ERROR RESILIENCE ANALYSIS OF MULTI-HYPOTHESIS MOTION COMPENSATED PREDICTION FOR VIDEO CODING

Wei-Ying Kung,† *Chang-Su Kim*‡ *and C.-C. Jay Kuo*†

† Integrated Media Systems Center and Department of Electrical Engineering
University of Southern California, Los Angeles, CA 90089-2564
‡ Department of Information Engineering, the Chinese University of Hong Kong

ABSTRACT

The relationship between the error propagation effect and the hypothesis coef£cients of the multi-hypothesis motion compensated prediction (MHMCP) is analyzed in this work. MHMCP enhances the error resilience of compressed video but demands a higher bit rate for the motion information. We study the rate-distortion performance of MHMCP in an error prone environment and then perform extensive experiments to con£rm the analytical results. Several design guidelines for MHMCP coders are drawn based on the analytical and experimental results.

1. INTRODUCTION

In an MHMCP coder, a macroblock (MB) can contain more than one motion vectors (MVs). Each MV speci£es a reference MB, which is called a hypothesis. Then, the target MB is estimated by linearly combining all hypotheses. By carefully selecting the hypotheses and their weighting coef£cients, a coding gain can be achieved. The coding ef£ciency of MHMCP was analyzed in [1].

Although MHMCP was originally proposed to improve coding ef£ciency, it has been observed recently that it can also enhance the error resilient property of compressed video. One application is to adopt the multi-hypothesis (MH) technique for error concealment at the decoder. In [2], we developed a multi-hypothesis error concealment method with adaptive weighting coef£cients to lower concealment errors as well as suppress propagating errors. An iterative MH error concealment algorithm for the latest standard H.26L was proposed in [3]. It is worthwhile to point out that the above methods work well even without MHMCP in the encoder. Actually, they did not address the error resilience of MHMCP at the encoder side. In [4], we discussed the general relationship between the error propagation effect and the hypothesis number. It was shown that MHMCP performs ef£ciently in the rate-distortion (R-D) sense, only when the hypothesis number is small. In this work, we investigate the impact of hypothesis coef£cients on the R-D performance while £xing the hypothesis number to 3.

The rest of this paper is organized as follows. The problem is formulated in Section 2. Then, we discuss the impact of the hypothesis coef£cients on the propagation error and the bit rate in Sections 3 and 4, respectively. The overall R-D tradeoff is discussed in Section 5. The performance of MHMCP is compared with that of the random intra-MB refreshing method in Section 6.

2. PROBLEM FORMULATION

Let us consider the effect of transmission errors in MHMCP. Suppose that a frame is corrupted during the transmission. The l-th frame after the erroneous one contains the propagation error, which can be written as

$$\varepsilon_l = \sum_{i=1}^{n} h_i \varepsilon_{l-i}, \tag{1}$$

where n denotes the hypothesis number and h_i denotes the hypothesis coef£cient corresponding to the previous i-th reference frame. This relation can be repeatedly applied to measure the error propagation effect quantitatively. It is obvious that the hypothesis number n and the hypothesis coef£cients h_i's, $i = 1, \ldots, n$, determine the amount of the propagating error. In this work, we focus on the effect of the hypothesis coef£cients h_1, h_2 and h_3, when $n = 3$. The sum of the hypothesis coef£cients is equal to 1 (*i.e.* $\sum_{i=1}^{3} h_i = 1$) so that no bias error occurs in ε_l.

3. IMPACT ON PROPAGATION ERROR

Let ε_0 denote the initial error in the 0th frame. Then, from (1), we can derive its error propagation to the following frames, given by

$$\varepsilon_l = \frac{1 + (1 - h_1 + h_3)\sum_{n=0}^{l} \alpha^{l-n}\beta^n + h_3 \sum_{n=0}^{l-1} \alpha^{l-1-n}\beta^n}{2 - h_1 + h_3} \varepsilon_0$$

where $\Delta = \sqrt{(1-h_1)^2 - 4h_3}$, $\alpha = \frac{(h_1-1)+\Delta}{2}$ and $\beta = \frac{(h_1-1)-\Delta}{2}$.

Transmission errors tend to attenuate as they propagate as a result of low pass £ltering operations in the prediction loop such as half-pixel motion compensation. By adopting the loop £lter model in [5], we can approximate the mean square error (MSE) of the l-th frame after the corrupted frame by

$$\text{MSE}_l = \frac{{\varepsilon_l}^2}{1+\gamma l}, \qquad (2)$$

where γ is a parameter that describes the effectiveness of the loop £lter to attenuate errors.

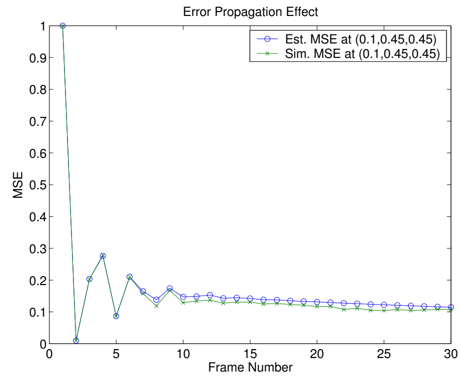
The H.264 reference codec of version JM6.1e has been modi£ed to provide the MHMCP functionality. We have tested the effect of a burst error with several combinations of hypothesis coef£cients and con£rmed that the estimated MSEs match simulated MSEs well in every combination. Fig. 1(a) demonstrates one example of error propagation effect when (h_1, h_2, h_3) = (0.1, 0.45, 0.45). Note that the propagation error oscillates with a decreasing amplitude and £nally converges to a non-zero value.

To further investigate the relationship between MSE and the hypothesis coef£cients, Fig. 1(b) depicts the average MSEs after a burst error, where the averaging range is 100 frames for the "Foreman" test sequence. The average MSEs are normalized with respect to the initial MSE ($= \varepsilon_0^2$) of the concealed frame. The two horizontal axes in Fig. 1(b) denote h_1 and h_2, respectively. Note that h_3 is given by $1 - h_1 - h_2$. The "stars" represent the estimated MSE values, while the lines represent the simulated MSE values. The largest MSE is observed when $(h_1, h_2, h_3) = (1, 0, 0)$, which is the conventional single-hypothesis motion compensated prediction (SHMCP). Also, the other two extreme cases, $(h_1, h_2, h_3) = (0, 1, 0)$ and $(0, 0, 1)$, provide relatively large MSEs. This indicates that MHMCP alleviates the propagation errors by mixing more than one prediction blocks. The minimum MSE is achieved around the point $(h_1, h_2, h_3) = (0, 0.18, 0.82)$. However, the hypothesis coef£cients also affect the prediction performance and, hence, the bit rate. The minimum MSE point results in a high bit rate due to less ef£cient motion prediction with zero value of h_1. The effects of hypothesis coef£cients on the bit rate are discussed in the next section.

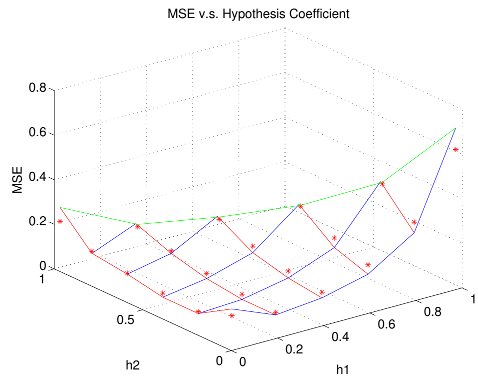
4. IMPACT ON BIT RATE

Assume that the prediction error of the i-th hypothesis is given by a Gaussian random variable with variance σ_i^2, $i = 1, 2, 3$. Then, the overall prediction error of MHMCP can be expressed as [4]

$$\sigma_{MH}^2 = \sum_{i=1}^{3}(h_i^2\sigma_i^2 + h_i \sum_{j=1, j\neq i}^{3} \rho_{i,j} h_j \sigma_i \sigma_j), \qquad (3)$$



(a) Comparison of the estimated MSE and the simulated MSE



(b) Normalized MSE vs. hypothesis coef£cients

Fig. 1. The impact of different hypothesis coef£cients on the propagation error.

where $\rho_{i,j}$ is a correlation coef£cient between the i-th and the j-th hypotheses. We can simplify the problem by assuming that $\rho_{i,j} = \rho$ for all i and j. Furthermore, we found experimentally from training sequences that $\sigma_2^2 = 1.18\sigma_1^2$ and $\sigma_3^2 = 1.385\sigma_1^2$.

Then, the overall change of the bit rate, as compared to the conventional SHMCP, can be expressed as the sum of the reduced bit rate due to the smaller prediction error and the increased bit rate due to two additional motion vectors [4], *i.e.*

$$\triangle R = \frac{1}{2}\log_2 \frac{\sigma_{MH}^2}{\sigma_1^2} + \frac{2c}{256}, \qquad (4)$$

where c is the average number of bits to represent an additional motion vector as well as the corresponding reference frame number. Parameter c depends on the coding method of the motion information. In this work, we predictively encode the motion vectors to exploit the correlation between motion vectors. First, the motion vector of each hypothesis is predicted from that of the corresponding hypothesis for the previous MB. Then, the predicted motion vectors of hypotheses 2 and 3 are further predicted from that of hypothesis 1. In our simulation, c is about 13.

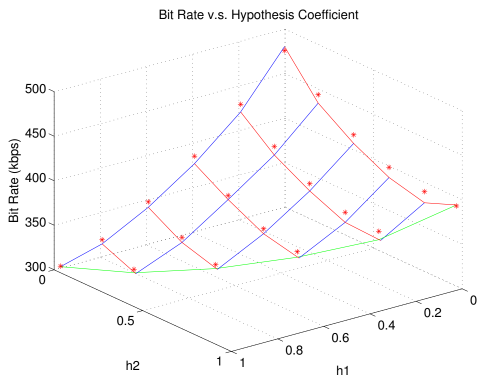


Fig. 2. The bit rate (bits per pixel) as a function of the hypothesis coef£cients.

Fig. 2 compares the simulated bit rates with the estimated bit rates. The stars represent the estimated bit rates while the lines represent the simulated bit rates. We can see that the formula in (4) approximates the experimental data well.

5. RATE-DISTORTION ANALYSIS

Using the MSE and bit rate models, we can plot the estimated R-D points in Fig. 3. Each point corresponds to a combination of (h_1, h_2, h_3). The three spikes are observed in the three extreme cases: $(h_1, h_2, h_3) = (1, 0, 0)$, and $(0, 1, 0)$ and $(0, 0, 1)$. We can see that the choice of hypothesis coef£cients affects the R-D performance considerably. From these data, it is found that the optimum R-D points on the convex hull have similar h_2 values, which are concentrated around $1/3$. The red curve corresponds to the combinations, in which h_2 is £xed to $1/3$. This indicates that for a video codec with triple-hypothesis motion compensation, we can £x h_2 to $1/3$ and change only h_1 and h_3 to meet the overall rate or distortion requirement. The adaptation of the hypothesis coef£cients can be made at the sequence or the frame level. However, due to the variation of video content and channel conditions, the performance of MHMCP can be further improved if the hypothesis coef£cients are adapted at the MB level.

We have modi£ed the H.264 codec to include the MH option and the MB level adaptation. The Lagrangian method is employed for greedy R-D optimization. Speci£cally, given the previously encoded MBs, we minimize the Lagrangian cost of a current MB, given by

$$J = D_s + D_c + \lambda(R + \triangle R), \tag{5}$$

where $\lambda = 0.85Q^2$ [6], D_s is the source distortion caused by the quantization of prediction residuals, D_c is the channel distortion caused by the concealment of erroneous MBs

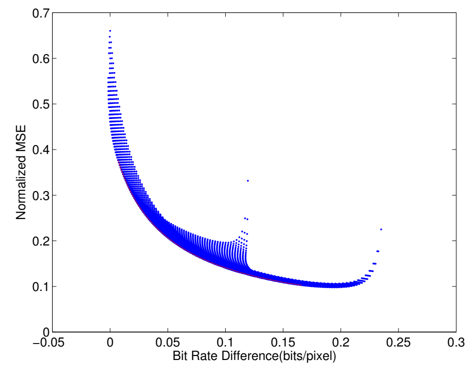


Fig. 3. The rate-distortion performance after a burst error.

and the accumulated propagation from previous frames, R is the bit rate for the original SH, and $\triangle R$ is the increased bit rate due to the MH option.

Let p_i denote the MB loss rate of the previous i-th MB. Then, D_c can be expressed as

$$D_c = \sum_{i=1}^{3} p_i \mathrm{MSE}_i,$$

where MSE_i denotes the mean square error of the i-th MB. Also, $\triangle R$ is given in (4). Thus, we can easily obtain the R-D pair for MHMCP, if we have the R-D pair for SHMCP that has been extensively studied.

When encoding each MB, the cost given by (5) is evaluated for various combinations of hypothesis coef£cients, and the optimal combination yielding the lowest cost is selected, where h_i varies from 0 to 1 with a step 0.2. Then, the index for the optimum combination is encoded as the header information for the MB.

Fig. 4 shows the PSNR performance of the proposed MB level adaptation scheme when 10% of video packets are randomly lost. We can see that the adaptive scheme provides up to 2.5 dB better performances than the £xed scheme with $(h_1, h_2, h_3) = (0.33, 0.33, 0.34)$.

6. COMPARISON WITH RANDOM INTRA-MB REFRESHING

Random intra-MB refreshing (RIR) [7] is a well known error resilient tool used in the video encoder to stop error propagation. In this section, we compare its error resilient capability with that of MHMCP.

In RIR, the amount of propagation error can be controlled by varying the refreshing rate $0 < p < 1$, which means the p percent of MBs in a frame should be intra-coded. Unlike ε_l in MHMCP, which converges to a nonzero value, ε_l in RIR converges to 0. Therefore, in the long run, RIR is better than MHMCP since RIR can stop the error

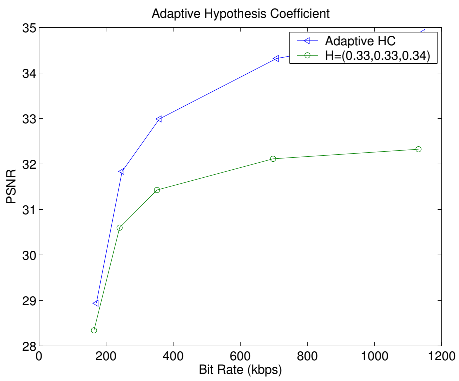
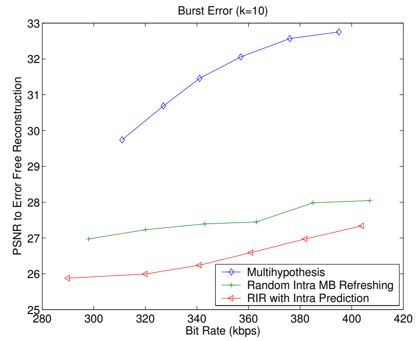
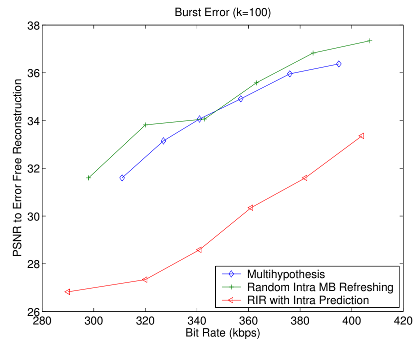


Fig. 4. The performance of the proposed adaptive coef£cient scheme compared with a £xed coef£cient scheme.



(a) Short term error suppression effect



(b) Long term error suppression effect

Fig. 5. Error suppression comparison between MHMCP (with various h_i values) and RIR after a burst error.

entirely. However, MHMCP suppresses error propagation more rapidly.

In Fig. 5, we compare MHMCP and RIR in terms of the PSNR of the reconstructed video and the bit rates. For MHMCP, we £x h_2 to 0.33 and change h_1 from 0.1 to 0.6. For comparison, random intra refreshing algorithm is also simulated. The refreshing rates are selected so that it generates the similar bit rate as MHMCP. To achieve fair comparison, the encoding parameters such as quantization parameters are set to be the same in both cases. The H.264 standard enforces the intra prediction, which reduces bit rates but deteriorates the error resilient capability. Both RIR with and without intra prediction are tested and depicted as red and green curves in Fig. 5.

The PSNR values in Fig. 5(a) are averaged in the short term (*i.e.* only the £rst 10 frames after an erroneous one) while the PSNR values in Fig. 5(b) are averaged in the long term (consisting of 100 frames in the average after an erroneous frame). MHMCP signi£cantly outperforms both RIR algorithms in the short term. In the long term, MHMCP still performs better than RIR with intra prediction, but achieves lower PSNR values than RIR without intra prediction. Therefore, MHMCP could be considered as a good error resilient tool in the H.264 codec. Also, even for H.263 and MPEG4 codecs without the enforced intra prediction, MHMCP may have its advantage provided that I-frames are regularly inserted.

7. CONCLUSION

In this work, we investigated the effect of error propagation in MHMCP and analyzed the rate-distortion performance in an error-prone environment in terms of hypothesis coef£cients (h_1, h_2, h_3). It was observed that the case $h_2 = 1/3$ provides the optimum R-D performance for the 3-HMCP codec. Also, it was shown that the performance of the proposed algorithm can be improved further, if the coef£cients are adapted at the MB level. Simulation results demonstrated that the proposed algorithm can suppress the short-term error propagation more effectively than the random intra refreshing method.

8. REFERENCES

[1] B. Girod, "Ef£ciency analysis of multihypothesis motion-compensated prediction for video coding," *IEEE Trans. Image Processing*, vol. 9, no. 2, Feb. 2000.

[2] W.-Y. Kung, C.-S. Kim, and C.-C. Jay Kuo, "A dynamic error concealment for video transmission over noisy channels," in *IEEE GlobeCom 2002*, Nov. 2002.

[3] Y. O. Park, C.-S. Kim, and S.-U. Lee, "Multi-hypothesis error concealment algorithm for H.26l video," in *ICIP 2003*, Sept. 2003.

[4] W.-Y. Kung, C.-S. Kim, and C.-C. Kuo, "Analysis of multi-hypothesis motion compensated prediction for robust video transmission," in *IEEE ISCAS2004*, May 2004.

[5] B. Girod and N. Farber, "Wireless video," in *Compressed video over networks*, M.-T. Sun, Ed. Marcel Dekker, 2000.

[6] G. J. Sullivan and T. Wiegand, "Rate-distortion optimization for video compression," *IEEE Signal Processing Magazine*, vol. 15, no. 6, pp. 74–90, nov 1998.

[7] *Video Coding for Low Bitrate Communication*, 1998, ITU-T Recommendation H.263.