Robust Kronecker-Decomposable Component Analysis for Low-Rank Modeling - Supplementary Material

Mehdi Bahri*[1], Yannis Panagakis†[1,2], and Stefanos Zafeiriou‡[1,3]

[1]*Department of Computing, Imperial College London, UK*
[2]*Department of Computer Science, Middlesex University London, UK*
[3]*Department of Computer Science and Engineering, University of Oulu, Finland*

November 6, 2017

Contents

*mehdi.b.tn@gmail.com
†i.panagakis@imperial.ac.uk
‡s.zafeiriou@imperial.ac.uk

1 Schatten norms and the Kronecker product

In this section we prove the identity used in section 2.2 on the Schatten norm of a Kronecker product.

Let us first remind the reader of the definition of the Schatten-p norm:

Definition 1. Let $\mathbf{A}$ a real-valued matrix, $\mathbf{A} \in \mathbb{R}^{m \times n}$. The Schatten-$p$ norm of $\mathbf{A}$ is defined as:

$$||\mathbf{A}||_p = (\sum_i s_i^p)^{1/p}$$

Where s_i is the i^{th} singular value of $\mathbf{A}$.

We argued the result stems from the compatibility of the Kronecker product with the singular value decomposition, that is:

Proposition 1. *Let* $\mathbf{A} = \mathbf{U}_A \mathbf{\Sigma}_A \mathbf{V}_A^\mathsf{T}$, $\mathbf{B} = \mathbf{U}_B \mathbf{\Sigma}_B \mathbf{V}_B^\mathsf{T}$ *two real-valued matrices given by their SVD, then:*

$$\mathbf{A} \otimes \mathbf{B} = (\mathbf{U}_A \otimes \mathbf{U}_B)(\mathbf{\Sigma}_A \otimes \mathbf{\Sigma}_B)(\mathbf{V}_A^\mathsf{T} \otimes \mathbf{V}_B^\mathsf{T})$$

Proof. See [1]. □

The identity we used is formally expressed in Theorem 1, of which we give the proof for completeness.

Theorem 1. *Let* $\mathbf{A} \in \mathbb{R}^{m \times n}$ *and* $\mathbf{B} \in \mathbb{R}^{p \times q}$*, then:*

$$\forall p > 0, ||\mathbf{A} \otimes \mathbf{B}||_p = ||\mathbf{A}||_p ||\mathbf{B}||_p$$

Where $||.||_p$ *denotes the Schatten-p norm.*

Proof. From Proposition 1, the singular values of $\mathbf{A} \otimes \mathbf{B}$ are the $\sigma_{A,i}\sigma_{B,j}$ so:

$$\begin{aligned}
||\mathbf{A} \otimes \mathbf{B}||_p &= (\sum_{i,j} (\sigma_{A,i}\sigma_{B,j})^p)^{1/p} \\
&= (\sum_{i,j} \sigma_{A,i}^p \sigma_{B,j}^p)^{1/p} \\
&= (\sum_i \sigma_{A,i}^p \sum_j \sigma_{B,j}^p)^{1/p} \\
&= (\sum_i \sigma_{A,i}^p)^{1/p} (\sum_j \sigma_{B,j}^p)^{1/p} \\
&= ||\mathbf{A}||_p ||\mathbf{B}||_p
\end{aligned}$$

□

2 Parameter tuning

In all experiments, we set r to its maximum possible size as given in Section 2.1 of the paper ($r = \min(m, n)$) and let the algorithm determine the rank.

Algorithm	First parameter tuned	Second parameter tuned
KDRSDL	λ	
Tensor RPCA (TRPCA '14)	λ	
Tensor RPCA (TRPCA '16)	λ	
NC TRPCA	Thresholding constant	Rank
HORPCA-S	λ	
Welsh ST	σ	λ
Cauchy ST	σ	λ
Matrix RPCA via ALM	λ	
Robust Non-Negative Dictionary Learning (RNNDL)	α	

Table 1: The algorithms tested and the parameters we tuned for each of them.

Algorithm	Range for first parameter	Range for second parameter
KDRSDL	$2 \cdot \text{logspace}(-4, -1, 30)$	
Tensor RPCA (TRPCA '14)	$\text{linspace}(0.5 \cdot 10^p, 1.5 \cdot 10^q, 40)$ $\max_p, \min_q 10^p < \frac{1}{\sqrt{\max(n_1,n_2)}} < 10^q$ $p, q \in \mathbb{Z}$	
Tensor RPCA (TRPCA '16)	$\text{linspace}(0.5 \cdot 10^p, 1.5 \cdot 10^q, 40)$ $\max_p, \min_q 10^p < \frac{1}{\sqrt{n_3 \max(n_1,n_2)}} < 10^q$ $p, q \in \mathbb{Z}$	
NC TRPCA	$[10, 1e2, 1e3, 1e4]$	Rank depending on $\min(n_1, n_2)$
HORPCA-S	$\text{logspace}(-1, 1, 30) \cdot \frac{1}{\sqrt{\max(n_1,n_2,n_3)}}$	
Welsh ST	$\text{logspace}(\log_{10}(0.05), \log_{10}(2), 20)$	$\lambda = \alpha \frac{\min(n_1,n_2)}{\sqrt{\max(n_1,n_2)}}$ $\alpha \in \text{linspace}(0, 1, 40)$
Cauchy ST	$\text{logspace}(\log_{10}(0.05), \log_{10}(2), 20)$	$\lambda = \alpha \frac{\min(n_1,n_2)}{\sqrt{\max(n_1,n_2)}}$ $\alpha \in \text{linspace}(0, 1, 40)$
RPCA ALM	$\text{linspace}(0.5 \cdot 10^p, 1.5 \cdot 10^q, 100)$ $\max_p, \min_q 10^p < \frac{1}{\sqrt{\max(n_1,n_2)}} < 10^q$ $p, q \in \mathbb{Z}$	
RNNDL	$0 : 0.5 : 5$	

Table 2: Range of parameters tested for each algorithm in Matlab notation.

For the Highway experiment, the range tested for matrix RPCA is $2 \cdot \text{logspace}(-4, -1, 20)$.

3 Numerical values for the experiments

To help understand Figures 6 and 8 of the paper, we present in this section the raw numerical values.

3.1 Yale

Algorithm	PSNR 10%	FSIM 10%	PSNR 30%	FSIM 30%	PSNR 60%	FSIM 60%
KDRSDL (proposed)	34.1806	**0.9835**	**31.9094**	**0.9568**	**26.6057**	**0.8956**
Cauchy ST	**36.8013**	0.9826	30.7079	0.9424	20.8269	0.8131
Welsh ST	31.6702	0.9539	25.8976	0.9008	16.5715	0.7335
TRPCA '14	19.1429	0.7902	17.1359	0.7627	15.0052	0.7201
TRPCA '16	35.145	0.9779	28.6409	0.9299	22.566	0.8427
HORPCA-S	34.3085	0.9698	28.6301	0.9179	22.141	0.8115
NC TRPCA	23.6362	0.8839	22.5348	0.8816	22.8502	0.8509
RNNDL (baseline)	20.074641	0.850326	17.197725	0.808746	12.994455	0.639077
RPCA (baseline)	30.16977	0.943962	25.766677	0.89536	17.929617	0.723441

Table 3: Mean values of the image quality metrics at the three noise levels on the Yale-B dataset.

Algorithm	PSNR 10%	FSIM 10%	PSNR 30%	FSIM 30%	PSNR 60%	FSIM 60%
KDRSDL (proposed)	3.393101	0.002874	3.162017	0.004965	2.026028	0.022767
Cauchy ST	3.9682	0.006831	2.334835	0.020514	1.315045	0.025015
Welsh ST	2.942964	0.004941	1.333094	0.013318	2.311143	0.064
TRPCA '14	2.604563	0.065118	3.244374	0.074961	2.830976	0.064604
TRPCA '16	3.315294	0.011233	2.579565	0.029985	2.905814	0.06481
HORPCA-S	3.761219	0.029682	3.290198	0.057666	3.141829	0.041267
NC TRPCA	3.263149	0.036687	3.25142	0.030743	4.19665	0.062005
RNNDL (baseline)	5.256621	0.094887	3.983758	0.097075	3.257458	0.126664
RPCA (baseline)	5.343386	0.097123	5.300932	0.111393	3.310981	0.095751

Table 4: Standard deviation of the image quality metrics at the three noise levels on the Yale-B dataset.

3.2 Facade

Algorithm	PSNR 10%	FSIMc 10%	PSNR 30%	FSIMc 30%	PSNR 60%	FSIMc 60%
KDRSDL (proposed)	**33.2469**	0.9935	27.8261	0.9699	**23.8064**	**0.9152**
Cauchy ST	31.7107	0.9924	**29.0492**	0.9724	22.315	0.882
Welsh ST	28.0183	0.9725	25.2946	0.9309	19.8508	0.8421
TRPCA '14	31.8137	0.9886	25.9989	0.9484	21.4069	0.8529
TRPCA '16	32.7462	**0.9955**	27.0979	**0.9764**	23.6552	0.9109
HORPCA-S	33.137	0.9938	27.52	0.9714	22.8811	0.906
NC TRPCA	27.8527	0.9744	21.2992	0.8881	19.4207	0.8121
RNNDL (baseline)	21.11498	0.926266	14.243917	0.804113	10.01287	0.668988
RPCA (baseline)	21.326577	0.940371	13.675052	0.814479	9.229267	0.663291

Table 5: Image quality metrics at the three noise levels on the Facade dataset.

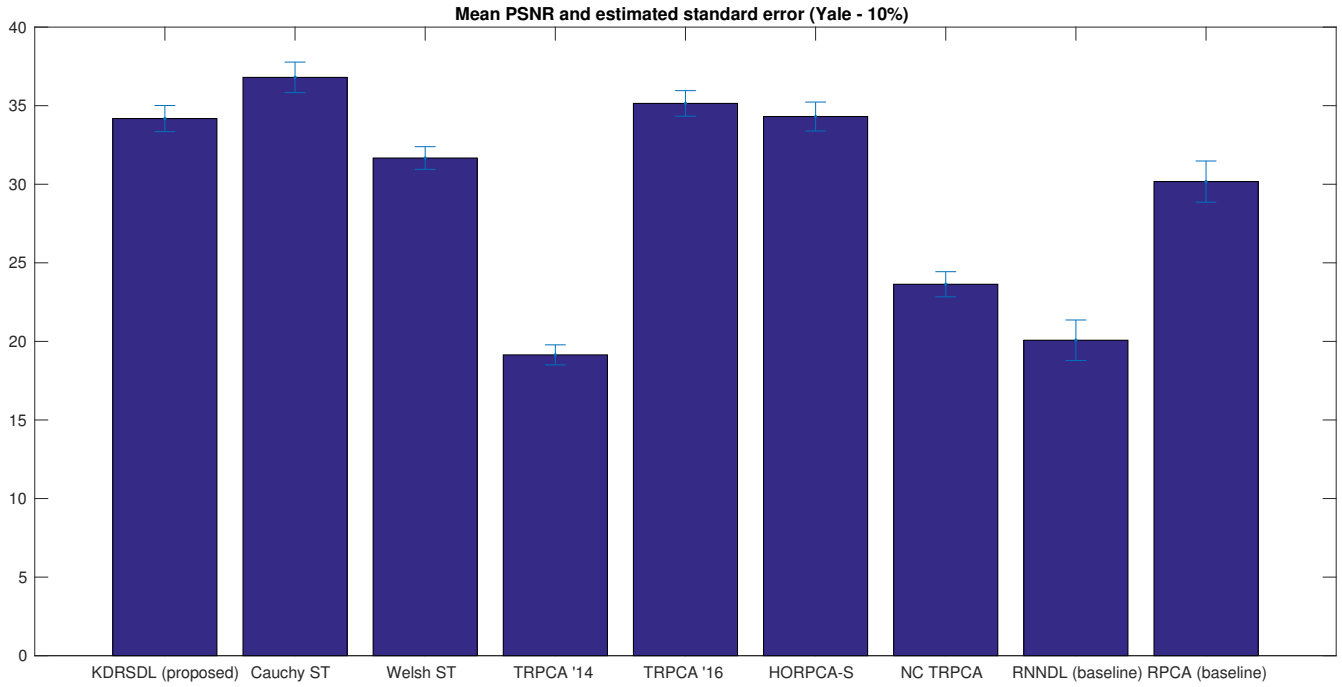


Figure 1: Mean PSNR and standard error at the 10% noise level.

4 Additional experimental results

In this section we provide additional information on the results reported in the paper, as well as supporting material.

4.1 Mean performance on the Yale dataset

On the Yale-B experiment, the measurements we present are average values over the 64 images of the first subject. There is therefore an uncertainty on the measures that we wish to address here.

We present in the following figures the mean values along with their estimated standard error. The formula we used for the standard error is

$$\text{STDERR} = 1.96 \frac{\hat{s}}{\sqrt{n}}$$

With $\hat{s}$ the sample standard deviation, and n the sample size (64).

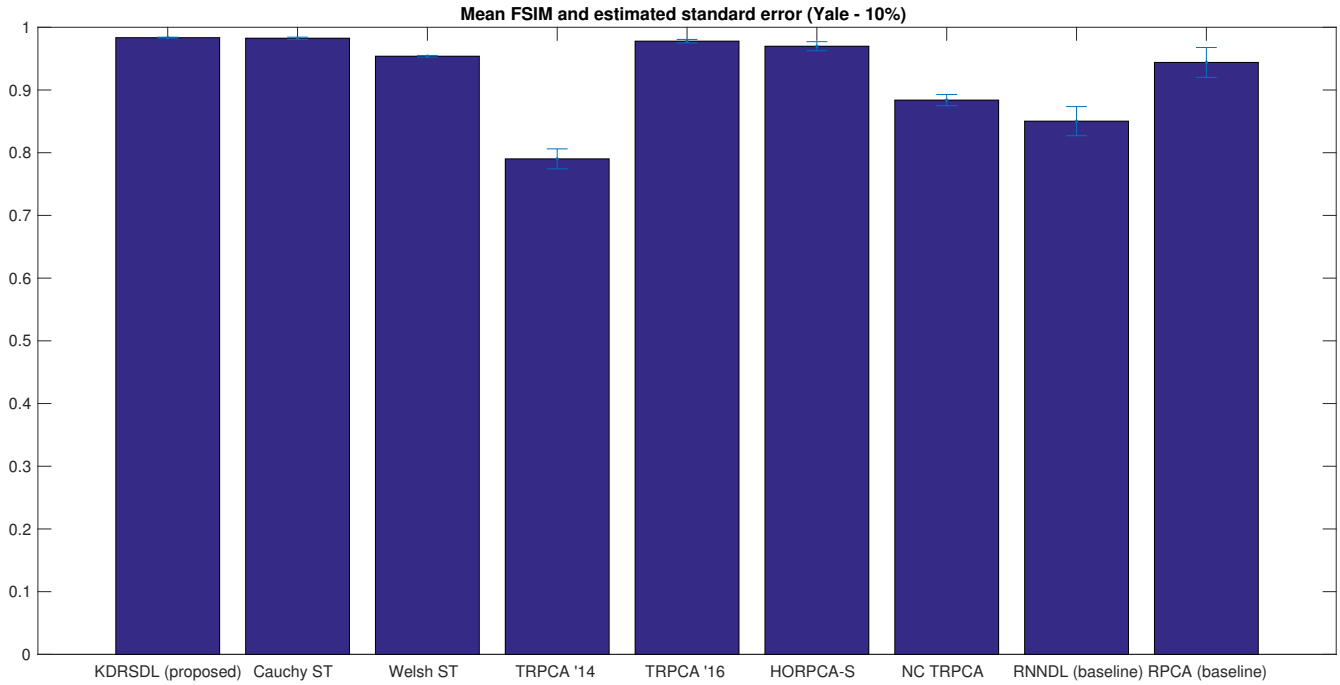


Figure 2: Mean FSIM and standard error at the 10% noise level.

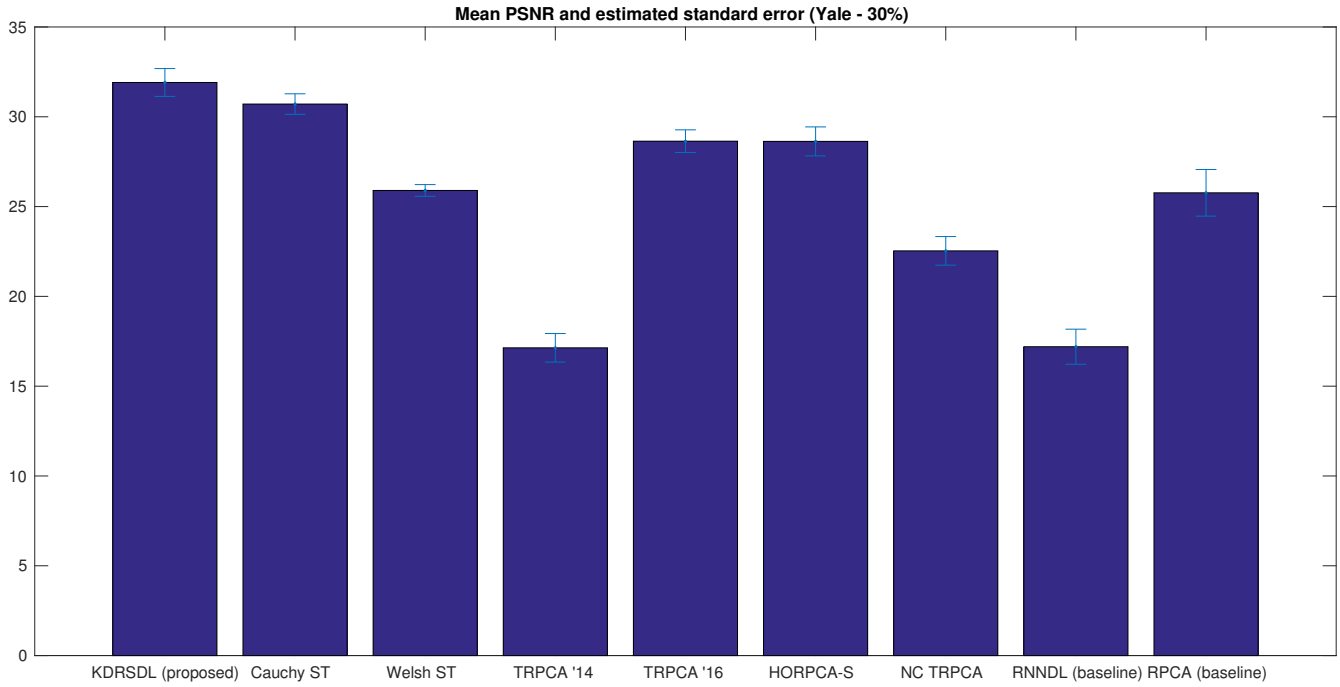


Figure 3: Mean PSNR and standard error at the 30% noise level.

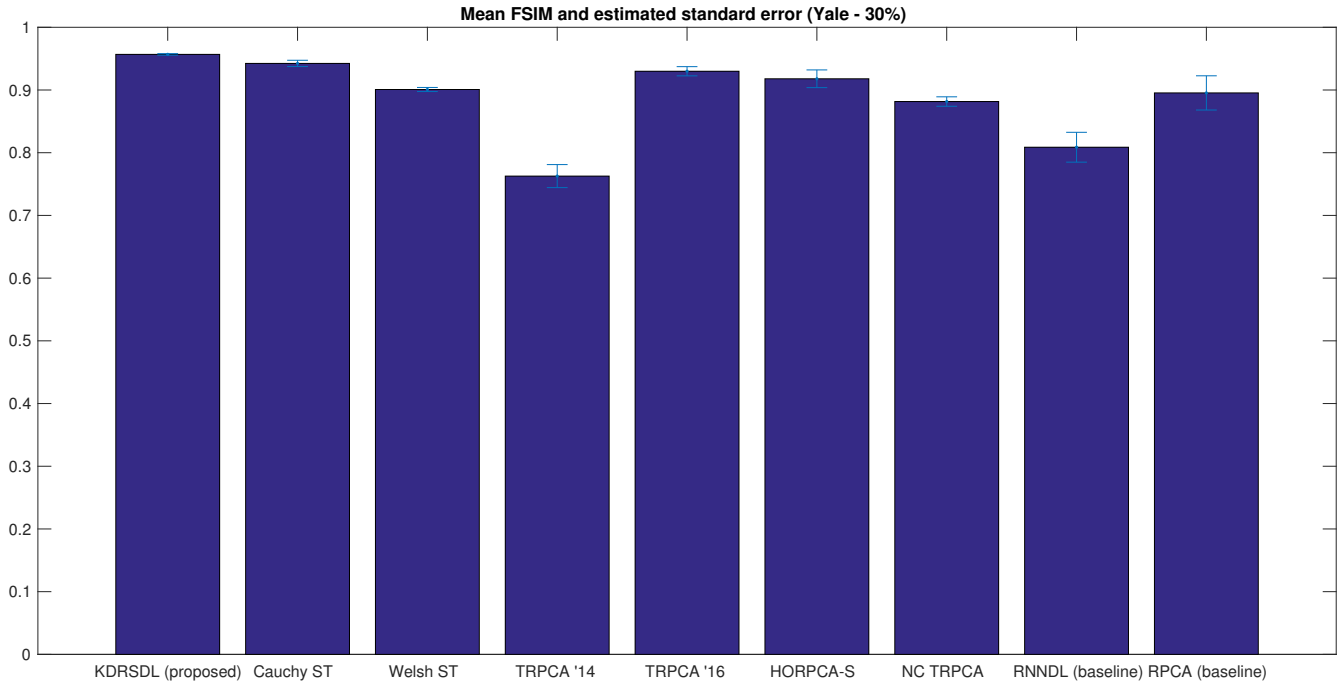


Figure 4: Mean FSIM and standard error at the 30% noise level.

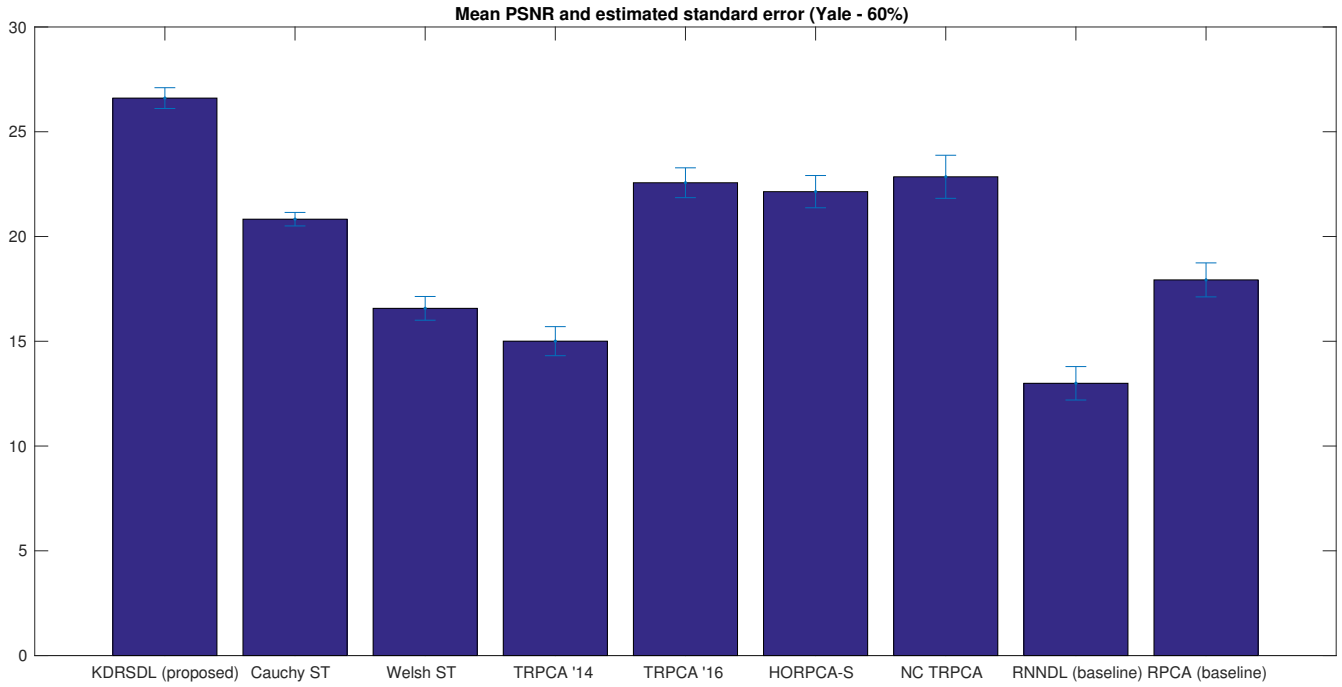


Figure 5: Mean PSNR and standard error at the 60% noise level.

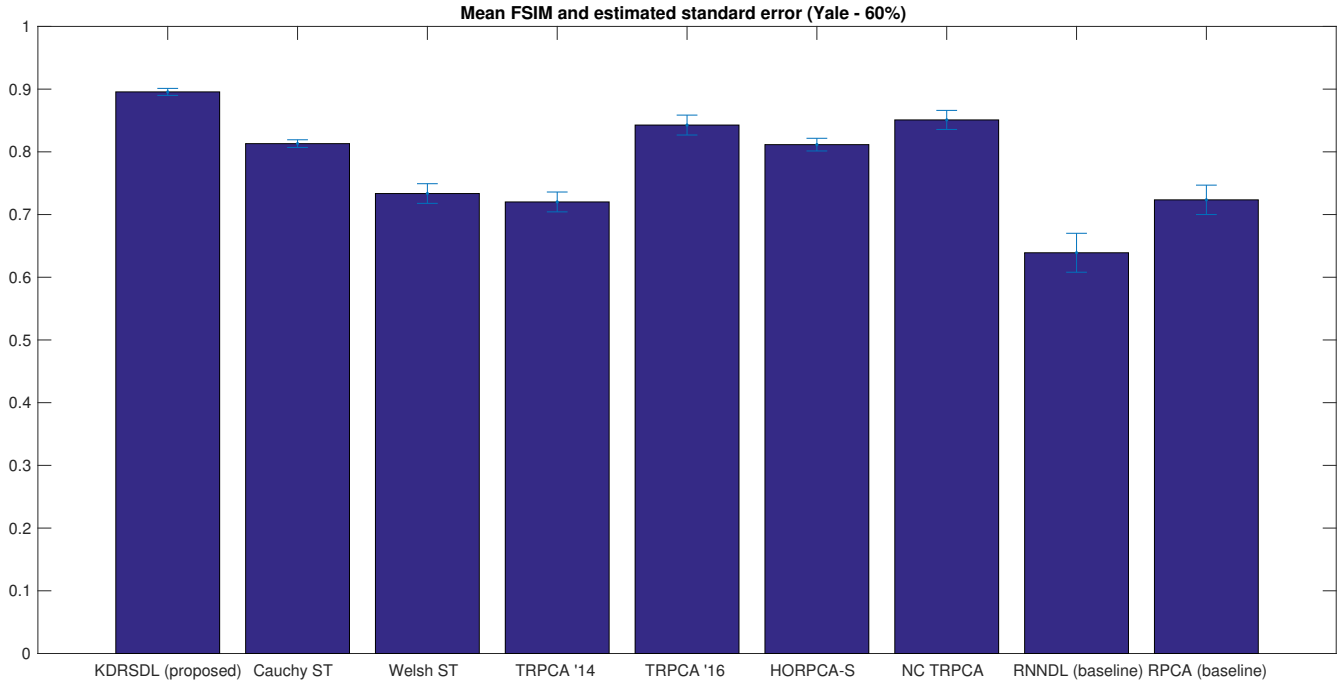
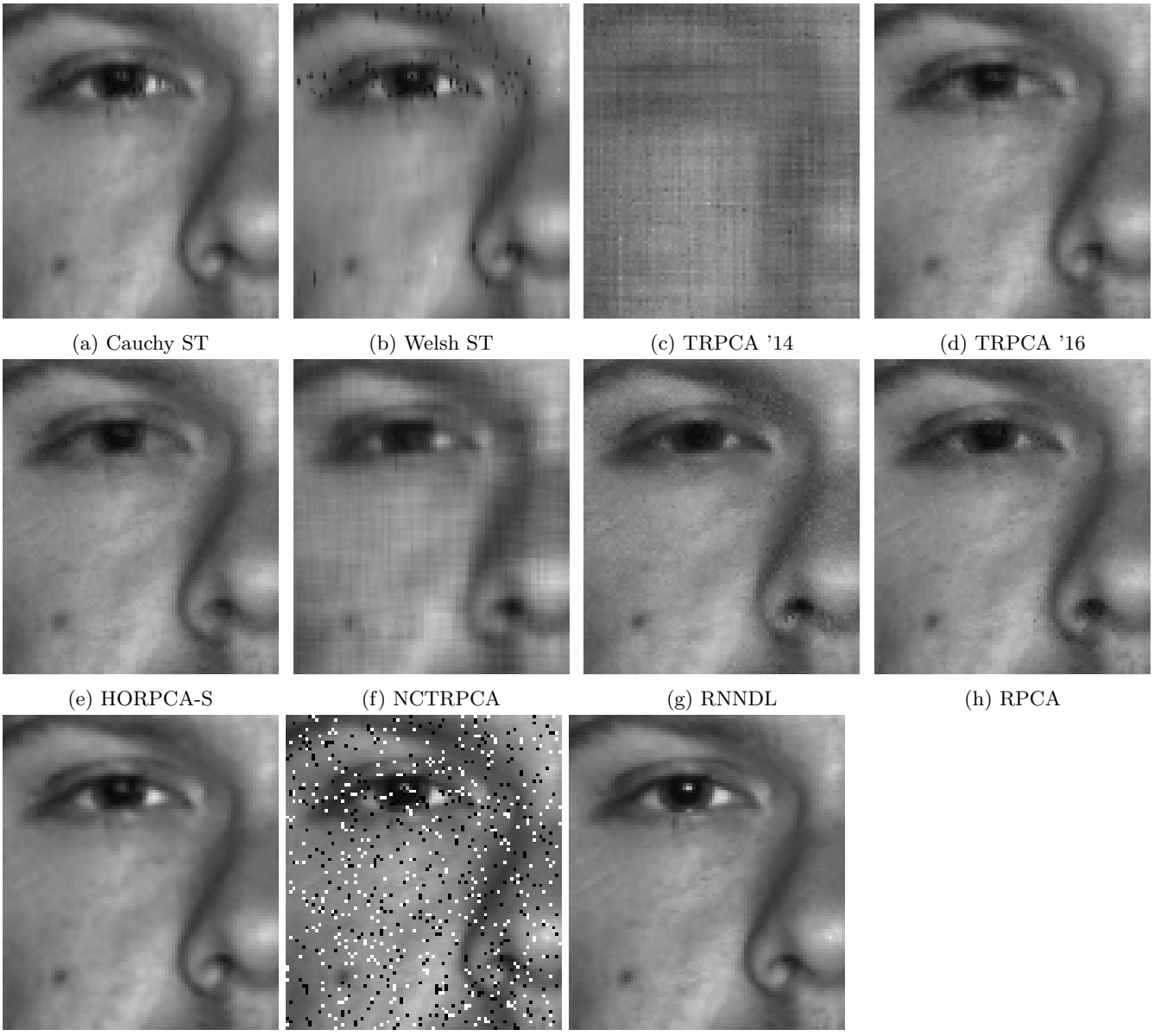


Figure 6: Mean FSIM and standard error at the 60% noise level.

4.2 Additional denoising results on Yale

We now present supplementary denoising results. In addition to the first illumination, we show the reconstruction obtained on the third illumination, and include the output of all algorithms.

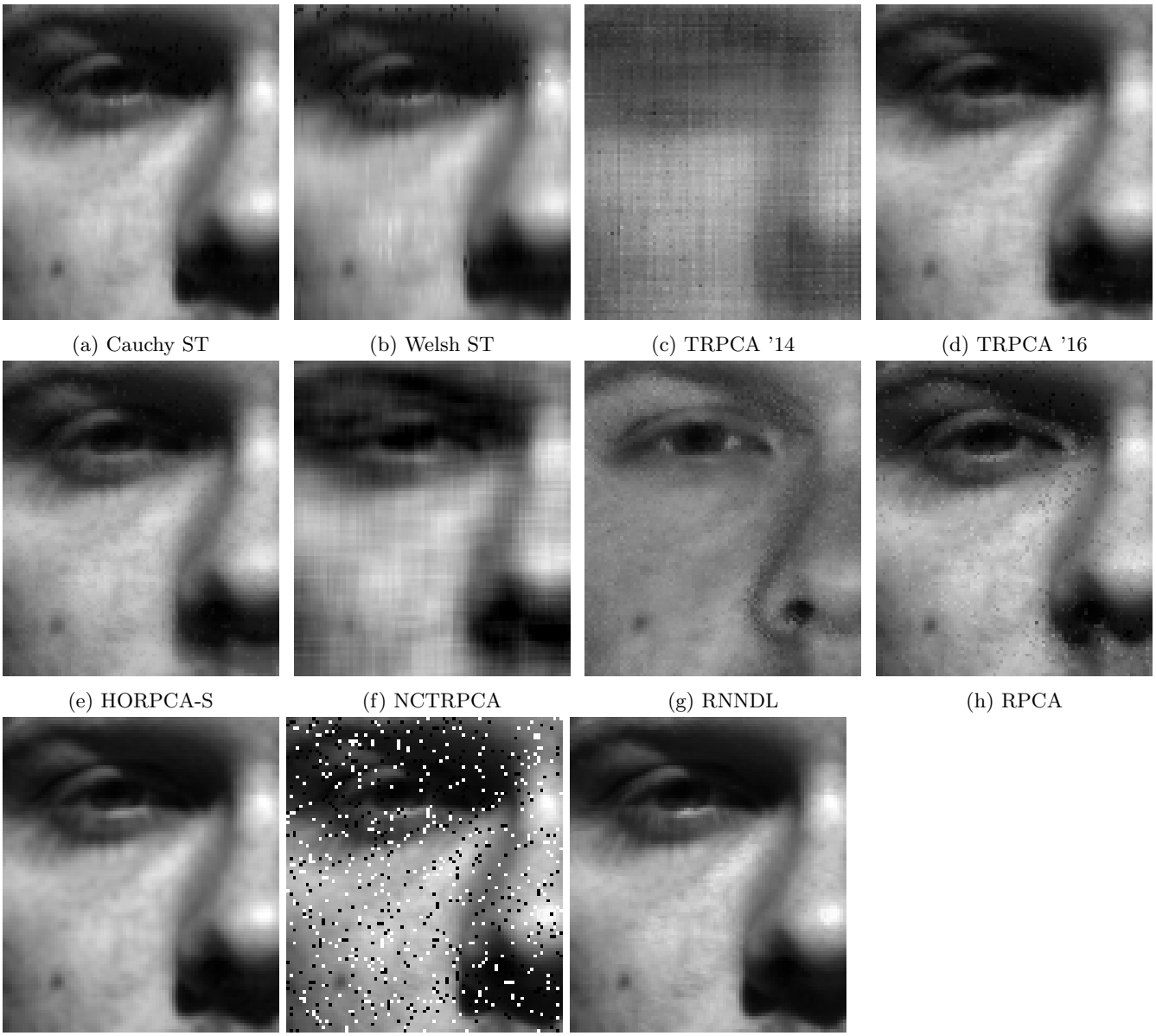


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 7: Comparative performance on the Yale benchmark with 10% salt & pepper noise - first illumination.

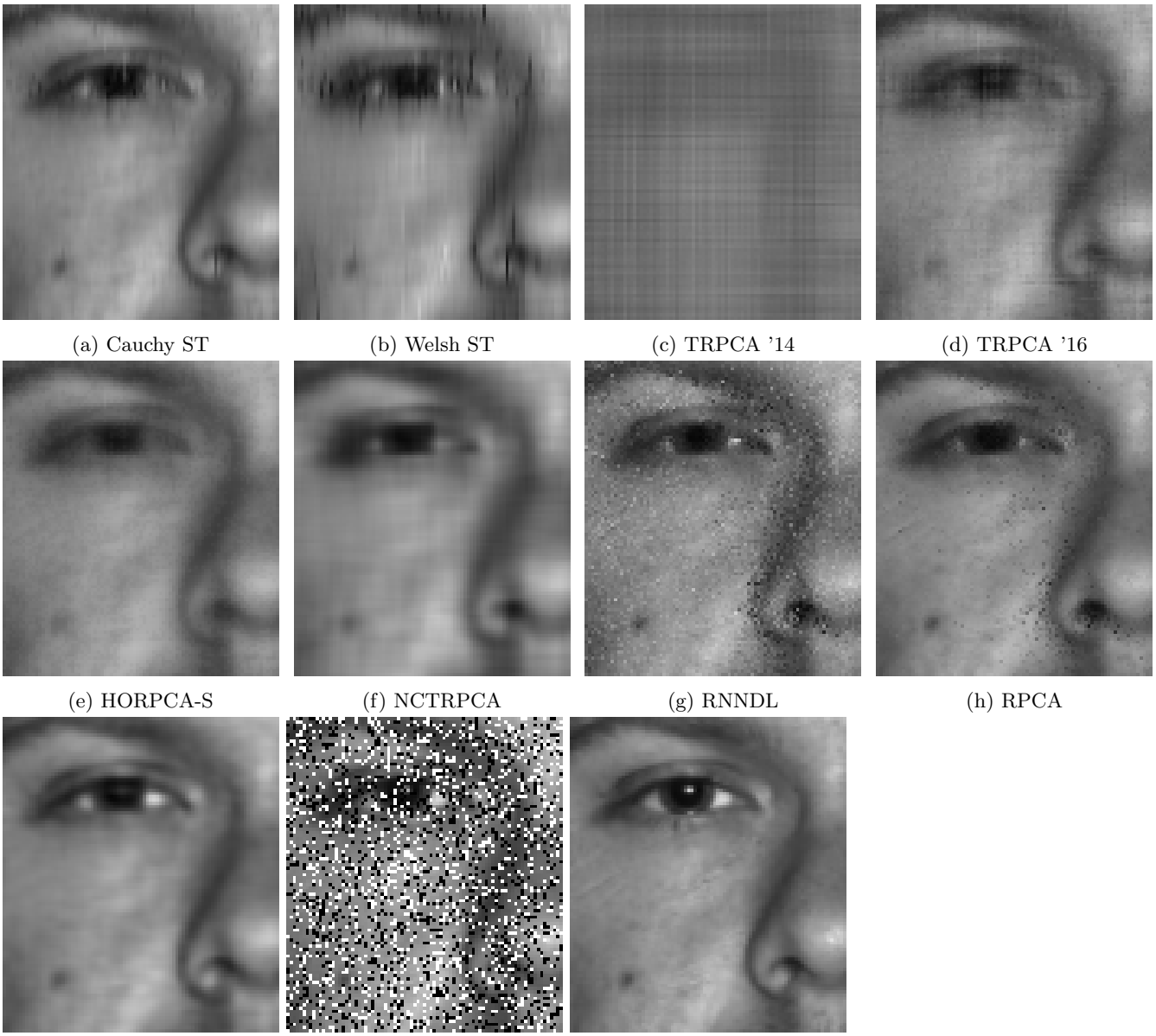


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 8: Comparative performance on the Yale benchmark with 10% salt & pepper noise - third illumination.

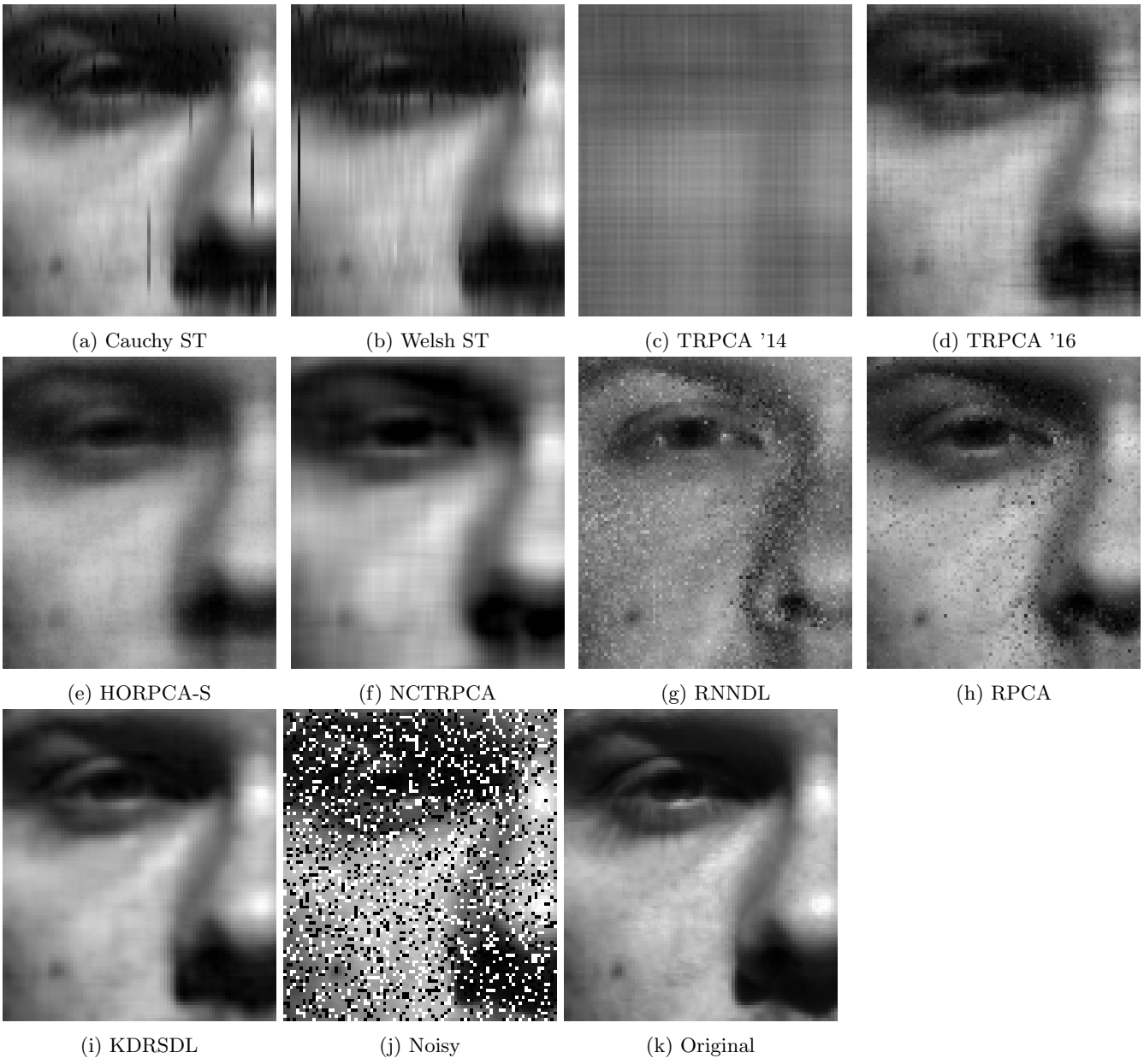


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 9: Comparative performance on the Yale benchmark with 30% salt & pepper noise - first illumination.

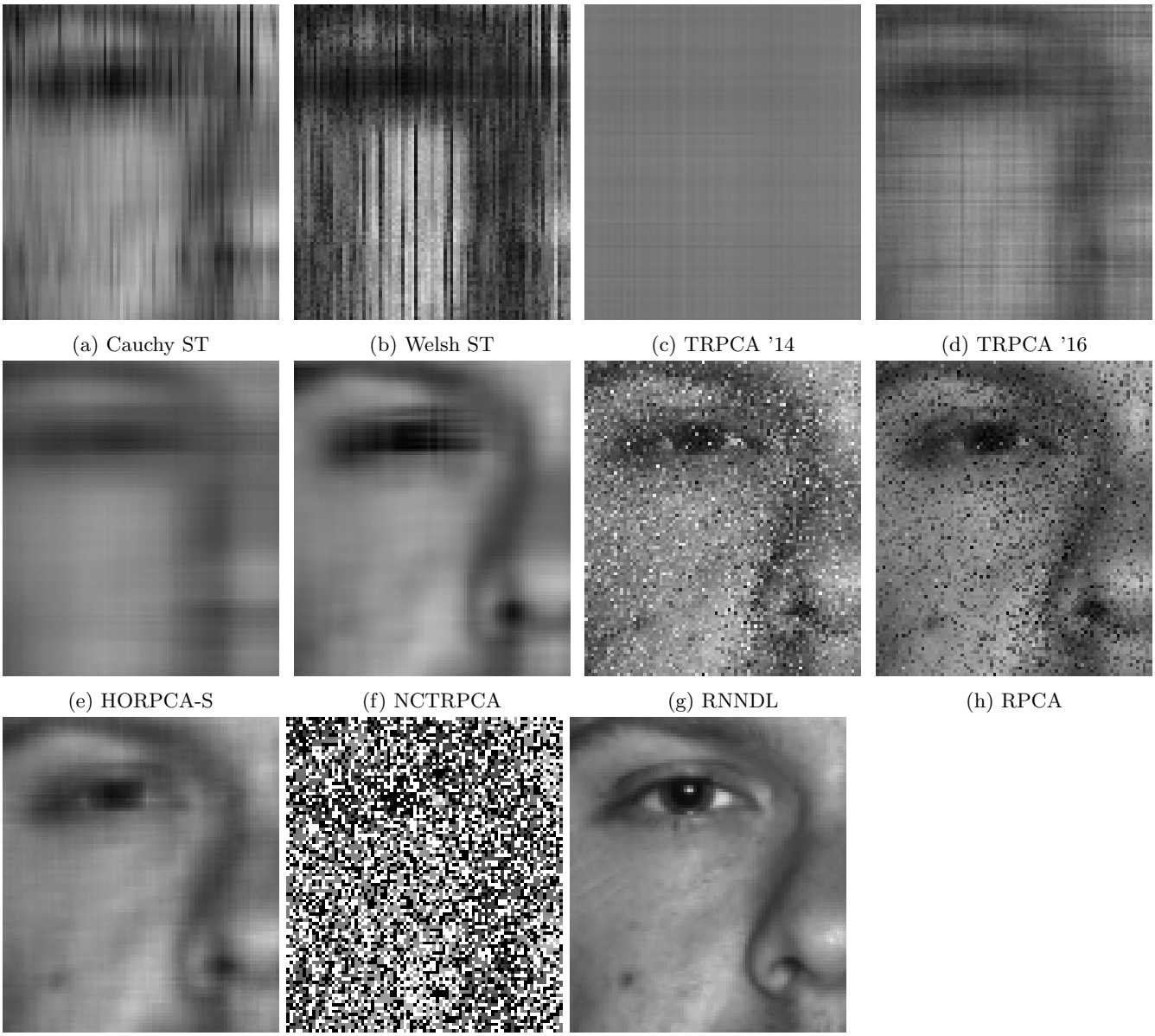


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 10: Comparative performance on the Yale benchmark with 30% salt & pepper noise - third illumination.

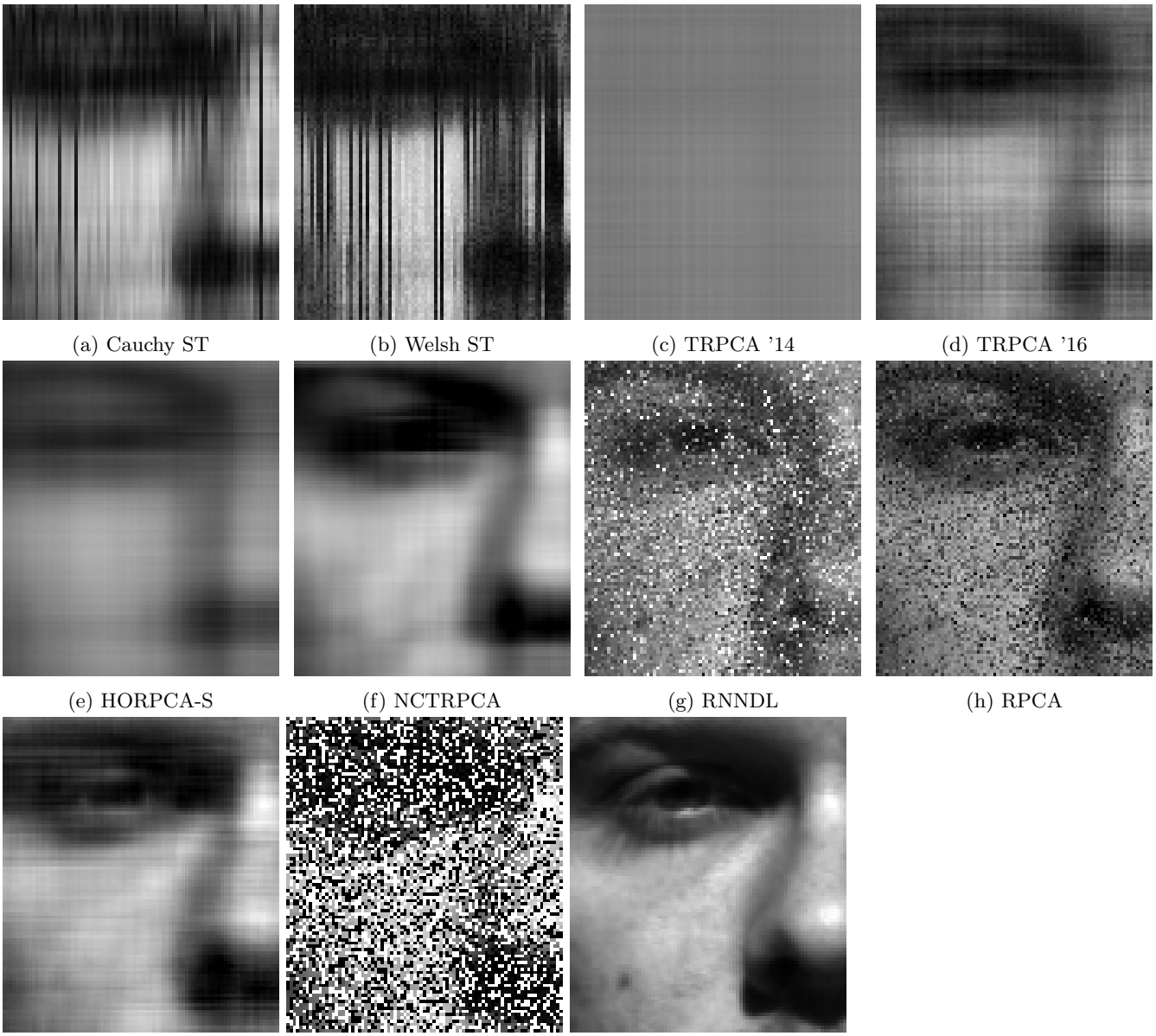


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 11: Comparative performance on the Yale benchmark with 60% salt & pepper noise - first illumination.



(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

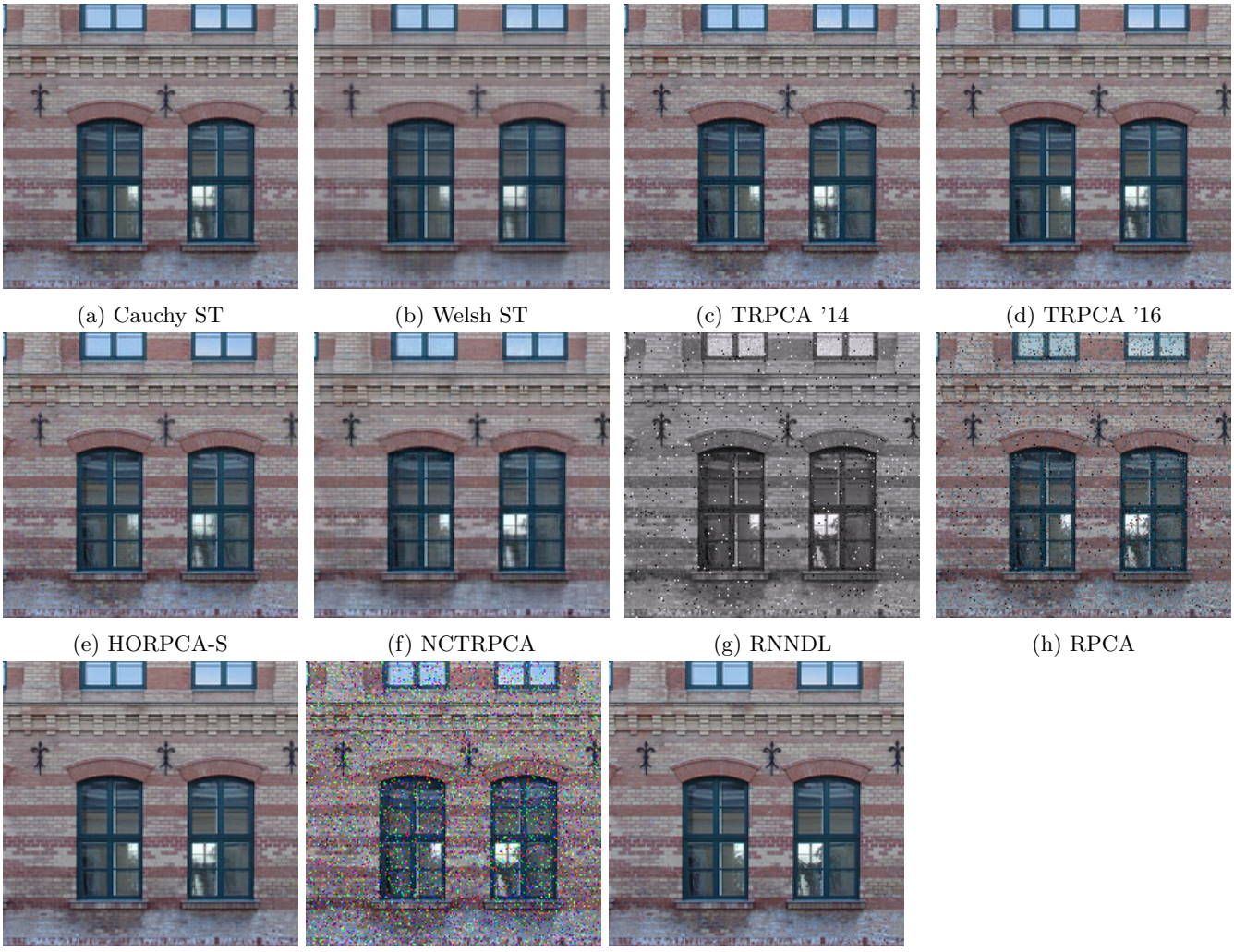
(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 12: Comparative performance on the Yale benchmark with 60% salt & pepper noise - third illumination.

4.3 Additional denoising results on Facade

We present the full denoising results at the 10% and 60% noise level, as well as the 30% level for completeness.

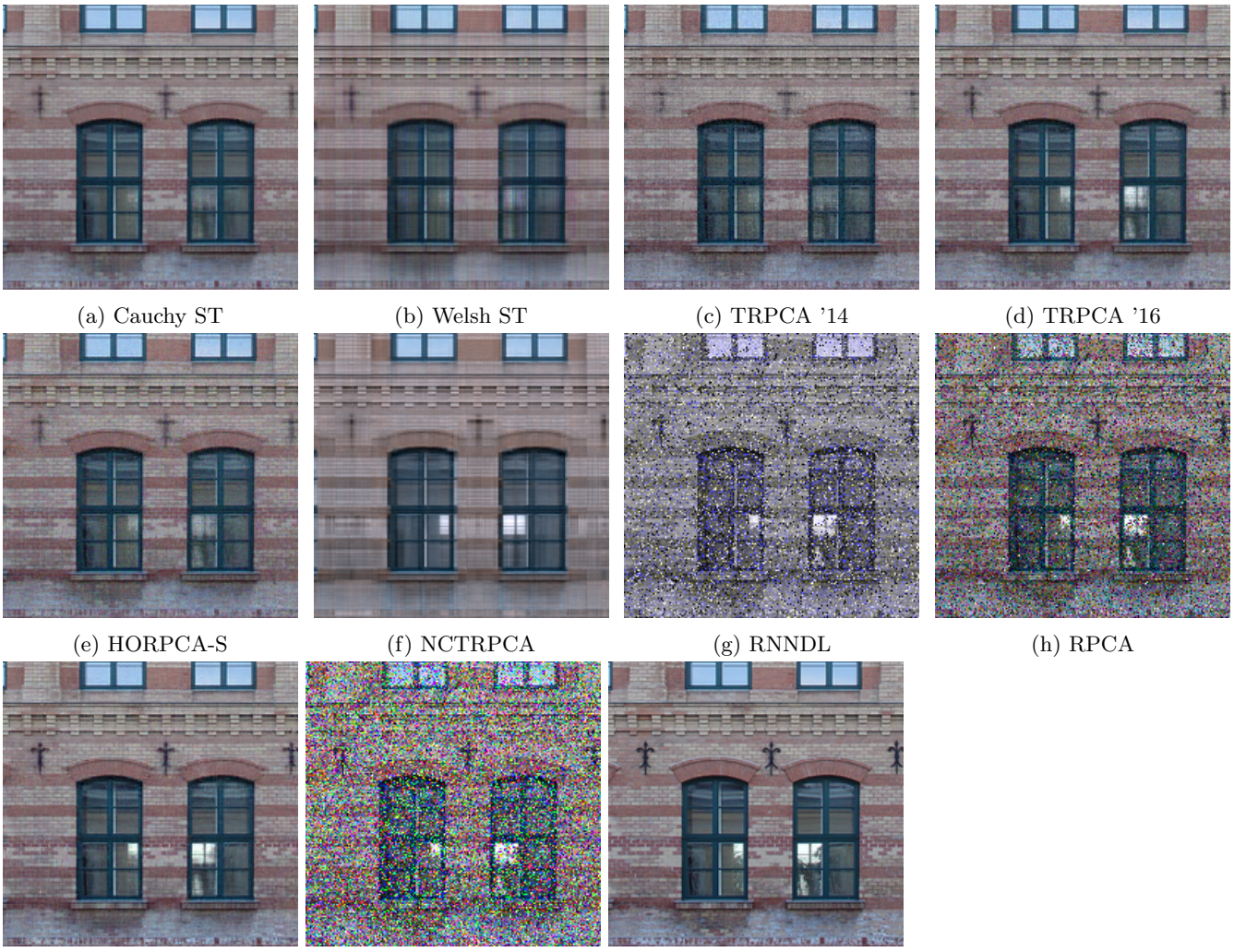


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 13: Comparative performance on the Facade benchmark with 10% noise.

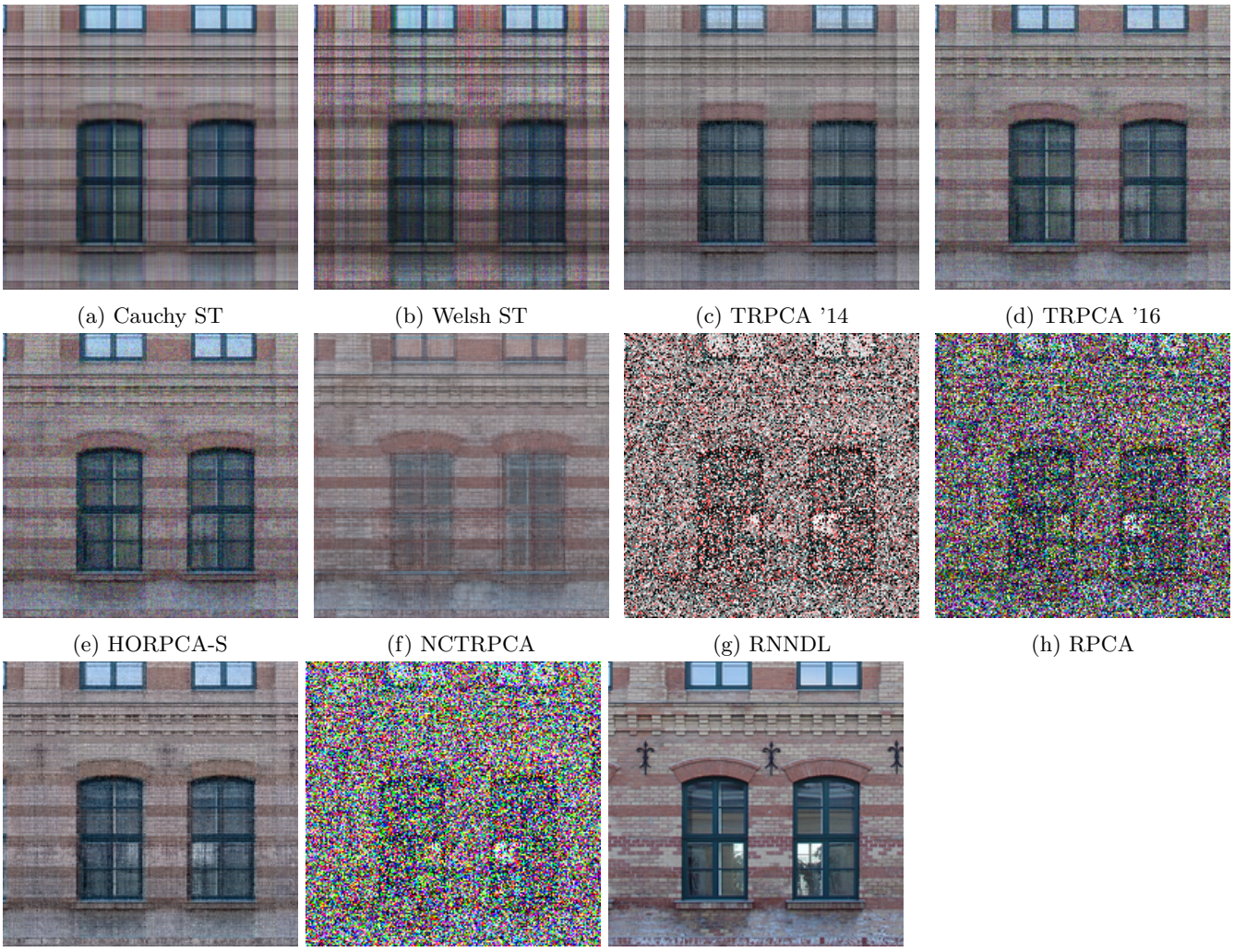


(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 14: Comparative performance on the Facade benchmark with 30% noise.



(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Noisy (k) Original

Figure 15: Comparative performance on the Facade benchmark with 60% noise.

4.4 Comparison with SeDiL

We present an experiment to assess the performance of SeDiL on the Yale-B benchmark. We chose this experiment because - as presented in the original paper [2] - SeDiL is designed for denoising grayscale images.

4.4.1 Design

The data and the noise are the same as in the Yale-B experimental results presented in the paper. We describe the tuning procedure employed.

We tuned in SeDiL: κ and λ for training the dictionary, and the parameter used in FISTA for denoising that we shall denote as λ_F. In the original paper, the authors choose $\kappa = \lambda = \frac{0.1}{4w^2}$ with w the dimension of the square image patches, and $\lambda_F = \frac{\sigma_{noise}}{100}$. We tuned the parameters via grid-search, choosing:

- $\kappa = \frac{\kappa_0}{4w^2}$, $\kappa_0 \in \text{linspace}(0.05, 0.5, 5)$
- $\lambda = \frac{\lambda_0}{4w^2}$, $\lambda_0 \in \text{linspace}(0.05, 0.5, 5)$
- $\lambda_F \in 5 * \text{logspace}(-4, 0, 15)$

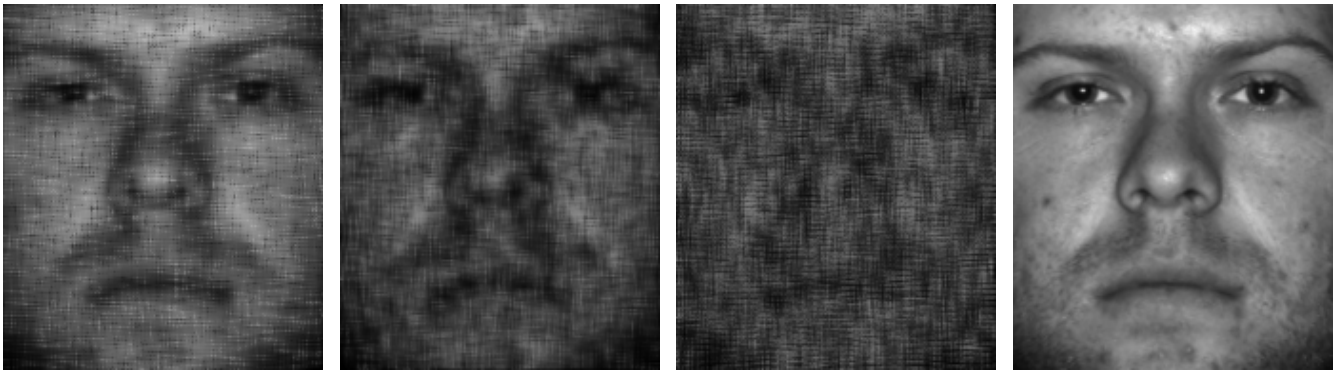
We kept $\rho = 100$ and patch sizes of 8×8 and extracted 40000 random patches for training from the 64 images of the first subject. We then followed the procedure described in the paper for denoising and extracted 8×8 patches in a sliding window with step size 1, denoised the patches with FISTA, and reconstructed the denoised image by averaging the denoised patches.

4.4.2 Results

We present below the results at the three noise levels: 10%, 30%, 60%. We report the mean PSNR and FSIM on the 64 images, the PSNR and FSIM on the first image, and show the reconstructions obtained.

Noise level	**Best PSNR face 1**	**Best mean PSNR**	**Best FSIM face 1**	**Best mean FSIM**
10%	23.937344	23.863287	0.840798	0.862061
30%	20.988780	19.249498	0.785824	0.807403
60%	17.480783	16.095463	0.703943	0.778116

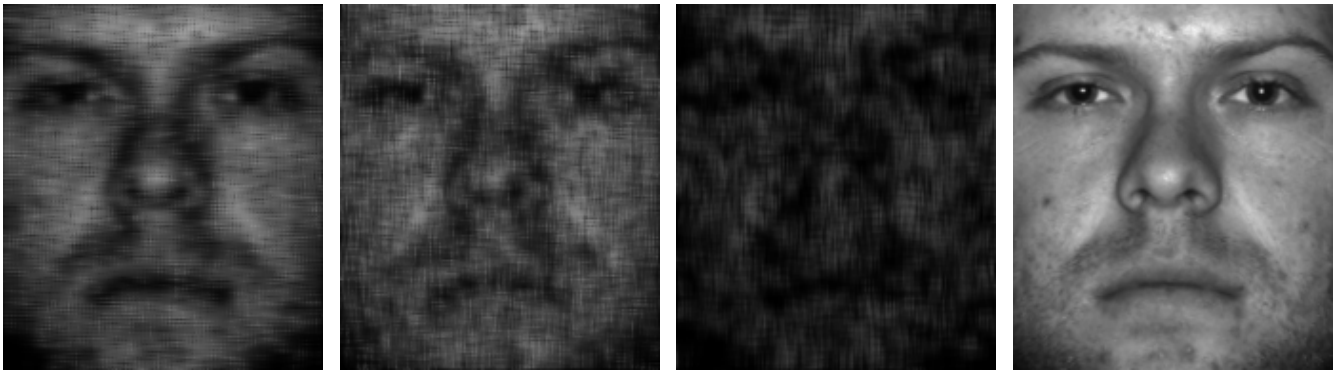
Table 6: Best PSNR and FSIM obtained on the first face and averaged on the 64 faces at the 3 noise levels.



(a) First face, 10% (b) First face, 30% (c) First face, 60% (d) Original face 1

(e) Third face, 10% (f) Third face, 30% (g) Third face, 60% (h) Original face 3

Figure 16: Best reconstructions obtained with SeDiL on the first and third illumination according to the best overall PSNR score.



(a) First face, 10% (b) First face, 30% (c) First face, 60% (d) Original face 1

(e) Third face, 10% (f) Third face, 30% (g) Third face, 60% (h) Original face 3

Figure 17: Best reconstructions obtained with SeDiL on the first and third illumination according to the best overall FSIM score.

4.4.3 Comments

As expected, the method is not robust to gross corruption. In all cases, the best results were obtained for $\kappa_0 = \lambda_0 = 0.5$ and $\lambda_F = 5$, indicative of the lack of robustness.

4.5 Background subtraction

In this section we present the DET curves obtained on the *Highway* and *Hall* experiments, and the backgrounds obtained with each algorithm on the *Highway* benchmark.

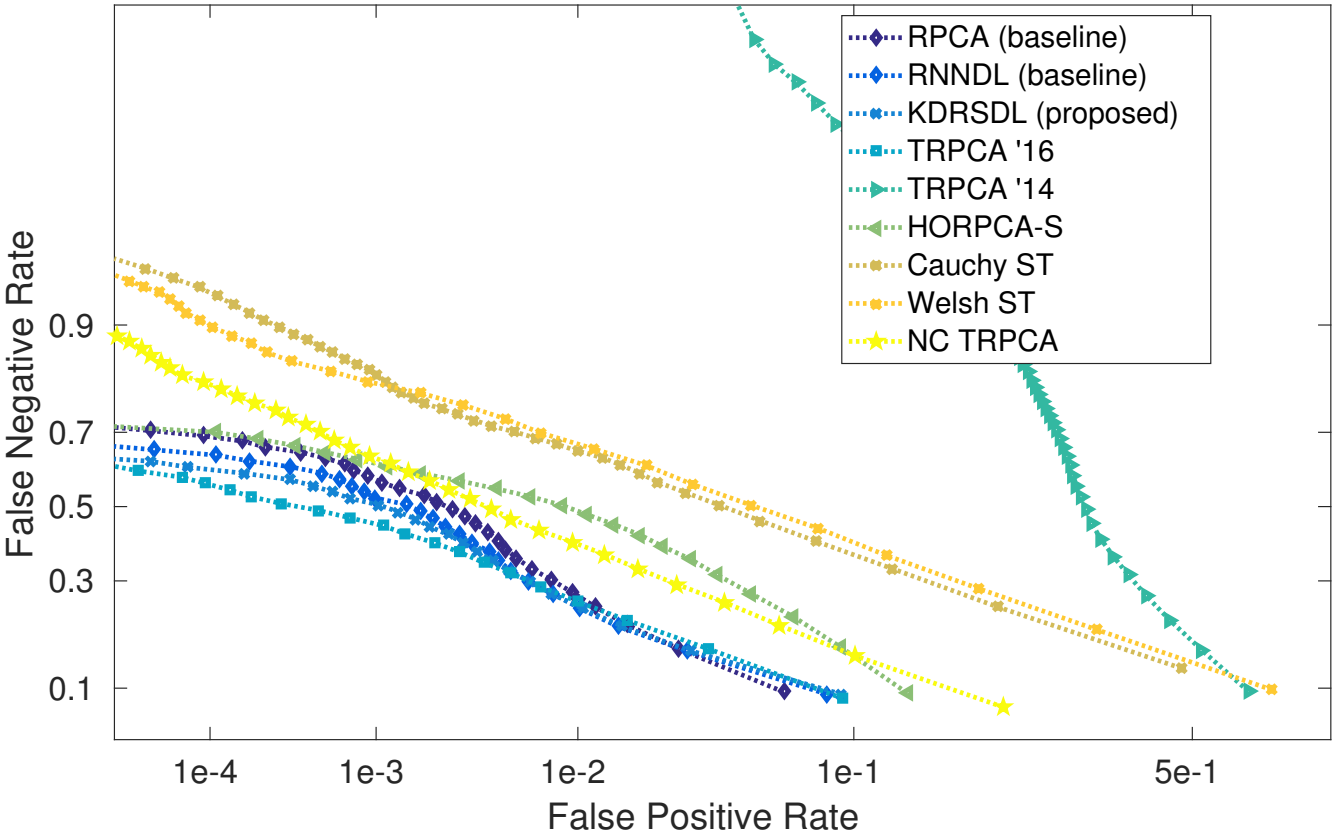


Figure 18: DET curves on the *Highway* dataset.

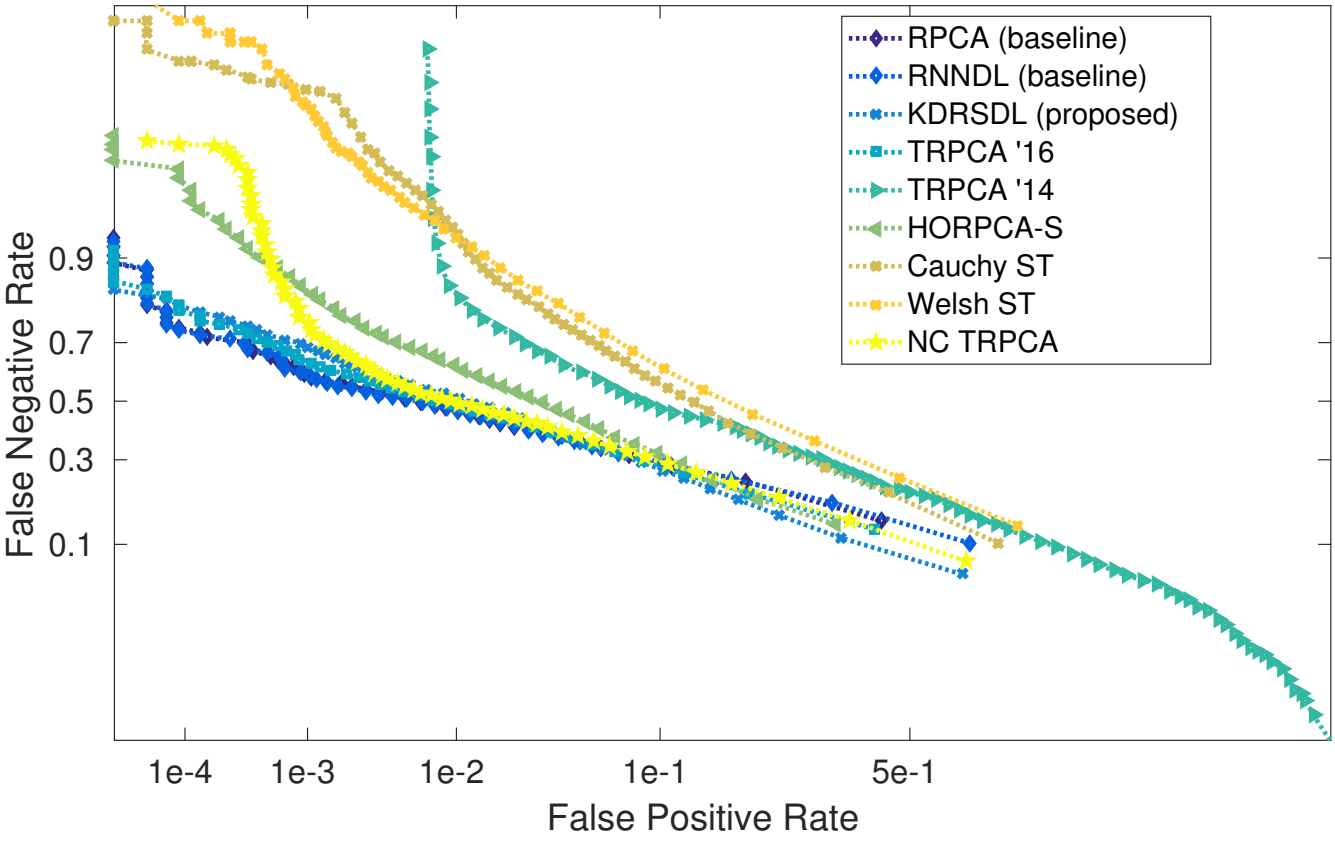
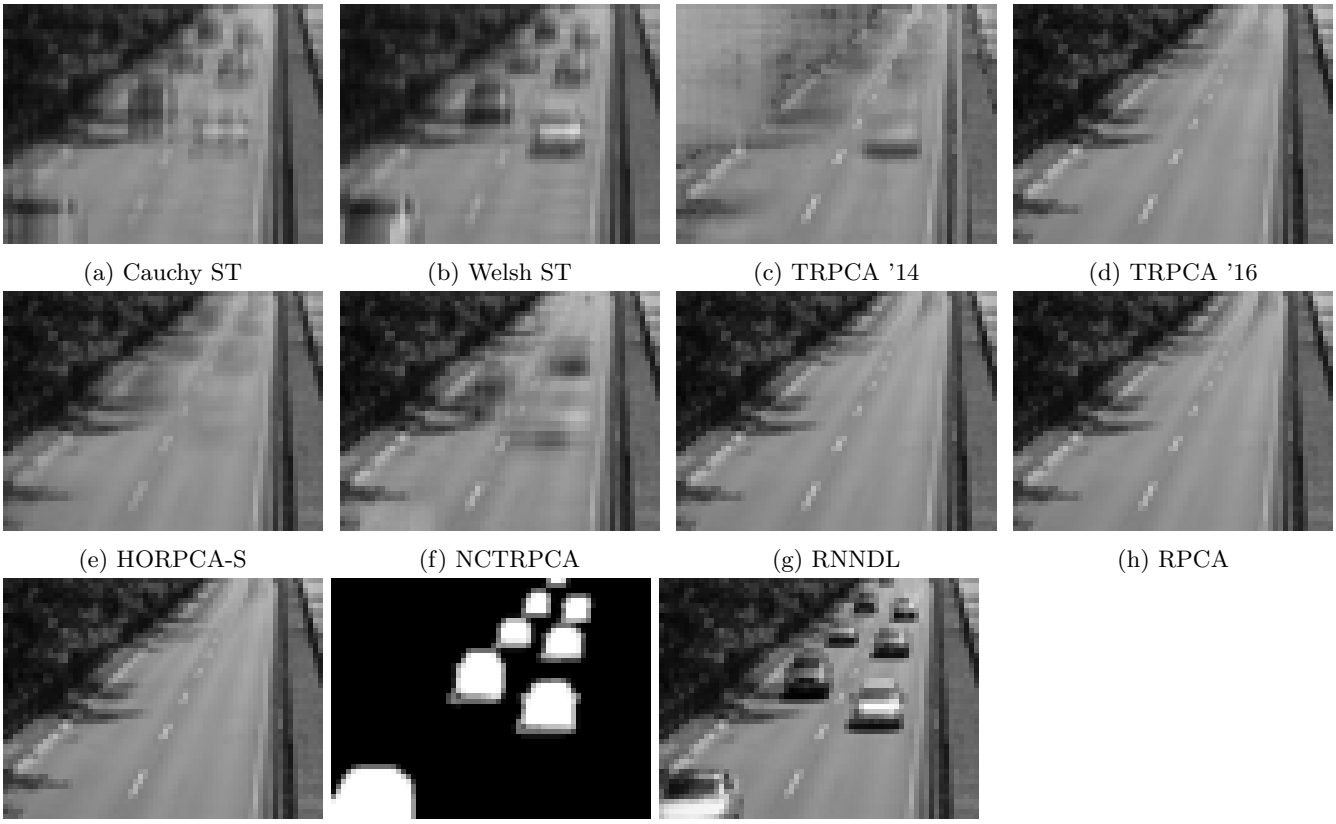


Figure 19: DET curves on the *Airport Hall* dataset.



(a) Cauchy ST (b) Welsh ST (c) TRPCA '14 (d) TRPCA '16

(e) HORPCA-S (f) NCTRPCA (g) RNNDL (h) RPCA

(i) KDRSDL (j) Ground truth (k) Original frame

Figure 20: Comparative performance on the Highway benchmark.

4.6 Synthetic data

In the paper, we presented curves of the reconstruction error on the low-rank tensor and of the density of the sparse component on synthetic data as a function of the parameter λ. Here we present additional results showing the reconstruction (ℓ_2) error on the sparse component in the 60% corruption case, and results in the 30% corruption case to support the claims made in section 3.1.

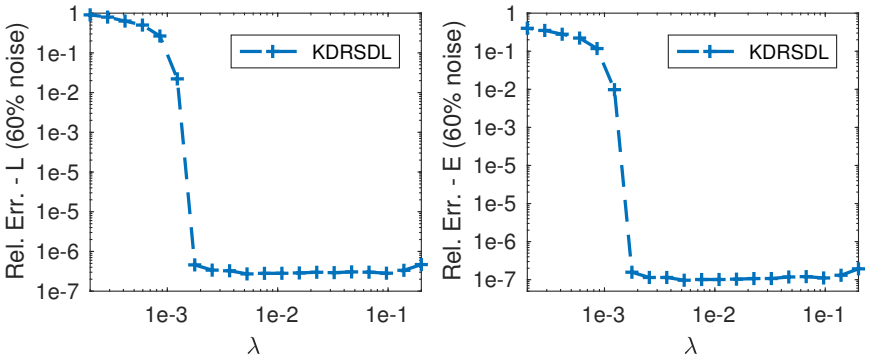


Figure 21: ℓ_2 error on the recovered dense (low-rank) and sparse components - 60% corruption.

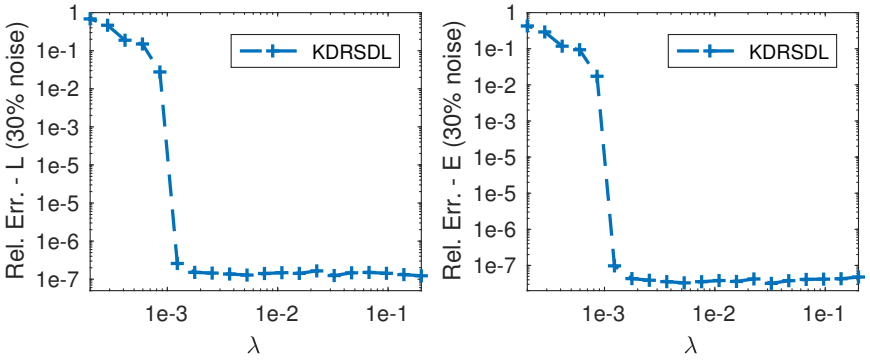


Figure 22: ℓ_2 error on the recovered dense (low-rank) and sparse components - 30% corruption.

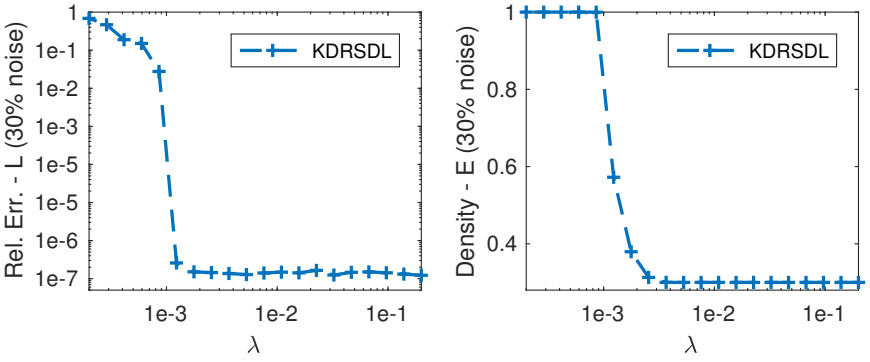


Figure 23: ℓ_2 error on the recovered dense (low-rank) and density of the sparse component - 30% corruption.

5 Original images

Figure 24: Original image - Facade dataset

Figure 25: Original image - Yale-B subject 1, pose 1, illumination 1

Figure 26: Original image - Yale-B subject 1, pose 1, illumination 3

6 Result remove because of numerical instability

The following is the reconstruction of the background provided by the algorithm known as "Tensor RPCA (TRPCA '14)" in the paper. Due to numerical instability, the method did not provide an actual separation of the background.

Figure 27: Background separated by "Tensor RPCA (TRPCA '14)" on the Airport Hall dataset

The reconstruction provided by "Tensor RPCA (TRPCA '14)" for the Yale dataset at 30% noise on is presented in Figure 9.

References

[1] Laub, Aj: *Matrix Analysis for Scientists and Engineers*. Society for Industrial and Applied Mathematics, 2005. http://dx.doi.org/10.1109/TSP.2007.893956. http://dx.doi.org/10.1109/TSP.2007.893956. – ISBN 0898715768

[2] Hawe, Simon ; Seibert, Matthias ; Kleinsteuber, Martin: Separable Dictionary Learning. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (2013), 3. http://arxiv.org/abs/1303.5244