

BLOCKY ARTIFACT REMOVAL WITH LOW-RANK MATRIX RECOVERY

Ming Yin* Junbin Gao^{†*} Yanfeng Sun[‡] Shuting Cai*

* School of Automation, Guangdong University of Technology
Guangzhou 510006, China. email: yiming@gdut.edu.cn, shutingcai@gdut.edu.cn

[†] School of Computing and Mathematics, Charles Sturt University
Bathurst, NSW 2795, Australia. email: jbgao@csu.edu.au

[‡] Beijing Municipal Key Lab of Multimedia and Intelligent Software Technology
Beijing University of Technology, Beijing 100124, China. email: yfsun@bjut.edu.cn

ABSTRACT

In this paper, a novel image blocky artifact removal scheme based on low-rank matrix recovery is proposed. The problem of suppressing blocky artifacts is formulated as recovering a low-rank matrix from corrupted observations. During the deblocking processing, we do not directly recover the whole clean image but only its high-frequency component and then synthesize the clean image by incorporating the low-frequency component of blocky image. To take advantage of the low-rank matrix recovery paradigm, we first cluster the similar patches of the high-frequency component of image via local pixel clustering, then the clean high-frequency component of image is recovered by formulating an optimization problem of the nuclear norm and ℓ_1 -norm. The experimental results show that the proposed algorithm can achieve competitive performance in terms of both quantitative and subjective quality.

Index Terms— blocky artifact, sparse representation, low-rank matrix recovery, patch based.

1. INTRODUCTION

Blocky artifact is a common problem in image coding and video coding based on block-based discrete cosine transform (BDCT) [16] and often occurs along the block boundaries in compressed images and seriously degrades the subjective image quality [15, 17], especially at a low bit rate. In order to attenuate these artifacts and achieve superior perceptual image quality, many postprocessing techniques, named image deblocking, have been proposed in the literature [4, 9, 10, 13]. Dabov *et al.* [4] proposed a collaborative image denoising scheme by patch matching and sparse 3D transform, which is called the BM3D algorithm, achieving pleasing denoising performance. Sun and Cham [13] modeled an image as a

high order Markov random field based on the fields of experts framework and then deblocked the blocky image by solving an inverse problem.

Recently, sparse representation strategy has been widely studied to solve image restoration [8], especially on image denoising [1, 6]. In terms of sparse representation, image patches can be represented as a linear combination of a few atoms from a certain set, called dictionary \mathbf{D} . That is, each image patch is well approximated by the linear combination of a small subset of patches in the dictionary. Jung *et al.* [8] proposed a new image deblocking method for B-DCT compressed images based on sparse representation, in which a general dictionary is trained by K-singular value decomposition (K-SVD) algorithm. Yeh *et al.* [3] proposed a self-learning-based image/video deblocking framework via sparse representation, where a morphological component analysis (MCA) is used for image decomposition problem.

One of conventional ways to remove image noise is to apply the traditional principal component analysis (PCA), however PCA is not robust in removing the gross errors or outliers. To overcome this drawback, Candes *et al.* [2] proposed the so-called robust PCA by utilizing a convex program that guarantees to recover a low-rank matrix despite gross sparse errors. Recently, motivated by the advance of sparse representation and robust PCA, Ji *et al.* [7] proposed a new video restoration scheme based on the joint sparse and low-rank matrix approximation. Ren *et al.* [12] presented an image blocking artifacts reduction approach via patch clustering and low-rank minimization, however the effectiveness of the approach needs to be further investigated.

Actually, a blocky artifact is a particular class of data error that is in general caused by the quantization processing in lossy data compression. That is, the blocky artifacts can be seen as outliers occurring in compressed images. These outliers may significantly degrade the performance of the mentioned denoising methods. Our goal in this paper is to develop a novel deblocking approach which is able to effectively recover the clean information from corrupted data. Differen-

*The work is supported by the Australian Research Council (ARC) through the grant DP130100364 and the National Nature Science Foundation of China (NSFC) under grant No. 61370119, No. 61201392 and No.61133003.

t from the existing methods, we do not directly recover the whole clean image from the blocky image but only its high-frequency component and then synthesize the clean image by integrating with the low-frequency component of the blocky image. The recovery of high-frequency image component is formulated as a problem of minimizing the nuclear norm of a matrix with ℓ_1 -norm constraint [5], which can be efficiently solved by the Augmented Lagrange Multiplier (ALM) algorithm [11].

2. NOTATION

Denote by x_{ij} the (i, j) -th entry of matrix $X \in \mathbb{R}^{m \times n}$ and $\|X\|_1 = \sum_{i=1}^m \sum_{j=1}^n |x_{ij}|$ the ℓ_1 -norm of a matrix. Let σ_i be the non-negative singular values of X and the nuclear norm of X is defined as $\|X\|_* = \sum_{i=1}^r \sigma_i$. As usual, $\|X\|_F^2$ is the matrix Frobenius norm.

3. IMAGE DEBLOCKING BASED ON LOW-RANK MATRIX RECOVERY

Our proposed scheme is shown in Fig. 1. The critical step is to recover the true high-frequency component of the image from a single blocky image. Specifically, our approach is composed of two steps. First, the blocky image is decomposed into the low-frequency and the high-frequency components. Then, the high-frequency component including the blocky artifact is recovered by solving an optimization problem of the nuclear norm and ℓ_1 -norm constraint. In the end, the deblocked image is obtained by combining the low-frequency component with the recovered high-frequency one.

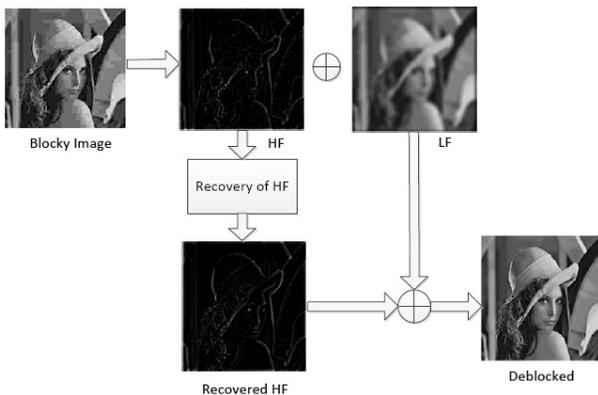


Fig. 1: Flowchart of our proposed deblocking scheme.

3.1. Problem Statement

It is easy to separate an image into its low-frequency and high-frequency components. The most basic information of

the image with blocky artifacts can be retained inside its low-frequency component while the blocky artifacts and the other edge/texture information are left in its high-frequency component. This procedure can be illustrated in Fig. 2. The procedure demonstrates that a high quality image can be recovered by synthesizing the low-frequency component of a blocky image and its high-frequency component of its clean image. When the blocky artifacts lying in the high-frequency component of blocky image can be well suppressed, the deblocking task can be handled successfully. Thus, our task for deblocking is transferred into recovering the high-frequency component information from a single blocky image. For simplicity, unless stated otherwise, we shall assume hereafter that the term *image* refers to its high-frequency component.



(a)



(b)

Fig. 2: Demonstration of using the true high-frequency image component for deblocking. a) deblocked image (left) and image with blockiness (right); b) high-frequency part of clean image (left) and blocky image (right).

3.2. Patch Matching Based on Local Pixel Clustering

Let Y be a blocky image, which is composed of non-blocky component X and the blocky artifact one E , i.e., $Y = X + E$. Consider a patch $\mathbf{p}_j = [p_{j1}, \dots, p_{jm}]^T$ centered at pixel j , then we search for N affinity patches $\{\mathbf{p}_j^n = [p_{j1}^n, \dots, p_{jm}^n]^T\}_{n=1}^N$ that are similar to \mathbf{p}_j in the whole image. The $m \times N$ patch matrix Y_j is obtained by arranging each patch vector \mathbf{p}_j^n into columns as follows,

$$Y_j = [\mathbf{p}_j^1, \dots, \mathbf{p}_j^N] \quad (1)$$

Thus, we can rewrite the image model into the form of patch matrices,

$$Y_j = X_j + E_j \quad (2)$$

Generally, the matrix constituted by the affinitive patches will have low-rank properties representing the underlying structure of the image. Thus, we apply block matching based on local pixel clustering to discriminate the similar patches. Assume that pixel j is centered in a $K \times K$ window, i.e., $m = K^2$. We want to find the similar patches in the training window $L \times L$ ($L > K$) containing the $K \times K$ window. Thus, there are in total $(L - K + 1)^2$ training samples for the patch centered at pixel j . Classifying the training samples can be easily computed as following,

$$\text{Err}_n(j) = \|\mathbf{p}_j - \mathbf{p}_j^n\|^2 = \frac{1}{m} \sum_{i=1}^m (p_{ji} - p_{ji}^n)^2 \quad (3)$$

In this paper, we set a threshold $\varepsilon\sigma^2$, in which σ is known as the standard deviation of artifacts, and the \mathbf{p}_j^n is picked out as the similar patch if $\text{Err}_n(j) < \varepsilon\sigma^2$ is satisfied.

3.3. Modeling Corruption as Sparse Error and Recovering

Mathematically, if we stack the patches from an image as column vectors to form a new matrix Y , then Y should be approximated by a low-rank matrix X and the residual part E can be modeled as sparse noise with most of its entries being zero. The low-rank matrix X representing the principal components is exactly the clean patches we want to restore from the blocky image Y . If $Y_j \in \mathbb{R}^{m \times N}$ is given as equation (1), we can surely find a sparse matrix E_j such that X_j has the lowest rank with high possibility. Thanks to the recent advances in the area of low-rank matrix recovery, we can achieve the goal by solving the following optimization problem for deblurring,

$$\min_{X_j, E_j} \|X_j\|_* + \lambda \|E_j\|_1 \quad \text{s.t. } Y_j = X_j + E_j \quad (4)$$

The above minimization problem will recover the low-rank matrix X from the blocky artifact image Y in which there exists not only sparse error but also quantization noise. As both $\|\cdot\|_*$ and $\|\cdot\|_1$ in (4) are convex, then the ALM method [11] can be easily adopted to efficiently solve problem (4). For problem (4), the augmented Lagrangian yields,

$$\mathcal{L}(X_j, E_j, \eta, \mu) = \|X_j\|_* + \lambda \|E_j\|_1 + \langle \eta, Y_j - X_j - E_j \rangle + \frac{\mu}{2} \|Y_j - X_j - E_j\|_F^2 \quad (5)$$

where $\eta \in \mathbb{R}^{m \times N}$ is a Lagrange multiplier matrix, $\mu > 0$, $\langle \cdot, \cdot \rangle$ represents the matrix inner-product.

The basic ALM algorithm can be given via iteratively estimating both the Lagrangian multiplier and the optimal solution,

$$\begin{aligned} (X_j^{k+1}, E_j^{k+1}) &= \underset{X_j, E_j}{\text{argmin}} \mathcal{L}(X_j, E_j, \eta_k, \mu_k) \\ \eta_{k+1} &= \eta_k + \mu_k (Y_j - X_j - E_j) \\ \mu_{k+1} &= \rho \cdot \mu_k \end{aligned} \quad (6)$$

where $\{\mu_k\}$ is a monotonically increasing positive sequence ($\rho > 1$). For the first subproblem of Eq (6), the solution is obtained by performing an alternating minimization process:

$$\begin{aligned} E_j^{k+1} &= \underset{E}{\text{argmin}} \lambda \|E\|_1 - \langle \eta_k, E \rangle + \frac{\mu_k}{2} \|Y_j - X_j^k - E\|_F^2 \\ X_j^{k+1} &= \underset{X}{\text{argmin}} \|X\|_* - \langle \eta_k, X \rangle + \frac{\mu_k}{2} \|Y_j - X - E_j^{k+1}\|_F^2 \end{aligned} \quad (7)$$

Now, the solution for E_j^{k+1} is achieved by applying the shrinkage operator,

$$E_j^{k+1} = \mathcal{S}_{\lambda\mu_k^{-1}}(Y_j - X_j^k + \mu_k^{-1}\eta_k) \quad (8)$$

where the shrinkage operator is defined as

$$\mathcal{S}_\alpha(X) = (|X| - \alpha)_+ \text{sgn}(X).$$

And the solution for X_j^{k+1} is iteratively obtained based on Accelerated Proximal Gradient (APG) [11] method as follows,

$$\begin{aligned} (U_i, S_i, V_i) &= \text{SVD}(Y_j + \mu_k^{-1}\eta_k - E_j + Z_i) \\ X_j^{k+1} &= U_i \mathcal{S}_{\mu_k^{-1}}[S_i] V_i^T \\ t_{i+1} &= \left(1 + \sqrt{1 + 4t_i^2}\right) / 2 \\ Z_{i+1} &= X_j^{k+1} + \frac{t_i - 1}{t_{i+1}} (X_j^{k+1} - X_j^k) \end{aligned} \quad (9)$$

Therefore, the main procedure of the optimization solving algorithm using ALM is summarized as follows.

Algorithm 1: Deblurring via Low-rank Matrix Recovery

Input: $Y_j \in \mathbb{R}^{m \times n}$, $\lambda > 0$, $\rho > 1$

Initialization: $\eta_1 = 0$; $X_j^1 = 0$; $E_j^1 = 0$; $\mu_1 = 1$; $\rho = 1.5$; $\lambda = 1/\sqrt{\max(m, n)}$.

Output: $(X_j^*, E_j^*) = (X_j^k, E_j^k)$

While not converged ($k = 1, 2, \dots$) **do**

$E_j^{k+1} = \mathcal{S}_{\lambda\mu_k^{-1}}[Y_j + \mu_k^{-1}\eta_k - X_j^k]$,

$t_1 = 1, Z_1 = X_j^k$

While not converged ($i = 1, 2, \dots$) **do**

$(U_i, S_i, V_i) = \text{SVD}(Y_j + \mu_k^{-1}\eta_k - E_j^{k+1} + Z_i)$,

$X_j^{k+1} = U_i \mathcal{S}_{\mu_k^{-1}}[S_i] V_i^T$,

$t_{i+1} = \left(1 + \sqrt{1 + 4t_i^2}\right) / 2$,

$Z_{i+1} = X_j^{k+1} + \frac{t_i - 1}{t_{i+1}} (X_j^{k+1} - X_j^k)$

End while

$\eta_{k+1} = \eta_k + \mu_k (Y_j - X_j - E_j)$,

$\mu_{k+1} = \rho \cdot \mu_k$,

End while

3.4. Estimation of Threshold

Actually, we can test the convergence by $\|Y_j - X_j - E_j\|_F^2 \leq \epsilon$ [11]. The iterative process can be stopped by a threshold denoted by Th , which is related to the standard deviation

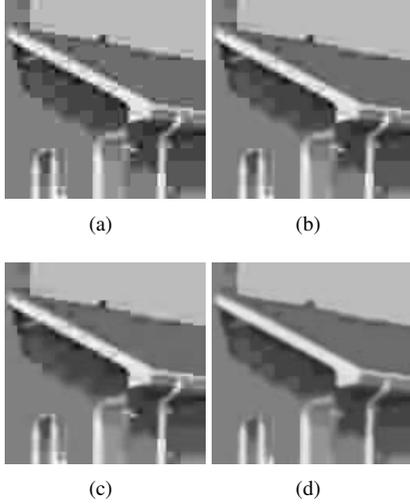


Fig. 3: The deblocking results of *House* subimage by different schemes at $q = 5$. (a) JPEG: PSNR=26.40dB, SSIM=0.75, GBIM=7.25; (b) BM3D: PSNR=27.35dB, SSIM=0.79, GBIM=3.94; (c) FoE: PSNR=27.56dB, SSIM=0.80, GBIM=3.78; (d) Proposed: PSNR=27.66dB, SSIM=0.80, GBIM=1.75

of the quantization noise. Here, we apply the intensity divergence between two boundary pixels to estimate the threshold, located on both sides of a boundary between two blocks [10]. In this paper, only the vertical block discontinuities are considered,

$$Th \propto |I(:, 8 \times j + 1) - I(:, 8 \times j)| \quad (10)$$

4. EXPERIMENTAL RESULTS

We perform the experiments on several grayscale images *House*, *Lena*, *Monarch*, *Barbara*, *Boat*, *Peppers*, *Baboon*, *Fruits* and *Cameraman*, whose sizes range from 256×256 to 512×512 pixels. JPEG standard quality factor (QF) is applied to measure the quality of the compressed images, and we perform the experiments with QF values from 0 to 20 where most of blocking artifacts occur in. As the prior work [8], three typical quantization tables, denoted as Q1, Q2 and Q3, have been commonly used for image compression, which correspond to QF parameters of $q = 11$, $q = 9$ and $q = 5$, respectively. In order to compare the quality of deblocking, three metrics are used in our paper, i.e., PSNR, SSIM (Structural SIMilarity)[14], GBIM (Generalized Block-edge Impairment Metric)[18]. Note that PSNR is a commonly used measure, SSIM index measures the similarity between two images and GBIM is an effective evaluation metric for blocky artifact. A smaller value of GBIM means a better degree of deblocking. In our experiments, we compare our proposed scheme to the state-of-the-art algorithms, including FoE (fields of experts) [13], BM3D (Block-matching and 3D

Table 1: Performance Comparison for test images at $512 \times 512, QF=11, Q1$

Image	FoE [13]			Sparse method[8]			PROPOSED		
	PSNR	SSIM	GBIM	PSNR	SSIM	GBIM	PSNR	SSIM	GBIM
Barbara	26.73	0.81	1.6	26.82	0.8	1.3	26.91	0.82	1.31
Lena	32.06	0.86	1.76	32.06	0.86	1.73	32.1	0.87	1.68
Boat	29.48	0.8	1.62	29.4	0.79	1.53	29.5	0.8	1.5
Peppers	31.9	0.84	1.94	31.7	0.83	1.92	31.82	0.83	1.53
Baboon	24.12	0.7	1.72	24.08	0.67	1.49	24.11	0.71	1.25
Fruits	31.48	0.84	2.0	31.43	0.84	1.78	31.46	0.86	1.32

Table 2: Performance Comparison for test images at $512 \times 512, QF=9, Q2$

Image	FoE [13]			Sparse method[8]			PROPOSED		
	PSNR	SSIM	GBIM	PSNR	SSIM	GBIM	PSNR	SSIM	GBIM
Barbara	26.18	0.79	1.61	26.29	0.78	1.29	26.32	0.8	1.13
Lena	31.32	0.85	1.92	31.34	0.85	1.81	31.38	0.86	1.52
Boat	28.79	0.78	1.73	28.74	0.77	1.59	28.81	0.81	1.35
Peppers	31.28	0.83	1.99	31.11	0.82	1.94	31.25	0.83	1.72
Baboon	23.6	0.66	1.78	23.56	0.63	1.51	23.74	0.71	1.47
Fruits	30.8	0.83	2.16	30.81	0.83	1.86	30.8	0.83	1.75

filtering) [4] and sparse representation[8]. For simplicity, we adopt a Gaussian filter with $\sigma = 11$ to decompose the blocky image and empirically select the size of patch as 7×7 with square neighborhood. The visual quality comparison of the proposed scheme is provided in Fig. 3, in which only sub-images of *House* are provided for visual comparison at $q = 5$. The results demonstrate that the proposed scheme achieves a very competitive deblocking performance with the fine image structure preservation, compared with the state-of-art image deblocking algorithms. In particular, the results of our method in terms of GBIM represent good blocky artifact removal and though there exists moderate results in terms of PSNR for some test images. According to the visual comparison, it can be seen that the proposed scheme is effective in this aspect. Please refer to Tables 1 to 3 for more details.

5. CONCLUSION

In this paper, a novel image deblocking algorithm based on low-rank matrix recovery is proposed, in which the high-frequency component of the blocky image is recovered by solving an optimization problem of the nuclear norm and ℓ_1 -norm constraint. According to the experimental results, the proposed scheme can achieve competitive deblocking performance with good image structure preservation.

Table 3: Performance comparison for test images at $256 \times 256, QF=5, Q3$

Image	FoE [13]			BM3D [4]			PROPOSED		
	PSNR	SSIM	GBIM	PSNR	SSIM	GBIM	PSNR	SSIM	GBIM
Lena	25.0	0.73	2.86	25.08	0.72	3.03	25.083	0.72	2.624
Monarch	24.11	0.82	3.12	24.27	0.81	2.74	24.23	0.81	2.4
Cameraman	23.3	0.82	3.31	23.34	0.81	3.25	22.89	0.79	3.08
House	27.56	0.8	3.78	27.35	0.79	3.94	27.66	0.8	1.75

6. REFERENCES

- [1] M. Aharon, M. Elad, and A. M. Bruckstein. The K-SVD: an algorithm for designing of over complete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 54:4311–4322, 2006.
- [2] Emmanuel Candes, Xiaodong Li, Yi Ma, and John Wright. Robust principal component analysis? *Journal of the ACM*, 58:11, 2011.
- [3] Yi-Wen Chiou, Chia-Hung Yeh, Li-Wei Kang, Chia-Wen Lin, and Shu-Jhen Fan-Jiang. Efficient image/video deblocking via sparse representation. In *V-CIP*, pages 1–6, 2012.
- [4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian. Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16:2080–2095, 2007.
- [5] Charles-Alban Deledalle, Joseph Salmon, and Arnak Dalalyan. Image denoising with patch based PCA: local versus global. In *Proc. British Machine Vision Conference 2011*, pages 1–10, 2011.
- [6] M. Elad and M. Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, 2006.
- [7] Hui Ji, Sibin Huang, Zuowei Shen, and Yuhong Xu. Robust video restoration by joint sparse and low rank matrix approximation. *SIAM Journal on Imaging Sciences*, 4:1122–1142, 2011.
- [8] Cheolkon Jung, Licheng Jiao, Hongtao Qi, and Tian Sun. Image deblocking via sparse representation. *Signal Processing: Image Communication*, 27:663–677, 2012.
- [9] Gopal Lakhani. Improved equations for JPEG’s blocking artifacts reduction approach. *IEEE Transactions on Circuits and System for Video Technology*, 7(6):930–934, 1997.
- [10] A.W.C. Liew and H. Yan. Blocking artifacts suppression in block-coded images using overcomplete wavelet representation. *IEEE Transactions on Circuits and System for Video Technology*, 14:450–461, 2004.
- [11] Zhouchen Lin, Minming Chen, and Yi Ma. The augmented Lagrange multiplier method for exact recovery of corrupted low-rank matrices. Technical report, UIUC Technical Report UILU-ENG-09-2215, 2009.
- [12] Jie Ren, Jiaying Liu, Mading Li, Wei Bai, and Zongming Guo. Image blocking artifacts reduction via patch clustering and low-rank minimization. In *2013 Data Compression Conference*, pages 516–516, 2013.
- [13] D. Sun and W.K. Cham. Postprocessing of low bit-rate block DCT coded images based on a fields of experts prior. *IEEE Transactions on Image Processing*, 16:2743–2751, 2007.
- [14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli. Image quality assessment: From error measurement to structural similarity. *IEEE Transactions on Image Processing*, 13(4):600–612, 2004.
- [15] Siu-Wai Wu and Allen Gersho. Improved decoder for transform coding with application to the jpeg baseline system. *IEEE Transactions on Communications*, 40(2):251–254, 1992.
- [16] Avideh Zakhor. Iterative procedures for reduction of blocking effects in transform image coding. *IEEE Transactions on Circuits and System for Video Technology*, 2(1):91–95, 1992.
- [17] T OG zcelik, J C Brailean, and A K Katsaggelos. Image and video compression algorithms based on recovery techniques using mean field annealing. In *Proc. IEEE*, volume 83, pages 304–315, 1995.
- [18] H.R. Wu and M. Yuen. A generalized block-edge impairment metric for video coding. *IEEE Signal Processing Letters*, 4(11):317–320, 1997.