A Novel Statistical Distortion Model Based on Mixed Laplacian and Uniform Distribution of Mpeg-4 FGS

Xie Li and Wenjun Zhang

Institute of Image Communication and Information Processing,
Shanghai Jiaotong University, 1954 Huashan Road, Shanghai, China. 200030
Email: lixie@sjtu.edu.cn, lixie2004@gmail.com

The Laplacian Distribution has been widely used to characterize the distribution of DCT coefficients in a great deal of research on statistical attributes of DCT for the reason that Laplacian distribution can achieve a good balance between the simplification of the model and fidelity to the real empirical data. But in our research, we found that there is a large discrepancy between the real data and the distortion model based on Laplacian distribution. Furthermore, the higher the rate, the larger the distortion. In this paper, we proposed a new distortion model based on mixed Laplacian and Uniform distributions. Experimental results show that the MLQ distortion model can approximate the real empirical data and has a higher degree of accuracy than the conventional distortion model based on Laplacian distribution. Compared with the conventional distortion model based on Laplacian distribution, the new model can decrease the MSE distortion by 2–6 units. The MLQ distortion model can represent the relationship of rate and distortion of the Mpeg-4 FGS enhancement layer with a higher degree of accuracy.

Keywords: Mpeg 4 FGS, Bit Plane, Scalable coding, Distortion, Statistical model

ACM Classification: I.4.2

1. INTRODUCTION

The research on the statistical attributes of Discrete Cosine Transform (DCT) coefficients began more than 20 years ago. Early on, DCT coefficients were widely regarded as having Gaussian distribution. However, soon plenty of experiments showed that DCT coefficients have Laplacian distribution instead of Gaussian distribution (Reininger and Gibson, 1983). Since then many researchers have tried to understand the statistical distribution of DCT coefficient and proposed a variety of possible distributions, such as Cauchy and general Laplacian distribution (Smart and Rowe, 1996; Yovanof and Liu, 1997). But those proposed distribution models only apply to one category of video sequences but fail on other categories of video sequences. Laplacian distribution has often been used to model the distribution of DCT coefficients for its simplification and fidelity to the real data.

Scalable coding is a new video encoding technology designed specifically for the Internet environment. Many researchers have demonstrated a great deal of interest in this research area since the Mpeg-2 was proposed and standardized. Ever since Mpeg-4 FGS technology was adopted into the Mpeg-4 Standards as a standardized scalable coding technology, Mpeg-4 FGS has been extensively implemented in many Internet applications and has become the de facto industry

Manuscript received: 24 June 2005
Communicating Editor: Geof McLachlan

standard. As is well known, the Mpeg-4 FGS encoder comprises two layers of encoders: the first layer encoder is a base layer encoder which will produce a low-quality bit stream and the second layer encoder is an enhancement layer encoder which encodes the residues from the base layer encoder. The enhancement layer encoder is a scalable encoder and the bit streams of the enhancement layer can be used to improve the output quality (MPEG-4, 2001). The relationship between rate and distortion of non-scalable codec has been intensively studied in the past two decades, but the research for scalable codec has been largely ignored. The relationship between rate and distortion of scalable codec is very complex because it is not only related closely to the implementation of the enhancement layer encoder but also influenced by the performance of the base layer encoder. Even now, we are still not very clear about it.

In this paper, we first assume that the DCT coefficient has MLQ distribution, and then propose a distortion model based on the MLQ distribution, mathematically analyze the relationship between rate and distortion of Mpeg-4 FGS enhancement layer. The results of experiments show that the novel distortion model based on mixed Laplacian and Uniform distribution can approximate the real empirical data very well. Compared to the conventional distortion model based on the Laplacian, the new model can decrease the MSE distortion by 2–6 units. The MLQ distortion model can represent the relationship of rate and distortion of the Mpeg-4 FGS enhancement layer with a higher degree of accuracy.

This paper is divided into the following parts: Section 2 is an overview of statistic models of signal sources; Section 3 explores the detail of implementation of Bit Plane encoding technology and the methods of quantisation, and then derives the new proposed MLU Rate-Distortion model. Section 4 gives the results of experiments; Section 5 presents our conclusions drawn from our research experiments and the follow-up research. The results of the experiments show that our proposed distortion model can approximate the real empirical data very well.

2. THE STATISTIC DISTRIBUTION MODEL OF DCT COEFFICIENTS

As mentioned in the first paragraph, DCT coefficients have a complex statistic distribution and many models have been proposed to model them. Unfortunately most of these distributions only apply to one category of video sequences but fail on other categories of video sequences. In normal situations, Laplacian distribution is used to model the statistic distribution of DCT coefficients because Laplacian distribution can achieve a good balance between the simplification of the model and fidelity to the real empirical data (Lam and Goodman, 2000). As shown in Figure 1, Laplacian

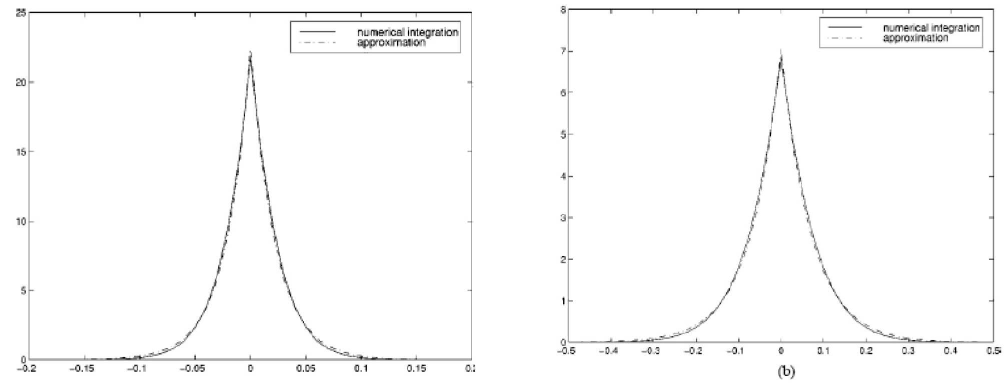
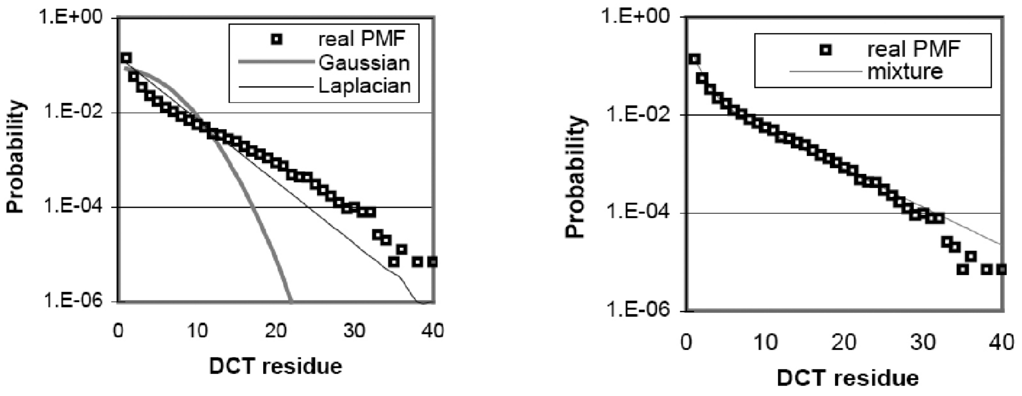


Figure 1: DCT Coefficients has Laplacian distribution



(a) Laplacian Model **(b) Mixed Laplacian Model**

Figure 2: Mixed Laplacian distribution models the real DCT residues with a higher accuracy.

distribution can fit the real distribution very well (Lam and Goodman, 2000; Wu and Cham, 1990; Guglielmo, 1986; Lakhani, 2000). However Laplacian distribution is widely applied simply because of its mathematical simplicity rather than its accuracy in modelling the actual data.

Similar to the distribution of DCT coefficients, the distribution of DCT residues is also very complicated and difficult to formulate simply. Recently some research has demonstrated that Laplacian distribution couldn't model the tail of the distribution of real DCT residues, as shown in Figure 2 (Dai, 2003; Dai, 2004). Therefore, a Mixed Laplacian distribution was proposed to model the distribution of DCT coefficients. Figure 2 shows that the double Laplacian distribution could model the real DCT coefficients and the tail of the distribution of the real DCT residues with a higher degree of accuracy.

In the paper, we use mixed Laplacian and Uniform distribution to model the real distribution of DCT residues. For the remainder of this paper, the distortion model based on mixed Laplacian and Uniform distribution will be referred to as the MLU Model and the distortion model based on Laplacian distribution as the Laplacian Model.

Suppose x is the random variable and x has a mixture of a Laplacian and Uniform distribution with the following PDF:

$$p(x) = \begin{cases} \lambda e^{-\lambda x}, x \in (0, \Delta] \\ const, L > x > \Delta \end{cases}$$

The MLU model is comprised of two different distributions: the first one is a Laplacian distribution modelling the distribution of DCT residue locating the $(0,\Delta]$ range. The second distribution is Uniform distribution modelling the real distribution of DCT residue outside of the range $(0,\Delta]$.

Our proposed MLU model is simple enough to be implemented easily. The results of the experiments show that the MLU model has a higher degree of accuracy than conventional Laplacian models.

3. THE DISTORTION MODEL OF MPEG-4 FGS ENHANCEMENT LAYER

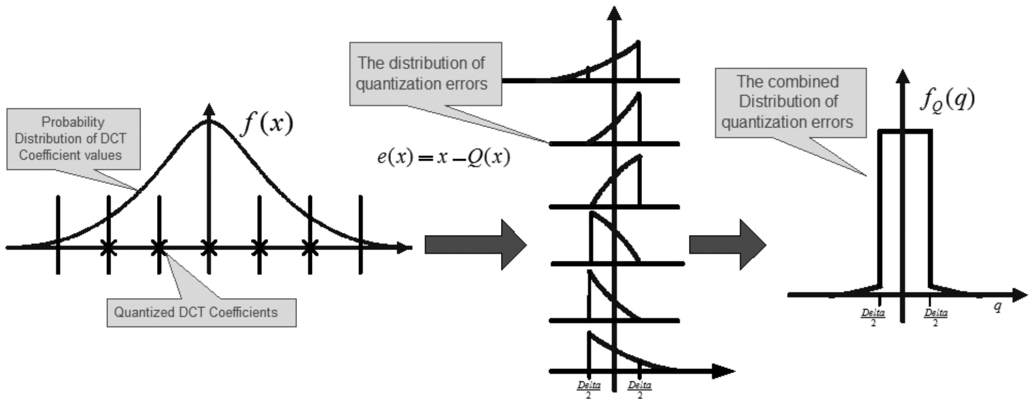
In this section, we are going to explore the implementation of Bit Plane encoding technology and the methods of quantisation, and then derive the new proposed MLU Rate-Distortion model.

3.1 The quantisation methods of Mpeg-4 FGS Enhancement Layer

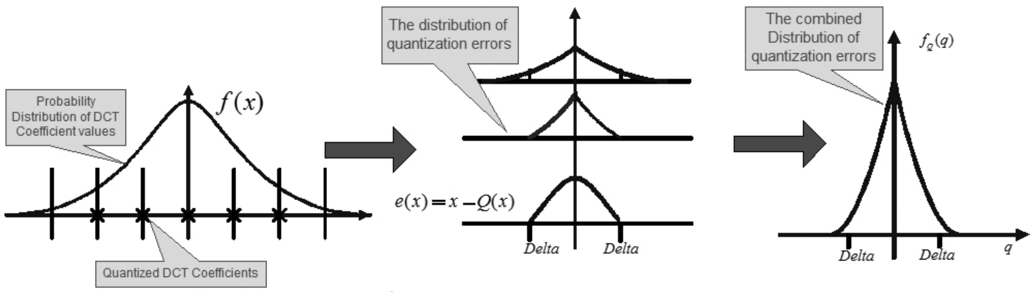
Bit Plane encoding technology is used to encode the DCT residues produced by the encoder of the base layers in the Mpeg-4 FGS encoders. The most important residues will be encoded first and placed at the beginning of the compressed bit streams, the less important residues will be placed at the end. The compressed bit streams can be truncated at any point in the bit stream, and that truncated bit stream can be decoded correctly by the decoder on the decoder side. The decoded data is used to improve the output video quality. The truncation process can be regarded as a special quantisation process that is one kind of embedded quantisation.

The approach of quantisation used to encode the enhancement layer of Mepg-4 FGS bit streams is different to the methods used in the conventional quantisation process. The Conventional quantisations always use the centroid of the quantisation bin as the reconstructed value, but the Mpeg-4 FGS quantisation process use the low boundary as the reconstructed value. See Figure 3.(Li and Zhang, 2005)

The conventional quantisation method is shown in Figure 3(a). The left-most figure illustrates the probability distribution of DCT coefficients, the intervals along the horizontal axis are the quantisation bins, and the centroid of each quantisation bin is the reconstructed DCT coefficients. The middle figure shows the distributions of quantisation errors in quantisation bins. The right-most



(a) Conventional Encoding



(b) FGS BitPlane Encoding

Figure 3: MPEG-4 FGS DCT Coefficient Distribution And DCT Residues Distribution

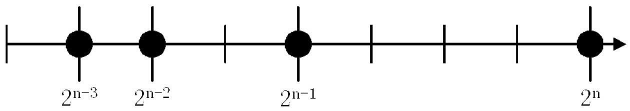


Figure 4: Embedded Quantisation

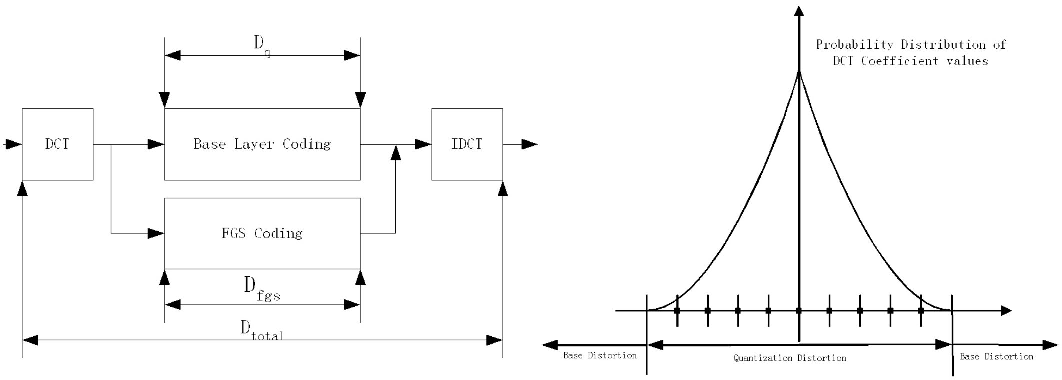
right figure shows the probability distribution of DCT residues. The compounded probability distribution of DCT residues becomes almost uniform.

The quantisation method used in the Mpeg-4 FGS encoding procedure is shown in Figure 3(b). The left-most figure shows the probability distribution of DCT coefficients, the middle figure shows the distributions of quantisation errors in quantisation bins, and the right-most figure shows the probability distribution of DCT residues. It is obvious that the compounded probability distribution of DCT residues does not have uniform distribution and the probability distribution in each quantisation bin is different to the others.

The distortion characteristic in the Mpeg-4 FGS enhancement layer is different to the distortion characteristic in the base layer. The difference is caused by the different quantisation methods used in the encoder of the enhancement layer. The quantisation method is one kind of embedded quantisation with the following feature: the quantisation error produced by coarse-scale quantisation can be quantised by finer-scale quantisation, and the reconstructed values produced by coarse-scale quantisation are distributed among the reconstructed values of the finer-scale quantisation. The feature is shown in Figure 4.

For example, the 2^n quantisation is a coarse-scale quantisation, the 2^{n-1},2^{n-2},… are finer-scale quantisations. The quantisation error produced by 2^n quantisation will be further quantized at the 2^{n-1},2^{n-2},… scales.

The distortion distribution in the Mpeg-4 FGS encoder is shown in Figure 5. *Dq* is the upper boundary of *Dfgs*. It is often the case that the difference between *Dtotal* and *Dfgs* is set to 0, –which would therefore imply that *baseDistortion* is 0. But in our research, we found that *baseDistortion* should not be set to 0. When all of the enhancement data are used to improve the output video quality, the output video quality of the Mepg-4 FGS decoder will reach the maximum. If *baseDistortion* is 0, there is no difference between the output video and original video. But in fact,



A. Distortion Distribution **B. Distribution in Probability Space**

Figure 5: Distribution of Distortion Location

it is not true. So *baseDistortion* should not be set to 0. *baseDistortion* is determined by the performance of the base layer encoder, and independent of the encoded video sequences. In our research, we set *baseDistortion* to 0.0802 based on the empirical data.

3.2 Distortion Model based on Laplacian Distribution

Suppose a random variable x has Laplacian distribution with the following PDF

$$p(x) = \frac{1}{2b} e^{-\frac{|x|}{b}} \tag{1}$$

The sign of each residue will only be considered when encoding the most significant bit of the DCT residues in the Bit Plane encoding process and the process can be regarded as performing absolution operations on DCT residues. Because the random variable x has Laplacian distribution, the random variable $|x|$ has exponential distribution, the random variable $|x|$ can be denoted by the following PDF

$$p(x) = \lambda e^{-\lambda x}, \quad \lambda > 0, x \geq 0 \tag{2}$$

Let x denote the random variables with exponential distribution. The quantisation distortion can be written as

$$D = \sum_{i=0}^{N-1} \int_{i*\Delta}^{i*\Delta+\Delta} (x - i*\Delta)^2 * \lambda e^{-\lambda x} dx \tag{3}$$

Where Δ is the quantiser step, N is the number of quantisation bins. Further simplify the equation and distortion can be rewritten as:

$$D = \frac{2}{\lambda^2}\left(1 - e^{-N\Delta\lambda}\right) - \frac{\Delta\lambda(2+\Delta\lambda)}{\left(-1+e^{\Delta\lambda}\right)\lambda^2}\left(1 - e^{-N\Delta\lambda}\right) \tag{4}$$

In most situations, the upper limit of $1 - e^{-N\Delta\lambda}$ approaches 1 when $N\Delta\lambda$ is large enough, and then we can rewrite the distortion as the follows.

$$D == \frac{2}{\lambda^2} - \frac{\Delta\lambda(2+\lambda\Delta)}{\left(e^{\Delta\lambda} - 1\right)\lambda^2} \tag{5}$$

We draw the $D \sim \Delta$ curve with $\lambda = 1$ in Figure 6. From Figure 6, we can see that:

1. The smaller the quantiser step, the smaller the distortion. When there is no quantisation, the distortion reaches the minimum, which means that all the data of the enhancement layer will be used to improve the video quality. In this case, the distortion of the Mpeg-4 FGS encoder is caused only by the DCT and Inverse Discrete Cosine Transform (IDCT) transforms the operation and the quantisation error is zero.
2. The maximum distortion is equal to the expectation of the square of x when the quantiser step is large enough to quantise all the DCT residues to 0. In this situation, all the data of the enhancement layer will be discarded; no data of the enhancement layer will be used to improve the output video quality.

Based on the distortion analysis in section 3.1, we can rewrite the distortion as the following:

$$D = baseDistortion + \frac{2}{\lambda^2} - \frac{\Delta\lambda(2+\lambda\Delta)}{\left(e^{\Delta\lambda} - 1\right)\lambda^2} \tag{6}$$

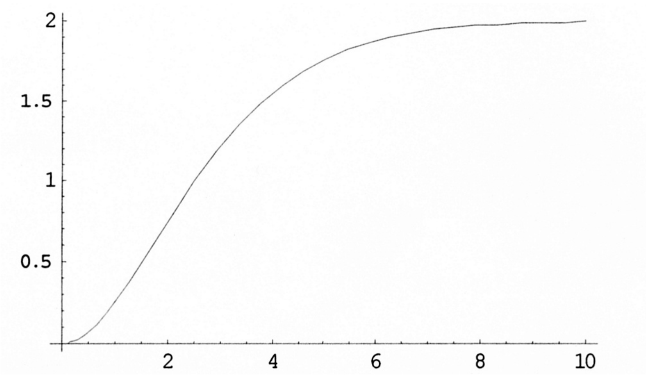


Figure 6: *D*~ Δ Curve. The relationship between distortion and quantiser step

3.3 MLU Distortion Model based on Mixed Laplacian and Uniform distribution

Let D denote the distortion and x denote the random variable, the distortion can be written as:

$$D = baseDistortion + \int_0^{\Delta} x^2 \lambda * e^{-\lambda x} dx + \sum_{i=1}^{N-1} \int_{i*\Delta}^{(i+1)*\Delta} c * (x - i * \Delta)^2 dx \tag{7}$$

baseDistortion is determined by the performance of the base layer encoder. What we are going to do is to analyze the distortions in two separate ranges.

1. [0,Δ) Range. DCT coefficients have exponential distribution in this range, and the distortion can be rewritten as:

$$D_{[0,\Delta)} = \int_0^{\Delta} x^2 \lambda * e^{-\lambda x} dx \tag{8}$$

Then we get

$$D_{[0,\Delta)} = 1 - e^{\Delta * \lambda} \tag{9}$$

2. $[\Delta, N * \Delta)$ Range. DCT coefficients have uniform distribution in this situation, and the distortion can be rewritten as

$$D_{[\Delta, N*\Delta)} = \sum_{i=1}^{N-1} \int_{i*\Delta}^{(i+1)*\Delta} c * (x - i * \Delta)^2 dx \tag{10}$$

Then we get

$$D_{[\Delta, N*\Delta)} = \frac{\Delta^2}{3}$$

Based on the above analysis, we can rewrite the distortion as:

$$D = baseDistortion + 1 - e^{\lambda\Delta} + \frac{\Delta^2}{3} \tag{11}$$

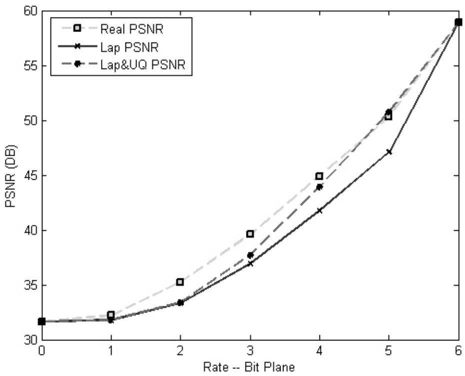
In our experiments, we set the *baseDistortion* to 0.0802

4. EXPERIMENTS AND RESULTS

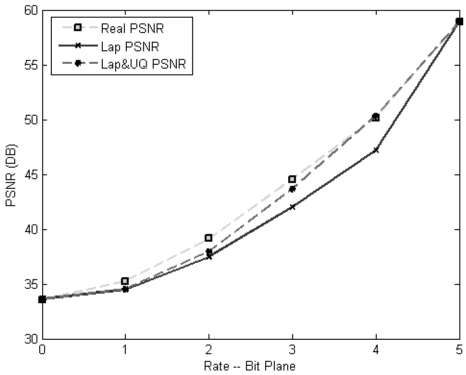
In our experiments, we use the Mpeg-4 reference software (MoMuSys with FGS support) to produce a compressed Mpeg-4 FGS bit stream. The two raw video sequences are a Foreman raw video sequence and a Mobile raw video sequence. They are both (define "CIF") CIFs and each contains 300 frames. In order to validate our proposed MLU distortion model of the Mpeg-4 FGS enhancement layer, we first collect the real empirical data during the encoding process, and then compare the output of different distortion models. We use PSNR as measurement of the distortion and use bit planes as the measurement of the rate. The results based on the Foreman CIF are shown in Figure 7 and the results based on the Mobile CIF are shown in Figure 8.

We randomly selected frames 2, 50, 135 and 272 from the 300 frames, and show the experiment results in Figure 8.

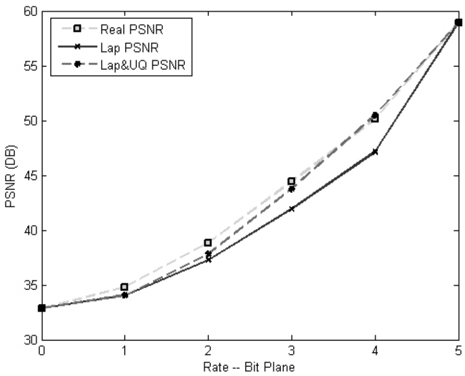
From Figure 7 and 8, we can see the Laplacian model can approximate the real *PSNR ~ Rate* curve very well when the rate is low, but has a big discrepancy with the real data when the rate is high. Meanwhile the MLU model can approximate the real *PSNR ~ Rate* curve very well throughout all the rates, be it low or high. Results of experiments show that the novel MLU model has a higher degree of accuracy to the real empirical data than the conventional distortion model based on the Laplacian model.



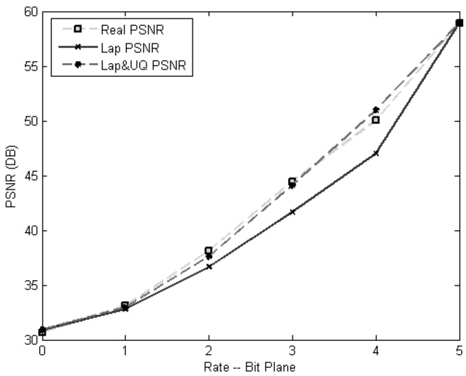
Frame 2



Frame 50

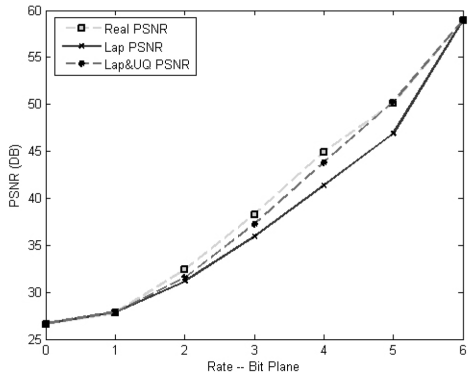


Frame 135

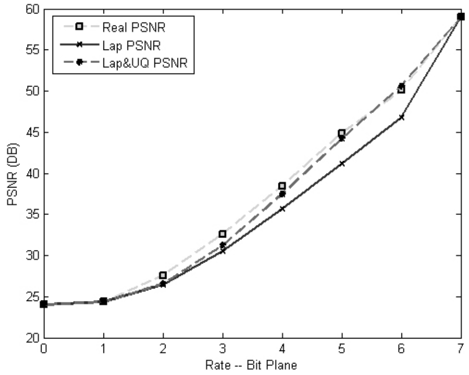


Frame 272

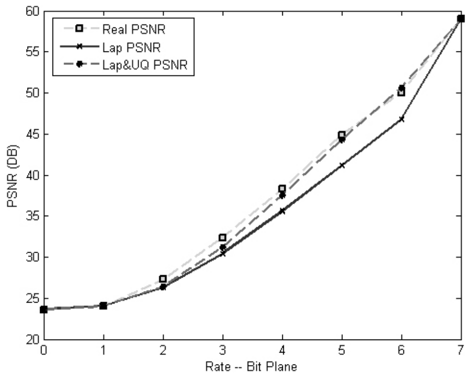
Figure 7: MLU Model for Foreman CIF



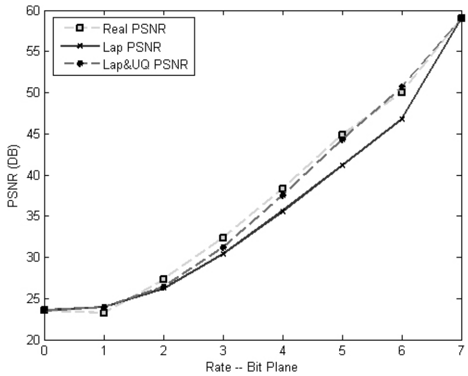
Frame 2



Frame 50

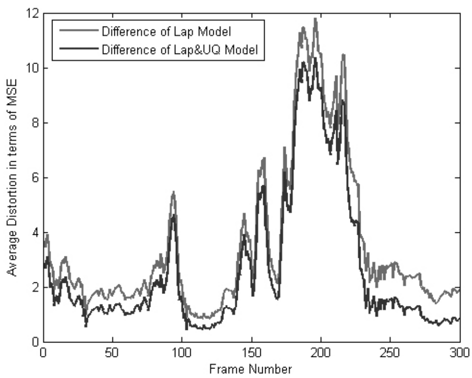


Frame 135

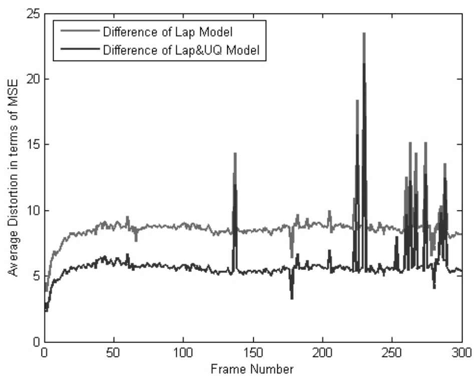


Frame 272

Figure 8: MLU Model for Mobile



(a) Foreman



(b) Mobile

Figure 9: The differences between two kinds of distortion model and real empirical data

To demonstrate the accuracy of our proposed new MLU distortion model, we compare the average absolute errors (measured in DB and averaged across the bit planes) between the MLU model and the Laplacian model for two CIF video sequences. Figure 9a shows the results in the Foreman CIF and Figure 9b shows the results in the Mobile CIF. Both results show that the MLU distortion model based on mixed Laplacian and Uniform distributions has a higher degree of accuracy to the real data than the distortion model based on Laplacian distribution.

Compared to the conventional distortion model based on Laplacian, the new model can decrease the MSE distortion by 2–6 units. The MLQ distortion model can model the relationship of rate and distortion of the Mpeg-4 FGS enhancement layer with a higher degree of accuracy.

5. CONCLUSIONS

The Laplacian Distribution has been widely used to characterize the distribution of DCT coefficients in a great deal of research on statistical attributes of DCT transform because Laplacian distribution can achieve a good balance between the simplification of the model and fidelity to the real empirical data. But in our research, we found that there are marked discrepancies between the real data and the distortion model based on Laplacian distribution. Furthermore, the higher the rate, the larger the distortion. In this paper, we proposed a new distortion model based on mixed Laplacian and Uniform distributions. Results of experiments show that the MLU distortion model can accurately approximate the real empirical data and has a higher degree of accuracy than the conventional distortion model based on Laplacian distribution. Compared to the conventional distortion model based on Laplacian, the new model can decrease the MSE distortion by 2–6 units. The MLU distortion model can represent the relationship of rate and distortion of the Mpeg-4 FGS enhancement layer with a higher degree of accuracy.

In the future, we will further modify and perfect our proposed distortion model and try to come up with a much deeper understanding of the rate distortion features in the Mpeg-4 FGS enhancement layer and design a new video transmission framework based on our proposed distortion model.

REFERENCES

DAI, M. (2003): Analysis and distortion modelling of Mpeg-4 FGS. *Proc. IEEE ICIP*, 14–17 September. Barcelona, Spain.

DAI, M. (2004): Rate-distortion modelling of scalable video coders. *IEEE ICIP*, 24–27 October, 2004, Singapore.

GUGLIELMO, M. (1986): An analysis of error behavior in the implementation of 2-D orthogonal transformations. *IEEE Transactions on Communications*, Vol. Com-34, No. 9, September.

LAKHANI, G. (2000): Distribution-based restoration of DCT coefficients. *IEEE Trans. Circuits and Systems for Video Technology*, 10(5): 819–823, August.

LAM, E. Y. and GOODMAN, J. W. (2000): A mathematical analysis of the DCT coefficient distributions for images. *IEEE Trans. Image Processing*, 9(10): 1661–1666, October.

LI, X. and ZHANG, W. (2005): A novel statistical distortion model of Mpeg-4 FGS, Emerging Technologies, 2005. *Proceedings of the IEEE Symposium* on 17–18 September. 1–6.

MPEG-4 (2001): Video Verification Model 18.0, ISO/IEC JTC1/SC29/WG11, N3908, January.

REININGER, R. and GIBSON, J. (1983): Distribution of the two-dimensional DCT coefficients for images. *IEEE Trans. Comm*. Vol. Com-31, June,

SMOOT, S. R. and ROWE, L. A. (1996): Study of DCT coefficient distributions. *Proc. SPIE*, 403–411, January.

WU, F. S. and CHAM, W. K. (1990): A comparison of error behaviour in the implementation of the DCT and the ICT. *IEEE Region 10 Conference on Computer and Communication Systems*, September, Hong Kong.

YOVANOF, G. S. and LIU, S. (1997): Statistical analysis of the DCT coefficients and their quantization error, in *Conf. Rec. 30th Asilomar Conf. Signals, Systems*, 1: 601–605.

BIOGRAPHICAL NOTES

Xie Li

Xie Li received the BSc and MSc degrees in physical, communication and information systems from Hubei University and Shanghai Normal University separately. He is currently working towards his PhD in communication and electronics systems at the Institute of Image Communication and Information Processing at Shanghai Jiaotong University, Shanghai, P.R.China. His research interests include digital video coding, video communication, embedded system and home network system.

Wenjun Zhang

As professor and chairman, Wenjun Zhang is now with the Institute of Image Communication and Information Processing at Shanghai Jiaotong University. His general interests include image processing, image communication and DTV.