

## Rate Allocation for FGS Coded Video Using Composite R-D Analysis

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

TR-2003-97    July 2003

### Abstract

In this paper, we propose a constant quality rate allocation algorithm for MPEG-4 FGS (Fine Granularity Scalability) coded video sequences. The rate allocation problem is formulated as a constrained minimization of quality fluctuation. The minimization is solved using a novel composite rate distortion analysis. For a set of video frames, a composite rate distortion curve is first computed and then used for computing the optimal rate allocation. The proposed algorithm is very efficient because it is neither iterative nor recursive. In addition, after the composite rate distortion curve is computed, it can be used to calculate optimal rate allocation for any rate budget. Therefore, it is suitable for FGS coded bitstreams, which need to be transmitted and decoded many times at many different rates. Moreover, the composite rate distortion curve can be updated efficiently over sliding windows. This further reduces the computational complexity. Experiments using both synthetic and real FGS coded videos have shown the effectiveness and the efficiency of the proposed algorithm.

This work may not be copied or reproduced in whole or in part for any commercial purpose. Permission to copy in whole or in part without payment of fee is granted for nonprofit educational and research purposes provided that all such whole or partial copies include the following: a notice that such copying is by permission of Mitsubishi Electric Research Laboratories, Inc.; an acknowledgment of the authors and individual contributions to the work; and all applicable portions of the copyright notice. Copying, reproduction, or republishing for any other purpose shall require a license with payment of fee to Mitsubishi Electric Research Laboratories, Inc. All rights reserved.

Proc. IEEE Int'l Conference on Multimedia and Expo, July 2003.



# RATE ALLOCATION FOR FGS CODED VIDEO USING COMPOSITE R-D ANALYSIS

*Hui Cheng<sup>1</sup>, Xi Min Zhang<sup>2</sup>, Yun Q. Shi<sup>2</sup>, Anthony Vetro<sup>3</sup>, Huifang Sun<sup>3</sup>*

<sup>1</sup> Sarnoff Corporation, Princeton, NJ 08543-5300, hcheng@sarnoff.com

<sup>2</sup> Dept. of ECE, NJ Institute of Technology, Newark, NJ 07102, {xxz9701,shi}@njit.edu

<sup>3</sup> Mitsubishi Electronic Research Lab, Murray Hill, NJ 07974, {avetro,hsun}@merl.com

## ABSTRACT

In this paper, we propose a constant quality rate allocation algorithm for MPEG-4 FGS (Fine Granularity Scalability) coded video sequences. The rate allocation problem is formulated as a constrained minimization of quality fluctuation. The minimization is solved using a novel composite rate distortion analysis. For a set of video frames, a composite rate distortion curve is first computed and then used for computing the optimal rate allocation. The proposed algorithm is very efficient because it is neither iterative nor recursive. In addition, after the composite rate distortion curve is computed, it can be used to calculate optimal rate allocation for any rate budget. Therefore, it is suitable for FGS coded bitstreams, which need to be transmitted and decoded many times at many different rates. Moreover, the composite rate distortion curve can be updated efficiently over sliding windows. This further reduces the computational complexity. Experiments using both synthetic and real FGS coded videos have shown the effectiveness and the efficiency of the proposed algorithm.

## 1. INTRODUCTION

With the growth of Internet and wireless communication, there is an increasing demand of video delivery over networks. Unlike digital video broadcasting systems, where bandwidth and channel characteristics are known, video delivery over networks has to meet constraints imposed by both users and network conditions. These constraints are often dynamic in nature and difficult to predict in advance. One of these constraints is the rate at which the video is transmitted. 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 can be encoded once, but transmitted and decoded many times at different rates, and can still take full advantage of the available rate.

FGS coding [1] and FGS temporal coding [2] have been developed and adopted as amendments to the MPEG-4 visual coding standards. The scalable coding with fine granularity is achieved by bit-plane coding of DCT coefficients in the enhancement layer(s). Therefore, 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 rate allocation schemes proposed in the literature [3,4]. The simplest one is constant bit-rate allocation (CBR). Although CBR is easy to compute and easy to implement, it often results in quality fluctuation in the reconstructed video, 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 by using an exponential model. In [6], Zhang and his colleagues 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 very close to the distortion of the constant quality rate allocation. The two distortions must be within the same R-D sample interval for all frames considered in the optimization.

In this paper, we propose a constant quality rate allocation algorithm for FGS using a novel composite rate-distortion (R-D) analysis. We also formulate the rate allocation as a constrained minimization of quality fluctuation. However, instead of minimizing the sum of absolute differences, we minimize the dynamic range of all distortions. The minimization is solved by first computing a composite R-D curve of all input frames. Then, for any

given rate budget, the constant quality which 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 very efficient and does not depend on initial guess. After the composite rate distortion curve is computed, it can be used to calculate the optimal rate allocation for any rate budget. Therefore, it is a desirable rate allocation scheme for FGS coded bitstreams, which needs to be transmitted and decoded at many times at many different rates. In addition, the composite rate distortion curve over sliding windows can be updated efficiently. This further reduces the computational complexity. Experimental results using both synthetic and real FGS coded videos confirm the effectiveness and the efficiency of the proposed algorithm.

## 2. PROBLEM FORMULATION

To measure the quality variation of  $N$  coded frames, we define the cost function  $C(\cdot)$  as the dynamic range of their distortions.  $C(\cdot)$  is a function of the rate allocated to each frame. Let  $r_j$  and  $D_j(r)$  be the rate and the R-D function of frame  $j$ , respectively. Then, the dynamic range of all distortions given their rates is defined as

$$C(r_0, r_1, \dots, r_{N-1}) = \max_j D_j(r_j) - \min_j D_j(r_j). \quad (1)$$

Dynamic range is a "strict" measure of variation among a data set. A data set can have a small variance, but a large dynamic range (not bounded by the variance, unless the number of samples is known). However, with a fixed dynamic range, the variance of a data set is bounded. Using dynamic range as the cost function, constant quality rate allocation can be formulated as a constrained minimization. That is

$$\min_{r_0, r_1, \dots, r_{N-1}} \left[ \left( \max_j D_j(r_j) \right) - \left( \min_j D_j(r_j) \right) \right], \quad (2)$$

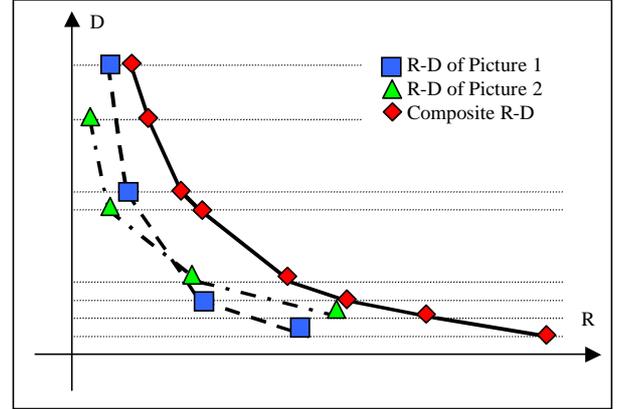
subject to the rate budget constraint

$$\sum_{j=0}^{N-1} r_j \leq R_T,$$

where  $R_T$  is the maximal average rate allowed. Although the constrained minimization defined in (2) consists of three non-linear functions, it can be solved using composite rate-distortion (R-D) analysis.

## 3. COMPOSITE RATE DISTORTION 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 will combine  $N$  R-D curves into one composite R-D curve.



**Figure 1.** Constant-quality-based composite R-D. At each R-D sample point with a distinct distortion, a composite R-D point is generated, which has the same distortion and the sum of rates of all frames at this distortion.

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

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

$\tilde{R}(D)$  is the accumulated rates when all frames have the same distortion  $D$ . Since all R-D curves are monotonic, we can show that for any given accumulated rate  $R_0$  less than or equal to the maximum accumulated rate, there is only one  $D_0$ , such that  $\tilde{R}(D_0) = R_0$ . Therefore, the inverse of  $\tilde{R}(D)$  exists. We denote the inverse of  $\tilde{R}(D)$  as  $\tilde{D}(R)$ , i.e.  $R = \tilde{R}(\tilde{D}(R))$ . Then, the solution to (2) is

$$r_j = R_j(\tilde{D}(R_T)). \quad (3)$$

When  $\tilde{D}(R) \leq \min_j D_{\max}(j)$ , Eq (3) results in constant quality over all frames, and the cost function (1) is zero. Therefore, (3) is the solution of (2). When

$\tilde{D}(R) > \min_j D_{\max}(j)$ , we can also prove that (3) is the solution of (2). However, we will omit the proof in this paper. In addition, when  $\tilde{D}(R) \leq \min_j D_{\max}(j)$ , (3) is also the solution to other measures of variation, such as variance and the sum of absolute differences.

#### 4. COMPUTING COMPOSITE R-D CURVES

Rate-distortion curves can be represented using both parametric and non-parametric representations. With parametric representations, the solution to Eq (2) can be calculated in closed-form by substituting the R-D representation in Eq (3).

However, as pointed out in [6], most parametric representations are not accurate over a large range of R. Since FGS is targeted to applications, where the desired bit rates can vary significantly, we will focus our discussion on non-parametric representation, specifically a 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  $(r_{i,j}^s, d_{i,j}^s)$ , where  $i = 0, \dots, M_j$  and  $M_j$  is the number of R-D samples of frame  $j$ . We also assume  $\max_i d_{i,j} = d_{0,j} < d_{1,j} < \dots < d_{M_j-1,j} = 0$ . Then,

$$R_j(D) = \begin{cases} r_{i-1,j}^s + \frac{D - d_{i-1,j}^s}{d_{i,j}^s - d_{i-1,j}^s} (r_{i,j}^s - r_{i-1,j}^s), & \text{if } d_{i,j}^s < D \leq d_{i-1,j}^s < d_{o,j}^s \\ R_{\min}(j), & \text{if } D > d_{o,j}^s \end{cases}$$

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  $(\sum_{k=0}^{N-1} R_k(d_{i,j}^s), d_{i,j}^s)$ .

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.

#### 5. EXPERIMENTAL RESULTS

We have conducted the following two experiments. For the first 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  $L = 1, \dots, 7$ . In Figure 2, we show three R-D curves computed from frames 1, 50 and 100. 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. The distributions of MSE are shown in Figure 3. For comparison, we also plot the distortion resulting from constant bitrate allocation. In addition, we list the dynamic range and the variance of the distortions in 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. One thing we want to point out is that our rate allocation algorithm guarantees constant quality when operating inside all R-D curves. Therefore, the remaining quality variation shown in Figure 3 is caused by both the limited window size and the accuracy of the R-D curve estimation. In Figure 4, we show the distribution of MSE computed from the input R-D curves of the test sequence. Since they are not the MSE of real decoded frames, Figure 4 shows the amount of quality fluctuation due to the limited window size. The differences between Figure 3 and Figure 4 are caused by R-D estimation.

To further stress our algorithm, we also generated a set of R-D curves that change dramatically from frame to frame. We repeat the above experiment with these synthetic R-d curves and the results are shown in Figure 5 and Table 1.

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 video sequences.

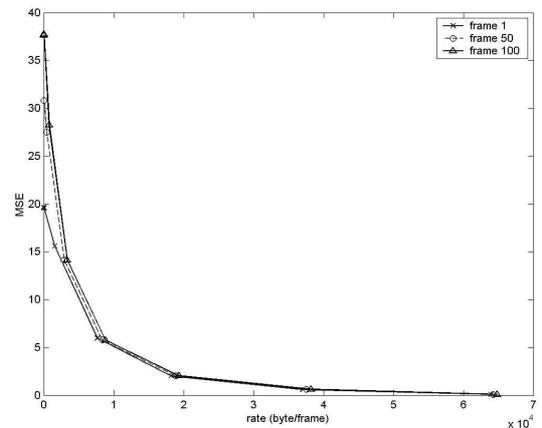
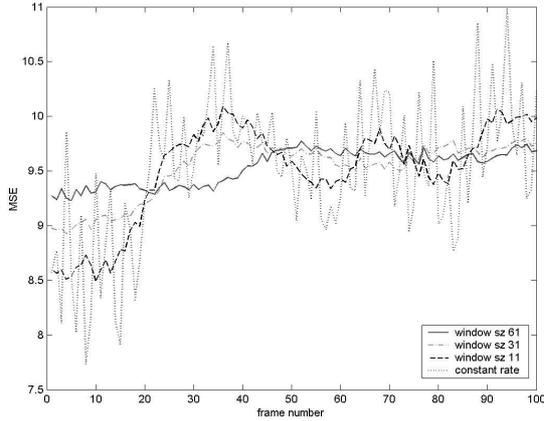


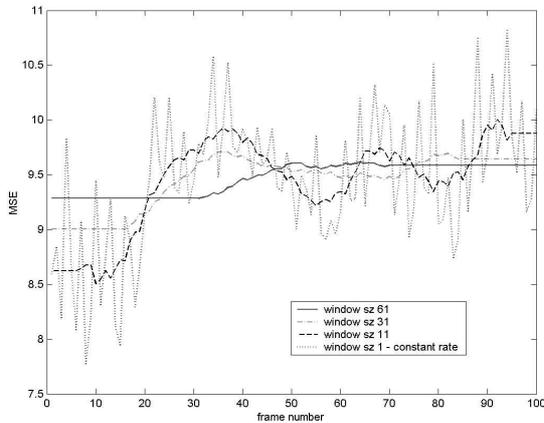
Figure 2. Three operational R-D curves.

## ACKNOWLEDGEMENT

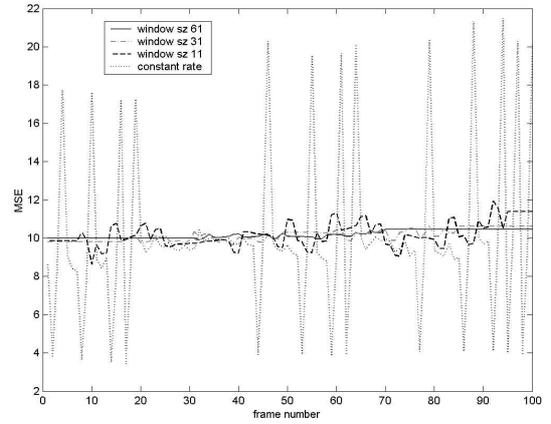
We would like to thank Dr. Michael A. Isnardi of Sarnoff for reviewing this article. This work was supported in part by National Institute of Health, grant no. NS-38494.



**Figure 3.** MSE distributions of the "carphone" sequence of four rate allocations with different window sizes. The MSE is computed from each frame decoded at its allocated rate.



**Figure 4.** MSE computed from R-D estimate, not decoded frames.



**Figure 5.** MSE distributions by applying the rate allocation to synthetic R-D curves. The MSE is computed from R-D estimates, not decoded frame.

## 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, 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, 2001.
- [3] J. Viéron, T. Turetli, 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, 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, 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. Note that rate allocation with window size 1 is the same as CBR allocation.

Window Size	Experiment with FGS Coded Test Sequence				Synthetic Experiment	
	MSE of decoded frames		MSE from estimated R-D curves		MSE from synthetic R-D curves	
	Dynamic Range (D.R.)	Variance	D. R.	Variance	D. R.	Variance
1 (CBR)	7.73	0.434	3.05	0.384	18.07	17.451
11	1.60	0.204	1.50	0.176	3.30	0.423
31	0.92	0.069	0.70	0.057	0.960	0.091
61	0.54	0.026	0.33	0.019	0.540	0.041