# Manifold Alignment of High-Dimensional Datasets

Sridhar Mahadevan (PI) & Rui Wang (co-PI)
Thomas Boucher, Clifton Carey, Stefan Dernbach,
Blake Foster, Hoa Vu, Chang Wang (IBM)

School of Computer Science
University of Massachusetts, Amherst

# Learning from Multiple Datasets

- In many applications, multiple “views” or multiple datasets are constructed
  - Bioinformatics
  - Activity recognition
  - Computer graphics
  - Scientific exploration (MARS rover)
  - Cross-lingual information retrieval
  - Spectral methods for learning latent variable models

# Canonical Correlation Analysis

(Hotelling, 1936)

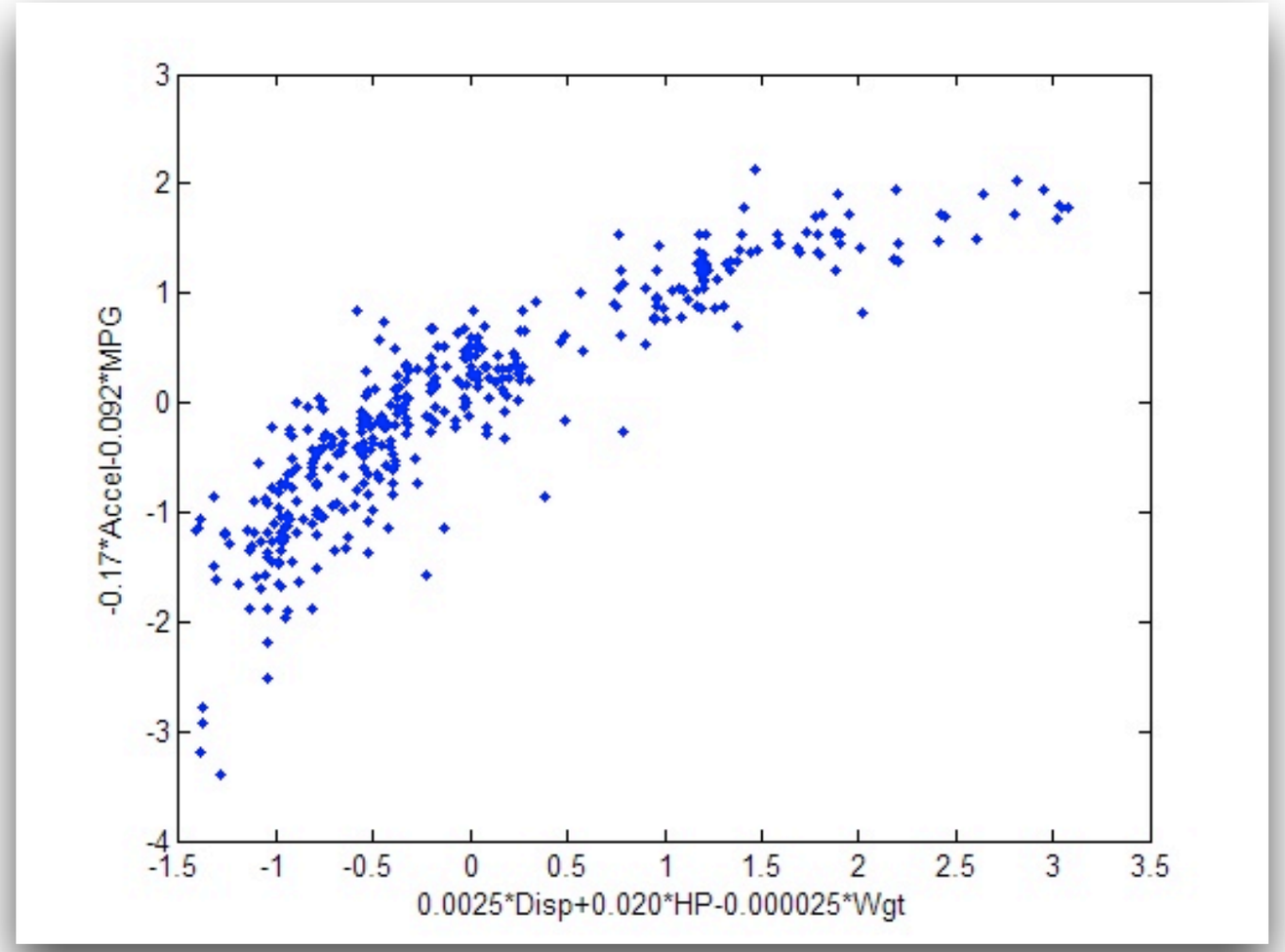
Acceleration MPG

Displacement Horsepower Weight

# Canonical Correlation Analysis

(Hotelling, 1936)

Acceleration MPG

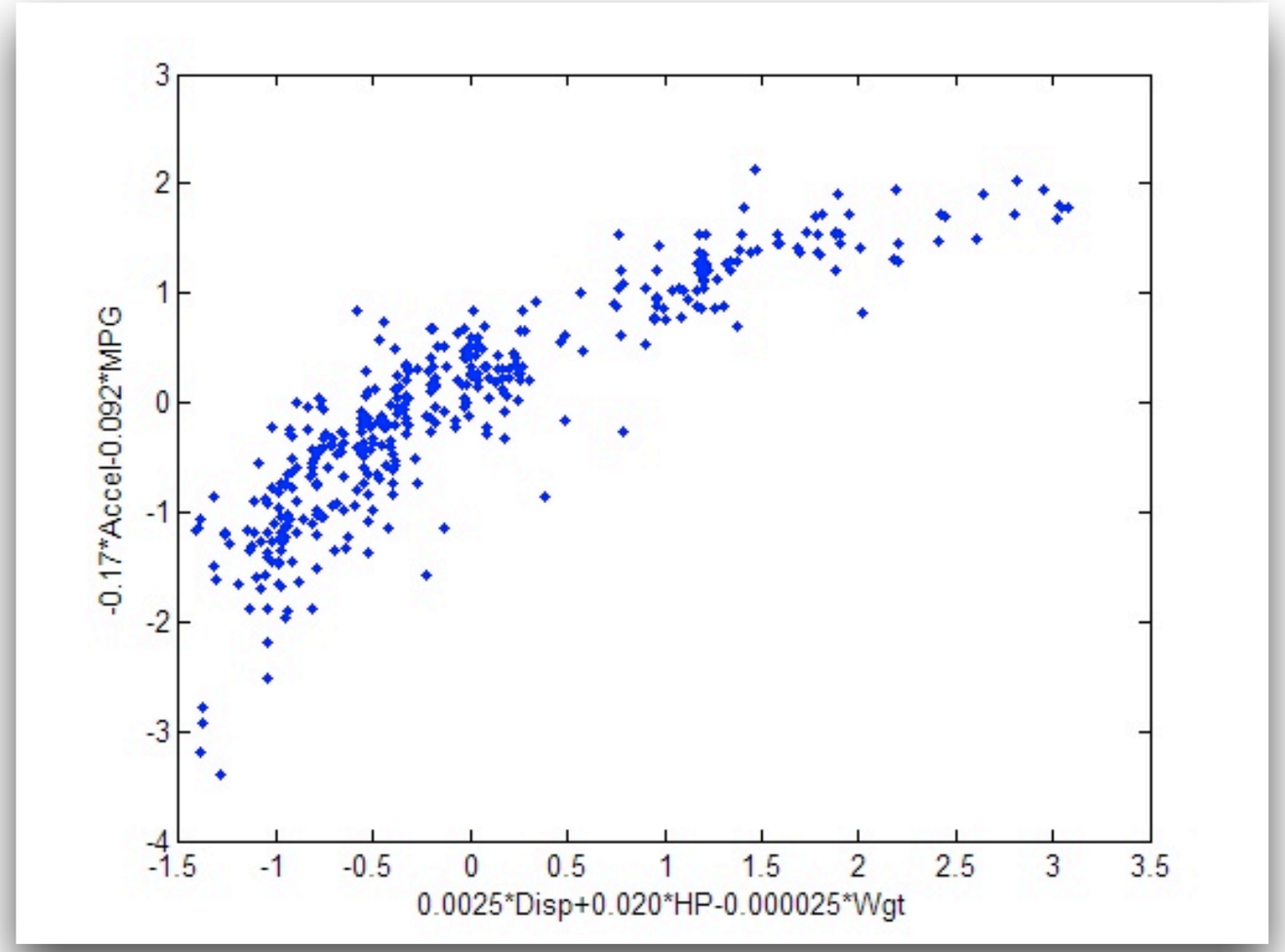



Displacement Horsepower Weight

# Canonical Correlation Analysis

(Hotelling, 1936)

Acceleration MPG



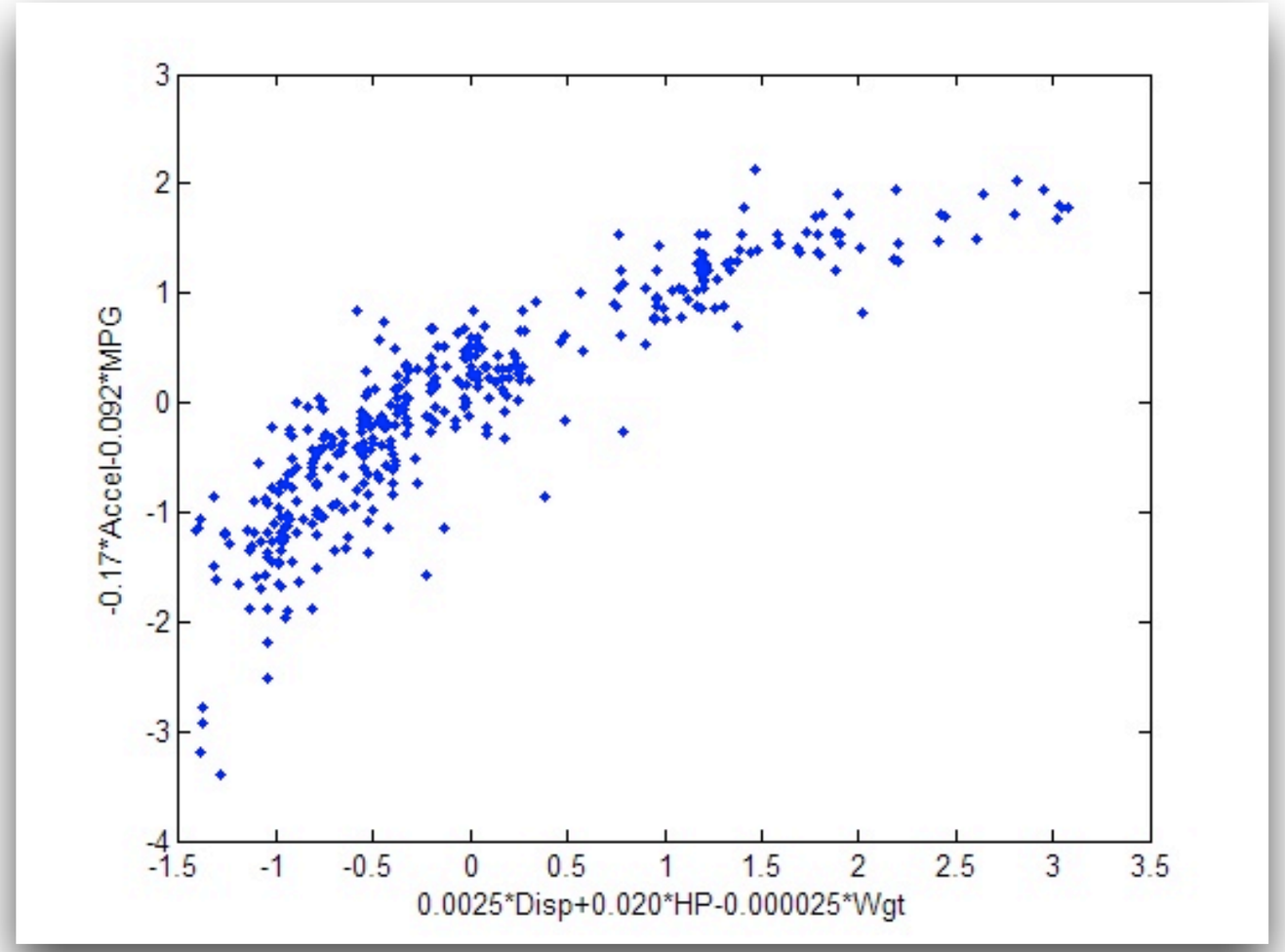

Displacement Horsepower Weight

$$\frac{u^T X^T Y v}{\sqrt{u^T X^T X u}\sqrt{v^T Y^T Y v}}$$

# Canonical Correlation Analysis

(Hotelling, 1936)

Acceleration MPG



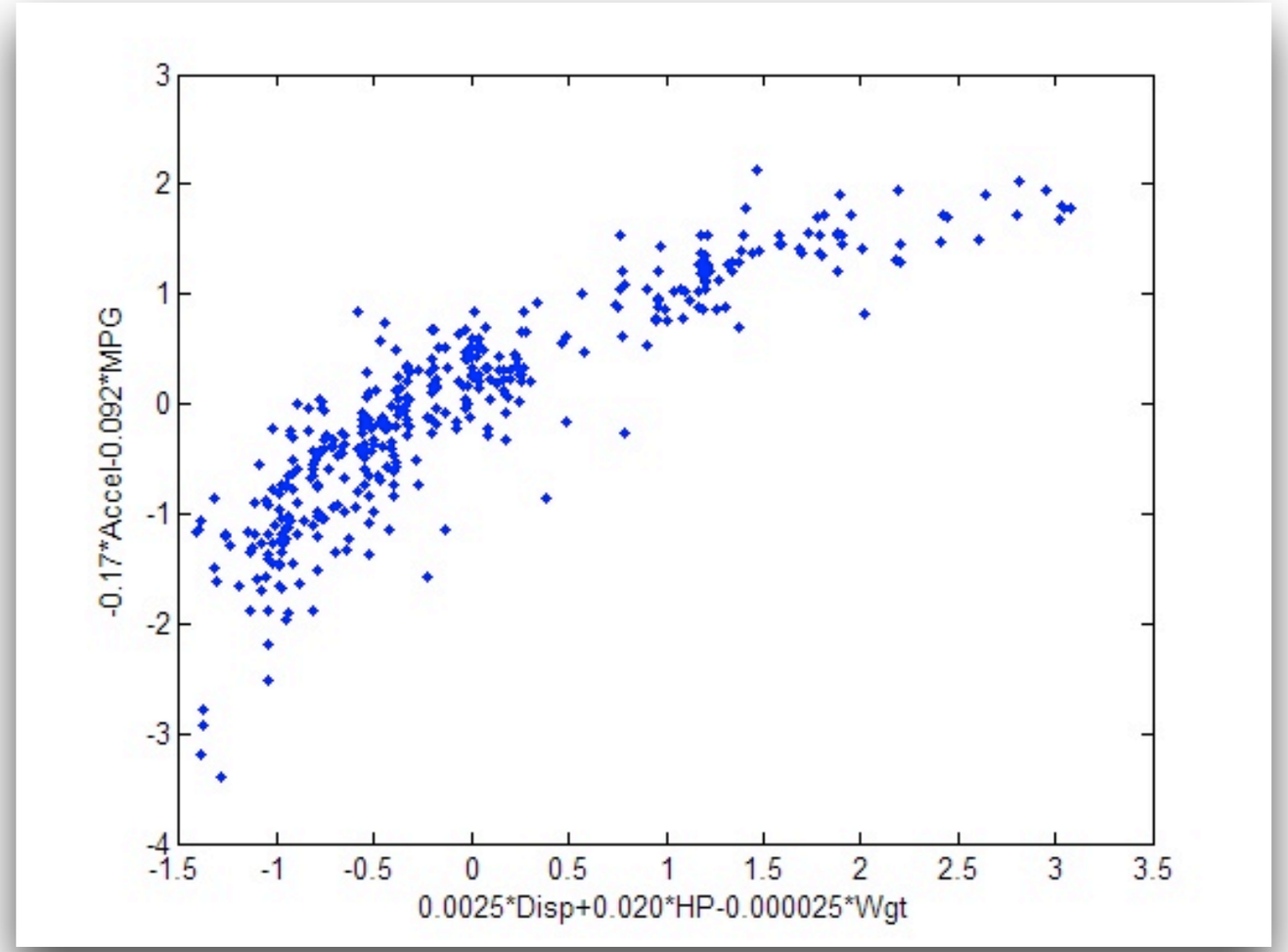

Displacement Horsepower Weight

Find u,v that maximizes $\frac{u^T X^T Y v}{\sqrt{u^T X^T X u}\sqrt{v^T Y^T Y v}}$

# Canonical Correlation Analysis

(Hotelling, 1936)

Acceleration MPG



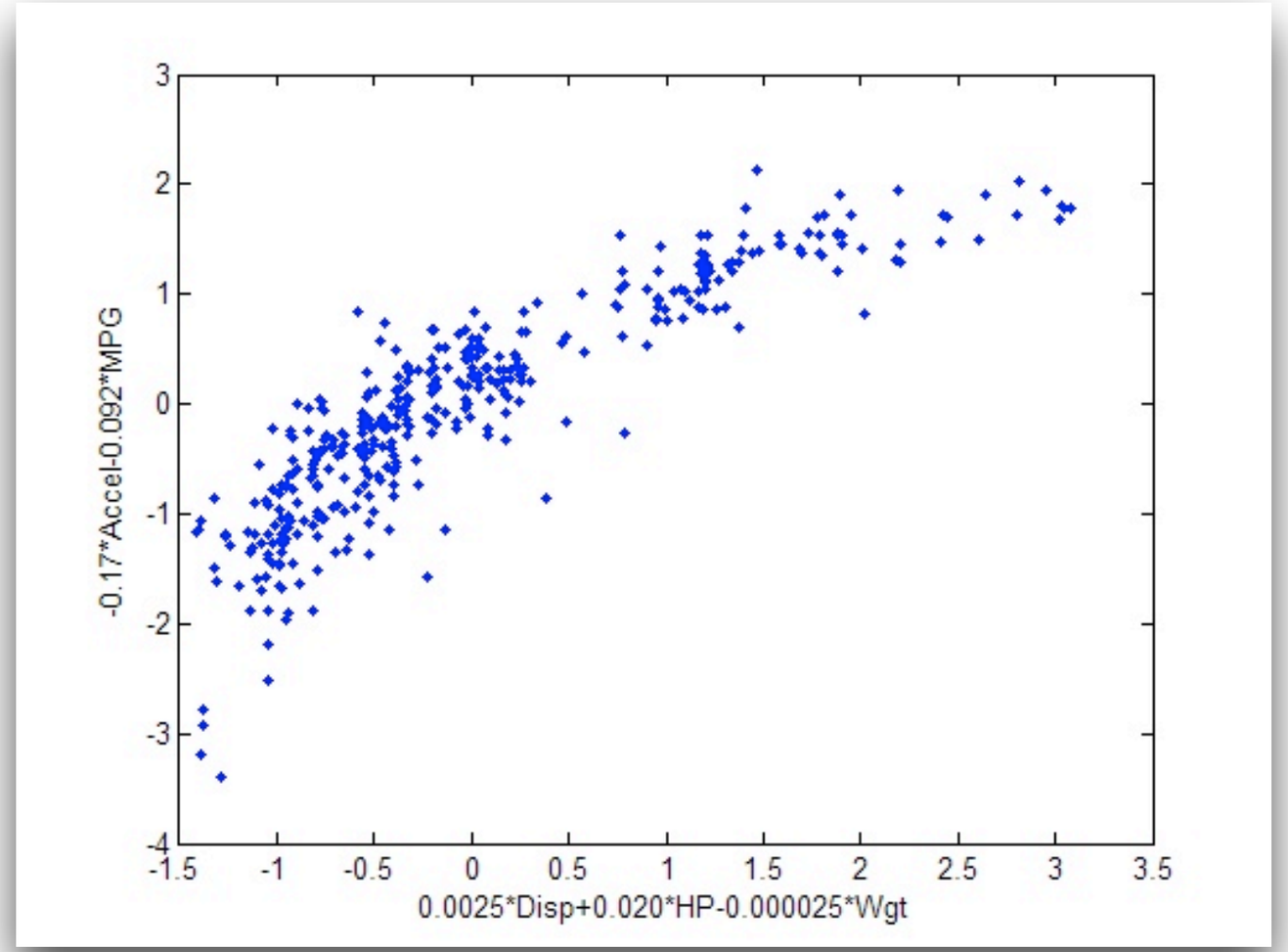

Displacement Horsepower Weight

Find u,v that maximizes $$\frac{u^T X^T Y v}{\sqrt{u^T X^T X u}\sqrt{v^T Y^T Y v}}$$

# Canonical Correlation Analysis

(Hotelling, 1936)

Acceleration MPG




Displacement Horsepower Weight

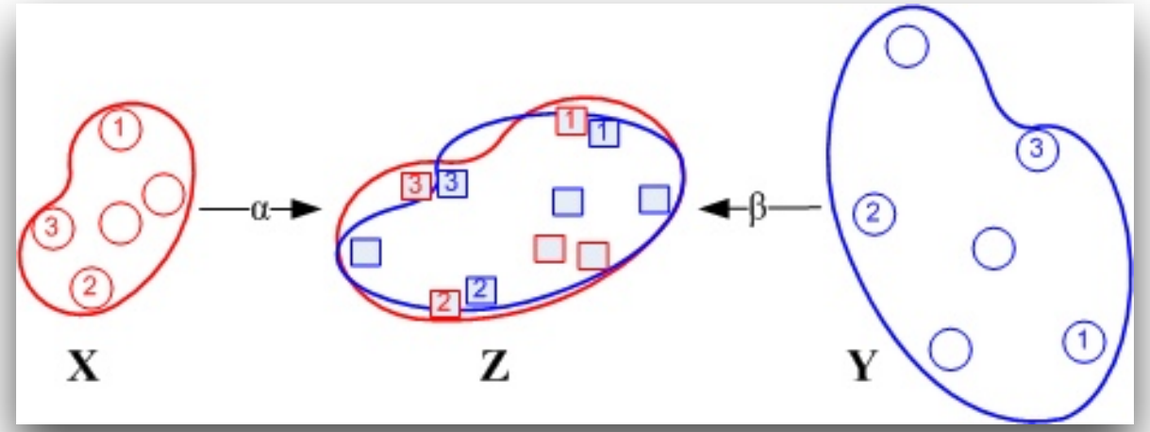
Pioneer of the first two statistics departments in the US!
UNC, Chapel Hill
Columbia University

Find u,v that maximizes $\frac{u^T X^T Y v}{\sqrt{u^T X^T X u}\sqrt{v^T Y^T Y v}}$

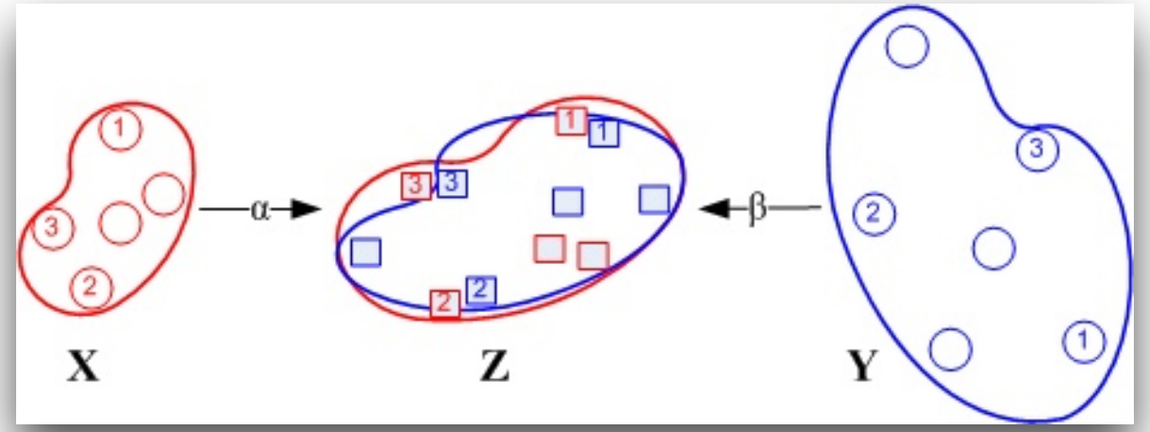
# FODAVA project: main contribution

- We developed a new class of methods, called manifold alignment, that outperforms CCA in many domains
  - Linear + Nonlinear
  - Local + Global
  - Supervised + Unsupervised
- If you use multiple datasets, you should try manifold alignment!

# Manifold Projections



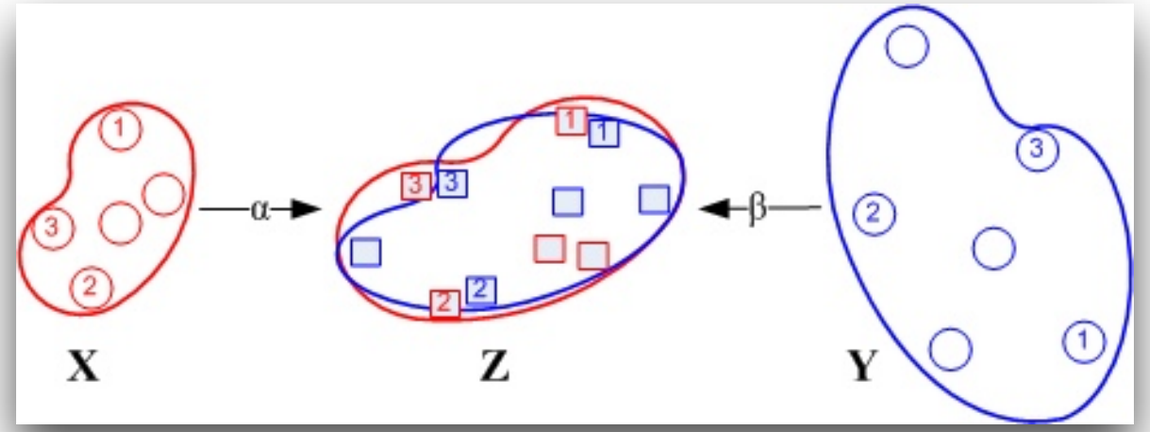
# Manifold Projections




We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

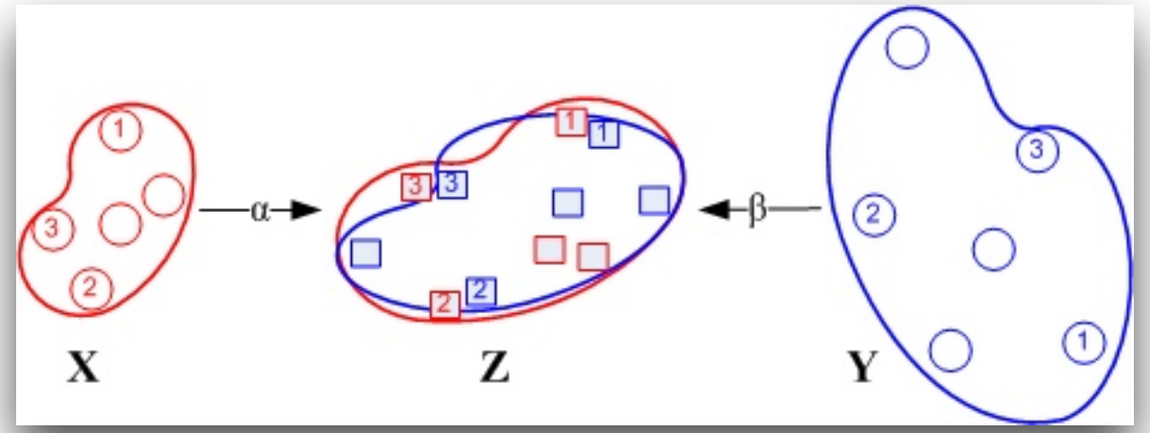
# Manifold Projections




We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

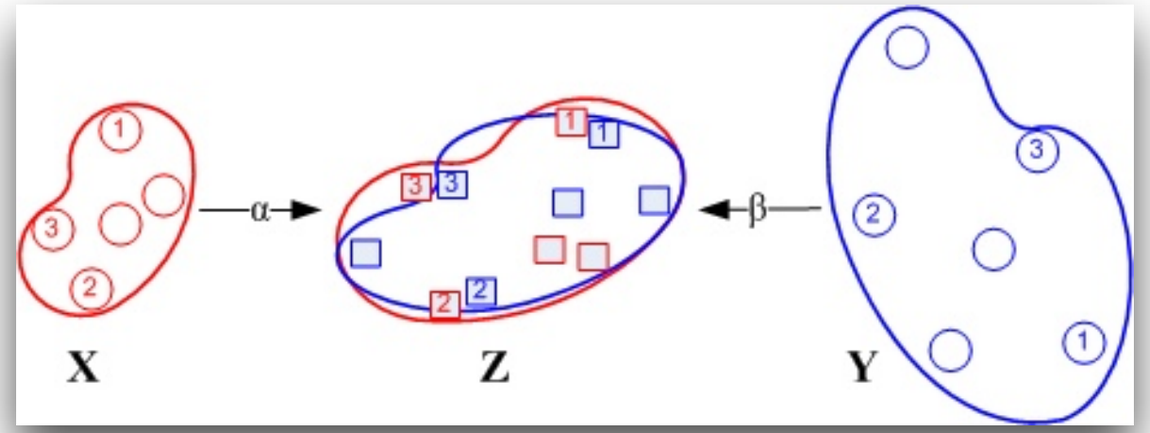
# Manifold Projections




We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

# Manifold Projections

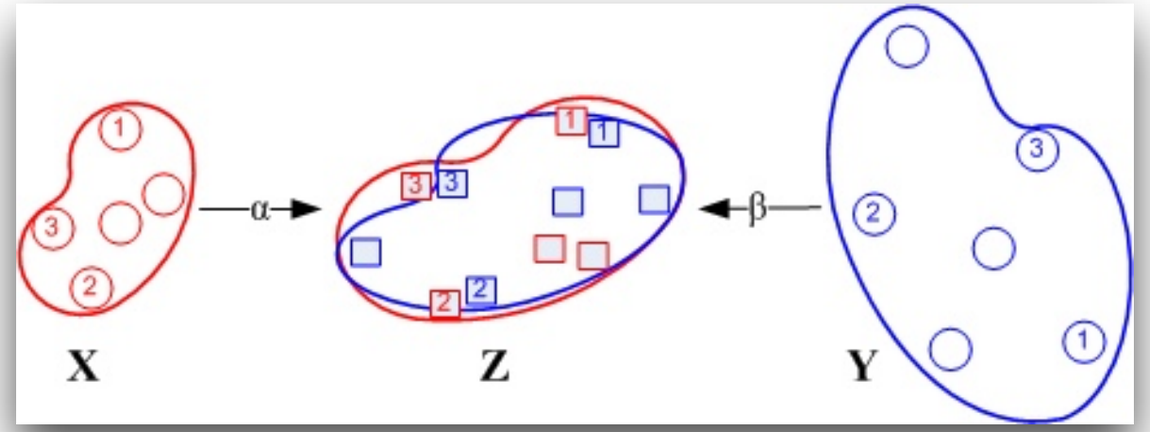



We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

Preserve correspondences

# Manifold Projections

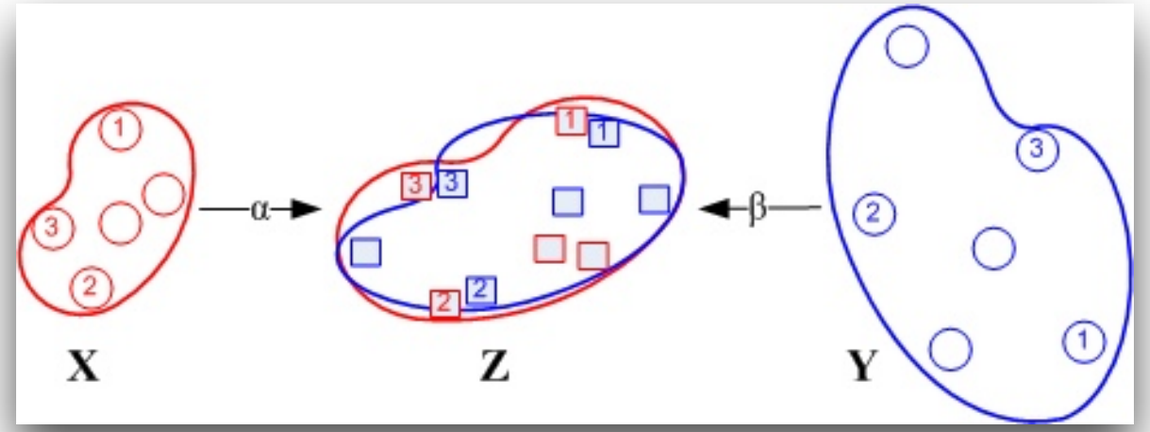



We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

Preserve correspondences

# Manifold Projections



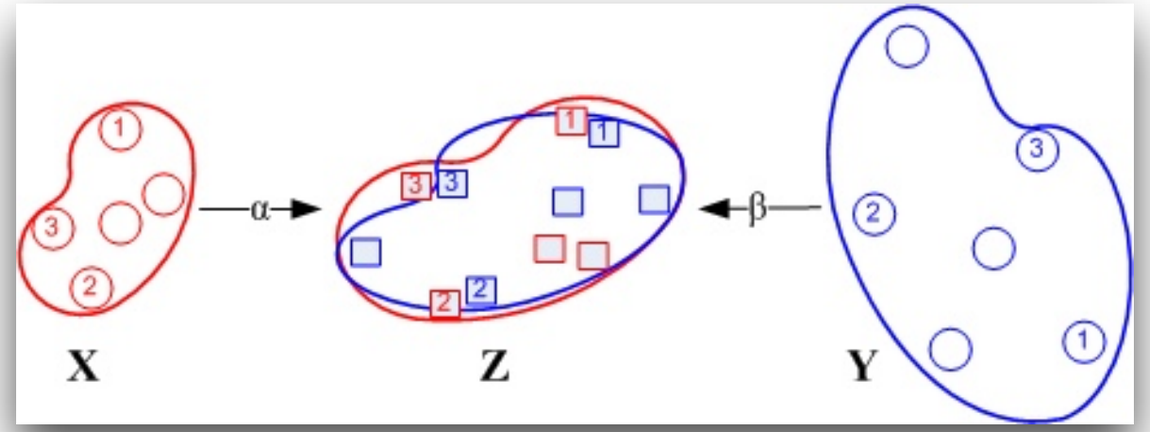

We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

Preserve correspondences

Preserve local geometry

# Manifold Projections




We want to find mapping functions $\alpha, \beta$ to minimize the cost function $C(\alpha, \beta)$, where

$$C(\alpha,\beta) = \mu \sum_i \sum_j (\alpha^T x_i - \beta^T y_j)^2 W^{i,j} + 0.5 \sum_{i,j} (\alpha^T x_i - \alpha^T x_j)^2 W_x^{i,j} + 0.5 \sum_{i,j} (\beta^T y_i - \beta^T y_j)^2 W_y^{i,j}$$

Preserve correspondences

Preserve local geometry

***(2) Theorem 1 :*** $\alpha, \beta$ ***to minimize*** $C(\alpha, \beta)$ ***are given by the eigenvectors corresponding to the smallest eigenvalues of*** $ZLZ^T\gamma = \lambda ZDZ^T\gamma$.

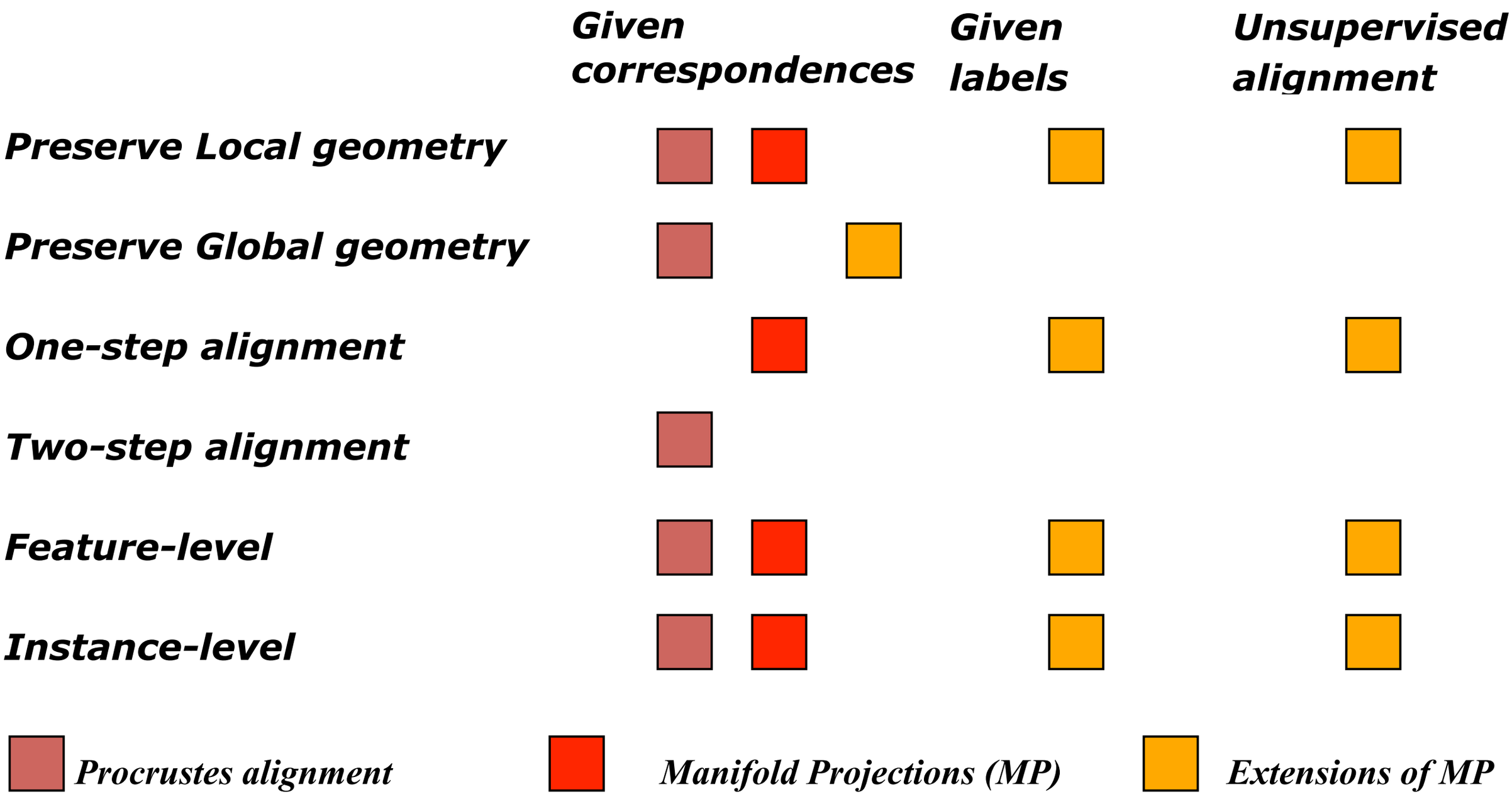
$D_x$ is a diagonal matrix: $D_x^{ii} = \sum_j W_x^{ij}$.
$L_x = D_x - W_x$.
$D_y$ is a diagonal matrix: $D_y^{ii} = \sum_j W_y^{ij}$.
$L_y = D_y - W_y$.
$\Omega_1$ is an $m \times m$ diagonal matrix, and $\Omega_1^{ii} = \sum_j W^{i,j}$.
$\Omega_2$ is an $m \times n$ matrix, and $\Omega_2^{i,j} = W^{i,j}$.
$\Omega_3$ is an $n \times m$ matrix, and $\Omega_3^{i,j} = W^{j,i}$.
$\Omega_4$ is an $n \times n$ diagonal matrix, and $\Omega_4^{ii} = \sum_j W^{j,i}$.

$$Z = \begin{pmatrix} X & 0 \\ 0 & Y \end{pmatrix}.$$

$$D = \begin{pmatrix} D_x & 0 \\ 0 & D_y \end{pmatrix}.$$

$$L = \begin{pmatrix} L_x + \mu\Omega_1 & -\mu\Omega_2 \\ -\mu\Omega_3 & L_y + \mu\Omega_4 \end{pmatrix}.$$

# A Summary of Manifold Alignment Approaches

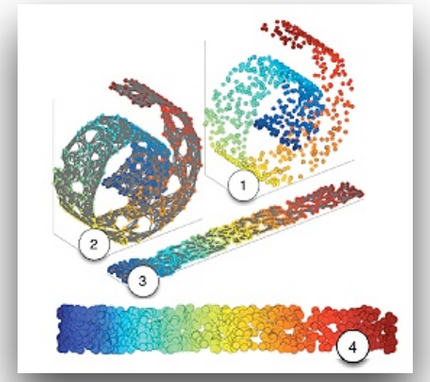


# Manifold Warping

(Hoa, Carey, Mahadevan: AAAI, 2012)

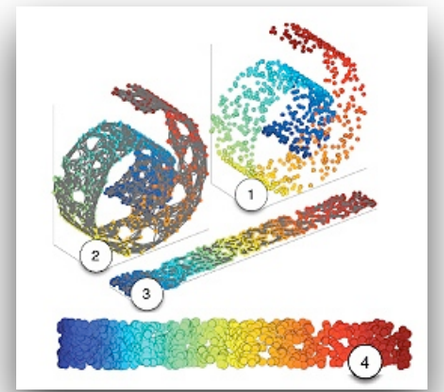
Dynamic Time Warping

+



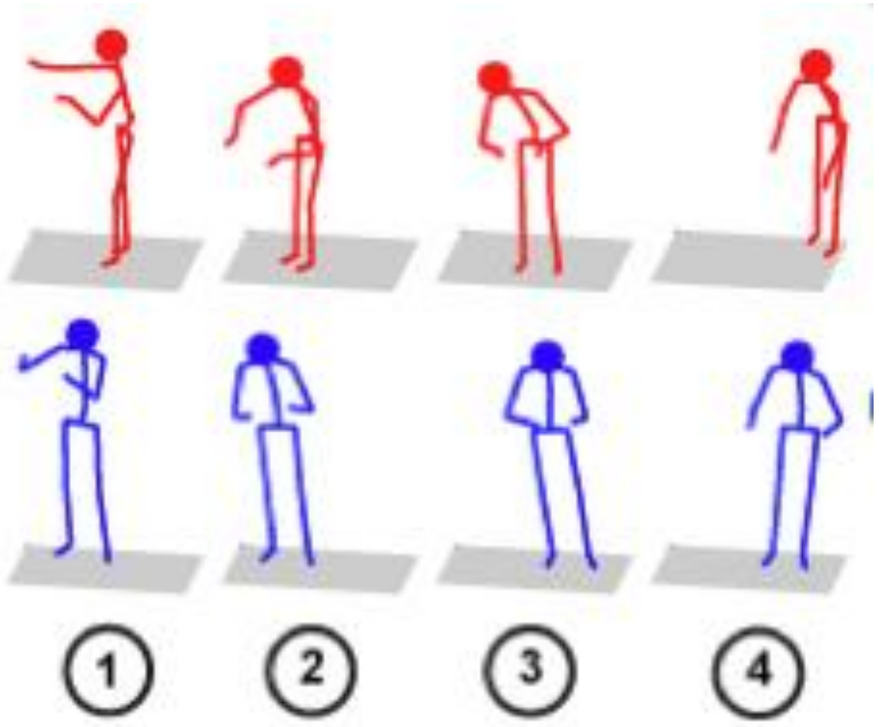

Iterate:

- Find projection to lower-dimensional space
- Find new set of correspondences

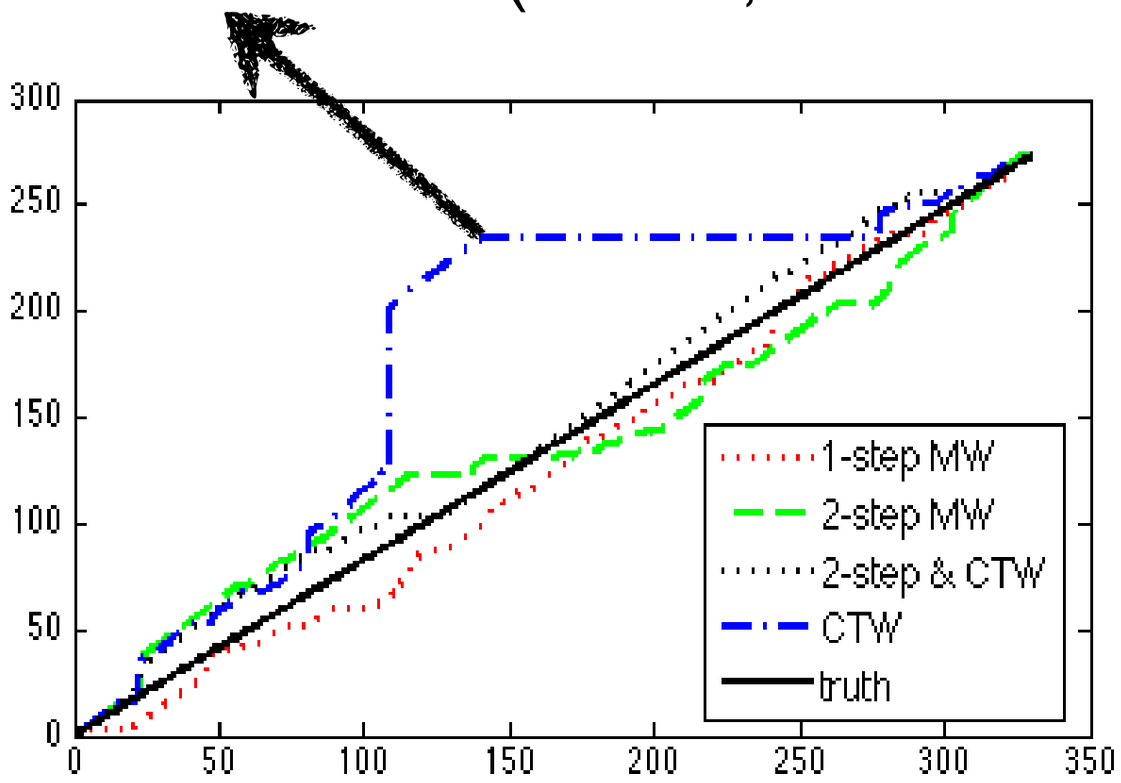



Manifold Alignment

# Activity Recognition




CCA+DTW (Zhou, NIPS 2009)




The resulted alignment path of manifold warping is much closer to the ground truth alignment

Vu, Carey, and Mahadevan, AAAI 2012

# Social Network Alignment

## Sparse Manifold Alignment

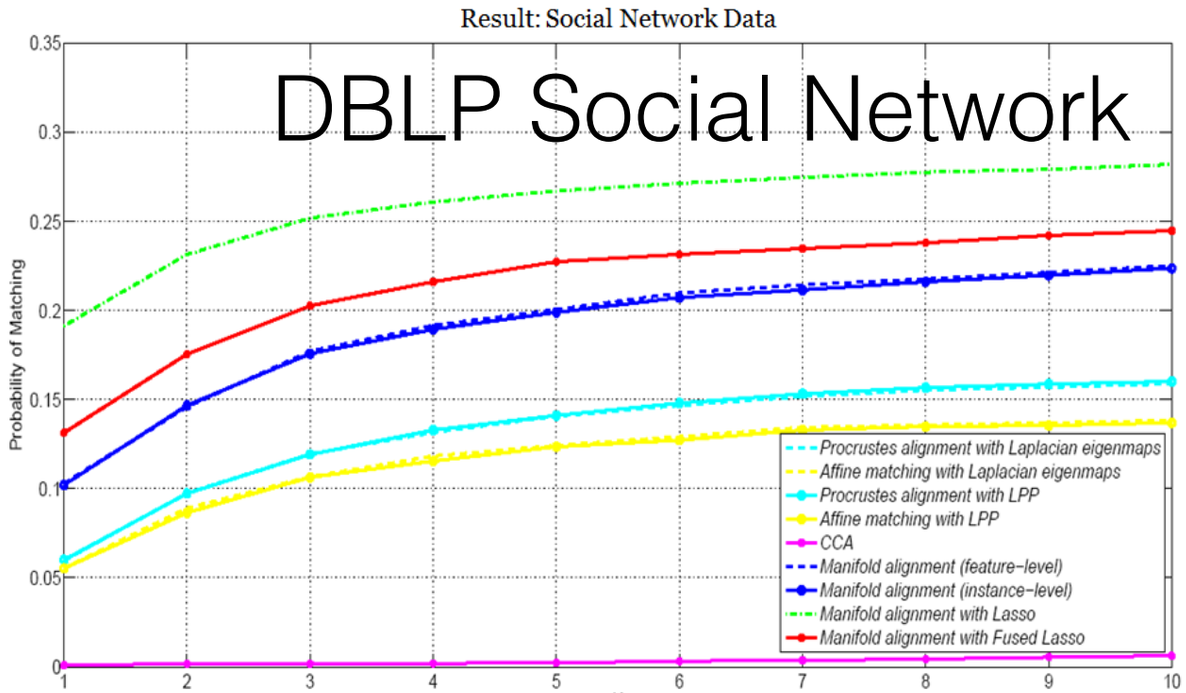
Use Lasso to find a sparse solution.

### Manifold Alignment with Lasso

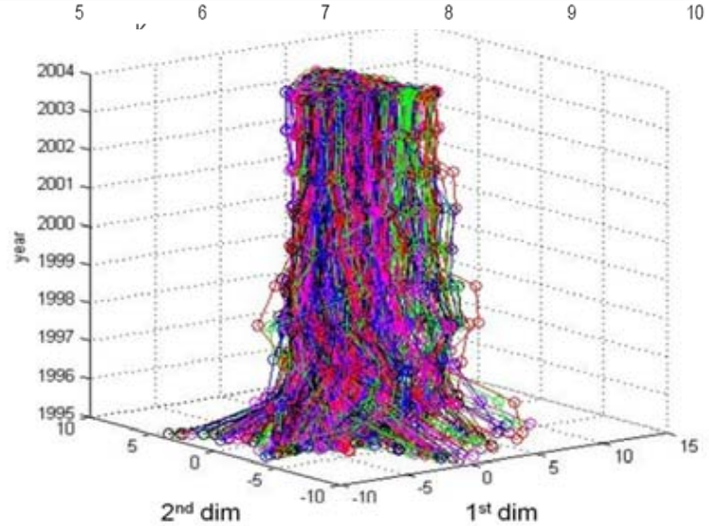
$$\|W^T Z - U_h^T Q^T\|_F^2 + \alpha \|W\|_{1,1}.$$

### Manifold Alignment with Fused Lasso

$$\|W^T Z - U_h^T Q^T\|_F^2 \quad + \quad \alpha \|W\|_{1,1} + \beta \sum_{j=1}^{h} \sum_{k=2}^{p+q} |w_{j,k} - w_{j,k-1}|.$$

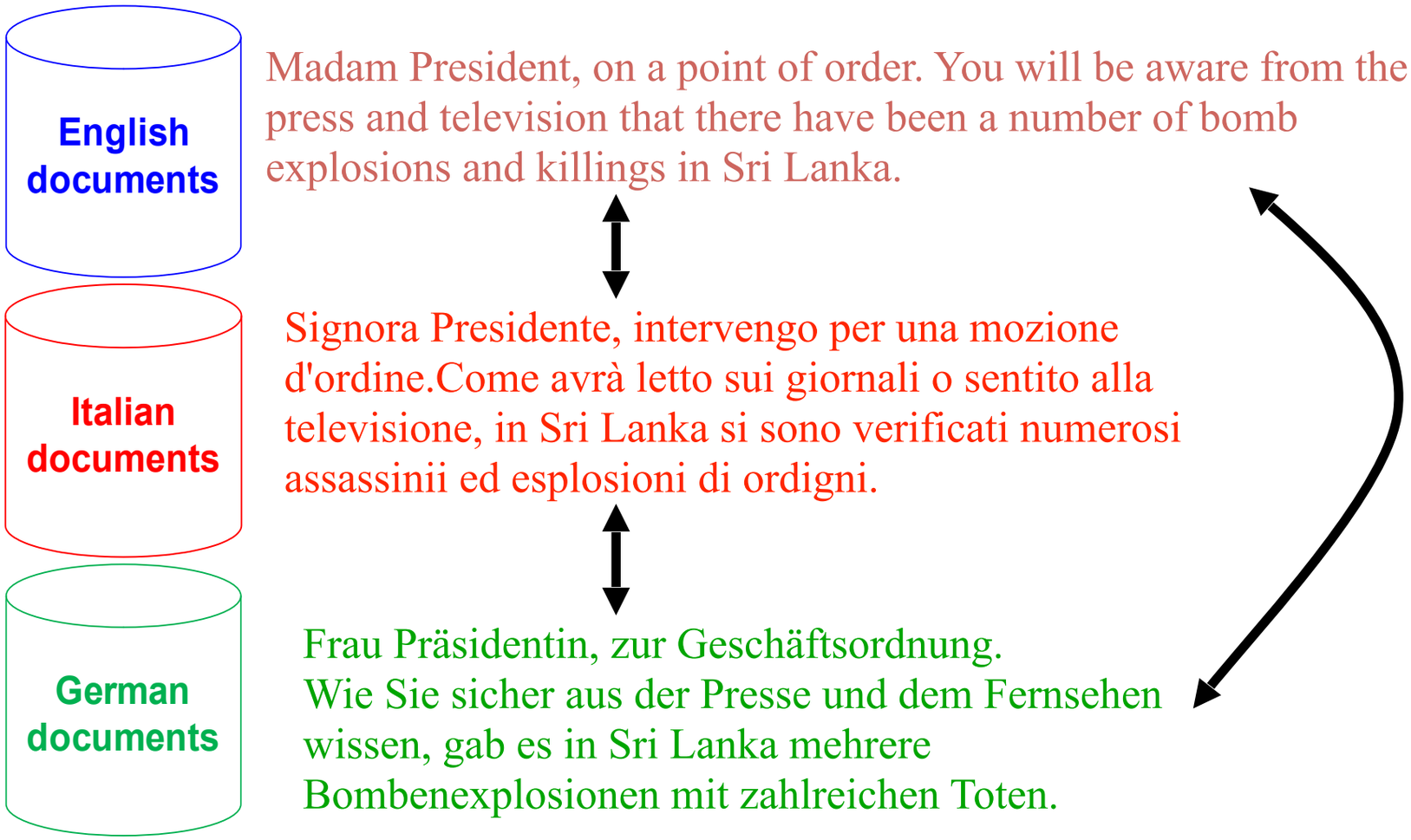



Wang, Liu, Vu, and Mahadevan, 2012

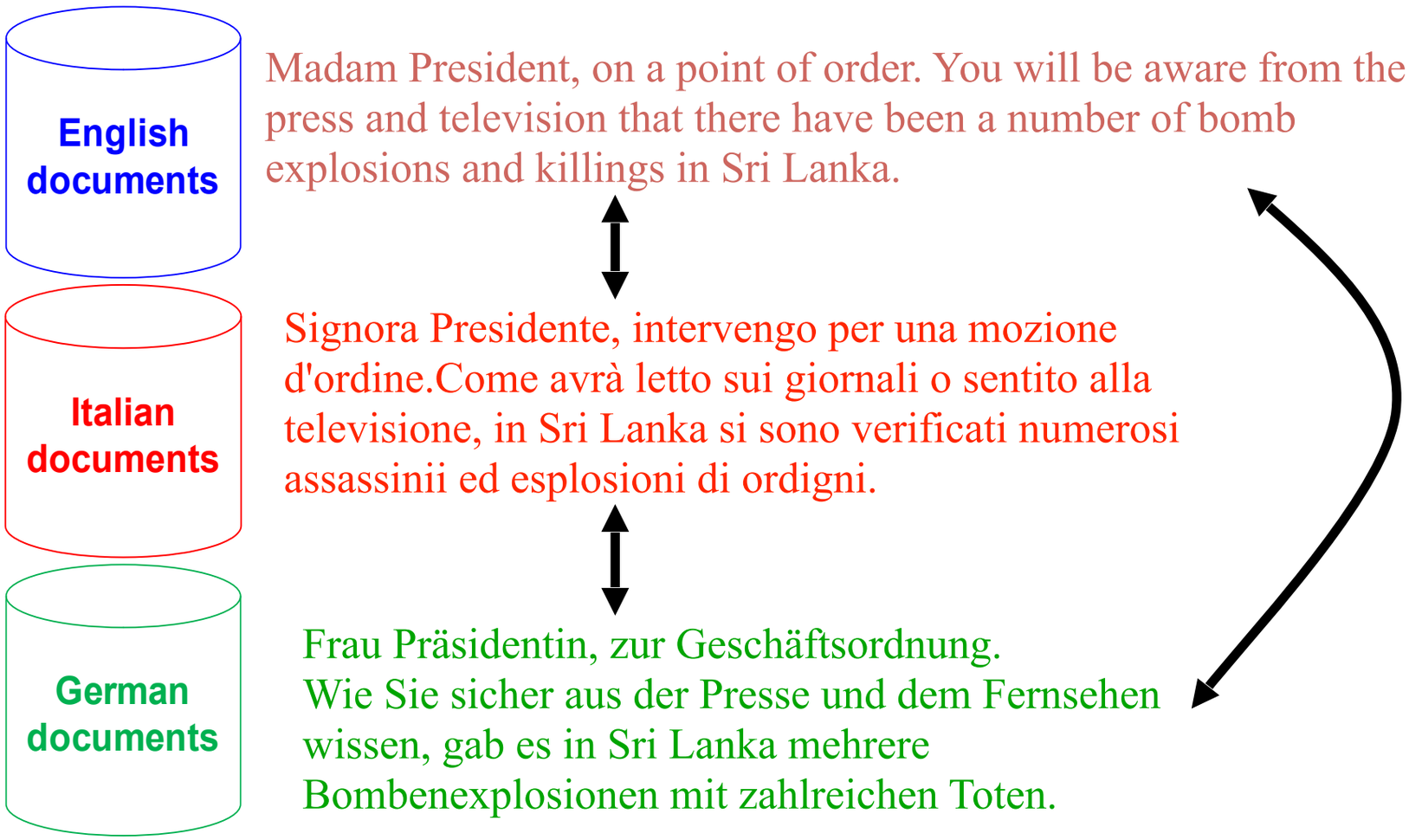


# Cross-Lingual Transfer in IR

# Cross-Lingual Transfer in IR



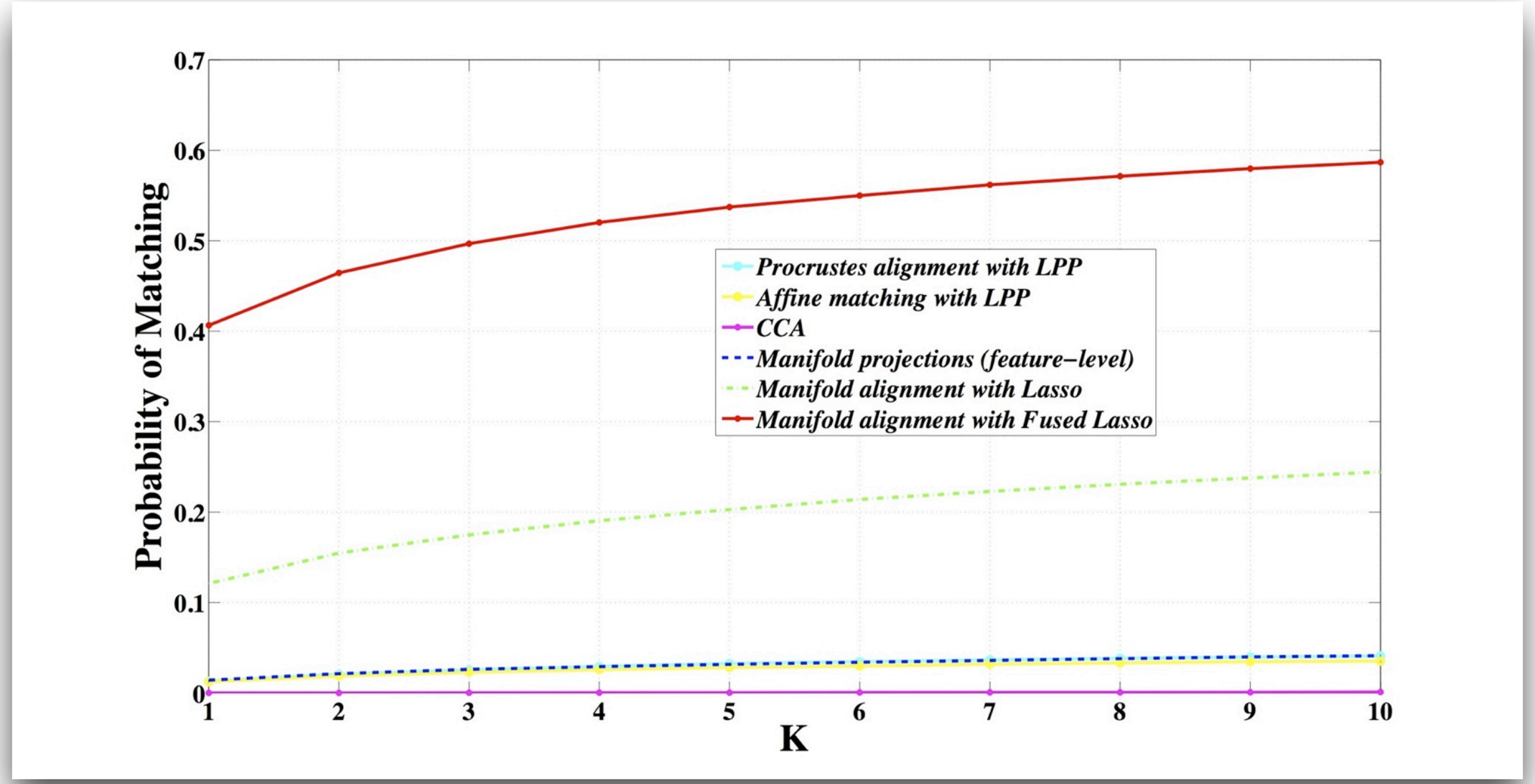
# Cross-Lingual Transfer in IR




Proceedings of the EU

# Cross-lingual IR

# Cross-lingual IR



# Impact of Work

# Impact of Work

- The most useful research I have done in 20 years!

# Impact of Work

- The most useful research I have done in 20 years!
- Led to several new collaborations

# Impact of Work

- The most useful research I have done in 20 years!
- Led to several new collaborations
  - Mars rover Curiosity (Darby Dyar, Mount Holyoke, NASA/ JPL scientific team)

# Impact of Work

- The most useful research I have done in 20 years!
- Led to several new collaborations
  - Mars rover Curiosity (Darby Dyar, Mount Holyoke, NASA/ JPL scientific team)
  - Proposals submitted to CDS&E and BIGDATA

# Impact of Work

- The most useful research I have done in 20 years!
- Led to several new collaborations
  - Mars rover Curiosity (Darby Dyar, Mount Holyoke, NASA/ JPL scientific team)
  - Proposals submitted to CDS&E and BIGDATA
  - Papers: 100+ citations on Google Scholar

# Impact of Work

- The most useful research I have done in 20 years!
- Led to several new collaborations
  - Mars rover Curiosity (Darby Dyar, Mount Holyoke, NASA/ JPL scientific team)
  - Proposals submitted to CDS&E and BIGDATA
  - Papers: 100+ citations on Google Scholar
  - Many many applications (bioinformatics, graphics, robotics, science, IR)

# Publications

Hoa Vu, CJ Carey, and Sridhar Mahadevan, " Manifold Warping: Manifold Alignment over Time " , Proceedings of the 26th Conference on Artificial Intelligence (AAAI), July 22-26, 2012, Toronto, Canada.

Chang Wang and Sridhar Mahadevan, " Manifold Alignment Preserving Global Geometry " , Technical Report, UMass Computer Science Department UM-CS-2012-031, 2012.

Chang Wang, Bo Liu, Hoa Vu, and Sridhar Mahadevan, " Sparse Manifold Alignment " , Technical Report, UMass Computer Science UM-2012-030, 2012.

Chang Wang and Sridhar Mahadevan, " Heterogeneous Domain Adaptation using Manifold Alignment " , Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), July 18-23, 2011, Barcelona, Spain.

Chang Wang and Sridhar Mahadevan, " Jointly Learning Data-Depdendent Label and Locality-Preserving Projections " , Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), July 18-23, 2011, Barcelona, Spain.

Blake Foster, Sridhar Mahadevan, and Rui Wang, " GPU-Based Approximate SVD Algorithm " , 9th International Conference on Parallel Programming and Mathematics, Torun, Poland, September 11-14, 2011 (also available as Technical Report UM-CS-2011-025, Univ. of Massachusetts, Amherst).

Chang Wang, Peter Krafft, and Sridhar Mahadevan, " Manifold Alignment ", appearing in Manifold Learning: Theory and Applications, Taylor and Francis CRC Press.

Chang Wang and Sridhar Mahadevan, "Multiscale Manifold Alignment" , Univ. of Massachusetts TR UM-CS-2010-049, 2010.

Chang Wang and Sridhar Mahadevan, "Learning Locality Preserving Discriminative Features" , Univ. of Massachusetts TR UM-CS-2010-048, 2010.