

Analysis-Based Nonlocal-Approximate Sparsity Representation in Image Processing

Xiaowei He¹, Li Zhang² and Jiping Xiong³

¹College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004, P.R. China

²College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004, P.R. China

³College of Mathematics, Physics and Information Engineering, Zhejiang Normal University, Jinhua 321004, P.R. China

¹jhhxw@zjnu.edu.cn, ²hnnzyzhangli@zjnu.net, ³xjping@zjnu.cn

Abstract

l_1 norm is a popular regularizer in various linear inverse problems including image processing, compressed sensing and machine learning. But the non-zero entries of the sparsity solution obtained by l_1 are independent with each other, which always leads to biased result to real solution. Actually, there always exist some different correlations among those non-zero entries in an image signal domain or various analysis domains. In this paper, based on a simple observation that the non-zero entries of the sparsity vector in various image analysis domains should be also approximate when the relevant signal values are proximate, we proposed a nonlocal-approximate sparsity regularizer in analysis domains by minimizing the sum of the l_2 norms of those vectors with the same nonzero pattern like signal vectors. This regularizer is applied to image denoising, edge detecting, inpainting and decomposition models successively. The numerical experiments demonstrate the effectiveness of our method in terms of PSNR, visual effect and edge preserving.

Keywords: sparsity representation; analysis-based sparsity model; inverse problem; image processing

1. Introduction

Image restoration is a very important task in image processing and it is usually formulated as a linear inverse problem. We denote a whole image as a vector by concatenating its columns. The objective of image restoration is to find an unknown true image $u \in R^n$ from the observed one $f \in R^m$:

$$f = Au + e \quad (1.1)$$

where e is white Gaussian noise, and $A \in R^{m \times n}$ is a linear operator which is typically the identity operator in image denoising, a projection in image inpainting or a convolution in image deconvolution. Images usually have the feature of sparsity representation (or approximation) in different domains, such as image signal domain [8], Fourier transform [22], Cosine transform [23], Wavelet transform [32], and discrete gradient transform [15]. Sparsity representation has become an important method for solving (1.1) in different domains, in which images can be sparsity represented. In these domains, l_1 norms always act as the regularizers of the corresponding minimization problems for (1.1). According to the different types of domain, those minimization problems can be divided into two sub-categories, namely the synthesis-based sparsity representation problems and the

analysis-based sparsity representation problems [1-2]. The former consider that the image signal u is sparse and is always applied for solving the minimization problem in image domain itself:

$$\min |d| \quad s.t. \quad ARd = f \quad (1.2)$$

where $|\cdot|$ might be the l_1 norm or its variants, d is the coefficient of the unknown image u , R is the synthesis operator such that $u = Rd$. The formulation (1.2) can be transformed to one unconstrained minimization problem:

$$\min |d| + H(Rd) \quad (1.3)$$

where $H(\cdot)$ is a smooth convex function acting as the data fidelity, *e.g.*, $H(\cdot)$ can be chosen as $\frac{1}{2}\|ARd - f\|_2^2$. The two formulations (1.2) and (1.3) cover many problems by selecting different A , R and H , *e.g.*, compressing sensing [3] and various inverse problems [4-5]. Synthesis-based sparsity problems with (1.2) and (1.3) are relatively easy to solve due to the non-smooth term $|d|$ being separable. The analysis-based sparsity problems are quite different and they always assume that the analysis coefficient $|Du|$ is sparse. These models can be formulated to find a solution with sparsity representation in the corresponding image analysis domain:

$$\min |Du| \quad s.t. \quad Au = f \quad (1.4)$$

where D is the analysis operator, $|\cdot|$ might be l_1 norm or one variant. The Formulation (1.4) can be transformed to one unconstrained minimization problem:

$$\min |Du| + H(u) \quad (1.5)$$

where $H(\cdot)$ is a smooth convex function acting as the data fidelity, *e.g.*, $H(\cdot)$ can be chosen as $\frac{1}{2}\|Au - f\|_2^2$. Many efficient methods were proposed for solving (1.3) and (1.5), such as dual methods [24], Bregman iteration [26], Linearized Bregman iterations [27] and Split Bregman iteration [1-25]. In literatures mentioned above and other papers, such as compressed sensing [26-27], sparse prediction [6] and frame-based inverse problems [5-7], l_1 norm is used as regularizer without exception due to its role of sparsity induction to its solution. Nevertheless, the sparse solution obtained by l_1 norm regularizer is always lack of correlation between its nonzero entries. Actually, relationships among nonzero entries of sparse solution need to be seriously considered in some image applications. For example, various structured sparse regularizer [9-12] extended from l_1 norm have been proven useful in different applications. In other related works [13-14], other group sparsity models are proposed for different goals. Recently, Jun Liu *et. al.*, [10] proposed an overlapping group sparsity total variation functional(OGS-TV) and an image restoration model. Liu *et. al.*, [11] proposed a new explicit thresholding/shrinkage formula for one class regularization problem with OGS-TV. The models by these regularizers can be uniformly formulated as the following two minimization problems: $\min H(u) + \alpha|u|_G$ and $\min H(u) + \alpha|Du|_G$, where $|u|_G$ or $|Du|_G$ are group sparse

regularizers designed for embodying special kind of relationship among non-zero entries in signal domain or analysis domains respectively, D is the analysis operator, non-negative scalar α is a regularization parameter controlling the sparsity of the solution, $H(x)$ is a smooth convex loss function. In this paper, we mainly concentrate on the analysis-based correlated sparsity problems $\min H(u) + \alpha \|Du\|_G$ in various image analysis domains. It is difficult to solve due to the non-differentiability of the term $\|Du\|_G$. Here, we introduce an auxiliary variable z for Du to separate the calculation of the non-differentiable term and the fidelity term. The model is thus equivalent to the following formulation:

$$\min H(u) + \alpha \|z\|_G \quad s.t. \quad z = Du \quad (1.6)$$

What kind of correlation among those nonzero entries of the obtained sparse solution is always decided by the nonzero structure of those atomic vectors in image analysis domain. According to those correlations, let all atomic vectors have similar features such as nonlocal approximate, local continuous or having the same support structure. Then, we take the formulation of minimizing the sum of their l_2 norms as the regularizer in the models for image inverse problems. In this way, some correlated sparsity solutions can be finally obtained due to the designed method. Most analysis coefficients should be approximate in similar points with approximate values in an analysis domain. In this paper, we pay more attention to this kind of nonlocal-approximate correlated sparsity in some analysis domains and propose a corresponding sparse regularizer, which was used in image restoration, decomposition, inpainting. More details are given in Section 2. In Section 3, we give a denoising and edge-detecting model by the nonlocal-approximate correlated sparsity term in image gradient domain and show restoration results for some real images. As an example, we apply the nonlocal-approximate correlated sparsity term to the analysis domain of a piecewise linear framework for cartoon part and to local discrete cosine transform domain for texture part respectively. In section 4 we combine them for image inpainting and decomposition. In the last section, conclusions and further ideas for correlated sparsity representation are shown.

2. Denoising and Edge-Detecting by Nonlocal-Approximate Correlated Sparsity Term

Inverse problems can be formulated as convex optimization problems in various fields. A model x can be written as a nonnegative combination of a few elements of an atomic set [16], i.e., $x \in R^p$ be formed as follows:

$$x = \sum_{i=1}^k c_i a_i, \quad a_i \in G, c_i \geq 0; \quad (2.1)$$

where G is a set of atoms which are used to combine unknown data or transform data x , k is a relatively small number. G could be the finite set of unit-term one-sparse vectors or an infinite set of unit-norm rank-one matrices. This model arises in many applications and attracted more attention recently [17-19]. The problem of recovering sparse vectors from limited measurements has also received more and more attentions, such as image processing and computer vision [20], machine learning [6] and so on. When G is a collection of one-sparse vectors, l_1 norm is always used to induce sparse solutions in many applications. Similarly, the atomic set G of k -support sparse vectors could be

viewed as unit-norm k -support vectors [6], which is applied to obtain sparse solutions with a group of non-zero variables. For more details about convex atoms see paper [14].

As we know, l_1 norm usually acts as the regularization term in a linear optimization inverse problems, those large or small non-zero entries in relevant analysis domains have the same effect when computing the number of non-zero entries in the sparsity representation. Namely, relationship among them is not considered at all. This is obviously unreasonable and might be extremely sensitive to the threshold when computing the proximal vector. Actually, if the values of two points in some analysis domain of an image are similar, then the two corresponding entries in the final sparse representation have a high possibility of being approximate. Based on this consideration, we classify all the points into different groups, according to value-similarity in every analysis domain and those points in every group might be scattered locally. In this paper, we show simulations in the analysis domains such as gradient domain, piecewise linear framework and local discrete cosine transform (LDCT). In these domains, we define a kind of nonlocal-approximate correlated term as a regularizer. Then, we apply it to different image tasks. We classify all points in every domain as follows: Firstly, the points in every analysis domain of one image are divided into different non-overlapping groups depending on the value-similarity of them. Then, all points of every obtained group are organized into some vectors dynamically with fix-length support set. Finally, those vectors are used to define the corresponding nonlocal-approximate correlated sparsity term. The generating algorithm for them can be described as follows (assume an $M \times N$ image, $d = M \times N$):

- 1 Classification: according to the value-similarity classification method as formulation(2.2), classify all points in every analysis domain $w_j, j=1, \dots, p$ into q_j groups $w_j^i, i=1, 2, \dots, q_j$.
- 2 Generating: reorganize all the points in every group w_j^i into fix-length set randomly or in sequence, which are the support sets of the atomic vectors $z_j^i = \bigcup_{k=1}^{N_j^i} z_k, z_k \in R^d, N_j^i$ is the number of atomic vectors in group w_j^i .

Algorithm 1: The Generating Method for Atomic Vectors z

$$w_j = \bigcup_{i=1}^{q_j} w_j^i \mid \forall g, v, w_j^g \cap w_j^v = \emptyset, \delta_{i-1} \leq |v(w_j^i)| \leq \delta_i, \delta_1 \leq \delta_2 \leq \dots \leq \delta_{q_j} \quad (2.2)$$

where w_j is an analysis domain, $v(w_j^i)$ are the values of the group w_j^i . According to the classification method, the support sets of those obtained atomic vectors in every domain do not overlap. Let the atom vector set $G_j (j=1, \dots, p)$ in the j -th analysis domain be formed as follows:

$$G_j = \bigcup_{i=1}^{q_j} \bigcup_{k=1}^{N_j^i} z_j^{i_k}, z_j^{i_k} \in R^d, \text{support}(z_j^{i_k}) \subset w_j^i \wedge \text{support}(z_j^{i_k}) \cap \text{support}(z_j^{i'}) = \emptyset, \forall k, g, v \quad (2.3)$$

where $z_j^{i_k}$ means the k -th vector with the support set i -th group in the j -th analysis domain. We write $z = \bigcup_{j=1}^p z_j = \bigcup_{j=1}^p \sum_{i=1}^{q_j} \sum_{k=1}^{N_j^i} z_j^{i_k}, z_j^{i_k} \in G_j$, Let $G = \bigcup_{j=1}^p G_j$, then the nonlocal-approximate correlated term in the analysis domains w is given by

$$\|z\|_G = \min \sum_{j=1}^p \|z_j\|_{G_j} = \min \sum_{j=1}^p \sum_{i=1}^{q_j} \sum_{k=1}^{N_j^i} \|z_j^{i_k}\|_2, z_j^{i_k} \in G_j \quad (2.4)$$

In this way, due to the non-overlap feature of all vectors in G_j and (2.4), the proximal vector of a given vector when minimizing a linear optimization problem by this term can

be easily obtained [28]. The important work for us to do is to partition every domain into different groups by some suitable methods [29-31] firstly. Here, we classify it by simple method (2.2). Let the data fidelity item be $H(u) = \frac{1}{2}\|u - f\|_2^2$, then the image denoising constrain model based on analysis domain is as follows:

$$\min_u \frac{1}{2}\|u - f\|_2^2 + \alpha\|z\|_G \quad s.t. \quad z = Du \quad (2.5)$$

where f the observed image, $\alpha > 0$ a regularization parameter, D the analysis operator, $Du = \bigcup_{j=1}^p Du_j$. The Augmented Lagrangian method for the constrained optimization problems (2.5) is formulated as

$$\min_{u,z} \max_{\mu} L_c(u, z, \mu) = \frac{1}{2}\|u - f\|_2^2 + \alpha\|z\|_G + \langle z - Du, \mu \rangle + \frac{r}{2}\|z - Du\|_2^2 \quad (2.6)$$

where $\mu = \bigcup_{j=1}^p \mu_j$ are the Lagrange multipliers, $r = \bigcup_{j=1}^p r_j$ are positive constants. The method is to seek a saddle point of the Augmented Lagrangian functional $L_c(u, z, \mu)$. To solve the problem (2.10), we separate it to the following series of sub-problems induced by optimized conditions:

$$\begin{cases} u^{k+1} = (1 + r)^{-1}(f + D^T(\mu^k) + rD^T(z^k)) \\ z_j^{k+1} = \arg \min_{z_j} \frac{r_j}{2}\|z_j - (D^T(u_j^{k+1}) - \mu_j^k/r_j)\|_2^2 + \alpha\|z_j^k\|_G, \quad j = 1, 2, \dots, p \\ \mu^{k+1} = \mu^k + r(z^{k+1} - \nabla u^{k+1}) \end{cases} \quad (2.7)$$

The proximal algorithm [21] for every sub-problem of the second line in (2.7) is given by

$$z_j^{k+1} = \begin{cases} (1 - 1/\|w\|_2)w, & \|w\|_2 > rr_j \\ 0, & \|w\|_2 \leq rr_j \end{cases} \quad (2.8)$$

Where $w = D^T(u_j^{k+1}) - \mu_j^k/r_j$.

The second line of (2.7) has been divided into p sub-problems and every one is designed for every domain Ω_j . The algorithm 2 illustrates the decomposition of the step two of (2.7).

- 1 for every domain Ω_j
- 2 for every group Ω_j^i in every domain Ω_j
- 3 obtain the proximal vectors by (2.8) for $w = D^T(u_j^{k+1}) - \mu_j^k/r_j$.
- 4 recover all proximal vectors exactly to the right positions of the domain Ω_j
- 5 end
- 6 obtain the new $z_j = \bigcup_{i=1}^{N_j} z_j^i$.
- 7 end

Algorithm 2: Decomposition of Step Two of (2.8)

Because all $\Omega_j^i (i=1, 2, \dots, N_j)$ are not overlapping, it is easy and fast for us to implement the Algorithm 2.

3. Inpainting and Decomposition Model by $\|\cdot\|_G$

The nonlocal-approximate correlated sparsity term described in (2.4) can be applied not only to image denoising but also to image decomposition and inpainting. It can even be applied to obtain the cartoon part and the texture part of one noisy image simultaneously. The cartoon part models the piecewise smooth part of one image, while the texture part represents the oscillating part of an image. In the following, we give the formulation for image decomposition by (2.4). The cartoon part is represented in piecewise linear frame, while the texture part is represented in local discrete cosine transform domain.

$$\min_{uc, ut} \left\{ \frac{1}{2} \|uc + ut - f\|_2^2 + \alpha_1 \|zc\|_G + \alpha_2 \|zt\|_G \right\} \quad s.t. \quad D_1(uc) = zc \wedge D_2(ut) = zt \quad (3.1)$$

where zc induces the cartoon part uc in the analysis domain based on D_1 (piecewise linear frame), zt induces the texture part ut in analysis domain D_2 (local discrete cosine transform). By the Augmented Lagrangian method for (3.1) and relevant optimality conditions, the following sub-problems are obtained immediately:

$$\begin{cases} uc^{k+1} = (r^2 + 2 \times r)^{-1} (r \times f - (1+r)D_1^T(\mu c^k + r \times zc^k) + rD_2^T(ut^k + r \times zt^k)) \\ ut^{k+1} = (r^2 + 2 \times r)^{-1} (r \times f - (1+r)D_2^T(\mu t^k + r \times zc^k) + rD_1^T(uc^k + r \times zc^k)) \\ zc^{k+1} = \arg \min_{zc} \|zc^k - (D_1(uc^{k+1}) - \mu c^k / r)\|_2^2 + r \|zc^k\|_G \\ zt^{k+1} = \arg \min_{zt} \|zt^k - (D_2(ut^{k+1}) - \mu t^k / r)\|_2^2 + r \|zt^k\|_G \\ \mu c^{k+1} = \mu c^k + r_1(zc^{k+1} - \nabla uc^{k+1}) \\ \mu t^{k+1} = \mu t^k + r_2(zt^{k+1} - \nabla ut^{k+1}) \end{cases} \quad (3.2)$$

The model for image inpainting and decomposition is as follows:

$$\min_{uc, ut} \left\{ \alpha_1 \|zc\|_G + \alpha_2 \|zt\|_G \right\} \quad s.t. \quad D_1(uc) = zc \wedge D_2(ut) = zt \wedge P_w(uc + ut) = P_w(f) \quad (3.3)$$

where zc and zt are the same as in (3.1), P_w is the project, which is a diagonal matrix with diagonals 1 if the corresponding pixel is known, or 0 otherwise. Then the following sub-problems are obtained:

$$\begin{cases} uc^{k+1} = (r^2 + 2 \times r)^{-1} P_w(rP_w f + (1+r)D_1^T(\mu c^k + r \times zc) + rD_2^T(ut^k + r \times zt^k) \\ \quad + (I - P_w)D_1^T(\mu c^k + r \times zc^k)) \\ ut^{k+1} = (r^2 + 2 \times r)^{-1} P_w(rP_w f + (1+r)D_2^T(\mu t^k + r \times zc) + rD_1^T(uc^k + r \times zc^k) \\ \quad + (I - P_w)D_2^T(\mu t^k + r \times zt^k)) \\ zc^{k+1} = \arg \min_{zc} \|zc^k - (D_1(uc^{k+1}) - \mu c^k / r)\|_2^2 + r_1 \|zc^k\|_G \\ zt^{k+1} = \arg \min_{zt} \|zt^k - (D_2(ut^{k+1}) - \mu t^k / r)\|_2^2 + r_2 \|zt^k\|_G \\ \mu c^{k+1} = \mu c^k + r_1(zc^{k+1} - \nabla uc^{k+1}) \\ \mu t^{k+1} = \mu t^k + r_2(zt^{k+1} - \nabla ut^{k+1}) \end{cases} \quad (3.4)$$

4. Experiment and Analysis

In this section, some experimental results are given to demonstrate the effectiveness of our model. To be fair, all experiments in this paper are carried out on Windows 7, 64-bit and Matlab

R2012a running on a desktop equipped with an Intel Core i5-4460 CPU 3.2 GHz and 4 GB of RAM. For convenience, parameters of all examples are illustrated as follows: Classnum represents the group number q_j in (2.2); according to the group number, only δ_1 and $\delta_{Classnum}$ are given in all examples, the others $\delta_2, \delta_3, \dots, \delta_{Classnum-1}$ in (2.2) are computed by mean iterative method; the different fix-length of Algorithm 1 are given for different examples. Simulation results of (2.7) (D is gradient operator) are given in Figure 1 and Figure 2, which show that (2.7) is a good method for obtaining the denoising and edge-detecting results simultaneously. If the classification result before image processing is good enough, the obtained denoising result is as good as the LPGPCA method [33] and the BM3D approach [34]. For example, Figure 1(e) and Figure 2(e) are the denoising results by (2.7) for two noisy images (Figure 1(b) and Figure 2(b)) based on the classification for the original images (Figure 1(a) and Figure 2(a)). Obviously, they are benefited from the satisfying classification.

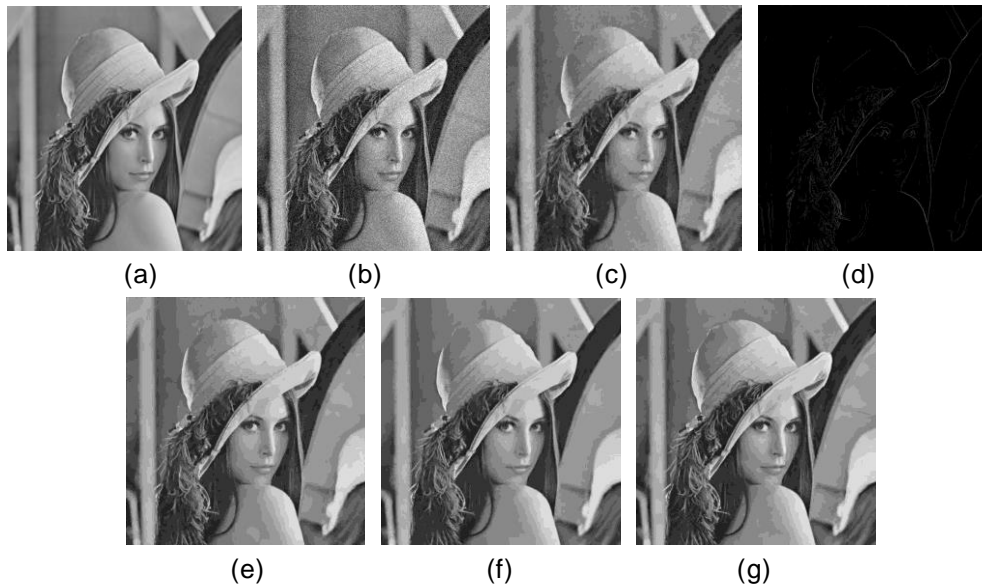
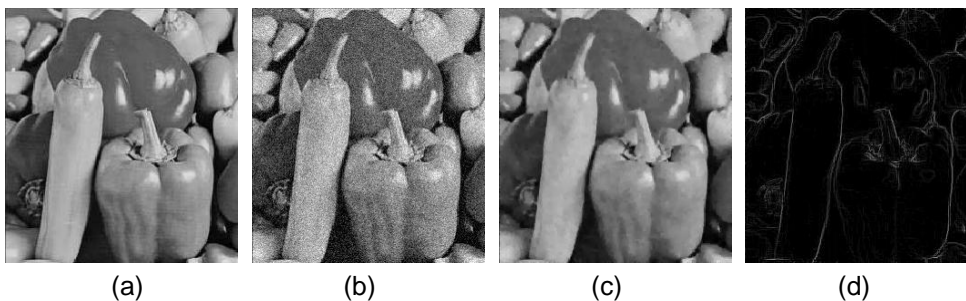


Figure 1. Denoising and Edge-Detecting Results for Lena512 Image in Gradient Domain. (a)Original Image; (b)Noisy Image, $\sigma = 20$, PSNR=22.07dB; (c)Classifying Gradient Domain Based on (b) and Denoising Result by (2.7), PSNR=31.3dB, ClassNum=8; Fix-Length=15; $\delta_8=0.1$; $\delta_1=0.01$; (d)Edge Detecting Result; (e)Classifying Gradient Domain Based on Original Image and Denoising Result by (2.7), PSNR=33.43dB, ClassNum=6; Fix-Length=55; $\delta_6=0.12$; $\delta_1=0.02$; (f)Denoising Result by LPGPCA[33], PSNR=32.64dB; (g)Denoising Result by BM3D[34], PSNR=33.2dB



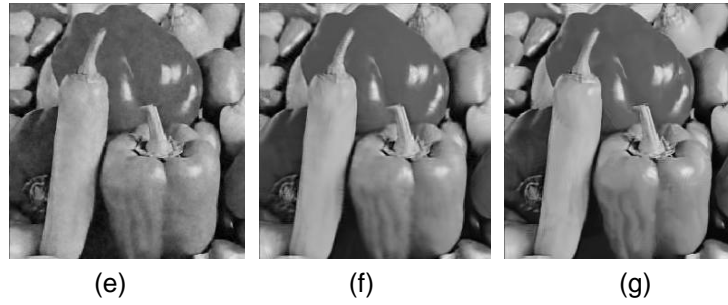


Figure 2. Denoising and Edge-Detecting Results for Vegetable Image in Gradient Domain. (a)Original Image; (b)Noisy Image, $\sigma=20$, PSNR=22.08dB; (c) Classifying (b) and Denoising Result by (2.7), PSNR=30.22dB, ClassNum=6; Fix-Length=12; $\delta_6=0.12$; $\delta_1=0.02$; (d)Edge Detecting Result; (e)Classifying Gradient Domain Based on Original Image and Denoising Result by (2.7), PSNR=32.3dB, ClassNum=6; fix-Length=40; $\delta_6=0.1$; $\delta_1=0.016$; (f)Denoising Result by LPGPCA [33], PSNR=30.8dB; (g)Denoising Result by BM3D[34], PSNR=31.56dB

Simulation results for noisy image decomposition by (3.2) are shown in Figure 3. These results are satisfying and the texture parts of them are pretty good. Simulation results for image inpainting and decomposition simultaneously by (3.4) and the Split Bregman for the ROF model are shown in Figure 4 and Figure 5 respectively. Comparing with the ROF model, our approach has absolute advantages, especially the texture parts of the images. Actually our approach has more potential due to the flexible choices for classification method and the fix-length of vector.



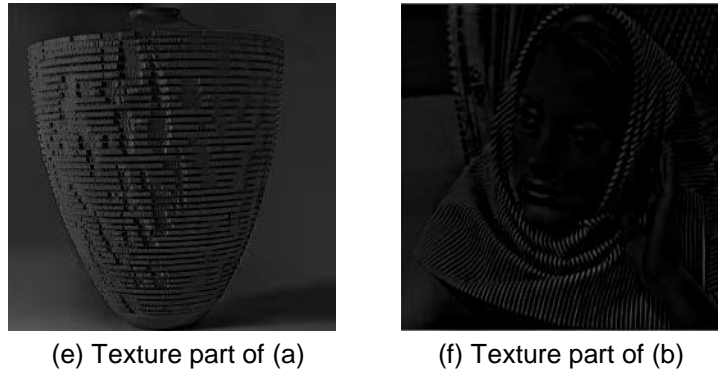
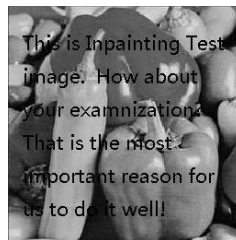


Figure 3. Cartoon and Texture Decomposition by (3.2). PSNR of (c)+(e): 33.4dB, PSNR of (d)+(f):33.2dB. ClassNum=8; Fix-Length=80; $\delta_8=0.2$; $\delta_1=0.01$



(a) Painting 256*256 Image, PSNR=17.3dB



(b) Cartoon part by (3.4)



(c) Texture part by (3.4)



(d) Cartoon+Texture by (3.4), PSNR=39.47dB



(e) Cartoon part by Split Bregman for ROF



(f) Texture part by Split Bregman for ROF



(g) Cartoon+Texture by Split Bregman for ROF, PSNR=38.7dB

Figure 4. Inpainting and Decomposition Simultaneously by (3.4) and Split l_1 , The Number of Iteration by (3.4) is 65 and CPU Time is 58s, the Latter is 95 and 36s. ClassNum=8; Fix-Length=80; $\delta_8=0.2$; $\delta_1=0.01$

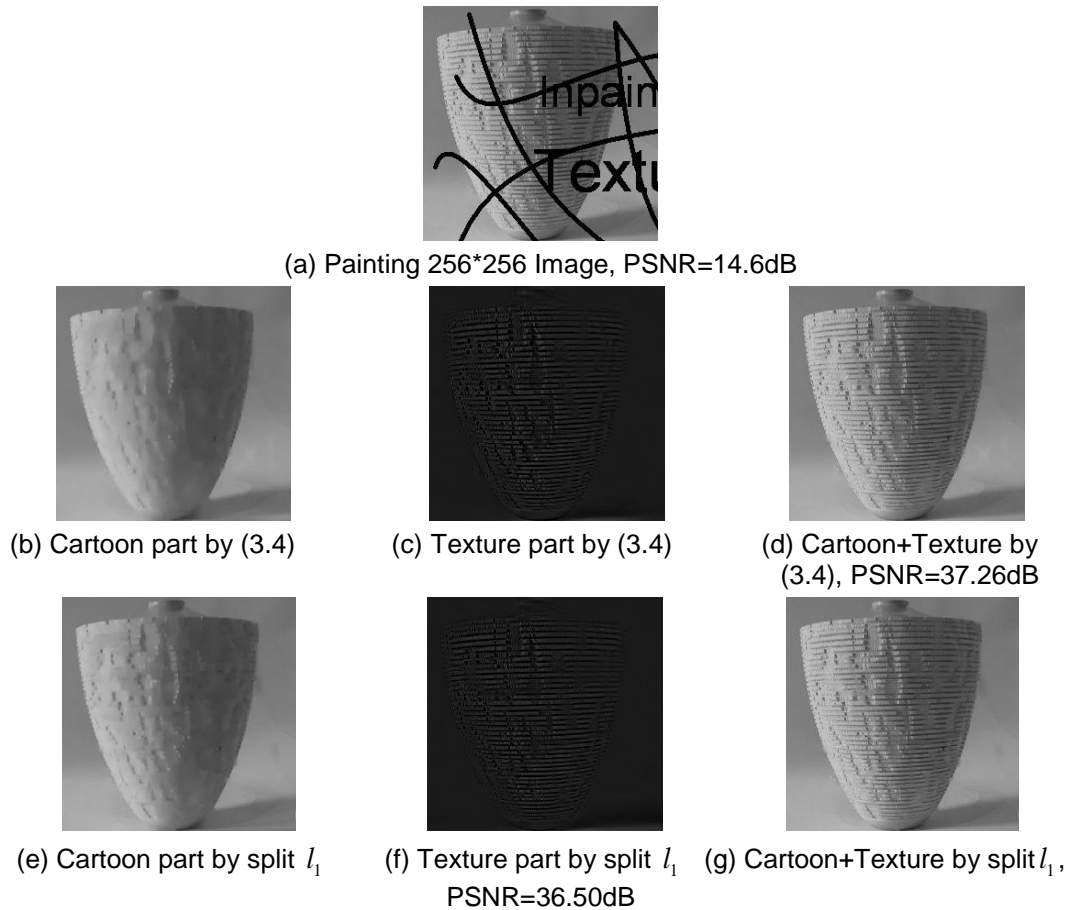


Figure 5. Inpainting and Decomposition Simultaneously by (3.4) and Split l_1 , The Number of Iteration by (3.4) is 79 and CPU Time is 69s, the Latter is 99 and 38s. ClassNum=8; Fix-Length=80; $\delta_8=0.2$; $\delta_1=0.01$.

5. Conclusions

To consider the relationships between non-zero entries of the sparsity vectors in image processing, we propose a correlated sparsity term based on a simple value-similarity classification method in analysis domains, namely nonlocal-approximate correlated sparsity term. Then relevant models for image restoration are obtained by it. In fact, how to organize atoms in a dictionary for correlated sparsity term in image domain and various analysis domains still need to be studied further and is easily extensible according to different image processing tasks.

Acknowledgments

This work is supported by Zhejiang Provincial Natural Science Foundation of China(grantno.LY14F010008) and NSFC(61572023). The authors would like to thank three anonymous reviewers for their valuable comments and constructive suggestions.

References

- [1] J. F. Cai, S. Osher and Z. Shen, "Split Bregman methods and frame based image restoration", Journal of Multiscale Modeling and Simulation, vol. 8, no. 2, (2009), pp. 337-369.

- [2] E. J. Candes, Y. C. Eldar, D. Needell and P. Randall, "Compressed sensing with coherent and redundant dictionaries", *Applied and Computational Harmonic Analysis*, vol. 31, no. 1, (2011), pp. 59-73.
- [3] D. L. Donoho, "Compressed sensing", *IEEE Transactions on Information Theory*, vol. 52, no. 4, (2006), pp. 1289-1306.
- [4] C. Chaux, P. L. Combettes, J. C. Pesquet and V. R. Wajs, "A variational formulation for frame-based inverse problems", *IOPscience in Inverse Problems*, vol. 32, no. 4, (2007), pp. 1495-1518.
- [5] I. Daubechies, G. Teschke and L. Vese, "Iteratively solving linear inverse problems under general convex constraints", *Inverse Problems and Imaging*, vol. 1, no. 1, (2007), pp. 29-46.
- [6] A. Argriou, R. Foygel and N. Srebro, "Sparse Prediction with the k-Support Norm", *Advances in Neural Information Processing Systems*, Lake Tahoe, (2012) December 8.
- [7] C. Chaux, P. L. Combettes, J. C. Pesquet and V. R. Wajs, "A variational formulation for frame-based inverse problems", *Inverse Problems*, vol. 23, no. 4, (2007), pp. 1533-1546.
- [8] M. Elad, "Sparse and Redundant Representation Modeling — What Next?", *IEEE Signal Processing Letters*, vol. 19, no. 12, (2012), pp. 922-928.
- [9] J. Friedman, T. Hastie, H. Hofling and R. Tibshirani, "Pathways coordinate optimization", *Annals of Applied Statistics*, vol. 1, no. 2, (2007), pp. 302-332.
- [10] J. Liu, T. Z. Huang, I. W. Selesnick, X. G. Lv and P. Y. Chen, "Image restoration using total variation with overlapping group sparsity", *Information Sciences*, vol. 295, no. 20, (2013), pp. 232-246.
- [11] G. Liu, T. Z. Huang, X. G. Lv and J. Liu, "New explicit thresholding/shrinkage formulas for one class of regularization problems with overlapping group sparsity and their applications", *Computer Vision and Pattern Recognition*, Portland, (2013) June 23-28.
- [12] L. Jacob, G. Obozinski and J. P. Vert, "Group lasso with overlap and graph lasso", *Proceedings of the 26th Annual International Conference on Machine Learning*, Montreal Quebec, Canada, (2009) June 14-18.
- [13] R. Jenatton, J. Mairal, G. Obozinski and F. Bach, "Proximal methods for hierarchical sparse coding", *Journal of Machine Learning Research*, vol. 9, no. 12, (2011), pp. 2297-2334.
- [14] R. Jenatton, J. Y. Audibert and F. Bach, "Structured variable selection with sparsity-inducing norms", *The Journal of Machine Learning Research*, vol. 12, no. 2, (2011), pp. 2777-2824.
- [15] L. I. Rudin, S. Osher and E. Fatemi, "Nonlinear total variation based noise removal algorithms", *Physica D: Nonlinear Phenomena*, vol. 60, no. 1, (1992), pp. 259-268.
- [16] P. A. Parrilo, A. S. Willsky, V. Chandrasekaran and B. Recht, "The convex geometry of linear inverse problems", *Foundations of Computational Mathematics*, vol. 12, no. 6, (2012), pp. 805-849.
- [17] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization", *Foundations of Computational mathematics*, vol. 9, no. 6, (2009), pp. 717-772.
- [18] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal l_1 norm solution is also the sparsest solution", *Communications on pure and applied mathematics*, vol. 59, no. 6, (2006), pp. 797-829.
- [19] B. Recht, M. Fazel and P. A. Parrilo, "Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization", *SIAM review*, vol. 52, no. 3, (2010), pp. 471-501.
- [20] S. Aja-Fernández, "Tensors in image processing and computer vision", New York, Springer, (2009).
- [21] P. L. Combettes and J. C. Pesquet, "Proximal splitting methods in signal processing", *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer New York, (2011), pp. 185-212.
- [22] B. Ghazi, H. Hassanieh, P. Indyk, D. Katabi, E. Price and L. Shi, "Sample-optimal average-case sparse fourier transform in two dimensions", *51st Annual Allerton Conference on Communication, Control, and Computing*, Urbana, USA, (2013) October 2-4.
- [23] B. J. Zhang, X. H. Yin, W. Wang and Q. Li, "Research on image processing with compressed sensing algorithm: Base on the improved layered discrete cosine transform", *IEEE International Conference on Communications*, Ottawa, Canada, (2012) June 10-15.
- [24] T. F Chan, G. H. Golub and P. Mulet, "A nonlinear primal-dual method for total variation-based image restoration", *SIAM Journal on Scientific Computing*, vol. 20, no. 6, (1999), pp. 1964-1977.
- [25] T. Goldstein and S. Osher, "The Split Bregman Algorithm for l_1 Regularized Problems", *SIAM Journal on Imaging Sciences*, vol. 2, no. 2, (2009), pp. 323-343.
- [26] W. Yin, S. Osher and D. Goldfarb, "Bregman iterative algorithms for l_1 -minimization with applications to compressed sensing", *SIAM Journal on Imaging Sciences*, vol. 1, no. 1, (2008), pp. 143-168.
- [27] J. F. Cai, S. Osher and Z. Shen, "Linearized Bregman iterations for compressed sensing", *Mathematics of Computation*, vol. 78, no. 267, (2009), pp. 1515-1536.
- [28] P. L. Combettes and J. C. Pesquet, "Proximal splitting methods in signal processing", *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, Springer New York, (2011), pp. 185-212.
- [29] I. Ram, M. Elad and I. Cohen, "Image denoising using NL-Means via smooth patch ordering", *The 38th International Conference on Acoustics, Speech, and Signal Processing*, Vancouver, Canada, (2013) May 26-30.

- [30] Y. Fan and S. Qin, "A new method of image classification based on local appearance and context information", *Neurocomputing*, vol. 119, no. 1, (2013), pp. 33-40.
- [31] M. Varma and A. Zisserman, "A statistical approach to material classification using image patch exemplars", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 31, no. 11, (2009), pp. 2032-2047.
- [32] J. F. Cai, B. Dong, S. Osher and Z. W. Shen, "Image restoration: Total variation, wavelet frames, and beyond", *Journal of the American Mathematical Society*, vol. 25, no. 4, (2012), pp. 1033-1089.
- [33] L. Zhang, W. S. Dong, D. Zhang and G. M. Shi, "Two-stage image denoising by principal component analysis with local pixel grouping", *Pattern Recognition*, vol. 43, no. 4, (2010), pp. 1531-1549.
- [34] K. Dabov, A. Foi, V. Katkovnik and K. Egirzarian, "BM3D image denoising with shape-adaptive principal component analysis", *Signal Processing with Adaptive Sparse Structured Representations*, Saint-Malo, France, (2009) April 6- 9.

Authors



Xiaowei He, was born in China, in April 1968. He received the B.S degree in Mathematics from Zhejiang Normal University, and the M.S degree in Computer Application from Zhejiang University, Jinhua, P.R. China in 1992 and Hangzhou in 2000 respectively. Now, he is an associate professor and also a Master Tutor in Computer Software and Theory in Zhejiang Normal University. His research interests include image processing, signal processing and inverse problems. E-mail: jhhxw@zjnu.edu.cn



Li Zhang, was born in China, in August 1990. She received the B.S. degree in Computer Science and technology from Xinyang Normal University in 2013, China. She is currently studying for a M.S. degree in Computer Science in Zhejiang Normal University. She main research interest is image processing including image denoising and reconstruction. E-mail: hnnyzhangli@zjnu.net



Jiping Xiong, received his B.S degree in electronics & communication engineering from University of Science and Technology of China in 2001, and Ph.D. degree in Communication and Information System from University of Science and Technology of China in 2006. Currently, he is an Associate Professor of Computer Engineering at Zhejiang Normal University of China. His research interests include compressive sensing, computer vision, networking, information security and signal processing. E-mail: xjping@zjnu.cn