Analysis of Multihypothesis Motion Compensated Prediction (MHMCP) for Robust Visual Communication

Wei-Ying Kung, Chang-Su Kim, *Senior Member, IEEE*, and C.-C. Jay Kuo, *Fellow, IEEE*

Abstract—A multihypothesis motion compensated prediction (MHMCP) scheme, which predicts a block from a weighted superposition of more than one reference blocks, is proposed and analyzed for error resilient visual communication in this research. By combining these reference blocks effectively, MHMCP can enhance the error resilient capability of compressed video as well as achieve a coding gain. In particular, we investigate the error propagation effect in the MHMCP coder and analyze the rate-distortion performance in terms of the hypothesis number and hypothesis coefficients. It is shown that MHMCP suppresses the short-term effect of error propagation more effectively than the intra-refreshing scheme.

Index Terms—Error propagation, error resilient video coding, multihypothesis motion compensated prediction (MHMCP).

I. INTRODUCTION

MODERN video coding standards, such as H.263, H.264, MPEG-2 and MPEG-4, are all based on the hybrid scheme of motion compensated prediction and transform coding. Each frame is partitioned into macroblocks (MB). Then, a motion estimator matches each MB to an MB in the reference frame, which has been previously encoded, to yield the smallest distance measure. Afterwards, the residual MB, obtained by subtracting the matched MB from the original MB, is transformed via the discrete cosine transform (DCT) for the purpose of energy compaction. Finally, quantization is applied to DCT coefficients followed by entropy coding to generate the desired bit stream.

Apparently, motion estimation (ME) plays an important role in the hybrid video coding. The basic idea of ME is to find the best MB in the reference frame to minimize the prediction error, thus, saving the bit rate. Several advanced ME techniques [1] have been proposed to improve the performance. One approach is the long-term memory motion compensation (LMMC) scheme [2], which enlarges the buffer size to store multiple reference frames. LMMC was originally proposed to stop error propagation [3], and has been adopted by H.263 [4]

W.-Y. Kung is with Motorola, San Diego, CA 92121 USA (e-mail: vivian.kung@motorola.com).

C.-S. Kim is with Department of Electronics Engineering, Korea University, Seoul, Korea (e-mail: cskim@ieee.org).

C.-C. J. Kuo is with Department of Electrical Engineering and Integrated Media Systems Center, University of Southern California, Los Angeles, CA 90089-2564 USA (e-mail: cckuo@sipi.usc.edu).
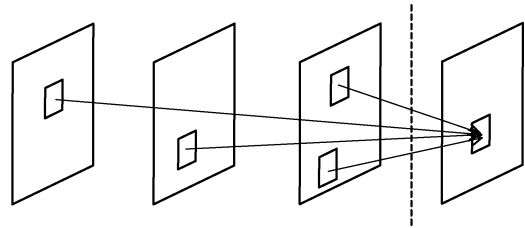
Fig. 1.　Illustration of the MHMCP scheme.

and MPEG-4 [5] as an optional mode. It was observed in recent research [1] that LMMC also achieves bit rate saving for cases such as scene cuts, uncovered background and texture with aliasing. Therefore, LMMC has been adopted by the new video coding standard H.264 [6] for high compression efficiency.

The concept of LMMC can be furthermore extended to obtain the multihypothesis motion compensated prediction (MHMCP) scheme [7], [8]. MHMCP predicts a block from a weighted superposition of multiple reference blocks stored in the frame buffer. Unlike LMMC, which generates only one motion vector (MV) for each MB, MHMCP uses more than one MVs (see Fig. 1). The current MB is estimated by linearly combining these hypotheses, each of which is specified by one MV. Sullivan [7] first proposed this idea in 1993 and claimed that the coding bit rate can be reduced by MHMCP. Girod [8] developed a theory on how MHMCP achieves high coding efficiency. Recently, MHMCP has been studied extensively for practical applications. An iterative hypothesis selection algorithm was proposed in [9]. Flierl *et al.* [10] showed that the hypothesis coefficients converge to $1/n$ regardless of their initial values, where n denotes the number of hypotheses. They further observed in [11] that two jointly estimated hypotheses provide a major portion of the achievable gain.

Although MHMCP was originally proposed to achieve high coding efficiency, it was also observed to enhance the error resilience of compressed video. Kim *et al.* [12] proposed the double-vector motion compensation (DMC), where each MB is predicted from two reference MBs. Even if one reference MB is corrupted, its error propagation can be effectively alleviated using the other reference MB for motion compensation. However, there is the encoder-decoder mismatch problem if only one reference MB is used in the decoder. This can be overcome by encoding mismatch signals [13], [14]. Also, it was shown that the performance of DMC can be further improved based on the minimum mean square error (MMSE) decoding in [15].

The multihypothesis (MH) concept can also be adopted for error concealment at the decoder. Al-Mualla *et al.* [16] proposed an MH error concealment scheme, which uses the weighted average of several concealed signals to replace corrupted MBs.

In [17], we developed an MH error concealment scheme that adaptively assigns weighting coefficients to lower concealment errors and suppress propagating errors. An MH error concealment method with iterative hypothesis search was proposed in [18]. It is worthwhile to point out that these three methods work well even without MHMCP in the encoder. Actually, none of them addressed the error resilience of MHMCP at the encoder side. Lin and Wang [19] discussed the error resilience property of double-hypothesis MCP. However, a thorough analysis of the error resilience property of MHMCP is still lacking up to now.

In this work, we investigate the error propagation effect in the MHMCP coder. First, the problem is formulated in Section II. In Section III, we discuss how the number of hypotheses influences error propagation and coding efficiency. In Section IV, we discuss the same issue in association with hypothesis coefficients. Concluding remarks are given in Section V.

II. ERROR PROPAGATION MODEL FOR MHMCP CODER

As shown in Fig. 1, a frame $\mathbf{f}_k$ at time index k in MHMCP can be expressed as

$$\mathbf{f}_k = \sum_{i=1}^{n} w_i \mathbf{h}_{k,i} + \mathbf{r}_k \tag{1}$$

where n is the number of hypotheses, $\mathbf{h}_{k,i}$ is the ith hypothesis image, w_i is a weighting (or hypothesis) coefficient, and $\mathbf{r}_k$ denotes the residual error after MHMCP. The sum of the hypothesis coefficients is equal to 1 (i.e., $\sum_{i=1}^{n} w_i = 1$) so that no bias error occurs in $\mathbf{r}_k$.

The hypothesis images $\mathbf{h}_{k,i}$'s are predicted blockwise from the reference frames using MVs. Ideally, the MVs should be jointly optimized to minimize the rate-distortion cost function. An iterative procedure [10] was proposed to select a locally optimal combination of MVs. In this case, the selected vectors are not necessarily the best ones for the conventional single-hypothesis motion compensated prediction (SHMCP) scheme, but achieves better performance by joint prediction.

In general, several hypotheses can be obtained from the same reference frame and overlap with one another in many cases. However, error resilience can be improved if each hypothesis image is obtained from a different reference frame. Then, even if a hypothesis is corrupted, the current frame can be reconstructed with an acceptable quality using the information in the other hypotheses as done in [12]. Therefore, we constrain the ith hypothesis image $\mathbf{h}_{k,i}$ to be predicted from $\mathbf{f}_{k-i}$ in this work. Then, we can rewrite (1) as

$$\mathbf{f}_k = \sum_{i=1}^{n} w_i P_{k,i}(\mathbf{f}_{k-1}) + \mathbf{r}_k \tag{2}$$

where $P_{k,i}$ denotes the motion compensation operator from $\mathbf{f}_{k-i}$ to $\mathbf{f}_k$, which is specified by the MVs.

Let us consider the effect of transmission errors in MHMCP. Suppose that the zeroth frame is corrupted during transmission and its error is denoted by $\mathbf{e}_0$. The error propagates to subsequent frames, and the propagation error $\mathbf{e}_k$ in the kth frame can be written as

$$\mathbf{e}_k = \sum_{i=1}^{n} w_i P_{k,i}(\mathbf{e}_{k-i}), \quad k \geq 0 \tag{2}$$

where $\mathbf{e}_k = \mathbf{0}$ for $k < 0$. As the error propagates, it tends to spread spatially due to nonzero MVs. Also, note that $\mathbf{e}_k$ is given by the weighted sum of error components $P_{k,i}(\mathbf{e}_{k-i})$'s. The weighted summation is equivalent to a lowpass filtering operation so that the error attenuates as it propagates. Generally speaking, the error $\mathbf{e}_k$ is attenuated more effectively when error components $P_{k,i}(\mathbf{e}_{k-i})$'s are less correlated.

If all MVs are zero, each motion compensation operator $P_{k,i}$ in (2) is the identity function. Then, we have

$$\begin{aligned} \mathbf{e}_k &= \sum_{i=1}^{n} w_i \mathbf{e}_{k-i} \\ &= \phi(k, n, \mathbf{w}) \mathbf{e}_0 \end{aligned} \tag{3}$$

where $\phi(k, n, \mathbf{w})$ is a scalar, called the MH attenuation factor, which depends on the number of hypotheses n and the set of hypothesis coefficients $\mathbf{w} = (w_1, w_2, \ldots, w_n)$. Let us consider an example with $n = 2$ and $\mathbf{w} = (1/2, 1/2)$. Since $\mathbf{e}_1 = (1/2)\mathbf{e}_0$, $\phi(1, n, \mathbf{w}) = 1/2$. Also, $\mathbf{e}_2 = (1/2)\mathbf{e}_1 + (1/2)\mathbf{e}_0 = (3/4)\mathbf{e}_0$ and $\phi(2, n, \mathbf{w}) = 3/4$. In general, it can be shown by solving the recursion that

$$\phi(k, n, \mathbf{w}) = \frac{2}{3} + \frac{1}{3}\left(-\frac{1}{2}\right)^k. \tag{4}$$

From (3), we see that each $\mathbf{e}_k$ is a scalar multiple of $\mathbf{e}_0$, and the correlation coefficient between two corresponding pixels in $\mathbf{e}_k$ and $\mathbf{e}_l$ is 1. Therefore, when all MVs are zero, there is no statistical attenuation due to the mixing of different random variables. In the above example in (4), as k approaches infinity, $\mathbf{e}_k$ does not vanish but converges to $(2/3)\mathbf{e}_0$.

Let $D_{k \leftarrow 0}$ denote the mean square error (MSE) of pixels in the kth frame due to the propagation of an initial error in the 0th frame. When all MVs are zero, we can obtain the following relationship from (3)

$$D_{k \leftarrow 0} = [\phi(k, n, \mathbf{w})]^2 D_{0 \leftarrow 0}, \tag{5}$$

where $D_{0 \leftarrow 0}$ is the mean square value of the initial pixel errors in the 0th frame. However, MVs are not fixed to zero in practice, and the initial error experiences a stronger attenuation. Even if MHMCP is not employed, the error is attenuated by the spatial filtering process in a typical video coder, such as sub-pixel accuracy motion compensation, overlapped block motion compensation and deblocking filtering. In [20], the attenuation due to the spatial filtering was analyzed and approximated by a decay factor $1/(1+\gamma k)$, where γ is a parameter describing the strength of the spatial filter. By incorporating the decay factor into (5), we obtain

$$D_{k \leftarrow 0} = \frac{[\phi(k, n, \mathbf{w})]^2}{1 + \gamma k} D_{0 \leftarrow 0}. \tag{6}$$

This is our model for error attenuation, which results from both MHMCP and spatial filtering. In this work, γ is trained from test sequences and set to 0.012.

Let us compute the MH attenuation factor $\phi(k, n, \mathbf{w})$ and discuss its effect on the error resilient property of the MHMCP coder in the following two sections.
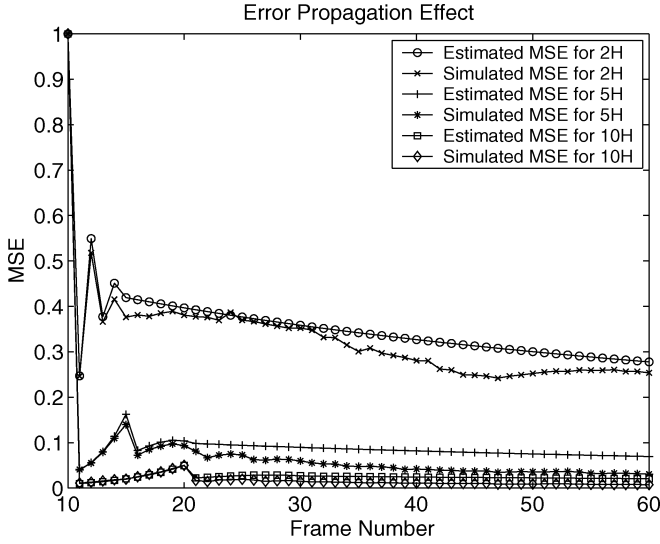
Fig. 2. Comparison of theoretical and experimental NMSE values after a burst error occurs in the tenth frame of the "Foreman" sequence.

III. EFFECT OF HYPOTHESIS NUMBER

In an error-prone environment, it is advantageous to employ MHMCP rather than SHMCP to reduce the error propagation. However, MHMCP needs more bits to represent the additional motion information. In this section, we discuss how the hypothesis number affects error resilience and coding efficiency.

A. Impact on Propagation Error

Let us analyze the relationship between the hypothesis number n and the propagation error. It is assumed that hypothesis coefficients are all the same. In other words, $w_i = 1/n$ for $1 \le i \le n$. Then, the solution to (3) can be obtained as

$$
\phi(k, n, \mathbf{w}) = \begin{cases} \frac{1}{n}\left(1 + \frac{1}{n}\right)^{k-1}, & 1 \le k \le n \\ \frac{1}{n}\left[\left(1 + \frac{1}{n}\right)^{k-1} \\ \quad -\frac{k}{n}\left(1 + \frac{1}{n}\right)^{k-n-2}\right], & n < k \le 2n \\ \vdots \\ \frac{2}{n+1}, & k \to \infty. \end{cases}
\tag{7}
$$

When $k = 1$, $\phi(k, n, \mathbf{w})$ achieves the lowest value $1/n$. In other words, the propagation error drops significantly right after the corrupted frame containing an initial error. But, the propagation error gently increases in the subsequent frames until the nth frame, in which the attenuation factor achieves the highest value $1/n(1 + (1/n))^{n-1}$. Then, the attenuation factor drops to the second lowest value $\phi(n + 1, n, \mathbf{w}) = 1/n[(1 + (1/n))^n - 1]$ and increases slowly until the $(2n)$th frame, and so on. The oscillating phenomenon becomes smoother and smoother, and the attenuation factor eventually converges to a nonzero value $2/(n + 1)$. This indicates that the propagation error cannot be totally eliminated by MHMCP alone, and the propagation error is more effectively suppressed when more hypotheses are used.

The H.264 reference codec of version JM6.1e has been modified to provide the MHMCP functionality. The "Foreman" QCIF (176×144) sequence at 25 frames/s is used in this test. A burst error causes three consecutive groups of blocks (GOBs)
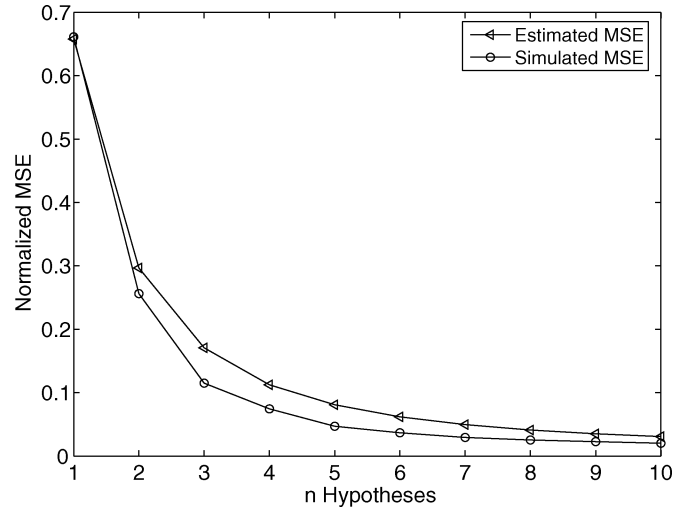

Fig. 3. Averaged NMSE as a function of the hypothesis number for the "Foreman" sequence.

lost in the 10th frame, and no more loss occurs in other frames. Fig. 2 shows the theoretical and experimental NMSE (Normalized MSE) values of each frame for three hypothesis numbers $n = 2$, 5 and 10. The theoretical NMSE is computed with the factor $[\phi(k, n, \mathbf{w})]^2/(1 + \gamma k)$ in (6) while the experimental NMSE is computed as the ratio of the actual MSE and the initial MSE of the 10th frame. We see that the theoretical and experimental NMSEs have a good match for the first 15 frames following the frame containing the error burst. After that, especially for the case $n = 5$, the simulated MSEs tend to attenuate more quickly than the estimated MSEs.

Under the same test condition, we also plot the NMSE performances with respect to the hypothesis number n in Fig. 3. Each NMSE value is averaged over 100 frames (from the 10th to 109th). As shown in the figure, when we employ a larger number of hypotheses, the averaged propagation error becomes smaller. However, a large number of hypotheses uses more bits to send additional MVs. The relationship between the hypothesis number and the bit rate will be studied in the next section.

B. Impact on Bit Rates

The hypothesis number n is related to the bit rate in two factors. On one hand, as more hypotheses are used, MHMCP yields smaller prediction errors, thus saving bits to encode residual DCT coefficients. On the other hand, more hypotheses demand more bits to encode a larger number of MVs. Thus, a balance has to be found between these two factors.

From (1), we see that MHMCP obtains the prediction $\hat{p}$ of each pixel p in the kth frame from n hypothesis pixels p_i, $1 \le i \le n$, via

$$
\hat{p} = \sum_{i=1}^{n} w_i p_i.
$$

Thus, the variance of the prediction error can be written as

$$
\sigma_{\text{MH}}^2 = E\left\{(p - \hat{p})^2\right\} \\
= \sum_{i=1}^{n}\left(w_i^2 \sigma_i^2 + w_i \sum_{j=i, j \ne i}^{n} w_j \rho_{i,j} \sigma_i \sigma_j\right)
\tag{8}
$$

where $\sigma_i^2 = E\{(p - p_i)^2\}$ and $\rho_{i,j}$ is the correlation coefficient between $(p - p_i)$ and $(p - p_j)$.

It is difficult to derive a closed-form solution to (8) in general. However, a simplification can be made by assuming that $\sigma_i^2 = \sigma_{\mathrm{SH}}^2$ and $\rho_{i,j} = \rho$ for all i and j. In other words, we assume that the variance of each individual prediction error is equal to the prediction error variance σ_{SH}^2 of SHMCP, and the correlation coefficients are fixed to ρ. Then, we have the minimum variance of the joint prediction error as

$$\sigma_{\mathrm{MH}}^2 = \frac{(1 + \rho(n - 1))\,\sigma_{\mathrm{SH}}^2}{n} \qquad (9)$$

when

$$w_i = \frac{1}{n}.$$

In this case, σ_{MH}^2 is bounded below by $\rho\sigma_{\mathrm{SH}}^2$ so that σ_{MH}^2 does not vanish even if we use as many hypotheses as possible. Also, note that hypotheses with lower correlations yield a smaller joint error variance.

To facilitate the following analysis, we assume that the prediction error is Gaussian-distributed. From the rate-distortion function for a Gaussian random variable [21], it can be shown that, as compared to SHMCP with prediction error variance σ_{SH}^2, the bit rate change for residual DCT coding in MHMCP can be written as

$$\triangle R_{\mathrm{residual}} = \frac{1}{2} \log_2 \frac{\sigma_{\mathrm{MH}}^2}{\sigma_{\mathrm{SH}}^2} \quad \text{(bits/pixel)}.$$

Further, we use (9) to compute $\triangle R_{\mathrm{residual}}$. The overall change in the bit rate, denoted by $\triangle R$, can be expressed as the sum of the reduced bit rate for residual coding and the increased bit rate for additional MVs. Thus, we have

$$\triangle R = \frac{1}{2} \log_2 \frac{1 + \rho(n - 1)}{n} + \frac{(n - 1)c}{256} \quad \text{(bits/pixel)} \quad (10)$$

where c is the number of bits to represent an additional MV and $256 = 16^2$ is the number of pixels per MB. Parameter c depends on the coding method of the motion information. We predictively encode MVs to exploit their correlation. For the "Foreman" sequence, the mean and the variance of c are 16.2 and 23.2, respectively. In this work, c is fixed to 16.

Fig. 4 compares the averaged bit rates for the 10th–110th frames of the "Foreman" sequence, when the hypothesis number n varies from 1 to 10. It is observed that the theoretical and experimental bit rates match closely.

C. Rate-Distortion Analysis

It was shown in [11] that a small number of hypotheses n is preferred for high coding efficiency. The optimal n is dependent on the correlation coefficients between hypotheses. In typical video sequences, hypotheses are highly correlated and two hypotheses are sufficient to provide good rate-distortion performance in an error-free environment. However, a larger value of n is beneficial to propagation error suppression. Generally speaking, the hypothesis number should be selected by considering both coding efficiency and error resilience.
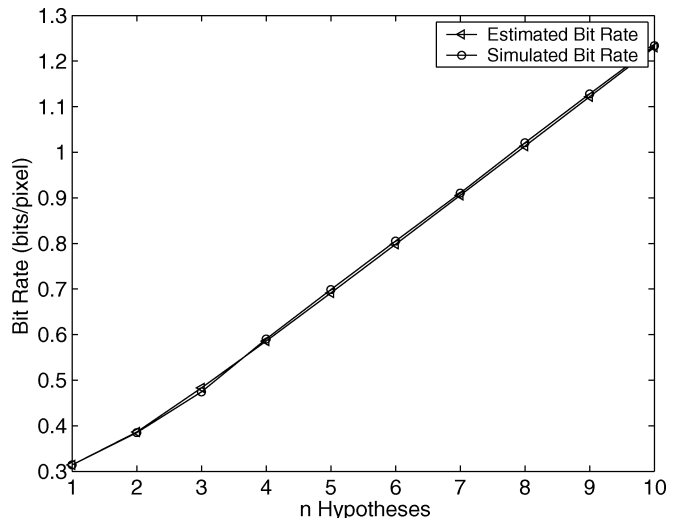

Fig. 4. Comparison of theoretical and experimental bit rates as a function of the hypothesis number for the "Foreman" sequence.
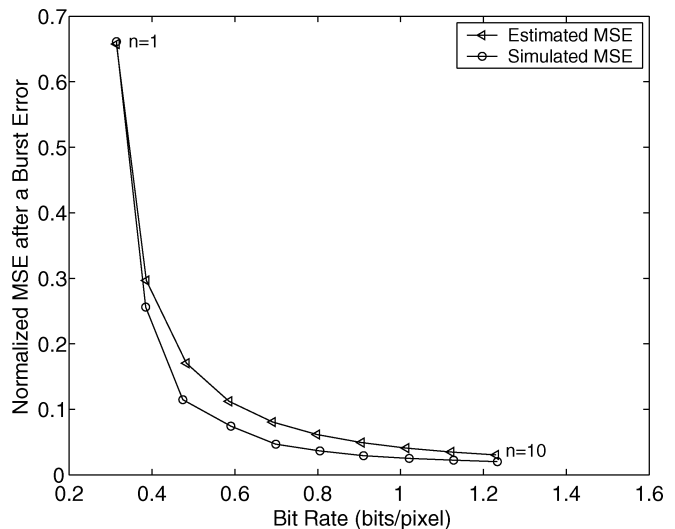

Fig. 5. Theoretical and experimental rate-distortion curves for the "Foreman" sequence caused by a burst error.

By combining the distortion data in Fig. 3 and the rate data in Fig. 4, we plot the theoretical and experimental rate-distortion (R-D) curves in Fig. 5. Each R-D point corresponds to one hypothesis number n, which takes values from 1 to 10 from the upper left position to the lower right position. Recall that these data were obtained by inserting a burst error at the 10th frame of the "Foreman" sequence. We see that the distortion decreases faster as the hypothesis number increases from 1 to 3. When $n \geq 4$, an increase in the hypothesis number gives only a marginal gain in image quality, but demands a much higher bit rate. Thus, for this error type, it is reasonable to confine the choice of consideration to be $n < 4$ or $n = 1, 2, 3$.

IV. EFFECT OF HYPOTHESIS COEFFICIENTS

In this section, we study how hypothesis coefficients w_i's are related to error propagation. As shown in the previous section, it is sufficient to use two or three hypotheses especially at low
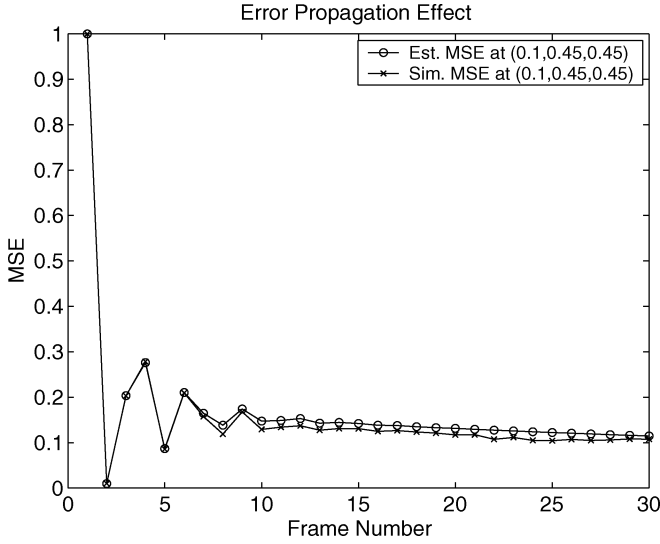
Fig. 6. Comparison of theoretical and experimental MSEs after a burst error in the triple-hypothesis MCP for the "Foreman" sequence with weighting coefficients $\mathbf{w} = (0.1, 0.45, 0.45)$.
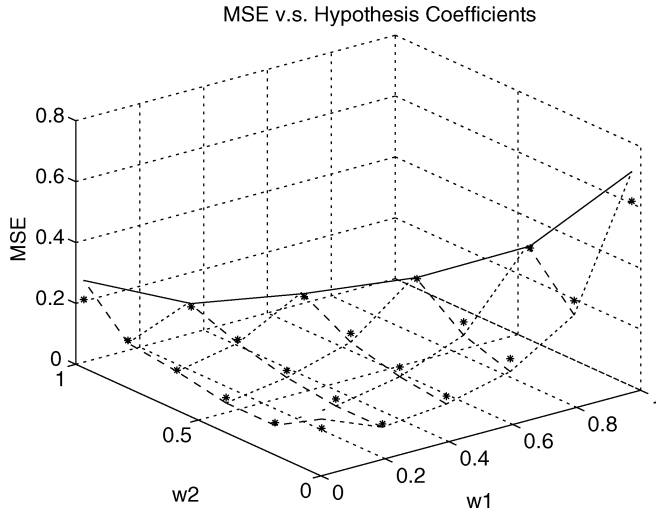

Fig. 7. Propagation error of the triple-hypothesis MCP is plotted as a function of hypothesis coefficients.
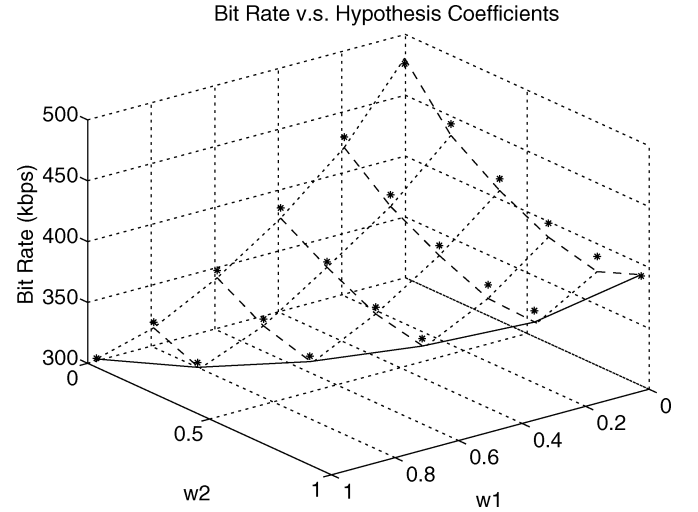

Fig. 8. Impact of hypothesis coefficients on the bit rates for the "Foreman" sequence in the triple-hypothesis MCP.

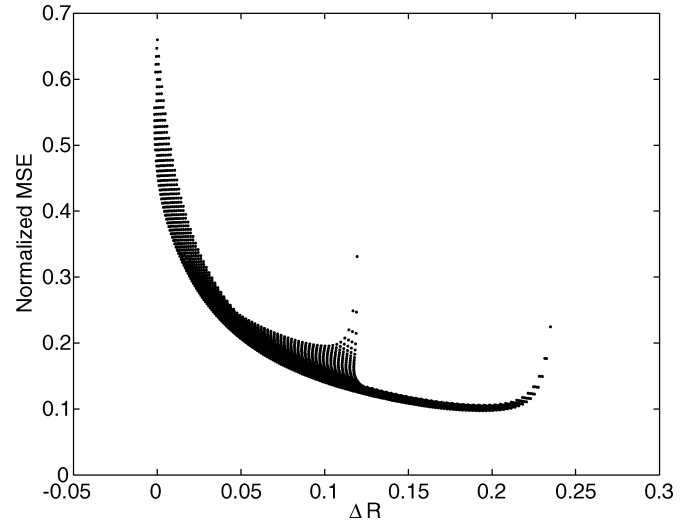Fig. 9. Rate-distortion performance of the triple-hypothesis MCP after a burst error.

bit rates. In the following analysis, we focus on the case $n = 3$. The case of $n = 2$ can be easily analyzed in a parallel fashion.

A. Impact on Propagation Error

When the number of hypotheses n is three and the hypothesis coefficients are given by the vector $\mathbf{w} = (w_1, w_2, w_3)$, the solution to (3) can be derived as

$$\phi\left(k, 3, (w_1, w_2, w_3)\right) =$$
$$\frac{1 + (1 - w_1 + w_3)\sum_{j=0}^{k} \alpha^{k-j}\beta^j + w_3 \sum_{j=0}^{k-1} \alpha^{k-1-j}\beta^j}{2 - w_1 + w_3}$$

(11)

where

$$\alpha = \frac{(w_1 - 1) + \sqrt{(1 - w_1)^2 - 4w_3}}{2}$$

$$\beta = \frac{(w_1 - 1) - \sqrt{(1 - w_1)^2 - 4w_3}}{2}.$$

By inserting (11) into (6), we can obtain the MSE of pixels in the kth frame after the corrupted frame.

We have tested the effect of a burst error with several combinations of hypothesis coefficients and confirmed that theoretical MSEs match experimental MSEs well in every combination. Fig. 6 demonstrates one example of the error propagation effect on the "Foreman" sequence, when $\mathbf{w} = (0.1, 0.45, 0.45)$. It is observed that the propagation error oscillates with a decreasing amplitude.

The average MSE is plotted as a two-dimensional surface in terms of hypothesis coefficients w_1 and w_2 in Fig. 7. Note that, since $w_1 + w_2 + w_3 = 1$, there exist only two independent weighting coefficients. The test condition is given below. A burst error is introduced in the 3rd frame of the "Foreman" sequence, and the MSEs are averaged over 100 frames (from the 3rd frame to the 102nd frame) and then normalized with respect to the initial MSE of the 3rd frame. The star points represent the theoretical MSE values, while the lines represent
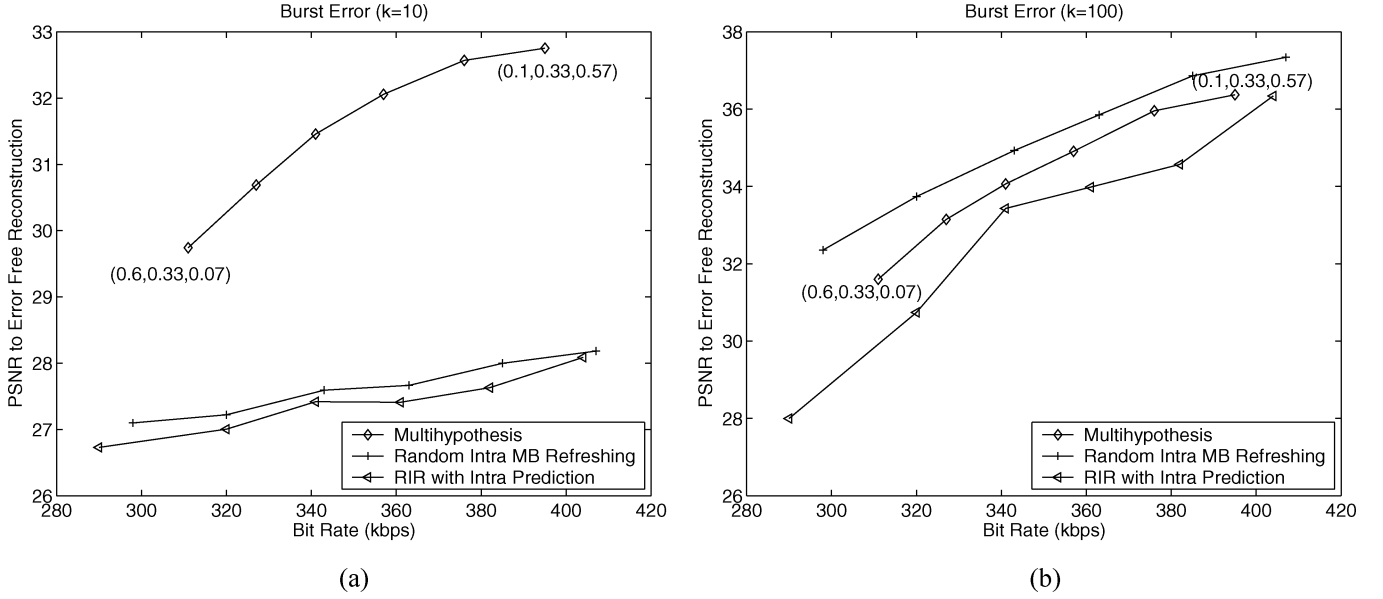
Fig. 10. Rate-PSNR performances of the triple-hypothesis MCP and IR for the "Foreman" sequence due to a burst error. (a) Next frame and (b) the average of the following 100 frames after the corrupted one.

the experimental MSE values. The largest MSE is observed when $\mathbf{w} = (1,0,0)$, which corresponds to the conventional SHMCP. Also, the other two extreme cases, $\mathbf{w} = (0,1,0)$ and $(0,0,1)$, provide relatively large MSEs. This indicates that MHMCP alleviates propagation errors by combining multiple prediction signals. The minimum MSE is achieved around $\mathbf{w} = (0, 0.18, 0.82)$. However, the hypothesis coefficients also affect the prediction performance and, hence, the bit rate. The point $\mathbf{w} = (0, 0.18, 0.82)$ results in a high bit rate due to less efficient motion prediction with zero value of w_1. The effect of hypothesis coefficients on the bit rate is discussed in the next section.

B. Impact on Bit Rates

The variance of prediction errors in the triple-hypothesis MCP can be expressed by replacing n with 3 in (8)

$$D_s = \sigma_{\mathrm{MH}}^2 = \sum_{i=1}^{3} \left(w_i^2 \sigma_i^2 + w_i \sum_{j=1, j \neq i}^{3} w_j \rho_{i,j} \sigma_i \sigma_j \right). \quad (12)$$

We found experimentally that the individual prediction error variance and the correlation coefficients can be approximated as

$$\sigma_2^2 = 1.18\sigma_1^2, \quad \sigma_3^2 = 1.385\sigma_1^2, \quad \rho_{1,2} = \rho_{1,3} = \rho_{2,3} = 0.95.$$

Similar to (10), the overall change of the bit rate is the sum of the reduced bit rate due to smaller prediction errors and the increased bit rate due to two additional MVs

$$\triangle R = \frac{1}{2} \log_2 \frac{\sigma_{MH}^2}{\sigma_1^2} + \frac{2c}{256} \quad \text{(bits/pixel)} \quad (13)$$

as compared to the conventional SHMCP with the prediction error variance σ_1^2.

We compare the experimental bit rates with the theoretical bit rates for the "Foreman" sequence in Fig. 8. The star points represent the theoretical bit rates while the lines represent the experimental ones. The highest bit rate is observed at $\mathbf{w} = (0,0,1)$,

which provides the poorest motion compensation performance. On the other hand, the lowest bit rate is achieved around $\mathbf{w} = (0.8, 0.2, 0)$, which yields smaller prediction errors by interpolating two hypotheses. This is consistent with the result in [11] that two hypotheses provide a major portion of the achievable coding gain in MHMCP.

C. Rate-Distortion Analysis

Using the MSE and the bit rate models, we plot theoretical rate-distortion (R-D) data points in Fig. 9 for the triple-hypothesis MCP after a burst error, where the x axis denotes the differential rate, which is required in addition to the SHMCP rate, and the y axis is the normalized MSE, which is averaged over 100 frames after the burst error. Each point corresponds to a certain value of $\mathbf{w} = (w_1, w_2, w_3)$. The three spikes are observed in the three extreme cases: $\mathbf{w} = (1,0,0)$, and $(0,1,0)$ and $(0,0,1)$. We see that the choice of hypothesis coefficients affects the R-D performance considerably. It is observed that the optimum R-D points on the convex hull have similar w_2 values, which are concentrated around 1/3. This indicates that, for a video codec with triple-hypothesis MCP, we can simply fix w_2 to 1/3 and change only w_1 and w_3 to meet the overall rate or distortion requirement.

D. Comparison With Random Intra-Refreshing

The random intra-refreshing (IR) scheme 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 MHMCP. Our discussion will lead to some useful guidelines for the MHMCP design.

In IR, an MB is intra-encoded with probability p_r, which is called the refreshing rate. If the zeroth frame is corrupted, an MB in the first frame is affected by the propagation error with a probability $(1 - p_r)$. In general, the probability that an MB in the kth frame is affected by the initial error in the zeroth frame can
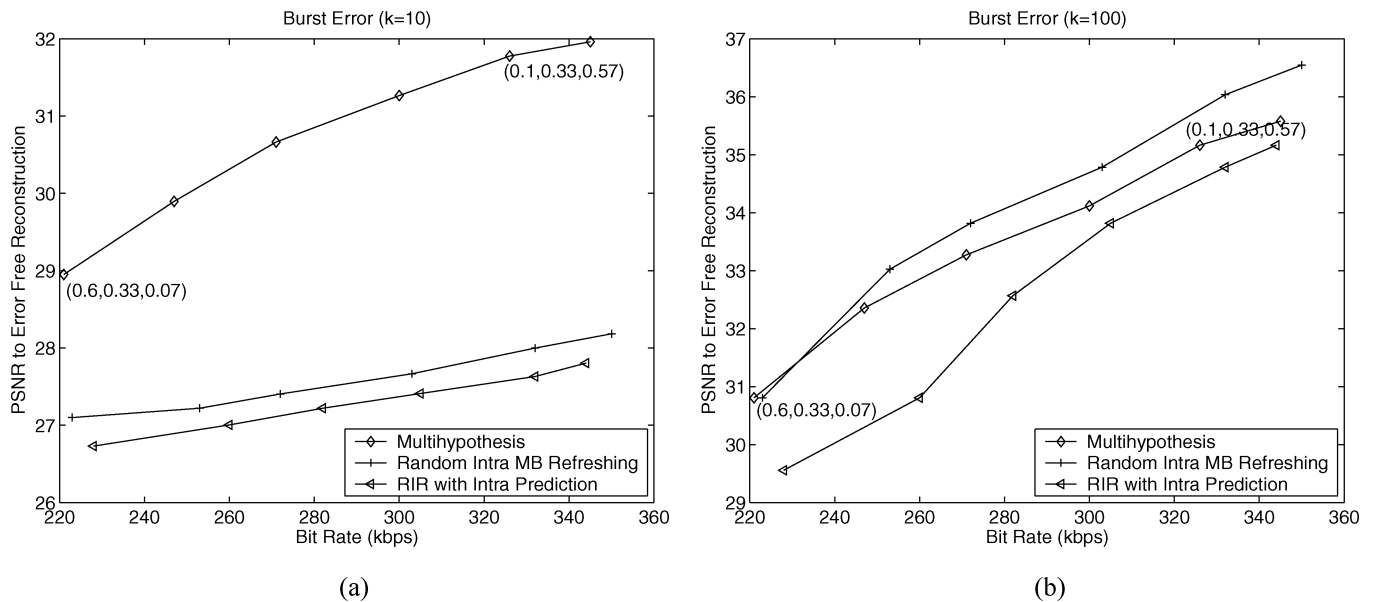
Fig. 11. Rate-PSNR performances of the triple-hypothesis MCP and IR for the "Carphone" sequence due to a burst error. (a) Next frame and (b) the average of the following 100 frames after the corrupted one.

be approximated by $(1 - p_r)^k$. Thus, as k approaches infinity, we see that IR can stop error propagation completely whereas the MHMCP attenuation factor in (11) does not vanish.

We compare the rate-PSNR performance of the triple-hypothesis MCP and IR for the "Foreman" and "Carphone" sequences in Figs. 10 and 11, respectively. For the triple-hypothesis MCP, we fix w_2 to 1/3 and vary w_1 from 0.1 to 0.6. For IR, the refreshing rates are varied to obtain a range of different bit rates. The H.264 standard adopts intra-prediction, which improves the coding gain but deteriorates the error resilience. IR with and without intra-prediction are both tested and depicted in Figs. 10 and 11.

Three consecutive GOBs are dropped at the 3rd frame. Fig. 10(a) and Fig. 11(a) show the PSNR performance of the first frame after the corrupted one, while Fig. 10(b) and Fig. 11(b) show the average PSNR performance over the 100 frames after the corrupted one. MHMCP significantly outperforms IR schemes for both test sequences in the short term. MHMCP still performs better than IR with intra-prediction in the long term, but achieves lower PSNR values than IR without intra-prediction. Therefore, MHMCP serves 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 can still be advantageous, if I-frames are periodically inserted to prevent the long-term error propagation.

V. CONCLUSION

The error propagation effect in MHMCP was examined and the relationship between the rate-distortion performance and the hypothesis number and hypothesis coefficients was thoroughly analyzed in this research. MHMCP can alleviate the effect of error propagation by combining several predictions for motion compensation, but requires a higher bit rate to represent the additional motion information. It was shown that a hypothesis

number no larger than three is suitable at low bit rates. Also, in the triple-hypothesis motion compensated prediction, the optimum R-D points are achieved when the second coefficient w_2 is set to 1/3. By comparing MHMCP with the IR scheme, we showed that MHMCP suppresses the short-term error propagation more effectively than IR.

REFERENCES

[1] T. Wiegand and B. Girod, *Multi-Frame Motion-Compensated Prediction for Video Transmission*. Norwell: Kluwer, 2001.

[2] T. Wiegand, X. Zhang, and B. Girod, "Long-term memory motion-compensated prediction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, no. 1, pp. 70–84, Feb. 1999.

[3] B. Girod and N. Farber, "Feedback-based error control for mobile video transmission," *Proc. IEEE*, vol. 87, pp. 1707–1723, Oct. 1999.

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

[5] *Information Technology—Coding of Audio-Visual Objects—Part 2: Visual*, ISO/IEC JTC 1/SC 29/WG11 N3056, 2000.

[6] *Study of Final Committee Draft of Joint Video Specification*, ITU-T Recommendation H.264, ISO/IEC 14 496-10 AVC, 2003.

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

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

[9] T. Wiegand, M. Flierl, and B. Girod, "Entropy-constrained linear vector prediction for motion-compensated video coding," in *IEEE Int. Symp. Inf. Theory*, vol. 3, Aug. 1998, p. 409.

[10] 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.*, Mar.–Apr. 1998, pp. 239–248.

[11] ——, "Rate-constrained multihypothesis prediction for motion-compensated video compression," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 11, pp. 957–969, Nov. 2002.

[12] C.-S. Kim, R.-C. Kim, and S.-U. Lee, "Robust transmission of video sequence using double-vector motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 11, no. 9, pp. 1011–1021, Sep. 2001.

[13] Y. Wang and S. Lin, "Error-resilient video coding using multiple description motion compensation," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, no. 6, pp. 438–452, Jun. 2002.

[14] S. Lin and Y. Wang, "Analysis and improvement of multiple description motion compensation video coding for lossy packet networks," in *IEEE Int. Conf. Image Processing*, vol. 2, Sep. 2002, pp. 185–188.

[15] C.-S. Kim, J. Kim, I. Katsavounidis, and C.-C. J. Kuo, "Robust mmse video decoding: theory and practical implementations," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 1, pp. 39–51, Jan. 2005.

[16] M. E. Al-Mualla, C. N. Canagarajah, and D. R. Bull, "Multiple-reference temporal error concealment," in *Proc. IEEE Int. Symp. Circuits Systems*, vol. 5, May 2001, pp. 149–152.

[17] W.-Y. Kung, C.-S. Kim, and C.-C. J. Kuo, "A dynamic error concealment for video transmission over noisy channels," in *Proc. IEEE GLOBECOM*, vol. 2, Nov. 2002, pp. 1769–1773.

[18] Y. O. Park, C.-S. Kim, and S.-U. Lee, "Multi-hypothesis error concealment algorithm for H.26l video," in *IEEE Int. Conf. Image Processing*, vol. 3, Sep. 2003, pp. 465–468.

[19] S. Lin and Y. Wang, "Error resilience property of multihypothesis motion-compensated prediction," in *Proc. IEEE Int. Conf. Image Processing*, Jun. 2002, pp. 545–548.

[20] N. Farber, K. Stuhlmuller, and B. Girod, "Analysis of error propagation in hybrid video coding with application to error resilience," in *IEEE Int. Conf. Image Processing*, Oct. 1999, pp. 550–554.

[21] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: Wiley, 1991.