A MODIFIED H.263 ALGORITHM USING BIT ALLOCATION BUFFER CONTROL ALGORITHM

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: *ktng@eee.hku.hk, **scchan@eee.hku.hk

ABSTRACT

Buffer control is an important problem in very low bit-rate video coding. In a recent work [11], the authors had proposed a new buffer control algorithm for motion-compensated hybrid DPCM/DCT coding. The algorithm is based on the use of bit allocation algorithm to determine the quantization scale factors in such coder to meet a given target bit rate. Simulation results showed that, using the proposed algorithm, the H.261 coder can achieve a higher PSNR and better visual quality than the coder using traditional buffer control algorithm. In this paper, we apply this buffer control algorithm to a modified version of the H.263 algorithm for very low bit-rate video coding. Comparing the performance of the modified H.263 codec with the TMN5 model also shows that better visual quality can be obtained at comparable PSNR values.

I. 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 is 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 quantization 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 variable length codes (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 bits. 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 [5-7].

For low bit-rate video compression, 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 sufficient 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 [8]. 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 a previous work [11], the authors had proposed a new buffer control algorithm for motion-compensated hybrid DPCM/DCT coding. The algorithm is based on the use of bit allocation algorithm to determine the quantization scale factors in such coder to meet a given target bit rate. The salient features of the 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. Simulation results showed that, using the proposed algorithm, the H.261 coder can achieve a higher PSNR and better visual quality than codec using traditional buffer control algorithm. In this paper, we apply this buffer control algorithm to a modified version of the H.263 algorithm for very low bit-rate video coding. Comparing the performance of

the modified H.263 codec with the TMN5 model also shows that better visual quality can be obtained at comparable PSNR values. The layout of the paper is as follows: In Section II, we shall briefly summarize the structure of the modified H.263 coder. The application of the bit allocation to the modified H.263 algorithm is discussed in Sections III. In Section IV, the simulation results are presented.

II. THE MODIFIED H.263 ALGORITHM

The basic configuration of the H.263 video coding algorithm [4] is based on the H.261 Recommendation. There are five standardised picture formats: sub-QCIF, QCIF, CIF, 4CIF and 16CIF. Unlike H.261, half pixel precision is used for motion compensation. In addition to the basic video source coding algorithm, four negotiable coding options are included for improved performance: Unrestricted Motion Vectors, Syntax-based Arithmetic Coding, Advanced Prediction, and P-B Frames. These options can be used together or separately.

Fig. 1 shows the generalized form of the H.263 source coder. The main elements are prediction, block transformation and quantization. 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), and Block Layers. The input frame is partitioned into macroblocks consisting of one luminance block of (16×16) pixels and two chrominance blocks of (8×8) pixels. The prediction is inter-picture and may be augmented by motion compensation (MC) (optional in the encoder) and a spatial filter.

For motion-compensation, the luminance component in each macroblock is used to predict the motion of the whole macroblock. The prediction error (INTER mode) or the input picture (INTRA mode) will then undergo the two dimensional (8×8) DCT followed by quantization and zigzag scanning. The run-length and quantization levels are jointly entropy coded. 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. Each of the other 31 quantizers uses equally spaced reconstruction levels with a central dead-zone around zero and with a step size of an even value in the range 2 to 62 [4]. The output bit-rate can be controlled by varying the quantization scale factors of the macroblocks.
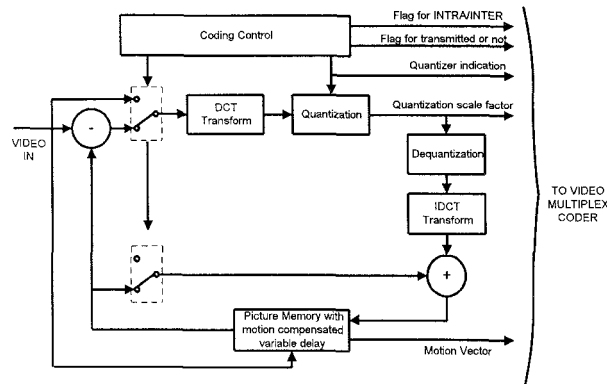

Figure 1: H.263 Source Coder.

A number of optimization has also been done in H.263 bit packing to make it more efficient as compared with H.261. For example, the End-Of-Block (EOB) is eliminated by specifying, in the VLC table, whether the transform coefficient is the last coefficient in the given block. Also, in H.263, the quantizer step size between consecutive macroblocks are constrained to reduce the number of bits needed to specify the quantizer scale. In fact, adjacent quantizer levels can only differ by ±2 (DQUANT). This is quite different from H.261 where the quantizer scale for each macroblock can take on any of the 31 different values.

The buffer control problem for H.263 is considerably more complicated than the H.261 [11]. This is due to the constraint in the quantizer, the highly coupled advanced prediction, and the use of PB-frames in the coder. To demonstrate the usefulness of the proposed buffer control algorithm in low bit-rate video coding, we modified the H.263 algorithm so that the quantizer constraints are removed. Also, we shall consider the default H.263 coder where the options are turned off. After performing the bit allocation, we allow the quantization scale factors to vary from 1 to 31 as in the H.261 algorithm. To save the number of bits to represent these scale factors, we modified the VLC codes in the macroblock type and coded block pattern to distinguish whether the difference in the quantizer scale factor is within the limit of ±2. If it is true, then we send the appropriate code for the macroblock type and coded block pattern together with the differential value to the receiver. Otherwise, we send the corresponding code together with the five bits quantization scale factor to specify one of the 31 quantizer that is going to be used in that block.

III. BIT ALLOCATION ALGORITHM

Consider encoding the residual image after motion compensation that consists of G GOB and N

macroblocks at a target bit-rate of R bits per frame using the modified H.263 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:

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

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 number of bits generated in coding the DCT coefficients of the k-th MB with quantization scale factor Q_k, and H^{Pict}, H^{GOB} and H_k^{MB} are the number of bits generated for the picture header, the GOB header and the k-th MB header, respectively. Information such as macroblock type & coded block pattern, quantizer information, and motion vector data is included in H_k^{MB}.

The minimization problem of (1) is very similar to the bit allocation problem [9,10]. 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.

We first compute the efficiency as follows:-

$$\lambda_k = \max_q \frac{-\Delta D|_{Q_k \to q}}{\Delta B|_{Q_k \to q}} \qquad (4)$$

for all values of k and then determine the blocks with the maximum efficiency:

$$\lambda = \max_k \lambda_k. \qquad (5)$$

$\Delta D|_{Q_k \to q}$ and $\Delta B|_{Q_k \to q}$ 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

$$\Delta D|_{Q_k \to q} = D_k(q) - D_k(Q_k) \qquad (6)$$

and

$$\Delta B|_{Q_k \to q} = [H_k^{MB}(q) + C_k(q)] - [H_k^{MB}(Q_k) + C_k(Q_k)] \qquad (7)$$

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

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.

IV. EXPERIMENTAL RESULTS

Computer simulation was done to evaluate the performance of the buffer control algorithm on the modified H.263 coder. All the four options in the H.263 were turned off. The Miss America (1-150) and Claire (340-489) sequences at QCIF format and 25 frames/sec were used for testing. The TMN5 algorithm and the proposed algorithm are encoded at 12.5 frames/sec with the same bit-rate. In the TMN5 algorithm [6] with buffer control, the codec will skip the current frame if it encounters a buffer overflow. Since the first picture is coded as an intra picture, large number of bits will be required. The first four frames after the intra frame were skipped by the codec. It was found that three more frames were skipped due to buffer overflow. For the proposed algorithm, the same number of bits allocated to each frame was chosen such that the averaged bit rate of the two algorithms were nearly identical. Only the first four frames were skipped and all the remaining 70 frames were encoded. Figure 2 and 3 show the PSNR comparison of the TMN5 algorithm and the proposed algorithm for the Miss America and Claire sequences. Table 1 shows the Mean PSNR of the TMN algorithm and the proposed algorithm. The PSNR value of the proposed coder is close to that of the TMN5 algorithm. A Viewstore 6000 real-time playback system with a 21 inch EIZO color monitor are used for comparing the visual quality between the sequences. We found that the proposed algorithm has a better visual quality, in addition to more steady frame rate, than the TMN5 algorithm. This is largely due to the use of the new buffer control algorithm and the relaxation of the constraints in the quantizer scale factor in the former

algorithm. This effectively distribute the error over the whole image and significantly reduced blocking artifacts that results from uneven distribution of quantization error over adjacent macroblocks.

Mean PSNR	Miss Amer. (TMN5)	Miss Amer. (proposed algorithm)	Claire (TMN5)	Claire (proposed algorithm)
Y	39.08	39.14	38.77	38.83
Cb	39.08	38.99	39.54	39.48
Cr	38.61	38.55	42.33	42.24

Table 1. Mean PSNR comparison of the TMN5 algorithm and the proposed algorithm.
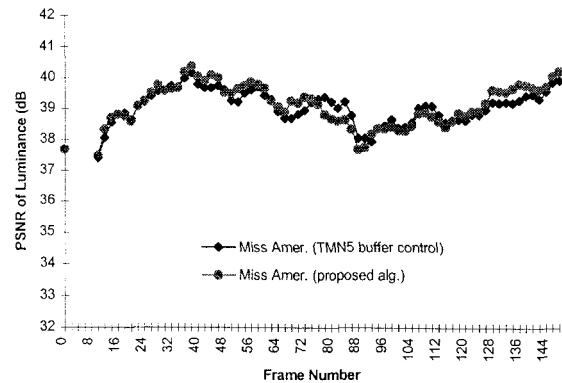
V. CONCLUSION

We have applied a new buffer control algorithm [11] to the modified H.263 algorithm for very low bit-rate video coding. Simulation results show that the modified H.263 codec has a better visual quality than the TMN5 model at comparable PSNR values. Since the rate-constrained adaptive quantization technique allows the user to have precise control of the bit-rate in the coder, it will be very useful in low bit-rate video coding in networking and storage applications. Further investigation will concentrate on the modification of the bit allocation to the more general coupled problem found in H.263 with all the options being turned on. Research in these area are now underway.

Figure 2: PSNR comparison of the TMN5 and the proposed algorithm for the Miss America sequence (24.4kbps).

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] ITU Telecommunication Standardization Sector LBC - 95. Study Group 15, Working Party 15/1. Expert's Group on Very Low Bitrate Visual Telephony. VIDEO CODEC TEST MODEL, TMN5.
[7] 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.
[8] M. T. Sun, K. Rijkse, D Schinkel and A Meijboom, "Coding and Interworking for Videotelephony," *Proc. ISCAS-93*, pp.20-23, 1993.
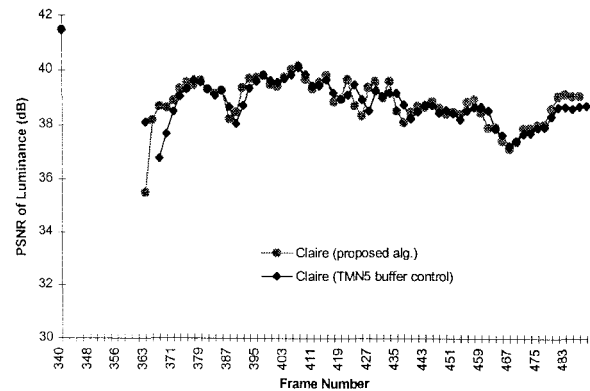
Figure 3: PSNR comparison of the TMN5 and the proposed algorithm for the Claire sequence (27.5kbps).

[9] Y. Shoham and A. Gersho, "Efficient Bit Allocation for an Arbitrary Set of Quantizers," *IEEE Trans. Acoust. Speech Signal Process.*, vol. ASSP-36, pp.1445-1453, September 1988.
[10] S. W. Wu and A. Gersho, "Rate-constrained picture-adaptive quantization for JPEG baseline coders," *Proc. IEEE ICASSP-93*, vol. V, pp. 390-392.
[11] K. T. Ng, S. C. Chan and T. S. Ng, "Buffer control algorithm for low bit-rate video compression," *Proc. IEEE Int. Conf. on Image Proc., ICIP'96*, Lausanne, 1996, Vol I, pp. 685-688.