Multimodal manifold analysis by simultaneous diagonalization of Laplacians

ICVSS 2014 Sicily ~ 13-19 July International Computer Vision Summer School

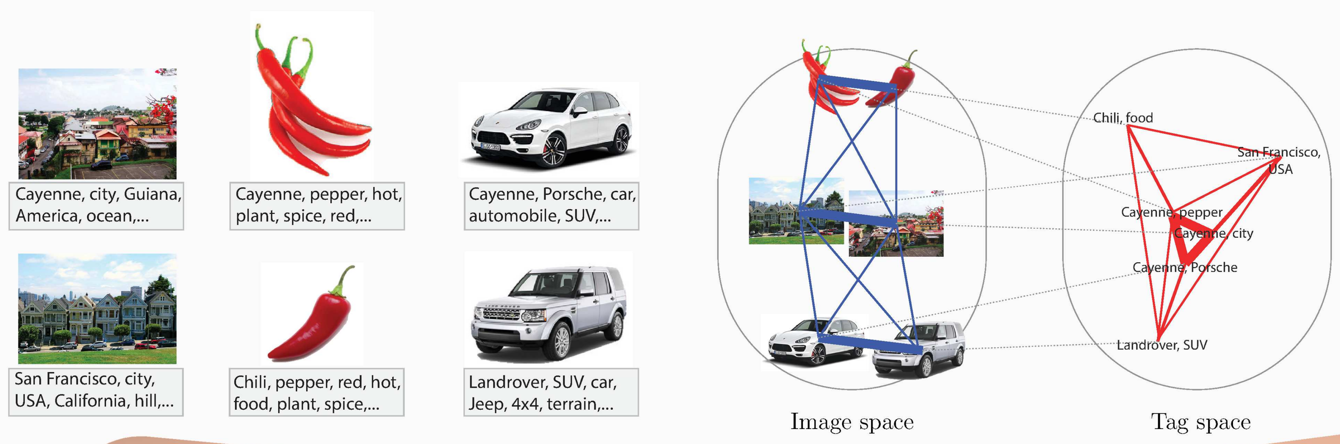
Eynard D., Kovnatsky A., Bronstein M., Glashoff K., and Bronstein A.

Institute of Computational Science, Faculty of Informatics, Università della Svizzera Italiana

Abstract

We construct an extension of spectral and diffusion geometry to multiple modalities through simultaneous diagonalization of Laplacian matrices. This naturally extends classical data analysis tools based on spectral geometry, such as diffusion maps and spectral clustering. We provide several synthetic and real examples of manifold learning, object classification, and clustering, showing that the joint spectral geometry better captures the inherent structure of multi-modal data.

Motivations and Objectives



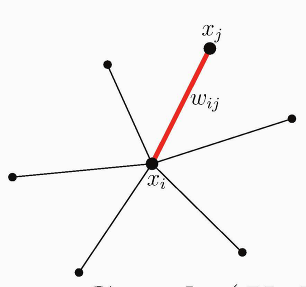
Data (e.g. colors, tags) are inherently ambiguous

Rely on multiple views of the same data to disambiguate it

References

[1] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," Neural Computation, vol. 15, pp. 1373–1396, 2002.
[2] R. Coifman and S. Lafon, "Diffusion maps," Applied and Computational Harmonic Analysis, vol. 21, pp. 5–30, 2006.
[3] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On spectral clustering: Analysis and an algorithm," in Proc. NIPS, 2001.
[4] J. F. Cardoso and A. Souloumiac, "Jacobi angles for simultaneous diagonalization," SIAM J. Matrix Anal. Appl, vol. 17, pp. 161âĂŞ164, 1996.
[5] A. Kovnatsky, M. M. Bronstein, A. M. Bronstein, K. Glashoff, and R. Kimmel, "Coupled quasi-harmonic bases," Computer Graphics Forum, vol. 32, no. 2, pp. 439–448, 2013.
[6] D. Eynard, K. Glashoff, M. Bronstein, and A. Bronstein, "Multi- modal diffusion geometry by joint diagonalization of laplacians," arXiv:1209.2295, 2012.

Image Laplacians



Graph (X, E)

- Discrete set of n vertices $X = \{x_1, \ldots, x_n\}$
- Gaussian edge weight

$$w_{ij} = \begin{cases} e^{-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}} & (i,j) \in E \\ 0 & \text{else} \end{cases}$$

- **Unnormalized Laplacian**

$$\mathbf{L} = \mathbf{D} - \mathbf{W}$$

$$\mathbf{D} = \mathrm{diag}(\textstyle\sum_{j \neq i} w_{ij})$$

- **Normalized Laplacian**

$$\mathbf{L}_{\mathbf{sym}} = \mathbf{D}^{-1/2}\mathbf{L}\mathbf{D}^{-1/2}$$

Laplacian Eigenmaps

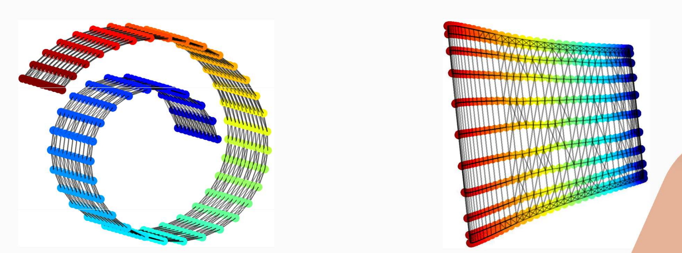
Eigenvalue problem:

$$\mathbf{L}\mathbf{\Phi} = \mathbf{\Phi}\mathbf{\Lambda}$$

- $\mathbf{\Lambda} = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ are the **eigenvalues** satisfying $0 = \lambda_1 \leq \lambda_2 \leq \ldots \lambda_n$
- $\mathbf{\Phi} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_n)$ are the orthonormal **eigenfunctions**

Laplacian Eigenmap [1] : m-dimensional embedding of X using the first m eigenvectors $\mathbf{U} = (\boldsymbol{\phi}_1, \ldots, \boldsymbol{\phi}_m)$ of the Laplacian

$$\mathbf{U} = \underset{\mathbf{U} \in \mathbb{R}^{n \times m}}{\mathrm{argmin}}\ \mathrm{tr}\,(\mathbf{U}^{\mathrm{T}}\mathbf{L}\mathbf{U}) \ \text{ s.t. } \ \mathbf{U}^{\mathrm{T}}\mathbf{U} = \mathbf{I}$$



X **U**

Laplacian eigenmap of the Swiss roll

Diffusion Distances

Heat diffusion on $\mathbf{X}$ is governed by the **heat equation:**

$$\mathbf{L}\mathbf{f}(t) + \frac{\partial}{\partial t}\mathbf{f}(t) = 0, \quad \mathbf{f}(0) = \mathbf{u},$$

where $\mathbf{f}(t)$ is the amount of heat at time t. The **heat kernel**

$$\mathbf{H}^t = e^{-t\mathbf{L}} = \mathbf{\Phi}e^{-t\mathbf{\Lambda}}\mathbf{\Phi}^{\mathrm{T}}$$

provides the solution of the heat equation $\mathbf{f}(t) = \mathbf{H}^t\mathbf{f}(0)$.

Diffusion distance [2]: crosstalk between heat kernels

$$d_t^2(x_p, x_q) = \sum_{i=1}^{n}((\mathbf{H}^t)_{pi} - (\mathbf{H}^t)_{qi})^2 = \sum_{i=1}^{n} e^{-2t\lambda_i}(\phi_{pi} - \phi_{qi})^2$$

equivalent to euclidean distance in the **diffusion map space**

$$\mathbf{U} = (e^{-t\lambda_1}\boldsymbol{\phi}_1, \ldots, e^{-t\lambda_m}\boldsymbol{\phi}_m)$$

Spectral clustering [3]: clustering obtained by running K-means in the Laplacian / diffusion map eigenspace

JADE

Eigendecomposition can be posed as the minimization problem

$$\min_{\mathbf{\Phi}^{\mathrm{T}}\mathbf{\Phi} = \mathbf{I}} \ \mathrm{off}(\mathbf{\Phi}^{\mathrm{T}}\mathbf{L}\mathbf{\Phi})$$

with **off-diagonality penalty** $\mathrm{off}(\mathbf{X}) = \sum_{i \neq j} x_{ij}^2$.

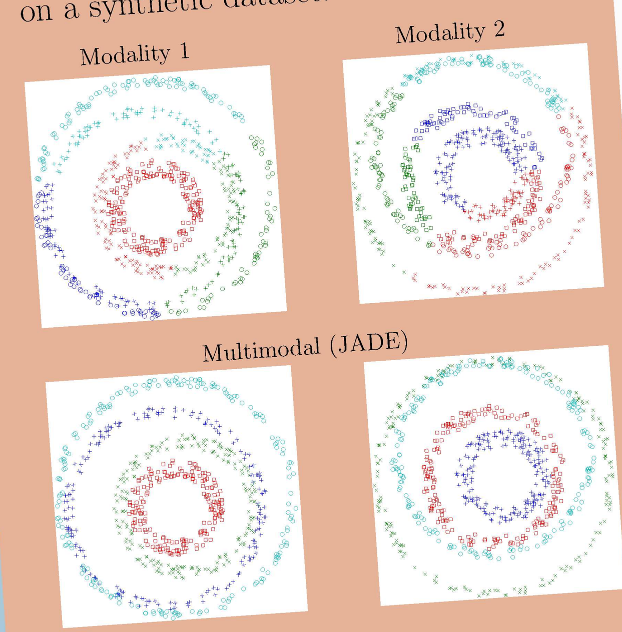
In JADE, Laplacians of two different modalities $\mathbf{X}$ and $\mathbf{Y}$ are diagonalized **simultaneously** [4]:

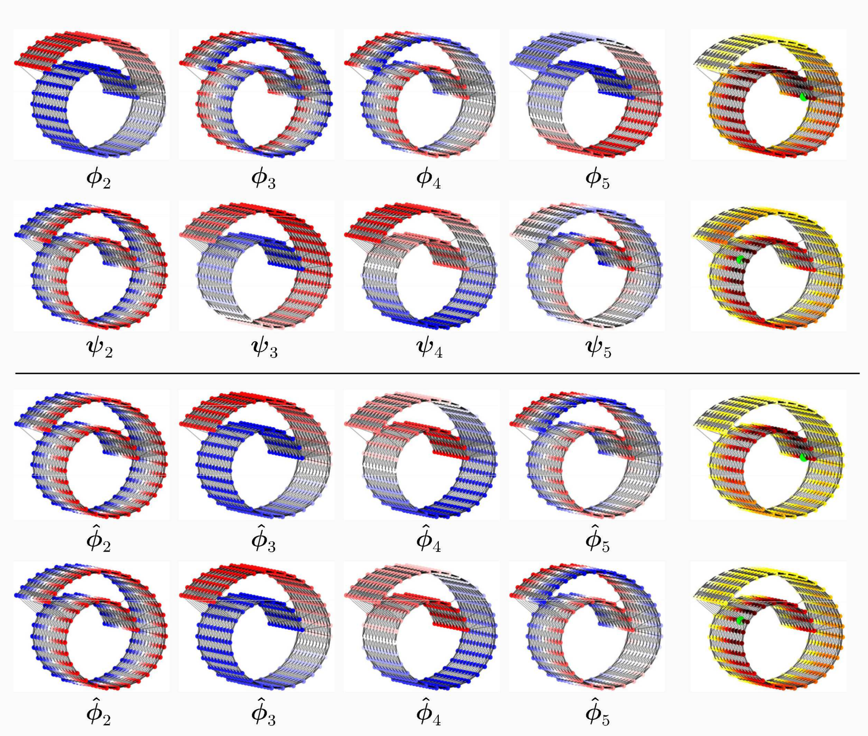
$$\min_{\hat{\mathbf{\Phi}}^{\mathrm{T}}\hat{\mathbf{\Phi}} = \mathbf{I}} \ \mathrm{off}(\hat{\mathbf{\Phi}}^{\mathrm{T}}\mathbf{L}_X\hat{\mathbf{\Phi}}) + \mathrm{off}(\hat{\mathbf{\Phi}}^{\mathrm{T}}\mathbf{L}_Y\hat{\mathbf{\Phi}})$$

Main drawbacks of JADE:

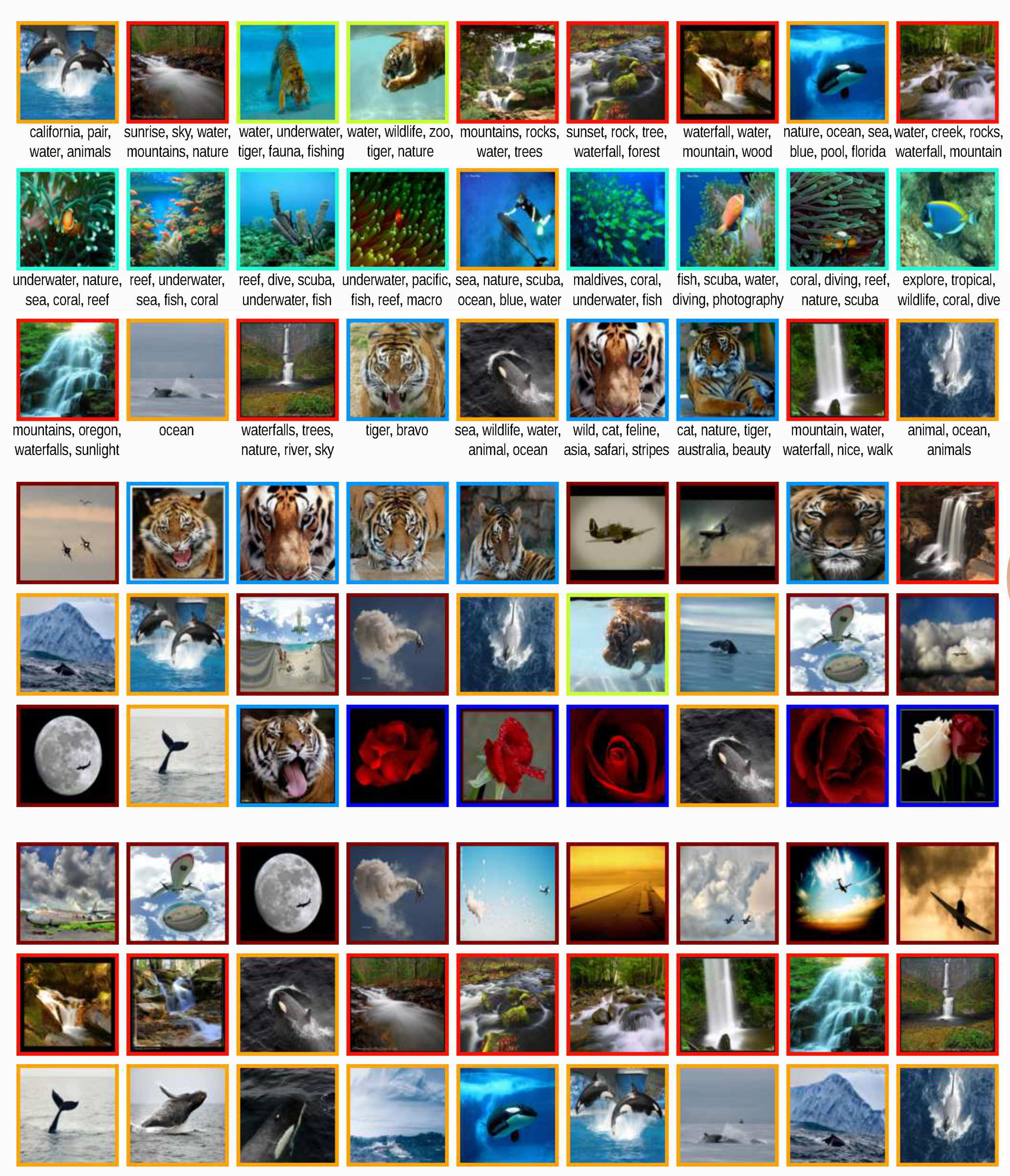
- Works on the whole basis, while often we need only the **first** $k \ll n$ eigenvectors
- Explicit assumption of orthonormality of the joint basis restricts Laplacian discretization to **symmetric matrices**
- Requires **bijective known correspondence**

Results of multimodal spectral clustering on a synthetic dataset:





Swiss rolls: Laplacian eigenvectors with different connectivity behave differently, as shown by their diffusion distances (right). Conversely, joint approximate eigenvectors behave in the same way.



NUS-WIDE dataset: annotated images belonging to 7 ambiguous classes. Modality 1: 1000-dimensional distributions of frequent **tags**. Modality 2: 64-dimensional **color histogram** image descriptors. Groundtruth clusters are shown in different colors.

Tag clusters

Color histogram clusters

Multimodal clusters

Coupled Diagonalization

Given wo discrete manifolds with **different number of vertices**, $X = \{x_1, \ldots, x_n\}$ and $Y = \{x_1, \ldots, x_m\}$, their Laplacians $\mathbf{L}_X, \mathbf{L}_Y$, and a set of **corresponding functions** $\mathbf{F} = (\mathbf{f}_1, \ldots, \mathbf{f}_q)$ and $\mathbf{G} = (\mathbf{g}_1, \ldots, \mathbf{g}_q)$, we look for **two sets** of **coupled approximate eigenvectors** [5] $\hat{\mathbf{\Phi}}, \hat{\mathbf{\Psi}}$

$$\min_{\hat{\mathbf{\Phi}}, \hat{\mathbf{\Psi}}} \ \mathrm{off}(\hat{\mathbf{\Phi}}^{\mathrm{T}}\mathbf{L}_X\hat{\mathbf{\Phi}}) + \mathrm{off}(\hat{\mathbf{\Psi}}^{\mathrm{T}}\mathbf{L}_Y\hat{\mathbf{\Psi}}) + \mu\|\mathbf{F}^{\mathrm{T}}\hat{\mathbf{\Phi}} - \mathbf{G}^{\mathrm{T}}\hat{\mathbf{\Psi}}\|_{\mathrm{F}}^2$$

$$\text{s.t.} \ \ \hat{\mathbf{\Phi}}^{\mathrm{T}}\hat{\mathbf{\Phi}} = \mathbf{I}, \quad \hat{\mathbf{\Psi}}^{\mathrm{T}}\hat{\mathbf{\Psi}} = \mathbf{I}$$

For approximately jointly diagonalizable Laplacians we know [4] that $\mathrm{span}\{\hat{\phi}_1, \ldots, \hat{\phi}_k\} \approx \mathrm{span}\{\phi_1, \ldots, \phi_k\}$. We can thus approximate the first k elements of the coupled bases as a linear combination of the first $k' \geq k$ eigenvectors of their respective Laplacians: $\hat{\mathbf{\Phi}} = \bar{\mathbf{\Phi}}\mathbf{S}$, $\hat{\mathbf{\Psi}} = \bar{\mathbf{\Psi}}\mathbf{R}$.
The optimization problem can be reformulated as:

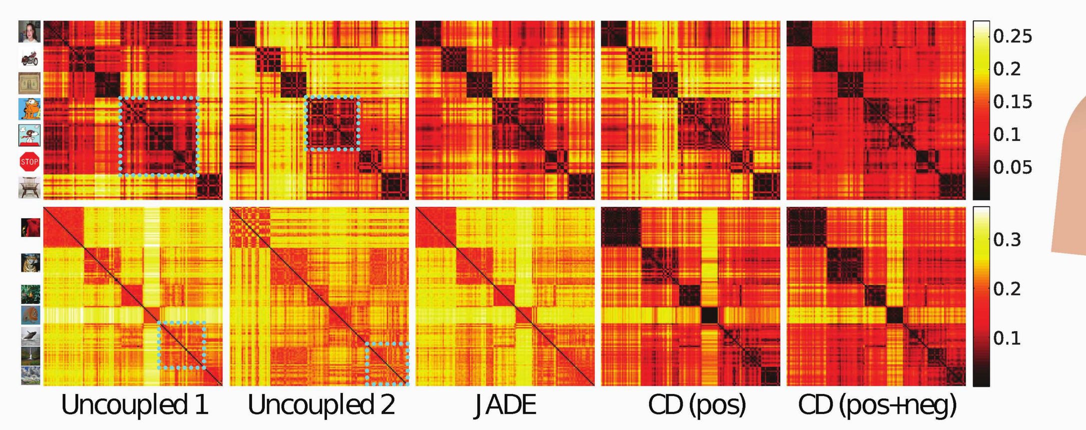
$$\min_{\mathbf{R}, \mathbf{S}} \ \mathrm{off}(\mathbf{R}^{\mathrm{T}}\bar{\mathbf{\Lambda}}_X\mathbf{R}) + \mathrm{off}(\mathbf{S}^{\mathrm{T}}\bar{\mathbf{\Lambda}}_Y\mathbf{S}) + \mu\|\mathbf{F}^{\mathrm{T}}\bar{\mathbf{\Phi}}\mathbf{R} - \mathbf{G}^{\mathrm{T}}\bar{\mathbf{\Psi}}\mathbf{S}\|_{\mathrm{F}}^2$$

$$\text{s.t.} \ \ \mathbf{R}^{\mathrm{T}}\mathbf{R} = \mathbf{I}, \quad \mathbf{S}^{\mathrm{T}}\mathbf{S} = \mathbf{I}$$

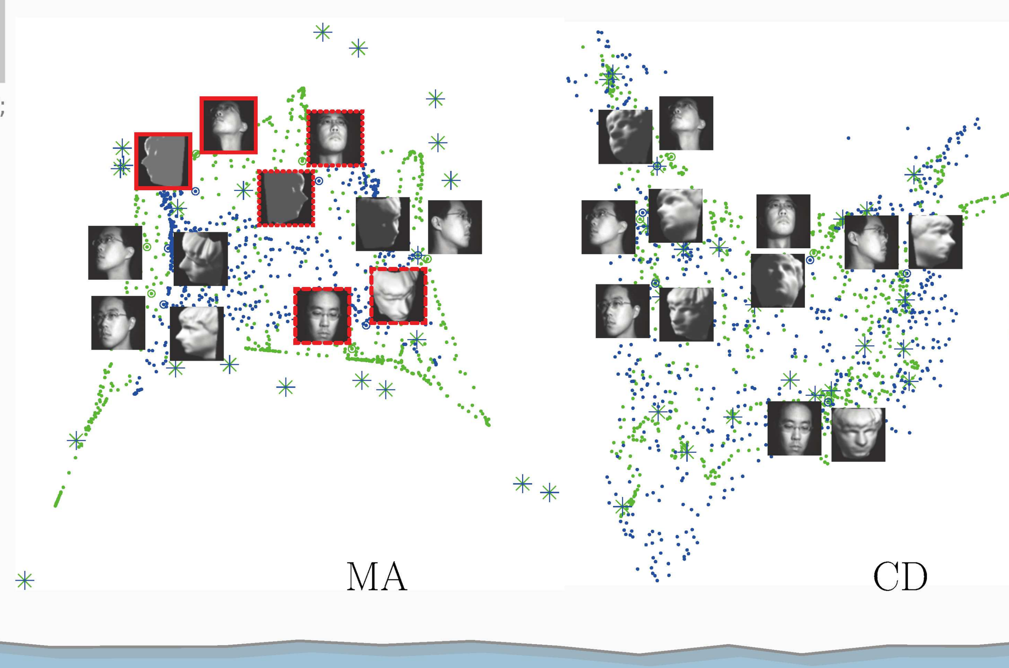
- Laplacians are **not used explicitly**: their first k' eigenfunctions $\bar{\mathbf{\Phi}}$, $\bar{\mathbf{\Psi}}$ and eigenvalues $\bar{\mathbf{\Lambda}}_X$, $\bar{\mathbf{\Lambda}}_Y$ are pre-computed.
- Problem size is $2k' \times k$, **independent** of the number of samples.
- **No bijective correspondence!**

Method			Accuracy (%) Circles	Text	Caltech	NUS	Digits	Reuters
#points			800	800	105	145	2000	600
Uncoupled			53.0	60.4	78.1	80.7	78.9	52.3
Harmonic Mean			95.6	97.2	87.6	89.0	87.0	52.3
Arithmetic Mean			96.5	96.9	87.6	95.2	82.8	52.2
Comraf			40.8	60.8	–	86.9	81.6	53.2
MVSC			95.6	97.2	81.0	89.0	83.1	52.3
MultiNMF			41.1	50.5	–	77.4	87.2	53.1
SC-ML			98.2	97.6	88.6	94.5	87.8	52.8
JADE			100	98.4	86.7	93.1	82.5	52.3
CD*	pos	10%	52.5	54.5	78.7	78.6	94.2	53.7
		20%	61.3	60.0	80.8	82.9	94.1	54.2
		60%	93.7	86.5	87.0	87.2	93.9	54.7
		100%	98.9	96.8	89.5	94.5	93.9	54.8
	pos+neg	10%	67.3	63.6	86.5	92.7	94.9	59.0
		20%	69.6	67.8	87.9	93.3	94.8	57.6
		60%	95.2	87.0	89.2	94.5	94.8	57.0

Methods: Eynard 2012; Bekkerman 2007; Cai 2011; Liu 2013; Dong 2013;
Data: Cai 2011; Chua 2009; Alpaydin 1998; Liu 2013; Amini 2009



Object recognition



Manifold alignment for cross-modal image retrieval