AN ADAPTIVE CONSTRAINED LEAST SQUARE APPROACH FOR REMOVING BLOCKING EFFECT

Sung-Wai Hong — *enhswai@hkpu01.polyu.edu.hk*
Yuk-Hee Chan — *enyhchan@hkpucc.polyu.edu.hk*
Wan-Chi Siu — *enwcsiu@hkpucc.polyu.edu.hk*

Dept. of Electronic Engineering, The Hong Kong Polytechnic University, Hong Kong

ABSTRACT

This paper presents a new adaptive objective function based on the regularized iterative block reduction technique for low-bit rate transform coded images. Also, a better initial estimate for the regularization approach is presented. Two types of prior knowledge are used: the first type bounds the maximum tolerable error (roughness), and the second type restricts the high-frequency content (smoothness) of the restorated images. Computer simulations showed that the new adaptive objective function with the proposed initial estimate performed better on both subjective and objective measures than did a previously proposed objective function.

1. INTRODUCTION

Using block-based transform coding (e.g. JPEG[1]) at low bit rates for image compression will result in "blocking effect." The blocking effect will lead to the perception of visible discontinuities between adjacent blocks. It is generally considered to be the most disturbing artifact in the reconstructed images.

Various approaches [2, 3, 4] to solving the blocking effect have been proposed in the past, but some of them suffer from high computational complexity or may have additional data overhead. Recently, Yang, Galatsanos and Katsaggelos proposed the use of regularization technique to tackle the blocking effect [5], which has created a new research direction. The new objective function presented in this paper is based on this recent proposal.

2. OVERVIEW OF THE CONSTRAINED LEAST SQUARE (CLS) METHOD

Image restoration [6] aims at recovering the best estimate from the degraded image. The degradation can be mathematically modeled as

$$y = Bf + n, \tag{1}$$

where vectors y and f (lexicographically ordered with size $N^2 \times 1$) are the degraded and the original images, respectively. The matrix B is the linear distortion operator, and vector n denotes additive Gaussian noise that is uncorrelated with the original image.

Regularization [7, 8] is an effective approach for converting the ill-posed problem to a well-posed one. It is simply by incorporating prior knowledge about the original image to constrain the possible sets of solutions. The regularized solution is found by minimizing the following objective function

$$J = \alpha||Sf||^2 + ||y - Bf||^2, \tag{2}$$

where S is the regularizing operator, which is generally a high-pass filter used to reduce the amount of noise (usually in the form of high-frequency) in the restorated image. $||\bullet||$ represents the Euclidean norm. Let ϵ_1^2 and ϵ_2^2 be the bounds for $||y-Bf||^2$ and $||Sf||^2$, respectively, i.e., $||y-Bf||^2 \leq \epsilon_1^2$, and $||Sf||^2 \leq \epsilon_2^2$. The former is the bound for the amount of error that can be tolerated, while the latter imposes a smoothness upperbound on the whole image. The ratio α ($= \frac{\epsilon_1^2}{\epsilon_2^2}$) is the regularization parameter that controls the degree of smoothness of the result.

Hence, the principle of regularization is to find the best estimate that can compromise these two constraints. A solution to the problem can be obtained by minimizing the objective function (2).

Iterative method, which has a number of advantages [9] over the direct inverse method, is commonly used to get the final solution.

3. FORMULATING THE OBJECTIVE FUNCTION

In our case, the point spread function (PSF) B represents a process that consists of the block DCT (BDCT) compression, quantization and decompression operations, and, y is the reconstructed blocky image that we have.

The objective function proposed in [5] is

$$J_{\alpha_1} = \alpha_1||Sf||^2 + ||f - y||^2. \tag{3}$$

For spatially varying images, this objective function provides a suboptimal solution only. Obviously, in an

This work was supported by The Hong Kong Polytechnic University Grant A/C No. 350.255.A3.420

0-7803-2127-8/94 $4.00 ©1995 IEEE

image, the region with low spatial activity (smooth region) and the region with high spatial activity (non-smooth region) should not be subjected to the same restoration conditions. This is based on the psychophysical considerations that

- Generally speaking, restored images with sharp edges will be perceived by human to have a better quality, and that
- Noise presence in non-smooth region is less disturbing to human than that in smooth region.

That means, in smooth region, noise suppression is more essential. The opposite of the above statement holds for non-smooth region. In order to have a general and spatially varying masking function, local mean $(\overline{y(i,j)})$ and local variance $(\sigma^2(i,j))$ at different image coordination (i,j) is proposed to be used, for the reason that it does not require prior knowledge of edge orientation or efforts for edge searching and classification. The local mean and local variance computed in a neighbourhood centered at (i,j) are as follows :

$$\overline{y(i,j)} = \frac{1}{(2W+1)(2H+1)} \sum_{m=i-W}^{i+W} \sum_{n=j-H}^{j+H} y(m,n). \tag{4}$$

$$\sigma^2(i,j) = \frac{1}{(2W+1)(2H+1)} \sum_{m=i-W}^{i+W} \sum_{n=j-H}^{j+H} (y(m,n) - \overline{y(i,j)})^2 \quad \text{where } 1 < i,j < N. \tag{5}$$

Here, $(2W+1)\times(2H+1)$ is the window size and the window should be symmetrical about the point (i,j).

Based on the idea we have mentioned, we propose our modified adaptive version as

$$J_{\alpha_2} = \alpha_2 ||LSf||^2 + ||R(f-y)||^2, \tag{6}$$

where R and L are both diagonal weighted matrices with size $N^2 \times N^2$. Specifically, R is defined as

$$R(m,m) = \frac{\sigma^2(i,j)}{\max\limits_{i,j=1\ldots N}(\sigma^2(i,j))} \quad \text{where } 1 \le i,j \le N, \quad \text{and } m = (i-1)\times N + j, \tag{7}$$

which is actually a normalized local variance matrix and L is defined as

$$L(i,i) = 1 - R(i,i) \quad \text{where } 1 \le i \le N^2. \tag{8}$$

In non-smooth region, each weighted coefficient of R will be assigned a relatively larger value than those in smooth region. Hence, contrast recovery will have an overwhelming effect over smoothing in non-smooth region. However, considerable amplification of high-frequency components may occur in non-smooth region. But from the psychophysical point of view, noise visibility in this region is not apparently increased. While in smooth region, R will have smaller coefficients to prevent noise amplification.

Besides, according to eqn.(8), L coefficients would have smaller magnitude in non-smooth region than in smooth region, so that excessive smoothing can be avoided and the sharpness of edges in the non-smooth region can be preserved. On the contrary, in smooth region, L coefficients should have larger values to constrain the high-frequency content and hence minimize the amount of noise.

Minimizing the objective function (6) with respect to f gives

$$\frac{\partial J_{\alpha_2}}{\partial f} = 2\alpha_2 S^t L^t LSf + 2R^t R(f-y) = 0, \tag{9}$$

where ^{t} is the matrix transpose operator. Eqn.(9) can now be written as

$$(R^t R + \alpha_2 S^t L^t LS)f = R^t Ry. \tag{10}$$

Eqn.(10) is then evaluated with the iterative equation:

$$f_{k+1} = f_k + \beta_2(R^t Ry - (R^t R + \alpha_2 S^t L^t LS)f_k), \tag{11}$$

where β_2 (the relaxation parameter) is a scalar, which has to be chosen to insure the convergence of the iterations. The range for β_2 that insures the convergence is given as

$$0 < \beta_2 < \frac{2}{||R^t R + \alpha_2 S^t L^t LS||}. \tag{12}$$

The iterations f_k will converge to an unique estimate of the original image.

4. NEW INITIAL ESTIMATE

In common practice, the initial estimate (f_0: f_k when $k=0$) can be chosen as a scaled version of the observed image y:

$$f_0 = \xi y, \quad \text{where } 0 \le \xi \le 1. \tag{13}$$

However, with such initial estimates, the admissible output images' objective and subjective quality are restricted. Better initial estimate which already have minimized block discontinuities should be used instead.

Based on the smoothness assumption, each block should be smoothly connected to its neighbouring blocks at the boundaries. Hence, we propose to get a better initial estimate by minimizing the following block-wise objective function first:

$$||Pf||^2 = \{||T_{m,n} - B_{m,n-1}||^2 + ||B_{m,n} - T_{m,n+1}||^2 + ||L_{m,n} - R_{m-1,n}||^2 + ||R_{m,n} - L_{m+1,n}||^2\}, \tag{14}$$

for $1 < m, n < \frac{N}{p}$ only, with each block of $p \times p$ pixels. The above objective function (eqn.(14)) confines the smoothness of block boundaries. In this formulation, P is a difference operator of the block boundaries. $T_{m,n}$, $B_{m,n}$, $L_{m,n}$ and $R_{m,n}$ are the 1-dimensional boundary vectors, while $B_{m,n-1}$, $T_{m,n+1}$, $R_{m-1,n}$ and $L_{m+1,n}$ are the 4 neighbouring block boundary vectors surrounding block i. Figure 1 shows their relationship.

By minimizing the objective function (14) with respect to f, we have

$$f_{k+1} = f_k - \gamma P f_k, \qquad (15)$$

and a better estimate can be obtained. Here γ (≤ 1) is a scalar constant that controls the amount of influence of the objective function on the output estimate. For typical applications, 3 iterations are enough for getting a good initial estimate for subsequent adaptive CLS process.

5. SIMULATION RESULTS

In our computer simulation, a set of 256 gray-level digital images of size 256×256 pixels are divided into 8x8 blocks and transform-coded with JPEG scheme to generate blocky images. The transform coefficients are quantized to 0.24 bit/pixel (bpp) with a uniform quantizer using the quantization table shown in Table 1. The blocky images are than restored with different approaches for comparison. The test images are further divided into 3 groups for analysis. The grouping are based on the amount of variation of the local spatial characteristics among blocks. Images in groups 1 and 3 have the least and the largest amount of variation respectively.

A 3×3 Laplacian filter is used as the regularization operator S. The regularization parameter is chosen to satisfy the bound: $\alpha = (\frac{\epsilon_1}{\epsilon_2})^2$. The PPSNR is used as an objective criterion of merit. $\frac{\|f_k - f_{k-1}\|^2}{\|f_k\|^2} \leq 4 \times 10^{-8}$ was used as the terminating criterion of the iterative process. The PPSNR is defined as

$$10 \, log_{10} \frac{(255)^2}{\sum_{i=1}^{256} \sum_{j=1}^{256} (g_{i,j} - x_{i,j})^2} \quad \text{dB}, \qquad (16)$$

where $g_{i,j}$ and $x_{i,j}$ are the (i,j)th pixels of the original and the processed images respectively.

For simplicity, we use CLS1 [5] and CLS2 as the abbreviations for the iterative methods derived from eqns.(3) and (6) respectively.

Table 2 lists the PPSNR improvements of a number of restored images after having been processed with CLS1 and CLS2, as well as the number of iterations required to converge to those PPSNRs. On the average, the objective improvement of CLS2 is approximately 4 times over that of CLS1 in terms of dB. Note both CLS1 and CLS2 have the maximum improvement for the 'Hat' image, which are about 0.2 dB and 0.9 dB respectively.

Figure 2 shows the magnified portion of a JPEG encoded 'Lena' ($PPSNR = 28.3573dB$). Figures 3 and 4 are the magnified portion of 'Lena' after having been processed by CLS1 and CLS2 respectively.

The findings show that CLS2 can further improve the image qualities on both objective and subjective measures over CLS1, especially for those images with highly uneven distribution of local spatial characteristics among blocks.

6. CONCLUSIONS

In this paper, we presented a new adaptive objective function for the constrained least square regularization approach and a new method for getting a better initial estimate to remove the blocking artifact. Findings reveal that the proposed objective function performed better on both objective and subjective measures.

7. REFERENCES

[1] W. B. Pennebaker and J. L. Mitchell, *"JPEG still image data compression standard"*, Van Nostrand Reinhold, New York, 1993.

[2] H. C. Reeve and J. S. Lim, *"Reduction of blocking effect in image coding"*, Optical Engineering, Vol. 23, No. 1, pp. 34-37, Jan./Feb. 1984.

[3] H. S. Malvar and D. H. Staelin, *"The LOT: transform coding without blocking effects"*, IEEE Trans. on Acoustics, Speech and Signal Processing, Vol. 37, No. 4, pp. 553-559, Apr. 1989.

[4] A. Zakhor, *"Iterative procedures for reduction of blocking effects in transform image coding"*, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 2, No. 1, pp. 91-95, Mar. 1992.

[5] Y. Yang, N. P. Galatsanos and A. K. Katsaggelos, *"Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images"*, IEEE Trans. on Circuits and Systems for Video Technology, Vol. 3, No. 6, pp. 421-432, Dec. 1993.

[6] H. C. Andrews and B. R. Hunt, *"Digital Image Restoration"*, Prentice Hall, New York, 1977.

[7] K. Stewart and T. S. Durrani, *"Constrained signal reconstruction - a unified approach"*, EURASIP Signal Processing III, pp. 1423-1426, 1986.

[8] A. K. Katsaggelos and R. M. Mersereau, *"A regularized iterative image restoration algorithm"*, IEEE Trans. on Signal Processing, Vol. 39, No. 4, pp. 914-929, Apr. 1991.

[9] A. K. Katsaggelos *"Iterative image restoration algorithms"*, Optical Engineering, Vol. 28, No. 7, pp. 735-748, July 1989.

Table 1: The quantization table used in JPEG encoding.

50	60	70	70	90	120	255	255
60	60	70	96	130	255	255	255
70	70	80	120	200	255	255	255
70	96	120	145	255	255	255	255
90	130	200	255	255	255	255	255
120	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255
255	255	255	255	255	255	255	255

Table 2: The experimental results for the test images.

Gp.	JPEG encoded image	ΔPPSNR (# of iterations)	
		CLS1 [5]	CLS2
1	Baboon	0.084dB(13)	0.139dB(26)
	Cameraman	0.096dB(11)	0.336dB(27)
	Sailboat	0.101dB(14)	0.338dB(21)
Average of group 1		0.09dB	0.27dB
2	F16	0.122dB(12)	0.423dB(19)
	Tiffany	0.098dB(12)	0.426dB(10)
	Germany	0.112dB(13)	0.447dB(28)
	Couple	0.122dB(13)	0.491dB(32)
	Peppers	0.121dB(13)	0.532dB(20)
	Lena	0.136dB(13)	0.533dB(22)
	House	0.149dB(12)	0.622dB(14)
	Girl	0.149dB(13)	0.634dB(21)
Average of group 2		0.13dB	0.51dB
3	Face	0.168dB(13)	0.708dB(5)
	Hat	0.210dB(14)	0.902dB(10)
Average of group 3		0.19dB	0.80dB

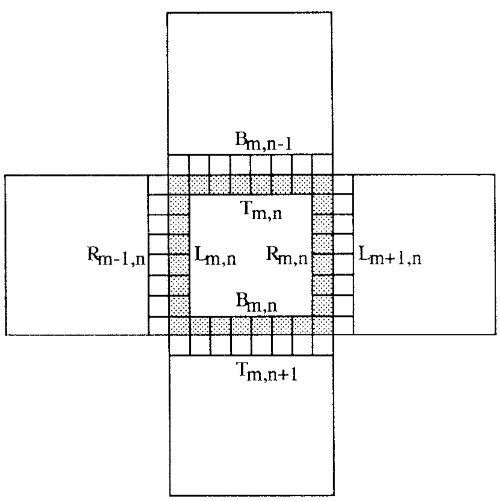

Figure 1: Shows the relationship of $T_{m,n}$, $B_{m,n}$, $L_{m,n}$, $R_{m,n}$ and $T_{m,n+1}$, $B_{m,n-1}$, $L_{m+1,n}$, $R_{m-1,n}$.

Figure 2: JPEG encoded 'Lena'.

Figure 3: CLS1 processed 'Lena'.

Figure 4: CLS2 processed 'Lena'.