An adaptive quantization algorithm in distributed video coding

Liang Cai [1, a *], Dengyin Zhang[1,b]

[1]Nanjing University of Posts and Telecommunications, Nanjing 210003, China

[a]cailiang0527@126.com, [b]zhangdy@njupt.edu.cn

Abstract. Quantizaion is the first step in converting the analog signal into digital signal. In the distributed video coding, a 4x4 quantization matrix is used to quantize each coefficient. However, such a scheme is a fixed quantization matrix and can not be adaptively adjusted according to the characteristics of the video stream. This paper proposes an adaptive quantization scheme(AQ). Under the premise of guaranteeing video quality, the algorithm adjusts the threshold quantitatively by estimating a correlation noise model on the encoding side. It can also improve the rate-distortion performance of video decoding and ensure the requirements of distributed video coding on the encoding side as simple as possible.

Introduction

In distributed video coding system, quantization is part of the bit rate control, and the important technology to realize bit allocation. Vijay Kumar Kodavalla[1] introduces the development and research hot issues of distributed video coding. Schwarz H, Wiegand T[2] introduces the standard of H.264 video coding, and expounds the problems and challenges faced by H.264 technology, which contains quantization. Sheng F[3] elaborates the quantization in H.264 standard. Weerakkody[4] proposes an adaptive quantized structure of DVC. Chien WJ[5]proposes a distributed video coding in transform domain according to adaptive quantization scheme of rate-distortion performance. CraveO[6] proposed a video compression scheme based on MDC and Wyner-Ziv (WZ) video coding schemes. Each frame is first subjected to wavelet transform, and then the transformed frequency band is quantized by a multi-description scalar quantizer.However, the decoding complexity is greatly increased. Yang CL[7]achieves the DVC system's adaptive quantization of transform coefficient through comparing and adjusting the quantization matrix. The algorithm is simple and effective to improve the rate-distortion performance of the system, but the threshold used in the experiment is empirical and have weak universality and portability.

Quantization in encoder

The simplest quantization method is uniform scalar quantization, which is to find an integer multiple of the fixed resolution closest to the original signal x. In the Wyner-Ziv coding of DCT domain, x represents a DCT coefficient of a frame image after 4x4 DCT transform. After DCT transform, an image is decomposed into 16 sub-bands in the frequency domain.In each 4x4 transform window, 0 is DC sub-band (DC sub-band), 1 to 15 are AC sub-band (AC sub-band) and arranged from high to low.

In the common quantization matrix, the quantization order of DC sub-band and high frequency AC sub-band is obviously higher than that of low frequency AC sub-band. During the image processing, the high frequency information determines the details of the image and the low frequency information determines the image contour. The DC information and the low frequency information have more influence on the quality of the image, so the quantization order of the high frequency sub-band is higher.Among the distributed video coding of DCT domain, the coefficients of DC sub-band are non-negative. While the coefficients of the AC sub-band are basically symmetric with respect to 0. Therefore, when calculating the quantization step size, the sub-band coefficient is D,and DC sub-band

is taken within $[0, D_{max}]$. While AC sub-band is taken $[-|D|_{max}, |D|_{max}]$ for quantization, and then divided into $2^M - 1$ intervals, so that we can get the quantization step size Q_i of i sub-band coding.

$$Q_i = \begin{cases} \dfrac{D_{max}}{2^{M_i}} & D \in DC \\ \dfrac{2|D|_{max}}{2^{M_i} - 1} & D \in AC \end{cases} \tag{1}$$

As shown in Fig.1, the quantization matrix is increased by value 2, the quantization coefficient is adjusted according to the sub-band characteristic of the coefficient band corresponding to the 2-valued quantization matrix.
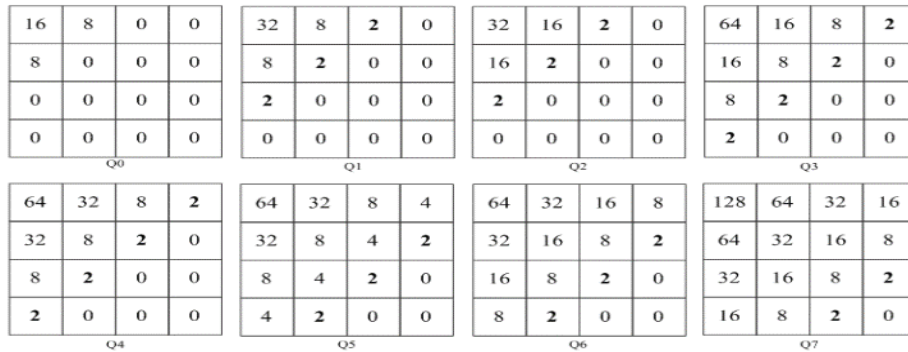
16	8	0	0
8	0	0	0
0	0	0	0
0	0	0	0

Q0

32	8	2	0
8	2	0	0
2	0	0	0
0	0	0	0

Q1

32	16	2	0
16	2	0	0
2	0	0	0
0	0	0	0

Q2

64	16	8	2
16	8	2	0
8	2	0	0
2	0	0	0

Q3

64	32	8	2
32	8	2	0
8	2	0	0
2	0	0	0

Q4

64	32	8	4
32	8	4	2
8	4	2	0
4	2	0	0

Q5

64	32	16	8
32	16	8	2
16	8	2	0
8	2	0	0

Q6

128	64	32	16
64	32	16	8
32	16	8	2
16	8	2	0

Q7

Fig.1 Quantitation matrix

Adaptive quantization algorithm

In order to simplify the calculation, the AC sub-bands of the DCT coefficients are first divided into six groups, namely {1,2}, {3,4,5}, {6,7,8,9}, {10,11,12}, {13,14}, {15}, the quantization order corresponding to the sub-band in each group is the same. Then, the thresholds corresponding to the group {3,4,5}, {6,7,8,9}, {10,11,12}, {13,14} are calculated, and the quantization order of the group is adjusted according to the threshold. Algorithm flow chart shown in Fig.2.
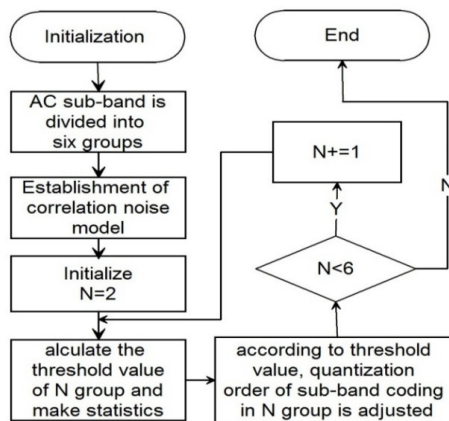

Fig.2 Improved adaptive quantization algorithm

Correlation noise model

The laplacian distribution model in the transform domain has a good balance between computational complexity and accuracy, so it has been widely used.

Step1. The residual frame of the WZ frame and side information is calculated by:

$$R_{XY}(x, y) = WZ(x, y) - SI(x, y) \tag{2}$$

Step2. The residual frame R_{XY} is subjected to a 4×4 DCT transform to obtain a DCT residual coefficient $C_{R_{XY}}^{b_k}$ of 16 sub-bands, $k \in \{0, \mathbf{L}, 15\}$ and $C_{R_{XY}}^{b_k}$ represents a set of coefficients of the sub-band b_k in the residual frame R_{XY}.

Step3. The distribution of the residual coefficients in sub-band b_k is described by the Laplacian probability density function with parameter a_{b_k}:

$$p\left(C_{R_{XY}}^{b_k}(m,v)\right) = \frac{a_{b_k}}{2} \exp\left(-a_{b_k} \left| C_{R_{XY}}^{b_k}(m,v) \right|\right) \tag{3}$$

The calculation of a_{b_k} is divided into variance estimation and maximum likelihood estimation respectively as follows

$$a_{b_k} = \sqrt{2/s_{b_k}^2}, s_{b_k}^2 = E\left(\left|C_{R_{XY}}^{b_k}\right|^2\right) - E\left(\left|C_{R_{XY}}^{b_k}\right|\right)^2 \tag{4}$$

$$a_{b_k} = N / \sum \left| C_{R_{XY}}^{b_k}(m,v) \right| \tag{5}$$

where $p(\cdot)$ denotes the Laplacian probability density function, $C_{R_{XY}}^{b_k}(m,v)$ denotes the DCT residual coefficient of the b_k sub-band located at (m,v), and $s_{b_k}^2$ is the variance of the b_k sub-band residual coefficient $C_{R_{XY}}^{b_k}(m,v)$.

Adaptive threshold calculation

In the AC sub-band, most of the coefficients are in the vicinity of value 0, while the real importance of those not quantized into 0 coefficients. If a majority of AC sub-band values do not need to participate in the quantization, but through the quantization matrix quantization is not actually quantified, so not only wasted the coding side of the computing performance, but also increase the bit rate. Therefore, this paper adjusts the quantization matrix by computing the residual model.

step1. The encoding side calculates the simple side information, the formula is as follows:

$$Y(x,y) = \frac{1}{2}\left(X_B(x,y) + X_F(x,y)\right) \tag{6}$$

where $X_B(x,y)$ and $X_F(x,y)$ present the sub-band coefficients of the preceding and succeeding frames, respectively, and $Y(x,y)$ represents the side information.

step2. Calculate the residuals of the WZ frame and side information

$$C(x,y) = \left| X(x,y) - Y(x,y) \right| \tag{7}$$

where $X(x,y)$ is the DCT sub-band coefficients of current WZ frame.

step3. Assuming that the sub-band coefficients are uniformly distributed over each quantization interval, i.e. $X \sim U(a,b)$ the mathematical expectation of the sub-band coefficients on each quantization interval is $E(X) = \dfrac{a+b}{2} = a + \dfrac{Q_i}{2}, (a \leq X \leq b)$ $\qquad$ (8)

Usually, the inverse quantization is performed by the following equation.

$$q^{-1}(X) = q(X) \times Q_i + \frac{Q_i}{2}, (a \leq X \leq b) \tag{9}$$

In the inverse quantization, the maximum allowable error is:

$$e = \left| q^{-1}(X) - X \right| \leq \frac{Q_i}{2}, (a \leq X \leq b) \tag{10}$$

step4. According to residual $C(x, y)$, the correlation noise model is calculated by maximum likelihood estimation method.

$$p\left(C(x, y)\right) = \frac{a}{2}\exp(-a\left|C(x, y)\right|) \tag{11}$$

The threshold T is calculated

$$T = \int_{C(x,y)<\frac{Q_i}{2}} p\left(C(x, y)\right)\mathbf{g}d\left(C(x, y)\right) \tag{12}$$

Adaptive adjustment of the sub-band quantization step

In AQ algorithm, sub-band quantization step size in the group {1,2} does not need to be adjusted, and the quantitative value given in figure 1 is directly used to qualify, and the value is set 2^{M^0}. The initial quantization step size of N sub-band group is set $2^{M_N^0}$, and the final adapted quantization step size of N sub-band is set 2^{M_N}. If $M^0 - M_N^0 = 1$, it means N sub-band group is adjacent to group {1,2}. If $2^{M_N^0} = 2$, then $2^{M_N} = 0$, or $M_N = M_N^0$. If $M^0 - M_N^0 \geq 2$, it means N group is not adjacent to group {1,2}, and it belongs to AC sub-band. If $p_q \leq T$, $M_N = M_N^0 + 1$, otherwise, if $2^{M_N^0} = 2$, then set $2^{M_N} = 0$. If $2^{M_N^0} \neq 2$, set $M_N = M_N^0$.

The AC sub-bands are divided into six groups, namely{1,2}, {3,4,5}, {6,7,8,9}, {10,11,12}, {13,14}, {15}, The quantization order corresponding to the sub-band within each group is the same. Then, the thresholds corresponding to the group{3,4,5}, {6,7,8,9}, {10,11,12}, {13,14} are calculated, and the quantization order of the group is adjusted according to the threshold.

Experimental results

The simulation experiments were performed on the soccer, hall, soccer, hall, foreman and coast guard sequences (176×144, Y: U: V 4: 2: 0, 15fps) in the standard video test library. The experimental environment is a CPU for the i7-4700, memory for the 8GB notebook. The software development environment is Matlab 2012b and Visual Studio 2010.

It can be seen from Fig.3 that the correlation noise model proposed is better than the fixed threshold algorithm in rate-distortion performance. At the same rate, for the video sequence Coast, the PSNR of this algorithm is about 0.4dB higher than that of the fixed threshold algorithm. For the video sequence Foreman, the PSNR of this algorithm is improved by 0.5dB compared with the fixed threshold algorithm. For the video sequence Hall, the PSNR of this algorithm is about 0.7 dB higher than that of the fixed threshold algorithm. For the video sequence Soccer, the PSNR of this algorithm is 0.3 dB higher than that of the fixed threshold algorithm.
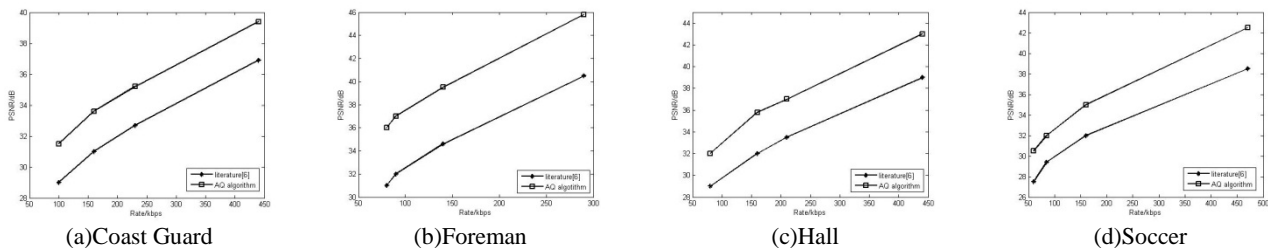

(a)Coast Guard (b)Foreman (c)Hall (d)Soccer

Fig.3 Rate-distortion performance comparison

The 129th standard sequences were selected to test in order to compare the decoding objective image quality of the AQ algorithm and fixed threshold algorithm. Images are shown in Fig.4.
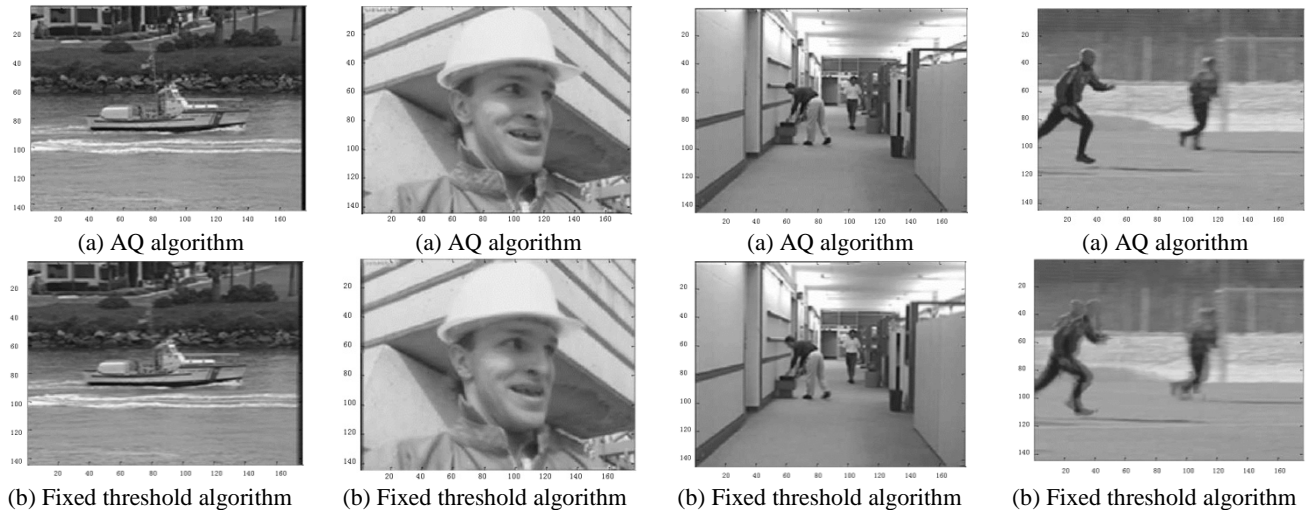
| (a) AQ algorithm | (a) AQ algorithm | (a) AQ algorithm | (a) AQ algorithm |

| (b) Fixed threshold algorithm | (b) Fixed threshold algorithm | (b) Fixed threshold algorithm | (b) Fixed threshold algorithm |

Fig.4 129th frames of test sequences

Since the original algorithm only provides a fixed empirical threshold for testing, we also use the threshold provided for each video sequence , but the threshold is for the Foreman sequence, so in the other sequence Which is not suitable. The AQ algorithm is based on the characteristics of real-time video sequence adjustment, so the image quality has been greatly improved.

Conclusion

The algorithm uses the simple correlation noise model which is a Laplacian distribution established on the encoding end to calculate the threshold value and adaptively adjust the quantization matrix. This paper redesigns the quantization matrix and increases the value of "2" before the quantization matrix does not exist, so as to realize the effect of dynamic adjustment. The experimental results show that the increase of coding complexity is within the acceptable range, and the AQ algorithm improves the overall rate-distortion performance of the system.

Acknowledgement

References

[1] Vijay Kumar Kodavalla. Challenges in practical deployment of distributed video coding. IEEE, 2016.

[2] Schwarz H, Wiegand T. The emerging JVT/H.264L video coding standard[J]. Proc.of IBC2002.

[3] Sheng F, Li-Wei Z,Ling H. An adaptive nested scalar quantization scheme for distributed video coding[C]. Signal Processing Systems,2007 IEEE Workshop on. IEEE:351-356.

[4] W.A.R.J.Weerakkody, W.A.C.Fernando,etal. Nonlinear Quantization for Pixel Domain Distributed Video Coding[J]. Electronics Letters,2009,45(5):261-262.

[5] Chien W J,Karam L J.AQT-DVC: adaptive quantization for transform-domain distributed video coding[C]. Image Processing (ICIP), 2009 16th IEEE International Conference on. IEEE, 2009: 1417-1420.

[6] Crave O, Pesquet-Popescu B, Guillemot C. Robust video coding based on multiple description scalar quantization with side information[J]. Circuits and Systems for Video Technology, IEEE Transactions on,2010, 20(6): 769-779.

[7] Yang CL,Xiao DQ,Po LM,etal. Adaptive quantization in DCT domain for distributed video coding[C]. Data Compression Conference (DCC), 2011. IEEE, 2011: 484-484.