Parameter Estimation for Image/Video Transcoding

Guo Zihua*, O.C. Au**, Member, IEEE, K.B. Letaief***, Senior Member, IEEE

zhguo@ee.ust.hk*, eeau@ee.ust.hk**, eekhaled@ee.ust.hk***

The Department of Electrical & Electronic Engineering
The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong

ABSTRACT

In most of the current image/video compression standard, such as JPEG, MPEG, H.263, etc, the DCT is used. The DCT coefficients for every AC frequency component usually have Laplacian distribution. Such a distribution is critical in the image or video transcoding system [5]. Therefore, the parameter in the distribution should be well estimated. Different from the normal parameter estimation, in the transcoding system, the original DCT value is unavailable, and only the dequantized value is available. In this paper, we will propose 3 estimation methods based on the dequantized value, specially, for the transcoding system. The simulation will demonstrate that they can yield very good results. In addition, under a fixed bit rate, they can work even better than estimation with the original DCT value in the sense of image/frame's PSNR.

1. INTRODUCTION

Image/video coding and the corresponding applications are becoming more and more pervasive now. However, different application may require different bit rate due to the different (limited) bandwidth. For example, suppose we have a MPEG-2 bitstream which is 4Mbps and we want to transmit it via a VOD channel whose capacity is 2Mbps, we must reduce the bitstream to be 2Mbps. The conversion from the 4Mbps bitstream to 2Mbps bitstream is called transcoding [1,2]. Similarly, for image, there is a need to change the file size which is also referred as transcoding. Normally, transcoding is associated with reducing the bitrate, that is, from $R_2 \rightarrow R_1$ with $R_2 < R_1$. Fully decoding the bitstream and reencoding are not needed, we can do it directly in the frequency domain (DCT domain) [1,2,3,4], the possible transcoding system for video is shown in Fig.1 [1].

The lower part of Fig.1 is mainly accounts for the motion re-estimation. While the upper part is one of the most important steps: requantization. That is, for the DCT coefficients, we first dequantize them with step size Q_1 , then requantize it with step size Q_2 . Requantization is the main reason to reduce the bit rate ($Q_2 > Q_1$). When we do requantization, there will be two cases for $Q_2 > Q_1$ which are shown in Fig.2.

Fig.1 The transcoding system for video

Fig. 2 Cell Location when requantization

The interesting problem is, in the cross-overlapping case in Fig.2, how to choose the value between the two candidate requantized values: m_1' and m_2' . In [5], the author proposed two requantization methods, named MMSE and MAP, respectively. For MMSE requantizer, it chose m_1' or m_2' with the criteria to minimize the mean square error between the original value and the selected new value. For MAP, it maximizes *a posteriori* probability of the original DCT value falling into the new cell. Both of these two requantizers are based on the proper DCT coefficient distribution model. Therefore, the estimation of the parameter in the distribution is critical to the requantizer.

Normally, the AC components of DCT coefficients have Laplacian distribution [1]. That is, the probability density function (pdf) can be given by

$$f(x) = 0.5\lambda e^{-\lambda|x|} \quad (1)$$

Where $\lambda = 1.414/\sigma$ and σ is the variance. Generally, different frequency component has different σ , the higher the frequency is, the smaller the σ will be. Now the problem is that, in the transcoding system, the original DCT value is unavailable, and only the dequantized value is available. Therefore, the usual Maximal Likelihood parameter estimation can not be applied. In this paper, we will propose 3 good estimation methods based on the dequantized value, specially, for the transcoding system.

The paper is organized as follows. In Section 2, we will propose 3 parameter estimation methods in detail. In Section 3, we will give some simulation results. Finally, the conclusion is presented in Section 4.

2. PARAMETER ESTIMATION

From (1) we can see that the only parameter we should estimate is λ . Usually, the Maximally Likelihood (ML) parameter estimation method can be used when all the original DCT values are known. While in the transcoding system, we only have the dequantized DCT values, but not the original ones. Therefore, we should modify the usual ML method to give a reasonable estimation.

Fig. 3 The distribution of DCT values

Fig. 3 gives the Laplacian distribution of DCT values, where $\pm c_1$ is threshold quantized to be 0 with Q_1 , $\pm c_2$ is the threshold quantized to be ± 1 with Q_1 . For TM5 quantizer, $c_1 = 0.625 * Q_1$, $c_2 = 1.625 * Q_1$.

We see that, since every original AC component of DCT values are Laplacian distributed, the probability that the original DCT value falls into the zero cells is given by:

$$p_1 = p(-c_1 < x < c_1) = 1 - e^{-c_1\lambda} \quad (2)$$

Similarly, the probability p_2 that the DCT value falls into the two ± 1 cells and the probability p_3 that the value falls into the other nonzero cells are given by, respectively

$$p_2 = p(c_1 < |x| < c_2) = e^{-c_1\lambda} - e^{-c_2\lambda} \quad (3)$$

$$p_3 = 1 - p_1 - p_2 \quad (4)$$

Then we assume:

- k_1 : number of 0;
- k_2 : number of ± 1 ;
- k_3 : number of all other value;
- n : the total number of this AC component values, that is, the total block number.

Therefore, k_1, k_2, k_3 are random variables, and only two of them are independent. In the following, we first consider two kinds of estimation, the first order and the second order.

2.1 The First Order Estimation

If we just consider the value: zero and nonzero, then:

$$p_1 = p(-c_1 < x < c_1) = 1 - e^{-c_1\lambda} \quad (5)$$

$$p_2 = p(|x| > c_1) = e^{-c_1\lambda} \quad (6)$$

Then the distribution of k_1 (the likelihood function) is found to be:

$$P(k_1) = \frac{n! p_1^{k_1} p_2^{n-k_1}}{k_1! (n-k_1)!} \quad (7)$$

Taking logarithm to get LL function and differentiating the LL function with respect to λ results in an analytical solution:

$$\lambda = -\frac{1}{c_1} \ln\left(\frac{n-k_1}{n}\right) \quad (8)$$

2.2 The Second Order Estimation

The joint pdf of k_1, k_2 (k_3) or the Likelihood function is found to be:

$$P(k_1, k_2) = \frac{n! p_1^{k_1} p_2^{k_2} p_3^{(n-k_1-k_2)}}{k_1! k_2! (n-k_1-k_2)!} \quad (9)$$

Take logarithm, getting the LL function:

$$LL = \ln(P(k_1, k_2)) \\ = c + k_1 * \ln(p_1) + k_2 * \ln(p_2) + (n - k_1 - k_2) * \ln(p_3) \quad (10)$$

Where c is a constant not related to λ , then differentiating with respect to λ resulting in:

$$\frac{c_1 * k_1}{1 - e^{-c_1\lambda}} + k_2 * \frac{-c_1 * e^{-c_1\lambda} + c_2 * e^{-c_2\lambda}}{e^{-c_1\lambda} - e^{-c_2\lambda}} - c_2 * (n - k_1 - k_2) = 0 \quad (11)$$

It can be solved with computer iteration with the initialization value from the first order estimation which we describe above.

2.3 The Estimation Based on Conditional pdf

Due to some special characteristics of requantization, we can do better when we estimate λ . There are two points we must pay attention to for requantization:

- Because we always requantize the 0 to be 0, that is, we do nothing with 0.
- From Fig. 4 which is composed of the blocks that make more than 2 error choices with TM5 quantizer for MAP method. We can see that most of the blocks are edge blocks where the high frequency components are opulent.

Therefore, these remind us to estimate the λ only with the nonzero dequantized DCT value, that is, with $\pm 1, \pm 2 \dots$. Now, refer to Fig. 3, because exponential pdf is a memoryless pdf, so Laplacian pdf is, i.e. it is still a double side exponential pdf. Therefore, with the similar method with the first order estimation mentioned above, we can derive the estimation of λ based on this conditional pdf. Then the final result for the first order conditional estimation is given by

$$\lambda = -\frac{1}{c_1} \ln\left(\frac{n - k_2}{n}\right) \quad (12)$$

But here n is the total number of all the nonzero dequantized DCT value for the specified AC component, and k_2 is the number of ± 1 .

Fig.4. Error image for $Q_1=14$, $Q_2=23$, MAP+TM5

3. EXPERIMENTAL RESULT

In order to have a look at the performance of our proposed algorithms, we apply them to the image Lena (512*512). In Fig.5, we give the performance (requantized image's PSNR, bit rate) comparison among the Maximum Likelihood estimation with the original DCT value, the first order unconditional estimation, the conditional estimation. The TM5 quantizer and the MAP requantizer are used in the simulation, and $Q_1=14$, $Q_2=15-29$. From (a) and (b), we can see that, with the same Q_2 , the estimation with the original DCT value gives the highest PSNR also

the highest bit rate. The unconditional estimation usually yields the lowest PSNR also the lowest bit rate. To have a more reasonable comparison, Fig.5(c) demonstrates the bit rate vs. PSNR. We can see that under a fixed bit rate, the estimation with the original value gives the lowest PSNR, while the conditional estimation usually has a little higher PSNR than the unconditional one. Therefore, our estimation can work even better than the estimation with the original value in this sense. This may be due to that the latter only optimize the PSNR but not the bit rate.

Fig. 5 Comparison for 3 parameter estimation methods

4. CONCLUSION

Due to the importance of the DCT coefficient distribution in the transcoding system, especially in the requantization processing, the parameter estimation of the distribution is critical. In this paper, we proposed 3 modified Maximum Likelihood estimation methods and compared with the usual Maximum Likelihood parameter estimation with the original DCT value. We have shown that they can give good result in the practical system, and in some sense, it can work better than the latter.

5. REFERENCES

- [1]. G. Keesman, R. Hellinghuizen, F. Hoeksema, G. Heideman, "Transcoding of MPEG bitstream", *Signal Processing: Image communication*, 8 (1996) pp. 481-500.
- [2]. Y. Nakajima, H. Hori, and T. Kanoh, "Rate conversion of MPEG coded video by re-quantization process", *Proc. ICIP'95*, Los Alamitos, CA, USA, 1995, pp. 408-411.
- [3]. S. Acharya, B. Smith, "Compressed Domain Transcoding of MPEG", *Proceedings. IEEE International Conference on Multimedia Computing and Systems*, Los Alamitos, CA, USA, 1998, pp.295-304.
- [4]. Ja-Ling Wu, Shiao-Juan Huang, Yuh-Ming Huang, Chiou-Ting Hsu, and Jiun Shiu, "An Efficient JPEG to MPEG-1 Transcoding Algorithm", *IEEE Transactions on Consumer Electronics*, vol.42, no.3, Aug. 1996, pp. 447-457.
- [5]. O. Werner, "Requantization for Transcoding of MPEG-2 Intraframes", *IEEE Transactions on Image Processing*, Vol. 8, No. 2, Feb. 1999, pp. 179-191.