# A Simple Algorithm for Constant Quality Reconstruction of Scalable Video Using a New Analytical R-D Model

[1,3]Jun Sun, [2]Debin Zhao, [3]Qingming Huang, [1,3]Wen Gao

[1]Institute of Computing Technology, Chinese Academy of Sciences
Beijing 100080, China
[2]Department of Computer Science, Harbin Institute of Technology
Harbin, 150001, China
[3]Graduate School of Chinese Academy of Science
Beijing, 100080, China

**Abstract.** In this paper, we first introduce the analytical R-D model of scalable video in our recent work. The R-D model is induced by the R-D analysis of Generalized Gaussian Distribution (GGD). Then we proposed a simple algorithm for constant quality reconstruction of scalable video using the analytical R-D model. It needs not any search operation. The computation complexity of the proposed algorithm is O(*N*), where *N* is the smoothed frames. Extensive experiments on MPEG-4 FGS video show the efficiency and effective of our algorithm. Since our R-D analysis and smooth algorithm are general, intuitively they can be used in SVC video.

***Index Terms***—MPEG-4 FGS, rate distortion function, constant quality reconstruction, video coding

## 1. INTRODUCTION

The Internet is experiencing explosive growth of video streaming. Since the Internet is a shared environment, it has been commonly accepted that video streaming should react to network congestion and match the video rate with the available network bandwidth [1]. Therefore, it is desirable to encode video with scalable technologies, so that it can be encoded once, but transmitted and reconstructed many times at different targeting rates. The scalability of MPEG-4 FGS [2] is achieved by bit-plane coding of DCT coefficients in the enhancement layer. The developing Scalable Video Coding (SVC) standardization project chooses the scalable extension of H.264/AVC as a start point, which realizes the scalability through motion compensated temporal filtering (MCTF) using a lifting framework [3].

To best utilize these scalable video during streaming, a rate allocation algorithm is needed to transfer the available network bandwidth into the rate assigned to each frame. The trivial constant bit-rate allocation usually results in significant quality fluctuation in the reconstructed video. Hence, some complex R-D based allocation algorithms should be employed to realize constant quality reconstruction at the allowed transmission rate by allocating rate according to the complexity of each frame.

In this paper, we first introduce the analytical R-D model of scalable video in our recent work. The R-D model is induced by the R-D analysis of Generalized Gaussian Distribution (GGD). Then we proposed a simple algorithm for constant quality reconstruction of scalable video using the analytical R-D model. It needs not any search operation. The computation complexity of the proposed algorithm is O(N), where N is the smoothed frames. Extensive experiments on MPEG-4 FGS video show the efficiency and effective of our algorithm. Since our R-D analysis and smooth algorithm are general, intuitively they can also be used in SVC video.

## 2. RATE DISTORTION ANALYSIS OF FGS EL WITH GENERALIZED GAUSSIAN DISTRIBUTION

### 2.1 Generalized Gaussian Distribution

Generalized Gaussian distribution is a nice model for the DCT coefficients [4] and wavelet coefficients [5].The PDF of zero-mean generalized Gaussian distributions can be described as follows:

$$p(x) = \frac{\alpha\eta(\alpha,\beta)}{2\Gamma(1/\alpha)}\exp\{-[\eta(\alpha,\beta)\,|\,x\,|]^{\alpha}\} \quad \text{(1a)}$$

with

$$\eta(\alpha,\beta) = \beta^{-1}\left[\frac{\Gamma(3/\alpha)}{\Gamma(1/\alpha)}\right]^{1/2}, \quad \text{(1b)}$$

where $\alpha > 0$ is the shape parameter describing the exponential rate of decay, $\beta$ is a positive quantity representing a scale parameter, and $\Gamma(\bullet)$ is the gamma function [11]. The variance of the random variable is expressed by $\sigma^2 = \beta^2$. For simplicity of denotation, zero-mean generalized Gaussian distribution is also called generalized Gaussian distribution in this paper.

Generalized Gaussian distributions cover a wide range of symmetric distributions. The distribution shape is controlled by the shape parameter $\alpha$. As we notice above, when $\alpha = 2$, the generalized Gaussian distribution corresponds to a Gaussian distribution. While for $\alpha = 1$, we have the Laplacian distribution. As $\alpha \to \infty$, the distribution approaches the uniform distribution in $[-\sqrt{3}\beta, \sqrt{3}\beta]$, and when $\alpha \to 0^{+}$ the distribution becomes a single point with $x = 0$ [12]. Generally generalized Gaussian distribution would be a better statistical model for DCT and wavelet coefficients than the Gaussian and Laplacian distributions because the extra shape parameter $\alpha$ can be tuned to the samples of DCT and wavelet coefficient.

## 2.2 Rate Distortion Analysis of GGD

In FGS video streaming, if the number of bitplanes in an EL frame is *n* and the last transmitted bitplane is *k*, then the quantizer can be considered as uniform quantization with step size $\Delta = 2^{n-k}$. For a standard FGS decoder, the quantization scheme can be described as follows:

$$T_i = i\Delta, i = \ldots, -N, \ldots, -1, 0, 1, \ldots, N, \ldots \qquad (2a)$$

and

$$\mathrm{R_i} = \begin{cases} T_i, & \text{if } 0 \le T_i \le \mathrm{x} < T_{i+1} \\ T_{i+1}, & \text{if } T_i < \mathrm{x} \le T_{i+1} \le 0 \end{cases} \qquad (2b)$$

where $T_i$ are called the quantization thresholds, $R_i$ are called the reconstruction levels, $\Delta$ is the quantization step size, and $x$ are the value of samples.

Under above quantization scheme, the distortion-rate function of GGD [10] can be described as

$$PSNR(R) = f_{\alpha}(R) + 20\log_{10}(255/\beta), \quad (3)$$

where $R$ is the bit rate of encoding, $PSNR(R)$ is the distortion with criterion PSNR at the bit rate $R$, $f_{\alpha}(R)$ is the function of bit rate $R$ with the parameter $\alpha$.

For the same shape parameter $\alpha$ and different root variance $\beta$, $d_{PSNR}/d_R$ is not relative to $\beta$. Fig. 1 shows the derivative comparison of the distortion-rate functions of the Laplacian distribution (shape parameter 1.0), the generalized Gaussian distribution with shape parameter 0.5 and the Gaussian distribution (shape parameter 2.0). Generally, the derivative of GGD distortion-rate function (PSNR criterion) decreases first and then increases to the traditional number 6.02. The inflexion of the derivative becomes larger if the shape parameter $\alpha$ is lower. For Laplacian distribution, the derivative decreases to about 4.34 at the bit rate 1.5 bits/sample and then increases to 6.02 gradually as the bit rate increases.

For actual FGS coding, most DCT coefficients have the shape value less than 1.0 in actual video frames [6] [10], thus we can assert that the derivative of actual distortion-rate function begins to increase at a comparatively high bit rate (higher than 1.5 bits/sample).

On the other hand, there are more bitplanes at a high bit rate, and the bitplanes are not independent from each other in fact. But these bitplanes are entropy coded independently. Therefore as bit rate increases, the efficiency of actual bitplane coding is decreased. This also delays and slows down the increase of the derivatives of actual distortion-rate function (PSNR criterion), and even flattens out the increase in some cases. So, for actual FGS coding (lower than 3 bits/sample), we can assume that the derivatives decrease continuously [6] [10].

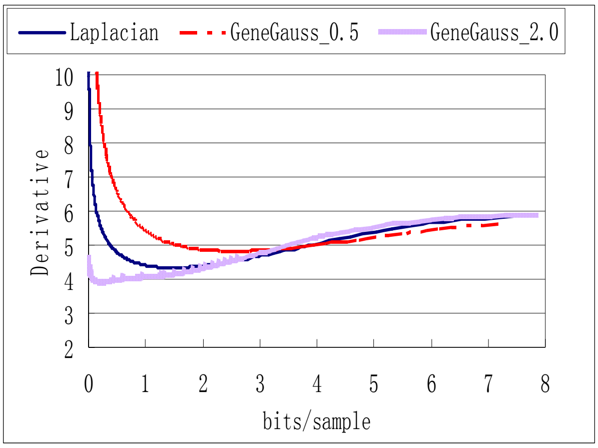



**Fig. 1.** Derivative comparison of the distortion-rate functions with the GGD shape parameter 0.5, 1.0 (Laplacian) and 2.0 (Gaussian).

## 2.3 Analytical R-D Model of MPEG-4 FGS Video

With above analysis, we can assume that the derivative of distortion-rate function (PSNR criterion) decreases continuously as the rate increases. Then we can get a heuristic R-D model of FGS video as follow [6] [10]:

$$PSNR(R) = a * R + A - (A - B)/(1 + b * R) \quad (4)$$

where *R* is the bit rate per sample, *B* is the PSNR of FGS BL coding, *a* and *A* is the asymptote parameter, *b* is the parameter controlling the approach of the actual RDF to the asymptote. The parameters *a* and *b* can also be set as constants for a coarse model.

## 3. CONSTANT QUALITY TRUNCATION OF MPEG-4 FGS VIDEO

Rate allocation between different video frames for constant quality reconstruction is one popular application of R-D models. The key problem is how to truncate the FGS EL to both match the available average bandwidth $\bar{R}$ and achieve a certain constant quality $D_{target}$ for each frame. Usually, to obtain the $D_{target}$, we need to calculate a combined function $C(D)$, which is constrained by $\bar{R}$:

$$C(D) = \frac{1}{N}\sum_{i=0}^{N-1} R_i(D) = \bar{R}. \quad (5)$$

where $R_i(D)$ is the rate-distortion function of frame $i$. $N$ is the number of smoothed frames. Through (5), we can obtain the target constant quality $D_{target} = C^{-1}(\bar{R})$. Then the allocated bit rate of frame $i$ can be calculated by $R_i(D_{target}) = R_i(C^{-1}(\bar{R}))$ [7] [8].

However it's difficult to get a closed-form solution of $C^{-1}$ for the known R-D models and a search algorithm for $D_{target}$ is a burden for streaming server since the $\bar{R}$ changes continually in the actual streaming [9]. Using the R-D model of the second experiment in Section IV-B, where $b$ is fixed as 1.5, we introduce a new simple algorithm for constant quality reconstruction (SACQR) as follows:

1) Calculate the average distortion $\bar{D}$ with uniform bit rate allocation $\bar{R}$:

$$\bar{D} = \Big[\bar{R} * \sum_{i=0}^{N-1} a_i + \sum_{i=0}^{N-1} A_i - \big(\sum_{i=0}^{N-1} A_i - \sum_{i=0}^{N-1} B_i\big)/(1+1.5*\bar{R})\Big] / N \quad (6)$$

where $a_i$, $A_i$ and $B_i$ are the corresponding values of (4) in frame $i$.

2) Calculate the initial bit rate allocation in frame $i$:

$$init\_rate_i = R_i(\bar{D}), \quad (7)$$

where $R_i(D)$ is the inverse function of (4) in frame $i$.

3) The average tune bit rate can be obtained by

$$tune\_rate = \frac{1}{N}\Big(\sum_{i=0}^{N-1} init\_rate_i\Big) - \bar{R}. \quad (8)$$

4) Calculate the tune weight of each frame $i$:

$$Tune\_Weight = N * [D_i'(init\_rate_i)]^{-1} \Big/ \sum_{i=0}^{N-1} [D_i'(init\_rate_i)]^{-1} \quad (9)$$

where $D_i'(init\_rate_i)$ is the derivative of distortion-rate function of frame $i$ at the bitrate $init\_rate_i$. $[D_i'(init\_rate_i)]^{-1} = R_i'(\bar{D})$ approximates the bitrate requirement of one unit distortion change at the distortion point $\bar{D}$ of frame $i$.

5) Then the transmitting bit rate can be calculated by

$$trans\_rate_i = init\_rate_i - tune\_rate * Tune\_Weight. \quad (10)$$

There may be some difference between the target $\bar{R}$ and the average transmitting bit rate $\frac{1}{N}\sum_{i=0}^{N-1} trans\_rate_i$. Usually the difference can be ignored. However if needed, it can also be decreased by recursively calculating (8-10). The computation complexity of the proposed SACQR is O($N$).

## 4. EXPERIMENTS

To validate the effectiveness of the methods that we have described, we have done some experiments in four sequences: 10fps Foreman, Tempete_ext (two concatenated 260_frame Tempete sequences), Stefan and Paris sequences. The base layer bitrate target is 128 kbps with TM5. The target EL bit rate $\bar{R}$ is 0.5 bits/sample (about 760kbits/s) and 1.0 bits/sample (about 1520 kbits/s). The smoothed frame is from 0 to 297. We apply two different algorithms in experiments. The first is the proposed SACQR. The uniform bit rate allocation is also applied for comparison.

Figs. 2-5 compare the proposed SACQR with uniform bitrate allocation algorithm in the four sequences respectively. It can be seen that the near-constant quality can be obtained with the proposed approach except for some special frames such as frame 278 in Fig. 2(a), where the base layer quality is larger than target constant quality, while the uniform bitrate allocation method contains significant variation.

## 5. CONCLUSION

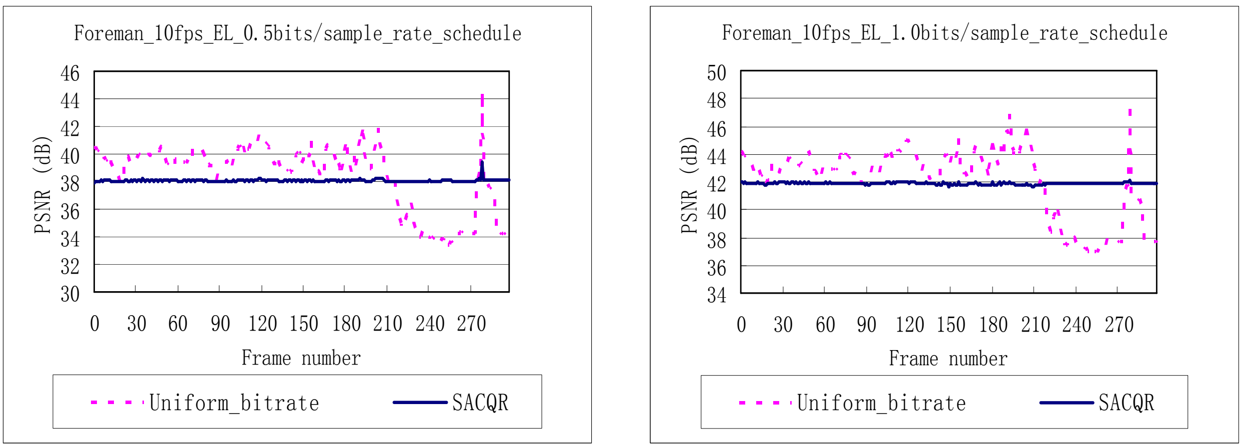
In this paper, we proposed a simple algorithm for constant quality reconstruction of scalable video using an analytical R-D model. It needs not any search operation. The computation complexity of the proposed algorithm is O(N). Extensive experiments on MPEG-4 FGS video show the efficiency and effective of our algorithm.


## 6. ACKNOWLEDGMENT

This work is partially supported by the National Natural Science Foundation of China under grant No. 60333020, the Natural Science Foundation of Beijing under grant No. 4041003, and the National Hi-Tech Development 863 Program of China under grant No. 2002AA119010.

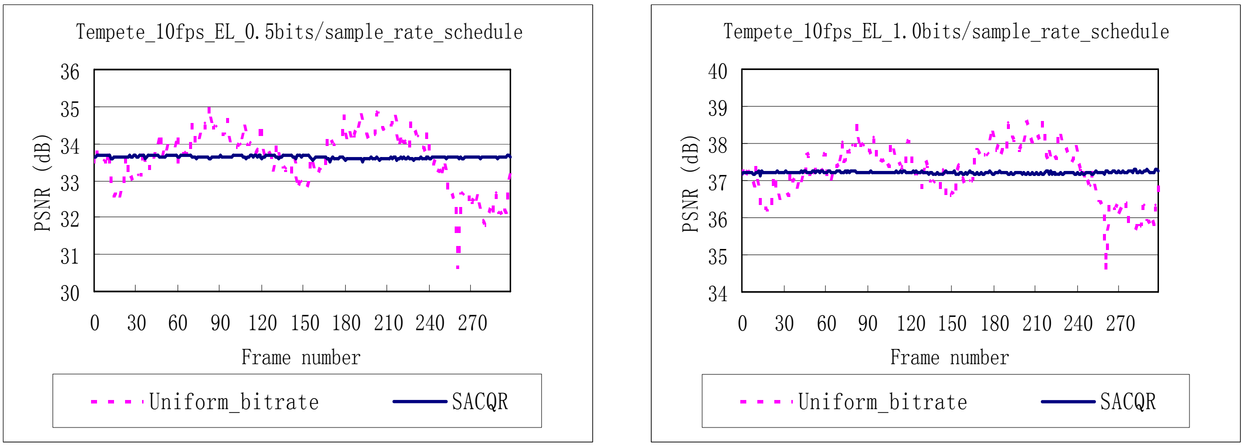
## REFERENCES


1. D. Wu, Y. Hou, W. Zhu, Y. Zhang, and J. Peha: Streaming video over the Internet: Approaches and directions. Vol. 11, no. 3 of *IEEE Trans. Circuits Syst. Video Technol*. IEEE Press, (Mar. 2001)
2. W. Li: Overview of fine granularity scalability in MPEG-4 video standard. Vol. 11, no. 3 of *IEEE Trans. Circuits Syst. Video Technol*. IEEE Press, (Mar. 2001)
3. Julien Reichel. ed.: Scalable Video Coding – Working Draft 1. Document JVT-N020, Joint Video Team of ITU-T VCEG and ISO/IEC MPEG (Jan. 2005)
4. E. Y. Lam, J. W. Goodman: A mathematical analysis of the DCT coefficient distributions for images. Vol. 9, no. 10 of *IEEE Trans. Image Processing*, IEEE Press, (Oct. 2000)
5. Mallat, S.G.: A theory for multiresolution signal decomposition: the wavelet representation. Vol. 11, no. 7 of *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Press, (1989)
6. J. Sun, W. Gao, D. Zhao and Q. Huang: Statistical Model, Analysis and Approximation of Rate Distortion Function in MPEG-4 FGS Videos. *VCIP, Bei Jing* (*2005*)
7. M. Dai, and D. Loguinov: Analysis of rate-distortion functions and congestion control in scalable Internet video streaming. *Proc. of NOSSDAV'03*, ACM Press, Monterey, CA, USA (Jun. 2003)
8. X. Zhang, A.Vetro, Y.-Q. Shi, and H. Sun: Constant quality constrained rate allocation for FGS-coded video. Vol. 13, no. 2 of *IEEE Trans. Circuits Syst. Video Technol*, IEEE Press, (Feb. 2003)
9. X. Zhao, Y. He, S. Yang, and Y. Zhong: Rate Allocation of Equal Image Quality for MPEG-4 FGS Video Streaming. *Packet Video Workshop*, April 2002.
10. J. Sun, W. Gao, D. Zhao and Q. Huang: Statistical Model, Analysis and Approximation of Rate Distortion Function in MPEG-4 FGS Videos. *To be printed in IEEE Trans. Circuits Syst. Video Technol*, IEEE Press
11. M. Abramowitz and I. A. Stegun: *Handbook of Mathematical Function*s. New York: Dover, (1970).
12. J. A. Domínguez-Molina, G. González-Farías and R. M. Rodríguez-Dagnino: A practical procedure to estimate the shape parameter in the generalized Gaussian distribution. Technique Report I-01-18, Sep. 2001, available: http://www.cimat.mx/reportes/enlinea/I-01-18_eng.pdf

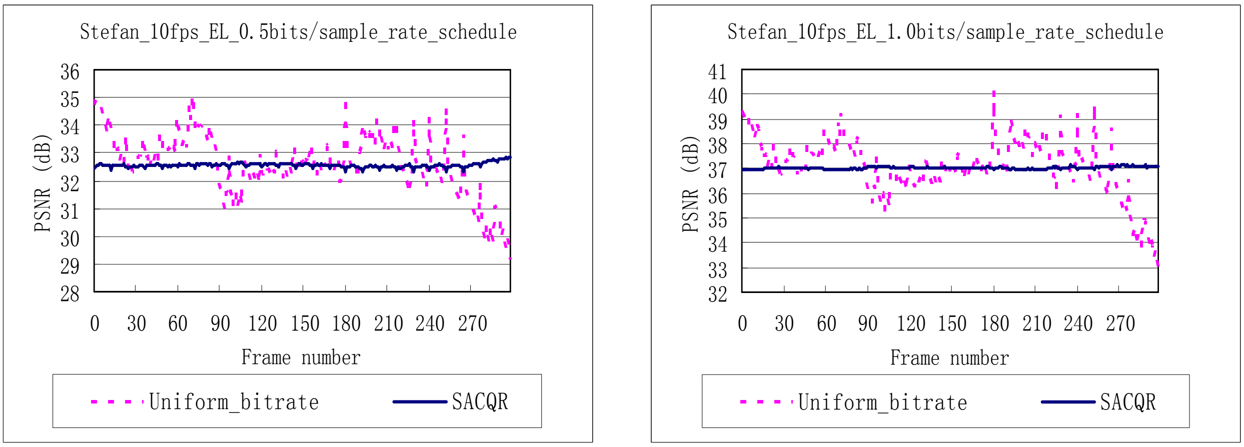

(a) (b)

**Fig. 2.** Comparison of quality variation between our SACQR algorithm and the uniform bit rate algorithm in the Foreman sequence. (a) Quality variation with bit rate 0.5 bits/sample, (b) Quality variation with bit rate 1.0 bits/sample.



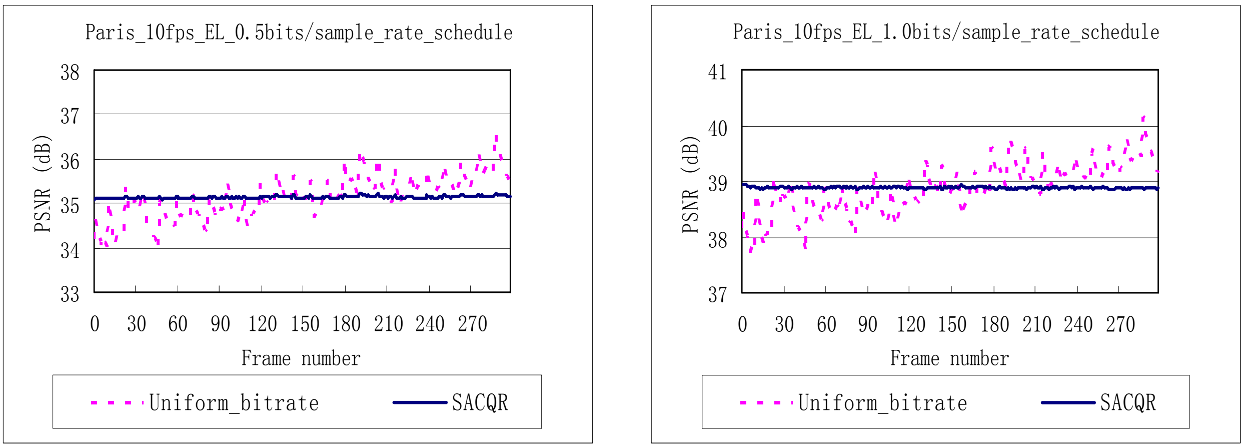

(a) (b)

**Fig. 3.** Comparison of quality variation between our SACQR algorithm and the uniform bit rate algorithm in the Tempete sequence. (a) Quality variation with bit rate 0.5 bits/sample, (b) Quality variation with bit rate 1.0 bits/sample.




(a) (b)

**Fig. 4.** Comparison of quality variation between our SACQR algorithm and the uniform bit rate algorithm in the Stefan sequence. (a) Quality variation with bit rate 0.5 bits/sample, (b) Quality variation with bit rate 1.0 bits/sample.




(a) (b)

**Fig. 5.** Comparison of quality variation between our SACQR algorithm and the uniform bit rate algorithm in the Paris sequence. (a) Quality variation with bit rate 0.5 bits/sample, (b) Quality variation with bit rate 1.0 bits/sample.