

An Improved Decoder Motion Estimation Algorithm in Distributed Video Coding

Cunying Chen^{1, a}, Hua Zhang^{2, b}

¹ Information Construction and Management Office, Hua Qiao University, Quanzhou 362021, China;

² College of Business Administration, Hua Qiao University, Quanzhou 362021, China.

^achencunying@hqu.edu.cn, ^bzhanghuav@hqu.edu.cn

Keywords: Unsupervised Learning of Motion Vector, Motion Estimation, Wyner-Ziv video coding, Side Information.

Abstract. This paper focuses on the theory of side information generation, and puts forward an improved scheme based on unsupervised learning of motion vector which improves system's performance. The scheme generates and optimizes side information through the introduction of rate control at the encoder and EM model at the decoder. Lastly, a DVC system is implemented based on the scheme, its principle of work is presented. A MATLAB simulation system has been developed for the DVC system. By comparing with typical side information generation methods, the outstanding performance of the proposed side information generation scheme is justified, and the validity and advantages and feasibility of the proposed DVC system are demonstrated.

1. Introduction

With the development of mobile communications, wireless networks and multimedia technology, video is increasingly being used under different situation [1][2][3]. Many video systems require the terminal to achieve not only low decoding complexity but also real-time encoding and transmission functions, such as mobile video phones, wireless cameras, wireless low-power surveillance systems. Since the constraints of equipment cost and battery capacity and many other factors, processing and storage capabilities of these terminal systems have been greatly restricted, which makes it difficult to perform complex video compression algorithm[4].

Distributed Video Coding (DVC for short in the following paragraph) is new technique which is totally different from the traditional structure of encoded video coding techniques, and meets the needs of solving the above problems and the conditions of limited resources. This technology belongs to the application of distributed source coding, which compresses by eliminating the redundancy between frames of video on the decoding side, in this way DVC thereby obtaining low complexity encoder [5]. DVC employs intra-frame coding and inter-frame decoding and transfers the complexity of the system from the encoding side to the decoding side, which better meet the requirements of video coding and communication.

Distributed encoding is based on channel coding with characteristics of simple coding, high error resilience, high compression efficiency, good robustness and so on [6]. Furthermore distributed encoding provides a new solution to achieve video signals transmission with high robustness. Distributed source coding theory provides a new coding framework, under which the coding complexity is superior to the conventional coding due to the more complex decoding [7]. In a distributed video coding framework, the decoder needs to use the redundant information of the video signal to decode effectively. To achieve a realistic codec, the key question is how the decoder takes motion estimation on decoding side. One possible approach is to use the decoded frame to adopt unsupervised learning through the motion vectors [8]. This method can be applied to the pixel domain coding and transform domain coding, at the same time, it is also the application of the maximum expected value algorithm. But the problem is that this unsupervised motion vector learning can directly lead to a huge computation by calculating the decoding speed [9][10].

This paper improved the implementation of the motion vector unsupervised learning system in distributed video coding by introducing the rate estimation module. The purpose of rate estimate is to roughly estimate the number of checking bits to successfully decode the bit-planes needed before sends the parity bits to decoder. Then the estimated number of checking bits once is transmitted to the decoder. In this way we can reduce the number of check bits feedback channel requests, thus greatly reduce the decoding time and improve operational efficiency of the codec system. This paper then elaborates the main principle of and improved decoding scheme. Finally we use a MATLAB simulation to do the further verification.

2. Model construction

2.1 Rate estimation module.

According to the distributed source coding theory proposed by Wyner and Ziv [11][12], if the statistical correlation among sources is completely known and some distortion is allowed, distributed encoding rate limit is:

$$R_{\min}(D) = \frac{1}{2} \log \frac{\sigma^2}{D} \quad (1)$$

In the actual DVC system we does not know the potential statistical relationship among data, that is, WZ frame with the corresponding side information could not be completely known. In WZ video coding framework based on the transform domain, in order to take advantage of side information to decode and reconstruct WZ frame, side information estimation residuals corresponding WZ frames usually used after Laplace distribution model:

$$f(WZ-SI) = \frac{\alpha}{2} e^{-\alpha|WZ-SI|} \quad (2)$$

Laplace distribution with a minimum rate parameter estimates are as follows:

$$R'_{\min}(D) = \frac{1}{2} \log \frac{2}{D\alpha^2} \quad (3)$$

2.2 The average number of requests feedback channel.

In WZ video coding framework based on the transform domain, the main task is to adjust the code rate according to the side information and statistics between the current changes of statistic correlation in decoded frame[13]. Therefore, some of the characteristics of the feedback channel for designing effective video encoding bit rate estimation algorithm greatly affected:

$$D_{Q_i,j} = \sum_{l=1}^N r_{bjl} / N \quad (4)$$

For low-motion frame, the feedback number of requests will get small. In contrast to the high-motion frames, the feedback number of requests will usually exceed the average, while the number of middle-motion frame will be close to the average.

2.3 The improved rate estimation algorithm.

Based on the above analysis, the greater the amount of motion video frame, the greater the variance of side information and WZ frame corresponding coefficient; the smaller the corresponding Laplace distribution parameters, the bigger the respective feedback request times. Generally, for low-motion frame, decoded bit-planes rate required transferred is relatively small. This paper introduces the rate estimation algorithm is described as follows:

$$R_{\min} = \begin{cases} \min(R'_{\min}, R_{\text{final}}) & \alpha_b \geq \alpha_{bs} \\ \text{mean}(R'_{\min}, R_{\text{final}}) & \alpha_b \leq \alpha_{bs} \end{cases}$$

$$\text{If } r_{bjl} < D_{Q_i,j} \quad (5)$$

$$R_{ini} = \begin{cases} \min(R'_{min}, R_{final}) & \alpha_b \geq \alpha_{bs} \\ \max(R'_{min}, R_{final}) & \alpha_b < \alpha_{bs} \end{cases}$$

If $r_{bjl} > D_{Q_i, j}$

(6)

$$R_{ini} = \begin{cases} R'_{min} & \alpha_b \geq \alpha_{bs} \\ \text{mean}(R'_{min}, R_{final}) & \alpha_b < \alpha_{bs} \end{cases}$$

If $R'_{min} < R_{final}$

(7)

The decoder works on the combination with the checking bite and side information from the encoder. If the decoding error probability is greater than a certain threshold, the decoder request additional parity bits through a feedback channel, this process is repeated until the desired decoding quality is achieved [14][15]. Using this rate control algorithm, both to prevent small movement sequence muzzle velocity transmission rate distortion caused by excessive decline, but will also prevent the initial rate high motion sequences transmitted is too small cause an increase in decoding complexity. Furthermore, rate estimation algorithm used in the direct use of the noise model parameter estimation exercise size distributions, no additional computation and parameter passing is needed. The initial rate estimated will use simple arithmetic. Therefore, under this circumstance no additional coding complexity is added, the number of requests in a channel is reduced, the operating speed of the system is improved, thereby the overall efficiency of the system is increased.

3. Results and Analysis

For the different GOP figures 1 and 2 show RD curve based on the improved unsupervised motion vectors learning algorithm. Figure 1 is Foreman sequence RD curve while Figure 2 is Carphone sequence RD curve. For the factor $Q = 0.5, 1, 2$, and 4. As shown from the results, with the GOP increases the RD performance is improved.

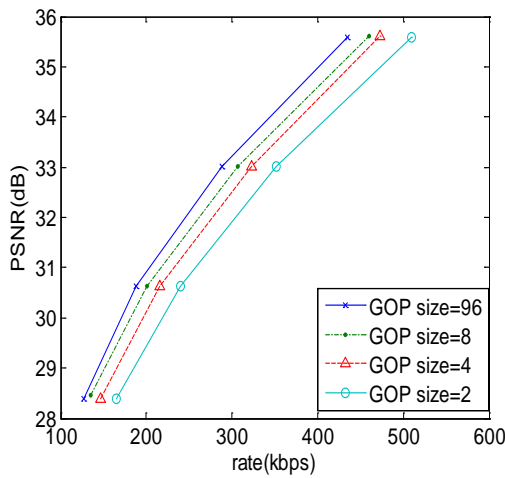


Figure 1. RD performance for Foreman sequence with GOP = 2, 4, 8, and 96.

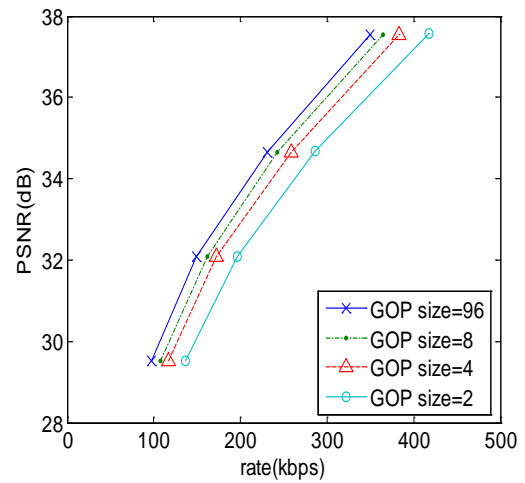


Figure 2. RD performance for Carphone sequence with GOP = 2, 4, 8, and 96.

When GOP = 2, $Q = 0.5, 1, 2, 4$, we adopt motion Oracle codec. And key frame and Wyner-Ziv frame alternately show up. In Wyner-Ziv video codec, we take three methods to decode Wyner-Ziv frame, and compare these three methods. the test results are given in Figure 3 and Figure 4 respectively.

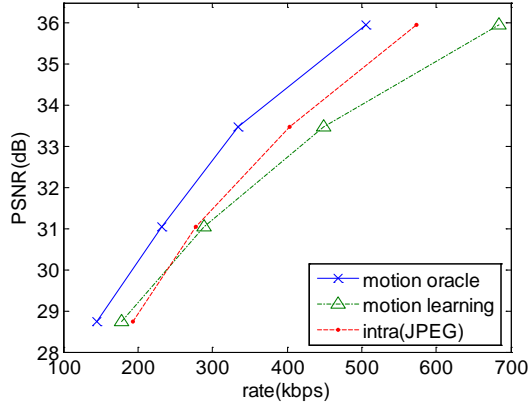


Figure 3. RD performance for Foreman Sequence with GOP = 2.

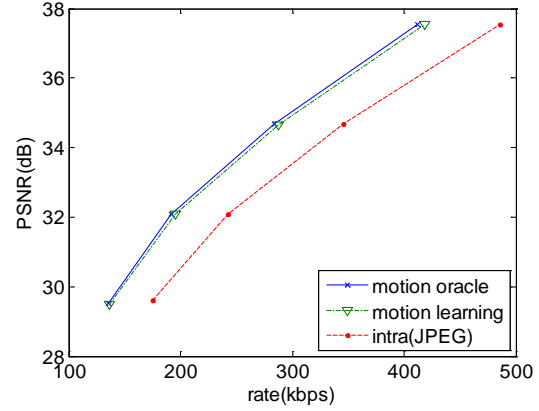


Figure 4. RD performance for Carphone sequence with GOP = 2.

When $GOP = 8$, we adopt three Wyner-Ziv Codec method: motion study, JPEG intra-frame coding and motion Oracle. Figure 5 and 6 respectively show the impact on the Foreman sequence and Carphone sequence RD performance curve. Obviously, for Foreman video sequence, there is only 0.25 dB lost in terms of motor learning method comparing to Oracle. When increasing the size of the GOP, motor learning codec performance increases.

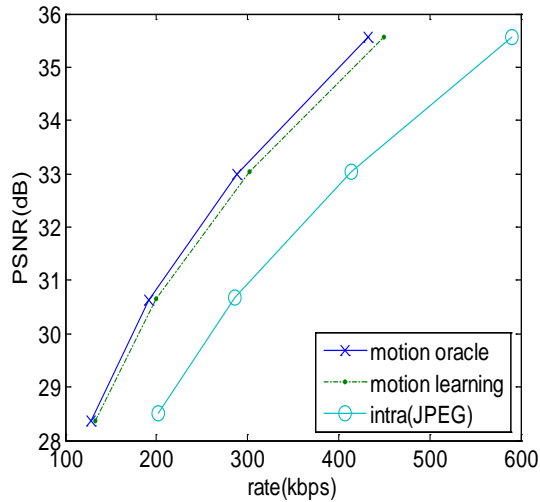


Figure 5. RD performance for Foreman sequence with GOP = 8.

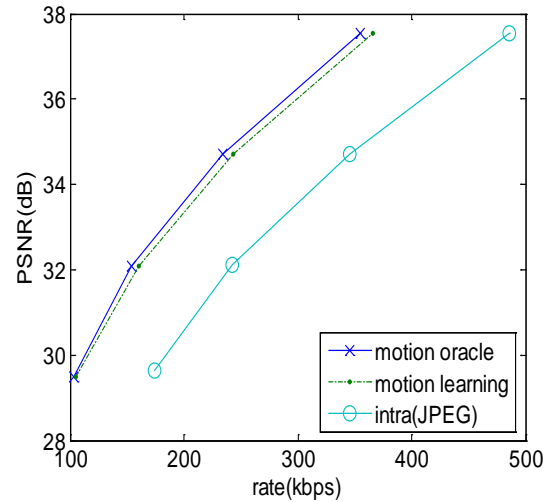


Figure 6. RD performance for Carphone sequence with GOP = 8.

4. Conclusion and discussion

Motion Oracle decoder realizes the optimal combination of motion vectors, the motion search is not required. What is more, unsupervised learning requires to determine the motion vector for each possible combinations, and each possible motion vectors have the same amount of time. The time is determined by the combination of a fixed period of motion estimation and an idle period components. When the block size of $K = 8$, each block has 99 QCIF motion vectors. Each possible motion vectors are 21^2 , the number of combinations is $(21^2)^{99}$. According to these data we can calculate the time taken:

$$F_{\text{learning}} = 0.24 \text{ seconds} \quad (8)$$

$$M_{\text{learning}} = 0.004 \text{ seconds} \quad (9)$$

The time cost of original Oracle and unsupervised learning algorithm motion vector decoder decodes a frame time are as followed:

$$T_{\text{oracle}} \approx 1045 \text{ seconds}$$

$$T_{\text{motion learning}} \approx 10895 \text{ seconds} \quad (10)$$

The improved motion vector unsupervised learning algorithm decoder decodes time-frame required are:

$$T_{\text{oracle}} \approx 477 \text{ seconds}$$

$$T_{\text{motion learning}} \approx 2093 \text{ seconds} \quad (11)$$

As can be seen, the improved motion vector learning algorithm greatly reduces the computational complexity of decoding.

Our improved unsupervised motion vector learning adopt the rate estimation and do simulation in transform domain. Simulation results show that the rate estimation modules greatly improved the speed of the system without reducing the video decoding quality, reduced the overall run time of the system. This paper presents an improved unsupervised motion vector learning system, and the run time is 1/5 of the original system but the overall efficiency of about 5 times.

References

- [1] Yang Y, Xiong Z. Wyner-Ziv coding based on TCQ and LDPC codes[Z]. Google Patents, 2008.
- [2] Fleming M, Zhao Q, Effros M. Network vector quantization[J]. IEEE Transactions on Information Theory. 2004, 50(8): 1584-1604.
- [3] Rebollo-monederio D, Zhang R, Girod B. Design of optimal quantizers for distributed source coding[C]. 2003.
- [4] Aaron A, Setton E, Girod B. Towards practical Wyner-Ziv coding of video[C]. Citeseer, 2003.
- [5] Aaron A, Rane S, Setton E, et al. Transform-domain Wyner-Ziv codec for video[C]. Citeseer, 2004.
- [6] Puri R, Ramchandran K. PRISM: A new robust video coding architecture based on distributed compression principles[C]. Citeseer, 2002.
- [7] Rebollo-monederio D, Aaron A, Girod B. Transforms for high-rate distributed source coding[C]. Citeseer, 2003.
- [8] Aaron A, Rane S, Girod B. Wyner-Ziv video coding with hash-based motion compensation at the receiver[C]. Citeseer, 2004.
- [9] Pradhan S S, Ramchandran K. Enhancing analog image transmission systems using digital side information: a new wavelet based image coding paradigm[C]. 2001.
- [10] Sehgal A, Ahuja N. Robust predictive coding and the Wyner-Ziv problem[C]. 2003.
- [11] Sehgal A, Jagmohan A, Ahuja N. Wyner-Ziv coding of video: an error-resilient compression framework[J]. IEEE Transactions on Multimedia. 2004, 6(2): 249-258.
- [12] Xiong Z, Liveris A D, Cheng S. Distributed source coding for sensor networks[J]. IEEE Signal Processing Magazine. 2004, 21(5): 80-94.
- [13] Girod B, Aaron A M, Rane S, et al. Distributed video coding[J]. Proceedings of the IEEE. 2005, 93(1): 71-83.
- [14] Barron R J, Chen B, Wornell G W. The duality between information embedding and source coding with side information and some applications[J]. IEEE Transactions on Information Theory. 2003, 49(5): 1159-1180.
- [15] Pishro-nik H, Rahnavard N, Fekri F. Nonuniform error correction using low-density parity-check codes[J]. IEEE transactions on information theory. 2005, 51(7): 2702-2714.