Improved Rate Allocation Algorithm for DVC without Feedback Channel

Yuan-yuan Wu[1], Rui Cai[2], Deng-yin Zhang[3]

Nanjing University of Posts and Telecommunication, Nanjing 210003, China

[1]wuyuanyuan1221@163.com, [2]champions1@163.com, [3]zhangdy@njupt.edu.cn

Abstract.

In the light of the high decoding complexity and large transmission delay caused by the bit rate allocation of Distributed Video Coding (DVC) scheme with feedback channel, this paper proposes an improved bit rate control algorithm without feedback. The algorithm uses the division of macro block to simplify the bit rate allocation, and switches Laplacian parameters in the Correlation Noise Model (CNM) between the block-level and frame-level to adjust rate based on the intensity of motion. The simulation results show that the proposed algorithm can accurately control coding bit rate and transmission delay only at the cost of a small amount of coding complexity, and effectively guarantee the rate-distortion performance of DVC system.

Keywords: Distributed Video Coding; Rate Allocation; Feedback Channel

Introduction

Most Distributed video coding (DVC) system allocates the proper bit rate to encode each video frame by using a feedback channel. Although the feedback channel is easy to implement, it will cause a large transmission delay, affect the real-time video transmission. Due to the influence of feedback channel of WZ encoding, some scholars proposed DVC scheme without feedback channel [1]. Ref. [2] predicts side information at the encoder, and then estimates the rate of each bit plane according to the statistical correlation between the side information and the original W frame.

In most DVC systems, whether it is based on the pixel domain or the transform domain, the whole frame is decoded using the same Laplacian parameters. That is clearly not practical. If the macro block’s motion is severe, underestimation of bit rate will cause decoding error of W frame; if the macro block’s motion is gentle, overestimation of bit rate will reduce the coding efficiency.

This paper proposes a low-complexity rate allocation algorithm without feedback channel. The algorithm uses division of macro block and switches the Laplacian parameter in the correlated noise model between block-level and frame-level to real-time adjust the rate according to the motion intensity of macro block. This algorithm guarantees the low bit rate and high accuracy.

Correlation Noise Model

There is a virtual correlation channel between the source *X* and the side information $\hat{S}$. The $\hat{S}$ can be regarded as an original signal (source *X*) plus the noise. The CNM in DVC can be expressed:

$$X(i,j) = \hat{S}(i,j) - D(i,j) \tag{1}$$

Where (i, j) is the coordinate of the pixel.

Laplacian model is more accurate than the Gaussian white noise model [3]. The residual between W frame and side information SI is regarded as $d(i,j) = W(i,j) - \hat{S}(i,j)$, the probability density function is:

$$p(d) = \frac{\alpha}{2}\exp(-\alpha\,|\,d\,|)\ \alpha = \sqrt{\frac{2}{\sigma^2}}, \sigma^2 = E(D^2) - [E(D)]^2 \tag{2}$$

Improved Rate Allocation Algorithm for DVC without Feedback Channel

Step 1 The encoder divides the image block into fast motion model and slow motion model by using macro block partition [5], which is based on the Sum of Absolute Difference (SAD) criteria.

$$R_{\mathrm{SAD}} = \sum_{x_B=1}^{N}\sum_{y_B=1}^{M}\left|B_W(x_B, y_B) - B_K(x_B, y_B)\right| \tag{3}$$

Among them, *M* and *N* respectively represent the length and the width of macro block B. $B_W(x_B, y_B)$ and $B_K(x_B, y_B)$ respectively represent the DCT coefficients of W frame and K frame.

For each macro block of original W frame, we calculate $R_{\text{SAD}}(n-1)$ at the corresponding position of the previous key frame K_{n-1}, as well as $R_{\text{SAD}}(n+1)$ of the post key frame K_{n+1}. Then compare R_{SAD} with the preset threshold T_0 (selected through the experiment):

$$H_{\text{mod}e} = \begin{cases} 0, & R_{\text{SAD}} < T_0 \\ 1, & R_{\text{SAD}} \geq T_0 \end{cases} \tag{4}$$

$H_{\text{mod}e}=0$ represents a gentle motion, and $H_{\text{mod}e}=1$ represents a severe motion.

Step 2 To calculate the correlation between the original frame W and the side information, and to ensure the low complexity of encoder in DVC system, the encoder uses a rapid prediction of side information $\hat{S}$. In this paper, the average value of K_b and K_f replaces the side information $\hat{S}$, as the Eq. 7.

$$\hat{S}(x,y) = \frac{1}{2}[K_b(x,y) + K_f(x,y)] \tag{5}$$

Step 3 Based on CNM from the previous section, this paper also assumes that the pixel values of noise $d \in D$ follow Laplacian distribution, defined as follows:

$$f_{W|\hat{S}}(d) = \frac{\alpha}{2} e^{-\alpha|d|} \tag{6}$$

Most of the literature uses the Laplacian parameters α_F on frame-level, in Eq. 3, $M = M_1$, $N = N_1$, and M_1 is the length of frame, N_1 is the width of frame, then we get α_F. In the interpolation process of the side information, the uneven intensity of video motion causes that different aspects of the frame have different noise models, so the Laplacian parameter estimation can be extended to the block level. Block-level Laplacian parameter α_B calculation is as follows:

in Eq.3, $M = m$, $N = n$, here m is the length of block, *n* is the width of block, then calculate α_B .

Considering the variance of a block tends to be 0 very likely, this paper revises α :

$$\alpha = \begin{cases} \alpha_B \,, \; \sigma_F^2 < \sigma_B^2 \\ \alpha_F \,, \; \sigma_F^2 \geq \sigma_B^2 \end{cases}$$

(7)

Step 4 Assuming the *C*-th coefficient tape is $X_C = \{x_0, x_1 ..... x_{L-1}\}$, x_n is the *n*-th coefficient. The coefficient of side information $\hat{S}$ is y_n . Assuming X_C has K bit-planes, the highest bit-plane is B_{K-1} . $b_{n,k}$ is the *k*-th bit-plane of x_n , $\hat{b}_{n,k}$ is the *k*-th bit-plane of y_n . For the *k*-th bit-plane, we can calculate $P(b_{n,k} = 0 \mid \hat{S}, b_{n,K-1}, b_{n,K-2}, ..., b_{n,k+1})$ and $P(b_{n,k} = 1 \mid \hat{S}, b_{n,K-1}, b_{n,K-2}, ..., b_{n,k+1})$.

Assuming $X_n^{k+1} = b_{n,K-1} \cdot 2^{K-1} + b_{n,K-2} \cdot 2^{K-2} + ... + b_{n,k+1} \cdot 2^{k+1}$, X_n^{k+1} is the lower bound of x_n which is calculated by first *K-k-1* bits: $P(b_{n,k} = \theta \mid y_n, X_n^{k+1})$, $\theta = 0 \; or \; 1$.

If $b_{n,k} = 0$, then $X_n^k = X_n^{k+1}$; if $X_n^k = X_n^{k+1} + 2^k$, based on Bayesian formula, we can calculate:

$$P(b_{n,k} = \theta \mid y_n, X_n^{k+1}) = \frac{P(b_{n,k} = \theta, X_n^{k+1} \mid y_n)}{P(X_n^{k+1} \mid y_n)} = \frac{P(X_n^k \mid y_n)}{P(X_n^{k+1} \mid y_n)}, \theta = 0 \; or \; 1$$

(8)

For the k -th bit-plane $b_{n,k}$ of x_n , because of the unknown values $b_{n,k-1}, b_{n,k-2}, ..., b_{n,0}$, x_n must belong to $[X_L, X_U]$, let X_L be the lower bound and X_U be the upper bound of x_n :

$$\begin{cases} X_L = b_{n,K-1} \cdot 2^{K-1} + b_{n,K-2} \cdot 2^{K-2} + ... + b_{n,k} \cdot 2^k = X_n^k \\ X_U = b_{n,K-1} \cdot 2^{K-1} + b_{n,K-2} \cdot 2^{K-2} + ... + b_{n,k} \cdot 2^k + 2^{k-1} + 2^{k-2} + ... + 1 = X_n^k + 2^k - 1 \end{cases}$$

(9)

The Eq.9 shows that, when $b_{n,k} = 0$, $X_L = X_n^{k+1}$ and $X_U = X_n^{k+1} + 2^k - 1$; when $b_{n,k} = 1$, $X_L = X_n^{k+1} + 2^k$ and $X_U = X_n^{k+1} + 2^{k+1} - 1$ .When $x_n \in [X_L, X_U]$, $P(X_n^k \mid y_n) = \int_{X_L\Delta}^{X_U\Delta} \frac{\alpha}{2} e^{-\alpha|x-y_n|} dx$, where Δ is the quantization step.

Then crossover probability between $b_{n,k}$ and $\hat{b}_{n,k}$ is:

$$P_n = \begin{cases} P(b_{n,k} = 0 \mid \hat{S}, b_{n,K-1}, b_{n,K-2}, ..., b_{n,k+1}),\ \hat{b}_{n,k} = 1 \\ P(b_{n,k} = 1 \mid \hat{S}, b_{n,K-1}, b_{n,K-2}, ..., b_{n,k+1}),\ \hat{b}_{n,k} = 0 \end{cases} \tag{10}$$

According to the crossover probability, the error probability can be calculated as: $P_k = \frac{1}{N}\sum_{n=1}^{N} P_n$.

At last, the minimum amount of information R_k needed by decoder to reconstruct the bits can be obtained through $H(W \mid S) = -P_k \times \log_2 P_k - (1-P_k) \times \log_2 (1-P_k)$ .The number of bits is $K \times R_k$.

Test results

We use three standard test video Foreman, Coastguard, Mother-daughter (176 × 144, Y: U: V is 4:2:0, 30 fps). Among them, Foreman has severe motion; Coastguard has rich texture detail, and moderate motion intensity; Mother-daughter has slight motion and static background. Experimental conditions are consistent with the experimental conditions in Ref. [4]

The experimental data of different video sequences by using this algorithm is shown in Table 1.

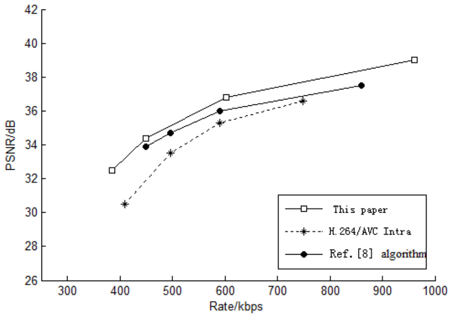
Table 1 The actual rate at different DVC's target rate (kbps)

Coastgard		Foreman		Mother-daughter	
Target Rate	Actual Rate	Target Rate	Actual Rate	Target Rate	Actual Rate
360	360.8	384	385.2	320	318.7
420	420.2	448	449.3	384	382.3
540	538.4	600	602.6	448	446.1
768	768.3	960	961.7	600	597.7

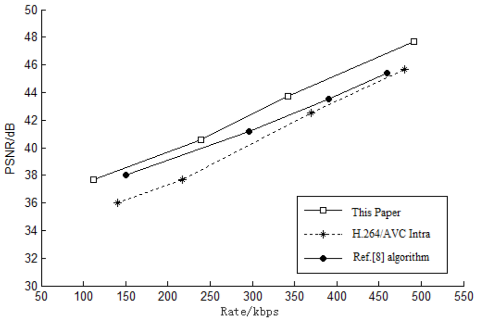
Experimental results show that, when Coastgard, Foreman and Mother-daughter these three video sequences are using Wyner-Ziv coding, bit rate error is less than 0.48%, which means that the actual rate is very similar to the target rate which is set in the experiment. Therefore, coding rate allocation algorithm in this paper can accurately allocate the rate in DVC system.

Foreman sequence and Mother-daughter sequence are encoded respectively with this rate allocation algorithm and the rate allocation algorithm [4] and H.264/AVC intra, Fig. 1 shows the rate distortion performance. Fig. 1(a) shows that when under the same encoding rate, this algorithm can improve the PSNR of decoded image by 1~2dB than H.264/AVC intra; this algorithm can also improve the PSNR by 1dB than Ref. [4]. In Fig. 1(b), this algorithm can improve the PSNR by 1.5 ~ 2.5dB than H.264/AVC intra; this algorithm can improve the PSNR up to 1.4dB than in Ref. [4].

From Fig. 1, under the same PSNR, for the Foreman, compared with the Ref. [4], this algorithm reduces 11%~14% rate; while saving about 17% rate in the Mother-daughter.



(a) Foreman



(b) Mother-daughter

Fig. 1 Comparision of Rate-distortion performance using different algorithm

The reason why our rate allocation algorithm is superior to Ref. [4] is that the Ref. [4] use a frame-level Laplacian parameters α_F to estimate virtual channel model, while we real-time adjust the α switched between the frame-level and block-level, making α more suitable for each video residuals, thus the transmission bit rate encoder need to send becomes more precise.

Conclusion

This paper proposes a rate allocation algorithm without feedback channel. We use the division of macro block to improve the efficiency of distributed video coding. When macro block motions gentlely, there is no need to perform the rate estimation algorithm, which is the coefficients of frame W in the decoder are replaced by the coefficients of side information directly without encoded; when macro block motions severely, we control the rate of frame W. Switching Laplacian parameter α between block-level and frame-level in CNM makes α more suitable for the distribution of video residuals, thus the transmission bit rate becomes more precise.

Acknowledgement

This research work is supported by the National Natural Science Foundations of P.R. China (NSFC) under Grant (61071093), National 863 Program (2010AA701202), Sweden-Asian International Cooperation Project (348-2008-6212), Jiangsu Province Major Technology Support Program (BE2012849), Jiangsu Province IOT application demonstration project (SJ212025), Jiangsu Province industry-university-research prospective joint research project (BY2014014), Jiangsu Province practice innovation plan (CXZZ13-0476), SRF for ROCS, and SEM (NJ209002).

References

[1] Artigas X., Torres L. Improved signal reconstruction and return channel suppression in distributed video coding systems[A]. Mustac I. International Symposium Electronics in Marine[C]. Zadar, Croatia: IEEE, 2005: 53-56.

[2] Morbee M, Prades-Nebot J, Pizurica,et al. Rate allocation algorithm for pix-domain for distribute video coding without feedback channel[C]. Proc of 2007 IEEE International Conference on Acoustics, Speech, and Signal Processing, Hawaii: IEEE, 2007:521-524.

[3] J.Skorupa, J. Slowack, S. Mys, et al.. Accurate correlation modeling for transform-domain Wyner-Ziv coding[C]. In Proceedings of Pacific-Rim Conference on Multimedia, 2008: 1-4.

[4] Brites C, Pereira F. An Efficient Encoder Rate Control Solution for Transform Domain Wyner-Ziv Video Coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(8):1278-1292.

[5] Zhang Dengyin, Wu Yuanyuan, Wan Mingxiang. Improved Side Information Generation Algorithm for Wyner-Ziv Video Coding. The Journal of China Universities of Posts and Telecommunications, 2014, 21(1): 109–115