Correlation based Universal Image/Video Coding Loss Recovery

Jinjian Wu[a], Weisi Lin[b,*], Guangming Shi[a], Jimin Xiao[c]

[a] *School of Electronic Engineering, Xidian University, China*
[b] *School of Computer Engineering, Nanyang Technological University, Singapore*
[c] *Department of Signal Processing, Tampere University of Technology, Finland*

Abstract

Coding artifacts are annoying in highly compressed signals. Most of the existing artifact reduction methods are designed for one specific type of artifacts, codecs, and bitrates, which are complex and exclusive for one type of artifact reduction. Since both the compressed image/video and the coding error contain information of the original signal, they are highly correlated. Therefore, we try to recover some lost data based on the correlation between the compressed signal and the coding error, and introduce a novel and universal artifact reduction method. Firstly, according to the spatial correlation among pixels, a pixel-adaptive anisotropic filter is designed to reconstruct the distorted signal. Next, a globally optimal filter is designed to further recover the coding loss. Experimental results demonstrate that within an extensive range of bitrates, the proposed method achieves about 0.8 dB, 0.45 dB, 0.3 dB, and 0.2 dB on average of PSNR improvement for JPEG, MPEG4, H.264/AVC, and HEVC compressed signals, respectively.

Keywords: Artifact Reduction, Coding Loss Recovery, Correlation, Structural Self-Similarity, Global Optimization

1. Introduction

Due to the limited available bandwidth and storage, image/video signals need to be compressed with reduced bitrates [1]. During the lossy image/video compression (e.g., JPEG, JPEG2000,

☆This work is supported by the Major State Basic Research Development Program of China (973 Program, No.2013CB329402), the Research Fund for the Doctoral Program of Higher Education (No. 20130203130001 and 20130203120009), the NSF of China (Nos. 61227004 and 61033004), and the Fundamental Research Funds for The Central Universities (No. JB140227).

*Corresponding author.

Email addresses: jinjian.wu@mail.xidian.edu.cn (Jinjian Wu), wslin@ntu.edu.sg (Weisi Lin), gmshi@xidian.edu.cn (Guangming Shi), xiaojimin1981@gmail.com (Jimin Xiao)

MPEGx, H.26x, etc.), the quantization process causes artificial discontinuity, such as blocking, ringing, mosquito noise, and so on [2, 3]. In order to ameliorate the qualities of compressed images/videos, appropriate postprocessing techniques are required at the decoder side [4, 5].

In the past decade, a series of coding artifact reduction methods have been introduced. In the early stage, according to the dominant blocking artifacts in low bit rate, many deblocking methods, which result in filtration (e.g., low-pass or average property) of signals [6–10], were proposed. In fact, these deblocking procedures are smoothing operators, which smooth out both the blocking artifacts and image details (e.g., texture and edge). In order to minimize the damage to the image details (e.g., texture and edge) and remove the blocking artifacts, many post-filtering schemes assume that most artifacts occur at some predictable positions [11–15]. However, for MPEGx or H.26x compression, the blocking artifacts may occur anywhere of a frame as a result of motion compensation [16–18]. Therefore, the positions where the blocking artifacts appear are unable to be predicted [13, 19] for MPEGx or H.26x compression. In addition, image regions with different types contents (i.e., plain, texture, edge, etc.) are firstly discriminated for adaptive processing [20–22]. However, the discrimination of image contents is extremely difficult, and it is also hard to accurately separate the blocking/ringing artifacts from the edge/texture information. Besides, these existing approaches are content and codec dependent, and too many models and parameters need to be decided. As a result, these existing approaches are restricted to certain types of codecs, artifacts, and bitrates. Therefore, a universal coding artifact reduction method is required.

In essence, the coding artifact reduction procedure can be viewed as an attempt of coding loss recovery (the error between the original and the compressed images) [19]. Therefore, we try to reduce the coding artifacts by image content reconstruction/recovery in the decoder side according to the correlations of image pixels. It is well known that image contents are highly correlated with their surroundings [23], and these correlated contents jointly present some self-similar structures [24]. The structural self-similarity is widely explored for image recovery. For example, according to the characteristics of structures, distortions can be accurately removed for image denoising [25, 26]; distorted structures can be efficiently recovered for image deblurring [27] and reconstruction [28, 29]. Moreover, in image compression, the compressed and lost parts usually present some common structural information of the original image. In other words, the coding loss is also correlated with the compressed image [19, 30]. Therefore, we can recover the coding loss according to its correlation with the compressed image.

In this paper, we try to recover the coding loss according to the correlation properties among image/video contents, and introduce an universal artifact reduction algorithm. Firstly, we analyze the relationship between the compressed image/video and the coding error. Secondly, the structural characteristics of the compressed image is explored, and a pixel-adaptive anisotropic filter is designed to reconstruct the distorted contents by exploiting these highly correlated pixels. Thirdly, based on the correlation between the compressed image and the coding error, a global optimal filter is designed to further recover the coding loss. Since the proposed algorithm recovers the lost data based on the correlation among image/video contents, it is not restricted to one specific artifact type. It is applicable to all codecs under any bitrate. Experimental results on JPEG, MPEG4, H.264/AVC, and HEVC codecs demonstrate the effectiveness of the proposed algorithm.

The rest of this paper is organized as follows. In Section 2, we explore and illustrate the correlation between the compressed image and the coding loss. The detailed recovery implementation of the proposed artifact reduction method is presented in Section 3. Then, in Section 4 the performance of the proposed method is demonstrated. Finally, we draw the conclusions in Section 5.

2. Image Correlation Analysis

During the lossy compression, some visual information is discarded through the coefficient quantization, which reduces the amount of data for transmission and storage [16, 17]. Let $I(x)$ denotes a pixel in the original image, and $I_c(x)$ denotes the pixel in the compressed image (e.g., with JPEG, JP2K, MPEG4, etc). The compression loss (discarded information, ΔI) can be expressed as,

$$\Delta I(x) = I(x) - I_c(x), \tag{1}$$

where $\Delta I(x)$ is the lost data of pixel $I(x)$.

Generally, the compressed image I_c possesses the main contents of the original image, and the compressed loss ΔI contains the residual information of the original image. Both of them

(a) (b) (c)

Figure 1: An example of correlation between the coded image and the lost data. (a) The original image. (b) JPEG coded image. (c) Lost data, in which the values are mapped into $[0, 255]$ for a better view.

possess the structural information of the original image, and jointly present the visual contents. An example of JPEG coding is shown in Fig. 1, in which (b) and (c) are the compressed and lost parts of (a). As can be seen, though Fig. 1 (b) is distorted by noticeable artifacts, we can acquire most of the visual information from Fig. 1 (b) as that provided by the original image (Fig. 1 (a)); meanwhile, Fig. 1 (c) presents the coarse contours of objects in Fig. 1 (a), with which we can imagine and recognize objects in the original image (i.e., peppers in Fig. 1 (a)). In summary, Fig. 1 (b) and (c) present some common structural information of Fig. 1 (a) (though Fig. 1 (b) provides more detailed structures and Fig. 1 (c) only shows the coarse structures of the original image). Therefore, the contents of the compressed image and the lost data are correlated at certain level.

In order to further illustrate the correlation between the compressed image and the lost data, a thorough analysis on image content similarity is made. Structural information represents the main content of an image, which is adapted to be extracted for image perception and understanding [31, 32]. Therefore, we try to extract structural information from images for content similarity analysis. From the view of image processing, global image structure is usually represented by local binary pattern based histogram [33], and local image structure is always represented by luminance difference [34] or local variance [31]. In this work, the local variance on luminance is

Table 1: Correlation between the compressed image and the lost data

image \ bpp	0.25	0.5	1	2	3
Pepper	0.7845	0.7418	0.6483	0.3562	0.1241
Boat	0.7863	0.7718	0.7407	0.6467	0.5168
Barbara	0.6959	0.6790	0.6756	0.6241	0.4733
Huts	0.6355	0.6050	0.5570	0.4264	0.2755

adopted for local structure analysis,

$$\mathcal{V}(x) = \frac{1}{(2r+1)^2} \sum_{\Delta x} (I(x+\Delta x) - \bar{I}(x))^2, \tag{2}$$

$$\bar{I}(x) = \frac{1}{(2r+1)^2} \sum_{\Delta x} I(x+\Delta x), \tag{3}$$

where $\mathcal{V}(x)$ is the variance of pixel $I(x)$, Δx is the location shift in the local region, $\bar{I}(x)$ is the local mean value of pixel $I(x)$, and r is the radius of the local region.

According to the Pearson correlation function [35], the structural correlation between the compressed image and the lost data can be computed as,

$$C(I_c, \Delta I) = \frac{1}{N} \sum_{x} \frac{\sum_x (\mathcal{V}_c(x) - \overline{\mathcal{V}}_c)(\mathcal{V}_d(x) - \overline{\mathcal{V}}_d)}{\sqrt{(\sum_x (\mathcal{V}_c(x) - \overline{\mathcal{V}}_c)^2)(\sum_x (\mathcal{V}_d(x) - \overline{\mathcal{V}}_d)^2)}}, \tag{4}$$

where $C(I_c, \Delta I)$ is the similarity coefficient between I_c and ΔI; N is the total pixel number of image I_c; $\mathcal{V}_c$ is the structural information of the compressed image I_c, and $\overline{\mathcal{V}}_c$ is the average value of $\mathcal{V}_c$; $\mathcal{V}_d$ is the structural information of the lost data ΔI, and $\overline{\mathcal{V}}_d$ is the average value of $\mathcal{V}_d$.

Four oft-used images (i.e., Pepper, Boat, Barbara, and Huts) are chosen for correlation analysis, which are mainly composed by smooth, edge, regular texture, and irregular texture, respectively. The four images are compressed with JPEG codec under five different bitrates (i.e., 0.25, 0.5, 1, 2, and 3 bpps), and the correlations between these compressed images and their corresponding lost data are computed with (4). The correlation coefficients of these images are listed in Tab. 1, and their corresponding statistical values are shown in Fig. 2. From Tab. 1 and Fig. 2 we can

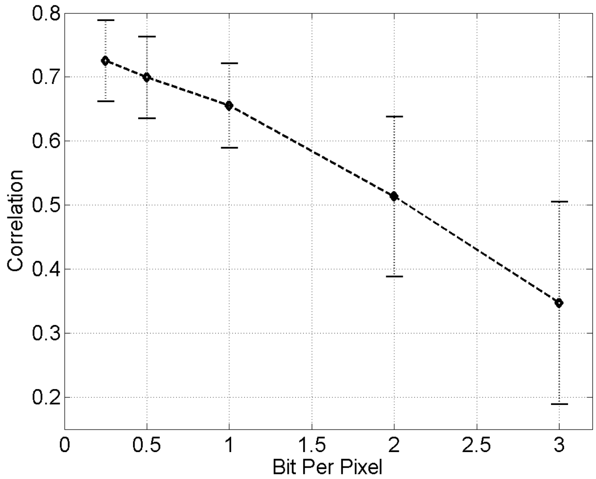


Figure 2: Correlation between the coded image and the lost data

see that under high bitrate (e.g., 3 bpp), the compressed images are less correlated with their corresponding lost data; whereas under low bitrate (e.g., 0.25 bpp), the correlations between the compressed images and their lost data are strong. This is because under high bitrate, almost all of the detail of the original image is preserved in the compressed image, and the lost data contains little structural information of the original image; when the bitrate is low, more content is lost during the compression, and the lost data contains much common structural information as that in the compressed image. In summary, the compressed image is correlated with its lost data, and the correlation between them will become stronger with the decrease of the bitrate.

According to the correlation analysis mentioned above, we suggest to separate the lost data into two portions,

$$\Delta I(x) = \Delta I_c(x) + \Delta I_n(x), \tag{5}$$

where ΔI_c represents the lost information related to I_c, and ΔI_n represents the lost information related to random noise. We regard $\Delta I_c(x)$ as an approximation of $\Delta I(x)$, namely, $\Delta I(x) \approx \Delta I_c(x)$, because only ΔI_c is predictable according to I_c. In the next section, we attempt to recover the lost portion ΔI_c based on I_c.

3. Coding Loss Recovery

As analyzed in the above section, there exists correlation between the compressed image and the lost data. In this section, we attempt to recover some coding loss based on the correlation. Firstly, according to the correlations among nearby pixels, some coding loss on image structures will be recovered through structural restoration. Then, a globally optimal filter is designed to further improve the coding loss recovery.

3.1. Structural Self-Similarity Based Loss Recovery

Natural scenes are redundant, in which nearby pixels are always highly correlated and usually present self-similar structures, e.g., all pixels in a plain region are highly similar; and these pixels with large luminance differences in a edge region jointly present the edge structure. This self-similarity characteristic is useful for image recovery and reconstruction. According to the structural self-similarity, these correlated pixels are found out to remove the noise [25, 26] and restore image contents [27, 29]. Here we adopt the structural self-similarity procedure, and recover some loss of a pixel based on the correlation with its surroundings.

Firstly, the correlations between the central pixel $I(x_o)$ and its surrounding pixels $I(y_i)$ are computed to find out these similar ones. The correlation coefficient between two pixels $I(x_o)$ and $I(y_i)$ is usually computed as the distance of two patches $\Omega(x_o)$ and $\Omega(y_i)$ (which centered at $I(x_o)$ and $I(y_i)$, respectively) [24, 25, 27].

$$\mathcal{S}(x_o, y_i) = \frac{1}{\alpha(x_o)} \exp\left(-\frac{d(x_o, y_i)}{2\sigma^2}\right), \tag{6}$$

$$d(x_o, y_i) = \left\|\Omega(x_o) - \Omega(y_i)\right\|_2^2, \tag{7}$$

$$\alpha(x_o) = \sum_{y_i} \exp\left(-\frac{d(x_o, y_i)}{2\sigma^2}\right), \tag{8}$$

where $\mathcal{S}(\cdot)$ is the correlation coefficient, $\alpha(\cdot)$ is a normalizing constant, and $d(\cdot)$ is the sum of squared differences ($\|\cdot\|_2^2$) between two patches. The parameter σ is a decay rate controller. As analyzed in the above section, the correlation between the compressed image and the lost data is related to the compression rate. Under high bitrate, only a little information is lost during the compression. So, the loss data is less correlated with the compressed image, and a fast decay rate (with small σ value) is required to pop out these highly correlated pixels for content restoration. When under low bitrate, much detailed information is lost during the compression, and therefore, the lost data are highly correlated with the compressed image. In this case, a slow decay rate (with large σ value) is required to find out these correlated information as much as possible for the lost data recovery. Therefore, the decay factor is related to the lost data during the coding, and we set the value of σ according to the energy of the lost data for the compressed image,

$$\sigma^2 = \left\| I - I_c \right\|_2^2, \tag{9}$$

where $\|\cdot\|_2^2$ is the mean squared difference between the original image I and the compressed image I_c. The value of σ is computed in the encoder side and transmitted to the decoder side within the overhead information.

At this stage, according to the correlations with the surrounding pixels, the value of the central pixel $I(x_o)$ is predicted and restored as following [25],

$$I_1(x_o) = \sum_{y_i \in \mathcal{R}} \mathcal{S}(x_o, y_i) I_c(y_i), \tag{10}$$

where $I_1(x_o)$ is the predicted value of pixel $I(x_o)$; and $\mathcal{R}$ is the searching window, in which these correlated pixels are used for restoration of the central pixel. So far, according to the correlations among pixels, the coding loss is recovered to some extent, and the restored image I_1 (composed with the restored values p_1) is acquired.

3.2. Global Optimization Based Loss Estimation

In the above subsection, the compressed image is restored based on the structural self-similarity. Since the lost data is correlated with the compressed image, we can directly use the correlation between the lost data and restored image I_1 to further improve the loss recovery, which can be expressed as,

$$\Delta I_c = t(I_1), \tag{11}$$

where $t(\cdot)$ represents the correlation between ΔI_c and I_1. As pixels are correlated to their neighborhood, a local filter is designed to predict the lost information of the central pixel,

$$\Delta I_c = F_{\mathcal{R}} \otimes I_1, \tag{12}$$

where $F_{\mathcal{R}}$ is a local filter (with size $\mathcal{R}$) for convolution whose elements indicates the contributions of the neighboring pixels in I_1, and the filter size $\mathcal{R}$ is chosen as a positive odd integer.

With (12), the predicted value of the original image I can be expressed as,

$$I_2 = I_1 + F_{\mathcal{R}} \otimes I_1 = H_{\mathcal{R}} \otimes I_1, \tag{13}$$

$$H_{\mathcal{R}} = F_{\mathcal{R}} + \delta \tag{14}$$

where I_2 is the recovery image, the filter $H_{\mathcal{R}}$ is the same as $F_{\mathcal{R}}$ except for the central element, and δ is Dirac delta function.

The determination of $H_{\mathcal{R}}$ can be regarded as a global optimization problem. Here we aim at the minimization of the coding loss error, so $H_{\mathcal{R}}$ can be set as,

$$\begin{aligned} \hat{H}_{\mathcal{R}} &= \arg\min_{H_{\mathcal{R}}} \|I - I_2\|_2^2 \\ &= \arg\min_{H_{\mathcal{R}}} \|I - H_{\mathcal{R}} \otimes I_1\|_2^2. \end{aligned} \tag{15}$$

Finally, according to (10) and (13), the coding loss in the compressed image I_c is recovered to some extent, and we acquire the recovery image I_2, within which we hope that the coding error $(I - I_2)$ is smaller than the original compressed loss ΔI. In the proposed method, the energy of the lost data σ and an optimal local filter $\hat{H}_{\mathcal{R}}$ are required to transmitted as the side information of the bitstream, which are used for structure recovery and content estimation, respectively.

4. Experimental Results

In this section, the performance of the proposed coding loss recovery method is demonstrated. First, the restored image is analyzed to demonstrate the effectiveness of the proposed coding loss recovery procedure. Next, a comprehensive analysis is given to illustrate the improvement of the proposed method in image/video quality. Finally, the proposed method is compared with the latest artifact reduction methods to further demonstrate its effectiveness. In this experiment, ten oft-used images [36] and five popular video sequences [37] are chosen, and the proposed method performs on four types of codecs (i.e., JPEG, MPEG4, H.264/AVC, and HEVC) with an extensive range of bitrates.

The performance of the proposed method varies according to several parameters. In all experiments given in this section, we use a same set of parameters. By considering the computational complexity, the radius r in (2) is set as 2, and the local window $\mathcal{R}$ in (10) and (12) is set as a 5×5 block in the experiment. Thus, only 26 coefficients (one for the loss energy σ^2 in (9) and 25 for the filter $H_{\mathcal{R}}$ coefficients in (14)) need to be packed and transmitted in the overhead.

4.1. Analysis of Loss Recovery

The proposed postprocessing method deems coding artifact reduction as coding loss recovery, which removes the artifacts and restores image structures based on the correlations among pixels. Therefore, the performance of the proposed method is highly related to the correlation between the coded signal and the coding error, regardless of codecs. In this subsection, we demonstrate the effectiveness of the proposed method on JPEG and MPEG4 codecs for image and video, respectively.

Fig. 3 shows an example of JPEG coding loss recovery on the *Barbara* image. Under the JPEG codec, the *Barbara* image is distorted with noticeable coding artifacts. As shown in Fig. 3 (a), the face and ground regions are disturbed with blockiness artifacts, the edges of the table and arm are contaminated with ringing artifacts, and so on. Though some structures of Fig. 3 (a) are distorted, the primary contents are still remained, e.g., we can still perceive the primary visual contents of the face, clothes, and table. In addition, as shown in Fig. 3 (c), the coding error map contains some contour information of the original image. Therefore, both the coded image and the coding error map contain some edges (e.g., the contours of the table and arm) and texture (e.g., the textures of the tablecloth and the kerchief of the original scene) of the original image and they are correlated. Moreover, according to (4), the correlation coefficient between Fig. 3 (a) and (c) is C=0.7172, which further confirms that they are highly correlated.

According to the correlation between the coding error and the coded signal, the proposed method can remove some coding artifacts and recover some distorted information. Fig. 4 shows four different patches (A, B, C, and D) from the *Barbara* image, where the first column is the original signals, the middle column is the coded signals, and the last column is the restored signals. The correlations between the four coded patches and their corresponding coding errors are

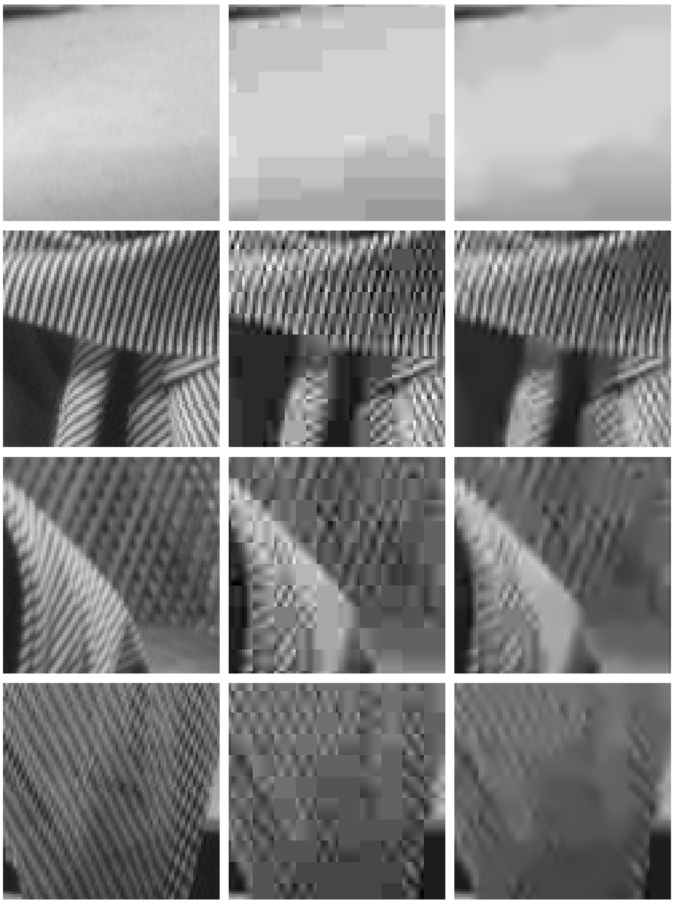
(a) Coded image with PSNR=24.77

(b) Restored image with PSNR=26.01

(c) Coding Error

(d) Restored residual

Figure 3: An example of coding loss recovery for JPEG codec. (a) The JPEG compressed image (PSNR=24.77). (b) The restored image of (a) (PSNR=26.01). (c) Coding loss between the original image and (a), and their correlation coefficient C=0.7172. (d) Residual between the original image and (b), and their correlation coefficient C=0.5866.



(a) Original (b) JPEG coded (c) Restored

Figure 4: An example of coding loss recovery for different image patches. From the top to the bottom rows, they are named as A, B, C, and D patches.

Table 2: Coding Loss Recovery for Image Patches with Different Correlation Coefficients.

Patch	C	PSNR$_c$	PSNR$_r$	ΔPSNR
A	0.7833	32.80	35.35	2.55
B	0.7380	21.41	22.84	1.43
C	0.4936	24.42	25.16	0.74
D	0.2870	23.93	24.36	0.43

listed in Table 2. Patch A locates at the smooth ground. Though A is distorted with obvious blockiness (PSNR=32.80 dB), the content in A is highly correlated with its corresponding coding error (C=0.7833). With the proposed method, the blockiness in A is effectively removed, as the restored patch shows in Fig. 4 (PSNR=35.35 dB). For the edge and texture patch B, the correlation between its coded content and coding error is also strong (C=0.7380) and some distorted contents are successfully restored (ΔPSNR=1.43 dB).

With the decrease of correlation between the coded signal and the coding error, less distorted information can be restored. For example, as patches C and D show in Fig. 4, the correlations between the coded signals and the coding errors (the correlation coefficients are 0.4936 and 0.2870 for C and D, respectively) are weaker than these of A and B. Since the proposed method recoveries coding error depending on the correlations, less information is restored for patches C and D (the ΔPSNR values for them are 0.74 dB and 0.43 dB, respectively). Moreover, if a coded image has a high correlation coefficient with the coding error but a local region is actually weakly correlates to its coding error, the distorted contents in such a local region may mislead the restoration of its coding loss. For example, as shown in patch D (the correlation coefficient of patch D is C=0.2870 and that of the whole image is C=0.7172), when removing the blockiness, some detailed information (especially for these pixels nearby the blockiness contents) is also removed.

In general, the proposed method can effectively remove the JPEG coding artifacts and restore some lost information. By comparing Fig. 3 (a) and (b), it can be seen that the restored image (PSNR=26.01 dB) has a better quality than the coded image (PSNR=24.77 dB). Moreover, the correlation between the error data Fig. 3 (d) and the recovery image Fig. 3 (b) (C=0.5866) is smaller than that between Fig. 3 (a) and (c), which further confirms that some correlated information is restored with the proposed method.

In addition, an example of MPEG4 coding loss recovery is given to demonstrate the effectiveness of the proposed method for video sequences. As shown in Fig. 5, where (a) is the MPEG4 compressed frame with PSNR=28.43 dB (we randomly select a frame from the foreman video sequence) and (b) is the restored frame with PSNR=29.10 dB. According to the correlations between the compressed image and the lost data, many lost contents are effectively recovered with the proposed method. By comparing Fig. 5 (a) with (b), it can be seen that the blockiness in the

(a) Coded frame (b) Restored frame

Figure 5: An example of coding loss recovery with MPEG 4 codec on foreman video sequence. (a) The MPEG4 compressed frame with PSNR=28.43 dB. (b) The restored frame of (a) with PSNR=29.10 dB.

face and wall is obviously removed, and the plain region in (b) is more smooth than that in (a). Moreover, the Gibbs effect in these regions with large luminance changes (e.g., around the edges of the building and the safety helmet) is restored. However, some detailed information is mislabeled as artifact, such as the texture in the building, and is removed in the proposed model. In summary, Fig. 5 (b) is much more similar with the original frame than Fig. 5 (a) (the PSNR value of (b) is higher than that of (a)).

According to the above analysis, we can conclude that the proposed postprocessing method recovers some coding loss based on the inner correlation of images, and it works for different types of codecs (e.g., JPEG and MPEG4) and signals (e.g., still image and video sequence).

4.2. Improvement in Image/Video Quality

In order to make a comprehensive analysis, we verify the proposed method with four types of codec (i.e., JPEG, MPEG4, H.264/AVC, and HEVC) under some oft-used databases (i.e., the USC-SIPI image database [36] and Derf's collection of video test media [37]). And an extensive range of bitrates is chosen in this experiment to demonstrate the quality improvement.

Table 3 shows the quality improvement (represented by the PSNR value) of the proposed

Table 3: The improvement of image quality during the loss recovery on JPEG image, where ΔPSNR means the quality (i.e., PSNR) increase of the recovered image against the coded image.

ΔPSNR (dB) \ bpp	3	1.5	1	0.5	0.3	0.2
Indian	0.38	0.60	0.69	0.67	0.87	0.94
Barbara	0.33	0.63	0.98	1.30	1.22	1.29
Lena	0.50	0.50	0.68	0.92	1.16	1.22
Pepper	0.60	0.25	0.47	0.80	1.22	1.35
Boat	0.55	0.74	0.97	0.94	1.07	1.07
Huts	0.58	0.52	0.58	0.60	0.83	0.96
Car	0.46	0.70	0.80	0.85	0.97	1.03
Airplane	0.55	0.67	0.83	0.95	1.17	1.21
Couple	0.61	0.41	0.47	0.69	1.01	1.11
Milk	0.56	0.36	0.65	1.02	1.06	1.19
Mean	0.51	0.54	0.71	0.87	1.06	1.14

Table 4: The improvement of quality during the loss recovery on MPEG4 coded video sequence, where ΔPSNR means the quality (i.e., PSNR) increase of the recovered video against the coded video.

kbps / dB	1024	512	256	192	128	64
Akiyo	0.20	0.26	0.53	0.51	0.59	0.72
Bowing	0.29	0.38	0.62	0.65	0.58	0.66
Foreman	0.21	0.33	0.38	0.34	0.55	0.76
Mobile	0.39	0.44	0.40	0.38	0.23	0.21
Mother-Daughter	0.37	0.35	0.45	0.39	0.46	0.65
Mean	0.29	0.35	0.48	0.45	0.48	0.60

method against JPEG coding. In this experiment, ten oft-used images from the USC-SIPI image database [36] are chosen. In order to conduct a comprehensive test on the performance of the proposed method, these images are coded in an extensive range of bitrates (0.2∼3.0 bpp; the bitrate of 0.2 bpp is for extremely low rate coding, while 3.0 bpp is far beyond the highest bitrate in common use).

Generally, little information is lost during the JPEG coding with high bitrate, and the correlation between the lost data and the compressed image is low. Thus, less information can be recovered, and the improvement on image quality (ΔPSNR) is limited, as shown in the second column (under 3 bpp) of Table 3. With the decrease of the bitrate, more structural information will leak into the lost data, and the correlation between the loss and the compressed image becomes higher. Therefore, more structural information can be recovered, and an obvious improvement on image quality is achieved, as shown in the last column (under 0.2 bpp) of Table 3. Moreover, the average values of ΔPSNR (as shown in the last row of Table 3) further demonstrate that more information can be recovered under lower bitrate for JPEG coded images. In summary, the improvement is obvious throughout all bitrates of JPEG codec, with an average of 0.51 dB improvement at very high bitrate (3 bpp) and 1.14 dB for low bitrate (0.2 bpp). It achieves about 0.8 dB of average PSNR improvement over all of the bitrates against the standard JPEG scheme.

Table 5: The improvement of quality during the loss recovery on H.264/AVC coded video sequence, where ΔPSNR means the quality (i.e., PSNR) increase of the recovered video against the coded video.

kbps / dB	1024	512	256	192	128	64
Akiyo	0.32	0.27	0.27	0.26	0.23	0.23
Bowing	0.32	0.32	0.37	0.44	0.55	0.49
Foreman	0.28	0.32	0.31	0.32	0.31	0.32
Mobile	0.30	0.30	0.28	0.27	0.22	0.14
Mother-Daughter	0.40	0.34	0.32	0.34	0.32	0.27
Mean	0.33	0.31	0.31	0.32	0.33	0.29

Table 6: The improvement of quality during the loss recovery on HEVC coded video sequence, where ΔPSNR means the quality (i.e., PSNR) increase of the recovered video against the coded video.

kbps / dB	1024	512	256	192	128	64
Akiyo	0.19	0.19	0.20	0.23	0.26	0.37
Bowing	0.13	0.14	0.19	0.20	0.23	0.29
Foreman	0.11	0.01	0.07	0.10	0.14	0.20
Mobile	0.15	0.17	0.18	0.19	0.19	0.18
Mother-Daughter	0.26	0.23	0.22	0.22	0.23	0.21
Mean	0.17	0.15	0.17	0.19	0.21	0.25

Besides JPEG codec, we verify the proposed method on MPEG4, H.264/AVC, and HEVC codecs for video sequences. Five oft-used video sequences (i.e., Akiyo, Bowing, Foreman, Mobile and Mother-Daughter) from Derf's collection of video test media [37] are chosen. The FFmpeg software [38] and HEVC codec [39] are adopted to compress these sequences with an extensive range of bitrates (64∼1024 kbps). IPPP GOP structure is used in the experiment, where the GOP length is 12 frames (the default value in FFmpeg software). The improvements in video quality (represented by increased PSNR) are listed in Tables 4, 5, and 6, respectively.

As shown in Table 4, under high bitrate (1024 kbps), the lost information during MPEG4 compression is less correlated with the coded sequence and little structural content is recovered (with an average improvement of 0.29 dB). While under low bitrate, much more information is lost during MPEG4 compression, and the correlation between the lost data and the compressed sequence is high. With the proposed correlation based loss recovery method, the lower the bitrate is, the more structures that can be restored. Under the lowest bitrate (64 kbps), an average of 0.60 dB improvement is achieved for these sequences. In summary, it achieves about 0.45 dB of average PSNR improvement over all of the bitrates against the standard MPEG4 scheme.

However, the coding loss recovery results on H.264/AVC and HEVC codecs are much different from that of MPEG4. As shown in Table 5 (Table 6), the improvement is less related to the bitrate, and an average of 0.3 dB (0.2 dB) improvement can be achieved for all bitrates. With further analysis we have found that during the H.264/AVC and HEVC compressions, a sophisticated deblocking filter is used and the blcokiness artifacts have already been removed [5, 40, 41]. This result further confirms that the proposed method is not only restrict to a certain type of artifact (blockiness).

In view of the above, the proposed correlation based coding artifact reduction method is regardless of codecs, which works well with JPEG, MPEG4, H.264/AVC, and HEVC. Furthermore, the proposed method can effectively reduce the coding artifacts within an extensive range of bitrates.

4.3. Performance Comparison

Finally, we compare the proposed method with two latest coding artifact reduction methods (i.e., Shao et al.'s method [3] (within which a training model for content classification is demanded

(a) Shao

(b) Zhai

Figure 6: Artifact reduction on JPEG coded image. (a) The output of Shao's model [3] with PSNR=25.16 dB. (b) The output of Zhai's model [13] with PSNR=25.71 dB. Their corresponding JPEG compressed image (with PSNR=24.77 dB) and the output of the proposed model (with PSNR=26.01 dB) are given in Fig. 3.

at the decoder side) and Zhai et al.'s method [13]) (in which a quality score of the decoded image is packed and transmitted in the overhead) to further demonstrate the effectiveness. Fig. 6 shows the artifact reduction results of Shao et al.'s model [3] and Zhai et al.'s model [13] on the *Barbara* image (compressed with JPEG codec), and the corresponding JPEG compressed image and the restored result of the proposed method are shown in Fig. 3.

In Shao et al.'s model, image contents are firstly classified into three types for discriminatively processing, namely, detailed, intermediate, and smooth regions. However, content classification is a difficult work even for uncontaminated image. Moreover, the classification thresholds are related to the content of the image and the energy of noise, however they are set as fixed values in Shao et al.'s model. Therefore, fixed thresholds cannot achieve better results. As shown in Fig. 6 (a), the floor and the wall are accurately classified into smooth regions, and the blockiness artifacts in these areas are effectively reduced. However, the face region is much more complicated and cannot be accurately classified. As a result, the deblocking effect on the face is not so well. Moreover, the artifacts in the edge of arm and scarf are still obvious. Though some artifacts are removed within Shao et al.'s model (the PSNR of Fig. 6 (a) is 25.16 dB, which is higher than that of the original compressed image with PSNR=24.77 dB), its quality is not as good as the proposed method (with PSNR=26.01 dB).

In Zhai et al.'s model, a shifted window is adopted to find out these similar blocks in the surroundings for central block smoothing. The threshold of the similarity determination comes from the quality of the compressed image. As shown in Fig. 6 (b), with Zhai et al.'s model, the blocking artifacts in the smooth regions are effectively removed. However, the luminance change in the edge and texture regions are large, and it is difficult to find out these similar blocks under a fixed threshold. As a result, the artifacts in the arms, edge of the chair, and the scarf cannot be fully removed. And the quality improvement of Zhai et al.'s model on Fig. 6 (b) (with PSNR=25.79 dB) is slightly lower than that of the proposed method (as shown in Fig. 3 (b) with PSNR=26.01 dB).

For a much comprehensive comparison, a large range of bitrate (0.2∼1.2 bpp) is used for test. In this experiment, ten oft-used images from the USC-SIPI image database [36] (same as that listed in Table 3) are chosen, and are compressed with JPEG codec under different bitrates. The

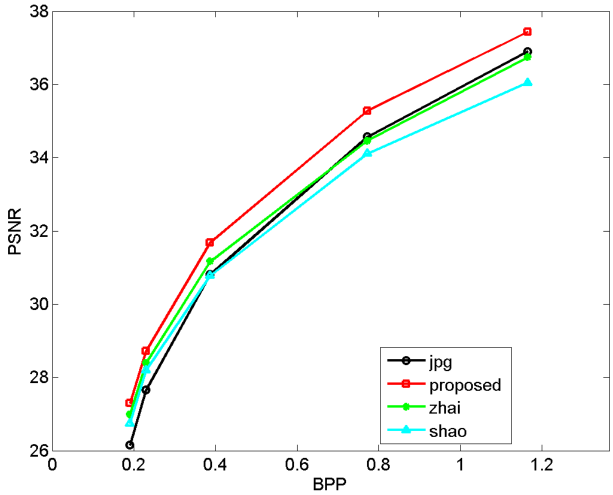


Figure 7: Comparison among JPEG artifact reduction methods.

average performances of these methods on these images are shown in Fig. 7. As can be seen, under low bitrate, all of the three methods achieve improvement on the JPEG coded images, and the proposed method performs slightly better than the other two methods. With the increase of the bitrate, Shao et al.'s method will over smooth the image, and disturb the detail of the original image; with the guide of the image quality factor, Zhai et al.'s method almost does not process the compressed image under high bitrate, which results in a similar PSNR value with the original compressed image; the proposed method can still improve the quality of the compressed image under high bitrate. In conclusion, the proposed method performs better than Shao et al.'s and Zhai et al.'s methods under different bitrates.

Furthermore, according to the inner mechanisms of Shao et al.'s and Zhai et al.'s methods, they are only effective for blocking artifact reduction. While as analyzed in the above subsection, the proposed method works well for all kinds of artifacts, such as blocking, ringing, mosquito, etc. And the proposed method can be used for a large variety of codecs (e.g., JPEG, MPEG4, H.264/AVC, HEVC, etc). Thus, the proposed method is more universal than the existing artifact reduction methods.

5. Conclusion

In this paper, we have introduced a novel postprocessing method for image/video coding artifact reduction. During the lossy compression, the main content of the original signal is preserved in the compressed part. Both of the compressed part and error signal contain the structural information of the original signal, and they are correlated (the correlation decreases when the bit rate increases). We adopted the structural similarity formulation to quantitatively analyze the correlation between them, and confirmed that they are highly correlated especially for low bitrate compression.

Since the compressed image/video and the coding loss data are correlated, we restored some loss data from the compressed image/video based on their correlation. Firstly, according to the correlations among pixels, a pixel-adaptive anisotropic filter was introduced, within which these highly correlated pixels were used to restore the distorted structures. Next, a globally optimal filter was designed to further improve the coding loss recovery. Since the proposed method reduces coding artifacts based on correlation, it can be universally used for postprocessing and works well for all types of codecs, artifacts, and bitrates. Experimental results on image/video coding loss recovery demonstrate that the proposed method achieves 0.8 dB, 0.45 dB, 0.3 dB, and 0.2 dB on average of PSNR improvement on JPEG, MPEG4, H.264/AVC, and HEVC compressed signals, respectively. Comparing with the existing artifact reduction methods, it is more universal and better performing.

[1] M. Shen, P. Xue, and C. Wang, "Down-sampling based video coding using super-resolution technique," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21, no. 6, pp. 755 –765, 2011.

[2] M. Shen and C. C. Jay Kuo, "Review of postprocessing techniques for compression artifact removal," *Journal of Visual Communication and Image Representation*, vol. 9, no. 1, pp. 2–14, 1998.

[3] L. Shao, "Simultaneous coding artifact reduction and sharpness enhancement for block-based compressed images and videos," *Signal Processing: Image Communication*, vol. 23, no. 6, pp. 463–470, 2008.

[4] T. P. O'Rourke and R. L. Stevenson, "Improved image decompression for reduced transform coding artifacts," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 5, no. 6, pp. 490–499, 1995.

[5] Y. Zhang, C. Yan, F. Dai, and Y. Ma, "Efficient parallel framework for H.264/AVC deblocking filter on many-core platform," *IEEE Transactions on Multimedia*, vol. 14, no. 3, pp. 510–524, Jun. 2012.

[6] Z. Xiong, M. Orchard, and Y. Zhang, "A deblocking algorithm for jpeg compressed images using overcomplete

wavelet representations," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 7, no. 2, pp. 433–437, 1997.

[7] H. Choi and T. Kim, "Blocking-artifact reduction in block-coded images using wavelet-based subband decomposition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 5, pp. 801–805, 2000.

[8] J. Mateos, A. K. Katsaggelos, and R. Molina, "A bayesian approach for the estimation and transmission of regularization parameters for reducing blocking artifacts," *IEEE Transactions on Image Processing*, vol. 9, no. 7, pp. 1200–1215, 2000.

[9] S. 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, 2002.

[10] A. Z. Averbuch, A. Schclar, and D. L. Donoho, "Deblocking of block-transform compressed images using weighted sums of symmetrically aligned pixels," *IEEE transactions on image processing*, vol. 14, no. 2, pp. 200–212, 2005.

[11] A. W. Liew and H. Yan, "Blocking artifacts suppression in block-coded images using overcomplete wavelet representation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 4, pp. 450 – 461, 2004.

[12] S. C. Tai, Y. Y. Chen, and S. F. Sheu, "Deblocking filter for low bit rate MPEG-4 video," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 15, no. 6, pp. 733–741, 2007.

[13] G. Zhai, W. Zhang, X. Yang, W. 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, 2008.

[14] G. Zhai, W. Zhang, X. Yang, and W. Lin, "Efficient deblocking with coefficient regularization, shape-adaptive filtering, and quantization constraint," *IEEE Transactions on Multimedia*, vol. 10, no. 5, pp. 735 –745, 2008.

[15] Q. B. Do, M. Luong, and A. Beghdadi, "Coding artifacts reduction method based on a perceptually-weighted variational approach," in *IEEE ISSPIT*, 2009, pp. 101 –106.

[16] J. L. Mitchell, *MPEG Video Compression Standard*. Springer, 1996.

[17] T. Wiegand, G. J. Sullivan, G. Bjontegaard, 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, 2003.

[18] C. Yan, Y. Zhang, F. Dai, and L. Li, "Highly parallel framework for HEVC motion estimation on many-core platform," in *Data Compression Conference (DCC), 2013*, Mar. 2013, pp. 63–72.

[19] W. Lin, J. Zheng, and L. Dong, "Recover image coding loss with LMS filtering," in *IEEE ISCAS*, 2008, pp. 3562 –3565.

[20] Y. L. Lee, H. C. Kim, and H. W. Park, "Blocking effect reduction of JPEG images by signal adaptive filtering," *IEEE Transactions on Image Processing*, vol. 7, no. 2, pp. 229–234, 1998.

[21] H. Paek, R. Kim, and S. Lee, "A dct-based spatially adaptive post-processing technique to reduce the blocking artifacts in transform coded images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10,

no. 1, pp. 36–41, 2000.

[22] L. Shao and I. Kirenko, “Coding artifact reduction based on local entropy analysis,” *IEEE Transactions on Consumer Electronics*, vol. 53, no. 2, pp. 691 –696, 2007.

[23] J. Wu, F. Qi, G. Shi, and Y. Lu, “Non-local spatial redundancy reduction for bottom-up saliency estimation,” *Journal of Visual Communication and Image Representation*, vol. 23, no. 7, pp. 1158–1166, 2012.

[24] J. Wu, F. Qi, and G. Shi, “Self-similarity based structural regularity for just noticeable difference estimation,” *Journal of Visual Communication and Image Representation*, vol. 23, no. 6, pp. 845–852, 2012.

[25] A. Buades, B. Coll, and J. M. Morel, “A non-local algorithm for image denoising,” in *IEEE CVPR*, 2005, pp. 60–65.

[26] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, “Image denoising by sparse 3-d transform-domain collaborative filtering,” *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 1080–1095, 2007.

[27] W. Dong, L. Zhang, G. Shi, and X. Wu, “Image deblurring and Super-Resolution by adaptive sparse domain selection and adaptive regularization,” *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, 2011.

[28] M. Protter, M. Elad, H. Takeda, and P. Milanfar, “Generalizing the Nonlocal-Means to Super-Resolution reconstruction,” *IEEE Transactions on Image Processing*, vol. 18, no. 1, pp. 36–51, 2009.

[29] J. Mairal, F. Bach, J. Ponce, G. Sapiro, and A. Zisserman, “Non-local sparse models for image restoration,” in *IEEE ICCV*, 2009, pp. 2272–2279.

[30] T. Frajka and K. Zeger, “Residual image coding for stereo image compression,” *Optical Engineering*, vol. 42, pp. 182–189, 2003.

[31] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[32] J. Wu, W. Lin, G. Shi, and A. Liu, “Perceptual quality metric with internal generative mechanism,” *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 43–54, 2013.

[33] T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, 2002.

[34] G. E. Legge and J. M. Foley, “Contrast masking in human vision,” *Journal of the Optical Society of America*, vol. 70, no. 12, pp. 1458–1471, 1980.

[35] J. L. Rodgers and W. A. Nicewander, “Thirteen ways to look at the correlation coefficient,” *The American Statistician*, vol. 42, no. 1, pp. 59–66, 1988.

[36] “The USC-SIPI image database,” http://sipi.usc.edu/database/database.php, [Online; accessed 2011].

[37] “Xiph.org video test media,” http://media.xiph.org/video/derf/, [Online; accessed 2012].

[38] “Ffmpeg,” http://www.ffmpeg.org/, [Online; accessed 2013].

[39] "Hevc," https://hevc.hhi.fraunhofer.de/, [Online; accessed 2012].

[40] P. List, A. Joch, J. Lainema, G. Bjontegaard, and M. Karczewicz, "Adaptive deblocking filter," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 614–619, 2003.

[41] G. J. Sullivan, J. 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, 2012.