BUFFER CONTROL ALGORITHM FOR LOW BIT-RATE VIDEO COMPRESSION

K.T.Ng, S.C.Chan and T.S.Ng

Department of Electrical and Electronic Engineering
The University of Hong Kong, Pokfulam Road, Hong Kong
E-mail: scchan@hkueee.hku.hk

ABSTRACT

In this paper, a new buffer control algorithm for motion-compensated hybrid DPCM/DCT coding (like H.261 and MPEG-1 I pictures) is presented. The algorithm uses the bit allocation algorithm to determine the quantization scale factor of each macroblock under a given target bit rate. An important advantage of the algorithm is that it has precise control of the buffer and avoids buffer overflow events which is a severe problem in low bit rate video coder. Furthermore, the coder is able to allocate bits to the picture as a whole, resulting in better rate-distortion trade-off. Simulation results show that the H.261 coder, using the proposed algorithm, can achieve a higher PSNR and better visual quality than codec using conventional buffer control algorithm.

1. INTRODUCTION

Motion-compensated hybrid DPCM/DCT coding is a commonly-used compression technique for digital video signals [1]-[4]. Motion estimation and compensation are used to reduce the temporal redundancies in natural video signal. While spatial redundancy in the original or residual error images are decorrelated by the block discrete cosine transform (DCT). After the DCT, the block energy is concentrated into a few low order coefficients and data compression is actually achieved by the quantization operation which maps the transformed coefficients into a finite set of reconstruction levels or symbols. To exploit the non-uniform distribution of the transformed coefficients, the quantized DCT coefficients are encoded using a variable length code (typically a Huffman/run-length code) which generates an output bit stream at a variable rate.

In order to transmit the variable rate compressed bit stream over a fixed rate channel, a channel buffer is required. To prevent the buffer from overflowing and underflowing, buffer control mechanism must be used to regulate the fluctuation of the output bit. This is usually achieved by adjusting the quantizer step size for the DCT coefficients. Typical channel buffering technique includes the size of the channel buffer, the frequency and degree of adjustment of quantization levels [5]-[8].

For low bit-rate video compression, due to the low channel rate, the end-to-end delay mainly depends on the transmission delay of the coded pictures stored in the encoder buffer. In order to limit the delay to an acceptable value, the size of the buffer needs to be scaled down. However, the frame-rate is considerably reduced to give each frame enough number of bits for reasonable picture quality, the average number of bits per frame will be relatively larger compared to the buffer size. This implies that the effectiveness of the buffer is more limited and the buffer regulation is more difficult [9]. In traditional buffer control algorithms, quantizer step size for the current block depends only on the current buffer status and sometimes on the activities of the block. Therefore, the quantization error of adjacent blocks can vary significantly especially at the boundary of moving objects. This is undesirable as it will generate significant blocking artifacts in the encoded images. In this paper, we formulate the buffer control mechanism as a bit allocation problem to regulate the bit stream generated by the motion-compensated hybrid DPCM/DCT video codec. The salient features of our scheme are that i) the quantization scale factors are determined using information of the whole picture; ii) it has precise control of the buffer; and iii) it tries to allocate the given number of bits as efficient as possible in a rate-distortion sense. In particular, the H.261 video codec is used for demonstrating the effectiveness of the algorithm.

2. CODEC DESCRIPTION

Fig. 1 shows the generalized form of the H.261 source coder. The main elements are prediction, block transformation and quantization. The input frame is partitioned into macroblocks consisting of one luminance

block of (16×16) pixels and two chrominance blocks of (8×8) pixels. For each macroblock, motion-compensation based on luminance component is used to predict the whole macroblock. The prediction error (INTER mode) or the input picture (INTRA mode) will then undergo the two dimensional DCT of size (8×8). The number of quantizers is 1 for the INTRA DC coefficient and 31 for all other coefficients. Within a macroblock, the same quantizer is used for all coefficients except the INTRA DC one. The INTRA DC coefficient is nominally the transform value linearly quantized with a step size of 8 and no dead-zone. Each of the other 31 quantizers is also nominally linear but with a central dead-zone around zero and with a step size of an even value in the range 2 to 62 [1].

The video multiplex is arranged in a hierarchical structure with four layers. From top to bottom the layers are Picture, Group of Block (GOB), Macroblock (MB), Block. By using MQUANT in the structure of macroblock layer, each macroblock can use its own quantization scale factor. The output bit-rate is controlled by varying the quantization scale factor. For example, in RM8, it is updated at every 11th MB.

3. PROBLEM FORMULATION

Consider encoding the residual image after motion compensation that consists of G GOB and N macroblocks at a target bit-rate of R using the H.261 coder. We want to find a set of quantization scale factor $\{\ Q_n\ :\ n = 1,...,N\}$ where ($Q_n \in \{1,2,...31\}$) for the N macroblocks that minimize the overall distortion

$$D = \sum_{n=1}^{N} D_n(Q_n) \qquad (1)$$

subject to the bit-rate constraint:

$$B = H^{Pict} + G * H^{GOB} + \sum_{n=1}^{N} [H_n^{MB} + C_n(Q_n)] \le R \qquad (2)$$

Here, $D_k(Q_k)$ is the reconstructed distortion of the k-th MB if it is quantized with quantization scale factor Q_k, $C_k(Q_k)$ is the no. of bits generated in coding the DCT coefficients of the k-th MB with quantization scale factor Q_k, and H^{Pict}, H^{GOB}, H_k^{MB} are the number of bits generated for picture header, GOB header and k-th MB header, respectively. Information such as macroblock address, coded block patterns, quantizer step sizes, and motion vectors are included in H_k^{MB}.
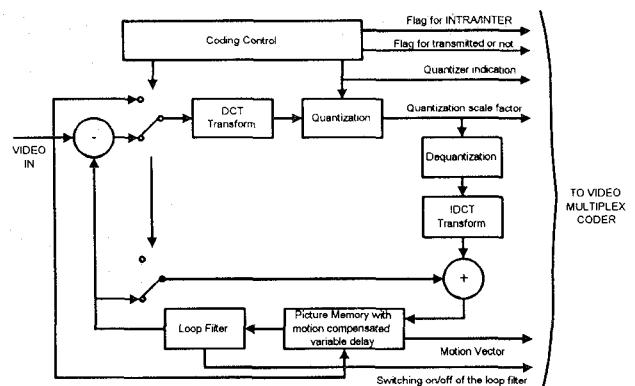

Fig. 1. H.261 Source Coder

The minimization problem of (1) is very similar to the bit allocation problem. Unfortunately, the Macroblock Address and Motion Vector Data in H_n^{MB} can either be coded directly or coded differentially depending on whether the previous blocks are skipped or not. This in turn depends on the quantizer step sizes of the previous blocks. Hence, it is very difficult to obtain an optimal solution. The problem can be solved by using the mean value to estimate those part of H_n^{MB} that will be affected by the nature of previous macroblocks. We can than calculate $\widetilde{H}_n^{MB}(Q_n)$ that only depend on Q_n. Eqn. (2) is then modified to:-

$$\widetilde{B} = \sum_{n=1}^{N} [\widetilde{H}_n^{MB}(Q_n) + C_n(Q_n)] \le \widetilde{R} \qquad (3)$$

where $\widetilde{B}$ is the number of bits used in coding the macroblocks and $\widetilde{R}$ is the target bit-rate for coding the macroblocks excluding bits that used by H^{Pict} and H^{GOB} which are constant for all picture.

Since the quantization of each macroblock is independent of each other, they can be viewed as allocating a given number of bits to N independent quantizers to minimize a total distortion measure. This problem has previously been solved by Gersho [10] & Riskin [11]. We shall show in the next section that how the bit allocation algorithm can be used to solve the buffer control problem in the H.261 coder.

4. QUANTIZATION SCALE FACTOR SELECTION ALGORITHM

The whole picture is first encoded with quantization scale factors of largest step size, corresponding to low bit-rate

and high distortion. Then, the algorithm decreases one of the quantization scale factors at a time until the target bit-rate is reached. In each step, the one with the best rate-distortion trade-off is selected. The quantization scale factors are updated in such a way that the ratio of decrease in distortion to increase in bit-rate is maximized over all possible reduced values for each quantization scale factor. Thus, we seek the values of value k and q that solve the maximization problem:

$$\max_{k} \max_{q} \frac{-\Delta D|_{Q_k \to q}}{\Delta \widetilde{B}|_{Q_k \to q}} \qquad (4)$$

where $\Delta D|_{Q_k \to q}$ and $\Delta \widetilde{B}|_{Q_k \to q}$ are respectively the change in distortion and the change in overall bit-rate used for all macroblocks when the quantization scale factor of k-th macroblock, Q_k, is replaced by q. As each quantizer is independent, these increments can be calculated as

$$\Delta D|_{Q_k \to q} = D_k(q) - D_k(Q_k) \qquad (5)$$

and

$$\Delta \widetilde{B}|_{Q_k \to q} = [\widetilde{H}_k^{MB}(q) + C_k(q)] - [\widetilde{H}_k^{MB}(Q_k) + C_k(Q_k)] \quad (6)$$

To simplify the computation, we split the maximization problem in (4) into two parts, where the first part computes

$$\lambda_k = \max_{q} \frac{-\Delta D|_{Q_k \to q}}{\Delta \widetilde{B}|_{Q_k \to q}} \qquad (7)$$

for all values of k and the second part solves

$$\lambda = \max_{k} \lambda_k \qquad (8)$$

To summarize, the algorithm is outlined below:

1. Calculate function $D_n(q)$ and $[\widetilde{H}_n^{MB}(q) + C_n(q)]$ for n = 1,...,N and q = 1,...,31.
2. Initialize all quantization scale factors to maximum value of Q_n.
3. Calculate λ_k and put them into a list in descending order of their values.
4. Update the quantization scale factor, Q_p that satisfies (8) (i.e. at the beginning of the list).
5. Calculate new value of λ_k and insert it into the list.
6. Repeat Steps 4 and 5 while $\widetilde{B} \le \widetilde{R}$.

It can be observed that our algorithm has precise control of the bit-rate for each image frame. Therefore, both the buffer requirement and the delay can be kept to minimal by using a constant bit-rate for each picture.

5. EXPERIMENTAL RESULTS

Computer simulations of H.261 codec operating at 64 kbps (12.5 frame per second) on QCIF image sequence were used to evaluate the proposed paradigm. Here, we compare the performance of our buffer control method with the PVRG-P64 codec from Stanford University, which is based on RM8. For fair comparison, both codecs start with an I picture using a uniform step size of 8. As different number of bits are used for each picture, the mean PSNR is measured. Table 1 shows the coding results for the test sequences Carphone, Miss America and Claire. The PSNR value of Miss America verses the frame number is illustrated in Fig. 2. It is observed that the proposed algorithm achieves a higher PSNR value as well as a better overall perceptual image quality than the PVRG-P64 coder at the same bit-rate.

The performance of the proposed algorithm operating at different bit-rate are shown in Table 2 and Fig. 3. It can be observed that the use of constant bit rate for each pictures together with the new buffer control algorithm can avoid rapid fluctuation in the quality of the final picture. This allows the codec to be scalable over a wide range of bit-rate.

6. CONCLUSION

We have presented an algorithm for the selection of quantization scale factors for H.261 coder. This algorithm efficiently allocates bits to each macroblock in the rate-distortion sense. The allocation is based on the nature of the input image and the quantization scale factors will adapt to the local activities of the picture instead of being determined by the status of the buffer as in conventional buffer control methods. As demonstrated in the simulation experiments, the bit allocation method can achieve much lower distortion at the same bit rate as compared with the traditional method. Furthermore, the rate-constrained adaptive quantization technique allows the user to have precise control of the bit-rate in H.261 coder. This feature is important in very low-bit-rate video coding where the effectiveness of the buffer is more limited.

Mean PSNR	Carphone (PVRG)	Carphone (proposed alg.)	Miss Amer. (PVRG)	Miss Amer. (proposed alg.)	Claire (PVRG)	Claire (proposed alg.)
Y Comp.	32.61	33.49	38.39	40.70	38.58	41.41
U Comp.	38.49	39.08	38.28	39.82	40.10	41.68
V Comp.	38.20	38.99	37.47	40.19	43.13	43.78

Table 1: Performance comparison.

Mean PSNR	128 kbps	64 kbps	16 kbps
Y Comp.	42.97	40.70	36.44
U Comp.	41.20	39.82	37.61
V Comp.	42.36	40.19	35.76

Table 2: Mean PSNR of Miss America sequence coded at different bit-rate using proposed algorithm.

REFERENCES

[1]   CCITT Recommendation H.261, "Video Coding for Audiovisual Services at px64 kbits/s," Genera, Aug. 1990.

[2]   ISO/IEC 11172-2, "Information Technology-Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to about 1,5 Mbit/s: Part 2 Video," Aug. 1993.

[3]   ITU-T Recommendation H.262|ISO/IEC 13818-2, "Information Technology-Generic Coding of Moving Pictures and Associated Audio Information: Video," Draft, Mar. 1994.

[4]   Draft ITU-T Recommendation H.263, "Video Coding for Low Bitrate Communication", July, 1995.

[5]   CCITT SG XV WP/1/Q4 Specialists Group on Coding for Visual Telephony, "Description of Reference Model 8 (RM8), June 1989, Document 525.

[6]   ISO/MPEG II, "Test Model 5," Document AVC-491, Apr. 1993.

[7]   ITU H.263, "Video Codec Test Model, TMN5", Jan. 1995

[8]   J. Zdepski, D. Raychaudhuri and K. Joseph, "Statistically Based Buffer Control Policies for a Constant Rate Transmission of the Compressed digital Video," IEEE Trans. on Commun., vol. 39, pp. 947-957, Jun. 1991.

[9]   M. T. Sun, K. Rijkse, D Schinkel and A Meijboom, "Coding and Interworking for Videotelephony," ISCAS, pp.20-23, 1993.

[10] Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers," IEEE Trans.
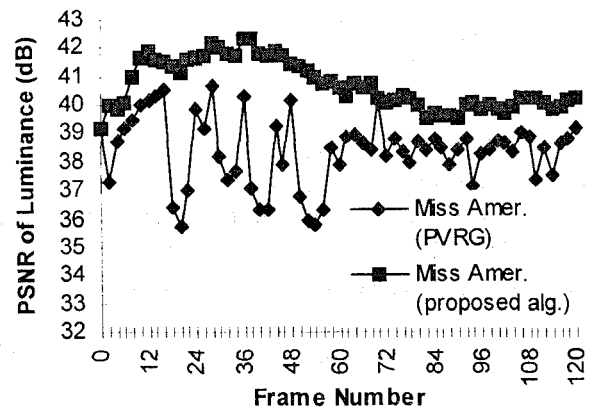
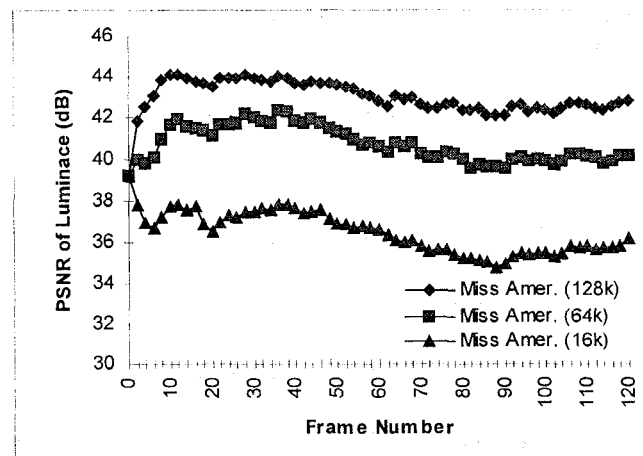Fig. 2. Quality comparison of proposed algorithm and px64 coder (at 64 kbps).

Fig. 3. PSNR of Luminance Component of Miss America verses Frame Number

Acoust. Speech Signal Process., vol. ASSP-36, pp.1445-1453, September 1988.

[11] E. A. Riskin, "Optimal Bit Allocation via the Generalized BFOS Algorithm." IEEE Trans. Inform. Theory, vol. 37, pp. 400-402, March 1991.