MODEL-BASED OPTIMAL DEPENDENT JOINT BIT ALLOCATION OF H.264/AVC STATISTICAL MULTIPLEXING

Chao Pang, Oscar C. Au, Feng Zou, Jingjing Dai

Department of Electronic and Computer Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong

ABSTRACT

In this paper, we address the dependent joint bit allocation problem in H.264/AVC statistical multiplexing. For most of the existing methods, in order to improve the overall visual quality, the bit allocation is performed upon the relative complexities of different video programs. However, due to the temporal prediction employed in H.264, the influence of current frame to the rate-distortion (R-D) performances of the future frames should be taken in to account as well. The contributions of the paper are two-fold. First, a simple but accurate inter-frame dependency model (IFDM) is introduced which can quantitatively measure the coding dependency between the current coding frame and its reference frame. Second, based on the IFDM, the dependent joint bit allocation problem is revisited, and both the frame complexity and inter-frame dependency are considered in the bit allocation process. Then, it is proved that the dependent joint bit allocation problem can actually be relaxed into a convex optimization problem which can be optimally and efficiently solved. Experimental results demonstrate that the proposed dependent bit allocation method achieves 33.45% and 11.10% bitrate reduction on average compared with the equivalent bit allocation (EBA) and the optimal independent joint bit allocation (OIJBA) methods respectively.

Index Terms— H.264 statistical multiplexing, inter-frame dependency model (IFDM), convex optimization, dependent joint bit allocation

1. INTRODUCTION

With the great advancement of the multimedia technologies and digital communications in the last few decades, the video sharing becomes increasingly popular. In many applications, such as direct broadcast satellite (DBS), video-on-command services, video streaming etc., in order to improve the bandwidth efficiency, multiple video streams are multiplexed and transmitted within a common channel. This is also known as statistical multiplexing. Due to the different characteristics of the video streams, how to efficiently allocate the shared bandwidth among the multiple video programs is a great challenge, and the joint rate allocation for statistical multiplexing has been extensively studied in the previous literature [1, 2].

Upon the different distortion criterion, the existing joint bit allocation methods can be divided into two categories: minimum distortion variance (MINVAR) oriented approaches and minimum average distortion (MINAVR) oriented approaches. However, the MINVAR oriented approaches are often achieved at the expense of the overall visual quality degradation. Different from the MINVAR oriented approaches, the MINAVR oriented approaches aims to enhance the average quality across multiple video programs. Given the R-D curve of each video program at each time instant, the average compression distortion can be achieved using standard optimization techniques like Lagrangian optimization. The bit allocation of this approach actually finds a point in each R-D curve where the slope is the same for all the R-D curves, and this is also known as the constant slope condition [1]. In [2], with the usage of not only the differing complexity of each stream at every moment but also the differing complexity of each stream over time, a competitive equilibrium bitrate allocation approach is proposed, and the resulting video streams perform better or at least as well as with individual encoding. Although the method is [2] is a good method, the basic unit for bit allocation is one group of pictures (GOP), and it can not be directly extended to the applications where the bit allocation is performed at frame-level. The reason is due to the temporal prediction used in the current video coding scheme. Because of the temporal dependency, the frame-level R-D function is coupled together, and approach proposed in [2] can not be directly used.

For most of the existing MIVAVR oriented approaches as [1, 2], one common drawback is that the bit allocation only depends on the frame complexities. Intuitively, although one frame is of less complexity, more bits should be assigned to it if this frame is very important for the temporal prediction of the future frames. Therefore, to achieve a better overall R-D performance, the inter-frame dependency should be em-

This work has been supported in part by the the Research Grants Council (RGC) of the Hong Kong Special Administrative Region, China. (GRF Project no. 610210)

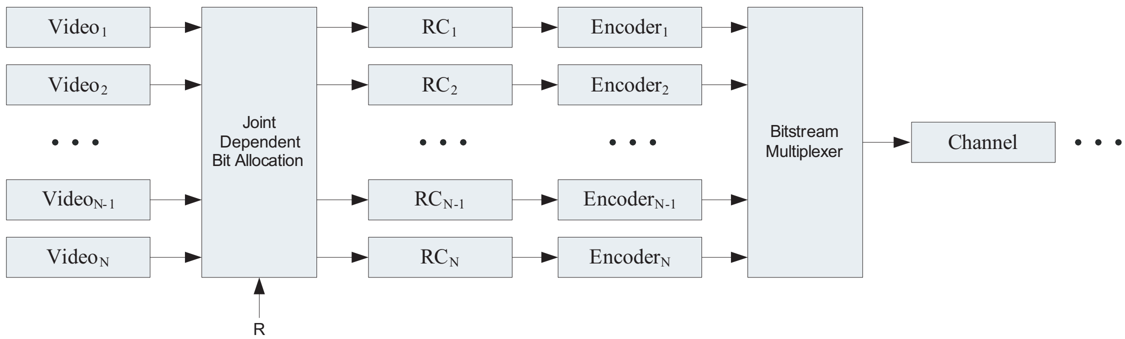

Fig. 1. structure of the proposed scheme

ployed during the joint bit allocation process. In this paper, we propose a model-based joint bit allocation approach to efficiently allocate the available network bandwidth to multiple video programs with the state-of-the-art H.264/AVC as the encoder. In our proposed scheme, at each time instant, both of the frame complexity and its influence to the coding of future frames are considered which aims to achieve the best overall visual quality under MINAVR measure.

The rest of this paper is organized as follows. In Section II, we give a general description of the statistical multiplexing system, and some key modules that would be used in the joint bit allocation process. In Section III, in addition to a quick review of the existing bit allocation methods, the proposed joint dependent bit allocation scheme is introduced. To verify the effectiveness of the proposed method, the experimental results are given, and the superiority over the existing methods is shown in Section IV. Section V concludes the paper.

2. SYSTEM MODEL

The system structure of the proposed scheme is shown in Fig. 1. In the statistical multiplexing system, the available bandwidth R is shared by N video programs. During the coding process, the centralizer will collect the coding information of the encoded frames and the future frames in the look-ahead window. Based on the collected information, each video program will be allocated with target bits for the current frame. Then, the following rate control (RC) module is used to tune the encoder parameters such that the generated bits are as close to the target bits as possible. After that, the output bitstreams of the N encoders will be multiplexed and transmitted through the channel. In this section, several key models, including the R-D model, rate-quantization (R-Q) model and the inter-frame dependency model (IFDM), used in the proposed system will be discussed in detail.

2.1. Rate-Distortion (R-D) Model

In a typical hybrid video coding scheme, intra and inter prediction are used to remove the spatial and temporal dependency respectively, and the residue signal can be obtained by subtracting the prediction signal from the original signal. It is well-known that probability distribution of the transformed coefficients can be well approximated with a zero-mean Laplacian distribution with the pdf

$$f_L(x) = \frac{\lambda}{e}^{-\lambda|x|} \tag{1}$$

$$\lambda = \frac{\sqrt{2}}{\sigma} \tag{2}$$

where x is the value of the transformed coefficients, and σ is its standard deviation. According to the classical R-D theory, the corresponding R-D function is

$$R_L(D) = a \cdot log\frac{\sigma^2}{D} \tag{3}$$

where a is the model parameter, and D is the compression distortion which is measured by the mean squared error (MSE) between the original signal and the corresponding reconstructed signal. Note that in hybrid video coding, even though the residue signal is zero, some bits are still needed for the other information, for instance, coding mode. Therefore, we add the offset b to compensate this effect. Consequently, the R-D model is changed into

$$R(D) = a \cdot log\frac{\sigma^2}{D} + b \tag{4}$$

2.2. Rate-Quantization (R-Q) Model

Accurate rate-quantization (R-Q) model plays an important role in model-based rate control algorithms. It aims to select the quantization parameter (QP) with which the encoder generates at approximately the target number of bits assigned for each coding unit. Note that one coding unit can be either an macroblock (MB), a slice, or a frame, and all the MBs in one coding unit share the same QP. By assuming that the transform coefficients are Laplacian distributed, the rate is estimated by a quadratic function of the quantization stepsize

(QStep) in [3], which can be written as

$$R = \alpha_1 \cdot \frac{MAD}{QStep} + \alpha_2 \cdot \frac{MAD}{QStep^2} \quad (5)$$

where α_1 and α_2 are model parameters. MAD serves as a complexity measure of the coding unit which is the mean absolute different(MAD) of the residue signal. Note that other more sophisticated R-Q models in the literature can also be employed. However, this is out of the focus of this article, and we utilize as the R-Q model in this paper due to its effectiveness and simplicity.

2.3. Inter-Frame Dependency Model (IFDM)

Note that in order to determine the R-D model in (4), the variance of the DCT coefficients, σ_n^2, needs to be estimated at first. As reported in our earlier work [4], it can be approximated by

$$\sigma_n^2 = \alpha \cdot D_{n-1} + \beta \cdot \sigma_{n,0}^2 + \gamma \cdot QStep_n \quad (6)$$

where $\sigma_{n,0}^2$ is the variance of the innovation signal, and $QStep_n$ is the quantization step size for the n^{th} frame. α, β and γ are model parameters. In the above equation , D_{n-1} is the compression distortion of the reference frame. Generally, the reference frames with larger distortion will lead to worse prediction. Hence, the variance of the residue tends to be large. For $\sigma_{n,0}^2$, it aims to measure the intrinsic difference , the so-called innovation signal, between the current frame and its reference frame in the clean video sequence, and in this paper this difference is represented with the variance of the innovational signal. The last term, in (6), that will influence the variance of the resulting residue signal is the quantization stepsize employed in the coding process. The main reason is due to the rate-distortion optimization (RDO) performed at the encoder side.

3. PROPOSED DEPENDENT JOINT BIT ALLOCATION

Based on the aforementioned models, our proposed model-based optimal dependent joint bit allocation algorithm will be discussed in this section. Different from [1, 2], both the frame complexity and the inter-frame dependency are considered in the proposed method. Due to the unpredictability of the future frames, the look-ahead approach is employed in our proposed statistical multiplexing system. Although the look-ahead approaches incurs certain delay and extra buffer to store the future frames, much more accurate R-D behaviors can be achieved. From the delay or buffer constraint in practical implementations, before the n^{th} frame is encoded, suppose that besides the n^{th} frame, $W-1$ future frames, i.e. $n+1, n+2....n+W-1$, are also available in the encoder buffer. Therefore, there are in total W frames can be analyzed for each video program. With the availability of the future frames in each video program, the influence of the n^{th} frame to them can be quantitatively measured by the proposed IFDM. Together with the proposed R-D model and IFDM, the joint dependent bit allocation problem can be formulated as

$$\begin{aligned}
\min_{R_t^{(m)}, D_t^{(m)}} \quad & \sum_{m=1}^{M} \sum_{t=n}^{n+W-1} w^{(m)} \cdot D_i^{(m)} \\
\text{s.t.} \quad & \sum_{m=1}^{M} R_n^{(m)} = R \\
& R_n^{(m)} = a^{(m)} \cdot log(\frac{\sigma_n^{2,(m)}}{D_n^{(m)}}) + b^{(m)} \\
& \sigma_n^{2,(m)} = \alpha^{(m)} \cdot D_{n-1}^{(m)} + \beta^{(m)} \cdot QStep_n^{(m)} + \gamma^{(m)} \cdot \sigma_{n,0}^{2,(m)}
\end{aligned} \quad (7)$$

In the optimization problem, since the values of $QStep_t^{(m)}$ ($t = n, n+1, ..., n+W-1$) are unknown before the coding process as we discussed in the previous section, it is approximated by the $QStep$ of the most recent encoded frame $QStep_{n-1}^{(m)}$. For $\sigma_{t,0}^{2,(m)}$, it is obtained from ME on the W frames in the encoder buffer. Then, both $QStep_t^{(m)}$ and $\sigma_{t,0}^{2,(m)}$ are known. We define

$$\begin{aligned}
Diff_t^{(m)} &= \beta^{(m)} \sigma_{t,0}^{2,(m)} + \gamma^{(m)} \cdot \cdot QStep_t^{(m)} \\
f_w(D_t^{(m)}) &= \sum_{t=n}^{w} a^{(m)} \cdot log(\frac{\alpha^{(m)} \cdot D_{t-1}^{(m)} + Diff_t^{(m)}}{D_t^{(m)}}) + b^{(m)}
\end{aligned} \quad (8)$$

One important property of $f_w(D_t^{(m)})$ is that it is a convex function of $D_t^{(m)}$. Then, the optimization problem can be relaxed into

$$\begin{aligned}
\min_{D_t^{(m)}} \quad & \sum_{m=1}^{M} \sum_{t=n}^{n+W-1} w^{(m)} \cdot D_i^{(m)} \\
\text{s.t.} \quad & \sum_{m=1}^{M} f_n(D_t^{(m)}) \leq R \\
& \sum_{m=1}^{M} f_{n+1}(D_t^{(m)}) \leq 2 \cdot R \\
& \ldots \\
& \sum_{m=1}^{M} f_{n+W-1}(D_t^{(m)}) \leq W \cdot R
\end{aligned} \quad (9)$$

Thus, the optimization problem in (9) can be effectively solved with the techniques introduced in [5]. For us, the scientific software CVX [6] is used to solve the problem, and the optimal bit allocation strategy can be achieved.

4. EXPERIMENTAL RESULTS

In order to evaluate the proposed dependent joint bit allocation method, we implement it using the H.264/AVC reference

Table 1. R-D performance comparison with different video program weights

Weight	BDPSNR(over EBA)	BDBR(over EBA)	BDPSNR(over OIJBA)	BDBR(over OIJBA)
(1.0, 1.0, 1.0, 1.0)	+1.69 dB	-32.52%	+0.58 dB	-12.01%
(1.6, 0.8, 0.8, 0.8)	+1.54 dB	-30.42%	+0.57 dB	-11.97%
(0.8, 1.6, 0.8, 0.8)	+1.43 dB	-28.36%	+0.48 dB	-10.49%
(0.8, 0.8, 1.6, 0.8)	+2.39 dB	-46.18%	+0.50 dB	-11.07%
(0.8, 0.8, 0.8, 1.6)	+1.58 dB	-29.76%	+0.48 dB	-9.95%
Average	+1.73 dB	-33.45%	+0.52 dB	-11.10%

software JM 11.0 as the video encoder. Meanwhile, the equal bit allocation (EBA) and optimal independent joint bit allocation (OIJBA) are also simulated for comparison reason. In our experiment, 4 video sequences of the resolution 352x288, (*akiyo*, *foreman*, *mobile* and *paris*) are selected as the test sequences. Note that the chosen sequence contain quite different video characteristics including slow and fast motion, smooth and complex scene. The configuration of the encoder is set as follows: all the sequences are encoded with the GOP structure of IPP..P, and the frame rate is 30 frames per second. CABAC is used as the entropy coding method, and the maximum search range for motion estimation is ± 16. The shared bandwidth can be selected to be 1.25, 1.5, 1.75 and 2.0 Mb/s.

The experimental results with different perceptual weights are summarized in Table 1. As shown in the table, the proposed bit allocation method has 1.73 dB and 0.52 dB BDPSNR improvement over EBA and OIJBA respectively. Or equivalently, up to 33.45% and 11.10% bitrate savings are achieved. When the shared bandwidth is 1.5 Mb/s, the evolutions of the frame-level allocated bits and PSNR are shown in Fig.2.

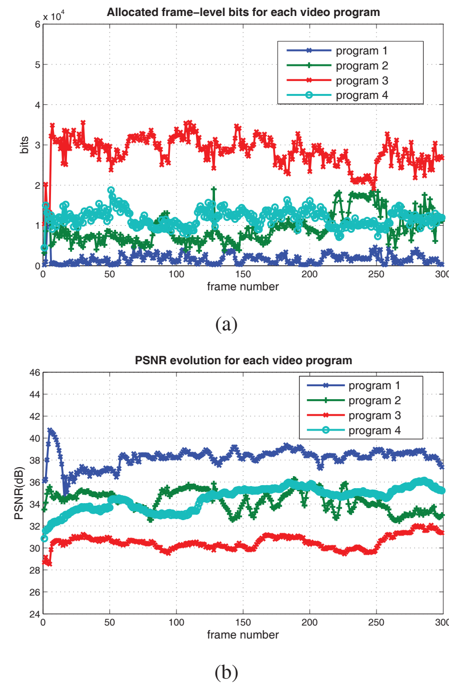

Fig. 2. The evolution of the frame-level bits and PSNR

5. CONCLUSION

In this paper, we address the dependent joint bit allocation problem in statistical multiplexing systems. Different from the previous methods where the bit allocation solely depends on the frame complexity, the inter-frame dependency, which is quantitatively measured by the proposed IFDM, is considered in the bit allocation process. In the end, the dependent joint bit allocation problem is formulated and relaxed into a convex optimization problem which can be optimally and efficiently solved. Experimental results demonstrate that the proposed method can have 1.73 dB and 0.52 dB BDPSNR improvement over EBA and OIJBA respectively.

6. REFERENCES

[1] Z. He and S. Mitra, "Optimum bit allocation and accurate rate control for video coding via ρ-domain source modeling," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 10, pp. 840–849, Oct. 2002.

[2] M. Tiwari, T. Groves, and P. C. Cosman, "Competitive equilibrium bitrate allocation for multiple video streams," *IEEE Trans. Image Process.*, vol. 19, no. 4, pp. 1009–1021, Apr. 2010.

[3] Z. Li, F. Pan, K. P. Lim, and S. Rahardja, "Adaptive rate control for H.264," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2004, pp. 745–748.

[4] C. Pang, O. C. Au, J. Dai, and F. Zou, "Frame-level dependent bit allocation via geometric programming," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, Rio de Janeiro, Brazil, May. 2011.

[5] S. Boyd and L. Vandenberghe, Eds., *Convex optimization*, Cambridge University Press, 2004.

[6] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," `http://cvxr.com/cvx`, Apr. 2011.