AN ANALYSIS OF THE DCT COEFFICIENT DISTRIBUTION WITH THE H.264 VIDEO CODER

Yucel Altunbasak and Nejat Kamaci

Center for Signal and Image Processing
Georgia Institute of Technology
Atlanta, GA 30332
email:{yucel,kamaci}@ece.gatech.edu

ABSTRACT

By extensive analysis with several video sequences, we observed that the statistical distribution of the DCT coefficients in typical video coding applications is closer to a Cauchy distribution than to a Laplacian distribution. We developed the rate and the distortion expressions as a function of the video coder quantization parameter based on this observation. Experiments with an H.264 codec demonstrate that the Cauchy distribution based expressions provide better estimates for the actual rate and distortion than those that are based on the Laplacian distribution.

1. INTRODUCTION

Most image and video coding standards use a block based two-dimensional discrete cosine transform (DCT) or an approximation to it as part of the coding algorithm [1, 2, 3]. The knowledge of the statistical behavior of the DCT coefficients is important in the design of such encoder algorithms as rate control and macroblock mode selection. Especially, the AC portion of the DCT coefficients (AC coefficients) is of great interest. Several studies on the statistical distribution of the AC coefficients have been proposed, in which the AC coefficients were conjectured to have Gaussian [4, 5], Laplacian [6, 7], or more complex distributions [8, 9]. Among these, the Laplacian distribution is probably the most popular, and used in practice. This is mainly due to the simplicity of the Laplacian probability density function (pdf) that makes it easy to derive mathematical formulations for the video coding algorithms.

However, the actual distribution of the AC coefficients in image and video applications differs from the Laplacian distribution in most cases. As a result, rate and distortion models based on this distribution sometimes fail to estimate the actual rate-distortion-coding parameter relations accurately. A more accurate, and simple approximation to the actual AC coefficient distribution would be very useful in image and video coding applications.

We claim that for most video sources, the actual AC coefficient distribution is closer to a Cauchy distribution than a Laplacian distribution. Using the Cauchy pdf, we can formulate the relations between the quantization parameter Q of a video coder and the output bit rate and the distortion caused by the quantization more accurately. In this paper, we show that the actual AC coefficient distribution is better approximated by a Cauchy distribution than a Laplacian distribution, by comparing the actual rate and distortion functions with the rate and distortion approximations obtained using the two densities.

2. ANALYSIS OF DCT COEFFICIENTS OF TYPICAL VIDEO SOURCES

The performance of a video coder in terms of its output bit rate, and the encoded video quality varies with the nature of the video source. Traditional video coders use block-based DCTs for compression[1]. For intra coding, the DCT is applied to the image itself; for non-intra coding, a residual image is obtained by performing a prediction, and the DCT is applied to this residual. In both cases, knowledge of the DCT coefficients distribution is valuable for optimizing the video coder. Fig. 1 shows a typical plot of the histogram of the AC coefficients for an 8×8 block based DCT of a video frame from the AKIYO sequence.

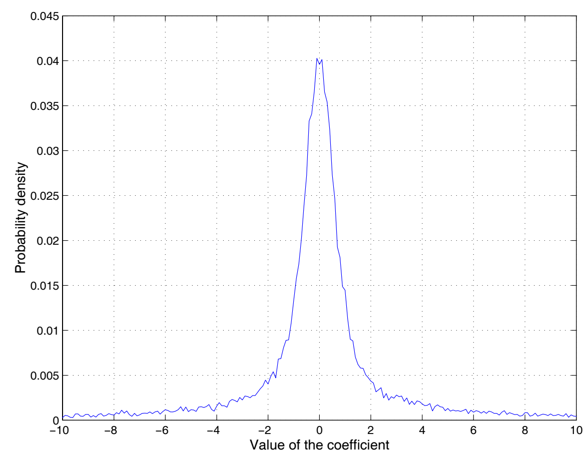


Fig. 1. Distribution of the AC DCT coefficients - for a video frame from the Akiyo Sequence (QCIF format).

As discussed in Section 1, the DCT coefficient distribution is traditionally approximated by a Laplacian pdf with parameter λ:

$$p(x) = \frac{\lambda}{2} \exp\{-\lambda |x|\}, \quad x \in \mathbf{R}. \tag{1}$$

The Laplacian pdf has an exponential form, leading to the property that the tail of the density decays very fast. In most cases, the actual DCT coefficient distribution has a considerable tail. A zero-mean Cauchy distribution with parameter μ, having the pdf

$$p(x) = \frac{1}{\pi} \frac{\mu}{\mu^2 + x^2}, \quad x \in \mathbf{R}, \tag{2}$$

[1]The H.264 coder uses an integer transform that is a close approximation to the DCT.

exhibits similar behavior as the actual AC distribution, for which the tail of the density decays slow. The parameter μ depends on the picture content and can be estimated using the histogram of the transform coefficients. Fig. 2 illustrates the accuracy of the fit for both the Cauchy and the Laplacian pdf's for the DCT coefficient distribution of a selected video frame from the TEMPETE sequence. In the figure, the first plot shows the actual distribution of the coefficients for the intra coding case; the second plot shows the actual distribution for the non-intra coding case. In both cases, the Cauchy pdf is a better fit to the actual distribution than the Laplacian pdf. Clearly, we need more experiments to support our claim. However, rather than comparing the histograms as we did in Fig. 2, we compare the actual rate and distortion functions with the estimates based on the Cauchy and the Laplacian distributions. That is, since the rate and the distortion functions that are derived from these pdfs are used in practice, we consider them as the basis for comparison.

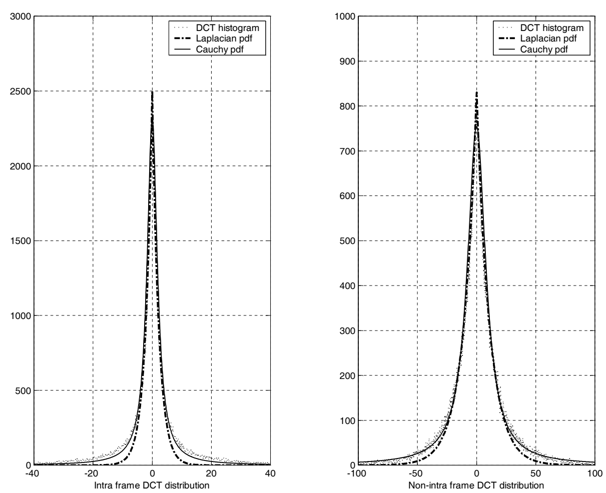


Fig. 2. Comparison of Laplacian vs Cauchy histograms - selected intra and non-intra frames from TEMPETE sequence

2.1. Cauchy-Based Rate Estimation

A more accurate estimate of the AC distribution will lead to more accurate estimation of the rate. Assume that the DCT coefficients are uniformly quantized with a quantization level Q. Let $P(iQ)$ be the probability that a coefficient is quantized to iQ. Then the entropy of the quantized DCT coefficients can be computed as

$$H(Q) = -\sum_{i=-\infty}^{\infty} P(iQ)\log_2[P(iQ)], \tag{3}$$

where

$$P(iQ) = \int_{(i-\frac{1}{2})Q}^{(i+\frac{1}{2})Q} f_X(x)\,dx.$$

For a Laplacian distribution, we have

$$P(iQ) = \begin{cases} \frac{1}{2}e^{-i\lambda Q}\left(e^{\frac{\lambda Q}{2}} - e^{-\frac{\lambda Q}{2}}\right) & \text{if } i > 0 \\ 1 - e^{-\frac{\lambda Q}{2}} & \text{if } i = 0 \\ \frac{1}{2}e^{i\lambda Q}\left(e^{\frac{\lambda Q}{2}} - e^{-\frac{\lambda Q}{2}}\right) & \text{if } i < 0 \end{cases}.$$

Therefore, the entropy as a function of Q for a Laplacian distribution is

$$H(Q) = -\left(1-e^{-\frac{\lambda Q}{2}}\right)\log_2\left(1-e^{-\frac{\lambda Q}{2}}\right) - 2e^{-\frac{\lambda Q}{2}}\left[\log_2\left(\frac{e^{\frac{\lambda Q}{2}}-e^{-\frac{\lambda Q}{2}}}{2}\right) + \frac{\lambda Q}{\left(1-e^{-\lambda Q}\right)\ln 2}\right].$$

For a Cauchy distribution,

$$P(iQ) = \begin{cases} \frac{1}{\pi}\tan^{-1}\left(\frac{\mu Q}{\mu^2+(i^2-1/4)Q^2}\right) & \text{if } i > 0 \\ \frac{2}{\pi}\tan^{-1}\left(\frac{Q}{2\mu}\right) & \text{if } i = 0 \\ \frac{1}{\pi}\tan^{-1}\left(\frac{\mu Q}{\mu^2+(i^2-1/4)Q^2}\right) & \text{if } i < 0 \end{cases}.$$

Therefore, the entropy function as a function of Q for a Cauchy distribution is

$$H(Q) = -\frac{2}{\pi}\tan^{-1}\left(\frac{Q}{2\mu}\right)\log_2\left(\frac{2}{\pi}\tan^{-1}\left(\frac{Q}{2\mu}\right)\right) - \frac{2}{\pi}\sum_{i=1}^{\infty}\tan^{-1}\left[\frac{1}{\frac{\mu}{Q}+\frac{(4i^2-1)Q}{4\mu}}\right]\log_2\tan^{-1}\left[\frac{1}{\frac{\mu}{Q}+\frac{(4i^2-1)Q}{4\mu}}\right] \tag{4}$$

These entropy functions based on the Laplacian and the Cauchy pdfs are computable, provided that the density parameters λ and μ are known. The Laplacian parameter λ can be computed using its relation to the variance of the AC coefficients (σ^2) as $\lambda = \sqrt{2}/\sigma$. Similarly, the Cauchy parameter μ can be computed using the histogram of the AC coefficients. Although the entropy of a quantized Cauchy source can be computed using Eq. (4), it would be preferable to use a simpler formula. After analyzing Eq. (4), one can find that there exists a very simple, yet accurate approximation to it, which is

$$R(Q) \approx aQ^{-\alpha}, \tag{5}$$

where $a, \alpha > 0$ are parameters that depend on μ [2]. To assess the accuracy of this approximation, we plot the entropy function and its approximation for different values of μ. As shown in Fig. 3, the approximation is accurate, especially for $\mu \le 1$.

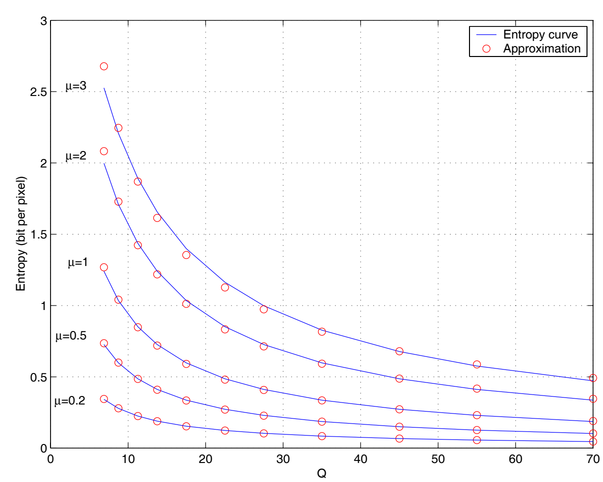
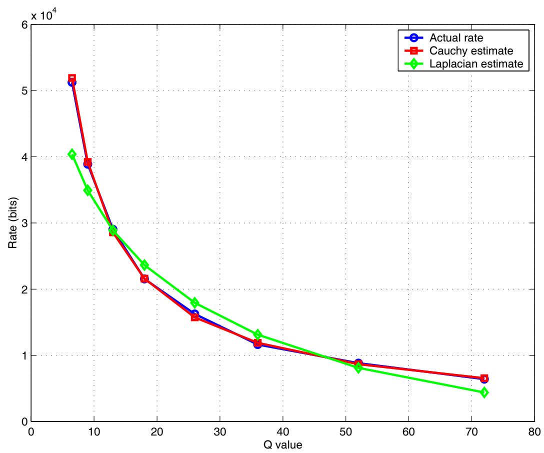


Fig. 3. Theoretical versus approximated entropy functions for five different values of the Cauchy distribution parameter μ.

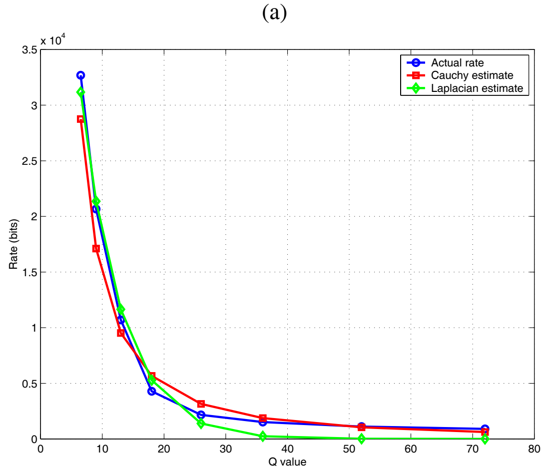
[2]There is no analytical expression that relates a and α to μ.

2.2. Rate Experiments

To show the effectiveness of the rate model given in Eq. (5), we conducted several experiments with a number of video sequences each in QCIF-format. We encoded an intra frame followed by a non-intra frame for each sequence at several quantization levels to obtain the actual rate as a function of the quantization level Q for both intra and non-intra coding. We used the H.264 reference software, version JM-6.0 [10] for our experiments[3]. The rate estimates are computed using the Laplacian entropy function and the Cauchy-based approximation given in Eq. (5). In Figs 4 and 5, the Laplacian-based rate estimates and Cauchy-based rate estimates are compared with the actual rate as a function of Q for both intra and non-intra coding.



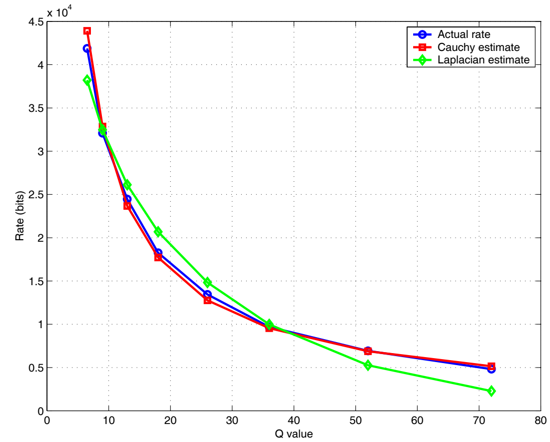
(a)



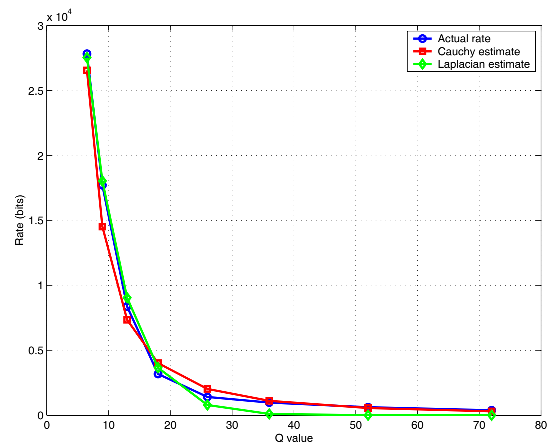
(b)

Fig. 4. The actual rate vs. Laplacian and Cauchy based rate estimates for (a) an intra frame, and (b) a non-intra frame from FOREMAN sequence.

Table 1 summarizes the rate estimation accuracy based on both distributions with several frames selected from a wide range of video sequences. The Cauchy-based rate estimation is significantly better than the Laplacian-based rate estimation for these



(a)



(b)

Fig. 5. The actual rate vs. Laplacian and Cauchy based rate estimates for (a) an intra frame, and (b) a non-intra frame from IRENE sequence.

video sequences. The Cauchy estimate matches the coder performance very accurately, especially for intra coding.

2.3. Cauchy-Based Distortion Estimation

The distortion due to quantization can also be estimated accurately based on the Cauchy pdf assumption. Assume that we have a uniform quantizer with step size Q. The distortion caused by quantization is given by

$$D\left(Q\right) = \sum_{i=-\infty}^{\infty} \int_{\left(i-\frac{1}{2}\right)Q}^{\left(i+\frac{1}{2}\right)Q} \left|x - iQ\right|^2 f_X\left(x\right) dx.$$

It can be shown that this infinite sum is convergent and bounded from above by $Q^2/4$. For a Cauchy source, this expression be-

[3]The entropy coder (CABAC) used in the H.264 video coder [3] is reported to be more advanced than the entropy coders used in the other video coders, so we expect its performance to be closer to the theoretical limits.

	Rate estimation error %			
	Intra coding		Non-intra coding	
Image	Laplace	Cauchy	Laplace	Cauchy
CARPHONE	0.140	0.023	0.360	0.115
CLAIRE	0.125	0.015	0.283	0.054
FOREMAN	0.130	0.017	0.410	0.287
MOB. & CAL.	0.062	0.047	0.282	0.106
IRENE	0.137	0.022	0.404	0.188
COASTGUARD	0.102	0.055	0.203	0.116
NEWS	0.080	0.036	0.133	0.078
TEMPETE	0.078	0.055	0.214	0.116

Table 1. Comparison of the rate estimation performance of the Laplacian-based rate model and the Cauchy-based rate model given in Eq. (5).

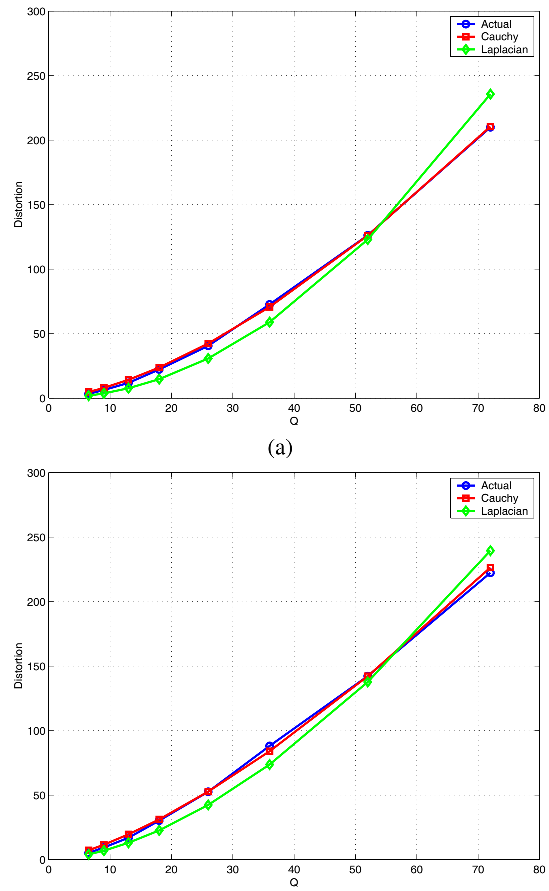
comes

$$
\begin{aligned}
D(Q) \quad = \quad & 2\sum_{i=1}^{M} \frac{\mu Q}{\pi} - \frac{i\mu Q}{\pi} \ln\left(\frac{\mu^2 + \left(i+\frac{1}{2}\right)^2 Q^2}{\mu^2 + \left(i-\frac{1}{2}\right)^2 Q^2}\right) \\
& -2\sum_{i=1}^{M} \frac{\mu^2 - i^2 Q^2}{\pi} \tan^{-1}\left(\frac{\mu Q}{\mu^2 + \left(i^2 - \frac{1}{4}\right) Q^2}\right) \\
& + \left[\frac{\mu Q}{\pi} - \frac{2\mu^2}{\pi} \tan^{-1}\left(\frac{Q}{2\mu}\right)\right].
\end{aligned}
$$

This equation suggests that the distortion depends on μ as well as Q. Although this equation is highly complex, for practical values of Q, it can be approximated simply as

$$D(Q) \approx bQ^{\beta}, \tag{6}$$

where $b, \beta > 0$ are parameters that depend on μ[4]. Fig. 6 shows the actual distortion functions for intra and non-intra coding cases for the MOBILE AND CALENDAR sequence and the estimated distortion functions using the Laplacian pdf and the Cauchy-based approximate rate model of Eq. (6).



(a)

(b)

Fig. 6. The actual distortion vs. Cauchy based distortion approximation for (a) an intra frame, and (b) a non-intra frame from MOBILE AND CALENDAR sequence.

3. SUMMARY AND CONCLUSION

In this paper, we claim that the AC coefficient distribution is better approximated by a Cauchy density than by a Laplacian density. The rate and distortion of a video coder as a function of its quantization parameter is derived under the Cauchy-distribution assumption. The accuracy of the Cauchy-based rate and distortion functions are examined against the traditional Laplacian-based functions using an H.264 coder. The experiments have indicated that Cauchy-based functions are more accurate than the Laplacian based functions in most cases. We expect that these new rate and distortion functions will give rise to better video encoding algorithms.

4. REFERENCES

[1] G. K. Wallace, "The JPEG still picture compression standard," *Commun. ACM,* vol. 34, pp. 30-44, Apr. 1991.

[2] "The MPEG-2 international standard," ISO/IEC, Reference number ISO/IEC 13818-2, 1996.

[3] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. on Image Proc.,*, vol. 13, no. 7, July 2003, pp. 560-576.

[4] W. K. Pratt, "Digital Image Processing," New York: Wiley-Interscience, 1978, ch. 10.

[5] A. N. Netravali and J. O. Limb, "Picture coding: A review," *Proc. IEEE,* vol. 68, pp. 7-12, Mar 1960.

[6] R. C. Reininger and J. D. Gibson, "Distributions of the two-dimensional DCT coefficients for images," *IEEE Trans. on Commun.,* vol. COM-31, pp 835-839, June 1983.

[7] S. R. Smooth and R. A. Lowe, "Study of DCT coefficients distributions," *Proc. SPIE* , pp. 403-311, Jan. 1996.

[8] E. Y. Lam and J. W. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Trans. on Image Proc.,* vol. 9, no. 10, Oct. 2000, pp. 1661-1666.

[9] T. Eude, R. Grisel, H. Cherifi, and R. Debrie, "On the distribution of the DCT coefficients," *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing* , vol. 5, Apr. 1994, pp. 365-368.

[10] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Joint Model Reference Software Version 6.0.

[4] As in the entropy rate case, there is no analytical expression that relates b and β to μ.