OPTIMAL WAVELET THRESHOLDING FOR VARIOUS CODING SCHEMES

Dong Wei *C. S. Burrus*
Department of Electrical and Computer Engineering
Rice University, Houston, TX 77251–1892, USA

ABSTRACT

We propose a nonlinear, universal method based on wavelet thresholding to efficiently improve the performance of various coding schemes. Coarse quantization of the transform coefficients often results in some undesirable artifacts, such as ringing effect, contouring effect and blocking effect, especially at very low bit rate. We perform the wavelet-domain thresholding on the decompressed image to attenuate the quantization noise effect while maintaining the relatively sharp features (e.g. edges) of the original image. Both the objective quality and the subjective quality of the reconstructed image are significantly improved with the reduction of coding artifacts. Experimental results show that de-noising using the undecimated discrete wavelet transform (DWT) achieves better performance than using the orthonormal DWT, with an acceptable computational complexity ($O(MN\log_2(MN))$ for an image of size $M \times N$).

Also Technical Report Rice University, CML TR95-13

1. INTRODUCTION

In the last decade, transform/subband coding and fractal coding have been demonstrated to be the efficient techniques for lossy image data compression. In order to achieve high compression ratios, the transform coefficients are coarsely quantized, which often result in some undesirable artifacts associated with the basis functions of the transform, such as the ringing effect and the contouring effect in the wavelet-transform/subband-coding compressed images, and the blocking effect in the JPEG/DCT compressed images, especially at very low bit rates. Similarly, there are also some annoying artifacts (e.g., blocking effect) in the fractal compressed images. In general, a decompressed image can be viewed as a noisy observation of the original image. Therefore, the task of post-processing or enhancing the decompressed image, which can be characterized as a typical statistical estimation problem, is then to extract the original image from the noisy observation of the form "signal plus additive noise". The reconstruction noise in the decompressed images is generally signal-dependent and spatially correlated, which makes the estimation problem very difficult. Both the dependence on the original image and the spatial correlation are affected by the various compression schemes.

*This work was supported in part by Texas ATP, BNR, and ARPA.

Many methods have been developed to deal with this image enhancement problem [1, 2, 3, 4, 5]. However, most of these methods have some of the following limitations: (*i*) lack of the ability to handle more than one type of coding artifacts, i.e., dependence on the coding scheme; (*ii*) lack of the ability to improve both the objective quality and the subjective quality; (*iii*) high computational complexity. To reduce the blocking effect in the JPEG/DCT algorithms, Gopinath *et al* [6] applied soft-thresholding ("shrink or kill") to the orthonormal DWT coefficients of the decompressed image, and obtained significant improvement in the visual quality. In this paper, we show that this method can also reduce other coding artifacts and improve the objective quality of the decompressed image. However, it has been argued [7] that de-noising with the orthonormal DWT sometimes exhibits visual artifacts such as the "pseudo-Gibbs phenomena" in the neighborhood of discontinuities due to the lack of translation invariance of the wavelet basis.

In this paper, we propose a more efficient method based on the "second-generation de-noising" technique [7, 8] to significantly improve both the objective quality and the subjective quality in the processed images and to avoid the "new" artifacts resulted from de-noising. We perform hard-thresholding ("keep or kill") on the undecimated DWT coefficients of the decompressed image to suppress the reconstruction noise due to quantization. Our method outperforms most of the above methods in terms of the objective measures if the numerical comparison is possible. The performance of our method is even better than that of the method in [6] in terms of both the peak signal-to-noise ratio (PSNR) and the reduction of coding artifacts. Another advantage of our method is that it is independent of coding schemes in the sense that it can attenuate several types of coding artifacts. For an image of size $M \times N$, the computational complexity of our method is $O(MN\log_2(MN))$.

The paper is organized as follows. In the next section we give a short review of the wavelet-domain thresholding method for noise reduction. We discuss the experimental results in section 3 and give the conclusions in section 4.

2. DE-NOISING BY WAVELET THRESHOLDING

A nonlinear method was proposed for reconstructing an unknown signal from noisy data [9]. The method attempts to reject noise by damping or thresholding in the orthogonal wavelet domain and has been proved to work well in many applications.

Suppose we wish to recover an unknown signal $\mathbf{x}$ from noisy data $\mathbf{y}$, $y_i = x_i + \sigma e_i$, $i = 0, 1, \ldots, n-1$, where $e_i \stackrel{iid}{\sim} \mathcal{N}(0,1)$ is a white Gaussian noise, and σ is the noise level. Let $\hat{\mathbf{x}}$ be the estimate of $\mathbf{x}$. Our goal is to optimize the mean-squared error $\frac{1}{n}E[\| \hat{\mathbf{x}} - \mathbf{x} \|_2^2]$. The simple wavelet-domain thresholding method has three steps:

1. Compute the orthonormal DWT of the noisy data $\mathbf{y}$, obtaining the wavelet coefficients;
2. Apply the soft-thresholding nonlinearity (shrinkage)

$$\eta_t(v) = \begin{cases} v - t & \text{for } v > t \\ 0 & \text{for } -t \le v \le t \\ v + t & \text{for } v < -t \end{cases}$$

   or the hard-thresholding nonlinearity

$$\eta_t(v) = \begin{cases} v & \text{for } |v| > t \\ 0 & \text{for } |v| \le t \end{cases}$$

   to the wavelet coefficients (except the coarsest level) with a specially-chosen threshold $t = t_n = \sqrt{2\log(n)}\sigma$;
3. Perform the inverse orthonormal DWT on the thresholded wavelet coefficients, recovering the estimate $\hat{\mathbf{x}}^*$.

The universal $\sqrt{2\log(n)}\sigma$ threshold was designed for the purpose of suppressing noise-induced spikes which spoil the smoothness of reconstructions. However, if one wants to only to measure performance by mean-squared error (MSE), then lower thresholds are better. One important qualitative feature of this method is that the relatively sharp features in $\mathbf{x}$ (e.g. edges) are maintained while the noise is suppressed [9].

It has been shown that the soft-thresholding is the MSE optimal nonlinear function to apply in the orthonormal wavelet domain if one requires the reconstructed signal to be at least as smooth as the original, noise-free one. The hard-thresholding, on the other hand, yields better estimate in the MSE sense but does not guarantee the smoothness property cited above. In fact, the de-noised signal using the hard-thresholding sometimes exhibits somewhat greater spurious oscillations in the vicinity of discontinuities than that using the soft-thresholding.

Though the above assumption on the additive noise is not valid in our applications, i.e., the reconstruction error in the decompressed image is generally neither white nor independently, identically distributed (i.i.d.), the wavelet-thresholding method can still be applied successfully to reduce the colored reconstruction noise [6].

De-Noising with the orthonormal DWT sometimes exhibits visual artifacts, such as the pseudo-Gibbs phenomena (alternating undershoot and overshoot of a specific target level) in the neighborhood of discontinuities, due to the lack of translation invariance of the wavelet basis [7]. These artifacts are related in some way to the precise alignments between features in the signal and features of basis elements; signal exhibiting similar features but with slightly different alignment in time or scale might generate fewer of the artifacts. One approach to correct unfortunate misalignments between features in the signal and features in a basis is to forcibly shift signals so that their features change positions which will overcome the mis-alignments, and to unshift the signal after analysis. However, when a signal contains several discontinuities, these may interfere with each other: the best shift for one discontinuity in the signal may also be the worst shift for another discontinuity. One method to overcome this difficulty is to average out the translation dependence [7]. For the range of all circulant shifts, one shifts the data, de-noises the shifted data, and then unshifts the de-noised data. Doing this for each of a range of shifts, and averaging the several results so obtained, produces a reconstruction subject to far weaker pseudo-Gibbs phenomena than the thresholding-based de-noising using the orthonormal DWT. For the data of size N, computation of the DWT of all circulant shifts can be accomplished by the undecimated (or translation-invariant) DWT in order of $N\log_2(N)$ time [7, 8].

Our experiments indicated that in general the hard-thresholding outperforms the soft-thresholding for noise reduction in terms of both the quantitative measures and the visual quality if the undecimated DWT is used. We attribute this to the fact that the translation-invariant approach will damp the spurious oscillations introduced by the hard-thresholding while maintaining a smaller MSE.

3. EXPERIMENTAL RESULTS

In our experiments, we choose four 256-gray-scale images: Lenna (512×512), Mandrill (480×480), Camera-Man (256×256) and Building (256×256). We denote the method using the orthonormal DWT and the soft-thresholding, and the method using the undecimated DWT and the hard-thresholding as "method 1" and "method 2", respectively. We apply both methods to post-process the images compressed by the JPEG algorithm, a subband coding (SBC) algorithm, and a fractal coding algorithm at different bit-rates.

In the wavelet-thresholding schemes, there are several parameters that needs to be determined: (*i*) the type of wavelets, (*ii*) the length of wavelet filters, and (*iii*) the number of wavelet-decomposition levels. Since there is no special advantage to use any particular wavelet for de-noising, we simply choose Daubechies' orthonormal wavelets [10]. While our experiments indicated that the length of the wavelet filter had little effect on the de-noising performance in terms of both the PSNR and the subjective measure, we also found that, for larger images, de-noising with longer filters would lead to slightly better results. Therefore, we use Daubechies' 12-tap wavelet filter for the "Lenna"image and the "Mandrill" image, and use Daubechies' 8-tap wavelet filter for the other two smaller images. Similarly, we found that the number of wavelet-decomposition levels had little effect on the de-noising performance if the number is larger than 3. Thus, we simply perform 5-level wavelet-decomposition on the "Lenna" image and the "Mandrill" image, and 4-level wavelet-decomposition on the other two images, to keep the subband images of the lowest resolution roughly the same size for all four test images.

Table 1~3 illustrate the de-noising performance of the two methods in terms of the PSNR for three types of decompressed images. We find that both methods can significantly improve the objective quality of the processed images, and the "method 2" outperforms the "method 1" in most cases. For the fractal-compressed image, we also compare our two methods with the post-processing method by Y. Fisher [11] and we find that the "method 2" is somewhat better than Fisher's method. Figure 1∼4 illustrate an example of the improvement in the visual quality of the JPEG-compressed "Lenna" image. Though both Figure 3 and Figure 4 have better visual quality than Figure 2, the blocking effect is removed more completely in Figure 3 than in Figure 4. More perceptual results can be found in [12].

4. CONCLUSIONS

A powerful, universal and relatively low-complexity method for improving the quality of the decompressed images resulting from various coding schemes has been presented. It is based on thresholding in the wavelet domain. In particular, using the undecimated DWT and the hard-thresholding to avoid the new perceptual artifacts, results in significant improvement in terms of both the PSNR and the visual quality in the sense that the various coding artifacts are greatly removed.

REFERENCES

[1] S. Minami and A. Zakhor. An optimization approach for removing blocking effects in transform coding. *IEEE Trans. on CAS for Video Tech.*, 5(2):74–82, April 1995.

[2] B. Ramamurthi and A. Gersho. Nonlinear space-variant postprocessing of block coded images. *IEEE Trans. ASSP*, 34(5):1258–1267, 1986.

[3] S.-W. Wu and A. Gersho. Improved decoder for transform coding with applications to JPEG baseline system. *IEEE Trans. Commun.*, 40(2):251–254, February 1992.

[4] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos. Regularized reconstruction to reduce blocking artifacts of blcok discrete cosine transform compressed images. *IEEE Trans. on CAS for Video Tech.*, 3(6):421–432, December 1993.

[5] A. Zakhor. Iterative procedures for reduction of blocking effects in image coding. *IEEE Trans. on CAS for Video Tech.*, 2(1):91–94, March 1992.

[6] R. A. Gopinath, M. Lang, H. Guo, and J. E. Odegard. Wavelet-based post-processing of low bit rate transform coded images. In *Proc. IEEE Int. Conf. Image Processing*, volume II, pages 913–917, Austin, TX, November 1994.

[7] R. R. Coifman and D. L. Donoho. Tranlation-invariant de-nosing. In Anestis Antoniadis, editor, *Wavelets and Statistics*. Springer-Verlag Lecture Notes, New York, NY, 1995. To appear.

[8] M. Lang, H. Guo, J. E. Odegard, C. S. Burrus, and R. O. Wells, Jr. Nonlinear processing of a shift-invariant DWT for noise reduction. In *SPIE conference on wavelet applications*, volume 2491, Orlando, FL, April 1995.

[9] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inform. Theory*, 41(3):613–627, May 1995.

[10] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, Philadelphia, PA, 1992. Notes from the 1990 CBMS-NSF Conference on Wavelets and Applications at Lowell, MA.

[11] Y. Fisher. *Fractal Image Compression: Theory and Application*. Springer-Verlag, New York, NY, 1995.

[12] D. Wei, M. Lang, H. Guo, J. E. Odegard, and C. S. Burrus. Quantization noise reduction using wavelet thresholding for various coding schemes. In *SPIE Math. Imagaging: Wavelet Applications in Signal and Image Processing*, volume 2569, San Diego, CA, July 1995.

Table 1. De-Noising performance for JPEG-coded images

test images	bpp	PSNR (dB)		
		JPEG	method 1	method 2
Lenna	0.65	35.80	36.08	36.19
	0.25	30.41	31.08	31.42
	0.18	27.33	28.30	28.57
Mandrill	1.84	25.64	25.85	25.74
	0.61	20.67	20.82	20.84
	0.35	19.01	19.24	19.29
Camera-Man	0.88	31.75	32.01	32.04
	0.34	26.44	26.69	26.84
	0.24	24.30	24.57	24.75
Building	0.92	31.33	31.48	31.46
	0.36	27.00	27.28	27.45
	0.27	24.75	25.16	25.34

Table 2. De-Noising performance for subband-coded images

test images	bpp	PSNR (dB)		
		SBC	method 1	method 2
Lenna	0.500	34.33	34.76	35.02
	0.250	31.30	31.72	32.01
	0.125	28.22	28.62	28.95
Mandrill	2.000	27.34	27.71	27.60
	1.000	22.64	23.03	23.03
	0.500	20.00	20.33	20.37
Camera-Man	1.000	32.73	33.20	33.26
	0.500	27.95	28.43	28.49
	0.250	24.65	24.98	25.09
Building	1.000	31.47	31.83	31.76
	0.500	28.78	29.14	29.12
	0.250	25.23	25.59	25.61

Table 3. De-Noising performance for the fractal-compressed "Lenna" images

bpp	PSNR (dB)			
	fractal	YF method	method 1	method 2
0.80	34.10	34.43	34.32	34.45
0.43	30.83	31.35	31.27	31.55
0.28	28.02	28.59	28.48	28.69

Figure 1. Original "Lenna" image (a part)

Figure 2. JPEG-compressed image at 0.25bpp

Figure 3. Processed image using "method 1"

Figure 4. Processed image using "method 2"