KSII TRANSACTIONS ON INTERNET AND INFORMATION SYSTEMS VOL. 11, NO. 3, Mar. 2017 1700
Copyright ⓒ2017 KSII

Post-Processing for JPEG-Coded Image Deblocking via Sparse Representation and Adaptive Residual Threshold

Liping Wang, Xiao Zhou*, Chengyou Wang, and Baochen Jiang
School of Mechanical, Electrical and Information Engineering, Shandong University, Weihai 264209, China
[e-mail: wlp@mail.sdu.edu.cn, zhouxiao@sdu.edu.cn, wangchengyou@sdu.edu.cn, jbc@sdu.edu.cn]
*Corresponding author: Xiao Zhou

Received June 6, 2016; revised September 27, 2016; revised November 14, 2016; accepted January 11, 2017; published March 31, 2017

Abstract

The problem of blocking artifacts is very common in block-based image and video compression, especially at very low bit rates. In this paper, we propose a post-processing method for JPEG-coded image deblocking via sparse representation and adaptive residual threshold. This method includes three steps. First, we obtain the dictionary by online dictionary learning and the compressed images. The dictionary is then modified by the histogram of oriented gradient (HOG) feature descriptor and K-means cluster. Second, an adaptive residual threshold for orthogonal matching pursuit (OMP) is proposed and used for sparse coding by combining blind image blocking assessment. At last, to take advantage of human visual system (HVS), the edge regions of the obtained deblocked image can be further modified by the edge regions of the compressed image. The experimental results show that our proposed method can keep the image more texture and edge information while reducing the image blocking artifacts.

Keywords: Image deblocking, sparse representation, adaptive residual threshold, orthogonal matching pursuit (OMP), human visual system (HVS)

This work was supported by the National Natural Science Foundation of China (No. 61201371), the Natural Science Foundation of Shandong Province, China (No. ZR2015PF004), the Research Award Fund for Outstanding Young and Middle-Aged Scientists of Shandong Province, China (No. BS2013DX022), and the Fundamental Research Funds of Shandong University, China (No. 2014ZQXM008).

https://doi.org/10.3837/tiis.2017.03.025 **ISSN : 1976-7277**

1. Introduction

Image and video have been becoming the main carriers of information. To adapt to the available and limited resource, such as a limited amount of transmitting bandwidth, images need to be compressed. Block discrete cosine transform (BDCT) has been widely used in current coding standards, such as JPEG [1], H.264/AVC [2], and H.265/HEVC [3]. However, BDCT coding technique usually produces blocking artifacts, especially at low bit rates. On the one hand, the blocking artifacts seriously affect the visual quality of the image. On the other hand, it affects the further image processing. Therefore, it is necessary to conduct image deblocking, especially for highly compressed images.

There are two main approaches for deblocking: in-loop processing and post-processing. In-loop processing conducts the process of image deblocking in the coding loop. Since post-processing is performed after image decoding, it is easily incorporated in any existing compression standards. Therefore, it is more practical for image deblocking. There are also two categories for post-processing methods [4], [5]: image enhancement based method and image restoration based method. The image deblocking method based on image enhancement does not consider the cause of blocking artifacts and only conducts image enhancement operation, such as spatial filtering [6], [7] and frequency domain filtering [8], [9].

For image deblocking method based on image restoration, we usually formulate the deblocking as an ill-posed inverse problem [10], such as projection onto convex sets (POCS) method [11], maximum a posteriori (MAP) estimation method [12], and the energy based method [13], [14]. Another likely method is using spare representation for deblocking [15]. The learning-based sparse representation has been widely and successfully used in Gaussian noise reduction [16], multiplicative noise removal [17], Poisson image deblurring [18], super-resolution [19], and so on. Similarly, many researchers have proposed image deblocking methods for the JPEG compressed images by sparse representation. Jung *et al*. [20] obtained a deblocking dictionary via the K-singular value decomposition (K-SVD) algorithm [21] and proposed an adaptive residual threshold for orthogonal matching pursuit (OMP) [22] in image deblocking. Instead of processing each image patch individually, Zhang *et al*. [10] proposed group-based sparse representation (GSR) and used GSR for image deblocking. Combined with the total variation regularization, Chang *et al*. [23] proposed a novel image deblocking method for the JPEG-coded images via sparse representation. These studies show that sparse representation is effective in image deblocking. Therefore, this paper will focus on image deblocking for JPEG-coded images via sparse representation.

For image deblocking via sparse representation, the proper over-completed dictionary and sparse coding algorithm are needed. An effective three-step algorithm for JPEG-coded images is propoesed in this paper. First, we obtain a over-completed dictionary for sparse representation using online dictionary learning algorithm [24] using the blocking image. The obtained dictionay is then decomposed into blocking sub-dictionary and non-blocking sub-dictionary via the K-means algorithm and histogram of oriented gradient (HOG) feature descriptor [25]. Second, for sparse coding algorithm, an adaptive residual threshold for OMP is proposed and used for sparse coding by combining blind image blocking assessment. At last, to take advantage of human visual system (HVS) [26], the edge regions of the obtained deblocked image can be further modified by the edge regions of the compressed image. Therefore, blocking artifacts are reduced by these steps.

In Section 2, we briefly show the JPEG compression model and review the concept of the sparse representation and dictionary learning techniques. In Section 3, we describe the proposed post-processing for JPEG-coded image via sparse representation and adaptive residual threshold in detail. Section 4 presents some experimental results and the corresponding analysis. Finally, we conclude the paper in Section 5.

2. Related Work

In this section, we show the model of JPEG compression and sparse representation. We also briefly review the concept of the over-completed dictionary learning techniques.

2.1 JPEG Compression Model

The image compression standard used here is JPEG [1], which contains three basic steps: BDCT, BDCT coefficient quantization, and Huffman entropy encoding. The decoding process is the inverse process of encoding. **Fig. 1** shows the basic processes of JPEG encoding and decoding.

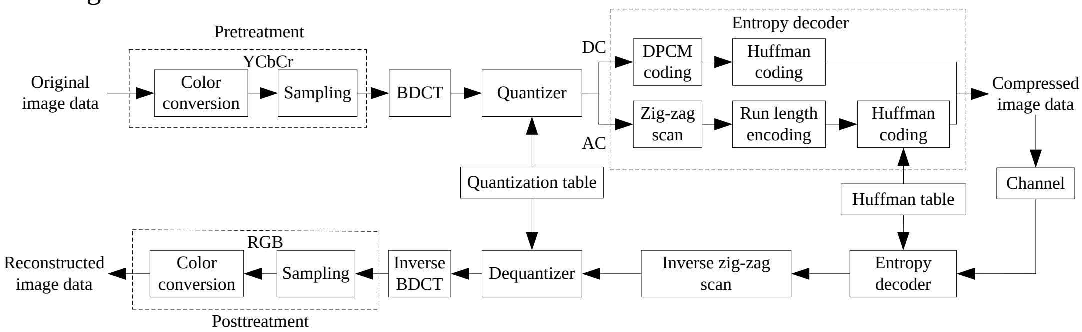


Fig. 1. Block diagram for JPEG compression.

The process of encoding is shown as follows:

Step 1. The input image is first converted into the YCbCr color space and then is 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 BDCT coefficients which are composed of one DC coefficient and 63 AC coefficients.

Step 4. Quantize each matrix block.

Step 5. For each block, the DC coefficient is performed a differential pulse code modulation (DPCM).The 63 coefficients are first converted into a 1-D zig-zag sequence, and then are conducted entropy encoding.

2.2 Sparse Representation Model

The sparse representation model is that every image patch $\boldsymbol{x} \in \boldsymbol{R}^n$ could be represented sparsely over a dictionary of size $\boldsymbol{D} \in \boldsymbol{R}^{n \times k}$,

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

where $\|\boldsymbol{\alpha}\|_0$ is the count of the nonzero entries in $\boldsymbol{\alpha}$. For the sparse representation model, the solution of Eq. (1) is indeed very sparse, $\|\hat{\boldsymbol{\alpha}}\|_0 \ll n$.

To make the sparse representation model more precise, we can replace the rough constraint $\boldsymbol{D\alpha} \approx \boldsymbol{x}$ with a more clear bounded representation error, $\| \boldsymbol{D\alpha} - \boldsymbol{x} \|_2 \leq \varepsilon$. The MAP estimator for image patch $\boldsymbol{y}$ denoising is established by solving

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \| \boldsymbol{\alpha} \|_0 \text{ subject to } \| \boldsymbol{D\alpha} - \boldsymbol{y} \|_2^2 \leq T, \quad (2)$$

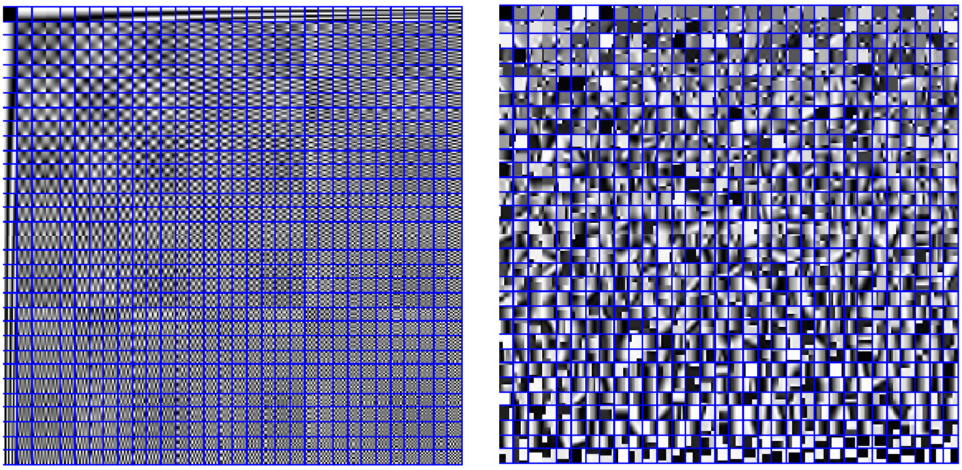
where $\boldsymbol{y}$ is obtained by the sparse signal $\boldsymbol{x}$ contaminated by the Gaussian noise with standard deviation σ. T is decided by ε and σ. The denoised image is obtained by $\hat{\boldsymbol{x}} = \boldsymbol{D}\hat{\boldsymbol{\alpha}}$.

In general, Eq. (2) is very hard to calculate. However, it can be efficiently solved by several available approximation techniques to get the sparse decomposition coefficients, such as basis pursuit (BP) [27], matching pursuit (MP) [28], and OMP [22]. In our work, we use OMP because of its simplicity and efficiency.

2.3 Over-Completed Dictionary Learning Techniques

Most recent algorithms for dictionary learning are iterative batch procedures such as the method of optimal directions (MOD) [29], the method proposed by Lewicki and Sejnowski [30], K-SVD [21], and the method proposed by Raina *et al*. [31]. To make a cost function smallest under some constraints, these methods use the whole training sets at each iteration. Therefore, these methods cannot handle very large training sets quickly and efficiently. Online dictionary learning algorithm [24] can solve this problem. Thus, in this paper, we adopt it to obtain the over-completed dictionary.

Fig. 2 shows the over-completed dictionaries of size 64×1024. **Fig. 2(a)** is the DCT dictionary, and **Fig. 2(b)** is obtained by the online dictionary learning which the input image is Lena image with compression factor $q = 10$.



(a) (b)

Fig. 2. The over-completed dictionaries: (a) DCT dictionary and (b) learned dictionary.

3. Our Proposed Method

The framework of our proposed method is shown in **Fig. 3**. Our method is divided into two parts: dictionary training processes and sparse coding with adaptive residual threshold.

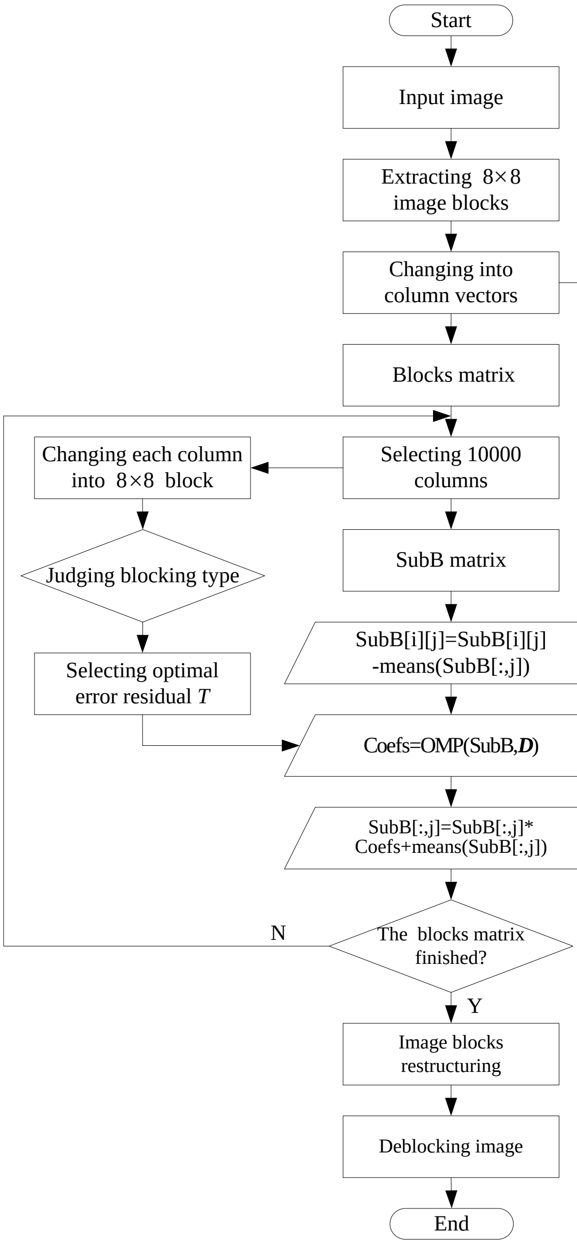


Fig. 3. Flow chart of our proposed algorithm.

3.1 Dictionary Training Processes

We first extract a series of overlapping image patches from the JPEG-coded images as the training samples for learning dictionary to obtain the over-completed dictionary $\boldsymbol{D}_{total}$. In general, the atoms in the dictionary have typical features in the original image. We obtain the dictionary by using the compressed images. Therefore, the learned dictionary will have the blocking artifacts. It is necessary to remove atoms including blocking artifacts and obtain dictionary without non-blocking component for the image deblocking.

The most striking feature of a blocking atom can be expressed by image gradient. In our paper, the modified HOG descriptor is used to describe the most striking feature of each atom in $\boldsymbol{D}_{total}$. Taking into account the characteristics of blocking artifacts, we only need to consider and use the vertical and horizontal gradients in each dictionary atom. Therefore, we use 72 bin histograms to count the gradient information of the atoms, and the 1st bin, 18th bin, 19th bin, 36th bin, 37th bin, 54th bin, 55th bin, and 72nd bin are gathered as the final HOG descriptor. That is, for horizontal HOG descriptor, we calculate $[355^{\circ}, 5^{\circ}]$ and $[175^{\circ}, 185^{\circ}]$; for the vertical HOG descriptor, we calculate $[85^{\circ}, 95^{\circ}]$ and $[265^{\circ}, 275^{\circ}]$. In this paper, we call this HOG descriptor as the blocking feature f_b.

The detailed process is described in the following parts and **Fig. 4** shows the $\boldsymbol{D}_{final}$ of the **Fig. 2(b)** by applying our proposed method.

Step 1. Obtain the over-completed dictionary $\boldsymbol{D}_{total}$ using the online dictionary learning algorithm;

Step 2. Calculate the blocking feature f_b for each atom in $\boldsymbol{D}_{total}$;

Step 3. To prevent dropping too much the texture and edge information of the image, we use the K-means algorithm to classify all of the atoms in $\boldsymbol{D}_{total}$ into three groups, $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, and $\boldsymbol{D}_3$, based on their blocking feature f_b;

Step 4. According to the centers of the three groups $\boldsymbol{D}_1$, $\boldsymbol{D}_2$, and $\boldsymbol{D}_3$, they can be divided into three categories: blocking sub-dictionary $\boldsymbol{D}_b$, little-blocking sub-dictionary $\boldsymbol{D}_{lb}$ and non-blocking sub-dictionary $\boldsymbol{D}_{nb}$. We combine $\boldsymbol{D}_{lb}$ and $\boldsymbol{D}_{nb}$ as the final learned dictionary $\boldsymbol{D}_{final}$.

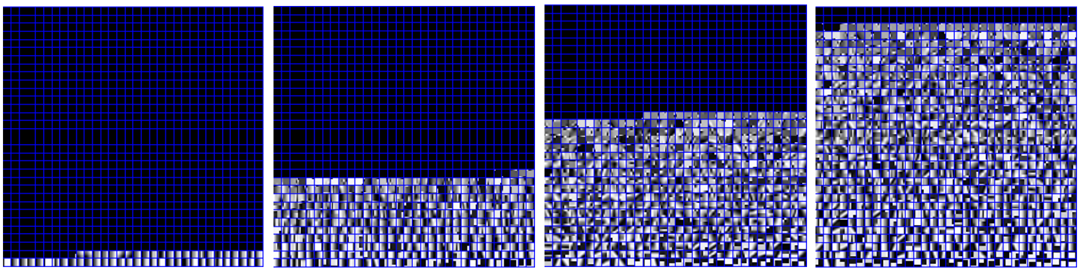

(a) (b) (c) (d)

Fig. 4. The dictionary learning and partition: (a) blocking sub-dictionary, (b) little-blocking sub-dictionary, (c) no-blocking sub-dictionary, and (d) final sub-dictionary.

3.2 Sparse Coding with Adaptive Residual Threshold

We use $\boldsymbol{D}_{final}$ for image deblocking. The objective function for image deblocking is shown by Eq. (2), but the optimization problem is a NP-hard problem. We usually use l_1 norm to approximately solve the problem as Eq. (3)

$$\hat{\boldsymbol{\alpha}} = \arg\min_{\boldsymbol{\alpha}} \| \boldsymbol{\alpha} \|_1 \text{ subject to } \| \boldsymbol{D\alpha} - \boldsymbol{y} \|_2^2 \leq T \,, \tag{3}$$

where $\boldsymbol{y}$ is the blocking image, and T is the residual threshold for OMP which is decided by ε and σ.

The residual threshold T is important to optimize Eq. (3). **Fig. 5** shows the objective experimental results which Lena and Barbara image at differernt compression factors are conducted image deblocking via sparse representation, where the dictionary is DCT dictionary of size 64×1024. **Fig. 6** and **Fig. 7** show some subjective experimental results of **Fig. 5**. From **Fig. 5**, we can find that: (i) different images have different optimal residual thresholds; (ii) one image with different compression factors has different optimal residual thresholds; (iii) the larger compression factor is, the smaller the optimal residual threshold would be. 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)** and **Fig. 7(f)**. If the residual threshold is too small, there are still some blocking artifacts in the images, as shown in **Fig. 6(c)** and **Fig. 7(c)**.

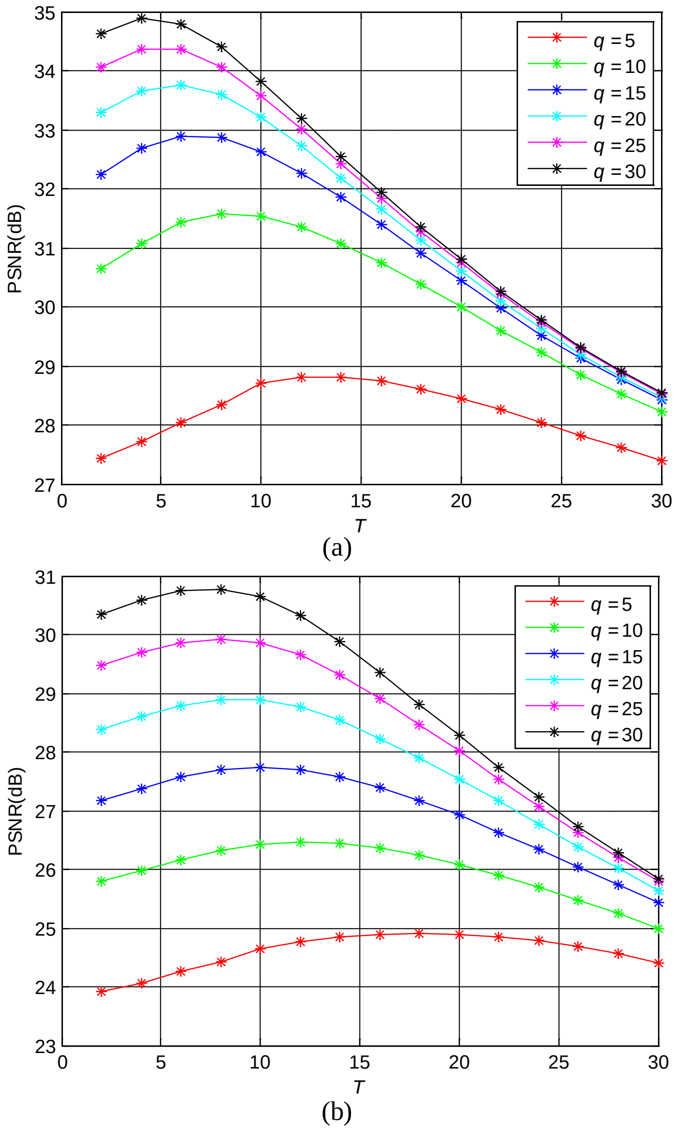


Fig. 5. The effects of residual threshold T on image deblocking: (a) Lena image and (b) Barbara image.

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.

In the previous work for image denoising [21], the T is decided by Eq. (4).

$$T = C \times \sigma, \tag{4}$$

where the σ is the standard deviation of quantization noise, and the noise gain C is set to be 1.15. However, the optimal residual threshold T_{block} for image deblocking is different from T for image denoising. To find the optimal residual threshold T_{block} for image deblocking, Jung *et al*. [20] modified Eq. (4) according to the compression factor q and the standard deviation of the entire image as follows:

$$T_{\text{Jung}} = C \times \sigma \times (\frac{a}{q+b} + c), \tag{5}$$

where a, b, and c are the control parameters, and their appropriate values are set to be 20, 10, and 0, respectively.

Fig. 7. The effects of residual threshold T on Barbara 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.

Eq. (5) is consistent with the experimental results of **Fig. 4**, that is, the larger the compression parameter is, the smaller the optimal residual threshold is. The compression factor is required to known in advance and we also need to obtain the standard deviation of quantization noise. To overcome these problems, we propose one pure adaptive residual threshold by using the blind image deblocking index [32] to modify Eq. (5). For method proposed in [32], it can detect the artifacts as shown in **Fig. 8**. The relationship between artifacts index score q_{index} and q is shown in **Fig. 9**. The higher quality of the image is, the smaller the artifacts index score q_{index} is. Therefore, we modify Eq. (5) as follows:

$$T_{base} = a_1 q_{index} + b_1 , \tag{6}$$

where a_1 and b_1 are the control parameters. Their appropriate values are given by experiments.

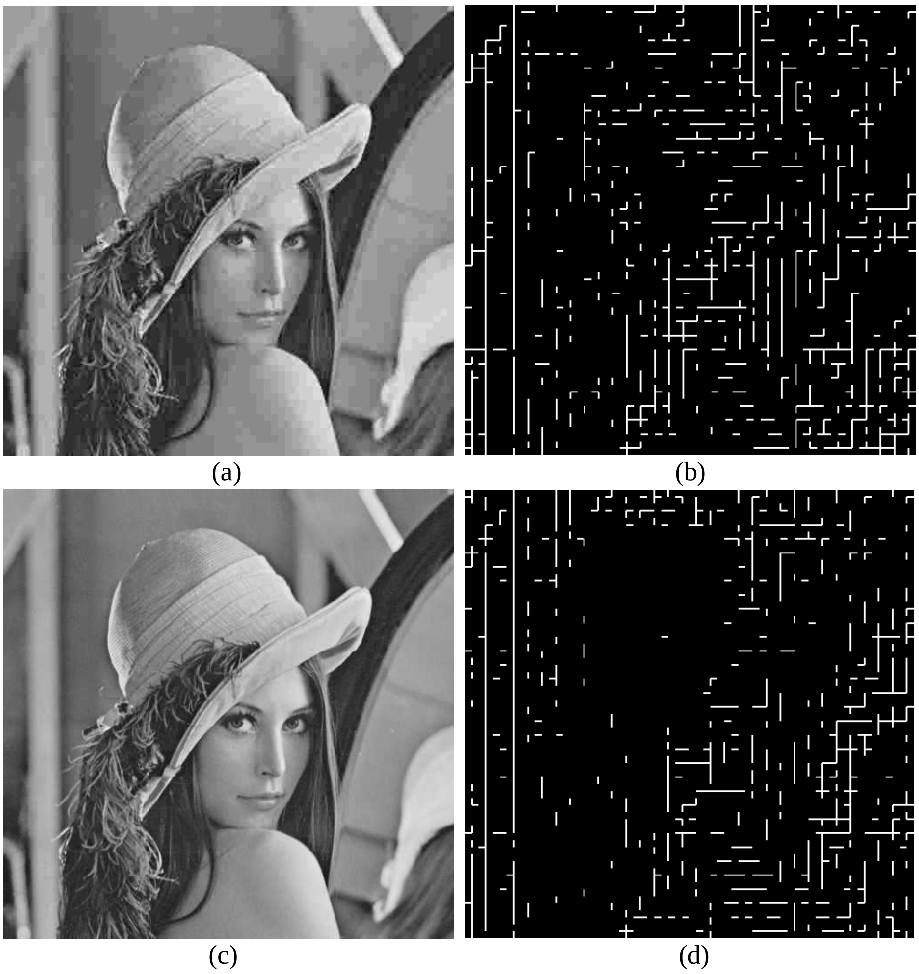


Fig. 8. Reconstructed Lena images with different q and blocking index images: (a) $q=10$, (b) the blocking image of (a), (c) $q=50$, and (d) the blocking image of (c).

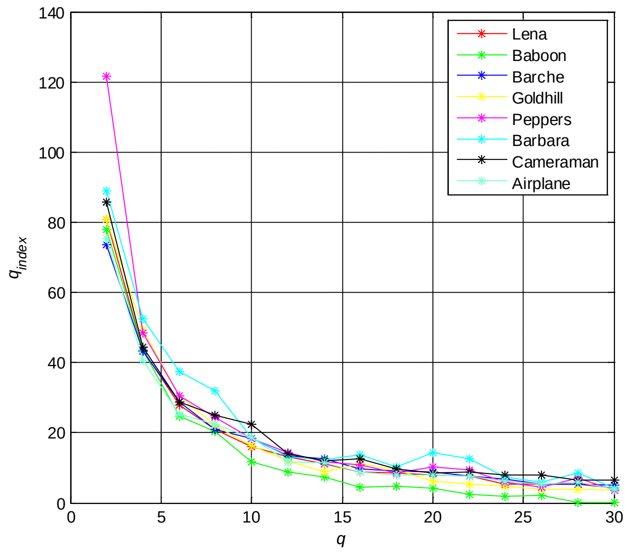


Fig. 9. The relationship between q and q_{index}.

HVS is the ultimate recipient of videos and images in our daily life. According to HVS models, our eyes can significantly perceive the artifacts such as blocking, and high-frequency noise in smooth regions. However, some of the texture distortion can be masked by texture. **Fig. 10** shows visual sensitivity based image classification. From **Fig. 10**, we can easily find the blocking artifacts in the image flat regions, but it is hard to find the blocking artifacts in the edge and texture regions. Therefore, based on these conceptions, the final edge regions of the deblocked image $I_{edgefinal}$ can be further modified by the edge regions of deblocked image $I_{deblocking}$ and the orginal image $I_{edgeblocking}$ as:

$$I_{edgefinal}(i,j) = w(i,j) I_{edgedeblocking} + (1 - w(i,j)) I_{edgeblocking} , \tag{7}$$

where w and $(1-w)$ denote the weights for $I_{edgedeblocking}$ and $I_{edgeblocking}$, respectively.

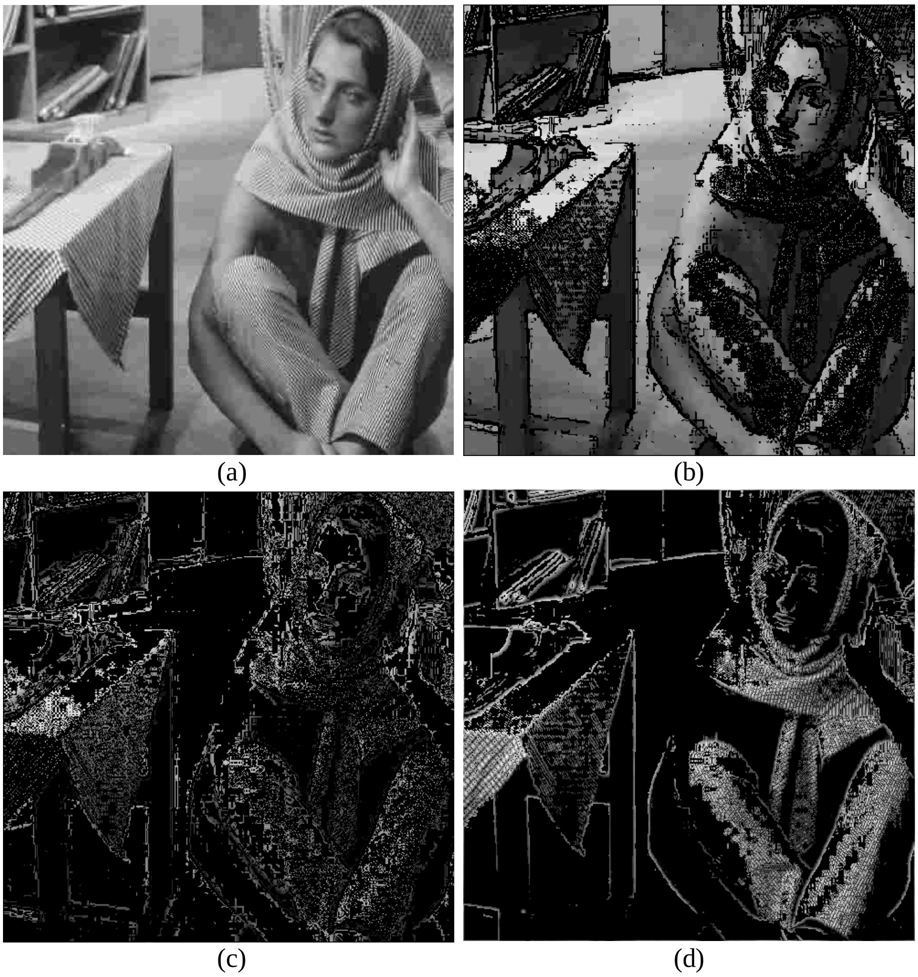


Fig. 10. Visual sensitivity based image classification: (a) the original image, (b) smooth-pixel image, (c) texture-pixel image, and (d) edge-pixel image.

4. Experimental Results and Analysis

In this section, we test the deblocking performance of our proposed scheme. These blocked images are JPEG-coded images. The intensity of blocking artifacts in JPEG-coded image is largely determined by the compression factor q, whose range is from 0 to 100. The larger of the compression factor is, the better the image quality will be. Typically, when the compression factor q is less than 20, the blocking artifacts become noticeable. In this paper, the performance of the proposed scheme and other methods are quantitatively measured by the following indexes: peak signal-to-noise ratio (PSNR), structural similarity (SSIM) index [33], PSNR-HVS-M [34], and FSIM [35]. For all of them, the larger the value is, the better it is. In this paper, hardware configuration of computer is Intel (R) Core (TM) i5-4690 4 core CPU, 3.50 GHz, 8 GB memory. The simulation is conducted with MATLAB 2014a.

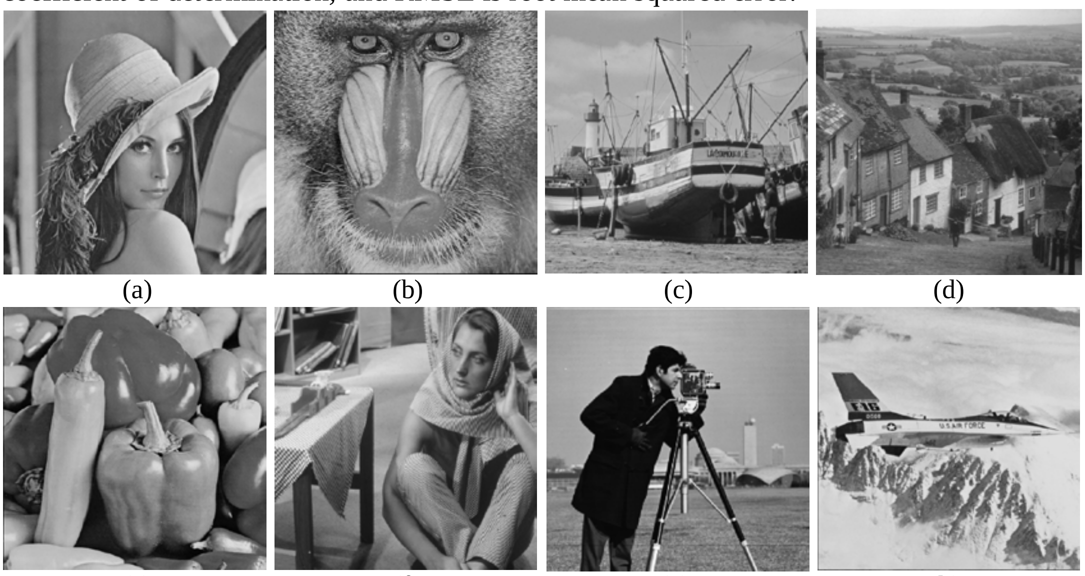
4.1 Automatic Estimation of the Residual Threshold

To find a suitable residual threshold for different types of contents in Eq. (3), we have conducted several tests on eight standard test images: Lena, Baboon, Barche, Goldhill, Peppers, Barbara, Cameraman, and Airplane, which are listed in **Fig. 11**.

First, we seek the relationship between q and q_{index} to obtain the optimal a_1 and b_1 for Eq. (6). As shown in **Fig. 9**, q and q_{index} are inversely proportional relationship. Therefore, in this paper, we use the linear fitting toolbox in MATLAB to obtain their expression:

$$q_{index} = \frac{170.8}{q + 0.1027}. \tag{8}$$

Second, to determine appropriate values a_1 and b_1, we have tested the PSNR performance of reconstructed images respect to different residual threshold T_{base} and compression factor q. The compression factors q are 5, 7, 10, 12, and 15. These experimental results are shown in **Table 1** - **Table 5**. In these tables, the bold numbers represent the best threshold values. Combined with Eq. (6) and Eq. (8), we use the best threshold values for eight images as shown in **Table 6** and the average of these values are used as the optimal $a_1 = 0.3203$ and $b_1 = 4.2123$. In **Table 6**, SSE is the sum of squares due to error mean squared error (MSE); R-square is coefficient of determination; Adjusted R-square is degree-of-freedom adjusted coefficient of determination; and RMSE is root mean squared error.



(a) (b) (c) (d)
(e) (f) (g) (h)

Fig. 11. Eight standard test images.

Table 1. The PSNR results of different T for $q=5$ ($q_{index}=33.4725$)

Image	$T=13$	$T=14$	$T=15$	$T=16$	$T=17$	$T=18$	$T=19$	**JPEG**
Lena	**28.82**	28.81	28.78	28.74	28.68	28.60	28.53	27.33
Baboon	22.08	22.09	**22.09**	22.09	22.07	22.05	22.01	21.52
Barche	**27.19**	27.18	27.16	27.11	27.06	26.99	26.92	26.03
Goldhill	**27.28**	27.25	27.21	27.13	27.07	26.99	26.90	26.16
Peppers	28.79	28.81	**28.81**	28.79	28.78	28.73	28.65	27.18
Barbara	24.81	24.85	24.88	24.90	24.91	**24.91**	24.91	23.86
Cameraman	24.95	24.97	24.99	25.01	25.01	**25.02**	25.01	24.21
Airplane	27.86	**27.88**	27.86	27.80	27.79	27.74	27.68	26.54

Table 2. The PSNR results of different T for $q=7$ ($q_{index}=23.9345$)

Image	$T=9$	$T=10$	$T=11$	$T=12$	$T=13$	$T=14$	$T=15$	**JPEG**
Lena	30.19	30.24	**30.25**	30.23	30.18	30.11	30.02	28.89
Baboon	22.89	22.92	22.93	22.94	**22.94**	22.93	22.90	22.47
Barche	28.47	28.51	**28.51**	28.48	28.43	28.345	28.26	27.46
Goldhill	28.33	**28.33**	28.28	28.21	28.12	28.03	27.92	27.43
Peppers	30.00	30.06	**30.08**	30.07	30.04	30.00	29.94	28.68
Barbara	25.45	25.51	25.56	25.59	25.61	25.62	**25.62**	24.76
Cameraman	25.93	25.97	25.60	26.01	**26.01**	26.01	26.00	25.33
Airplane	29.26	29.34	29.37	**29.39**	29.37	29.34	29.21	28.11

Table 3. The PSNR results of different T for $q=10$ ($q_{index}=16.9064$)

Image	$T=7$	$T=8$	$T=9$	$T=10$	$T=11$	$T=12$	$T=13$	**JPEG**
Lena	31.53	**31.57**	31.57	31.53	31.46	31.36	31.23	30.41
Baboon	23.74	23.78	23.81	**23.82**	23.81	23.79	23.76	23.43
Barche	29.76	29.81	**29.81**	29.77	29.70	29.58	29.44	28.86
Goldhill	**29.39**	29.38	29.31	29.23	29.12	28.98	28.84	28.65
Peppers	31.24	31.28	31.29	**31.30**	31.24	31.16	31.06	30.14
Barbara	26.25	26.32	26.38	26.43	26.46	**26.47**	26.47	25.70
Cameraman	26.76	26.80	26.85	26.87	**26.87**	26.84	26.80	26.26
Airplane	30.78	30.86	30.92	**30.93**	30.92	30.81	30.63	29.77

Table 4. The PSNR results of different T for $q=12$ ($q_{index}=14.1126$)

Image	$T=5$	$T=6$	$T=7$	$T=8$	$T=9$	$T=10$	$T=11$	**JPEG**
Lena	31.98	32.10	32.16	**32.17**	32.13	32.05	31.93	31.09
Baboon	24.13	24.18	24.23	24.26	**24.27**	24.27	24.25	23.90
Barche	30.23	30.35	30.41	**30.43**	30.39	30.31	30.18	29.53
Goldhill	29.86	**29.92**	29.91	29.84	29.74	29.61	29.45	29.23
Peppers	31.65	31.77	31.84	**31.85**	31.84	31.79	31.71	30.79
Barbara	26.65	26.75	26.82	26.89	26.94	26.97	**26.98**	26.26
Cameraman	27.27	27.35	27.41	27.45	27.47	**27.48**	27.46	26.92
Airplane	31.31	31.45	31.55	31.62	**31.63**	31.58	31.48	30.50

Table 5. The PSNR results of different T for $q=15$ ($q_{index}=11.3092$)

Image	$T=4$	$T=5$	$T=6$	$T=7$	$T=8$	$T=9$	$T=10$	JPEG
Lena	32.68	32.82	32.89	**32.90**	32.86	32.77	32.6	31.95
Baboon	24.68	24.74	24.78	24.82	24.84	**24.84**	24.83	24.50
Barche	30.94	31.08	31.16	**31.20**	31.17	31.09	30.95	30.36
Goldhill	30.48	30.57	**30.59**	30.53	30.42	30.27	30.08	29.95
Peppers	32.23	32.36	32.43	**32.45**	32.43	32.38	32.23	31.54
Barbara	27.38	27.48	27.57	27.64	27.70	27.73	**27.75**	27.05
Cameraman	27.93	28.01	28.09	28.14	28.16	**28.17**	28.16	27.65
Airplane	32.11	32.28	32.41	32.47	**32.48**	32.43	32.29	31.44

Table 6. The parameters for the eight test images

Image	a_1	b_1	SSE	R-square	Adjusted R-square	RMSE
Lena	0.2786	3.844	0.6241	0.9752	0.967	0.4561
Baboon	0.297	5.276	0.8662	0.9699	0.9599	0.5373
Barche	0.269	4.235	0.2888	0.9876	0.9834	0.3103
Goldhill	0.3397	1.624	0.6552	0.9824	0.9765	0.4673
Peppers	0.3448	3.323	1.153	0.9703	0.9604	0.6201
Barbara	0.3669	5.881	0.1516	0.9965	0.9953	0.2248
Cameraman	0.3971	4.28	0.8645	0.983	0.9773	0.5368
Airplane	0.269	5.235	0.2888	0.9876	0.9834	0.3103

4.2 Image Deblocking Results

Several JPEG-decoded images are used to test the performance of the proposed deblocking algorithm. Two other methods are used for comparisons with our method. The two deblocking methods are: the over-completed DCT dictionary and the over-completed learned K-SVD dictionary. The deblocking results for the JPEG-coded Lena, Barbara, Peppers, Baboon, Barche, and Fruits according to different compression factors are as shown in **Table 7** and **Table 8**.

From **Table 7** and **Table 8**, our proposed method is better than the method of over-completed DCT dictionary and the over-completed learned K-SVD dictionary in terms of the PSNR, SSIM, PSNR-HVS-M, or FSIM.

We also give deblocked results of the Monarch and Parrots in **Fig. 12** and **Fig. 13**, respectively. The visual quality of the compressed image is greatly improved in the three deblocking methods. In **Fig. 12** and **Fig. 13**, we add blue rectangle to show the image deblocking results for smooth regions. For the red rectangle, it indicates the image deblocking results for non-smooth regions. For the smooth regions, the blockness can be removed by three methods. For non-smooth regions, the detail information can be protected. However, our proposed method not only smooths the blocking, but also can protect the edge information better. Our proposed method can keep more information of the original images and our method produces more natural-looking images. For example, the feathers of parrot are more vividly as shown in **Fig. 13(d)**.

Table 7. Performance evaluation results on test images (Lena, Barbara, and Peppers)

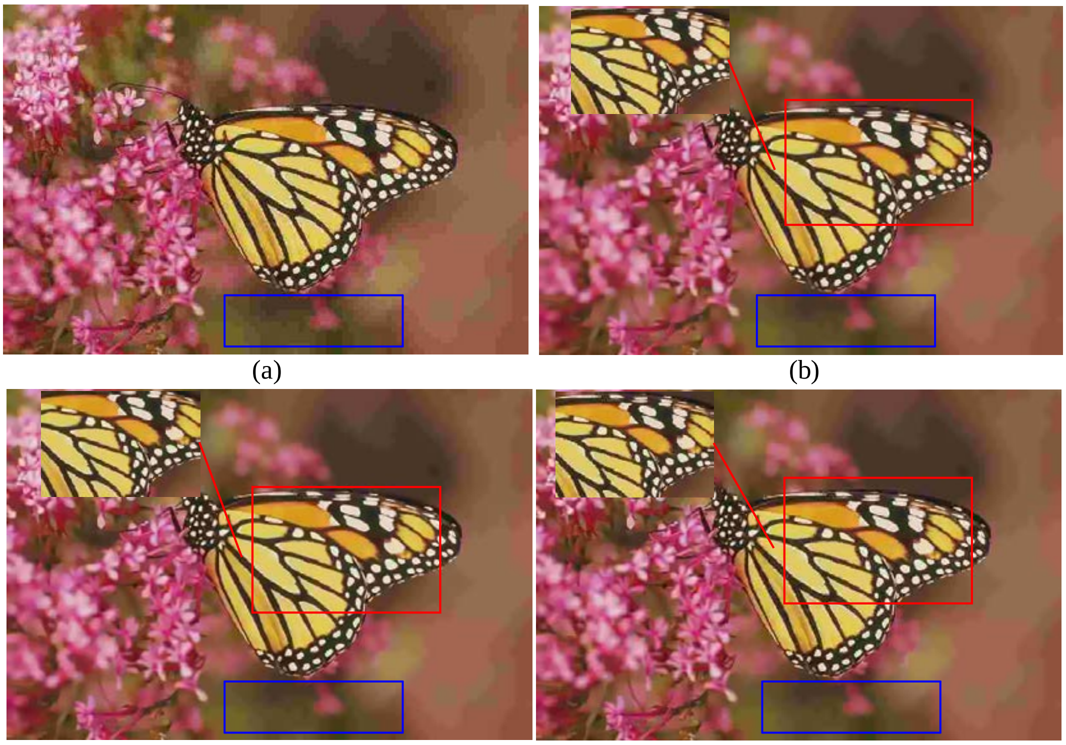
Image	Factor	Metric	JPEG	Over-completed DCT	K-SVD method	Proposed without weights	Proposed
Lena	q =5	PSNR (dB)	27.32	28.75	28.89	28.74	28.62
		SSIM	0.7660	0.8252	0.8286	0.8258	0.8260
		PSNR-HVS-M (dB)	23.92	25.66	25.88	25.69	25.61
		FSIM	0.8614	0.9138	0.9143	0.9128	0.9131
	q =9	PSNR (dB)	29.94	30.61	31.07	31.19	31.07
		SSIM	0.8590	0.8782	0.8835	0.8964	0.8962
		PSNR-HVS-M	28.14	28.47	29.17	29.93	29.85
		FSIM	0.9265	0.9404	0.9431	0.9496	0.9495
	q =11	PSNR (dB)	30.77	31.01	31.55	31.96	31.85
		SSIM	0.8879	0.8897	0.8954	0.9160	0.9159
		PSNR-HVS-M (dB)	29.71	29.05	29.84	31.47	31.42
		FSIM	0.9440	0.9444	0.9472	0.9597	0.9597
Barbara	q =5	PSNR (dB)	23.86	24.90	25.01	25.01	24.87
		SSIM	0.7681	0.8116	0.8156	0.8083	0.8092
		PSNR-HVS-M (dB)	22.88	24.60	24.87	24.57	24.50
		FSIM	0.8590	0.9145	0.9160	0.9114	0.9115
	q =9	PSNR (dB)	25.41	26.17	26.46	26.40	26.26
		SSIM	0.8674	0.8717	0.8801	0.8882	0.8889
		PSNR-HVS-M (dB)	27.10	27.53	28.23	28.61	28.59
		FSIM	0.9255	0.9395	0.9430	0.9478	0.9480
	q =11	PSNR (dB)	25.98	26.60	26.97	26.73	26.61
		SSIM	0.8924	0.8836	0.8928	0.9135	0.9131
		PSNR-HVS-M (dB)	28.81	28.30	29.17	30.41	30.32
		FSIM	0.9414	0.9437	0.9477	0.9579	0.9576
Peppers	q =5	PSNR (dB)	27.18	28.81	29.01	28.87	29.46
		SSIM	0.7788	0.8510	0.8532	0.8519	0.8575
		PSNR-HVS-M (dB)	24.01	26.32	26.57	26.39	27.23
		FSIM	0.8564	0.9255	0.9260	0.9249	0.9295
	q =9	PSNR (dB)	29.69	30.52	31.10	31.07	31.31
		SSIM	0.8678	0.9002	0.9044	0.9100	0.9115
		PSNR-HVS-M (dB)	28.24	29.08	29.91	30.48	30.97
		FSIM	0.9249	0.9478	0.9516	0.9558	0.9575
	q =11	PSNR (dB)	30.47	30.90	31.59	31.70	31.63
		SSIM	0.8901	0.9091	0.9135	0.9223	0.9222
		PSNR-HVS-M (dB)	29.77	29.66	30.60	31.79	31.76
		FSIM	0.9428	0.9518	0.9555	0.9631	0.9631
Average		PSNR (dB)	27.85	28.70	29.07	29.07	29.08
		SSIM	0.8419	0.8689	0.8741	0.8814	0.8823

	PSNR-HVS-M (dB)	26.95	27.63	28.25	28.82	28.92
	FSIM	0.9091	0.9357	0.9383	0.9426	0.9433

Table 8. Performance evaluation results on test images (Baboon, Barche, and Fruits)

Image	Factor	Metric	JPEG	Over-completed DCT	K-SVD method	Proposed without weights	Proposed
Baboon	q =5	PSNR (dB)	21.52	22.09	22.12	22.05	21.98
		SSIM	0.7328	0.7260	0.7317	0.7379	0.7393
		PSNR-HVS-M (dB)	20.95	21.99	22.09	22.02	21.91
		FSIM	0.8577	0.8671	0.8685	0.8730	0.8743
	q =9	PSNR (dB)	23.14	23.40	23.52	23.50	23.45
		SSIM	0.8585	0.8115	0.8211	0.8638	0.8635
		PSNR-HVS-M (dB)	25.77	26.19	26.64	26.64	26.55
		FSIM	0.9219	0.9034	0.9082	0.9330	0.9329
	q =11	PSNR (dB)	23.67	23.80	23.95	24.02	23.96
		SSIM	0.8880	0.8290	0.8391	0.8909	0.8907
		PSNR-HVS-M (dB)	27.70	26.04	26.61	28.51	28.44
		FSIM	0.9453	0.9112	0.9170	0.9472	0.9471
Barche	q =5	PSNR (dB)	26.03	27.13	27.27	27.17	27.06
		SSIM	0.7801	0.8133	0.8186	0.8199	0.8201
		PSNR-HVS-M (dB)	23.33	24.67	24.92	24.84	24.73
		FSIM	0.8634	0.8966	0.8984	0.8989	0.8992
	q =9	PSNR (dB)	28.43	28.76	29.21	29.48	29.37
		SSIM	0.8724	0.8650	0.8726	0.9010	0.9011
		PSNR-HVS-M (dB)	27.74	27.29	27.98	29.23	29.18
		FSIM	0.9326	0.9252	0.9296	0.9492	0.9494
	q =11	PSNR (dB)	29.19	29.13	29.68	30.21	30.12
		SSIM	0.8899	0.8739	0.8823	0.9168	0.9170
		PSNR-HVS-M (dB)	29.39	27.86	28.68	30.80	30.79
		FSIM	0.9459	0.9271	0.9320	0.9578	0.9579
Fruits	q =5	PSNR (dB)	27.19	28.52	28.77	28.62	28.55
		SSIM	0.7249	0.7937	0.7988	0.7926	0.7932
		PSNR-HVS-M (dB)	24.24	26.03	26.32	26.07	26.05
		FSIM	0.8270	0.8927	0.8949	0.8914	0.8920
	q =9	PSNR (dB)	29.56	30.06	30.68	30.80	30.73
		SSIM	0.8332	0.8462	0.8547	0.8727	0.8728
		PSNR-HVS-M (dB)	28.13	28.34	29.07	29.94	29.92
		FSIM	0.9083	0.9180	0.9234	0.9357	0.9359
	q =11	PSNR (dB)	30.28	30.41	31.12	31.44	31.38
		SSIM	0.8605	0.8563	0.8654	0.8889	0.8890
		PSNR-HVS-M	29.50	28.80	29.64	31.08	31.08

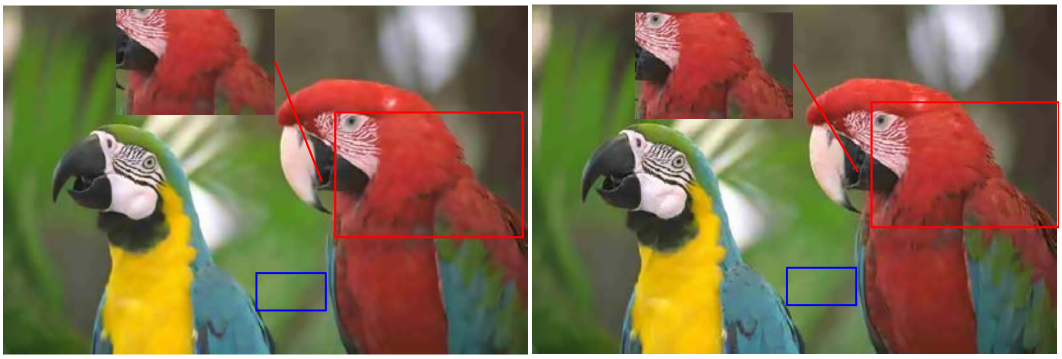
		(dB)					
		FSIM	0.9270	0.9220	0.9278	0.9447	0.9448
Average		PSNR (dB)	26.56	27.03	27.37	27.48	27.40
		SSIM	0.8267	0.8239	0.8316	0.8538	0.8541
		PSNR-HVS-M (dB)	26.31	26.36	26.88	27.68	27.63
		FSIM	0.9032	0.9070	0.9111	0.9257	0.9259



(a) (b) (c) (d)

Fig. 12. Deblocked results of Monarch: (a) the blocked image (PSNR: 26.85 dB, SSIM: 0.8372, PSNR-HVS-M: 23.23 dB, FSIM: 0.9315), (b) the over-completed DCT method (PSNR: 27.51 dB, SSIM: 0.8808, PSNR-HVS-M: 24.11 dB, FSIM: 0.9517), (c) K-SVD method (PSNR: 27.86 dB, SSIM: 0.8834, PSNR-HVS-M: 24.24 dB, FSIM: 0.9562), and (d) the proposed method (PSNR: 27.71 dB, SSIM: 0.8845, PSNR-HVS-M: 24.00 dB, FSIM: 0.9582).

(a) (b)



(c) (d)

Fig. 13. Deblocked results of Parrots: (a) the blocked image (PSNR: 31.85 dB, SSIM: 0.8642, PSNR-HVS-M: 28.72 dB, FSIM: 0.9659), (b) the over-completed DCT method (PSNR: 32.01 dB, SSIM: 0.8858, PSNR-HVS-M: 28.90 dB, FSIM: 0.9606), (c) K-SVD method (PSNR: 32.32 dB, SSIM: 0.8878, PSNR-HVS-M: 29.23 dB, FSIM: 0.9623), and (d) the proposed method (PSNR: 32.67 dB, SSIM: 0.8981, PSNR-HVS-M: 29.60 dB, FSIM: 0.9729).

The computational complexity of image deblocking is an important issue for decoders. For image deblocking using spare representation method, the proper over-compeleted dictionary and spare coding are important. For spare coding, OMP is all adopted by three methods. For the over-completed dictionary, the over-completed DCT dictionary, K-SVD dictionary, and online dictionary are used. Therefore, to compare the computational complexity, we give the time of generating the over-completed dictionaries based on processing time in **Table 9**. The less time is, the better the method is. All of the test images are uncompressed images. Our proposed method has lower computational complexity.

Table 9. Computation complexity based on processing time

Image	Size of dictionary	Processing time (s)		
		Over-completed DCT	K-SVD method	Proposed
Lena	64×256	4.89×10^{-3}	73.57	19.43
	64×512	5.21×10^{-3}	154.54	52.49
	64×1024	5.55×10^{-3}	417.98	149.92
Barbara	64×256	4.97×10^{-3}	73.88	16.72
	64×512	5.55×10^{-3}	154.74	43.44
	64×1024	5.36×10^{-3}	419.24	123.31
Peppers	64×256	4.91×10^{-3}	72.87	19.72
	64×512	5.71×10^{-3}	154.59	55.98
	64×1024	5.22×10^{-3}	416.47	157.43

5. Conclusion

In this paper, we propose an effective three-step deblocking algorithm for JPEG-coded image via sparse sepresentation and adaptive residual threshold. The first step involves dictionary learning and the second step involves error threshold constraint for the sparse coding algorithm. First, we create a over-completed dictionary for sparse representation using online dictionary learning algorithm based on the blocking image. The obtained dictionay is then decomposed into blocking sub-dictionary and non-blocking sub-dictionary via the K-means algorithm and HOG feature descriptor, and the non-blocking sub-dictionary is used for image deblocking. Second, an adaptive residual threshold for OMP is proposed and used for sparse coding by combining blind image blocking assessment. At last, to take advantage of HVS, the edge regions of the obtained deblocked image can be further modified by the edge regions of the compressed image.

In comparison with state-of-the-art techniques for image deblocking, the proposed method does not need to know compression factor in advance, which means that it is a pure post-processing approach. Our experimental results show that the proposed method achieves better performance on image deblocking. However, compared with some other image deblocking methods, it will spend a longer time. Therefore, we can improve the efficiency by parallel computing based on graphic processing unit (GPU) in the future.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (No. 61201371), the Natural Science Foundation of Shandong Province, China (No. ZR2015PF004), the Research Award Fund for Outstanding Young and Middle-Aged Scientists of Shandong Province, China (No. BS2013DX022), and the Fundamental Research Funds of Shandong University, China (No. 2014ZQXM008).

References

[1] G. K. Wallace, "The JPEG still picture compression standard," *Communications of the ACM*, vol. 34, no. 4, pp. 30-44, Apr. 1991. Article (CrossRef Link).

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

[3] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012. Article (CrossRef Link).

[4] M.-Y. Shen and C.-C. J. Kuo, "Review of postprocessing techniques for compression artifact removal," *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2-14, Mar. 1998. Article (CrossRef Link).

[5] C.-H. Yeh, L.-W. Kang, Y.-W. Chiou, C.-W. Lin, and S.-J. F. Jiang, "Self-learning-based post-processing for image/video deblocking via sparse representation," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 891-903, Jul. 2014. Article (CrossRef Link).

[6] G. T. Zhai, W. J. Zhang, X. K. Yang, W. S. Lin, and Y. Xu, "Efficient image deblocking based on postfiltering in shifted windows," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 18, no. 1, pp. 122-126, Jan. 2008. Article (CrossRef Link).

[7] 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. Article (CrossRef Link).

[8] G. T. Zhai, W. J. Zhang, X. K. Yang, W. S. Lin, and Y. Xu, "Efficient deblocking with coefficient regularization, shape-adaptive filtering, and quantization constraint," *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 735-745, Aug. 2008. Article (CrossRef Link).

[9] J. Kim, "Adaptive blocking artifact reduction using wavelet-based block analysis," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 2, pp. 933-940, May 2009. Article (CrossRef Link).

[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. of the IEEE International Conference on Image Processing*, Quebec City, QC, Canada, Sept. 27-30, pp. 306-310, 2015. Article (CrossRef Link).
[11] Y. Kim, C.-S. Park, and S.-J. Ko, "Fast POCS based post-processing technique for HDTV," *IEEE Transaction on Consumer Electronics*, vol. 49, no. 4, pp. 1438-1447, Nov. 2003. Article (CrossRef Link).
[12] H. Noda and M. Niimi, "Local MAP estimation for quality improvement of compressed color images," *Pattern Recognition*, vol. 44, no. 4, pp. 788-793, Apr. 2011. Article (CrossRef Link).
[13] K. Bredies and M. Holler, "A total variation-based JPEG decompression model," *SIAM Journal on Imaging Sciences*, vol. 5, no. 1, pp. 366-393, Jan. 2012. Article (CrossRef Link).
[14] K. Bredies and M. Holler, "Artifact-free decompression and zooming of JPEG compressed images with total generalized variation," in *Proc. of the 7th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, Rome, Italy, Feb. 24-26, pp. 242-258, 2012. Article (CrossRef Link).
[15] G. M. Farinella and S. Battiato, "On the application of structured sparse model selection to JPEG compressed images," in *Proc. of the 3rd International Workshop on Computational Color Imaging*, Milan, Italy, Apr. 20-21, 2011, pp. 137-151. Article (CrossRef Link).
[16] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736-3745, Dec. 2006. Article (CrossRef Link).
[17] Y.-M. Huang, L. Moisan, M. K. Ng, and T. Y. Zeng, "Multiplicative noise removal via a learned dictionary," *IEEE Transactions on Image Processing*, vol. 21, no. 11, pp. 4534-4543, Nov. 2012. Article (CrossRef Link).
[18] L. Y. Ma, L. Moisan, J. Yu, and T. Y. Zeng, "A dictionary learning approach for Poisson image deblurring," *IEEE Transactions on Medical Imaging*, vol. 32, no. 7, pp. 1277-1289, Jul. 2013. Article (CrossRef Link).
[19] J. C. Yang, J. Wright, T. Huang, and Y. Ma, "Image super-resolution as sparse representation of raw image patches," in *Proc. of the 26th IEEE Conference on Computer Vision and Pattern Recognition*, Anchorage, AK, USA, Jun. 23-28, 8 pages, 2008. Article (CrossRef Link).
[20] 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. Article (CrossRef Link).
[21] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *IEEE Transactions on Signal Processing*, vol. 54, no. 11, pp. 4311-4322, Nov. 2006. Article (CrossRef Link).
[22] J. A. Tropp and A. C. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4655-4666, Dec. 2007. Article (CrossRef Link).
[23] H. B. Chang, M. K. Ng, and T. Y. Zeng, "Reducing artifacts in JPEG decompression via a learned dictionary," *IEEE Transactions on Signal Processing*, vol. 62, no. 3, pp. 718-728, Feb. 2014. Article (CrossRef Link).
[24] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online learning for matrix factorization and sparse coding," *Journal of Machine Learning Research*, vol. 11, no. 1, pp. 19-60, Jan. 2010. Article (CrossRef Link).
[25] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Diego, CA, USA, Jun. 20-25, vol. 1, pp. 886-893, 2005. Article (CrossRef Link).
[26] Y. N. Zhang, B. T. Pham, and M. P. Eckstein, "The effect of nonlinear human visual system components on performance of a channelized Hotelling observer in structured backgrounds," *IEEE Transactions on Medical Imaging*, vol. 25, no. 10, pp. 1348-1362, Oct. 2006. Article (CrossRef Link).
[27] 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. Article (CrossRef Link).

[28] S. G. Mallat and Z. F. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Transactions on Signal Processing*, vol. 41, no. 12, pp. 3397-3415, Dec. 1993. Article (CrossRef Link).

[29] K. Engan, S. O. Aase, and J. H. Husoy, "Frame based signal compression using method of optimal directions (MOD)," in *Proc. of the IEEE International Symposium on Circuits and Systems*, Orlando, FL, USA, May 30 - Jun. 2, vol. 4, 4 pages, 1999. Article (CrossRef Link).

[30] M. S. Lewicki and T. J. Sejnowski, "Learning overcomplete representations," *Neural Computation*, vol. 12, no. 2, pp. 337-365, Feb. 2000. Article (CrossRef Link).

[31] R. Raina, A. Battle, H. Lee, B. Packer, and A. Y. Ng, "Self-taught learning: Transfer learning from unlabeled data," in *Proc. of the 24th International Conference on Machine Learning*, Corvalis, OR, USA, Jun. 20-24, pp. 759-766, 2007. Article (CrossRef Link).

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

[33] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transaction on Image Processing*, vol. 13, no. 4, pp. 600-612, Apr. 2004. Article (CrossRef Link).

[34] N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, and V. Lukin, "On between-coefficient contrast masking of DCT basis functions," in *Proc. of the 3rd International Workshop on Video Processing and Quality Metrics for Consumer Electronics*, Scottsdale, AZ, USA, Jan. 25-26, 4 pages, 2007. Article (CrossRef Link).

[35] L. Zhang, L. Zhang, X. Q. Mou, and D. Zhang, "FSIM: A feature similarity index for image quality assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378-2386, Aug. 2011. Article (CrossRef Link).

Liping Wang 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 computer vision (image dehazing, image deblocking, image quality assessment, etc.).

Xiao Zhou 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, image communication, and computer vision.

Chengyou Wang 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, demosaicking, deblocking, dehazing, image quality assessment, etc.), computer vision (tamper detection, content-based image retrieval, etc.), pattern recognition and machine learning, and wireless communication technology.

Baochen Jiang received his B.S. degree in radio electronics from Shandong University, China, in 1983 and his M.E. degree in communication and electronic systems from Tsinghua University, China, in 1990. He is currently a professor and supervisor of postgraduate students with Shandong University, Weihai, China. His current research interests include signal and information processing, digital image/video processing and analysis, and smart grid technology.