

Hybrid Model and Split Bregman Iteration Algorithm for Image Denoising

Yibin Yu ^a, Jialin Zhang

School of Information Engineering, Wuyi University, Jiangmen 529020, China

^ayuyibintony@163.com

Keywords: hybrid model; difference of convex algorithm; split Bregman iteration; image denoising.

Abstract. Although difference of convex model has attracted many research efforts due to its superior performance for image processing, no attention has focused on robust data fidelity for this model. In this paper, we propose a novel model, which combines the l_1 and l_2 fidelity terms with a weighted difference of anisotropic and isotropic total variation (TV). Since our model takes a new difference form of convex terms, we employ difference of convex algorithm (DCA). In this paper, we adopt split Bregman iteration (SBI) to solve each DCA subproblem of the proposed model. Image denoising verifies the convergence of optimal solution and monotone decreasing of objective function. Experimental results on image denoising demonstrate that the proposed methods outperform other competing methods in terms of quantitative criteria and perceptual quality.

Introduction

Image restoration is an inverse problem in which the observed data f is obtained approximately by applying a linear operator K on the clean data u with additive noise n . Mathematically, a linear invariant imaging system can be expressed as

$$f = Ku + n \quad (1)$$

where $K \in \mathbb{R}^{MN \times MN}$ is the matrix representation of a direct convolution operation for deblurring and K is an identity matrix for denoising. As is common in image processing, we utilize the vector notation for images, where the pixels of an $M \times N$ image are stacked into an MN vector. $u \in \mathbb{R}^{MN \times 1}$ is a vector denoting the unknown image (potentially sharp) with size $M \times N$, $f \in \mathbb{R}^{MN \times 1}$ is a vector representing of the observed image, $n \in \mathbb{R}^{MN \times 1}$ is a vector denoting the noise.

The task of image restoration is to recover u from the observed image f . In most scenarios, solving u from $Ku = f$ is insufficient to get a unique solution. For this reason, we need to introduce a regularization term to constrain the solution to the minimization problem, with the prior knowledge of the observed image. A general model for inverse problem is

$$\min_u \Phi_{\text{reg}}(u) + m\Phi_{\text{fid}}(u, f) \quad (2)$$

where $\Phi_{\text{reg}}(u)$ is a regularization term, $\Phi_{\text{fid}}(u, f)$ is a data fidelity term which measures some type of deviation of u from the observation f , m is a positive parameter which balances these two terms.

Rudin-Osher-Fatemi (ROF) model with TV regularization has the priorities of preserving sharp edges and object boundaries, which was proposed by Rudin [1]. TV regularization can either be the anisotropic or isotropic form as follows:

$$\|u\|_{\text{ani-TV}} = \|Du\|_1 = \|D_x u\|_1 + \|D_y u\|_1 \quad (3a)$$

$$\|u\|_{\text{iso-TV}} = \|Du\|_{2,1} = P \sqrt{|D_x u|^2 + |D_y u|^2} P_1 \quad (3b)$$

where the operators D_x , D_y are the forward finite difference operators along the horizontal and vertical direction, respectively.

In recent years, some researchers proposed nonconvex TV regularization terms include l_p for $p \in (0,1)$ [2] and $l_1 - l_2$ [3]. Lou *et al.* [4] proposed a new regularization method which is a weighted difference of anisotropic and isotropic TV norm, namely, $l_1 - 0.5l_2$. In addition, TV or its variants

regularization terms are widely used in image processing, such as denoising [5, 6], super-resolution [7].

It is well known that the common data fidelity for the gaussian noise is $\Phi_{\text{fid}}(u, f) = \|Ku - f\|_2^2$, which is widely used in various of inverse problems [1, 8]. However, when the images corrupted by salt-and-pepper noise, $\Phi_{\text{fid}}(u, f) = \|Ku - f\|_2^2$ does not yield satisfactory restoration while $\Phi_{\text{fid}}(u, f) = \|Ku - f\|_1$ was suggested in [9], which remove salt-and-pepper noise successfully. Recently, Michael proposed a combined l_1/l_2 data fidelity with total variation model [10], which dealt with the images corrupted simultaneously by the gaussian noise and salt-and-pepper noise very well and preserved details better than the pure l_2 -TV data fidelity model and it does not suffer from a sudden loss of image features like the l_1 -TV model. Note that, we use hybrid data fidelity and combined l_1/l_2 data fidelity interchangeably throughout this paper.

Contributions and Organization. The contributions of this paper are summarized as follows: The authors of [4] proposed a weighted difference of anisotropic and isotropic total variation model, which regularization term approximates the heavy-tailed distribution of gradient. In [10], they proposed a combined l_1 and l_2 data fidelity with TV regularization model, which can preserves details better than l_2 data fidelity with TV. Motivated by them, we propose to combine hybrid data fidelity with the weighted difference TV model.

We apply split Bregman iteration to solve the subproblem of DCA for our model. In addition, we compare our proposed method with other representative methods. The experimental results show that the proposed methods can achieve highly performance.

The remainder of this paper is organized as follows. In Section 2, we give the related work which include weighted difference of anisotropic and isotropic total variation model and hybrid data fidelity model. The proposed model and algorithms are presented in Section 3. Section 4 describes the experiments in detail, including parameter setting, image denoising. A concluding remark is given in Section 5.

Related Work

Weighted Difference of Anisotropic and Isotropic Total Variation Model. It is well known that most of natural images are sparse in gradient domain, l_1 is the best convex relaxation of l_0 to bypass the NP-hard l_0 . l_1 norm on the gradient, called anisotropic TV, will yield “blocky” artifacts. Lou *et al.* [4] proposed a weighted difference of anisotropic and isotropic total variation model for image restoration. The formula of weighted difference of anisotropic and isotropic total variation regularization is presented as following

$$J(u) = \|u\|_{\text{ani-TV}} - a \|u\|_{\text{iso-TV}} + \|D_x u\|_1 = \|D_y u\|_1 - a \|\sqrt{|D_x u|^2 + |D_y u|^2}\|_1 \quad (4)$$

where a is a weighted parameter to balance the anisotropic and isotropic TV terms. Since the gradient distribution follows the heavy-tailed distribution, the authors of [4] derive the value of $a = 0.5$ and it is consistent with the choice of hyper-Laplacian [2] for image processing. It is an effective method especially for piecewise image. In addition, Based on the difference of convex algorithm, Lou *et al.* applied the split Bregman iteration for $l_1 - 0.5l_2$ minimization and prove that it converges to a stationary point satisfying the first order optimality condition. Although the model is a nonconvex function, it satisfies Lipschitz regularity and guarantees convergence via the DCA.

Hybrid Data Fidelity Model. As aforementioned description, the data fidelity term guarantees the consistency between the recovered and the observed image. In general, the type of noise contained in the observed image determines the choice of the data fidelity term. As is known, the l_2 data fidelity

term is appropriate for image contained gaussian noise, while salt-and-pepper noise can be removed successfully via non-smooth l_1 data fidelity term [9].

Michael *et al.* [10] have found that mixed l_1/l_2 data fidelity is superior to l_1 data fidelity or l_2 data fidelity when the regularization term is TV model. Mixed l_1/l_2 data fidelity avoids suffering from a sudden loss of image features, not like l_1 data fidelity and preserves the details better than l_2 data fidelity. Image restoration with mixed l_1/l_2 is the following minimization problem

$$\min_{u \in L^2(\Omega)} J_{a_1, a_2}(u) := j(|Du|)(\Omega) + a_1 \|K_1 u - f_1\|_{L^1(\Omega)} + a_2 \|K_2 u - f_2\|_{L^2(\Omega)}^2 \quad (5)$$

where $K_i : L^2(\Omega) \rightarrow L^2(\Omega)$ is a bounded linear operator, $f_i \in L^2(\Omega)$ is a given datum, $a_i \geq 0$ for $i = 1, 2$, with $a_1 + a_2 > 0$, $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, and $j(|\cdot|)$ is a convex function of measures representing regularization.

3 Proposed Model and Algorithms

Proposed Model. Inspired by the reference [4] and [10], we propose a new effective image restoration model which employs a weighted difference of anisotropic and isotropic total variation with a hybrid data fidelity. Then, the objective function of our model is

$$F(u) := \frac{m}{2} \|Ku - f\|_2^2 + r \|Ku - f\|_1 + \|D_x u\|_1 + \|D_y u\|_1 - 0.5 \|\sqrt{|D_x u|^2 + |D_y u|^2}\| \quad (6)$$

where $m > 0$ and $r > 0$ are coefficients of $\|Ku - f\|_2^2$ and $\|Ku - f\|_1$ respectively.

Since our model belongs to difference of two convex functions form. *DCA* is appropriate when the objective function can be decoupled into difference of two convex functions, therefore, we employ the *DCA* to solve our model. Tao and An proposed *DCA* [11, 12] to minimize the cost function $F(x) = G(x) - H(x)$ on the space \mathbb{R}^n , where $G(x)$ and $H(x)$ are lower semicontinuous proper convex functions. G and H are *DC* components of F .

We need to construct two sequences $\{x^k\}$ and $\{y^k\}$ for *DCA*, the candidates for optimal solutions of primal and dual problems respectively. The *DCA* can be done via iterative computation

$$\begin{cases} y^k \in \partial H(x^k) \\ x^{k+1} = \arg \min_{x \in \mathbb{R}^n} G(x) - (H(x^k) + \langle y^k, x - x^k \rangle) \end{cases} \quad (7)$$

where $y^k \in \partial H(x^k)$ indicates that y^k is a subgradient of $H(x)$ at x^k .

Corresponding difference-of-convex components of our model are

$$\begin{cases} G(u) = \|D_x u\|_1 + \|D_y u\|_1 + c \|u\|_2^2 + \frac{m}{2} \|Ku - f\|_2^2 + r \|Ku - f\|_1, \\ H(u) = 0.5 \|\sqrt{|D_x u|^2 + |D_y u|^2}\| + c \|u\|_2^2 \end{cases} \quad (8)$$

where $c \in \mathbb{R}$ is a positive constant which ensures strong convexity of G and H . After linearizing H term, we have

$$u^{n+1} = \arg \min_u \|D_x u\|_1 + \|D_y u\|_1 + c \|u\|_2^2 + \frac{m}{2} \|Ku - f\|_2^2 + r \|Ku - f\|_1 - 0.5 \langle Du, q^n \rangle - 2c \langle u, u^n \rangle \quad (9)$$

where $q^n = (D_x u^n, D_y u^n) \cdot (|D_x u^n|^2 + |D_y u^n|^2)^{\frac{1}{2}} = (q_x^n, q_y^n)$ at step u^n . Note that q^n is a point-wise calculation. In (9), the $G(u)$ of *DCA* is a l_1 TV minimization problem. This problem can be solved by the split Bregman iteration technique.

In what follows, we briefly review the SBI for the sake of completeness. The SBI is useful when l_1 minimization problem is following form

$$u = \arg \min \|f(u)\|_1 + E(u) \quad (10)$$

where E and f are convex. It can be solved easily when we use alternating method to split the problem into subproblem. Therefore, we have

$$(u, d) = \arg \min \|d\|_1 + E(u), \quad \text{s.t. } f(u) = d \quad (11)$$

It can be simplified as following

$$\begin{cases} u^{k+1} = \arg \min E(u) + \frac{I}{2} \|d^k - f(u) - b^k\|_2^2 \\ d^{k+1} = \arg \min \|d\|_1 + \frac{I}{2} \|d - f(u^{k+1}) - b^k\|_2^2 \\ b^{k+1} = b^k + f(u^{k+1}) - d^{k+1} \end{cases} \quad (12)$$

DCASBI Algorithm. To address (10), we introduce intermediate variables d_x, d_y, d_r as well as auxiliary variables b_x, b_y, b_r , and turn it into an equivalent problem

$$\begin{aligned} u^{n+1} = \arg \min_{u, d_x, d_y, d_r} & \|d_x\|_1 + \|d_y\|_1 + r \|d_r\|_1 + c \|u\|_2^2 - 2c \langle u, u^n \rangle - 0.5(d_x^T \cdot q_x^n + d_y^T \cdot q_y^n) \\ & + \frac{m}{2} \|Ku - f\|_2^2 + \frac{I}{2} \|d - Du - b\|_2^2 + \frac{t}{2} \|d_r - Ku + f - b_r\|_2^2 \end{aligned} \quad (13)$$

where $\|d - Du - b\|_2^2$ can be written as $\|d_x - D_x u - b_x\|_2^2$ plus $\|d_y - D_y u - b_y\|_2^2$. I and t are regularization parameters, which are associated with the quadratic penalty term $\|d - Du - b\|_2^2$ and $\|d_r - Ku + f - b_r\|_2^2$ respectively.

In what follows, we use the alternating minimization to deal with the following sub-problems iteratively:

$$u^{k+1} = \arg \min_u \frac{m}{2} \|Ku - f\|_2^2 + c \|u\|_2^2 - 2c \langle u, u^n \rangle + \frac{I}{2} \|d - Du - b\|_2^2 + \frac{t}{2} \|d_r - Ku + f - b_r\|_2^2 \quad (14)$$

$$\begin{cases} d_x^{k+1} = \arg \min_{d_x} \|d_x\|_1 - 0.5d_x^T \cdot q_x^n + \frac{I}{2} \|d_x - D_x u - b_x\|_2^2 \\ d_y^{k+1} = \arg \min_{d_y} \|d_y\|_1 - 0.5d_y^T \cdot q_y^n + \frac{I}{2} \|d_y - D_y u - b_y\|_2^2 \\ d_r^{k+1} = \arg \min_{d_r} r \|d_r\|_1 + \frac{t}{2} \|d_r - Ku + f - b_r\|_2^2 \end{cases} \quad (15)$$

$$\begin{cases} b_x^{k+1} = b_x^n + D_x u^{n+1} - d_x^{n+1} \\ b_y^{k+1} = b_y^n + D_y u^{n+1} - d_y^{n+1} \\ b_r^{k+1} = b_r^n + K u_r^{n+1} - f - d_r^{n+1} \end{cases} \quad (16)$$

Now we can deal with these sub-problems one by one.

1) **u**-subproblem: By dropping the indices n , the solution of (15) can be obtained by setting the gradient to zero, leads to the requirement that

$$(mK^T K + I D^T D + t K^T K + 2c I_{MN}) u = mK^T f + I D^T (d - b) + t K^T (d_r + f - b_r) + 2cu^n \quad (17)$$

where $D = [D_x^T, D_y^T]^T$, left multiplying $(mK^T K - I \Delta + t K^T K + 2c I_{MN})^{-1}$, then we get the following formula

$$u = (mK^T K - I \Delta + t K^T K + 2c I_{MN})^{-1} (mK^T f - I \text{div}(d - b) + t K^T (d_r + f - b_r) + 2cu^n) \quad (18)$$

where $\Delta = -(D_x^T D_x + D_y^T D_y)$ and $div = -D^T$. This equation can be solved in the Fourier domain. We employ a discretized version of the Laplacian and a Gauss-Seidel scheme in [14].

$$U = (m|\hat{K}|^2 + t|\hat{K}|^2 - I\Re(\hat{\Delta}) + 2cI_{MN})^{-1} [m\hat{K}F - I\text{div}(\hat{d} - b) + t\hat{K}(\hat{a}_r + f - b_r) + 2c\hat{K}^H] \quad (19)$$

where the hat symbol represents the Fourier transform, \Re is the real part, the inverse and multiplications are pointwise.

2) d_x, d_y, d_r -subproblem: The problem (16) can be updated via the shrinkage formula [6], we can get d_x, d_y , and d_r

$$\begin{cases} d_x &= \text{shrink}(D_x u + b_x + 0.5q_x / I, 1/I) \\ d_y &= \text{shrink}(D_y u + b_y + 0.5q_y / I, 1/I) \\ d_r &= \text{shrink}(Ku - f + b_r, r/t) \end{cases} \quad (20)$$

The overall algorithm for solving (7) is summarized in Algorithm 1.

Algorithm 1 Pseudo-code of DCASBI Algorithm

Require:

Input data \mathbf{f}, \mathbf{K} , MaxDCA, tol and MaxBregman

Input parameters μ, λ, ρ, τ and c .

1: Initialize $\mathbf{u} = \mathbf{q}_x = \mathbf{q}_y = 0, k = 0$

2: **for** 1 to MaxDCA **do**

$$\mathbf{b}_r = \mathbf{b}_x = \mathbf{b}_y = 0$$

3: **for** 1 to MaxBregman **do**

4: Solve the \mathbf{u} -subproblem (14) by using (19).

5: Solve $\mathbf{d}_x, \mathbf{d}_y, \mathbf{d}_r$ -subproblem (15) by using (20).

6: Solve $\mathbf{b}_x, \mathbf{b}_y, \mathbf{b}_r$ -subproblem (16).

7: **end for**

$$(\mathbf{q}_x, \mathbf{q}_y) = (\mathbf{D}_x \mathbf{u}, \mathbf{D}_y \mathbf{u}) / \sqrt{|\mathbf{D}_x \mathbf{u}|^2 + |\mathbf{D}_y \mathbf{u}|^2}$$

8: **Check stop criterion**

$$\text{if } \frac{\|\mathbf{u}^{k+1} - \mathbf{u}^k\|_2}{\|\mathbf{u}^k\|_2} < \text{tol} \text{ then}$$

break

end if

9: **end for**

Experiments

In this section, we show an applications of the proposed method, image denoising. Since the matrix K of these examples can be diagonalized by Fourier transform, therefore our algorithms can be effectively carried out. We compare our methods with l_1 [14], $l_1 - l_2$ [3], and $l_1 - al_2$ [4].

In order to verify the proposed methods thoroughly, we use 14 test images from [15] including synthesis images and natural images which are shown in Fig. 1. Note that these images are gray-scale images. Here, the size of images (a), (d), (f), (g), (h), (j), and (l) are 512×512 , image (b) is of size 128×128 , image (c), (e), (i), (m), and (n) are of size 256×256 , the size of image (k) is 748×500 .

Evaluation Criteria. In order to evaluate the performance of the proposed method, we use structural similarity index ($SSIM$) and peak-signal-to-noise ratio ($PSNR$) as quantitative measure for restoration image quality. If the $SSIM$ value is closer to 1, the characteristic (edges and textures) of restored images is more similar to the original image.

Parameters Setting. In image denoising, input data of DCASBI, the stopping criterion tol , inner iteration $MaxBregman$, and outer iteration DCA are fixed to 10^{-4} , 200, and 2 respectively, which are the same as reference [4]. The optimal value of the parameters m, l, r and t are 10, 5, 3 and 0.043 respectively. Similarly, we do experiments for other test images. Due to limit space, we do not

present all experimental results. Based on experiments and analysis for other test images, the parameters l , r and t are fixed 5, 3 and 0.04 respectively, which are associated with the term $\|d - Du - b\|_2^2$, $\|Ku - f\|_1$ and $\|d_r - Ku + f + b_r\|_2^2$. As regard parameter m , the range of value m is from 5 to 20 to adapt different images. During experiment, we find that we should increase the value m when there are more periodic textures in image.

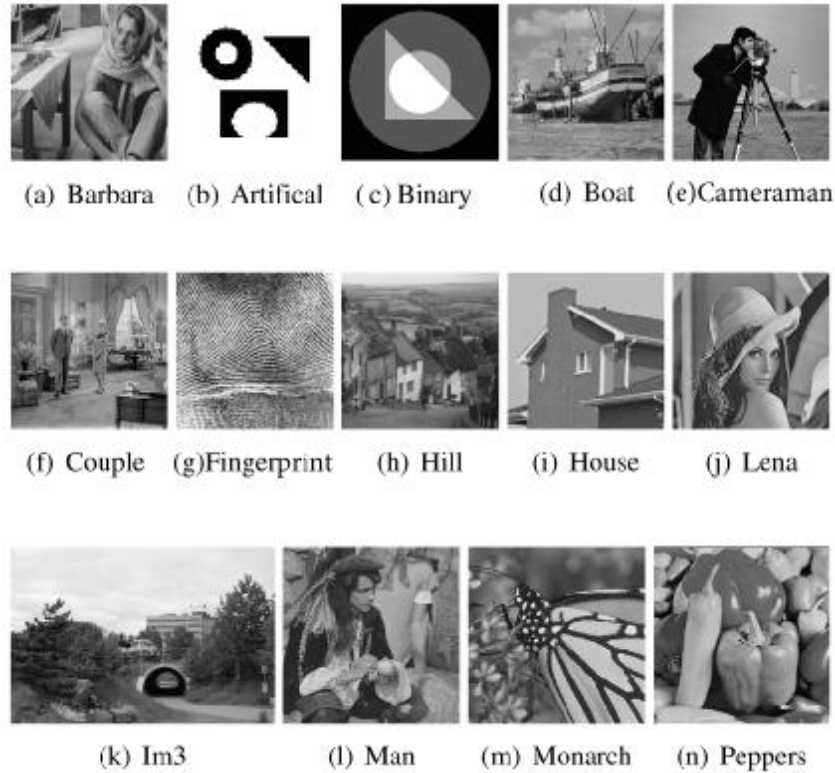


Figure 1. Test images [15]

As regards parameter c in (9), $c > 0$ guarantees the strong convexity and convergence of DCA . While $c = 0$, the objective function becomes monotone non-increasing, thus cannot ensure that converges to zero. Considering that $c = 0$ can achieve higher PSNR and SSIM for image denoising, we choose $c = 0$ for our proposed method.

In reference [4], the authors considered that 2 or 10 is a good stopping criterion as the outer iteration. In this paper, we chose 2 as the outer iteration of stopping criterion.

Image Denoising. We add zero-mean additive white Gaussian noise with standard deviation 0.05 to the test images. In terms of perceptual quality, Fig. 2 demonstrates that our methods are on par with $l_1 - 0.5l_2$ [4]. Furthermore, Table 1 shows that our methods are competitive with $l_1 - 0.5l_2$ in terms of quantitative measure (PSNR and SSIM). Besides, we compare our methods with other methods on the basis of (SSIM, PSNR and TIME) in the case of the noise variance from 0.01 to 0.1 for test image: Barbara. In terms of quantitative criteria (PSNR and SSIM), the proposed method is more efficient than the other methods.

Conclusion

In this paper, we propose a novel image restoration model which incorporates the hybrid data fidelity into the weighted difference of anisotropic and isotropic total variation. We employ DCA to solve the proposed model since it belongs to the difference form. We apply SBI to address the subproblem of DCA . This method is very effective and simple. Experiments show that the proposed model and algorithm can obtain more valid solution than competing methods in terms of objective criteria.

Acknowledgements

The research is supported by Characteristic Innovation Project (Natural Science) of the Education Department of Guangdong Province (Grant No. 2015KTSCX148), Guangdong Natural Science Fund of China (Grant No. 2015A030313646).

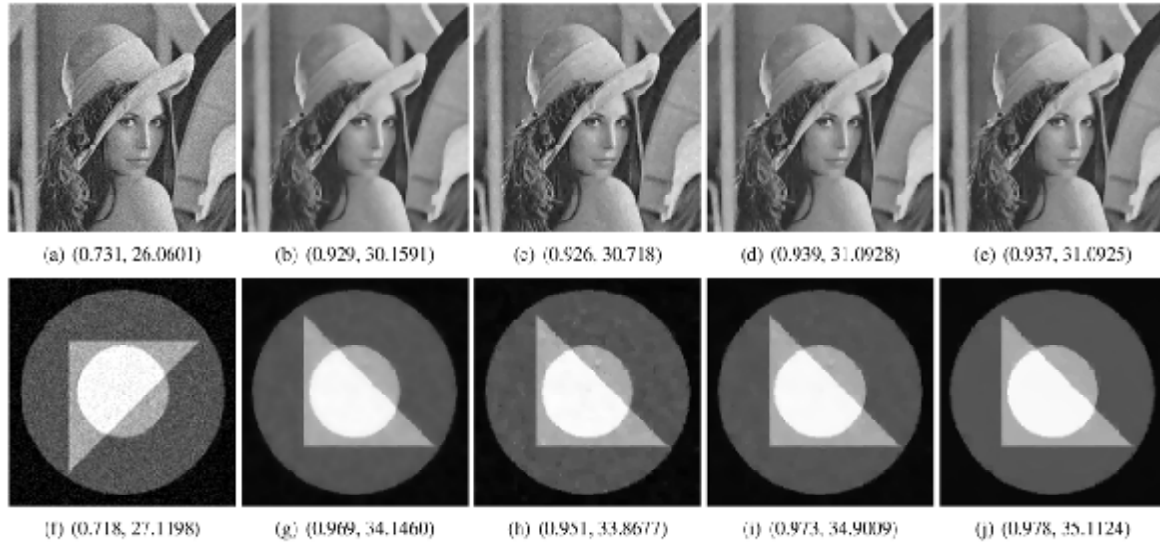


Figure 2. Denoising performance comparison for Lena and Binary. The numbers in parentheses denotes (SSIM, PSNR (dB)). (a)(f): noisy image. (b)(g): output of l_1 [14]. (c)(h): output of $l_1 - l_2$ [3]. (d)(i): output of $l_1 - 0.5l_2$ [4]. (e)(j): output of DCASBI.

Table 1. PSNR and SSIM for noisy images and restored images by applying different methods. Here, $DCA = 2$ in [4]. All numbers are average of five experiments in our experiments. The best results are highlighted in bold.

Image	l_1 [14]		$l_1 - l_2$ [3]		$l_1 - 0.5l_2$ [4]		DCASBI	
	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR
Barbara	0.9005	28.6847	0.9093	29.5159	0.9143	29.5712	0.9158	29.8230
Boat	0.8913	28.7490	0.9140	30.0350	0.9155	30.0207	0.9189	30.3654
Cameraman	0.9264	29.0375	0.9296	30.7694	0.9418	30.5949	0.9413	30.7327
Couple	0.8807	28.6988	0.9151	30.2453	0.9125	30.1333	0.9183	30.5219
Fingerprint	0.8625	23.7294	0.9426	26.4997	0.9204	25.6681	0.9507	27.2276
Hill	0.8835	29.0663	0.9070	30.2136	0.9086	30.1402	0.9134	30.4936
House	0.9455	30.8107	0.9286	31.5390	0.9536	32.0952	0.9560	32.0062
Lena	0.9286	30.1114	0.9256	30.7155	0.9384	31.0366	0.9365	31.0373
Man	0.8908	28.4032	0.9100	29.5972	0.9124	29.4182	0.9163	29.7220
Binary	0.9706	34.3267	0.9510	34.1107	0.9741	35.1562	0.9796	35.4856
Artificial	0.9828	29.5553	0.9834	31.5272	0.9886	30.7964	0.9905	30.9627
Monarch	0.9367	27.3518	0.9509	29.0206	0.9524	28.8182	0.9538	29.2325
Peppers	0.9370	28.0192	0.9417	29.5388	0.9472	29.2907	0.9464	29.5874
Im3	0.9009	29.7388	0.9113	30.6456	0.9200	30.7598	0.9228	30.9769

References

- [1] L. I. Rudin, S. Osher, E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 1992, 60(1): 259-268.
- [2] Z. Xu, X. Chang, F. Xu, H. Zhang. Regularization: A thresholding representation theory and a fast solver, *IEEE Transactions on neural networks and learning systems*, 2012, 23(7):1013-1027.

- [3] Y. Lou, M. Yan. Fast l_1 - l_2 minimization via proximal operator. arXiv preprint, 2016, 1609.09530.
- [4] Y. Lou, T. Zeng, S. Osher, et al. A weighted difference of anisotropic and isotropic total variation model for image processing. *SIAM Journal on Imaging Sciences*, 2015, 8(3): 1798-1823.
- [5] A. Beck, M. Teboulle. Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems. *IEEE Transactions on Image Processing*, 2009, 18(11): 2419-2434.
- [6] C. Li. An efficient algorithm for total variation regularization with applications to the single pixel camera and compressive sensing. PhD thesis, Citeseer, 2009.
- [7] A. Marquina, S. J. Osher. Image super-resolution by tv-regularization and bregman iteration. *Journal of Scientific Computing*, 2008, 37(3): 367-382.
- [8] C. R. Vogel, M. E. Oman. Fast, robust total variation-based reconstruction of noisy, blurred images. *IEEE Transactions on Image Processing*, 1998, 7(6): 813-824.
- [9] S. Alliney. A property of the minimum vectors of a regularizing functional defined by means of the absolute norm. *IEEE Transactions on Signal Processing*, 1997, 45(4): 913-917.
- [10] M. Hintermüller, A. Langer. Subspace correction methods for a class of nonsmooth and nonadditive convex variational problems with mixed l_1/l_2 data-fidelity in image processing. *SIAM Journal on Imaging Sciences*, 2013, 6(4): 2134-2173.
- [11] P. D. Tao, L. T. H. An. Convex analysis approach to dc programming: theory, algorithms and applications. *Acta Mathematica Vietnamica*, 1997, 22(1): 289-355.
- [12] P. D. Tao, L. T. H. An. A dc optimization algorithm for solving the trust-region subproblem. *SIAM Journal on Optimization*, 1998, 8(2): 476-505.
- [13] T. F. Chan, S. Esedoglu. Aspects of total variation regularized l_1 function approximation. *SIAM Journal on Applied Mathematics*, 2005, 65(5): 1817-1837.
- [14] T. Goldstein, S. Osher. The split bregman method for l_1 -regularized problems. *SIAM journal on imaging sciences*, 2009, 2(2): 323-343.
- [15] W. Zuo, D. Meng, L. Zhang, et al. A generalized iterated shrinkage algorithm for non-convex sparse coding. in *Proceedings of the IEEE international conference on computer vision*, 2013, 217-224.