

Image Similarities for Learning Video Manifolds

Selen Atasoy

MICCAI 2011 Tutorial

Image Spaces

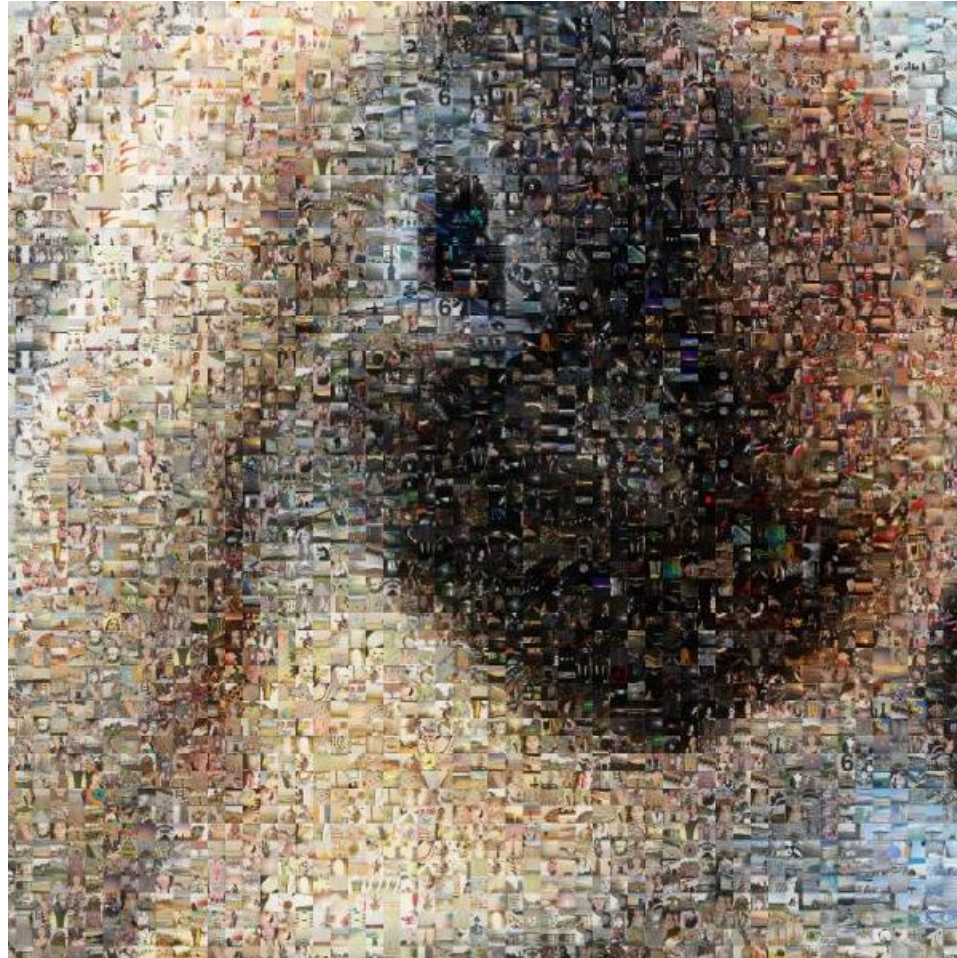
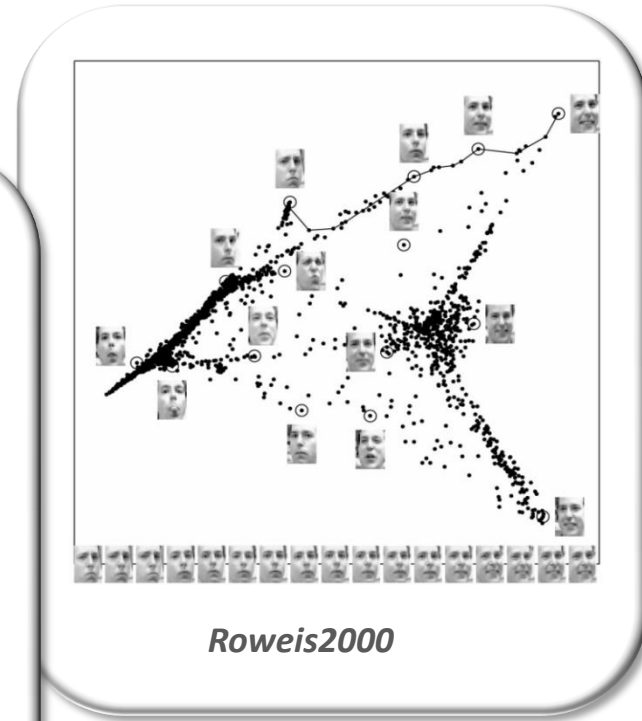
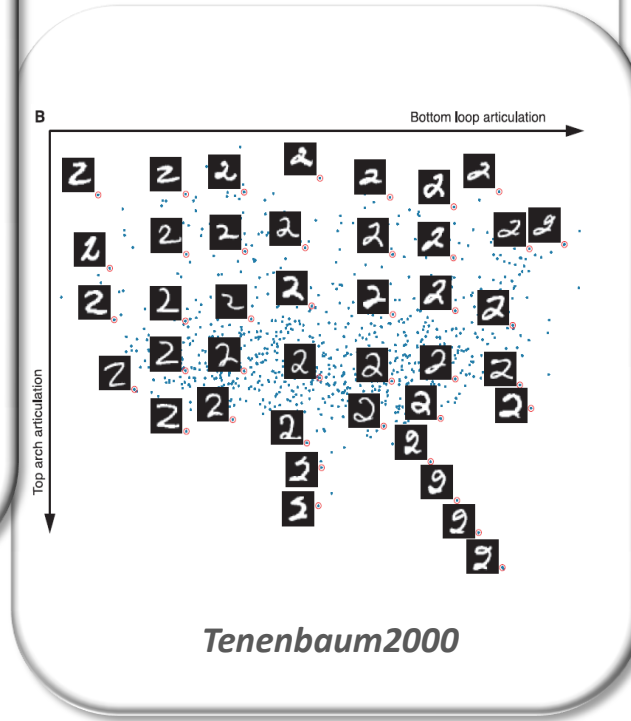
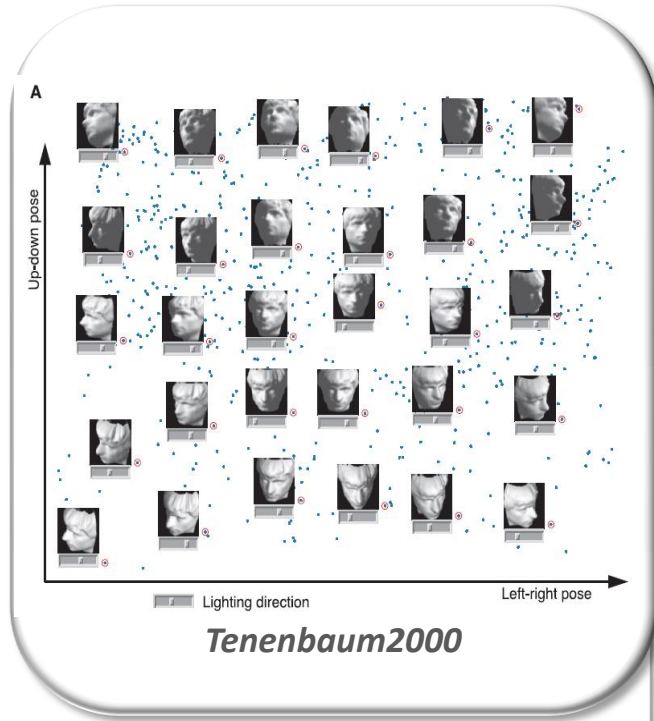


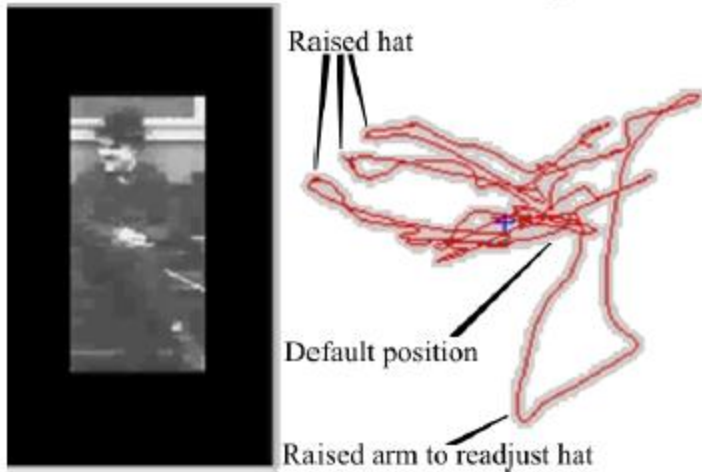
Image Manifolds



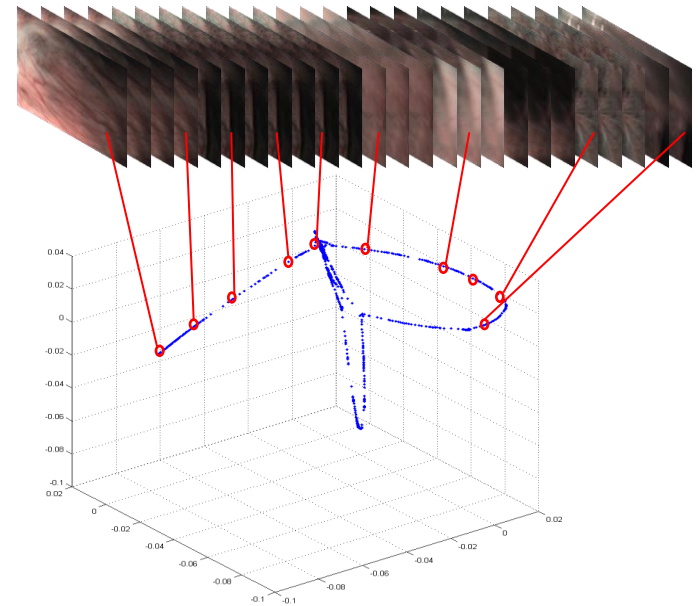
[**Tenenbaum2000**: J. B. Tenenbaum, V. Silva, J. C. Langford: A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500), 2000.]

[**Roweis2000**: S. T. Roweis, L. K. Saul: Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2000]

Video Manifolds



Pless2003



Atasoy2010

[*Pless2003*: R. Pless: Using Isomap to Explore Video Sequences: *ICCV*, 2003.]

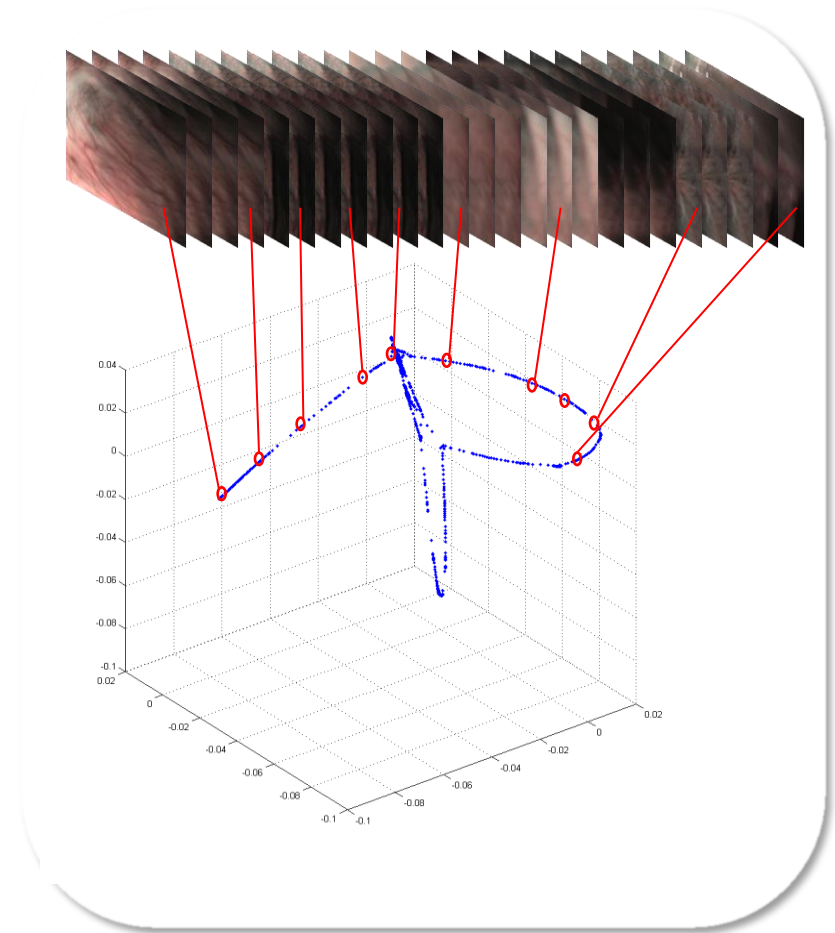
[*Atasoy2010*: S. Atasoy, D. Mateus, J. Lallemand, A. Meining, G.Z. Yang, N. Navab: Endoscopic Video Manifolds, *MICCAI*, 2010.]

[*Atasoy2011*: S. Atasoy, D. Mateus, A. Meining, G.Z. Yang, N. Navab: Targeted Optical Biopsies for Surveillance Endoscopies, *MICCAI*, 2011.]

Theoretical Background

Manifold Learning

- High dimensional data points lying on or near a manifold
 $\mathbf{x} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \in \mathcal{M} \subset \mathbb{R}^{(w \times h)}$
- Low dimensional representation
 $\mathbf{y} = \{\mathbf{y}_1, \dots, \mathbf{y}_n\} \in \mathbb{R}^d$
 $d \ll (w \times h)$
- Find a mapping $\mathbf{y} = f(\mathbf{x})$
 $f : \mathcal{M} \rightarrow \mathbb{R}^d$
 that best preserves ???



Manifold Learning

A General Recipe

1. Define a matrix based on the relations between data points

$$A_{i,j} = \mathcal{A}(\mathbf{x}_i, \mathbf{x}_j)$$

2. Compute the eigenvectors & eigenvalues

$$A\mathbf{u}_0 = \lambda_0\mathbf{u}_0$$

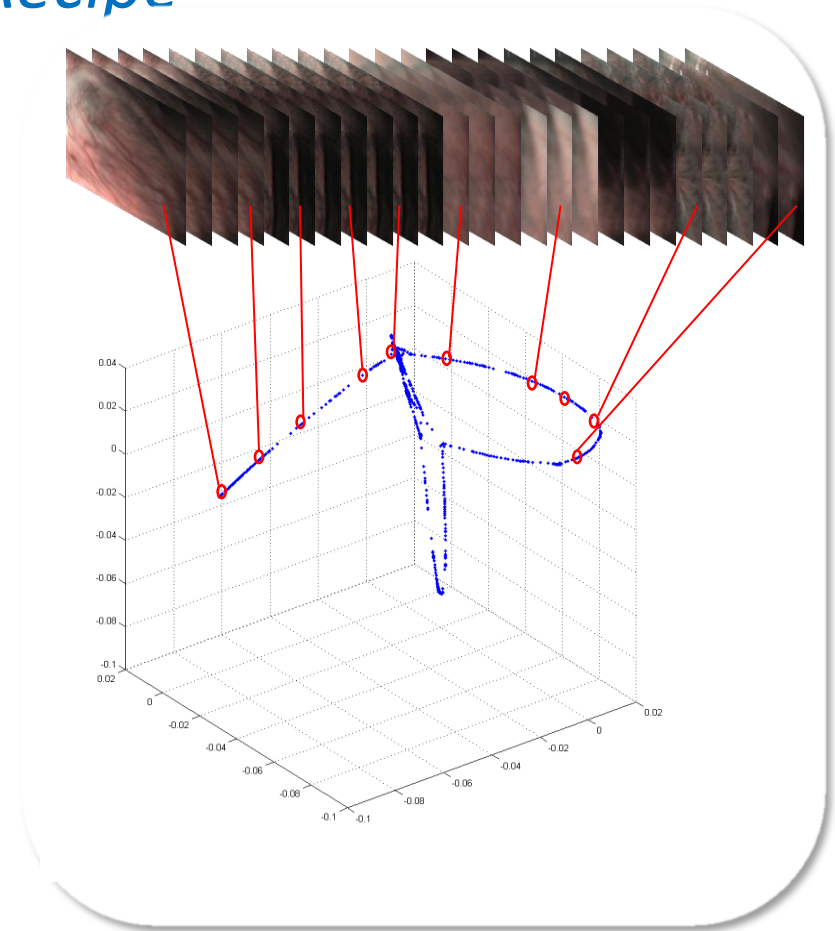
$$A\mathbf{u}_1 = \lambda_1\mathbf{u}_1$$

⋮

$$A\mathbf{u}_d = \lambda_d\mathbf{u}_d$$

3. Embed each sample \mathbf{x}_i

$$f(\mathbf{x}_i) = \mathbf{y}_i = [\mathbf{u}_1(i), \dots, \mathbf{u}_d(i)]^\top$$



Manifold Learning

A General Recipe

Method	Operator/Matrix	Preserved	Objective Function
PCA	Covariance matrix	Variance of the dataset / Euclidean distances between data points	$\mathbf{u}^\top \Sigma \mathbf{u}$
Laplacian Eigenmaps	Graph Laplacian	Distances within the local neighbourhood of each data point	$\mathbf{u}^\top L \mathbf{u}$
ISOMAP	Geodesic distance matrix	Geodesic distances between data points	$\mathbf{u}^\top D_G \mathbf{u}$
LLE	Reconstruction weights	Reconstruction weights within the local neighbourhood of each data point	$\mathbf{u}^\top W \mathbf{u}$

Manifold Learning

Why does it work?

- Rayleigh-Ritz Theorem:

$$\arg \min_{\mathbf{y}} \frac{\mathbf{y}^\top \mathbf{H} \mathbf{y}}{\mathbf{y}^\top \mathbf{y}}$$

$$\mathbf{H} \quad \underbrace{\{\lambda_1, \dots, \lambda_n\}}_{\text{eigenvalues}} \quad \underbrace{\{\mathbf{y}_1, \dots, \mathbf{y}_n\}}_{\text{eigenvectors}}$$

- Recall:

- Scalar product: $\mathbf{f}^\top \mathbf{g} = \langle \mathbf{f}, \mathbf{g} \rangle$

- Scalar product in H : $\mathbf{f}^\top \mathbf{H} \mathbf{g} = \langle \mathbf{f}, \mathbf{g} \rangle_H = \langle \mathbf{f}, \mathbf{H} \mathbf{g} \rangle_{L_2}$

- Norm: $\|\mathbf{f}\|_{L_2}^2 = \langle \mathbf{f}, \mathbf{f} \rangle$

- Norm in H : $\|\mathbf{f}\|_H^2 = \langle \mathbf{f}, \mathbf{f} \rangle_H = \langle \mathbf{f}, \mathbf{H} \mathbf{f} \rangle_{L_2}$

Manifold Learning

Why does it work?

Discrete Domain

- $\mathbf{f}, \mathbf{g} \in \mathbb{R}^n$ vectors

$$\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}^\top \mathbf{g}$$

$$\langle \mathbf{f}, \mathbf{g} \rangle_H = \mathbf{f}^\top \mathbf{H} \mathbf{g}$$

Continuous Domain

- $f, g : \mathbb{R} \rightarrow \mathbb{R}$ functions

$$\langle f, g \rangle = \int f(t)g(t)dt$$

$$\langle f, g \rangle_H = \int f(t) [\mathcal{H}g](t) dt$$

Schwarz's Kernel Theorem: Each linear operator is given as an integration against a **unique** kernel. That kernel is the impulse response of the linear system to an impulse (a delta function).

$$[\mathcal{H}g](t) = \int \mathcal{K}_H(t, s)g(s)ds$$

Manifold Learning

Why does it work?

Discrete Domain

- $\mathbf{f}, \mathbf{g} \in \mathbb{R}^n$ vectors
- $\langle \mathbf{f}, \mathbf{g} \rangle = \mathbf{f}^\top \mathbf{g}$
- $\langle \mathbf{f}, \mathbf{g} \rangle_H = \mathbf{f}^\top \mathbf{H} \mathbf{g}$

Continuous Domain

- $f, g : \mathbb{R} \rightarrow \mathbb{R}$ functions
- $\langle f, g \rangle = \int f(t)g(t)dt$
- $\langle f, g \rangle_H = \int f(t)[\mathcal{H}g](t)dt$
 $= \int f(t) \int \mathcal{K}_H(t, s)g(s) ds dt$
 $= \int \int \mathcal{K}_H(t, s)f(t)g(s) ds dt$

Manifold Learning

Why does it work?

Discrete Domain

- $\mathbf{f}, \mathbf{g} \in \mathbb{R}^n$ vectors
- $\langle \mathbf{f}, \mathbf{f} \rangle_H = \mathbf{f}^\top \mathbf{H} \mathbf{f}$
- $\langle \mathbf{f}, \mathbf{H} \mathbf{f} \rangle_{L_2}$

Continuous Domain

- $f, g : \mathbb{R} \rightarrow \mathbb{R}$ functions
- $\langle f, f \rangle_H = \int f(t) [\mathcal{H}f](t) dt$
- $= \int \int \mathcal{K}_{\mathcal{H}}(t, s) f(t) f(s) ds dt$

$\|f\|_H$

The matrix \mathbf{H} defines:

- *which operator is applied*
- *which (Hilbert) space we are working in*
- *which quantity will be conserved*

Laplacian Eigenmaps

Manifold Learning

Laplacian Eigenmaps

- Solve

$$\arg \min_{\mathbf{y}} \frac{\mathbf{y}^\top \mathbf{L} \mathbf{y}}{\mathbf{y}^\top \mathbf{y}}$$

- Find the eigenvectors of the graph Laplacian

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \quad \mathbf{A}(i, j) = \begin{cases} 1 & \text{if } i \in \mathcal{N}_j^{\text{sim}} \\ 0 & \text{otherwise} \end{cases} \quad \mathbf{D} = \sum_{i=1}^n \mathbf{A}(i, i)$$

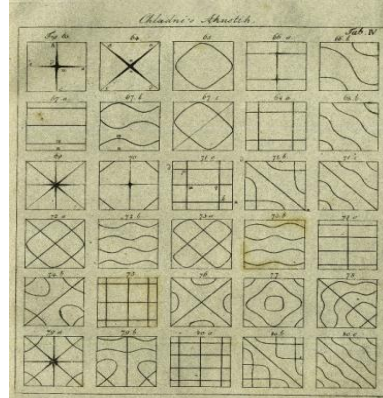
- Equivalent to solving the Helmholtz Equation

$$\mathbf{L} \mathbf{y} = \lambda \mathbf{y} \quad \Delta_{\mathcal{M}} f = \lambda f$$

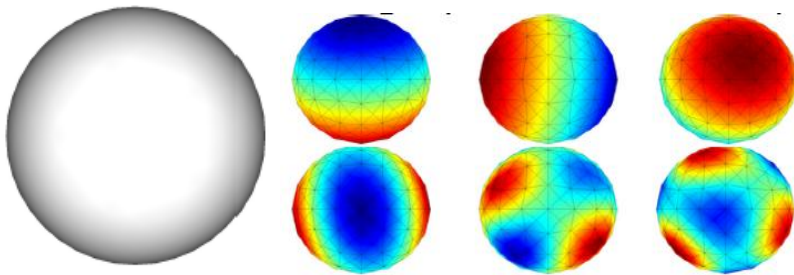
[*Belkin2003: M. Belkin, P. Niyogi: Laplacian eigenmaps for dimensionality reduction and data representation. Neural computation, 15(6), 1373-1396. MIT Press, 2003*]

Manifold Learning

Laplacian Eigenmaps - Interpretation



[Chladni1787]



[Levy2010]



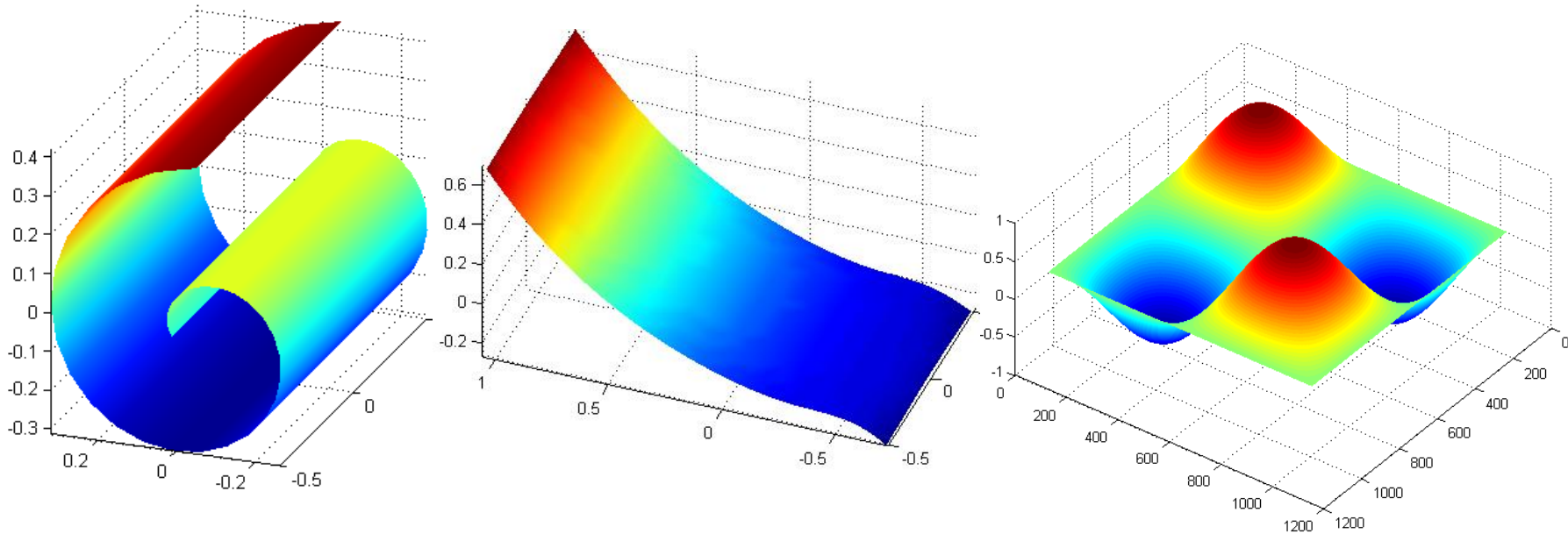
[Levy2010]

[Chladni1787: E. Chladni: Discoveries in the Theory of Sound, 1787.]

[Levy2010: B. Levy: Spectral Geometry Processing: ACM SIGGRAPH Course Notes, 2010.]

Non-linear Manifold Learning

Laplacian Eigenmaps - Interpretation



- Manifold learning as bending, stretching without cutting or creating wholes
- Vibrational modes are preserved while bending the manifold

Endoscopic Video Manifolds (EVMs)

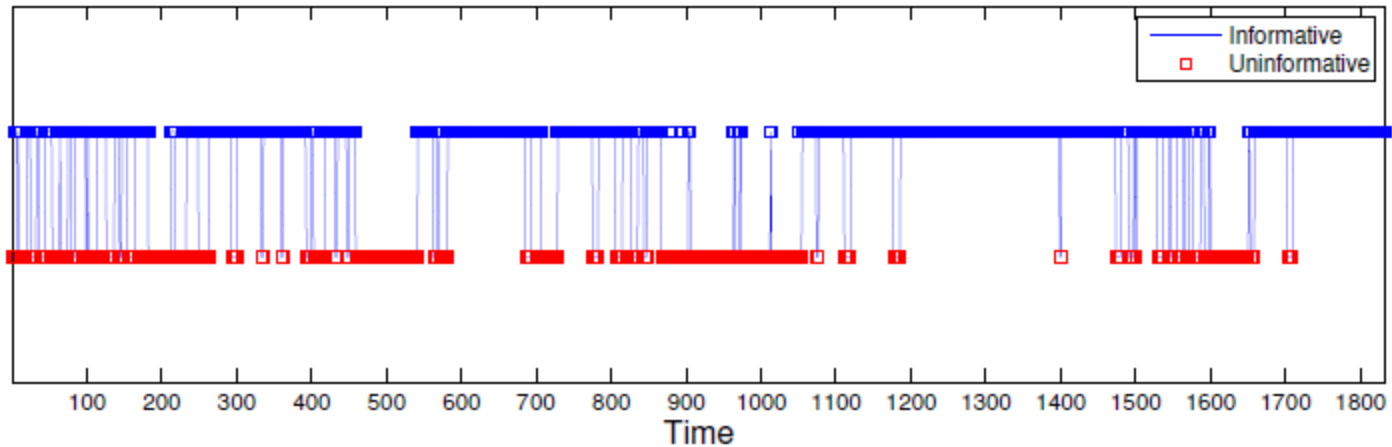
Endoscopic Video Manifolds

Challenges

- Clustering Uninformative Frames

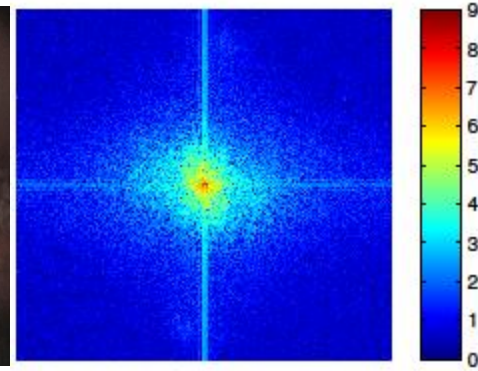
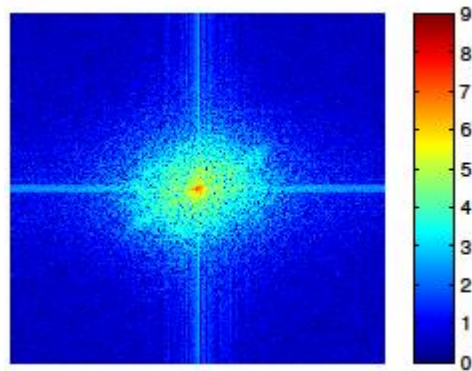
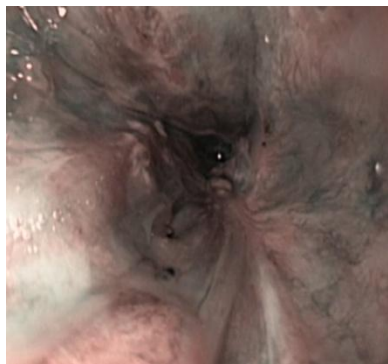
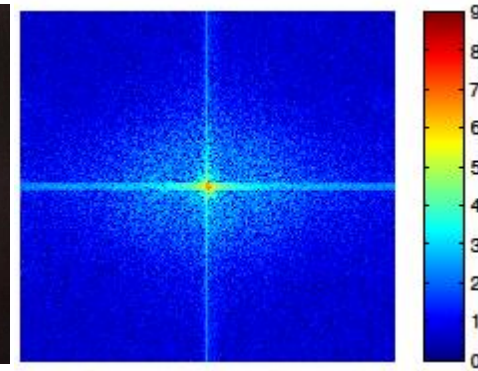
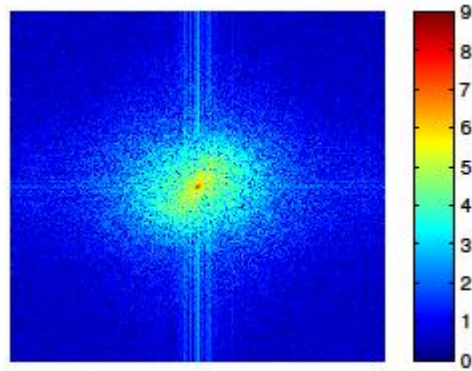
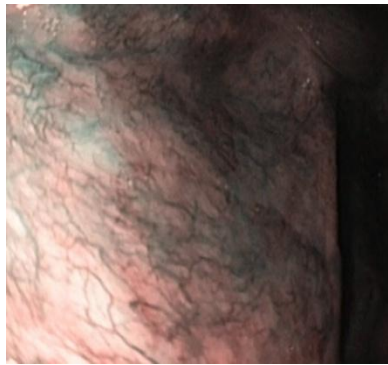


Informative/Uninformative Frame Distribution over Time



Endoscopic Video Manifolds

Clustering Uninformative Frames



Endoscopic Video Manifolds

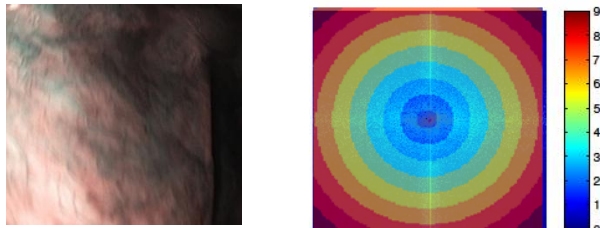
Clustering Uninformative Frames



Informative frame & power spectrum

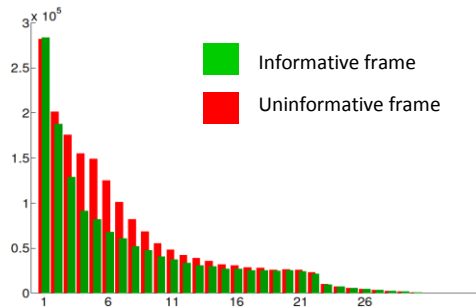
$$\mathcal{F}_i(\omega) = \sum_{\theta=0}^{2\pi} f_i(\omega, \theta)$$

$\{\gamma_1, \dots, \gamma_\delta, \dots, \gamma_b\}$



Uninformative frame & power spectrum

$$\mathcal{F}_j(\omega) = \sum_{\theta=0}^{2\pi} f_j(\omega, \theta)$$

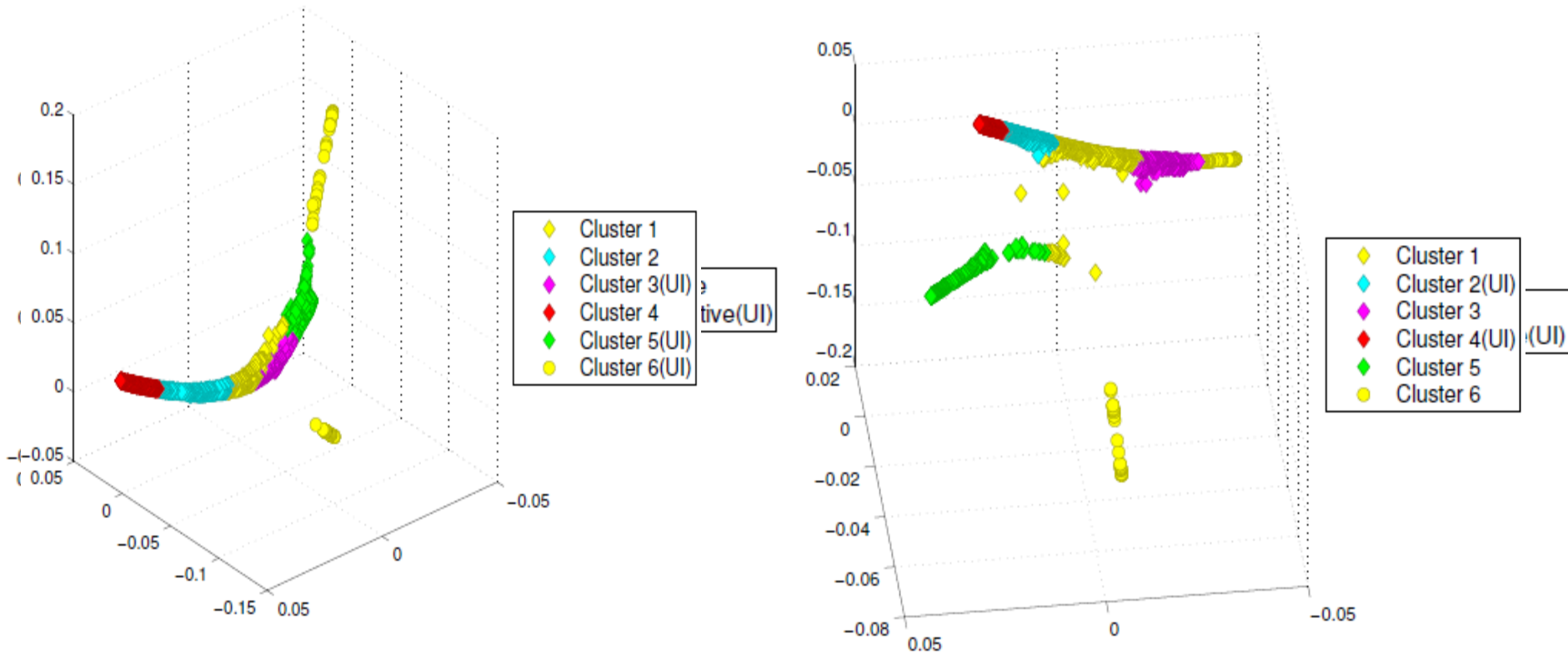


$$W(\mathcal{I}_i, \mathcal{I}_j) = W_{\text{EH}}(\mathcal{I}_i, \mathcal{I}_j)$$

$$W_{\text{EH}}(\mathcal{I}_i, \mathcal{I}_j) = 1 - \left(\frac{\langle \mathbf{h}_i, \mathbf{h}_j \rangle}{\|\mathbf{h}_i\| \cdot \|\mathbf{h}_j\|} \right)$$

Endoscopic Video Manifolds

Clustering Uninformative Frames

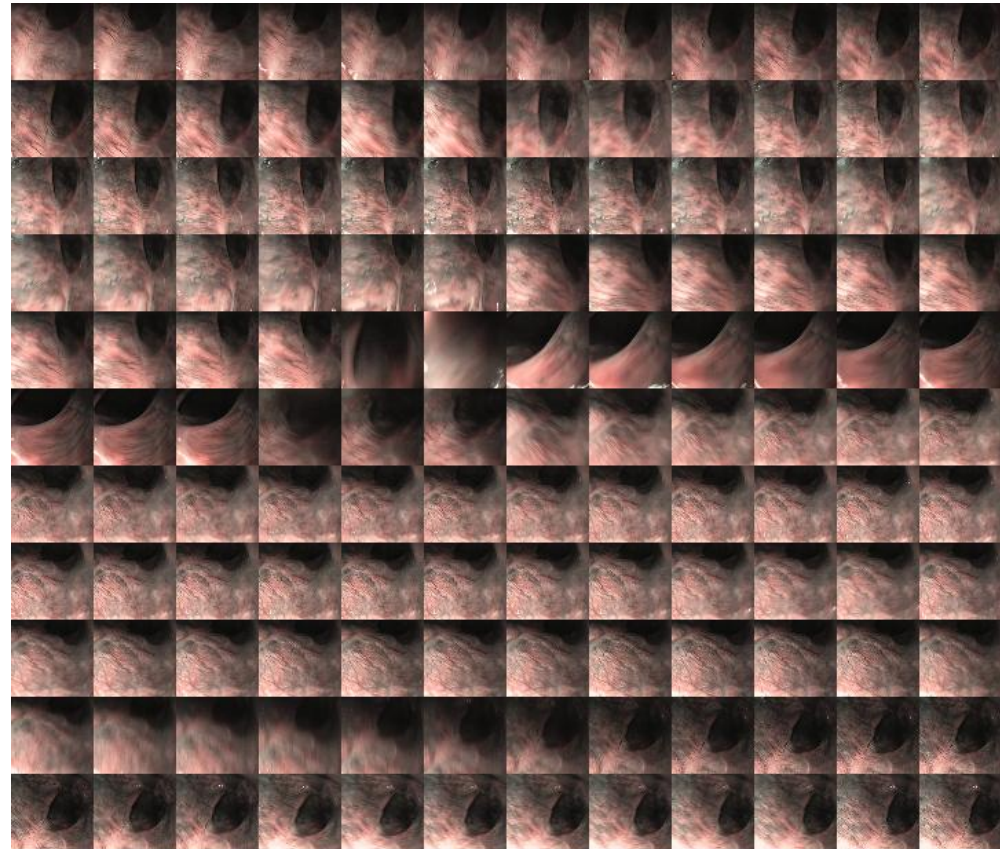


[Atasoy2010: S. Atasoy, D. Mateus, J. Lallemand, A. Meining, G.Z. Yang, N. Navab: Endoscopic Video Manifolds, MICCAI, 2010.]

Endoscopic Video Manifolds

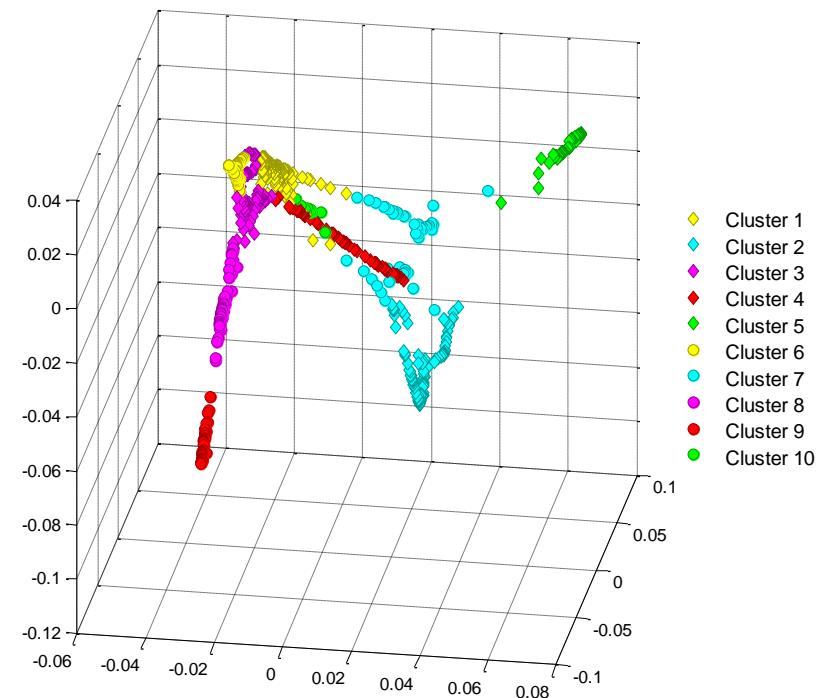
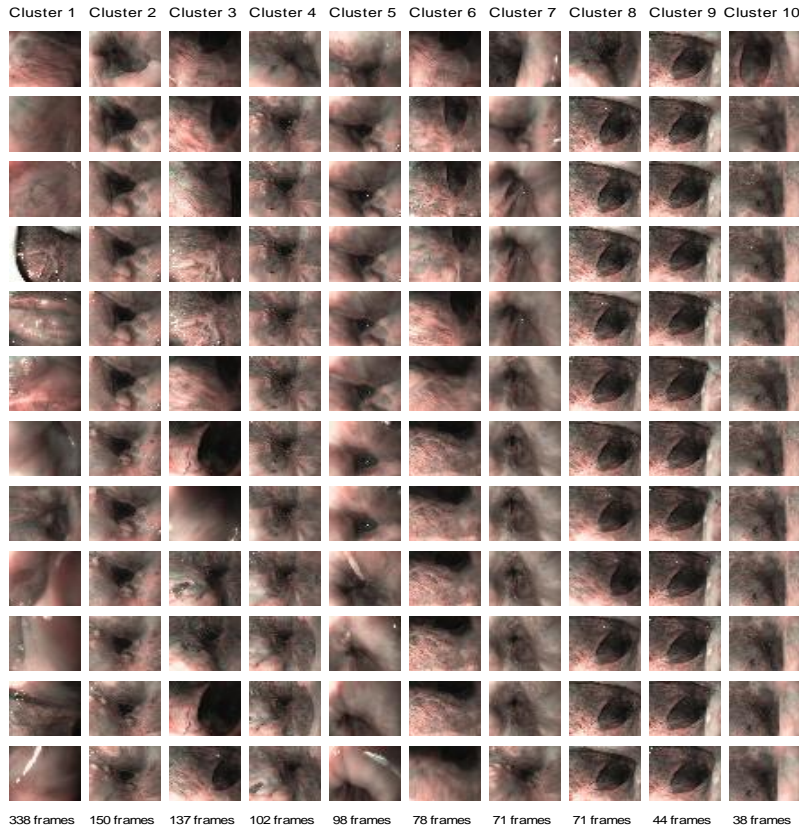
Challenges

- Significant change in endoscope viewpoint
- Small overlap between frames showing the same scene
- Scenes do not necessarily contain distinctive features



Endoscopic Video Manifolds

Clustering Endoscopic Scenes – Euclidean Distance



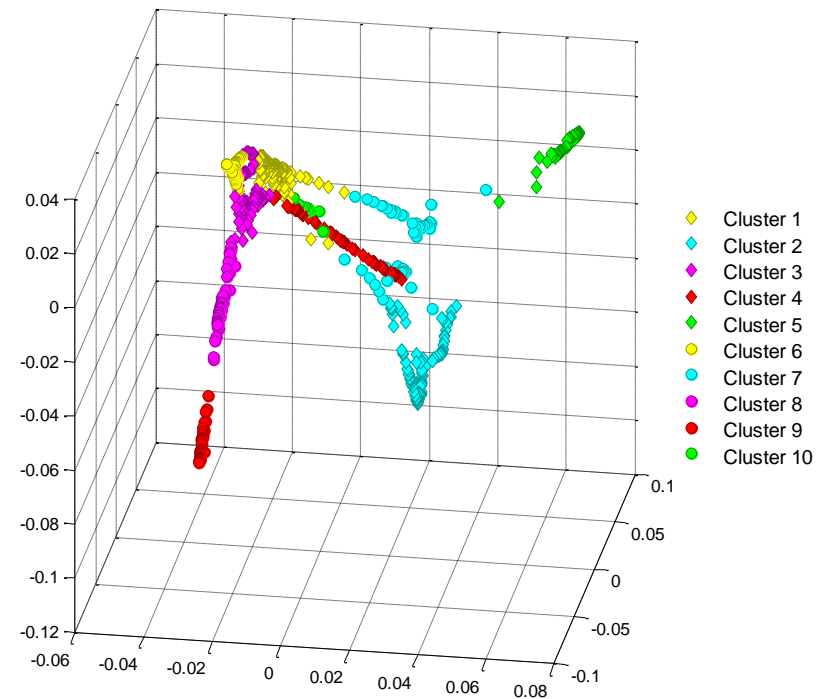
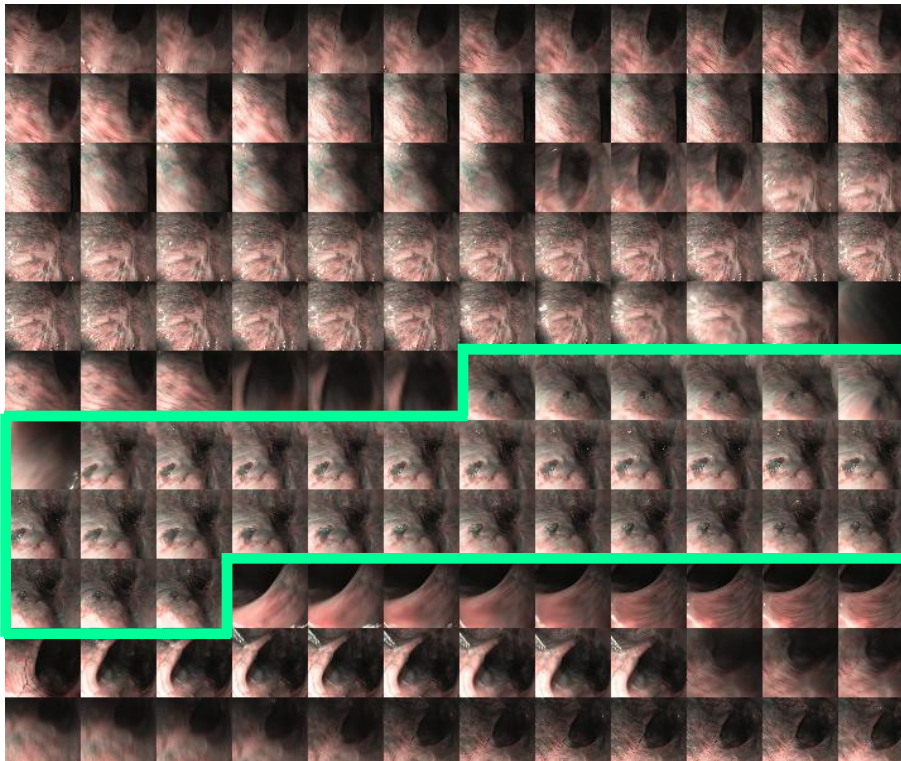
[Belkin2003: M. Belkin, P. Niyogi: Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, 15(6), 1373-1396. MIT Press, 2003]

[Atasoy2010: S. Atasoy, D. Mateus, J. Lallemand, A. Meining, G.Z. Yang, N. Navab: Endoscopic Video Manifolds, *MICCAI*, 2010.]

Endoscopic Video Manifolds

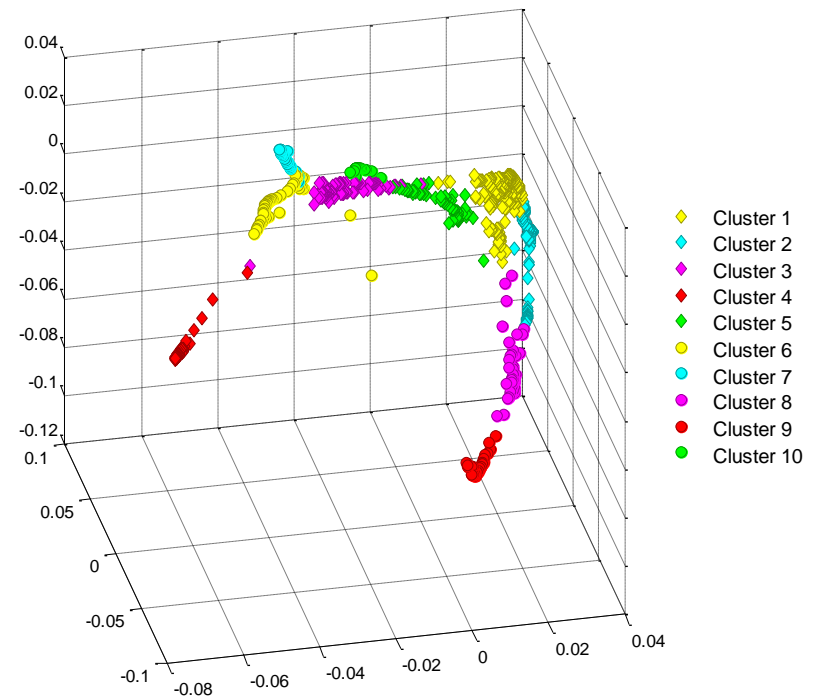
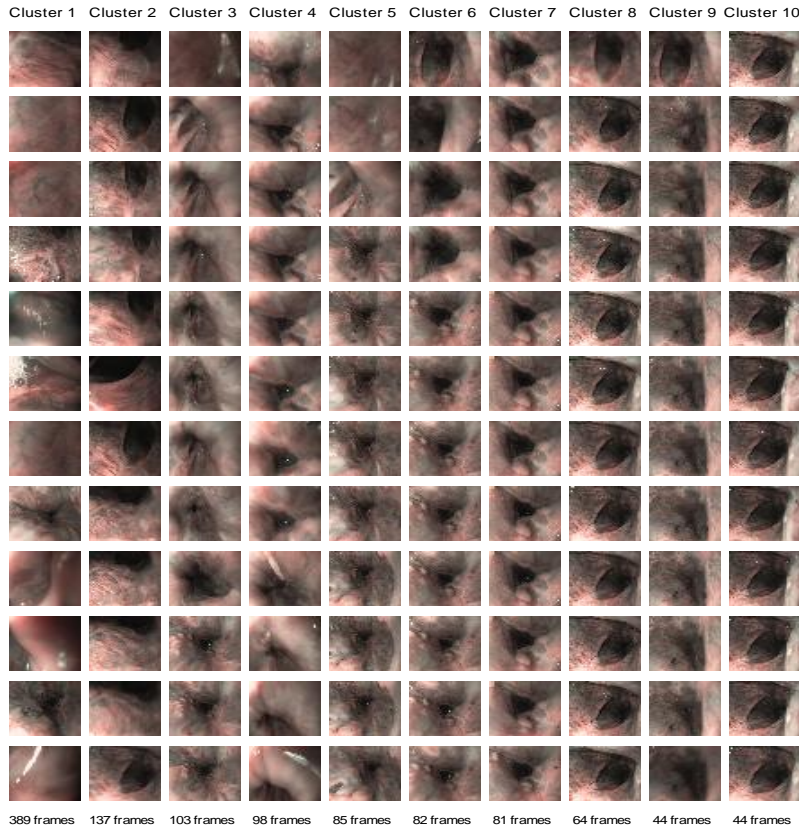
Clustering Endoscopic Scenes – Euclidean Distances

Cluster 3



Endoscopic Video Manifolds

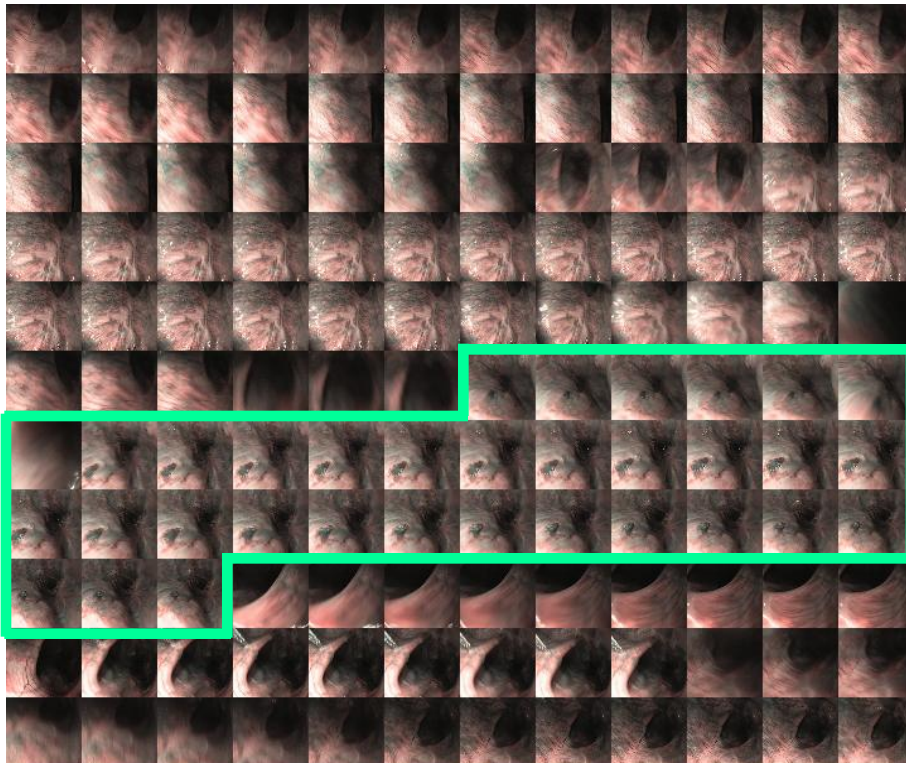
Clustering Endoscopic Scenes - NCC



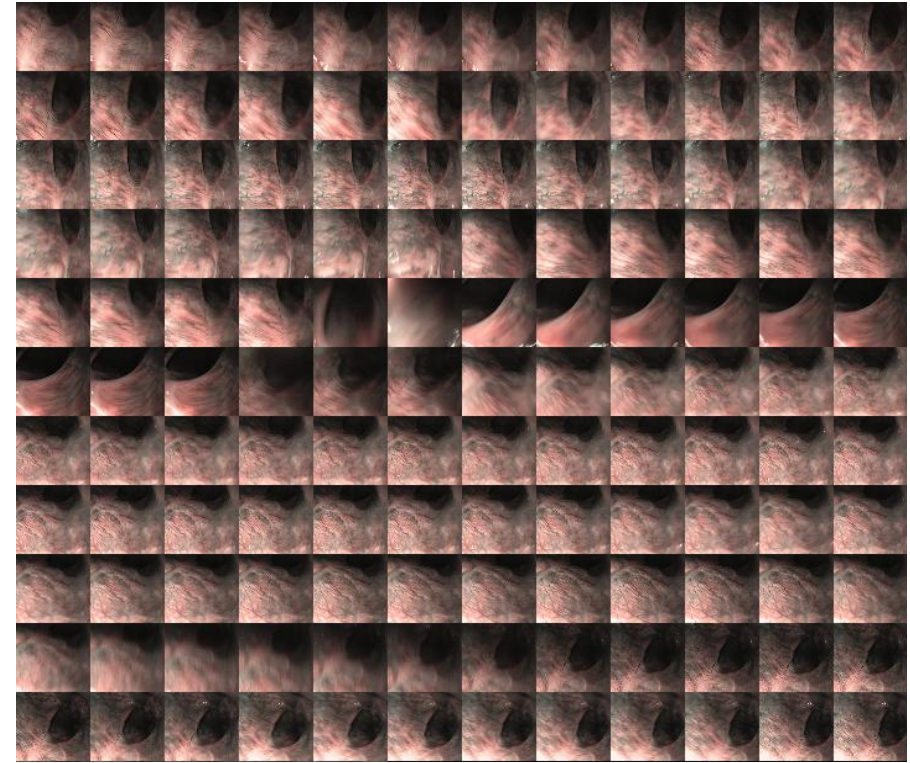
[Atasoy2010: S. Atasoy, D. Mateus, J. Lallemand, A. Meining, G.Z. Yang, N. Navab: Endoscopic Video Manifolds, MICCAI, 2010.]

Endoscopic Video Manifolds

Clustering Endoscopic Scenes - NCC



Euclidean Distance



Normalized Cross Correlation

[Atasoy2010: S. Atasoy, D. Mateus, J. Lallemand, A. Meining, G.Z. Yang, N. Navab: Endoscopic Video Manifolds, MICCAI, 2010.]

Endoscopic Video Manifolds

Clustering Endoscopic Scenes - NCC

Patient 1
Endoscopic Video Manifolds

Endoscopic Video Manifolds

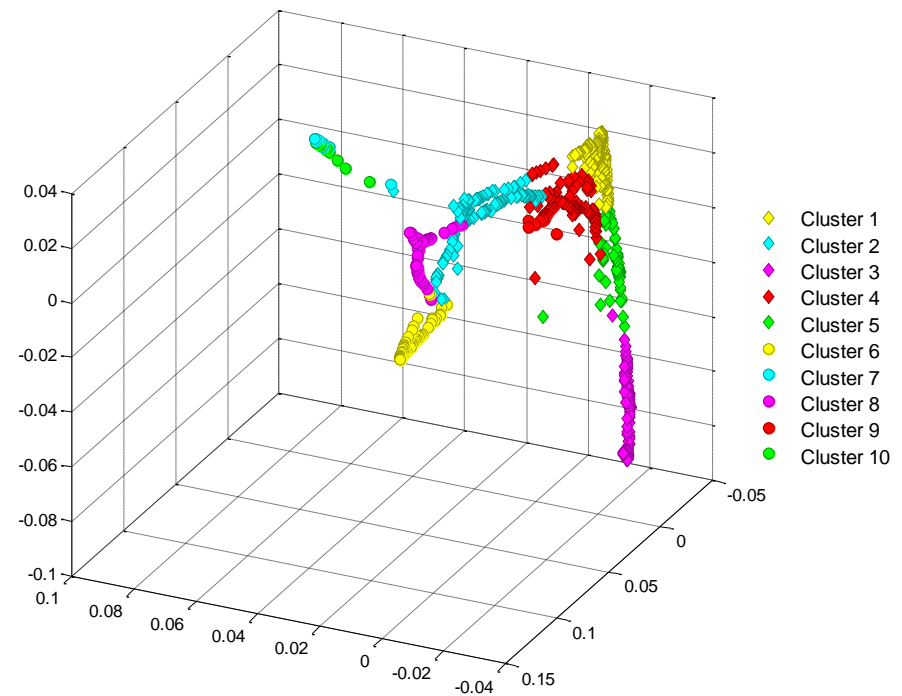
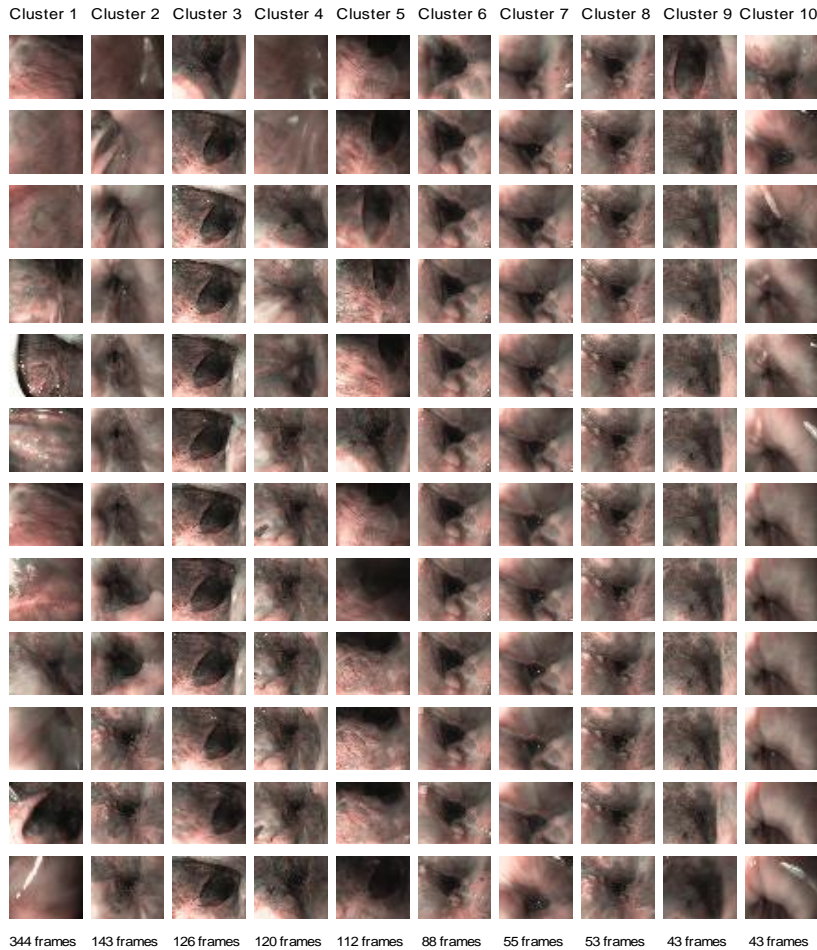
Clustering Endoscopic Scenes with Temporal Constraints

- Change the adjacency matrix to include temporal constraints

$$\mathbf{A}(i, j) = \begin{cases} 1 & \text{if } i \in \mathcal{N}_j^{\text{sim}} \\ 0 & \text{otherwise} \end{cases}$$

Endoscopic Video Manifolds

Clustering Endoscopic Scenes with Temporal Constraints



Patient 1

Endoscopic Video Manifolds

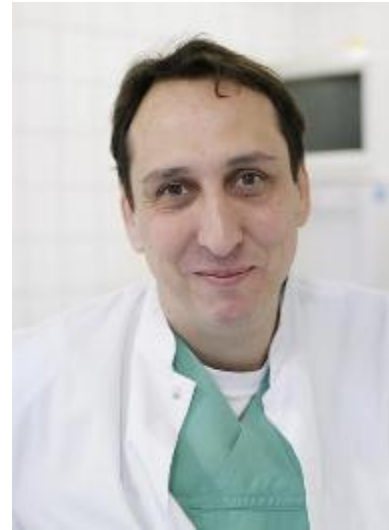
Acknowledgements



Prof. Nassir
Navab



Prof. Guang-Zhong
Yang



Prof. Alexander
Meining



Dr. Diana
Mateus

Thank you for your attention!!!

