CHEMICAL ENGINEERING TRANSACTIONS

VOL. 46, 2015

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.aidic.it/cet

Guest Editors: Peiyu Ren, Yancang Li, Huiping Song

ISBN 978-88-95608-37-2; **ISSN** 2283-9216

DOI: 10.3303/CET1546030

An Improved Sparse Representation Model for Robust Image Denoising

Zhi Cui*, Xianpu Cui

School of Communication and Electronic Engineering, Hunan City University, Yiyang, 413000, China.
zhicui@yeah.net

Though the sparse representation has demonstrated to be a very effective tool to de-noise the images with low levels of noise, it usually losses the power to well preserve structural features in images with high levels of noise. In this paper, we propose an improved de-noising method for images with low signal-to noise ratio. Specifically, the proposed method takes the histogram structural similarity (HSSIM) as similarity factor to replace the reconstruction error as the new fidelity term, and finds the most appropriate sparse coefficients by using the modified orthogonal matching pursuit (OMP) algorithm which enables structures in the reconstructed image run as close as possible to the ideal image. In addition, the proposed method adaptively trains the initialized dictionary by using the K-singular value decomposition (K-SVD) algorithm based on HSSIM to assure the image structures can be well reconstructed under the high noise circumstance. Experiment results have shown that the proposed method is better than some well-known de-noising methods in terms of PSNR and edge-preserved index (EPI) in high noise condition.

1. Introduction

De-noising is a fundamental work in image processing. In the past few years, many approaches have been investigated to against this problem such as curvelet transformation (Ma and Plonka (2007)), total variation diffusion (Easley, *et al* (2009)), wavelet threshholding (Donoho (1995)), quaternion wavelet (Yin, *et al* (2012)), and inter-scale orthonormal wavelet thresholding (Luisier and Blu (2008)).
Recently, sparse and redundant representation has attracted more and more attention. It is based on the theory that the given non-zero vector can be linear combined by minimum column in a predefined row-full-rank matrix with constraint of row number is much less than column number (Mallat and Zhang (1993)). Generally, the row-full-rank matrix is called dictionary and the columns in this matrix are called atoms. Sparse and redundant representation model has broken up the limit of orthogonal basis and maximize the advantages of redundant basis, and has been widely used in image fusion (Li, *et al* (2013)), image restoration (Ning, *et al* (2013)), and image Super-resolution (Kim and Kwon (2010)).
Application of sparse representation in image de-noising has also garnered considerable attention. Because of the satisfaction of sparsity and separability for signal to noise, the sparse representation method based on redundant dictionary can achieve better de-noised effect than the transform domain based method (Li, *et al* (2012)). One of the most successful technique is the K-SVD algorithm reported by Elad and Aharon (2006). This approach extends the work by reported by Aharon and Elad (2006) in many ways, and the results have demonstrated that a dictionary which can adaptively updated would cause improved de-noising performance. On the foundation of work by Elad and Aharon (2006), Z. Zhou, *et al* (2012) proposed a K-least mean square algorithm (K-LMS) for the dictionary learning and image representation, and the de-noising could be achieved by combine the adaptive image sparse decomposition in a learned over-complete dictionary and the threshold process. However, the fixed step length adopted in the sparse decomposition would cause a large level of steady-state error.
From the research course of the image de-noising based on the sparse representing, it is known that most of the current de-noising methods consider the reconstruction error of the images before and after de-noising as the fidelity term. However, affected by work environment and monitoring objects, image collected in the real condition always contains rich details and low signal-to-noise ratio. If we continue using the above criteria, the noise components introduced in the reconstructing process will have a larger impact on the image

Please cite this article as: Cui Z., Cui X.P., 2015, An improved sparse representation model for robust image denoising, Chemical Engineering Transactions, 46, 175-180 DOI:10.3303/CET1546030

reconstruction accuracy, and the image's structural similarity also could not be well preserved. In order to keep the structural features information as much as possible when the image is effectively de-noised, in this paper, we focus on the removal of additive Gaussian white noise from standard images and propose an improved method based on the foundational work by Elad and Aharon (2006). Our method has two specific parts. The first is to replace the reconstructed error with the histogram structural similarity as the new fidelity term. This is motivated by the conclusion that HSSIM aligns to the features of visual cortex and thus tends to produce better results that agree with the human visual system (Wang and Bovik (2004)). The second relates to the training of redundant dictionary. With our method, the initialized dictionary is trained by using K-SVD with constraint of HSSIM so that the image structures can be well reconstructed under the high noise circumstance.

2. Related Work

Fang *et al* (2012) reported that any ideal image could be described by the model of $\mathbf{y} \in \mathbf{R}^{n \times N}$ in which the image is separated into $\sqrt{n} \times \sqrt{n}$ parts. According to the sparse representation, a dictionary matrix $\mathbf{D} \in \mathbf{R}^{n \times N}$ is defined to represent all the image parts as follows:

$$\hat{\mathbf{a}} = \arg\min \|\mathbf{a}\|_0 \quad s.t. \quad \|\mathbf{D}\mathbf{a} - \mathbf{x}\|_2^2 \le \varepsilon \tag{1}$$

where $\mathbf{a} \in \mathbf{R}^n$ is the sparse representation coefficient, $\|\mathbf{a}\|_0$ is the number of the non-zero values which means the sparsity of $\mathbf{a}$, $\mathbf{D}$ is the dictionary.
Transform the constraints to the penalty term. According to the regularization optimization, the equation (1) can be changed to:

$$\{\hat{\mathbf{a}}_{ij}, \hat{\mathbf{y}}\} = \underset{\mathbf{a}_{ij}, \mathbf{y}}{\operatorname{argmin}} \, \lambda \|\mathbf{y} - \mathbf{x}\|_2^2 + \sum_{i,j} \mu_{ij} \|\mathbf{a}_{ij}\|_0 + \sum_{i,j} (\mathbf{D}\mathbf{a}_{ij} - \mathbf{R}_{ij}\mathbf{y}) \tag{2}$$

where λ is the Lagrange multiplier, μ_{ij} is a coefficient, the first component represent the integral similarity between the image with noise and the clear image, and it should be less than the convolution of $C * \sigma^2$ (*C* is a constant). The second one is the sparse constraints. $\mathbf{R}_{ij}$ is the $n \times N$ matrix to pick up the image block in (i, j) from the image of. $\mathbf{D}$ is the over-completed dictionary. Finally, the denoised image $\mathbf{y}$ is updated by

$$\hat{\mathbf{y}} = \underset{\mathbf{y}}{\operatorname{argmin}} \, \lambda \|\mathbf{y} - \mathbf{x}\|_2^2 + \sum_{i,j} \mu_{ij} \|\mathbf{D}\mathbf{a}_{ij} - \mathbf{R}_{ij}\mathbf{y}\|_2^2 \tag{3}$$

3. The proposed method

In this section, we describe the proposed method. The core of our method is to replace the reconstruction error with histogram structure similarity and make it as a new fidelity term. In addition, when the ways of initial over-complete dictionary and orthogonal matching pursuit algorithm are applied in sparse decomposition, we incorporate HSSIM into the constraint condition.
In practical terms, the demonstration of the proposed method can be shown in Fig. 1.

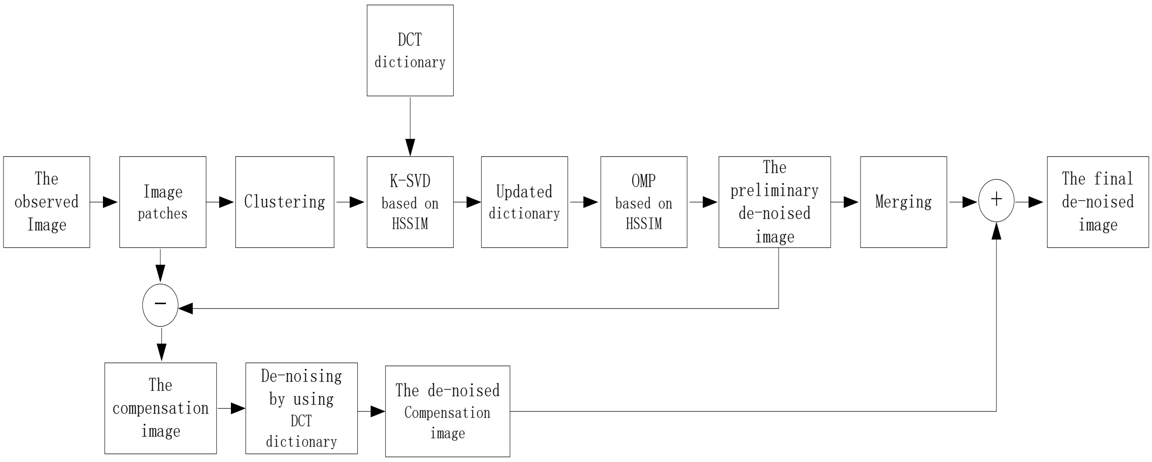

Figure 1: Flowchart of the proposed method

Here is the improved sparse representation de-noising model that we have proposed:

$$\{\hat{\mathbf{a}}_{ij}, \hat{\mathbf{y}}\} = \underset{\mathbf{a}_{ij}, \mathbf{y}}{\arg\min}\, \lambda \|\mathbf{y}-\mathbf{x}\|_2^2 + \sum_{i,j} \mu_{ij} \|\mathbf{a}_{ij}\|_0 + \sum_{i,j} (1 - \mathrm{H}(\mathbf{D}\mathbf{a}_{ij} - \mathbf{R}_{ij}\mathbf{y})) \tag{4}$$

In equation (4), the first and the second parts on the right side are the constraints, and the third one is the similarity factor which replaces the reconstruction error of K-SVD as the new computational fidelity term.
HSSIM is defined as follows:

$$H(x,y) = l(x,y)q(x,y)h(x,y) \tag{5}$$

where $l(x,y)$, $q(x,y)$ and $h(x,y)$ represents intensity, contrast and ambiguity, respectively.
Assume that the initial condition is $\mathbf{y}=\mathbf{x}$, the sparse coefficient of each image patch can be calculated as follows:

$$\hat{\mathbf{a}}_{ij} = \underset{\mathbf{a}_{ij}}{\arg\min}\, \mu_{ij} \|\mathbf{a}_{ij}\|_0 + (1 - \mathrm{H}(\mathbf{D}\mathbf{a}_{ij} - \mathbf{R}_{ij}\mathbf{y})) \tag{6}$$

Then the result of de-noising can be acquired by solving $\mathbf{y}$ partial derivatives of equation (5):

$$\hat{\mathbf{y}} = (\lambda \mathbf{I} + \sum_{i,j} \mathbf{R}_{ij}^{\mathrm{T}} \mathbf{R}_{ij})^{-1} (\lambda \mathbf{x} + \sum_{i,j} \mathbf{R}_{ij}^{\mathrm{T}} \mathbf{D} \hat{\mathbf{a}}_{ij}) \tag{7}$$

From the above analysis, the procedure of our method is summarized as follows:
(1) We adopt the completed DCT dictionary as the initial dictionary, satisfying $\mathbf{y}=\mathbf{x}$, and cluster each sample into $\mathbf{K}$ sub-groups.

(2) We divide the observed image into overlapping blocks, and take these overlapping image patches as the training data set for the K-SVD algorithm. Specifically, we take $1-\mathrm{HSSIM}(\mathbf{D}\mathbf{a}_{ij} - \mathbf{R}_{ij}\mathbf{y})$ to replace $\|\mathbf{D}\mathbf{a}_{ij} - \mathbf{R}_{ij}\mathbf{y}\|_2^2$ as the new fidelity term in order to well preserve the image structural features. Through equation (13), we can obtain all the sparse coefficients.

(3) In this stage, we define the error matrix as $E_k = \mathbf{y} - \sum_{j \neq k} \mathbf{d}_j \mathbf{a}_j^{\mathrm{T}}$, and define the atom used for the sparse separation in the dictionary as $\omega_i = \{i \,|\, 1 \le i \le k, \mathbf{a}_j^{\mathrm{T}}(i) \neq 0\}$, where $\mathbf{d}_j$ is the first column of dictionary atom, and $\mathbf{a}_j^{\mathrm{T}}$ is the first row of the sparse coefficient matrix.. Then the dictionary updating is switched to the following model:

$$\min_{\mathbf{d}_j,\mathbf{a}_j}(1-H(\mathbf{y},\sum_{j\neq k}\mathbf{d}_j\mathbf{a}_j^{\mathrm{T}}))\quad s.t.\quad \mathbf{a}_j^{\mathrm{T}}\subseteq\omega_i \tag{8}$$

Repeat the stage (3) J times until the stop criterion.

(4) When the above stages are completed, we can now obtain the preliminary de-noised image $\mathbf{y}$. However, the details such as edge in the image might be lost. To solve it, we first obtain the compensation image $\mathbf{y}_{dif}$ by calculating the difference between the preliminary de-noised image $\mathbf{y}$ and the noisy image $\mathbf{x}$, then obtain the de-noised compensation image $\hat{\mathbf{y}}_{dif}$ by using the DCT redundant dictionary, after that we add the de-noised compensation image $\hat{\mathbf{y}}_{dif}$ to the preliminary de-noised image $\mathbf{y}$ and obtain the final de-noised image $\mathbf{y}_t$. The solution is formulated as follows:

$$\mathbf{y}_t=\mathbf{y}+\hat{\mathbf{y}}_{dif} \tag{9}$$

4. Experiments

In this section, we carry out experiments with the purpose of testing the performance of the proposed method. We used other three different methods for a fair comparative analysis. The parameters of the proposed method are set as $J=10$ and $n=8\times8$, respectively. Since in [25], Elad and Aharon set λ and C as $30/\sigma$ and 1.15 respectively, we adopt the same value, but default value for the parameters in the other methods. Considering the objectivity of the effect evaluation, PSNR and EPI were adopted as the objective indices to evaluate the quality of de-noised images. All the experiments are promoted on a Core i5(R) 2.6 GHz PC with 4 GB RAM.

EPI is defined as follows:

$$E(x,y)=\frac{\sum\left|p_{x_i}-p_{x_j}\right|}{\sum\left|p_{y_i}-p_{y_j}\right|} \tag{10}$$

where p_x and p_y represents the gray level in observed image and denoise image.

First, we compared the EPI and PSNR results obtained by the proposed method and some other methods, such as contourlet, K-SVD, and K-LMS. As is shown that, the tests were performed on three images: “Man”, “Babara” and “Boat”. AWGN is superimposed on them with $\sigma=$20, 30 and 40. The results are given in Tab. 1.

Table 1: Performance of the de-noising methods by HSSIM and PSNR

Image	Noise σ	EPI				PSNR/dB			
		Contourlet	K-SVD	K-LMS	**Proposed**	Contourlet	K-SVD	K-LMS	**Proposed**
Man	20	0.546	0.603	0. 617	**0.689**	26.45	28.74	28.83	**28.91**
	30	0.527	0.582	0.594	**0.632**	25.97	28.13	28.24	**28.47**
	40	0.503	0.541	0.545	**0.599**	24.44	26.69	26.71	**27.85**
Babara	20	0.573	0.612	0.633	**0.681**	24.42	27.48	27.54	**27.83**
	30	0.554	0.593	0.605	**0.624**	23.57	26.60	26.63	**26.97**
	40	0.539	0.567	0.575	**0.591**	21.31	24.26	24.39	**25.65**
Boat	20	0.534	0.601	0.627	**0.673**	22.63	25.54	25.60	**25.89**
	30	0.516	0.587	0.593	**0.616**	21.48	24.21	24.59	**24.70**
	40	0.489	0.551	0.572	**0.584**	20.19	22.48	22.57	**23.94**

The differences of de-noising results obtained by the proposed method and the others for the same image with the same σ are significant. Specifically, by adding the white noise with the variance of 30 to “Man”, the EPI

result obtained by the proposed method is 0.632, with the increased ratio by 19.92%%, 8.59% and 6.39% respectively compared with the values by using Contourlet, K-SVD, and K-LMS, and the PSNR result obtained by the proposed method is 28.47 dB, with increased value by 2.50dB, 0.34dB and 0.23dB respectively compared with the values in Contourlet, K-SVD and K-LMS. It also can be found that the results obtained by the proposed method are better than the others for different images with different σ. For instance, the PSNR result obtained by the proposed method is increased by 3.31dB, 0.67dB and 0.56dB on average respectively than in Contourlet, K-SVD and K-LMS, and the EPI result obtained by the proposed method is increased by 19.11%, 8.64% and 6.14% on average respectively than in Contourlet, K-SVD and K-LMS.
The comparison of the de-noising effects obtained by the proposed method and the others for all the three images are given in Fig. 2, with the high noise variance of 70, because under the low noise condition ($\sigma < 20$), the de-noising results are almost perfect, just with some slight differences. From the aspect of subjective visual effect, it can be found that the proposed method has some advantages in keeping the structure features compared with the other methods.

Figure 2: Visual comparison of the reconstructed results on three images (Man, Babara and Boat), with σ=70. The first column: noise-free images. The second column: reconstructed results obtained by Contourlet. The third column: reconstructed results obtained by K-SVD. The fourth column: reconstructed results obtained by K-LMS. The fifth column: reconstructed results obtained by the proposed method.

5. Conclusions

In this paper, we have proposed an improved de-noising method for low-SNR images. The proposed method makes use of structure information in image in order to keep the de-noising result meet the human visual system. Traditional sparse representation based methods take the reconstruction error as fidelity term, which leads to the shortage of structure preserving. This regret motivated us to replace the reconstruction error with similarity factor as a new fidelity term. Detail compensation is also used for improve the effect of de-noising. Experiment results show that the proposed method outperforms Contourlet, K-SVD and K-LMS in terms of EPI and PSNR. In the future, we will consider the dictionary learning method.

Acknowledgments

This work is supported by Scientific Research Fund of Hunan Provincial Education Department under Grant of 13C122.

References

Donoho, D. L., 1995, De-noising by soft threshholding, IEEE Transactions on Information Theory, 41(3), 613-627, 10.1109/APCCAS.2000.913631

Easley, G. R., Labate, D., Colonna, F., 2009, Shearlet based total variation diffusion for denoising, IEEE Transactions on Image Processing, 18(2), 260-268, 10.1109/tip.2008.2008070

Elad, M., Aharon, M., 2006, Image denoising via learned dictionaries and sparse representation, Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, 895-900, 10.1109/CVPR.2006.142

Fang, L. Y., Li, S. T., Nie, Q., Izatt, J. A., Toth, C. A., Farsiu, S., 2012, Sparsity based denoising of spectral domain optical coherence tomography images, Biomedical Optics Express, 3(5), 927-942, 10.1364/boe.3.000927

Kim, K. I., Kwon, Y. H., 2010, Single-image super-resolution using sparse regression and natural image prior, IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(6), 1127-1133, 10.1109/tpami.2010.25

Li, S.T., Yin, H.T., Fang, L.Y., 2013, Remote sensing image fusion via sparse representations over learned dictionaries, IEEE Transactions on Geoscience and Remote sensing, 51(9), 4779–4789, 10.1109/tgrs.2012.2230332

Li, S. T., Fang, L.Y., Yin, H. T., 2012, An efficient dictionary learning algorithm and its application to 3-D medical image denoising, IEEE Transactions on Biomedical Engineering, 59(2), 417-427, 10.1109/tbme.2011.2173935

Luisier, F., Blu, T., 2008, SURE-LET multichannel image denoising: Interscale orthonormal wavelet thresholding, IEEE Transactions on Image Processing,17(4), 1057-7149,

M. Aharon, M. Elad, A.M. Bruckstein, 2006, K-SVD: An algorithm for designing of overcomplete dictionaries for sparse representation, IEEE Transactions on Signal Processing, 54(11), 4311-4322, 10.1109/tsp.2006.881199

Ma, J., Plonka, G., 2007, Combined curvelet shrinkage and nonlinear anisotropic diffusion, IEEE Transactions on Image Processing, 16(9), 2198-2206, 10.1109/tip.2007.902333

Mallat, S., Zhang, Z., 1993, Matching pursuits with time-frequency dictionaries, IEEE Transactions on Image Processing, 41(12), 3397–3415, 10.1109/tip.2008.919370

Wang, Z., Bovic, A. C., 2009, Mean square error: Love it or leave it? A new look at signal fidelity measures, IEEE Signal Processing Magazine, 26(1), 98-117, 10.1109/msp.2008.930649

Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P., 2004, Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing, 13(4), 600–612, 10.1109/tip.2003.819861

Ning, Q., Chen, K., Yi, L., Fan, C. C., Lu, Y., Wen, J. T., 2013, Image super-resolution via analysis sparse prior, IEEE Signal Processing Letters, 20(4), 399-402, 10.1109/lsp.2013.2242198

Yin, M., Liu, W., Shui, J., Wu, J. M., 2012, Quaternion wavelet analysis and application in image denoising, Mathematical Problems in Engineering, 2012, 1-21, 10.3724/sp.j.1187.2012.00338

Zhou, Z., Luo, L. M., 2012, Research on Image Denoising Algorithm based on Adaptive Overcomplete Sparse Representation Theories, Journal of Convergence and Information Technology, 7(16), 315-321, 10.4156/jcit.vol7.issue16.38