Endoscopic Video Manifolds

Atasoy S.[1,2], Mateus D.[1], Lallemand J.[1], Meining A.[1], Yang G.Z.[2] and Navab N.[1]
Technical University of Munich[1], Imperial College London[2]
{atasoy, mateus, lalleman, navab}@in.tum.de, {catasoy, g.z.yang}@imperial.ac.uk

ICVSS 2010
Registration and Video Analysis

Abstract

In this work, we address two tasks: clustering of *poor-quality frames* and *different scenes* in endoscopic videos. For each task, our method creates a manifold representation using an appropriate inter-frame similarity measure and then performs a clustering on the created endoscopic video manifolds (EVMs). The introduced EVMs enable the clustering of poor-quality frames and grouping of different segments of the endoscopic video in an unsupervised manner. Furthermore, we present two novel inter-frame similarity measures for manifold learning to create structured manifolds from complex endoscopic videos.

Problem Statement

- An endoscopic video $\mathcal{I}$ can be represented by the set of its n individual frames $\{\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_n\}$.
- Each frame is a data point in the high dimensional input space $\mathcal{I}_1, \mathcal{I}_2, \cdots \mathcal{I}_n \in \mathbb{R}^{w \times h}$, where w and h are the width and height of the frames, respectively. Thus, the number of degrees of freedom (DoF) is equal to $w \times h$.
- Due to the continuity of the video sequence the actual DoF is much smaller. So, the high dimensional data points actually lie on a lower dimensional manifold $\mathcal{I}_1, \mathcal{I}_2, \cdots, \mathcal{I}_n \in \mathcal{M}$, where $\mathcal{M}$ is a manifold embedded in $\mathbb{R}^{w \times h}$.

We propose to represent the endoscopic video by this embedded low dimensional manifold and perform the two tasks of clustering poor-quality frames and endoscopic scenes on the introduced EVMs.

Methods

Defining the similarities: For each pair $(\mathcal{I}_i, \mathcal{I}_j)$, of the given n data points $i, j \in \{1, \cdots, n\}$, first a similarity measure is defined $W : \mathcal{I} \times \mathcal{I} \to \mathbb{R}$.

Computing the adjacency graph: Given the similarity matrix W, first, k-nearest neighbours of each data point are computed. Then, the adjacency graph is created as:

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

where $\mathcal{N}_j^k$ states the k-nearest neighbours of the j-th data point.

Learning the manifold: The intrinsic low dimensional sstructure is learned using the Laplacian Eigenmaps (LE) method [1].

Clustering on the manifold: Finally, the clustering is using the K-means algorithm [2].

Inter-frame Similarities

1. Clustering Poor-Quality Frames:
For this task, EVM is created using the following similarity measure:

$$W_{\text{EH}}(\mathcal{I}_i, \mathcal{I}_j) = 1 - \frac{acos\left(\langle(\mathcal{F}_i, \mathcal{F}_j\rangle\right)}{||\mathcal{F}_i|| \cdot ||\mathcal{F}_j||}\ , \qquad (2)$$

where $\mathcal{F}_i$ states the rotation invariant energy histogram computed from the power spectrum of frame $\mathcal{I}_i$.

2. Clustering Endoscopic Scenes:
For this task we propose two different similarity measures:

$$W_{\text{DOFF}}(\mathcal{I}_i, \mathcal{I}_j) = 1 - \frac{\psi_i^j}{\max(\psi_i^j)}, \qquad (3)$$

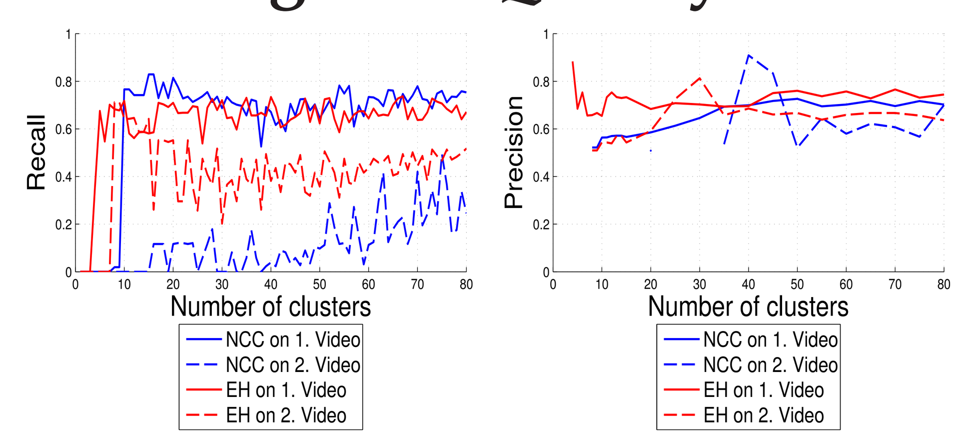
$$\psi_i^j = \sum_{x=1}^{w} \sum_{y=1}^{h} |\nabla \Phi_i^j(x,y)|,$$

where $\Phi_i^j(x,y)$ states the optical flow from frame $\mathcal{I}_i$ to frame $\mathcal{I}_j$ and,

$$W_{\text{NCC}}(\mathcal{I}_i, \mathcal{I}_j) = NCC(\mathcal{I}_i, \mathcal{I}_j) \qquad (4)$$

Experiments

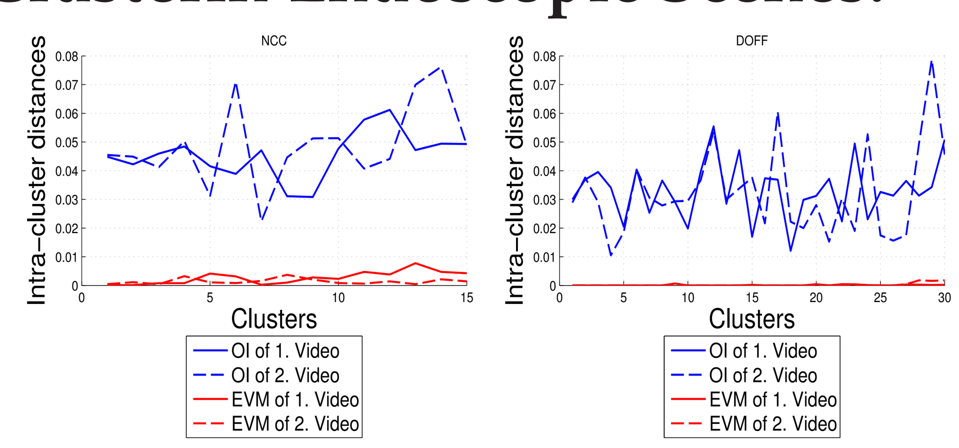
Clustering Poor-Quality Frames:



(a) (b)

Figure 1: a) Recall and b) precision values for clustering uninformative frames.

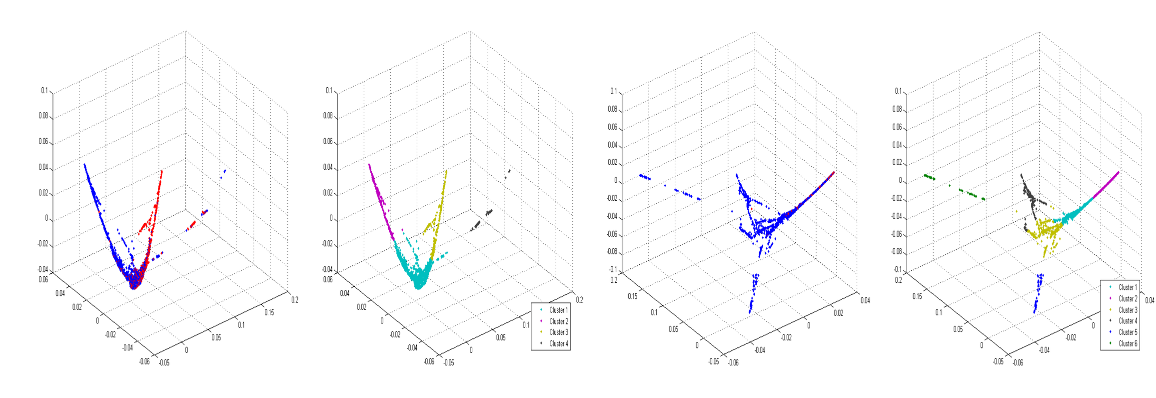
Clusterin Endoscopic Scenes:



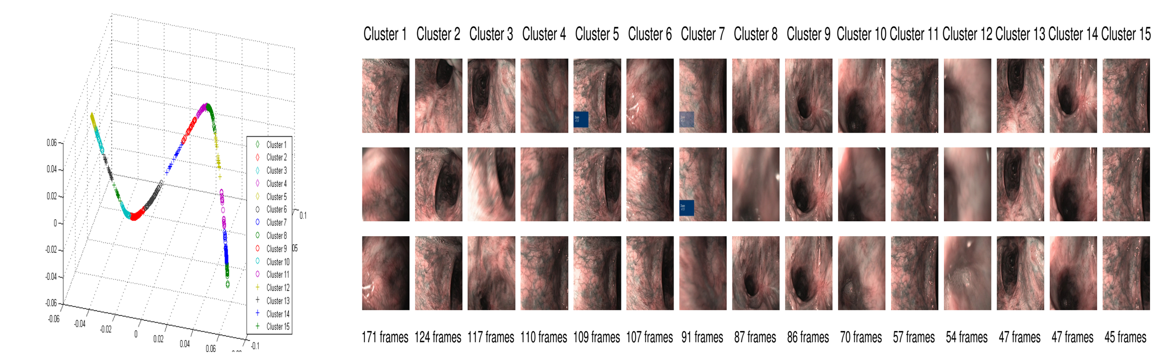
(a) (b)

Figure 2: Intra-cluster distances for clustering endoscopic scenes. a) W_{NCC} and b) W_{DOFF}.

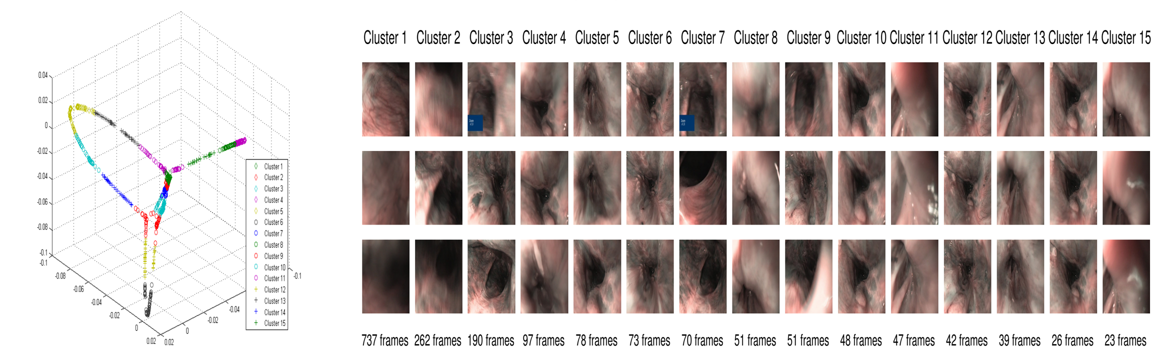
Results



(a)



(b)



(c)

Figure 3: Clustering of (a) poor-quality frames using W_{EH} and (b-c) endoscopic scenes using W_{DOFF} and W_{NCC}, respectively.

References

[1] M. Belkin, P. Niyogi, Laplacian eigenmaps for dimensionality reduction and data representation. *Neural computation*, (2003)
[2] J. Hartigan, M. Wong, A k-means clustering algorithm. *JR Stat. Soc.*, (1979)