DC Coefficient Distributions for P-Frames in H.264/AVC

Wei Wu and Bin Song

In this letter, the distributions of direct current (DC) coefficients for P-frames in H.264/AVC are analyzed, and the distortion model of the Gaussian source under the quantization of the dead-zone plus-uniform threshold quantization with uniform reconstruction quantizer is derived. Experimental results show that the DC coefficients of P-frames are best approximated by the Laplacian distribution and the Gaussian distribution at small quantization step sizes and at large quantization step sizes, respectively.

Keywords: Distortion model, Gaussian source, DC coefficient, P-frames, H.264/AVC.

I. Introduction

In hybrid video coding, the statistical distributions of the discrete cosine transform (DCT) coefficients are the basis in designing rate control and rate-distortion optimization (RDO) schemes. Assuming the coefficients are Laplacian-distributed, the most popular quadratic rate-distortion (RD) formula was proposed in [1] which is used in rate control for MPEG-4 and H.264/AVC, and an RDO scheme is presented in [2] which has a better performance than the uniform distribution-based RDO scheme recommended in the H.264/AVC reference software.

DCT coefficients include direct current (DC) and alternating current (AC) coefficients. For the AC coefficient in video coding, the Laplacian distribution is often utilized to model its distribution [3], [4]. However, the DC coefficient distribution in video coding is not yet discussed. If the DC coefficient in video coding does not follow the Laplacian distribution, the RD formula in [1] and the RDO scheme in [2] are not accurate enough. More accurate rate control and RDO schemes can be designed based on the DC coefficient distribution. To design more accurate rate control and RDO schemes, it is necessary for the DC coefficient in video coding to analyze its distribution. Since H.264/AVC is the state-of-the-art video coding standard, we focus on DC coefficient distributions for P-frames in H.264/AVC.

II. Probability Density Functions of DC Coefficients

In this section, the probability density functions (PDFs) of the DC coefficients for P-frames are analyzed. Some preliminary experiments are implemented to examine their distribution characteristics. In the experiments, H.264/AVC reference software JM 16.0 [5] is used.

Figure 1 shows the DC coefficient distributions at two various quantization parameters (QPs) for the 'Mobile' video sequence. In H.264/AVC, there is a deterministic relation between QP and quantization step size. The distributions of the DC coefficients are plotted and approximated by the Gaussian and the Laplacian distributions, respectively.

As shown in Fig. 1, the DC coefficient distributions at small and large quantization step sizes are different. From the figure, we observe that the statistics of the DC coefficients are best approximated by the Laplacian and the Gaussian distributions at small and large quantization step sizes, respectively.

When DC coefficients are quantized with a quantizer, the distortion is decided by the distribution of the coefficient and the quantization and reconstruction rules of the quantizer. In JM, the dead-zone plus-uniform threshold quantization (DZ-

Manuscript received Dec. 7, 2010; revised Apr. 1, 2011; accepted Apr. 18, 2011.

This work was supported by the Fundamental Research Funds for the Central Universities (No. JY10000901015), National Natural Science Foundation of China (No. 60802032), the 111 Project (B08038), Program for Changjiang Scholars and Innovative Research Team in University (IRT0852), and also supported by ISN State Key Laboratory.

Wei Wu (phone: +86 15109210470, email: wwu@xidian.edu.cn) and Bin Song (email: bsong@mail.xidian.edu.cn) are with the State Key Laboratory of Integrated Services Networks, Xidian University, Xi'an, China.

http://dx.doi.org/10.4218/etrij.11.0210.0461

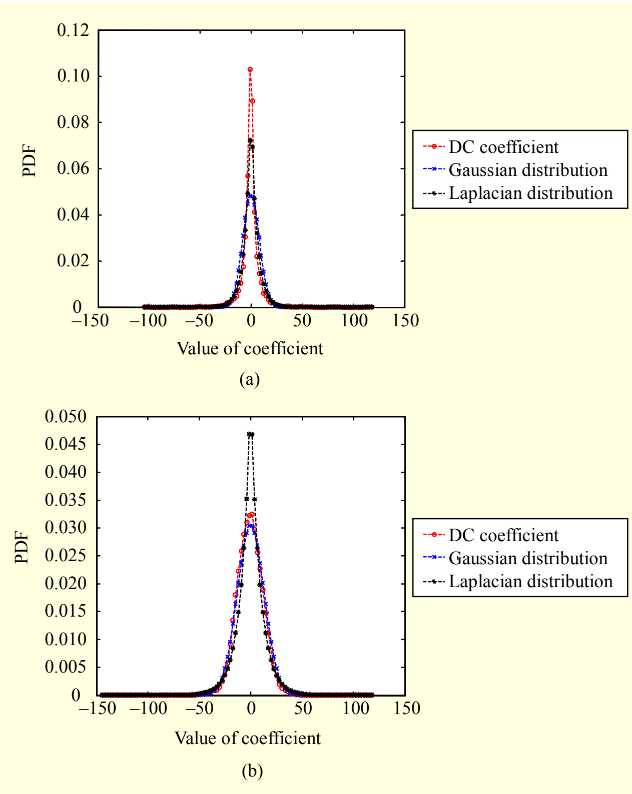

Fig. 1. PDFs of DC coefficients and approximations by Gaussian and Laplacian distributions of 'Mobile' CIF video sequence at (a) QP=10 and (b) QP=36.

UTQ) with uniform reconstruction quantizer (URQ) [6] is adopted. Turaga and others [7] derive the distortion model of the Laplacian source under the quantization of DZ-UTQ with URQ. For the Gaussian source under the quantization of DZ-UTQ with URQ, its distortion model is derived in section III.

III. Gaussian Source Distortion Model for H.264/AVC

The quantization and reconstruction rules of DZ-UTQ with URQ quantizer are described as follows:

Quantization rule. Supposing y is the value to be quantized, q is the quantization step size, and k is the corresponding quantization index. The quantization rule is

$$k = \mathrm{sign}(y)\cdot \max\left(0, \left\lfloor \frac{|y|}{q} + \alpha \right\rfloor\right), \tag{1}$$

where α is the quantizer dead-zone parameter which is set to be 1/6 for P-frames in JM.

Reconstruction rule. Let y^{reco} be the reconstructed value of y. URQ is implemented as

$$y^{\mathrm{reco}} = k\cdot q. \tag{2}$$

Let x be the Gaussian-distributed random variable. After the quantization of DZ-UTQ with URQ quantizer, the distortion of x, in terms of mean square error, can be written as

$$\begin{aligned}\varepsilon^2 &= \sum_{k=-\infty}^{\infty} (x - x^{\mathrm{reco}})^2 P(x^{\mathrm{reco}} = kq) \\ &= \frac{1}{\sqrt{2\pi}\sigma}\sum_{k=-\infty}^{-1}\int_{(k+\alpha-1)q}^{(k+\alpha)q} (x-kq)^2 e^{-\frac{x^2}{2\sigma^2}}dx \\ &+ \frac{1}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}\int_{(k-\alpha)q}^{(k-\alpha+1)q} (x-kq)^2 e^{-\frac{x^2}{2\sigma^2}}dx \\ &+ \frac{1}{\sqrt{2\pi}\sigma}\int_{-(1-\alpha)q}^{(1-\alpha)q} x^2 e^{-\frac{x^2}{2\sigma^2}}dx, \end{aligned}\tag{3}$$

where $P(\cdot)$ denotes the probability.

Due to symmetry, the first and second terms may be combined into one term. Hence,

$$\begin{aligned}\varepsilon^2 &= \frac{2}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}\int_{(k-\alpha)q}^{(k-\alpha+1)q}(x-kq)^2 e^{-\frac{x^2}{2\sigma^2}}dx + \frac{1}{\sqrt{2\pi}\sigma}\int_{-(1-\alpha)q}^{(1-\alpha)q} x^2 e^{-\frac{x^2}{2\sigma^2}}dx \\ &= \varepsilon_1^2 + \varepsilon_2^2, \end{aligned}\tag{4}$$

where

$$\begin{aligned}\varepsilon_1^2 &= \frac{2}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}\int_{(k-\alpha)q}^{(k-\alpha+1)q}(x-kq)^2 e^{-\frac{x^2}{2\sigma^2}}dx \\ &= \frac{2}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}\int_{(k-\alpha)q}^{(k-\alpha+1)q}(x^2 - 2kqx + k^2q^2) e^{-\frac{x^2}{2\sigma^2}}dx, \end{aligned}\tag{5}$$

$$\varepsilon_2^2 = \frac{1}{\sqrt{2\pi}\sigma}\int_{-(1-\alpha)q}^{(1-\alpha)q} x^2 e^{-\frac{x^2}{2\sigma^2}}dx. \tag{6}$$

We know that

$$\int x e^{-\frac{x^2}{2\sigma^2}}dx = -\sigma^2 e^{-\frac{x^2}{2\sigma^2}}, \tag{7}$$

$$\int x^2 e^{-\frac{x^2}{2\sigma^2}}dx = -\sigma^2 x e^{-\frac{x^2}{2\sigma^2}} + \sigma^2\int e^{-\frac{x^2}{2\sigma^2}}dx. \tag{8}$$

Substituting (7) and (8) into (5) and (6), we get

$$\begin{aligned}\varepsilon_1^2 &= \frac{2}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}\left(-\sigma^2 x e^{-\frac{x^2}{2\sigma^2}} + 2kq\sigma^2 e^{-\frac{x^2}{2\sigma^2}}\right)\Bigg|_{(k-\alpha)q}^{(k-\alpha+1)q} \\ &+ \frac{2}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}\int_{(k-\alpha)q}^{(k-\alpha+1)q}(\sigma^2 + k^2q^2) e^{-\frac{x^2}{2\sigma^2}}dx \\ &= \underbrace{\frac{2\sigma q}{\sqrt{2\pi}}\sum_{k=1}^{\infty}\left((k+\alpha-1)e^{-\frac{(k-\alpha+1)^2q^2}{2\sigma^2}} - (k+\alpha)e^{-\frac{(k-\alpha)^2q^2}{2\sigma^2}}\right)}_{\varepsilon_{11}^2} \\ &+ \underbrace{\frac{2}{\sqrt{2\pi}\sigma}\sum_{k=1}^{\infty}(\sigma^2 + k^2q^2)\int_{(k-\alpha)q}^{(k-\alpha+1)q} e^{-\frac{x^2}{2\sigma^2}}dx}_{\varepsilon_{12}^2}, \end{aligned}\tag{9}$$

$$\varepsilon_2^2 = \frac{1}{\sqrt{2\pi}\sigma}\left[-\sigma^2 x e^{-\frac{x^2}{2\sigma^2}} \Big|_{-(1-\alpha)q}^{(1-\alpha)q} + \sigma^2 \int_{-(1-\alpha)q}^{(1-\alpha)q} e^{-\frac{x^2}{2\sigma^2}} dx \right]$$

$$= \frac{-2(1-\alpha)\sigma q}{\sqrt{2\pi}} e^{-\frac{(1-\alpha)^2 q^2}{2\sigma^2}} + \frac{\sigma}{\sqrt{2\pi}} \int_{-(1-\alpha)q}^{(1-\alpha)q} e^{-\frac{x^2}{2\sigma^2}} dx. \quad (10)$$

In (9) and (10), there is a definite integral $\int_b^c e^{-\frac{x^2}{2\sigma^2}} dx$, where the interval of integration is [b, c]. The function $e^{-\frac{x^2}{2\sigma^2}}$ is not an integrable function, so we adopt the following method called composite trapezoidal rule to compute $\int_b^c e^{-\frac{x^2}{2\sigma^2}} dx$:

$$\int_b^c f(x)dx = \frac{h}{2}\left(f(b) + f(c) + 2\sum_{i=1}^{n-1} f(b+ih) \right), \; h = \frac{(c-b)}{n}, \quad (11)$$

where n is the number of subintervals and its value should be larger than 1. Suppose the error of the composite trapezoidal rule is the difference between the value of the integral and the numerical result. The larger the value of n is, the smaller the error is.

Using (11), ε_{12}^2 and ε_2^2 can be solved and expressed as

$$\varepsilon_{12}^2 = \frac{q}{\sqrt{2\pi}n\sigma} \sum_{k=1}^{m} \left((\sigma^2 + k^2 q^2)(e^{-\frac{(k-\alpha)^2 q^2}{2\sigma^2}} + e^{-\frac{(k-\alpha+1)^2 q^2}{2\sigma^2}} + 2\sum_{i=1}^{n-1} e^{-\frac{(k-\alpha+\frac{i}{n})^2 q^2}{2\sigma^2}}) \right),$$

$$m = \infty, \quad (12)$$

$$\varepsilon_2^2 = \frac{2(1-\alpha)(1-n)\sigma q}{\sqrt{2\pi}n} e^{-\frac{(1-\alpha)^2 q^2}{2\sigma^2}} + \frac{2(1-\alpha)\sigma q}{\sqrt{2\pi}n} \sum_{i=1}^{n-1} e^{-\frac{(1-\alpha)^2 q^2 (\frac{2i}{n}-1)^2}{2\sigma^2}}. \quad (13)$$

From (9), ε_{11}^2 can be rewritten as

$$\varepsilon_{11}^2 = \frac{2\sigma q}{\sqrt{2\pi}} \sum_{k=1}^{m} \left((k+\alpha-1) e^{-\frac{(k-\alpha+1)^2 q^2}{2\sigma^2}} - (k+\alpha) e^{-\frac{(k-\alpha)^2 q^2}{2\sigma^2}} \right),$$

$$m = \infty. \quad (14)$$

When $k \to \infty$, we can get

$$\left((k+\alpha-1) e^{-\frac{(k-\alpha+1)^2 q^2}{2\sigma^2}} - (k+\alpha) e^{-\frac{(k-\alpha)^2 q^2}{2\sigma^2}} \right) \to 0, \quad (15)$$

$$\left((\sigma^2 + k^2 q^2)(e^{-\frac{(k-\alpha)^2 q^2}{2\sigma^2}} + e^{-\frac{(k-\alpha+1)^2 q^2}{2\sigma^2}} + 2\sum_{i=1}^{n-1} e^{-\frac{(k-\alpha+\frac{i}{n})^2 q^2}{2\sigma^2}}) \right) \to 0, \quad (16)$$

so m can be selected a constant to estimate ε_{11}^2 and ε_{12}^2. Substituting (12), (13), and (14) into (4), the distortion ε^2 is

$$\varepsilon^2 = \frac{2\sigma q}{\sqrt{2\pi}} \sum_{k=1}^{m} \left((k+\alpha-1) e^{-\frac{(k-\alpha+1)^2 q^2}{2\sigma^2}} - (k+\alpha) e^{-\frac{(k-\alpha)^2 q^2}{2\sigma^2}} \right)$$
$$+ \frac{q}{\sqrt{2\pi}n\sigma} \sum_{k=1}^{m} \left((\sigma^2 + k^2 q^2)(e^{-\frac{(k-\alpha)^2 q^2}{2\sigma^2}} + e^{-\frac{(k-\alpha+1)^2 q^2}{2\sigma^2}} + 2\sum_{i=1}^{n-1} e^{-\frac{(k-\alpha+\frac{i}{n})^2 q^2}{2\sigma^2}}) \right)$$
$$+ \frac{2(1-\alpha)(1-n)\sigma q}{\sqrt{2\pi}n} e^{-\frac{(1-\alpha)^2 q^2}{2\sigma^2}} + \frac{2(1-\alpha)\sigma q}{\sqrt{2\pi}n} \sum_{i=1}^{n-1} e^{-\frac{(1-\alpha)^2 q^2 (\frac{2i}{n}-1)^2}{2\sigma^2}}, \quad (17)$$

where m and n are two constants.

IV. Experimental Results

If DC coefficients follow some distribution, for example, Gaussian, using the corresponding distortion model should achieve the closer estimated distortion to the actual one than using the models of other distributions. Thus, at small and large quantization step sizes, using the distortion models of the Laplacian and the Gaussian sources, respectively, can achieve more accurate estimated distortions.

To validate the DC coefficient distributions of P-frames, we employ JM 16.0. Ten video sequences in CIF format are tested. The test frame rate is 30 fps. Each of the ten sequences has 118 frames to be encoded in which the first frame is I-frame followed with subsequent P-frames. In the experiment, I-block coding in P-frames is disabled. The QP values are respectively adopted from 4 to 42 to encode each sequence. At a QP, the variance σ^2 of the DC coefficient is calculated over all the P-frames in a sequence. Substitute α, σ^2, and q into (17) to estimate the distortion. In addition, the distortion model of the

Table 1. Mean values of estimation errors.

Video sequence	Small step size (QP=4:25)		Large step size (QP=26:42)	
	Laplacian	Gaussian	Laplacian	Gaussian
Mobile	**1.6283**	2.0785	72.3869	**25.7153**
Foreman	**0.5715**	0.9223	39.1718	**3.4252**
News	**0.3858**	1.1312	24.1949	**10.6923**
Bus	**1.6215**	1.8993	42.9006	**16.0328**
Carphone	**0.4801**	1.1831	24.6643	**7.4868**
Coastguard	**1.4760**	1.9292	56.3593	**12.1272**
Mother	**0.3939**	0.8274	38.5926	**4.9837**
Flower	**1.7040**	2.2363	23.8592	**20.9831**
Akiyo	**0.3916**	0.6614	26.1706	**2.1674**
Tempete	**1.5384**	1.8629	66.9318	**17.5470**

Table 2. Effect of n on computational complexity and estimation error of 'Mobile' sequence at QP=30.

	n=10	n=20	n=30	n=40
Running time (ms)	0.0382	0.0671	0.1017	0.1536
Estimation error	12.8866	12.7173	12.6859	12.6749

Laplacian source in [7] is also used:

$$\varepsilon^2 = \sigma^2 - ((1-2\alpha)q + \sqrt{2}\sigma) \times \frac{e^{-\sqrt{2}(1-\alpha)q/\sigma} q}{1-e^{-\sqrt{2}q/\sigma}}. \quad (18)$$

In the experiment, we found that when $n \geq 20$, the error of the composite trapezoidal rule is close to 0, which shows that twenty subintervals are enough to compute the value of the integral. To reduce computational complexity, we select n=20 in this letter. In addition, we found that when $m \geq 20$, ε_{11}^2 and ε_{12}^2 keep stable, which shows that the probability of $x^{\text{reco}} = mq$ $(m \geq 20)$ is very small because x follows the Gaussian distribution. So we select m=20 in this letter.

In order to quantitatively validate the DC coefficient distributions, the estimation errors between the actual distortions and the estimated distortions by using (17) and (18) are computed at QP=4:42 for each sequence, respectively. Table 1 shows the mean values of the estimation errors at small and large step sizes. According to the results of the table, we observe that at small step sizes, the DC coefficient can achieve the smaller estimation errors by using (18) for all the sequences. At large step sizes, the estimated distortions are more accurate by using (17) for all the sequences. The experimental results indicate that at small and large quantization step sizes, the DC coefficients are best approximated by the Laplacian and the Gaussian distributions, respectively.

For evaluating the effect of n on computational complexity and estimation error, the running times and the estimation errors are computed. Our platform is Intel Core i3 2.93G CPU and 2 GB RAM. The results are shown in Table 2. From the table, we can see that the estimation errors are similar among n=20, 30, and 40, which indicates n can be selected to be 20. Also, it can be seen that the running time of the proposed distortion model at n=20 is 0.0671 ms, which is small.

V. Method to Select the Distribution Adaptively

Since the DC coefficient distribution is related to QP, and from the experiments we found that the distribution changes from Laplacian distribution to Gaussian distribution with the increase of QP, a method to select the distribution adaptively during CBR H.264 encoding is proposed in the following.

- For the initial l frames in a sequence, after encoding each frame, the distribution is determined by using (17) and (18) based on the encoding results. The QPs used for the l frames are classified into different QP groups according to different distributions, where l is an empirical value and set to be 20.
- For each of the subsequent frames, before encoding the frame, the distribution can be predicted. When there is only one QP group, the distribution is selected to be the corresponding distribution of the group. When there are two QP groups, the QP used for the frame, QP_c, is compared with the maximum QP, $QP_{\max,L}$, in the Laplacian group and the minimum QP, $QP_{\min,G}$, in the Gaussian group. If $|QP_c - QP_{\max,L}| < |QP_c - QP_{\min,G}|$, the distribution is selected to be Laplacian distribution; else, the distribution be Gaussian distribution. After encoding the frame, the distribution is determined based on the encoding results and QP_c is classified.

VI. Conclusion

In this letter, the DC coefficient distributions for P-frames in H.264/AVC were analyzed, and it was found that the DC coefficients of P-frames are best approximated by the Laplacian and the Gaussian distributions at small and large quantization step sizes, respectively. Then, a mathematic derivation for the distortion of the Gaussian source under the quantization of DZ-UTQ with URQ quantizer was proposed. Experimental results validate the DC coefficient distributions. Furthermore, a method to select the distribution adaptively during CBR H.264 encoding was proposed.

References

[1] A. Vetro, H. Sun, and Y. Wang, "MPEG-4 Rate Control for Multiple Video Objects," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, Feb. 1999, pp. 186-199.

[2] X. Li et al, "Laplace Distribution Based Lagrangian Rate Distortion Optimization for Hybrid Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, Feb. 2009, pp. 193-205.

[3] S.R. Smoot and L.A. Rowe, "Study of DCT Coefficient Distributions," *Proc. SPIE*, 1996, pp. 403-411.

[4] Z.Z. Chen and K.N. Ngan, "Towards Rate-Distortion Tradeoff in Real-Time Color Video Coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 2, Feb. 2007, pp. 158-167.

[5] H.264/AVC Reference Software Version JM 16.0. Available: http://iphome.hhi.de/suehring/tml/

[6] G.J. Sullivan and S. Sun, "On Dead-Zone Plus Uniform Threshold Scalar Quantization," *Proc. SPIE Visual Commun. Image Process.*, 2005, pp. 1041-1052.

[7] D.S. Turaga, Y. Chen, and J. Caviedes, "No Reference PSNR Estimation for Compressed Pictures," *Signal Process. Image Commun.*, vol. 19, no. 2, Sept. 2004, pp. 173-184.