The Augmented Lagrange Multipliers Method for Matrix Completion from Corrupted Samplings with Application to Mixed Gaussian-Impulse Noise Removal

Fan Meng^{1,2}, Xiaomei Yang^{1*}, Chenghu Zhou¹

1 Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, **2** University of Chinese Academy of Sciences, Beijing, China

Abstract

This paper studies the problem of the restoration of images corrupted by mixed Gaussian-impulse noise. In recent years, low-rank matrix reconstruction has become a research hotspot in many scientific and engineering domains such as machine learning, image processing, computer vision and bioinformatics, which mainly involves the problem of matrix completion and robust principal component analysis, namely recovering a low-rank matrix from an incomplete but accurate sampling subset of its entries and from an observed data matrix with an unknown fraction of its entries being arbitrarily corrupted, respectively. Inspired by these ideas, we consider the problem of recovering a low-rank matrix from an incomplete sampling subset of its entries with an unknown fraction of the samplings contaminated by arbitrary errors, which is defined as the problem of matrix completion from corrupted samplings and modeled as a convex optimization problem that minimizes a combination of the nuclear norm and the l_1 -norm in this paper. Meanwhile, we put forward a novel and effective algorithm called augmented Lagrange multipliers to exactly solve the problem. For mixed Gaussian-impulse noise removal, we regard it as the problem of matrix completion from corrupted samplings, and restore the noisy image following an impulse-detecting procedure. Compared with some existing methods for mixed noise removal, the recovery quality performance of our method is dominant if images possess low-rank features such as geometrically regular textures and similar structured contents; especially when the density of impulse noise is relatively high and the variance of Gaussian noise is small, our method can outperform the traditional methods significantly not only in the simultaneous removal of Gaussian noise and impulse noise, and the restoration ability for a low-rank image matrix, but also in the preservation of textures and details in the image.

Citation: Meng F, Yang X, Zhou C (2014) The Augmented Lagrange Multipliers Method for Matrix Completion from Corrupted Samplings with Application to Mixed Gaussian-Impulse Noise Removal. PLoS ONE 9(9): e108125. doi:10.1371/journal.pone.0108125

Editor: Xi-Nian Zuo, Institute of Psychology, Chinese Academy of Sciences, China

Received: April 7, 2014; **Accepted:** August 25, 2014; **Published:** September 23, 2014

Copyright: © 2014 Meng et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability: The authors confirm that all data underlying the findings are fully available without restriction. In the field of image processing, some classical images such as Lena, Boats and Barbara, widely used by the researchers, are publicly available on the Internet. All the data in our manuscript can be found at http://www.imageprocessingplace.com/root_files_V3/image_databases.htm.

Funding: This work was partially supported by grants from the National High Technology Research and Development Program of China (2013AA122901, 2012AA121201) (<http://www.863.gov.cn>) and the National Natural Science Foundation of China (40971224) (<http://www.nsfc.gov.cn>). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* Email: yangxm@lreis.ac.cn

Introduction

Image denoising is highly demanded in the field of image processing, since noise is usually inevitable during the process of image acquisition and transmission, which significantly degrades the visual quality and makes subsequent high-level image analysis and understanding very difficult. There exist two types of most common and extensively studied noise: additive Gaussian noise and impulse noise. Additive Gaussian noise is usually generated during image acquisition and characterized by adding each image pixel a random value from the Gaussian distribution with zero mean and standard deviation σ , while impulse noise is very different in nature from Gaussian noise. Impulse noise can often be introduced in both image acquisition and transmission process by malfunctioning pixels in camera sensors, faulty memory locations in hardware, or transmission in a noisy channel, which mainly includes two models, namely salt-pepper noise and

random-valued impulse noise [1]. For the former model, each gray value is replaced with a given probability p by the extreme value $g_{\min}$ or $g_{\max}$, leaving remaining pixels unchanged, where $[g_{\min}, g_{\max}]$ denotes the dynamic range of an image and p determines the level of the salt-pepper noise; as for the latter case, the intensity values of contaminated pixels are taken place by random values identically and uniformly distributed in $[g_{\min}, g_{\max}]$. In this paper, we will focus on the first model of impulse noise. A fundamental target of image denoising is to effectively remove noises from an image while preserving image details and textures.

For Gaussian noise removal, the non-local means (NL-means) proposed by Buades A. et al. [2] is quite an efficient method in suppressing Gaussian noise while keeping details and structures in the image intact, which better exploits the redundancy of natural images and calculates the weighted average of all the pixels in the image based on similarity of neighborhoods, to achieve a satisfying filtering result. Inspired by nonlocal concept, many more nonlocal

methods such as BM3D [3] and K-SVD [4] have been introduced and get state-of-the-art denoising performance for Gaussian noise. In addition, some filters based on wavelet transform [5] and partial differential equations (PDE) [6,7] including nonlinear total variation (TV) are also powerful tools for it. The traditional approaches for impulse noise removal operate locally and nonlinearly on images. Among them, let us mention the classical median filter and its variants like adaptive median filter (AMF), progressive switching median filter (PSMF), and direction weighting median filter (DWMF). This kind of nonlinear filters can remove impulse noise effectively but cannot preserve image details well. In order to better preserve edge structures in images, variational approaches have been developed. In [8], a data-fidelity term of l_1 -norm was first introduced to achieve a significant improvement to remove outliers like impulse noise. Some authors [9,10,11] proposed two-phase schemes to better preserve image details: the main idea is to detect the location of noisy pixels corrupted by impulse noise using median-like filters, followed by some variational methods to estimate the gray values for those contaminated pixels.

However, most existing image denoising methods can only deal with a single type of noise, which violates the fact that most of the image noise we encounter in real world can normally be represented by a mixture of Gaussian and impulse noise. Effectively eliminating mixed noise is difficult due to the distinct characteristics of both types of degradation processes. It should be noted that the two-phase approaches [12–16] also show exuberant vitality for mixed noise removal, namely detecting or estimating the outliers before proceeding with the restoration phase. Garnett et al. [12] introduced a universal noise removal algorithm that first detects the impulse corrupted pixels, estimates its local statistics and incorporates them into the bilateral filter, resulting in the trilateral filter. In [13] somewhat similar ideas were used to establish a similarity principle which in turn supplies a simple mathematical justification for the nonlocal means filter in removing Gaussian noises. This is also the case of [15] where outliers are first detected by a median-like filter and then a K-SVD dictionary learning is performed on impulse free pixels to solve a l_1-l_0 minimization problem, and the case of [16] where a median-type filter is used to remove impulse noise first and then NL-means based method is applied to remove the remaining Gaussian noise. Liu et al. [18] proposed a general weighted l_2-l_0 norms energy minimization model to remove mixed noise, which was built upon maximum likelihood estimation framework and sparse representations over a trained dictionary. While in [19], weighted encoding with sparse nonlocal regularization was put forward for mixed noise removal, where there was not an explicit step of impulse pixel detection, and soft impulse pixel detection via weighted encoding was used to deal with impulse and Gaussian noise simultaneously instead. See [17,34,35] for more approaches to mixed noise removal. Although these denoising methods above were proposed specially for mixed Gaussian-impulse removal and indeed can alleviate the impact on image visual effect brought by mixed noises to some extent, most of them will erase details, and cannot better preserve regularly geometrical textures and fine structures.

In recent years, low-rank matrix reconstruction has become a research hotspot in many scientific and engineering domains, with myriad applications ranging from web search to machine learning to computer vision and image analysis, which mainly involves the problem of matrix completion (MC) [20,21,26,27] and robust principal component analysis (RPCA) [22–25,28], namely recovering a low-rank matrix from an incomplete but accurate sampling subset of its entries and from an observed data matrix with an

unknown fraction of its entries being arbitrarily corrupted, respectively. Luckily, there usually exist many regularly geometrical textures and fine structures in both natural image and Remote Sensing (RS) image, due to the self-similarity and redundancy of image, which makes the grayscale matrix of the image possess low-rank features. In light of this, we propose a novel denoising framework for better preserving details and low-rank features in images while removing mixed Gaussian-impulse noise, based on the theory of low-rank matrix reconstruction.

In this paper, we consider the problem of recovering a low-rank matrix from an incomplete sampling subset of its entries with an unknown fraction of the samplings contaminated by arbitrary errors, which is defined as the problem of matrix completion from corrupted samplings (MCCS) and modeled as a convex optimization problem that minimizes a combination of the nuclear norm and the l_1 -norm. To exactly solve the problem, we introduce an effective algorithm called augmented Lagrange multipliers (ALM). By detecting impulse noises in a noisy image and treating the impulse free entries of the image matrix as available samplings first, and then regarding the Gaussian noises underlying the samplings as arbitrary errors, we can exploit the proposed algorithm to remove mixed noises and to recover the image matrix with low-rank or approximately low-rank features. To sum up, the main contributions of the paper include modeling the problem of low-rank matrix recovery from incomplete and corrupted samplings of its entries, solving the convex optimization problem via the proposed ALM algorithm and creatively applying it to mixed Gaussian-impulse noise removal.

The rest of the paper is organized as follows. We start in Section 2 by giving some introductions about the theory of low-rank matrix recovery including the problem of MC and Robust PCA. We put forward the ALM algorithm for the problem of MCCS and describe some implementation details and parameter settings for our methods in Section 3, in which we also analyze its powerful performance for exact recovery of low-rank matrices with erasures and errors. In Section 4, we will present in details our full denoising schemes for the impulse detector and mixed Gaussian-impulse noise removal. Experiments and comparisons with recent approaches are demonstrated in Section 5. Finally, we conclude this paper in Section 6.

Low-Rank Matrix Reconstruction

Low-rank matrix plays a central role in large-scale data analysis and dimensionality reduction, since low-rank structure is usually used either to approximate a general matrix, or to recover corrupted or missing data [25]. From the mathematical viewpoint, these practical problems come down to the theory of low-rank matrix reconstruction, which mainly involves the problem of MC and Robust PCA at present.

1 Matrix completion

To describe the problem of MC, suppose to simplify that the unknown matrix $M \in \mathbb{R}^{n \times n}$ is square, and that one has m sampled entries $\{M_{ij} : (i,j) \in \Omega\}$ where Ω is a random subset of cardinality m . The recovery of complete matrix M from these incomplete samplings of its entries is the MC problem.

The issue is of course that this problem is extraordinarily ill-posed because, with fewer samples than entries, there are infinitely many completions. Therefore, it is apparently impossible to identify which of these candidate solutions is indeed the “correct” one without some additional information. In many instances, however, the matrix we wish to recover has low rank or approximately low rank, which may change the property of the

problem, and make the search for solutions feasible since the lowest-rank solution now tends to be the right one.

Candes and Recht [20] showed that matrix completion is not as ill-posed as once thought. Indeed, they proved that most low-rank matrices can be recovered exactly from most sets of sampled entries even though these sets have surprisingly small cardinality, and more importantly, they proved that this can be done by solving a simple convex optimization problem

$$\begin{aligned} \min \|X\|_* \\ \text{s.t. } X_{ij} = M_{ij}, (i,j) \in \Omega, \end{aligned} \quad (1)$$

provided that the number of samples obeys

$$m \geq Cn^{6/5}r \log n \quad (2)$$

for some positive numerical constant C , where r is the rank of matrix M . In (1), the notation $\|\bullet\|_*$ denotes the nuclear norm of a matrix, which is the sum of its singular values. The optimization problem (1) is convex and can be recast as a semidefinite program. In some sense, this is the tightest convex relaxation of the NP-hard rank minimization problem

$$\begin{aligned} \min \text{rank}(X) \\ \text{s.t. } X_{ij} = M_{ij}, (i,j) \in \Omega. \end{aligned} \quad (3)$$

Another interpretation of Candes and Recht's result is that under suitable conditions, the rank minimization program (3) and the convex program (1) are formally equivalent in the sense that they have exactly the same unique solution. The state-of-the-art algorithms to solve the MC problem include the accelerated proximal gradient (APG) approach and the singular value thresholding (SVT) method [21].

2 Robust principal component analysis

Principal component analysis (PCA), as a popular tool for high-dimensional data processing, analysis, compression and visualization, has wide applications in scientific and engineering fields. It assumes that the given high-dimensional data lie near a much lower-dimensional linear subspace. To large extent, the goal of PCA is to efficiently and accurately estimate this low-dimensional subspace.

Suppose that the given data are arranged as the columns of a large matrix $D \in \mathbb{R}^{m \times n}$. The mathematical model for estimating the low-dimensional subspace is to find a low-rank matrix $A^{m \times n}$, such that the discrepancy between A and D is minimized, leading to the following constrained optimization:

$$\min_{A,E} \|E\|_F, \text{ s.t. } \text{rank}(A) \leq r, D = A + E, \quad (4)$$

where $r < \min(m,n)$ is the target dimension of the subspace and $\|\bullet\|_F$ is the Frobenius norm, which corresponds to assuming that the data are corrupted by i.i.d. Gaussian noise. This problem can be conveniently solved by first computing the Singular Value Decomposition (SVD) of D , and then projecting the columns of D onto the subspace spanned by the r principal left singular vectors of D .

As PCA gives the optimal estimate when the corruption is caused by additive i.i.d. Gaussian noise, it works well in practice as long as the magnitude of noise is small. However, it breaks down

under large corruption, even if that corruption only affects very few of the observations. Therefore, it is necessary to study whether a low-rank matrix A can still be efficiently and accurately recovered from a corrupted data matrix $D = A + E$, where some entries of the additive error matrix E may be arbitrarily large.

Recently, Wright et al. [23] have shown that under rather broad conditions the answer is affirmative: provided the error matrix E is sufficiently sparse, one can exactly recover the low-rank matrix A from $D = A + E$ by solving the following convex optimization problem:

$$\min_{A,E} \|A\|_* + \lambda \|E\|_1, \text{ s.t. } D = A + E, \quad (5)$$

where $\|\bullet\|_1$ denotes the sum of the absolute values of matrix entries, and λ is a positive weighting parameter. Due to the ability to exactly recover underlying low-rank structure in the data, even in the presence of large errors or outliers, this optimization is referred to as Robust PCA. So far, RPCA has been successfully introduced into several applications such as background modeling and removing shadows and specularities from face images. Ganesh et al. [31] proposed two new algorithms for solving the problem (5), which are in some sense complementary to each other. The first one is an accelerated proximal gradient algorithm applied to the primal, which is a direct application of the FISTA framework introduced by [30], coupled with a fast continuation technique; the other one is a gradient-ascent approach applied to the dual of the problem (5). Recently, the augmented Lagrange multiplier method was introduced by Lin et al. for exact recovery of corrupted low-rank matrices, which shows rather promising performance when dealing with the problem (5).

Problem of MCCS and the ALM Algorithm

1 MCCS and its optimization model

Since the MC problem is closely connected to the RPCA problem, we may formulate the MC problem in the same way as RPCA

$$\min_A \|A\|_*, \text{ s.t. } A + E = D, \pi_\Omega(E) = 0, \quad (6)$$

where $\pi_\Omega : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^{m \times n}$ is a linear operator that keeps the entries in Ω unchanged and sets those outside Ω (i.e., in $\bar{\Omega}$) zeros. As E will compensate for the unknown entries of D , the unknown entries of D are simply set as zeros. Then the partial augmented Lagrangian function of (6) is

$$L(A, E, Y, \mu) = \|A\|_* + \langle Y, D - A - E \rangle + (\mu/2) \|D - A - E\|_F^2, \quad (7)$$

where μ is a positive scalar. For updating E , the constraint $\pi_\Omega(E) = 0$ should be enforced when minimizing $L(A, E, Y, \mu)$.

The problem of restoring matrix from corrupted entries (i.e., RPCA) is less studied than the simpler MC problem when the available entries are not corrupted. However, in many practical applications, there usually exists the case when missing entries and corrupted entries simultaneously concur in an observed data matrix, which urgently drives researchers to pay more and more attention to the problem of recovering low-rank matrix from its observed data matrix with erasures and errors.

For observed matrix $D = \pi_\Omega(A^* + E^*)$ where the ordered pair (A^*, E^*) denotes the true solutions to low-rank matrix and sparse error matrix respectively, suppose that one can only know the entries of D on subset Ω (the set of indices of known entries),

and that one aims to recover low-rank matrix A^* from the observed matrix D with erasures and errors. In this paper, we define the problem as matrix completion from corrupted samplings or robust matrix completion, which is the problem of restoring low-rank matrix from incomplete and corrupted samplings.

Similarly, as the entries of matrix E on set $\bar{\Omega}$ will compensate for the unknown entries of D and the entries of E on set Ω are sparse, we can split E into $E = E_{\Omega} + E_{\bar{\Omega}}$, where $E_{\Omega} = \pi_{\Omega}(E)$, $E_{\bar{\Omega}} = \pi_{\bar{\Omega}}(E)$ and E_{Ω} denotes sparse error components for the known entries. Now we can formulate the problem of MCCS as follows:

$$\begin{cases} \min_{A,E} \|A\|_* + \lambda \|E_{\Omega}\|_1 \\ s.t. A + E = D = \pi_{\Omega}(A^* + E^*) \end{cases} \quad (8)$$

To recover the low-rank matrix for the MCCS problem (8), we propose a novel and effective algorithm based on the ALM method, whose objective function is as follows:

$$\begin{aligned} L(A, E_{\Omega}, E_{\bar{\Omega}}, Y, \mu) &:= \|A\|_* + \lambda \|E_{\Omega}\|_1 + \langle Y, D - A - E \rangle + \frac{\mu}{2} \|D - A - E\|_F^2 \\ &= \|A\|_* + \lambda \|E_{\Omega}\|_1 + \langle Y, D - A - E_{\Omega} - E_{\bar{\Omega}} \rangle + \frac{\mu}{2} \|D - A - E_{\Omega} - E_{\bar{\Omega}}\|_F^2 \\ &= \|A\|_* + \lambda \|E_{\Omega}\|_1 + \langle Y_{\Omega}, D_{\Omega} - A_{\Omega} - E_{\Omega} \rangle + \frac{\mu}{2} \|D_{\Omega} - A_{\Omega} - E_{\Omega}\|_F^2 \\ &\quad + \langle Y_{\bar{\Omega}}, D_{\bar{\Omega}} - A_{\bar{\Omega}} - E_{\bar{\Omega}} \rangle + \frac{\mu}{2} \|D_{\bar{\Omega}} - A_{\bar{\Omega}} - E_{\bar{\Omega}}\|_F^2 \end{aligned} \quad (9)$$

In (9), as E_{Ω} and $E_{\bar{\Omega}}$ are complementary to each other, we can ignore one when solving another one, and finally combine the two to obtain E . Unlike the technique presented in [29], we separate E into two parts in the objective function, which leads to the splitting of Y, D, A afterwards, and aims to achieve subsection optimization.

2 The proposed ALM algorithm and parameter settings

The ALM algorithm for the Robust MC problem proposed in the paper is described in Algorithm 1. It is apparent that computing the full SVD for the Robust MC problem is unnecessary, so we only need those singular values that are larger than a particular thresholding and their corresponding singular vectors. Firstly, we should predict the dimension of principal singular space, and in the paper, we exploit the following prediction rule:

$$sv_{k+1} = \begin{cases} sv_k + 1, & \text{if } sv_k < sv_k \\ \min(sv_k + \text{round}(0.05d), d), & \text{if } sv_k = sv_k \end{cases} \quad (10)$$

where $d = \min(m, n)$, sv_k is the predicted dimension and sv_k denotes the number of singular values in the sv_k singular values that are larger than μ_k^{-1} , and $sv_0 = 5$. In the paper, we choose $\lambda = 1/\sqrt{\text{tr}(p * \max(m, n))}$, $p = |\Omega|/(mn)$, and set $\mu_0 = 0.5/\|D\|_2$. The following conditions are chosen as the stopping criterion:

$$\frac{\|D - A^k - E^k\|_F}{\|D\|_F} < \varepsilon_1 \text{ and } \frac{\mu_{k-1} \|E^k - E^{k-1}\|_F}{\|D\|_F} < \varepsilon_2 \quad (11)$$

For convenience, we introduce the following soft-thresholding (shrinkage) operator:

$$\mathcal{S}_{\varepsilon}[x] \doteq \begin{cases} x - \varepsilon, & \text{if } x > \varepsilon \\ x + \varepsilon, & \text{if } x < -\varepsilon \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

where $x \in \mathbb{R}$ and $\varepsilon > 0$. This operator can be extended to vectors and matrices by applying it element-wise. In Algorithm 1, $J(D) = \max(\|D\|_2, \lambda^{-1} \|D\|_{\infty})$, where $\|\bullet\|_{\infty}$ is the maximum absolute value of the matrix entries, and $\|\bullet\|_2$ denotes the maximum singular value of a matrix.

Algorithm 1(MC from Corrupted Samplings via the ALM Method)

Input : Observation matrix $D \in \mathbb{R}^{m \times n}$, Ω , λ .

$Y_{\Omega}^0 = D_{\Omega}/J(D)$, $Y_{\bar{\Omega}}^0 = 0$, $A^0 = E^0 = 0$, $\mu_0 > 0$, $\rho > 1$, $k = 0$.

While not converged do

//Solve $E^{k+1} = \arg \min_E L(A^k, E, Y^k, \mu_k)$.

$E_{\Omega}^{k+1} = \mathcal{S}_{\lambda \mu_k^{-1}}[D_{\Omega} - A_{\Omega}^k + \mu_k^{-1} Y_{\Omega}^k]$;

$E_{\bar{\Omega}}^{k+1} = D_{\bar{\Omega}} - A_{\bar{\Omega}}^k + \mu_k^{-1} Y_{\bar{\Omega}}^k$;

//Solve $A^{k+1} = \arg \min_A L(A, E^{k+1}, Y^k, \mu_k)$.

$(U, S, V) = \text{svd}(D - E^{k+1} + \mu_k^{-1} Y^k)$;

$A^{k+1} = U \mathcal{S}_{\mu_k^{-1}}[S] V^T$;

$Y^{k+1} = Y^k + \mu_k (D - A^{k+1} - E^{k+1})$;

Update μ_k to $\mu_{k+1} = \min(\rho \mu_k, \bar{\mu})$;

$k \leftarrow k + 1$.

End while

Output : (A^k, E^k) .

3 Performance analysis of the proposed algorithm

In this subsection, we first demonstrate the recovery accuracy of the proposed ALM algorithm by simulation experiments; and then for clarifying how the erasure rate (the fraction of the unknown entries in observed matrix D , i.e., $1 - d(\Omega)$, where $d(\Omega) = |\Omega|/(m * n)$) and the error probability (the fraction of randomly corrupted entries in the known entries available for restoration, i.e., $\|E_{\Omega}\|_0/|\Omega|$, where $\|\bullet\|_0$ denotes the sparsity of a matrix) affect the performance of our method, we carry out the phase transition analysis for the proposed algorithm. We generate the rank- r matrix A^* as a product LR^T , where L and R are independent $m \times r$ matrices whose elements are i.i.d. Gaussian random variables with zero mean and unit variance, and set a fraction of entries chosen randomly in A^* as zeros; and then, treat a fraction of non-zero entries also chosen randomly in A^* as corrupted entries after adding arbitrarily large errors E^* to them; finally, we get the observed data matrix $D^{m \times m}$ with erasures valued zeros.

We exploit different data matrices and various combinations (ρ_{Ω}, ρ_s) to do the experiments, where $\rho_{\Omega} = d(\Omega)$ and ρ_s stands for the error probability, and the reconstruction results are presented in Table 1. We observe that the proposed algorithm can accurately recover the low-rank matrices and sparse errors even when erasure rate reaches 30% and 20% of the available entries are corrupted.

Table 1. Recovery results of low-rank matrices with different (ρ_Ω, ρ_s) using the proposed ALM.

$\rho_r = \text{rank}(A^*)/m$	$\rho_\Omega = \Omega /m^2$	$\rho_s = \ E^*\ _0/ \Omega $	$\ A^* - \hat{A}\ _F / \ A^*\ _F$	$\text{rank}(\hat{A})$	$\ \hat{E}_\Omega\ _0$
20/500	225000 250000	0.1	1.20e-6	20	22500
		0.2	2.47e-6	20	45000
	200000 250000	0.1	1.38e-6	20	20000
		0.2	2.22e-6	20	39999
	175000 250000	0.1	1.27e-6	20	17500
		0.2	2.54e-6	20	35109
60/1000	9×10^5 10^6	0.1	6.20e-7	60	90000
		0.2	4e-4	60	181894
	8×10^5 10^6	0.1	6.82e-7	60	80000
		0.2	2e-4	60	163523
	7×10^5 10^6	0.1	4.84e-7	60	70233
		0.2	7.47e-4	60	154281

doi:10.1371/journal.pone.0108125.t001

As for the experiments of phase transition analysis, we fix $m = 512$, set $r = 8, 16, 64$ respectively, and vary ρ_Ω and ρ_s between 0 and 1. For each r and each (ρ_Ω, ρ_s) pair, we generate 5 pairs (A^*, E^*) to obtain 5 observed matrices as described above. We deem the recovery successful if the recovered $\hat{A}$ satisfies $\|\hat{A} - A^*\|_F / \|A^*\|_F < 0.01$ for all the 5 observed data matrices. The curves colored red, green and blue in Figure 1 define “phase transition” bounds for the case of $r = 8, 16, 64$, respectively. In Figure 1, the horizontal coordinate indicates ρ_Ω , while the vertical coordinate denotes ρ_s . At the points of these curves, all 5 trials were accomplished with success, whereas for the points above the curves, at least one attempt failed. It is assumed that the regions under the curves are “regions of success”. Our experiments show that for fixed m , the smaller rank r is, the larger area the region of

success will have. In fact, when other factors keep unchanged, the recovery performance of the algorithm is inversely proportional to $\text{rank}(A^*)/m$. In addition, we can also observe that erasure rate and error probability keep the relationship of interacting.

Our Denoising Scheme for Mixed Gaussian-Impulse Noise

Image denoising is a research hotspot in the area of image processing all the way. However, in real world, images are typically contaminated by more than one type of noise during image acquisition and transmission process. As a matter of fact, we often encounter the case where an image is corrupted by both Gaussian and impulse noise. Such mixed noise could occur when an image that has already been stained by Gaussian noise in the procedure of image acquisition with faulty equipment suffers impulsive corruption during its transmission over noisy channels successively.

As we know, there usually exist many regularly geometrical textures and similar structures in both natural image and Remote Sensing image, due to the self-similarity and redundancy of image, which makes the grayscale matrix of the image possess low-rank features. Matrix completion can accurately reconstruct low-rank information by exploiting reliable pixels after detecting the isolated noises, which is quite suitable for impulse noise (especially for salt-pepper noise) removal. Robust PCA can recover low-rank matrix from observed matrix with sparse errors, which may deal with the problem of Gaussian noise removal to some extent, since for Gaussian noise, corruptions with large magnitude are rarely distributed while most of the corruptions are located near zero. Luckily, the proposed ALM algorithm for the Robust MC problem, also supplies a powerful technique for removing mixed Gaussian-impulse noise from images with low-rank features.

In this paper, we adopt the two-phase scheme for mixed noise removal: By detecting impulse noises in a noisy image and treating the impulse free entries of the image matrix as available samplings first, and then regarding the Gaussian noises underlying the samplings as arbitrary errors, we can exploit the proposed algorithm to remove mixed noises and to recover the grayscale matrix of the image with low-rank or approximately low-rank features. It should be noted that, we cannot simply use MC for

Figure 1. Phase transition with regard to (ρ_Ω, ρ_s) of the proposed ALM algorithm. The curves colored red, green and blue define “phase transition” bounds for the case of $r = 8, 16, 64$, respectively.

doi:10.1371/journal.pone.0108125.g001

Figure 2. Comparative results on Mondrian image with $p=0.1, \sigma=15/255$. The PSNR (dB) and SSIM for (b) ~ (f) are (13.681, 0.883), (27.704, 0.981), (23.542, 0.966), (23.376, 0.966), (**30.158, 0.985**), respectively. Figure (a): Original Image; (b): Noisy Image; (c): Trilateral Filter [12]; (d): Xiao et al. [15]; (e): Xiong and Yin's [16]; (f): Proposed ALM.
doi:10.1371/journal.pone.0108125.g002

removing impulse noises after detecting them, since the remaining Gaussian noises no longer make the image possess low-rank features; in addition, neither can we directly employ Robust PCA to restore the noisy image at first, because impulse noises destroy the sparsity of large errors, which may greatly decrease the accuracy of the restoration.

The success of two-phase approaches for noise removal relies on the accurate detection of impulse noise. Many impulse noise detectors are proposed in the literatures, e.g. AMF is used to detect salt-pepper noise, while adaptive center-weighted median filter (ACWMF) and rank-ordered logarithmic difference (ROLD) [32] are utilized to detect random-valued impulse noise. In this paper, study on detectors for random-valued impulse noise is beyond the scope of the topic.

For salt-pepper noise, we employ the following strategy to detect it after considering its features taken on in the image: (1) Grayscale thresholding. Specifically, set a fixed thresholding α , and regard those pixels whose gray-levels are distributed in the interval $[255-\alpha, 255]$ or $[0, \alpha]$ as candidate noises. (2) Median detecting. For current candidate noise, we consider the absolute difference between its gray-level and the median gray-level of its neighborhood pixels. If the absolute difference is larger than another given thresholding β , we then treat current candidate noise as true noise. Empirically, when the density of salt-pepper noise namely noise level is high, the window size of the neighborhood for median detecting should be large.

Experiments and Discussion

This section is devoted to the experimental analysis and discussion of the denoising scheme introduced in the previous

section. We compare our ALM algorithm with some recently proposed approaches [12,15,16] for dealing with the mixture of impulse and Gaussian noise. Comparison results are presented both under the form of statistical index tables and that of visual effects. To evaluate the quality of the restoration results, peak signal to noise ratio (PSNR) and structural similarity (SSIM) are employed for objective evaluation and for subjective visual evaluation, respectively.

Given an image $x \in [0, 255]^{m \times n}$, the PSNR of the restored image $\hat{x}$ is defined as follows:

$$PSNR(x, \hat{x}) = 10 \log_{10} \frac{255^2}{MSE}, \quad MSE = \|x - \hat{x}\|_F^2 / (m \times n). \quad (13)$$

The larger the value of PSNR is, the better the quality of restoration will be. We can formulate the SSIM between original image and the recovered image as

$$SSIM(x, \hat{x}) = \frac{(2\mu_x \mu_{\hat{x}} + C_1)(2\sigma_{x\hat{x}} + C_2)}{(\mu_x^2 + \mu_{\hat{x}}^2 + C_1)(\sigma_x^2 + \sigma_{\hat{x}}^2 + C_2)} \quad (14)$$

(see [33] for more details). The dynamic range of SSIM is $[0, 1]$, which means a better recovery performance with the SSIM closer to the value 1.

As the proposed algorithm is suitable for restoring images with low-rank information, we first choose the classical Mondrian and Barbara image in the field of image processing to do the experiments. In all experiments, parameters of each method have been tuned to yield the best results. On image Mondrian, we

Table 2. Comparisons of statistical indices under different (p, σ) on Barbara Image.

(p, σ)	Statistical Index	Noisy Image	ROAD-Trilateral [12]	Xiao et al. [15]	Xiong and Yin's [16]	Proposed ALM
$(0.1, \frac{5}{255})$	PSNR(dB)	15.132	21.528	27.027	26.472	27.163
	SSIM	0.956	0.962	0.983	0.984	0.986
$(0.2, \frac{5}{255})$	PSNR(dB)	12.184	19.712	26.325	25.713	26.386
	SSIM	0.913	0.949	0.980	0.981	0.982
$(0.3, \frac{5}{255})$	PSNR(dB)	10.407	17.652	25.217	24.688	25.287
	SSIM	0.870	0.933	0.975	0.976	0.977
$(0.4, \frac{5}{255})$	PSNR(dB)	9.179	14.854	24.017	23.438	24.237
	SSIM	0.827	0.906	0.969	0.970	0.971
$(0.5, \frac{5}{255})$	PSNR(dB)	8.223	12.096	22.977	22.289	23.315
	SSIM	0.785	0.857	0.964	0.962	0.965
$(0.1, \frac{15}{255})$	PSNR(dB)	14.910	20.726	26.041	25.892	25.385
	SSIM	0.938	0.954	0.978	0.978	0.975
$(0.2, \frac{15}{255})$	PSNR(dB)	12.044	19.432	25.542	25.280	25.067
	SSIM	0.896	0.945	0.975	0.976	0.973
$(0.3, \frac{15}{255})$	PSNR(dB)	10.349	16.838	24.574	24.256	24.182
	SSIM	0.855	0.928	0.971	0.971	0.970
$(0.4, \frac{15}{255})$	PSNR(dB)	9.131	13.945	23.555	23.076	23.307
	SSIM	0.814	0.893	0.965	0.965	0.965
$(0.5, \frac{15}{255})$	PSNR(dB)	8.206	11.604	22.492	21.955	22.684
	SSIM	0.776	0.846	0.961	0.958	0.962

doi:10.1371/journal.pone.0108125.t002

Figure 3. Comparison of visual effect on Barbara image when $p=0.3, \sigma=5/255$. Figure (a): Original Image; (b): Noisy Image; (c): Trilateral Filter [12]; (d): Xiao et al. [15]; (e): Xiong and Yin's [16]; (f): Proposed ALM.

doi:10.1371/journal.pone.0108125.g003

Figure 4. Comparison of visual effect on RS image when $p=0.5, \sigma=5/255$. Figure (a): Original Image; (b): Noisy Image; (c): Trilateral Filter [12]; (d): Xiao et al. [15]; (e): Xiong and Yin's [16]; (f): Proposed ALM.
doi:10.1371/journal.pone.0108125.g004

Table 3. Comparisons of statistical indices under different (p, σ) on RS Image.

(p, σ)	Statistical Index	Noisy Image	ROAD-Trilateral [12]	Xiao et al. [15]	Xiong and Yin's [16]	Proposed ALM
$(0.1, \frac{5}{255})$	PSNR(dB)	15.620	33.976	34.459	36.079	36.246
	SSIM	0.954	0.994	0.994	0.996	0.996
$(0.2, \frac{5}{255})$	PSNR(dB)	12.509	30.145	33.993	34.918	35.170
	SSIM	0.910	0.989	0.993	0.995	0.995
$(0.3, \frac{5}{255})$	PSNR(dB)	10.691	27.797	32.550	33.037	33.691
	SSIM	0.861	0.983	0.992	0.992	0.993
$(0.4, \frac{5}{255})$	PSNR(dB)	9.559	24.160	30.941	31.603	32.317
	SSIM	0.828	0.976	0.988	0.990	0.991
$(0.5, \frac{5}{255})$	PSNR(dB)	8.590	19.342	29.151	29.097	30.892
	SSIM	0.787	0.948	0.984	0.985	0.988
$(0.1, \frac{15}{255})$	PSNR(dB)	15.144	29.618	29.893	31.581	28.385
	SSIM	0.932	0.985	0.986	0.989	0.981
$(0.2, \frac{15}{255})$	PSNR(dB)	12.393	27.657	29.809	30.978	27.943
	SSIM	0.892	0.981	0.985	0.988	0.979
$(0.3, \frac{15}{255})$	PSNR(dB)	10.739	26.382	29.277	30.168	27.481
	SSIM	0.851	0.976	0.984	0.987	0.977
$(0.4, \frac{15}{255})$	PSNR(dB)	9.563	23.601	28.385	29.103	26.995
	SSIM	0.829	0.969	0.980	0.985	0.975
$(0.5, \frac{15}{255})$	PSNR(dB)	8.543	16.873	26.993	27.606	26.617
	SSIM	0.774	0.937	0.975	0.977	0.974

doi:10.1371/journal.pone.0108125.t003

Figure 5. Comparative results on Lena image with $p=0.3, \sigma=5/255$. The PSNR (dB) and SSIM for (b) ~ (d) are (10.669, 0.887), (27.237, 0.987), **(28.070, 0.989)**, respectively. Figure (a): Original Lena; (b): Noisy Lena; (c): Trilateral Filter; (d): Proposed ALM.
doi:10.1371/journal.pone.0108125.g005

successively add Gaussian noise with $\sigma=15/255$ and salt-pepper noise with $p=0.1$. The visual comparison of the four approaches is presented in Figure 2, from where we will see that our method can simultaneously remove Gaussian noise and salt-pepper noise in the restoration phase, and well reconstruct the original image with low-rank structures; above all, it can better preserve details and edges in the image. By contrast, the other three approaches either cannot thoroughly remove the mixed noises or will destroy fine structures. The values of PSNR and SSIM also demonstrate that the ALM-based approach outperforms other methods significantly not only in objective index, but also in subjective visual effect.

In the following experiment, image Barbara abundant in geometrical textures, is corrupted by Gaussian noise with zero mean and different standard deviations $\sigma=5/255, 15/255$ first, and then we add salt-pepper noise with different levels $p=0.1, 0.2, 0.3, 0.4, 0.5$ on the image. Comparisons of statistical indices under different combinations of (p, σ) are provided in Table 2, and the comparison of visual effect when $p=0.3, \sigma=5/255$ is also shown in Figure 3. From Figure 3, we also observe that our method can preserve regularly geometrical textures perfectly while removing the mixed Gaussian-impulse noises from the image. However, other approaches will more or less blur the textured information consisting in the image.

Meanwhile, we can see that the differences of recovery performance between the proposed ALM algorithm and the other three methods will be more remarkable when the level of salt-pepper noise is relatively high while the variance of Gaussian noise is small.

To better demonstrate the excellent performance of our method, we do a further comparison experiment of the four approaches in removing mixed Gaussian-impulse noise from Remote Sensing image with low-rank features. As in the previous experiment, we add Gaussian and salt-pepper noise with different values of p and σ on RS image. The experimental results are presented in Figure 4 and Table 3. Again, it shows that our method can better remove the mixed noise and restore the image especially when the level of impulse noise is relatively high while Gaussian noise is small.

The above experiments mainly involve images with significant low-rank features such as similar structures and regular textures, which violates the fact that most of images in the real world cannot possess globally low-rank features. However, most local parts in the image will meet low-rank or approximately low-rank condition, due to the self-similarity and spatial correlation of images. Consequently, we can utilize the proposed algorithm for image de-noising via block processing-based technique. The

Figure 6. Comparative results on Boat image with $p=0.3, \sigma=5/255$. The PSNR (dB) and SSIM for (b) ~ (d) are (10.696, 0.909), (23.737, 0.973), **(27.192, 0.985)**, respectively. Figure (a): Original Boat; (b): Noisy Boat; (c): Trilateral Filter; (d): Proposed ALM. doi:10.1371/journal.pone.0108125.g006

following experiments demonstrate the results of removing mixed Gaussian-impulse noise from Lena and Boat images, compared to trilateral filter method. Both of the noisy images are of size 512×512 pixels, either divided into many image blocks of size 16×16 pixels. The de-noising results for Lena and Boat are presented in Figure 5 and Figure 6 respectively, which also show the better performance of our method over trilateral filter. We can see that the details of hairs are not preserved well in Figure 5c, while in Figure 6c, some structures of the masts on the boats are lost.

Conclusion

In this paper, regarding the problem of low-rank matrix recovery from incomplete and corrupted samplings of its entries namely the Robust MC problem, we construct a mathematical model based on convex optimization, and put forward a novel and effective ALM algorithm, to solve this kind of optimization

problem which can be considered as the extension of the Robust PCA and the MC problem. Experiments on performance analysis and mixed Gaussian-impulse noise removal demonstrate the reliability and practicability of the proposed algorithm, which also show that our method can well preserve details and textures and keep the consistency of structures, while simultaneously removing mixed noises from images with low-rank features. By virtue of the novelty and powerful advantages, the approach will bring promising application value in the fields of data mining, image processing, machine learning, RS information processing and so forth.

Author Contributions

Conceived and designed the experiments: FM CZ XY. Performed the experiments: FM. Analyzed the data: FM XY. Contributed reagents/materials/analysis tools: FM XY. Wrote the paper: FM.

References

1. Pok G, Liu JC, Nair AS (2003) Selective removal of impulse noise based on homogeneity level information. *IEEE Transactions on Image Processing* 12(1): 85–92.
2. Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. *IEEE International Conference on Computer Vision and Pattern Recognition*: 60–65.

3. Dabov K, Foi A, Katkovnik V, Egiazarian K (2007) Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing* 16(8): 2080–2095.
4. Elad M, Aharon M (2006) Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing* 15(12): 3736–3745.
5. Xu J, Osher S (2007) Iterative regularization and nonlinear inverse scale space applied to wavelet-based denoising. *IEEE Transactions on Image Processing* 16(2): 534–544.
6. You YL, Kaveh M (2000) Fourth-order partial differential equations for noise removal. *IEEE Transactions on Image Processing* 9(10): 1723–1729.
7. Rudin L, Osher S, Fatemi E (1992) Nonlinear total variation based noise removal algorithms. *Physica D* 60: 259–268.
8. Nikolova M (2004) A variational approach to remove outliers and impulse noise. *Journal of Mathematical Imaging and Vision* 20(1–2): 99–120.
9. Chan R, Dong Y, Hintermüller M (2010) An efficient two-phase L1-TV method for restoring blurred images with impulse noise. *IEEE Transactions on Image Processing* 19(7): 1731–1739.
10. Cai JF, Chan R, Nikolova M (2010) Fast two-phase image deblurring under impulse noise. *Journal of Mathematical Imaging and Vision* 36(1): 46–53.
11. Chan R, Hu C, Nikolova M (2004) An iterative procedure for removing random-valued impulse noise. *IEEE Signal Processing Letters* 11(12): 921–924.
12. Garnett R, Huegerich T, Chui C, He W (2005) A universal noise removal algorithm with an impulse detector. *IEEE Transactions on Image Processing* 14(11): 1747–1754.
13. Li B, Liu Q, Xu J, Luo X (2011) A new method for removing mixed noises. *Science China Information Sciences* 54(1): 51–59.
14. Cai JF, Chan R, Nikolova M (2008) Two-phase approach for deblurring images corrupted by impulse plus Gaussian noise. *Inverse Problems and Imaging* 2(2): 187–204.
15. Xiao Y, Zeng T, Yu J, Ng MK (2011) Restoration of images corrupted by mixed Gaussian-impulse noise via l_1-l_0 minimization. *Pattern Recognition* 44(8): 1708–1720.
16. Xiong B, Yin Z (2012) A universal denoising framework with a new impulse detector and nonlocal means. *IEEE Transactions on Image Processing* 21(4): 1663–1675.
17. Delon J, Desolneux A (2013) A patch-based approach for removing impulse or mixed Gaussian-impulse noise. *SIAM Journal on Imaging Sciences* 6(2): 1140–1174.
18. Liu J, Tai XC, Huang H, Huan Z (2013) A weighted dictionary learning model for denoising images corrupted by mixed noise. *IEEE Transactions on Image Processing* 22(3): 1108–1120.
19. Jiang J, Zhang L, Yang J (2014) Mixed noise removal by weighted encoding with sparse nonlocal regularization. *IEEE Transactions on Image Processing* 23(6): 2651–2662.
20. Candes EJ, Recht B (2009) Exact matrix completion via convex optimization. *Foundations of Computational Mathematics* 9(6): 717–772.
21. Cai JF, Candes EJ, Shen Z (2010) A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization* 20(4): 1956–1982.
22. Candes EJ, Li X, Ma Y, Wright J (2011) Robust principal component analysis? *Journal of the ACM*. doi: 10.1145/1970392.1970395.
23. Wright J, Ganesh A, Rao S, Ma Y (2009) Robust principal component analysis: exact recovery of corrupted low-rank matrices via convex optimization. *Proceedings of Advances in Neural Information Processing Systems*: 2080–2088.
24. Xu H, Caramanis C, Sanghavi S (2012) Robust PCA via outlier pursuit. *IEEE Transactions on Information Theory* 58(5): 3047–3064.
25. Chen Y, Jalali A, Sanghavi S, Caramanis C (2013) Low-rank matrix recovery from errors and erasures. *IEEE Transactions on Information Theory* 59(7): 4324–4337.
26. Liu Z, Vandenberghe L (2009) Interior-point method for nuclear norm approximation with application to system identification. *SIAM Journal on Matrix Analysis and Applications* 31(3): 1235–1256.
27. Ji H, Liu C, Shen Z, Xu Y (2010) Robust video denoising using low rank matrix completion. *IEEE International Conference on Computer Vision and Pattern Recognition*: 1791–1798.
28. Tao M, Yuan X (2011) Recovering low-rank and sparse components of matrices from incomplete and noisy observations. *SIAM Journal on Optimization* 21(1): 57–81.
29. Liang X, Ren X, Zhang Z, Ma Y (2012) Repairing sparse low-rank texture. *Computer Vision-ECCV*: 482–495.
30. Beck A, Teboulle M (2009) A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences* 2(1): 183–202.
31. Ganesh A, Lin Z, Wright J, Wu L, Chen M, et al. (2009) Fast algorithms for recovering a corrupted low-rank matrix. 3rd IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP): 213–216.
32. Dong Y, Chan RH, Xu S (2007) A detection statistic for random-valued impulse noise. *IEEE Transactions on Image Processing* 16(4): 1112–1120.
33. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP (2004) Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing* 13(4): 600–612.
34. Li YR, Shen L, Dai DQ, Suter BW (2011) Framelet algorithms for de-blurring images corrupted by impulse plus Gaussian noise. *IEEE Transactions on Image Processing* 20(7): 1822–1837.
35. Lopez-Rubio E (2010) Restoration of images corrupted by Gaussian and uniform impulsive noise. *Pattern Recognition* 43(5): 1835–1846.