

RESEARCH ARTICLE

Compressive Sensing via Nonlocal Smoothed Rank Function

Ya-Ru Fan¹, Ting-Zhu Huang^{1*}, Jun Liu², Xi-Le Zhao¹

1 School of Mathematical Sciences/Research Center for Image and Vision Computing, University of Electronic Science and Technology, Chengdu, Sichuan, 611731, P. R. China, **2** School of Mathematics and Statistics, Northeast Normal University, Changchun, Jilin, 130024, P. R. China

* tingzhuhuang@126.com

Abstract

Compressive sensing (CS) theory asserts that we can reconstruct signals and images with only a small number of samples or measurements. Recent works exploiting the nonlocal similarity have led to better results in various CS studies. To better exploit the nonlocal similarity, in this paper, we propose a non-convex smoothed rank function based model for CS image reconstruction. We also propose an efficient alternating minimization method to solve the proposed model, which reduces a difficult and coupled problem to two tractable subproblems. Experimental results have shown that the proposed method performs better than several existing state-of-the-art CS methods for image reconstruction.



OPEN ACCESS

Citation: Fan Y-R, Huang T-Z, Liu J, Zhao X-L (2016) Compressive Sensing via Nonlocal Smoothed Rank Function. PLoS ONE 11(9): e0162041. doi:10.1371/journal.pone.0162041

Editor: Pew-Thian Yap, University of North Carolina at Chapel Hill, UNITED STATES

Received: May 31, 2016

Accepted: August 16, 2016

Published: September 1, 2016

Copyright: © 2016 Fan et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper.

Funding: This work is supported by the 973 Program (2013CB329404), National Science Foundation of China (NSFC) (61370147, 61402082), the Fundamental Research Funds for the Central Universities (ZYGX2013Z005, ZYGX2013J106).

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

Introduction

Compressive sensing (CS) [1, 2] allows us to reconstruct high dimensional data with only a small number of samples or measurements, and captures only useful information and has the potential of significantly improving the energy efficiency of sensors in the real-world applications. The key idea behind CS is that the majority of real-world signals including images and videos, can be sparsely represented by given some appropriate basis. Due to the positive theoretical and experimental results, many CS-based imaging methods have been proposed and applied to the various areas such as magnetic resonance imaging [3, 4] in medicine, compressed spectral and hyperspectral imaging [5, 6] in industry, neural network [7] in biotechnology.

CS also makes it possible to well restore corrupted signals at a fast speed and the small memory cost. Conventional CS recovery uses ℓ_1 norm to characterize the sparsity of a signal, and the resulting convex optimization problems are tractable. Although there are several methods can be used to efficiently solve the ℓ_1 regularization based model for signal recovery [8–10], they only achieve suboptimal recovery performance due to their relaxation of the ℓ_0 norm based sparse optimization [11]. More recently, structured or group sparsity based methods [12–14] and nonlocal sparsity based methods [15, 16] have provided better results for CS recovery. Intuitively, the structured or group sparsity can reduce the degrees of freedom in the

solution, and the nonlocal sparsity explicitly exploits self-similarities of the signal, thereby obtaining more accurate recovery performance than the common sparsity.

In CS studies, a number of works have suggested that non-convex optimization based approach often yields better results than convex ones though costing higher computational complexity. Therefore, we propose a CS recovery model by considering a non-convex smoothed function to approximate the rank, denoted as smoothed rank function (SRF), as a low-rank regularization. The emphasis behind the proposed model is the utilization of nonlocal sparsity by image patch grouping. Concretely, we group a set of similar image patches to form a matrix \mathbf{X}_i for i -th exemplar image patch, which is extracted from the test image. Then the matrix \mathbf{X}_i is low-rank since similar patches have similar structure. To solve the low-rank optimization problem, we minimize the SRF function to approximate the rank minimization problem, and the resulting problem is a non-convex optimization problem. In order to avoid getting trapped in local solutions, we initialize a rough approximation of the rank, and gradually improve the approximation as the iteration proceeds.

The basic idea of using nonlocal sparsity for image patches has already been used in [16–18] with the very impressive results. There have been an abundant research literatures using the patch-based low-rank as plenty with nonlocal sparsity [19–22]. In [20, 21], the low-rank problem is solved by minimizing the nuclear norm of the low-rank matrix, which leads to a convex minimization problem with many efficient methods available. In [22], a more accurate approximation for rank is proposed by exploiting the logdet function, which can be derived as a weighted nuclear norm. Specifically, compared with previous surrogates for the rank such as nuclear norm and weighted nuclear norm, our surrogate is differentiable and can approximate the rank adaptively.

The outline of this paper is organized as follows. Firstly, we briefly review the background of CS. Secondly, the SRF function is introduced and the model based on SRF for CS image reconstruction is proposed. Thirdly, the optimization algorithm is presented to solve the proposed model. To demonstrate the effectiveness of the proposed method, we show some numerical results on several test images in the following section. In the end, the conclusion and future work are given.

Background

The CS recovery problem aims to find the sparsest solution $\mathbf{x} \in \mathbb{R}^N$ from the underdetermined linear system $\mathbf{y} = \Phi \mathbf{x}$, where $\mathbf{y} \in \mathbb{R}^M$ is the measurements and $\Phi \in \mathbb{R}^{M \times N}$, $M < N$ is the measurement matrix. It can be formulated as follows:

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad s.t. \quad \mathbf{y} = \Phi \mathbf{x}, \tag{1}$$

where $\|\cdot\|_0$ is the ℓ_0 norm counting the number of nonzero elements of \mathbf{x} . In practical applications, such as signal reconstruction problem, the measurement noise is unavoidable. Then the noisy CS recovery problem is formulated as

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad s.t. \quad \|\mathbf{y} - \Phi \mathbf{x}\|_2 \leq \epsilon, \tag{2}$$

where $\|\cdot\|_2$ is the ℓ_2 norm and ϵ is the residual error. However, since [problem \(2\)](#) is NP-hard, it is infeasible to solve it directly. Thus it is proposed to replace the ℓ_0 norm by the convex ℓ_1 norm, namely

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad s.t. \quad \|\mathbf{y} - \Phi \mathbf{x}\|_2 \leq \epsilon, \tag{3}$$

which is the well-known basis pursuit problem [2]. The [problem \(3\)](#) is the convex relaxation of

problem (2) and easy to solve by some methods including iterative shrinkage algorithm [8], alternating direction method of multipliers (ADMM) [23] and Bregman split algorithm [24]. Although problems (2) and (3) are fundamentally different, they return the same solution in many interesting situations [25]. By using an appropriate regularization parameter λ , we can equivalently rewrite the problem (3) as follows:

$$\min_{\mathbf{x}} \quad \|\mathbf{y} - \Phi\mathbf{x}\|_2^2 + \lambda\|\mathbf{x}\|_1. \tag{4}$$

By replacing the ℓ_1 norm by the non-convex ℓ_p ($0 > p > 1$) norm, in [26], the non-convex optimization problem based on ℓ_p norm can achieve the more exact CS reconstruction result than the convex one based on the ℓ_1 norm. Moreover, it has been shown in [18, 27] that the nonlocal sparsity exploiting the self-similarity of the natural image leads to the state-of-the-art performance. The nonlocal sparsity is a significant prior for CS recovery. In this work, we will propose a non-convex CS image recovery approach, which exploits the nonlocal sparsity.

Nonlocal smoothed rank function

In this section, we present a procedure of patch grouping that uses image self-similarity and leads to the low-rank problem. Hence one underlying assumption is the image exhibits abundant self-similarities. For a test image, we select some exemplar patches of size $\sqrt{n} \times \sqrt{n}$, reordered into the column vector lexicographically and denoted as $\hat{\mathbf{x}}_i \in \mathbb{R}^n$. For each exemplar patch, the underlying assumption makes it possible to find a number of similar patches. Specially, we employ a variant of k-nearest-neighbor search in a local window for each exemplar patch $\hat{\mathbf{x}}_i$ to find its similar patches as follows:

$$\mathbf{G}_i = \{j \mid \|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j\|_2 < c\}, \tag{5}$$

where c is a pre-defined threshold and \mathbf{G}_i indexes the positions of corresponding similar patches. After patch grouping, we align the similar patches as column vectors to form a matrix $\mathbf{X}_i = [\hat{\mathbf{x}}_i, \hat{\mathbf{x}}_{i_1}, \dots, \hat{\mathbf{x}}_{i_{m-1}}]$, $\mathbf{X}_i \in \mathbb{R}^{n \times m}$ for the i -th exemplar patch $\hat{\mathbf{x}}_i$. Under the above assumption, each matrix \mathbf{X}_i has the low-rank property since the similar patches have the similar structures. In this way, the nonlocal sparsity leads to a rank minimization problem in our image reconstruction approach.

Since the resulting matrix \mathbf{X}_i also contains the noise during patch grouping, in order to obtain a clean and clear image to well match the ground truth, let $\mathbf{X}_i = \mathbf{L}_i + \mathbf{W}_i$, where \mathbf{L}_i denotes the low-rank matrix and \mathbf{W}_i represents the noise matrix. We can obtain the low-rank matrix \mathbf{L}_i by solving the following problem:

$$\min_{\mathbf{L}_i} \text{rank}(\mathbf{L}_i) \quad \text{s.t.} \quad \|\mathbf{X}_i - \mathbf{L}_i\|_F^2 \leq \epsilon^2, \tag{6}$$

where $\|\cdot\|_F$ is the Frobenius norm and ϵ is the residual error. Unfortunately, problem (6) is NP-hard. A popular heuristic method is to adopt the nuclear norm (sum of the singular values) to replace the rank, i.e.,

$$\min_{\mathbf{L}_i} \|\mathbf{L}_i\|_* \quad \text{s.t.} \quad \|\mathbf{X}_i - \mathbf{L}_i\|_F^2 \leq \epsilon^2, \tag{7}$$

where $\|\cdot\|_*$ is the nuclear norm. The problem (7) is a convex surrogate of the rank and can be solved efficiently by the singular value thresholding (SVT) algorithm [28]. However, it is in practice suboptimal due to equally treating each singular value. In [19], Gu et al. have demonstrated that non-convex low-rank approximations adaptively treating the singular values at different scales yield better results than those convex ones.

In [29, 30], they have been shown that for a low-rank matrix $\mathbf{X} \in \mathbb{R}^{n \times m}$, the rank can be approximated by the following SRF function:

$$G_\delta(\mathbf{X}) = \ell - \sum_{j=1}^{\ell} e^{-\sigma_j^2(\mathbf{X})/2\delta^2}, \tag{8}$$

where $\sigma_j(\mathbf{X})$ is the j -th singular value of \mathbf{X} , $\ell = \min\{n, m\}$ and δ is an adjustable parameter. Although the problem of minimizing $G_\delta(\cdot)$ function using a small δ will lead to many local minima, the solution of minimizing $G_\delta(\cdot)$ function converges to the minimum rank solution as δ goes to zero [30] as the red curve shown in Fig 1. In order to avoid trapping in local solutions, we initialize a large δ and gradually decrease δ to improve the degree of approximation for the rank. With the decreasing δ , the rank can be better approximated by SRF Eq (8). This technique refines the minimizer of the SRF minimization problem by considering the minimizer of the previous iteration (large δ) as the new initial point of current iteration (small δ), which makes the solution of the SRF minimization problem get closer to the minimum rank solution during the iterations. Fig 1 intuitively illustrates the comparison of SRF, the rank, the nuclear norm and logdet function in the case of a scalar. We observe that SRF can better approximate the rank than the nuclear norm and logdet function. Therefore, we speculate that the problem of minimizing SRF toward rank minimization problem could achieve better performance. Then, we consider the SRF function as the low-rank regularization for CS image recovery.

For the low-rank matrix \mathbf{L}_i , the rank minimization problem can be replaced to

$$\min_{\mathbf{L}_i} G_\delta(\mathbf{L}_i) \quad s.t. \quad \|\mathbf{X}_i - \mathbf{L}_i\|_F^2 \leq \epsilon^2, \tag{9}$$

which can be reformulated with an appropriate parameter λ ,

$$\mathbf{L}_i = \arg \min_{\mathbf{L}_i} \frac{1}{2} \|\mathbf{X}_i - \mathbf{L}_i\|_F^2 + \lambda G_\delta(\mathbf{L}_i). \tag{10}$$

Obviously, the problem (10) is smoothed and differentiable, and we can adopt the gradient

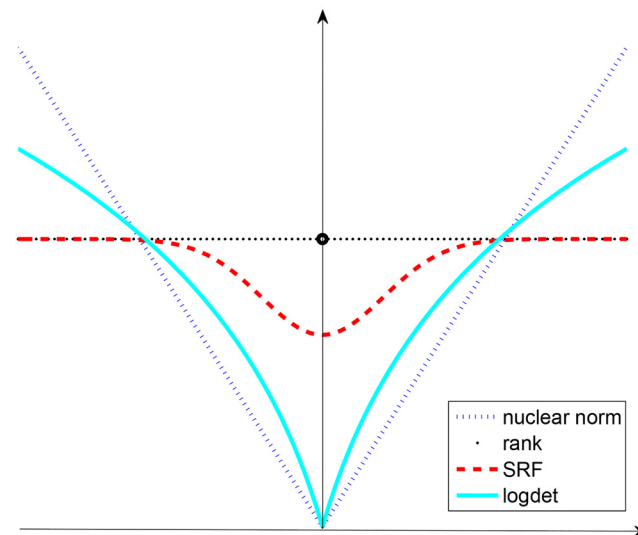


Fig 1. Performance of SRF = $G_\delta(x)$, rank = $\|x\|_0$, the nuclear norm = $\|x\|_1$ and logdet = $\log(|x|+\epsilon)$ in the case of a scalar, where $\delta = 1$ and $\epsilon = 1$.

doi:10.1371/journal.pone.0162041.g001

descent method to solve it. For each matrix \mathbf{X}_i , the same method can be used to obtain the corresponding low-rank matrix \mathbf{L}_i .

For CS image reconstruction problem, based on above presented $G_\delta(\mathbf{L}_i)$, we propose the global model as follows:

$$(\hat{\mathbf{x}}, \hat{\mathbf{L}}_i) = \arg \min_{\mathbf{x}, \mathbf{L}_i} \|\mathbf{y} - \Phi\mathbf{x}\|_2^2 + \eta \sum_i \{ \|\mathbf{P}_i\mathbf{x} - \mathbf{L}_i\|_F^2 + \lambda G_\delta(\mathbf{L}_i) \}, \quad (11)$$

where η is a regularization parameter and $\mathbf{P}_i\mathbf{x} = [\mathbf{P}_{i_0}\mathbf{x}, \mathbf{P}_{i_1}\mathbf{x}, \dots, \mathbf{P}_{i_{m-1}}\mathbf{x}]$ is the matrix constituted by the set of similar patches for each exemplar patch \mathbf{x}_i . The [model \(11\)](#) exploits the non-local sparsity of the image patches and non-convexity of the SRF function $G_\delta(\cdot)$. Therefore, we conjecture the proposed method can achieve the good performance. To solve the [model \(11\)](#), we develop an efficient alternative minimization method in the next section.

Optimization algorithm

It is difficult to directly solve the global [model \(11\)](#) since \mathbf{x} and \mathbf{L}_i is coupled. We use the alternative minimization method to decouple \mathbf{x} and \mathbf{L}_i as follows:

- For \mathbf{L}_i -subproblem, its optimization problem is

$$\hat{\mathbf{L}}_i = \arg \min_{\mathbf{L}_i} \|\mathbf{P}_i\mathbf{x} - \mathbf{L}_i\|_F^2 + \lambda G_\delta(\mathbf{L}_i). \quad (12)$$

- For \mathbf{x} -subproblem, its optimization problem is

$$\hat{\mathbf{x}} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \Phi\mathbf{x}\|_2^2 + \eta \sum_i \|\mathbf{P}_i\mathbf{x} - \mathbf{L}_i\|_F^2. \quad (13)$$

In the following subsections we will present the optimization algorithm to solve subproblems Eqs [\(12\)](#) and [\(13\)](#) in detail.

Low-rank matrix optimization algorithm

Let $f(\mathbf{L}_i) = \|\mathbf{P}_i\mathbf{x} - \mathbf{L}_i\|_F^2 + \lambda G_\delta(\mathbf{L}_i)$. Then the low-rank matrix \mathbf{L}_i can be obtained via the gradient descent method, i.e., $\mathbf{L}_i^{(k+1)} \leftarrow \mathbf{L}_i^{(k)} - \mu^{(k)} \nabla f(\mathbf{L}_i^{(k)})$ where $\mu^{(k)}$ is the step size in the k -th iteration and $\nabla f(\cdot)$ is the gradient of $f(\cdot)$. Before deriving $\nabla f(\cdot)$, we show the gradient of SRF $G_\delta(\mathbf{L}_i)$ at \mathbf{L}_i [\[30\]](#) as follows:

$$\nabla G_\delta(\mathbf{L}_i) = -\mathbf{U} \text{diag} \left(-\frac{\sigma_1}{\delta^2} e^{-\sigma_1^2/2\delta^2}, \dots, -\frac{\sigma_\ell}{\delta^2} e^{-\sigma_\ell^2/2\delta^2} \right) \mathbf{V}^T, \quad (14)$$

where matrixes \mathbf{U} and \mathbf{V} and singular values σ_i ($i = 1, \dots, \ell$) come from the singular value decomposition (SVD) of \mathbf{L}_i , which is obtained in previous iteration,

$$\mathbf{L}_i = \mathbf{U} \text{diag}(\sigma_1, \dots, \sigma_\ell) \mathbf{V}^T. \quad (15)$$

Then, it is easy to derive the gradient of the function $f(\mathbf{L}_i)$,

$$\nabla f(\mathbf{L}_i) = -2(\mathbf{P}_i\mathbf{x} - \mathbf{L}_i) - \lambda \mathbf{U} \text{diag} \left(-\frac{\sigma_1}{\delta^2} e^{-\sigma_1^2/2\delta^2}, \dots, -\frac{\sigma_\ell}{\delta^2} e^{-\sigma_\ell^2/2\delta^2} \right) \mathbf{V}^T. \quad (16)$$

According to the gradient descent method, at each iteration, one has

$$\mathbf{L}_i \leftarrow \mathbf{L}_i - \mu \nabla f(\mathbf{L}_i). \tag{17}$$

In addition, the step size μ in formula (17) should be set in a decreasing order to return a better result. Following the papers [30, 31], we set $\mu = \delta^2$ to decrease μ proportional to δ^2 . Then, above formula (17) incorporating the Eq (16) can be rewritten as

$$\mathbf{L}_i \leftarrow \xi \mathbf{L}_i + (1 - \xi) \mathbf{P}_i \mathbf{x} - \lambda \mathbf{U} \text{diag}(\sigma_1 e^{-\sigma_1^2/2\delta^2}, \dots, \sigma_i e^{-\sigma_i^2/2\delta^2}) \mathbf{V}^T, \tag{18}$$

where $\xi = 1 - 2\delta^2$.

Image reconstruction via alternating direction method of multipliers

After obtaining the low-rank matrix \mathbf{L}_i , we reconstruct the image \mathbf{x} via solving the problem (13). It is clear that the problem (13) is a quadratic optimization problem with a closed-form solution,

$$\mathbf{x} = (\Phi^H \Phi + \eta \sum_i \mathbf{P}_i^T \mathbf{P}_i)^{-1} (\Phi^H \mathbf{y} + \eta \sum_i \mathbf{P}_i^T \mathbf{L}_i). \tag{19}$$

However, the inverse of the matrix $(\Phi^H \Phi + \eta \sum_i \mathbf{P}_i^T \mathbf{P}_i)$ is very large and difficult to compute. Therefore, we consider to reconstruct the image \mathbf{x} in the framework of ADMM, which also leads to the closed-form solutions for each subproblem. The ADMM method is often used to image restoration [32–36].

Based on the definition of ADMM, we present the augmented Lagrangian function of the problem (13) as follows:

$$\mathcal{L}_d(\mathbf{x}, \mathbf{z}) = \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \beta \|\mathbf{x} - \mathbf{z} + \frac{\mathbf{d}}{2\beta}\|_2^2 + \eta \sum_i \|\mathbf{P}_i \mathbf{z} - \mathbf{L}_i\|_F^2, \tag{20}$$

where $\mathbf{z} = \mathbf{x}$ is the auxiliary variable, \mathbf{d} is the Lagrangian multiplier, and $\beta > 0$ is a scalar. With respect to the variables \mathbf{x} , \mathbf{z} and \mathbf{d} , they are decoupled in the framework of ADMM, thus, can be solved separately, leading to the following iterations:

$$\mathbf{z}^{(k+1)} = \arg \min_{\mathbf{z}} \beta \|\mathbf{x}^{(k)} - \mathbf{z} + \frac{\mathbf{d}^{(k)}}{2\beta}\|_2^2 + \eta \sum_i \|\mathbf{P}_i \mathbf{z} - \mathbf{L}_i\|_F^2, \tag{21}$$

$$\mathbf{x}^{(k+1)} = \arg \min_{\mathbf{x}} \|\mathbf{y} - \Phi \mathbf{x}\|_2^2 + \beta \|\mathbf{x}^{(k)} - \mathbf{z}^{(k+1)} + \frac{\mathbf{d}^{(k)}}{2\beta}\|_2^2, \tag{22}$$

$$\mathbf{d}^{(k+1)} = \mathbf{d}^{(k)} + \beta(\mathbf{x}^{(k+1)} - \mathbf{z}^{(k+1)}). \tag{23}$$

Clearly, both subproblems Eqs (21) and (22) are quadratic optimization problem and have the closed-form solutions. For the subproblem Eq (21), its explicit solution is:

$$\mathbf{z}^{(k+1)} = (\eta \sum_i \mathbf{P}_i^T \mathbf{P}_i + \beta \mathbf{I})^{-1} \left(\beta \mathbf{x}^{(k)} + \frac{\mathbf{d}^{(k)}}{2} + \eta \sum_i \mathbf{P}_i^T \mathbf{L}_i \right). \tag{24}$$

For the subproblem Eq (22), according to its first-order derivation for \mathbf{x} , we can derive the following equation:

$$(\Phi^H \Phi + \beta \mathbf{I}) \mathbf{x}^{(k+1)} = \Phi^H \mathbf{y} + \beta \mathbf{z}^{(k+1)} - \frac{\mathbf{d}^{(k)}}{2}. \tag{25}$$

Considering measurement matrix Φ is a partial Fourier transform matrix, we can transform above problem from image space to Fourier space to efficiently obtain \mathbf{x} . Concretely, we let $\Phi = \mathbf{DF}$, where \mathbf{D} is the down-sampling matrix and \mathbf{F} is the fourier transform matrix. It is easy to achieve

$$\mathbf{F}((\mathbf{DF})^H \mathbf{DF} + \beta \mathbf{I}) \mathbf{F}^H \mathbf{F} \mathbf{x}^{(k+1)} = \mathbf{F}(\mathbf{DF})^H \mathbf{y} + \mathbf{F} \left(\beta \mathbf{z}^{(k+1)} - \frac{\mathbf{d}^{(k)}}{2} \right), \quad (26)$$

and then $\mathbf{x}^{(k+1)}$ can be drawn

$$\mathbf{x}^{(k+1)} = \mathbf{F}^H (\mathbf{D}^T \mathbf{D} + \beta \mathbf{I})^{-1} \left(\mathbf{D}^T \mathbf{y} + \mathbf{F} \left(\beta \mathbf{z}^{(k+1)} - \frac{\mathbf{d}^{(k)}}{2} \right) \right). \quad (27)$$

We simultaneously obtain the low-rank matrix \mathbf{L}_i and the image \mathbf{x} by the alternative minimization method, and the overall procedure is summarized below as **Algorithm 1**. Empirically we have found that **Algorithm 1** is convergent, but in theory the convergence analysis of **Algorithm 1** is difficult to give due to the non-convex subproblem. Although there are some papers have proved the convergence of their non-convex optimization problem, these proofs hold in a few unrealistic and rigorous assumptions. Specially, to save computational complexity, we set $J = K = 1$ in **Algorithm 1**. Moreover, we estimate an image $\hat{\mathbf{x}}$ using the discrete cosine transform (DCT) method as the initial solution for a better initial point, which has been seen in [22]. As iteration increases, the high accuracy results will be achieved.

The complexity of **Algorithm 1** is $O([T_s + \log N]N)$ (N is the total number of image pixels and T_s is the average complexity to compute similar patches per exemplar patch), which is mainly generated by the DCT method in 1 step of initialization and the Fourier transform in 4 (b) step of inner loop. The complexity of SVD in 3 (a) step of inner loop, i.e. $O(n \times m^2)$, can be ignored due to $n, m \ll N$. Therefore, the proposed method is practical feasible and promising.

Algorithm 1 CS via the SRF function

Initialization:

1. Estimate an initial image $\hat{\mathbf{x}}$ using the discrete cosine transform (DCT) method;
2. Set regularization parameters λ and η , and parameter $\beta > 0$;
3. Let $\delta^{(0)} = \bar{\delta}$, $\mu^{(0)} = (\bar{\delta}^{(0)})^2$, $\mathbf{x}^{(1)} = \hat{\mathbf{x}}$, $\mathbf{d}^{(1)} = 0$;
4. Find a set of similar patches by the method Eq (5) for each exemplar patch using $\mathbf{x}^{(1)}$;

Outer loop: for $s = 1, 2, \dots, S$ do

1. Match similar patches into a matrix \mathbf{X}_i for each exemplar patch using $\mathbf{x}^{(s)}$;
2. Set $\mathbf{L}_i^{(0)} = \mathbf{X}_i$;
3. **Inner loop** (solving the \mathbf{L}_i -subproblem Eq (12)): for $j = 1, 2, \dots, J$ do
 - (a) Compute the SVD of $\mathbf{L}_i^{(j-1)} = \mathbf{U} \mathbf{diag}(\sigma_1, \dots, \sigma_t) \mathbf{V}^H$;
 - (b) Compute $\nabla f(\mathbf{L}_i^{(j-1)})$ by Eq (16);
 - (c) Compute $\mathbf{L}_i^{(j)} = \mathbf{L}_i^{(j-1)} - \mu^{(j-1)} \nabla f(\mathbf{L}_i^{(j-1)})$;
 - (d) Update $\delta^{(j)} = c \delta^{(j-1)}$, $0 < c < 1$;
 - (e) Update $\mu^{(j)} = (\delta^{(j)})^2$;
 - (f) Output $\mathbf{L}_i = \mathbf{L}_i^{(j)}$ until $j = J$.

End for

4. **Inner loop** (solving the \mathbf{x} -subproblem Eq (13)): for $k = 1, 2, \dots, K$ do

- (a) Update $z^{(k+1)}$ by the Eq (24);
- (b) Update $\mathbf{x}^{(k+1)}$ by the Eq (27);
- (c) Update $\mathbf{d}^{(k+1)}$ by the Eq (23);
- (d) Output $\mathbf{x}^{(s)} = \mathbf{x}^{(k+1)}$ until $k = K$.

End for

5. Output the reconstructed image $\hat{\mathbf{x}} = \mathbf{x}^{(s)}$ until $s = S$.

End for

Numerical results

Here, we present the experimental results of our approach for CS image reconstruction based on the SRF regularization. We generate the CS measurements by random and pseudo-radial sampling of the Fourier transform coefficients of test images respectively. The number of measurements is $M = rate * N$, where *rate* is the sampling rate. We use peak signal to noise ratio (PSNR) and Structural SIMilarity (SSIM) index [37] as the quantitative measures in our numerical experiments, and the PSNR is defined as

$$PSNR = 20 \log_{10} \frac{MAX_f}{\sqrt{MSE}},$$

where MAX_f is the maximum possible pixel value of the image and MSE is the mean squared error, defined as

$$MSE = \frac{1}{N} \sum_{j=1}^N [f(j) - g(j)]^2,$$

where f and g are the original image and the restored image, respectively.

In all experiments, the main parameters of the **Algorithm 1** are set as follows: patch size $\sqrt{n} = 6$; the number of similar patches for each exemplar patch $m = 45$ (the more similar patches form a more low-rank matrix but lead to high computational complexity); initialize value $\delta^{(0)} = \bar{\delta}$, where $\bar{\delta}$ is a constant around two times of the largest singular value of initial L , [30]; the parameter c is set as 0.08 experimentally in formula $\delta^{(j)} = c\delta^{(j-1)}$; the outer loop iteration number $S = 400$ (that is selected based on the convergence rate of **Algorithm 1**). To achieve better performance, we select exemplar patch in each 5 pixels along both horizontal and vertical directions.

In addition, the total variation (TV) method [38], the BM3D based CS method (BM3D-CS) [16], the nuclear norm method (denoted as NLR-CS-baseline) [22] and the logdet function method (denoted as NLR-CS) [22] for CS image recovery are compared with the proposed SRF approach (called as SRF-CS). The TV method just considered the underlying sparsity of an image, and the BM3D-CS method used the nonlocal sparsity with an outstanding performance. The NLR-CS-baseline method exploited the nuclear norm to replace the rank for solving the rank minimization problem. The NLR-CS approach adopted the logdet function to achieve a more accurate surrogate than the nuclear norm for the rank. The source codes of above methods [16, 22, 38] are publicly downloaded from the authors's websites. These methods are significantly state-of-the-art CS algorithms for image recovery. We have carefully tuned their parameters to obtain the best results for fair comparison. Both noiseless and noisy experiments are presented to demonstrate the performance of the proposed approach for CS recovery. The test images (256×256) are exhibited in Fig 2, where Fig 2(a), 2(c), 2(e) and 2(f) can be publicly downloaded from the test image net <http://decsai.ugr.es/cvg/dbimágenes/g256.php>, Fig 2(b), 2(d) and 2(g) can be publicly downloaded from the author's homepage of [22] http://see.xidian.edu.cn/faculty/wsdong/NLR_Exps.htm.

Noiseless experiments for CS recovery

We reconstruct the images from less $M = 0.05N$ and $M = 0.1N$ measurements by randomly sampling the Fourier transform coefficients of the test images. Fig 3 shows the restored images by TV [38], NLR-CS-baseline [22], NLR-CS [22] and the proposed methods for CS image recovery. The TV method for CS recovery can't work well due to less measurements. The close-ups of Number image and Book image obtained by the TV method are indistinct and

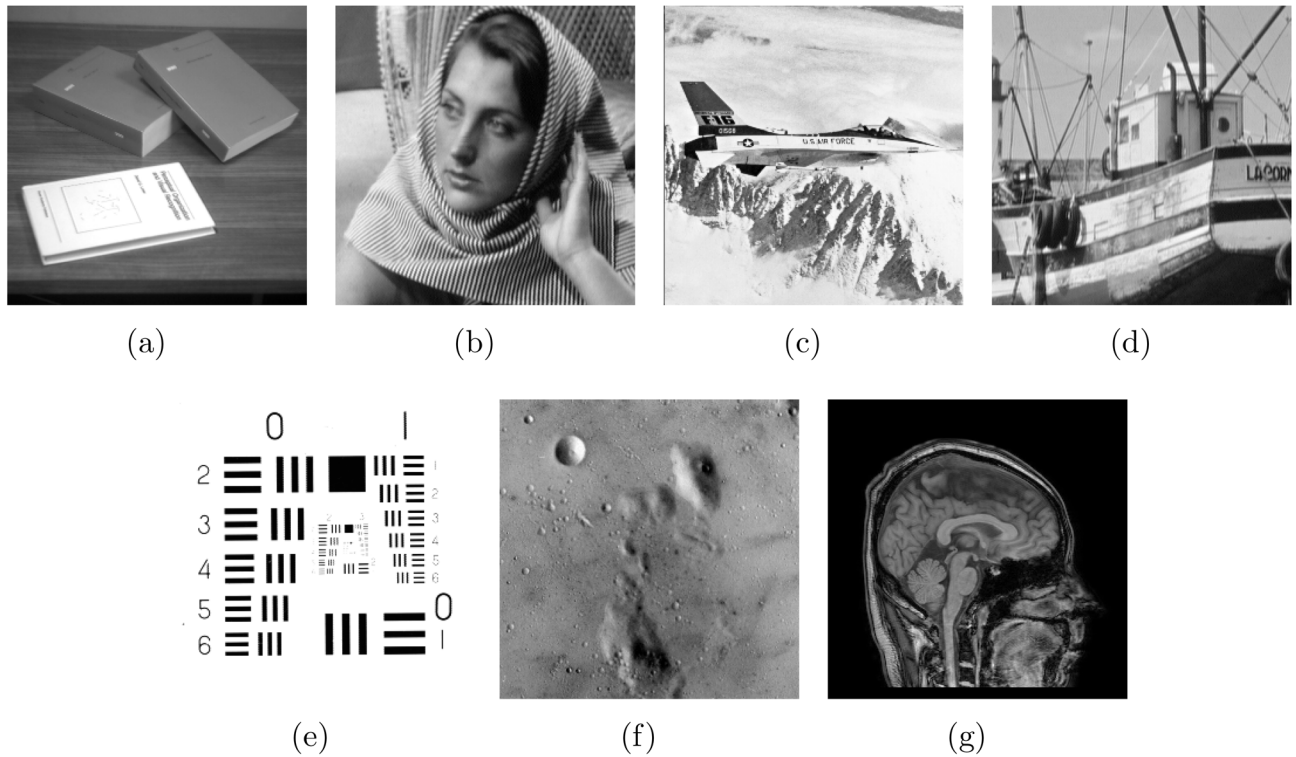


Fig 2. The test images: (a) Book; (b) Barbara; (c) Plane; (d) Boat; (e) Number; (f) Moonscape; (g) Head.

doi:10.1371/journal.pone.0162041.g002

lose much structured information. The NLR-CS-baseline method gives rise to some artificial shadow like fog as shown in the close-up of Number image. It is well known that the NLR-CS method is a very competitive method for CS recovery due to its remarkable performance. For these test images in Fig 3, however, the performance of the NLR-CS method is undesirable. Obviously, our approach achieves the best visual quality and highest PSNR value among all test methods. Noteworthy, the better performance of our approach compared with the NLR-CS-baseline method verifies that the SRF function Eq (8) is a more accurate surrogate for the rank than the nuclear norm. Moreover, Table 1 displays the PSNR values and the SSIM values obtained by all test methods.

Although the proposed method leads to a non-convex optimization problem, we can see that the proposed method for CS image recovery converges in a few iterations in Fig 4. In the following subsection, we will verify that the proposed method is also robust in a noisy situation.

Noisy experiments for CS recovery

In the noisy experiments, we add Gaussian noise with 0.05 standard deviation to the original images. Firstly, we recover the Head image from $M = 0.15N$ measurements and the Boat image from $M = 0.2N$ measurements by randomly sampling their Fourier transform coefficients respectively. The restoration results by the TV [38], BM3D-CS [16], NLR-CS-baseline [22], NLR-CS [22] and the proposed methods are presented in Fig 5. Clearly, the restored Head image becomes blocky obtained by the TV method, and contains heavy noise obtained by the NLR-CS-baseline method. Meanwhile, some edges in the close-up of the restored Head image achieved by the NLR-CS method are unclear, and the close-up of Head image generated by the BM3D-CS method is blurring. However, compared with other methods, the reconstructed

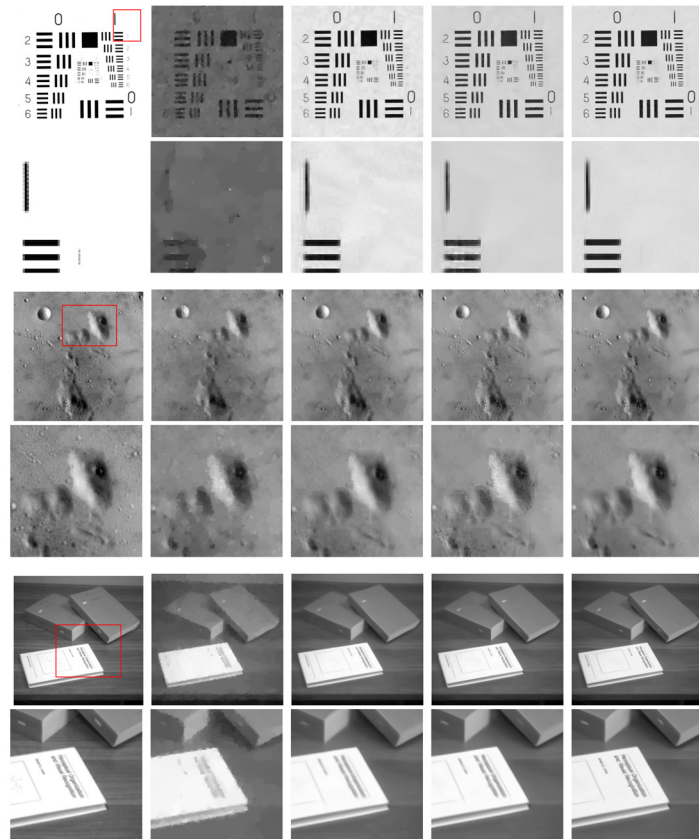


Fig 3. Reconstructed images by randomly sampling. From top to bottom for each column: Number image, close-up of the red rectangle of Number image, Moonscape image, close-up of the red rectangle of Moonscape image, Book image, close-up of the red rectangle of Book image. From left to right for each row: the original image, the TV method, the NLR-CS-baseline method, the NLR-CS method, the proposed method.

doi:10.1371/journal.pone.0162041.g003

Table 1. The PSNR (dB) and SSIM values of the restoration results achieved by the TV method [38], the NLR-CS-baseline method [22], the NLR-CS method [22] and the proposed method for noiseless images.

Image	Number of measurements	Method	PSNR (dB)	SSIM
Number	M = 0.05N	TV	24.10	0.6545
		NLR-CS-baseline	26.29	0.8970
		NLR-CS	26.70	0.9407
		proposed	29.04	0.9497
Moonscape	M = 0.1N	TV	26.41	0.5450
		NLR-CS-baseline	30.19	0.7037
		NLR-CS	29.92	0.7004
		proposed	30.95	0.7143
Book	M = 0.05N	TV	32.41	0.7304
		NLR-CS-baseline	37.41	0.9371
		NLR-CS	38.44	0.9430
		proposed	39.77	0.9528

doi:10.1371/journal.pone.0162041.t001

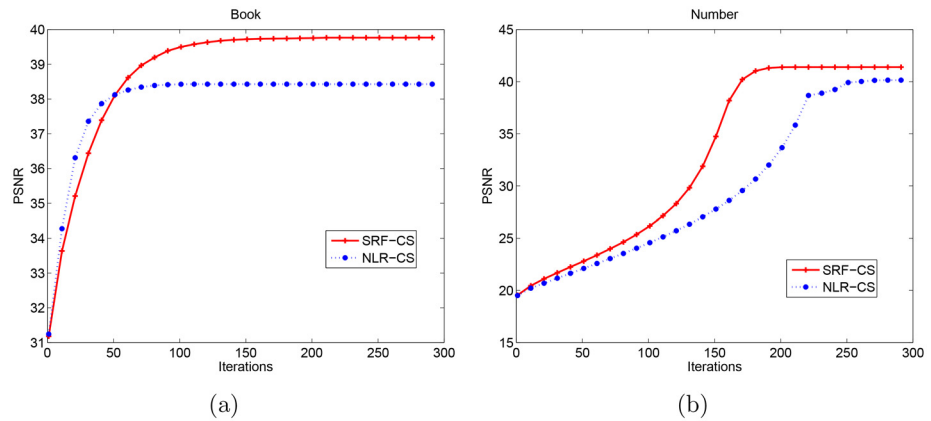


Fig 4. PSNR value vs. Iteration plots for the proposed method and the NLR-CS method. (a) Book image, measures number $M = 0.05N$ (b) Number image, measurements number $M = 0.1N$.

doi:10.1371/journal.pone.0162041.g004

Head image by the proposed approach returns the clearer edges and less noise. For the Boat image, we can see that the proposed approach also outperforms the other test methods with the better visual quality, the higher PSNR and SSIM values. Thus the proposed method with less measurements is robust in noisy situation.

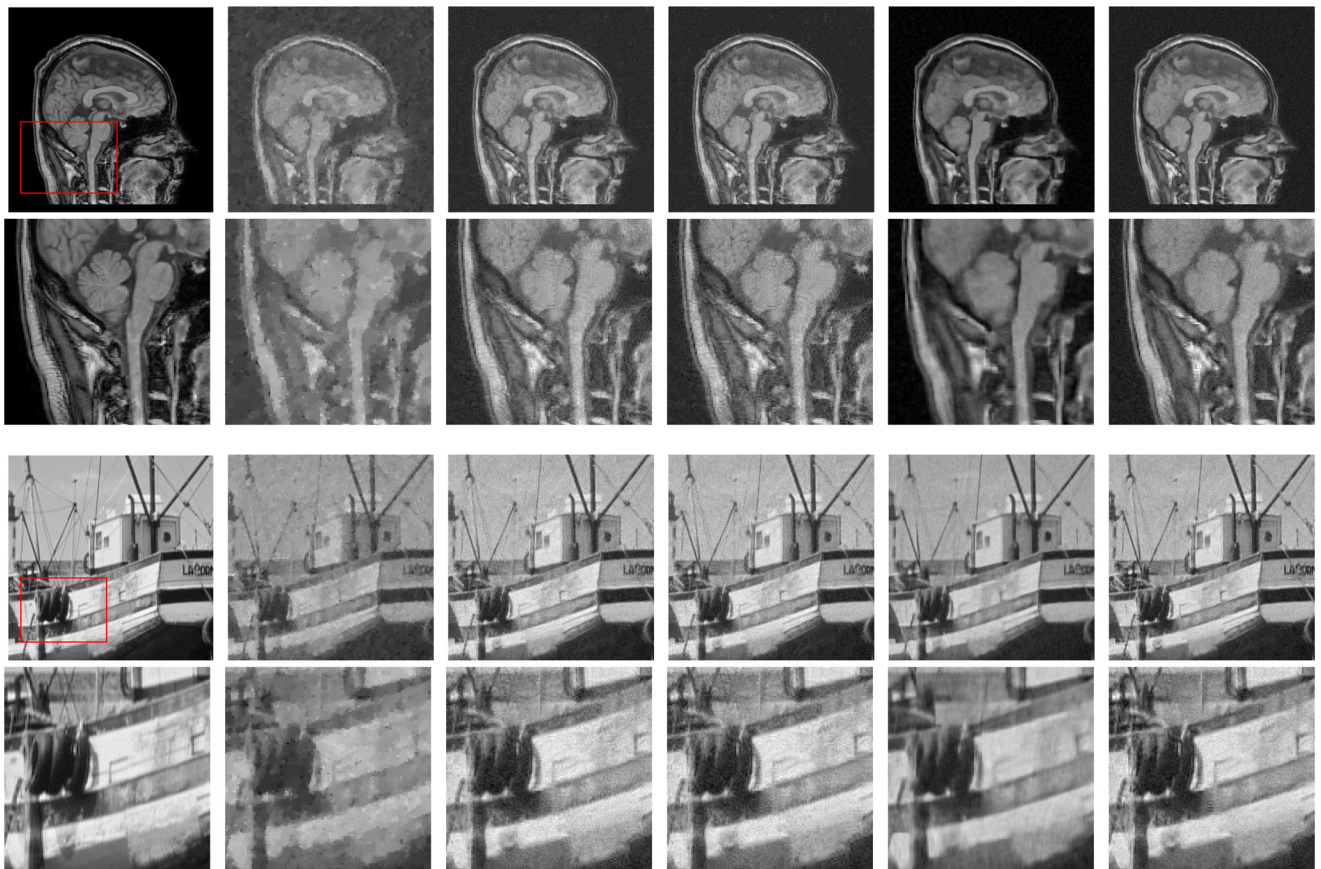


Fig 5. Reconstructed images by randomly sampling. From top to bottom for each column: Head image, close-up of the red rectangle of Head image, Boat image, close-up of the red rectangle of Boat image. From left to right for each row: the original image, the TV method, the NLR-CS-baseline method, the NLR-CS method, the BM3D-CS method, the proposed method.

doi:10.1371/journal.pone.0162041.g005



Fig 6. Reconstructed images by pseudo-radial sampling. From top to bottom for each column: Plane image, close-up of the red rectangle of Plane image, Barbara image, close-up of the red rectangle of Barbara image. From left to right for each row: the original image, the TV method, the NLR-CS-baseline method, the NLR-CS method, the BM3D-CS method, the proposed method.

doi:10.1371/journal.pone.0162041.g006

Secondly, we recover the Boy image from $M = 0.24N$ (65 radial lines) measurements and the Barbara image from $M = 0.29N$ (80 radial lines) measurements by pseudo-radial sampling their Fourier transform coefficients. The pseudo-radial sampling way generates streaking artifacts leading to more difficult image reconstruction than randomly sampling way. We illustrate the performance results of all the test methods in Fig 6. Visually, the close-ups of Plane image generated by the TV, NLR-CS-baseline, NLR-CS and BM3D-CS methods are blurring. Nevertheless, the close-up of Plane image obtained by our approach is good with less noise and higher PSNR value. The close-ups of Barbara image achieved by the NLR-CS-baseline method and the NLR-CS method create some artificial strips, which do not exist in the original Barbara image. The bad performance may be caused by the pseudo-radial sampling way or noise. The texture of restored Barbara image by BM3D method is lost. However, the proposed approach removes the artifacts and a lot of noise returning a better result, which demonstrates its robustness. In addition, Table 2 displays the PSNR and SSIM values obtained by all test methods. To sum up, the proposed approach outperforms other competing methods for CS image reconstruction by both randomly sampling and pseudo-radial sampling schemes.

Table 2. The PSNR (dB) and SSIM values of the restoration results obtained by the TV method [38], the NLR-CS-baseline method [22], the NLR-CS method [22], the BM3D-CS method [16] and the proposed method for noisy images.

Image	Number of measurements	Sampling scheme	Method	PSNR (dB)	SSIM
Head	M = 0.15N	random	TV	23.07	0.5489
			NLR-CS-baseline	27.15	0.6508
			NLR-CS	27.20	0.6515
			BM3D-CS	27.56	0.7203
			proposed	27.76	0.6615
Boat	M = 0.2N	random	TV	22.90	0.5313
			NLR-CS-baseline	27.60	0.7091
			NLR-CS	27.62	0.7093
			BM3D-CS	27.75	0.7281
			proposed	27.98	0.7397
Plane	M = 0.2N	pseudo-radial	TV	25.86	0.6410
			NLR-CS-baseline	28.64	0.8657
			NLR-CS	28.98	0.8866
			BM3D-CS	28.55	0.8203
			proposed	29.38	0.9014
Barbara	M = 0.29N	pseudo-radial	TV	24.09	0.6131
			NLR-CS-baseline	27.79	0.7721
			NLR-CS	28.12	0.7980
			BM3D-CS	28.62	0.8093
			proposed	29.20	0.8231

doi:10.1371/journal.pone.0162041.t002

Conclusion

To better exploit the nonlocal sparsity of similar patches and non-convexity of rank minimization, in this paper, we use the non-convex SRF function surrogating the rank as a low-rank regularization for CS image recovery. This SRF function can better approximate the rank than the standard nuclear norm and the logdet function. We propose an efficient algorithm in the framework of alternative minimization method, which divides this CS problem into the SRF minimization subproblem and the least square subproblem. With respect to the minimization subproblem of the SRF function, we adopt the gradient descent method to solve it since it is differentiable. Simultaneously, the clear image is reconstructed by solving the least square subproblem using the ADMM method. Both noiseless and noisy numerical experiments demonstrate that the proposed approach achieves the better performance and vision quality under the lower sampling rate situation. In the future, we would like to explore better surrogates for the rank to improve performance, and solve other practical problems.

Author Contributions

Formal analysis: YRF JL XLZ.

Project administration: TZH XLZ.

Resources: JL.

Software: YRF.

Validation: TZH.

Writing – original draft: YRF.

Writing – review & editing: YRF TZH JL XLZ.

References

1. Donoho DL. Compressed sensing. *IEEE Transactions on Information Theory*. 2006; 52(4):1289–1306. doi: [10.1109/TIT.2006.871582](https://doi.org/10.1109/TIT.2006.871582)
2. Candès EJ, Romberg J, Tao T. Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*. 2006; 52(2):489–509. doi: [10.1109/TIT.2005.862083](https://doi.org/10.1109/TIT.2005.862083)
3. Lustig M, Donoho DL, Santos JM, Pauly JM. Compressed sensing MRI. *IEEE Signal Processing Magazine*. 2008; 25(2):72–82. doi: [10.1109/MSP.2007.914728](https://doi.org/10.1109/MSP.2007.914728)
4. Lustig M, Donoho D, Pauly JM. Sparse MRI: The application of compressed sensing for rapid MR imaging. *Magnetic Resonance in Medicine*. 2007; 58(6):1182–1195. doi: [10.1002/mrm.21391](https://doi.org/10.1002/mrm.21391) PMID: [17969013](https://pubmed.ncbi.nlm.nih.gov/17969013/)
5. Gehm M, John R, Brady D, Willett R, Schulz T. Single-shot compressive spectral imaging with a dual-disperser architecture. *Optics Express*. 2007; 15(21):14013–14027. doi: [10.1364/OE.15.014013](https://doi.org/10.1364/OE.15.014013) PMID: [19550674](https://pubmed.ncbi.nlm.nih.gov/19550674/)
6. Du Q, Fowler JE. Hyperspectral image compression using JPEG 2000 and principal component analysis. *IEEE Geoscience and Remote Sensing Letters*. 2007; 4(2):201–205. doi: [10.1109/LGRS.2006.888109](https://doi.org/10.1109/LGRS.2006.888109)
7. Dony RD, Haykin S. Neural network approaches to image compression. *Proceedings of the IEEE*. 1995; 83(2):288–303. doi: [10.1109/5.364461](https://doi.org/10.1109/5.364461)
8. Beck A, Teboulle M. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*. 2009; 2(1):183–202. doi: [10.1137/080716542](https://doi.org/10.1137/080716542)
9. Zibulevsky M, Elad M. L1-L2 optimization in signal and image processing. *IEEE Signal Processing Magazine*. 2010; 27(3):76–88. doi: [10.1109/MSP.2010.936023](https://doi.org/10.1109/MSP.2010.936023)
10. Wang Y, Yin W. Sparse signal reconstruction via iterative support detection. *SIAM Journal on Imaging Sciences*. 2010; 3(3):462–491. doi: [10.1137/090772447](https://doi.org/10.1137/090772447)
11. Koltchinskii V. Sparsity in penalized empirical risk minimization. *Annales de l'IHP Probabilités et statistiques*. 2009; 45(1):7–57.
12. Huang J, Zhang T, Metaxas D. Learning with structured sparsity. *The Journal of Machine Learning Research*. 2011; 12:3371–3412.
13. Deng W, Yin W, Zhang Y. Group sparse optimization by alternating direction method. In: *SPIE Optical Engineering + Applications*. International Society for Optics and Photonics; 2013. p. 88580–88580.
14. Liu J, Huang TZ, I W S, Lv XG, Chen P. Image restoration using total variation with overlapping group sparsity. *Information Sciences*. 2015; 295:232–246. doi: [10.1016/j.ins.2014.10.041](https://doi.org/10.1016/j.ins.2014.10.041)
15. Dong W, Shi G, Li X, Zhang L, Wu X. Image reconstruction with locally adaptive sparsity and nonlocal robust regularization. *Signal Processing: Image Communication*. 2012; 27(10):1109–1122.
16. Egiazarian K, Foi A, Katkovnik V. Compressed sensing image reconstruction via recursive spatially adaptive filtering. In: *2007 IEEE International Conference on Image Processing*. vol. 1. IEEE; 2007. p. 1–549.
17. Buades A, Coll B, Morel JM. Image denoising by non-local averaging. In: *ICASSP (2)*; 2005. p. 25–28.
18. Dabov K, Foi A, Katkovnik V, Egiazarian K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*. 2007; 16(8):2080–2095. doi: [10.1109/TIP.2007.901238](https://doi.org/10.1109/TIP.2007.901238) PMID: [17688213](https://pubmed.ncbi.nlm.nih.gov/17688213/)
19. Gu S, Zhang L, Zuo W, Feng X. Weighted nuclear norm minimization with application to image denoising. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; 2014. p. 2862–2869.
20. Han D, Kong W, Zhang W. A partial splitting augmented Lagrangian method for low patch-rank image decomposition. *Journal of Mathematical Imaging and Vision*. 2015; 51(1):145–160. doi: [10.1007/s10851-014-0510-7](https://doi.org/10.1007/s10851-014-0510-7)
21. Ji H, Liu C, Shen Z, Xu Y. Robust video denoising using low rank matrix completion. In: *CVPR*. Cite-seer; 2010. p. 1791–1798.
22. Dong W, Shi G, Li X, Ma Y, Huang F. Compressive sensing via nonlocal low-rank regularization. *IEEE Transactions on Image Processing*. 2014; 23(8):3618–3632. doi: [10.1109/TIP.2014.2329449](https://doi.org/10.1109/TIP.2014.2329449) PMID: [24951688](https://pubmed.ncbi.nlm.nih.gov/24951688/)

23. He B, Tao M, Yuan X. Alternating direction method with Gaussian back substitution for separable convex programming. *SIAM Journal on Optimization*. 2012; 22(2):313–340. doi: [10.1137/110822347](https://doi.org/10.1137/110822347)
24. Zhang X, Burger M, Bresson X, Osher S. Bregmanized nonlocal regularization for deconvolution and sparse reconstruction. *SIAM Journal on Imaging Sciences*. 2010; 3(3):253–276. doi: [10.1137/090746379](https://doi.org/10.1137/090746379)
25. Romberg J. Imaging via compressive sampling [introduction to compressive sampling and recovery via convex programming]. *IEEE Signal Processing Magazine*. 2008; 25(2):14–20.
26. Chartrand R. Exact reconstruction of sparse signals via nonconvex minimization. *IEEE Signal Processing Letters*. 2007; 14(10):707–710. doi: [10.1109/LSP.2007.898300](https://doi.org/10.1109/LSP.2007.898300)
27. Mairal J, Bach F, Ponce J, Sapiro G, Zisserman A. Non-local sparse models for image restoration. In: 2009 IEEE 12th International Conference on Computer Vision. IEEE; 2009. p. 2272–2279.
28. Cai J, Candès EJ, Shen Z. A singular value thresholding algorithm for matrix completion. *SIAM Journal on Optimization*. 2010; 20(4):1956–1982. doi: [10.1137/080738970](https://doi.org/10.1137/080738970)
29. Ghasemi H, Malek-Mohammadi M, Babaie-Zadeh M, Jutten C. SRF: Matrix completion based on smoothed rank function. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE; 2011. p. 3672–3675.
30. Malek-Mohammadi M, Babaie-Zadeh M, Amini A, Jutten C. Recovery of low-rank matrices under affine constraints via a smoothed rank function. *IEEE Transactions on Signal Processing*. 2014; 62(4):981–992. doi: [10.1109/TSP.2013.2295557](https://doi.org/10.1109/TSP.2013.2295557)
31. Mohimani H, Babaie-Zadeh M, Jutten C. A fast approach for overcomplete sparse decomposition based on smoothed ℓ_0 norm. *IEEE Transactions on Signal Processing*. 2009; 57(1):289–301. doi: [10.1109/TSP.2008.2007606](https://doi.org/10.1109/TSP.2008.2007606)
32. Deng LJ, Huang TZ, Zhao XL. Exemplar-based image inpainting using a modified priority definition. *Plos One*. 2015; 10(10):e0141199. doi: [10.1371/journal.pone.0141199](https://doi.org/10.1371/journal.pone.0141199) PMID: [26492491](https://pubmed.ncbi.nlm.nih.gov/26492491/)
33. Zhao XL, Wang F, Ng MK. A new convex optimization model for multiplicative noise and blur removal. *SIAM Journal on Imaging Sciences*. 2014; 7(1):456–475. doi: [10.1137/13092472X](https://doi.org/10.1137/13092472X)
34. Liu G, Huang TZ, Liu J, Lv XG. Total variation with overlapping group sparsity for image deblurring under impulse noise. *Plos One*. 2015; 10(4):e0122562. doi: [10.1371/journal.pone.0122562](https://doi.org/10.1371/journal.pone.0122562) PMID: [25874860](https://pubmed.ncbi.nlm.nih.gov/25874860/)
35. Chen C, Ng MK, Zhao XL. Alternating direction method of multipliers for nonlinear image restoration problems. *IEEE Transactions on Image Processing*. 2015; 24(1):33–43. doi: [10.1109/TIP.2014.2369953](https://doi.org/10.1109/TIP.2014.2369953) PMID: [25398177](https://pubmed.ncbi.nlm.nih.gov/25398177/)
36. Zhao XL, Wang W, Zeng TY, Huang TZ, Ng MK. Total variation structured total least squares method for image restoration. *SIAM Journal on Scientific Computing*. 2013; 35(6):B1304–B1320. doi: [10.1137/130915406](https://doi.org/10.1137/130915406)
37. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*. 2004; 13(4):600–612. doi: [10.1109/TIP.2003.819861](https://doi.org/10.1109/TIP.2003.819861) PMID: [15376593](https://pubmed.ncbi.nlm.nih.gov/15376593/)
38. Software ℓ_1 -magic. available: <http://www.l1-magic.org>.