 IEEE Transactions on Consumer Electronics, Vol. 42, No. 3, AUGUST 1996

JOINT RATE CONTROL FOR MULTI-PROGRAM VIDEO CODING

Limin Wang and André Vincent
Communications Research Centre
3701 Carling Avenue
P.O. Box 11490, Station H
Ottawa, Canada, K2H 8S2

Abstract

This paper presents a joint coding scheme for multi-program video transmission. By exploiting the statistical variations of program content, the joint coding achieves a more optimal bit allocation, which results in a better and more uniform picture quality. The bit allocation strategy assigns the same quantization parameter to all the programs, and rate control adjusts this quantization parameter from frame to frame so that the aggregate bit rate for the programs is within a given bit budget.

1 Introduction

With recent advances in video compression (e.g. MPEG-2 [1]), digital television programs can be compressed by a great degree while still maintaining an acceptable picture quality. For example, digital NTSC programs can be compressed to a bit rate of 4-8 Mbits/s. On the other hand, digital transmission technology can support a payload of up to 19 Mbits/s in 6 MHz terrestrial broadcast channels and 27/38 Mbits for Cable TV channels. Hence, it is now possible to deliver several digitally compressed television programs in the same bandwidth presently occupied by a single analog television channel.

Fig. 1, shows a multi-program video transmission environment in which several programs are compressed, and the resulting bit streams are then multiplexed at the transport layer and transmitted over the same channel. Typically, programs are encoded independently, each at a fixed bit rate. The channel capacity can be shared equally among programs, if there is no a priori knowledge of program content, or it may be allocated based on the type of programs (e.g. news, sports, or movie). There are however two main problems associated with independent coding at constant rate (CBR). First, program content may vary significantly from program to program. As a result, equal sharing of channel capacity will most likely lead to a large variation of picture quality among the programs. Second, the scene content may also change within a program as a function of time. Hence, if the bit rate for a program is set sufficiently high to provide an adequate picture quality for the most demanding segments, the channel capacity is effectively wasted during the less challenging periods.

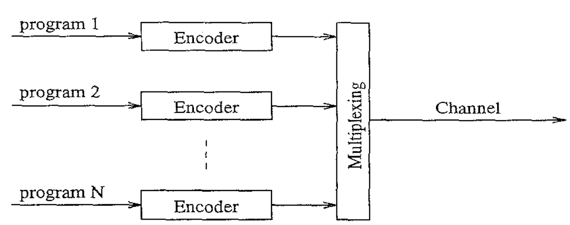


Figure 1: Multi-program environment: several programs are compressed and transmitted over a single channel.

This paper presents a joint coding scheme for multi-program video transmission which addresses these problems with independent coding by dynamically distributing the channel capacity among programs according to their complexities. With this scheme, more bits are allocated to the more demanding (from a coding point of view) programs

Manuscript received June 10, 1996

than to the easier programs, thereby resulting in a relatively uniform picture quality among programs. The program-dependent bit allocation is achieved by using the same quantization parameter for all the programs at each frame. The use of the same quantization parameter also implies a similar quality level for all the programs. To maintain the aggregate bit rate at a constant value, the joint coding implements a joint rate control which controls the sum of bit rates for the programs, instead of the individual bit rates.

In the following section, we present the joint coding with joint rate control for multi-program transmission. In section 3, we report the simulation results.

2 Joint Coding and Rate Control

In this paper, the programs are coded using the MPEG-2 video coding standard which inherently generates a variable rate bit stream [1]. If the variable rate bit stream needs to be transmitted over a fixed rate channel, a channel buffer is usually required to smooth out the bit stream. To prevent the buffer from overflowing and underflowing, rate control has to be implemented. Typically, the information on buffer occupancy is fedback to the encoder and used to adjust the coding parameters, such as quantization. For example, if the buffer is getting full, the quantization parameter is increased, which reduces the bit rate. On the other hand, if the buffer becomes empty, the quantization parameter is decreased, which increases the bit rate. Rate control plays a very important role in determining the final picture quality.

In multi-program video transmission, the programs may be coded independently, each at a constant rate. Each program will therefore need a separate rate control to maintain its bit rate at that constant value, as shown in Fig. 2. The proposed joint coding however implements a joint rate control which controls the sum of the bit rates for the programs, instead of each individual bit rate, as shown in Fig. 3. In other words, the bit rate for each individual program is allowed to vary even though the aggregate bit rate is maintained at a constant value.

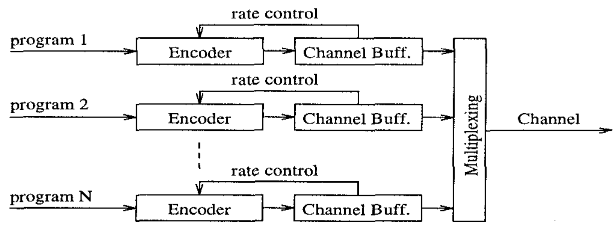


Figure 2: Independent coding: the programs are coded independently; each has a separate rate control.

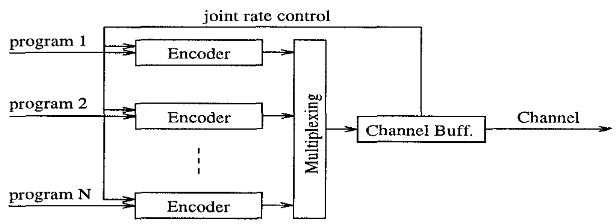


Figure 3: Joint coding: the programs are coded jointly; a joint rate control is implemented.

2.1 Quantization Parameter

In MPEG-2 video coding standard, the quantization parameter is allowed to change from macroblock to macroblock [1]. The bit rate control scheme described in Test Model 5 [2] actually adjusts the quantization parameter at each macroblock according to the channel buffer occupancy. As an example, Fig. 4 shows the quantization parameter versus macroblock number for three picture types I (intra), P (predictive) and B (bi-directional predictive), derived from the sequence *Mobile* using TM5 rate control. As seen, the quantization parameter varies across each picture in order to accommodate the variation of channel buffer occupancy. The variation in quantization parameter results in nonuniform distortion over the picture. For example, two identical macroblocks may be assigned two different quantization parameters due to different buffer occupancy at different times, and therefore would be rendered at different quality levels. To maintain a relatively uniform quality over each picture, the same quantization parameter should be applied to all the macroblocks in each frame [3]. This guarantees that all the macroblocks in a frame are treated equally.

This concept is now extended to multi-program video coding. In the proposed joint coding, the same quantization parameter is used for all the programs at each frame. In order to accommodate local activity variations, the quantization parameter can be further modulated by the local activity, as in TM5 [2]. Clearly, since the same quantization parameter is applied, the more complex programs will generate more bits and hence, use more of the channel capacity than the less complex programs. Furthermore, the use of the same quantization parameter guarantees the same upper bound of the coding distortion for all the programs [4].

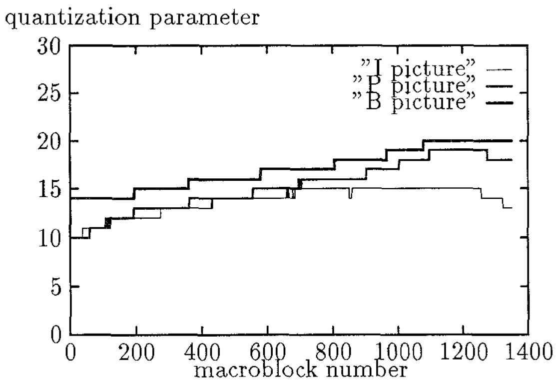


Figure 4: Quantization parameter versus macroblock number for frames of type I, P and B from the sequence *Mobile* by using TM5 rate control [2].

2.2 Rate Control

We now describe how to select the quantization parameter. Assume there are N programs in a multi-program transmission system. We define a super frame as a collection of N frames, one from each of the N programs at the same time instant. For the sake of simplification, it is further assumed that the N frames in a super frame are of the same picture type. Rate control is achieved by adjusting the quantization parameter from frame to frame at the super frame level. It consists of two steps. The first step is to set a target number of bits for a super frame before the super frame is coded. This step is similar to the 'Picture Target Setting' described in Test Model 5 (TM5) [2], except that the 'global complexity' is now measured over a super frame. Let $Q_{i,p,b}$ be the quantization parameter applied to a super frame of picture type I, P or B and $S_{i,p,b}$ be the number of bits generated from the super frame using $Q_{i,p,b}$. Here, the subscripts i,p,b corresponds to the picture type I, P or B. After encoding a super frame of picture type I, P, or B, the respective "global complexity" is updated as

$$X_{i,p,b} = S_{i,p,b} Q_{i,p,b} \tag{1}$$

The target number of bits for the next super frame of picture type I, P or B is then computed as

$$\begin{cases} T_i &= \max\{\frac{R}{1+\frac{N_p X_p}{K_p X_i}+\frac{N_b X_b}{K_b X_i}}, \frac{channel\_rate}{8 \times picture\_rate}\} \\ T_p &= \max\{\frac{R}{N_p+\frac{N_b K_p X_b}{K_b X_p}}, \frac{channel\_rate}{8 \times picture\_rate}\} \\ T_b &= \max\{\frac{R}{N_b+\frac{N_p K_b X_p}{K_p X_b}}, \frac{channel\_rate}{8 \times picture\_rate}\} \end{cases} \tag{2}$$

where K_p and K_b are constants, R is the remaining number of bits assigned to the current group of pictures (GoP) [1], N_p and N_b are the number of P and B pictures remaining in the current GoP, and *channel_rate* and *picture_rate* are the channel rate in bits/sec and the frame rate in frames/sec., respectively. Note that before encoding each GoP, R is updated as $R = R + R_G$, where

$$R_G = \frac{N_G \times channel\_rate}{picture\_rate} \tag{3}$$

and N_G is the total number of frames in a GoP.

Once the target number of bits for a super frame, $T_{i,p,b}$, is set, the next step is to select a quantization parameter for the super frame. The joint rate control algorithm chooses the one among all the possible quantization parameters that minimizes the difference between the target and the actual number of bits generated for the super frame. A binary tree search algorithm described in [3] is used to determine such a quantization parameter for a super frame given a target number of bits, $T_{i,p,b}$. Assume l bits are assigned to represent a quantization parameter. There can be $L = 2^l$ possible quantization parameters. The L quantization parameters are ordered according to their values, i.e.,

$$Q(0) < Q(1) < \ldots < Q(L-1) \tag{4}$$

The binary tree search starts by trying the one in the central position of the set of quantization parameters, i.e.,

$$q_0 = Q(L/2) \tag{5}$$

to the super frame. If the number of bits generated from the super frame is smaller than the target number of bits, $T_{i,p,b}$, the quantization parameter, q_0, is deemed to be too large. So, the search is moved to the lower half of the set. On the other hand, if the number of bits is greater than the target number of bits, the search is moved to the higher half of the set. In either case, the new set is only about half the size of the old one. Again, the quantization parameter located in the centre of the new set is applied to the super frame and the resulting number of bits is compared with the target number of bits, $T_{i,p,b}$. The procedure continues until there is only one quantization parameter left in the ordered set. Clearly, the binary search requires a fixed number of trials, l. Each trial involves only the quantization and variable length coding operations. The quantization parameter used at each trial k can be written in a more general form as

$$q_k = Q(L/2 \cdot \sum_{i=0}^{k} \pm \frac{1}{2^i}) \qquad k = 0, 1, \ldots, L-1 \tag{6}$$

where the $\pm$ sign depends upon whether the search is moving to the upper or lower half of the ordered set at stage i ($\leq k$). Among all the l trials of quantization parameters, the one that results in a bit rate closest to the target bit rate is selected. MPEG assigns $l = 5$ bits to each quantization parameter [1]. The quantization parameter for a super frame can therefore be located by five binary searches.

3 Simulations

Simulations were carried out using joint coding for six programs at a total bit rate of 18 Mbits/s. The six programs consisted of three video sequences and three film sequences, each was 5 seconds in length. All the programs had a spatial resolution of 720x480 and a frame rate of 30 frames/s, interlaced with a color sampling ratio of 4:2:2. For comparative purpose, the six programs are also coded independently, each at a fixed rate of 3 Mbits/s.

Table 1 presents the bit rates for the six programs using both independent and joint coding. As seen, by joint coding, the six programs were coded with quite different number of bits depending upon their complexities. For example, the sequence *Mobile* was coded with more than twice as many bits as *Film2*.

Table 1. Bit Rate in Mbits/sec.

Coding	Flower	Mobile	Tennis	Film1	Film2	Film3
Ind.	3.00	3.00	3.00	3.00	3.00	3.00
Joint	3.84	4.17	2.80	2.90	1.67	2.62

Fig. 5 illustrates the buffer occupancy and the difference between the target and the actual number of bits at each super frame. Here, the buffer occupancy at each super frame is defined as

$$R_{buffer\_occupancy} = \sum_i (R_i - \frac{channel\_rate}{picture\_rate}) \tag{7}$$

where R_i is the number of bits generated from super frame i. Note that in joint coding, there is only one channel buffer (Fig. 3). The buffer occupancy varies regularly according to the picture type. The maximum buffer occupancy in this example is 1.4 Mbits, less than 10% of the channel rate of 18 Mbits/s. In practice, the channel buffer has to be able to cover the variation of buffer occupancy. The difference between the target and the actual number of bits is seen to fluctuate around zero within a very small margin. The average absolute difference is only 23478 per super frame, or 3913 bits per program frame.

Table 2 shows the maximum buffer occupancies for the six programs using independent coding. To cover the variations of these buffer occupancies, the total size of the six buffers will be at least 1.875 Mbits which is 34% larger than 1.4 Mbits required for the joint coding. Table 3 presents the average absolute differences between the target and the actual number of bits for the six programs using independent coding. These differences are much greater

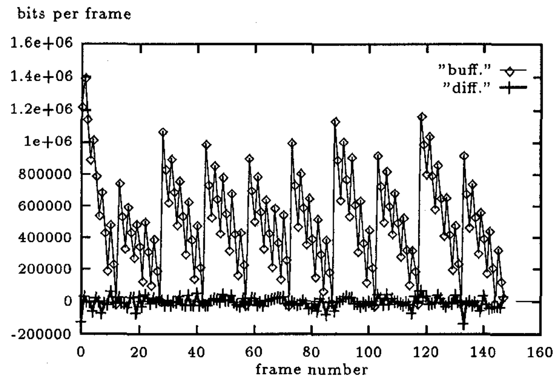


Figure 5: The buffer occupancy and the difference between the target and the actual number of bits at each super frame.

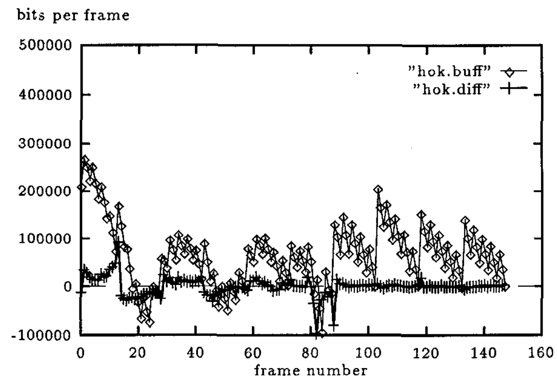


Figure 6: The buffer occupancy and the difference between the target and the actual number of bits at each frame for *Film3*.

than the 3913 bits difference obtained with joint coding. As an example, Fig. 6 shows the buffer occupancy and the bit difference for *Film3* using independent coding where the buffer occupancy is less regular and the fluctuation of bit difference is greater.

Table 2. Maximum Buffer Occupancy (bits)

Flower	Mobile	Tennis	Film1	Film2	Film3
346701	417456	349483	254593	238764	268467

Table 3. Ave. Absolute Difference (bits)

Flower	Mobile	Tennis	Film1	Film2	Film3
7585	9477	12873	7965	8431	10907

Table 4 lists the Peak Signal-to-Noise Ratio (PSNR) for the six programs using both independent and joint coding. It is seen that joint coding improves the quality of the more complex programs. For example, the PSNRs for the sequences *Flower* and *Mobile* were increased by 2.45 dB and 2.92 dB, respectively. The average PSNR over the six programs was actually increased by about 1 dB per program by using joint coding (Table 4). From Table 4, it should also be noted that the quality variation among the six programs for joint coding was within a much smaller margin than for independent coding.

Table 5 gives the variances of the PSNR for the six programs. For the sequences with substantial temporal variability in PSNR, joint coding significantly reduces this variability. For example, joint coding gives a much smaller variance in PSNR for *Tennis, Film1, Film2* and *Film3*. For sequences where the variability was already low with independent coding, joint coding had a relatively smaller impact on the variability of PSNR.

Table 4. PSNR (dB)

Coding	Flower	Mobile	Tennis	Film1	Film2	Film3
Ind.	28.32	26.70	30.42	32.67	36.35	33.42
Joint	30.77	29.62	31.25	33.47	34.83	33.74

Table 5. Variance of PSNR (dB)

Coding	Flower	Mobile	Tennis	Film1	Film2	Film3
Ind.	0.41	0.53	8.76	3.22	2.91	5.76
Joint	0.70	1.02	4.97	1.19	0.67	1.79

Fig. 7 and 8 show the PSNR versus the frame number for three film sequences using independent and joint coding, respectively. Clearly, joint coding results in a relatively more uniform picture quality among programs and within a program.

Finally, the less demanding film materials, such as *Film2*, were allocated many less bits and therefore coded at a slightly lower PSNR level. However, informal observation showed that their subjective quality was not decreased appreciably. On the other hand, by using additional bits saved from coding the less demanding programs, the more complex video programs showed a significant subjective quality improvement, as judged using informal viewing.

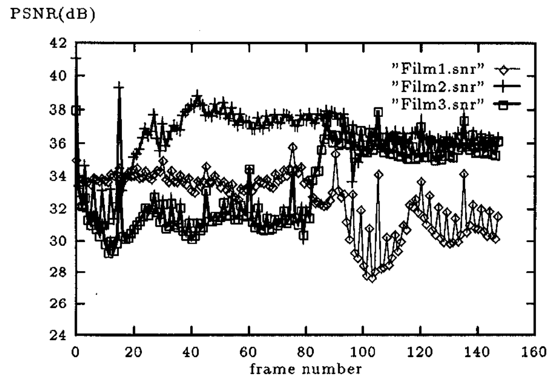


Figure 7: The PSNR for three film sequences using independent coding.

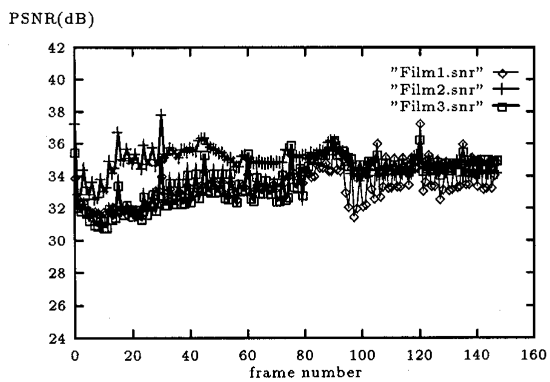


Figure 8: The PSNR for three film sequences using joint coding.

4 Conclusions

This paper presents a joint coding scheme for multi-program transmission which dynamically distributes the channel capacity over the programs according to their complexities; more complex programs are allocated more bits and less complex programs less bits. The program-depended bit allocation is achieved by using the same quantization parameter for all the programs at each frame. The joint coding implements a joint rate control which maintains the aggregate bit rate for the programs at a constant value, but allows the bit rate for each individual program to vary. The simulation results demonstrate that joint coding results in a relatively uniform picture quality among programs as well as within a program.

5 References

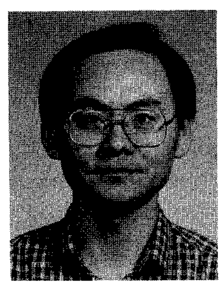
1. ISO/IEC (MPEG-2), "Generic coding of moving pictures and associated audio", March, 1994.

2. ISO/IEC (MPEG-2), "Test Model 5", April, 1993.

3. Limin Wang, "Rate control for MPEG video coding", *Proc. SPIE. on Visual Communications and Image Processing*, pp. 53-64, Taiwan, May, 1995.

4. Limin Wang and Andre Vincent, "Joint coding for multi-program transmission", *IEEE Conf. on Image Processing*, Lausanne, Sept. 1996.

Biographies



Limin Wang received his Ph.D. degree in Electrical Engineering from University of Ottawa in 1988. From 1989-1990, he worked for Medical Communications Research Centre at University of Ottawa. From 1990-1991, he was with MPR Teltech as Member of Technical Staff. In 1991, he joined Communications Research Centre as Research Scientist. His current research interests are in the general area of image processing and visual communications.

André Vincent received his B.Sc.A in Electrical Engineering from L'Ecole Polytechnique de Montréal in 1975. Between 1975 and 1979, he worked on the design and development of mobile radio communications systems. He joined the Communications Research Centre in 1979, where he conducted research in the areas of Teletext, data transmission, television channel characterization and digital mobile radio. Since 1986, he is manager of the Video Systems group, and is involved in research in video processing.