Statistical Model, Analysis and Approximation of Rate-Distortion Function in MPEG-4 FGS Videos

Jun Sun, Wen Gao, *Member, IEEE*, Debin Zhao, and Qingming Huang

***Abstract*—In this paper, the generalized Gaussian distribution is employed first to model the DCT coefficients of image data from MPEG-4 fine-granularity scalability (FGS) frame. Then, according to the quantization theory, the distortion-rate function of the generalized Gaussian model is analyzed and it is concluded that the derivative of the distortion-rate function first decreases, and then increases up to the boundary of 6.02 as the bit rate increases. For actual FGS coding, the derivative of actual distortion-rate function usually decreases as the rate increases, and then begins to increase slowly at a comparatively high bit rate. Finally, based on above observations, a rate-distortion (R-D) model is proposed to approximate the actual distortion-rate function. Experiments show that the proposed R-D model is accurate and flexible.**

***Index Terms*—Fine granularity scalability (FGS), MPEG-4, quantization theory, rate-distortion (R-D), statistical model.**

I. Introduction

THE bit-plane-based fine granularity scalability (FGS) is a standard coding tool of MPEG-4 for main streaming video applications [1]. It consists of one ordinarily encoded base layer (BL) and a bit-plane-encoded enhancement layer (EL). Typically, the BL is coded in constant bit rate (CBR) and targeted for minimum user network capacity. The EL, through the bit-plane coding of discrete cosine transform (DCT) residues, can be truncated at any point, allowing the quality of reconstructed frames to be continuously promoted as the received EL bits increase. Since the amount of information in video sequences is inherently variable, the BL generated by the popular CBR-based algorithms such as TM5 [2] usually exhibits significant quality fluctuation between frames [3], [4]. Therefore, some rate-distortion (R-D) based allocation algorithms should be employed to scale the EL during transmission so as to smooth the fluctuating quality of the BL. In this regards, it should be helpful to know the proper R-D function of each EL frame.

There are three main means to obtain R-D functions: analytical, empirical and heuristic. In analytical approaches, both coding system and source data are first analyzed and decomposed into components of known mathematical models. Then the R-D function of each component is acquired and combined together to form a complete R-D function for the whole coding system [5]. In empirical methods, several R-D data are first measured from the encoded video frame, then the R-D function of each video frame is obtained by smooth interpolation between the sampled R-D points [3], [6]. In heuristic approaches, the R-D model with several control parameters is first derived by the known R-D formula in R-D theory. Then the control parameters are estimated by the observed R-D data of the current coding frame or the previous coding frames [7].

Our research in this paper can be categorized into the analytical and heuristic approaches. With analytical method, we first decompose a video frame into 64 $8 * 8$ DCT independent components and employ the generalized Gaussian distribution (GGD) to model each DCT coefficient. Then according to the quantization scheme of FGS coding, the R-D function of the generalized Gaussian model is analyzed. It is concluded that the derivative of the distortion-rate function first decreases, and then grows up to the boundary of 6.02 as the bit rate increases. Finally, considering the distributions of DCT coefficients in actual FGS coding, we conclude that the derivative of the actual distortion-rate function usually decreases continuously as the bit rate increases, and then begins to increase slowly at a comparatively high bit rate. Guided by the above analysis and observations, we heuristically design a R-D model with three control parameters to approximate the actual R-D function. All the three control parameters of the heuristic model can be calculated by the actual R-D data of FGS EL to obtain an accurate approximation. Furthermore, the three control parameters can also be reduced to two or one for a rough approximation.

The rest of this paper is organized as follows. In Section II, the GGD is proposed to characterize the DCT coefficients. Section III analyzes the R-D function of FGS EL with the generalized Gaussian model. In Section IV, a simple heuristic R-D model is proposed to approximate the actual R-D function and extensive experiments are also provided. The conclusions are given in the last section.

II. Statistical Model of DCT Coefficients

A. Zero-Mean GGD

There have been several different assumptions on the distributions of the DCT coefficients. Generally the dc coefficient is modeled as a Gaussian distribution, and the ac coefficients are modeled as Laplacian distributions [8]. These two models are usually more popular due to their simplicity for mathematical tractability rather than their accuracy for describing the real source. The GGD was proposed by Müller [9] as an accurate model for the DCT coefficients of natural images. In this paper, we employ zero-mean GGDs to model the 64 $(8 * 8)$ DCT coefficients in a residue frame of FGS EL. The PDF of the distri-

Manuscript received June 7, 2005; revised November 30, 2005. This work was supported by National Science Foundation of China (60333020). This paper was recommended by Associate Editor H. Sun.

J. Sun and W. Gao are with the Institute of Computing Technology, Chinese Academy of Sciences, and also with the Graduate School of Chinese Academy of Science, Beijing 100080, China (e-mail: jsun@jdl.ac.cn; wgao@jdl.ac.cn).

D. Zhao is with the Department of Computer Science, Harbin Institute of Technology, Harbin 150001, China (e-mail: dbzhao@jdl.ac.cn).

Q. Huang is with the Graduate School of Chinese Academy of Science, Beijing 100080, China (e-mail: qmhuang@jdl.ac.cn).

Digital Object Identifier 10.1109/TCSVT.2006.871394

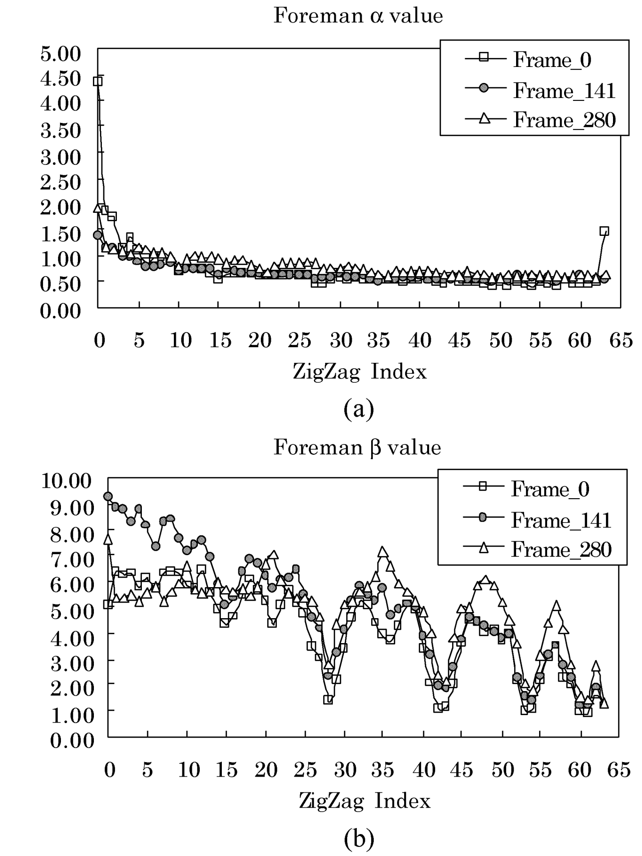

Fig. 1. GGDs' parameter estimation of 64 DCT coefficients in the I-frame 0, P-frame 141, and B-frame 280 of the Foreman sequences. (a) Shape parameter a estimate. (b) Root variance β estimate.

bution can be described as follows:

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

with

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

where $\alpha > 0$ is the shape parameter describing the exponential rate of decay, β is the variance of the random variable, and $\Gamma(\bullet)$ is the gamma function. For simplicity of denotation, zero-mean GGD is also called GGD in this paper.

B. Experimental Results of the Generalized Gaussian Model for the DCT Coefficients of FGS EL

To illustrate the distribution of the DCT coefficients, we present here some simulation results for the "Foreman" sequences in CIF (352 × 288) format. The test sequence is coded with the FGS reference software (MoMuSys version 2). The BL is coded with TM5 [2] at 128 kbits/s, and the frame rate is 10 frames/s (fps). The intra-period is 12 with the GOP mode "IBBP." All the DCT data including Y, Cr, and Cb in an EL frame are considered together.

Fig. 1 shows the estimates of α and β for all the 64 DCT coefficients respectively in the I-frame 0, P-frame 141, and B-frame 280 of the Foreman sequence. It can be found from Fig. 1 that not only the root variance β decreases typically, but also the shape parameter α shows the same trend of decrease in the zigzag order. This means that the coefficient distributions generally become sharper in the zigzag order. And most DCT coefficients have the shape value less than 1.0.

III. R-D Analysis of the FGS EL

In this section, we first present the quantization scheme of FGS EL. Then we analyze the R-D function of generalized Gaussian model and obtain some properties of the R-D function. At last, some conclusions of the actual R-D function of FGS EL are provided.

A. Quantization Scheme of FGS EL

In FGS video streaming, if the number of bit planes in an EL frame is n and the last transmitted bit plane 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:

$$T_i = i\Delta, i = \ldots, -N, \ldots, -1, 0, 1, \ldots, N, \ldots \tag{2a}$$

and

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

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

B. R-D Analysis of GGD

GGD is a nice statistical model to describe the DCT coefficients. It is important to analyze the R-D function of the distribution under the above quantization scheme of FGS EL. For simplicity of deduction, we rewrite the PDF of the GGD (1) as following:

$$p(x) = \frac{k_1(\alpha)}{\beta} \exp\left(-\left[k_2(\alpha)\frac{|x|}{\beta}\right]^{\alpha}\right) \tag{3}$$

where $k_1(\alpha) = \alpha[\Gamma(3/\alpha)]^{1/2}/2[\Gamma(1/\alpha)]^{3/2}$ and $k_2(\alpha) = [\Gamma(3/\alpha)/\Gamma(1/\alpha)]^{1/2}$. For an arbitrary variable $\delta > 0$, let the quantization step size $\Delta_\beta = \beta\delta$. If the reconstruction probability at the level $k\Delta_\beta$ is $P_k(\Delta_\beta)$, then we can have

$$P_k(\Delta_\beta) = \begin{cases} \int_{-\beta\delta}^{\beta\delta} \frac{k_1(\alpha)}{\beta} \exp\left(-\left[k_2(\alpha)\frac{|x|}{\beta}\right]^{\alpha}\right) d_x \\ = \int_{-\delta}^{\delta} k_1(\alpha) \exp(-[k_2(\alpha)|x|]^{\alpha})\, d_x, & \mathrm{k}=0 \\ \int_{k\beta\delta}^{(k+1)\beta\delta} \frac{k_1(\alpha)}{\beta} \exp\left(-\left[k_2(\alpha)\frac{|x|}{\beta}\right]^{\alpha}\right) d_x \\ = \int_{k\delta}^{(k+1)\delta} k_1(\alpha) \exp(-[k_2(\alpha)|x|]^{\alpha})\, d_x, & \mathrm{k} \ge 1 \\ P_{-k}(\Delta_\beta), & \mathrm{k} \le -1. \end{cases} \tag{4}$$

For a specific shape parameter α, it is easy to see that $P_k(\Delta_\beta)$ is just determined by the variable δ. The entropy rate can be calculated by

$$H(\Delta_\beta) = \sum_{k=-\infty}^{\infty} -P_k(\Delta_\beta)\log_2 P_k(\Delta_\beta). \tag{5}$$

So for the same shape parameter α, the entropy rate is also just determined by δ. That is $H(\Delta_\beta) = f_\alpha(\delta)$. From the meaning of quantization, it is easy to see that $f_\alpha(\delta)$ is a decreasing function. So we can have

$$\delta = f_\alpha^{-1}\left(H(\Delta_\beta)\right). \tag{6}$$

For the GGD, the quantization scheme in (2) is symmetric about zero, so the mean square error (MSE) with the quantization step size Δ_β can be computed as

$$D(\Delta_\beta) = 2\sum_{k=0}^{+\infty} D_k(\Delta_\beta) \tag{7}$$

where

$$\begin{aligned} D_k(\Delta_\beta) = & \int_{k\beta\delta}^{(k+1)\beta\delta} (x - k\beta\delta)^2 \frac{k_1(\alpha)}{\beta} \\ & \times \exp\left(-\left[k_2(\alpha)\frac{|x|}{\beta}\right]^\alpha\right) d_x \\ = & \int_{k\delta}^{(k+1)\delta} (\beta x - k\beta\delta)^2 k_1(\alpha) \exp\left(-\left[k_2(\alpha)|x|\right]^\alpha\right) d_x \\ = & \beta^2 \int_{k\delta}^{(k+1)\delta} (x - k\delta)^2 k_1(\alpha) \exp\left(-\left[k_2(\alpha)|x|\right]^\alpha\right) d_x. \end{aligned} \tag{8}$$

Then if we use peak signal-to-noise ratio (PSNR) as the distortion criteria, we can obtain

$$\begin{aligned} & \mathrm{PSNR}(\Delta_\beta) \\ & = 20\log_{10} 255 - 10\log_{10} D(\Delta_\beta) \\ & = 20\log_{10} 255 - 10\log_{10} 2\sum_{k=0}^{+\infty} D_k(\Delta_\beta) \\ & = 20\log_{10}\left(\frac{255}{\beta}\right) - 10\log_{10}(2 \\ & \quad \times \sum_{k=0}^{+\infty} \int_{k\delta}^{(k+1)\delta} (x - k\delta)^2 k_1(\alpha) \exp\left(-\left[k_2(\alpha)|x|\right]^\alpha\right) d_x). \end{aligned} \tag{9}$$

The second part in (9) can be denoted as $g_\alpha(\delta)$. Using (6), (9) can be rewritten as

$$\begin{aligned} & \mathrm{PSNR}(\Delta_\beta) \\ & = 20\log_{10}\left(\frac{255}{\beta}\right) + g_\alpha(\delta) \\ & = 20\log_{10}\left(\frac{255}{\beta}\right) + g_\alpha\left(f_\alpha^{-1}\left(H(\Delta_\beta)\right)\right). \end{aligned} \tag{10}$$

From (10), we can see that for the same shape parameter α and different root variance β, the rate distortion curves of GGD have the same shape, that is, $d_{\mathrm{PSNR}(\Delta_\beta)}/d_{H(\Delta_\beta)}$ is not relative to β.

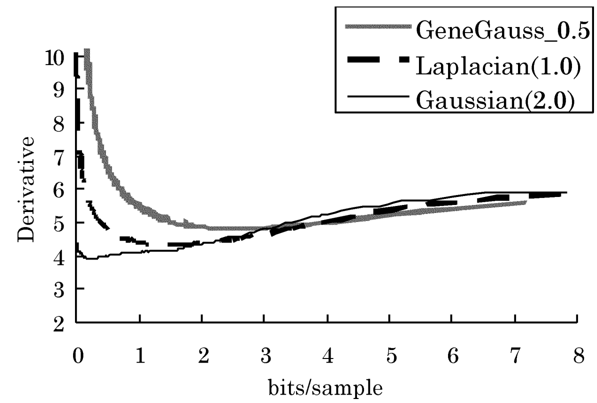

Fig. 2. Derivative comparison of the distortion-rate with the GGD shape parameter 0.5, 1.0 (Laplacian), and 2.0 (Gaussian).

We have done some computer simulations to calculate the derivative of the distortion-rate function of the GGD with different shape parameter α. Fig. 2 shows the derivative comparison of the distortion-rate functions of the GGD with shape parameter 0.5, the Laplacian distribution (shape parameter 1.0) and the Gaussian distribution (shape parameter 2.0). Generally, the derivative of the distortion-rate functions (PSNR criterion) decrease first and then increase to the traditional number 6.02. The inflexion of the derivative becomes larger if the shape parameter α 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.

C. *Properties of the Actual Distortion-Rate Function*

In Section III-B, we have analyzed that, for Laplacian distribution, the derivative of the distortion-rate function (PSNR criterion) decreases to about 4.34 and then increases to 6.02 gradually as the coding rate increases. Most DCT coefficients have the shape value less than 1.0 in actual video frames (see Fig. 1), 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 bit planes at a high bit rate, and the bit planes are not independent from each other in fact. But these bit planes are entropy coded independently. Therefore, as bit rate increases, the efficiency of actual bit-plane coding is decreased [10]. This also delays and slows down the increase of the derivatives of actual distortion-rate function, 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.

Table I shows the derivative of some frames' actual distortion-rate functions in the 10 fps Foreman and Tempete_ext sequences, where the Tempete_ext sequence is the concatenation of two 260-frame Tempete sequences. On some points of very low bit rate, the derivative of the distortion-rate function is smaller than that of the following point just because the header information occupies much of the bit rate compared with the coefficient information in these points. The derivatives of the Foreman frame 0 and 141 decrease continuously. It also can be seen from the Foreman frame 280 and the Tempete_ext frame

TABLE I
DERIVATIVE OF ACTUAL DISTORTION-RATE FUNCTION IN THE FOREMAN AND TEMPETE EXT SEQUENCES

Sequences \ Frames		Frame 0			Frame 141			Frame 280		
		Rate	PSNR	Derivative	Rate	PSNR	Derivative	Rate	PSNR	Derivative
Foreman Bitplanes beginning	0	0.0000	35.258	-	0.0000	33.430	-	0.0000	33.941	-
	1	0.0044	35.285	5.931	0.0173	33.781	20.353	0.0120	34.093	12.652
	2	0.1180	36.681	12.304	0.1798	35.887	12.954	0.1557	35.560	10.204
	3	0.4875	40.368	9.974	0.6113	39.634	8.685	0.6330	38.955	7.113
	4	1.1333	45.205	7.491	1.3634	44.672	6.698	1.5403	44.092	5.662
	5	2.1030	51.536	6.529	2.4199	51.406	6.374	2.7466	51.262	5.944
Tempete_ext Bitplanes beginning	0	0.0000	28.184	-	0.0000	28.070	-	0.0000	26.786	-
	1	0.0059	28.228	7.446	0.0078	28.152	10.398	0.0029	26.844	19.653
	2	0.1455	30.145	13.734	0.1467	29.965	13.054	0.0137	26.859	1.429
	3	0.5285	33.743	9.396	0.5509	33.609	9.017	0.0406	27.063	7.583
	4	1.1795	38.418	7.181	1.2214	38.296	6.990	0.2387	29.468	12.142
	5	2.0759	43.978	6.203	2.1364	43.941	6.170	0.6832	33.425	8.903
	6	3.2418	51.257	6.243	3.2863	51.274	6.377	1.3699	38.214	6.973
	7	-	-	-	-	-	-	2.3012	43.864	6.068
	8	-	-	-	-	-	-	3.4729	51.226	6.283

0, 141, 280 that the derivatives decrease at first and begin to increase at a high bit rate, and that the increase of the derivative is very slow and can be neglected.

IV. APPROXIMATION OF THE ACTUAL R-D FUNCTION

A. Proposed R-D Model

What we have known allows us to construct a new heuristic R-D model to approximate the actual distortion-rate function. The main idea is to utilize the properties in Section III-C, that is, we can assume that the derivative of distortion-rate function (PSNR criterion) decreases continuously as the bit rate increases in the actual FGS coding. Heuristically we can suppose that the derivative decreases to a constant. It is evident that the model to be constructed should have an asymptote $a * R + A$, where R is the bit rate per sample, A and a are the asymptote parameters. Since the derivative of distortion-rate function decreases to a as the bit rate increases, we can use a simple power function $-1/R$ to model the decreasing property ($(-1/R)' = 1/R^2$ decreases to zero as R increases). And from Fig. 2, we also notice that the derivative decreases slower as the bit rate increases. That is, the second-order derivative of the distortion-rate function should increase to zero as the bit rate increases. $-1/R$ is also good enough to model the second-order derivative ($(-1/R)'' = -2/R^3$ increases to zero as R increases). Since the domain of bit rate definition should have zero, we change $-1/R$ to $-1/(1 + b * R)$, where $b > 0$ is a parameter controlling the approach of the actual distortion-rate function to the asymptote. Finally, our model can be described as follows:

$$\mathrm{PSNR}(R) = a * R + A - \frac{k}{(1 + b * R)}. \tag{11}$$

When the bit rate $R = 0$, the $\mathrm{PSNR}(0)$ is the PSNR of the BL coding B. Then we can obtain $k = A - B$. So the complete formula of the proposed R-D model can be described as follows:

$$\mathrm{PSNR}(R) = a * R + A - \frac{(A - B)}{(1 + b * R)}. \tag{12}$$

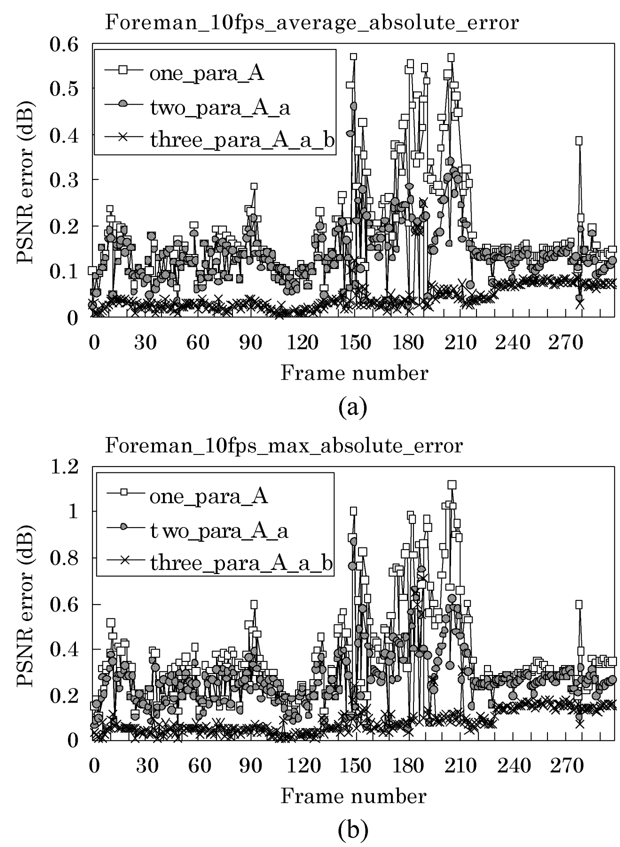

Fig. 3. Average and maximum absolute estimate error of our model in the three experiments of the 10 fps Foreman CIF sequence. (a) Average absolute estimate error. (b) Maximum absolute estimate error.

Generally, the parameter a is between 3.0 and 6.5, and the parameter b is between 0.5 and 6.0. In the approximation of the actual distortion-rate function, a and b could also be selected to be constants for a coarse approximation. Within the context of video coding, the parameters can also be estimated through the parameters of the previous frames. In this section, to verify the R-D model, the control parameters are all estimated through the R-D data of the current video frame.

TABLE II
AVERAGE OF ABSOLUTE ESTIMATE ERROR AND THE AVERAGE OF MAXIMUM ABSOLUTE ESTIMATE ERROR ACROSS ALL CODED FRAMES

Average across all frames	Foreman			Stefan			Tempete			Paris		
	Exp1	Exp2	Exp3	Exp1	Exp2	Exp3	Exp1	Exp2	Exp3	Exp1	Exp2	Exp3
Average error (dB)	0.18	0.14	0.04	0.41	0.06	0.03	0.16	0.09	0.06	0.18	0.04	0.02
Maximum error (dB)	0.36	0.27	0.09	0.77	0.13	0.08	0.38	0.18	0.12	0.37	0.09	0.04

B. Experimental Verification of our R-D Model

The proposed R-D Model has three parameters (a, b and A) that need to be estimated through the actual R-D data of the coded video frame. In this subsection, we also use the MoMuSys MPEG-4 codec to obtain the actual R-D information as in Section II-B. The BL bitrate target is 128 kbps with TM5 [2]. The frame rate is 10 fps in the experiments. The four test videos are Foreman, Stefan, Tempete_ext, and Paris CIF sequences. The four sequences were selected because of their different characteristics, i.e., Foreman and Stefan have rapid motion and changing background while Tempete_ext and Paris have little motion and stationary background. The R-D information of each frame is obtained at the beginning of each bit plane of the EL frame.

To verify our R-D model, we design three experiments. In the first experiment, we set a as the constant 5.5 and b as the constant 1.5. The parameter A is gained by the nonlinear least-squares data fitting. In the second experiment, we set b as the constant 1.5. The other two parameters (a and A) of the asymptote are calculated using the same nonlinear least-squares data fitting. In the third experiment, all the three parameters (a, b and A) of the R-D model are acquired through the nonlinear least-squares data fitting.

Fig. 3 shows the average absolute estimate error of the R-D model (over all the R-D data, that is, the R-D points at the beginning of each bit plane in a frame) for the 10 fps Foreman. Table II shows the mean of the maximum and average absolute estimate error across all coded frames for the Foreman, Stefan, Tempete_ext, and Paris CIF sequences respectively. It is observed that the average absolute estimate error of our complete model on all the test videos is just about 0.0375 dB in the third experiment. It is accurate enough for almost all R-D related application. It also can be noted that the maximum absolute estimate error of the model in all the three experiments is just about two times than the average absolute estimate error. From the experimental results, we can see that generally, the proposed R-D model is accurate and flexible, and usually, two-control-parameter model is good enough to balance the accuracy and the complexity of the R-D model.

V. CONCLUSION

In this paper, we analyze the R-D function for MPEG-4 FGS video and present an approximation of the R-D function. There are three points we have contributed in this work. First, we analyze the distribution of DCT coefficients using the generalized Gaussian model. Second, we obtain some properties of actual R-D function of FGS EL through a thorough R-D analysis of the generalized Gaussian model. Generally speaking, the derivative of distortion-rate function (PSNR criterion) of FGS EL first decreases as the bit rate increases, and then begins to increase slowly at a comparatively high bit rate. Third, we present a new R-D model to accurately and flexibly approximate the actual R-D function. All the above analysis and models bring us much insight into the FGS coding and its R-D function.

REFERENCES

[1] W. Li, "Overview of fine granularity scalability in MPEG-4 video standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 3, pp. 301–317, Mar. 2001.
[2] *MPEG Test Model 5*, ISO/IES JTC/SC29/WG11 Document, Apr. 1993.
[3] X.-M. Zhang, A. Vetro, Y.-Q. Shi, and H. Sun, "Constant quality constrained rate allocation for FGS-coded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 2, pp. 121–130, Feb. 2003.
[4] M. Dai and D. Loguinov, "Analysis of rate-distortion functions and congestion control in scalable Internet video streaming," in *Proc. NOSSDAV'03*, Monterey, CA, Jun. 2003, pp. 60–69.
[5] H.-M. Hang and J.-J. Chen, "Source model for transform video coder and its application. I. Fundamental theory," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 2, pp. 287–298, Apr. 1997.
[6] L. J. Lin and A. Ortega, "Bit-rate control using piecewise approximated rate-distortion characteristics," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 8, no. 4, pp. 446–459, Aug. 1998.
[7] T. Chiang and Y. Q. Zhang, "A new rate control scheme using quadratic rate distortion model," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, no. 1, pp. 246–250, Feb. 1997.
[8] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. Image Process.*, vol. 9, no. 10, pp. 1661–1666, Oct. 2000.
[9] F. Müller, "Distribution shape of two-dimensional DCT coefficients of natural images," *Electron. Lett.*, vol. 29, no. 22, pp. 1935–1936, Oct. 1993.
[10] C. E. Shannon, "A mathematical theory of communication," in *Claude Elwood Shannon: Collected Papers*, N. J. A. Sloane and A. D. Wyner, Eds. Piscataway, NJ: IEEE Press, 1993, pp. 5–83.