WAVELET-BASED POST-PROCESSING OF LOW BIT RATE TRANSFORM CODED IMAGES

*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

ABSTRACT

In this paper we propose a novel method based on wavelet thresholding for enhancement of decompressed transform coded images. Transform coding at low bit rates typically introduces artifacts associated with the basis functions of the transform. In particular, the method works remarkably well in "deblocking" of DCT compressed images. The method is nonlinear, computationally efficient, and spatially adaptive and has the distinct feature that it removes artifacts yet retain sharp features in the images. An important implication of this result is that images coded using the JPEG standard can efficiently be postprocessed to give significantly improved visual quality in the images. The algorithm can use a conventional JPEG encoder and decoder for which VLSI chips are available.

Also Technical Report Rice University, CML TR94-15

1. INTRODUCTION

By applying lossy compression methods for image compression it is possible to achieve compression ratios of 10 and more (gray scale images) with only a small noticeable perceptual deterioration [11]. For considerably higher compression ratios the perceptual image quality decreases rapidly. There are essentially two undesired effects: (*i*) loss of image details and (*ii*) blocking. The former is a result of disregarding high frequency components during the quantization process and is hence lost information. The latter is due to the DCT coding of 8×8 blocks in the JPEG standard.

Several algorithms have been developed to deal with these coding artifacts. Niss [10] proposed 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 [11]). 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 [13, 14]. 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. However, it is interesting to notice that the performance of their method is sensitive to the choice of the training set [13].

2. IMAGE ENHANCEMENT BY DENOISING

Let y be the decompressed version of the original image x. A model for the image enhancement problem is then given by $y = x + e$ where e is the reconstruction error. The artifacts in y associated with the compression procedure at the given bit rate is incorporated in e. The goal is to obtain a procedure, $\hat{x}(y)$, to generate an estimate of the original image which "smoothes" out the effects of e on y, yet retain the important features in x. A measure of performance could be the mean squared error (mse).

Assuming that the error e is Gaussian and uncorrelated (both spatially and with x), then finding x from y is 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 find $\hat{x}$ such that $\sup_{x \in X} \|\hat{x}(y) - x\|$ is minimized. If one wants to avoid spurious oscillations or, equivalently, maintain the smoothness of the signal x one has to impose the condition [3]

$$|W\hat{x}(y)| \leq |Wx| \,. \tag{1}$$

This classical problem has a solution which is asymptotically near optimal (simultaneously) for a wide variety of classes, X [3]. A procedure, $\hat{x}(y)$, satisfying the above minimax problem with the given smoothness condition is given by soft-thresholding in the wavelet domain where the threshold depends on the variance of e. Let W and W^{-1} denote the wavelet and inverse wavelet transform operators respectively. Then the nonlinear procedure is given by

$$\hat{x}(y) = W^{-1} T_\delta(Wy) \tag{2}$$

where the nonlinear soft thresholding operation is defined by

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

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

The threshold δ is typically obtained through an estimation from the observed image y. Two important qualitative features of the method are: (i) the reconstruction is noise-free in the sense that no spurious oscillations are introduced (other than in the data y) and (ii) relatively sharp features in y are maintained [3].

It has been shown [3] that wavelet soft-thresholding is (a) nearly a minimax MSE procedure (no currently known better one) and (b) there is no better estimator satisfying the smoothness condition in Eqn. 1. 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 [3].

In our problem of image enhancement it is evident that e is correlated both spatially and with x. Consequently, the stochastic assumptions on e are not valid. However, we can still successfully apply the algorithm for the following reasons (heuristic). The wavelet transform tends to whiten the data and hence while e is not uncorrelated, We might be. Secondly, Donoho [3] also shows that if the error is bounded (which is clearly the case for most signal processing problems) then soft-thresholding is optimal. Finally, we propose to apply nonuniform/adaptive soft-thresholding to compensate for data correlation.

3. ENHANCEMENT ALGORITHM

Fig. 1 schematically shows the basic steps of our algorithm for image enhancement. The input y of our procedure is a

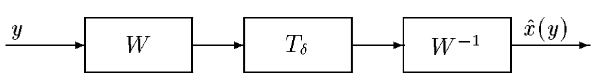

Figure 1. Algorithm for image enhancement.

decompressed image resulting from a standard JPEG codec. In attempting to implement this algorithm there are several parameters that needs to be determined: (i) a suitable error measure, (ii) the type of the transform W like M-band wavelets [4], wavelet packets [2], space-varying wavelets [1, 5], (iii) given the kind of transform W we need to determine the degree and the type of wavelet filter (i.e., Daubechies' filters, optimal l_2 filters, etc.), (iv) the number of levels to use, (v) the best thresholding scheme and (vi) the choice of the threshold parameter δ.

It is obvious that optimizing over all parameters is not feasible. Thus we restrict our attention to those parameters which experimentally seems to have the greatest impact on the performance.

Error measure: An error measure which shows high correlation to perceptual image quality is desirable but in reality nonexistent at present time. Although there are numerous approaches for which advances have been made in recent years [8], most of them are highly specific. Hence, we have chosen to use a simple two step approach. First we search for the optimum parameter set minimizing the mean squared error. We then refine the parameters by using subjective criteria such as perceptual quality.

Transform: To simplify the search we limited the analysis to 2-band wavelet bases.

Choice of wavelet: There is a number of possible choices for the wavelet analysis. From the standpoint of implementation the wavelet filters to be used should be short to keep the computational burden low and furthermore using long filters tends to smear image details. Furthermore, experimentation shows that regularity of the wavelet basis plays an important role for removing the blocking artifacts. Thus we chose to optimized over a very small subset — Daubechies' filters of length 4, 6, and 8.

Number of levels: We limited the investigation to consider only 5 level wavelet expansions. Experimentation indicates that the important scales are those which reduce the unwanted artifacts to roughly single pixels.

Thresholding scheme: The choice with the greatest impact on the performance of the postprocessing algorithm is in our opinion the thresholding scheme and the corresponding parameters. As we pointed out in Section 2 and is confirmed by our experiments the shrinkage condition in Eqn. 1 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.

Threshold: It is relatively simple to find a uniform threshold, δ, in the case of uncorrelated (Gaussian) noise e. Threshold values $\delta = 1.5 \ldots 3\sigma$ where σ is the standard deviation of e yield excellent results [7]. Only the noise level (i.e., the standard deviation of e) remains to be determined. This is effectively done as follows. Denote by C_{HH_l} the operator which when applied to y picks out the high/high (HH) portion of y on the lth level of the wavelet transform. Operators to extract low/high (LH) and high/low (HL) are analogously defined. As a simple and very reliable estimate for the variance of We one can use

$$\hat{\sigma}^2(y) = \frac{1}{N-1} \sum (C_{HH_1} y - m)^2 \qquad (4)$$

where N is the number of data points and m is the mean of $C_{HH_1} y$. This approach exploits the fact that the variance of $C_{HH_1} x$ is small in most images. The noise, however, is uniformly distributed over all scales and consequently shows up most clearly in $C_{HH_1} y$.

Saito [12] gives an alternative approach for parameter selection based on an entropy related criterion called minimum description length (MDL). 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. However, our limited experiments have shown that this method tends to give too large threshold values which result in oversmoothed images. However, a more in detail study is necessary for this method – and in particular one needs to consider in details the best basis selection proposed by Saito [12].

Nason [9] recently proposed a method for threshold selection based on cross validation. This method for selecting

*Artifacts occur if one use hard-thresholding (i.e., one keep all values above a threshold and sets all other to zero).

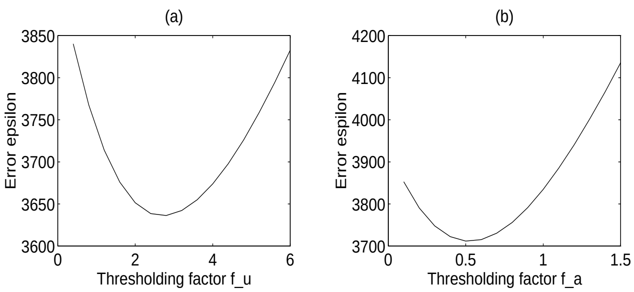

Figure 2. Dependence of ϵ on the threshold factor for (a) uniform thresholding and (b) scale adaptive thresholding.

the threshold has been shown to offer better noise rejection for signals corrupted by handle heavy tailed noise.

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, further complicating the estimation of a sensible threshold. In the next section we investigate how some of these problems could be solved.

4. EXPERIMENTS AND RESULTS

We used four classical gray scale images 1) Lenna 512×512, 2) Mandrill 480×480, 3) Camera-man 256×256, and 4) Building 256×256). For the JPEG codec we used a quality factor of 10 resulting in compression ratios of 32 (0.25bpp), 13.3 (0.60bpp), 23.4 (0.34), and 22.1 (0.36bpp), respectively. The optimization was carried out by determining the minimum mse over Daubechies' filters of length 4, 6, 8, soft-thresholding of $C_{HH_{(\cdot)}}y$, $C_{HL_{(\cdot)}}y$, $C_{LH_{(\cdot)}}y$ at the finest 3 to 5 levels. Each of these experiments was then carried out for both a uniform threshold δ and a scale adaptive threshold δ.

Uniform Threshold: In this experiment we estimated the error variance using Eqn. 4. We then determined the minimum mse over the parameters as described above and $\delta = f_u \hat{\sigma}$, $f_u \in \{0.4, 0.8, \ldots, 6.0\}$. Fig. 2a shown a typical plot of the dependence between the the error energy

$$\epsilon = \sqrt{\sum_i e_i^2} \qquad (5)$$

and the thresholding factor $f_u = \delta/\hat{\sigma}$. There is a clear minimum for a factor $f_u = 2.8$. Table 1 shows the set of optimal parameters when using this minimum mse criterion for each of the test images.

From Table 1 we observe that the mean squared error compared to the decompressed image is 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

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

	JPEG	Uniform Thr.				Adaptive Thr.			
	$\epsilon/10^3$	$\epsilon/10^3$	L	N	f_u	$\epsilon/10^3$	L	N	f_a
Lenna	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

unnecessary to optimize over 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 Lenna, the blocking effects are hardly removed and vice versa one gets an oversmoothed image. Due to space limitations we can only include one set of images in this paper and we have chosen to use a part of the Lenna image since it tends to best illustrate most of the properties of the algorithm. For a more complete set of images see Gopinath et al. [6]. The image corresponding to the Fig. 4. Clearly, the blocking has been considerably reduced while at the same time only few details are lost subject to smoothing.

Adaptive threshold: Next consider the dependence of the noise level on the scale. As previously mentioned the error for this type of images are clearly correlated – both spatially and with the original image. To better deal with correlated noise we altered the thresholding scheme and applied a scale adaptive threshold. The intuition for scale adaptive thresholding is that we should threshold "harder" on the scales where the artifacts are reduced to single pixels. To further quantify why a scale adaptive thresholding scheme might work consider Fig. 3. Fig. 3 shows a plot of the standard deviation of $C_{HH_1}e, \ldots, C_{HH_5}e$ for all four images. Notice that the standard deviation increases for

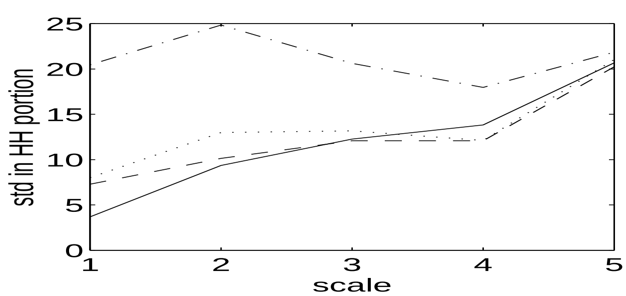

Figure 3. Standard deviation of the HH portion for different scales and images Lenna (—), Mandrill (– · –), Camera-man (· · ·), and Building (– –).

coarser scales with exception of the image Mandrill where it seems to be nearly constant. The behavior for the LH and HL are similar. Hence, we applied the following a scale adaptive threshold

$$\delta_{b_s}(e) = f_a \hat{\sigma}_{b_s}(e) = f_a \sqrt{\frac{1}{N-1} \sum (C_{b_s} e - m)^2} \qquad (6)$$

where s denotes the interesting scale (1,2,...,5)), b the interesting portion (HH, HL, or LH), m the corresponding mean and f_a a constant factor to be optimized. The other parameters are chosen from the same range as in the previous experiment. We optimized over values $f_a \in \{0.1, \ldots, 1.5\}$.

A typical plot showing $\epsilon(f_a)$ is given in Fig. 2b. The optimal parameters for all four images can be found in Table 1. Again the mean squared error is reduced compared to the JPEG case. However, it is slightly larger than using uniform thresholding for 3 of the images (Lenna, Camera-man and Building). 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 was used for this scale adaptive thresholding. The image corresponding to the optimal set of parameters are depicted in the lower right corners of the Fig. 4. Again, the blocking effects have been considerably reduced and in our opinion the perceptual quality is slightly better (less smoothing of details) compared to uniform thresholding, albeit the larger error values ϵ.

Blocking can further be reduced for both the fixed thresholding scheme and the adaptive thresholding scheme without loss of significant image detail. By using subjective evaluation a perceptually better images can be obtained by changing the optimal threshold factor. Since we currently have not performed any extensive study of perceptual acceptability of these manually optimized images and the improvement on the test images at the given bit rate are marginal we have not included any samples [6].

5. CONCLUSION AND FUTURE WORK

A powerful method for removing blocking artifacts introduced by the JPEG codec at high compression ratios have presented. 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. Further improvement can be obtained by scale adaptive thresholding. However, while the uniform thresholding algorithm can be applied to the decoded image without knowledge of the original image, the scale adaptive algorithm require that the JPEG encoder appends a few bits for encoding information related to the standard deviation of the error image e at various scales.

We are currently working on finding a data dependent way of computing the optimal factors f_u/f_a such that oversmoothing is avoided. Furthermore, we are investigating how scale adaptive thresholding can be improved and be applied without knowing the true values of the variance of We at the interesting scales. We are also currently investigating how encoded information in the JPEG data stream can be applied intelligently for improving the performance.

In future work we will consider the usage of various wavelet analyses, continue the investigation of the method of minimal description length [12], investigate methods for threshold selection by cross validation [9] and investigate several error measures based on perceptual criteria.

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. De-noising by soft-thresholding. *IEEE Trans. Inform. Theory*, 1994. Also Tech. Report, Department of Statistics, Stanford University,1992.

[4] 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.

[5] R. A. Gopinath and C. S. Burrus. Factorization approach to unitary time-varying filter banks. *IEEE Trans. SP*, submitted May 1993. Also Tech Report No. CML TR92-23, December 1992.

[6] R. A. Gopinath, M. Lang, H. Guo, and J. E. Odegard. Enhancement of decompressed images at low bit rates. In *SPIE Math. Imagaging: Wavelet Applications in Signal and Image Processing*, volume 2303, San Diego, CA, July 1994. Also Tech report CML TR94-05, Rice University, Houston, TX.

[7] 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.

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

[9] G. P. Nason. Wavelet regression by cross-validation. Technical report, Department of Mathematics, University of Bristol, Bristol, U.K., March 1994. Also Tech. Report.

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

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

[12] 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.

[13] 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.

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

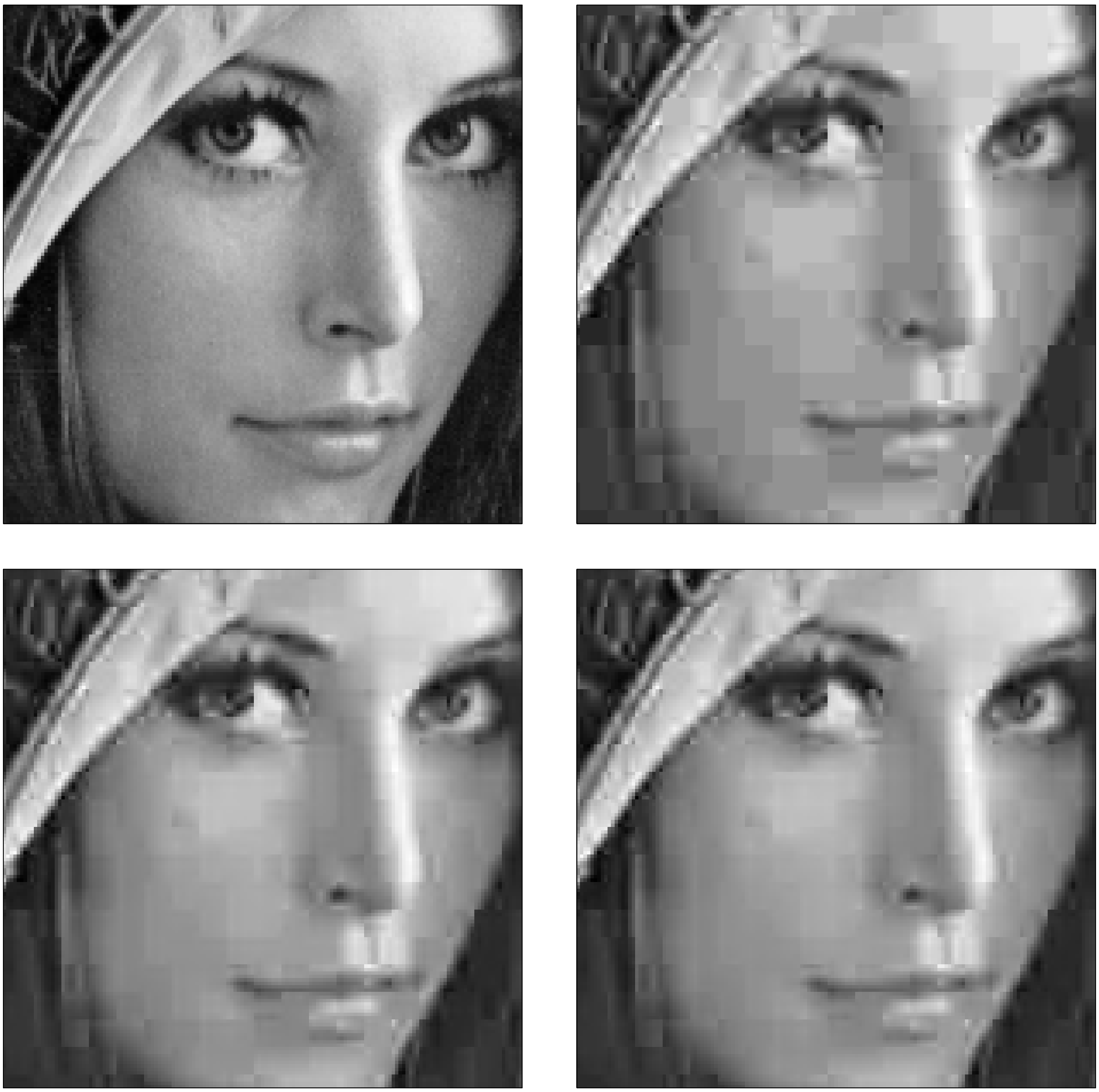

Figure 4. Image Lenna beginning in the top left corner: (tl) original image; (tr) image after JPEC codec with quality 10; (bl) image after optimal thresholding (uniform); (br) image after optimal thresholding (scale adaptive)