H.264 ERROR RESILIENCE CODING BASED ON MULTIHYPOTHESIS MOTION COMPENSATED PREDICTION[+]

Yuh-Chou Tsai and Chia-Wen Lin

Department of Computer Science and Information Engineering
National Chung Cheng University
Chiayi 621, Taiwan

ABSTRACT

In this paper, we propose efficient schemes for enhancing the error robustness of multi-hypothesis motion-compensate predictive (MHMCP) coder without sacrificing the coding efficiency significantly. The proposed schemes utilize the concept of reference picture interleaving and data partitioning to make the MHMCP-coded video more resilient to channel errors, especially for burst channel error. Besides, we also propose a scheme of integrating adaptive intra-refresh into the proposed MHMCP coder to further improve the error recovery speed. Extensive simulation results show that the proposed methods can effectively and quickly mitigate the error propagation and the penalty on coding efficiency for clean channels due to the inserted error resilience features is rather minor.

1. INTRODUCTION

Multimedia applications, such as communications, entertainments, and surveillance, are enabled and getting more and more popular by the evolution of technologies in personal computers and networks. In multimedia applications, compression is a crucially important issue since raw video/audio data demand tremendous storage space and network bandwidth. To reduce the temporal redundancy in raw video data, motion-compensated prediction (MCP) has been widely exploited in the contemporary video coding standards. With MCP, incoming video frames are predicted from previously reconstructed frames stored in frame buffers by spatial displacement vectors (namely, motion vectors). Since only the motion vectors and the quantized prediction errors are coded in the bitstream, the more accurate the prediction, the more compression gain can be achieved. To achieve better prediction accuracy, long-term memory motion compensation (LMMC) was proposed to predict the video frames more effectively among multiple reconstructed frames, leading to bitrate saving of up to 30% using 50 reference frames. The emerging H.264/MPEG-4 AVC standard adopts LMMC as a coding tool to achieve high coding efficiency.

The Concept of MHMCP was first introduced in [1], which reported that the prediction error reductions between 18% and 40% (0.9-2.2 dB) can be obtained with MHMCP. Subsequent research works [2-4] have been dedicated to the theoretical analysis of the compression improvement and the integration of the MHMCP into a standard codec. In [3], an optimal hypothesis selection and predictor coefficient selection algorithm was proposed. A rate-constrained mode decision algorithm was also proposed in [4]. The combination of LMMC and MHMCP is natural and can achieve better coding efficiency. Simulation results in [4] show that the combination of MHMCP (2 hypotheses), variable block-size coding, and long-term memory (10 reference frames) saves up to 30% bit-rate for the same video quality. In [5], an integration of MHMCP into H.264 was presented, which includes syntax adjustment, rate-constrained mode decision, entropy coding model adjustment (CABAC), and multi-hypothesis motion estimation issues.

Besides the compression issues in the multimedia applications, the robustness of multimedia files against transmission errors is also essential, especially for video streaming over unreliable networks. Due to the variable length coding (VLC) adopted in most video coder, a small data error will result in loss of synchronization. Coded data succeeding to the loss will become undecodable even if they are correctly received. Furthermore, due to the MCP exploited in most existing coding standards, the distortion of one video frame which results from transmission errors will propagate to its subsequent frames. These reasons make compressed video data very sensitive to channel errors and cause serious quality degradation. Thus, enhancing the resilience of video streams has become a crucial problem.

The error robustness of an MHMCP codec against transmission errors has been investigated recently. Lin and Wang [6] presented a decoder distortion model (given channel conditions) and a simplified encoder distortion model of a 2HMCP coder. Their proposed algorithm selects the optima hypothesis weighting factors which minimize the total distortion. In [7,8], mathematical models are developed to characterize the error propagation effect caused by packet losses in an MHMCP coder. The error resiliency of MHMCP is also analyzed and compared to the well-known random intra-refresh approach. It is concluded in [7,8] that MHMCP is preferred for low bit rate, whereas intra-refresh outperforms the MHMCP for higher bit rate condition.

In this work, we shall propose efficient methods for further enhancing the error resiliency of the MHMCP coder with two forward hypotheses involved in the prediction process. We also integrate the intra-refresh mechanism with the proposed methods to further improve the error recovery speed.

---

2. MULTIHYPOTHESIS MOTION-COMPENSTAED PREDICTIVE CODING

In a traditional predictive coder, a MB prediction is reconstructed by a given spatial displacement vector (motion vector) and a temporal displacement index (reference frame index). With these two parameters, decoder can find the prediction of a MB/sub-block among the frame buffer which is reconstructed during the decoding process of previous frames. In an MHMCP coder, MB/sub-block predictions for the kth frame are formed by linear combinations of multiple hypotheses, each is a block of pixels specified by a spatial displacement vector and a temporal displacement index as formulated in [6,7]. Note that if the weight for each hypothesis varies, then the weighting information should also be signaled in the video bitstream.

$$\hat{\psi}(k) = \sum_{i=1}^{n} h_i \tilde{\psi}(n, n\text{-}i), \text{ where } \sum_{i=1}^{n} h_i = 1 \qquad (1)$$

In this work, we focus on the 2HMCP codec, in which macroblocks (MBs) are predicted by the linear combination of two forward hypotheses, which are selected from the (k-1)th and (k-2)th reference frames, respectively. The actual prediction for the mth MB is formed by linearly combining two hypothesis using two weighting factors, h and 1-h, as expressed in (2).

$$\hat{\psi}(k) = h \cdot \tilde{\psi}(k, k\text{-}1) + (1\text{-}h) \cdot \tilde{\psi}(k, k\text{-}2) \qquad (2)$$
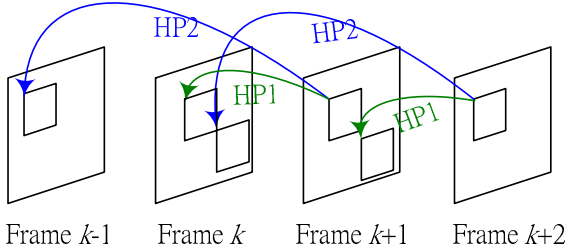

Fig. 1. MB predictions for the 2HMCP.

Fig. 1 illustrates the 2HMCP coder. Due to the leaky prediction adopted in the MHMCP coder. not all the concealment distortion of the corrupted frame will be propagated to the next frame since part of the MB prediction (that is, another hypothesis) comes from the previous error-free reference frame. As a result, the distortion would attenuate and the error propagation is thereby suppressed. A distortion model for the 2HMCP encountering a single-frame loss is presented in (3) which is proposed in [7].

$$D_d(k) = (\frac{1\text{-}(h\text{-}1)^{k+1}}{2\text{-}h})^2 \frac{\theta}{1+\gamma_d k} D_d(0) \qquad (3)$$

where γ_d and θ are constant model parameters, h and 1-h are the weights for the first and second hypotheses, respectively. In this model, the distortion (MSE) value of the kth frame after the lost one is derived by that of the lost frame ($D_d(0)$).

3. PROPOSED ERROR RESILIENCE CODING SCHEMES FOR MHMCP

In a traditional single-hypothesis MCP coder, due to error propagation, frames which are correctly received at the decoder may still suffer from the loss of previous frames. It is because that a prior reconstructed frame is used to predict the current

encoded frame, and so on. With the predictive coding mechanism, when a video frame is lost during transmission, the decoder reconstructs the lost frame by using certain error concealment method, leading to reconstruction distortion. Such reconstruction distortion will result in the mismatch between the encoder's and the decoder's reconstructed frames, thereby leading to drifting error. The drifting error will propagate to the succeeding frames all the way until reaching an Intra-refresh point. In the 2HMCP coder, the error propagation still exists since the predictive coding scheme is exploited as well. To mitigate the error propagation due to packet loss, we propose efficient error resilience coding methods at the encoder and decoder sides, respectively.

3.1. Encoder-Side Error Resilience Tools

A. Hypothesis parameter separation

In the 2HMCP coder, as mentioned above, each hypothesis requires a motion vector and a reference picture index to obtain the predicted block from the corresponding frame buffer. In what follows, the term "hypothesis parameters" is referred to as the motion vectors, the reference picture indices, and other information data required to form the prediction for the two hypotheses. Though the 2HMCP can suppress the error propagation well, the distortion of a lost frame may still be serious. The frame loss cannot be easily recovered by error concealment in the decoder. To reduce the distortion of a lost frame and its error propagation, we propose a hypothesis parameter separation scheme.

For the mth MB of the kth frame, there are two sets of hypothesis parameters required for specifying the two hypotheses used for prediction. These parameters are both coded and embedded into the locations between the MB header and the transform coefficients of residues of the mth MB in the bitstream. When a frame is lost during transmission, all the data, including hypothesis parameters and residue signals, would also be lost since they are usually all encapsulated in a same packet. This is why a packet loss may result in significant quality degradation.
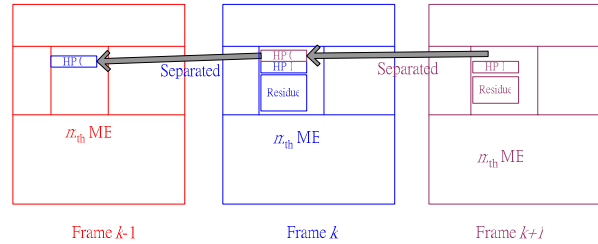

Fig. 2. Hypothesis parameter separation.

In the proposed approach, these two sets of hypothesis parameters are separated in the bitstream to greatly reduce the impact of a single-frame loss. The hypothesis parameters of the mth MB of the kth frame is separated and embedded into the mth MBs of the kth frame and the (k-1)th frame, respectively, as illustrated in Fig. 2. The motivation of the scheme is to mitigate the impact caused by a single-frame loss by separating the most important data into two partitions. As a result, when a single frame is lost, only half of the hypothesis parameters and all the transform coefficient of residues are lost. The decoder can reconstruct the erroneous frame more precisely with the other half of the hypothesis parameters rather than reconstruct the lost

frame by copying the prior reconstructed frame (i.e., zero-MV error concealment). In this way, the distortion of a lost frame would be significantly reduced.

Although the frame after the erroneous one also loses half of its hypothesis parameters, which is not the case that would happen to the original 2HMCP coder, the decoder could still reconstruct this frame using another half of the hypothesis parameters. Only a little drifting error would be introduced due to the prediction mismatch between the encoder and the decoder.

The merit of the hypothesis parameter separation is to mitigate the distortion caused by frame loss without embedding any additional overhead. Besides, this approach does not require a feedback channel between the sender and the receiver sides. A drawback of this method is the extra encoding latency introduced. While separating the parameters, encoder should keep the data of current encoding frame until the hypothesis parameters for the next frame are available. This latency may require a small spare memory in the encoder. Since bitstream data for one frame normally will not occupy too much memory space, this should not be a critical issue.

B. Reference interleaving

The design of 2HMCP shown in Fig. 1 may suffer from burst losses a lot, since both hypotheses will be lost in the original 2HMCP when a burst-loss of length larger than two happens. The next frame after the burst-loss frames will lead to error propagation which is difficult to recover. We propose a "Reference Interleaving" scheme to mitigate the impact of a burst-loss with a length of 2 or more. In our method, we consider only the cases with bur-length of 2 or less because it is shown in [8] that the average burst length caused by random packet loss is around 1.5 to 2. For a burst loss with a long burst-length caused by slow signal fading, packet interleaving can be used to effectively spread out the long burst loss into short individual packet losses. Fig. 3 illustrates the Reference Interleaving method with an interleaving distance of two to combat against packet losses with a length of 2 or less. In this figure, the two hypotheses are selected from the $(k$-1)th and $(k$-3)th reference frames, respectively. This reference selection can effectively mitigate the error propagation due to a burst-loss, since an error-free hypothesis still exists when two consecutive frames are lost.
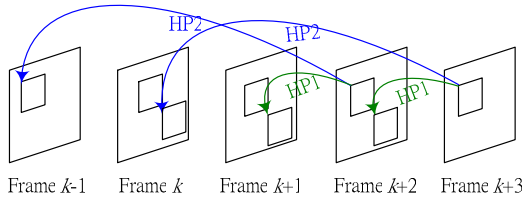

Fig. 3. Reference interleaving.

3.2. Decoder-Side Error Resilience Tool

Clean hypothesis formation

At the decoder, we propose an efficient method, namely "Clean Hypothesis Formation" as illustrated in Fig. 4, to further mitigate the error propagation effect due to random and burst packet losses. Assume that frame k is lost during transmission. At the decoder side, the predictions of MBs within subsequent frames (k and k+1) which contain the hypothesis predicted from frame

k-1 will only be estimated with the error-free hypothesis. That is, the corrupted hypotheses are discarded when constructing the predictions of the subsequent frame at the decoder side.
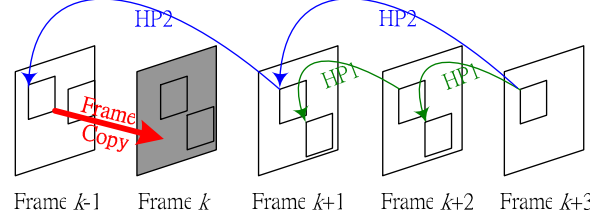

Fig. 4. Clean hypothesis formation.

With this method, the distortion of the frames after the erroneous one can be significantly reduced compared to the original 2HMCP method. Although using a different prediction at the decoder side may cause a little drifting error due to the prediction mismatch between the encoder and decoder, the overall resilience against frame loss is improved, especially for single-frame loss situations.

4. EXPERIMENTAL RESULTS

In our experiments, the H.264 JM7.3 reference codec is modified to support 2HMCP. Several QCIF (176x144) standard test sequences with a frame rate of 30 frames/s are used for simulation. In a GOP, the first frame is coded in intra-mode (I-frame) and the second and third frames are coded in inter-mode (P-frames). The rest frames in the sequence are then all coded in the 2HMCP mode (M-frames). Each frame is encoded with various fixed quantization parameters, ranging from 12 to 36. Since the error resilience performance is the central focus of this work, for simplicity but without loss of generality, all MBs are coded in fixed 8x8 block size with integer-pel precision ME. In the experiments, each frame is assumed to be encapsulated into a single packet. For channel condition simulations, a two-state Markov model, which adopted a simplified Gilbert channel, is used to generate packet-loss patterns with a packet loss rate (PLR) of 5% and a mean burst-error length of two packets. Ten different patterns under this configuration are generated for simulations.

The random MB intra-refresh, which is widely used for error resilience coding, is compared with the proposed method. The refresh rate is set to 10% and 20%. The proposed method is also integrated with random intra-refresh with a lower intra-refresh rate of 5% and 10% to further improve error resiliency. Fig. 5 shows the rate-PSNR curves for the *Foreman* and *Carphone* sequences, respectively, where the PSNR is the average PSNR value among the 300 frames for 10 error patterns (PLR = 5%). The results show that the proposed 2HMCP with 5% IR outperforms the Inter-MCP with 10% IR by about 1 dB on average. It can even beat the Inter-MCP with 20% IR for low to medium bit-rates. Moreover, the coding efficiency of the proposed 2HMCP with 5% IR is significantly better than the Inter MCP with 10% and 20% IR. Therefore, the proposed 2HMCP with a lower-rate IR can significantly outperform Inter MCP with the same or higher rate IR for most types of sequences in terms of both coding efficiency for clean channels and error resilience for lossy channels. Fig. 6 illustrates the frame-by-frame performance comparisons for the *Forman*

sequence with error pattern #7 and QP = 24. The proposed method evidently yields a significant improvement on error resilience over the original 2HMCP coder. The improvement ranges from 3 dB to 5 dB on average for the same bitrate and intra-refresh rate. From Fig. 6, we can observe the differences between the characteristics of intra-refresh and 2HMCP in terms of error resilience and coding efficiency. The 2HMCP reacts to error quickly but is usually not able to stop the error propagation entirely; whereas the intra-refresh scheme reacts to channel more slowly but can recover from error propagation almost entirely in the long run. Combining the proposed method with IR can do a better job in overcoming the drawbacks of the 2HMCP and further enhancing the error robustness of the 2HMCP coder.
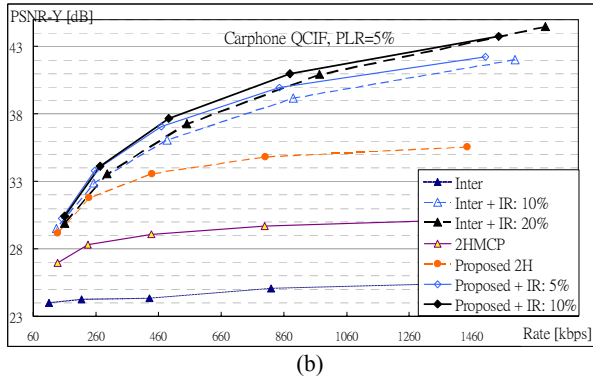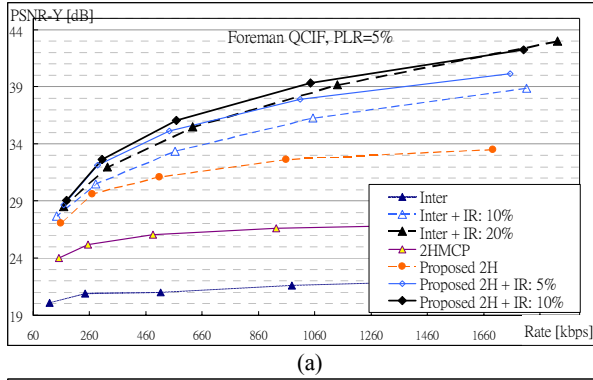

(a)

(b)

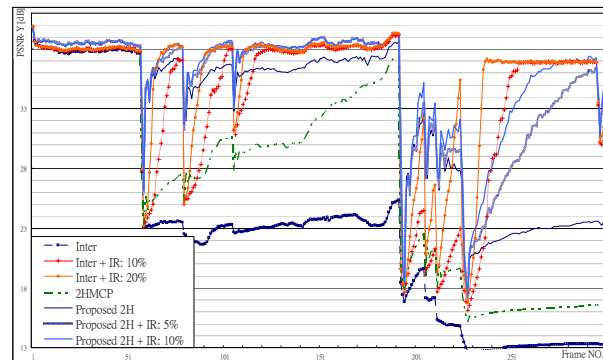Fig. 5. Average rate-PSNR performance for (a) *Foreman*, (b) *Carphone* (PLR=5%, 10 error patterns).

Fig. 6. Frame by frame PSNR comparison of Inter-MCP, 2HMCP, and proposed 2HMCP with three IR rates for *Foreman* (PLR = 5%, QP = 24, pattern #7)

5. CONCLUSION

MHMCP which has proven to be more compression efficient than the traditional motion-compensated prediction has also been addressed its error resilience against channel errors. We proposed efficient schemes for enhancing the error resilience capability of 2HMCP, especially for burst error conditions. Experimental simulation results show that the proposed methods can effectively and quickly mitigate the error propagation and the penalty on coding efficiency for clean channels due to the inserted error resilience features is rather minor. We have shown that combining the proposed 2HMCP with random MB intra-refresh can achieve faster error recovery speed. The proposed 2HMCP with a lower-rate intra-refresh can achieve better or comparable error resilience compared to Inter MCP with a much higher-rate intra-refresh, while the coding performance of our method would be significantly higher for clean channels.

REFERENCES

[1] G. J. Sullivan, "Multi-hypothesis motion compensation for low bit-rate video coding," in *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing,* vol. 5, pp. 437-440, Apr. 1993.

[2] B. Girod, "Efficiency analysis of multihypothesis motion-compensated prediction for video coding," *IEEE Trans. Image Processing,* vol. 9, no. 2, pp.173-183, Feb. 2000.

[3] M. Flierl, T. Wiegand, and B. Girod, "A locally optimal design algorithm for block-based multi-hypothesis motion-compensated prediction," in *Proc. Data Compression Conf.*, pp. 239-248, Apr. 1998.

[4] M. Flierl, T. Wiegand, and B. Girod, "Rate-constrained multihypothesis prediction for motion-compensated video compression," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 12, no. 11, pp. 957-969, Nov. 2002.

[5] M. Flierl, and B. Girod, "Generalized B pictures and the draft H.264/AVC video-compression standard," *IEEE Trans. Circuits Syst. Video Technol.,* vol. 13, no. 7, pp. 587-597, July 2003.

[6] S. Lin and Y. Wang, "Error resilience property of multihypothesis motion-compensated prediction," in *Proc. IEEE ICIP,* vol. 3, pp. 545-548, Sept. 2002, Rochester, NY.

[7] W.-Y. Kung, C.-S. Kim and C.-C. J. Kuo, "Analysis of multi-hypothesis motion compensated prediction for robust video transmission," in *Proc. IEEE ISCAS,* vol. 3, pp. 761-764, May 2004, Vancouver, Canada.

[8] W.-Y. Kung, C.-S. Kim and C.-C. J. Kuo, "Error resilience analysis of multi-hypothesis motion compensated prediction for video coding with multi-hypothesis motion compensated prediction," in *Proc. IEEE ICIP,* pp. 821-824, Oct. 2004, Singapore.

[9] S. Gnavi. M. Grangetto, E. Magli, and G. Olmo, "Comparison of rate allocation strategies for H.264 video transmission over wireless lossy correlated networks," in *Proc. IEEE ICME,* vol. 2, pp. 517-520, July 2003, Baltimore, Maryland.