Robust Principal Component Analysis on Graphs

Nauman Shahid

Vassilis Kalofolias, Xavier Bresson, Michael Bronstein & Pierre Vandergheynst

EPF Lausanne, Switzerland

SPARS 2015, Cambridge UK

What is this talk about?

Sparsity?

SPARSITY!!!!

NO!

True for most of the people in this auditorium 

Unfortunately "sparsity" in this talk represents "gross errors"

This talk is about Principal Component Analysis / Low-rank representation: The most widely used tool for linear dimensionality reduction & clustering!

Outline

Principal Component Analysis
Robust Principal Component Analysis

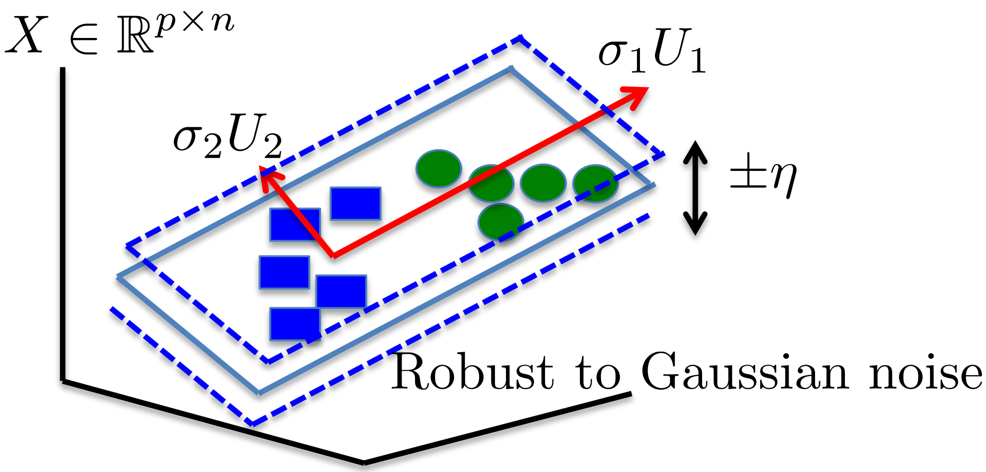
Robust Principal Component Analysis on Graphs
How and Why?

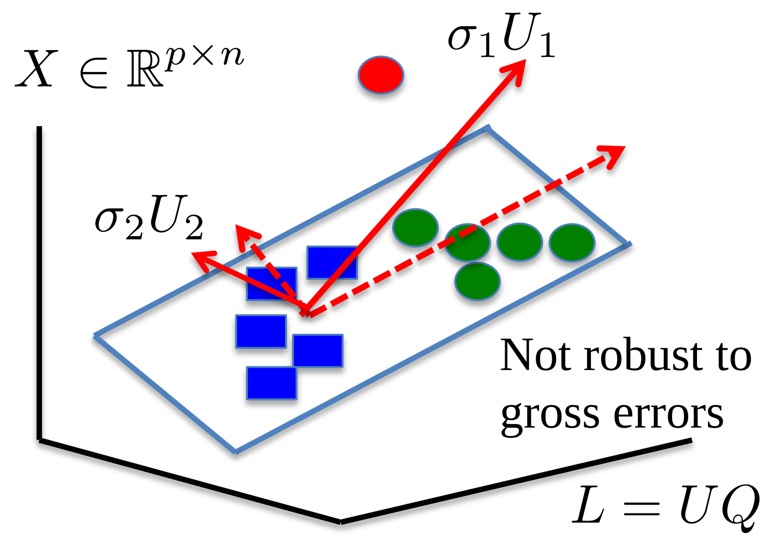
Solution using ADMM

Results:
Clustering in low-dimensional space
Low-rank recovery

Robust Principal Component Analysis

Original data space





$$\min_{U,Q} \|X - UQ\|_F^2 \quad \text{s.t.} \quad U^\top U = I$$

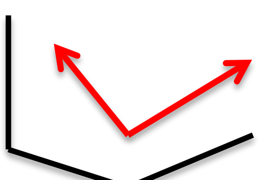
solution: $SVD(X) = U\Sigma Q'$

$$U = [U_1 \mid U_2] \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix} Q' = \begin{bmatrix} Q_1'^\top \\ Q_2'^\top \end{bmatrix}$$

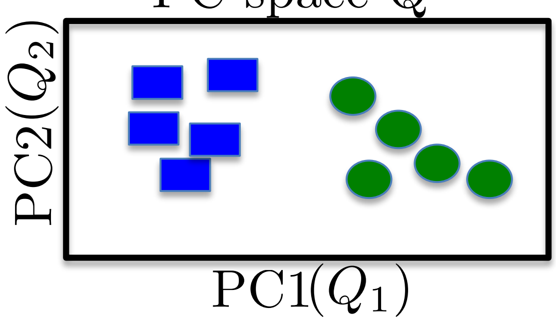
$$\min_{L,S} \|L\|_* + \lambda\|S\|_1 \quad \text{s.t.} \quad X = L + S$$

solved using ADMM

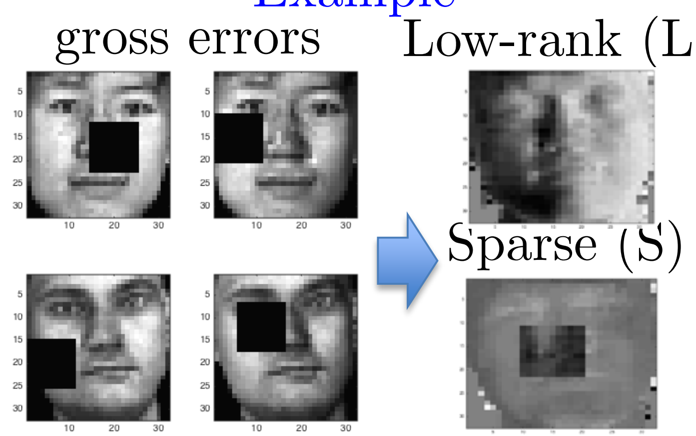
Principal directions / basis



PC space Q

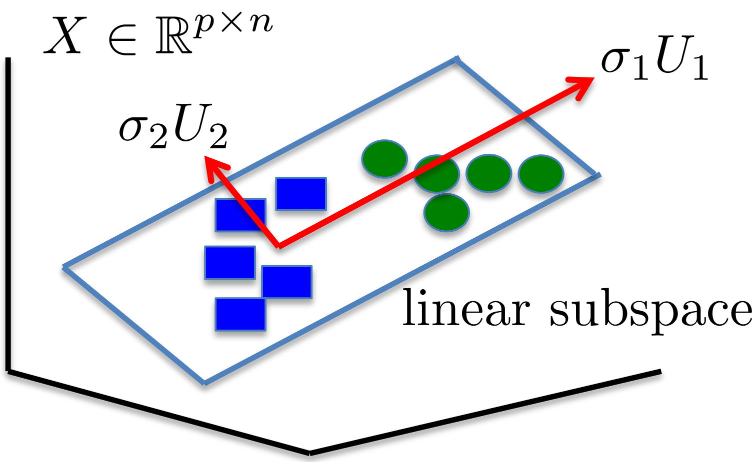


Example



Robust Principal Component Analysis on Graphs

Original data space



standard form

$$\min_{U,Q} \|X - UQ\|_F^2$$
$$\text{s.t. } U^\top U = I$$

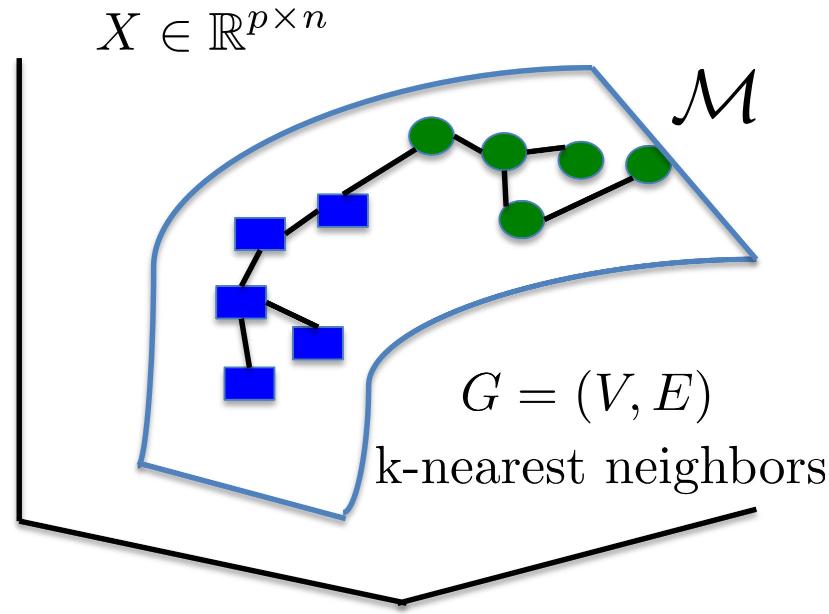
Robust PCA

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1$$
$$\text{s.t. } X = L + S$$

$$L = UQ$$

How & why?

Original data space



non-linear smooth manifold

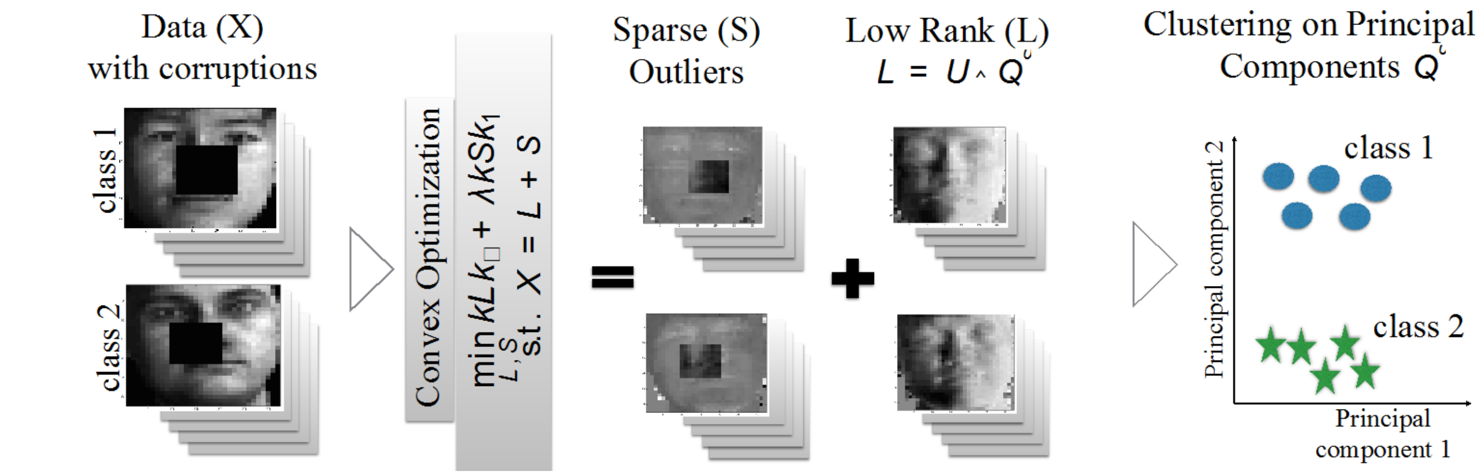
characterized via a discrete graph G

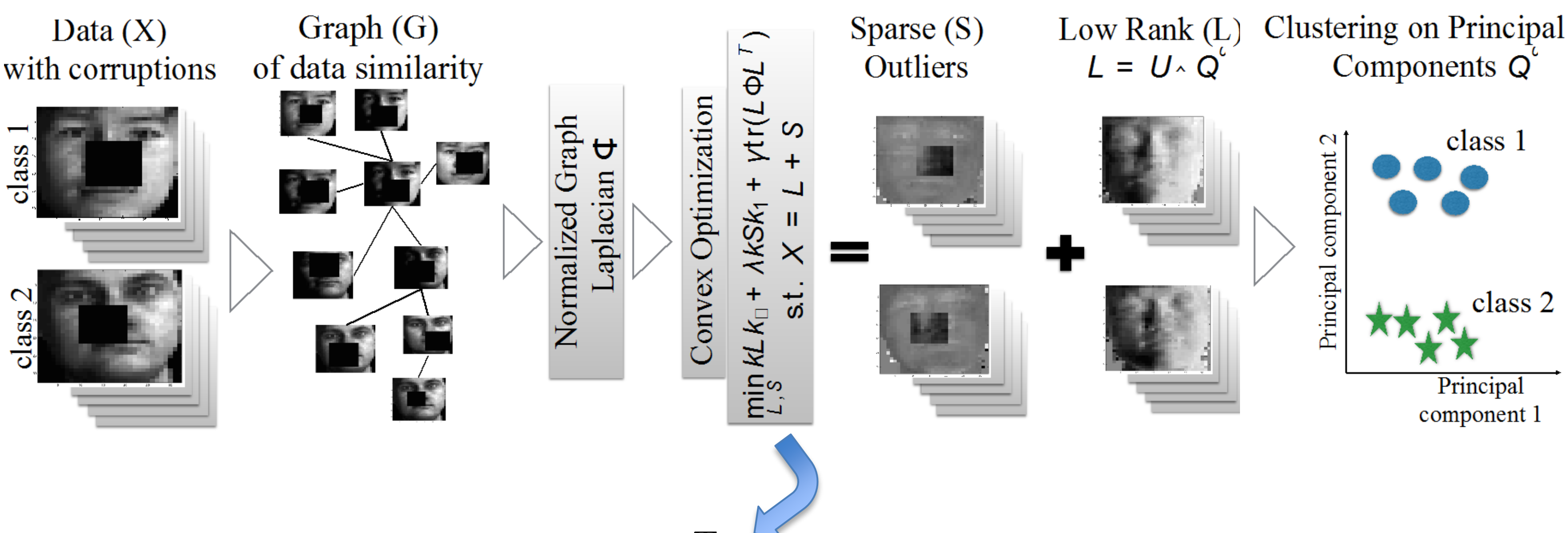
Edge weights: $W_{i,j} = \exp(-\frac{\|x_i - x_j\|^2}{\sigma^2})$

$\Phi = D - W$ unnormalized Laplacian

$\Phi = I - D^{-1/2} W D^{-1/2}$ normalized

Robust Principal Component Analysis on Graphs: How?





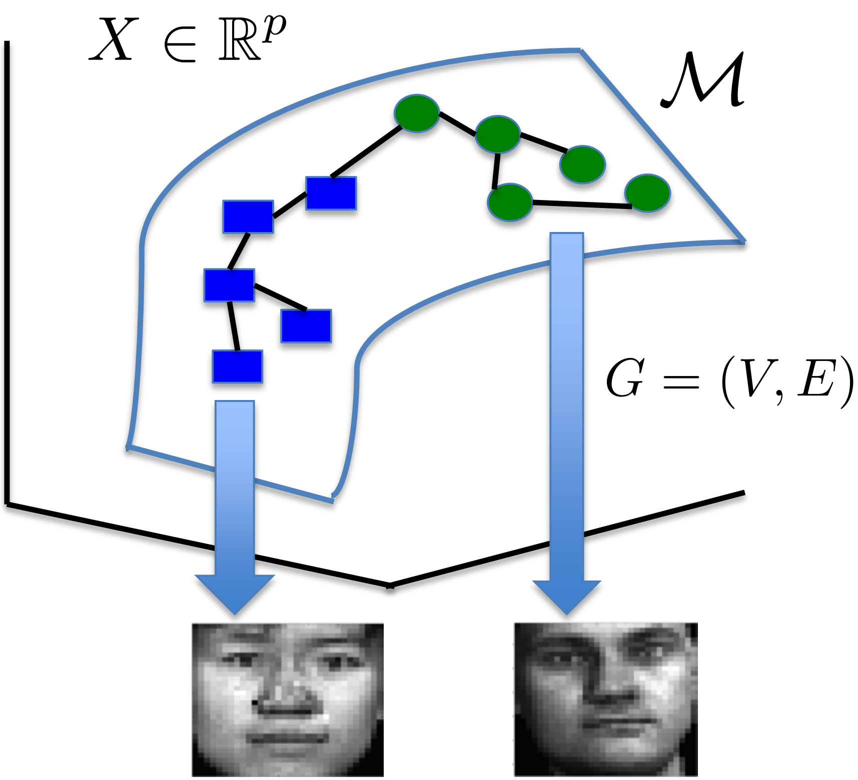
$$\min_{L,S} \|L\|_* + \lambda\|S\|_1 + \gamma \mathrm{tr}(L\Phi L^T) \quad \text{s.t.} \quad X = L + S$$

Robust Principal Component Analysis on Graphs

Why?

$$\min_{L,S} \|L\|_* + \lambda \|S\|_1 + \gamma \operatorname{tr}(L \Phi L^T) \quad \text{s.t.} \quad X = L + S$$

Original data space



non-linear smooth manifold

Exploit local similarity information

- Enhanced low-rank representation, smooth on the manifold
- Better clusters in low-dimensional space

Two important applications of PCA!

IMPORTANT: Smooth low-rank, NOT only principal components

convexity

Solution using ADMM

$$\min_{L,S} \|L\|_* + \lambda\|S\|_1 + \gamma \mathrm{tr}(L\Phi L^T) \qquad \text{s.t.} \;\; X = L + S$$

Augmented Lagrangian

$$(L,S,U)^{k+1} = \min_{L,S,U} \|L\|_* + \lambda\|S\|_1 + \gamma \mathrm{tr}(U\Phi U^T) + \langle Z_1, X - L - S\rangle + \frac{r_1}{2}\|X - L - S\|_F^2 + \langle Z_2, U - L\rangle + \frac{r_2}{2}\|U - L\|_F^2$$

$$Z_1^{k+1} = Z_1^k + r_1(X - L - S)^{k+1}, Z_2^{k+1} = Z_2^k + r_2(U - L)^{k+1}$$

$$L^{k+1} = \mathrm{prox}_{1/(r_1+r_2)\|L\|_*}\left(\underbrace{\frac{r_1(X - S + Z_1/r_1) + r_2(U + Z_2/r_2)}{r_1 + r_2}}_{SVD\,:\,A\Omega B}\right)$$

singular value soft-thresholding

$O(pn^2)$

$$S^{k+1} = \mathrm{prox}_{\frac{\lambda}{r_1}\|S\|_1}\left(\underbrace{X - L^{k+1} + \frac{Z_1}{r_1}}\right)$$

constant

element wise soft-thresholding

$$U^{k+1} = \underbrace{r_2(\gamma H + r_2 I)^{-1}\left(L^{k+1} - \frac{Z_2}{r_2}\right)}$$

$O(I|E|n)$

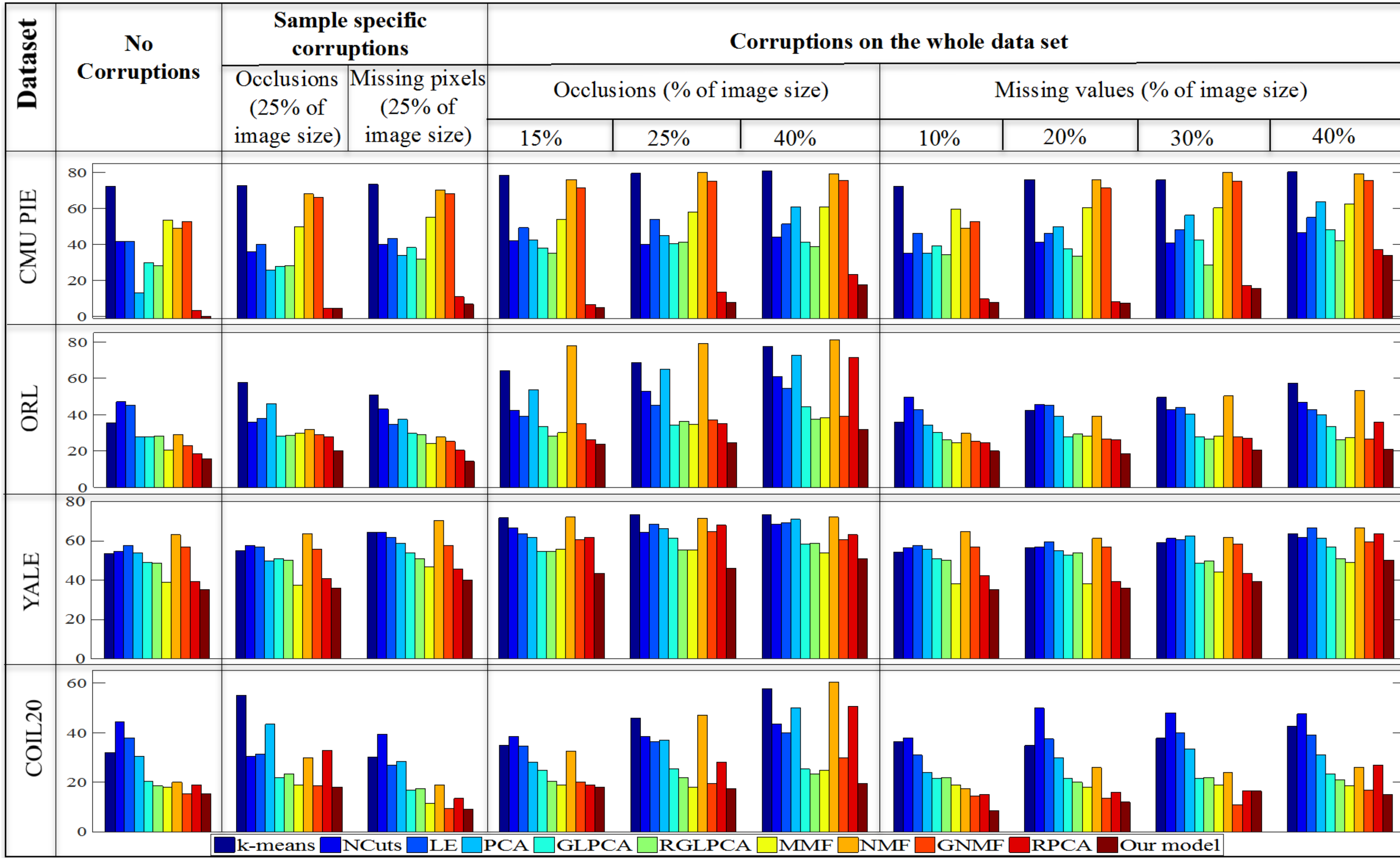
system of linear equations

$p < n$

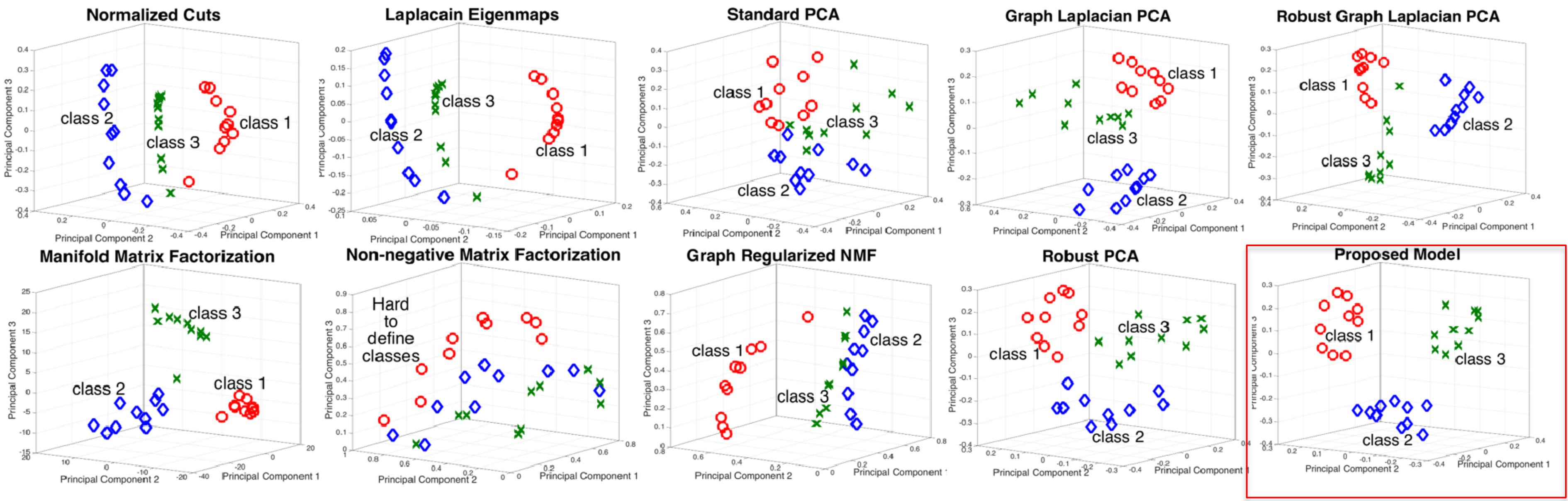
Results

- Clustering in low-dimensional space
- Static background separation from dynamic foreground

Results: Clustering



Results: Principal Components for 3 classes of ORL dataset



Results: Background Separation from Videos

Please have a look at the
full version of the paper on arXiv
Demo & code available at:
https://lts2.epfl.ch/blog/nauman/recent-projects/

Fast version and its code will be available in 2 weeks.

Future Work: Theoretical investigation