Fine Granular Scalability Video over Wireless Communication Networks

Tao FANG; Lap-Pui CHAU

School of Electrical and Electronic Engineering
Nanyang Technological University

Abstract

In this paper, we address the problem of joint source-channel coding (JSCC) for fine-granular-scalability (FGS) video over wireless packet-erasure channels in two aspects. First, we study the impact of JSCC applying equal error protection (EEP) to bit-planes of the enhancement-layer and we compare optimal protection resulted from the scheme with fixed protection. Second, we present JSCC scheme using unequal error protection (UEP) between the bit-planes of the enhancement-layer, in which unequal amounts of protection are optimally allocated to progressive streaming to provide a graceful degradation of video quality as packet loss rate varies. The first scheme is easy and simple to be implemented and the second one achieves better performance. Theoretical analysis and experimental results both demonstrate the advantages of the proposed algorithms.

1. Introduction

With the development of broadband wireless networks, such as IEEE 802.11b wireless LAN, delivering video over wireless networks has recently gained increasing attention. Although broadband wireless networks can transmit video data at high bit rates, there are still major challenges existing, such as fluctuations in channel quality and high bit-error rates compared with wired links [1]. Transmission errors, together with lossy source coding techniques, lead to the distortion of the video sequences at the decoder. Hence, proper allocation of system resources to minimize the distortion is required to transmit video efficiently.

JSCC has been proven to be very promising to resolve this problem. In [2]-[8], JSCC schemes over different types of channels, such as AWGN channel, Rician Channel and Rayleigh fading channel, have been discussed. Yet, to the best of our knowledge, JSCC for the scalable video over wireless packet-erasure channels has not been studied extensively. One important contribution of this paper is the introduction of packet-erasure channel model for wireless LAN channels, which allows us to develop JSCC algorithms that take into account the varying channel conditions.

An effective JSCC scheme takes advantage of the differential sensitivities of the output bitstream of video encoder. Srinivasan and Chelleppa [2] focus JSCC on subband coding. They allocate the source and channel rates based on the unequal importance of subbands in reconstruction of the source. Motion-compensated DCT-based SNR scalable coding is discussed by Kondi, et al. [6]. The available bit rate is allocated between all scalable layers and between source and channel coding within each layer. The particular source code algorithm that Cheung and Zakhor [8] use is 3D subband coding. The problem of optimal source and channel allocation between different quantization layers of different subbands is discussed. In this paper, the scalable video coder we apply is FGS video coding, which has been recently adopted by the MPEG-4 as the video coding technique for streaming applications [9], [10] due to its compensation for the unpredictability and variability in bandwidth between sender and receiver over the communication networks. For FGS video, if the most significant bit-plane (MSB) is lost, the other bit-planes will not be decodable. Therefore, on one hand, we adjust the number of bit-planes and the protection bits to control the source and channel rates according to the varying channel condition. On the other hand, we discuss the problem of optimal source and channel allocation between bit-planes of the enhancement-layer.

Because the bit-planes in the enhancement-layer are not equally important, an obvious way of protecting such a bit-stream is to apply UEP to add more protection to the bit-plane that will impact the quality most, such as the MSB. UEP has been addressed by many researchers. Horn, et al. [11] focus on a two-layer scalable video combined with UEP. The proposed approach does not require support from the network and is suited for applications that no feedback channel can be employed. Mohr, et al. [12] applies UEP on Set Partitioning in Hierarchical Trees (SPIHT) coder and find an algorithm for optimizing the amount of protection bits used to protect progressive data. UEP has been also used by to FGS video in [13]. Schaar and Radha study the impact of applying UEP between base- and enhancement-layer, as well as the impact of applying fine-grained loss protection (FGLP) to the enhancement-layer. However, to the best of our knowledge, there has been no work applying UEP in JSCC for FGS over wireless packet-erasure channels. In this paper, we introduce a new JSCC method that combines the scalable advantages of FGS coding with properties of UEP. The method can optimally allocate source and channel rates between the bit-planes of the enhancement-layer based on the time-varying channel condition to perform a graceful degradation of real-time video streaming over wireless LANs.

The rest of the paper is organized as follows. In Section 2, we present our JSCC algorithm using EEP for the bit-planes of the enhancement-layer for video transmission over
wireless packet-erasure channel. In Section 3, we formulate the problem of JSCC using UEP between bit-planes of the enhancement-layer. Finally, we draw the conclusion in Section 4.

2. JSCC with EEP of the Enhancement-Layer

In contrast to the existing JSCC schemes that change the quantization parameter to control the source bit-rate, in our proposed scheme for FGS video, we adjust the number of bit-planes to control the source bit-rate according to the varying channel condition when allocating the source and channel rates. In this section, we discuss the proposed scheme with EEP within the enhancement-layer.

2.1 Optimizing Problem with EEP within the Enhancement-Layer

![Fig. 1. Scheme of Transmission of an FGS stream with EEP in this work.](image)

Without loss of generality, we focus on the error protection for bit-planes in the enhancement-layer, where it is assumed that the base-layer is received correctly and timely. Due to the fine-granularity in FGS, the preceding part in each bit-plane can be decoded to increase the quality of the video even if the tail part in the bit-plane is missing. Therefore, we can packet the data in the way shown in Fig. 1. We first partition the data of the enhancement-layer into $L$ fine bit-planes (note that the fine bit-plane here is not the actual bit-plane in FGS streaming, but the overall information data in each horizontal row as shown in Fig. 1. In the rest of the paper, the bit-plane will mean fine bit-plane without specific explanation.) with the length $K$ and then add $N-K$ FEC codes to each bit-plane. The $L$ rows of data are partitioned vertically into $N$ packets eventually. $B(i,j)$ or FEC in Fig. 1 represents a byte of data.

Below, we discuss the problem of optimally allocating rates of source and channel encoding over all the bit-planes of the enhancement-layer, we want to optimally allocate bits such that the total distortion $D_{S+C}$ is minimized, that is,

$$\text{Min } D_{S+C}, \text{ subject to } R_{S+C} \leq R_{\text{budget}} - R_{BL}$$

(1)

$R_{S+C}$ is defined as

$$R_{S+C} = \frac{R_s}{R_c}$$

with $R_s = R_s(K)$ being the individual source rate and $R_c = R_c(N-K)$ being the individual channel rate, where $N = \left[\frac{R_{\text{budget}} - R_{BL}}{L}\right]$. The objective therefore of the optimization is to find $K$, so that the overall rate does not exceed $R_{\text{budget}}$ while minimizing the overall expected distortion, which can be expressed as

$$D_{S+C} = \Delta D_{S+C}(K) \times P_{\text{Dec}}(K)$$

where $\Delta D_{S+C}(K)$ is defined as the differential improvement due to the conclusion of the enhancement-layer in the reconstruction. Thus, the distortion will be negative since inclusion of the enhancement-layer improves the quality. In our experiment, we generate $\Delta D_{S+C}(K)$ by interpolating the discrete Rate-Distortion values. $P_{\text{Dec}}(K)$ denotes the probability that the enhancement-layer is decodable. The estimate of $P_{\text{Dec}}(K)$ is given by probability mass function (PMF) $P(m,N),m=0,1,\cdots,N$, which is the probability that $n$ packets are lost when transmitting $N$ packets. Then the probability that $n$ packets or fewer packets are lost, can be written as $P(n) = \sum_{i=0}^{n} P(i,N)$.

Thus, the decodable probability of the enhancement-layer is $P_{\text{Dec}}(K) = c(N-K)$. We calculate $P(m,N)$ based on a two-state Markov packet loss model.

2.2 Simulation Results of JSCC for FGS Streams Under EEP

In the experiments, the base-layer is compress using the same QP 24. The bit-stream in the enhancement layer of one frame is transmitted using 100 packets with 114 bytes per packet, i.e., $N = 100$ and $L = 114$. Experiments were performed to transmit enhancement layer of the video sequences over a Two-state Markov channel. The end-to-end performance (average PSNR) over a range of average packet loss rate $P_1$ (from 0 to 40%) with RS channel coding is illustrated in Fig. 2. The value of PSNR in Fig. 2 is the average value of 200 frames of Foreman QCIF. The result using our optimal algorithm is compared with the performance of four fixed-protection solutions. In contrast,
the proposed JSCC solution exhibits excellent performance over a very wide range of average packet loss rate and always performs at or above the performance of any fixed-protection solution. This illustrates the advantage of this method to obtain maximal performance at all times under varying channel conditions.

3. JSCC with UEP of the Enhancement-Layer

B(1,1) B(1,2) B(1,j) B(2,k) B(2,j) B(2,1) B(2,2) B(i,1) B(i,k) B(i,j) B(i,2) B(L,1) B(L,k) B(L,j) B(L,2) B(i,k+1) B(L,N)

Bit-plane 1 Bit-plane 2 Bit-plane i Bit-plane L

Packet 1 Packet 2 Packet j Packet K Packet K+1

Fig. 3. Transmission of an FGS stream with UEP.

From its inception, the FGS structure was designed to be packet-loss resilient, especially with UEP [13]. As the enhancement-layer consists of many bit-planes, UEP could be employed for the bit-planes to allocate system resources more efficiently and provide a good quality of service.

3.1 Optimizing Problem with UEP

We packet the data in the way shown in Fig. 3. We first partition the data of the enhancement-layer into $L$ bit-planes with the length $K_i$ and then add $N - K_i$ FEC codes to each bit-plane. The $L$ rows are partitioned vertically into $N$ packets at last. Below, we discuss the problem of optimally allocating rates of source and channel coding as described in Subsection 2.1 (1).

For $L$ bit-planes, $R_{S+C}$ is defined as

$$R_{S+C} = \sum_{l=1}^{L} R_{S+C,l}$$

where $R_{S+C,l} = \frac{R_{S,l}}{R_{C,l}}$ is the combined source and channel rates for bit-plane $l$. $R_{S,l} = R_{S,l}(K_i)$ is the individual source rate and $R_{C,l} = R_{C,l}(N - K_i)$ is the individual channel rate. The objective therefore of the optimization is to find $K_i$, so that the overall rate does not exceed $R_{budget}$ while minimizing the overall expected distortion.

In this paper, we define distortion per bit-plane as the differential improvement due to the inclusion of this bit-plane in the reconstruction, which is similar with [6]. Therefore, in the absence of the channel errors, only the distortion for the base-layer would be positive and the distortion for all the bit-planes in the enhancement-layer would be negative since inclusion of these bit-planes improves the quality of the video.

As mentioned in Section 2, the distortion of one bit-plane depends not only on the rate of the bit-plane but also on the rates of the previous bit-planes. Due to this dependency, we express the overall distortion as

$$D_{S+C} = \sum_{l=1}^{L} \Delta D_{S+C,l}(K_i | K_{i-1}; \cdots ; K_i) \times P_{Dec}(K_i | K_{i-1}; \cdots ; K_i)$$

In our experiment, we generate $\Delta D_{S+C}(K_i | K_{i-1}; \cdots ; K_i)$ as follows,

$$\Delta D_{S+C,l}(K_i | K_{i-1}; \cdots ; K_i) = \Delta D_{S,C} \left( \sum_{i=4}^{l} K_i \right) - \Delta D_{S,C} \left( \sum_{i=1}^{l-1} K_i \right)$$

where $\Delta D_{S+C}$ is the differential improvement due to the conclusion of the enhancement-layer in the reconstruction, defined in Subsection 2.1. $P_{Dec}(K_i | K_{i-1}; \cdots ; K_i)$ denotes the probability that bit-plane $l$ is decodable. As we can see from Fig. 3, if the first bit-plane is lost, all the following bit-planes will be useless. Thus, more protection should be added to the former bit-planes resulting in $K_1 \geq K_2 \geq \cdots \geq K_L$.

At last, we use the local search hill-climbing algorithm in [12] to solve the optimization problem. The algorithm makes limited assumptions about the data, but is computationally tractable.

3.2 Simulation Results for FGS Streams Under UEP
To evaluate the packet-loss resilience of the proposed JSCC scheme using UEP in the enhancement-layer, the same experiment conditions described in Subsection 2.2 were employed. For comparison, the result of JSCC scheme with EEP in the enhancement-layer in Section 2 is also given. The result of UEP degrade gracefully as the packet loss rate varies whereas that of EEP has a sharp transition at some particular value of packet loss rate, which can be seen from Fig. 4. Evidently the quality level of result by JSCC scheme with EEP drops deeply to a worst case from time to time since all the bit-planes are treated equally for this channel condition. In contrast, the result by JSCC with UEP stays at a relatively stable level when the channel condition is seriously bad, i.e., a significant number of packets are lost. The scheme with UEP alleviates the sharp drops of PSNR to some extent. As expected, UEP outperforms EEP most of the time for this example.

4. Conclusion

In this paper, we proposed two JSCC schemes for FGS video: 1) JSCC using EEP between the bit-planes of the enhancement-layer. The proposed solution efficiently combines progressive source coding with FEC coding. An optimal tradeoff has been found between the bit-planes and the channel protection. 2) JSCC using UEP between the bit-planes of the enhancement-layer. The advantage of the method is that it can optimally allocate source and channel rates between the bit-planes of the enhancement-layer based on the time-varying channel condition to perform a graceful degradation of real-time video streaming over wireless LANs.

As illustrated by our extensive simulation results, the algorithms can provide significant resilience under a wide range of average packet-erasure rate (0–40%).

References