Deblocking Scheme for JPEG-Coded Images Using Sparse Representation and All Phase Biorthogonal Transform

Liping Wang, Chengyou Wang, and Xiao Zhou
School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China
Email: sduwlp@163.com; {wangchengyou, zhouxiao}@sdu.edu.cn

Abstract —For compressed images, a major drawback is that those images will exhibit severe blocking artifacts at very low bit rates due to adopting Block-based Discrete Cosine Transform (BDCT). In this paper, a novel deblocking scheme using sparse representation is proposed. A new transform called All Phase Biorthogonal Transform (APBT) was proposed in recent years. APBT has the similar energy compaction property with Discrete Cosine Transform (DCT). It has very good column properties, high frequency attenuation characteristics, low frequency energy aggregation, and so on. In this paper, we use it to generate the over-completed dictionary for sparse coding. For Orthogonal Matching Pursuit (OMP), we select an adaptive residual threshold by combining blind image blocking assessment. Experimental results show that this new scheme is effective in image deblocking and can avoid over-blurring of edges and textures. We can obtain deblocked images in the receiver.

Index Terms—Image deblocking, sparse representation, All Phase Biorthogonal Transform (APBT), Orthogonal Matching Pursuit (OMP)

I. Introduction

With the development and progress of technology, the resolution of the imaging device is higher and higher, which leads to the amount of image data increasing exponentially. To solve the contradiction between the data rate and channel bandwidth, the massive data must be compressed to meet the needs of image transmission and storage. The Block-based Discrete Cosine Transform (BDCT) has good features of regularity and simplicity for hardware implementation. Therefore, BDCT is widely utilized by several image and video international coding standards, such as JPEG [1], H.264/AVC [2], and H.265/HEVC [3]. However, at lower bit rates, the blocking artifacts are caused by quantization error. The blocking artifacts make image's visual quality degenerated. Besides, it also has a negative impact on the further image processing. Therefore, image deblocking is necessary and important, especially for highly compressed images.

The image deblocking algorithm aims to alleviate blocking artifacts and improve visual quality of compressed images. Deblocking methods can be mainly divided into two categories: in-loop processing methods and post-processing methods. The in-loop deblocking processing methods not only avoid the propagation of blocking artifacts between adjacent frames, but also can enhance coding efficiency. For example, H.264/AVC [2] adopts the in-loop deblocking processing. Some researchers have designed the in-loop deblocking filter [4]. Numerous experimental results manifest that the in-loop deblocking filter can provide both objective and subjective improvement compared with the compressed images. The image deblocking methods which is performed after decompressed is called post-processing. Therefore, post-processing is more practical and used for image deblocking.

For removing and eliminating the blocking artifacts, numerous image deblocking approaches based on post-processing have been proposed, such as filtering methods [5] and Projection Onto Convex Sets (POCS) [6]. Recently, learning-based sparse representation has been successfully used for image deblocking. Jung *et al*. [7] obtained an over-completed dictionary using the K-Singular Value Decomposition (K-SVD) algorithm [8], and proposed new method to automatically estimate the residual threshold for Orthogonal Matching Pursuit (OMP) [9]. Instead of processing each image patch individually, Zhang *et al*. [10] proposed Group-based Sparse Representation (GSR), and a novel image deblocking method is proposed based on GSR and Quantization Constraint (QC) prior. Similarly, Zhao *et al*. [11] proposed Structural Sparse Representation (SSR). And combined with QC prior, a novel algorithm for image deblocking was proposed. All of them demonstrated that sparse representation can effectively remove image deblocking and ensure the image's visual quality. Our work has been inspired by the image deblocking results based on sparse representation. In this paper, we will mainly focuses on image deblocking method for JPEG-coded images.

For image deblocking method based on sparse representation, a proper dictionary and sparse coding algorithm are needed. Because All Phase Biorthogonal Transform (APBT) [12] has very good column properties, high frequency attenuation characteristics, low frequency energy aggregation, and so on. The dictionary based on

Manuscript received August 28, 2016; revised December 20, 2016.
This work was supported by the National Natural Science Foundation of China (No. 61201371), the promotive research fund for excellent young and middle-aged scientists of Shandong Province, China (No. BS2013DX022), and the Natural Science Foundation of Shandong Province, China (No. ZR2015PF004).
Corresponding author email: wangchengyou@sdu.edu.cn.
doi:10.12720/jcm.11.12.1095-1101

APBT can reveal the sparsity of image better than the dictionary based on the commonly used orthogonal basis. In this paper, we use the APBT to produce the over-completed dictionary. For sparse coding algorithm, an adaptive residual threshold for OMP is used by combining blind image blocking assessment. Experimental results show that our proposed scheme is better than other methods.

The rest of this paper is organized as follows. Section II concisely introduces the sparse representation and JPEG model. Section III presents the image deblocking scheme for JPEG compressed images using sparse representation and APBT. In Section VI, experimental results and comparisons are presented. Finally, conclusion and further research are given in Section V.

II. Sparse Representation and JPEG Model

A. Sparse Representation Model

For image patches of size $\sqrt{n}\times\sqrt{n}$, they are converted into column vectors $\boldsymbol{x}\in\mathbb{R}^n$ in lexicographical order. In order to construct sparse-land model, a dictionary of size $\boldsymbol{D}\in\mathbb{R}^{n\times k}$ (with $k>n$, implying that it is redundant) is defined. In general, the dictionary is assumed to be known and fixed. According to the sparse representation model, every image patch $\boldsymbol{x}$ could be represented sparsely over this dictionary $\boldsymbol{D}$,

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \text{ subject to } \boldsymbol{D\alpha}\approx\boldsymbol{x} \tag{1}$$

The $\|\boldsymbol{\alpha}\|_0$ represents the number of the nonzero entries in $\boldsymbol{\alpha}$. In general, $\|\hat{\boldsymbol{\alpha}}\|_0 \ll n$. The basic principle of sparse representation is that every signal can be considered as a linear combination of few columns of the dictionary $\boldsymbol{D}$.

If we substitute a specific requirement of an allowable bounded representation error $\|\boldsymbol{D\alpha}-\boldsymbol{x}\|_2\le\varepsilon$ for the rough constraint $\boldsymbol{D\alpha}\approx\boldsymbol{x}$, this sparse representation model will become more precise. For the sparse-land input signal $\boldsymbol{x}$ and the noisy output signal $\boldsymbol{y}$ which is contaminated by white Gaussian noise with standard deviation σ, the Maximum A Posteriori (MAP) estimator for denoising this noisy signal is represented as

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{\alpha}\|_0 \text{ subject to } \|\boldsymbol{D\alpha}-\boldsymbol{y}\|_2^2\le T \tag{2}$$

where T is residual threshold, and it is decided by ε and σ. The optimization task of (2) can be also changed into

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \|\boldsymbol{D\alpha}-\boldsymbol{y}\|_2^2 + \mu\|\boldsymbol{\alpha}\|_0 \tag{3}$$

For a proper value of μ, the (2) and (3) are equivalent. At last, the denoised image can be obtained by $\hat{\boldsymbol{x}}=\boldsymbol{D}\hat{\boldsymbol{\alpha}}$.

Generally, (2) or (3) is very hard to solve. However, we can efficiently solve this problem using several available approximation techniques, such as OMP [9], Basis Pursuit (BP) [13], and Matching Pursuit (MP) [14]. In our work, we use OMP due to its simplicity and efficiency.

B. JPEG Model

The frame of the image compression standard used here is JPEG [1], which contains three basic steps: DCT, DCT coefficient quantization, and Huffman entropy encoding. The decoding process is inverse process of encoding. The process of encoding is shown as follows:

Step 1. The input image is first converted into the YCrCb color space and then grouped into blocks of size 8×8.

Step 2. Before carrying out a BDCT, the input image data are shifted from unsigned integers to signed integers.

Step 3. Each block is transformed by BDCT. Each block will include 64 DCT coefficients composed of one DC coefficient and 63 AC coefficients.

Step 4. Quantize each matrix block.

Step 5. After quantization, the DC coefficient of each block is conducted a Differential Pulse Code Modulation (DPCM) coding. Before entropy encoding, the 63 coefficients are scanned into a 1-D zig-zag sequence.

III. Proposed Scheme

The framework of our proposed method is shown in Fig. 1. Our proposed method includes two parts: the generation of over-completed dictionary and sparse coding with adaptive residual thresholds.

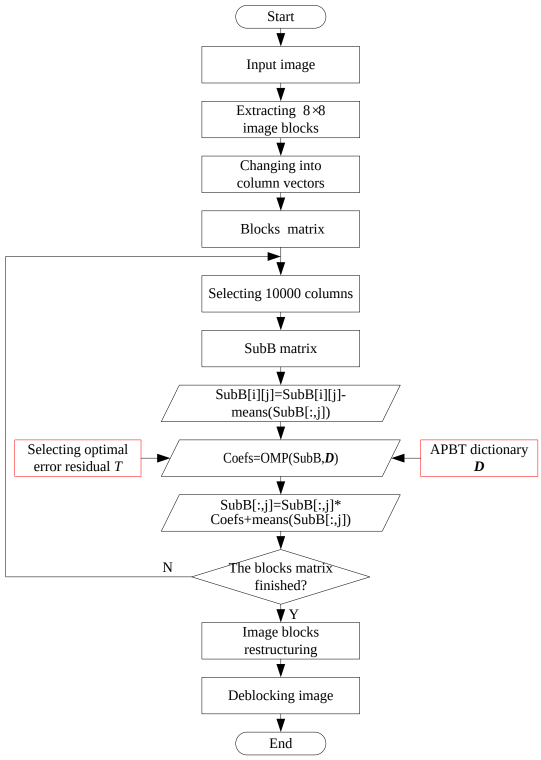


Fig. 1. Flow chart of our proposed method

A. The Over-Completed APBT Dictionary

According to all phase digital filtering, there are three kinds of APBT based on WHT, DCT, and IDCT [12]. In this paper, we will focus on the APBT based on DCT and

IDCT in order to compare with the DCT. Similar to DCT matrix, APBT can also be used to generate the over-completed dictionary for image spares representation. Taking APDCBT (APBT based on DCT) for example, the process of two-dimensional APBT is shown as follows. For a signal sequence with N points $\boldsymbol{x}$ and the APDCBT matrix $\boldsymbol{V}$ with size of $N \times N$, the transform coefficients $\boldsymbol{y}$ can be denoted by (4) and (5) after two-dimensional APDCBT transform.

$$\boldsymbol{y} = \boldsymbol{V}\boldsymbol{x} \tag{4}$$

$$\boldsymbol{V}(m,n) = \begin{cases} \dfrac{N-m}{N^2}, & m = 0,1,\cdots,N-1,\ n = 0, \\ \dfrac{1}{N^2}\left[(N-m)\cos\dfrac{mn\pi}{N} - \csc\dfrac{n\pi}{N}\sin\dfrac{mn\pi}{N}\right], & m = 0,1,\cdots,N-1,\ n = 1,2,\cdots,N-1. \end{cases} \tag{5}$$

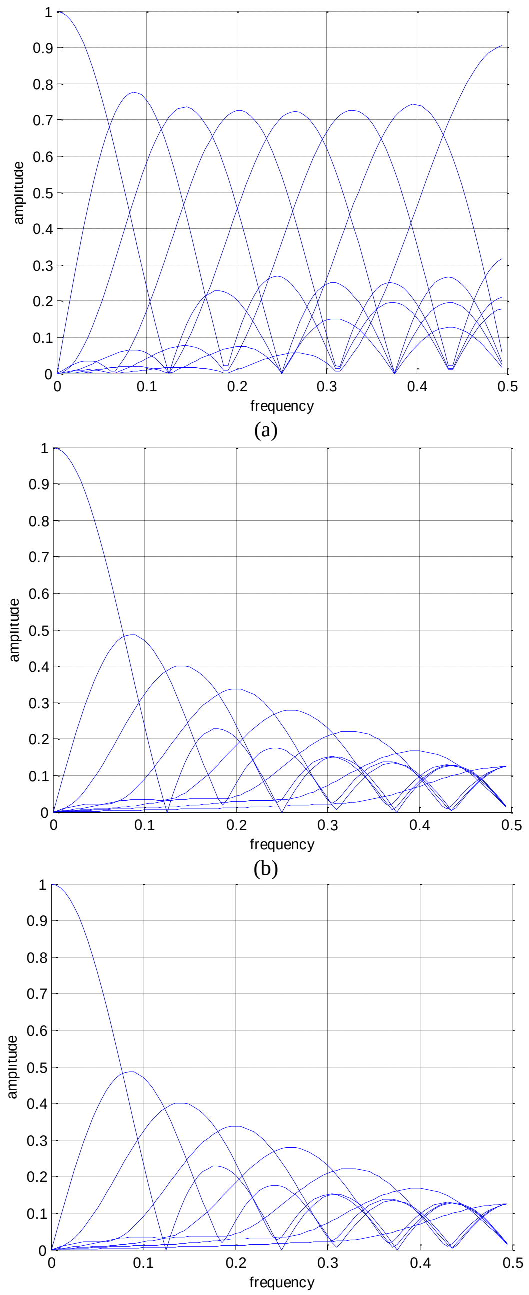


Fig. 2. The normalized amplitude frequency response of each filter: (a) DCT, (b) APDCBT, and (c) APIDCBT.

Similarly, APIDCBT (APBT based on IDCT) can be denoted by

$$\boldsymbol{V}(m,n) = \begin{cases} \dfrac{1}{N}, & m = 0,1,\cdots,N-1,\ n = 0, \\ \dfrac{N-m+\sqrt{2}-1}{N^2}\cos\dfrac{m(2n+1)\pi}{2N}, & m = 0,1,\cdots,N-1,\ n = 1,2,\cdots,N-1. \end{cases} \tag{6}$$

Fig. 2 gives the normalized amplitude frequency response of each filter of DCT matrix and APBT matrix. From the point of view of frequency domain analysis, the energy of APBT transform is more concentrative than that of DCT transform. The APBT transform can make more image energy concentrate on the low frequency part and gather them. That is, different frequencies are weighted differently in the process of transformation, and it can better reveal the sparse property of an image.

The Lena image of size 64×64 is conducted transformation by APBT and DCT. According to the order from big to small, we select 50 transform coefficients, and the distribution curves of 50 transform coefficients are shown in Fig. 3. The transform coefficients of the signal by using APBT are more concentrated than DCT, and coefficient distribution curve of APBT has better attenuation characteristics. Therefore, in this paper, we use the APBT to generate the over-completed dictionary for the image deblocking. Fig. 4 shows the different over-completed dictionaries.

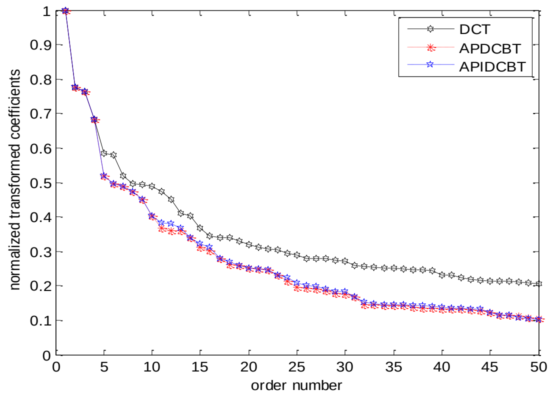


Fig. 3. Coefficients distribution curves of Lena image.

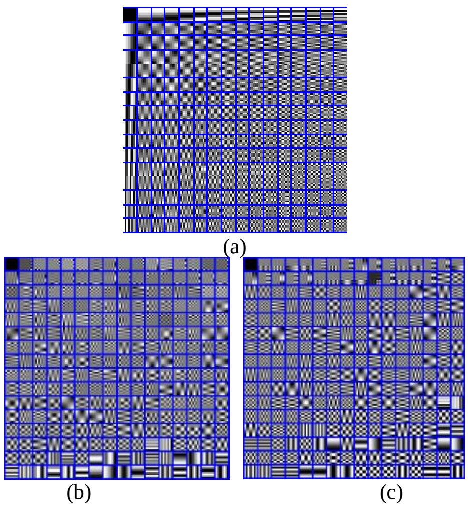


Fig. 4. The dictionaries with size of 64×256: (a) DCT, (b) APDCBT, and (c) APIDCBT.

B. An Adaptive Residual Threshold

The residual threshold T is important to optimize (2). Fig. 5 shows the objective experimental results in which the Lena image with different quality factors q are conducted image deblocking via sparse representation. The dictionary is the APDCT dictionary of size 64×1024.

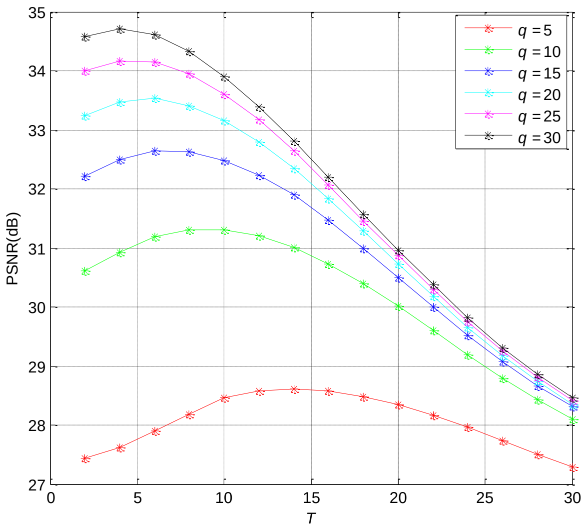


Fig. 5. The effects of residual threshold on image deblocking.

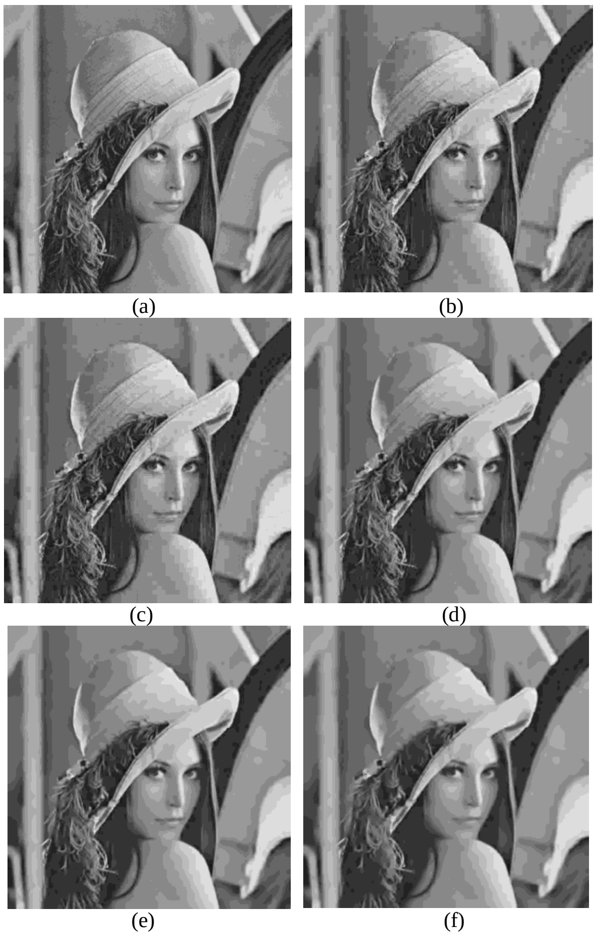


Fig. 6. The effects of residual threshold T on Lena image deblocking: (a) the uncompressed image; (b) the compressed image with $q = 10$; (c), (d), (e), and (f) are the deblocked image with $T = 2$, 8, 14, and 20, respectively.

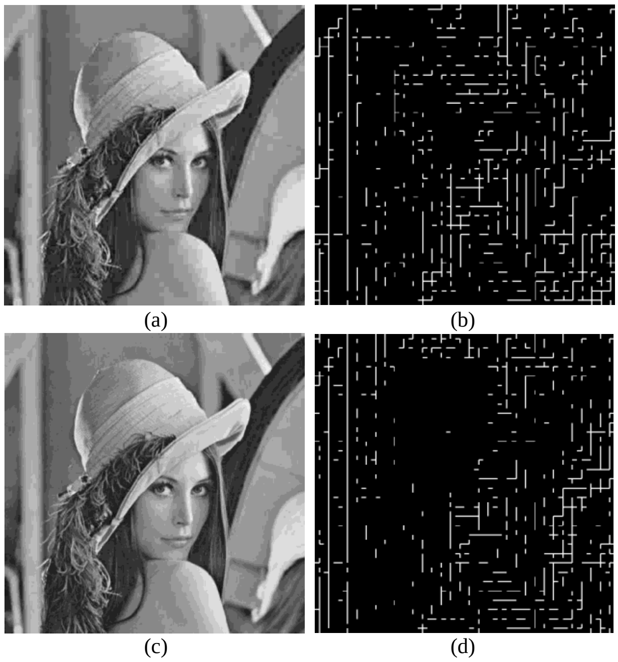


Fig. 7. Reconstructed Lena images with different q and blocking index images: (a) $q = 10$, (b) the blocking image of (a), (c) $q = 50$, and (d) the blocking image of (c).

Fig. 6 shows some subjective experimental results of Fig. 5. From Fig. 5, we can conclude that: (i) different images have different optimal residual thresholds; (ii) the larger quality factor is, the smaller the optimal residual threshold is. If we select a larger residual threshold, the deblocked image would be too smooth and some of the important information of edge and texture will lose as shown in Fig. 6(f). If the residual threshold is too small, there are still some blocking artifacts in the images as shown in Fig. 6(c).

To overcome this problem, we propose one post-processing deblocking method by using the blind image deblocking index [15]. For method proposed in [15], it can detect the blocking artifacts shown in Fig. 7. The higher quality of the image is, the smaller the artifacts index score is. After a lot of experiments, we select three kinds of values for the residual thresholds: 6, 8, and 10. For the blocked image, the proposed scheme with different residual thresholds is conducted, respectively. The deblocked image with the smallest artifacts index score is the final output image.

IV. Experimental Results

In this section, the deblocking performance of our proposed scheme is tested. All blocked images used in the test are compressed using the JPEG compression standard. The severity of blocking artifacts is mainly up to the quality factor q ranging from 0 to 100. The higher of the compression factor is, the better image quality would be. In general, the blocking artifacts become obviously visible when q is below 20. So four quality factors q: 10, 15, 20, and 25 are used. The proposed scheme is compared with two deblocking methods using conventional over-completed DCT dictionary and learned over-completed dictionary by K-SVD [8], respectively. The residual threshold of these two deblocking methods is 8. The size of dictionary is 64×256. In this paper, the performance of the proposed scheme and other methods are quantitatively measured by Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM) index.

The deblocking results for the JPEG-coded Lena, Barbara, and Peppers under different quality factors are shown in Table I and Table II. In Table I and Table II, JPEG represents the JPEG-coded images. DCT and K-SVD represent the image deblocking method by over-completed DCT dictionary and K-SVD dictionary, respectively. APDCBT and APIDCBT represent our proposed method. From Table I and Table II, our proposed method is better than the method of over-completed DCT dictionary or K-SVD dictionary in terms of PSNR or SSIM. The APDCBT and APIDCBT have almost the similar effects on image deblocking.

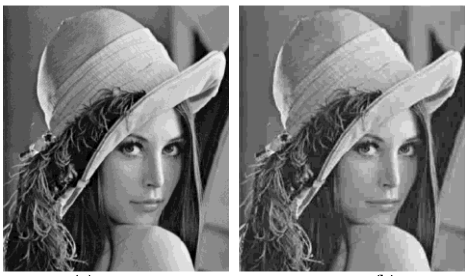
TABLE I: PSNR COMPARISONS OF DEBLOCKING METHODS

Images		**PSNR (dB)**				
		JPEG	**DCT**	**K-SVD**	**APDCBT**	**APIDCBT**
Lena	$q = 10$	30.41	31.58	31.78	31.64	31.65
	$q = 15$	31.95	32.86	33.10	32.99	32.99
	$q = 20$	32.96	33.59	33.80	33.85	33.85
	$q = 25$	33.70	34.06	34.26	34.41	34.42
Barbara	$q = 10$	25.70	26.33	26.45	26.43	26.43
	$q = 15$	27.05	27.71	27.86	27.75	27.74
	$q = 20$	28.25	28.91	29.12	28.98	28.98
	$q = 25$	29.31	29.95	30.21	30.02	30.03
Peppers	$q = 10$	30.14	31.29	31.51	31.37	31.37
	$q = 15$	31.54	32.42	32.69	32.50	32.51
	$q = 20$	32.43	33.01	33.30	33.19	33.19
	$q = 25$	33.06	33.43	33.69	33.67	33.68

TABLE II: SSIM COMPARISONS OF DEBLOCKING METHODS

Images		**SSIM**				
		JPEG	**DCT**	**K-SVD**	**APDCBT**	**APIDCBT**
Lena	$q = 10$	0.8760	0.9102	0.9126	0.9105	0.9105
	$q = 15$	0.9168	0.9292	0.9324	0.9367	0.9367
	$q = 20$	0.9380	0.9361	0.9393	0.9451	0.9451
	$q = 25$	0.9504	0.9408	0.9441	0.9508	0.9508
Barbara	$q = 10$	0.8811	0.9044	0.9055	0.8987	0.8988
	$q = 15$	0.9231	0.9330	0.9341	0.9315	0.9315
	$q = 20$	0.9460	0.9463	0.9478	0.9456	0.9457
	$q = 25$	0.9585	0.9536	0.9552	0.9582	0.9582
Peppers	$q = 10$	0.8809	0.9187	0.9198	0.9190	0.9191
	$q = 15$	0.9184	0.9341	0.9359	0.9381	0.9381
	$q = 20$	0.9381	0.9400	0.9422	0.9456	0.9457
	$q = 25$	0.9502	0.9442	0.9466	0.9506	0.9506

For comparison of visual quality, these deblocked results on partial enlarged Lena image are shown in Fig. 8. For color image, we use Peppers image, and the deblocked results are shown in Fig. 9. From Fig. 8 and Fig. 9, visual quality of the blocked image is improved by three deblocked methods. However, we can find that our method makes deblocked images more natural-looking. For example, the hairs are more vivid as shown in Fig. 8(e) and Fig. 8(f).



(a) (b)

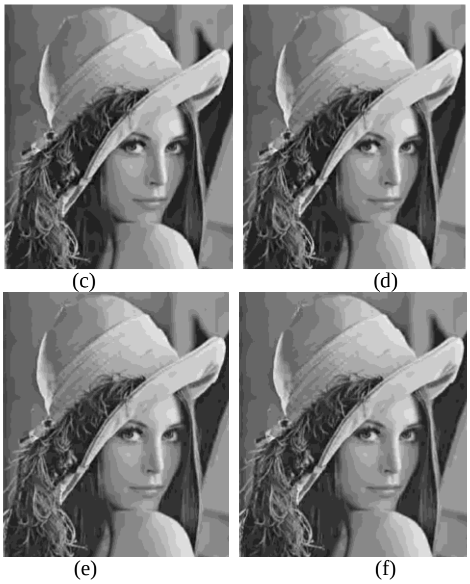


Fig. 8. Deblocked results of Lena: (a) the uncompressed image; (b) the compressed image with $q = 15$; (c), (d), (e), and (f) are the deblocked image by DCT, K-SVD, APDCBT, and APIDCBT, respectively.

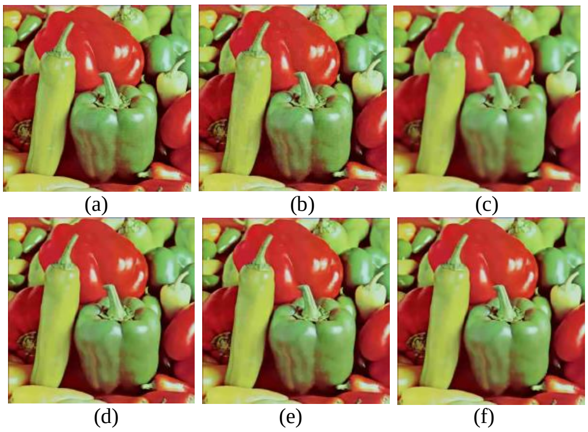


Fig. 9. Deblocked results of Peppers: (a) the uncompressed image; (b) the compressed image with $q = 15$; (c), (d), (e), and (f) are the deblocked image by DCT, K-SVD, APDCBT, and APIDCBT, respectively.

V. Conclusion

In this paper, we propose a new post-processing for JPEG-coded image deblocking via sparse representation and APBT. We use the APBT to generate the over-completed dictionary. By combining the blind image quality assessment, we select the adaptive residual thresholds. Compared with the learned over-completed dictionary by K-SVD, our proposed scheme has better performance. For the future work, we can improve the speed of our method using Graphic Processing Unit (GPU).

Acknowledgment

This work was supported by the National Natural Science Foundation of China (No. 61201371), the promotive research fund for excellent young and middle-aged scientists of Shandong Province, China (No. BS2013DX022), and the Natural Science Foundation of Shandong Province, China (No. ZR2015PF004). The authors thank Heng Zhang and Yunpeng Zhang for their kind help and valuable suggestions in revising this paper. The authors also thank the anonymous reviewers and the editors for their valuable comments to improve the presentation of the paper.

References

[1] G. K. Wallace, "The JPEG still picture compression standard," *IEEE Trans. on Consumer Electronics*, vol. 38, no. 1, pp. 18-34, Feb. 1992.

[2] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, Jul. 2003.

[3] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.

[4] A. Norkin, G. Bjøntegaard, A. Fuldseth, M. Narroschke, M. Ikeda, K. Andersson, M. H. Zhou, and G. Van der Auwera, "HEVC deblocking filter," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1746-1754, Dec. 2012.

[5] C. Wang, J. Zhou, and S. Liu, "Adaptive non-local means filter for image deblocking," *Signal Processing: Image Communication*, vol. 28, no. 5, pp. 522-530, May 2013.

[6] J. J. Zou and H. Yan, "A deblocking method for BDCT compressed images based on adaptive projections," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 3, pp. 430-435, Mar. 2005.

[7] C. Jung, L. C. Jiao, H. T. Qi, and T. Sun, "Image deblocking via sparse representation," *Signal Processing: Image Communication*, vol. 27, no. 6, pp. 663-677, Jul. 2012.

[8] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Trans. on Signal Processing*, vol. 54, no. 11, pp. 4311-4322, Nov. 2006.

[9] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Trans. on Information Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007.

[10] J. Zhang, S. W. Ma, Y. B. Zhang, and W. Gao, "Image deblocking using group-based sparse representation and quantization constraint prior," in *Proc. IEEE Int. Conf. on Image Processing*, Quebec City, QC, Canada, Sept. 27-30, 2015, pp. 306-310.

[11] C. Zhao, J. Zhang, S. W. Ma, X. P. Fan, Y. B. Zhang, and W. Gao, "Reducing image compression artifacts by structural sparse representation and quantization constraint prior," *IEEE Trans. on Circuits and Systems for Video Technology*, pp. 1-14, Jun. 2016. (in Press).

[12] Z. X. Hou, C. Y. Wang, and A. P. Yang, "All phase biorthogonal transform and its application in JPEG-like image compression," *Signal Processing: Image Communication*, vol. 24, no. 10, pp. 791-802, Nov. 2009.

[13] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Review*, vol. 43, no. 1, pp. 129-159, Mar. 2001.

[14] S. G. Mallat and Z. F. Zhang, "Matching pursuit with time-frequency dictionaries," *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp. 3397-3415, Dec. 1993.

[15] S. Z. Liu and A. C. Bovik, "Efficient DCT-domain blind measurement and reduction of blocking artifacts," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 12, no. 12, pp. 1139-1149, Dec. 2002.

Liping Wang was born in Shandong province, China, in 1990. She received her B.E. degree in communication engineering from Shandong University, Weihai, China, in 2014. She is currently pursuing her M.E. degree in signal and information processing at Shandong University, China. Her current research interests include digital image processing and analysis techniques (image dehazing, image deblocking, image quality assessment) and computer vision.

Chengyou Wang was born in Shandong province, China, in 1979. He received his B.E. degree in electronic information science and technology from Yantai University, China, in 2004, and his M.E. and Ph.D. degrees in signal and information processing from Tianjin University, China, in 2007 and 2010, respectively. He is currently an associate professor and supervisor of postgraduate students with Shandong University, Weihai, China. His current research interests include digital image/video processing and analysis (transform coding, digital watermarking, deblocking, demosaicking, dehazing, image quality assessment, etc.), computer vision (tamper detection, content-based image retrieval, etc.), pattern recognition, and machine learning.

Xiao Zhou was born in Shandong province, China, in 1982. She received her B.E. degree in automation from Nanjing University of Posts and Telecommunications, China, in 2003; her M.E. degree in information and communication engineering from Inha University, Korea, in 2005; and her Ph.D. degree in information and communication engineering from Tsinghua University, China, in 2013. She is currently a lecturer and supervisor of postgraduate students with Shandong University, Weihai, China. Her current research interests include wireless communication technology, digital image processing, and computer vision.