Robust Kronecker-Decomposable Component Analysis for Low-Rank Modeling

Mehdi Bahri, Yannis Panagakis, and Stefanos Zafeiriou

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 $\mathbf{D} = [\mathbf{B} \otimes \mathbf{A} \quad \mathbf{I}] \in \mathbb{R}^{mn \times d}$ with representations $\mathbf{y}_i = [\mathbf{r}_i \quad \mathbf{e}_i]^\mathsf{T} \in \mathbb{R}^d$ through a two-level structured regularized sparse dictionary learning problem:

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

- The $\mathbf{e}_i \in \mathbb{R}^{mn}$ model the presence of outliers (gross corruption)
- The codes $\mathbf{r}_i \in \mathbb{R}^{r^2}$ are learnt with respect to a Kronecker dictionary $\mathbf{B} \otimes \mathbf{A}$ with $\mathbf{A} \in \mathbb{R}^{m \times r}, \mathbf{B} \in \mathbb{R}^{n \times r}$

Robust tensor factorization 1

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
- Allows to solve matrix equations efficiently instead of quadratically larger linear systems

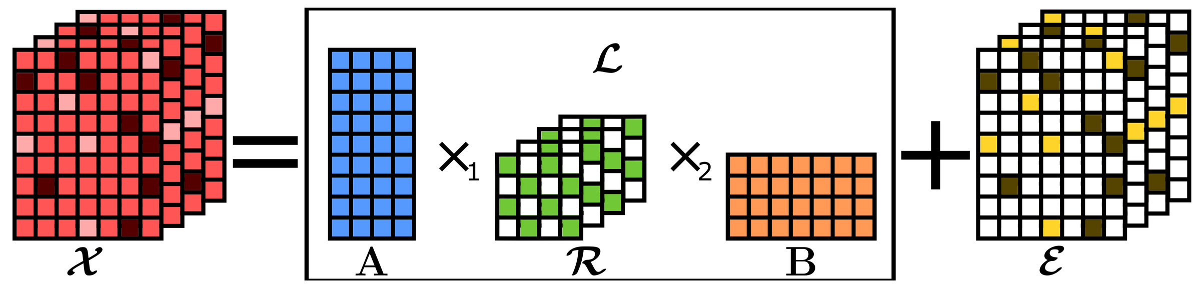
Let $\mathbf{r}_i = \mathrm{vec}(\mathbf{R}_i), \mathbf{R}_i \in \mathbb{R}^{r \times r}$, (1) becomes:

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

Or equivalently, as a structured tensor factorization:

$$\begin{aligned} \min_{\mathbf{A},\mathbf{B},\boldsymbol{\mathcal{R}},\boldsymbol{\mathcal{E}}} \quad & \lambda||\boldsymbol{\mathcal{R}}||_1 + \lambda||\boldsymbol{\mathcal{E}}||_1 + ||\mathbf{B} \otimes \mathbf{A}||_\mathrm{F} \\ \text{s.t} \quad & \boldsymbol{\mathcal{X}} = \boldsymbol{\mathcal{R}} \times_1 \mathbf{A} \times_2 \mathbf{B} + \boldsymbol{\mathcal{E}} \end{aligned} \qquad (3)$$

Robust tensor factorization 2

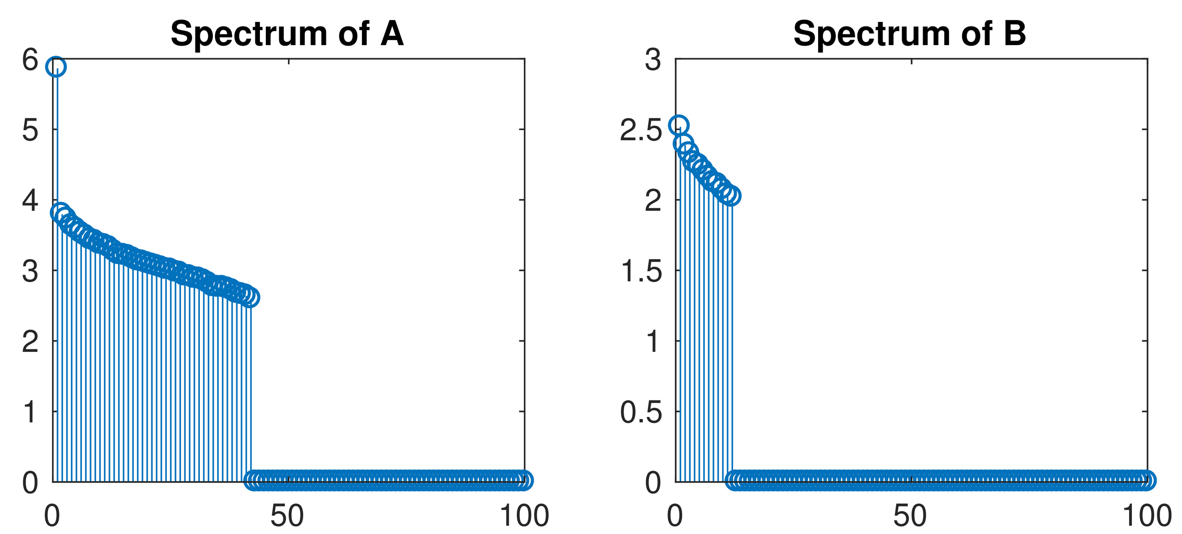


- $\mathbf{X}_i, \mathbf{R}_i$, and $\mathbf{E}_i$ concatenated as the frontal slices of 3-way tensors
- $r \leq \min(m, n)$ bounds the mode-1 and mode-2 ranks of $\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{R}} \times_1 \mathbf{A} \times_2 \mathbf{B}$

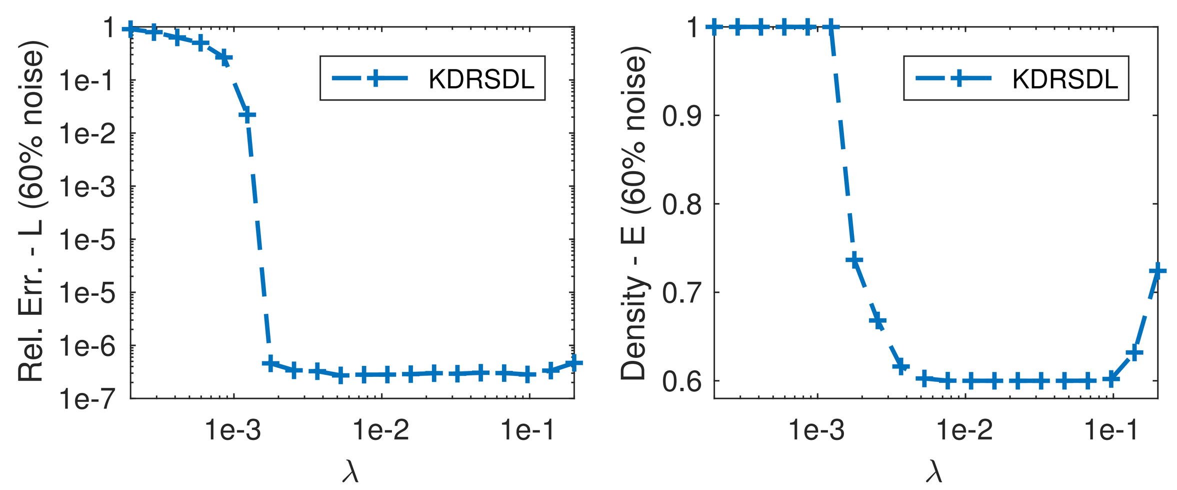
Min. bound via non-convex ADMM with splitting and $||\mathbf{B} \otimes \mathbf{A}||_\mathrm{F} = ||\mathbf{A}||_\mathrm{F}||\mathbf{B}||_\mathrm{F} \leq \frac{1}{2}\left(||\mathbf{A}||_\mathrm{F}^2 + ||\mathbf{B}||_\mathrm{F}^2\right)$.

Validation on synthetic data

Our algorithm successfuly recovers the components of (3) on synthetic data.

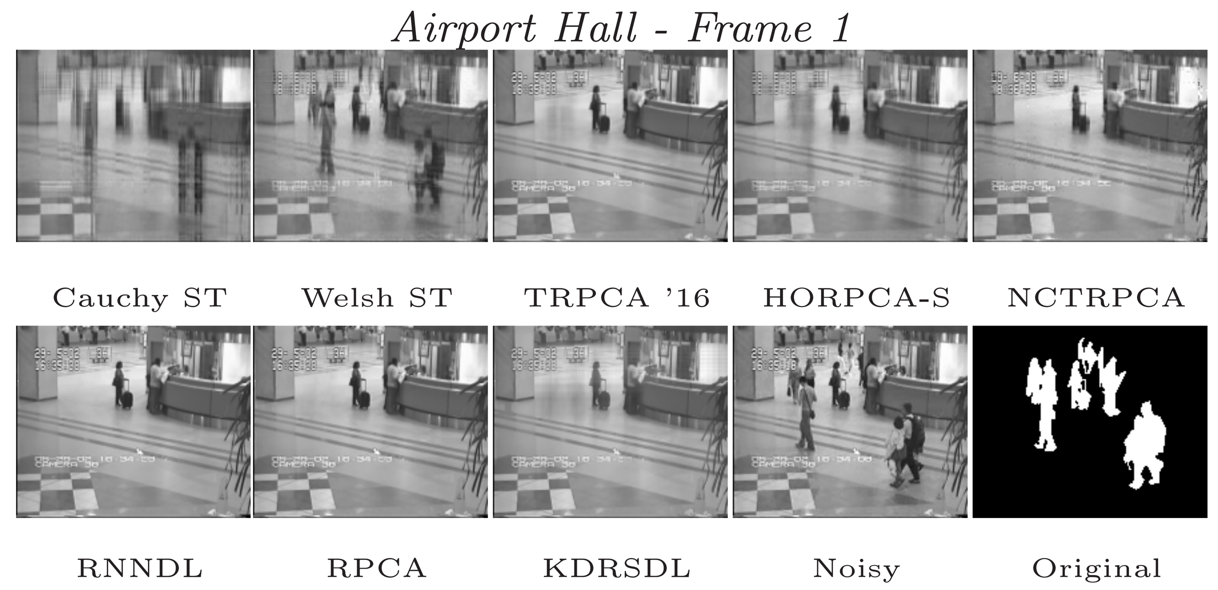


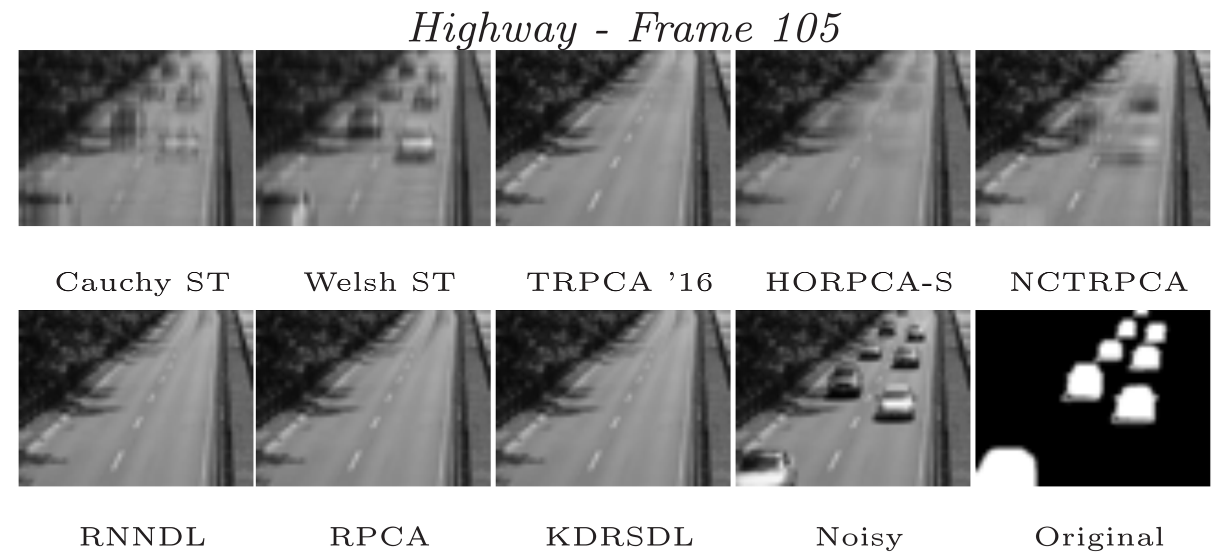
True ranks (42, 12) recovered, $r = 100$



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

Background subtraction experiments



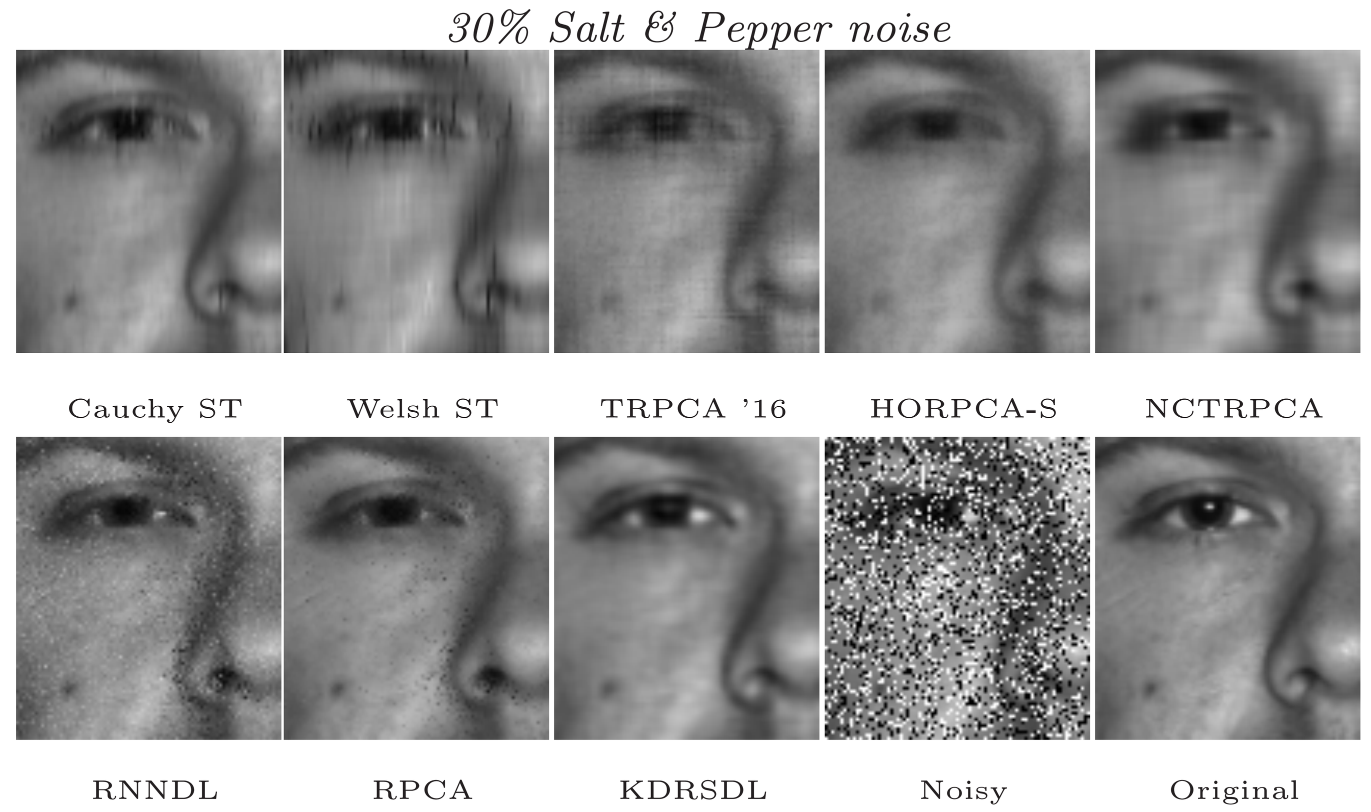


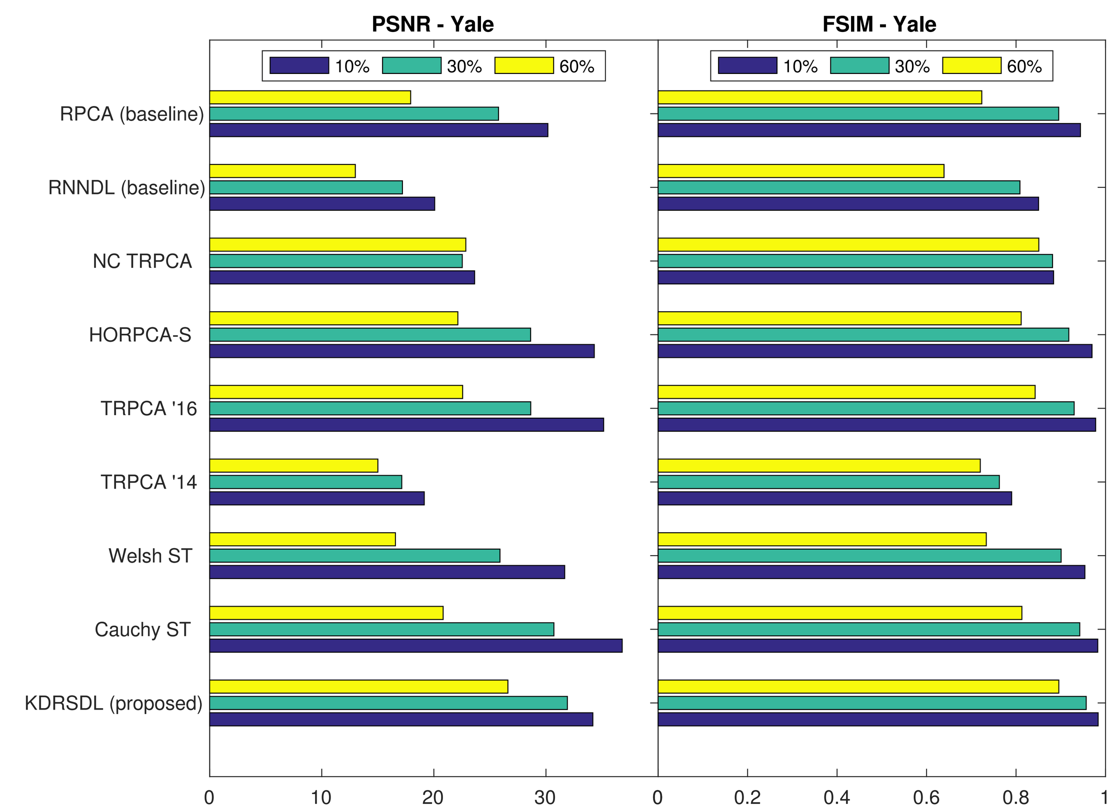
Procedure and results. Panel of recent algorithms for robust component analysis compared on excerpts of the *Highway* dataset [1], and of the *Airport Hall* dataset [2]. We report the AUC scores.

Our model matched the best performance on the *Highway* dataset and provided the highest performance on the *Hall* benchmark.

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 experiments on the Yale-B dataset





Procedure and results. On the 64 illuminations of first subject [3], data tensor low-rank on all 3 modes.

At noise $\geq 30\%$, we achieved markedly higher quantitative scores and noticeably better reconstructions.

Differences best seen on:

- Skin texture, white of eye
- Reflected light (pupil, skin)

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, 11 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, 6 2001.