On the Robust Transmission Technique for H.263 Video Data Stream Over Wireless Networks

Han-Seung Jung
School of Electrical Eng.
Seoul Nat’l Univ.
Seoul, Korea
jhs@claudia.snu.ac.kr

Rin-Chul Kim
School of Info. and Comp. Eng.
Hansung Univ.
Seoul, Korea
rin@hsel.hansung.ac.kr

Sang-Uk Lee
School of Electrical Eng.
Seoul Nat’l Univ.
Seoul, Korea
sanguk@sting.snu.ac.kr

Abstract

In this paper, we propose an error-resilient transmission technique for the H.263 compatible video data stream, based on the data partitioning technique. And the proposed algorithm employs the bit rearrangement technique in each layer, which provides the unequal error protection against the channel errors, without requiring additional side information. In addition, we propose the recovery algorithm for the lost or erroneous motion vectors. The proposed algorithm is implemented, based on the H.263 standard, and evaluated through intensive computer simulation. The experimental results demonstrate that the proposed algorithm provides acceptable performance both subjectively and objectively at various bit error rate and burst length.

1 Introduction

Most video source coders employ the MC-DCT coding scheme, which adopts the variable length codes (VLCs), such as the Huffman codes. But these techniques are very vulnerable to the bit errors, causing the error propagation both in the spatial and the temporal domains. In this case, the forward error correction (FEC) and the error concealment (EC) techniques are widely used [1, 2]. And, in order to prevent the corruption of the consecutive VLCs, due to the bit errors, the error-localization technique, such as the error-resilience entropy coding (EREC) [3], is alternative to the insertion of sync-codes. However, the FEC technique does not only degrade the compression efficiency, but also fails with increased bit error rate (BER) or longer bursts. On the other hand, the EC technique, which does not increase the transmission bandwidth, yields much poor performance, when information relating to adjacent blocks is not available. Moreover, even in the EREC technique, the errors in one slot can be propagated to others, though it alleviates the degradation of quality, by packetizing all bits to the fixed length slots with least redundancy.

In this paper, we propose an error-resilient transmission technique for the H.263 [4] compatible video data stream, based on the data partitioning (DP) technique, where each layer is protected against the channel errors, using the hierarchical synchronization technique. In addition, the recovery technique for the corrupted motion vectors is also proposed, based on the proposed error protection technique. The paper is organized as follows. Section 2 describes the data partitioning technique and the hierarchical synchronization technique to provide unequal error protection (UEP). Section 3 presents the recovery algorithm for the corrupted motion vectors. Section 4 examines the performance of the proposed algorithm through intensive computer simulation, demonstrating the effectiveness of the proposed algorithm. Section 5 presents the conclusion of this paper.

2 Hierarchical Synchronization Technique

In this paper, an advanced error-localization technique is proposed, and by applying it to the data partitioning, the unequal error protection (UEP) can be achieved effectively in each layer, without degrading the coding efficiency.

In the data partitioning, the DCT coefficients are divided into the base and enhancement layers, by the priority break point (PBP). The relatively important information (the macroblock header, the motion vectors, and the low frequency DCT coefficients) is transmitted in the base layer and the remnant (the high frequency DCT coefficients) in the enhancement layer. As shown in Fig. 1, the DCT coefficients in each coded block of a macroblock (MB) are separated into two layers, and the coded block information (MCBPC and
CBPY) is regenerated in each layer. Thus, each layer has the same structure as in the H.263, except that there is no information on the motion vector in the enhancement layer.

It is generally assumed that no loss occurs in the base layer [3], or the base layer is loss-free, using sufficient redundancy in FEC codes [6]. But it is not always true for the given transmission bandwidth in the wireless network, in which each layer bit stream is subject to the equi-probable losses, implying that the performance mostly depends on the effects of the bit errors in the base layer. Thus, the base layer, as well as the enhancement layer, should be highly protected against the channel errors at realistic bit rate, which can be achieved by the proposed synchronization technique. That is, to provide the UEP, without increasing the bit rate, and alleviate the degradation in the quality, the error-localization technique is employed hierarchically in each layer, based on the H.263 compatible bit stream structure.

The EREC technique is capable of synchronizing the consecutive VLCs or variable length blocks (VLBs), by packetizing all bits to the fixed length slots with least redundancy [3]. Let \(N\) VLBS be synchronized by the EREC, and assume that the bit errors occur in \(p\) slots, then the EREC can protect the rest \((N - p)\) VLBS at most. Similarly, the EREC can be employed in the H.263 structure, but a simple application to some VLB level is not practical. Notice that each group of block (GOB) is already synchronized by the GOB sync-code, and thus it can be done by simply applying the EREC to the GOB layer. Moreover, since the MBs in a GOB and the blocks in an MB are equally important, the EREC in each level provides more effective protection. So, it is believed that the repeated application of the EREC provides much better performance both objectively and subjectively.

Figure 1: The macroblock structure of the base and the enhancement layer in the data partitioning

Figure 2: The structure of the bit rearrangement in the MB and the block levels, where \(M_k\), \(B_i\), and \(HD\) are the \(k\)-th MB, the \(i\)-th encoded block, and the MB information, respectively.

Firstly, the GOB layer is synchronized by simple application of the EREC, based on the information of total used bits in a frame. Secondly, the bit rearrangement is repeated in the MB level, based on the GOB sync-position. And then the coded blocks in each MB are rearranged within the average MB size\(^1\), using the coded block pattern (CBP) information as shown in Fig. 2, where \(M_k\), \(B_i\), and \(HD\) are the \(k\)-th MB, the \(i\)-th encoded block, and the MB information (the CBP information and the motion vectors), respectively. Since the header information \(HD\) is relatively much more important than the coded block data in the MB level, the hierarchical structure of the bit rearrangement protects the header information in each MB effectively, and provides faster synchronization in the MB and the block levels, alleviating the degradation in video quality. Finally, the synchronization technique is applied to the DCT coefficients level. As a result, the effects of the errors can be confined within each level, since the errors in a slot affect only the high frequency DCT coefficients in others.

Let \(s_{MB}\), \(T\) and \(N\) be the average MB size, the number of bits in a frame, and the number of MBs, respectively, and also let \(M_i\) and \(s_{PB}^i\) be the \(i\)-th MB and the average block size of \(i\)-th MB, then the encoding procedure is summarized as follows:

1. Calculate \(s_M = \left\lceil \frac{T}{N} \right\rceil\), where \(\lceil \cdot \rceil\) is the ceiling function.

\(^1\)the size is defined as the number of used bits
2. Calculate $s_{B}^{1} = \frac{s_{M}}{\# \text{ of blocks}}$, based on the CBP in each MB.

3. Classify all MBs into the set $A = \{M_{i} : \text{length} \leq s_{M}\}$ and the complementary set $A^{c}$.

4. For $M_{i} \in A$, rearrange the DCT coefficients by $s_{B}^{1}$ and synchronize each block.

5. For $M_{i} \in A^{c}$, classify the coded blocks of $M_{i}$ by $s_{B}^{1}$, similar to the step 3. Rearrange the bits for the DCT coefficients within $s_{B}^{1}$ and synchronize each block. And for the blocks, whose length is larger than $s_{B}^{1}$, rearrange the remaining bits in the blocks sequentially at the tail of $M_{i}$.

6. Repeat the step 4 and 5 until all bits in the GOB level are completely rearranged.

Assuming that the picture-sync-code is transmitted reliably, the total bits $T_{k}$ used in the $k$-th frame is given by $T_{k} = F_{k} - F_{k-1}$, where $F_{i}$ is the position of $k$-th picture-sync-code. So, the sync-positions in each level can be derived easily from total bits used for the frame and the CBP information.

By rearranging the VLBs to the fixed length slots, the VLBs, whose length is longer than the average size, are spread to other slots (rearranged in the backward direction). Thus, the VLBs are apt to be corrupted by the errors with high probability. In most cases, however, since only the high frequency DCT coefficients are included in the dependent set\(^2\), the degradation in the video quality is not noticeable. Moreover, the base layer is encoded at the constant bit rate (CBR), by controlling the PBP. That is, the MB and the coded block size can be made nearly constant in each frame, making the size of the dependent set decreased. Notice that, in the EREC structure, as the variation in the size of VLBs increases, the errors are more severely propagated and vice versa [3]. Therefore, the error propagation can be almost prevented in the base layer, and the UEP can be achieved without the loss of the coding efficiency.

3 Recovery of The Erroneous Motion Vectors

The conventional H.263 encodes the motion vectors employing the differential pulse code modulation (DPCM). Thus, even an error in one motion vector could damage the subsequent motion vectors significantly. But, since the differences of the motion vectors and DCT coefficients of the following MBs are successfully decoded in our approach, the corrupted motion vectors can be recovered, by extending the boundary matching algorithm [7]. If the motion vector of the $i$-th macroblock $(m_{x}, m_{y})_{i}$ is corrupted, then it can be recovered, by selecting possible motion vector, minimizing the sum of $\sum_{j < i \leq i_{\text{max}}} F\{D(M_{j})\}$, which is the sum of differences between the boundary pixel value of the $j$-th decoded macroblock $D(M_{j})$ and that of its neighborhood, given by

$$(m_{z}, m_{y}) = \arg \min_{(m_{x}, m_{y}) \in \mathcal{M}} \sum_{k \in G'_{i}} F\{D(M_{k})\}$$

$$d_{U} = \sum_{x \in M_{k}} |M_{k}(x, y_{\text{min}}) - M_{k}(x, y_{\text{min}} - 1)|$$

$$d_{R} = \sum_{y \in M_{k}} |M_{k}(x_{\text{max}}, y) - M_{k}(x_{\text{max}} + 1, y)|$$

$$d_{L} = \sum_{x \in M_{k}} |M_{k}(x, y_{\text{max}}) - M_{k}(x, y_{\text{max}} + 1)|$$

where $\mathcal{M}$ is the set of possible motion vectors, or $\mathcal{M} = \{(m_{x}, m_{y}) : -16.5 \leq m_{x}, m_{y} < 16\}$, which is of half-pixel accuracy, and $G'_{i}$ is the set of MBs affected by the erroneous motion vector, respectively. If the $i$-th motion vector $\tilde{V}_{i}$ is erroneous, then the possible motion vector $\tilde{V}_{k}$ of $\Delta_{j}(\Delta_{j} = \tilde{V}_{j} - \tilde{V}_{j-1}, i < j \leq k)$ yield the $\tilde{V}_{k}$, given by $\tilde{V}_{k} = \tilde{V}_{i} + \sum_{j=i+1}^{k} \Delta_{j}$. For the given motion vector $\tilde{V}_{i}$, the distortion in the boundary area is increased, due to the accumulation of the erroneous motion vectors. Therefore, the corrupted motion vectors can be recovered with more accuracy as the larger $G'_{i}$ is available, even if perfect recovery cannot be guaranteed.

4 Simulation Results

The proposed algorithm is evaluated on the real “Foreman” and “Carphone” sequences, which are in the standard QCIF format with the frame rate of 8.33 frames/sec. The base layer is encoded at CBR and the enhancement at VBR, with the quantizer step size 10. The simulation is performed at various BER, and average burst length varies from 0.1ms to 10ms. Fig. 3 shows the PSNR performance at various BER for average burst length 1.0ms. The overall bit rate of the “Foreman” sequence is 60.68kbps (29.66kbps for the base layer and 31.02kbps for the enhancement layer) and that of the “Carphone” sequence is 42.49kbps (20.31kbps for the base layer and 22.18kbps for the enhancement layer). In comparison, we consider the con-
over error-prone channels. The proposed algorithm provides more efficient protection to the relatively important information. Moreover, the erroneous motion vectors can be recovered effectively. The experimental results demonstrated that the proposed algorithm yields good image quality both subjectively and objectively, in spite of transmission errors, and prevents the error propagation both in the spatial and the temporal domain efficiently.

Acknowledgments
This work was supported by Daewoo Electronics co.

References


Figure 3: The PSNR performance at various BER for average burst length 1.0 ms.

5 Conclusion
This paper presented an error-resilient transmission technique for the H.263 compatible video data stream...