Correspondence

Comments on "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering"

Yingkun Hou, Chunxia Zhao, Deyun Yang, and Yong Cheng

***Abstract*—In order to resolve the problem that the denoising performance has a sharp drop when noise standard deviation reaches 40, [1] proposed to replace the wavelet transform by the DCT. In this comment, we argue that this replacement is unnecessary, and that the problem can be solved by adjusting some numerical parameters. We also present this parameter modification approach here. Experimental results demonstrate that the proposed modification achieves better results in terms of both peak signal-to-noise ratio and subjective visual quality than the original method for strong noise.**

***Index Terms*—BM3D, high levels noise, image denoising.**

I. Introduction

Recently, a powerful image denoising strategy, block-matching and 3-D filtering (BM3D), was proposed in [1]. It is considered to be current state-of-the-art in image denoising [2]. To further improve the denoising performance of BM3D, the authors of [1] also developed other algorithms [3], [4]. [3] improves denoising performance when the noise level is low; however, for relatively high noise levels, the denoising performance in [3] is even worse than [1]. [4] only gave the denoising results for noise standard deviations less than 35. Two main techniques of BM3D are grouping and collaborative filtering. Given a reference image block, similar blocks are found by block-matching. These blocks are stacked to form a 3-D array. Collaborative filtering is then performed on each 3-D array with a 3-D transform. The 3-D transform is actually separable composition of a 2-D transform for every block and 1-D transform for each column of the 3-D array. Transform-domain shrinkage attenuates the noise and the inverse of the 3-D transform produces estimates of all grouped blocks. Finally, return all estimated blocks back to their original position by weighted averaging.

For a noisy image z, given a reference block Z_{x_R} within a neighborhood X in z, grouping finds a set of blocks Z_x in X which are similar to Z_{x_R} by the ℓ^2-distance as

$$d^{\mathrm{noisy}}(Z_{x_R}, Z_x) = \frac{\|Z_{x_R} - Z_x\|_2^2}{(N_1^{\mathrm{ht}})^2} \tag{1}$$

where $\|\cdot\|_2$ denotes the ℓ^2-norm, N_1^{ht} is the block size and x is the coordinate of the top-left corner of the block. Using the d-distance (1), one can find a set $S_{x_R}^{\mathrm{ht}}$ containing image blocks Z_x, all of which are similar to a reference block Z_{x_R}

$$S_{x_R}^{\mathrm{ht}} = \left\{ x \in X : d^{\mathrm{noisy}}(Z_{x_R}, Z_x) \leq \tau_{\mathrm{match}}^{\mathrm{ht}} \right\} \tag{2}$$

where $\tau_{\mathrm{match}}^{\mathrm{ht}}$ is the maximum d-distance for which a block is similar to the reference block. The parameter $\tau_{\mathrm{match}}^{\mathrm{ht}}$ is selected only by an empirical value. Because $d(Z_{x_R}, Z_{x_R}) = 0$, the cardinality $|S_{x_R}^{\mathrm{ht}}|$ of $S_{x_R}^{\mathrm{ht}}$ is always greater than 1. A group is formed by stacking the matched noisy blocks $Z_{x \in S_{x_R}^{\mathrm{ht}}}$ to form a 3-D array of size $N_1^{\mathrm{ht}} \times N_1^{\mathrm{ht}} \times |S_{x_R}^{\mathrm{ht}}|$, which is denoted as $Z_{S_{x_R}^{\mathrm{ht}}}$. The collaborative filtering of $Z_{S_{x_R}^{\mathrm{ht}}}$ is realized by hard-thresholding in the 3-D transform domain.

Generally, denoising performance should gradually weaken with growing noise levels. However, in the case of [1], denoising performance has a sharp drop when noise standard deviation reaches 40. The authors ascribed this to erroneous grouping, and proposed measuring the block-distance using coarse prefiltering. This prefiltering is realized by applying a normalized 2-D linear transform on both blocks and then hard-thresholding the obtained coefficients, which results in (4) in [1]

$$d(Z_{x_R}, Z_x) = \frac{\left\| \Upsilon'\left(\mathcal{T}_{2-\mathrm{D}}^{\mathrm{ht}}(Z_{x_R})\right) - \Upsilon'\left(\mathcal{T}_{2-\mathrm{D}}^{\mathrm{ht}}(Z_x)\right) \right\|_2^2}{\left(N_1^{\mathrm{ht}}\right)^2} \tag{3}$$

where Υ' is the hard-thresholding operator with threshold $\lambda_{2-D}\sigma$ and $\mathcal{T}_{2-\mathrm{D}}^{\mathrm{ht}}$ denotes the normalized 2-D linear transform. [1] also proposed changing $\mathcal{T}_{2-\mathrm{D}}^{\mathrm{ht}}$ from the biorthogonal wavelet transform (Bior1.5) to the DCT and increase N_1^{ht} from 8 to 12. Although these adjustments solved this problem in a certain extent, there are some potential drawbacks. Although prefiltering can attenuate the noise, some true signals are removed by filtering as well. Coarse prefiltering cannot achieve correct grouping, and it even often worsens the grouping results. The replacement of the 2-D-Bior1.5 by the 2-D-DCT in Table I of [1] is especially unnecessary, as it only makes matters worse by destroying the mathematical continuity of the algorithm. One of the advantages of BM3D (as claimed by [1]) is that less artifacts are introduced. However, a number of periodic artifacts will be introduced when the DCT is used on 2-D blocks, as the transform is local in this case. Furthermore, increasing N_1^{ht} from 8 to 12 will increase time complexity.

In this comment, we argue that the change of 2-D transform can be avoided. We also propose a more effective approach (in terms of PSNR and visual quality) than what is used in [1]. Our approach involves increasing the maximum number of blocks N_2^{ht} in a group, increasing $\tau_{\mathrm{match}}^{\mathrm{ht}}$, changing the DCT transform and decreasing N_1^{ht}. All these modifications not only achieve slightly better results in terms of both peak signal-to-noise ratio and subjective visual quality than the original method but also make the algorithm more consistent. Due to the decrease in N_1^{ht}, the time complexity is also lower than the original method.

II. Proposed Modification

The maximum number of grouped blocks was restricted in [1] to be a power-of-two integer, denoted N_2^{ht}, and $N_2^{ht} = 16$ was used by default. In fact, N_2^{ht} should take a greater value when the noise level is relatively higher, e.g., 32. $\tau_{\mathrm{match}}^{\mathrm{ht}}$ should accordingly take a greater value to ensure that there are enough blocks in 3-D arrays to improve the denoising performance.

Manuscript received January 19, 2010; revised March 26, 2010 and May 17, 2010; accepted May 21, 2010. Date of publication June 14, 2010; date of current version December 17, 2010. This work was supported in part by the National Science Foundation of China under Grants 61072148, 90820306, and 60875010. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. James E. Fowler.

Y. Hou, C. Zhao, and Y. Cheng are with the School of Computer Science and Technology, Nanjing University of Science and Technology, Nanjing 210094, China (e-mail: njusthyk@163.com).

Y. Hou and D. Yang are with the School of Information Science and Technology, Taishan University, Taian 271021, China.

Y. Cheng is with the School of Communication Engineering, Nanjing Institute of Technology, Nanjing 211167, China.

Digital Object Identifier 10.1109/TIP.2010.2052281

(a) (b)

Fig. 1. Grayscale house denoising results comparison, $\sigma = 50$. (a) Denoised result of BM3D. (b) Result of the BM3D with the proposed modification. (a) PSNR 29.37 dB. (b) PSNR 29.69 dB.

(a) (b)

Fig. 2. Color house denoising results comparison, $\sigma = 75$. (a) Denoised result of BM3D. (b) Result of the BM3D with the proposed modification. (a) PSNR 29.37 dB. (b) PSNR 29.69 dB.

Fig. 3. Denoised grayscalce fragments, $\sigma = 50$. Left column: the denoised results of BM3D. Right column: the results of the BM3D with the proposed modification.

In order to make the algorithm more consistent and avoid introducing more periodic artifacts, the Bior1.5 transform on 2-D blocks should be still used just the same as the situation with the low noise level. Due to the orthogonality of the Bior1.5 transform and to the matrix transform employed in [1] on each image block, the size of image blocks should be 2^n, so we change N_1^{ht} from 12 to 8. The number of blocks in a neighborhood does not change, but the size of blocks decreases, so this change will obtain much lower time complexity.

Because the coarse prefiltering can not achieve the optimal grouping and often even worsens the grouping results, we change $\lambda_{2-\mathrm{D}}$ from 2.0 to 0.

To sum up, the parameters for $\sigma > 40$ in Table I in [1] are amended as follows: $\mathcal{T}_{2-\mathrm{D}}^{\mathrm{ht}} : 2\mathrm{D}-\mathrm{DCT} \rightarrow 2\mathrm{D}-\mathrm{Bior}1.5; N_1^{\mathrm{ht}} : 12 \rightarrow 8; N_2^{\mathrm{ht}} : 16 \rightarrow 32; \lambda_{2-\mathrm{D}} : 2.0 \rightarrow 0; \tau_{\mathrm{match}}^{\mathrm{ht}} : 5000 \rightarrow 25000$.

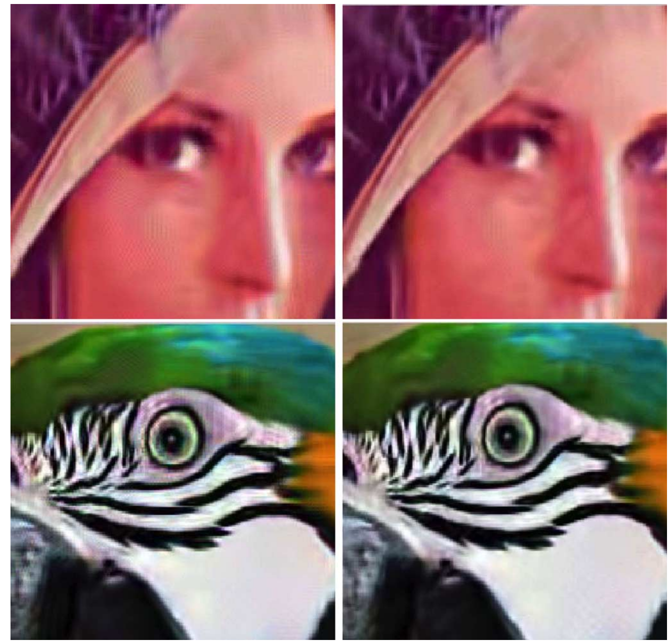

Fig. 4. Denoised color fragments, $\sigma = 75$. Left column: the denoised results of BM3D. Right column: the results of the BM3D with the proposed modification.

TABLE I
GRAYSCALE-IMAGE DENOISING: OUTPUT PSNR COMPARISON BETWEEN [1] AND THE PROPOSED METHOD

σ/PSNR	*C.man*	*House*	*Peppers*	*Montage*	*Lena*	*Barbara*
50 / 14.16	25.84	29.37	26.41	27.35	28.86	27.17
	26.12	**29.69**	**26.68**	**27.90**	**29.05**	**27.23**
75 / 10.63	24.05	27.20	24.48	25.04	27.02	25.10
	24.33	**27.51**	**24.73**	**25.52**	**27.26**	**25.12**
100 / 8.14	22.81	25.50	22.91	23.38	25.57	23.49
	23.07	**25.87**	**23.39**	**23.89**	**25.95**	**23.62**

III. EXPERIMENTAL RESULTS

We conducted a number of comparison experiments between BM3D and the improved method on the images provided by the authors on the BM3D website. All experiments were conducted on a Core2 2.20 GHz Duo. Figs. 1–4 show that less artifacts are introduced and more details can be reserved by the proposed method than by the original method. Tables I and II give an objective PSNR evaluation; The PSNR values are moderately increased by the proposed method. Because N_1^{ht} is decreased, so the time complexity of the proposed modification is lower than in [1]; Table III gives the running time comparison. From Table III, we can see that the average running time is significantly decreased by the proposed modification.

TABLE II
COLOR-IMAGE DENOISING: OUTPUT PSNR COMPARISON BETWEEN [1] AND THE PROPOSED METHOD

σ/PSNR	*Lena*	*Peppers*	*Baboon*	*F16*	*House*
50 / 14.15	29.72	28.68	23.14	29.41	30.22
	29.88	**28.93**	**23.15**	**29.79**	**30.47**
75 / 10.63	28.19	27.12	21.71	27.60	28.33
	28.36	**27.38**	**21.76**	**27.97**	**28.69**

TABLE III
RUNNING TIME COMPARISON BETWEEN [1] AND THE PROPOSED METHOD

Time(second)	256×256	512×512
Grayscale image	2.8	12.1
	2.0	**8.3**
Color image	10.1	36.4
	4.2	**18.3**

IV. CONCLUSION

The proposed modification in this comment achieves better denoising results than [1], especially the changing of the DCT to make the algorithm more consistent and introduce less artifacts than original method. Due to the decrease in block size, the time complexity is also lower in the proposed method.

ACKNOWLEDGMENT

The authors would like to thank X. Qu (Xiamen Univesity) and A. Christodoulou (University of Illinois at Urbana-Champaign) for their helpful discussions and the anonymous reviewers for their insightful reviews.

REFERENCES

[1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[2] S. Lansel, D. Donoho, and T. Weissman, "DenoiseLab: A standard test set and evaluation method to compare denoising algorithms," Stanford University, Stanford, CA [Online]. Available: http://www.stanford.edu/slansel/DenoiseLab/

[3] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "A nonlocal and shape-adaptive transform-domain collaborative filtering," in *Proc. Int. Workshop Local and Non-Local Approx. Image Process.*, Lausanne, Switzerland, Aug. 2008, pp. 179–186.

[4] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "BM3D image denoising with shape-adaptive principal component analysis," in *Proc. Workshop Signal Process. Adaptive Sparse Structured Represent.*, Saint-Malo, France, Apr. 2009.

An Improved Image Compression Algorithm Using Binary Space Partition Scheme and Geometric Wavelets

Garima Chopra and A. K. Pal

***Abstract*—Geometric wavelet is a recent development in the field of multivariate nonlinear piecewise polynomials approximation. The present study improves the geometric wavelet (GW) image coding method by using the slope intercept representation of the straight line in the binary space partition scheme. The performance of the proposed algorithm is compared with the wavelet transform-based compression methods such as the embedded zerotree wavelet (EZW), the set partitioning in hierarchical trees (SPIHT) and the embedded block coding with optimized truncation (EBCOT), and other recently developed "sparse geometric representation" based compression algorithms. The proposed image compression algorithm outperforms the EZW, the Bandelets and the GW algorithm. The presented algorithm reports a gain of 0.22 dB over the GW method at the compression ratio of 64 for the Cameraman test image.**

***Index Terms*—Binary space partition scheme, image compression, geometric wavelets, piecewise polynomial approximation and sparse geometric representations.**

I. INTRODUCTION

Digital images are ubiquitous in many application areas as diverse as internet browsing, medical sciences, astronomy and remote sensing. Once personal computers gained the capacity to display sophisticated pictures as digital images, people started to seek methods for efficient representation of these digital pictures in order to simplify their transmission and save disk space. At this point image compression became very important and highly applicable and since then it has been the researchers favorite. The field of image compression has a wide spectrum ranging from classical lossless techniques and popular transform approaches to the more recent segmentation based (or second generation) coding methods. Further, compression techniques can be classified into lossless and lossy techniques [1]. The lossless techniques allow to compress an image without losing any information while the images reproduced by the lossy techniques are not very perfect.

In the past decades, the discrete cosine transform (DCT) has been the most popular for compression because it provides optimal performance and can be implemented at a reasonable cost. Several compression algorithms, such as the JPEG standard [2] for still images and the MPEG standard [3] for video images are based on DCT [4]. However, the EZW [5], the SPIHT [6], the SPECK [7], the EBCOT [8] algorithms and the current JPEG 2000 [9] standard are based on the discrete wavelet transform (DWT) [10]–[12]. DWT has the ability to solve the blocking effect introduced by DCT, it also reduces the correlation between the neighboring pixels and gives multi scale sparse representation of the image. In spite of providing excellent results in terms of rate-distortion compression, the transform-based coding methods do not take an advantage of the underlying geometry of the edge singularities in an image. From

Manuscript received December 11, 2009; revised April 14, 2010; accepted June 15, 2010. Date of publication July 08, 2010; date of current version December 17, 2010. This work was supported in part by G. B. Pant University of Agriculture and Technology, Pantnagar, India, and a Junior Research Fellowship provided by the Council of Scientific and Industrial Research, Human Resource Development Group, New Delhi, India. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Eero P. Simoncelli.

The authors are with the College of Basic Sciences and Humanities, G. B. Pant University of Agriculture and Technology, 263145 Pantnagar, India (e-mail: garima.chopra@gmail.com; arun_pal1969@yahoo.co.in).

Digital Object Identifier 10.1109/TIP.2010.2056378