


## Article

# Salt and Pepper Noise Removal with Multi-Class Dictionary Learning and L0 Norm Regularizations

Di Guo <sup>1,\*</sup> , Zhangren Tu <sup>1,2</sup>, Jiechao Wang <sup>3</sup>, Min Xiao <sup>2</sup>, Xiaofeng Du <sup>1</sup> and Xiaobo Qu <sup>3</sup>

<sup>1</sup> School of Computer and Information Engineering, Xiamen University of Technology, Xiamen 361024, China; zhangrentu@stu.xmut.edu.cn (Z.T.); xfd@xmut.edu.cn (X.D.)

<sup>2</sup> School of Opto-Electronic and Communication Engineering, Xiamen University of Technology, Xiamen 361024, China; xiaomin\_ced@xmut.edu.cn

<sup>3</sup> Department of Electronic Science, Xiamen University, Xiamen 361005, China; wangjch@stu.xmu.edu.cn (J.W.); quxiaobo@xmu.edu.cn (X.Q.)

\* Correspondence: guodi@xmut.edu.cn, Tel.: +86-592-6291920

Received: 1 December 2018; Accepted: 23 December 2018; Published: 25 December 2018



**Abstract:** Images may be corrupted by salt and pepper impulse noise during image acquisitions or transmissions. Although promising denoising performances have been recently obtained with sparse representations, how to restore high-quality images remains challenging and open. In this work, image sparsity is enhanced with a fast multiclass dictionary learning, and then both the sparsity regularization and robust data fidelity are formulated as minimizations of  $L_0$ - $L_0$  norms for salt and pepper impulse noise removal. Additionally, a numerical algorithm of modified alternating direction minimization is derived to solve the proposed denoising model. Experimental results demonstrate that the proposed method outperforms the compared state-of-the-art ones on preserving image details and achieving higher objective evaluation criteria.

**Keywords:** impulse noise removal; sparse representation; dictionary learning;  $L_0$  norm

## 1. Introduction

Images may be corrupted by the salt and pepper noise when they are acquired by imperfect sensors or transmitted in unideal channels [1]. This noise introduces the sharp and sudden disturbances in the images and the noise value usually equals to either the minimal or maximal pixel value. To remove the impulse noise, traditional methods include spatial domain approaches such as median [1] or adaptive median filtering (AMF) [2], and transform domain approaches such as wavelet denoising [3,4]. The former ones try to distinguish noise from meaningful image structures with some pre-defined filtering, owing fast computation but limited performance. The latter ones explore the image sparsity in transform domains, e.g., wavelets and contourlets, to distinguish noise but has to carefully choose appropriate basis functions to represent different image features. More comparisons can be found in a recent review on impulse noise removal [5]. Recently, the transform domain sparsity is incorporated into sparse image reconstruction models [6–8] to significantly improve the image quality. The improvement on denoising, however, is still unsatisfactory since the sparsity is limited by pre-defined dictionaries or transforms which may not capture different image structures [8–11].

The state-of-the-art approaches introduce the adaptive dictionary learning [8–11] to provide sparser representations, leading to boosted performance for salt and pepper noise removal. However, the commonly used redundant dictionary learning algorithm, K-SVD [12], is relatively time consuming [13], which may slow down the iterative image reconstruction [8,9,14,15] or lose optimal sparsity when partial image patches are used for fast training [12]. By exploring the intrinsic self-similarity in images, patch-based nonlocal operator (PANO) [16], which is originated from

the well-known block-matching and 3D filtering [17], has been shown to provide adaptive sparse representations and achieve high quality image reconstructions in impulse noise removal [10,11,16]. Yet, PANO may be suboptimal since similar patches are sparsified with pre-defined wavelets. Therefore, it is highly expecting to find new dictionary (or transform) that enables adaptive sparse representation and fast computation while performs well in impulse noise removal.

Even with a proper dictionary, the sparse image reconstruction model for impulse noise removal should be carefully defined. Generally, a restored image is obtained by minimizing a model that balance the sparsity and data fidelity [10,11,18]. For instance, as a convex approximation of the sparsity, the commonly used  $L_1$  norm may bias to penalize small sparse coefficients than large ones, and, thus, suffers from losing weak signal details in the reconstruction problems, such as image reconstruction [19–21], biological spectrum recovery [22–24], and fault detection [25]. Therefore, it is meaningful to incorporate other norms to restore more weak signals.

Beyond the sparsity constraint in the sparse image reconstruction models [8–11,16], the data fidelity, which models the impulse noise with the  $L_1$  norm, may not be appropriate. It was recently found that the  $L_0$  norm, that model the impulse noise, can be interpreted from the Bayesian view of maximum a posteriori probability [18] and can further improve the denoising performance.

To overcome these limitations of current approaches, in this paper, salt and pepper noise removal is improved from two aspects: (1) sparser representation with fast transform; and (2) better regularization of the sparsity and data fidelity. Accordingly, three contributions are made: (1) adaptive sparse image representations with fast orthogonal dictionary learning will be introduced to improve the denoising performance with updated reference images; (2)  $L_0$  norms are introduced to regularize not only sparsity but also the data fidelity, resulting in stronger sparse constraint for images and robustness to outliers introduced by salt and pepper noise; and (3) the impulse denoising model is solved with a feasible numerical algorithm.

The rest of this paper is organized as follows: The typical  $L_1$ - $L_1$  regularization model is reviewed in Section 2.1. The proposed method is presented in Section 2.2. Experimental results are analyzed in Section 3. Discussions are made in Section 4. Finally, Section 5 presents the conclusions.

## 2. Methods

### 2.1. Typical Sparsity-Based Impulse Noise Removal Method

The target image is usually assumed to be sparse under a certain dictionary or transform representation. The sparsity is usually indicated by measuring the  $L_1$  norm of the representation coefficients [6–8] and this norm can be used as feature selection criteria in the sparse representation-based image fusion [26–28]. In addition, the salt and pepper noise are often treated as the outlier in images and the  $L_1$  norm is employed to constraint the data consistency [29]. Thus, a  $L_1$ - $L_1$  regularization model is defined to remove impulse noise according to:

$$\min_x \|\Psi x\|_1 + \lambda \|y - x\|_1 \quad (1)$$

where  $\lambda$  balances the sparsity of an image  $x$  (under the transform  $\Psi$ ) and the data fidelity that is robust to outliers in the noisy image  $y$ .

How to sparsify the image with a proper  $\Psi$  heavily affects the denoising results. A typical choice is a fixed sparse representation, e.g. the finite difference, a basic form of total variations (TV) [18,30] that models the piece-wise constant signals, and wavelets which characterize the piecewise smooth signals [6]. To better captures the image features, a dictionary [8,9] or transform [10,11] may be adaptively trained from the noisy image itself. Unlike the typical dictionary learning, which is time consuming in the iterative image reconstruction [8,9], the recently-proposed PANO not only saves the training time (only several seconds) but also provides adaptively sparse representation to the image by learning the self-similarity [11]. Similar patches are grouped into 3D cubes and then sparsified with 3D wavelets. For removing salt and pepper noise, PANO significantly improved the denoising

performance [11] and obtained better results with impulse detection [10]. With PANO, Equation (1) is turned into:

$$\min_x \sum_{j=1}^J \|A_j x\|_1 + \lambda \|W(y - x)\|_1 \quad (2)$$

where  $\sum_{j=1}^J \|A_j x\|_1$  promotes the sparsity on  $J$  groups of similar patches and  $W$  is a diagonal matrix whose entries stands for weights on pixels. A reasonable weight penalizes the noisy pixels much heavier than the noise-free ones [10]. It is worth noting that the penalty function  $\sum_{j=1}^J \|A_j x\|_1$  promotes the intrinsic group sparsity [31], which is not only important in image denoising [32,33], but also in other applications [34].

Although the model in Equation (2) has shown promising performance in salt and pepper noise removal, it has two limitations: (1) sparsity is insufficient since 3D wavelets in PANO are still pre-defined basis; and (2) the  $L_1$  norm term is only approximation of the sparsity which may lose weak image structures, e.g., small edges, in the reconstruction [19–21] or lack robustness to impulse noise [18].

## 2.2. Proposed Method

In this work, we propose an approach to remove the impulse noise with fast adaptive dictionary learning to provide sparsity of images and formulate the denoising problems with the  $L_0$  norms regularizations.

The flowchart of the whole scheme is summarized in Figure 1. First, a reference image is initially obtained from a noisy image using the AMF. Then, the geometric directions are learnt from the reference image to get the adaptive sparse transform is via fast dictionary training. Next, a denoised image is reconstructed using the proposed  $L_0$  norms regularization model.

In the following, the essential part of the approach will be given in more details.

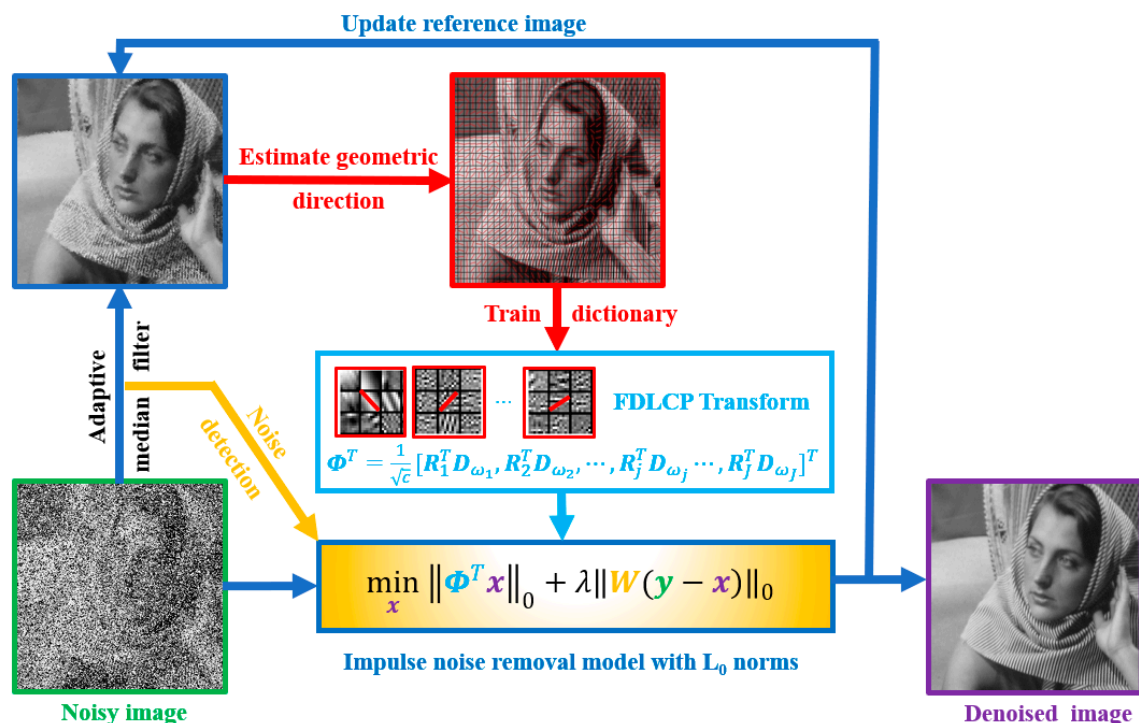
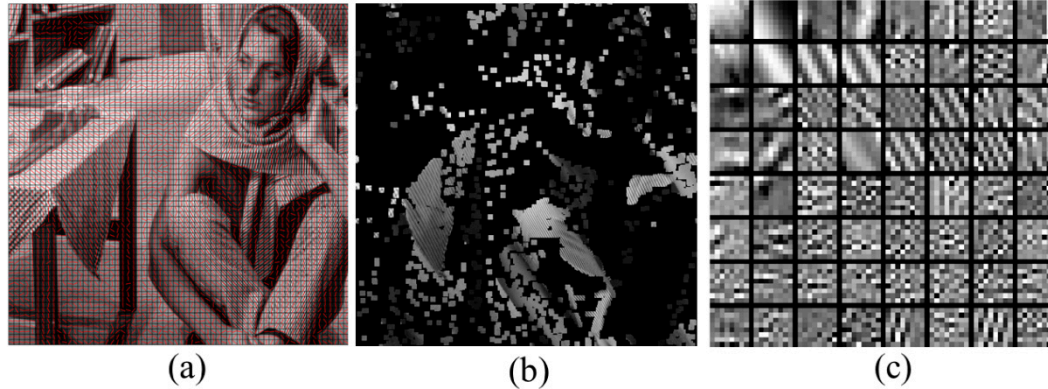


Figure 1. A flowchart of the proposed method on impulse noise removal.

### 2.2.1. Adaptive Dictionary Learning in Salt and Pepper Noise Removal

We first introduce the Fast Dictionaries Learning Method on Classified Patches (FDLCP) [13] into salt and pepper noise removal. FDLCP not only inherits fast learning to form orthogonal

dictionaries, but also provides sparser representations by training comprehensive dictionaries for different geometrical image patches. As shown in Figure 2, patches that share the same geometric direction are grouped together to form one class, and then this class is used to learn a specific dictionary. Therefore, the dictionary of each class effectively captures the underlying image structures (Figure 2c).



**Figure 2.** Adaptive representations learnt with FDLCP. (a) Geometric directions estimated from the Barbara image, (b) one class of image patches that share the same geometric direction, (c) one dictionary learnt from the class of patches in (b). Note: red lines in (a) indicate geometric directions of image patches.

Mathematically, within a class of patches  $X_\omega$  that shares a same geometric direction  $\omega$ , an orthogonal dictionary is trained by solving the optimization problem [15]:

$$\min_{D_\omega, Z} \|X_\omega - D_\omega Z\|_F^2 + \eta^2 \|Z\|_0 \quad s.t. \quad D_\omega^T D_\omega = I \quad (3)$$

where  $D_\omega$  is the dictionary and  $\|\cdot\|_F$  is the Frobenius norm. We choose the orthogonal dictionary learning because it enables the significant reduction of computation [13,15] than the commonly redundant dictionary trained in the K-SVD. For example, it has been shown that the computational time of learning orthogonal dictionary will be approximately 1% than that of the redundant dictionary in K-SVD [13].

Equation (3) can be fast solved by alternatively computing the sparse coding  $Z$  and updating the dictionary  $D_\omega$  in each iteration [13,15] as follows:

- (1) Given the current dictionary  $D_\omega$ , update sparse coefficients  $Z$  via hard thresholding:

$$Z = H_c(D_\omega^T X_\omega) \quad (4)$$

where the hard threshold  $H_c(Q)$  for a matrix  $Q$  is an element-wise operator that performs on the element  $Q_{ij}$  according to  $H_c(Q_{ij}) = \begin{cases} 0, & \text{if } \text{abs}(Q_{ij}) \leq c \\ Q_{ij}, & \text{if } \text{abs}(Q_{ij}) > c \end{cases}$  with a threshold  $c$ .

- (2) Given the coefficients  $Z$ , compute dictionary  $D_\omega$  according to:

$$\min_{D_\omega} \|X_\omega - D_\omega Z\|_F^2 \quad s.t. \quad D_\omega^T D_\omega = I, \quad (5)$$

where  $\eta$  is a parameter to decide the sparsity and  $I$  is an identity matrix. The solution of Equation (5) is:

$$D_\omega = PV^T, \quad (6)$$

where  $P$  and  $V$  are orthogonal matrices of the following singular value decomposition:

$$X_\omega Z^T = P\Lambda V^T, \quad (7)$$

where  $T$  denotes the matrix transpose. The solutions in Equations (7) and (8) are called the orthogonal Procrustes solutions, which were first considered in dictionary learning before [35]. In our implementation, the parameter  $\eta$  is set to 0.2 in all experiments.

A good feature of the FDLCP [15] is that it balances the fast learning via orthogonal dictionary, comparing with the relatively slow learning in the classic K-SVD [12], and sparse representation for each classes, comparing with the sole orthogonal dictionary for all image patches [13]. FDLCP has been observed to outperform K-SVD in the medical image reconstruction [15], but not yet applied to impulse denoise so far.

To further derive the mathematical property of FDLCP, let  $R_j$  denote the operator that extract a patch owned with  $j$ th geometric direction from an image  $x$ , and then the FDLCP transform is simplified as  $\Psi^T = [R_1^T D_{\omega_1}, R_2^T D_{\omega_2}, \dots, R_J^T D_{\omega_J}]^T$  for the total  $J$  geometric directions. The forward and inverse transform satisfies:

$$\Psi^T \Psi = [R_1^T D_{\omega_1}, R_2^T D_{\omega_2}, \dots, R_J^T D_{\omega_J}]^T [R_1^T D_{\omega_1}, R_2^T D_{\omega_2}, \dots, R_J^T D_{\omega_J}] = cI, \quad (8)$$

where  $I$  is an identity matrix and the variable  $c$  is the overlapping factor for image patches [15]. By further setting  $\Phi^T = \frac{1}{\sqrt{c}} \Psi^T$ , the transform  $\Phi$  obeys:

$$\Phi \Phi^T = I, \quad (9)$$

meaning that the modified FDLCP transform  $\Phi$  is a tight frame for image reconstruction [36].

Then, how to train the representations with FDLCP in the salt and pepper noise removal? We suggest learning the FDLCP from a reasonable pre-reconstructed image, e.g., which is obtained with AMF. This strategy has been found to achieve sparse representations under fast computing not only in impulse noise removal [10,11], but also other image reconstruction problems [15,16,37].

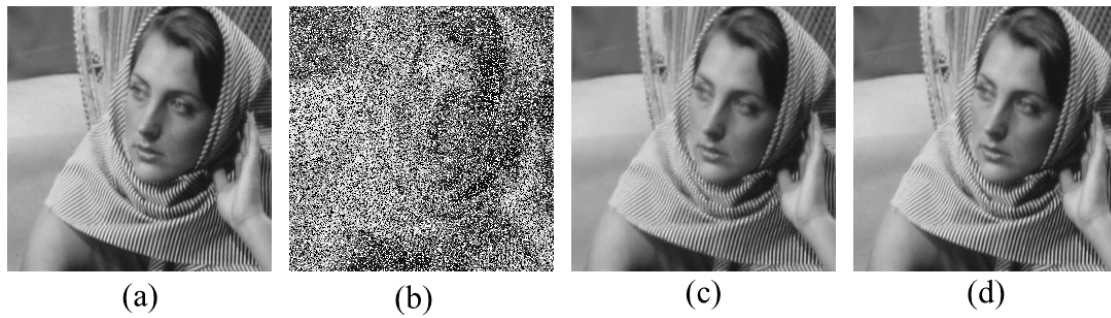
## 2.2.2. FLCLP-Based Image Reconstruction with The $L_0$ Norm Regularizations

Two  $L_1$  norm terms are defined in Equation (2) for impulse noise removal, and they play different roles and can be replaced by two  $L_0$  norms for a better reconstruction. The first term enforces the image sparsity, and the  $L_1$  norm may lose weak signals in the reconstruction, thus, researchers tend to minimize the sparsity by the  $L_0$  norm. This modification has been observed to improve the reconstruction [8,15,20,21] including impulse denoise removal [18]. The second term maintains the data fidelity, and  $L_1$  norm may reduce robustness to impulse noise [6,29,30], while using the  $L_0$  norm has been observed to significantly improve rejecting the outliers and good edge preserving in the reconstruction [18]. Therefore, both the  $L_0$  norm terms on the sparsity and data fidelity are adopted in this work.

The FLCLP-based image reconstruction with the  $L_0$  norm regularizations is proposed as follows:

$$\min_x \|\Phi^T x\|_0 + \lambda \|W(y - x)\|_0 \quad (10)$$

where the  $\|\alpha\|_0$  is the total number of nonzero entries in the  $\alpha$ . For simplicity, we name the proposed method as FDLCP- $L_0$  in the following descriptions. The two  $L_0$  norms significantly boost the FDLCP-based impulse noise removal as shown in Figure 3.



**Figure 3.** Comparisons on the denoised Barbara images using different norms. (a) is the noise-free image, (b) is the noisy image, (c) and (d) are denoised images using FDLCP with  $L_1$  and  $L_0$  norm minimizations, respectively. Note: The noise level is 0.5, meaning that 50% pixels are contaminated by salt and pepper noise.

To better understand the model, we will illustrate  $L_0$  norm regularizations from the Bayesian view.

Given the observed noisy image  $\mathbf{y}$ , the maximize posterior probability for the reconstructed image  $\mathbf{x}$  can be obtained by maximize posterior probability:

$$\max_{\mathbf{x}} \{\log p(\mathbf{y}|\mathbf{x}) + \log p(\mathbf{x})\}. \quad (11)$$

Since a given  $\mathbf{x}$  corresponds to the unique transform coefficients  $\Phi^T \mathbf{x}$ , then  $p(\mathbf{y}|\Phi^T \mathbf{x}) = p(\mathbf{y}|\mathbf{x})$ . By plugging in the salt and pepper noise model:

$$\mathbf{y}_j = \begin{cases} 0, & \text{with probability } \frac{q}{2} \\ 1, & \text{with probability } \frac{q}{2} \\ x_j, & \text{with probability } 1 - q \end{cases} \quad (12)$$

where  $q$  is the probability that  $j$ th pixel is contaminated by noise, then the first term in Equation (11) is modified as [18]:

$$p(\mathbf{y}|\mathbf{x}) = e^{1-q} = 1 - \frac{\|\mathbf{y} - \mathbf{x}\|_0}{N} \quad (13)$$

Therefore, ignoring the constant in Equation (13), maximizing the posterior probability of  $\mathbf{x}$  means minimizing  $\|\mathbf{y} - \mathbf{x}\|_0$ , which is the same as our data consistency term in Equation (10) by setting  $\mathbf{W}$  as the identity matrix.

We then use a transform domain prior as  $p(\mathbf{x}) = \frac{1}{v} e^{-g[\Phi^T \mathbf{x}]}$  where  $\boldsymbol{\alpha} = \Phi^T \mathbf{x}$  and  $g(\boldsymbol{\alpha}_j) = \begin{cases} 1 - h, & \text{if } \boldsymbol{\alpha}_j = 0 \\ h, & \text{otherwise} \end{cases}$  and  $h$  is a probability of a coefficient being in nonzero value. Then, maximizing the log function of  $p(\mathbf{x})$  leads to minimizing  $\|\Phi^T \mathbf{x}\|_0$  which is the same as our sparsity constraint term.

The above analysis implies that the  $L_0$  norm in data consistency specifically works for impulse noise removal, and the  $L_0$  norm that directly counts the number of non-zero coefficients in sparsity constraint evaluates the sparsity better the  $L_1$  norm.

### 2.2.3. Numerical Algorithm

Directly solving Equation (10) is very hard since its  $L_0$  norm terms are non-smooth and non-differential. In this work, we adopt the alternating direction minimization with continuation algorithm (ADMC) [10,16,37] to solve Equation (10). ADCM is chosen since it enables to obtain the solution with much easier sub-problems that have analytical solutions as discussed below.

By introducing the continuation parameter  $\beta$  and augmented variables  $\alpha$  and  $d$ , a relax form of Equation (10) is:

$$\min_{\alpha, d, x} \left( \|\alpha\|_0 + \frac{\beta}{2} \|\alpha - \Phi^T x\|_2^2 \right) + \lambda \left( \|d\|_0 + \frac{\beta}{2} \|d - W(y - x)\|_2^2 \right) \quad (14)$$

where  $\alpha$  and  $d$  are two augmented variables. When  $\beta \rightarrow \infty$ , the solution of Equation (14) approaches that of Equation (10) since any non-zero values of  $\|\alpha - \Phi^T x\|_2$  (or  $\|d - W(y - x)\|_2$ ) will lead  $\frac{\beta}{2} \|\alpha - \Phi^T x\|_2^2$  (or  $\frac{\beta}{2} \|d - W(y - x)\|_2^2$ ) to be infinite. Therefore, in order to minimize the cost function of Equation (14), both  $\alpha = \Phi^T x$  and  $d = W(y - x)$  must be simultaneously satisfied, meaning that Equation (14) is equivalent to Equation (10) when  $\beta \rightarrow \infty$ . In the implementation, we gradually increase  $\beta$ .

While  $\beta$  is fixed at the  $k$ th iterations, the solution of Equation (14) is obtained by alternatively solving the following sub-problems where each one has the analytical solution.

(1) Fixed  $d^{(k)}$  and  $x^{(k)}$ , solve:

$$\alpha^{(k+1)} = \operatorname{argmin}_{\alpha} \|\alpha\|_0 + \frac{\beta}{2} \|\alpha - \Phi^T x^{(k)}\|_2^2, \quad (15)$$

whose solution is achieved by hard thresholding according to:

$$\alpha^{(k+1)} = H_{\sqrt{\frac{2}{\beta}}}(\Phi^T x). \quad (16)$$

(2) Fix  $x^{(k)}$  and  $\alpha^{(k+1)}$ , solve:

$$d^{(k+1)} = \operatorname{argmin}_d \|d\|_0 + \frac{\beta}{2} \|d - W(y - x^{(k)})\|_2^2, \quad (17)$$

whose solution is also achieved by hard thresholding as follows

$$d^{(k+1)} = H_{\sqrt{\frac{2}{\beta}}}(W(y - x^{(k)})) \quad (18)$$

(3) Fix  $\alpha^{(k+1)}$  and  $d^{(k+1)}$ , solve:

$$x^{(k+1)} = \operatorname{argmin}_x \|\alpha^{(k+1)} - \Phi^T x\|_2^2 + \lambda \|d^{(k+1)} - W(y - x)\|_2^2, \quad (19)$$

and have the solution as follows:

$$x^{(k+1)} = (I + \lambda W^T W)^{-1} (\Phi \alpha^{(k+1)} + \lambda W^T (W y - d^{(k+1)})). \quad (20)$$

The algorithm is summarized in Algorithm 1. It is worth noting that the existing ADMC algorithms solve either a reconstruction model with minimization of  $L_1$  norm (sparsity regularization) and  $L_2$  norm (data fidelity) in compressive image recovery [16,38,39], or a denoising model with minimization of  $L_1$  norm (sparsity regularization) and  $L_1$  norm (data fidelity) for impulse noise removal [10,11]. The ADMC algorithm derived in the work solves the denoising model with minimization of  $L_0$  norm (sparsity regularization) and  $L_0$  norm (data consistency) as Equation (14) and has totally different sub-problems and solutions from previous ones.

**Algorithm 1.** ADMC Algorithm for the  $L_0$  Norm Minimizations.

Input: The noise image  $\mathbf{y}$ , diagonal matrix  $\mathbf{W}$ , adaptive dictionaries  $\Phi$ , and the regularization  $\lambda$

Initialize:  $\beta = 1$ ,  $\beta_{max} = 2^{10}$ ,  $k = 1$ , and  $\mathbf{x}^{(1)} = \mathbf{y}$ .

Main:

While  $\beta \leq \beta_{max}$ , repeat steps (a)~(d) until convergence:

(a) Update  $\mathbf{a}^{(k+1)}$  by computing  $\mathbf{a}^{(k+1)} = H_{\sqrt{\frac{2}{\beta}}}(\Phi^T \mathbf{x}^{(k)})$  ( $H_c(\cdot)$  denotes the hard thresholding operator with a threshold  $c$  [20].) according to Equation (16).

(b) Update  $\mathbf{d}^{(k+1)}$  by solving  $\mathbf{d}^{(k+1)} = H_{\sqrt{\frac{2}{\beta}}}(\mathbf{W}(\mathbf{y} - \mathbf{x}^{(k)}))$  according to Equation (18).

(c) Compute  $\mathbf{x}^{(k+1)}$  via Equation (20).

(d) Evaluate the difference of successive reconstruction  $\xi = \frac{\|\mathbf{x}^{(k+1)} - \mathbf{x}^{(k)}\|_2}{\|\mathbf{x}^{(k)}\|_2}$ ,

(e) If  $\xi > 10^{-4}$ , go to (a); else set  $\beta \leftarrow 2\beta$  and  $k \leftarrow k + 1$ .

End While

Output: Reconstructed image  $\mathbf{x}$ .

**3. Results**

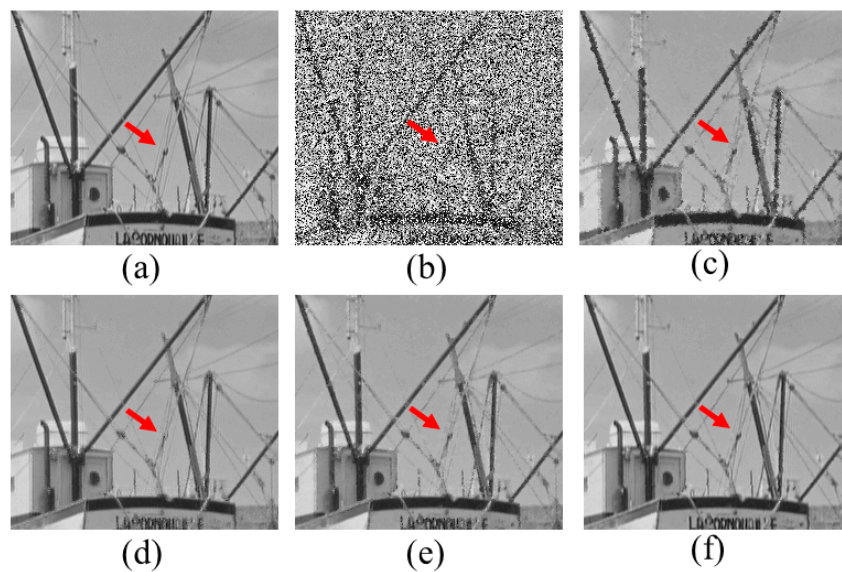
The proposed method, FDLCP- $L_0$ , is compared with basic AMF [2], the state-of-the-art PANO- $L_1$  [9] and total variation (TV)- $L_0$  methods [17]. Two objective evaluation criteria, the peak signal-to-noise ratio (PSNR) and the mean measure of structural similarity (MSSIM) [40], which are most widely used in image denoising and reconstruction [15,17,37,41,42], are adopted to quantitatively measure the denoising performance here. The PSNR measures average pixel distortion of the denoised image and the MSSIM specifically cares about the preserved image structures, e.g., local luminance and contrast, which are important for human visual systems [40]. Higher PSNR and MSSIM mean better denoising performance. All the computations were performed on four cores 3.6 GHz CPU desktop computer with 16 GB RAM. The typical reconstruction time of a  $512 \times 512$  image with AMF, PANO- $L_1$ , TV- $L_0$ , and FDLCP- $L_0$  are 0.25 s, 55.7 s, 6.74 s, and 473.22 s, respectively. The test images are downloaded from website [43].

**3.1. Denoising Performance under a Fixed Noise Level**

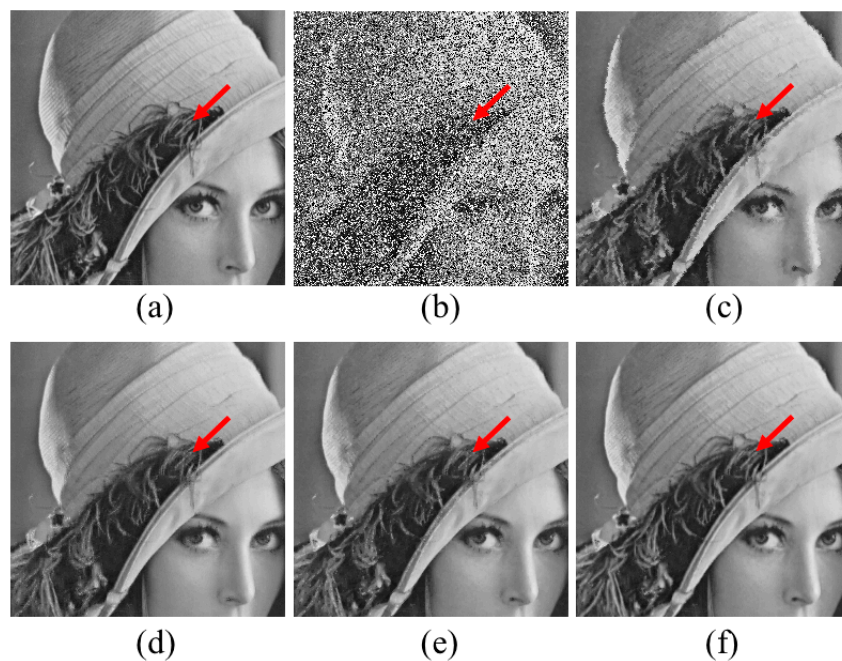
At the typical noise level of 0.5 (50% pixels are contaminated by salt and pepper noise), both PANO- $L_1$  and TV- $L_0$  remove noise much better than the classic AMF. However, the two state-of-the-art approaches still lose some straight lines (Figure 4d,e) or weaken some textures (Figure 5d,e). The proposed method reconstructs these image features much better than others. As listed in Table 1, evaluation criteria including both PSNR and MSSIM indicate that the proposed method reconstructs the most consistent images to the noise-free ones.

**Table 1.** Quantitative measures at the salt and pepper impulse noise level of 0.5.

Images	Quantitative Measure	Methods			
		AMF	PANO- $L_1$	TV- $L_0$	FDLCP- $L_0$
Boat	PSNR	36.32	37.95	38.29	<b>40.26</b>
	MSSIM	0.8715	0.9375	0.9307	<b>0.9631</b>
Lena	PSNR	37.42	38.87	39.48	<b>41.01</b>
	MSSIM	0.8878	0.9360	0.9382	<b>0.9531</b>
Barbara	PSNR	34.61	36.28	35.63	<b>39.44</b>
	MSSIM	0.8080	0.9293	0.8746	<b>0.9644</b>



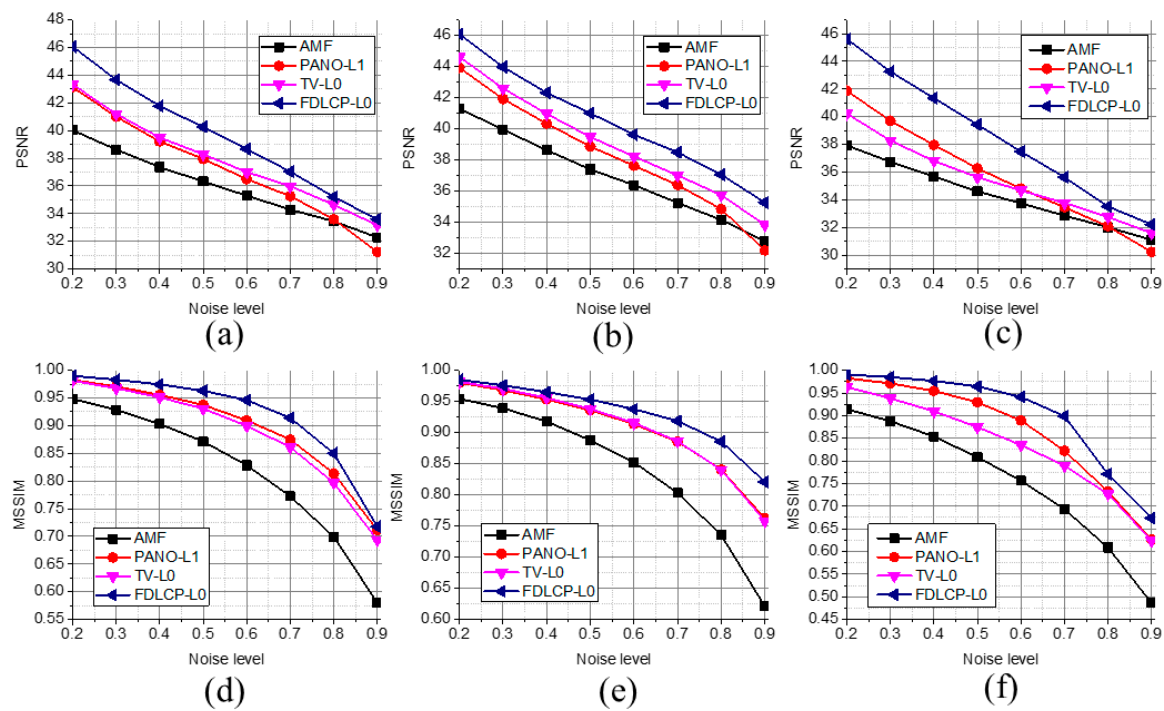
**Figure 4.** Denoised boat images at the noise level of 0.5. (a) is the noise-free image, (b) is the noisy image, (c)–(f) are denoised images using AMF, PANO- $L_1$ , TV- $L_0$ , and the proposed FDLCP- $L_0$ , respectively.



**Figure 5.** Denoised Lena images at the noise level of 0.5. (a) is the noise-free image, (b) is the noisy image, (c)–(f) are denoised images using AMF, PANO- $L_1$ , TV- $L_0$ , and the proposed FDLCP- $L_0$ , respectively.

### 3.2. Denoising Performance under Different Noise Levels

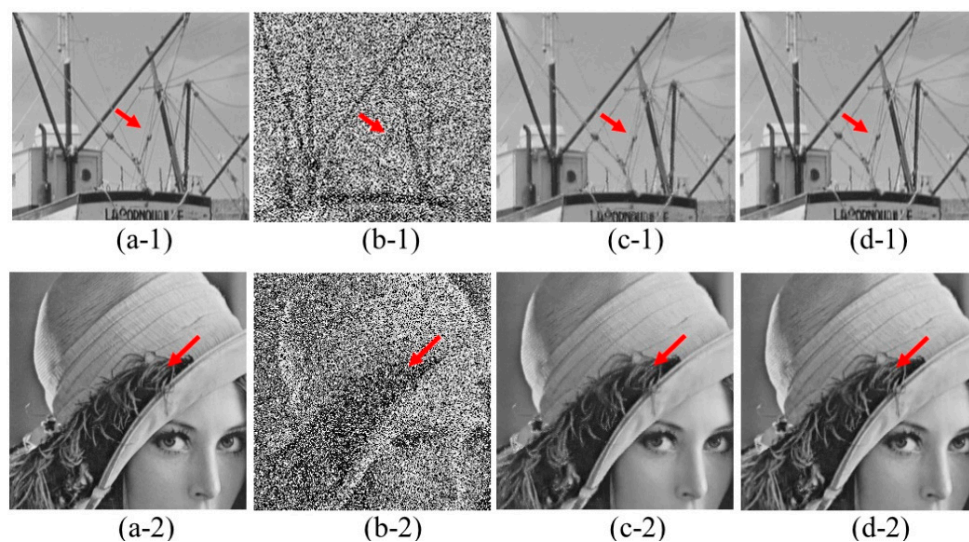
The denoising performances under different noise levels were evaluated in Figure 6. The AMF is inferior to other compared methods. The TV- $L_0$  outperforms the PANO- $L_1$  although PANO provides adaptive sparse representation. This observation implies that the  $L_0$  norm is more robust than the  $L_1$  norm to the outliers introduced by the impulse noise. The proposed method not only makes use of adaptive sparsity in the representation but also incorporates the robust  $L_0$  norm to better remove outliers in the data fidelity term. Therefore, improved denoising performance by the proposed method are consistently observed in all experiments with different noise levels.



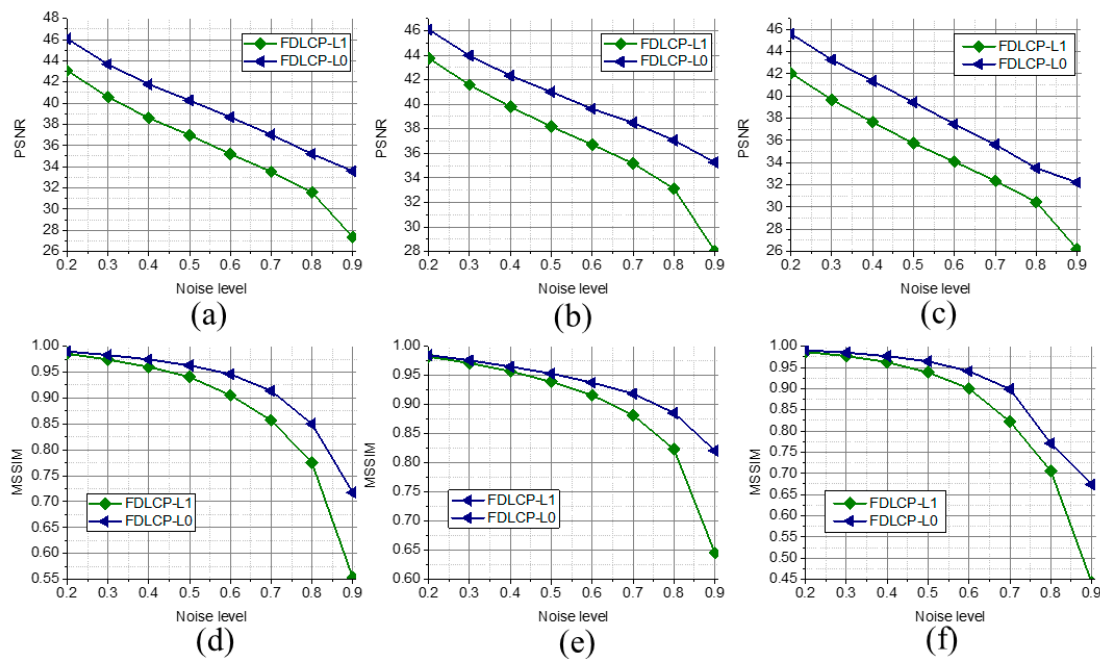
**Figure 6.** Quantitative measure at different noise levels. (a)–(c) are PSNRs of *Boat*, *Lena*, and *Barbara* images, respectively; (d)–(f) are MSSIMs of *Boat*, *Lena*, and *Barbara* images, respectively.

#### 4. Discussions

To further confirm the advantage of using the  $L_0$  norm over the  $L_1$  norm, FDLCP- $L_1$  is compared with the proposed method FDLCP- $L_0$ . Typical denoised images are shown in Figure 7 and quantitative measures are shown Figure 8, Tables 2 and 3. Image features are preserved much better using FDLCP- $L_0$  than that using FDLCP- $L_1$ . Consistent improvements at all noise levels are observed for FDLCP- $L_0$ . Therefore, replacing the  $L_1$  norm with the  $L_0$  norm is valuable to improve the quality of the images that are contaminated by the salt and pepper noise.



**Figure 7.** Comparisons on denoised images using FDLCP- $L_1$  and FDLCP- $L_0$ . (a) and (b) are noise-free and noisy images, (c) and (d) are denoised images using FDLCP- $L_1$  and FDLCP- $L_0$ , respectively. The first row and the second row are *Boat* and *Lena* images. Note: The salt and pepper noise level is 0.5 and the denoised *Barbara* images is shown in Figure 3.



**Figure 8.** Comparisons on quantitative measure using FDLCP-L<sub>1</sub> and FDLCP-L<sub>0</sub>. (a)–(c) are PSNRs of *Boat*, *Lena*, and *Barbara* images, respectively; (d)–(f) are MSSIMs of *Boat*, *Lena*, and *Barbara* images, respectively.

**Table 2.** PSNR performance under the salt and pepper noise using FDCLP with L<sub>1</sub> or L<sub>0</sub> norms.

Images	Methods	Noise Level							
		0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>Boat</i>	L <sub>1</sub>	43.07	40.58	38.60	36.99	35.20	33.54	31.58	27.35
	L <sub>0</sub>	46.07	43.68	41.78	40.26	38.67	37.04	35.21	33.59
<i>Lena</i>	L <sub>1</sub>	43.75	41.58	39.80	38.17	36.71	35.18	33.12	28.04
	L <sub>0</sub>	46.09	43.97	42.32	41.01	39.63	38.49	37.07	35.27
<i>Barbara</i>	L <sub>1</sub>	42.06	39.67	37.65	35.74	34.08	32.38	30.45	26.21
	L <sub>0</sub>	45.62	43.27	41.35	39.44	37.48	35.63	33.52	32.22

**Table 3.** MSSIM performance under the salt and pepper noise using FDCLP with L<sub>1</sub> or L<sub>0</sub> norms.

Images	Methods	Noise Level							
		0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
<i>Boat</i>	L <sub>1</sub>	0.9856	0.9743	0.9601	0.9407	0.9055	0.8571	0.7752	0.5537
	L <sub>0</sub>	0.9900	0.9830	0.9745	0.9631	0.9461	0.9139	0.8501	0.7178
<i>Lena</i>	L <sub>1</sub>	0.9822	0.9711	0.9569	0.9391	0.9159	0.8819	0.8232	0.6458
	L <sub>0</sub>	0.9849	0.9759	0.9650	0.9531	0.9376	0.9188	0.8856	0.8208
<i>Barbara</i>	L <sub>1</sub>	0.9867	0.9769	0.9624	0.9385	0.9001	0.8220	0.7058	0.4427
	L <sub>0</sub>	0.9909	0.9849	0.9766	0.9644	0.9412	0.8987	0.7709	0.6736

## 5. Conclusions

A new salt and pepper impulse noise removal method is proposed by simultaneously exploring: (1) The adaptively sparse representation of images; (2) better regularizing the sparsity of representation and the data fidelity under impulse noise. The former is accomplished via fast orthogonal dictionary learning within multiclass geometric image patches while the latter is enforced by regularizing L<sub>0</sub> norms on both the sparsity and data fidelity constraints. Experimental results demonstrate that this

proposed approach outperforms the compared ones in terms of better preserving image structures and higher objective denoising performance evaluation criteria. The theoretical proof of the advantage of  $L_0$  norm over  $L_1$  norm in impulse noise removal would be an interesting future work.

**Author Contributions:** Conceptualization, D.G.; Data curation, Z.T. and J.W.; Formal analysis, Z.T. and J.W.; Funding acquisition, D.G. and X.Q.; Investigation, D.G. and X.D.; Methodology, D.G.; Project administration, D.G.; Resources, M.X.; Software, J.W.; Supervision, X.Q.; Writing—original draft, D.G.; Writing—review & editing, D.G., Z.T., X.D. and X.Q.

**Funding:** This work was supported in part by National Natural Science Foundation of China (No. 61871341, 61672335, and 61601276), the Science and Technology Program of Xiamen (No. 3502Z20183053), the Natural Science Foundation of Fujian Province of China (No. 2016J05205 and 2018J06018), the Fundamental Research Funds for the Central Universities (No. 20720180056).

**Acknowledgments:** Di Guo is grateful to Ming-Ting Sun for hosting the visit at University of Washington when writing this work and China Scholarship Council (No. 201808350010) to support the visit.

**Conflicts of Interest:** The authors declare that there is no conflict of interests regarding the publication of this article.

## References

- Gonzalez, R.C.; Richard, E. *Digital Image Processing*; Prentice Hall Press: Upper Saddle River, NJ, USA, 2002.
- Hwang, H.; Haddad, R. Adaptive median filters: New algorithms and results. *IEEE Trans. Image Process.* **1995**, *4*, 499–502. [[CrossRef](#)] [[PubMed](#)]
- Sree, P.S.J.; Kumar, P.; Siddavatam, R.; Verma, R. Salt-and-pepper noise removal by adaptive median-based lifting filter using second-generation wavelets. *Signal Image Video Process.* **2013**, *7*, 111–118. [[CrossRef](#)]
- Adeli, A.; Tajeripoor, F.; Zomorodian, M.J.; Neshat, M. Comparison of the Fuzzy-based wavelet shrinkage image denoising techniques. *Int. J. Comput. Sci.* **2012**, *9*, 211–216.
- Mafi, M.; Martin, H.; Cabrerizo, M.; Andrian, J.; Barreto, A.; Adjouadi, M. A comprehensive survey on impulse and Gaussian denoising filters for digital images. *Signal Process.* **2018**. [[CrossRef](#)]
- Huang, S.; Zhu, J. Removal of salt-and-pepper noise based on compressed sensing. *Electron. Lett.* **2010**, *46*, 1198–1199. [[CrossRef](#)]
- Wang, X.-L.; Wang, C.-L.; Zhu, J.-B.; Liang, D.-N. Salt-and-pepper noise removal based on image sparse representation. *Opt. Eng.* **2011**, *50*, 097007. [[CrossRef](#)]
- Xiao, Y.; Zeng, T.; Yu, J.; Ng, M.K. Restoration of images corrupted by mixed Gaussian-impulse noise via  $l1-l0$  minimization. *Pattern Recognit.* **2011**, *44*, 1708–1720. [[CrossRef](#)]
- Wang, S.; Liu, Q.; Xia, Y.; Dong, P.; Luo, J.; Huang, Q.; Feng, D.D. Dictionary learning based impulse noise removal via  $L1-L1$  minimization. *Signal Process.* **2013**, *93*, 2696–2708. [[CrossRef](#)]
- Guo, D.; Qu, X.; Du, X.; Wu, K.; Chen, X. Salt and pepper noise removal with noise detection and a patch-based sparse representation. *Adv. Multimed.* **2014**, *2014*. [[CrossRef](#)]
- Guo, D.; Qu, X.; Wu, M.; Wu, K. A modified iterative alternating direction minimization algorithm for impulse noise removal in images. *J. Appl. Math.* **2014**, *2014*, 595782. [[CrossRef](#)]
- Aharon, M.; Elad, M.; Bruckstein, A. K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Trans. Signal Process.* **2006**, *54*, 4311–4322. [[CrossRef](#)]
- Cai, J.-F.; Ji, H.; Shen, Z.; Ye, G.-B. Data-driven tight frame construction and image denoising. *Appl. Comput. Harmonic Anal.* **2014**, *37*, 89–105. [[CrossRef](#)]
- Ravishanker, S.; Bresler, Y. MR image reconstruction from highly undersampled k-space data by dictionary learning. *IEEE Trans. Med. Imaging* **2011**, *30*, 1028–1041. [[CrossRef](#)] [[PubMed](#)]
- Zhan, Z.; Cai, J.F.; Guo, D.; Liu, Y.; Chen, Z.; Qu, X. Fast multiclass dictionaries learning with geometrical directions in MRI reconstruction. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 1850–1861. [[CrossRef](#)] [[PubMed](#)]
- Qu, X.; Hou, Y.; Lam, F.; Guo, D.; Zhong, J.; Chen, Z. Magnetic resonance image reconstruction from undersampled measurements using a patch-based nonlocal operator. *Med. Image Anal.* **2014**, *18*, 843–856. [[CrossRef](#)] [[PubMed](#)]
- Dabov, K.; Foi, A.; Katkovnik, V.; Egiazarian, K. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Trans. Image Process.* **2007**, *16*, 2080–2095. [[CrossRef](#)] [[PubMed](#)]

18. Ganzhao, Y.; Ghanem, B. L0TV: A new method for image restoration in the presence of impulse noise. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition-CVPR 2015, Boston, MA, USA, 7–12 June 2015; pp. 5369–5377.
19. Trzasko, J.; Manduca, A. Highly undersampled magnetic resonance image reconstruction via homotopic L0-minimization. *IEEE Trans. Med. Imaging* **2009**, *28*, 106–121. [[CrossRef](#)]
20. Ning, B.; Qu, X.; Guo, D.; Hu, C.; Chen, Z. Magnetic resonance image reconstruction using trained geometric directions in 2D redundant wavelets domain and non-convex optimization. *Magn. Reson. Imaging* **2013**, *31*, 1611–1622. [[CrossRef](#)]
21. Qu, X.; Cao, X.; Guo, D.; Hu, C.; Chen, Z. Compressed sensing MRI with combined sparsifying transforms and smoothed L0 norm minimization. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing-ICASSP 2010, Dallas, TX, USA, 14–19 March 2010; pp. 626–629.
22. Qu, X.; Qiu, T.; Guo, D.; Lu, H.; Ying, J.; Shen, M.; Hu, B.; Orekhov, V.; Chen, Z. High-fidelity spectroscopy reconstruction in accelerated NMR. *Chem. Commun.* **2018**, *54*, 10958–10961. [[CrossRef](#)]
23. Qu, X.; Guo, D.; Cao, X.; Cai, S.; Chen, Z. Reconstruction of self-sparse 2D NMR spectra from undersampled data in the indirect dimension. *Sensors* **2011**, *11*, 8888–8909. [[CrossRef](#)]
24. Qu, X.; Cao, X.; Guo, D.; Chen, Z. Compressed sensing for sparse magnetic resonance spectroscopy. In Proceedings of the 18th Scientific Meeting on International Society for Magnetic Resonance in Medicine-ISMIRM 2010, Stockholm, Sweden, 1–7 May 2010; p. 3371.
25. Li, Q.; Liang, S. Weak fault detection of tapered rolling bearing based on penalty regularization approach. *Algorithms* **2018**, *11*, 184. [[CrossRef](#)]
26. Zhu, Z.; Yin, H.; Chai, Y.; Li, Y.; Qi, G. A novel multi-modality image fusion method based on image decomposition and sparse representation. *Inf. Sci.* **2018**, *432*, 516–529. [[CrossRef](#)]
27. Zhu, Z.; Qi, G.; Chai, Y.; Chen, Y. A novel multi-focus image fusion method based on stochastic coordinate coding and local density peaks clustering. *Future Internet* **2016**, *8*, 53. [[CrossRef](#)]
28. Qi, G.; Zhu, Z.; Chen, Y.; Wang, J.; Zhang, Q.; Zeng, F. Morphology-based visible-infrared image fusion framework for smart city. *Int. J. Simul. Process Modell.* **2018**, *13*, 523–536. [[CrossRef](#)]
29. Nikolova, M. A variational approach to remove outliers and impulse noise. *J. Math. Imaging Vis.* **2004**, *20*, 99–120. [[CrossRef](#)]
30. Cai, J.-F.; Chan, R.; Nikolova, M. Fast two-phase image deblurring under impulse noise. *J. Math. Imaging Vis.* **2010**, *36*, 46–53. [[CrossRef](#)]
31. Huang, J.; Zhang, T. The benefit of group sparsity. *Ann. Stat.* **2010**, *38*, 1978–2004. [[CrossRef](#)]
32. Wang, L.; Chen, Y.; Lin, F.; Chen, Y.; Yu, F.; Cai, Z. Impulse noise denoising using total variation with overlapping group sparsity and Lp-pseudo-norm shrinkage. *Appl. Sci.* **2018**, *8*, 2317. [[CrossRef](#)]
33. Chen, P.; Selesnick, I.W. Group-sparse signal denoising: Non-convex regularization, convex optimization. *IEEE Trans. Signal Process.* **2014**, *62*, 3464–3478. [[CrossRef](#)]
34. Alpago, D.; Zorzi, M.; Ferrante, A. Identification of sparse reciprocal graphical models. *IEEE Control Syst. Lett.* **2018**, *2*, 659–664. [[CrossRef](#)]
35. Lesage, S.; Gribonval, R.; Bimbot, F.; Benaroya, L. Learning unions of orthonormal bases with thresholded singular value decomposition. In Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing-ICASSP'05, Philadelphia, PA, USA, 18–23 March 2005; Volume 5, pp. 293–296.
36. Liu, Y.; Zhan, Z.; Cai, J.F.; Guo, D.; Chen, Z.; Qu, X. Projected iterative soft-thresholding algorithm for tight frames in compressed sensing magnetic resonance imaging. *IEEE Trans. Med. Imaging* **2016**, *35*, 2130–2140. [[CrossRef](#)] [[PubMed](#)]
37. Lai, Z.; Qu, X.; Liu, Y.; Guo, D.; Ye, J.; Zhan, Z.; Chen, Z. Image reconstruction of compressed sensing MRI using graph-based redundant wavelet transform. *Med. Image Anal.* **2016**, *27*, 93–104. [[CrossRef](#)] [[PubMed](#)]
38. Yang, J.; Zhang, Y.; Yin, W. A fast TVL1-L2 Minimization Algorithm for Signal Reconstruction from Partial Fourier Data. Rice University. 2009. Available online: <ftp://ftp.math.ucla.edu/pub/camreport/cam09-24.pdf> (accessed on 1 October 2018).
39. Yang, J.; Zhang, Y.; Yin, W. A fast alternating direction method for TV l1-l2 signal reconstruction from partial fourier data. *IEEE J. Sel. Top. Signal Process.* **2010**, *4*, 288–297. [[CrossRef](#)]
40. Zhou, W.; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. Image quality assessment: From error visibility to structural similarity. *IEEE Trans. Image Process.* **2004**, *13*, 600–612.

41. Portilla, J.; Strela, V.; Wainwright, M.J.; Simoncelli, E.P. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Trans. Image Process.* **2003**, *12*, 1338–1351. [[CrossRef](#)] [[PubMed](#)]
42. Elad, M.; Aharon, M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Trans. Image Process.* **2006**, *15*, 3736–3745. [[CrossRef](#)] [[PubMed](#)]
43. Computational Imaging and Visual Image Processing. Available online: <http://www.io.csic.es/PagsPers/JPortilla/image-processing/bls-gsm/63-test-images> (accessed on 1 October 2018).



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).