

# Iterative Projection and Matching: Finding Structure-preserving Representatives and Its Application to Computer Vision

Alireza Zaeemzadeh\*, Mohsen Joneidi\*, Nazanin Rahnavard, and Mubarak Shah

\*indicates shared first authorship.

## Overview

- Goal:** Finding Structure-preserving representatives from a set of data.
- Main characteristics of IPM:**
  - ✓ Linear complexity w.r.t. the number of data.
  - ✓ No parameters to be tuned.

## Proposed Algorithm: IPM

- Given  $M$  data points  $\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M \in R^N$ ,  

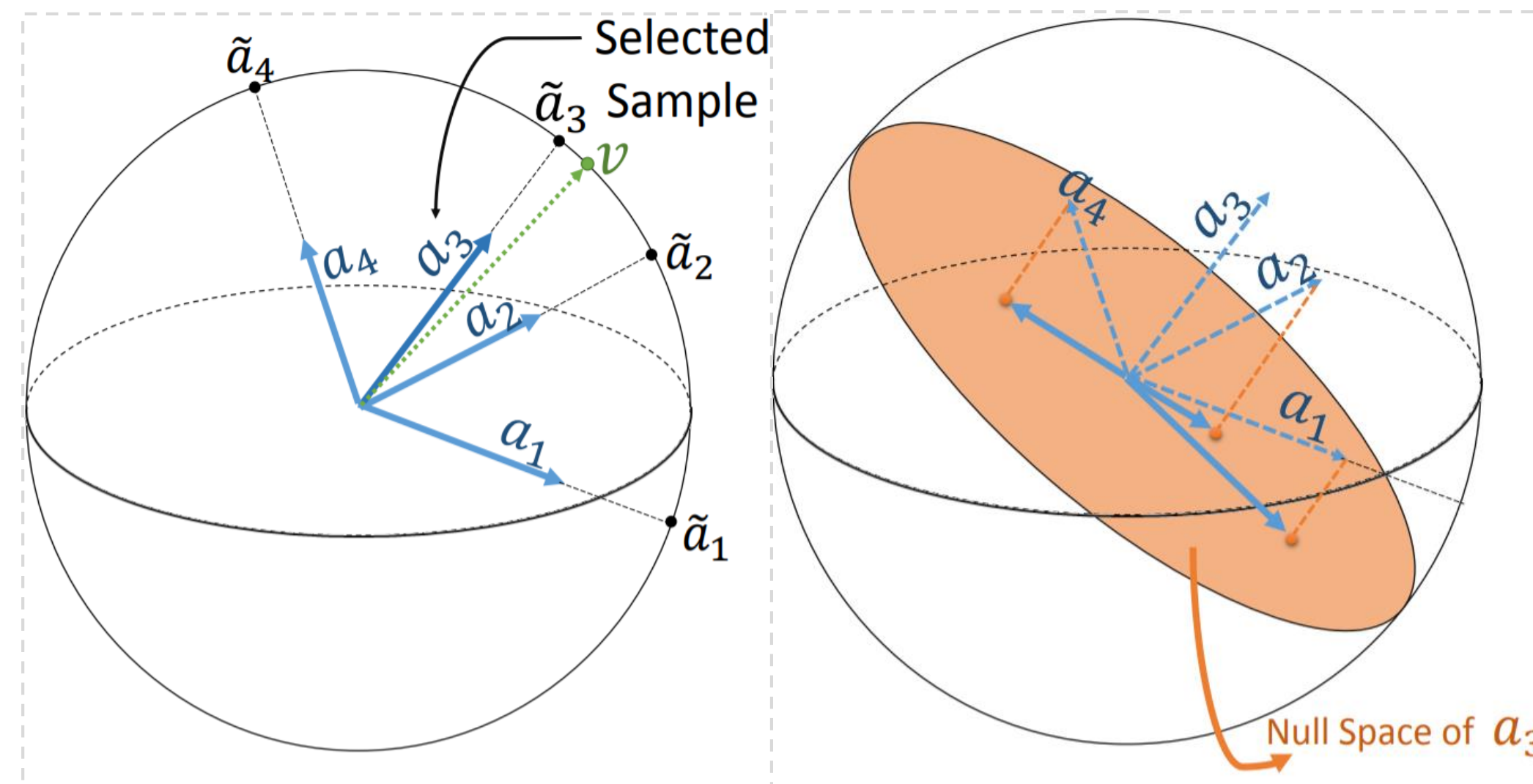
$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_M^T \end{bmatrix}$$
- Projection onto the subspace spanned the  $K$  rows:  

$$\arg\min_{\mathbf{U}, \mathbf{V}} \|\mathbf{A} - \mathbf{UV}^T\|_F^2 \text{ s.t. } v_k \in A$$
- Selecting only 1 data point:

$$(\mathbf{u}, \mathbf{v}) = \arg\min_{\mathbf{u}, \mathbf{v}} \|\mathbf{A} - \mathbf{uv}^T\|_F^2 \text{ s.t. } \|\mathbf{v}\| = 1, \quad (1)$$

$$\mathbf{m}^{(1)} = \arg\max_m |\mathbf{v}^T \mathbf{a}_m|. \quad (2)$$

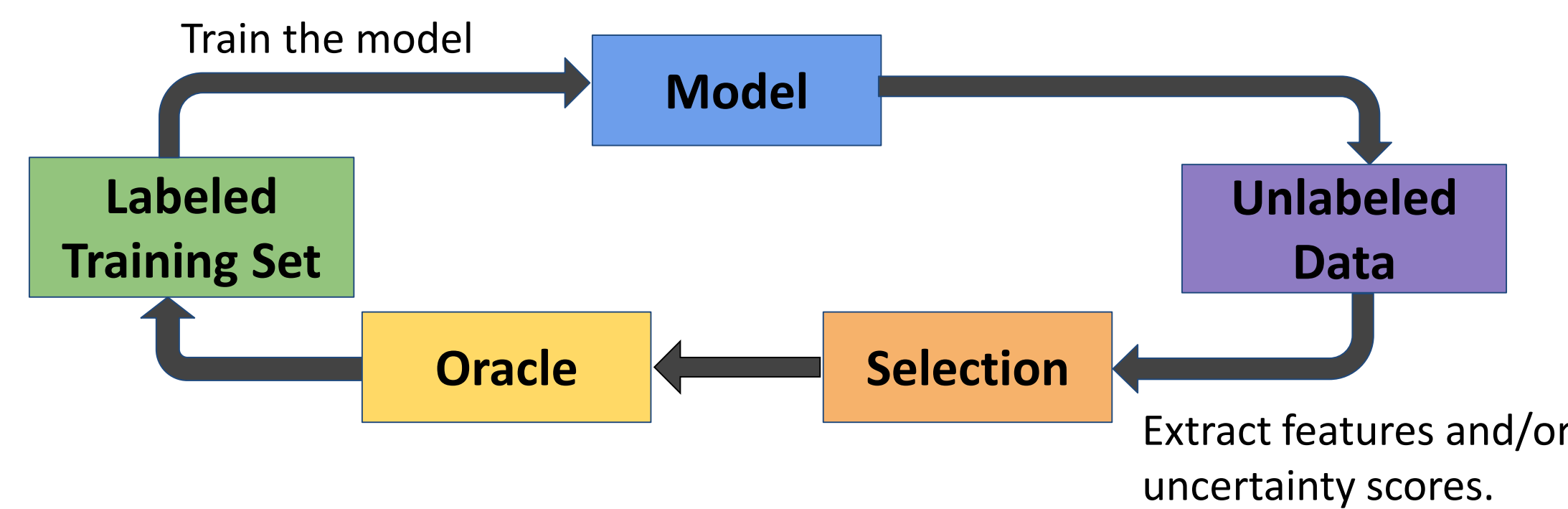
- The captured information is neglected by projection on the null-space of previously selected samples



## Theoretical Guarantees

- Proposition 1** There exists at least one data point such that its correlation with the first right singular vector of is greater than or equal to  $\frac{\sigma_1}{\|\mathbf{A}\|_F}$ .
- Proposition 2** If the gap between consecutive eigenvalues of a matrix is decreasing, then its first eigenvector is the most robust spectral component against changes in the data.

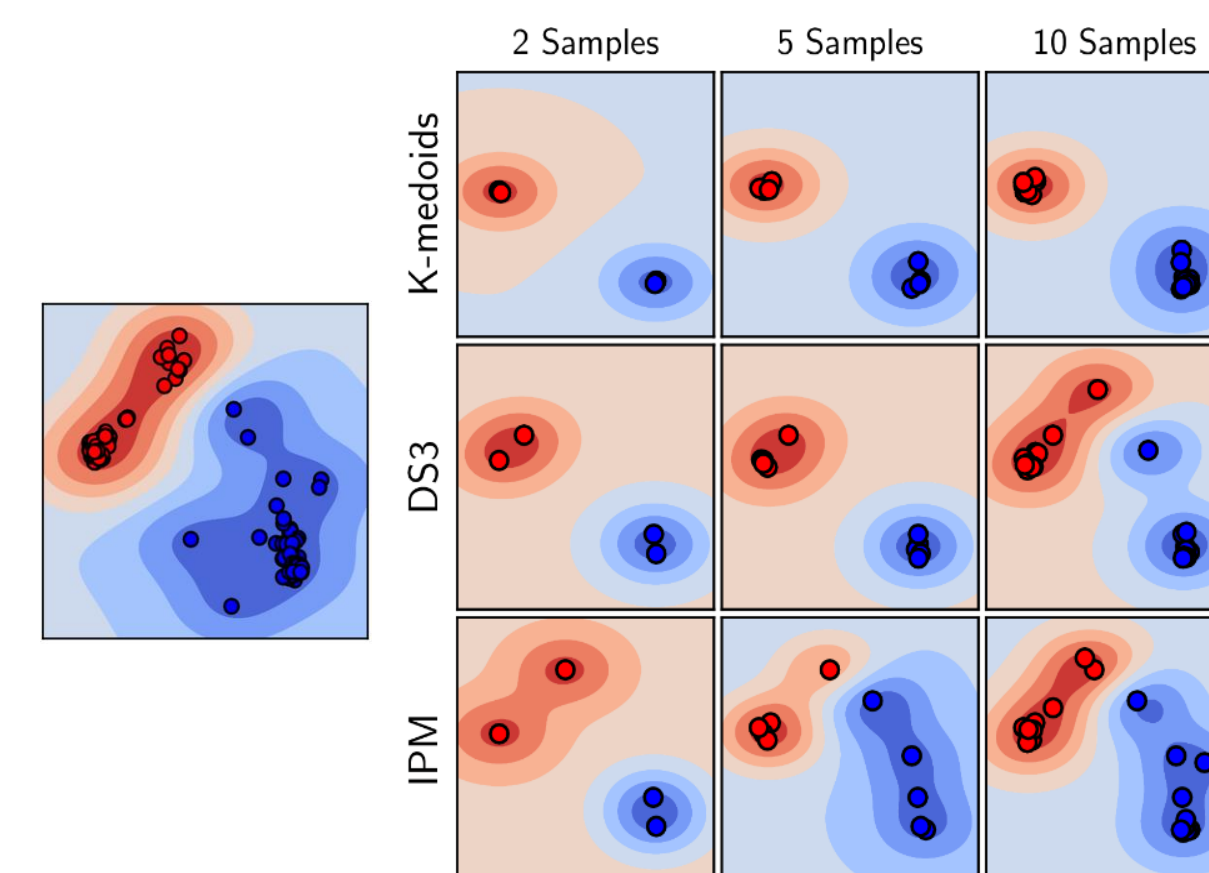
## Active Learning



### Dataset: UCF-101

Average Sample per class	2	3	4	5	6	7
Random	60.1	65.1	68.2	69.9	71.7	73.0
K-medoids	60.1	65.3	68.4	69.2	72.3	73.6
DS3 [1]	64.0	66.5	67.8	68.3	69.6	70.9
Uncertainty [3]	59.5	66.7	69.4	71.5	73.9	75.5
IPM	<b>64.6</b>	68.7	72.2	73.4	74.3	74.7
IPM + Uncertainty	64.3	<b>79.4</b>	<b>72.8</b>	<b>73.8</b>	<b>76.2</b>	<b>76.3</b>

T-SNE visualization of two randomly selected classes of UCF-101 dataset.



## Learning Using Representatives

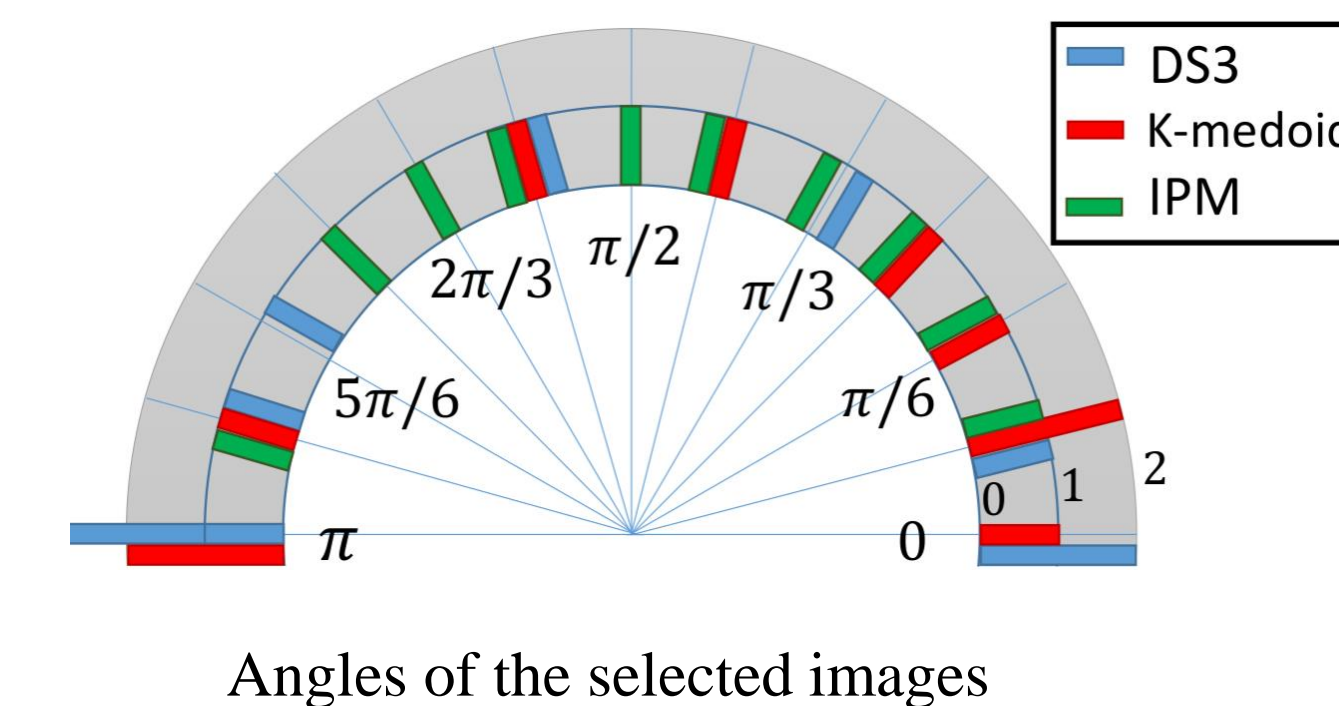
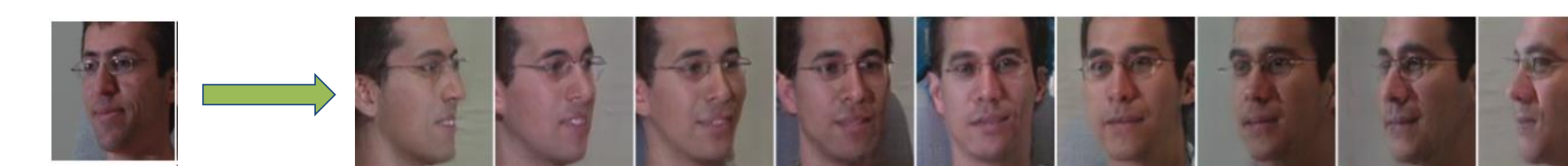
- Dataset: ImageNet**
- Selection methods based on convex-relaxation fail to run in a tractable time.



Classification accuracy (%) using k-NN.				
Images per class	1	5	10	50
Random	3.1	8.7	12.9	25.6
K-medoids	11.7	17.0	17.5	26.8
IPM	<b>12.5</b>	<b>21.6</b>	<b>25.2</b>	<b>30.7</b>

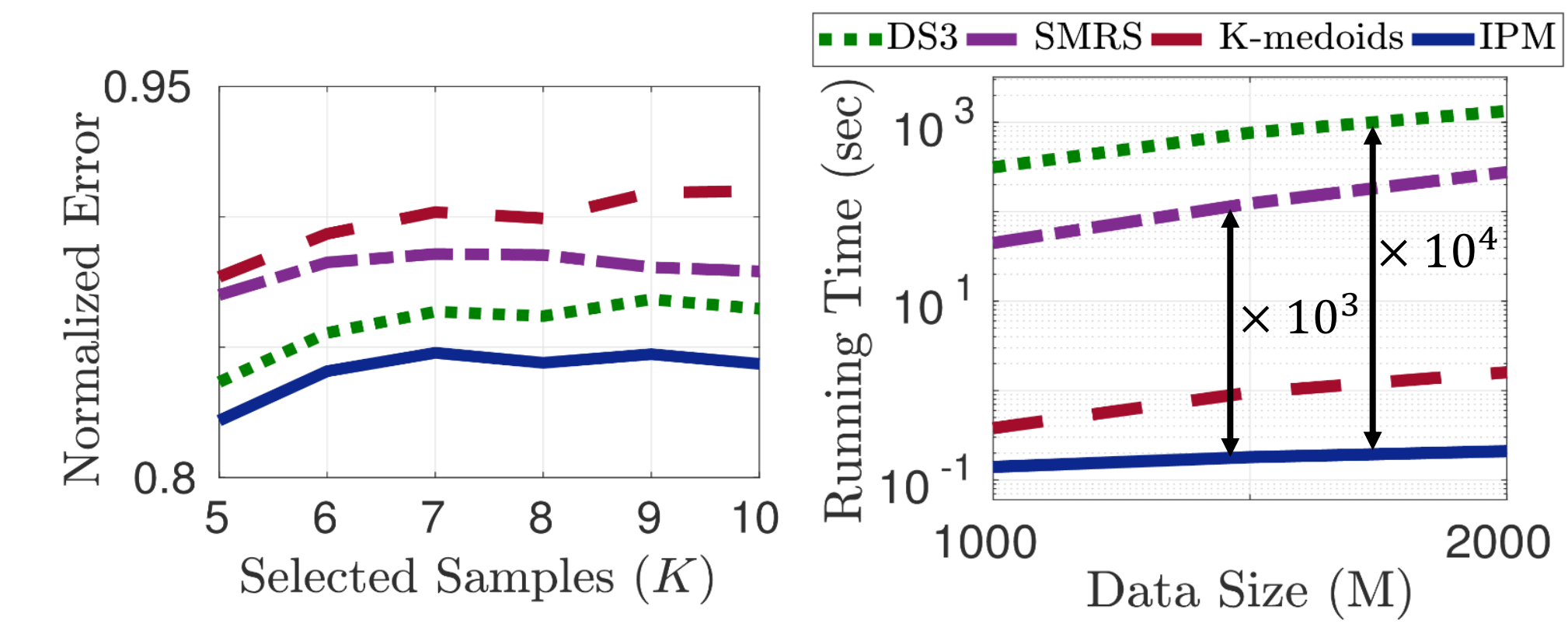
### Dataset: Multi-PIE face dataset

- Multi-view face generation using only 9 images per subject.



Identity dissimilarities between real and generated images

Method	Random	K-medoids	DS3[1]	IPM
9 images/subject	0.561	0.599	0.602	<b>0.553</b>
360 imaged/subject	0.536			



## Video Summarization

- Dataset: UT Egocentric**
- IPM is a close competitor to the supervised methods.



F-measure and recall scores using ROUGE-SU metric

Method	F-measure	Recall
Selection Methods (Unsupervised)		
Random	26.3	23.7
Uniform	28.7	25.8
K-medoids	30.1	27.3
DS3 [1]	30.1	27.3
IPM	<b>31.53</b>	<b>29.1</b>
Supervised Methods		
SeqDPP	28.9	26.8
Sub-Mod	29.3	27.4
Sub-Mod+	34.1	31.6

## References

- [1] E. Elhamifar et. al. "Dissimilarity based sparse subset selection". PAMI 2016.
- [2] E. Elhamifar et. al. "See all by looking at a few: Sparse modeling for finding representative objects". CVPR 2012
- [3] Y. Gal, R. Islam et. al. "Deep Bayesian Active Learning with Image Data". PMLR 2017