JOINT MPEG-2 CODING FOR MULTI-PROGRAM BROADCASTING OF PER-RECORDED VIDEO

Irene Koo[†], Panos Nasiopoulos[†], Rabab Ward[†]*

[†]The University of British Columbia, Vancouver, Canada
[*]Daikin US Comtec, Vancouver, Canada

ABSTRACT

We developed a cost-effective operational system suitable for digital TV, video on demand, and high definition TV broadcast over satellite networks with limited bandwidth. This MPEG-2 based system is easy to implement and allows the joint video coding of multiple video programs. Compared to present broadcast operation and for the same level of picture quality, our system greatly increases the number of video streams transmitted in each channel. As a result, either a large number of transponders can be freed up to carry real-time broadcasting or the level of the transmitted picture quality can be significantly increased. By switching from tape storage to video server technology, the need for numerous (expensive) playback VTR systems at the headend is eliminated. In addition, the majority of the complete MPEG-2 encoders are replaced by much less complex MPEG-2 transcoders. All this means significant savings for the broadcast stations. In addition to the gain in bandwidth and the reduction in cost, our system speeds up the encoding process by six folds.

1. INTRODUCTION

Advancements in multimedia technology and digital communications enabled the broadcasting of multiple programs in a channel, which was used to transmit a single analog program. The trade-off between picture quality and bandwidth is an extremely important issue in analog and digital broadcasting. The level of picture quality and the efficiency in bandwidth usage in a multi-program environment are strongly influenced by:

- the quality of the video encoder,
- the choice of the multiplexing method and the ability to optimally distribute bandwidth to the video streams based on content and complexity,
- the available bandwidth, and
- the number of programs to be transmitted simultaneously.

A typical digital satellite broadcast system consists of three parts: uplink earth stations, distribution satellites, and downlink earth stations. A satellite has the bandwidth to support multiple transmission channels, and each transmission channel can carry multiple digital video streams. The uplink earth station is where the encoding and multiplexing of video streams take place. Besides news, sports, and some other live shows, the majority of the broadcast programs such as movies, commercials, and music videos, are pre-recorded. The latter also applies to all video on demand materials. The video encoding process consists of playing back pre-recorded video programs, feeding each play-back video stream and each real-time video program (e.g., sports, news, etc.) into an encoder, and setting the appropriate bit-rate for each video stream according to a certain quality of service (QoS). The bit-rate assigned to a video stream does not necessarily offer high picture quality but rather meets a minimum guaranteed quality. The assigned bit-rate stays constant for a long period of time and thus does not reflect the complexity (activity, amount of motion) of the video stream at a given moment. Because the bit-rate is constant, the quality varies dependently on the complexity of the specific video program as well as the complexities of the coexisting video programs in the same channel. Previous works have already proven that such constant bit-rate (CBR) coding is inherently inefficient in terms of bandwidth usage [1, 2]. In order to use the bandwidth effectively, a digital broadcast system should take advantage of the variable bit-rate nature of the compressed video streams and the capability of a channel carrying multiple compressed video streams. Non-statistical bit-rate assignments for real-time video programs cannot be avoided since the video characteristics of the video programs cannot be foreseen. However, a broadcast system performance could be significantly improved if it takes advantage of the fact that most materials are pre-recorded. In this case, the video characteristics can be analyzed beforehand in order to have bit-rate assignments that closely relate to the video characteristics.

Some work has been done on joint bit-rate control, which is one of the techniques used in handling bit-rate assignment for multiple video programs [1, 2, 3]. The idea of joint bit-rate control is to allow the bit-rate of each individual video program to vary according to some video characteristic such as picture complexity (activity, motion), while the sum of the bit-rates of all the video programs in one channel remains constant. However, due to the conceptual nature of this research, it never found its application in commercial broadcast systems.

We developed a joint bit-rate coding system, which can be easily implemented for commercial practical use. Our system is a two-stage digital video encoding solution to multi-program transmission. The first stage is composed of an MPEG-2 encoder(s) and a video server (Figure 1). It performs multi-channel video encoding in two stages. During the first stage, the system performs high bit-rate – high quality video coding on the video sources and gathers all the relevant statistics on the complexities of the video streams. The encoded video along with the gathered statistics are stored in the video server for later use. When it is time to transmit the video signals, the second stage of our system becomes operational. The second stage is composed of a joint bit-rate controller and a set of transcoders (Figure 1). The controller performs the bit allocation and the transcoders convert the video streams, which were previously encoded at high quality levels, to video streams with bit-rates determined by the controller. Bit allocation in the system is performed on a GOP basis, and the number of bits allocated to each source depends upon the relative GOP complexities of the video streams. This system will be shown to guarantee the efficient use of the available bandwidth. Compared to existing applications and for the same level of contracted QoS, our system significantly increases the number of transmitted video streams per channel. In addition, this implementation offers a very cost-effective solution: 1) since the video programs could be encoded at any earlier time, the number of playback systems as well as the number of complete MPEG-2 encoders needed at the broadcasting headend are greatly reduced and 2) since more programs are squeezed into a channel, many transponders are freed up for real-time broadcasting use.

The rest of this paper is organized as follows: in Section 2, we discuss the first stage encoding in details. In Section 3, we describe our joint bit-rate control process. In Section 4, we introduce the concept of transcoding and describe how transcoding works in our solution to give the desired video bit-rates. In Section 5, we present the simulation results.

2. FIRST STAGE ENCODER

Video encoding is the first stage of the two-stage joint bit-rate coding process. The aim here is to encode each video stream in an off-line fashion for later use. All operations of the MPEG-2 encoding process, which require lengthy computations, are carried here and will not be repeated. At this stage, one of the main objectives is to ensure a trade-off between storage cost for the video streams and overall performance of our system. Due to the fact that a re-quantization process always introduces some deterioration in picture quality (when compared with a one-pass-one-step quantization process at the same bit-rate), our goal was to determine a maximum fixed quantizer parameter value, which minimizes this quality difference. Performance evaluations have shown that for quantizer parameter values less than or equal to 7, the difference in picture quality between the one-step quantization and the re-quantization methods is kept at acceptable levels. This difference increases significantly for values greater than 7. Therefore, we chose 7 to be the maximum fixed quantizer parameter value, which guarantees a trade-off between the cost of storage and picture quality performance for our system. Another important process during the first stage is the gathering of statistics about the complexity levels of the pictures in each video source. This information is recorded in data files, which we call complexity files. More specifically, ***picture complexity*** and ***GOP complexity*** are recorded in a complexity file. Picture complexity is defined as the number of bits used in coding a picture during the first stage. GOP complexity is defined as the sum of the complexities of the pictures in a GOP, which is also the number of bits used in coding all pictures of a GOP. Both the high quality encoded video streams and their corresponding complexity files are stored in the video server.

3. JOINT BIT-RATE CONTROLLER

The second stage of our encoding system consists of a joint bit-rate controller that oversees the bit allocation operations of the video streams to be multiplexed in a single channel. This controller receives the encoded video streams and their corresponding complexity files, which were generated during the first encoding stage. Based on this information and the given channel bandwidth, the controller determines a GOP bit target for each video source and sends these values to a set of transcoders. The controller assumes that all videos have the same N and *frame_rate*, where N is the number of frames in a GOP. For a given bandwidth, the aim of our joint bit-rate controller is to offer appropriate bit allocation and consistent picture quality for each GOP of the video streams.

The joint bit-rate controller allocates a fraction of the channel bandwidth to the current GOP of each video stream. The fraction assigned to each video stream is proportional to the relative GOP complexity of the video streams. That is,

$$GOP\_target_i = \frac{\sqrt{GOP\_Complexity_i}}{\sum_{k=1}^{Np}\sqrt{GOP\_Complexity_k}} * aggregrate\_GOP\_target \qquad (1)$$

where, N_p is the number of video streams to be multiplexed; *aggregrate_GOP_target* is number of bits available to transmit the GOP's of all video streams. *aggregrate_GOP_target* is derived from the available channel bandwidth by

$$aggregrate\_GOP\_target = channel\_bandwidth / (\frac{\text{\# of GOP's}}{\text{sec}}) \qquad (2)$$

Our experimental performance evaluations have shown that assigning GOP bit targets according to the square root function compresses the complexity ratio between the high-complexity streams and the low-complexity streams into a reasonable range of values and improves the bit allocation between different sources. This is important, especially when the complexities of the video streams reach extreme high levels. In this case, using the ratio $\frac{GOP\_Complexity}{\sum_{k=1}^{Np} GOP\_Complexity_k}$ (and not their square roots) results in allocating a large proportion of the *aggregrate_GOP_target* to the high-complexity streams, leaving insufficient number of bits for the low-complexity streams.

Simulations have shown that the GOP bit target assignments of each video stream closely match the GOP complexities of the video stream. In addition, each individual video sequence is assigned with variable GOP bit targets, and the sum of the assigned GOP bit targets or the *aggregate_GOP_target* at each instance is constant.

Our joint bit-rate controller uses the GOP complexities previously recorded in the complexity files to determine GOP bit targets for each video stream. Then, this information is sent to the transcoders. Each transcoder, upon receiving its corresponding GOP bit target, distributes the target bits amongst the pictures in the GOP. The joint bit-rate controller can also help the transcoders in optimizing their performance by providing them with picture bit targets derived from the picture complexities previously recorded in complexity files.

By using the picture complexities recorded in complexity files and the determined GOP bit target, the joint bit-rate controller can actually perform the picture bit allocation for the transcoders. The picture bit target of a picture is given as follows:

$$Picture\_Target_i = \frac{Picture\_Complexity_i}{\sum_{j=1}^{N} Picture\_Complexity_j} * GOP\_Target \qquad (3)$$

where, N is the number of pictures in a GOP; *Picture_Complexity* is the picture complexity of each picture recorded in the complexity file; *GOP_Target* is the GOP bit target determined by the joint bit-rate controller.

By sending the picture bit targets to the transcoders, we eliminate the need for each transcoder to perform picture bit allocation. As a result, the complexity of a transcoder can be reduced and the encoding time of the transcoding stage can be shortened.

4. TRANSCODERS

Transcoding is the final step of the joint bit-rate coding process. It is performed immediately before multiplexing and transmission of the video streams. The first half of the transcoding process partially decodes the video stream up to the stage where all DCT coefficients of macroblocks are accessed. The latter half of the transcoding process re-quantizes the DCT coefficients and assembles the video stream back together. Thus, our transcoding process involves variable length decoding, inverse scanning, inverse quantization, re-quantization, forward scanning, and variable length encoding of the incoming video stream.

A transcoder, essentially, consists of a cascaded decoder and encoder [4]. The complexity of a transcoder can range from the extreme, where it comprises a complete decoder and a complete encoder, to the simplest, where it is just a re-quantizer. Our transcoder implementation takes on the simplest approach. This is possible because the objective of our transcoding process is to compress the video stream from a high bit-rate to a lower bit-rate suitable for transmission. That is, no other reformatting such as re-sampling is involved. Since re-quantization is the sole purpose of our transcoding process, changing the picture coding types, or the coding decisions, or the re-estimation of motion vectors is not required. Therefore, all these required information, which were already obtained during the first stage, can now be used. By reusing the set of coding decisions and the set of motion vectors, we significantly reduce the transcoder’s complexity and cost and the processing time.

In Section 3, we mentioned that the joint bit-rate controller could send picture bit targets (computed from the calculated GOP bit targets and the recorded picture complexities) to the transcoders to improve the transcoders' performance. Therefore, each transcoder receives two sets of information from the joint bit-rate controllers:

1. the GOP bit target for the current GOP and
2. a set of picture bit targets for the pictures within this GOP.

The transcoder, upon receiving this information, determines the appropriate value for the quantizer scale parameter of each macroblock within each picture and re-quantizes each macroblocks using its corresponding quantizer parameter value. The appropriate quantizer parameter for each macroblock is determined using adaptive quantization [5].

5. EXPERIMENTAL RESULTS AND DISCUSSIONS

An experiment was carried out to evaluate the performance of the joint bit-rate controller. A set of five video sequences, which consists of highly complex scenes as well as simple scenes, was encoded at a total bandwidth of 18 Mbps. Each video sequence is 10 seconds in length and has a spatial resolution of 720x480, which is double that used in cable TV. The sequences were encoded at 30 frames/sec, with a fixed quantizer scale of 6, a GOP size of 12, and a GOP pattern IBBPBBPBBPBB using our first stage encoder. Figure 2(a) illustrates the GOP complexities of one of the video sequences. Figure 2(b) traces the GOP bit targets assigned to the same sequence. It can be seen that there is a high degree of resemblance between the GOP complexity plot and the GOP bit target plot of the sequence. Thus, the GOP bit targets determined by the joint bit-rate controller truly reflect the complexity levels of the video stream. Figure 2(c) shows the sum of the GOP bit targets for each bit allocation decision. We observe that individual video sequence is assigned with variable GOP bit targets, and the sum of the assigned GOP bit targets is constant. For the same contracted picture quality, the corresponding constant bit-rates for the five sequences should be 6.90 Mbps, 6.64 Mbps, 7.20 Mbps, 7.20 Mbps, and 6.50 Mbps. This means that only 2.6 out of the 5 streams would be transmitted using the 18 Mbps channel and CBR encoding.

6. CONCLUSIONS

In this paper, we present an efficient operating procedure for multi-program encoding. The procedure can be easily implemented for commercial use in digital video broadcast applications. The system uses a two-stage coding approach. During the first stage, the video streams are encoded off-line at any time prior to their transmission; however, it is here required that the encoding should result in high picture quality. Complexity files recording statistics of the video streams are also generated and stored. The second stage becomes operational during the actual broadcasting of the video programs. Knowing the complexities of the video streams, the necessary number of bits needed to encode each video stream are here determined and a set of transcoders executes the bit allocation decisions.

Simulation results have shown that our joint bit-rate coding system increases the number of video streams supported in a channel from 2.6 to 5. That is, a 92% increase. This is a significant improvement since a large number of the transponders can be freed up to carry real-time video programs or to provide other communication services. Switching from tape storage to video server technology constitutes another advantage of our proposed system. Playback systems (each costing over $100,000) are eliminated since video streams can be directly accessed via the video server. This also implies fewer devices to operate during the transmission time of the video programs. Also, since the first-stage video encoding of our system is an off-line process, fewer complete (also expensive) MPEG-2 encoders are required for our system. Presently, an encoder is needed for each video stream to be transmitted. Another advantage is that the simpler structure of a transcoder makes the manufacturing of the transcoder hardware significantly cheaper than the manufacturing of the encoder hardware. The last three properties translate to a lower cost of operation and of hardware. In addition to the gain in bandwidth and the reduction in cost, simulation results have also shown that our system speeds up the encoding process by six folds.

7. REFERENCES

[1] Limin Wang and André Vincent. "Joint Rate Control for Multi-Program Video Coding," *IEEE Trans. on Consumer Electronics*, vol. 42, n. 3. August 1996. p. 300-305

[2] Sanghoon Lee, Seong Hwan Jang, and Jeong Su Lee. "Dynamic Bandwidth Allocation for Multiple VBR MPEG Video Sources," *IEEE International Conference on Image Processing*, vol.1. 1994. p.268-272

[3] Gertjan Keesman. "Multi-Program Video Compression Using Joint Bit-Rate Control," *Philips Journal of Research*, vol. 50. no. ½. 1996. p.21-45

[4] P. N. Tudor and O. H. Werner. "Real-Time Transcoding of MPEG-2 Video Bit Streams," *IEE Conference Publication*, n 447. 1997. p.298-301

[5] MPEG-2 Test Model 5 (TM5). *Doc. ISO/IEC JTC1/SC29/WG11/93-400*. Test Model Editing Committee. April 1993

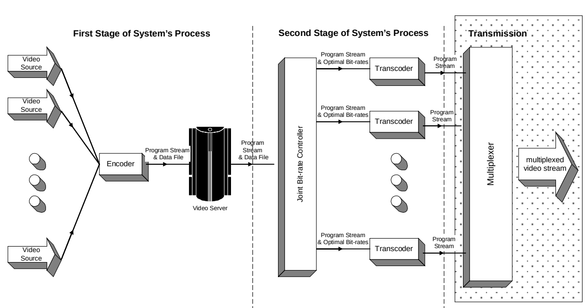
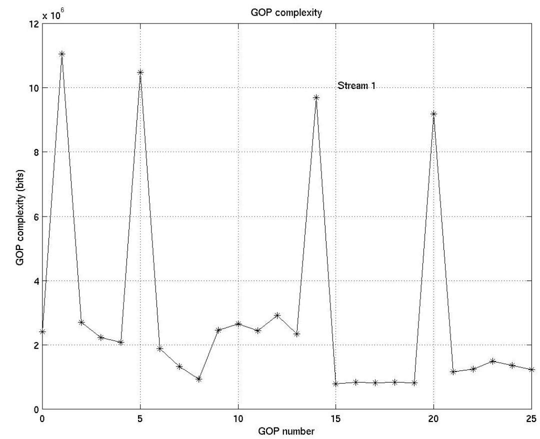
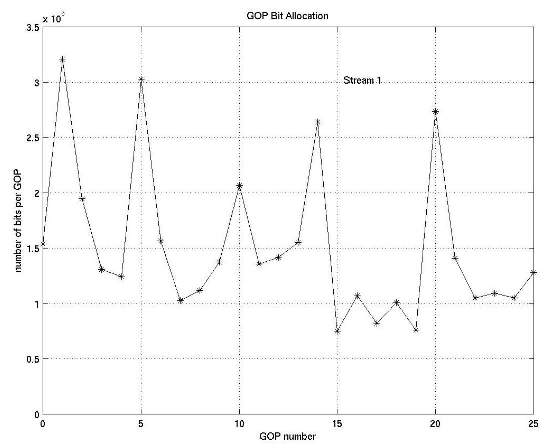


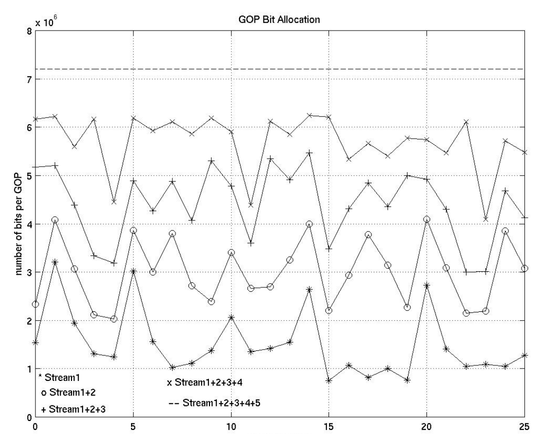
Figure 1: Block diagram of the two-stage joint bit-rate coding system.



(a) GOP complexity of SEQUENCE 1.



(b) GOP targets assigned to SEQUENCE 1.



(c) Sum of GOP bit targets assigned to all video sequences.

Figure 2: GOP bit allocation performed by the joint bit-rate controller for SEQUENCE 1, SEQUENCE 2, SEQUENCE 3, SEQUENCE 4, and SEQUCNE 5.