Embedded Coring in MPEG Video Compression

Peter M. B. van Roosmalen, Reginald L. Lagendijk, *Senior Member, IEEE*, and Jan Biemond, *Fellow, IEEE*

***Abstract*—Coring is a well-known technique for removing noise from images. The mechanism of coring consists of transforming a noise-degraded image into a frequency-domain representation, followed by a reduction of the image transform coefficients by a (nonlinear) coring function. After inverse transforming the cored coefficients, the noise-reduced image is obtained. In this paper, we show that coring can be embedded into MPEG encoders with relatively little additional complexity. We exploit statistical properties of the DCT coefficients to find the optimal Bayesian coring function for each of the DCT coefficients. Experimental results show the effectiveness of the MPEG-embedded coring function on compressing noisy image sequences.**

***Index Terms*—Bayesian optimization, coring, DCT coefficient quantization, MPEG, noise filtering, noise reduction, video compression, wavelets.**

I. Introduction

REMOVING noise from digital image sequences is not only important for improving the visual quality of these sequences, but also for increasing the efficiency of MPEG video encoders. A property of (white) noise is that the noise energy spreads out evenly over all the DCT coefficients in an MPEG encoder. Therefore, the presence of noise in an image sequence leads to fewer DCT coefficients that are zero, decreasing the efficiency of zero-runlength encoding of quantized DCT coefficients. Furthermore, on the average, the amplitudes of the remaining DCT coefficients are larger than in the noise-free case, which also leads to a loss of compression efficiency.

In recent years, the technique of *coring* has gained great popularity for denoising degraded signals, and in particular for denoising 2-D images [1]–[4]. Coring is a (nonlinear) noise reduction technique in which each frequency component of an observed noisy signal is adjusted according to a certain characteristic: the *coring function*. Originally, coring was developed as a heuristic technique for crisping television pictures [5]. In the early 1980s, coring was first applied in the digital domain for noise reduction [3], [4], [6]. The technique of coring received a great deal of attention after Donoho and Johnstone applied it successfully in the wavelet transform domain in 1994 [1], [2]. The transformation to the wavelet domain is especially useful to separate locally concentrated signal energy and the (white) noise contributions because, unlike the Fourier transform, scale-space representations allow local signal characteristics at different scales to be taken into account. In more recent work, coring has also been applied to video sequences, for instance, in the 3-D wavelet domain [7], [8].

Manuscript received October 19, 2000; revised December 14, 2001. This paper was recommended by Associate Editor S. Chen.

P. M. B. van Roosmalen was with the Information and Communication Theory Group, Department of Mediamatics, Faculty of Information Technology and Systems, Delft University of Technology, 2628 CD Delft, The Netherlands.

R. L. Lagendijk and J. Biemond are with the Information and Communication Theory Group, Department of Mediamatics, Faculty of Information Technology and Systems, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail: R.L.Lagendijk@its.tudelft.nl; J.Biemond@its.tudelft.nl).

Publisher Item Identifier S 1051-8215(02)02812-4.

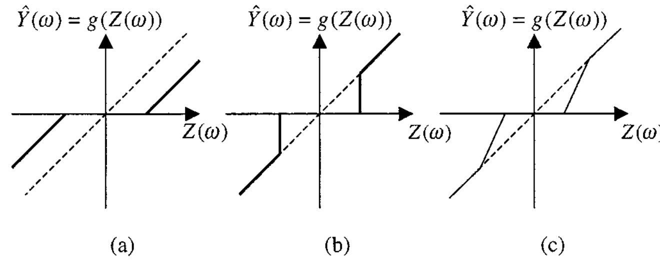


Fig. 1. Coring functions $g(\,.\,)$ where $Z(\omega)$ is the noisy signal (wavelet, DCT, Fourier) coefficient. (a) Soft thresholding. (b) Hard thresholding. (c) Piecewise linear coring.

In general, coring functions leave transform coefficients with high amplitudes unaltered, and coefficients with low amplitudes are shrunk toward zero. Coefficients with large amplitudes are reliable and should not be altered because they are not significantly influenced by noise. Coefficients with small amplitudes carry relatively little information and are easily influenced by noise. Therefore, these coefficients are unreliable, and their contribution to the observed data should be reduced.

Following these intuitive guidelines, three well-known coring or *thresholding* functions have been proposed (see Fig. 1).

1) *Soft-thresholding* [1], [3], illustrated in Fig. 1(a)

$$\hat{Y}(\boldsymbol{\omega}) = \begin{cases} \mathrm{sgn}((Z(\boldsymbol{\omega}))\,(|Z(\boldsymbol{\omega})| - T)), & \text{if } |Z(\boldsymbol{\omega})| > T \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Here, $Z(\boldsymbol{\omega})$ is one of the (Fourier, wavelet, DCT) transform coefficients of the noisy signal $z(\boldsymbol{i})$, where $\boldsymbol{i}$ and $\boldsymbol{\omega}$ represent 2-D spatial coordinates and 2-D frequency coordinates, respectively. Further, $\hat{Y}(\boldsymbol{\omega})$ is the cored (noise-reduced) transform coefficient, and T is a coring threshold that depends on the amount of the noise in the signal $z(\boldsymbol{i})$.

2) *Hard-thresholding* [1], [3], illustrated in Fig. 1(b)

$$\hat{Y}(\boldsymbol{\omega}) = \begin{cases} Z(\boldsymbol{\omega}), & \text{if } |Z(\boldsymbol{\omega})| > T \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

3) *Piecewise linear coring* [3], illustrated in Fig. 1(c)

$$\hat{Y}(\boldsymbol{\omega}) = \begin{cases} Z(\boldsymbol{\omega}), & \text{if } |Z(\boldsymbol{\omega})| > T_1 \\ \mathrm{sgn}(Z(\boldsymbol{\omega}))\dfrac{|Z(\boldsymbol{\omega})| - T_0}{T_1 - T_0}\,T_1, & \text{if } T_0 \le |Z(\boldsymbol{\omega})| \le T_1 \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

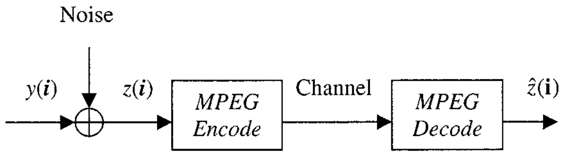


Fig. 2. MPEG encoding of noisy image sequences.

Here, T_0 and T_1 are noise-dependent thresholds. Piecewise linear coring is a compromise between hard and soft thresholding. It attempts to reduce the ringing artifacts from hard thresholding, while avoiding the loss of low-contrast picture detail common to soft-thresholding.

In the case of video sequences, noise reduction is often carried out as a preprocessing step prior to MPEG compression. If a coring technique is to be used, the noise-reduction preprocessor must carry out a forward and an inverse signal transform, such as a wavelet or DCT transform. This is rather inefficient since internally in the MPEG compression algorithm a similar (DCT) transform is carried out. Embedding the noise reduction step into the MPEG compression algorithm may therefore greatly reduce the complexity and processing delay of the overall system.

In this paper, we consider the embedding of the coring functionality into the MPEG video-encoding algorithm. Furthermore, instead of using one of the three coring functions shown in Fig. 1, we will derive Bayesian optimal coring functions for the DCT coefficients. In Section II, we first describe the problem of coring DCT coefficients, and describe the formal solution to this problem. Then, in Section III, we consider how to apply coring to the three different picture types that exist within MPEG, namely I, P, and B frames. In Section IV, we find the optimal Bayesian coring functions for the DCT coefficients using a model for the probability density function of the DCT coefficients. Then, in Section V, we evaluate the proposed MPEG encoder with embedded coring. Conclusions are drawn in Section VI.

II. Coring and Quantization of DCT Coefficients

We consider a broadcasting environment in which noisy image sequences are digitally broadcasted after passing through an MPEG compression system. Fig. 2 establishes the notational conventions. MPEG compression systems try to minimize the quantization errors between input frame $z(\boldsymbol{i})$ and output frame $\hat{z}(\mathbf{i})$ for a given bit rate. However, in the case of noisy video sequences, what they should be doing is minimizing the quantization errors between the original noise-free frame $y(\boldsymbol{i})$ and the compressed frame $\hat{z}(\mathbf{i})$. When the encoder does so, it can be considered a device for simultaneous noise reduction and video compression.

Let $\varepsilon(\boldsymbol{i})$ denote the overall error between $y(\boldsymbol{i})$ and $\hat{z}(\boldsymbol{i})$. This error consists of the sum of the quantization noise resulting from the compression of the DCT coefficients and the (reduced) amount of noise present in $z(\boldsymbol{i})$. The variance of this overall error can be expressed in terms of original and quantized DCT coefficients

$$E[\varepsilon^2(\boldsymbol{i})]=E\left[(y(\boldsymbol{i})-\hat{z}(\boldsymbol{i}))^2\right]=\sum_{k=1}^{64} E\left[\left(Y_k(\boldsymbol{n})-\hat{Z}_k(\boldsymbol{n})\right)^2\right]. \tag{4}$$

Here, $Y_k(\boldsymbol{n})$ and $\hat{Z}_k(\boldsymbol{n})$ (with $k=1,\ldots,64$) represent the 64 original and quantized DCT coefficients, respectively, of each 8×8 DCT block within a video frame. The column, row, and frame numbering of the DCT blocks is indicated in shorthand by $\boldsymbol{n}$. Clearly, an optimal *joint noise-reduction and quantization strategy* minimizes (4). Since we assume that the DCT coefficients are mutually uncorrelated, we confine ourselves to *scalar* estimation of $\hat{Z}_k(\boldsymbol{n})$ in the remainder of this paper.

Two approaches can be followed to minimize equation (4), although it was already shown in [9] that these approaches yield the same result. The first approach directly minimizes $E[(y(\boldsymbol{i})-\hat{z}(\boldsymbol{i}))^2]$ for each of the DCT coefficients individually. Essentially, this approach determines optimal (nonuniform) quantizers for the DCT coefficients of a *noisy* signal. With proper knowledge of the signal and noise distributions, it is not difficult to find the relations that the quantizer representation and decision levels should satisfy. In fact, the resulting expressions are highly akin to the ones found for Lloyd–Max quantizers [8], [10].

The second approach splits the minimization of the overall error variance (4) into two successive minimizations. The *first step* is the noise reduction step, which for a moment ignores the subsequent quantization of the DCT coefficients. This step, therefore, consists of the minimization of

$$E[\varepsilon^2(\boldsymbol{i})]=E\left[(y(\boldsymbol{i})-\hat{y}(\boldsymbol{i}))^2\right]=\sum_{k=1}^{64} E\left[\left(Y_k(\boldsymbol{n})-\hat{Y}_k(\boldsymbol{n})\right)^2\right]. \tag{5}$$

If the minimization of (5) is done per DCT coefficient (i.e., we derive a scalar estimator), this leads to the following well-known estimation result:

$$\hat{Y}_k(\boldsymbol{n})=E[Y_k(\boldsymbol{n})|Z_k(\boldsymbol{n})]. \tag{6}$$

The noise-reduced DCT coefficients are found as the conditional expectation of the true DCT coefficient, given the observed noisy DCT coefficient. Note that (6) assumes that the probability density function of the observed noisy DCT coefficients is known. The *second step* now simply consists of designing a scalar quantizer that is optimal for the original noise-free DCT coefficients. Notice that in this step, we can ignore that the DCT coefficients to be quantized are in fact noise-reduced coefficients.

In the particular case of an MPEG encoder, the advantage of the second approach is that the quantizer in the MPEG encoder is already optimized for encoding noise-free DCT coefficients. It is, therefore, not necessary to redesign new quantization tables as is required for the first approach. In the second approach, all that needs to be done is to core the DCT coefficients prior to the usual DCT coefficient quantization. In fact, MPEG encoders implicitly core noisy DCT coefficients to some extent by incorporating a dead zone in the quantizers for the DCT coefficients of the nonintra-coded frames [11], [12]. As a result of the dead zone, DCT coefficients with small magnitudes are mapped to zero. We note, however, that the use of dead zones is suboptimal for noise reduction because quantizers with dead zones are not applied to all frames, and because they do not address the noise on DCT coefficients with larger amplitudes.

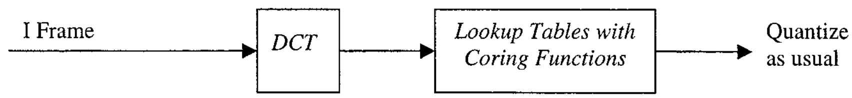


Fig. 3. Coring of the DCT coefficients of I frames in an MPEG-2 encoder.

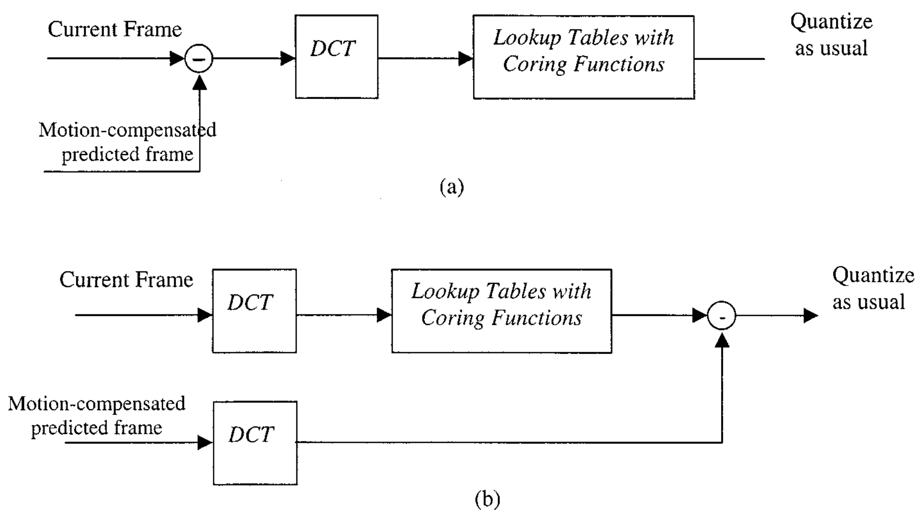


Fig. 4. (a) Coring function applied to DCT of frame differences. (b) Alternative configuration in which B and P frames can be cored similar to I frames. Note that the predicted frame is extracted from a coded frame that has already been noise reduced and need not be cored again.

III. Coring of the MPEG I, P, and B Frames

In this section, we first consider the case of coring MPEG I frames. This solution is then extended to MPEG P and B frames.

A. Coring of I Frames

MPEG I frames are compressed by dividing the frames in 8×8 blocks, applying the DCT transform to each of the blocks, and quantizing the DCT coefficients. Two approaches can be followed toward coring the DCT coefficients of I frames. The first approach is to estimate the probability density function (PDF) for each of the DCT coefficient from the observed data for each frame. Given the estimated PDFs, the conditional expectation of each DCT coefficient can be computed according to (6). Finally, the noise reduction can take place by replacing the observed DCT coefficients with their conditional expectations. Computing the optimal coring functions for *each I frame* of a video sequence is, however, difficult to realize in real-time MPEG encoders.

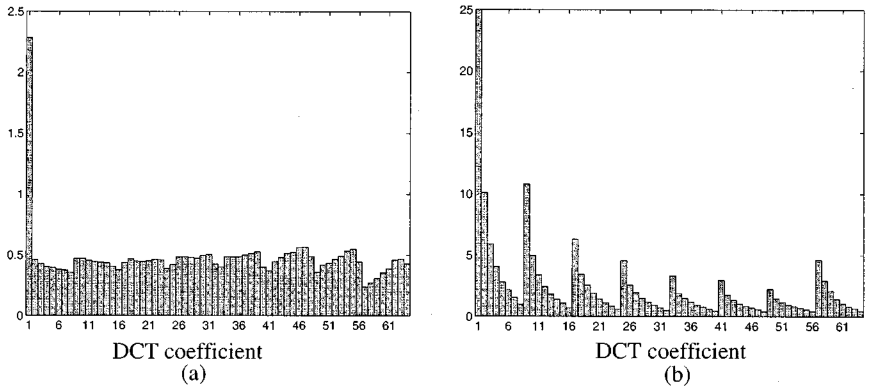
The second approach does not optimize the coring functions for each frame. Instead, fixed sets of coring functions are computed and stored in the MPEG encoder as lookup tables (Fig. 3). These coring functions are computed *off-line* from a large set of frames, so that on average, the encoder gives the best results that can possibly be achieved under the condition of static lookup tables. A different set of coring functions is computed for a number of noise levels. Upon MPEG compression of a video sequence, only the noise level has to be estimated (possibly per frame), after which the proper coring function lookup table can be selected. In Section IV, we will discuss how the DCT coring functions for MPEG I-frames can be determined.

B. Coring of B and P Frames

MPEG B and P frames are (bi-directionally) predicted from frames coded previously. The frame differences between the predicted and current frames are encoded like I frames, i.e., by applying the DCT transform and scalar quantization. This process is illustrated in Fig. 4(a). Finding the optimal coring coefficients is, however, more difficult for B and P frames because the PDFs of the frame differences and the noise are much harder to find than in the case of I-frames. This is because the PDFs now depend on the nonlinear coring and quantization of the reference frames, and because the PDFs are highly dependent on the quality of the motion estimator used for the prediction of the B and P frames.

We, therefore, propose using the alternative strategy shown in Fig. 4(b). Instead of coring the DCT coefficients of frame differences, the DCT transform and coring are performed *prior* to subtracting the current frame from its (motion-compensated) prediction. Note that the coring functions in Fig. 4(a) and (b) are different because they are applied on signals with entirely different statistical properties. Furthermore, the results obtained by the two different approaches are generally not identical.

Two aspects of the scheme in Fig. 4(b) are noteworthy. First, the *predicted* frames have already been compressed and hence they have already been noise-reduced earlier. Therefore, it is not necessary to core these predicted frames again. Second, the optimal coring characteristics are identical to those computed



1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16
17	18	19	20	21	22	23	24
25	26	27	28	29	30	31	32
33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48
49	50	51	52	53	54	55	56
57	58	59	60	61	62	63	64

(c)

Fig. 5. (a) Shape parameters and (b) standard deviations estimated for the DCT coefficient. Numbering of DCT coefficient is according to (c).

before for the I frames. Therefore, only one set of lookup tables is required for the I, P, and B frames.

IV. Bayesian Coring Functions for Noisy DCT Coefficients

This sections deals with computing the coring functions for the I, P, and B frames. As indicated in the previous section, the coring functions are computed from a large set of video sequences, so that the encoder gives the best results that can be achieved on average with static lookup tables. Computing the coring functions consists of two steps. First, the PDFs of the intensities of the frames and the noise have to be determined. Next, the coring functions can be computed using (6).

The noise corrupting the DCT coefficients of the video sequences is assumed to be additive white zero-mean Gaussian noise with known variance σ_W^2. Therefore, the noise variance does not depend on the index of the DCT coefficient. As the amount of noise may be time-variant, its variance has to be estimated per frame.

The PDF of each of the 63 DCT ac coefficients is modeled by a Laplacian distribution [13], [14]. In practice, the *generalized* Gaussian PDF is more accurate [15], [16], which includes the Laplacian as a special case. The DCT dc coefficient is not cored because it is usually insignificantly influenced by noise.

The generalized Gaussian probability density function is defined as follows:

$$f_X(x) = a \exp(-b|x|^c) \tag{7}$$

with

$$a = \frac{bc}{2\Gamma(1/c)} \quad \text{and} \quad b = \frac{1}{\sigma}\sqrt{\frac{\Gamma(3/c)}{\Gamma(1/c)}}. \tag{8}$$

Here, $\Gamma(\,.\,)$ is the gamma function and σ is the standard deviation of the DCT coefficient under consideration. It can be seen from (7) and (8) that the generalized Gaussian PDF is completely determined by the shape parameter c and the variance σ^2. An efficient method for estimating the shape parameter c from a set of image sequences based on second-order statistics is given in [16]. Let Y_k denote DCT coefficients with coefficient number $k = 1, \ldots, 64$. The mean μ_k and the variance σ_k^2 of a set of observed DCT coefficients with coefficient number k can be estimated directly from the observed data. Let ρ_k be defined as follows:

$$\rho_k = \frac{\sigma_k^2}{E^2[|Y_k - \mu_k|]}. \tag{9}$$

The shape parameter c_k for the PDF of the kth DCT coefficient is then found by solving

$$\frac{\Gamma(1/c_k)\Gamma(3/c_k)}{\Gamma^2(2/c_k)} = \rho_k. \tag{10}$$

Equation (10) can be solved using numerical methods. Fig. 5 shows the resulting shape parameter c_k and the standard deviation σ_k that are estimated from the DCT coefficients obtained from a set of 18 different video frames [numbering of the DCT coefficients is according to Fig. 5(c)]. Except for the DCT DC component, it can be seen that c_k is in the range of 0.4–0.5. The standard deviation of the DCT coefficients decreases with increasing frequency, which is consistent with the well-known fact that natural images contain less energy in high frequencies than in low frequencies.

We can now proceed to calculate the coring functions using (6). To this end, we calculate the coring function $g(Z_k(\boldsymbol{n}))$ as the Bayesian estimator given in (6) as follows:

$$\begin{aligned} g(Z_k(\boldsymbol{n})) &= \hat{Y}_k(\boldsymbol{n}) \\ &= E[Y_k(\boldsymbol{n})|Z_k(\boldsymbol{n})] \\ &= \int_{-\infty}^{\infty} y_k(\boldsymbol{n}) f_{Y_k(\boldsymbol{n})|Z_k(\boldsymbol{n})}(y_k(\boldsymbol{n}))\, dy_k(\boldsymbol{n}). \end{aligned} \tag{11}$$

In order to solve (11), we need the *a posteriori* PDF $f_{Y_k(\boldsymbol{n})|Z_k(\boldsymbol{n})}(y_k(\boldsymbol{n}))$, i.e., the *a posteriori* PDF of the kth DCT coefficient $Y_k(\boldsymbol{n})$ given the observed noisy DCT coefficient $Z_k(\boldsymbol{n})$. Using Bayes' rule and the fact that the noise $W_k(\boldsymbol{n})$ on the kth DCT coefficient is additive, we obtain

$$\begin{aligned} f_{Y_k(\boldsymbol{n})|Z_k(\boldsymbol{n})}(y_k(\boldsymbol{n})) &= \frac{f_{Z_k(\boldsymbol{n})|Y_k(\boldsymbol{n})}(z_k(\boldsymbol{n}))f_{Y_k(\boldsymbol{n})}(y_k(\boldsymbol{n}))}{f_{Z_k(\boldsymbol{n})}(z_k(\boldsymbol{n}))} \\ &= \frac{f_{W_k}(z_k(\boldsymbol{n}) - y_k(\boldsymbol{n}) f_{Y_k(\boldsymbol{n})}(y_k(\boldsymbol{n})))}{f_{Z_k(\boldsymbol{n})}(z_k(\boldsymbol{n}))} \end{aligned} \tag{12}$$

where $f_{W_k}(w_k(\boldsymbol{n}))$ is the PDF of the noise on the kth DCT coefficient. We have assumed that this PDF is Gaussian with

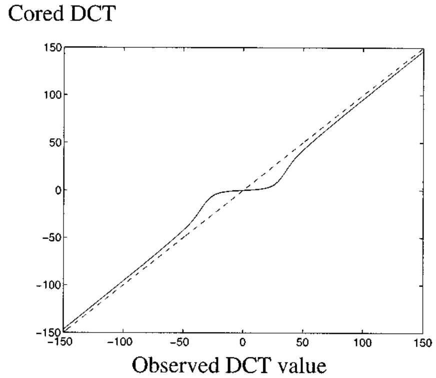


Fig. 6. Coring function for DCT coefficient 8, computed for noise with variance 100 corrupting the image.

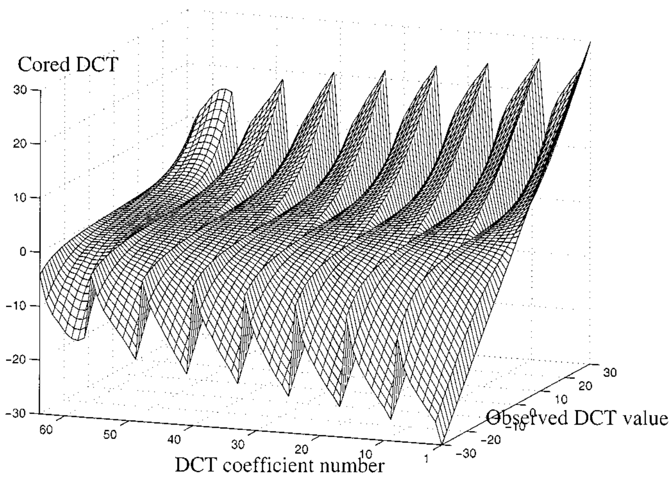


Fig. 7. Plot of part of the coring functions for all 64 DCT coefficients, computed for noise with variance 100 corrupting the image.

zero-mean and variance $\sigma^2_{W_k}$. All terms in (11) and (12) are now known, so that the coring function $g(Z_k(\boldsymbol{n}))$ for each of the DCT coefficients can be determined.

As an illustration, Fig. 6 shows the coring function computed for DCT coefficient number 8 for noise with variance of 100, i.e., $\sigma^2_{W_8} = 100/64$. In this figure, small values are cored toward zero, whereas larger values are altered less. This agrees with the intuitive conclusion that data with small amplitudes are noisy and unreliable, and they should therefore be discarded. Fig. 7 shows the coring functions for all 64 DCT coefficients, again for noise with variance 100 corrupting the image. The dc coefficient is not cored, hence the 45-degree line for this DCT coefficient. It can be seen that coefficients representing higher spatial frequencies are cored toward zero more strongly than coefficients representing lower spatial frequencies. This, again, matches well with the fact that natural images contain less energy in high frequencies than in low frequencies.

The coring functions must clearly depend on the noise variance. This can be seen from (12), which depends on the PDF, and therefore variance, of the noise. In a practical MPEG compression system, coring functions can be computed in advance for different noise variances to avoid computational overhead during the actual video compression.

V. Experimental Results

We have experimented with the embedded coring functions within the context of the standard *Test Model 5* (TM5) MPEG-2 encoder [17]. This section describes two sets of experiments. The first experiment evaluates the performance of the embedded coring functions for varying amounts of noise in terms of PSNR. The second experiment compares the performance of the embedded coring functions with the application of a “standard” wavelet-based noise-reduction coring prefilter entirely outside the MPEG encoder [7], [8], [18].

The first experiment evaluates the performance of the MPEG encoder with embedded coring functions in terms of the PSNR when applied to video sequences with varying amounts of noise. Fig. 8 shows the configuration used for measuring the PSNR of the compressed video sequences. Note that in this simulation, we measure the PSNR with respect to the *noise-free* video sequence. Of course, in the actual coring or quantization of the noisy video sequence, no use is made of the original video sequence. Fig. 9(a) (test sequence *Plane*) and Fig. 9(b) [test sequence *MobCal* (Mobile-Calendar)] show the measured PSNRs for bit rates ranging from 2 to 15 Mbit/s. The results show that the PSNRs of the cored and coded sequences are considerably higher than those of the noisy input sequences. Clearly, MPEG compression with embedded coring improves the quality of the video sequences rather than degrading it as in the typical case with noise-free video sequences.

The PSNRs of the corrected sequences increase more rapidly with increasing bit rate at low bit rates than at high bit rates. Specifically, the curves for test sequences with noise variance 100 and 225 are quite flat over the range from 4 to 15 Mbit/s. This contrasts with the PSNRs for noise-free sequences, which increase steadily with increasing bit rate. This implies that there is an “early” *saturation point* for the bit rate in compressing noisy video sequences. Encoding with bit rates above this saturation point gives only marginal improvements in coded image quality.

The second experiment investigates whether the MPEG encoder with embedded coring performs better than the standard MPEG encoder in combination with a prefilter. It could be imagined that even though a highly optimized prefilter outperforms the embedded coring functions in the MPEG encoder, its superior quality may be lost due to quantization errors introduced by the subsequent MPEG encoding. The prefilter used in these experiments is the coring of wavelet coefficients obtained from the temporally extended version of the 2-D Simoncelli pyramid [18] that is described in [7] and [8]. This coring technique is called the *3-D pyramid filter*. Similar results can be obtained with other state-of-the-art wavelet and coring-based noise-reduction techniques.

A moderate amount of noise (variance 100) was added to the *Plane* and *MobCal* sequences. Fig. 10 shows the PSNR as a function of the bit rate of the noisy video sequences after encoding by the standard TM5 MPEG encoder with and without prefiltering by the 3-D pyramid filter. The PSNRs that result from applying the MPEG coder with embedded coring to the noisy image sequences are also shown in this figure. For the purpose of reference, the PSNRs of the compressed original (noise-free) image sequences are plotted in the same figure.

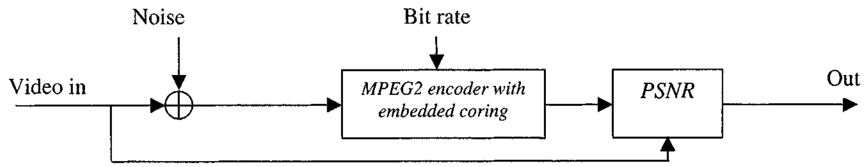


Fig. 8. Scheme for measuring the PSNR of cored and coded noisy image sequences.

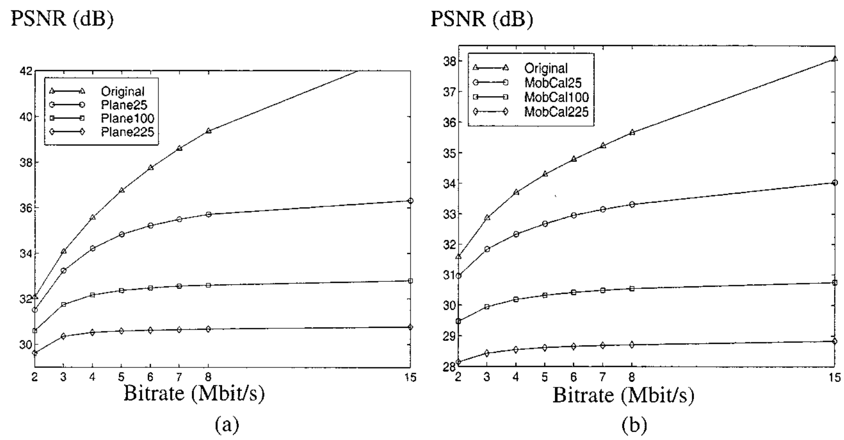


Fig. 9. Measured PSNRs for (a) Plane and (b) MobCal sequence using an MPEG encoder with embedded coring. The noise variance in the noisy sequences was 25, 100, and 225, which correspond to PSNRs of 33.0, 27.0, and 23.5 dB, respectively.

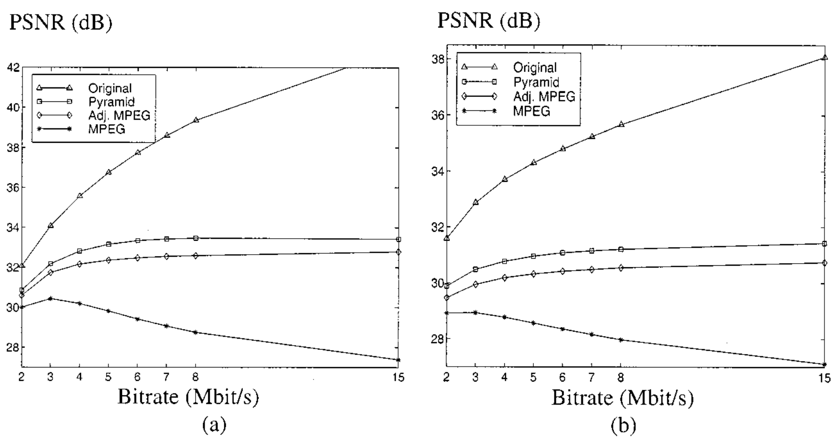


Fig. 10. (a) PSNR versus bit rate for the original Plane sequence, noisy Plane sequence (noise variance 100), and noise-reduced Plane sequence (filtered by the 3-D pyramid filter) encoded by the standard TM5 MPEG encoder. Also shown is the PSNR of the noisy Plane sequence that was encoded by the MPEG encoder with embedded coring function. (b) As in (a) but now for the MobCal sequence.

Fig. 10 shows that prefiltering video sequences with a moderate amount of noise prior to encoding with the standard MPEG encoder gives a PSNR that is maximally 1-dB higher than when the coring is embedded into the MPEG encoder. It can also be seen that the standard MPEG encoder (without coring and without prefilter) already operates as a noise reducer at low bit rates. At 3 Mbit/s, the PSNR of the coded noisy *Plane* sequence is about 3.5-dB higher than that of the noisy original (30.5 dB compared to 27.0 dB). This figure is 1.5 dB for the *MobCal* sequence (28.5 dB compared to 27.0 dB). The PSNR decreases at higher bit rates. This behavior is not surprising. The encoder applies a coarse quantization at low bit rates, removing a lot of noise energy as a result of the dead zone. Since the encoder is capable of encoding the signal and therefore also the noise more accurately at higher bit rates, the noise in the signal is preserved better, reducing the PSNR with respect to the noise-free original video sequences. In the limiting case, at very high bit rates, the noisy sequence is encoded without quantization errors, yielding a PSNR of 27.0 dB: the PSNR of the noisy sequence.

VI. Conclusions

In this paper, we have shown that noise reduction can take place within an MPEG encoder by embedding a coring function. The coring takes place on each of the DCT coefficients. Coring modules have to be placed between the DCT transform and the DCT coefficient quantizers. The additional costs of the embedded coring functions compared to a standard MPEG encoder consist of 63 noise-dependent (switchable) lookup tables and separate 8×8 DCTs on the B- and P-frames. This is a cheaper solution than preceding the MPEG encoder with a separate noise-reduction prefilter, while the performance is quite comparable, giving a loss of at most 1 dB over a wide range of bit rates at a moderate noise level of 27 dB.

REFERENCES

[1] D. L. Donoho, "De-noising by soft-thresholding," *IEEE Trans. Inform. Theory*, vol. 41, pp. 613–627, 1995.

[2] D. L. Donoho and I. M. Johnstone, "Ideal spatial adaptation via wavelet shrinkage," *Biometrika*, vol. 81, pp. 425–455, 1994.

[3] P. G. Powell and B. E. Bayer, "A method for the digital enhancement of unsharp, grainy photographic images," in *Proc. IEE Int. Conf. Electronic Image Processing*, 1982, pp. 179–183.

[4] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden, "Pyramid methods in image processing," *RCA Engineer*, vol. 29, no. 6, pp. 33–41, 1984.

[5] P. C. Goldmark and J. M. Hollywood, "A new technique for improving the sharpness of television pictures," *J-SMPTE*, vol. 57, pp. 382–396, 1951.

[6] J. P. Rossi, "Digital techniques for reducing television noise," *J-SMPTE*, vol. 87, pp. 134–140, 1978.

[7] P. M. B. van Roosmalen, S. J. P. Westen, R. L. Lagendijk, and J. Biemond, "Noise reduction for image sequences using an oriented pyramid thresholding technique," in *Proc. ICIP '96*. Lausanne, Switzerland: IEEE, 1996, vol. I, pp. 375–378.

[8] A. Hanjalic, G. C. Langelaar, P. M. B. van Roosmalen, J. Biemond, and R. L. Lagendijk, *Image and Video Databases: Restoration, Watermarking and Retrieval*, ser. Advances in Image Communications. Amsterdam, The Netherlands: Elsevier, 2000, vol. 8.

[9] J. K. Wolf and J. Ziv, "Transmission of noisy information to a noisy receiver with minimum distortion," *IEEE Trans. Inform. Theory*, vol. IT-16, pp. 406–411, 1970.

[10] N. S. Jayent and P. Noll, *Digital Coding of Waveforms: Principles and Applications to Speech and Video*. Englewood Cliffs, NJ: Prentice-Hall, 1984.

[11] J. L. Mitchell, W. B. Pennebaker, C. E. Fog, and D. J. LeGall, *MPEG Video Compression Standard*. London, U.K.: Chapman & Hall, 1996.

[12] B. K. Natarajan, "Filtering random noise from deterministic signals via data compression," *IEEE Trans. Signal Processing*, vol. 43, pp. 2595–2605, Nov. 1995.

[13] R. L. Joshi and T. R. Fisher, "Comparison of generalized Gaussian and Laplacian modeling in DCT image coding," *IEEE Signal Processing Lett.*, vol. 2, pp. 81–82, May 1995.

[14] R. C. Reiniger and J. D. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. Commun.*, vol. COM-31, pp. 835–839, June 1983.

[15] M. Barni, F. Bartolini, A. Piva, and F. Rigacci, "Statistical modeling of full frame DCT coefficients," in *Proc. Signal Processing IX*, vol. 3, Rhodes, Greece, 1998, pp. 1513–1516.

[16] K. Sharifi and A. Leon-Garcia, "Estimation of shape parameter for generalized Gaussian distributions in subband decomposition of video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 5, pp. 52–56, Jan. 1995.

[17] *Test Model 5*, Test Model Editing Committee, IEC/ISO JTC1/SC29/WG11/93-400, 1993.

[18] E. P. Simoncelli and E. H. Adelson, "Noise removal via Bayesian wavelet coring," in *Proc. ICIP '96*, vol. I, Lausanne, Switzerland, 1996, pp. 379–382.

Peter M. B. van Roosmalen received the M.Sc. and Ph.D. degrees in electrical engineering from the Information and Communication Theory Group, Delft University of Technology, Delft, The Netherlands, in 1994 and 1999, respectively. His research was carried out in the context of the European research program Automated Restoration of Original Video and Film Archives (AURORA).

He was a Visiting Scientist at the Institut National d'Audiovisuel (INA), Paris, France, during his studies. He is co-author of the book *Image and Video Databases: Restoration, Watermarking, and Retrieval* (Amsterdam, The Netherlands: Elsevier, 2000).

Reginald L. Lagendijk (S'87–M'90–SM'97) received the M.Sc. and Ph.D. degrees in electrical engineering from the Delft University of Technology (TU-Delft), Delft, The Netherlands, in 1985 and 1990, respectively.

He became an Assistant Professor and Associate Professor at TU-Delft in 1987 and 1993, respectively. He was a Visiting Scientist in the Electronic Image Processing Laboratories, Eastman Kodak Research, Rochester, NY, in 1991, and a Visiting Researcher at Microsoft Research, Beijing, China, in 2000. Since 1999, he has been Full Professor in the Information and Communication Theory Group, TU-Delft. He is author of the book *Iterative Identification and Restoration of Images* (Norwell, MA: Kluwer, 1991), and co-author of the books *Motion Analysis and Image Sequence Processing* (Norwell, MA: Kluwer, 1993), and *Image and Video Databases: Restoration, Watermarking, and Retrieval* (Amsterdam, The Netherlands: Elsevier, 2000). At present, his research interests include signal processing and communication theory, with emphasis on visual communications, compression, analysis, searching and watermarking of image sequences. He has been involved in the European Research projects DART, SMASH, STORit, DISTIMA, and CERTIMARK. He is currently leading several projects in the field of wireless communications, among which the interdisciplinary research program Ubiquitous Communications at TU-Delft.

Dr. Lagendijk has served as Associate Editor of the IEEE TRANSACTIONS ON IMAGE PROCESSING, and he is currently Area Editor of Eurasip's *Journal Signal Processing: Image Communication*. He is a member of the IEEE Signal Processing Society's Technical Committee on Image and Multidimensional Signal Processing.

Jan Biemond (M'80–SM'87–F'92) was born in De Kaag, The Netherlands. He received the M.Sc. and Ph.D degrees in electrical engineering from Delft University of Technology (TU-Delft), Delft, The Netherlands, in 1973 and 1982, respectively.

Currently, he is a Professor in the Information and Communication Theory Group and Head of the Mediamatics Department of the Faculty of Information Technology and Systems at TU-Delft. His research interests include image and video processing (restoration, compression, analysis, retrieval) and motion estimation with applications in digital TV, 3D TV, HDTV, multimedia, digital libraries, scan-rate conversion, and computer vision. He has published extensively in these fields. In 1983, he was a Visiting Professor at Rensselaer Polytechnic Institute, Troy, NY, and at Georgia Institute of Technology, Atlanta, GA. He is the Scientific Editor of a series of books on Advances in Image Communication with Elsevier Science BV.

Dr. Biemond is President Elect of the IEEE Benelux Section in Region 8 and a member of the Nominations and Appointments Committee of the IEEE Signal Processing (SP) Society. He served the SP Society as a Distinguished Lecturer from 1993 to 1994, and is a former member of the Administration Committee of the European Association for Signal Processing (EURASIP), the Technical Committee on Image and Multidimensional Signal Processing, and the Board of Governors of the IEEE SP Society. He served as the General Chairman of the 5th IEEE SP/EURASIP Workshop on Multidimensional Signal Processing, held at Noordwijkerhout, The Netherlands, in 1987, as the General Chairman of the 1997 Visual Communication and Image Processing Conference (VCIP '97), San Jose, CA, and as the Chairman of the 21st Symposium on Information Theory in Benelux in Wassenaar, The Netherlands, in 2000. He received the Dutch Telecom Award "Vederprijs" in 1986 for his contributions in the area of digital image processing, in particular in image restoration and subband coding.