

# Partial Sum Minimization of Singular Values in RPCA for Low-Level Vision \*

Tae-Hyun Oh\*  
KAIST

Hyeonwoo Kim  
KAIST

Yu-Wing Tai  
KAIST

Jean-Charles Bazin  
ETHZ

In So Kweon  
KAIST

## Abstract

*Robust Principal Component Analysis (RPCA) via rank minimization is a powerful tool for recovering underlying low-rank structure of clean data corrupted with sparse noise/outliers. In many low-level vision problems, not only it is known that the underlying structure of clean data is low-rank, but the exact rank of clean data is also known. Yet, when applying conventional rank minimization for those problems, the objective function is formulated in a way that does not fully utilize a priori target rank information about the problems. This observation motivates us to investigate whether there is a better alternative solution when using rank minimization.*

*In this paper, instead of minimizing the nuclear norm, we propose to minimize the partial sum of singular values. The proposed objective function implicitly encourages the target rank constraint in rank minimization. Our experimental analyses show that our approach performs better than conventional rank minimization when the number of samples is deficient, while the solutions obtained by the two approaches are almost identical when the number of samples is more than sufficient. We apply our approach to various low-level vision problems, e.g. high dynamic range imaging, photometric stereo and image alignment, and show that our results outperform those obtained by the conventional nuclear norm rank minimization method.*

## 1. Introduction

Robust Principal Component Analysis (RPCA) [5, 3] aims to recover a low-rank matrix  $\mathbf{A} \in \mathbb{R}^{m \times n}$ , from corrupted observations  $\mathbf{O} = \mathbf{A} + \mathbf{E}$ , where  $\mathbf{E} \in \mathbb{R}^{m \times n}$  represents errors with arbitrary magnitude and distribution. The rank minimization approach [21, 23, 3, 4] assumes  $\mathbf{E}$  is sparse in its distribution and formulates the problem as:

$$\arg \min_{\mathbf{A}, \mathbf{E}} \|\mathbf{A}\|_* + \lambda \|\mathbf{E}\|_1, \quad \text{s.t. } \mathbf{O} = \mathbf{A} + \mathbf{E}, \quad (1)$$

where  $\|\mathbf{A}\|_* = \sum_i \sigma_i(\mathbf{A})$  is the nuclear norm of  $\mathbf{A}$ ,  $\|\mathbf{E}\|_1$  is the  $l^1$ -norm of  $\mathbf{E}$  and  $\lambda$  is the relative weight between

the two terms. Eq. (1) can be solved effectively by various methods [19, 24].

Wright *et al.* [23] and Candés *et al.* [3] proved that, under broad conditions, the solution of Eq. (1) is unique<sup>1</sup>. Yet, when the number of samples in  $\mathbf{O}$  is very limited, we observe that the unique solution of Eq. (1) might include some outliers as inliers and vice versa. The converged solution can be degenerated. For instance, in the photometric stereo problem [25], the solution of  $\mathbf{A}$  might have a rank lower than the theoretical rank of 3. Such limited observations are not uncommon in many computer vision problems due to practical reasons. For example, in High Dynamic Range (HDR) context, often only 2-4 differently exposed images are captured and photometric stereo requires only 3 input images in theory.

In this paper, based on the prior knowledge about the rank of  $\mathbf{A}$ , we propose an alternative objective function which minimizes the partial sum of singular values of  $\mathbf{A}$ :

$$\arg \min_{\mathbf{A}, \mathbf{E}} \sum_{i=N+1}^{\min(m,n)} \sigma_i(\mathbf{A}) + \lambda \|\mathbf{E}\|_1, \quad \text{s.t. } \mathbf{O} = \mathbf{A} + \mathbf{E}, \quad (2)$$

where  $N$  is the target rank of  $\mathbf{A}$  which can be derived from problem definition, e.g.  $N = 1$  for background subtraction,  $N = 3$  for photometric stereo. Eq. (2) minimizes the rank of residual errors of  $\mathbf{A}$ , instead of the nuclear norm. Although Eq. (2) is non-convex, we observe in our experiments that Eq. (2) encourages the resulting low-rank matrix to have the rank close to  $N$  even with deficient observations.

We empirically studied the proposed objective function in many low-level vision problems, e.g. HDR imaging, photometric stereo, and image alignment, where the theoretical rank of  $\mathbf{A}$  is known and the number of observations is limited. Our experimental analyses show that our formulation, described in Eq. (2), converges to a solution more robust to outliers than the solution obtained by the conventional objective function in Eq. (1) in rank minimization, when the number of observations is limited. Empirically, we also found that the solutions of Eq. (1) and Eq. (2) are almost identical when there are more than the sufficient number of observed samples.

\*Acknowledgement: This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MEST) (No. 2010-0028680).

<sup>1</sup>Uniqueness is subject to the choice of  $\lambda$ . Wright *et al.* [23] suggested  $\lambda = O(m^{-1/2})$  where  $m$  is dimension of  $\mathbf{O}$ .

In short summary, our contributions are as follows:

- We present a partial sum objective function and its corresponding minimization method for RPCA.
- We empirically study the partial sum objective function and claim that it outperforms conventional rank minimization when the number of observed samples is very limited.
- We apply our technique on various low-level vision problems and demonstrate superior results over previous works. We also demonstrate how these low-level vision problems can be formulated into our framework.

## 2. Related Works

In this section, we briefly review early works related to RPCA, then we discuss some recent advances in RPCA and its applications in computer vision. We invite readers to refer to Candés *et al.* [3] for a thoughtful review of RPCA.

In conventional PCA [17], the goodness-of-fit of data is evaluated by the  $l^2$ -norm which is very sensitive to outliers. Early works in RPCA tried to reduce the effects of outliers by random sampling [10], robust M-estimator [5, 6], or alternating minimization [18] to identify outliers or to penalize data with large errors. These methods share some limitations either they are sensitive to the choice of parameters or their algorithms are not polynomial in running time.

Recent advances in RPCA showed that the heuristic nuclear norm solution [21, 23, 3], which minimizes the rank of the data matrix, converges to a solution which is robust to sparse outliers. Candés *et al.* [3] formulated the RPCA problem as:

$$\min_{\mathbf{A}, \mathbf{E}} \text{rank}(\mathbf{A}) + \lambda \|\mathbf{E}\|_0, \quad \text{s.t. } \mathbf{O} = \mathbf{A} + \mathbf{E}. \quad (3)$$

where  $\|\cdot\|_0$  denotes the  $l^0$ -norm. Since Eq. (3) is an NP-hard problem, Candés *et al.* [3] proved that the convex relaxation version in Eq. (1) approximates the solution of Eq. (3) as long as  $\mathbf{E}$  is randomly sparse and the underlying  $\text{rank}(\mathbf{A})$  is lower than a certain upper bound. To solve Eq. (1), various methods have been proposed [19, 24]. Among them, the inexact augmented Lagrange multiplier (ALM) [19] has shown to be good in terms of computational efficiency.

The robustness and scalability of the rank minimization algorithm for RPCA [3, 19, 24] have inspired many applications in computer vision, such as background subtraction [3], image and video restoration [16], image alignment [20], regular texture analysis [26], and robust photometric stereo [25]. These applications are based on the observation that the underlying structures of clean data are linearly correlated, which forms a low-rank data matrix. As briefly mentioned in the introduction, in some applications, such as background subtraction [3] and photometric stereo [25], the rank of clean data can be determined by the problem definition. In practice, rank minimization proposed

by Candés *et al.* [3] is general in the sense that it does not require to know a priori the rank of clean data.

The success of rank minimization based RPCA comes from the blessing of dimensionality [8, 23], implying large amount of observations. However, when the number of observations is limited, which is common in practice, results from RPCA might be degenerated, *e.g.* correct samples might be considered as outliers and vice versa.

The impetus of this work is to introduce an alternative objective function that can efficiently deal with deficient examples in rank minimization problem. Seeing the limitation of rank minimization as an addendum, the proposed alternative objective function can control the lower bound of the rank with a simple and efficient minimizer. We demonstrate the effectiveness of our proposed objective function through thoughtful experiments.

## 3. Partial Sum Minimization by Partial Singular Value Thresholding Operator

Our partial sum objective function in Eq. (2) is a constrained optimization problem. To solve problems of this type, Lin *et al.* [19] proposed the augmented Lagrange multipliers (ALM) method which is known to converge fast and be scalable. The augmented Lagrangian function of Eq. (2) is formulated by:

$$L(\mathbf{A}, \mathbf{E}, \mathbf{Z}, \mu) = \sum_{i=N+1}^{\min(m,n)} \sigma_i(\mathbf{A}) + \lambda \|\mathbf{E}\|_1 + \langle \mathbf{Z}, \mathbf{O} - \mathbf{A} - \mathbf{E} \rangle + \frac{\mu}{2} \|\mathbf{O} - \mathbf{A} - \mathbf{E}\|_F^2, \quad (4)$$

where  $\mu$  is a positive scalar, and  $\mathbf{Z} \in \mathbb{R}^{m \times n}$  is an estimate of the Lagrange multiplier. Minimizing the Lagrangian function directly might be particularly challenging. According to a recent development of alternating direction minimization [19], Eq. (4) can be solved by minimizing each variable alternatively while fixing the other variables. This is not equivalent to the exact minimization, but in practice the converged solution is very close to the solution of the original problem while still satisfying the constraints. From the Lagrangian function in Eq. (4), the optimization problem can be divided into two sub-problems:

$$\begin{aligned} \mathbf{A}^* &= \arg \min_{\mathbf{A}} L(\mathbf{A}, \mathbf{E}_k, \mathbf{Z}_k, \mu_k) \\ &= \arg \min_{\mathbf{A}} \sum_{i=N+1}^{\min(m,n)} \sigma_i(\mathbf{A}) \\ &\quad + \langle \mathbf{Z}_k, \mathbf{O} - \mathbf{A} - \mathbf{E}_k \rangle + \frac{\mu_k}{2} \|\mathbf{O} - \mathbf{A} - \mathbf{E}_k\|_F^2 \\ &= \arg \min_{\mathbf{A}} \mu_k^{-1} \sum_{i=N+1}^{\min(m,n)} \sigma_i(\mathbf{A}) \\ &\quad + \frac{1}{2} \|\mathbf{A} - (\mathbf{O} - \mathbf{E}_k + \mu^{-1} \mathbf{Z}_k)\|_F^2, \end{aligned} \quad (5)$$

$$\begin{aligned}
\mathbf{E}^* &= \arg \min_{\mathbf{E}} L(\mathbf{A}_k, \mathbf{E}, \mathbf{Z}_k, \mu_k) \\
&= \arg \min_{\mathbf{E}} \lambda \|\mathbf{E}\|_1 \\
&\quad + \langle \mathbf{Z}_k, \mathbf{O} - \mathbf{A}_k - \mathbf{E} \rangle + \frac{\mu_k}{2} \|\mathbf{O} - \mathbf{A}_k - \mathbf{E}\|_F^2 \\
&= \arg \min_{\mathbf{E}} \lambda \mu_k^{-1} \|\mathbf{E}\|_1 \\
&\quad + \frac{1}{2} \|\mathbf{E}_k - (\mathbf{O} - \mathbf{A}_k + \mu_k^{-1} \mathbf{Z}_k)\|_F^2. \tag{6}
\end{aligned}$$

where  $k$  indicates the iteration index (see Alg. 1).

**Solving  $\mathbf{A}^*$**  To minimize Eq. (5), we first define the Partial Singular Value Thresholding (PSVT) operator  $\mathbb{P}_{N,\tau}[\cdot]$ .

Let  $\tau > 0$ ,  $l = \min(m, n)$  and  $\mathbf{X}, \mathbf{Y} \in \mathbb{R}^{m \times n}$  which can be decomposed by SVD and  $\mathbf{Y}$  can be considered as the sum of two matrices,  $\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2 = \mathbf{U}_{Y1} \mathbf{D}_{Y1} \mathbf{V}_{Y1}^T + \mathbf{U}_{Y2} \mathbf{D}_{Y2} \mathbf{V}_{Y2}^T$ , where  $\mathbf{U}_{Y1}, \mathbf{V}_{Y1}$  are the singular vector matrices corresponding to the largest singular values from the first to the  $N$ -th, and  $\mathbf{U}_{Y2}, \mathbf{V}_{Y2}$  from the  $(N+1)$ -th to the last from SVD. Eq. (5) can be re-written in the form of the following minimization problem:

$$\arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \tau \sum_{i=N+1}^l \sigma_i(\mathbf{X}), \tag{7}$$

It can be shown that the solution of such problem is given by the following PSVT operator:

$$\begin{aligned}
\mathbb{P}_{N,\tau}[\mathbf{Y}] &= \mathbf{U}_Y (\mathbf{D}_{Y1} + \mathcal{S}_\tau[\mathbf{D}_{Y2}]) \mathbf{V}_Y^T \\
&= \mathbf{Y}_1 + \mathbf{U}_{Y2} \mathcal{S}_\tau[\mathbf{D}_{Y2}] \mathbf{V}_{Y2}^T, \\
&\quad \text{where } \mathbf{D}_{Y1} = \text{diag}(\sigma_1, \dots, \sigma_N, 0, \dots, 0), \\
&\quad \mathbf{D}_{Y2} = \text{diag}(0, \dots, 0, \sigma_{N+1}, \dots, \sigma_l), \tag{8}
\end{aligned}$$

where  $\mathcal{S}_\tau[x] = \text{sign}(x) \cdot \max(|x| - \tau, 0)$  is the soft-thresholding (shrinkage) operator [12], and  $x \in \mathbb{R}$ . This operator can be extended to vectors and matrices by applying it element-wisely. The detail proof and derivation of the PSVT operator can be found in the supplementary material. Notice that the proposed PSVT operator provides the closed-form solution of the sort of Eq. (8) (e.g. Eq. (5)).

To clarify PSVT, when  $\tau = \infty$ , the optimal solution of Eq. (7) is a low-dimensional projection of  $\mathbf{Y}$  known as singular value projection [15] which enforces the target rank constraint through projection. When  $\sigma_i < \tau$  for  $1 \leq i \leq N$ , conventional SVT [2] projects  $\sigma_i$  to zero resulting in a more deficient rank of  $\mathbf{A}$  than the target rank while PSVT does not. Hence, PSVT implicitly encourages the resulting matrix  $\mathbf{A}$  to meet the target rank even when all the  $\sigma_i$  are small. This situation occasionally happens when the number of observed samples is limited. Since Eq. (5) is a non-convex function, PSVT only guarantees a local minimum for the sub-problem. Yet, PSVT monotonically decreases the value

---

**Algorithm 1** Partial sum of singular values minimization via the ALM method

---

**Input :**  $\mathbf{O} \in \mathbb{R}^{m \times n}$ ,  $\lambda > 0$ , the constraint rank  $N$ .

- 1: Initialize  $\mathbf{A}_0 = \mathbf{E}_0 = \mathbf{0}$ ,  $\mathbf{Z}$  as suggested in [19],  $\mu_0 > 0$ ,  $\rho > 1$  and  $k = 0$ .
- 2: **while** not converged **do**
- 3:     **while** not converged **do**
- 4:         Update  $\mathbf{A}$  by Eq. (10).
- 5:         Update  $\mathbf{E}$  by Eq. (10).
- 6:     **end while**
- 7:      $\mathbf{Z}_{k+1} = \mathbf{Z}_k + \mu_k (\mathbf{O} - \mathbf{A}_{k+1} - \mathbf{E}_{k+1})$ .
- 8:      $\mu_{k+1} = \rho \mu_k$ .
- 9:      $k = k + 1$ .
- 10: **end while**

**Output :**  $(\mathbf{A}_k, \mathbf{E}_k)$ .

---

of the objective function of the sub-problem, and the experiments in the next section show that it provides satisfying results.

**Solving  $\mathbf{E}^*$**  As suggested by Hale *et al.* [12], the solution to the sub-problem in Eq. (6) can be solved as the following form:

$$\mathcal{S}_\tau[\mathbf{Y}] = \arg \min_{\mathbf{X}} \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|_F^2 + \tau \|\mathbf{X}\|_1, \tag{9}$$

Despite its simple structure, shrinkage methods are shown to be very effective in minimizing the  $l^1$ -norm and the proximity term and guarantee that the solution is the global minimum [12, 2].

**Updating  $\mathbf{A}^*$  and  $\mathbf{E}^*$**  At each iteration,  $\mathbf{A}_k$  and  $\mathbf{E}_k$  can be updated with the operators  $\mathcal{S}_\tau[\cdot]$  and  $\mathbb{P}_{N,\tau}[\cdot]$  as:

$$\begin{aligned}
\mathbf{A}_{k+1} &= \mathbb{P}_{N,\mu_k^{-1}}[\mathbf{O} - \mathbf{E}_k + \mu_k^{-1} \mathbf{Z}_k] \\
\mathbf{E}_{k+1} &= \mathcal{S}_{\lambda \mu_k^{-1}}[\mathbf{O} - \mathbf{A}_{k+1} + \mu_k^{-1} \mathbf{Z}_k]. \tag{10}
\end{aligned}$$

We also found that updating  $\mathbf{A}_k$  and  $\mathbf{E}_k$  just once when solving the sub-problem is sufficient to converge to the solution of Eq. (2), called inexact ALM [19] which is used for computational efficiency. The iteration is terminated when the equality constraint is satisfied (in all the experiments,  $\frac{\|\mathbf{O} - \mathbf{A} - \mathbf{E}\|_F}{\|\mathbf{O}\|_F} < 1e^{-7}$ ). For more details, one can refer to the report of Lin *et al.* [19].

## 4. Experiment Results

We compare the performance of the proposed method against RPCA [3] with synthetic data sets and real world applications. We use the Matlab implementation of RPCA provided by Wright *et al.* [24]. In all the experiments including synthetic and real-world data, we use the default parameters recommended by Wright *et al.* [24] ( $\lambda =$

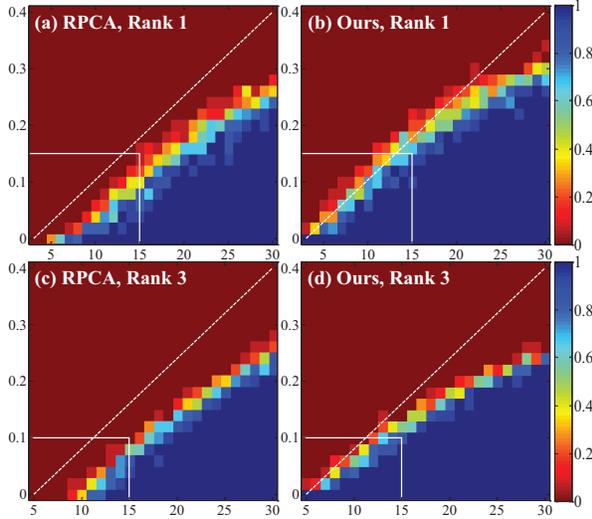


Figure 1. Success ratio for synthetic data with varying numbers of columns  $n$ . Comparison between RPCA and ours for the rank-1 (a,b) and rank-3 case (c,d). The Y-axis represents the corruption ratio  $r \in [0, 0.4]$ . The X-axis represents the column size  $n \in [3, 30]$  for the rank-1 case in (a,b),  $n \in [5, 30]$  for the rank-3 case in (c,d). The color magnitude represents success ratio  $[0, 1]$ .

$1/\sqrt{\max(m, n)}$ , where  $m$  and  $n$  are the row and column size of the matrix  $\mathbf{O}$ ). We applied our method for real world data sets in the contexts of HDR, photometric stereo and batch image alignment. Further results can be found in the supplementary materials.

#### 4.1. Synthetic Dataset

We compare our method with RPCA on synthetic data for success ratio and convergence behaviors. To synthesize a ground-truth low-rank matrix  $\mathbf{A}_{GT} \in \mathbb{R}^{m \times n}$  of rank  $N$ , we perform a linear combination of  $N$  arbitrary orthogonal basis vector. The weight vector used to span each column vector of  $\mathbf{A}_{GT}$  is randomly sampled from the uniform distribution  $U[0, 1]$ . To generate sparse outliers, we select  $m \times n \times r$  entries from  $\mathbf{A}_{GT}$ , where  $r$  denotes the corruption ratio. Larger  $r$  means more outlier entries. The selected entries are corrupted by random errors from  $U[0, 1]$ . We ran each of the tests listed in the following, over 50 trials, and report the average errors of overall trials.

**Comparison for Success Ratio** To verify the robustness of our method to outliers, we test how the performances of RPCA and the proposed method are affected by the number of observations, data dimension and the corruption ratio from  $\mathbf{E}$ . We examine whether the estimation is successful or not by counting the number of successes. If the recovered  $\hat{\mathbf{A}}$  satisfies  $\frac{\|\mathbf{A}_{GT} - \hat{\mathbf{A}}\|_F}{\|\mathbf{A}_{GT}\|_F} < 0.01$ , where  $\hat{\mathbf{A}}$  is the estimated result from RPCA or our method, we consider that  $\mathbf{A}$  and  $\mathbf{E}$  are successfully estimated. We compare the success ratio with varying the column size  $n$  (i.e. the number of obser-

Table 1. Accuracy comparison among SVP, Zheng *et al.* [27] and ours.

No. inputs ( $n$ )	Mean accuracy		
	SVP	[27]	Ours
5	0.1628	0.1456	<b>0.0686</b>
10	0.1267	0.0407	<b>0.0145</b>
15	0.1073	0.0052	<b>0.0042</b>
20	0.0985	0.0022	<b>0.0014</b>

variations), and varying the row size  $m$  (i.e. data dimension). The magnitude in Fig. 1 indicates the percentage of success. A larger blue area indicates a more robust performance of the algorithm.

We performed experiments where we fixed  $m = 10000$  and varied  $n$  and  $r$ . The comparisons between RPCA and our method with rank-1 and -3 constraint are shown in Fig. 1-(a-d). As  $n$  decreases (i.e. the number of observations decreases), the success ratio of RPCA decreases more rapidly than our method. When more observations are available (over  $n = 25$ ), RPCA and the proposed method show similar behavior. Additional results for the other rank cases and the varying column  $n$  cases can be found in the supplementary material.

We notice that, even though we use the same parameter  $\lambda$  in all the experiments varying the matrix size, rank and corruption ratio, our method shows a more broad success range compared to RPCA.

**Comparison with Other Approaches** We provide another comparisons with the singular value projection (SVP) as a baseline method and a low-rank matrix approximation approach by matrix factorization. General matrix factorization methods enforce the target rank  $N$  constraint of data matrix ( $M = UV$ ) by factorizing it into a product of rank  $N$  basis ( $U$ ) and coefficient ( $V$ ) as hard constraint. Out of matrix factorization based methods, we compare with Zheng *et al.* [27] as a state-of-the-art method. We use the default parameters for Zheng *et al.*

The measured accuracy  $\left(\frac{\|\mathbf{A}_{GT} - \hat{\mathbf{A}}\|_F}{\|\mathbf{A}_{GT}\|_F}\right)$  results are displayed in Table 1. We generate the synthetic matrix with the row size  $m = 10000$ , rank 3 and the corruption ratio  $r = 0.05$ . While our objective function is also non-convex, it is closer to the original convex function of RPCA with nuclear norm by definition of the partial sum of singular values which allow our method to converge to a better solution comparing to Zheng *et al.* In our analysis, we found that Zheng *et al.* often provides a local minimum due to the highly nonlinear bilinear form ( $UV$ ).

**Convergence Behavior** To examine the convergence behavior of both RPCA [19] and our method, we plot the evolution of the relative errors  $\frac{\|\mathbf{A}_{GT} - \hat{\mathbf{A}}\|_F}{\|\mathbf{A}_{GT}\|_F} + \frac{\|\mathbf{E}_{GT} - \hat{\mathbf{E}}\|_F}{\|\mathbf{E}_{GT}\|_F}$  and termination criteria  $\frac{\|\mathbf{O} - \mathbf{A} - \mathbf{E}\|_F}{\|\mathbf{O}\|_F}$  over the iterations in Fig. 2.

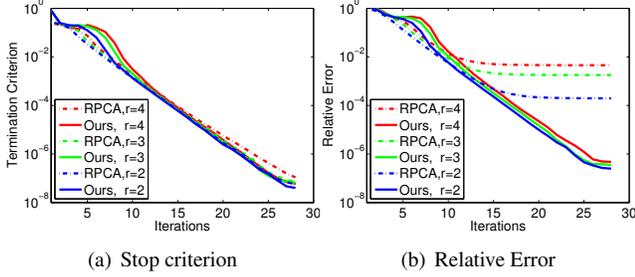


Figure 2. Convergence behavior of RPCA [19] and our method for the rank 2,3 and 4 cases.

We randomly generated matrices with  $m = 5000$  rows and  $n = 40$  columns for the rank 2, 3, 4 cases, and the average value over the trials is computed. Fig. 2-(b) also shows a better accuracy than RPCA as well as a gradual convergence shown in Fig. 2 under the same termination criterion, which shows that our algorithm converges. In Sec. 5, we will discuss the convergence in further details.

## 4.2. Real-world Applications

### 4.2.1 Background Modeling and High Dynamic Range (HDR) Imaging

We apply the proposed method for modeling a background scene and a ghost-free HDR composition. We assume that differently exposed images  $I_i$  are aligned and the camera response function (CRF) is calibrated (or linear). Then, the captured images can be represented as  $I_i = kR\Delta t_i$ , where  $R$  denotes the sensor irradiance,  $\Delta t_i$  is the exposure time for the  $i$ -th image, and  $k$  is a positive scalar. We construct the observed intensity matrix  $\mathbf{O} \in \mathbb{R}^{m \times n} = [\text{vec}(I_1) \cdots \text{vec}(I_n)]$  by stacking the vectorized input images. In this application,  $m$  and  $n$  are the number of pixels and the number of images respectively.

Ideally, the observed intensity  $I_i$  is linearly related to irradiance  $R$ , which means that  $\mathbf{O}$  is a rank-1 matrix. However, in practice, the rank of  $\mathbf{O}$  is higher than 1 due to moving objects, saturation or other artifacts (shown in Fig. 3). We model these artifacts and the background scene as a summation of a low-rank matrix (essentially a rank-1 matrix) and sparse outliers,  $\mathbf{O} = \mathbf{A} + \mathbf{E}$ . To compose an HDR image without ghost artifacts, we first estimate a Low Dynamic Range (LDR) background scene from the low-rank matrix  $\mathbf{A}$ , then the weighted sum is applied to the LDR images. In the background estimation step, Eq. (2) is used as objective function with a target rank  $N = 1$ . We apply RPCA and our method to each color channel independently.

We use the Arch. sequence from Gallo *et al.* [11]. It contains differently exposed images as HDR composition dataset, which are well aligned, but include some moving objects without overlaps and the brightness of images is varying. The estimated background as low-rank matrix and sparse outlier results from the RPCA and the proposed

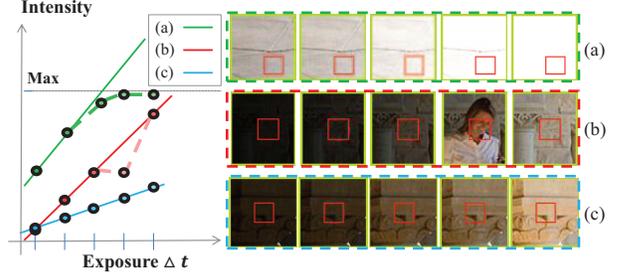


Figure 3. Illustration of the observed intensity values for (a) saturation region, (b) moving object, and (c) consistent cases. Solid lines denote ideal relationship between intensity and exposure, and dots and dotted lines denote the observed intensities.

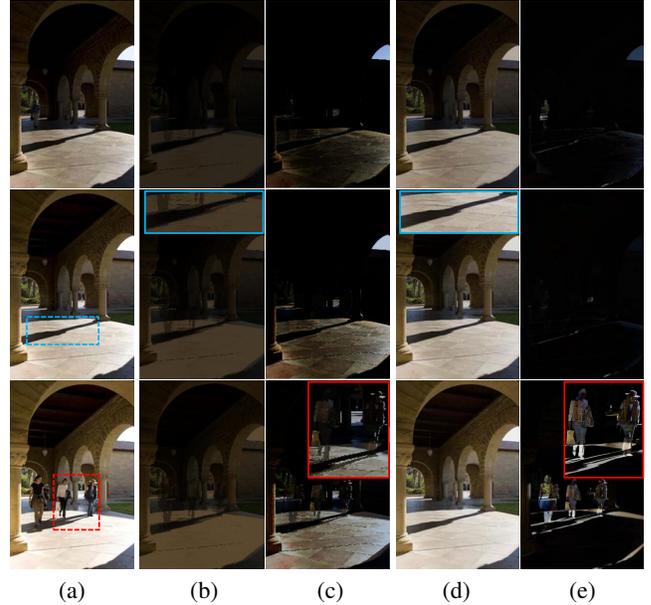


Figure 4. Comparison of the low-rank matrix and sparse error results between RPCA and ours. (a) Input multi-exposure images. Low-rank (b,d) and sparse error (c,e) results, respectively obtained by RPCA (b,c) and the proposed approach (d,e).

method are shown in Fig. 4. The example in Fig. 4-(a) consists of only 5 images which is very limited. Ideally, a decomposed low-rank matrix  $\mathbf{A} = [\text{vec}(A_1) \cdots \text{vec}(A_n)]$  consists of relative intensities of the background scene from which moving objects or saturation artifacts are removed (see Fig. 4-(b,d)). RPCA returns a low-rank matrix whose magnitude differs drastically from the input image, as shown in Fig. 4-(b). Moreover it yields a dense non-zero entries in  $\mathbf{E}$ , instead of being sparse, as shown in Fig. 4-(c). On the other hand, our proposed method shows well modeled background scene and successfully detects outlier regions, as respectively shown in Fig. 4-(d,e). For Fig. 4-(c,e), each color component (R,G,B) is set with  $(|\mathbf{E}_R|, |\mathbf{E}_G|, |\mathbf{E}_B|)$ , where  $\mathbf{E}_{\{R,G,B\}}$  denotes the sparse error matrix for each channel. More results and comparisons can be found in our supplemental materials.

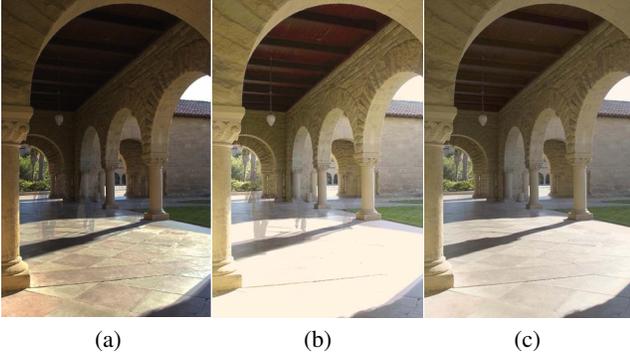


Figure 5. HDR composition results from Debevec *et al.* [7] (a), the RPCA (b) and the proposed method (c).

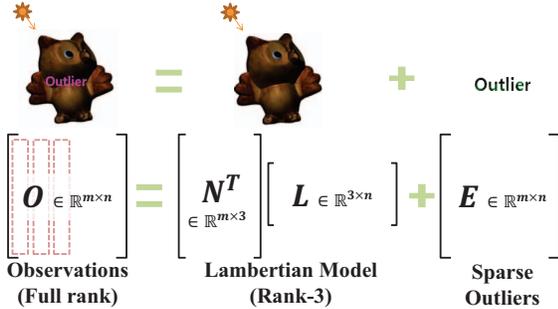


Figure 6. Photometric stereo illustration for the used model.

As mentioned above, we can combine the low-rank images to create a more informative HDR image by  $H(x) = \sum_{i=1}^n W_i(x) \cdot A_i(x) / \Delta t_i$ , where  $n$  represents the number of input exposures,  $W_i(x)$  is the saturation weight of the pixel located at  $x$  in the  $i^{\text{th}}$  exposure, and  $H(x)$  denotes the composited HDR image. Experiments showed that a uniform weight for  $W_i(x)$  is sufficient. For display,  $H(x)$  is normalized and tone-mapped by a simple gamma function. The final HDR results obtained by different methods are compared in Fig. 5. The standard method of Debevec *et al.* [7], in Fig. 5-(a), leads to ghost-artifact due to the presence of moving objects. In Fig. 5-(b), moving objects still remain in the HDR obtained by RPCA. In contrast, our HDR result correctly captures the background scene with high dynamic range, as shown in Fig. 5-(c).

#### 4.2.2 Outlier Rejection for Photometric Stereo

Traditional photometric stereo is based on Lambertian model,  $\mathbf{O} = [\text{vec}(O_1)] \cdots [\text{vec}(O_n)] = \mathbf{N}^T \mathbf{L}$ , where  $\mathbf{O} \in \mathbb{R}^{m \times n}$ ,  $\mathbf{N} \in \mathbb{R}^{3 \times m}$  and  $\mathbf{L} \in \mathbb{R}^{3 \times n}$  denote intensity, normal and light direction matrix, respectively, and where  $m$  and  $n$  are the number of pixels and images. Hayakawa *et al.* [13] show that the intensity matrix lies in a subspace of rank 3, as illustrated in Fig. 6. However, this constraint is hardly satisfied in real situations due to shadow from self-occlusion, saturation and some object materials which do not exactly follow the Lambertian diffuse model. Considering the rank-3 constraint from Hayakawa *et al.* [13], the

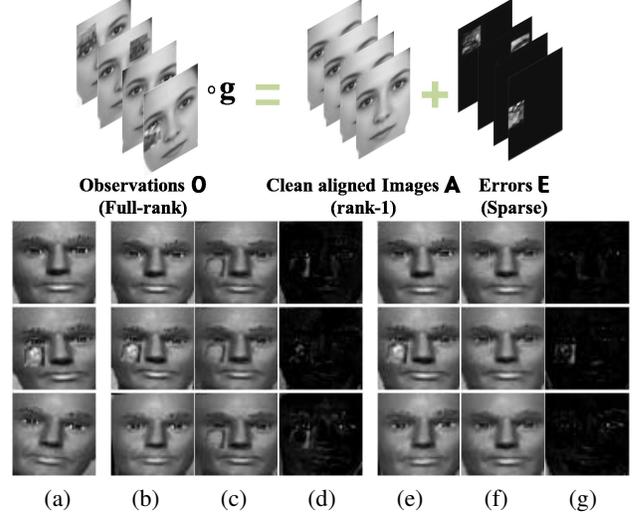


Figure 8. Top row: Illustration of the transformed low-rank structure of batch images. Bottom row: batch image alignment experiments. (a) Three input images. (b-d) The aligned, low-rank, sparse results from Peng *et al.* [20]. (e-g) The aligned, low-rank, sparse results from the proposed method.

artifacts mentioned above could be modeled as sparse outliers and we get a low-rank structure as  $\mathbf{O} = \mathbf{N}^T \mathbf{L} + \mathbf{E}$ .

The robust photometric stereo with outlier rejection can be formulated into a RPCA problem as suggested by Wu *et al.* [25]. They proposed the blind and the non-blind methods, which respectively assume that the locations of the corrupted region are known or unknown. We replace the formulation of the blind method by our partial sum objective function. It means that any prior information for outlier regions is not required. For this experiment, we set rank  $N = 3$  in Eq. (2).

The proposed method is compared with the standard least square (LS) method [22] and RPCA by Wu *et al.* [25]. The LS based photometric stereo estimates the normals by minimizing  $\|\mathbf{O} - \mathbf{N}^T \mathbf{L}\|_F^2$ . We use the *Bunny* dataset [14] generated using the Cook-Torrance reflectance model and consisting of 40 different lighting conditions. A representative data image is shown in the top of Fig. 7-(a). The average ratio of specular and shadow regions in *Bunny* are 8.4% and 24% respectively, which act as outliers. For quantitative evaluations shown in Table 2, we vary the number of images and add 5% of uniformly distributed corruption. Each value in Table 2 is averaged over 20 randomly selected test sets. Fig. 7 shows an example result for qualitative evaluation. Wu *et al.* [25] produce degenerated results (see top of Fig. 7-(c)), as the rank of resulting matrix is lower than 3 due to the lack of supports from the observations. When more input images are available, RPCA returns more satisfying results (see bottom of Fig. 7-(c)), but still the accuracy is lower than the LS method. In contrast, our method provides robust results for both limited observations and sufficient observations.

Table 2. Photometric stereo results of *Bunny* with 5% corruption ratio, additional specular and shadow.

No. Image	Mean error (in degrees)			Max error (in degrees)			Standard deviation		
	LS	[25]	Ours	LS	[25]	Ours	LS	[25]	Ours
5	8.53	27.88	<b>7.06</b>	159.72	130.77	<b>120.78</b>	14.48	16.45	<b>12.30</b>
8	9.03	13.34	<b>5.87</b>	142.45	139.07	<b>85.48</b>	11.24	10.96	<b>9.62</b>
10	9.24	11.14	<b>5.70</b>	148.05	110.12	<b>79.54</b>	9.91	9.77	<b>8.06</b>
12	8.96	9.95	<b>5.09</b>	130.04	80.21	<b>76.86</b>	9.17	9.02	<b>7.59</b>

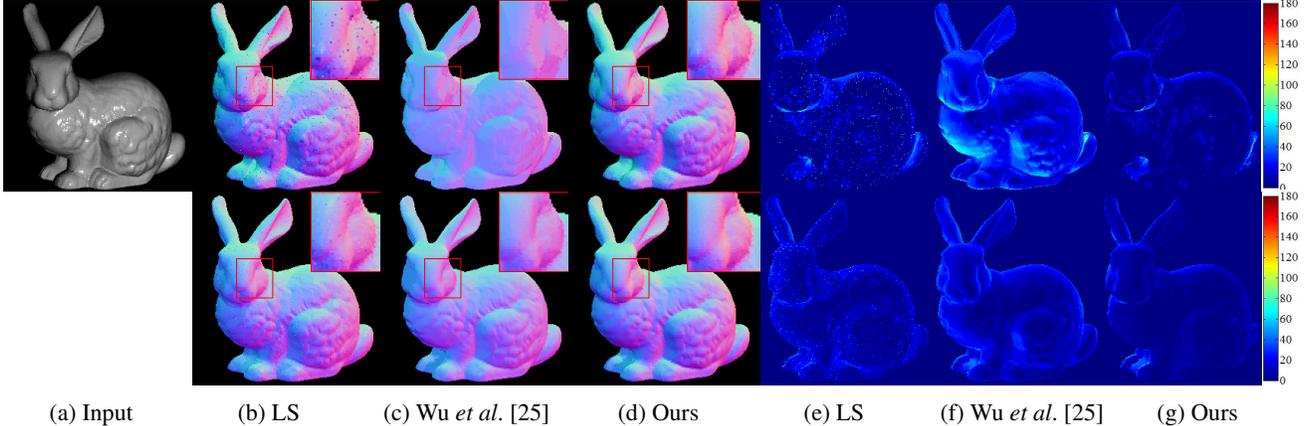


Figure 7. Photometric stereo results from 5 (top) and 12 (bottom) images of *Bunny* dataset with corruption. (b-d) Recovered surface normal by LS, Wu *et al.* [25] and ours. (e-g) Corresponding error maps for each algorithm.

### 4.2.3 Batch Image Alignment

Given several images of an object of interest (e.g. face), the batch image alignment task aims to align them to a fixed canonical template [1, 20]. The rank minimization approach has led to impressive results for robust alignment of linearly correlated images [20]. In addition to the applications of the previous sections, we also search for a transformation  $g_i$  for each image  $I_i$  to make the images linearly correlated (cf. the top row in Fig. 8). We note  $\mathbf{g}$  the set of transformations:  $\mathbf{g} = \{g_1, \dots, g_n\}$  where  $n$  is the number of images and write  $\mathbf{O} \circ \mathbf{g} = [\text{vec}(I_1 \circ g_1) \cdots \text{vec}(I_n \circ g_n)]$  as illustrated at the top of Fig. 8. Contrary to the formulation of Peng *et al.* [20], we consider the partial sum of singular values. This can be mathematically formulated as follows:

$$\arg \min_{\mathbf{A}, \mathbf{E}, \mathbf{g}} \sum_{i=N+1}^{\min(m,n)} \sigma_i(\mathbf{A}) + \lambda \|\mathbf{E}\|_1, \text{ s.t } \mathbf{O} \circ \mathbf{g} = \mathbf{A} + \mathbf{E}. \quad (11)$$

We applied our approach to a head dataset acquired under varying pose (cf. Fig 8-(a)) [20]. For linearly correlated noise-free batch images, the rank must be  $N = 1$ , when the transformations for exact image alignment are estimated. Our results of alignment, low-rank estimation and error sparsity are shown in Fig 8-(e,f,g). Compared to the results obtained by RPCA, our method can correctly detect the outliers (Fig 8-(c) v.s. Fig 8-(f)).

## 5. Discussion and Conclusion

In this paper, we revisited the rank minimization method in RPCA for low-level vision problems. When the target rank is known, we show that, by modifying the objective

function from the nuclear norm to the partial sum of singular values, we can achieve a better control of the target rank of the low-rank solution, even when the number of observations is limited. The beauty of our solution is that it can be easily utilized in existing algorithms, e.g. inexact ALM [19], and the efficient properties still hold. The generality of our approach and the effectiveness are supported through our encouraging experiments on both synthetic examples and several real-world applications which outperform the conventional nuclear norm objective function. An interesting direction of future work is the mathematical analysis of the properties, e.g. the necessary and the sufficient conditions [21] of our partial sum objective function compared to the nuclear norm solution. In the following, we discuss some open questions related to our paper.

**Sufficient number of samples versus minimum number of samples** In our experimental analysis, we found that our solution is more robust than the nuclear norm solution when facing a limited number of samples. Defining  $K$  as the minimum number of samples for processing, e.g. 2 images for HDR, 3 images for photometric stereo, our approach requires more than  $K$  samples for a robust model estimation and outlier rejection. We believe that the number of needed additional samples depends on the problem setting, e.g. the shape of feature space or the distribution of the samples.

**Convergence** The proof of convergence of the exact and inexact ALM with an alternating scheme has been established by [19, 9]. In contrast, to the best of our knowledge, the convergence property of inexact ALM alternating for non-convex (solving  $\mathbf{A}^*$ ) and convex (solving  $\mathbf{E}^*$ ) programming has not been answered yet. Despite the absence of theoret-

ical guarantee for the convergence, we conducted extensive experiments and found that our algorithm converges in all our experiments and the objective of the sub-problem for  $\mathbf{A}$  is always decreased by PSVT. Since the objective function of the partial sum of singular values is a non-convex function, the global optimal solution cannot be guaranteed. Nevertheless extensive experiments showed that our solution is very closed to the nuclear norm solution when the number of observations is more than sufficient and converges to a better solution when the number of observations is limited. This is again based on our empirical study through many experimental testings. Further mathematical analysis is left as future work.

**Target rank** While our formulation implicitly encourages a target rank constraint in the resulting matrix, this constraint is not hardly enforced. We discuss here two possible scenarios can produce the resulting matrix having a rank different from the target rank. A first scenario is when a very limited number of samples are observed. In such case, PSVT can produce a deficient rank lower than the target rank when the span of the observed samples is less than the target rank, but this case is a fundamental limitation of under-sampling rather than a conceptual limitation of our approach. Another scenario is due to too much noise (especially for Gaussian noise that does not follow the sparsity property) in the observed samples which results in large singular values in the residual ranks. In this case, a solution to satisfy the rank constraint is to increase  $\tau$  in Eq. (8). When  $\tau$  is equal to infinity, our solution is close to the result using singular value projection [15]. However, the projection method enforcing target rank could produce an over-fitting solution due to the mentioned noise effects.

## References

- [1] L. G. Brown. A survey of image registration techniques. In *ACM Computing Surveys*, 1992.
- [2] J.-F. Cai, E. J. Candès, and Z. Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*, 2010.
- [3] E. Candès, X. Li, Y. Ma, and J. Wright. Robust principal component analysis? *Journal of the ACM*, 2011.
- [4] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Rank-sparsity incoherence for matrix decomposition. *SIAM Journal on Optimization*, 2011.
- [5] F. De la Torre and M. Black. Robust principal component analysis for computer vision. In *ICCV*, 2001.
- [6] F. De la Torre and M. Black. A framework for robust subspace learning. *IJCV*, 2003.
- [7] P. Debevec and J. Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH*, 1997.
- [8] D. L. Donoho. High-dimensional data analysis: the curses and blessings of dimensionality. In *American Mathematical Society Conf. Math Challenges of the 21st Century*. 2000.
- [9] J. Eckstein and D. Bertsekas. On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Math. Programming*, 1992.
- [10] M. Fischler and R. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 1981.
- [11] O. Gallo, N. Gelfand, W. Chen, M. Tico, and K. Pulli. Artifact-free high dynamic range imaging. *ICCP*, 2009.
- [12] E. T. Hale, W. Yin, and Y. Zhang. Fixed-point continuation for  $l_1$ -minimization: Methodology and convergence. *SIAM Journal on Optimization*, 2008.
- [13] H. Hayakawa. Photometric stereo under a light source with arbitrary motion. *Journal of the Optical Society of America A*, 1994.
- [14] S. Ikehata, D. Wipf, Y. Matsushita, and K. Aizawa. Robust photometric stereo using sparse regression. In *CVPR*, 2012.
- [15] P. Jain, R. Meka, and I. Dhillon. Guaranteed rank minimization via singular value projection. In *NIPS*, 2010.
- [16] H. Ji, C. Liu, Z. Shen, and Y. Xu. Robust video denoising using low rank matrix completion. In *CVPR*, 2010.
- [17] I. T. Jolliffe. Principal Component Analysis. In *Springer-Verlag*. 1986.
- [18] Q. Ke and T. Kanade. Robust  $l_1$  norm factorization in the presence of outliers and missing data by alternative convex programming. In *CVPR*, 2005.
- [19] Z. Lin, M. Chen, and Y. Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. Technical Report UILU-ENG-09-2215, UIUC, 2009.
- [20] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma. RASL: Robust alignment by sparse and low-rank decomposition for linearly correlated images. *IEEE Trans. on PAMI*, 2012.
- [21] B. Recht, W. Xu, and B. Hassibi. Necessary and sufficient conditions for success of the nuclear norm heuristic for rank minimization. In *IEEE Conference on Decision and Control*, 2008.
- [22] R. Woodham. Photometric method for determining surface orientation from multiple images. In *Optical Engineering*, 1980.
- [23] J. Wright, A. Ganesh, S. Rao, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In *NIPS*, 2009.
- [24] J. Wright, A. Ganesh, Z. Zhou, K. Min, S. Rao, Z. Lin, Y. Peng, M. Chen, L. Wu, Y. Ma, E. Candès, and X. Li. Low-rank matrix recovery and completion via convex optimization. <http://perception.csl.illinois.edu/matrix-rank/home.html>.
- [25] L. Wu, A. Ganesh, B. Shi, Y. Matsushita, Y. Wang, and Y. Ma. Robust photometric stereo via low-rank matrix completion and recovery. In *ACCV*, 2010.
- [26] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma. TILT: Transform invariant low-rank textures. *IJCV*, 2012.
- [27] Y. Zheng, G. Liu, S. Sugimoto, S. Yan, and M. Okutomi. Practical low-rank matrix approximation under robust  $l_1$ -norm. In *CVPR*, 2012.