A New Approach for Restoring Block-Transform Coded Images with Estimation of Correlation Matrices

Steven S. O. Choy, Yuk-Hee Chan and Wan-Chi Siu
Department of Electronic Engineering
The Hong Kong Polytechnic University
Hong Kong, China

Abstract

This paper presents a new restoration approach to reduce coding artifacts in block-transform image coding. Different from conventional restoration techniques, the proposed one is non-iterative and requires light computational cost, yet can reconstruct objectively and subjectively better images. This good performance is achieved because of the following advantages the proposed approach has: (i) efficient incorporation of the solution bound into restoration; and (ii) effective exploitation of local image properties and statistical knowledge about the quantizers used.

1 Introduction

Block-transform coding has been the most widely used technique for lossy image compression, thanks to the widespread adoption of JPEG and MPEG image and video coding standards. However, such a coding technique produces undesirable block boundaries discontinuity in encoded images at high compression ratios. These blocking artifacts are visually not pleasing and they pose the major problem in low-bit rate visual communication applications.

In order to remove these blocking artifacts while maintaining compatibility with current industrial standards, various postprocessing techniques have been proposed. Depending on the angle from which the problem is viewed, the postprocessing deblocking techniques can be broadly divided into two distinct approaches, either image enhancement or image restoration. In general, deblocking through enhancement techniques [1, 2] can be implemented at a low computational cost, but the images they produced may be unfaithful to the original data. On the other hand, deblocking through restoration techniques [3, 4] can reconstruct images to a better quality, but an expensive computational cost has to be paid. There exists a tradeoff between the deblocking performance and the implementation cost.

In this paper, we propose a new technique that can offer a better tradeoff. We tackle the coding artifacts as a restoration problem and apply a general regularization approach to formulate the improved reconstruction. The proposed technique has the following advantages: (i) efficient incorporation of the solution bound formed by the known 'step-size' information about the quantizers; (ii) effective exploitation of local image properties and statistical knowledge about the quantizers used; (iii) light computational cost in realization, which makes real-time application possible. Our experimental results show that the proposed technique could reconstruct images with superior perceived quality as well as significant objective PSNR improvement, as compared with some well-known deblocking techniques [1, 3].

2 Regularized restoration of block-transform coded images

Throughout our formulation, a digital $N\times N$ image, as well as its block transform, is viewed as an $N^2\times 1$ vector in the space R^{N^2} by lexicographic ordering. Block transformation is then represented by a $N^2\times N^2$ matrix which carries out linear transformation from R^{N^2} to R^{N^2}. Then, for an image $\mathbf{x}$, its transform version is $\mathbf{X} = \mathbf{T}_B\mathbf{x}$, where $\mathbf{T}_B$ is the block transformation matrix. The compressed data $\mathbf{Y}$ is obtained by applying a quantization Q to $\mathbf{X}$, which can be written as $\mathbf{Y} = Q[\mathbf{T}_B\mathbf{x}]$. The encoded image with blocking artifacts is given by $\mathbf{y} = \mathbf{T}_B^t Q[\mathbf{T}_B\mathbf{x}]$. The operator Q basically performs a many-to-one mapping from R^{N^2} to R^{N^2}. Therefore, there are many images whose encoded versions are $\mathbf{y}$. Let X_c denote the set of all these images, i.e.

$$X_c = \{\mathbf{z} : Q[\mathbf{T}_B\mathbf{z}] = \mathbf{Y}\}. \qquad (1)$$

This set gives us a constraint for seeking the original image $\mathbf{x}$.

0-8186-8183-7/97 $10.00 © 1997 IEEE

Although the encoded image $\mathbf{y}$ provides the maximum fidelity to the compressed data $\mathbf{Y} = Q[\mathbf{T}_B\mathbf{x}]$, that $\mathbf{y}$ contains blocking artifacts conflicts with our prior knowledge about the solution. Low-pass filtering $\mathbf{y}$ can easily remove the blocking artifacts, but it at the same time corrupts the image details such as edges. Therefore, the problem we are faced with is how to remove the blocking artifacts without distorting the image too much. We employ the idea of regularization in functional spaces [5] as the framework for resolving this problem. Its basic feature is the introduction of a compromise between the fidelity to the data and the fidelity to some prior knowledge about the solution. This compromise is usually measured with a single objective functional.

Denote the *a priori* mean of $\mathbf{x}$ as $\check{\mathbf{x}}$. We propose applying the following objective functional to formulate the regularized solution $\hat{\mathbf{x}}$

$$\hat{\mathbf{x}} = \arg\min_{\hat{\mathbf{x}} \in X_c} \left\{ \|\hat{\mathbf{x}} - \mathbf{y}\|^2_{\mathbf{R}} + \|\hat{\mathbf{x}} - \check{\mathbf{x}}\|^2_{\mathbf{S}} \right\}. \tag{2}$$

Here, $\|\hat{\mathbf{x}} - \mathbf{y}\|^2_{\mathbf{R}}$ and $\|\hat{\mathbf{x}} - \check{\mathbf{x}}\|^2_{\mathbf{S}}$ denote respectively $(\hat{\mathbf{x}} - \mathbf{y})^t\mathbf{R}(\hat{\mathbf{x}} - \mathbf{y})$ and $(\hat{\mathbf{x}} - \check{\mathbf{x}})^t\mathbf{S}(\hat{\mathbf{x}} - \check{\mathbf{x}})$. The former is a weighted quadratic distance applied to measure the fidelity of the solution to the encoded data, while the latter measures the fidelity to the prior information.

Using an objective functional to formulate a regularized restoration of block-transform coded images was also suggested by Yang *et. al.* [3]. The objective functional they proposed is $\hat{\mathbf{x}} = \arg\min\{\alpha\|\hat{\mathbf{x}} - \mathbf{y}\|^2 + \|\mathbf{L}\hat{\mathbf{x}}\|^2\}$, where α is a nonnegative scalar and $\mathbf{L}$ is a high-pass operator. According to that formulation, the regularized solution is specified by only a scalar and a high-pass operator. Our proposed objective functional (2) is more general in the sense that the flexibility in specifying the regularized solution is introduced by means of two additional matrices.

Through the definition of objective functional (2), we have established a general mechanism for finding a regularized reconstruction from the compressed image. The matrices $\mathbf{R}$ and $\mathbf{S}$ determine the final solution we have. It can be derived that if we aim at the $\hat{\mathbf{x}}$ that minimizes the mean-square-error (MSE) from the original image $\mathbf{x}$, i.e. the $\hat{\mathbf{x}}$ that minimizes $E\{\|\hat{\mathbf{x}} - \mathbf{x}\|\}$, we should have $\mathbf{R} = \mathbf{C}_n^{-1}$ and $\mathbf{S} = \mathbf{C}_x^{-1}$, where

$$\mathbf{C}_n = E\left\{(\mathbf{x} - \mathbf{y})(\mathbf{x} - \mathbf{y})^t\right\}, \tag{3}$$

$$\mathbf{C}_x = E\left\{(\mathbf{x} - \check{\mathbf{x}})(\mathbf{x} - \check{\mathbf{x}})^t\right\}. \tag{4}$$

3 Approach for estimating correlation matrices

The estimation of correlation matrices $\mathbf{C}_n$ and $\mathbf{C}_x$ plays a crucial role in our formulation of MMSE restoration using objective functional (2). In the following we propose an approach to estimate these two matrices.

Let $\mathbf{T}_x$ be a unitary transform that decorrelates $\mathbf{x}' = \mathbf{x} - \check{\mathbf{x}}$. Here, decorrelating $\mathbf{x}'$ means making elements of $\mathbf{T}_x\mathbf{x}'$ much less correlated than those of $\mathbf{x}'$. The correlation matrix $E\{(\mathbf{T}_x\mathbf{x}')(\mathbf{T}_x\mathbf{x}')^t\}$, which is equivalent to $\mathbf{T}_x\mathbf{C}_x\mathbf{T}_x^t$, will turn out to be very sparse and only the diagonal elements, $E\{[\mathbf{T}_x\mathbf{x}']_i^2\}$, are of significant values. In view of that, $\mathbf{T}_x\mathbf{C}_x\mathbf{T}_x^t$ can be well approximated by a diagonal matrix $\mathbf{\Lambda}_{Tx}$ defined by

$$\mathbf{\Lambda}_{Tx} = \mathrm{diag}\left(E\left\{[\mathbf{T}_x\mathbf{x}']_i^2\right\}\right). \tag{5}$$

We then have

$$\mathbf{C}_x = \mathbf{T}_x^t(\mathbf{T}_x\mathbf{C}_x\mathbf{T}_x^t)\mathbf{T}_x \approx \mathbf{T}_x^t\mathbf{\Lambda}_{Tx}\mathbf{T}_x. \tag{6}$$

Based on this approximation, the estimation of the bulk matrix $\mathbf{C}_x$ is simplified to the estimation of the local variance of $\mathbf{T}_x\mathbf{x}'$ through a transformation. This reduces to a large extent the computational requirements of the estimation.

Similarly, we have

$$\mathbf{C}_n = \mathbf{T}_n^t(\mathbf{T}_n\mathbf{C}_n\mathbf{T}_n^t)\mathbf{T}_n \approx \mathbf{T}_n^t\mathbf{\Lambda}_{Tn}\mathbf{T}_n \tag{7}$$

for estimating $\mathbf{C}_n$, where $\mathbf{T}_n$ is a unitary transform that decorrelates $\mathbf{n} = \mathbf{x} - \mathbf{y}$ and $\mathbf{\Lambda}_{Tn}$ is a diagonal matrix defined by

$$\mathbf{\Lambda}_{Tn} = \mathrm{diag}\left(E\left\{[\mathbf{T}_n\mathbf{n}]_i^2\right\}\right). \tag{8}$$

4 Realization algorithm

To estimate $\mathbf{C}_n$ and $\mathbf{C}_x$ with the aforementioned approach, we have to deal with (i) the selections of two unitary transforms, $\mathbf{T}_n$ and $\mathbf{T}_x$, that will be used respectively to decorrelate $\mathbf{n}$ and $\mathbf{x}'$; (ii) the estimation of local variances $E\left\{[\mathbf{T}_n\mathbf{n}]_i^2\right\}$ and $E\left\{[\mathbf{T}_x\mathbf{x}']_i^2\right\}$. From a practical point of view, both $\mathbf{T}_n$ and $\mathbf{T}_x$ should satisfy the following three requirements: (i) they should decorrelate their corresponding signals to a large extent; (ii) they should facilitate the estimation of the local variances of their corresponding transform components; and (iii) they should have fast realization algorithms. These three requirements may sometimes happen to be in conflict, and one has to make a compromise according to practical constraints.

In the following, our selections of $\mathbf{T}_n$ and $\mathbf{T}_x$, and the estimation of $E\{[\mathbf{T}_n\mathbf{n}]_i^2\}$ and $E\{[\mathbf{T}_x\mathbf{x}']_i^2\}$ are presented. A deblocking algorithm is then proposed.

A. Selection of $\mathbf{T}_n$ and estimation of $E\{[\mathbf{T}_n\mathbf{n}]_i^2\}$

The noise $\mathbf{n} = \mathbf{x} - \mathbf{y}$ originates in the transform domain when $\mathbf{X}$ undergoes a quantization into $\mathbf{Y}$. The

Table 1: PSNR IMPROVEMENTS OF THE IMAGES RECONSTRUCTED WITH VARIOUS ALGORITHMS.

JPEG Encoded Image	bpp	PSNR	PSNR Improvement		
			LPF [1]	POCS [3]	Proposed
Baboon	0.449	21.244	-0.026	0.085	0.190
Cameraman	0.315	26.442	-0.288	0.119	0.405
Peppers	0.323	27.686	0.374	0.361	0.655
House	0.244	30.512	0.404	0.432	0.767
Lenna	0.318	27.879	0.447	0.389	0.734
Girl	0.231	30.502	0.633	0.486	0.822
Germany	0.238	29.287	0.487	0.357	0.645
Couple	0.233	30.597	0.433	0.479	0.649
Sailboat	0.374	25.693	0.222	0.287	0.491
Tiffany	0.232	29.328	0.477	0.398	0.715
Face	0.282	30.171	0.922	0.680	1.073
Hat	0.352	29.399	1.164	0.779	1.394

error introduced is $\mathbf{N} = \mathbf{X} - \mathbf{Y} = \mathbf{T}_B(\mathbf{x} - \mathbf{y})$. Since transform coefficients are basically uncorrelated and each coefficient X_i is quantized separately, the signal $\mathbf{N}$ is highly uncorrelated. Moreover, statistics about $\mathbf{N}$ can be easily obtained with the quantizers on hand. For example, if each X_i is quantized by a uniform quantizer of step-size q_i, the variance of N_i is $q_i^2/12$ by assuming a uniform probability density function. Since $\mathbf{T}_B$ can satisfy our requirements of $\mathbf{T}_n$, we set $\mathbf{T}_n$ to be $\mathbf{T}_B$, and $E\left\{[\mathbf{T}_n\mathbf{n}]_i^2\right\}$ is then given as $q_i^2/12$.

B. Selection of $\mathbf{T}_x$ for image decorrelation

Theoretically, the optimal choice of $\mathbf{T}_x$ is the Karhunen-Loeve transform (KLT). However, from a practical point of view, applying KLT is not recommended as performing KLT is computationally very expensive. Appropriate alternatives are those transforms which have fast realization algorithms and yet can decorrelate images to a reasonable extent. In transform image coding, block-transform $\mathbf{T}_B$ is in fact a fast unitary transform for image decorrelation. Therefore, one very reasonable selection of $\mathbf{T}_x$ is $\mathbf{T}_B$.

Particular advantage can be gained when $\mathbf{T}_x$ is set to be $\mathbf{T}_B$. The formulation for $\hat{\mathbf{x}}$, when $\mathbf{T}_x = \mathbf{T}_n = \mathbf{T}_B$, can be greatly simplified as

$$\hat{\mathbf{x}} = \check{\mathbf{x}} + \mathbf{T}_B^t \mathbf{W} \mathbf{T}_B(\mathbf{y} - \check{\mathbf{x}}), \tag{9}$$

where $\mathbf{W} = [\mathbf{\Lambda}_{Tx} + \mathbf{\Lambda}_{Tx}]^{-1}\mathbf{\Lambda}_{Tx}$ and it is diagonal.

C. Estimation of local statistics about image

The *a priori* mean $\check{\mathbf{x}}$ is set to be the local spatial mean obtained from $\mathbf{x}$, which is given by

$$\bar{\mathbf{x}} = \frac{1}{(2L+1)^2} \sum_{m=-L}^{L} \sum_{n=-L}^{L} \mathbf{x}^{<m,n>}, \tag{10}$$

where $\mathbf{x}^{<m,n>}$ denotes the shift version of $\mathbf{x}$ obtained by shifting all its elements m steps up and n steps right in the spatial domain.

The local variance $E\left\{[\mathbf{T}_x\mathbf{x}']_i^2\right\}$ is set to be $\sigma_{X_i}^2$ given by

$$\sigma_{X_i}^2 = \frac{1}{(2L+1)^2} \sum_{m=-L}^{L} \sum_{n=-L}^{L} \left[\mathbf{T}_x(\mathbf{x}^{<m,n>} - \bar{\mathbf{x}})\right]_i^2. \tag{11}$$

In above, $(2L+1)^2$ is the extent of the analysis window. Since $\mathbf{x}$ is unavailable in practical case, both $\bar{\mathbf{x}}$ and $\sigma_{X_i}^2$ are estimated from $\mathbf{y}$ as follows

$$\bar{\mathbf{x}} = \bar{\mathbf{y}} \tag{12}$$

$$\sigma_{X_i}^2 = \sigma_{Y_i}^2 - \frac{q_i^2}{12}. \tag{13}$$

D. Incorporation of solution bound

By substituting the above estimates of $E\left\{[\mathbf{T}_n\mathbf{n}]_i^2\right\}$, $E\left\{[\mathbf{T}_x\mathbf{x}']_i^2\right\}$ and $\check{\mathbf{x}}$ into (9), we have

$$\hat{\mathbf{x}} = \bar{\mathbf{y}} + \mathbf{T}_B^t \mathbf{W} \mathbf{T}_B(\mathbf{y} - \bar{\mathbf{y}}), \tag{14}$$

where w_i, the i-th diagonal element of $\mathbf{W}$, is given as $(\sigma_{Y_i}^2 - q_i^2/12)/\sigma_{Y_i}^2$. Finally, in order to confine $\hat{\mathbf{x}}$ into the quantization constraint X_c, we have

$$|\hat{X}_i - Y_i| < \frac{q_i}{2}, \tag{15}$$

where $\hat{X}_i = [\mathbf{T}_B\hat{\mathbf{x}}]_i$. Combining (14) and (15) yields

$$w_i > 1 - \frac{q_i/2}{|[\mathbf{T}_B(\mathbf{y} - \bar{\mathbf{y}})]_i|} \tag{16}$$

We therefore have

$$w_i = \max\left\{\frac{\sigma_{Y_i}^2 - q_i^2/12}{\sigma_{Y_i}^2}, 1 - \frac{q_i/2}{|[\mathbf{T}_B(\mathbf{y} - \bar{\mathbf{y}})]_i|}\right\}. \quad (17)$$

5 Experimental results

Experiments were carried out to evaluate the performance of the proposed algorithm. In order to show its robustness, we used a number of *de facto* standard 256 gray-level test images of size 256×256 each. At first the test images were encoded with the block DCT based JPEG compression algorithm, where the quantization table used is the same as that in [3]. The proposed algorithm was then applied to reconstruct the encoded data. For comparative study, two typical deblocking techniques were also applied in reconstruction. They are the lowpass filtering (LPF) [1] and the projection onto convex sets (POCS) algorithms [3]. The former is an enhancement technique while the latter is a restoration one. Images reconstructed with these algorithms were then compared with each other. Table 1 shows their PSNR improvements, with respect to the JPEG-encoded images. These figures reveal that the proposed algorithm outperforms the two other algorithms in terms of objective PSNR improvement. It was also found that the proposed algorithm can provide reconstructed images of subjectively better quality. Figure 1 shows a magnified portion of the JPEG-encoded 'Lenna' at 0.318 bpp, and Figure 2 shows the image reconstructed by the proposed algorithm.

6 Conclusions

In this paper, a new approach for restoring block-transform coded images was proposed. Practical issues in the realization of the proposed restoration approach were addressed and a deblocking algorithm was then devised. The proposed algorithm can reconstruct an objectively and subjectively better image at a lower computational effort, as compared with the conventional methods [1, 3]. This was achieved by the efficient incorporation of a solution bound into the restoration, and the effective exploitation of statistics about the quantizers as well as some local properties about images.

References

[1] H. C. Reeves and J. S. Lim, "Reduction of blocking effects in image coding," *Opt. Eng.*, vol. 23, pp. 34-37, Jan./Feb. 1984.

[2] Z. Xiong, M. T. Orchard and Y.-Q. Zhang, "A deblocking algorithm for JPEG compressed images using overcomplete wavelet representations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 433-437, Apr. 1997.

[3] Y. Yang, N. P. Galatsanos and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 3, pp. 421-432, Dec. 1993.

[4] T. P. O'Rourke and R. L. Stevenson, "Improved image decompression for reduced transform coding artifacts," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 490-499, Dec. 1995.

[5] G. Demoment, "Image reconstruction and restoration: overview of common estimation structures and problems," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, pp. 2024-2036, Dec. 1989.

Figure 1: Normal JPEG reconstruction

Figure 2: Reconstruction by the proposed method