A New Multiple Description Layered Coding Method Over Ad-hoc Network*

Song Xiao   Chengke Wu   Fang Zhang

ISN National Key Lab., Xidian University, Xi'an 710071 China

ABSTRACT

A new combined multiple description and layered coding method based on adaptive frame insertion (AFI-MDLC) for transmission of video bit stream over wireless ad-hoc network is proposed in this paper. The main idea of the method is to adaptively insert the transition frames according to the relative motion between two neighboring frames, and then divide the video sequence into two descriptions with independent prediction loops. Afterwards, the base layer and the enhancement layer bit stream for each description are generated and transmitted across wireless network, which is modeled with finite state Markov process. A multiple-path transmitting strategy is presented for the AFI-MDLC data packets according to the characteristics of ad hoc multi-hop wireless network. The experimental results show that the method can help the decoder recover from packet loss more quickly compared with previous methods, and provides more stable and better quality for the reconstruction of video sequence.

1. INTRODUCTION

An ad-hoc network is a dynamically re-configurable wireless network with no fixed infrastructure[1]. In such networks, besides having quite high transmission bit error rates during fading period, the topology of it may change frequently and unpredictably, which makes video transmission over ad-hoc networks more challenging than over conventional wireless network. Since all nodes in an ad-hoc network can be connected dynamically in an arbitrary manner, it is usually possible to establish more than one path between a source and a destination given their mesh topology. This characteristic of multiple paths can be used for a video coding and transmission scheme to combat transmission errors.

There are two main approaches that have been proposed for video transmission over lossy networks. One is Layered Coding (LC) [2-3], in which a video sequence is coded into a base layer and one or more enhancement layers. The base layer guarantees a basic display quality, and each enhancement layer refines the video quality. Users can subscribe up to a certain layer according to their capacities. Enhancement layer packets can be dropped at the congested node to let the base layer get through. So it's highly adaptive to the dynamic variation of network bandwidth. However, the enhancement layer cannot be decoded if the base layer is lost. Thus LC requires the base layer to be transmitted in an essentially error free channel, realized via either strong FEC or ARQ schemes.

Another approach is Multiple Description Coding (MDC) [4-5]. Similar to LC, MDC also generates several sub-bit streams called descriptions. What differentiates MDC from LC is that it does not impose any dependency among its descriptions such that each extra successfully received description improves the quality further regardless of what has been received so far. A signal of acceptable quality can still be reconstructed if only one description is available. This property makes the bit stream of MDC more robust than that of LC when transmitting over lossy networks.

The different characteristics of LC and MDC motivate us to look for an adaptive approach to combine their advantages so as to provide reliable video communication over wireless ad-hoc network. Consequently, we propose a new combined multiple description and layered coding method based on adaptive frame insertion (AFI-MDLC) for video transmission in this paper, and a multiple-path transmitting strategy is also presented for the AFI-MDLC data packets over ad hoc multi-hop wireless network.

This paper is organized as follows. Section 2 presents the system architecture of the proposed method and describes the new AFI-MDLC method in detail. In section3, the performance of the proposed scheme is examined compared to other algorithms. The conclusion of the paper will be presented in Section 4.

2. PROPOSED SYSTEM ARCHITECTURE

2.1. System overview

* This project is supported by the National Natural Science Foundation of China (No. 60372043 and No. 60272050).

The proposed system architecture is shown in figure 1. The raw video stream is coded into two separate bit streams by multiple description generation based on adaptive frame insertion (AFI). Each stream forms a description with different prediction and passes the layer encoder to generate base layer and enhancement layer bit stream. Then the streams are packed and transmitted by multi-path scheduling over different ad-hoc path undergoing independent error effects. The receiver sends negative acknowledgments (ARQ requests) back to report packet losses, which are assumed to be error free. The size of such feedback is small enough so that the transmission time for these packets is negligible. If there are no errors and both descriptions are reconstructed correctly, then both of them are decoded to frames which are interleaved for final display. If one description has an error, the other independently decodable description can still be accurately and straightforwardly decoded to produce usable video at the cost of the reduction of the frame rate.

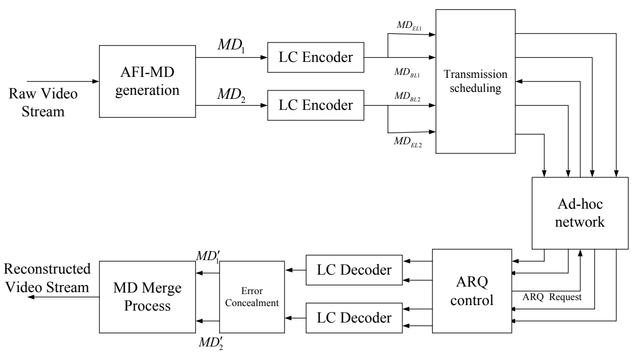

Figure 1. System architecture of the proposed method

2.2. AFI-based multiple description generation

Previous method[5] averagely divided original video stream into two subsequences with even frames and odd frames to form two descriptions, which is efficient when there exists only slow moving of object or background. While for fast moving sequence, the multiple description separation will further increase the magnitude of motion vectors of macroblocks between two neighboring frame, which may cause the encoder to frequently select intra mode to encode the macroblock, thus lead to drastic fluctuate of the coding rate. Furthermore, if an error occurs in the large motion vector, the reconstructed video quality will decrease drastically, and it will be very difficult for the decoder to recover form the error. As a result, a new multiple description generation method for transmission of video bit stream over lossy network is proposed in this paper. The method inserts transition frames according to the relative motion between two neighboring frames, and then divides the video sequence into two descriptions. The larger the relative motion is, the more transition frames are inserted in the video sequence. The method increases the relativity between frames by inserting the transition frame, which can not only improve the reconstructed quality of video with only one correctly received description due to bad channel conditions at the receiver, but also overcome the problem of drastic fluctuate of coding rate when there exists fast motion in original video sequence, thus may decrease the realization complexity of post processing application such as rate control. The detailed method for multiple description generation is described as follows.

Firstly, a criterion is determined for the frame insertion.

Define SAD as the parameter to describe the relative motion degree between two neighboring frames, which is

$$SAD_i(x,y)=\sum_{x=0}^{16}\sum_{y=0}^{16}\left|X(x,y,t)-\hat{X}(x,y,t-1)\right| \qquad (1)$$

In order to take into account both global motion and local motion, the macroblock is used as the unit to the calculation of SAD. So SAD_i denotes the relative motion of the i^{th} macroblock between two frames. $X(x,y,t)$ and $\hat{X}(x,y,t-1)$ denote the pixel value of point (x,y) in current and previous frame respectively.

Given a threshold T_{SAD} of SAD, there are two criterions that can reflect the relativity between two neighboring frames. They are

$$\max_{i\in(1,N_{MB})} SAD_i > T_{SAD} \qquad (2)$$

and

$$N_{SAD_i>T'_{SAD}} / N_{MB} > T_f \qquad (3)$$

Where $N_{MB}=(M\cdot N)/(16\cdot 16)$, which denotes the number of macroblocks contained in one frame with M in vertical resolution and N in horizontal resolution. $N_{SAD_i>T'_{SAD}}$ denotes the number of macroblocks whose value of SAD are larger than a threshold of T'_{SAD}. T_f denotes another threshold of fraction.

The equation 2 reflects the local motion of some macroblock while the equation 3 reflects the global motion of large part of macroblocks. A transition frame will be inserted when either of two equations is satisfied.

Secondly, the algorithm of calculating the inserting frame is determined.

Considering the complexity of the algorithm, there are three approaches for inserting the frame.

1. inserting the average value of two neighboring frames.
2. inserting a copy of previous frame.
3. inserting a copy of next frame.

Since the multiple description separation is finished before the video coding process of motion estimation, there is no information about the motion of macroblocks in the whole video sequence. At the same time, the motions between frames are often very complicated. (e.g. Scene change, single object moving, multiple objects moving etc.) Therefore, inserting the frame by directly averaging two neighboring frames may not be suitable. Sometimes, it may introduce additional error to video encoding process. Consequently, the second or the third approach may be used according to the criterion in our method.

Assuming F_{MVj} denotes the macroblock motion fraction of the j^{th} frame relative to the $(j-1)^{th}$ frame, and $F_{MVj} = N_{SAD_i > T'_{SAD}} / N_{MB}$. Then the algorithm can be described as follows:

If $F_{MVj} > F_{MV(j+1)}$ then

inserting a copy of the $(j-1)^{th}$ frame before the j^{th} frame;

Else

inserting a copy of the j^{th} frame before the j^{th} frame.

End if.

2.3. Multiple-path transmission scheduling

There are four kinds of bit streams generated by the proposed AFI-MDLC scheme. They are the base layer bit stream MD_{Bl1} and the enhancement layer bit stream MD_{El1} of the first description and the base layer bit stream MD_{Bl2} and the enhancement layer bit stream MD_{El2} of the second description. Four different paths of ad-hoc network can be selected to transmit them, which are defined as P_1, P_2, P_3 and P_4 respectively. Since there exists no fixed infrastructure in an ad-hoc network, $P_i (i = 1,2,3,4)$ here refers to four paths selected randomly by router protocol regardless any dynamically change of link when transmitting different packets. Assuming $P_1(n), P_2(n), P_3(n)$ and $P_4(n)$ represent the n^{th} packet transmitted through four paths respectively, the multiple-path transmission scheduling strategy can be described as follows:

1. transmitting MD_{Bl1} and MD_{El1} with P_1 and P_2 , meanwhile transmitting MD_{Bl2} and MD_{El2} with P_3 and P_4.

2. if $P_2(n-1)$ or $P_4(n-1)$ is lost, which means there exist mistakes when receiving MD_{El1} or MD_{El2} , the strategy will not be adjusted since MD_{Bl1} and MD_{Bl2} are received correctly, and video bit streams are still decodable.

3. if $P_1(n-1)$ or $P_3(n-1)$ is lost, then stop transmitting MD_{El1} or MD_{El2} and adjust to transmit $P_1(n-1)$ of MD_{Bl1} or $P_3(n-1)$ of MD_{Bl2} with P_2 or P_4.

4. Since four paths are randomly selected, it is possible that there exist continuous packet loss in each path. Assuming that the maximum burst error is limited within two packets length (properly packet length chosen can satisfying the assumption), if two packets loss occurred in MD_{Bl1} or MD_{Bl2}, eg. $P_1(n-1)$ and $P_1(n-2)$ are lost, then stop transmitting MD_{El1} and MD_{El2} in P_2 and P_4 , and transmit $P_1(n-1)$ with P_2 and $P_1(n-2)$ with P_4.

3. SIMULATION RESULTS

The effectiveness of the proposed multiple description layered video coding system is examined for football and mobile (352×240 pixel/f, 30f/s) sequences over wireless channel. The FGS coding technique in MPEG-4 is used to generate BL and EL streams. The first frame of each description is I VOP, followed by P VOPs. The total bit rate of the base layer of two descriptions is 256kbps with 128kbps for each description. Rate control algorithm of TMN8 is applied to the base layer stream. A finite state discrete time Markov process is used to model a wireless link in the Ad-hoc network. Suppose a path consists of h links, the n^{th} packet with length l_n is being transmitted along the path and its first bit is transmitted on the first link at time t . Assuming the BER of the link models is $p(t)$, then the loss probability for the n^{th} packet over first link is $P_1^n(t) = 1 - \prod_{i=1}^{l_n} [1 - p(t+i-1)]$, and the loss probability over the entire path for the n^{th} packet starting at time t is $P^n(t) = 1 - \prod_{k=1}^{h} [1 - P_k^n(t + (k-1) \times l_n)]$.

Assuming four paths composed of three links are used for transmitting AFI-MDLC packets over Ad-hoc network. A three-state Markov model was used for the links on each path. The loss probabilities for the states are: $p_0 = 1, p_1 \in [10^{-5}, 10^{-3}]$ and $p_2 = 0$, respectively, which implies that each link can be down, error free, or up but with loss probability of p_1.

In order to help the decoder recover from errors, the proposed system estimated the lost frame by computing the motion between the neighboring correctly received (previous and future) frames in other description and performing a motion compensated interpolation, in contrast to conventional single description system, which estimates the lost frame as the last correctly decoded frame.

Figure 2 shows the performances of MDLC and the proposed AFI-MDLC system when one frame is lost. We can see that all the schemes suffer from severe quality degradation when the frame is lost. But our method recovers from this degradation and stops the error propagation more quickly, and can obtain at most 2dB's gain in PSNR compare to MDLC method. Although the transition frames inserting in our method add additional redundancy to the MDLC, they decrease the relative motion between two neighboring frames, thus only have slightly drop in compression ratio compared to previous method. The redundancy of two methods is shown in table 1.

The performances under the same experiment condition of LC, MDC[5], MDLC and the proposed AFI-MDLC system are shown in figure3. The value of PSNR is obtained by averaging that of 100 coding frames.

It is shown that our method outperforms other methods when the effective bandwidth of network is low, which shows that our scheme is much more resilient to the packet loss than the comparison scheme. However, when the effective bandwidth of network is high, which means the packet loss rate is low, the performance of our method is lower than that of MDC but higher that those of LC and MDLC.

4. CONCLUSION

A new combined multiple description and layered coding method based on adaptive frame insertion (AFI-MDLC) for transmission of video bit stream over wireless ad-hoc network is proposed in this paper. A multiple-path transmitting strategy is also presented for the AFI-MDLC data packets according to the characteristics of ad hoc multi-hop wireless network.

The experimental results show that the method can help the decoder recover from packet loss more quickly compared with previous methods, and provides more stable and better quality for the reconstruction of video sequence.

5. REFERENCES

[1] Pearlman, M.R.; Haas, Z.J.; Sholander, P.; Tabrizi, S.S. "On the impact of alternate path routing for load balancing in mobile ad hoc networks", *Proceedings of MobiHOC*. 2000, pp.3-10, Aug. 2000.

[2] Hayder M. Radha, Mihaela van der Schaar and Yingwei Chen, "The MPEG-4 fine-grained scalable video coding method for multimedia streaming over IP", *IEEE trans. on multimedia*,3(1): pp.53-68, Mar. 2001.

[3] Shiwen Mao; Shunan Lin; Panwar, S.S.; Yao Wang, "Reliable transmission of video over ad-hoc networks using automatic repeat request and multipath transport", *Proceedings of VTC 2001*, vol. 2, pp. 615-619, Oct. 2001

[4] A.Reibman, H.Jafarkhani, Y.Wang, M.Orchard, and R.Puri, "Multiple description coding for video using motion compensated prediction," *IEEE Inter.Conf.Image Processing*, vol.3, pp.837-841, Oct. 1999.

[5] J.Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *Proc. SPIE Conf. Visual Comm. Image Process,* pp.392-409. Jan. 2001.

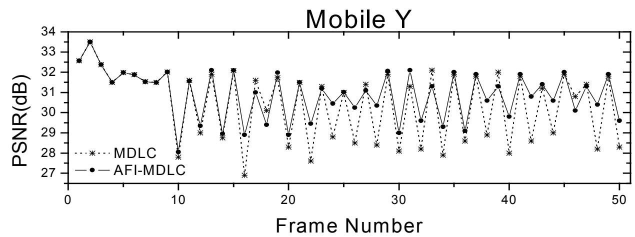

Figure 2. PSNR vs frame number of MDLC and AFI-MDLC

	Football	Mobile
MDLC	2429344 bits	3133824 bits
AFI-MDLC	2469456 bits	3180608 bits
Algorithm redundancy	1.65%	1.49%

Table 1. Redundancy of MDLC and AFI-MDLC

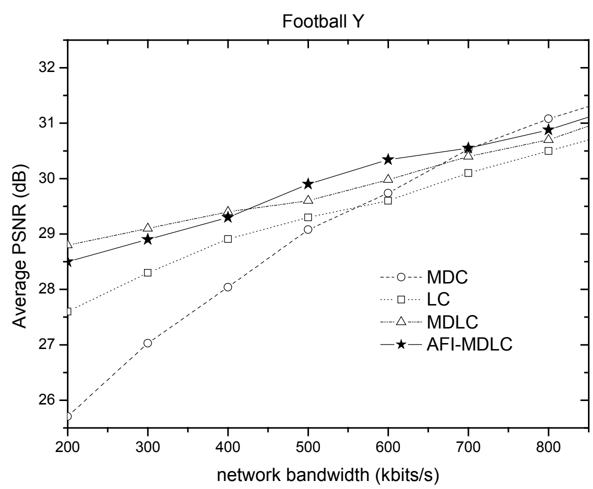

Figure 3. The performance of LC, MDC, MDLC and proposed AFI-MDLC