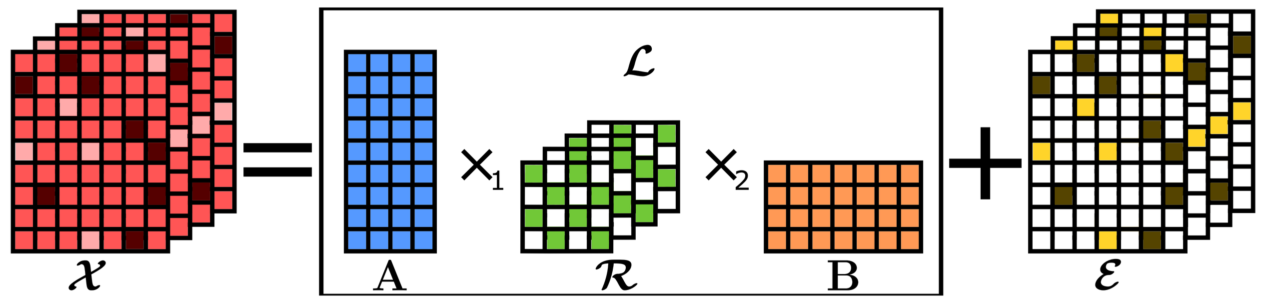


Robust Kronecker-Decomposable Component Analysis for Low-Rank Modeling

Mehdi Bahri, Yannis Panagakis, and Stefanos Zafeiriou

Imperial College London

mehdi.b.tn@gmail.com, i.panagakis@imperial.ac.uk, s.zafeiriou@imperial.ac.uk

Contribution: Robust Dictionaries and Tensor Factorization

We present a method for learning compact sparse representations in a noisy setting that combines ideas from dictionary learning and robust low-rank modeling. By imposing a separable dictionary we achieve scalability and exhibit links with tensor factorizations. Experimental assessment shows improvements of up to 16% on image denoising benchmarks, and competitive background subtraction performance.

Structured dictionaries

We decompose N observations $\mathbf{x}_i \in \mathbb{R}^{mn}$ on $\mathbf{C} \in \mathbb{R}^{mn \times d}$ with representations $\mathbf{y}_i \in \mathbb{R}^d$ through a standard regularized sparse dictionary learning problem:

$$\min_{\mathbf{C},\mathbf{Y}} \sum_i ||\mathbf{x}_i - \mathbf{C}\mathbf{y}_i||_2^2 + \lambda \sum_i ||\mathbf{y}_i||_1 + ||\mathbf{C}||_{\mathrm{F}} \qquad (1)$$

~~With $d = r_1 r_2 + mn$,~~ we impose a two-level structure:

$$\mathbf{y}_i = \begin{bmatrix} \mathbf{r}_i \\ \mathbf{e}_i \end{bmatrix} \in \mathbb{R}^d \quad \mathbf{C} = \begin{bmatrix} \mathbf{B} \otimes \mathbf{A} & \mathbf{I} \end{bmatrix} \in \mathbb{R}^{mn \times d} \qquad (2)$$

~~The $\mathbf{e}_i \in \mathbb{R}^{mn}$ model the presence of outliers (gross corruption). The codes $\mathbf{r}_i \in \mathbb{R}^{r_1 r_2}$ are learnt with respect to a Kronecker dictionary $\mathbf{B} \otimes \mathbf{A}$ with $\mathbf{A} \in \mathbb{R}^{m \times r_1}$, $\mathbf{B} \in \mathbb{R}^{n \times r_2}$.~~

Robust tensor factorization

In vision, observations are often vectorized matrices $\mathbf{x}_i = \mathrm{vec}(\mathbf{X}_i), \mathbf{X}_i \in \mathbb{R}^{m \times n}$. Matrix form preserves the spatial structure of images, and allows us to solve matrix equations efficiently instead of quadratically-larger linear systems. ~~Without loss of generality, we choose $r_1 = r_2 = r$ and~~ $\mathbf{r}_i = \mathrm{vec}(\mathbf{R}_i), \mathbf{R}_i \in \mathbb{R}^{r \times r}$, ~~and recast the~~ problem as:

$$\min_{\mathbf{A},\mathbf{B},\mathcal{R},\mathcal{E}} \sum_i ||\mathbf{X}_i - \mathbf{A}\mathbf{R}_i\mathbf{B}^{\mathsf{T}} - \mathbf{E}_i||_{\mathrm{F}}^2 + \lambda \sum_i ||\mathbf{R}_i||_1 + \lambda \sum_i ||\mathbf{E}_i||_1 + ||\mathbf{B} \otimes \mathbf{A}||_{\mathrm{F}} \qquad (3)$$

~~Enforcing equality constraints shows the problem is a structured tensor factorization:~~

$$\begin{aligned} \min_{\mathbf{A},\mathbf{B},\mathcal{R},\mathcal{E}} \quad & \lambda||\mathcal{R}||_1 + \lambda||\mathcal{E}||_1 + ||\mathbf{B} \otimes \mathbf{A}||_{\mathrm{F}} \\ \text{s.t} \quad & \mathcal{X} = \mathcal{R} \times_1 \mathbf{A} \times_2 \mathbf{B} + \mathcal{E} \end{aligned} \qquad (4)$$

$\mathbf{X}_i$, $\mathbf{R}_i$, and $\mathbf{E}_i$ are concatenated as the frontal slices of 3-way tensors. ~~We set $r \leq \min(m, n)$ as a natural upper bound on the rank of the frontal slices of $\mathcal{L} = \mathcal{R} \times_1 \mathbf{A} \times_2 \mathbf{B}$, and on its mode-1 and mode-2 ranks.~~

Non-convex ~~ADMM~~

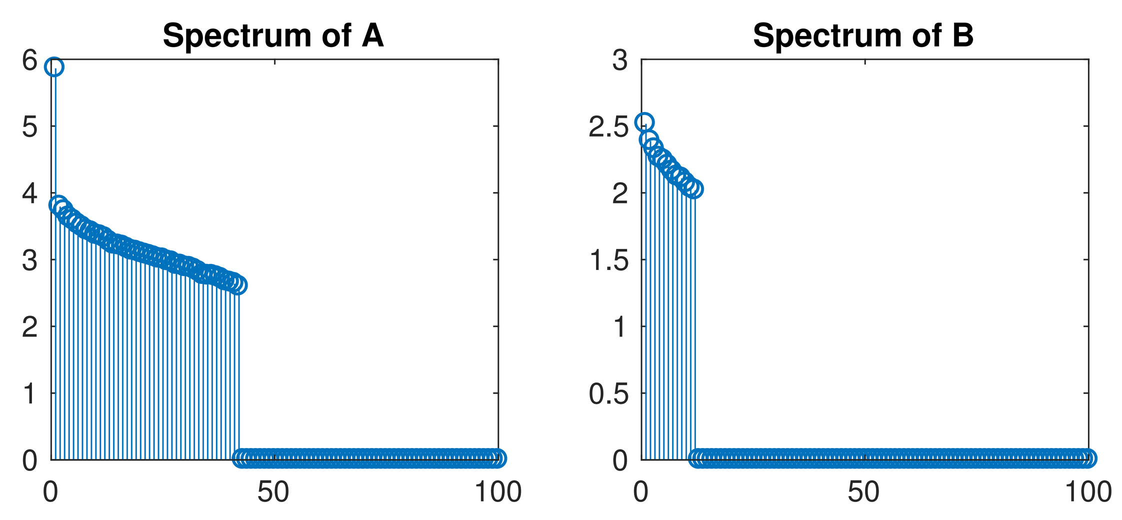
Problem (4) is convex in each component.
We optimize an upper bound on (4) with $||\mathbf{B} \otimes \mathbf{A}||_{\mathrm{F}} = ||\mathbf{A}||_{\mathrm{F}}||\mathbf{B}||_{\mathrm{F}} \leq \frac{||\mathbf{A}||_{\mathrm{F}}^2 + ||\mathbf{B}||_{\mathrm{F}}^2}{2}$ for scalability and tractability. We introduce split variables $\mathbf{K}_i$ such that $\forall i$, $\mathbf{K}_i = \mathbf{R}_i$ and solve:

$$\begin{aligned} \min_{\mathbf{A},\mathbf{B},\mathcal{R},\mathcal{K}\mathcal{E}} \quad & \lambda||\mathcal{R}||_1 + \lambda||\mathcal{E}||_1 + \tfrac{1}{2}(||\mathbf{A}||_{\mathrm{F}}^2 + ||\mathbf{B}||_{\mathrm{F}}^2) \\ \text{s.t} \quad & \mathcal{X} = \mathcal{K} \times_1 \mathbf{A} \times_2 \mathbf{B} + \mathcal{E} \\ \text{s.t} \quad & \mathcal{R} = \mathcal{K} \end{aligned} \qquad (5)$$

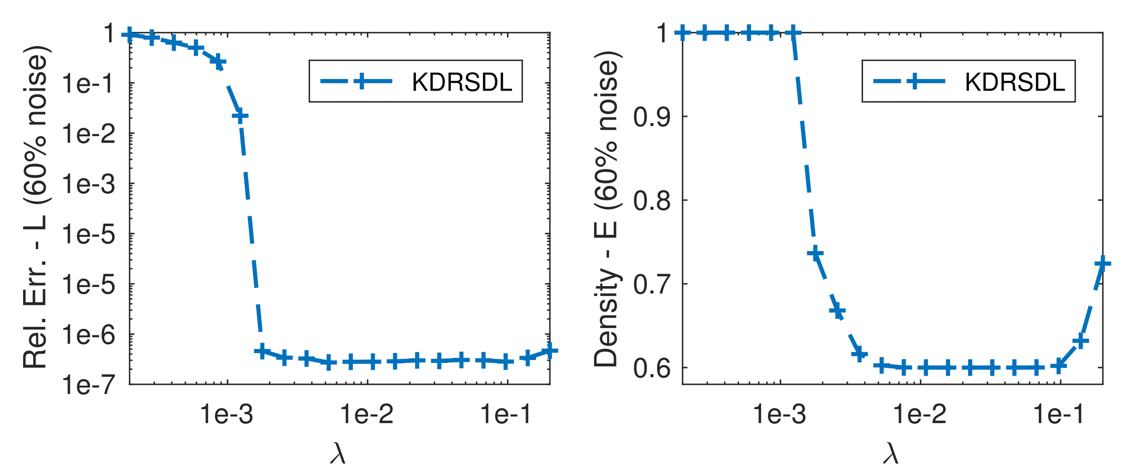
~~Spliting allows for closed-form proximal operators and no fixed-point updates. The dual steps are bounded for convergence.~~

Validation on synthetic data

Our algorithm successfuly recovers the components of (4) on synthetic data.

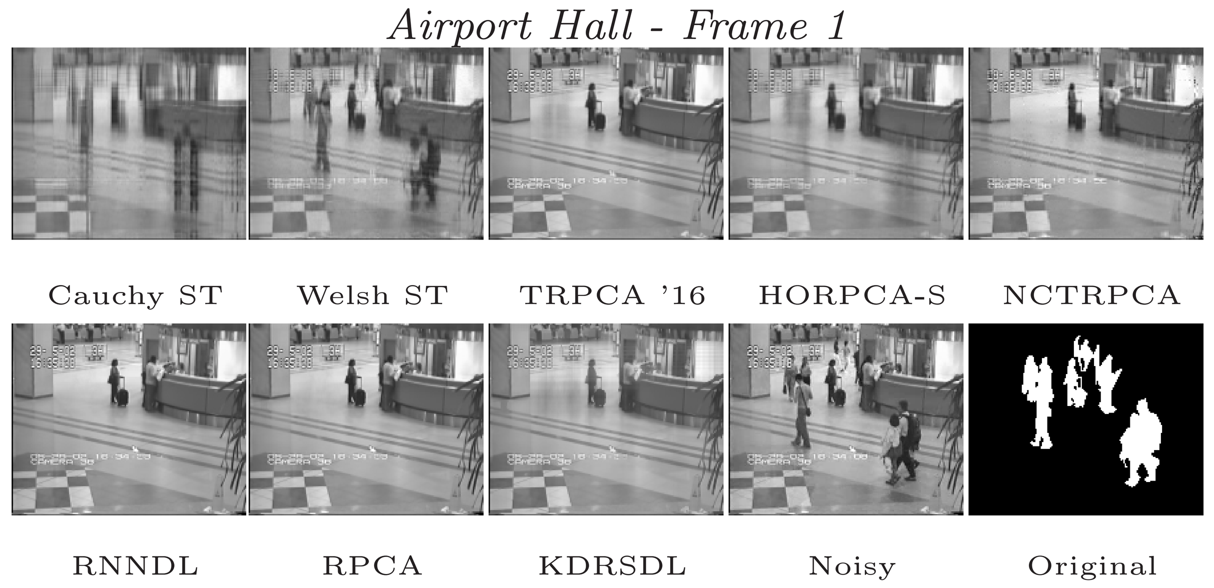


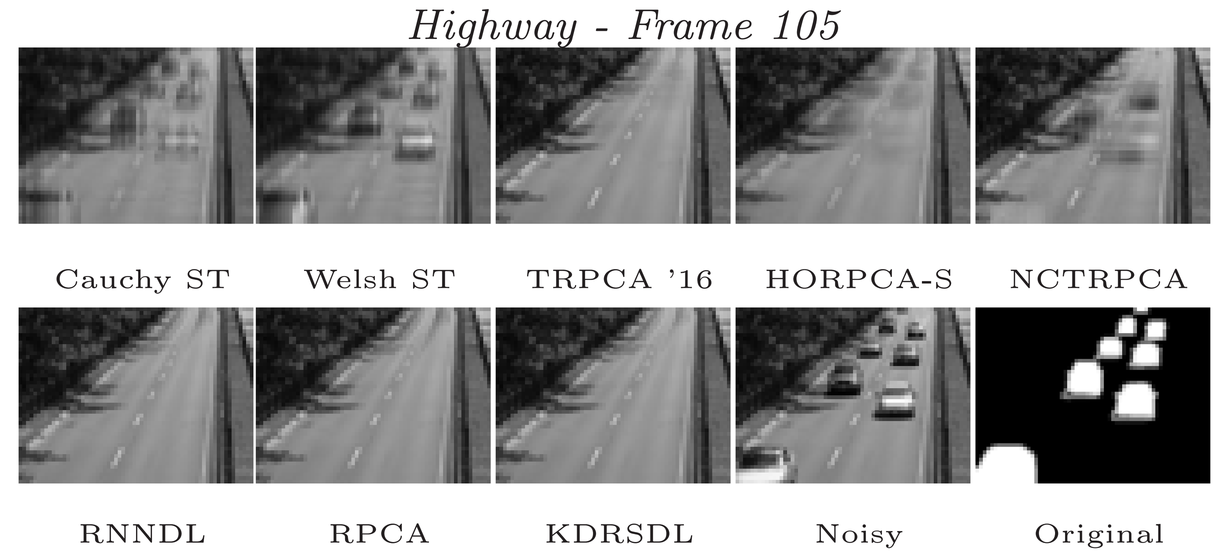
True ranks (42, 12) recovered, $r = 100$.



ℓ_2 error on $\mathcal{L}$, density of $\mathcal{E}$, 60% corruption

Background subtraction experiments

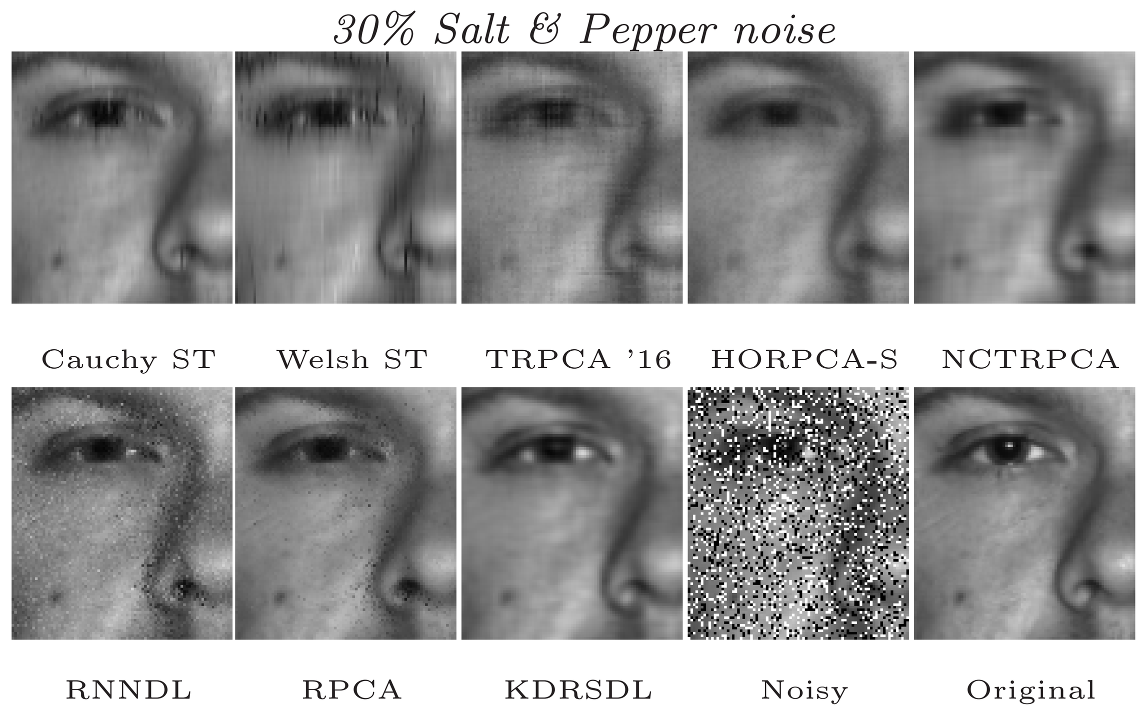


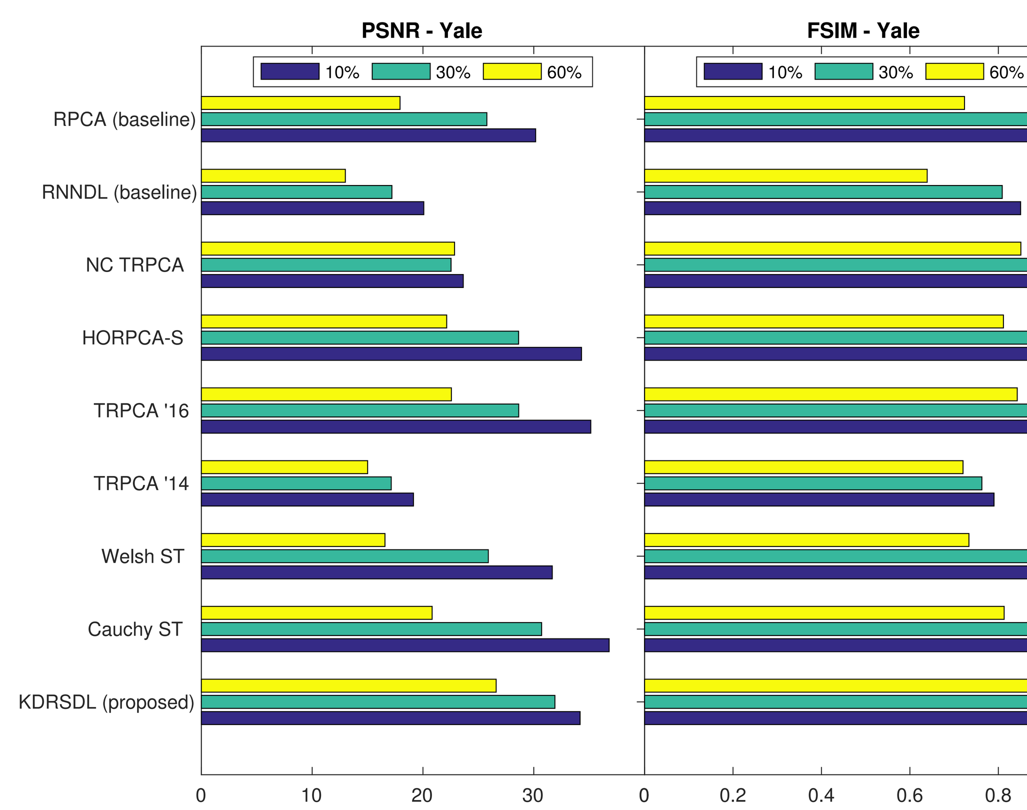


Procedure and results. We compare a panel of recent algorithms for robust component analysis on excerpts of the *Highway* dataset [1], and of the *Airport Hall* dataset ([2]). We report the AUC scores ranked by mean performance. Our model matched the best performance on the *Highway* dataset and provided significantly higher performance than the other on the *Hall* benchmark. ~~An additional step for robust estimation of the mean image was used.~~

Algorithm	Highway	Hall
KDRSDL (proposed)	0.94	0.88
TRPCA '16	0.94	0.86
NC TRPCA	0.93	0.86
RPCA (baseline)	0.94	0.85
RNNDL (baseline)	0.94	0.85
HORPCA-S	0.93	0.86
Cauchy ST	0.83	0.76
Welsh ST	0.82	0.71
TRPCA '14	0.76	0.61

Image denoising - ~~example~~ on the Yale-B dataset





Procedure and results. Keeping the 64 illuminations of the first subject [3], we expect the data tensor to be low-rank on all 3 modes. At 30% noise and above, our method showed markedly higher quantitative scores and noticeably better reconstructions.

We invite the reader to look at the texture of the skin, the white of the eye, and at the reflection of the light on the subject's skin and pupil. The latter, in particular, is very close in nature to the white pixel corruption of the salt & pepper noise.

References

[1] N. Goyette, P. M. Jodoin, F. Porikli, J. Konrad, and P. Ish-war. changedetection.net: A new change detection benchmark dataset. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2012.

[2] L. Li, W. Huang, I.-H. Gu, and Q. Tian. Statistical Modeling of Complex Backgrounds for Foreground Object Detection. In IEEE Transactions on Image Processing, 2004

[3] A. Georghiades, P. Belhumeur, and D. Kriegman. From few to many: illumination cone models for face recognition under variable lighting and pose. In IEEE Transactions on Pattern Analysis and Machine Intelligence, 2001.