Enhancement of Decompressed Images at Low Bit Rates

R. A. Gopinath*
IBM T. J. Watson Research, #J2-J05, 30 Saw Mill River Road, Hawthorne, NY 10542

M. Lang[†] H. Guo[‡] and J. E. Odegard[§]
Department of Electrical and Computer Engineering, Rice University, Houston, TX-77251

Paper presented and published in the Proceedings of SPIE, *Mathematical Imaging: Wavelet Applications in Signal and Image Processing II (2303-30)*, July 27-29, 1994, San Diego, CA. **This is also Technical Report Rice University, CML TR94-05**

ABSTRACT

Transform coding at low bit rates introduces artifacts associated with the basis functions of the transform. For example, decompressed images based on the DCT (discrete cosine transform) - like JPEG[16] - exhibit blocking artifacts at low bit rates. This paper proposes a post-processing scheme to enhance decompressed images that is potentially applicable in several situations. In particular, the method works remarkably well in "deblocking" of DCT compressed images. The method is non-linear, computationally efficient, and spatially adaptive - and has the distinct feature that it removes artifacts while yet retaining sharp features in the images. An important implication of this result is that images coded using the JPEG standard can be efficiently post-processed to give significantly improved visual quality in the images.

Keywords: image enhancement, wavelet shrinkage, jpeg, image restoration, image compression

1 INTRODUCTION

Effective compression of still images such as the JPEG standard[16] is important for many applications which use bandwidth-limited media such as radio channels or storage media with steadily growing yet limited capacity. Using these (lossy) compression methods it is possible to achieve compression ratios of 10 and more with only a small noticeable perceptual deterioration.[16] However, for considerably higher compression ratios the perceptual quality decreases rapidly. There are essentially two undesired effects, the loss of image details and blocking. The former is a result of disregarding high frequency components during the quantization process. The latter is due to the DCT coding of 8×8 blocks in the JPEG standard.

*This research was supported by BNR

†This research was partially supported by the Alexander von Humboldt foundation and by AFOSR under grant F49620-1-0006 funded by DARPA

‡This research was partially supported by TI and by AFOSR under grant F49620-1-0006 funded by DARPA

§This research was supported by AFOSR under grant F49620-1-0006 funded by DARPA

Several approaches have been proposed to overcome these coding artifacts. Niss[13] presents a prediction scheme for the low-frequency AC coefficient from the DC coefficient changes within a 3×3 array of blocks. This method considerably reduces blocking effects in smooth areas but might introduce large artifacts elsewhere, e.g., at the edges.[16] In another approach Wu and Gersho propose a suboptimal decoder jointly minimizing the least squared reconstruction error in the actual image block as opposed to minimizing the individual errors separately.[18,19] Assuming that the encoder is given they use a table lookup for the set of indices similar to that in a vector quantizer. This approach requires a novel decoder based on the design of suitable codevectors. Wu and Gersho report an improvement of the peak SNR between 1.7% and 4.6% for selected images outside the training set using relatively small compression ratios of 5.8–12.3. Furthermore they state that "... most images ... have better perceptual quality in some areas compared to those reconstructed by the conventional decoder".[18] In particular, the mosquito noise along lines and edges is reduced. Unfortunately, they do not give sample images showing that their smaller SNR values correspond to better perceptual quality which need not necessarily be the case. It is interesting to notice that the performance of their method is sensitive to the choice of the training set.[18]

We propose a novel method based on wavelet thresholding for postprocessing of decompressed images. It is especially suited for high compression rates where the blocking effect is clearly to be seen. With our method these artifacts can be reduced while sharp features in the image are maintained. Thus, the postprocessed images show a considerably improved perceptual quality as well as smaller SNR values compared to the conventional decoder. Our new algorithm can use a conventional JPEG encoder and decoder for which VLSI chips are available. Moreover, it is computationally very efficient.

2 EXPOSITION OF THE PROBLEM

The model for the problem of image enhancement is as follows: y is a decompressed version of the image x and e is the reconstruction error:

$$y = x + e. \tag{1}$$

y has artifacts which depend both on the compression procedure P and the bit rate and is incorporated in e. One post-processes the data with a procedure to generate $\hat{x}(y)$ which "smoothes" out the effects of e on y, while yet retaining the important features in x. A measure of performance could be the root mean-squared error $\|\hat{x}(y) - x\|$.

If the error e is Gaussian and uncorrelated (both spatially and with x), then this becomes a classical statistical estimation problem. Since the procedure $\hat{x}(y)$ should work for a variety of images $x \in X$ the goal is to

$$\text{minimize } \sup_{x \in X} \|\hat{x}(y) - x\| \text{ with respect to } \hat{x}. \tag{2}$$

A procedure for which (2) holds is called minimax mean-squared error (mse) procedure. This classical problem has a solution that is asymptotically near optimal (simultaneously) for a wide variety of classes X.[3,5,4] The procedure $\hat{x}(y)$ is based on soft-thresholding in the wavelet domain.[4] Two important qualitative features of the method are a) the reconstruction is noise-free in the sense that no spurious oscillations are introduced (other than in the data y) and b) relatively sharp features in y are maintained.

In our problem it is evident that e is correlated both spatially and with x. Consequently, the stochastic assumptions on e (uncorrelated, Gaussian) are not valid. To understand how the soft-thresholding approach can still be succesfully applied we give some heuristic arguments why it works in the case of uncorrelated Gaussian noise. There the since e is uncorrelated the error signal We in the wavelet domain is uncorrelated too. W denotes the wavelet transform under consideration. On the other hand the transformed signal Wx is more concentrated in relatively few coefficients, typically at the coarser scales. There are several possible ways to utilize this fact. A special approach called soft-thresholding with threshold parameter δ given by the operator

$$T_\delta(s) = \operatorname{sgn}(s)(|s| - \delta)_+, = \begin{cases} s - \delta & \text{for } s > \delta \\ 0 & \text{for } -\delta \leq s \leq \delta \\ s + \delta & \text{for } s < -\delta \end{cases} \tag{3}$$

has the following desirable properties. If one wants to avoid spurious oscillations or, equivalently, maintain the smoothness of the signal x one has to impose the condition[4]

$$|W\hat{x}(y)| \leq |Wx|. \tag{4}$$

It can be shown[4] that wavelet soft-thresholding is 1) a nearly minimax mse procedure (with no currently known better one) and 2) there is no better estimator satisfying the smoothness condition (4). Thus, the small coefficients (we assume here that the error is relatively small) corresponding to the error signal are simply set to zero and the larger ones are shrinked towards zero. The threshold δ is related to the actual noise level which is constant according to the assumption of e being uncorrelated and is assumed to be known. The Gaussian assumption is needed only for computing an asymptotically optimal value δ depending on the variance and the data size.[4]*

3 DESCRIPTION OF THE PROCEDURE

Figure 1 schematically shows the basic steps of our method. It is mathematically given by

$$\hat{x}(y) = W^{-1}T_\delta(Wy), \tag{5}$$

where W^{-1} denotes the inverse wavelet transform and T_δ a thresholding scheme with the (possibly adaptively chosen) threshold δ. The input y of our procedure is a decompressed image resulting from a standard JPEG

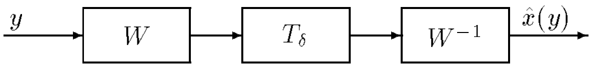

Figure 1: Blockdiagram of our image enhancement method.

codec. There are several choices to be made such as

- a suitable error measure
- the degree and the type of wavelet filter like Daubechies' filters, optimal l_2 filters, etc.
- the type of the transform W like M-band wavelets,[6] wavelet packets,[2] space-varying wavelets[1,7]
- the number of levels to use
- the best thresholding scheme
- the choice of the threshold parameter δ.

In the following we describe our approach of finding a sensible set of choices. It is obvious that an optimization over all parameters listed above is barely practicable. Thus we restrict our attention to those parameters which seem to have the greatest impact on the performance.

*Our experience has shown that the threshold should depend only on the error variance and *not* on the data size to avoid 'oversmoothing' for larger signal lengths.

Error measure: As Jayant et al. state[11] the choice of an error measure which shows high correlation to perceptual judgments is a principal problem for coding of signals at low bit rates. Although there are numerous approaches for which advances have been made in recent years, most of them are highly specific. They depend, e.g., on the actual coder type or the signal environment.[11] Moreover, the role of perception in various stages of the coding decoding scheme needs to be better understood. Consequently, we use a simple and therefore practicable two step approach. First we search for the optimum parameter set minimizing the root mean-squared error. Then we refine it using the subjective perceptual quality. This approach is further justified by the following facts: soft thresholding minimizes the l_2 error over several smoothness classes of the signal x and is known to yield perceptually good results.

Choice of wavelet: As is clear from the list above there is a vast number of possible choices for the wavelet analysis. From the standpoint of implementation the wavelet filters to be used should be small to keep the computational burden low. Furthermore, it seems that regularity of the wavelet basis functions plays an important role for removing the blocking artifacts. Thus we optimized only over a very small subset — Daubechies' filters of length 4, 6, and 8. Tests have shown that Daubechies' filter with larger lengths tend to yield oversmoothed images with considerably lower perceptual quality. We are currently investigating the effect of not maximally flat (at frequency zero) scaling filters like those in.[15,8,14,12]

Thresholding: The choice with the greatest impact on the performance of the post-processing method is, in our opinion, that of the thresholding scheme and the corresponding parameters. As we pointed out in Section 2 and is confirmed by our experiments the shrinkage condition (4) is essential to avoid artifacts introduced by the post processing scheme.† We consequently use soft-thresholding throughout all our experiments. There remains the problem of finding an optimal threshold and possibly adapt it to the local error level.

As stated before, it is relatively simple to find a uniform threshold in the case of uncorrelated (Gaussian) noise e. Threshold values $\delta = 1.5 \ldots 3\sigma$ (σ is the standard deviation of e) yield excellent results.[10,9] This choice for δ corresponds to the suppression of $86.6\% \ldots 99.7\%$ (for Gaussian noise) of all noise values. Only the noise level, i.e., the standard deviation of e remains to be determined. This is effectively done as described in the following. Let us denote with $C_{1_{HH}}y$ the operator that picks out the high/high (index HH) portion of y after the first stage (index 1) of the wavelet transform. The corresponding operators for the low/high portion, the high/low portion, and coarser scales (indices $2, 3, \ldots$) are analogously defined. As a simple and very reliable estimate for the variance of We one can use

$$\hat{\sigma}^2(y) = \frac{1}{N-1} \sum (C_{1_{HH}}y - m)^2 \tag{6}$$

where N is the number of data points and m the mean of $C_{1_{HH}}y$. This approach exploits the fact that the variance of $C_{1_{HH}}x$ is small in most images. The noise, however, is uniformly distributed over all scales (see discussion above) and consequently shows up most clearly in $C_{1_{HH}}y$.

Saito[17] gives an alternative approach for parameter selection based on an entropy related criterion called minimum description length. The appealing property of this method is that it finds a data dependent set of parameters and could thus be used for finding a proper threshold. Our experiments have shown that this method tends to yield too large threshold values which result in oversmoothed images.

As we discussed above, the assumption that the error We is Gaussian and uncorrelated does not hold for images processed by the JPEG codec at low bit rates. We have shown that the relevant assumption for the applicability of uniform thresholding is that of an uncorrelated error signal. We therefore have to consider the dependence of the noise level of e on different scales and bins (HH, HL, and LH). Moreover, it is not clear that the noise is Gaussian even at a fixed scale and level. That further complicates the determination of a sensible threshold. In the next section we investigate how these problems could be solved.

†Artifacts occur, e.g., if one uses the so-called hard-thresholding, i.e., one keeps all values above a threshold and sets all other to zero.

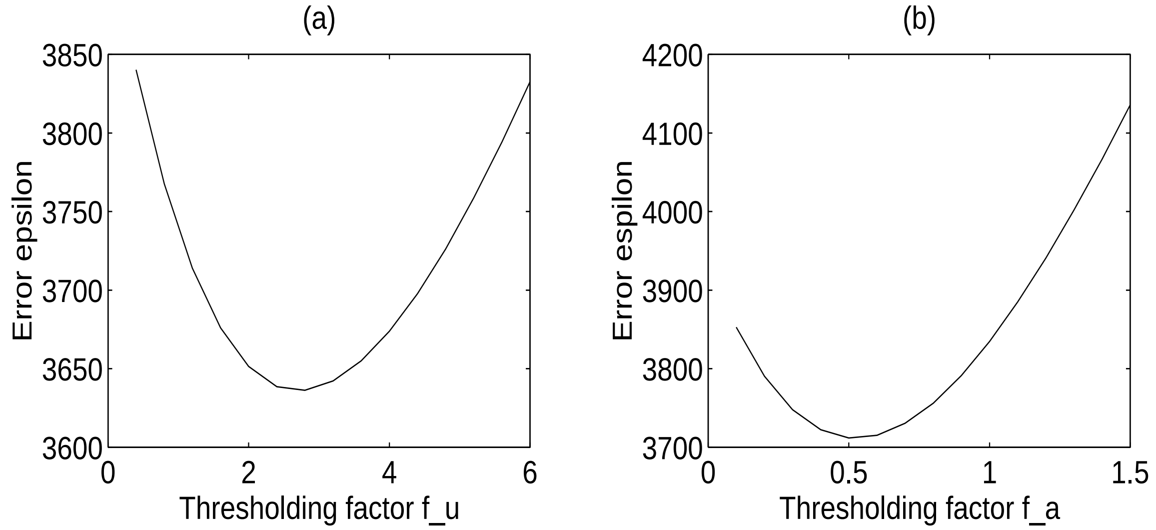

Figure 2: Dependence of ϵ on the threshold factor for (a) uniform thresholding and (b) scale adaptive thresholding.

4 EXPERIMENTS AND RESULTS

For our experiments we used four images (1=Lena 512×512, 2=Mandrill 480×480, 3=Camera 256×256, and 4=Building 256×256), three of which are depicted in the figures 4–6[‡]. For the JPEG codec we used a quality factor of 10 resulting in compression ratios of 32 (0.25bit per pixel), 13.3 (0.60bpp), 23.4 (0.34), and 22.1 (0.36bpp), respectively. The upper left and right picture in the figures 4–6 show the original and decompressed image, respectively. The blocking artifacts and the loss of details are clearly to be seen. The optimization was carried out by determining the minimum mse over Daubechies' filters of length 4, 6, 8, soft-thresholding of $C_{\cdot HH}y$, $C_{\cdot HL}y$, $C_{\cdot LH}y$ at the finest level up to the finest 5 levels and various fixed and scale adaptively chosen values δ.

Fixed Threshold: In this experiment we estimated the error variance using Eq. (6) and determined the minimum mse over the parameters as described above and values $\delta = 0.4\hat{\sigma}(y), 0.8\hat{\sigma}(y), \ldots, 6.0\hat{\sigma}(y)$. A typical plot showing the dependence of the root of the error energy

$$\epsilon = \sqrt{\sum_i e_i^2} \tag{7}$$

on the thresholding factor $f_u = \delta/\hat{\sigma}$ is shown in figure 2(a) (Lena). There is a clear minimum for a factor 2.8. Table 1 shows the set of optimal parameters when using the minimum mse criterion.

We observe that the error compared to the decompressed image can be reduced in all cases by proper soft-thresholding. The optimum number of thresholded scales L varies between 3 and 5, the optimum filter length N between 4 and 8. It turns out, however, that the choice of L and N is not critical. Thus, it is possible to choose $L = 3$ and $N = 4$ in all cases with only a minor change of ϵ and perceptual quality. It seems therefore unnecessary to optimize these parameters. This is not true for the thresholding factor f_u. If one applies, e.g., the factor f_u of the image Mandrill to the image Lena, the blocking effects are hardly removed and one gets an oversmoothed image vice versa. The images corresponding to the optimal set of parameters are depicted in the lower left corner of figures 4–6. Clearly, the blocking has been considerably reduced. At the same time only few details are lost subject to smoothing.

[‡]Notice that due to space restrictions, figure 4, 6, and 7 show only a part of the full image.

	JPEG	Uniform Thr.				Adaptive Thr.			
	$\epsilon/10^3$	$\epsilon/10^3$	L	N	f_u	$\epsilon/10^3$	L	N	f_a
Lena	3.94	3.64	5	8	2.8	3.70	3	8	0.6
Mandrill	11.3	11.2	3	8	0.8	11.1	3	8	0.3
Camera	3.11	3.02	3	4	1.6	3.02	3	4	0.5
Building	2.92	2.81	5	4	2.4	2.82	3	4	0.5

Table 1: Optimal parameters (L = number of levels, N = filter length, f = threshold factor)

Adaptive threshold: We now investigate the dependence of the noise level on the scale and try to utilize it for the thresholding scheme. Figure 3 shows a plot of the standard deviation of $C_{1_{HH}}e, \ldots, C_{5_{HH}}e$ for all four images. It increases for coarser scales with exception of the image Mandrill where it seems to be nearly constant. The behaviour for the LH and the HL portion is similar. We therefore use in a second optimization a scale adaptive threshold

$$\delta_{s_b}(e) = f_a \hat{\sigma}_{s_b}(e) = f_a \sqrt{\frac{1}{N-1} \sum (C_{s_b}e - m)^2} \tag{8}$$

where s denotes the interesting scale, b the interesting portion (HH, HL, or LH), m the corresponding mean and f_a a constant factor. The other parameters are chosen from the same range as in the previous experiment. We optimized over values $f_a = 0.1, \ldots, 1.5$. A typical plot showing $\epsilon(f_a)$ is given in figure 2(b). The optimal parameters for all four images can be found in table 1. Again the error is reduced compared to the JPEG case. However, it is slightly larger than using uniform thresholding for the images 1, 3, 4. Just as in the previous experiment the value of ϵ is not very sensitive to the choice of the parameters L and N. Thus, $L = 3$ and $N = 4$ is a convenient choice. In contrast to uniform thresholding the range of the optimal threshold factor is considerably smaller. It appears that a value $f_a = 0.5$ yields good results for all four images. Notice, however, that the variance of the true error signal at all interesting scales has to be known for this scale adaptive thresholding to work. The pictures corresponding to the optimal set of parameters are depicted in the lower right corners of the figures 4–6. Again, the blocking effects have been considerably reduced. In our opinion the perceptual quality is slightly better (less smoothing of details), albeit the larger error values ϵ.

Figures 7(a,b) give examples for further reduced blocking effects for uniform and adaptive thresholding, respectively. This is achieved by using a threshold twice as large as in the optimal case.

5 CONCLUSION AND FUTURE WORK

We have presented a powerful method for removing blocking artifacts introduced by the JPEG codec at high compression ratios. It is based on soft-thresholding in the wavelet domain. It is also very efficient since the computation of the wavelet transform requires only $O(N)$ floating point operations. Although uniform thresholding is optimal only for uncorrelated error signals it yields excellent results, assuming the optimal factor f_u is known. Potentially better results can be obtained by scale adaptive thresholding. In a first approach where we use the (in a practical application not available) knowledge of the error variance at the interesting scales we achieved minor perceptual improvements.

We are currently working on finding a data dependent way of computing the optimal factor f_u such that oversmoothing is avoided. Furthermore, we investigate how the scale adaptive thresholding can be improved and

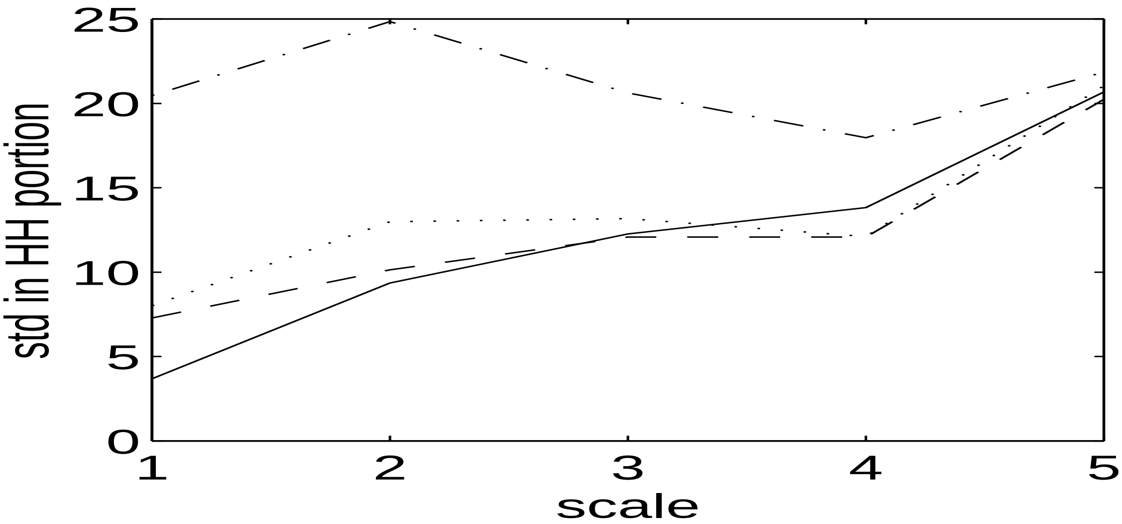

Figure 3: Standard deviation of the HH portion for different scales and images 1 (—), 2 (– · –), 3 (· · ·), and 4 (– –).

be applied without knowing the true values of the variance of We at the interesting scales. Also, we currently investigate different ways of including the knowledge about the compression ratio at hand for adapting the threshold factors and the impact of different (higher and lower) compression ratios on the performance.

In future work we will consider the usage of various wavelet analyses, the development of an improved information theoretic criterion, and the investigation of several error measures based on perceptual criteria.

Figure 4: Image Lena clockwise beginning in the upper left corner: (ul) original image; (ur) image after JPEC codec with quality 10; (ll) image after optimal thresholding (uniform); (lr) image after optimal thresholding (scale adaptively)

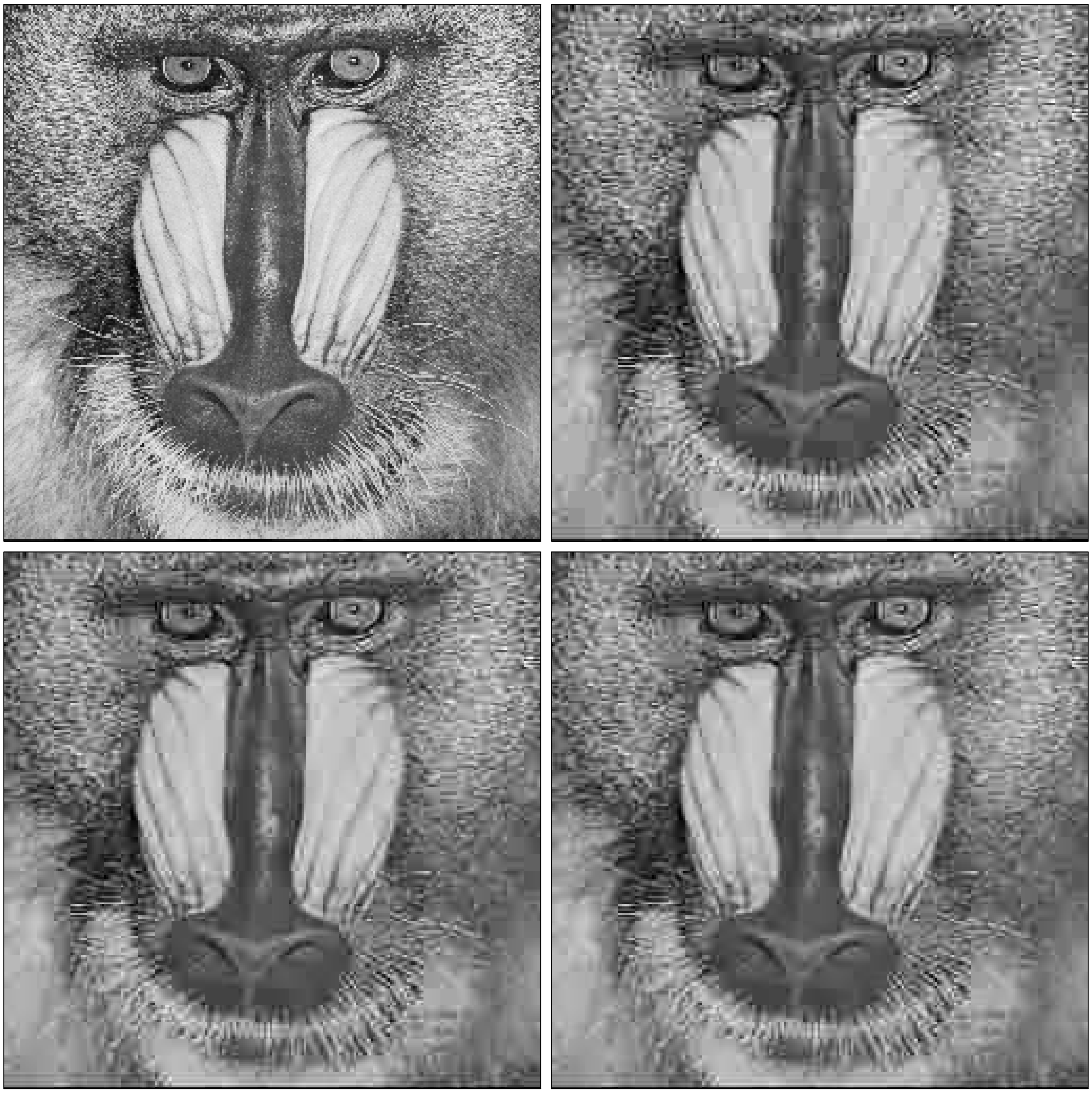

Figure 5: Image Mandrill clockwise beginning in the upper left corner: (ul) original image; (ur) image after JPEC codec with quality 10; (ll) image after optimal thresholding (uniform); (lr) image after optimal thresholding (scale adaptively)

Figure 6: Image Camera clockwise beginning in the upper left corner: (ul) original image; (ur) image after JPEC codec with quality 10; (ll) image after optimal thresholding (uniform); (lr) image after optimal thresholding (scale adaptively)

Figure 7: Image Camera with reduced blocking using larger thresholds: (left) uniform thresholding (right) scale adaptive thresholding

6 REFERENCES

[1] A. Cohen, I. Daubechis, and P. Vial. Wavelets on the interval and fast wavelet transforms. *Applied and Computational Harmonic Analysis*, 1(1):54–81, December 1993.

[2] R. R. Coifman and M. V. Wickerhauser. Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, 38(2):1713–1716, 1992.

[3] D. L. Donoho. Unconditional bases are optimal bases for data compression and for statistical estimation. *Applied and Computational Harmonic Analysis*, 1(1):100–115, December 1993.

[4] D. L. Donoho. De-noising by soft-thresholding. *IEEE Trans. Inform. Theory*, 1994. Also Tech. Report, Department of Statistics, Stanford University,1992.

[5] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika.* To appear. Also Tech. Report, Department of Statistics, Stanford University,1992.

[6] R. A. Gopinath and C. S. Burrus. Wavelets and filter banks. In C. K. Chui, editor, *Wavelets: A Tutorial in Theory and Applications*, pages 603–654. Academic Press, San Diego, CA, 1992. Also Tech. Report CML TR91-20, September 1991.

[7] R. A. Gopinath and C. S. Burrus. A tutorial overview of wavelets, filterbanks and interrelations. In *Proc. of the ISCAS*, Chicago, IL, May 1993. IEEE.

[8] R. A. Gopinath, J. E. Odegard, and C. S. Burrus. Efficient design of modulated filter banks and modulated wavelet tight frames. Technical Report CML TR92-20, Computational Mathematics Laboratory, Rice University, Houston, TX, 1992.

[9] H. Guo, J. E. Odegard, M. Lang, R. A. Gopinath, I. Selesnick, and C. S. Burrus. Speckle reduction via wavelet shrinkage with application to SAR based ATD/R. In *SPIE Math. Imagaging: Wavelet Applications in Signal and Image Processing*, volume 2303, San Diego, CA, July 1994. Also Tech report CML TR94-03, Rice University, Houston, TX.

[10] H. Guo, J. E. Odegard, M. Lang, R. A. Gopinath, I. Selesnick, and C. S. Burrus. Wavelet based speckle reduction with application to SAR based ATD/R. In *Proc. Int. Conf. Acoust., Speech, Signal Processing*, Austin, TX, November 1994. IEEE. Also Tech report CML TR94-02, Rice University, Houston, TX.

[11] N. Jayant, J. Johnston, and R. Safranek. Signal compression based on models of human perception. *Proc. IEEE*, 81(10):1385–1422, October 1993.

[12] M. Lang, I. Selesnick, J. E. Odegard, and C. S. Burrus. Constrained FIR filter design for 2-band filter banks and orthonormal wavelets. In *Sixth Digital Signal Processing Workshop*, October 1994.

[13] B. Niss. *Prediction of AC Coefficients from the DC Values.* ISO/IEC JTC1/SC2/WG8 N745, May 1988.

[14] J. E. Odegard, R. A. Gopinath, and C. S. Burrus. Design of linear phase cosine modulated filter banks for subband image compression. Technical Report CML TR94-06, Computational Mathematics Laboratory, Rice University, Houston, TX, February 1994. Submitted to ICIP 1994, Austin, TX.

[15] J. E. Odegard, M. Lang, and C. S. Burrus. Design of filter banks and wavelets using lagrange multipliers. Technical Report CML TR94-07, Computational Mathematics Laboratory, Rice University, Houston, TX 77251, 1994.

[16] W. B. Pennebaker and J. L. Mitchell. *JPEG - Still Image Data Compression Standard.* Van Nostrand Reinhold, New York, 1993.

[17] Naoki Saito. Simultaneous noise suppresion and signal compression using a library of orthonormal bases and the minimum description length criterion. In E. Foufoula-Georgiou and P. Kumar, editors, *Wavelets in Geophysics.* Academic Press, 1994.

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

[19] S-W. Wu and A. Gersho. Enhancement of transform coding by nonlinear interpolation. *Proc. SPIE Visual Commun. Image Process.*, 1605:487–498, 1991.