

Picture-group-based priority sorting method for H.264/AVC video transmission over wireless channels

Nina Feng, Yilin Chang, Fuzheng Yang, Zhaoyang Lu, and Peng Li

Xidian University, State Key Lab of Integrated Services Networks, Xi'an 710071, Shaanxi China
E-mail: ninafeng@mail.xidian.edu.cn

Abstract. We propose a priority sorting method that is suitable for H.264/AVC. The proposed method takes into account the unequal importance of encoded video data both in one slice and in one group of pictures. The importance is evaluated by the average distortion over one group of pictures, determined by recursive distortion derivation. A heuristic algorithm is presented to determine the ultimate priority sorting. Experimental results demonstrate that the proposed sorting method provides better error resilience and adaptability. © 2010 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.3479933]

Subject terms: H.264/AVC; group of pictures; priority sorting.

Paper 100317LR received Apr. 16, 2010; revised manuscript received Jul. 12, 2010; accepted for publication Jul. 19, 2010; published online Aug. 19, 2010.

1 Introduction

Layered coding combined with unequal error protection (UEP) is a promising method for video transmission over error-prone channels. Data partitioning¹ in the H.264/AVC standard² is an effective layered coding technique and is utilized in many UEP schemes, owing to its lower overhead and better error resilience.^{3,4} With data partitioning, each partition in one slice has unequal importance for the reconstructed video quality due to their different dependency relationships.

In one group of pictures (GOP), which consists of one intraframe [I-frame or instantaneous decoder refresh (IDR)-frame] and a set of interframes (P-frames and B-frames), the intraframe is the predictive reference frame of subsequent interframes, thus the intraframe is more important than the interframes. Furthermore, previous P-frames are more important than succeeding P-frames, and B-frames are the least important in the event that hierarchical B-frames are not applied. Based on this fact, some video transmission schemes provide unequal error protection to frames with different encoding types.⁵

From the prior analysis, the importance measurement at different coded levels, e.g., slice level, frame level, or GOP level, is a prerequisite for designing an effective UEP scheme. In this work, we develop an algorithm for deriving distortion as a metric to evaluate importance, jointly considering the unequal importance of different partitions in one slice and frames with different positions in one GOP.

Moreover, a priority sorting method, which employs the characteristics of wireless channels, is proposed.

2 Distortion Derivation

According to the definition of data partitioning,^{1,2} the encoded video data of one slice is rearranged into three different partitions, namely, partitions A, B, and C. Without loss of generality, we adopt a simple slice mode, i.e., one slice per frame, thus the encoded data of one frame is divided into three partitions. The average distortion over one GOP is selected as the metric to assess the importance of different partitions in different frames. A GOP is composed of one I-frame followed by $N_G - 1$ P-frames, where N_G denotes the total number of frames in one GOP. Due to its importance, the I-frame is assumed to be perfectly protected and correctly received.

Let f_n^i denote the i th pixel in frame n , and \hat{f}_n^i and \tilde{f}_n^i denote its corresponding reconstructed pixel at the encoder and the decoder, respectively. Moreover, let d_n^i denote the distortion in terms of the mean squared error (MSE) of f_n^i , which can be calculated by:

$$d_n^i = (f_n^i - \tilde{f}_n^i)^2. \quad (1)$$

To analyze the importance of different partitions in frame n , we derive the distortion of each partition in frame n under the assumption that only the partition under discussion is lost and other partitions in the current frame and all partitions in subsequent frames are error-free. As for the error concealment scheme, the modified temporal replacement is adopted, in which the intracoded macroblock (MB) or intercoded MB with lost motion information is replaced by the colocated MB in the previous frame.

To compute d_n^i using Eq. (1), \tilde{f}_n^i is required. For the determination of \tilde{f}_n^i , three cases are considered, depending on which partition in frame n is lost. We refer to an intracoded MB as I-MB, and an intercoded MB as P-MB.

Partition A is lost. When partition A in frame n is lost, the whole frame is corrupted, because partition B and partition C are dependent on it. Hence, for all pixels in the current frame, the reconstructed pixel \tilde{f}_n^i can be expressed as:

$$\tilde{f}_n^i = \tilde{f}_{n-1}^i, \quad i \in \text{I-or P-MB}. \quad (2)$$

Partition B is lost. When partition B in frame n is lost, the pixels in an I-MB are affected, but the reconstructed pixels in a P-MB can be obtained by motion compensated prediction with the correct motion vector, reference frame, and residual data. Let \hat{f}_{n-1}^k denote the pixel from which f_n^i is predicted, and \hat{e}_n^i refers to the quantized prediction error. Accordingly, \tilde{f}_n^i can be formulated as:

$$\tilde{f}_n^i = \begin{cases} \tilde{f}_{n-1}^i, & i \in \text{I-MB} \\ \hat{f}_{n-1}^k + \hat{e}_n^i, & i \in \text{P-MB} \end{cases}. \quad (3)$$

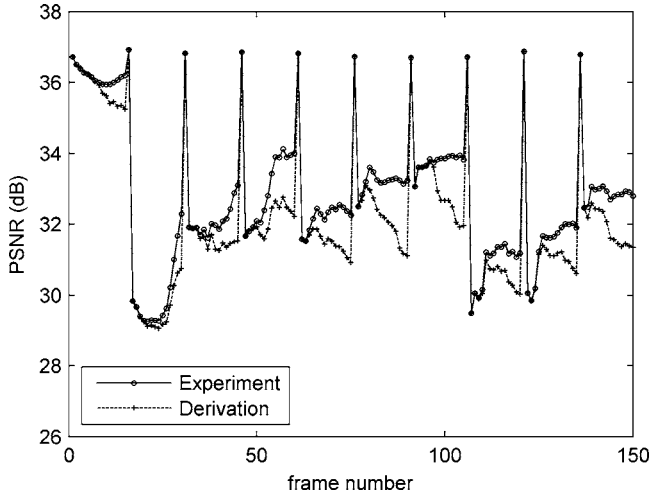


Fig. 1 Comparison of derived distortion and actual distortion for Silent.

Partition C is lost. When partition C in frame n is lost, the reconstructed pixels in an I-MB will not be influenced when constrained intraprediction is utilized. Therefore, the reconstructed pixel \tilde{f}_n can be computed as:

$$\tilde{f}_n = \begin{cases} \hat{f}_n^i, & i \in \text{I-MB} \\ \tilde{f}_{n-1}^k, & i \in \text{P-MB} \end{cases} \quad (4)$$

For frames from frame $n+1$ to frame N_G in the same GOP, a reconstructed pixel in I-MB will be \hat{f}_{n+1}^i , and a reconstructed pixel in P-MB will be $\tilde{f}_n^k + \hat{e}_{n+1}^i$, in which \tilde{f}_n^k can be computed recursively by Eqs. (2)–(4).

As analyzed before, \tilde{f}_n in frame n and successive frames within one GOP can be obtained when any previously mentioned case has occurred. Moreover, let $\bar{D}_{n,1}$, $\bar{D}_{n,2}$, and $\bar{D}_{n,3}$ denote the average distortion for cases 1, 2, and 3, respectively, and L stand for the total pixels in one frame. Substituting \tilde{f}_n^i into Eq. (1) to compute d_n^i , and then the corresponding average distortions, which will be used to assess the importance of the different partitions, are given by:

$$\bar{D}_{n,k} = \frac{1}{N_G \times L} \sum_{i=n}^{N_G} \sum_{j=1}^L d_i^j, \quad k = 1, 2, 3. \quad (5)$$

3 Priority Sorting

Based on the analysis of $\bar{D}_{n,1}$, $\bar{D}_{n,2}$, and $\bar{D}_{n,3}$ for frames from frame 2 to frame N_G , we present a priority sorting method jointly taking into consideration the unequal importance of partitions in one frame and P-frames at different positions in one GOP. That is, all partitions of the first m P-frames and partition A of the middle n P-frames following the preceding m P-frames are labeled as high priority (HP), and partition B and partition C of the n P-frames and all partitions of the remaining p P-frames in one GOP are labeled as lower priority (LP).

For efficient delivery of the coded video data over wireless channels, the coded video data with HP must be delivered to the decoder with small probability of being affected by bit errors, providing basic reconstruction video quality. For each subcarrier in wireless channels employing orthogonal frequency division multiplexing (OFDM) technology, the extent of attenuation subject to fading is different, which can be utilized to categorize the subcarriers into two groups. The subcarriers with the signal-to-noise ratio (SNR) above S_0 are assigned to the high quality (HQ) subchannel group, otherwise the subcarriers are assigned to the low quality (LQ) subchannel group. Thus, the threshold S_0 used for grouping subchannels is determined by the bit error rate (BER) requirement of the coded video data with HP. Let C_H and C_L denote the available transmission capacity provided by HQ and LQ subchannel groups during $N_G - 1$ P-frame intervals, respectively. And C_0 refers to the summation of all partition A for total P-frames. Moreover, let $b_{i,1}$, $b_{i,2}$, and $b_{i,3}$ signify the total coded bits for partitions A, B, and C in frame i , respectively, and $b_{i,0}$ be the summation of $b_{i,2}$ and $b_{i,3}$ for simplicity. Therefore, m , n and p for determining the priority of the total coded video data in all P-frames should satisfy the following constraints:

$$\begin{cases} m + n + p = N_P \\ C_0 \leq \sum_{i=1}^m (b_{i,1} + b_{i,0}) + \sum_{j=m+1}^{m+n} b_{j,1} \leq C_H \\ \sum_{j=m+1}^{m+n} b_{j,0} + \sum_{k=N_P-p+1}^{N_P} (b_{k,1} + b_{k,0}) \leq C_L \end{cases}, \quad (6)$$

where N_P , i.e., $N_P = N_G - 1$, is the total number of P-frames in one GOP.

Given C_H , C_L , and video encoding parameters, m , n , and p can be determined by a heuristic algorithm summarized as follows:

1. Calculate C_0 and subtracted it from C_H , and then initialize m and p to one.
2. Compare $\sum_{i=1}^m b_{i,0}$ with the remaining transmission capacity supported by the HQ subchannel group, i.e., $C_H - C_0$. If $\sum_{i=1}^m b_{i,0}$ is less than or equal to the latter, increase m by one and repeat step 2; otherwise go to step 3.
3. Similarly, compare $\sum_{i=1}^m b_{i,0}$ with $C_H - \sum_{k=1}^{N_P-p} b_{k,1}$. If $\sum_{i=1}^m b_{i,0}$ is less than or equal to the latter, increase m by one and repeat step 3; otherwise go to step 4.
4. Increase p by one and compare $\sum_{i=1}^m b_{i,0}$ with the rest of the transmission capacity provided by the HP subchannel group. If $\sum_{i=1}^m b_{i,0}$ is less than or equal to the latter, increase m by one and repeat step 4; otherwise the procedure is finished.

4 Experimental Results

We first examine the performance of the distortion derivation algorithm. The sequences ‘‘Grandma’’ and ‘‘Silent’’ in QCIF format with a frame rate 30 Hz were encoded using H.264/AVC reference software JM15.1. The GOP consisted of one I-frame and $N_G - 1$ P-frames without B-frames,

Table 1 Average PSNR comparison of the proposed scheme against data partitioning.

Sequence	BER for LP data is 10^{-3}			BER for LP data is 10^{-2}		
	Original	Proposed	Δ PSNR	Original	Proposed	Δ PSNR
Mother and Daughter	36.66	36.89	+0.23	35.62	36.29	+0.67
Grandma	36.62	36.66	+0.04	35.98	36.20	+0.22
Salesman	35.02	35.25	+0.23	33.72	34.62	+0.90
Silent	32.19	32.74	+0.55	30.63	31.72	+1.09
Foreman	29.95	31.47	+1.52	28.66	31.09	+2.43

where N_G was set to 15. In addition, the UseConstrained-IntraPred option was set up, and the quantization parameter (QP) was fixed at 28 for all experiments.

Figure 1 shows the distortion of all subsequent frames versus frame index when partition A of the first P-frame is lost for “Silent.” From Fig. 1, we can observe that the distortion derived accords with the distortion simulated to a large extent, so the distortion determined by Eq. (5) can be used as the metric to evaluate the importance of different partitions of P-frames in one GOP.

Next, the performance of the priority sorting method is verified. Five video sequences of 300 frames, “Mother&Daughter,” “Grandma,” “Salesman,” “Silent,” and “Foreman,” were encoded with the same parameters used in the first set of experiments. Assume that the BER of HP coded video data is 10^{-6} , the BER of LP data is 10^{-3} and 10^{-2} , and C_H and C_L satisfied $C_H:C_L=1:2$. For comparison, the ratios of m , n , and p to N_p were fixed as 1/6, 2/3, and 1/6, respectively.

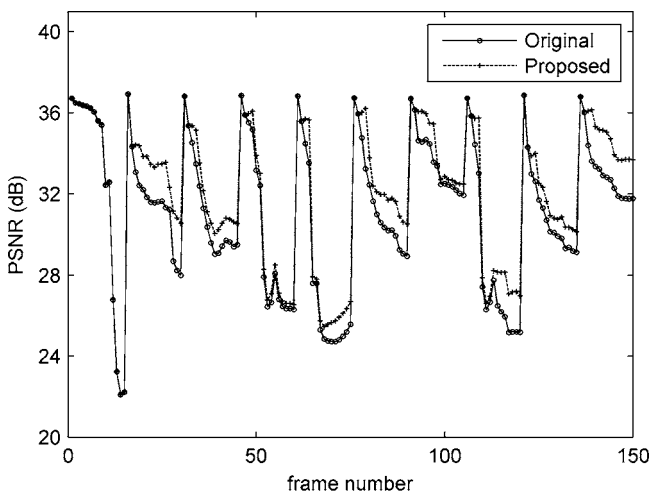


Fig. 2 PSNR comparison of the proposed scheme against original data partitioning for Silent.

Comparing the proposed priority sorting scheme with data partitioning in H.264/AVC at the frame level, the average peak signal-to-noise ratio (PSNR) for all sequences and the PSNR versus the frame number for sequence “Silent” are given in Table 1 and Fig. 2, respectively. From Table 1 and Fig. 2, we can see that the proposed method can improve the reconstructed quality due to its better error resilience and adaptability, especially for sequences with medium motion.

5 Conclusions

A distortion derivation model is proposed by jointly considering the unequal importance of different partitions in one slice and each frame in one GOP. Furthermore, a heuristic algorithm is designed to determine priority sorting. Simulation results show that the distortion derivation model matches the actual distortion well, and the introduced priority sorting method improves the reconstructed video quality under error-prone environments.

Acknowledgments

This research was supported by the National Natural Science Foundation Research Program of China, numbers 60772134 and 60902081, and the 111 project (B08038). We would like to thank the editors and anonymous reviewers for their valuable comments and suggestions.

References

1. S. Wenger, “H.264/AVC over IP,” *IEEE Trans. Circuits Syst. Video Technol.* **13**(7), 645–656 (2003).
2. Telecommunication Standardization Sector of ITU, “Advanced video coding for generic audiovisual services,” Recommendation ITU-T H.264 (2005).
3. B. Barnada, M. M. Ghandi, E. V. Jones, and M. Ghanbari, “Prioritized transmission of data partitioned H.264 video with hierarchical QAM,” *IEEE Signal Process. Lett.* **12**(8), 577–580 (2005).
4. A. Ksentini, M. Naimi, and A. Gueroui, “Toward an improvement of H.264 video transmission over IEEE 802.11e through a cross-layer architecture,” *IEEE Commun. Mag.* **44**(1), 107–114 (2006).
5. C. L. Huang and S. Liang, “Unequal error protection for MPEG-2 video transmission over wireless channels,” *Signal Process. Image Commun.* **19**(1), 67–79 (2004).