Proceedings, IEEE International Symposium on Circuits and Systems, London, England, May-June 1994.

Error Accumulation of Repetitive Image Coding

Shih-Fu Chang and Alexandros Eleftheriadis

Department of Electrical Engineering & Center for Telecommunications Research
Columbia University
New York, NY 10027
sfchang, eleft@ctr.columbia.edu

ABSTRACT

Repetitive image coding is encountered in applications such as video transcoding and image/video servers. This paper studies the effect of error accumulation in repetitive image coding. We describe the conditions under which *zero-error-accumulation* property holds. We study the impact of intermediate operations (like linear filtering, scaling, and shifting) on image quality in repetitive coding. Experiments with DCT, MC, and MPEG coding algorithms are reported.

1. Introduction

Most image/video compression algorithms assume that the image data will go through the encoding/decoding process once only and performance is usually optimized based on this assumption. However, there are several situations which require repetitive image coding [6], that is, encoding and decoding of the same bit stream more than once. One example is video signal processing within the network. Input video signals are decoded, processed, and then re-coded before being transmitted to final destinations. Another example is image databases and video servers. The stored image/video signals (in coded form) are retrieved, processed, and re-coded before they are re-stored back to the database or sent to users. One important issue here is how this repetitive image coding process affects the image quality. This paper studies this issue for two popular compression algorithms — Discrete Cosine Transform (DCT) [2] and Motion Compensation (MC) [3]. We will describe conditions under which the *zero-error-accumulation* property can be achieved. We will also describe the effect of intermediate image processing (e.g., linear filtering and scaling) on the error accumulation effect.

2. Repetitive Transform Coding and Quantization

The error-accumulation property of the DCT algorithm with quantization is investigated in this section. Similar approaches can be applied to other transform coding algorithms, like Discrete Sine Transform (DST). Figure 1(a) shows a diagram illustrating the procedure of repetitive DCT coding plus quantization. Assume u_1 is a one-dimensional (1D) image vector. The first decoded image u_1' is further processed by some intermediate operation (denoted as H) to produce image vector u_2. This new image vector is then encoded, quantized, and decoded again to produce the second decoded image u_2'. The same procedure may be repeated in a multi-pass repetitive coding situation. For comparison, Figure 1(b) shows the situation when only a single-pass coding is used and the required image processing function is done at the source.

We ignore the computing roundoff error here. The *zero-error-accumulation* property is achieved when

$$u_i = u_i', \quad v_i = v_i', \quad i \geq 2 \tag{1}$$

where v_i is the DCT transform of image vector u_i. In other words, there is no additional noise introduced besides the first coding pass. However, this property holds only for few cases. One example is when the intermediate function H is null (i.e., output = input) and the second coding pass uses the same quantizer as that for the first pass. In general, there will be additional noise introduced when the second quantizer differs from the first one. For example, to transcode a JPEG encoded bit stream [5] to a lower quality level, we can use a coarser quantizer in the 2nd pass, that is, to increase the quantizer multiplicative factor (denoted as *M*). In JPEG, the quantization process can be described as

$$Q(v(k)) = \left\lfloor \frac{v(k)}{q(k) \cdot M} \right\rfloor \cdot (q(k) \cdot M) \tag{2}$$

where $\lfloor x \rfloor$ stands for the floor function, Q(.) the quantizer and q(k) the quantization step size. If the second quantizer uses a different multiplicative factor M', then the second quantized DCT coefficients become

$$Q(Q(v(k))) = \left\lfloor \frac{\left\lfloor \frac{v(k)}{q(k) \cdot M} \right\rfloor}{M'/M} \right\rfloor \cdot (q(k) \cdot M') \tag{3}$$

The *zero-error-accumulation* property holds only when

$$\left\lfloor \frac{v(k)}{q(k) \cdot M} \right\rfloor = n \cdot M'/M \tag{4}$$

where n is an arbitrary integer. However, this condition usually does not hold and therefore the second coding pass usually introduces extra noise.

The intermediate operation H between two consecutive coding passes also has significant impact on the image quality. It may reshape the DCT spectrum and therefore the second coding pass will treat the output of H as a new signal and in general add new quantization noise. One typical operation is linear filtering, including scaling. Because of the block structure of DCT, the effect of linear filtering on the DCT

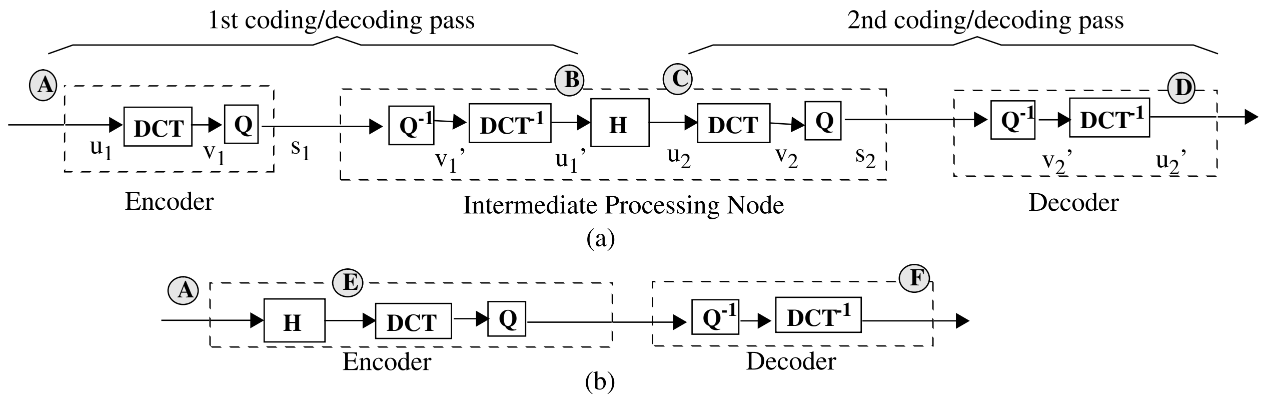

Fig.1 (a) DCT-based repetitive image coding. H represents intermediate operations between two consecutive coding passes. (b)Single-pass coding in which required image operation H is performed before encoding.

spectrum shape is not straightforward. As derived in [1,8,9], linear filtering can be expressed in a block-wise form:

$$\bar{u} = \sum_{i \in S} H_i u_i \tag{5}$$

where u_i's are input blocks which have contributions to the filtered output block $\bar{u}$, and the summation range S is determined by the filter kernel length. The DCT of the filter output can be obtained by [1,8,9]

$$DCT(\bar{u}) = \sum_{i \in S} DCT(H_i)\, DCT(u_i) \tag{6}$$

To study the effect of linear filtering on the DCT spectrum, we assume the input image has a first-order stationary Markov model. Its variance matrix is as follows,

$$R_u = \rho^{|i-j|} \tag{7}$$

From equation 5, we can derive the variance matrix of the filtered image vector as follows,

$$R_{\bar{u}} = \sum_{i, j \in S} H_i R_{u_i u_j} H_j^t \tag{8}$$

and the variance matrix of its DCT as follows,

$$\begin{aligned} R_{\bar{v}} &= A\left(\sum_{i, j \in S} H_i R_{u_i u_j} H_j^t\right) A^t \\ &= \sum_{i, j \in S} DCT(H_i)\, R_{v_i v_j}\, DCT\left(H_j^t\right) \end{aligned} \tag{9}$$

where A is the DCT transform matrix. To illustrate, we study how the down-scaling operation (e.g., 2:1 down scaling) affects the DCT spectrum. We analyze two cases — $\rho = 0.95$ and $\rho = 0$. The former represents highly correlated images and the latter uncorrelated signals. We compute the variance of each DCT coefficient (i.e. the diagonal elements of the DCT variance matrix) and compare their energy distribution and compactness. A three-tap LPF ({0.25, 0.5, 0.25}) is used. The results in Figure 2 show that for the highly-correlated Markov source (i.e. ρ=0.95), the down-scaling operation spreads out the DCT spectrum. The spectrum *flatness* figure (defined as the ratio of geometric mean to arithmetic mean [7]) is increased from 0.13 to 0.15. For the totally uncorrelated source (i.e. ρ=0), the down-scaling operation reduces the overall energy by more than one half. This is because that the higher half-band energy is eliminated.

We also ran experiments on real images to verify the theoretical analysis. Figure 3 shows a typical result by using one non-uniform quantizer (DC coefficient is not quantized). The PSNR of the reconstructed images at different stages (labeled as in Figure 1) are shown. Some observations can be made from this example:

Observation 1. The 2nd coding pass does introduce extra noise. This can be verified by the moderate PSNR between C & D (denoted as $PSNR_{CD}$). It can also be verified by the fact that single-pass coding always produces higher PSNR

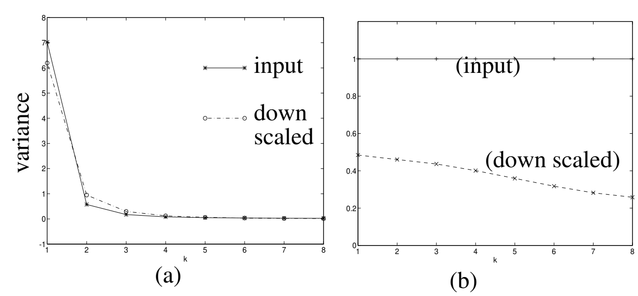

Fig.2 Effect of down-scaling on the DCT spectrum of Markov image sequences. (a) ρ=0.95 (b) ρ=0.

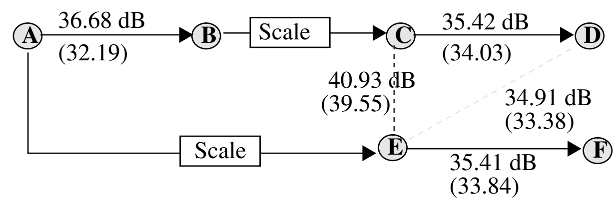

Fig.3 Image quality at different stages in a repetitive coding system, labeled as in Figure 1. Reported values are PSNR for DCT coding. Values within parenthesis are PSNR for MPEG2 coding, described in Section 3.1.

than two-pass coding ($PSNR_{AB}$, $PSNR_{EC}$, and $PSNR_{EF}$ are all higher than $PSNR_{ED}$).

Observation 2. The down-scaled image suffers more quality degradation during coding than the full-size image. This is mainly due to the more wide-spread spectrum shape of the down-scaled image, as opposed to the concentrated spectrum of the original full-size image. As shown in Figure 3, $PSNR_{EF}$ is lower than $PSNR_{AB}$ by 1.2 dB.

Observation 3. The down-scaling operation can effectively remove the quantization noise from the previous coding pass ($PSNR_{EC}$ is much higher than $PSNR_{AB}$). It is mainly due to the noise reduction capability of the down-scaling operation described earlier. The quantization noise of different DCT coefficients can be approximately considered as independent components. As shown in Figure 2(b), for a totally uncorrelated random sequence, the down-scaling operation reduces the noise energy by more than one half.

The last observation has significant implications for practical applications. If the same quantizer is used, a coding pass followed by a down-scaling (or in general low-pass filtering) produces higher image quality than the opposite approach, that is, down-scaling before coding. As shown in Figure 3, $PSNR_{EC}$ is much higher than $PSNR_{EF}$. However, it should be noted that the required bit rate is different.

Besides scaling, we also tested different intermediate operations such as shifting and overlapping. *Observation 1* still holds, but not the other two. The second coding pass still introduces extra noise (at least 1.2 dB in our simulations) since the DCT coefficients are more or less modified by the intermediate operations. However, the spectrum shape is approximately maintained so that the execution order between the coding process and the intermediate operation no longer matters (opposed to that described in *Observation* 2). Also, these operations do not provide the same noise reduction capability as that for the down-scaling operation (described in *Observation 3*).

Figure 4 shows the image quality evolution in a multi-pass coding scenario. Because the down-scaling operation makes the image DCT spectrum flatter and usually higher-order DCT coefficients are coded more coarsely, each coding pass causes more quality loss than its previous one. This can be verified by the decreasing PSNR values along the coding path in Figure 4(a). However, the down-scaling operation also has the capability of filtering out the higher-band quantization noise, as discussed in *Observation 3*. The actual image quality (compared to the down-scaled images from the un-coded path) has much higher PSNR than that shown in each coding pass. For the shifting operation, to the contrary, although every coding pass introduces approximately the same level of noise (the PSNR values along the coding path remain approximately constant), the quantization noise is accumulated and thus the final image quality is much lower than that in the down-scaling case.

3. Repetitive MC Coding

Unlike DCT, the MC coding algorithm may accumulate errors even when there are no intermediate operations between consecutive coding passes. This property originates from the fact that the MC algorithm, unlike DPCM, does not use static predictions. In DPCM, the prediction comes from the same position (either spatially or temporally). In MC, the motion estimation procedure searches for the most similar pixels in the previous frame as the prediction. The prediction errors are directly quantized or transformed into DCT domain and then quantized. After quantization, it is no longer guaranteed that the original prediction pixels still have the minimal distance from the current pixels. Figure 5 shows an example where a new prediction (originally a sub-optimal choice) is closer to the reconstructed pixels than the old prediction (originally the optimal choice). New prediction errors need to be calculated in the second coding pass, and hence the reconstructed pixels from the second pass will be different from the reconstructed pixels from the first pass. Once this happens, the whole effect may ripple through future frames in the video sequence because current reconstructed frames will be used as predictions for future frames.

One way to rectify this error accumulation problem for the MC algorithm is to use a quantizer that always pulls the reconstructed pixels closer towards the original prediction pixels, that is, the quantized prediction error should be "smaller" than the original prediction error. One example of this type of quantizer is as follows,

$$\begin{aligned} \hat{x} &= t_{i-1} \quad \text{when} \quad x \in [t_{i-1}, t_i) \;\&\; t_{i-1} \geq 0 \\ &= t_i \quad\quad \text{when} \quad x \in (t_{i-1}, t_i] \;\&\; t_i \leq 0 \end{aligned} \tag{10}$$

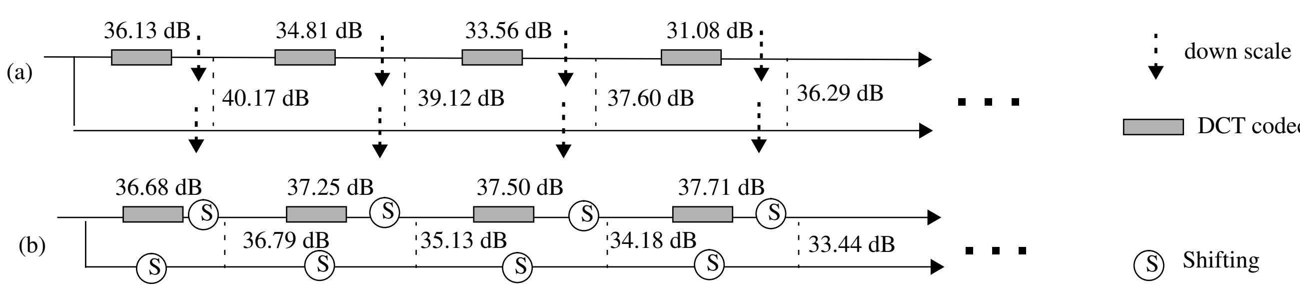

Fig.4 Image quality evolution in repetitive DCT-based coding with different intermediate operations (a)down-scaling (b)shifting

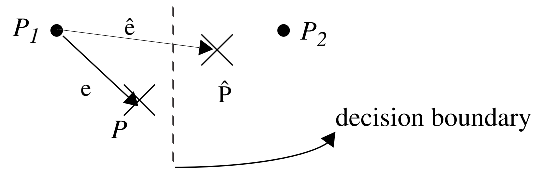

Fig.4 Quantization of a prediction error may make the reconstructed pixel closer to an originally sub-optimal prediction point. *P* is the original pixel, P_1 the optimal prediction pixel in the first coding pass, *e* prediction error, ê quantized prediction error, and P̂ reconstructed pixel. P_2 will be chosen to be the new prediction in the 2nd coding pass.

where $\hat{x}$ is quantized reconstruction value of x and t_i's are quantizer decision levels.

However, quite often the motion compensated prediction errors are quantized in the transform domain. It is difficult to design transform coefficient quantizers in a way that the reconstructed pixels can be pulled inward towards the prediction pixels. We have simulated a simple MC-only repetitive coding algorithm without any intermediate operations. The test sequence is the "pingpong" video sequence. We used the non-linear quantizer for DPCM suggested in [10]. The error accumulation effect does appear but seems to be insignificant (about 0.1 dB). However, as described above, the overall accumulated error depends on the quantizer, the specific test sequence, and the intermediate operations (if any).

Using quantizers like that described above (Equation (10)) we can avoid error accumulation and maintain the image quality in repetitive coding. However, due to its sub-optimality, the initial video quality (i.e. the result of first coding pass) may be inferior to that using optimal quantizers (but causing error accumulation in repetitive coding). Thus, a tradeoff exists between the initial image quality and the long-term image quality.

3.1 Experiment with MC-DCT coded video

We also ran the repetitive coding experiment of Figure 3 with the MC-DCT-based hybrid coding algorithm, such as the MPEG2 coding standard [4]. The test sequence is "Flower Garden" (704 pixel by 480 pixel, YUV interlaced) and we use 8 Mbps for the first coding pass and 2 Mbps for the second pass (since the input video is scaled down to a quarter size). All I,P,B-frame modes are used and MC is field-based. Video quality (PSNR) at different stages are shown in Figure 3 (values within the parenthesis). Similar to the results for DCT coding, the second coding pass causes an extra quality loss of about 0.5 dB ($PSNR_{ED}$ vs. $PSNR_{EF}$). The down-scaling operation effectively reduces the quantization noise introduced in the first coding pass ($PSNR_{EC}$ much higher than $PSNR_{AB}$). It should be noted that with MPEG rate-constrained coding, it is hard to control to use the same quantizers during different coding passes.

4. Conclusions

Repetitive image coding will be encountered in applications like image databases, video servers, video transcoding, and video processing within networks. We have shown that two popular video compression algorithms, DCT and MC, are in general not error-accumulation-free in repetitive coding. The actual accumulated error amount depends on the quantizer design and the intermediate operations between consecutive coding passes. We have also described the conditions under which the accumulated errors can be avoided. However, error accumulation usually cannot be avoided whenever there are intermediate operations. Our experiment using DCT repetitive coding with an intermediate down-scaling operation shows that an additional quality degradation of at least 0.5 dB ($PSNR_{ED}$ vs. $PSNR_{EF}$ in Figure 3) is introduced in the second coding pass. If we apply a shifting operation between consecutive coding passes, the second coding pass introduces an extra quality loss of at least 1.2 dB. The reason causing this difference is partly because that down-scaling can filter out noise energy in the high-frequency band and improve the overall image quality. For repetitive MC coding, the error accumulation effect seems to be less significant. For the hybrid coding algorithm (e.g., MPEG2), error accumulation in multi-pass coding is also observed (about 0.5 dB extra quality loss).

We also derive the formula for finding the DCT spectrum after linear filtering. We show that down-scaling tends to spread out the DCT energy distribution for highly correlated images. Down-scaling also substantially reduces the uncorrelated quantization noise introduced in the coding process.

5. References

[1] S.-F. Chang and D. G. Messerschmitt, "A New Approach to Decoding and Compositing Motion Compensated DCT-Based Images," ICASSP '93, Minneapolis, Minnesota, pp. V421-V424, April, 1993.

[2] N. Ahmed, T. Natarajan, and K.R. Rao, "Discrete Cosine Transform," IEEE Trans. on Computers, Vol. C-23, pp.90-93, Jan. 1974.

[3] A.N. Netravali and J.D. Robbins, "Motion-Compensated Television Coding: Part I," The Bell System Technical Journal, Vol. 58(3), pp.631-670, Mar. 1979.

[4] Standard Draft, MPEG Video Committee Draft, Nov., 1993.

[5] Standard Draft, JPEG-9-R7, Feb. 1991

[6] A. Lippman, "Feature Sets for Interactive Images," Communications of ACM, pp. 93-102, Vol. 34, No. 4., April, 1991.

[7] J. Woods, *Subband Image Coding*, Kluwer Academic Publisher, 1991.

[8] S.-F. Chang and D. G. Messerschmitt, "Manipulation and Compositing of MC-DCT compressed Video," submitted to IEEE J. of Selected Areas in Communications, 1993.

[9] J.B. Lee and B.G. Lee, "Transform Domain Filtering Based on Pipelining Structure," IEEE Trans. on Signal Processing, pp. 2061-4, Vol. 40, No. 8, Aug. 1992.

[10] A.N. Netravali and B.G. Haskell, *Digital Pictures: Representation and Compression*, New York, Plenum Press, 1988.