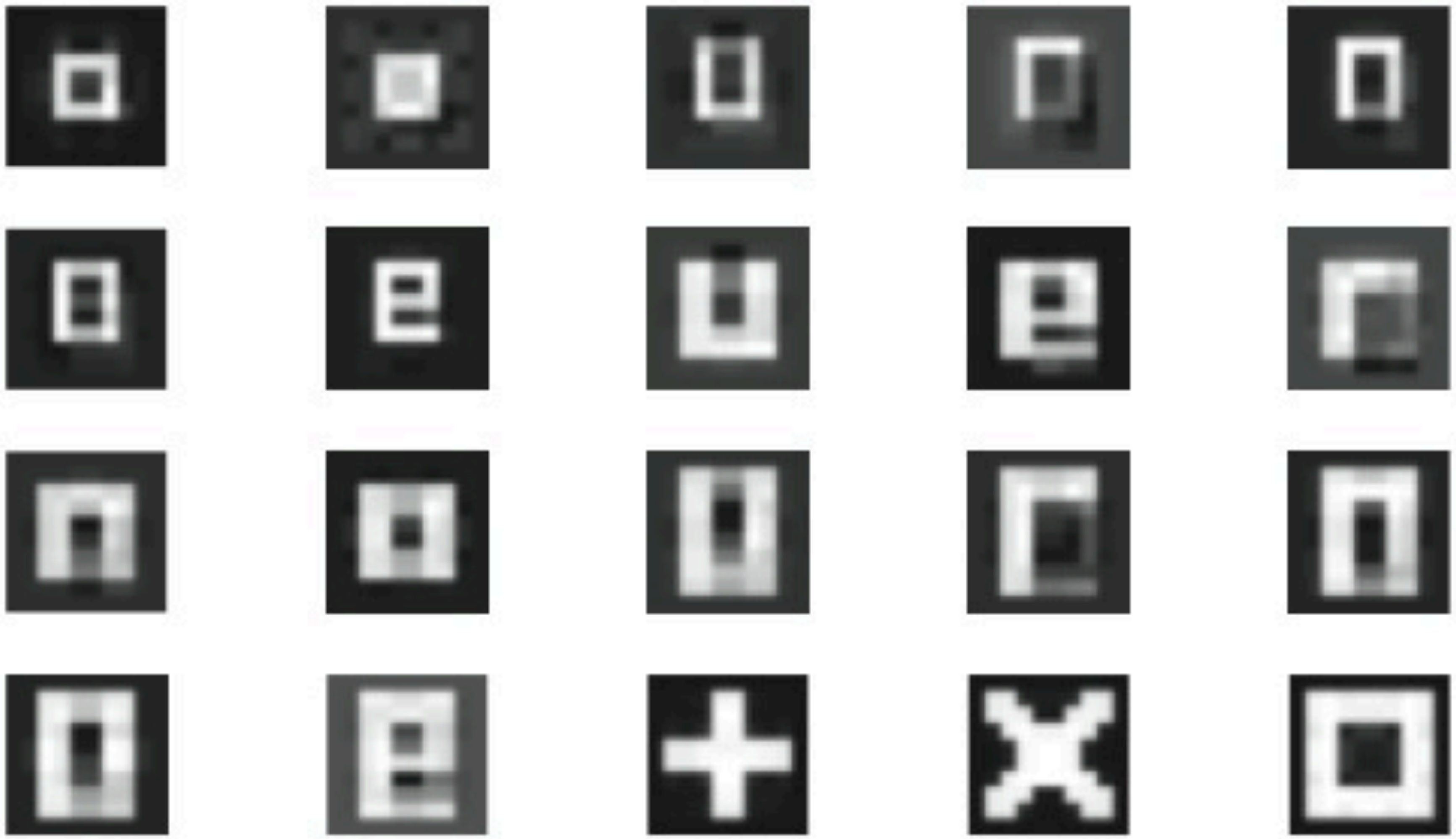
Experiment

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

Learned features



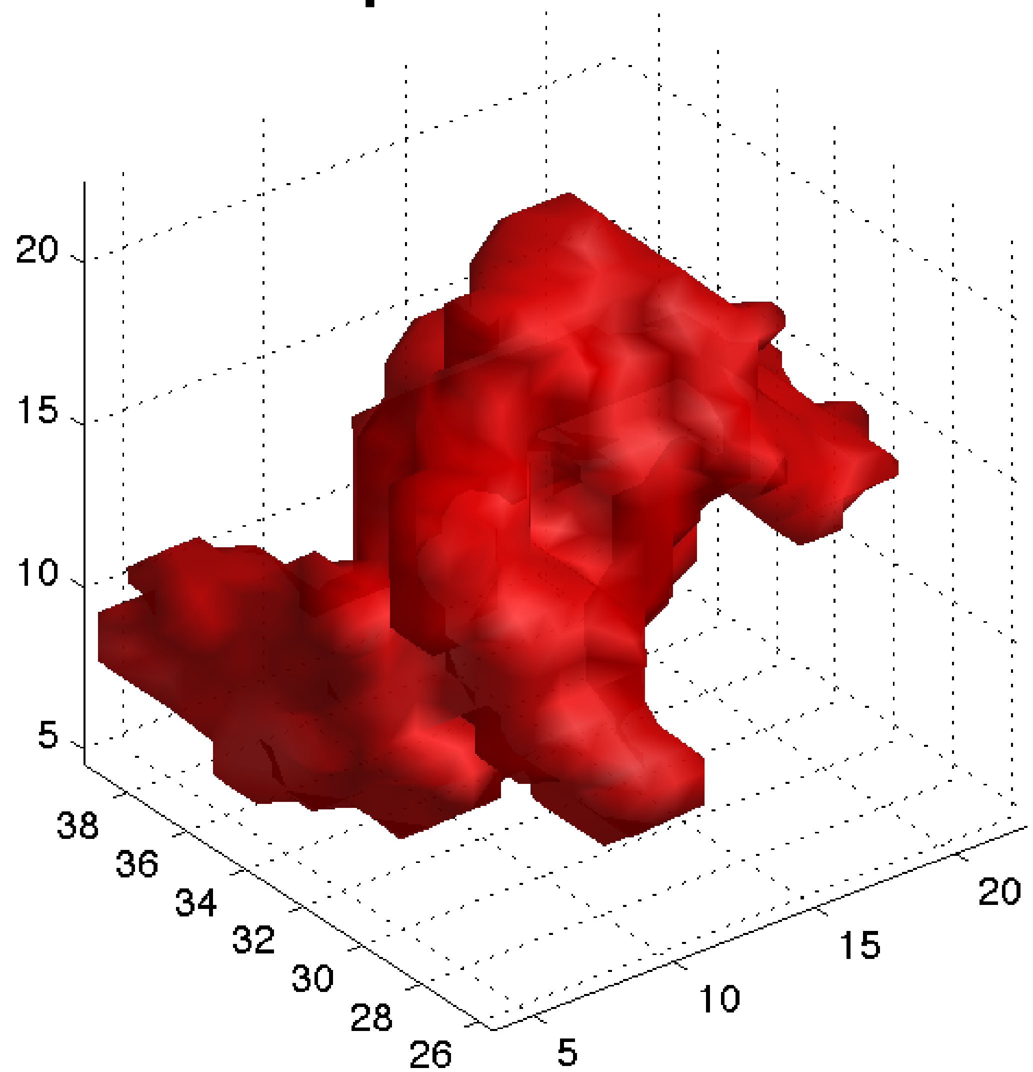
Learned features (rows of the matrix **Q**) are similar to principal components of the original images but change as a function of ν



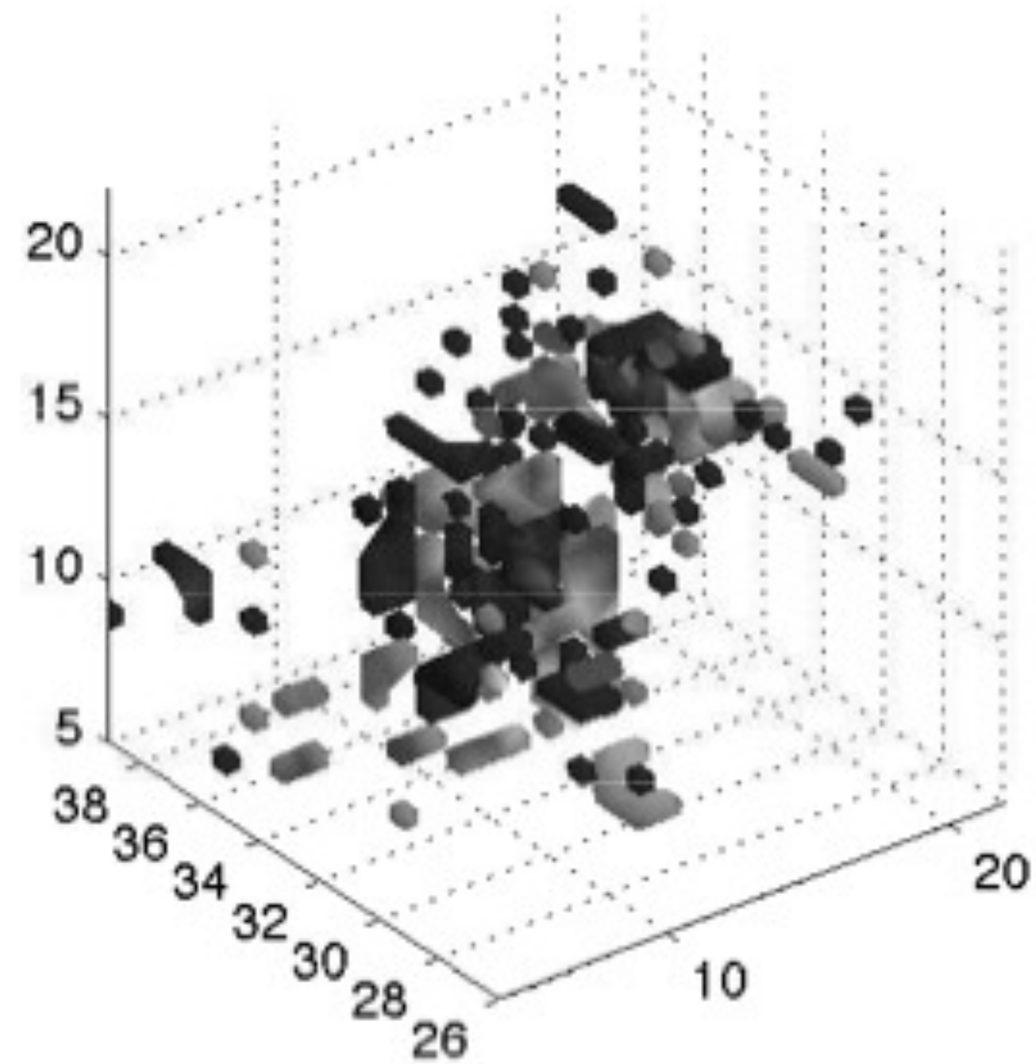
Reconstructions on hold-out data

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

input voxels

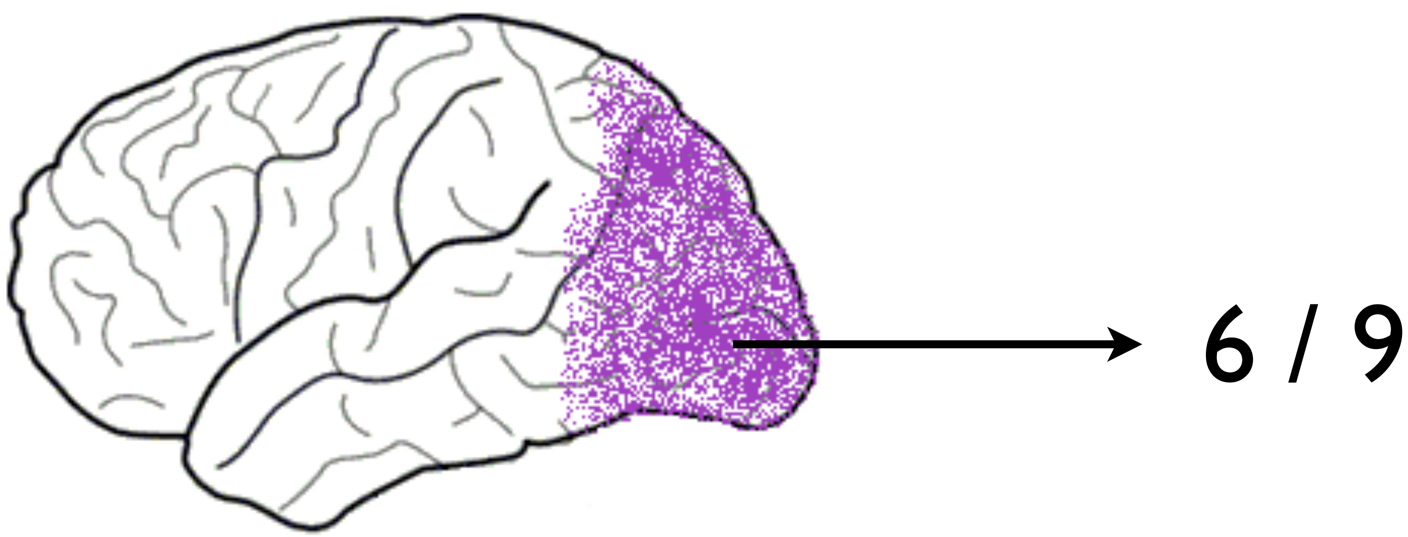


selection



For ν=0.01, 80% of the parameters are set to zero

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



We require:

$$p(\boldsymbol{\theta} \mid \mathbf{D}) \propto p(\mathbf{D} \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})$$

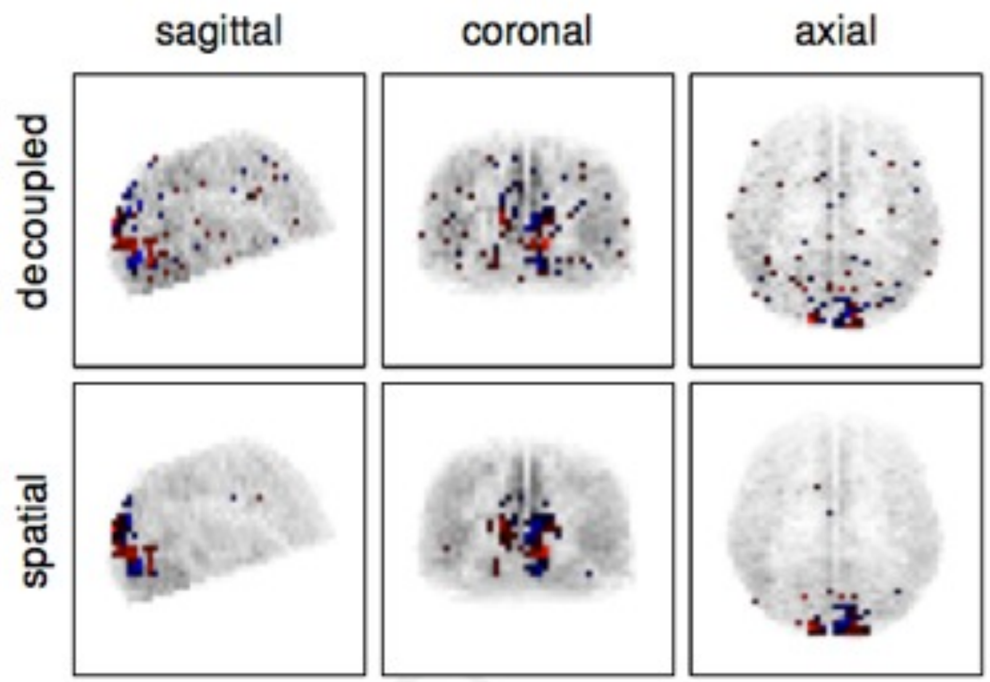
likelihood: *logistic regression model*

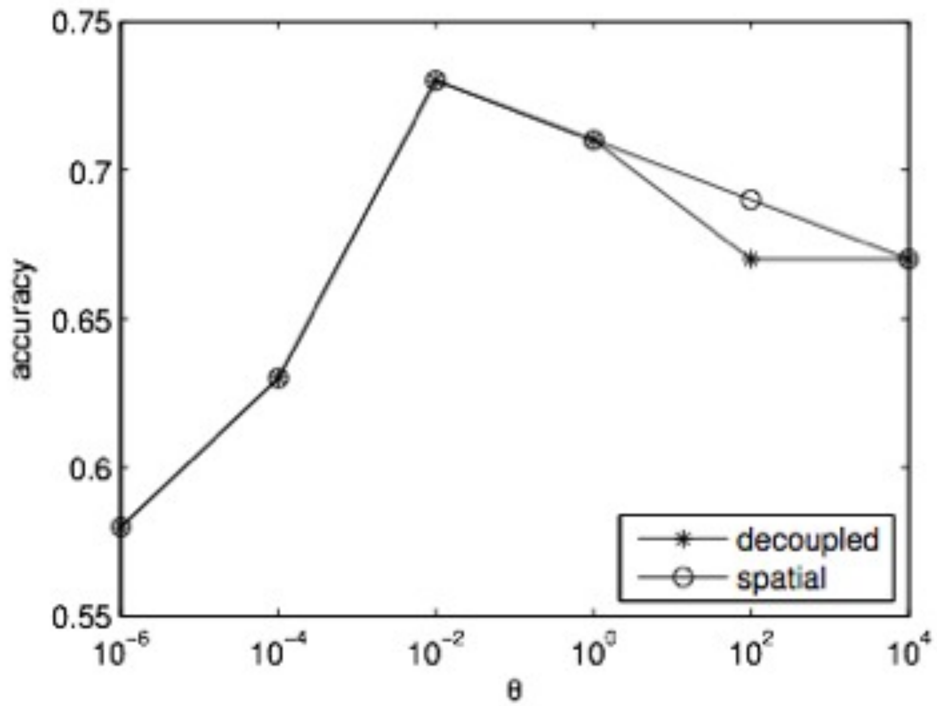
prior: *sparse and smooth solutions*

$$p(\boldsymbol{\theta}) = \int d\mathbf{uv} \left(\prod_k \mathcal{N}(\theta_k; 0, u_k^2 + v_k^2) \right) \mathcal{N}(\mathbf{u}; \mathbf{0}, \mathbf{Q}) \mathcal{N}(\mathbf{v}; \mathbf{0}, \mathbf{Q})$$

- preference for small regression coefficients
- large magnitude in one coefficient reduces regularization of coupled coefficients

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





van Gerven MAJ, Cseke B, de Lange FP, Heskes T. Efficient Bayesian Multivariate fMRI analysis using a sparsifying spatio-temporal prior. Neuroimage. 2010;50(1):150–161.

van Gerven MAJ, Cseke B, Oostenveld R, Heskes T. Bayesian source localization with the multivariate Laplace prior. In: Bengio Y, Schuurmans D, Lafferty J, Williams CKI, Culotta A, editors. Neural Information Processing Systems 23. 2009. p. 1901–1909.

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

- 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

- 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

- 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

- 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

- 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:

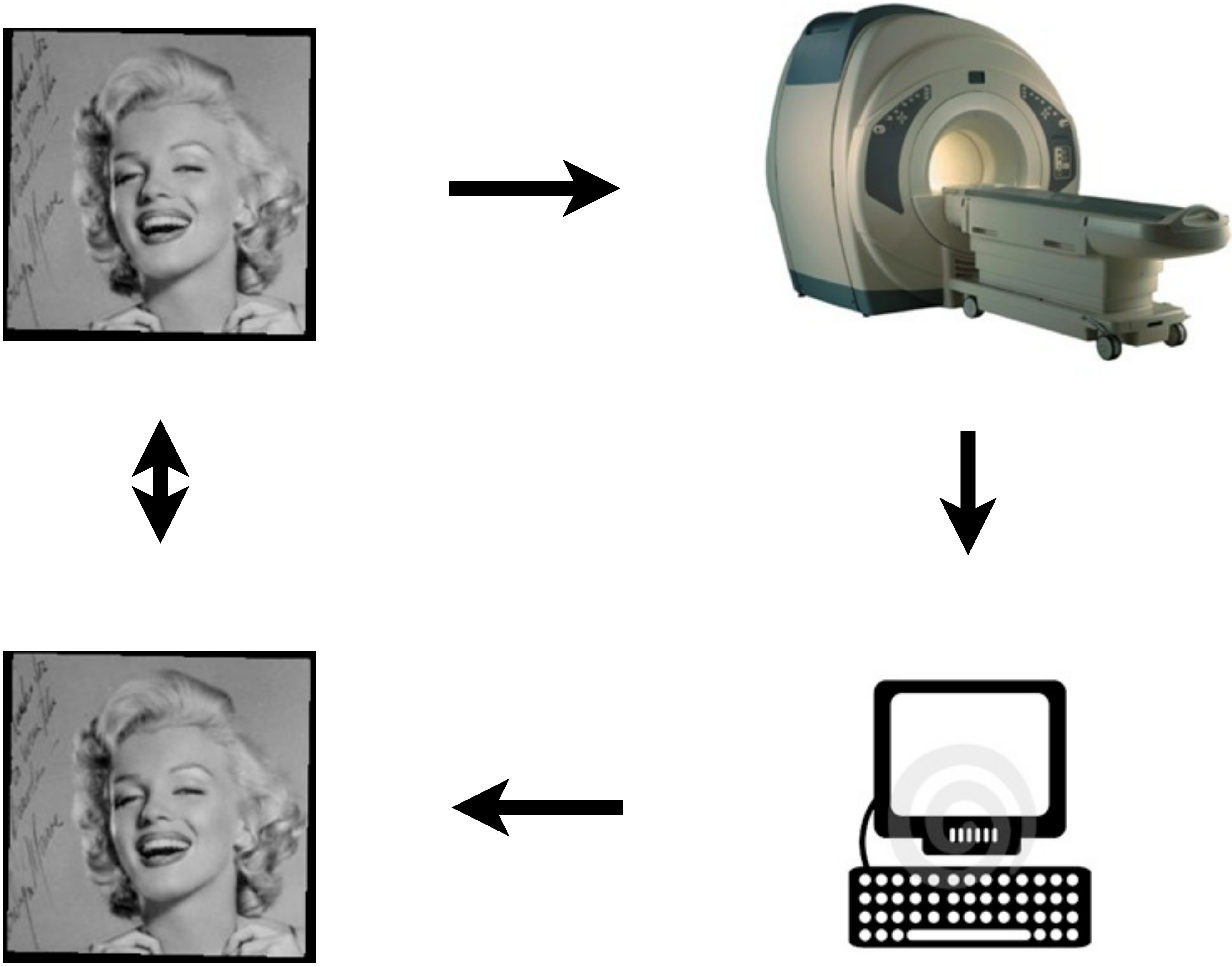
- other regularizers (e.g., total variation)
- multitask formulation of SPLS
- Bayesian formulation of SPLS
- combine low-level and high-level information
- focus on reactivation during imagery and memory retrieval
- use intracranial EEG and high-field fMRI

Acknowledgements:

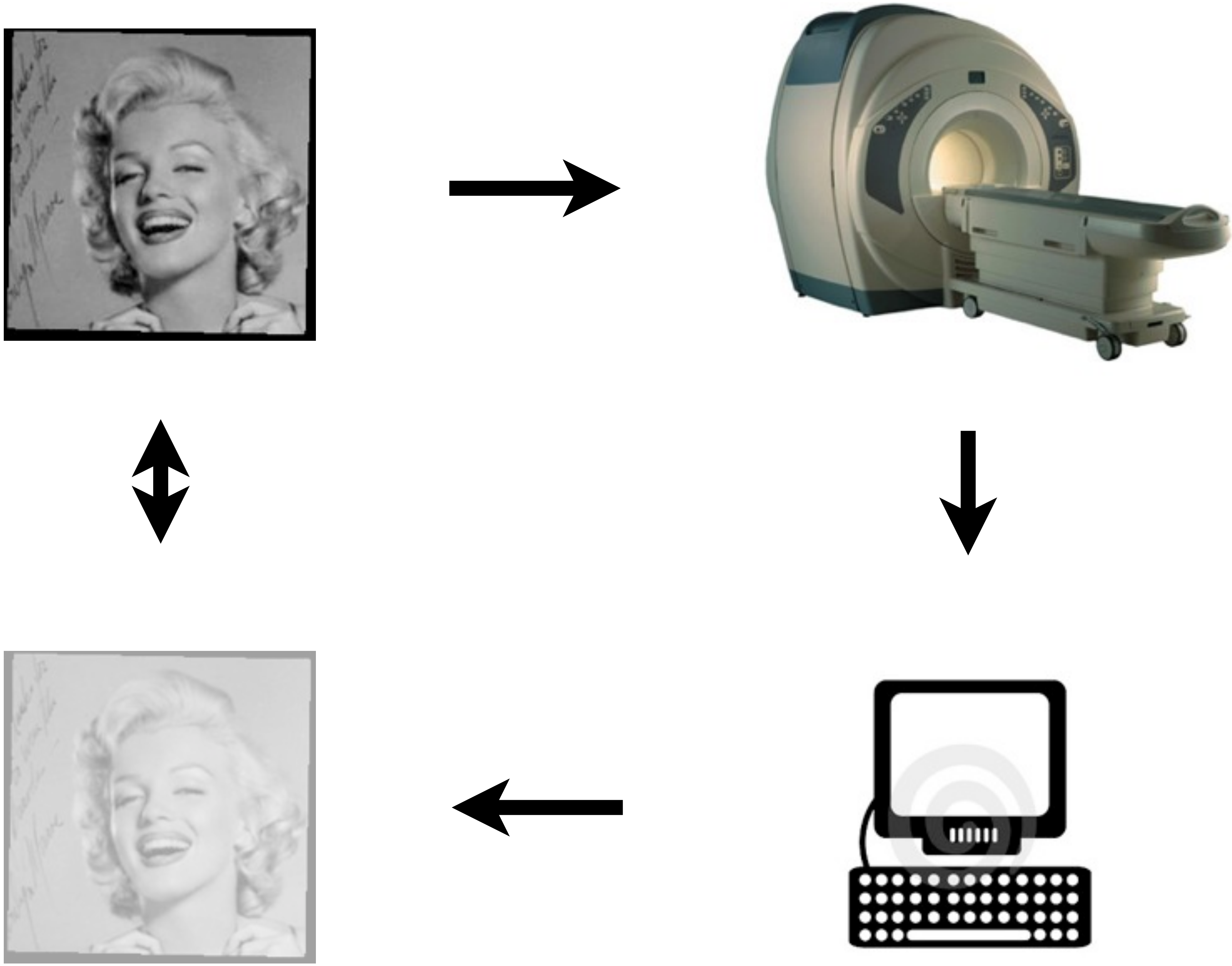
Tom Heskes, Floris de Lange, Eric Maris

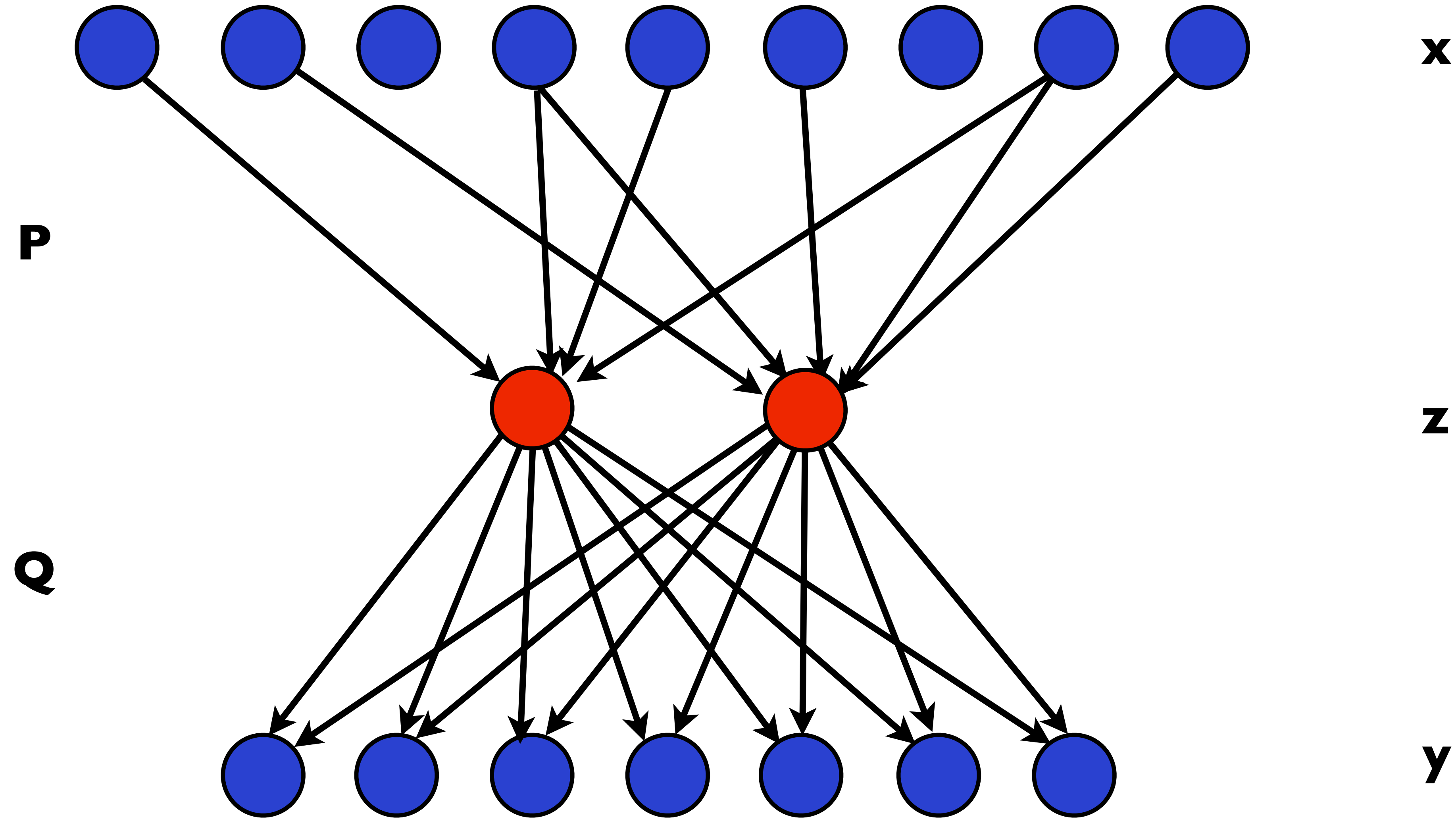
methods part of FieldTrip multivariate module (http://fieldtrip.fcdonders.nl)

Face decoding

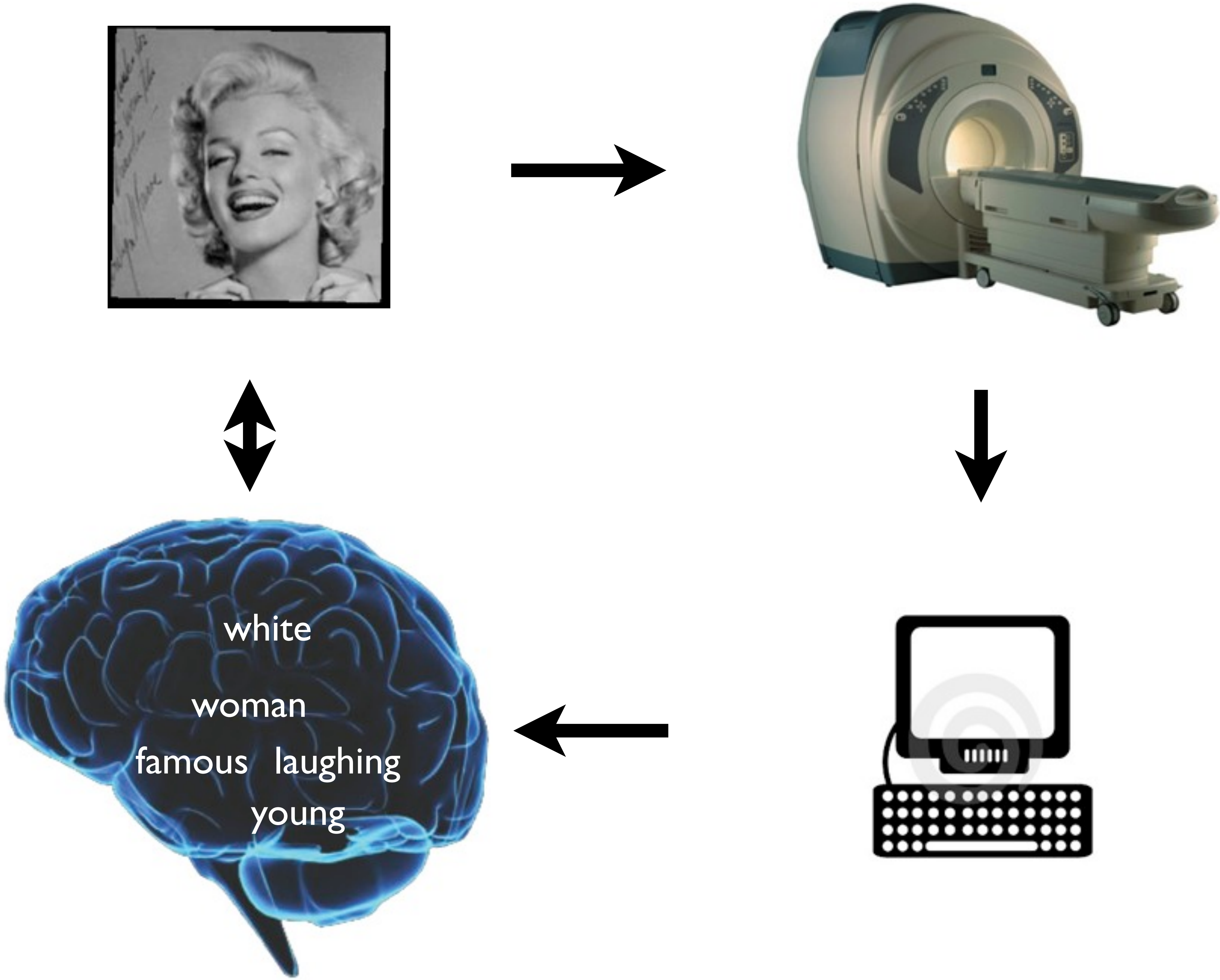


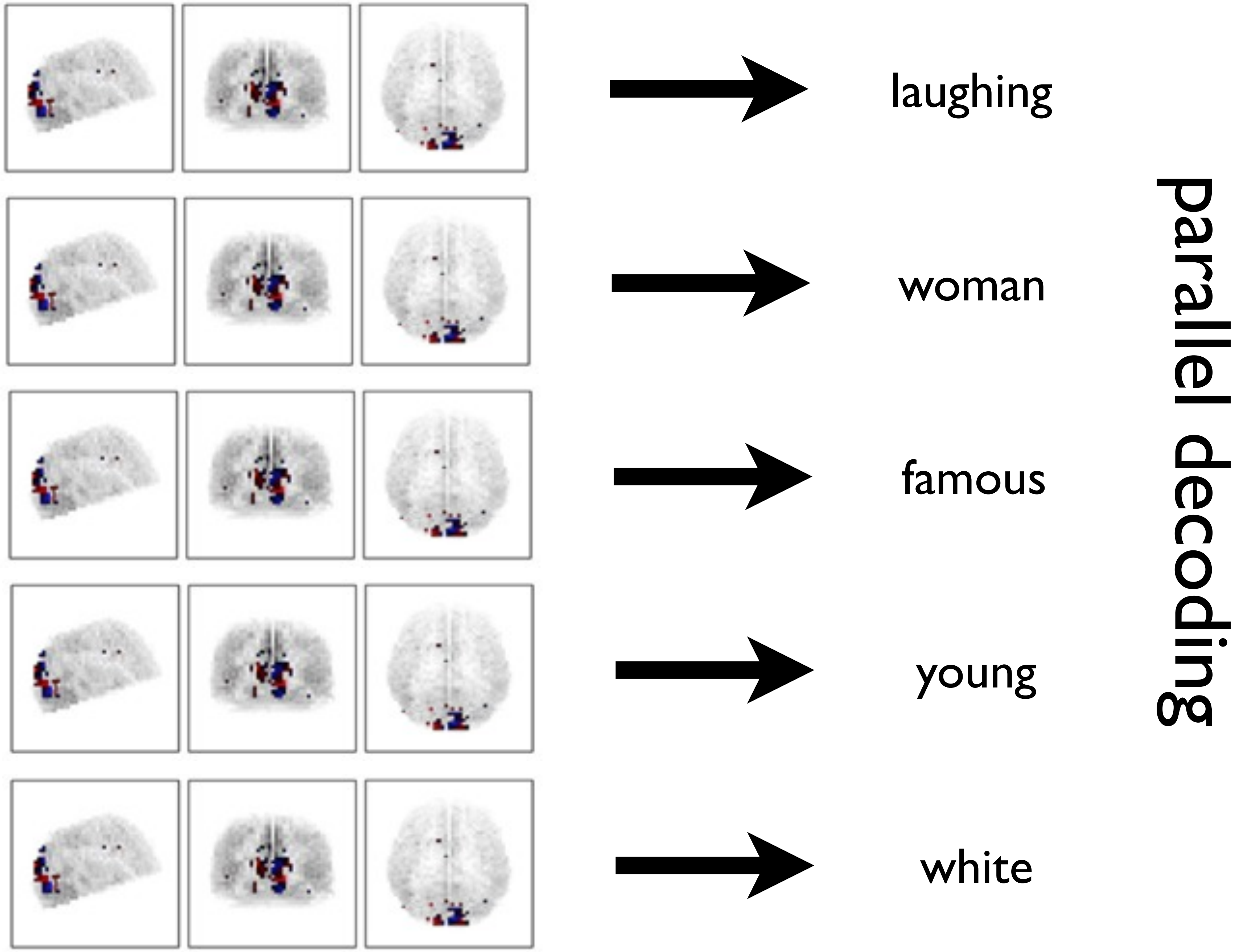
Spatial statistics




x
P
z
Q
y

Semantics





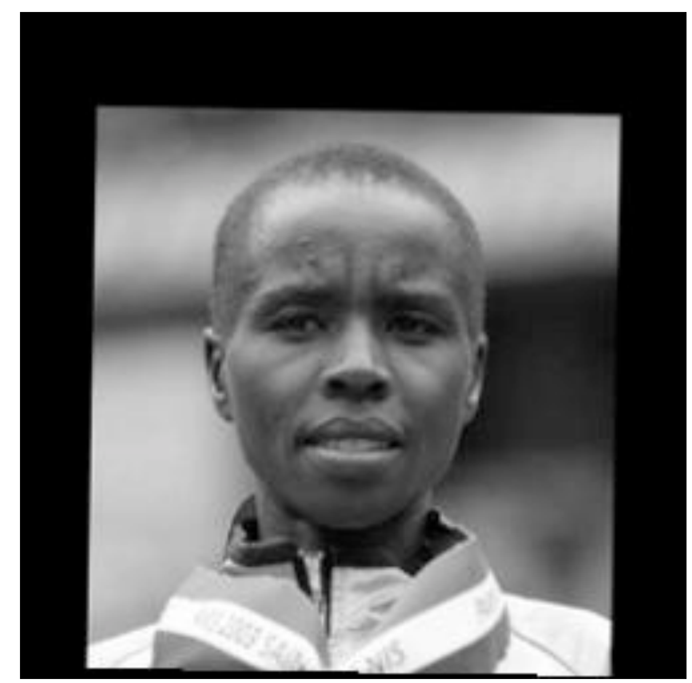
M. A. J. van Gerven, B. Cseke, F. P. de Lange, and T. Heskes. Efficient Bayesian multivariate fMRI analysis using a sparsifying spatio-temporal prior. *NeuroImage*, 50(1):150–161, 2010.

spatial statistics

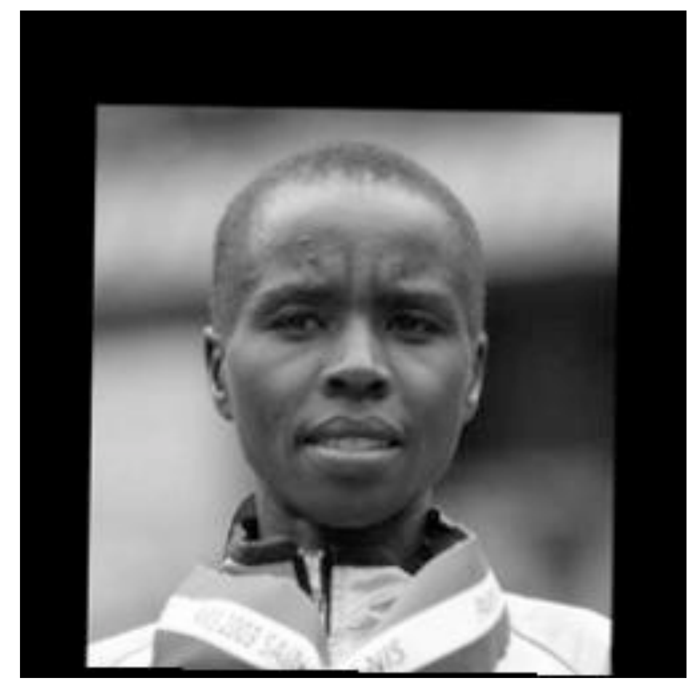
semantics

decoding of gender 80% correct

Face identification



Face identification



Face identification

laughing

laughing

young

laughing

young

woman

Imagined face identification