MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Constant Quality Rate Allocation for FGS Coding Using Composite R-D Analysis

Hui Cheng, Xi Min Zhang, Yun Q. Shi, Anthony Vetro, Huifang Sun

TR2006-019 April 2006

Abstract

In this correspondence, we propose a constant quality rate allocation algorithm for Fine Granularity Scalability (FGS) coded videos. The rate allocation problem is formulated as a contrained minimization of quality fluctuation. The rate allocation problem is formulated as a constrained minimization of quality fluctuation. The minimization is solved using a composite rate distortion (R-D) analysis. For a set of video frames, a composite R-D curve is first computed and then used for computing the optimal rate allocation. This algorithm is efficient because it is neither iterative nor recursive. After the composite R-D curve is computed, it can be used for optimal rate allocation of any rate budget. Moreover, the composite R-D curve can be updated efficiently over sliding windows. Experimental results have shown both the effectiveness and the efficiency of the proposed algorithm.

IEEE Transactions on Multimedia

CONSTANT QUALITY RATE ALLOCATION FOR FGS CODING USING COMPOSITE R-D ANALYSIS

Hui Cheng [1*], *Xi Min Zhang* [2], *Yun Q. Shi* [2], *Anthony Vetro* [3], *Huifang Sun* [3]

[1] Sarnoff Corporation, 201 Washington Rd, Princeton, NJ 08540,

Phone: (609) 734-2492, Fax: (609) 734-2662, Email: hcheng@sarnoff.com

[2] Dept. of ECE, NJ Institute of Technology, Newark, NJ 07102-1982,

Phone: (973) 596-3501, Fax: (973) 596-5680, Emails: {xxz9701,shi}@njit.edu

[3] Mitsubishi Electronic Research Lab, 201 Broadway, Cambridge, MA 02139

Phone: (617) 621-7500, Fax: (617) 621-7550, Emails: {avetro,hsun}@merl.com

ABSTRACT

In this correspondence, we propose a constant quality rate allocation algorithm for Fine Granularity Scalability (FGS) coded videos. The rate allocation problem is formulated as a constrained minimization of quality fluctuation. The minimization is solved using a composite rate distortion (R-D) analysis. For a set of video frames, a composite R-D curve is first computed and then used for computing the optimal rate allocation. This algorithm is efficient because it is neither iterative nor recursive. After the composite R-D curve is computed, it can be used for optimal rate allocation of any rate budget. Moreover, the composite R-D curve can be updated efficiently over sliding windows. Experimental results have shown both the effectiveness and the efficiency of the proposed algorithm.

EDICS: 1-CPRS, 5-QOSV, 5-STRM

1. INTRODUCTION

With the growth of Internet and wireless communication, there is an increasing demand of video delivery over networks. Unlike digital broadcasting systems, where bandwidth and channel characteristics are known, video delivery over networks has to meet constraints imposed by both users and network conditions. The same video may need to be delivered under dramatically different rate constraints, from dedicated high-speed networks to cable modems, DSL, dial-up modems or even wireless networks. Therefore, it is desirable to encode video with Fine Granularity Scalability (FGS), so that it is encoded once, but transmitted many times at different rates, and can still take full advantage of the available rate.

FGS codings [1,2] have been developed and adopted as amendments to the MPEG-4. The scalable coding with fine granularity is achieved by bit-plane coding of DCT coefficients in the enhancement layer(s). The bitstream of an enhancement layer can be truncated at any point, allowing the quality of the reconstructed video to be continuously improved as the number of bits received increases.

To best utilize FGS encoding, a rate allocation algorithm is needed to transfer the rate constraint into the rate assigned to each frame, and at the same time, to maximize the visual quality. There are a number of schemes proposed in the literature [3-6]. The simplest one is constant bit-rate allocation (CBR). However, CBR often results in quality fluctuation, hence, significantly degrades the overall quality. To solve this problem, variable bit-rate (VBR) allocation is proposed for constant quality reconstruction by allocating rate according to the complexity of each frame. Wang, et al. [5] proposed an optimal rate allocation using an exponential model. Zhang and his colleagues [6] proposed a constant quality rate allocation by minimizing the sum of absolute differences of qualities between adjacent frames under the rate constraint. The solution is computed by solving a set of linear equations. However, the optimality of this approach depends on the initial condition, which is computed based on the assumption that the average distortion of CBR rate allocation is close to the distortion of the constant quality rate allocation. In fact, the two distortions must be within the same R-D sample interval for all frames, in order to have a valid solution to the set of linear equations.

In this correspondence, we propose a constant quality rate allocation algorithm for FGS using a novel composite rate-distortion (R-D) analysis. We formulate the rate allocation as a constrained minimization of quality fluctuation measured by the dynamic range of all distortions. The minimization is solved by first computing a composite R-D curve of all frames in the processing window. Then, for any given rate budget, the constant quality that can be achieved is calculated from the composite R-D curve. Finally, this constant quality is used to allocate the rate for each video frame.

The proposed rate allocation has a number of advantages over existing rate allocation algorithms. It is the true optimal solution, not an approximation. It is neither iterative nor recursive, so it is efficient and does not need an initial guess. After the composite R-D curve is computed, it can be used for the rate allocation of any rate budget. Therefore, it is suitable for FGS coded bitstreams that need to be transmitted at many different rates. In addition, the composite R-D curve over sliding windows can be updated efficiently, which further reduces the computational complexity. Experimental results using real FGS coded videos confirm both the effectiveness and the efficiency of this algorithm.

2. CONSTANT QUALITY RATE ALLOCATION

To measure the quality variation of N coded frames, we define the cost function $C(\cdot)$ as the dynamic range of their distortions. Let r_j and $D_j(r)$ be the rate and the R-D function of frame j, respectively. Then, the dynamic range of distortions is defined as $C(r_0, r_1, \cdots r_{N-1}) = \max_j D_j(r_j) - \min_j D_j(r_j)$. With this cost function, constant quality rate allocation is formulated as a constrained minimization:

$$\min_{r_0, r_1, \cdots r_{N-1}} \left[\left(\max_j D_j(r_j) \right) - \left(\min_j D_j(r_j) \right) \right], \qquad \text{subject to} \qquad \sum_{j=0}^{N-1} r_j \le R_T, \tag{1}$$

where R_T is the bit budget. Although the constrained minimization defined in (1) consists of three non-linear functions, it can be solved using composite R-D analysis.

The source complexity with respect to an encoding system, such as an FGS encoder, is measured by its R-D curve. To allocate rate in a window of N frames, N R-D curves are needed. In this section, we will propose a composite R-D analysis that combines N R-D curves into one composite R-D curve.

For frame j, we denote its R-D function as $R_j(D)$, i.e. $R_j(D) = D_j^{-1}(R)$, and its maximal distortion as $D_{\max}(j)$. The maximal distortion is achieved when the corresponding rate is either 0 or the minimal rate allowed by the system, $R_{\min}(j)$. $R_j(D)$ is set to $R_{\min}(j)$, if $D > D_{\max}(j)$. Then, we define a constant-quality-based composite R-D curve, $\tilde{R}(D)$.

$$\tilde{R}(D) = \sum_{j=0}^{N-1} R_j(D).$$

As shown in Figure 1, $\tilde{R}(D)$ is the accumulated rate when all frames have the same distortion D. Since all R-D curves are monotonic, $\tilde{R}(D)$ is also monotonic. Therefore, the inverse of $\tilde{R}(D)$, denoted as $\tilde{D}(R)$, exists. $R = \tilde{R}\left(\tilde{D}(R)\right)$. Then, the solution to (1) is

$$r_j^* = R_j\left(\tilde{D}(R_T)\right). \tag{2}$$

When $\tilde{D}(R) \le \min_j D_{\max}(j)$, Eq (2) results in constant quality over all frames, and the cost is zero. When $\tilde{D}(R) > \min_j D_{\max}(j)$, we can also prove that (2) is the solution of (1).

3. COMPUTING COMPOSITE R-D CURVES

R-D curves can be represented using both parametric and non-parametric representations. With parametric representations, the solution to Eq (1) is in closed-form. However, as pointed out in [6], most parametric representations are not accurate over a large range of rates. Since FGS is targeted to applications, where the desired bit rates can vary significantly, we will focus our discussion on non-parametric representation, specifically the piecewise linear model using operational R-D samples.

Using the piecewise linear model, the R-D curve of the j-th frame $R_j(D)$ is modeled as a piecewise linear function that is defined by a set of R-D points, denoted as $(r^s_{i,j}, d^s_{i,j})$, where $i = 0,\cdots,M_j$ and M_j is the number of R-D samples of frame j. We also assume $\max_i d_{i,j} = d_{0,j} < d_{1,j} < \cdots < d_{M_j-1,j} = 0$. Then,

$$R_j(D) = \begin{cases} r^s_{i-1,j} + \dfrac{D - d^s_{i-1,j}}{d^s_{i,j} - d^s_{i-1,j}}\left(r^s_{i,j} - r^s_{i-1,j}\right), & \text{if } d^s_{i.j} < D \le d^s_{i-1.j} < d^s_{o.j} \\ R_{\min}(j), & \text{if } D > d^s_{o.j} \end{cases}.$$

It is easy to show that the composite R-D curve is also piecewise linear, and it is determined by the set of R-D points $\left(\sum_{k=0}^{N-1} R_k(d^s_{i,j}), d^s_{i,j}\right)$, $i = 0,\cdots,M_j$ and $j = 0,\cdots,N$.

When the rate allocation is performed over sliding windows, the set of R-D points defining the composite R-D curve can be efficiently updated by "subtracting" the R-D curve of the frame moving out of the window and "adding" the one moving into the window.

4. EXPERIMENTAL RESULTS

In the experiment, we use 100 frames of the "carphone" sequences at CIF resolution. The sequence is first encoded using an FGS encoder with default quantization values recommended in the reference model. For each frame, 8 R-D points are computed. They are the R-D points of the base-layer and the base-layer plus the first L most important bit planes of the enhancement layer, where $L = 1,\cdots,7$. The proposed rate allocation is applied with a sliding window of three different sizes: 11, 31 and 61 frames. The average bit rate is set to 1.44Mbps. For each window size, we decode each frame using the allocated rate and then compute the distortion in terms of MSE[1]. The distributions of MSE are shown in Figure 2. For comparison, we also plot the distortion resulting from constant bitrate allocation, which is the same as using a window of size 1. In addition, we list the dynamic range and the variance of the distortions in

[1] The proposed rate allocation algorithm can work in conjunction with any distortion metrics. In the experimental results presented in this paper, we used MSE as the distortion measure because of its simplicity and wide adoption in related works. Although the correlation between MSE and subjective quality ratings is not very high, it allows us to show the effectiveness of the proposed rate allocation algorithm for FGS coded videos.

Table 1. It clearly shows that the quality fluctuation is significantly reduced by using our rate allocation algorithm, even when the window size is as small as 11. Our rate allocation algorithm guarantees constant quality when operating inside all R-D curves. Therefore, the remaining quality variation shown in Figure 2 is caused by both the limited window size and the inaccuracy of the R-D curve estimation. In Figure 3, we plot the distributions of MSE computed from the estimated R-D curves. Figure 3 shows the amount of quality fluctuation due to the limited window size. The differences between Figure 2 and Figure 3 are caused by the error in R-D estimation.

In conclusion, we propose a constant quality rate allocation using a novel composite R-D analysis. This approach is proven to be both effective and efficient for rate allocation of FGS coded videos.

ACKNOWLEDGEMENT

We would like to thank Dr. Michael A. Isnardi of Sarnoff for reviewing this article. This work was supported in part by National Institute of Health, grant no. NS-38494.

REFERENCES

[1] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. on Circuits, Systems and Video Technology*, vol. 11, no. 3, pp. 302-317, March 2001.

[2] M. van der Schaar and H. Radha, "A hybrid temporal-SNR fine-granular scalability for internet video," *IEEE Trans. on Circuits, Systems and Video Technology*, vol. 11, no. 3, pp. 302-317, March 2001.

[3] J. Viéron, T. Turletti, X. Hénocq, C. Guillemot and K. Salamatian, "TCP-compatible rate control for FGS layered multicast video transmission based on a clustering algorithm," *Proc. of the IEEE Int'l Symposium on Circuits, Systems (ISCAS)*, vol. 1, pp.453-456, Scottsdale, AZ, May 2002.

[4] M. van der Schaar and H. Radha, "Temporal-SNR rate-control for Fine-Granular Scalability," *Proc. of IEEE Int'l Conf. on Image Peocessing*, Vol. 2, pp. 1037-1040, Thessaloniki, Greece, Oct 2001.

[5] Q. Wang, F. Wu, S. Li, Z. Xiong, Y.-Q. Zhang and Y. Zhong, "A new rate allocation scheme for progressive fine granular scalable coding," *Proc. of the IEEE Int'l Symposium on Circuits, Systems (ISCAS)*, vol. 2, pp.397-400, Sydney, NSW, Australia, May 2001.

[6] X. Zhang, A. Vetro, Y. Q. Shi and H. Sun, "Constant quality rate allocation for FGS video coded bitstreams," *Proc. of the SPIE/IS&T Int'l Conf. on Video Comm. and Image Processing*, pp.817-827, San Jose, CA, Jan. 2002.

Table 1. Dynamic range and variance of MSE resulting from the proposed rate allocation algorithm with different window sizes. Note that rate allocation with window size 1 is the same as CBR allocation.

Window Size	MSE computed from decoded frames		MSE computed from R-D curves	
	Dynamic Range	Variance	Dynamic Range	Variance
1 (CBR)	7.73	0.43	3.05	0.38
11	1.60	0.20	1.50	0.18
31	0.92	0.07	0.70	0.06
61	0.54	0.04	0.33	0.02

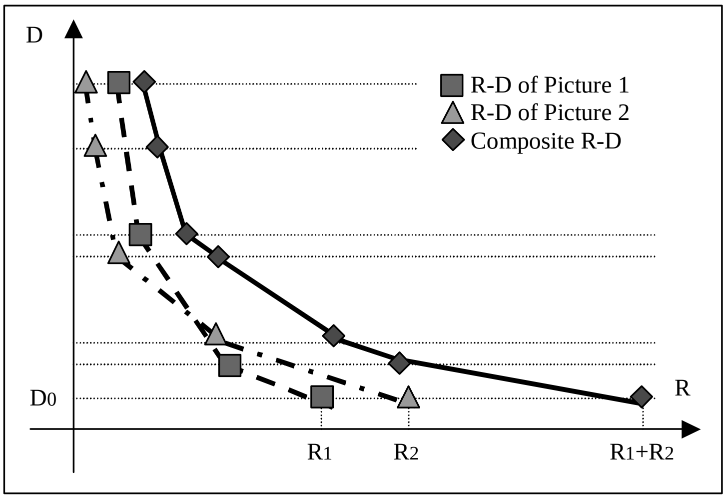

Figure 1. Constant-quality-based composite R-D curve. At each R-D sample point with a distinct distortion D_0, a composite R-D point is generated. The distortion at the composite R-D point is also D_0. But the rate at the composite R-D point is the sum of rates of all pictures at the distortion D_0.

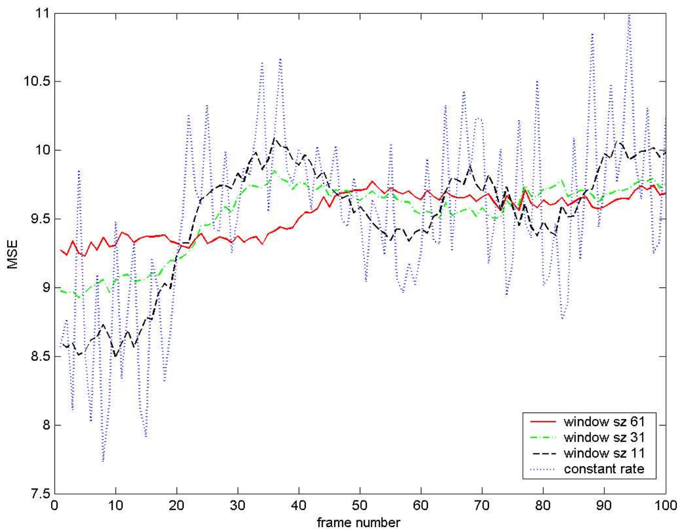

Figure 2. MSE distributions of the "carphone" sequence resulting from proposed rate allocations with four different window sizes. The MSE is computed from each decoded frame at its allocated rate.

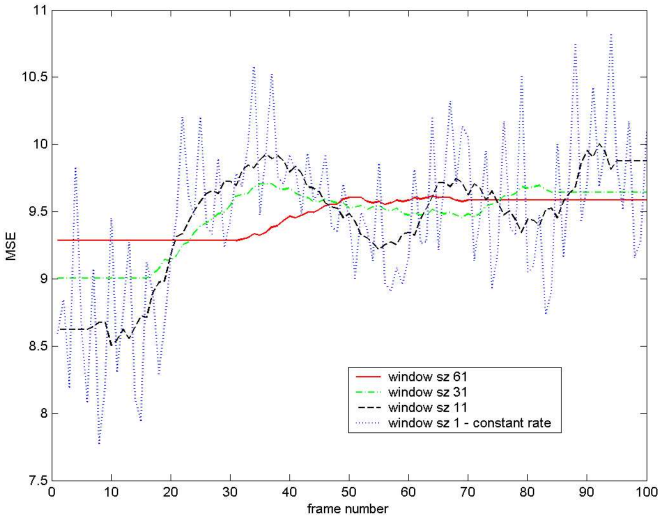

Figure 3. MSE distributions of the "carphone" sequence resulting from proposed rate allocations with four different window sizes. The MSE is computed from R-D estimated for that frame using the linear R-D model at its allocated rate.

Table 1. Dynamic range and variance of MSE resulting from the proposed rate allocation algorithm with different window sizes. Note that rate allocation with window size 1 is the same as CBR allocation.

Figure 1. Constant-quality-based composite R-D curve. At each R-D sample point with a distinct distortion D_0, a composite R-D point is generated. The distortion at the composite R-D point is also D_0. But the rate at the composite R-D point is the sum of rates of all pictures at the distortion D_0.

Figure 2. MSE distributions of the "carphone" sequence resulting from proposed rate allocations with four different window sizes. The MSE is computed from each decoded frame at its allocated rate.

Figure 3. MSE distributions of the "carphone" sequence resulting from proposed rate allocations with four different window sizes. The MSE is computed from R-D estimated for that frame using the linear R-D model at its allocated rate.