Compression Artifact Reduction for Low Bit-Rate Images Based on Non-Local Similarity and Across-Resolution Coherence

Jing Mu, Ruiqin Xiong, Xiaopeng Fan and Siwei Ma
Institute of Digital Media, Peking University, Beijing 100871, China
Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China
Email: {jmu, rqxiong, swma}@pku.edu.cn, fxp@hit.edu.cn

Abstract— **This paper proposes a method to estimate coefficients for blocking artifact reduction at low bit rate. Across-resolution coherence that low and high resolution image are similar is introduced to preserve signal continuity. Non-local similarity is used to provide samples for estimation by searching similar blocks of reference block. We have two sources of estimation. One source is exploiting non-local similarity to estimate coefficients of low resolution of decoded image, and interpolating the low resolution image to high resolution. We obtain the coefficients estimation for high resolution image based on the coherence across different resolutions. The other source of estimation is the quantization coefficients. These estimations are fused by their reliability respectively. Experimental results demonstrate that the proposed algorithm outperforms some recently presented methods in terms of both objective and subjective qualities of the reconstruction images.**

I. INTRODUCTION

Block-based Discrete Cosine Transform (BDCT) has been widely used in compression standards. The reconstructed images suffer from blocking artifacts when aggressively quantized. Therefore, post-processing is proposed to improve the visual quality of decoded images and videos. Choy et al. [1] estimated the DCT coefficients from bit stream with local statistic of coefficients. Buades et al. [2, 3] proposed the non-local means filter to estimate the intensity of a pixel as a weighted linear combination of the other pixels. The weights of the combination are determined by the similarity between the neighborhoods of the target and source pixels. But this method does not take advantage of the quantization information and the filtering strength is difficult to determine. In order to reduce the JPEG compression artifacts, Zhai et al. [4] proposed that averages the estimated block and its similar blocks. The similar blocks are selected according to compression quality factors. Zhang et al. [5] and [6] adaptively fused the non-local estimation and the quantization coefficients based on their reliability to reduce compression artifacts.

This work was supported by the National Natural Science Foundation of China (61370114, 61421062), Beijing Natural Science Foundation (4132039), Research Fund for the Doctoral Program of Higher Education (20120001110090) and also by Cooperative Medianet Innovation Center.

Traditional compression method processes each block separately. The continuity between blocks is destroyed. We introduce the coherence across different resolutions of image, which means the low resolution and high resolution of the decoded image are highly similar. We observe that the low resolution image of compressed image has a good visual quality with invisible blocking artifacts, and the discontinuity from block to block is not obvious. High resolution of decoded image is obtained from low resolution image via interpolation. Interpolation is a low-pass filter which preserves signal continuity well. So high resolution image preserves signal coherence via interpolation from low resolution image.

Inspired by the idea mentioned above, this paper proposes an approach to reduce compression artifacts by estimating the DCT coefficients. We exploit non-local similarity of low resolution of decoded image to estimate the coefficients of low resolution. Coefficients of each resolution are estimated by overlapped blocks. For each block, we search its non-local similar blocks and average the similar blocks as an estimation. We estimate the coefficients of high resolution image from low resolution via interpolation iteratively. We also regard the quantization coefficients as an estimation. The different sources of estimation are adaptively fused according to their reliability respectively. Our approach outperforms the existing approaches in terms of both objective and subjective qualities of the reconstruction results.

The remainder of this paper is organized as follows. Section II discusses non-local similarity and across-resolution coherence of image. Section III describes the coefficient estimation method from different prediction sources according to their reliability. Experimental results are reported in Section IV and Section V concludes the paper.

II. NON-LOCAL SIMILARITY AND ACROSS-RESOLUTION COHERENCE FOR COEFFICIENT ESTIMATION

In order to estimate the coefficients, we exploit the non-local similar blocks of current block to provide statistical samples for the current block coefficient estimation and utilize the coherence across different resolutions to preserve the signal continuity. In this section, we will reveal the rationality of non-local similarity and across-resolution coherence.

A. Non-Local Similarity for Coefficient Estimation

978-1-4799-8391-9/15/$31.00 ©2015 IEEE

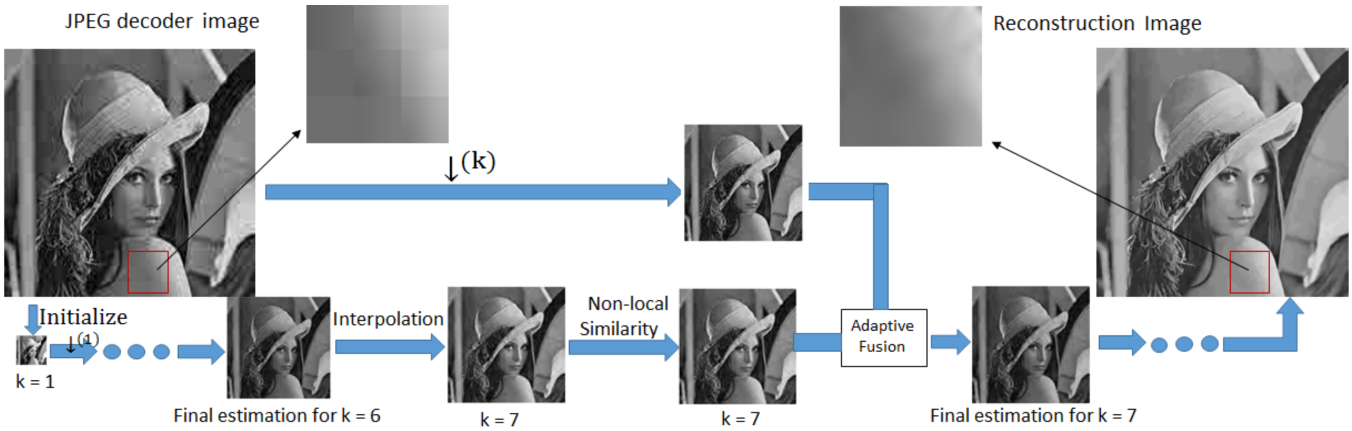


Fig. 1 Illustration of proposed method. k represents the resolution index and $\downarrow^{(k)}$ is the down-sampling operator, which extracts top left $k\times k$ bands of each block with the scale factor to form a new image.

The key idea of the non-local means filter is that a given image is filtered by Eq. (1),

$$u(x)=\int w_f(x,y)f(y)dy \tag{1}$$

where $w_f \propto \exp\left(-\frac{\|f_x-f_y\|^2}{h^2}\right)$. The weight w depends on the similarity measure of two blocks.

According to the property of natural image that the coefficients in similar blocks usually have similar statistical properties, we infer the coefficient distribution of a pixel utilizing the coefficients of pixels in similar blocks in a non-local area of the image. Coefficients of each pixel can be overlapped estimated based on Eq. (1).

B. Across-Resolution Coherence for Coefficient Estimation

In DCT domain, it is easy to realize resolution conversion. Fig. 2 shows the period of frequency bands when the number of sample points changes from N to M in the [1-1/2, N+1/2] interval. According to the definition of DCT and IDCT, we get the high resolution coefficients $x(t)$ from low resolution coefficients $z(\hat{t})$ with a scale factor as Eq. (2).

$$x(t)=z(\hat{t})\times\sqrt{\frac{M}{N}} \tag{2}$$

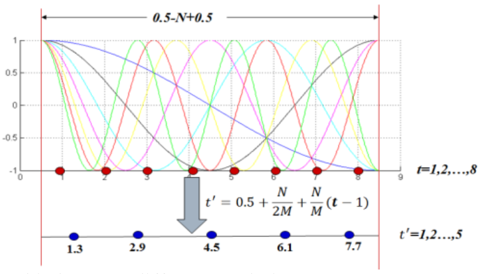


Fig. 2 Relationship between different resolutions.

In image processing, interpolation is the process of sample rate conversion. Referring to the definition of interpolation which is a low-pass filter, interpolation can preserve the signal continuity. High resolution image is obtained from low resolution image via interpolation, and the high resolution image is highly similar to low resolution image. The high frequency of each block is produced. Medium frequency is modified, but they still keep up with the low resolution coefficients. Finally, the low frequency will be preserved mostly.

III. The Proposed Algorithm

Let $\boldsymbol{x}$ represents the original data of image I. We use $\boldsymbol{x_B}$ and $\boldsymbol{X_B}$ to represent the pixel intensity and transform coefficients of block B. They are related by DCT – T, $\boldsymbol{X_B}=T(\boldsymbol{x_B})$. $\boldsymbol{Y_B}$ is the reconstructed coefficients gained by $\boldsymbol{Y_B}=Q(\boldsymbol{X_B})$, and Q is quantization matrix. X and Y are BDCT coefficients of image I.

A. Framework of Algorithm

As shown in Fig.1, the coefficients of high resolution image are estimated from low resolution of decoded images iteratively. We extract $k\times k$ coefficients of each block to form a new coefficient matrix, and the matrix is estimated by overlapped $k\times k$ blocks. For each block, we find its non-local similar blocks and average the similar blocks as an estimation of the $k\times k$ block coefficients. The quantization information is used as another estimation. The two sources of estimation are adaptively fused as the k^{th} resolution coefficient estimation. We get a low resolution image with resolution index k of the decoded image by transforming the coefficient matrix to image domain. The k+1$^{\text{th}}$ resolution image can be obtained by interpolation of the k^{th} resolution image.

By introducing two distance term D_1 and D_2, the proposed compression artifact reduction method is formulated as the following optimization problem.

$$X^*=\arg\min_x \sum_{B\in\Omega}\left(D_1\left(X_B,Y_B\right)+\sum_{k=2}^{blocksize}D_2\left(X_{B^{(k)}},\left\{Y_{B^{(k)'}}\right\}_{B^{(k)'}\in N\left(B^{(k)}\right)}\right)\right) \tag{3}$$

subject to the quantization constraint:

$$X_B(u,v)\in\left[Y_B^{\min}(u,v),Y_B^{\max}(u,v)\right]. \tag{4}$$

where $B^{(k)}$ is block of the k^{th} resolution and the $N(B^{(k)})$ is the nonlocal block set used to predict the target coefficient block of the relative resolution k.

In Eq. (3), the first term D_1 measures the distance between the estimated coefficients and the reconstructed coefficients from decoded image, which can be regarded as the data fidelity. The second term measures the distance between the estimated coefficients and the statistic of non-local similar blocks in each resolution.

In order to obtain X^*, we solve Eq.(3) from resolution k=2 to block size. Eq. (3) can be written as (5).

$$X^* = \arg\min_x \sum_{k=2}^{blocksize} \left(\sum_{B\in\Omega} \left(D_1\left(X_{B^{(k)}}, Y_{B^{(k)}}\right) + D_2\left(X_{B^{(k)}}, \left\{Y_{B^{(k)'}}\right\}_{B^{(k)'}\in N\left(B^{(k)}\right)} \right) \right) \right) \tag{5}$$

To solve Eq. (5), we compute X of each resolution k separately. The result $X^{(k)*}$ is used for the computation of $X^{(k+1)*}$.

B. Reliability for Coefficient Estimation

As the accuracy of coefficient estimation is affected by the presence of quantization error, the coefficients in low frequency bands are more accurate. Based on such discussion, the reliability of the estimations should not be calculated by block, but by frequency band. The low frequency bands which are not sensitive to noise have high reliability, and the high frequency bands have low reliability relatively. The reliability can be adaptively computed. According to the discussion in Part A, we use Eq. (6) for the data fidelity term,

$$D_1\left(X_B, Y_B\right) = W_1 \left\| X_B - Y_B \right\|_2^2 \tag{6}$$

$$W_1 \propto \frac{1}{\sigma_Q^2} \tag{7}$$

where W_1 is the reliability of the quantization coefficients , which is inversely proportional to the variance of quantization error.

For the D_2, we exploit non-local similarity in each resolution. And the coherence across different resolutions is implemented by interpolation. Due to image structure variation, different coefficient blocks play different roles in the prediction process. Therefore, the coefficients of non-local similar blocks are weighted averagely. The expectation is used as an estimation of coefficients. The aggregation measure of similar blocks, i.e. variance of similar block coefficients, is used to reflect the reliability of the prediction. Based on the above discussion, D_2 can be formulated as,

$$D_2\left(X_{B^{(k)}}, \left\{Y_{B^{(k)'}}\right\}_{B^{(k)'}\in N\left(B^{(k)}\right)} \right) = W_2 \left\| X_{B^{(k)}} - \left\{Y_{B^{(k)'}}\right\}_{B^{(k)'}\in N\left(B^{(k)}\right)} \right\|_2^2 \tag{8}$$

$$W_2 \propto \frac{1}{\sigma^2_{N\left(B^{(k)}\right)}} \tag{9}$$

$$\sigma^2_{N\left(B^{(k)}\right)} = \sum_{B^{(k)'}\in N\left(B^{(k)}\right)} w_{B^{(k)'}} \left(Y_{B^{(k)'}} - \sum_{B^{(k)'}\in N\left(B^{(k)}\right)} w_{B^{(k)'}} Y_{B^{(k)'}} \right) \tag{10}$$

where $w_{B'(k)}$ is to weight the similarity between the similar blocks and reference block as mentioned in Section II.A. The weight W_2 is inversely proportional to the variance of coefficients of similar blocks in each resolution.

IV. Experimental Results

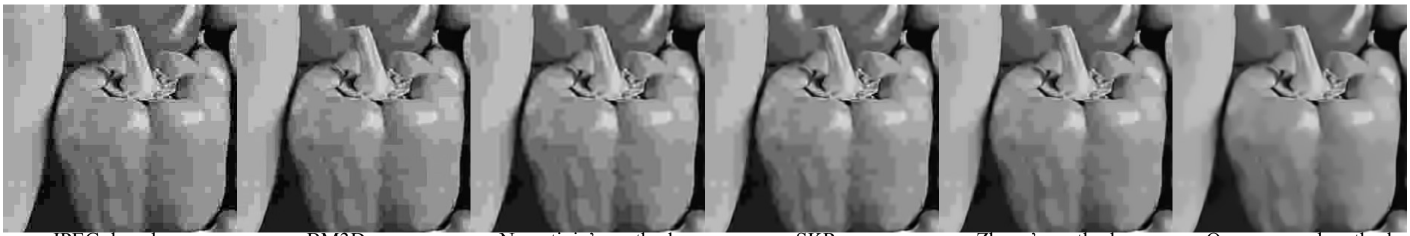
To evaluate the performance of our proposed scheme, we compare it with some recently presented methods. For our proposed method, we set block size as 8×8 and use k from 2 to block size. In order to get a better performance, we introduce overlapped blocks and estimate them independently as many methods before. We consider both coding blocks and non-coding blocks at the same time. We only take the overlapped step size as half of the block size to reduce the complexity. To get high resolution image from low resolution of decoded image, we use the '*lanczos3*' method to interpolate.

The PSNR and SSIM results of the reconstruction algorithms for 5 classic test images at low bit rate (low quality factor) are given. Due to limited space, only QF = 3 and QF = 5 are listed in Table I and II, where we have highlighted the best method for each image at every QF with underlines. It can be found that the proposed algorithm is consistently better compared to recently presented algorithms (JPEG decoder [7], BM3D [8], Nosratinia's method [9], SKR [10], and Zhang's method [5] which is the state-of-the-art method) in all the testing scenarios. The best average PSNR of the proposed algorithm is found to outperform all competitors with an average PSNR gain of about 0.14 dB over the best methods of tested algorithms set when QF = 5. Average PSNR gain is about 0.36 dB when QF = 3.

The processing results of the image reconstruction algorithms of the test images *Peppers*(512×512), *Lena* (512×512) and *Elaine*(512×512) with QF=5 are shown in Fig. 3, 4 and 5. In order to observe the visual quality, we extract part of the whole image. Obviously, our proposed method both preserves sharp edges and smooth areas, showing much clearer and better visual results than the other competing methods at very low bit rate. The high performance of our proposed method is attributed to the coefficient estimation which exploits non-local similarity of low resolution of decoded image to estimate the low resolution coefficients and interpolates the low resolution image to high resolution to estimate the high resolution coefficients.

V. Conclusions

In order to reduce blocking artifacts at low bit rate, this paper proposes an approach to estimate the DCT coefficients from two sources. Non-local similarity is used to provide similar blocks for estimating the current block. Across-resolution coherence that low and high resolution image are similar is introduced to preserve signal continuity. One source is exploiting non-local similarity to estimate coefficients of low resolution of decoded image, and interpolating the low resolution image to high resolution. We obtain the coefficients estimation for high resolution image based on the coherence across different resolutions. Quantization coefficients are the other source. The estimations are adaptively fused by their reliability. The proposed method achieves a remarkably better reconstruction quality than recently presented methods in both subjective and objective quality.



JPEG decoder PSNR=27.17 dB | BM3D PSNR=28.16 dB | Nosratinia's method PSNR=28.44 dB | SKR PSNR=28.60 dB | Zhang's method PSNR=28.78 dB | Our proposed method PSNR=28.97 dB

Fig. 3 Part of the reconstructed images with different methods. The test image, *Peppers*, is compressed by JPEG at *QF* = 5 with bpp= 0.1782 (bits per pixel).

JPEG decoder PSNR=27.33dB | BM3D PSNR=28.30 dB | Nosratinia's method PSNR=28.47 dB | SKR PSNR=28.79 dB | Zhang's method PSNR=28.85 dB | Our proposed method PSNR=28.96 dB

Fig. 4 Part of the reconstructed images with different methods. The test image, *Lena*, is compressed by JPEG at *QF* = 5 with bpp= 0.1749 (bits per pixel).

JPEG decoder PSNR=27.50dB | BM3D PSNR=28.42 dB | Nosratinia's method PSNR=28.63 dB | SKR PSNR=29.03 dB | Zhang's method PSNR=29.02 dB | Our proposed method PSNR=29.06 dB

Fig. 5 Part of the reconstructed images with different methods. The test image, *Elaine*, is compressed by JPEG at *QF* = 5 with bpp= 0.1577 (bits per pixel).

TABLE I. PSNR quality (in dB) and SSIM quality of reconstruction images using different methods for test images compressed with QF = 3.

Image	JPEG	BM3D	Nosritinia's	SKR	Zhang's	Proposed
Lena	24.83/0.672	25.60/0.704	25.89/0.724	26.10/0.734	26.34/0.739	**26.80/0.757**
Pepper	24.83/0.645	25.66/0.683	26.12/0.713	26.14/0.718	26.51/0.728	**26.85/0.739**
Parrot	25.94/0.741	26.58/0.766	26.84/0.787	26.90/0.802	27.43/0.807	**27.96/0.828**
Elaine	25.26/0.566	26.02/0.601	26.42/0.629	26.67/0.637	26.86/0.640	**27.02/0.649**
Caps	25.92/0.727	26.46/0.747	26.76/0.762	26.83/0.769	27.03/0.772	**27.32/0.805**
Avg	25.36/0.670	26.06/0.700	26.41/0.723	26.53/0.732	26.83/0.737	**27.19/0.756**

TABLE II. PSNR quality (in dB) and SSIM quality of reconstruction images using different methods for test images compressed with QF = 5.

Image	JPEG	BM3D	Nosritinia's	SKR	Zhang's	Proposed
Lena	27.33/0.737	28.30/0.774	28.47/0.787	28.79/0.800	28.85/0.797	**28.96/0.803**
Pepper	27.17/0.708	28.16/0.751	28.44/0.769	28.60/0.780	28.78/0.779	**28.97/0.791**
Parrot	28.36/0.779	29.13/0.814	29.35/0.830	29.44/0.845	29.86/0.844	**30.06/0.849**
Elaine	27.50/0.620	28.42/0.661	28.63/0.677	29.03/0.690	29.02/0.685	**29.06/0.690**
Caps	27.90/0.758	28.56/0.784	28.69/0.792	28.82/0.802	29.01/0.803	**29.13/0.808**
Avg	27.65/0.724	28.51/0.757	28.72/0.771	28.94/0.783	29.10/0.782	**29.24/0.789**

REFERENCES

[1] S. S. O. Choy, Y.-H. Chan, and W.-C. Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients," *IEEE Signal Process*. Lett., vol. 4, no. 1, pp. 5–7, Jan. 1997.

[2] A. Buades, B. Coll, and J.-M. Morel, "A non-local algorithm for image denoising," in *Proc. IEEE Int. Conf. CVPR*, vol. 2. Jun. 2005, pp. 60–65.

[3] A. Buades, B. Coll, and J. M. Morel, "A review of image denoising algorithms, with a new one," *SIAM J. Multiscale Model. Simul.*, vol. 4, no. 2, pp. 490–530, Jan. 2005.

[4] G. Zhai, W. Zhang, X. Yang, W. Lin and Y. Xu, "Efficient Image Deblocking Based on Postfiltering in Shifted Windows," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 18, no.1, pp.122-126, Jan. 2008.

[5] X. Zhang, R. Xiong, X. Fan, S. Ma, and W. Gao, "Compression artifact reduction by overlapped-block transform coefficient estimation with block similarity," *IEEE Trans. Image Process*., vol. 22, no. 12, pp. 4613–4626, Dec. 2013.

[6] X. Zhang, R. Xiong, S. Ma and W. Gao, "Reducing Blocking Artifacts in Compressed Images via Transform-Domain Non-local Coefficients Estimation," *IEEE International Conference on Multimedia and Expo (ICME)*, pp.836-841, Jul. 2012.

[7] (2007) JPEG Encoder and Decoder. [Online]. Available: http://www.ijg.org/

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

[9] A. Nosratinia, "Enhancement of JPEG-compressed image by reapplication of JPEG," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 27, no. 2, pp. 69–79, Feb. 2001.

[10] H. Takeda, S. Farsiu, and P. Milanfar, "Kernel regression for image processing and reconstruction," *IEEE Trans. Image Process*., vol. 16, no. 2, pp. 349–366, Feb. 2007.