A Nonconvex Projection Method for Robust PCA*

Aritra Dutta⋆, Filip Hanzely⋆, Peter Richtárik⋆,Ω,α

King Abdullah University of Science and Technology⋆, University of EdinburghΩ, MIPTα

Robust Principal Component Pursuit

Robust principal component pursuit is a robust matrix decomposition model in which we wish to decompose A into the sum of a low-rank matrix L and an error matrix S: A = L + S.

Data matrix   Low-rank matrix   Sparse matrix

The celebrated principal component pursuit uses the ℓ_0 norm to address the sparsity and one obtains

$$\min_{L,S} \operatorname{rank}(L) + \lambda\|S\|_{\ell_0} \quad \text{subject to} \quad A = L + S,$$

Different RPCA formulations by using surrogate constraints and optimization objective functions. Wright et al. 2009; Lin, Chen, and Ma 2010; Cande`s et al. 2011 proposed the celebrated Robust Principal Component Analysis as follows:

RPCA

$$\min_{L,S} \|L\|_* + \lambda\|S\|_{\ell_1} \quad \text{subject to} \quad A = L + S.$$

Moving the constraint to the objective as a penalty, together with adding explicit constraints on the target rank r and target sparsity s leads to the following formulation (Zhou and Tao 2011) as:

Matrix Completion

$$\min_{L,S} \|A - L - S\|_F^2 \quad \text{subject to} \quad \operatorname{rank}(L) \le r \text{ and } \|S\|_0 \le s.$$

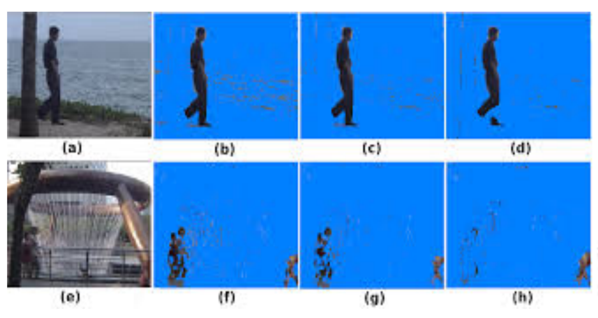
Cherapanamjeri, Gupta, and Jain 2017; Chen et al. 2011; Tao and Yang 2011 proposed the robust matrix completion (RMC) problem as:

Robust Matrix Completion

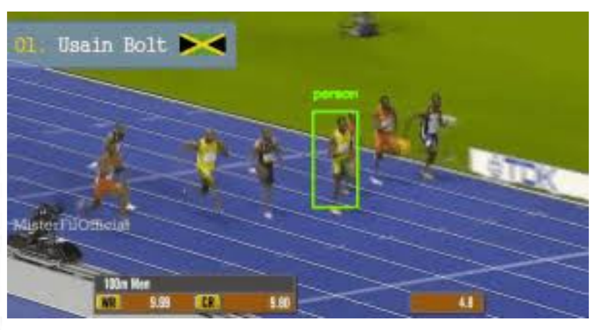
$$\min_{L,S} \|\mathcal{P}_\Omega(A - L - S)\|_F^2 \quad \text{subject to} \quad \operatorname{rank}(L) \le r \text{ and } \|\mathcal{P}_\Omega(S)\|_0 \le s'.$$

where $(\mathcal{P}_\Omega[X])_{ij} = \begin{cases} X_{ij} & (i,j) \in \Omega \\ 0 & \text{otherwise.} \end{cases}$

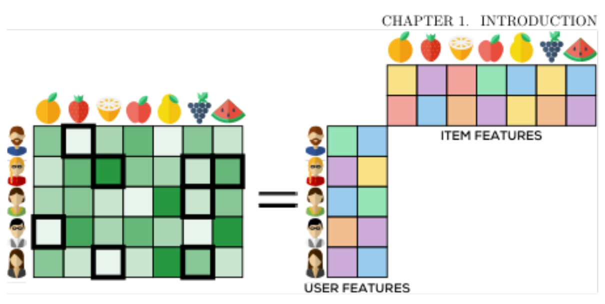
Applications to Real-World Problems



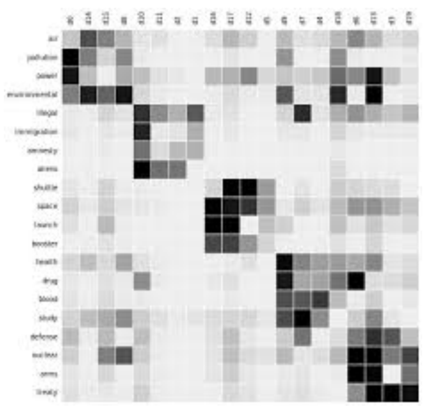
Background estimation



Tracking



Recommender systems



Latent Semantic Indexing

And many more…

1. EEG images
2. fMRI images
3. Sparse Subspace clustering
4. Hyperspectral Image denoising
5. Contrast-filled vessels extraction
6. …

Nonconvex Feasibility and Alternating Projections

Set feasibility:
Find point in the intersection of closed sets

$$\text{Find } x \in \mathcal{X}; \quad \mathcal{X} = \cap_{i=1}^m \mathcal{X}_i$$

A natural alternating projection algorithm to solve set feasibility is Algorithm 1.

Now we consider the following reformulation of RPCA:

$$\text{Find} \quad M \stackrel{\text{def}}{=} [L, S] \in \mathcal{X} \stackrel{\text{def}}{=} \cap_{i=1}^3 \mathcal{X}_i \neq \emptyset,$$

where

$$\mathcal{X}_1 \stackrel{\text{def}}{=} \{M \mid L + S = A\}$$
$$\mathcal{X}_2 \stackrel{\text{def}}{=} \{M \mid \operatorname{rank}(L) \le r\}$$
$$\mathcal{X}_3 \stackrel{\text{def}}{=} \{M \mid \|S_{(i,:)}\|_0 \le \alpha n \text{ and } \|S_{(:,j)}\|_0 \le \alpha m \text{ for all } i \in [m], j \in [n].\}$$

Algorithm 1: Alternating projection method for set feasibility

Input : $\Pi_i(\cdot)$ – Projector onto $\mathcal{X}_i$ for each $i \in \{1, \dots m\}$, starting point x_0
for $k = 0, 1, \dots$ **do**
  Choose via some rule i (e.g., cyclically or randomly)
  $x_{k+1} = \Pi_i(x_k)$
end
Output : x_{k+1}

Algorithm 2: Alternating projection method for RPCA

Input : $A \in \mathbb{R}^{m \times n}$ (the given matrix), rank r, sparsity level $\alpha \in (0, 1]$
Initialize : L_0, S_0
for $k = 0, 1, \dots$ **do**
  $\tilde{L} = \frac{1}{2}(L_k - S_k + A)$
  $\tilde{S} = \frac{1}{2}(S_k - L_k + A)$
  $L_{k+1} = H_r(\tilde{L})$
  $S_{k+1} = \mathcal{T}_\alpha(\tilde{S})$
end
Output : L_{k+1}, S_{k+1}

Projection of $[L_0, S_0]$ onto set of $[L, S]$ such that $L + S = A$
$$[\tfrac{1}{2}(L_0 - S_0 + A), \tfrac{1}{2}(S_0 - L_0 + A)]$$

Projection of L_0 onto rank r constraint
Operator H_r – rank r SVD of L_0

Projection of S_0 on the sparsity constraint (approximate)
Keep largest α fraction of values in each row and each column

The next Algorithm is proposed to solve the RMC problem:

Algorithm 3: Alternating projection method for RMC

Initialize : L_0, S_0
for $k = 0, 1, \dots$ **do**
  $\tilde{L} = \frac{1}{2}\mathcal{P}_\Omega(L_k - S_k + A)$
  $\tilde{S} = \frac{1}{2}\mathcal{P}_\Omega(S_k - L_k + A)$
  $L_{k+1} = H_r(\tilde{L} + \mathcal{P}_{\Omega^c}(L_k))$
  $S_{k+1} = \mathcal{T}_\alpha(\tilde{S})$
end
Output : L_{k+1}, S_{k+1}

Projection of $[L_0, S_0]$ onto set of $[L, S]$ such that $L_\Omega + S_\Omega = A_\Omega$
Analogous to RPCA

Convergence of Algorithm 2

Local linear convergence   Euclidean distance to optimum

$$d_{\mathcal{X}_1 \cap \mathcal{X}_\cap}([L_k \;\; S_k]) < c^k d_{\mathcal{X}_1 \cap \mathcal{X}_\cap}([L_0 \;\; S_0])$$

Cosine of angle between tangent spaces of $\mathcal{X}_1$, $\mathcal{X}_\cap$ Define $\mathcal{X}_\cap = \mathcal{X}_2 \cap \mathcal{X}_3$

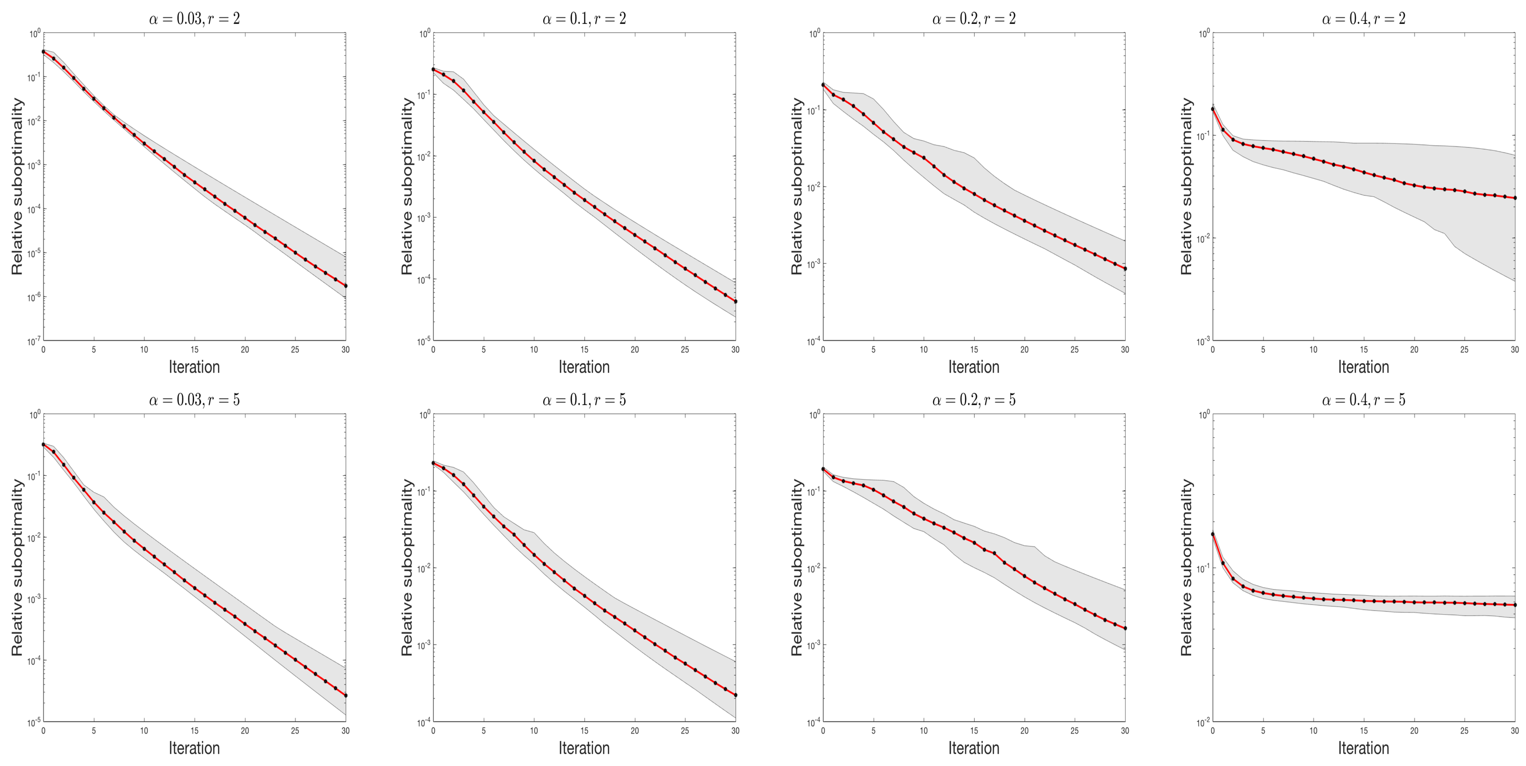
Convex relaxation (l1 norm for sparsity, nuclear norm instead of low rank)

Global linear rate

Numerical Results

We demonstrate our results on both synthetic and real data.

Sensitivity of Algorithm 2 on the initialization.
The best, the worst, and the median case are plotted for each iteration.



White denotes success and black denotes failure

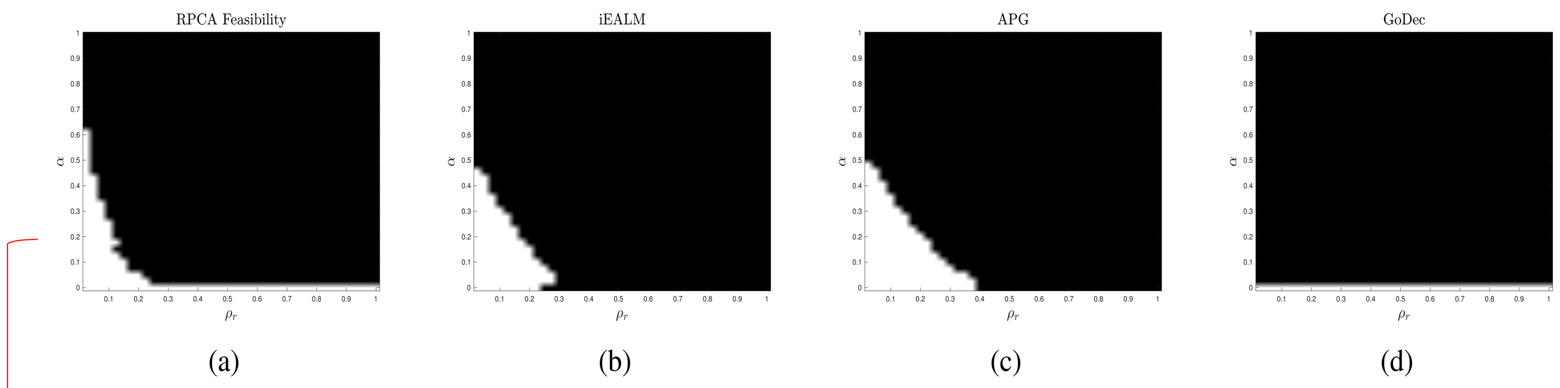


Figure 1: Phase transition diagram for RPCA F, iEALM, APG, and GoDec with respect to rank and error sparsity. Here, $\rho_r = \operatorname{rank}(L)/m$ and α is the sparsity measure. We have $(\rho_r, \alpha) \in (0.025, 1] \times (0, 1)$ with $r = 5 : 5 : 200$ and $\alpha = \texttt{linspace}(0, 0.99, 40)$. We perform 10 runs of each algorithm.

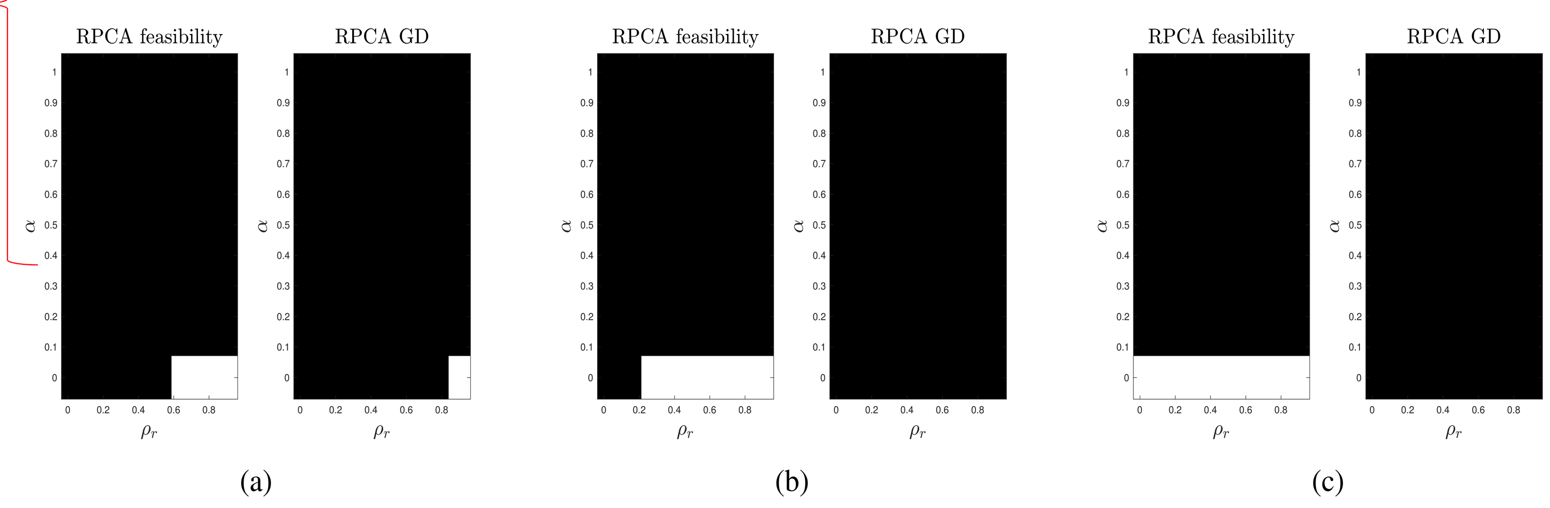


Figure 2: Phase transition diagram for Relative error for RMC problems: (a) $|\Omega^C| = 0.5(m.n)$, (b) $|\Omega^C| = 0.75(m.n)$, (c) $|\Omega^C| = 0.9(m.n)$. Here, $\rho_r = \operatorname{rank}(L)/m$ and α is the sparsity measure. We have $(\rho_r, \alpha) \in (0.025, 1] \times (0, 1)$ with $r = 5 : 25 : 200$ and $\alpha = \texttt{linspace}(0, 0.99, 8)$.

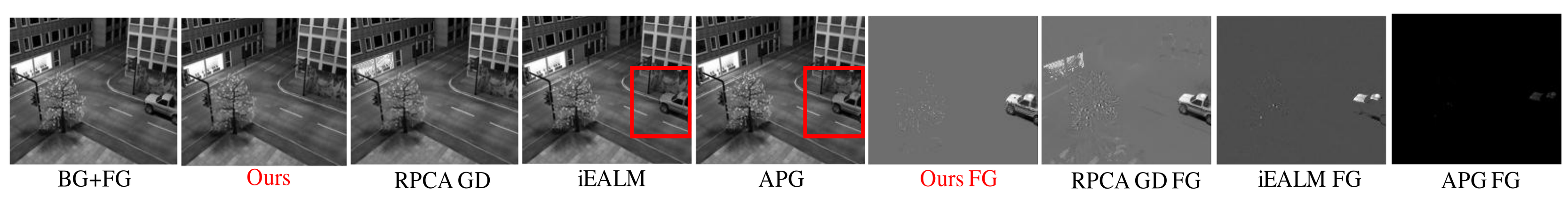


Figure 3: Background and foreground separation on Stuttgart dataset Basic video. Except RPCA GD and our method, all other methods fail to remove the static foreground object.

Figure 4: Background and foreground separation on Stuttgart dataset Basic video. We used 90% sample. GRASTA forms a fragmentary background and exhausts around 540 frames to form a stable video. We also note that RPCA GD has more false positives in the foreground.

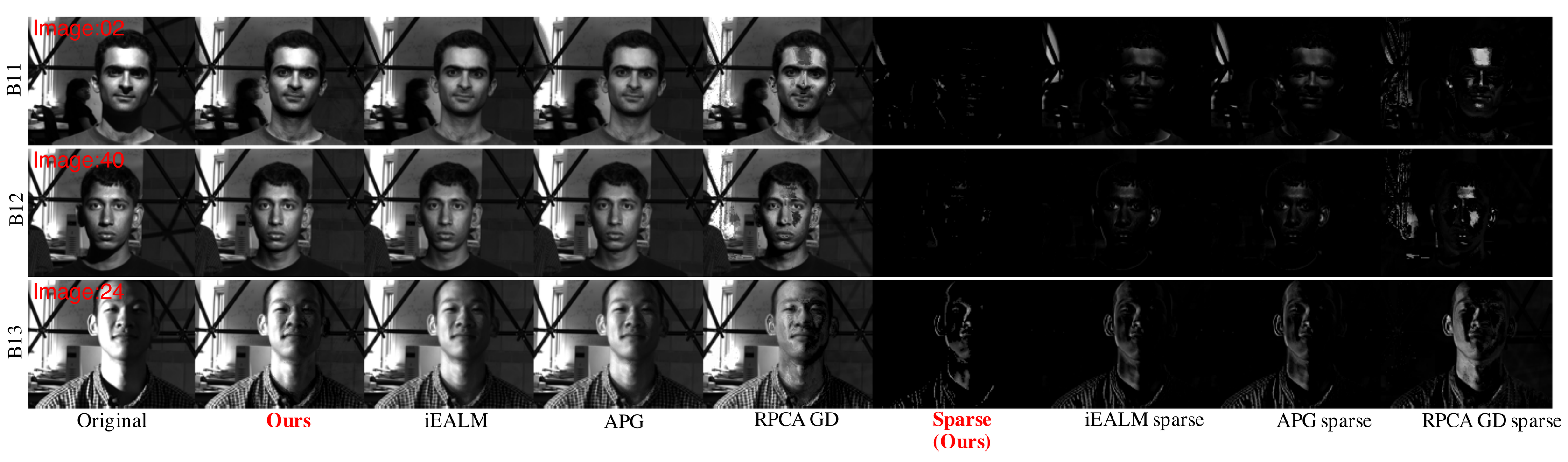


Figure 5: Shadow and specularities removal from face images captured under varying illumination and camera position. Our feasibility approach provides comparable reconstruction to that of iEALM and APG.

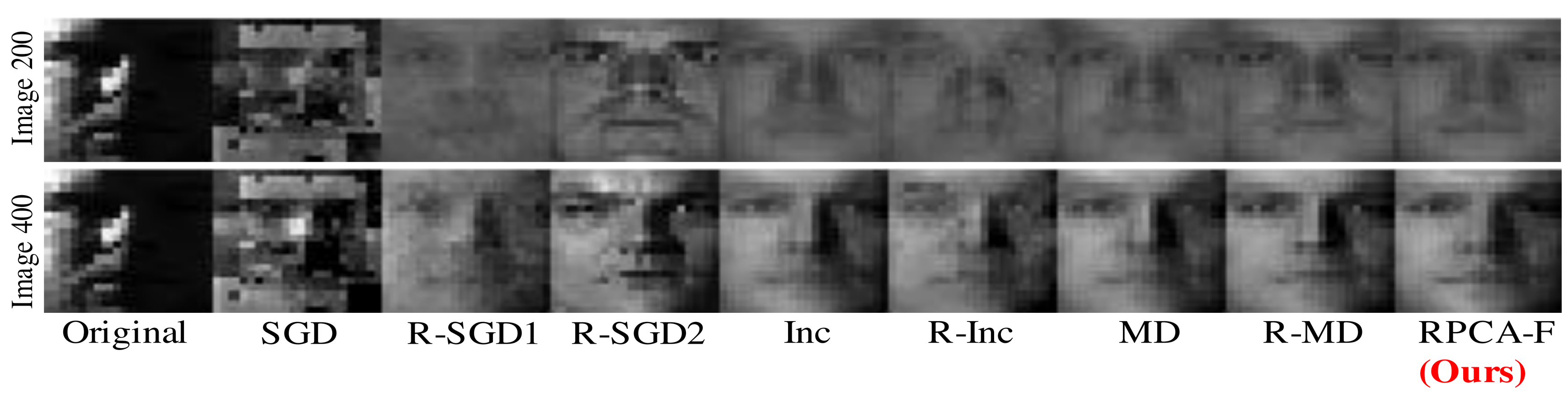


Figure 13: Inliers and outliers detection. Face images captured in different lighting conditions are inliers. We project different faces to 9 dimensional subspaces found by different methods.

*To appear in 2019 Thirty Third AAAI conference on Artificial Intelligence (acceptance rate 16.8% out of 7700 submissions). Designated for oral presentation.