Compression Artifact Reduction with Adaptive Bilateral Filtering

Ming Zhang and Bahadir Gunturk
Department of Electrical and Computer Engineering
Louisiana State University
Baton Rouge, LA 70803

ABSTRACT

In this paper, we present a spatially adaptive method to reduce compression artifacts observed in block discrete cosine transform (DCT) based image/video compression standards. The method is based on the bilateral filter, which is very effective in denoising images without smoothing edges. When applied to reduce compression artifacts, the parameters of the bilateral filter should be chosen carefully to have a good performance. To avoid over-smoothing texture regions and to effectively eliminate blocking and ringing artifacts, in this paper, texture regions and block boundary discontinuities are first detected; these are then used to control/adapt the spatial and intensity parameters of the bilateral filter. Experiments show that the proposed method improves over the standard non-adaptive bilateral filter visually.

Keywords: Compression artifacts removal, Bilateral filter

1. INTRODUCTION

Block-based discrete cosine transform (DCT) adopted by widely used image/video compression standards, such as JPEG, MPEG-x, and H.26x, is considered as an asymptotic approximation of the optimal Karhunen-Loeve transform (KLT) due to its high energy compaction and low computational complexity. One problem associated with the block based processing is the blocking artifacts, the discontinuities along the block boundaries caused by the coarse quantization of the coefficients in DCT. The blocking artifacts and other compression artifacts, such as the ringing problem, often truncate the high-frequency DCT coefficients and become more severe with increasing compression rates.

There are numerous methods proposed to reduce compression artifacts, such as blocking artifacts and ringing artifacts. The post-processing methods, which do not require any codec changes can be categorized into two: enhancement based algorithms and restoration based algorithms. Another way of categorizing the deblocking methods is spatial domain vs. transform domain, depending on the what domain the image is processed. There are methods that use both domains. An example of the enhancement based algorithms is made by Apostolopoulos et al.,[1] where the blockiness is first detected based on the number of zero DCT coefficients in each block, and then applying 1D median filter to reduce block discontinuities and 2D median filter to reduce mosquito artifacts. Examples of restoration based algorithms are the Bayesian approach by Mateos et al.,[2] the projection onto convex sets (POCS) method of Weerasinghe et al.,[3] the minimum mean squared error method of Triantafyllidis et al.,[4] the post-filter using the DCT coefficients of shifted blocks to deblock and preserve the details by Qiu et al.,[5] the fast and blind measurement of detection and reduction to the blocks in the DCT domain by Bovik et al.,[6] and the DCT-domain Markov Random Field model by Delp et al.[7]

In this paper, we present an enhancement method that is based on the bilateral filter.[8] The bilateral filter does a weighted spatial averaging, where the weights depend on both spatial distances and intensity distances. In this way, edge preserving smoothing is achieved. Bilateral filtering was recently utilized to reduce compression artifacts; for example, Pham et al.[9] presents a separable implementation of the bilateral filter to reduce the edge jaggedness in videos. Yu et al.[10] decomposes an image into its low and high frequency components, and then applies low-pass filtering to the low-frequency parts and bilateral filtering to the high-frequency parts (to reduce ringing artifacts).

The method we present here is a spatially adaptive version of the bilateral filter. The parameters of the bilateral filter should be chosen to reduce block discontinues and ringing artifacts effectively while avoiding to

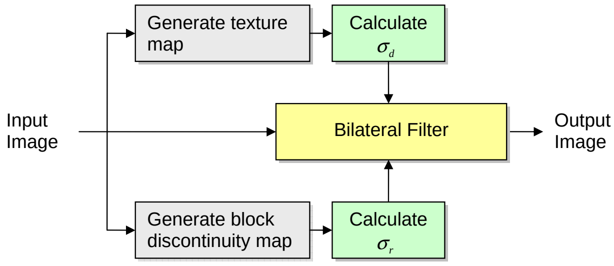

Figure 1. The block diagram of the proposed method. Discontinuity and texture detection modules produce space varying maps that are used to compute the range (intensity) and domain parameters of the bilateral filter. The bilateral filter is applied on to the image b

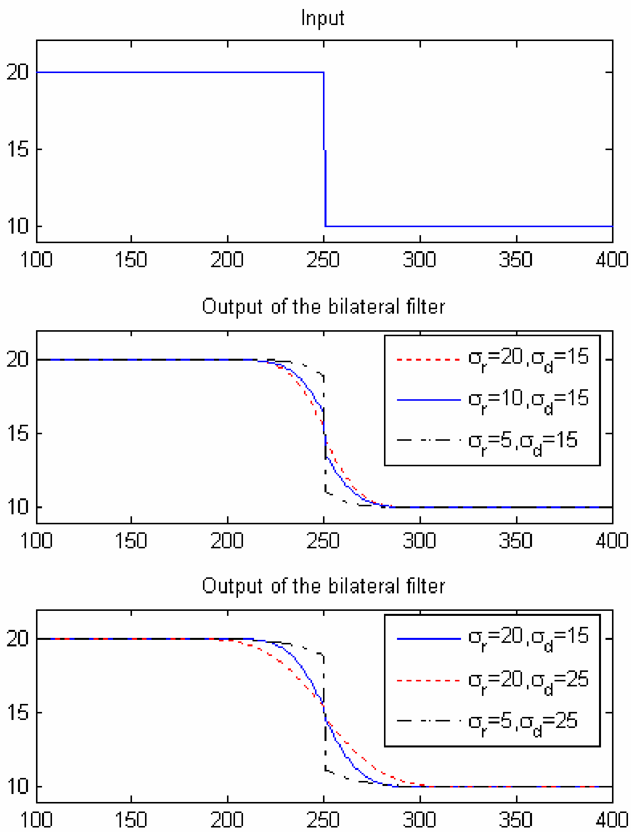

Figure 2. Effects of the values of the bilateral filter parameters σ_r and σ_d on a block discontinuity are illustrated.

over-smooth texture regions. Our method adapts the parameters to achieve that. In Section 2, we overview the bilateral filter and its parameters. Based on some simulations, we evaluate the influence of its parameters for compression artifact reduction applications. In Section 3, the details of the proposed method are given. In Section 4, we compare the proposed method with original (non-adaptive) bilateral filter.[8] The experiment results show that the adaptive bilateral filter works better perceptually.

2. PARAMETERS OF THE BILATERAL FILTER

The bilateral filter, presented by Tomasi et al.,[8] is achieved by combining two Gaussian filters; one filter works in spatial domain, the other filter works in intensity domain. Therefore, not only the spatial distance but also

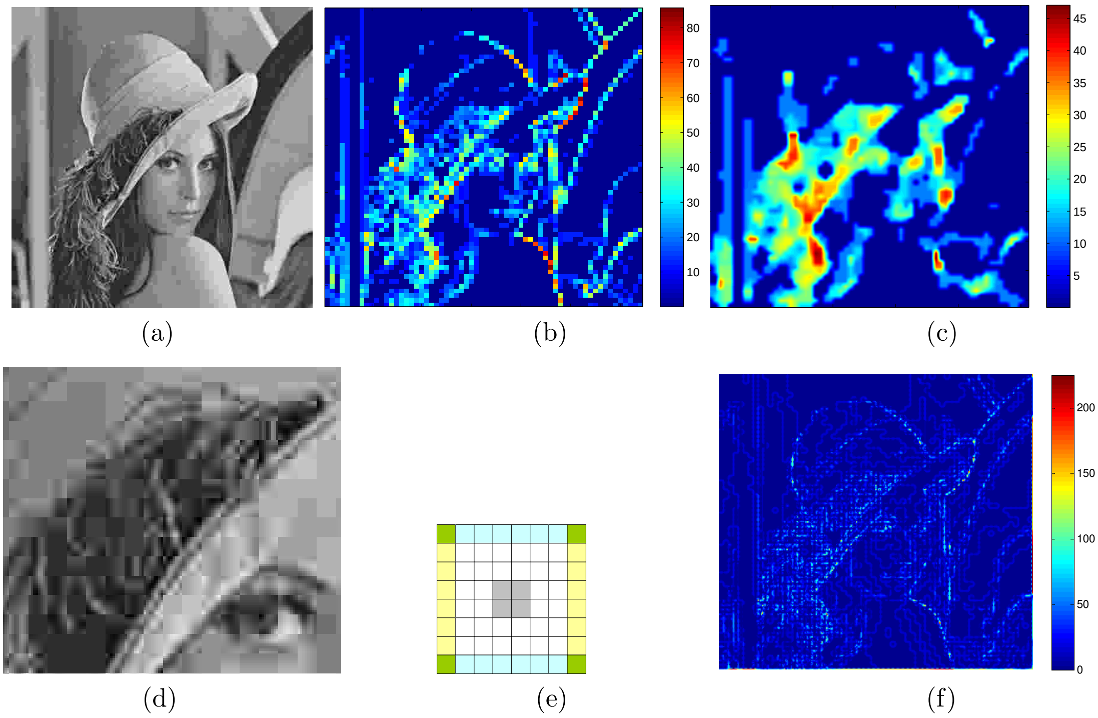

Figure 3. Texture and discontinuity detection steps for the compressed Lena image of size 512 × 512. The compression method is JPEG, with quality factor 6 in Matlab. (a) Input image. (b) Local standard deviation for texture detection. (c) Median filtered local standard deviation. (d) A zoomed in region of compressed Lena (compare with Figure 8.) (e) Interpolation of the block discontinuities at each block. (f) Block discontinuity map.

the intensity distance is important for the determination of the weights. At a pixel location $\mathbf{x} = (x_1, x_2)$, the output of a bilateral filter can be formulated as follows:

$$\tilde{I}(\mathbf{x}) = \frac{1}{C} \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} e^{\frac{-\|\mathbf{y}-\mathbf{x}\|^2}{2\sigma_d^2}} e^{\frac{-|I(\mathbf{y})-I(\mathbf{x})|^2}{2\sigma_r^2}} I(\mathbf{y}), \tag{1}$$

where σ_d and σ_r are parameters controlling the fall-off of weights in spatial and intensity domains, $\mathcal{N}(\mathbf{x})$ is a spatial neighborhood of pixel $I(\mathbf{x})$, and C is the normalization constant:

$$C = \sum_{\mathbf{y} \in \mathcal{N}(\mathbf{x})} e^{\frac{-\|\mathbf{y}-\mathbf{x}\|^2}{2\sigma_d^2}} e^{\frac{-|I(\mathbf{y})-I(\mathbf{x})|^2}{2\sigma_r^2}}. \tag{2}$$

There are two parameters that control the behavior of the bilateral filter. Referring to (1), σ_d and σ_r characterizes the spatial and intensity domain behaviors, respectively. In case of compression artifact reduction, these parameters should be chosen carefully. Figure 2 illustrates this on a one dimensional signal. The first subplot in that figure shows an edge signal; the edge discontinuity is 10. The second subplot displays the outputs of the bilateral filter for different values of σ_r. When the σ_r value is less than the discontinuity amount, the filter is basically useless against eliminating the discontinuity. When σ_r is larger than the discontinuity amount, the discontinuity can be eliminated. In the third subplot, it is demonstrated that the extent of the smoothing can be controlled by the σ_d value. The larger the σ_d value, the wider the extent of smoothing. On the other hand, if σ_r value is less than the discontinuity amount, elimination of the discontinuity is impossible no matter the value of σ_d.

These observations tell us to measure the discontinuity amount along the block boundaries and adapt the value of σ_r accordingly. On the other hand, we would like to avoid over-smoothing texture regions. This could be done by first estimating the texture regions (through, for example, estimating the local variances), and then control the extent of smoothing by adapting the σ_d value. For a smooth region, the value of the σ_d can be large; otherwise, it should be small. Our specific method is explained in the next section.

3. PROPOSED METHOD

As we will show shortly, the non-adaptive application of bilateral filter creates some problems: if strong parameters are chosen to eliminate blockiness, it over-blurs the texture details; if weaker parameters are chosen, the blocking artifacts are not completely removed. To address these issues, we present an adaptive bilateral filtering framework, whose block diagram is given in Figure 1. In the light of discussion of the previous section, we included two modules in the framework. One module detects the block discontinuities and adjusts the value of σ_r accordingly; the other module detects smoothness of local regions and adjusts the value of σ_d accordingly.

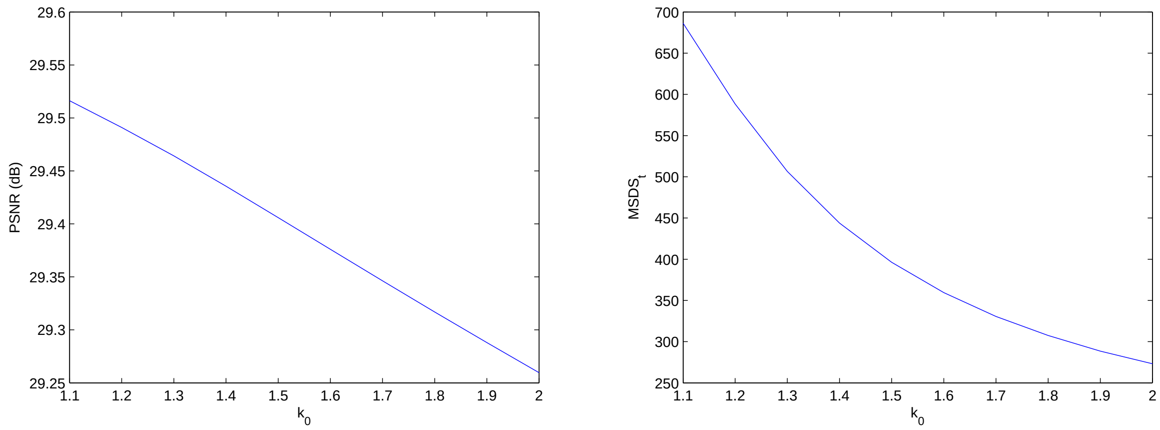

Figure 4. Relationship between the k_0 and the PSNR, $MSDS_t$ of the proposed method for the Lena test image under bitrate=0.18. The image size is 512×512. $k_1 = 3$.

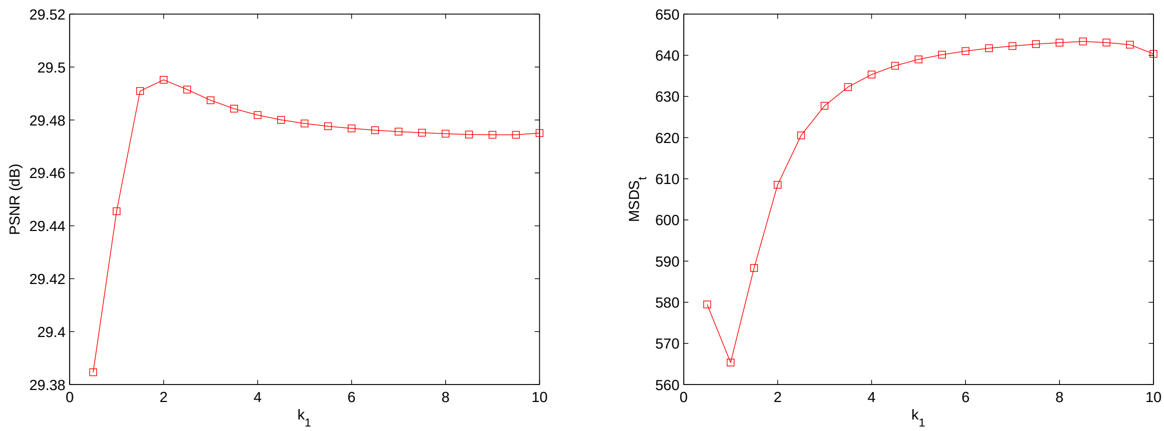

Figure 5. Relationship between the k_1 and the PSNR, $MSDS_t$ of the proposed method for the Lena test image under bitrate=0.18. The image size is 512×512. $k_0 = 1.2$.

To detect block discontinuities, the input image is filtered with $[-1, 0, 1]$ (for vertical boundaries) and with $[-1, 0, 1]^T$ (for horizontal boundaries) along the block boundaries, and then absolute values of the results are taken. The σ_r value should be at least equal to these values to be effective. The discontinuities are detected

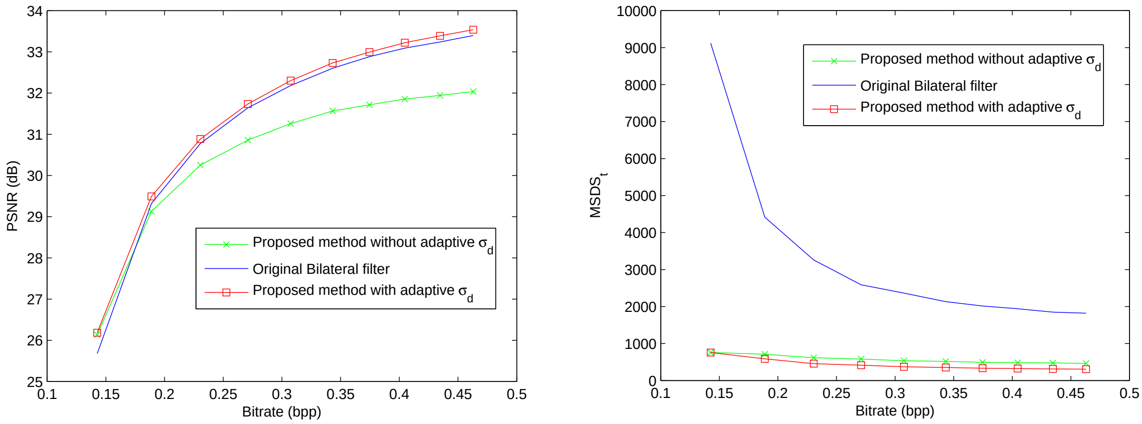

Figure 6. Comparison of original Bilateral with the proposed method with and without the adaptive σ_d for the Lena test image under different bit-rate. The image size is 512×512. $\sigma_r = 20$.

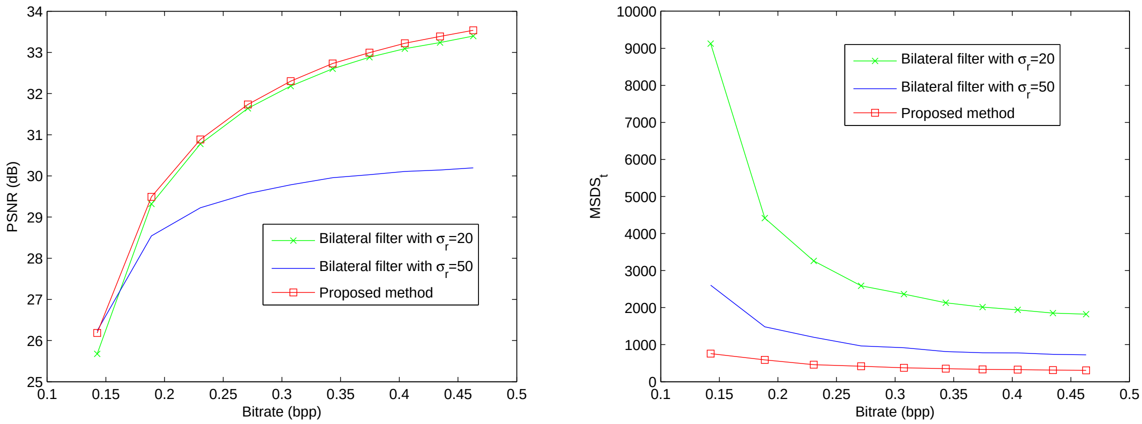

Figure 7. Comparison of original Bilateral filter with σ_r of 20 and 50, and the proposed method for the Lena test image under different bit-rate. The image size is 512×512. $\sigma_d = 3$.

along the block boundaries; however, if the bilateral filter is applied along the boundaries only, the blockiness cannot be eliminated. Consider a single block; if the bilateral filter is applied along the boundaries only, the discontinuity moves further inside the block. To eliminate the blockiness effectively, the bilateral filter should be applied to the entire block. Thus, the discontinuities along the boundaries should be diffused into the blocks. One approach is as follows: Referring Figure 3(e), the center four pixels inside a block is set to zero; the horizontal and vertical discontinuities along the boundaries are kept except for the corner pixels, where the larger of the horizontal/vertical discontinuities is chosen; and then the rest of the pixels are interpolated linearly. This is repeated for all blocks to obtain the block discontinuity map $M_b(\mathbf{x})$. The block discontinuity map $M_b(\mathbf{x})$ for the JPEG compressed Lena image is shown Figure 3(f). Once $M_b(\mathbf{x})$ is calculated, the adaptive $\sigma_r(\mathbf{x})$ is calculated as

$$\sigma_r(\mathbf{x}) = max(\sigma_{r,min}, k_0 M_b(\mathbf{x})), \tag{3}$$

where $\sigma_{r,min}$ is the minimum value of $\sigma_r(\mathbf{x})$, and k_0 is a scale factor. The reason we use such a minimum value is that we would like to apply a minimal filtering to the entire image; if this was not done, other compression artifacts, such as the mosquito artifact, could not be eliminated and some sort of spatial unevenness appear in the final image. (Note that a slightly different implementation is to set the center values of each block to $\sigma_{r,min}$ instead of zero during the block discontinuity interpolation process.)

Input Image	Bit-rate(bpp)	JPEG	H.263	MPEG4	POCS[3]	PDCT[5]	FDCT[6]	BLT[8]	Proposed
Lena 512 × 512	0.22	29.47	30.20	30.02	30.23	30.32	30.27	30.37	30.49
Peppers 512 × 512	0.22	29.21	30.02	30.04	29.85	29.95	29.93	30.59	30.59
Goldhill 512 × 512	0.23	27.90	28.50	28.31	28.46	28.51	28.40	28.38	28.53
Mandrill 512 × 512	0.30	22.05	22.35	22.15	22.44	22.49	22.39	22.46	22.26
Airplane 512 × 512	0.24	28.72	29.34	29.32	29.34	29.39	29.33	29.87	29.81

Table 1. Comparison of the proposed method in PSNR with JPEG Coded, H.263, MPEG4, POCS,[3] Post DCT,[5] Fast DCT[6] and original bilateral filter. The parameters of bilateral filter is that σ_r=20, σ_d=3, window size is 6, $k_0 = 1.0$, $k_1 = 10.0$, and $k_2 = 1.0$.

Input Image	Bit-rate(bpp)	JPEG Coded	TSD-MRF[7]	Bilateral	Proposed Method
Lena 512 × 512	0.20	6674	2229	3841	526
	0.30	4384	1641	2365	374
Boat 512 × 512	0.20	10947	3844	5744	879
	0.30	8695	3969	4095	747
Cameraman 256 × 256	0.20	5164	2904	4512	438
	0.30	4276	2554	2616	351
Peppers 512 × 512	0.20	6341	2212	5622	2157
	0.30	3524	1322	4525	2193

Table 2. Comparison of the proposed method in $MSDS_t$ with JPEG Coded, TSD-MRF[7] and original bilateral filter. The parameters of bilateral filter is that σ_r=20, σ_d=3, window size is 6, $k_0 = 1.0$, $k_1 = 10.0$, and $k_2 = 1.0$.

To detect high-frequency texture regions, we compute the standard deviation of each block. The standard deviation is used as an indicator of texture and to adapt value of σ_d to preserve the texture information. Figure 3(b) shows the standard deviation of each block for the compressed Lena image. Note that the edge regions are highlighted in addition to the texture regions. However, we would like to apply strong bilateral to edge regions as well to eliminate ringing type of artifacts. One solution is to apply a median filter to eliminate edge regions from the texture map. In our experiments, the standard deviation of each 8×8 block is calculated, a 3×3 median filter is applied, and then the resulting image is interpolated to obtain the texture map $M_t(\mathbf{x})$. (For the Lena image of Figure 3(a), the texture map is shown in Figure 3(d).) The value of the σ_d should be inversely proportional to $M_t(\mathbf{x})$. One way of calculating $\sigma_d(\mathbf{x})$ is

$$\sigma_d(\mathbf{x}) = max(\sigma_{d,min}, \frac{k_1}{1 + M_t(\mathbf{x})}), \tag{4}$$

where k_1 is a constant parameter controlling the mapping from $M_t(\mathbf{x})$ to $\sigma_d(\mathbf{x})$, and $\sigma_{d,min}$ is minimum value of $\sigma_d(\mathbf{x})$. Such a minimum $\sigma_{d,min}$ is introduced again to do a minimum level filtering to the entire image.

4. EXPERIMENTS AND ANALYSIS

There have several measurements for the blocking artifacts. The most popular two are PSNR and MSDS(Mean Squared Difference of Slope).[11] MSDS involves the intensity gradient (slope) of the pixels close to the boundary of two blocks. It is based on the empirical observation that quantization of the DCT coefficients of two neighboring blocks increases the MSDS between the neighboring pixels on their boundaries. Consider an 8×8 block of the input image and four blocks w, s, e, n horizontally adjacent to f. The MSDS is defined by

$$\epsilon_w = \sum_{m=0}^{7} (d_1(m) - d_2(m))^2, \tag{5}$$

where $d_1(m)$ is the intensity slope across the boundary between the f and w blocks, defined by

$$d_1(m) = f(m, 0) - w(m, 7) \tag{6}$$

and $d_2(m)$ is the average between the intensity slope of f and w blocks close to their boundaries, defined by

$$d_2(m) = \frac{w(m,7) - w(m,6)}{2} + \frac{f(m,1) - f(m,0)}{2} \tag{7}$$

Then, the MSDS which involves both horizontal and vertical adjacent blocks is given by

$$MSDS_1 = \epsilon_w + \epsilon_s + \epsilon_e + \epsilon_n \tag{8}$$

$MSDS_t$ is proposed in,[12] which extends the definition of MSDS by involving the four diagonally adjacent blocks. If nw is a block diagonally adjacent to f , then define

$$\epsilon_{nw} = (g_1(m) - g_2(m))^2, \tag{9}$$

where

$$g_1(m) = f(0,0) - w(7,7) \tag{10}$$

and

$$g_2(m) = \frac{nw(7,7) - nw(6,6)}{2} + \frac{f(1,1) - f(0,0)}{2} \tag{11}$$

If ne, ns, nw and nn are the four blocks diagonally adjacent to f; the MSDS involving only the diagonally adjacent blocks is

$$MSDS_2 = \epsilon_{nw} + \epsilon_{ns} + \epsilon_{ne} + \epsilon_{nn} \tag{12}$$

The total $MSDS_t$ considered for the intensity slopes of all the adjacent blocks is

$$MSDS_t = MSDS_1 + MSDS_2 \tag{13}$$

We test this proposed adaptive bilateral filter for blocking artifacts reduction for some standard images, such as "Lena", "Cameraman", "Boat", "Airplane", "Mandrill", "Peppers" and "Goldhill". We compare our results in PSNR, with JPEG Coded, H.263, MPEG-4, POCS, Method in,[5] Method in[6] and original bilateral filter. We also compare our method in $MSDS_t$ with JPEG Coded, TSD-MRF[7] which is considered as the best in $MSDS_t$ among the previous methods and original bilateral filter. Table1 and 2 show the results with the parameters cited in the caption. From Table1 and table2 we notice that the bilateral filter and the proposed method has the best performance in PSNR, while the proposed method performs best in $MSDS_t$.

We also test the parameters in our proposed method. Figure 4 shows how k_0, which controls σ_r along the blocking boundaries according to equation 3, influences the PSNR and $MSDS_t$. Obviously, when the k_0 increases, the PSNR will be undermined while the $MSDS_t$ can be improved. Therefore, we have to choose a proper k_0 to make the balance between the PSNR and $MSDS_t$. Figure 6 shows that under different bit-rate, using the proposed adaptive σ_d for the texture detected, Both of the PSNR and $MSDS_t$ can be improved a lot. In the equation 3 we use the k_1 to control the intensity of the σ_d along the blocking boundary. Figure 5 illustrates the relation ship between k_1 and PSNR, $MSDS_t$, which implies only the moderate k_1 can make both metric best.

In order to differentiate the performance of bilateral filter and proposed method, we plot the comparison of bilateral filter using different intensity parameter σ_r with the proposed method. From Figure 7, we can find that when σ_r increases, the PSNR of the bilateral filter becomes worse, although the $MSDS_t$ decreases. It proves that intensity parameter of bilateral filter can control the blocking reduction as well as blur the details. Only the proposed method can reduce the blocking artifacts as well as saving the details with the best PSNR and $MSDS_t$. Visually, we can find the same conclusion from in Figure 8. where a portion of the result for the Lena image is shown. The parameters are included in the caption of that figure. Figure 9 displays the visual comparison of the Bilateral filter and our proposed method under different bit-rate. We can clearly find that the proposed adaptive method can eliminate the blocks effectively, compared to the residual blocking artifacts in the bilateral filter. Figure 10 presents the full test image after bilateral filter and proposed adaptive method. It is obvious that for the texture part such as the shoulder of 'Lena', the proposed method can preserve the details smoothly. We also test for the video sequences such as "Foreman", shown as Figure 11. The results indicates that our method works best for saving the details in the texture part as well as removing the artifacts.

Figure 8. A region from our experiment with the Lena image is displayed. Image size is 128 × 128; and the bit-rate for the compressed image is 0.18. From left to right: (a) The standard bilateral filter with $\sigma_r = 20$, $\sigma_d = 3$; (b) The standard bilateral filter with $\sigma_r = 50$, $\sigma_d = 3$; (c) The adaptive bilateral filter with $\sigma_{r,min} = 20$, $\sigma_{d,min} = 3$, $k_0 = 1.0$, $k_1 = 10.0$.

5. CONCLUSION

In this paper, we present an spatial adaptive method for the blocking artifact reduction, which is manifested as an automatic detection for the texture and the discontinuity in the image so that we can apply different spatial and intensity parameters of the bilateral filter upon them. The value of the parameters is determined by the local mapping from the index assigned by the detection. From the experiment, the proposed method has the best performance in PSNR and $MSDS_t$. At the same time, the visual quality of the results show that the proposed method can eliminate the blocking artifacts better and keep more texture details than the original bilateral filter.

In our preliminary experiments, the parameters were selected after some trial and error. As a next task, we will do a further analysis of these parameters. The preliminary results indicate that the adaptive method reduces the blockiness effectively while keeping the texture. Further improvement can be achieved by applying the smoothing process repeatedly. Another possible approach to improve the results is constrain the DCT coefficients of the resulting image. The upper and lower bounds for each DCT coefficient are available at the decoder side. By iterating the processes of projecting the resulting image onto these bounds in the DCT domain and applying adaptive bilateral filtering in spatial domain, a better reconstruction can be achieved.

REFERENCES

[1] Apostolopoulos, J. G. and Jayant, N. S., "Postprocessing for very low bit-rate video compression," *IEEE Transactions on Image Processing* **8**, 1125–1129 (August 1999).

[2] J. Mateos, A. K. K. and Molina, R., "A bayesian approach for the estimation and transmission of regularization parameters for reducing blocking artifacts," *IEEE Trans. Image Processing* **9**, 1200–1215 (July 2000).

[3] C. Weerasinghe, A. W. C. L. and Yan, H., "Artifact reduction in compressed images based on region homogeneity constraints using the projection onto convex sets algorithm," *IEEE Transactions on Circuits and Systems for Video Technology* **12**, 891– 897 (October 2002).

[4] G. A. Triantafyllidis, D. Tzovaras, M. G. S., "Blocking artifact reduction in frequency domain," *International Conference on Image Processing* **1**, 269 – 272 (2001).

[5] T. Chen, H. R. W. and Qiu, B., "Adaptive postfiltering of transform coefficients for the reduction of blocking artifacts," *IEEE Trans. Circuits and Systems for Vide Technology* **11**, 594–602 (May 2001).

[6] Liu, S. and Bovik, A. C., "Efficient dct-domain blind measurement and reduction of blocking artifacts," *IEEE Trans. Circuits and Systems for Vide Technology* **12** (December 2002).

[7] Li, Z. and Delp, A. J., "Block artifact reduction using a transform-domain markov random field model," *IEEE Transactions on Circuits and Systems for Video Technology* **15** (December 2005).

[8] Tomasi, C. and Manduchi, R., "Bilateral filtering for gray and color images," in [*Proc. Int. Conf. Computer Vision*], 839–846 (1998).

Figure 9. A region from our experiment with the Lena image is displayed. Image size is 64×64; The first row shows the compressed image, the second row shows the original Bilateral filter, and the third row presents the proposed method. From left to right the bit-rate is 0.18, 0.22, 0.24. $k_0 = 1.2$, $k_1 = 10.0$, $\sigma_{r,min} = 20$, $\sigma_{d,min} = 1.5$, $\sigma_r = 20$, $\sigma_d = 3$.

[9] T. Q. Pham, L. J. V., "Separable bilateral filtering for fast video preprocessing," *IEEE International Conference on Multimedia and Expo* (July 2005).

[10] C. Wang, P. Xue, W. L. W. Z. and Yu, S., "Fast edge-preserved postprocessing for compressed images," *EEE Transactions on Circuits and Systems for Video Technology* **16**, 1142 – 1147 (September 2006).

[11] Minami, S. and Zakhor, A., "An optimization approach for removing blocking effects in transform coding," *IEEE Transactions on Circuits and Systems for Video Technology* **5** (April 1995).

[12] G. A. Triantafyllidis, D. T. and Strintzis, M., "Blocking artifact detection and reduction in compressed data," *IEEE Transactions on Circuits and Systems for Video Technology* **12** (October 2002).

Figure 10. The experiment with the Lena image is displayed. Image size is 512×512; and the bit-rate for the compressed image is 0.18. (a) Original image; (b) Input compressed image; (c)The standard bilateral filter with $\sigma_r = 20$, $\sigma_d = 3$; (d) The adaptive bilateral filter with $\sigma_{r,min} = 20$, $\sigma_{d,min} = 3$, $k_0 = 1.0$, $k_1 = 10.0$.

Figure 11. One frame from our experiment with the video sequence "Foreman" is displayed. (a)Uncompressed frame; (b)Compressed frame; (c)Bilateral filter; (d)Proposed method. Image size is 144×137; The input bit-rate is 0.495, $k_0 = 1.2$, $k_1 = 10.0$, $\sigma_{r,min} = 20$, $\sigma_{d,min} = 1.5$, $\sigma_r = 20$, $\sigma_d = 3$.