

# Manifold Learning: Applications in Neuroimaging

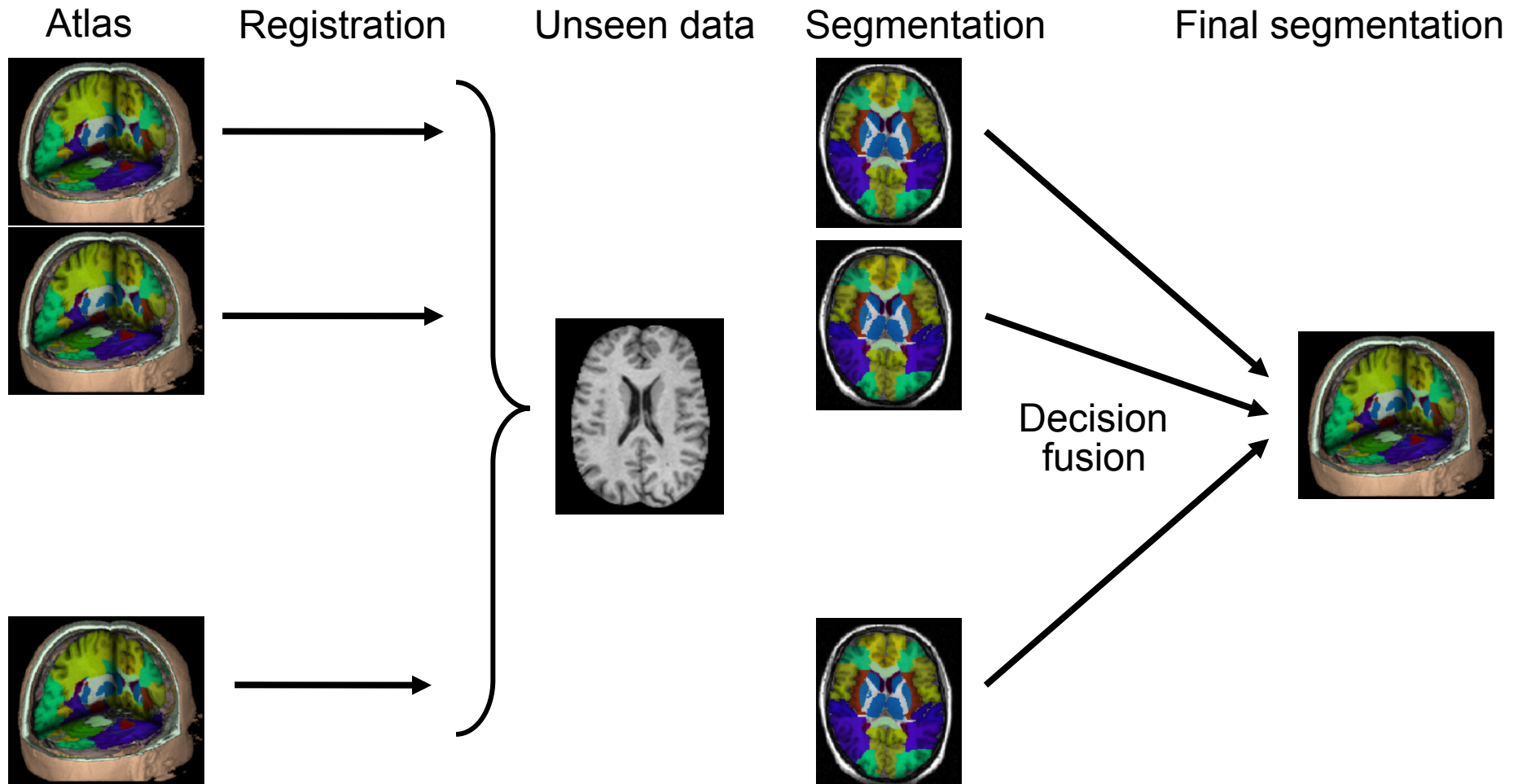
Robin Wolz

23/09/2011

# Overview

- Manifold learning for Atlas Propagation
  - Multi-atlas segmentation
  - Challenges
  - LEAP
- Manifold learning for classification
  - Cross-sectional data
  - Longitudinal data
  - Metadata
- Conclusions

# Segmentation using multi-atlas fusion

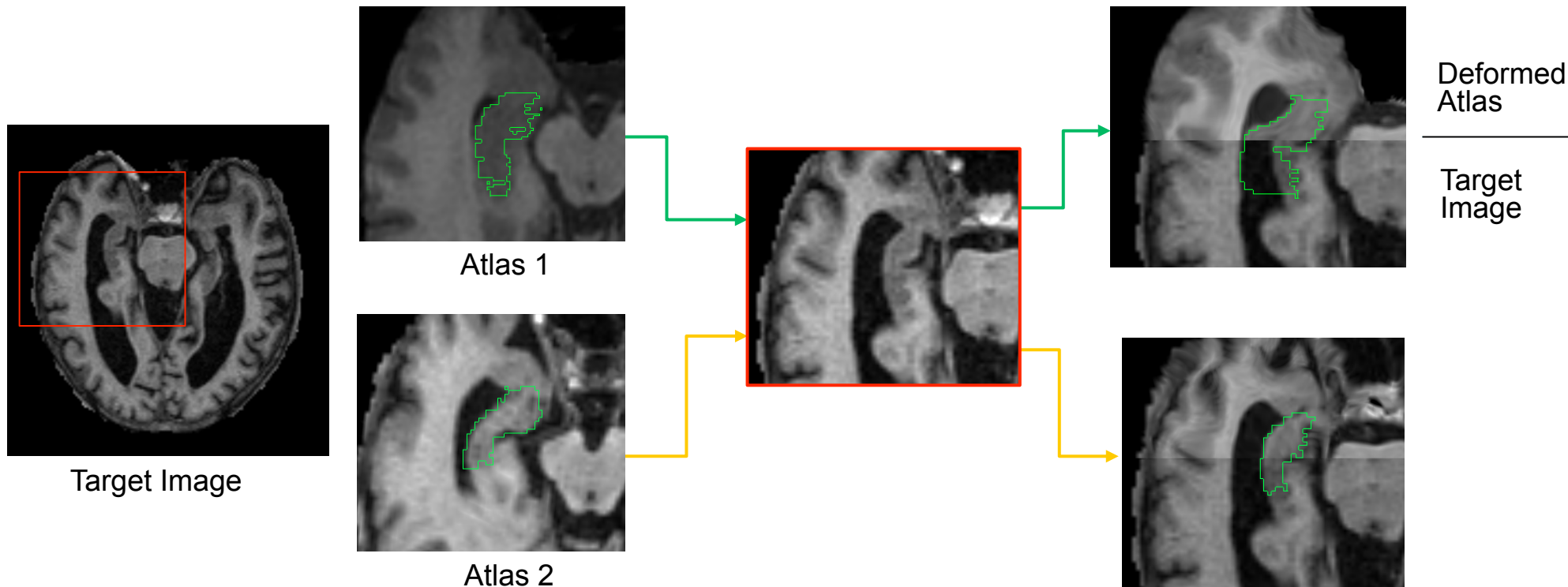


Heckemann et al., Neuroimage 2006

# Segmentation using multi-atlas fusion

## Problems:

- Number of atlases is typically limited
- Changing population characteristics or disease may necessitate new atlases



# Segmentation using multi-atlas fusion

## **Problems:**

- Number of atlases is typically limited
- Changing population characteristics or disease may necessitate new atlases

## **Solutions:**

- Can we bootstrap or learn atlases from the population directly?
- Use manifold learning to model characteristics of a population of images

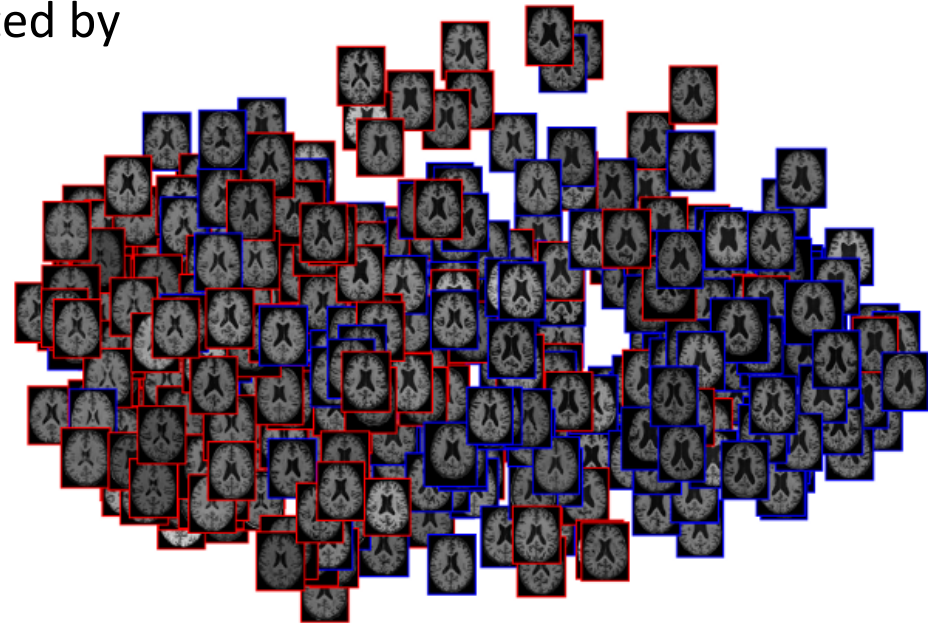
# Population modelling

- Space of brain MR images is typically very high-dimensional ( $D > 10^6$ )
- The natural variation of images may be described in a space with much lower dimension  $d$
- Manifold learning aims at establishing this low-dimensional space
- $N$  input images are represented by intensity vectors

$$\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_N\} \in \mathbb{R}^D$$

- Manifold coordinates are of dimension  $d$

$$\mathbf{Y} = \{\mathbf{y}_1, \dots, \mathbf{y}_N\} \in \mathbb{R}^d$$



# How to measure similarities

- A similarity measure can be defined based on the application:

## **Shape-based measures**

- Distances extracted from the deformation
- Deformation magnitude
- Jacobian determinant
- Other measures extracted from the deformation field

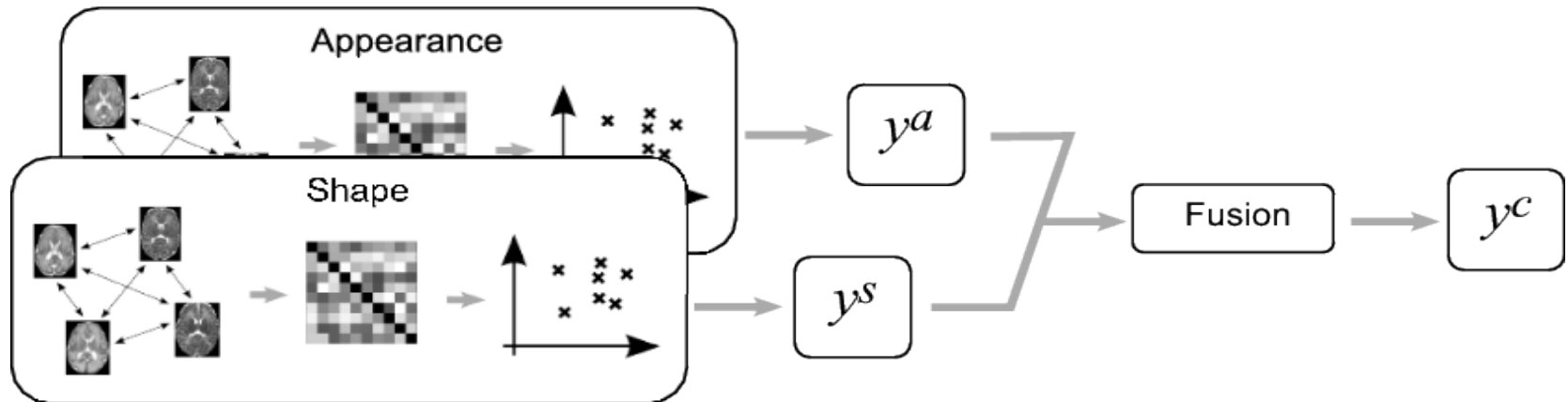
## **Appearance-based measures**

- Similarities extracted from image intensities
- Sums of squared differences (SSD)
- Cross-correlation
- Mutual information

- A weighted measure combining shape and appearance captures both aspects
- Similarities  $S_{ij}$  can be transformed to distances  $D_{ij}$  and vice-versa

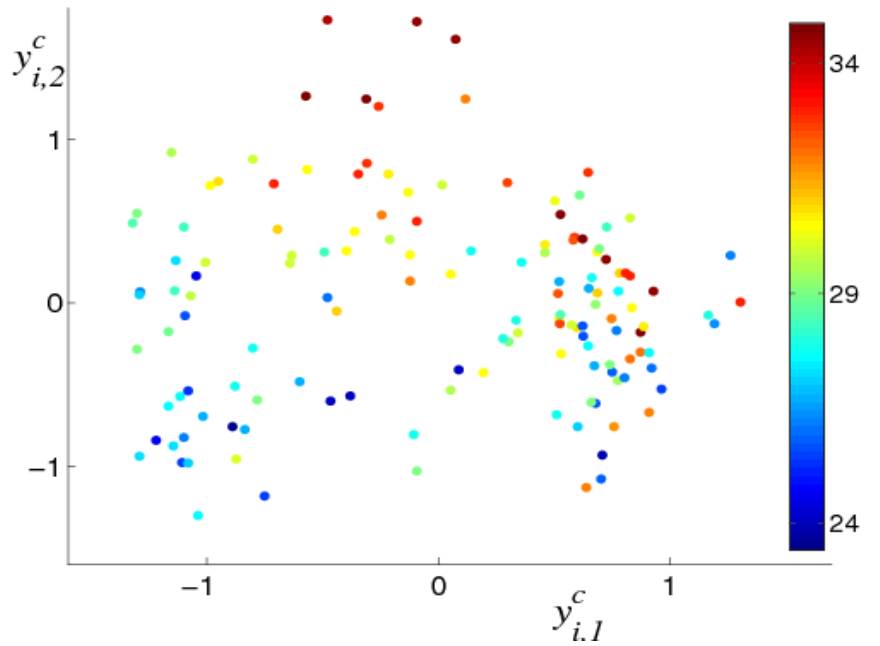
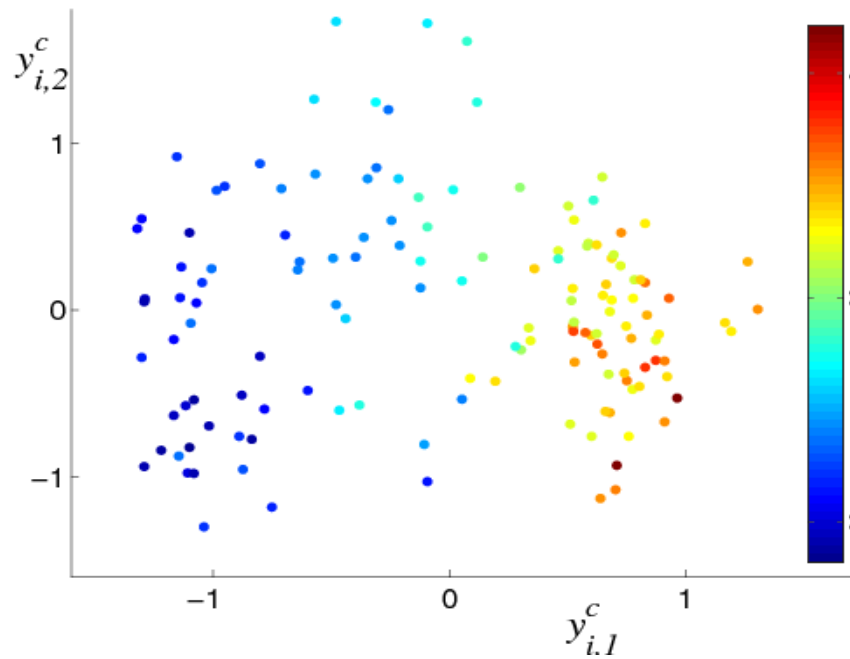
# How to measure similarities

- Application to neonatal data
- Multiple tailored measures
  - Shape and MR appearance



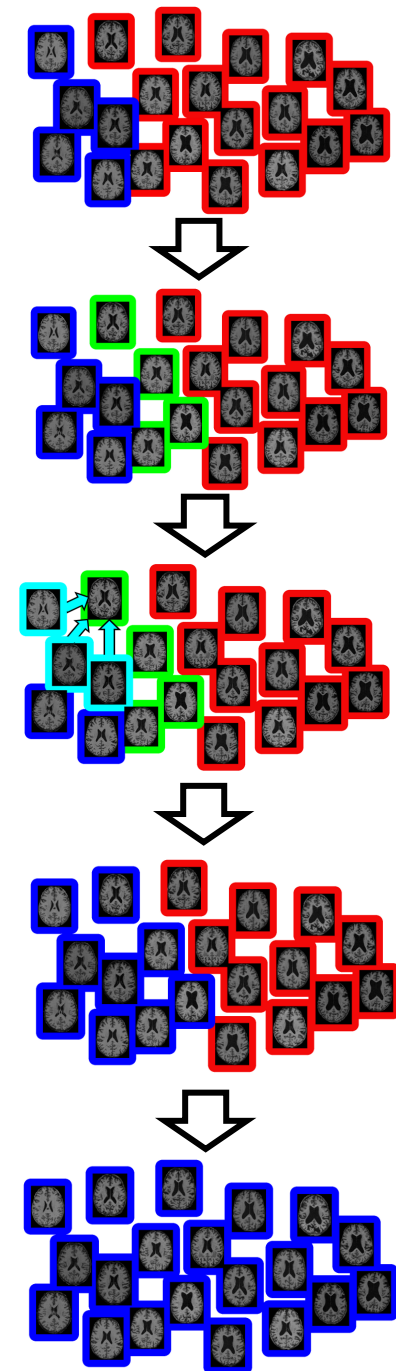


# Linking to infant data



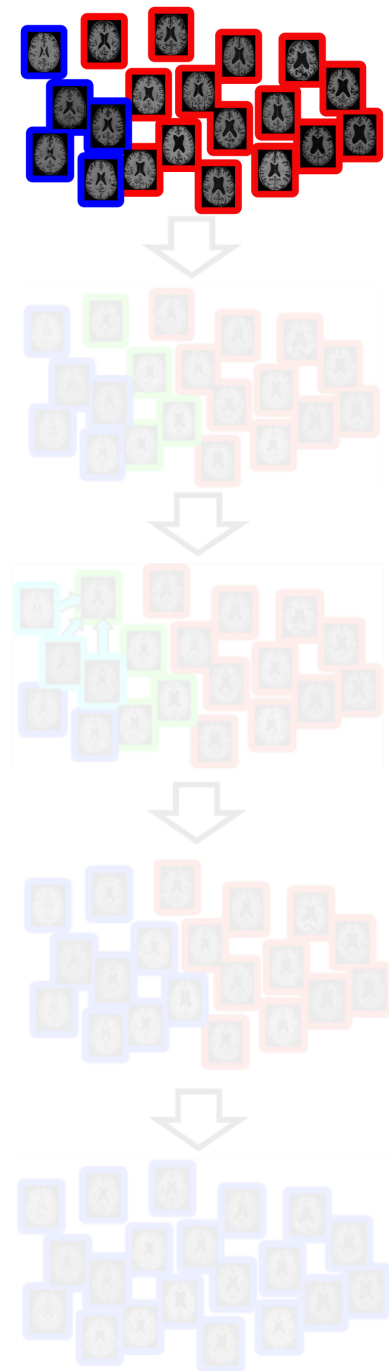
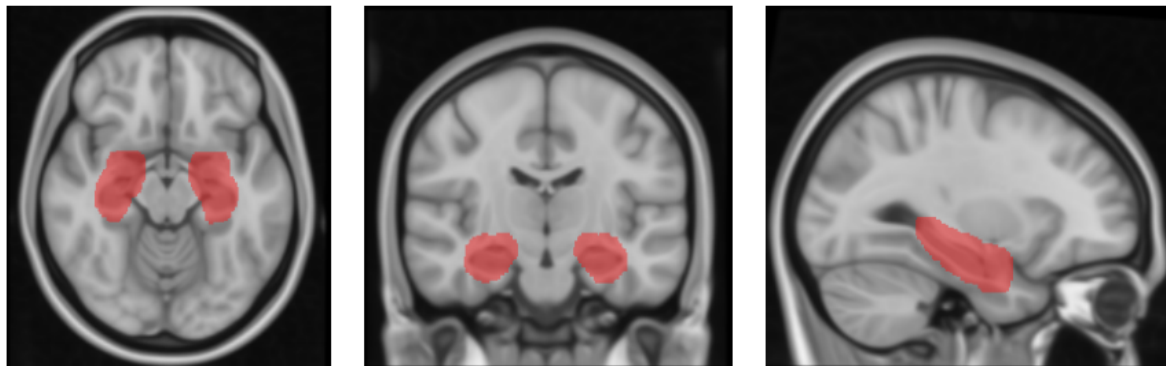
# LEAP

- LEAP aims at segmenting diverse image datasets by **L**earning **E**mbeddings for **A**tlas **P**ropagation
- Learns new representation for all images
- Neighbourhoods are defined by image similarities
- Initial small set of atlases is propagated throughout the data
- Atlases are propagated to ‘nearby’ images
- Labelled images are used as *bootstrapped* atlases thereafter



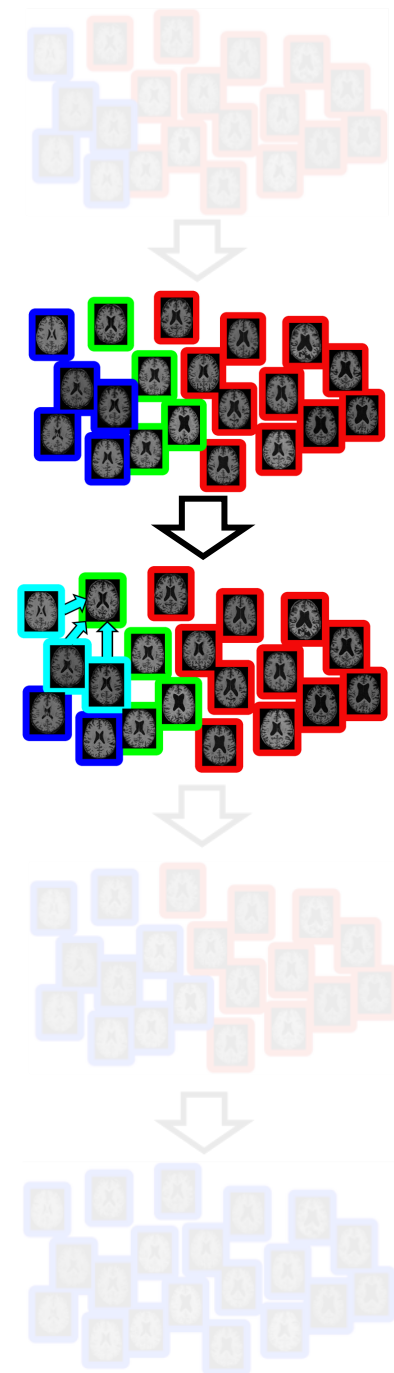
# Intensity-based similarities

- Here, we use intensity differences estimated in a template space
- All N images are registered to the MNI152-template
- The level of registration can be adapted to the size of the structure of interest
- Pair-wise similarities can be estimated over the whole brain or in a region of interest



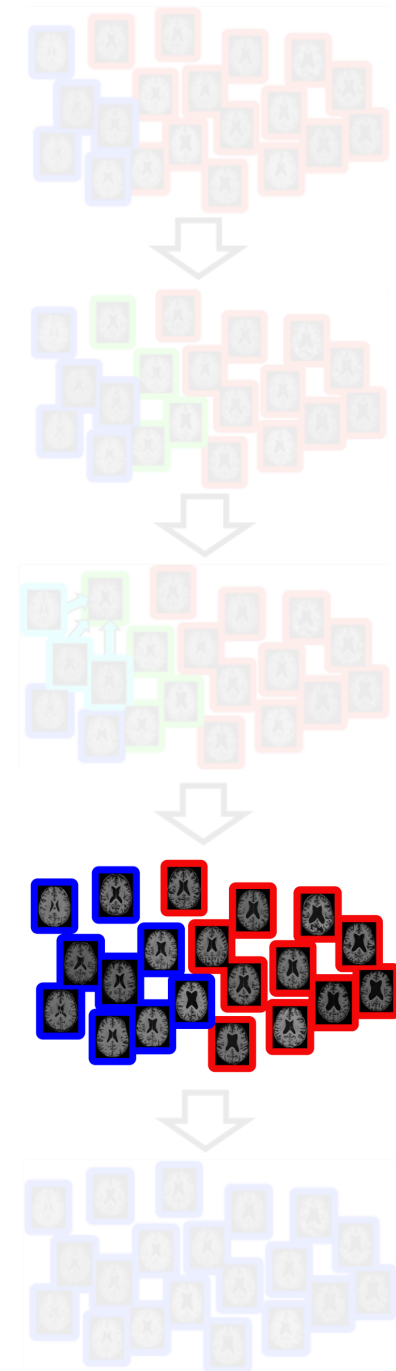
# LEAP propagation

- Distances in the learned manifold are used to identify atlas propagation steps
- The  $N$  unlabelled images that are closest to the set of labelled images are selected for segmentation
- For each selected images, the  $M$  closest labelled images are selected as atlases
- All selected atlas images are accurately registered to a target image



# LEAP propagation (2)

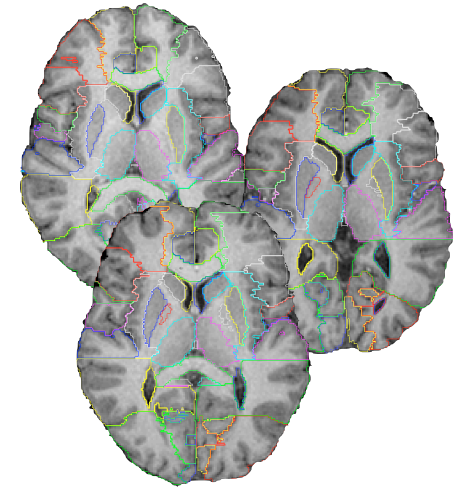
- A spatial prior is generated from multiple atlases
- An intensity model is estimated from the target image
- The target segmentation is estimated based on both models



# Application to the segmentation of ADNI

## Available set of atlases:

- 30 atlases from young, healthy subjects
- Manually delineated into 83 structures of interest

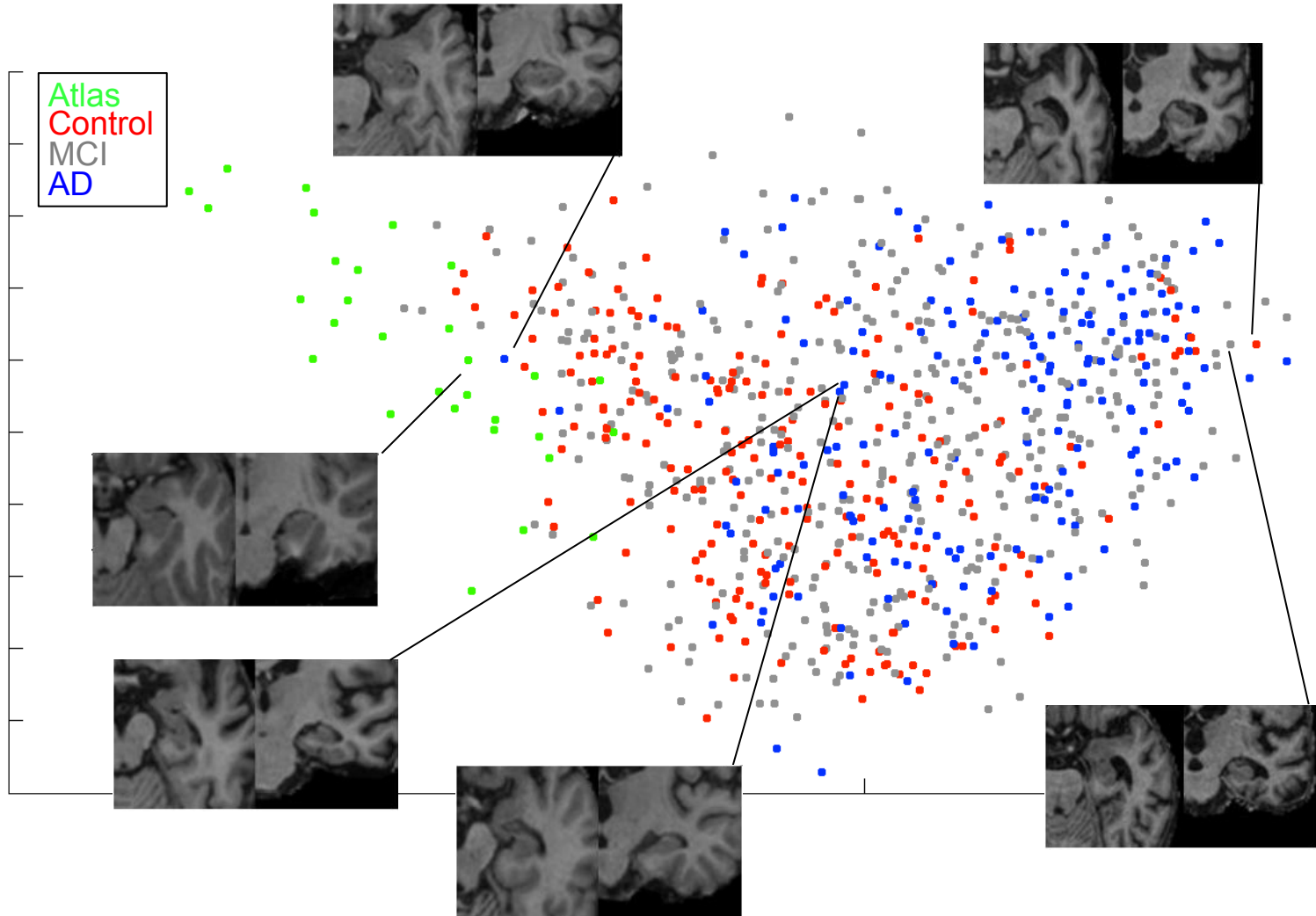


## ADNI dataset:

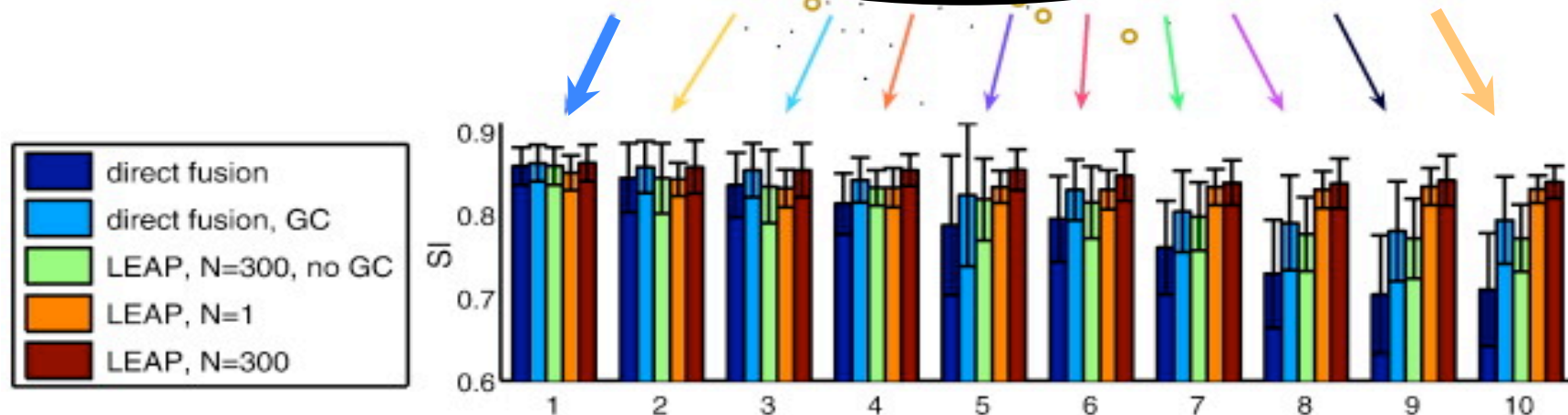
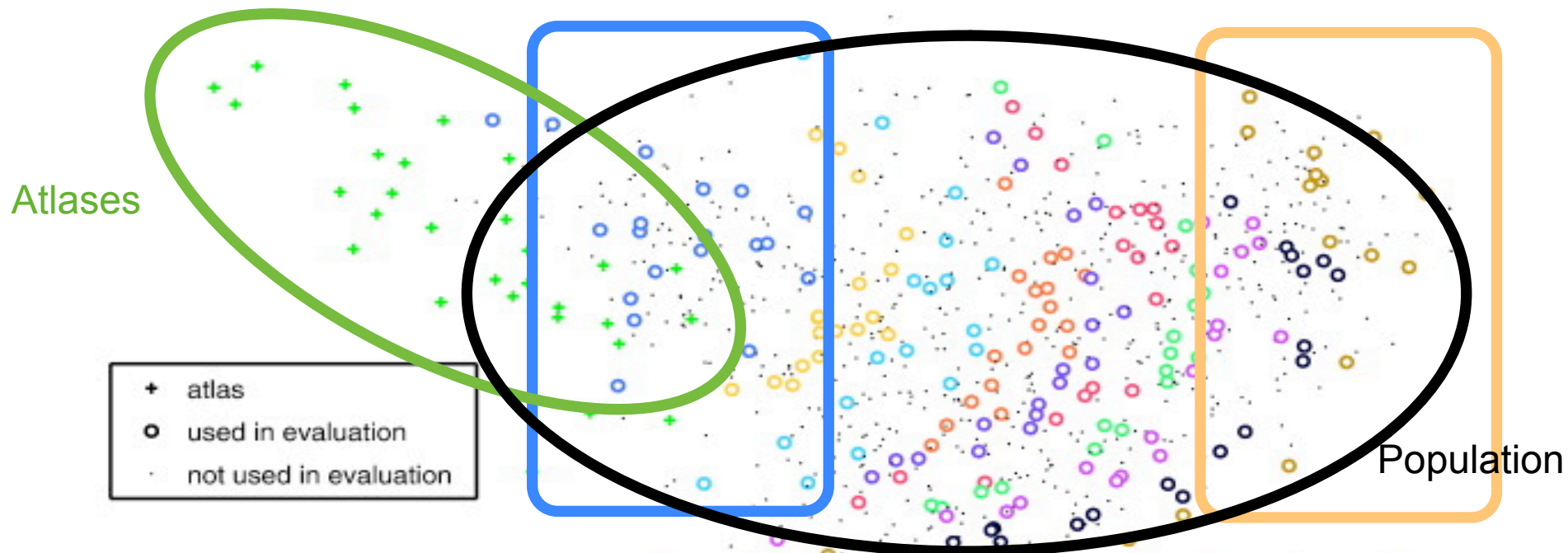
- 838 images from elderly subjects with dementia and age-matched healthy controls
- Strong pathology due to ageing and disease progression



# Hippocampal segmentation



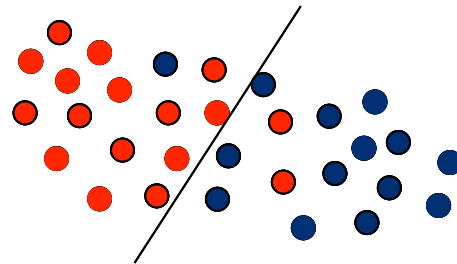
# Manifold learning for multi-atlas segmentation: Results





# Manifold Learning: classification

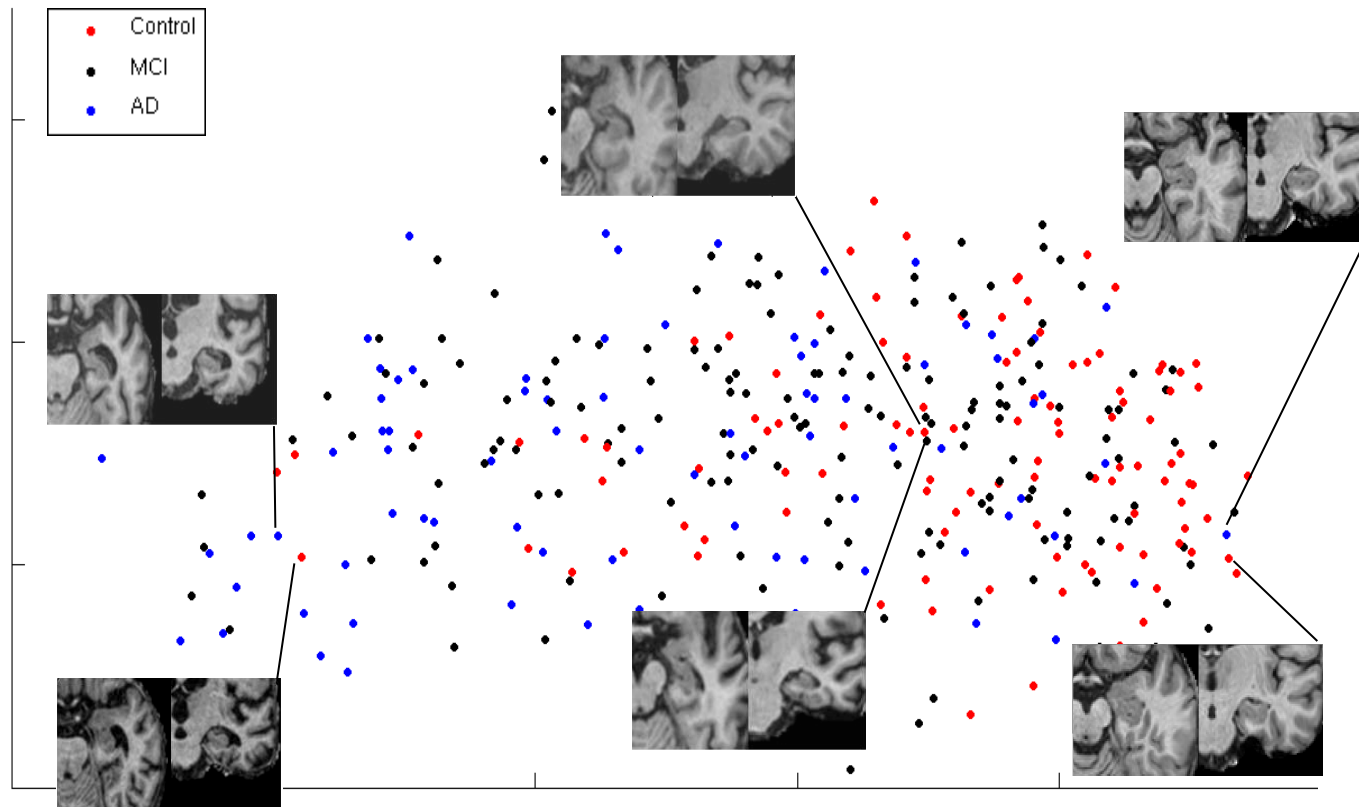
- Manifold coordinates can be directly used to extract information
- Assuming, a clinical label is available for a subset of images, manifold coordinates can be used to classify the unlabelled subjects



2D-embedding

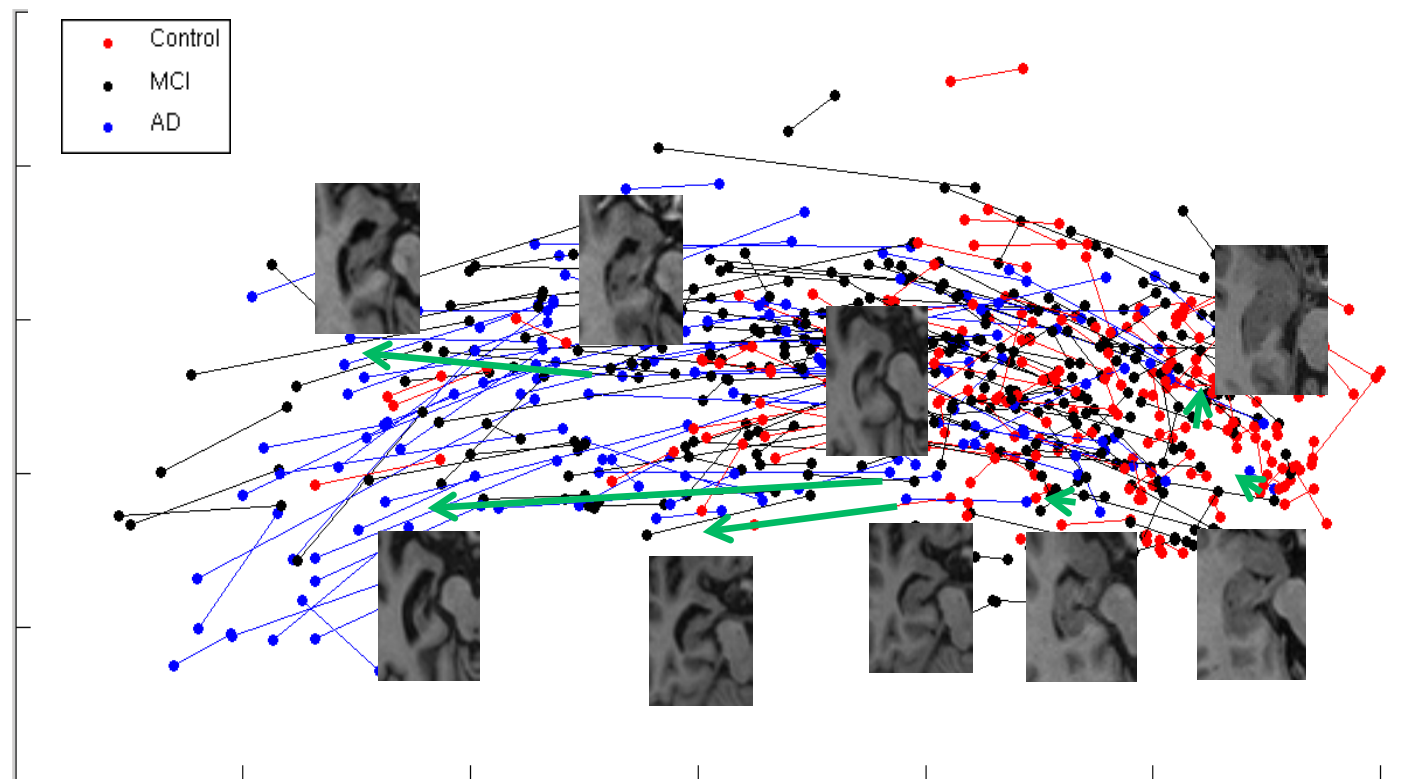
# Embedding of baseline images

- 2D embedding of baseline images
- principal axis resembles disease progression



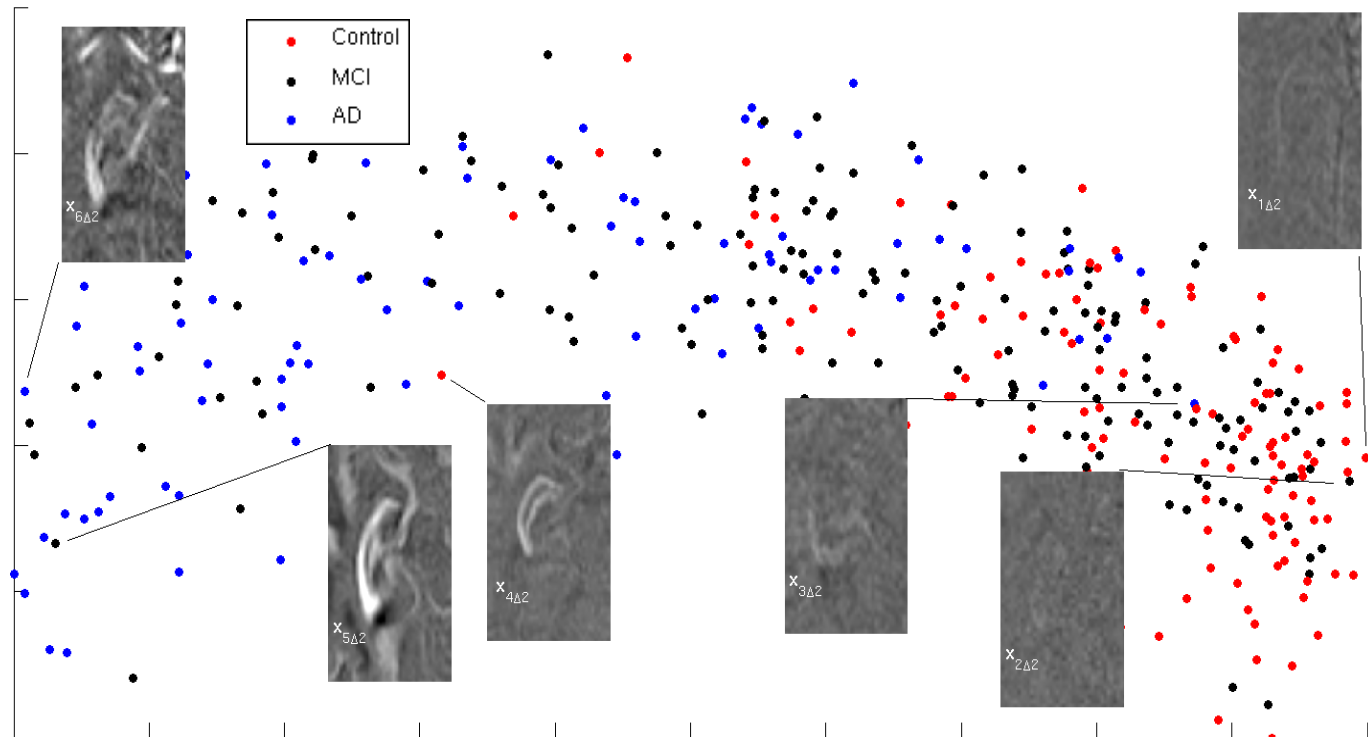
# Combined embedding

- Single manifold is learned from subjects at two timepoints
- Subjects “move” along principal axis
- More atrophied subjects move “faster”



# Embedding of intra-subject variation

- Image similarities are based on difference images between baseline and follow-up scans
- Features can be combined with embedding of baseline scans



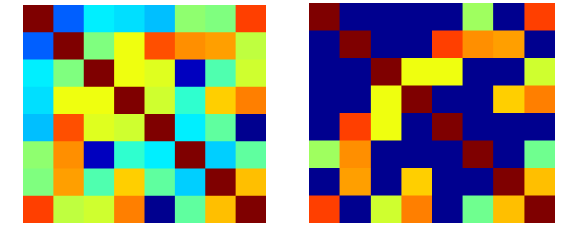
Wolz et al, MICCAI MLMI 2010

# Laplacian Eigenmaps

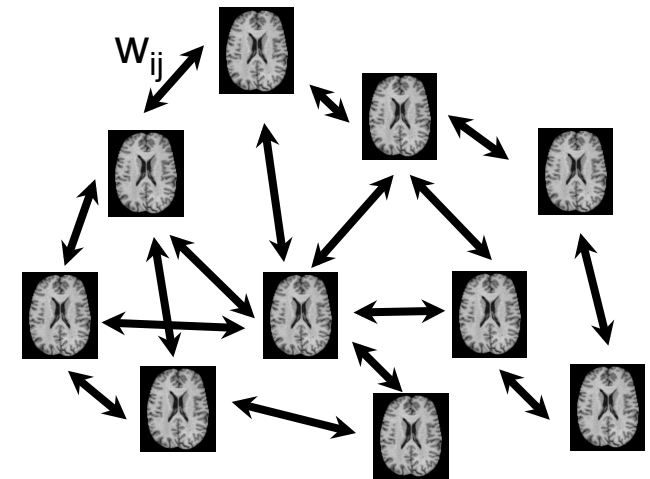
- All images are represented in a  $k$ -nn graph
- Every subject is connected to its  $n$  closest neighbours
- Edge weights  $w_{ij}$  are defined by image similarities and form a weight matrix  $W$
- Subjects that are similar in input space are close in manifold space with the objective function

$$\sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 w_{ij}$$

- Defining the graph Laplacian from the weight matrix  $W$  allows a closed form solution [1]



Full similarity matrix     $k$ -nn similarity matrix

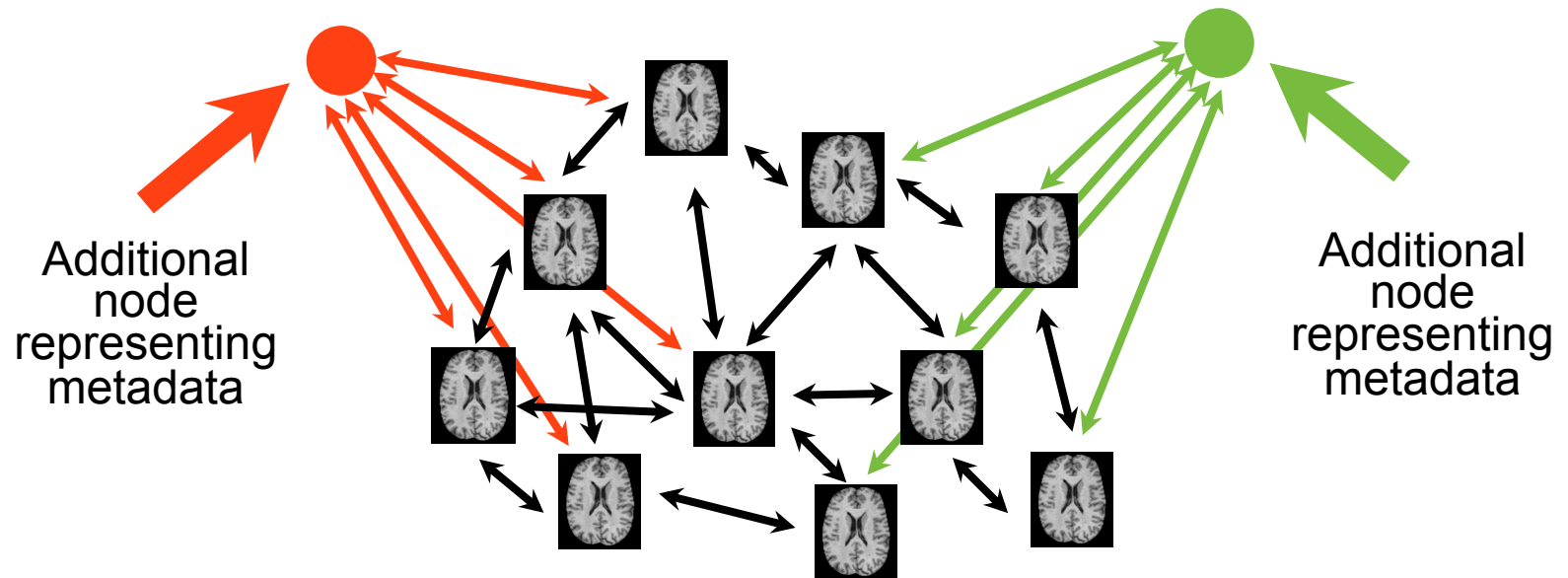


$k$ -nn neighbourhood graph

[1] Belkin and Niyogi, 2003, Neur. Comp.

# Extended similarity graph

- Laplacian eigenmaps only considers image similarities
- Subject metadata (e.g. age, genotype) gives additional information to compare subjects
- An extension of the similarity graph by additional nodes allows to consider such information

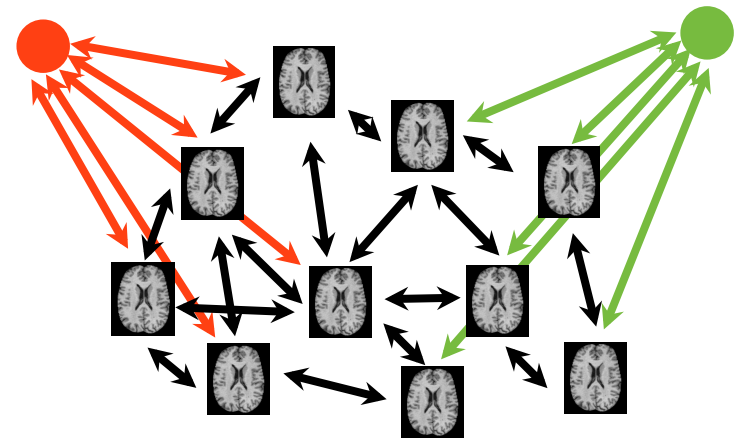


# Extended objective function

- In the extended similarity graph,  $M$  additional nodes represent  $M$  groups of metadata
- Weights  $\hat{w}_{im}$  can be defined discrete or continuously
- An extended objective function can be defined

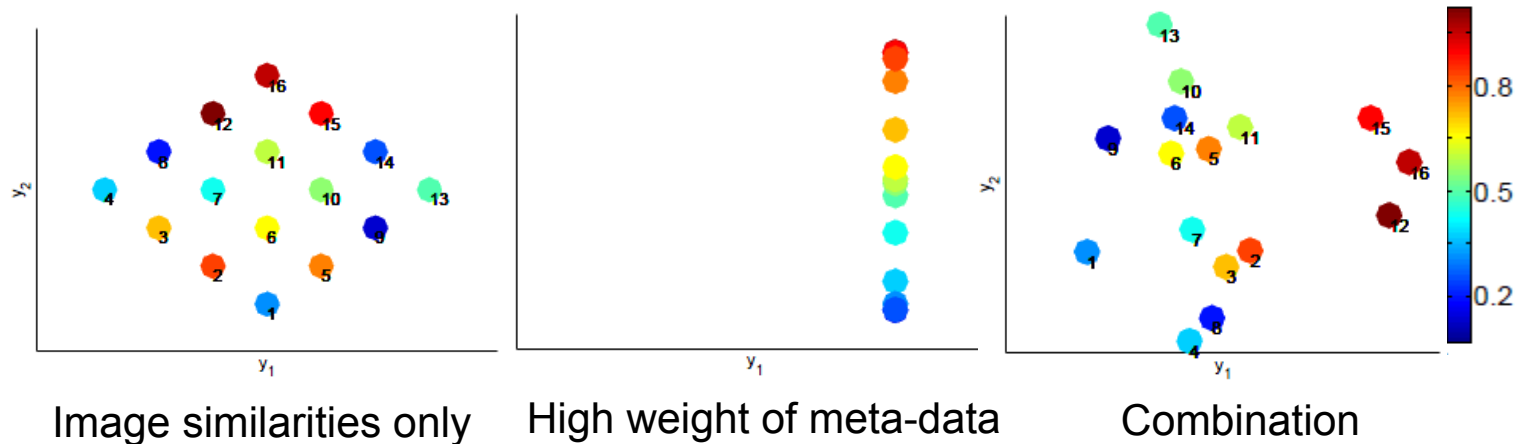
$$\gamma \sum_{ij} (\mathbf{y}_i - \mathbf{y}_j)^2 w_{ij} + \sum_{im} (\mathbf{y}_i - \hat{\mathbf{y}}_m)^2 \hat{w}_{im}$$

- Subjects with similar metadata values are clustered in embedding space
- $\gamma$  defines the influence of metadata on the final embedding



# Illustrative example

- Every node has some meta-information with a value between 0 and 1
- Three additional nodes are introduced in the similarity graph and weights to every image are defined by the metadata
- Changing the influence of the meta-information leads to different embedding results





# Image data and meta-information

- ADNI baseline images were used for evaluation of the method
- Used non-imaging metadata:
  - CSF concentration of beta amyloid A $\beta$ -42 (continuous)
  - APOE-genotype (discrete)
- Derived imaging metadata:
  - Hippocampal volume (continuous)
- The 420 subjects for which the CSF biomarker was available were used:

	N (F)	MMSE	A $\beta$ -42	$\epsilon$ 2/ $\epsilon$ 4 carriers	Hippo. Vol.
CN	116 (56)	29.1+/-1.0	202+/-58	16/28	4.53+/-0.55
S-MCI	112 (36)	27.2+/-1.8	179+/-62	9/49	4.26+/-0.59
P-MCI	89 (33)	26.6+/-1.8	146+/-46	1/52	3.93+/-0.65
AD	83 (44)	23.6+/-1.9	148+/-46	4/63	3.92+/-0.73

# Composite similarity measure

- Pairwise image similarities are based on a combined similarity measure incorporating deformation energy and intensity differences

$$\mathcal{D}_{ij} = \alpha \mathcal{D}_{ij}^{\text{def}} (I_i, I_j, \phi_{ij}) + (1 - \alpha) \mathcal{D}_{ij}^{\text{int}} (I_i, I_j, \phi_{ij})$$

- Deformation energy is based on the deformation magnitude resulting from registering two images

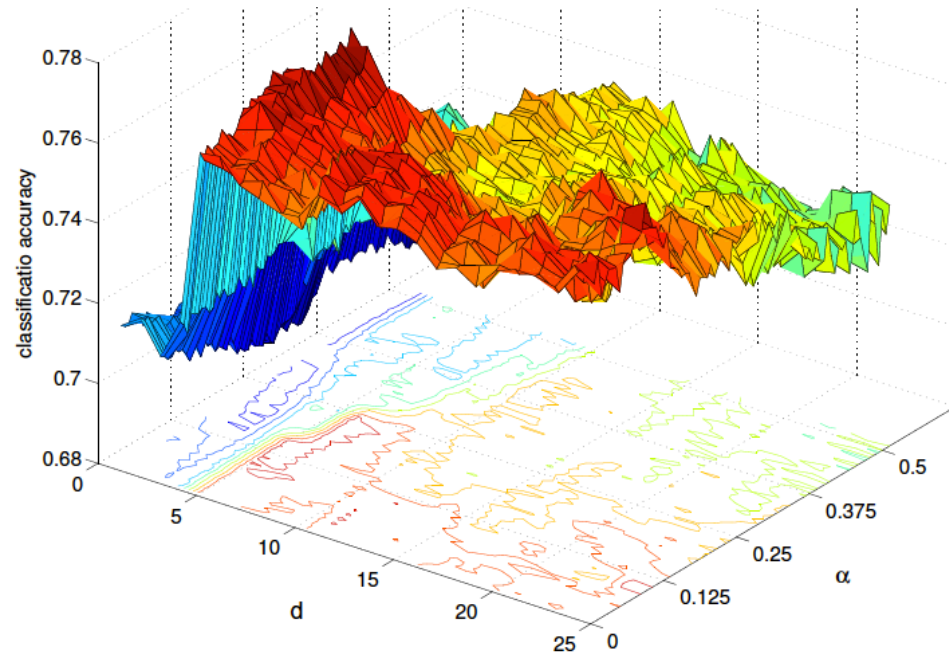
$$\mathcal{D}_{ij}^{\text{def}} (I_i, I_j, \phi_{ij}) = \sum_{x \in \Omega} \|u_{ij}(x)\|_2^2 dx$$

- Sums of squared intensity differences are used to represent the residual difference

$$\mathcal{D}_{ij}^{\text{int}} (I_i, I_j, \phi_{ij}) = \sum_{x \in \Omega} (I_i(x) - I_j(\phi_{ij}(x)))^2$$

# Parameter setting

- The weighting factor  $\gamma$  defines the influence of image similarities and metadata
- Classification results on a training data set show a good performance of the similarity-based measure
- Using a 5-10 dimensional manifold leads to stable classification results



# Classification results

- Manifold coordinates are corrected for age
- 1,000 leave-25%-out runs are performed to obtain classification rates

	AD vs CN	P-MCI vs S-MCI	P-MCI vs CN
Laplacian Eigenmaps	86%	63%	82%
& ApoE	83%	69%	81%
& A $\beta$ -42	87%	68%	84%
& Hippo. Vol.	86%	66%	83%
& A $\beta$ -42 / Hippo. Vol.	88%	67%	87%
& A $\beta$ -42 / Hippo. Vol. / ApoE	88%	69%	87%

Classification accuracy using manifold learning

# Conclusions

- Manifold learning allows to model the characteristics of a large population of brain images
- In LEAP, the defined metric space is used to propagate a set of manually labelled atlas images in several steps through the whole manifold
- An improved segmentation and classification accuracy shows the benefit of the manifold-based approach
- Manifold coordinates can be directly used to infer from subjects with a clinical label to unlabelled subjects
- An approach to incorporate metadata into Laplacian eigenmaps was described