Adaptive Multi-path Prediction for Error Resilient H.264 Coding

Xiaosong Zhou and C.-C. Jay Kuo
Integrated Media Systems Center and Department of Electrical Engineering
University of Southern California, Los Angeles, California, U.S.A.

***Abstract*— An adaptive reference selection (ARS) scheme is proposed to enhance error resilience performance of H.264 video in this work, where multiple prediction paths can be created in the compressed video stream at the macroblock level without a large amount of bit rate overhead. We first develop a method to measure the expected distortion at the decoder when the H.264 video is transmitted through erroneous channels. Then, we use an updated rate-distortion cost function to incorporate this measurement into the mode decision process. The best prediction for each macroblock is selected with the objective to achieve the highest expected rate-distortion performance of the GOP in the received video stream. It is shown by experimental results that error propagation is largely reduced and the quality of received video stream is improved significantly by the proposed scheme.**

Topic area - Multimedia Processing.

I. INTRODUCTION

Digital video communication has found a broader range of applications. When the communication channel fails to provide guaranteed quality of service (QoS), video data packets may be lost, corrupted or delayed during transmission. A single transmission error can degrade the PSNR and subjective quality of received video over a long period due to error propagation. Even though the emerging video coding standard, H.264 [1], has excellent performance in the rate-distortion performance, there is not much work reported in the literature on its error resilient coding to reduce error propagation by exploring new features of this new standard.

Various error resilient tools have been introduced in the past [2] to improve the error resilience of coded video such as error resilient entropy coding [3] and unequal error protection by scalable or layered coding [4]. They can be implemented to help reconstruction of corrupted video data. However, as the reconstruction is not perfect, mismatch exists and error propagation cannot be stopped properly. Additionally, these methods may not be compatible with the H.264 standard. One way to reduce error propagation in previous standards is to insert intra macroblocks in temporally coded (P or B) video frames, which is called intra refreshing. An optimal intra/inter mode selection method was proposed in [5] with excellent error resilient coding performance. However, as intra macroblocks in H.264 are coded by intra predictions based on pixels of neighboring areas, the insertion of intra macroblocks may fail to stop error propagation since its neighboring blocks could be corrupted or affected by transmission errors. Another problem of intra refreshing is that intra macroblocks have much lower coding efficiency than inter macroblocks so that the overall coding efficiency of the video stream may be degraded significantly if a large number of intra macroblocks are inserted.

In H.264, multiple reference (long-term reference) motion compensation predictive coding (LTMCP) is used to enhance the coding efficiency of the compressed video stream. This or other similar coding feature proposed before has been investigated for error resilience. For example, a reference selection scheme was proposed in [6] for real-time encoders to select the best reference frame by evaluating the expected reconstruction calculated based on the error feedback and an error propagation model. Recently, we proposed a feedback-based error resilient scheme called ”alternative macroblock coding (AMC)” in [7] and [8] that is able to largely stop error propagation for off-line coded video. One of the methods to create alternative error-resilient macroblocks is to utilize the long-term reference frames.

In this work, we further explore the LTMCP feature to develop an error resilient video coding scheme that is able to generate off-line coded video with improved error resilience performance without the requirement of a feedback channel. We observe that quality degradation caused by error propagation in coded H.264 video can be largely reduced using multiple prediction paths. An adaptive reference selection (ARS) scheme is developed to select the best reference frame for each macroblock to create multiple prediction paths and achieve the highest expected rate-distortion performance of the received video in erroneous environment. The ARS scheme can also be tailored to other adaptive mode decision problems in an erroneous environment.

The rest of the paper is organized as follows. A new error resilient video coding scheme is proposed in Sec. II, where the concept of multi-path predictive coding and the ARS scheme is discussed in detail. The method of calculating expected decoder distortion for H.264 is also illustrated in this section. It is followed by experimental results in Sec. IV and concluding remarks in Sec. V.

II. ADAPTIVE MULTI-PATH PREDICTION FOR H.264

A. Observation and Research Motivation

When a compressed video stream is transmitted over erroneous channel, a correctly received macroblock may not be perfectly reconstructed if its reference used in predictive coding is corrupted by errors. Similarly, blocks in subsequent frames that use part of this macroblock as reference will be affected as well. This phenomenon is called error propagation, which degrades the quality of received video over a long while. Fig. 1 (a) shows how video frames are usually coded using motion compensation in a group of pictures (GOP). Each P frame uses its immediate preceding frame as the reference. Based on this prediction pattern, a single transmission error

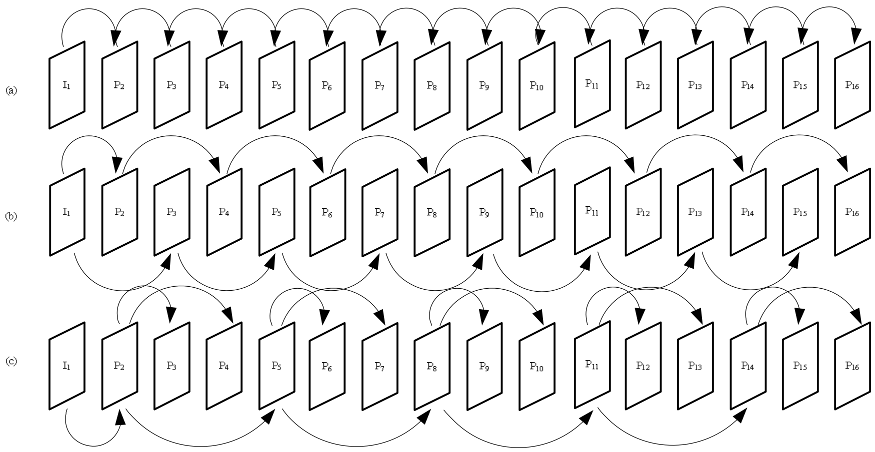

Fig. 1. Different patterns in multi-path predictive coding.

may cause error propagation in all subsequent frames. It is obvious that video frames that are further away from the I frame, which is the first frame in the figure, are more vulnerable to error propagation since any transmission error that occurs in any of its previous frames may affect its reconstruction.

In H.264, the new coding feature LTMCP allows the encoder to choose the best prediction from a number of reference frames. It improves coding efficiency of motion compensated predictive video coding since the best reference for the prediction of some blocks may exist in a long-term reference frame. However, only a small potion of blocks falls in this category, and the prediction shown in Fig. 1 (a) is still common for H.264 coded video.

In this work, we argue that it is actually convenient to create multiple prediction paths in the coded video stream by utilizing long term reference frames for the purpose of error resilience. As shown in Fig. 1 (b) and (c), video frames are encoded into certain predictive patterns to create multiple prediction paths. For a given P frame, say, P_{10}, the number of frames in its prediction path is reduced by about one half and one third in patterns (b) and (c) as compared to that in pattern (a). As a result, these two patterns are much less vulnerable to error propagation. In particular, for every frame in pattern (b), the number of frames in its prediction path reduces to one half. Errors that occur in even-indexed P frames will not affect the reconstruction of odd-indexed P frames, and vice versa. For pattern (c), transmission errors that occur in frames P_{3n} and P_{3n+1} are less critical than errors in frames P_{3n-1}. Thus, the error resilience performance of the coded video stream is improved significantly using such alternative prediction patterns.

It is possible to create multiple prediction paths in the coded video stream at the frame level as shown in Fig. 1. However, as most video frames are forced to use a distant reference frame, coding efficiency is likely to be sacrificed if a video stream is encoded into these fixed prediction patterns. The coding performance of predictions generated from the same reference frame may vary significantly for different macroblocks since it is largely dependent on the video content. It is difficult to design a fixed prediction pattern at the frame level to realize multi-path predictive coding without sacrificing coding efficiency significantly.

In H.264, reference frame selection is done at the macroblock level. Motion estimation is processed with respect to the macroblock of the target frame, and rate-distortion costs of various macroblock predictions in multiple reference frames are compared. The macroblock that offers the best prediction is chosen as the final reference. It is desirable to incorporate the idea of multi-path predictive coding at the macroblock level to seek the best balance between error resilience and coding efficiency. To realize this idea, different levels of error resilience performance of predictions need to be measured, which is discussed in the next subsection.

B. Computation of Expected Decoder Distortion

Based on examples given in Fig. 1, it is clear that multiple reference frames can provide different error resilience degrees. It is important to measure the error resilience degree and incorporate it in the reference selection process.

The expected decoder distortion D_e, which measures the distortion between the expected reconstruction of the video stream after being transmitted through an erroneous channel and its reconstruction in an error-free environment, can be used as a measure for this purpose. Even though a number of results were reported in the literature [5], [9] to compute expected decoder distortion in previous standards, some of new coding features in H.264, *e.g.* quarter-pel motion resolution and in-loop de-blocking filters, affects the error propagation behavior. Thus, a new method to calculate the expected decoder distortion for H.264 accurately is needed.

We propose a direct method to compute the expected decoder distortion. An error map is created and maintained for each allowed reference frame in the reference buffer. The error map stores the absolute value of the expected error e of every pixel in the frame, which is the difference between the expected decoder reconstruction $\tilde{f}$ in erroneous environment and the reconstruction $\hat{f}$ in error-free environment. We have

$$e = |\hat{f} - \tilde{f}|. \quad (1)$$

As $\tilde{f}$ cannot be obtained directly, for a given pixel in the n^{th} frame, we update its value of e by

$$e_n = (1 - p_e)e_{p,n} + p_e e_{c,n}, \quad (2)$$

where p_e is the channel error rate, $e_{p,n}$ is the expected error from error propagation and $e_{c,n}$ is the expected error from error concealment when the pixel is lost and has to be concealed at the decoder.

To calculate $e_{c,n}$, the mismatch caused by the reconstruction of a certain error concealment scheme is calculated. Here, we consider a simple error concealment method to illustrate the basic idea. That is, a block in intra-coded pictures is concealed by copying pixels from the boundary of the correctly reconstructed block above the target block. A block in inter-coded pictures is concealed by copying the block from the same position in the previous frame. Then, the mismatch can be calculated using the concealed reconstruction and the original reconstruction under an error-free environment. To compute the concealed reconstruction, the expected errors of pixels used to conceal the block should be taken into account. The de-blocking operation should also be considered in the distortion calculation since the error generated by error concealment is attenuated by the in-loop de-blocking filter. Fig. 2 shows an example of pixels at different boundaries of 4×4 block, M.

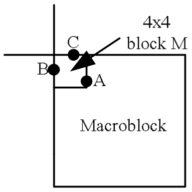

Fig. 2. The in-loop de-blocking operation applied to concealed blocks.

These pixels can be grouped into the following three categories and they should be treated differently.

- Pixels on the inner boundary of a macroblock.

  One such example is Pixel A in Fig. 2. It is on the boundary of M with another block that resides in the same macroblock. As all blocks in a macroblock are lost and concealed together, the de-blocking operation should be processed using the concealed reconstruction of the corrupted block itself and its neighboring block to calculate the new reconstruction value of A. Then, the value e_c of A can be updated accordingly.
- Pixels on the vertical outer boundary of a macroblock.

  Pixel B in Fig. 2 provides an example for this case. When the macroblock is lost or corrupted, its neighboring macroblock in the same row is usually lost or corrupted as well as. Therefore, the concealed reconstruction of neighboring macroblocks should be used in the de-blocking filter.
- Pixels on the horizontal outer boundary of a macroblock

  We refer to Pixel c in Fig. 2 as an example. Since vertically neighboring macroblocks may not be grouped in the same packet as the current macroblock, they can be correctly reconstructed often so that the original reconstruction of the neighboring macroblocks should be used in this case.

To calculate $e_{p,n}$, the motion vector of the block that contains the pixel is retrieved and the propagating error is calculated from the error map of its reference frame. Since quarter-pel motion resolution is used, the error is attenuated. To get the precise value of the propagating error, the same interpolation method used for motion compensation is processed in the error map. For example, for the luminance component, the 6-tap filter $(1, -5, 20, 20, -5, 1)$ is used to create half samples while quarter samples are created by averaging full and half samples in the error map.

Furthermore, the propagating error value of pixels at the boundaries of 4×4 blocks will go through another round of de-blocking. These values are updated by applying the same de-blocking filters to the propagating error of its neighboring pixels. The overall attenuated propagation error e_p of the n^{th} frame can be written of the following form:

$$e_{p,n} = \alpha_{n-1} e_{n-1}, \quad (3)$$

where e_{n-1} is the error value of the pixel that is closest to the position where the motion vector points in the reference frame and α_{n-1} is the attenuation factor of the propagating error from the $(n-1)^{st}$ frame to the n^{th} frame. α_{n-1} can be calculated to evaluate the overall attenuation of propagation error.

With the above method, the expected propagation error and concealment error of each pixel can be calculated in a straightforward fashion. The expected decoder distortion introduced for a given prediction is retrieved from the error map of the prediction block by taking the square of the error values as

$$D_{e,n} = e_n^2. \quad (4)$$

III. ADAPTIVE REFERENCE SELECTION (ARS) SCHEME

In LTMCP, motion search is performed with respect to each of allowed reference frames, and multiple predictions are created accordingly. As shown in Fig. 3, predictions A and B generated from different reference frames can be used to encode block M. As pixels in each prediction go through different prediction paths, they have a different impact on error propagation when they are chosen to predict M. Generally speaking, by choosing prediction A, block M will be less vulnerable against error propagation since fewer frames exist in this prediction path. Such a difference can be measured by

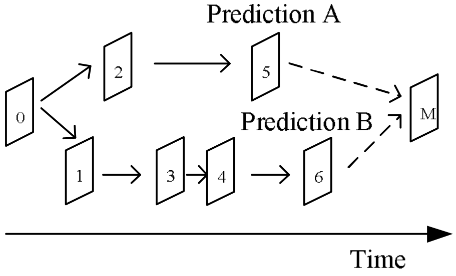

Fig. 3. Two predictions generated by two paths.

the expected decoder distortion D_e of block M when using different predictions as discussed in the previous section.

We propose an adaptive reference selection (ARS) scheme that incorporates the expected decoder distortion in the reference selection process. The predictions generated from multiple reference frames are evaluated based on both their coding and error resilience performance. With the ARS scheme, multiple prediction paths for each macroblock can be created and selected automatically during the encoding process.

Propagating errors caused by a single transmission error are confined within a group of pictures (GOP) since the next GOP starts with a intra picture that is immune to error propagation from previous GOPs. Thus, we choose the GOP as one coding unit. It contains a number of frames, where each frame contains a number of macroblocks. Let x_N be the set of all N macroblocks in a GOP and $x_N = (X_1, \cdots, X_N)$. Macroblock X_i is encoded with mode M_i by choosing from all possible modes that include the macroblock type, the reference frames, the motion estimation block sizes, etc. We use mode vector $m_N = (M_1, \cdots, M_N)$ to denote all modes selected by each macroblock. The quantization parameters used to encode these macroblocks are also expressed in vector form as $q_N = (Q_1, \cdots, Q_N)$.

Then, the overall mode decision problem is to find a combination of coding modes and quantization parameters for each macroblock to achieve the best coding performance at a given rate constraint of GOP R_g. It can be formulated mathematically as

$$\min_{m_N, q_N} D(x_N, m_N, q_N) \text{ subject to } R(x_N, m_N, q_N) \leq R_g. \quad (5)$$

where $D(x, m, q)$ and $R(x, m, q)$ are the sums of distortions and bit rates of all macroblocks in the GOP. The problem can be converted to an unconstraint optimization problem using the Lagrange Multiplier method [10]. That is, we have

$$\min_{m_N, q_N} \sum_{i=1}^{N} J(X_i, m_N, q_N), \quad (6)$$

where

$$J(X_i, m_N, q_N) = D(X_i, m_N, q_N) + \lambda R(X_i, m_N, q_N) \quad (7)$$

is the Lagrange cost function for macroblock X_i.

In erroneous environment, the above problem needs to be updated by minimizing the expected distortion of the received video stream under the same constraint. As discussed, D_e measures the expected distortion between the expected decoder reconstruction $\tilde{f}$ in erroneous environment and the reconstruction $\hat{f}$ in error-free environment. Suppose the original value of the pixel is f, then the overall expected distortion of the reconstruction at the decoder end after transmitted through an erroneous channel can be written as

$$\begin{aligned} E(D') &= E[(f - \tilde{f})^2] \\ &= E(f - \hat{f})^2 + E(\hat{f} - \tilde{f})^2 \\ &\quad + 2E(\hat{f} - \tilde{f})(f - \hat{f}). \end{aligned} \quad (8)$$

The term $\hat{f} - \tilde{f}$ is the expected error that has mean zero and it is uncorrelated to the encoder error $f - \hat{f}$. Therefore, we have

$$E(D') = E(f - \hat{f})^2 + E(\hat{f} - \tilde{f})^2 = D + D_e. \quad (9)$$

We can re-write Eqn. 6 and Eqn. 7 as

$$\min_{m_N, q_N} \sum_{i=1}^{N} J'(X_i, m_N, q_N), \quad (10)$$

where

$$\begin{aligned} J'(X_i, m_N, q_N) &= D(X_i, m_N, q_N) + D_e(X_i, m_N) \\ &\quad + \lambda R(X_i, m_N, q_N). \end{aligned} \quad (11)$$

is the updated Lagrange rate-distortion cost function.

In words, the problem is to find the best combination of coding modes and parameters to get the best overall rate-distortion performance of the reconstructed video stream through erroneous channels. Here, we only consider different error resilience performance caused by the reference frame selection. Thus, $D_e(X_i, m_N)$ is only affected by the choice of reference frames.

Eqn. 10 is difficult to solve if we consider the dependency among macroblocks. Some methods have been proposed to solve a simplified version of Eqn. 10 using optimization techniques by assuming that the impact on rate and distortion of any give macroblock is limited to its neighbors spatially [11]. However, a high computational complexity is still required to solve the problem.

A practical solution to this problem is often adopted by assuming that the mode decision of a given macroblock is not affected by other macroblocks. Then, the mode decision problem of each macroblock can be solved independently. Mathematically, we can re-write Eqn. 10 as

$$\min_{m_N, q_N} \sum_{i=1}^{N} J'(X_i, m_N, q_N) = \sum_{i=1}^{N} \min_{M_i, Q_i} J'(X_i, M_i, Q_i). \quad (12)$$

This approach works well for mode decision without considering the error resilience performance. However, it does not work well when error resilience is considered. This can be explained by the observation that the weights of expected decoder distortions for macroblocks in mode decision are identical in Eqn. 12. However, in a given GOP, the error resilience performance of macroblocks in earlier frames should have more impact to the GOP than that in later frames due to error propagation.

To determine the mode and the quantization parameter for macroblock X_i to achieve the best result for the whole GOP in an erroneous environment, we impose the simplifying assumption; namely, rate $R(X_i, M_i, Q_i)$ and distortion $D(X_i, M_i, Q_i)$ of macroblock i have no impact to other macroblocks. However, expected decoder distortion, $D_e(X_i, M_i)$, does affect macroblocks in subsequent frames due to error propagation. Such an impact can be estimated by calculating the total propagation distortion in the GOP introduced by $D_e(X_i, F_i)$. Since both D and R are additive terms, the local mode decision cost function for macroblock X_i satisfying the objective function in Eqn. 10 can be written as

$$\begin{aligned} \min_{M_i,Q_i} \quad & J"(X_i, M_i, Q_i) \\ = \quad & \min_{M_i,Q_i} [D(X_i, M_i, Q_i) + (1+\beta) D_e(X_i, M_i) \\ & + \lambda R(X_i, M_i, Q_i)], \end{aligned} \tag{13}$$

where $\beta D_e(X_i, M_i)$ is the expected overall distortion of other parts of the GOP introduced by $D_e(X_i, F_i)$ due to error propagation and

$$\beta = \sum_{j=1}^{M} \alpha^{2j}, \tag{14}$$

and where α is the attenuation factor and M is the expected number of frames in the future prediction path of the pixel. The method to calculate M is shown in Fig. 4.

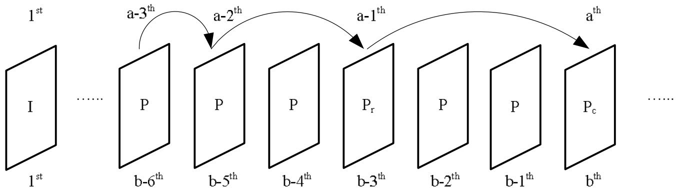

Fig. 4. Illustration of M calculation.

As shown in the figure, when evaluating the prediction generated from the reference frame, P_r, for a macroblock in the current frame, P_c, the previous prediction path is examined and the number of frames in the prediction path, denoted by a, is calculated. Suppose that the total number of frames in a GOP is N and the current frame is the b^{th} frame in the GOP. Then, by assuming that the same prediction path is going to be repeated in future frames, the expected number of frames in its future prediction path is estimated as

$$M = \frac{a}{b}(N - b). \tag{15}$$

With this updated cost function, the mode decision of each macroblock can still be processed independently while the objective of achieving the best rate-distortion performance of the received video stream through erroneous channels is still met. It can be used directly to select the best reference frame.

In summary, the reference frame for a given macroblock can be selected adaptively by considering both coding efficiency and error resilience using the proposed ARS scheme. A sub-optimal solution can be derived to achieve a good R-D performance of the received video stream through erroneous channels. The performance of the proposed ARS scheme is evaluated in the next section.

TABLE I

THE LIST OF TEST CONDITIONS.

Test Condition	Error Type
T1	1×10^{-3} Random Packet Loss
T2	5×10^{-3} Random Packet Loss
T3	1×10^{-2} Random Packet Loss
T4	5×10^{-2} Random Packet Loss
T5	1×10^{-3} Burst Packet Loss (Length = 3)
T6	5×10^{-3} Burst Packet Loss (Length = 3)
T7	1×10^{-2} Burst Packet Loss (Length = 3)
T8	5×10^{-2} Burst Packet Loss (Length = 3)

TABLE II

PSNR (DB) OF THE RECEIVED FOREMAN QCIF SEQUENCE.

Test Cond.	Proposed Scheme	Benchmark	PSNR Gain
T1	35.82	35.60	+0.22
T2	35.64	35.13	+0.51
T3	35.43	34.11	+1.32
T4	33.03	30.66	+2.37
T5	35.85	35.60	+0.25
T6	35.67	35.12	+0.55
T7	35.45	34.03	+1.42
T8	33.23	30.70	+2.53

IV. EXPERIMENTAL RESULTS

In our experiments, the H.264 reference codec JM 10.2 [12] was modified to include the proposed error resilient coding scheme as a module. Its error resilience performance was evaluated under test conditions listed in Table I. We simulated the network packet loss caused by both random errors and burst errors with different error rates. The proposed scheme is compared with the standard H.264 video coding, which serves as the performance benchmark.

Table II shows the PSNR values of received video using the proposed adaptive multi-path selection scheme and the benchmark, where the Foreman QCIF sequence was used in the experiment. The default error concealment tools were used in the decoder to conceal corrupted video data. As shown in the table, the proposed scheme outperforms the benchmark significantly and improves the quality of received video under all test conditions. The quality gain of the proposed scheme increases when the error rate becomes higher.

With the proposed ARS scheme and the updated cost function, multiple predictions paths can be created automatically at the macroblock level. To show that multiple prediction paths can be created effectively by the proposed scheme, the averaged reference frame indices used by all macroblocks in a

GOP at different packet loss rates are shown in Fig. 5, where the preceding frame has index 0 and the index for long-term (distant) reference frames increases by 1 when the distance becomes longer by one frame in the temporal domain. When the loss rate is 0, the original reference frame selection scheme that considers only coding efficiency is used. We see that only a small number of blocks use long-term reference frames in this case. When the loss rate becomes higher, the average reference frame index of the proposed scheme becomes larger. Thus, the proposed ARS scheme generates multiple prediction paths automatically during the mode decision process.

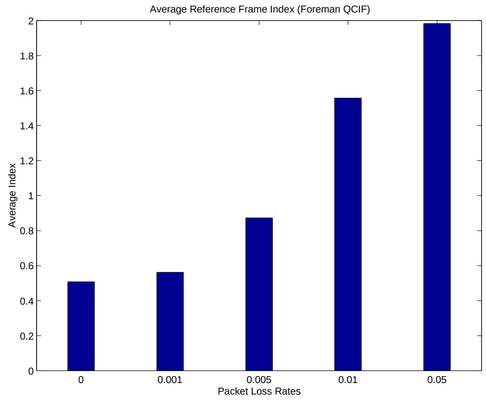

Fig. 5. The averaged frame index.

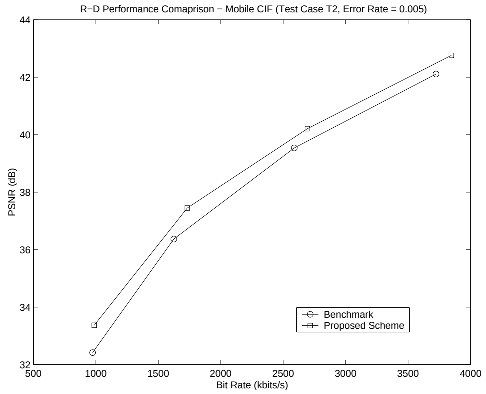

Fig. 6. The R-D performance comparison under test condition T2.

To further demonstrate the performance of the proposed error resilient video coding scheme, the R-D performance of received video for mobile CIF sequence is shown in Fig. 6 and Fig. 7. We see from these two figures that the proposed scheme has much better R-D performance than the benchmark. It confirms the claim that the proposed scheme is able to enhance the error resilience performance of transmitted video without significantly degrading its coding efficiency.

V. CONCLUSION AND FUTURE WORK

An adaptive prediction selection scheme was proposed in this work to create multiple prediction paths in the compressed video stream. An accurate method was developed specifically for H.264 to calculate the expected decoder distortion generated by different predictions to support the proposed scheme. It was demonstrated by experimental results that the existence of multiple prediction paths improves the error resilience performance of H.264 coded video significantly. The proposed scheme is able to maintain good coding efficiency of the compressed stream while serve as an effective error resilience tool in visual communication applications. In the future, we plan to develop a new model to simplify the calculation of the expected decoder distortion to reduce the complexity of the proposed scheme.

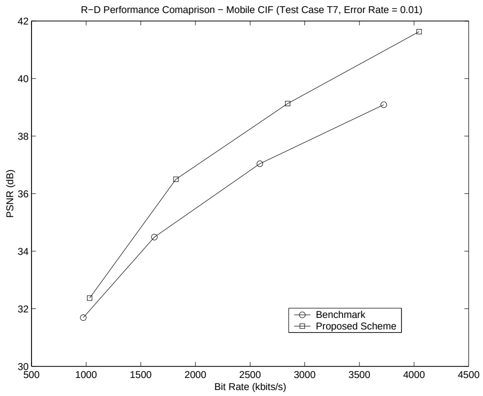

Fig. 7. The R-D performance comparison under test condition T7.

REFERENCES

[1] ITU-T Rec. H.264 - ISO/IEC 14496-10 AVC, Document JVT-G050, 8th Meeting: Geneva, Switzerland, May 2003.

[2] Y. Wang, S. Wenger, J. Wen, and A. K. Katsaggelos, “Error resilient video coding techniques,” *IEEE Signal Processing Magazine*, pp. 61-82, July 2000.

[3] J. Wen and J. Villasenor, “A class of reversible variable length codes for robust image and video coding,” in *Proc. IEEE Int. Conf. on Image Processing*, vol. 2, pp. 65-68, 1997.

[4] M.Ghanbari, “Two-layer coding of video signals for VBR networks,” *IEEE J. Select. Areas Commun.*, vol. 7, pp. 771-781, 1989.

[5] R. Zhang, S. L. Regunathan, and K. Rose, ”Video coding with optimal inter/intra-mode switching for packet loss resilience,” *IEEE J. Select. Areas Commun.*, vol. 18, pp. 966-976, June 2000.

[6] T. Wiegand and N. Färber and K. Stuhlmüller and B. Girod, ”Error-resilient video transmission using long-term memory motion-compensated prediction”, *IEEE J. Select. Areas Commun.*, vo.18, pp 1050-1062, June 2000.

[7] Xiaosong Zhou, Wei-ying Kung and C.-C. Jay Kuo, ”A robust H.264 video streaming scheme for portable devices”, *IEEE International Symposium on Circuits and Systems*, Kobe, Japan, May 2005.

[8] Xiaosong Zhou, Wei-ying Kung and C.-C. Jay Kuo, ”Error resilient H.264 video with SP/SI coded macroblocks”, *The 14th International Packet Video Workshop*, December 2004.

[9] Chang-Su Kim, JongWon Kim, Ioannis Katsavounidis and C.-C. Jay Kuo, ”Robust MMSE video decoding: theory and practical implementations ,” in *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 15, no. 1, pp. 39-44, January 2005.

[10] H. Everett III, ”Generalized Lagrange multiplier method for solving problems of optimum allocation of resources,” in *Operations Res.*, vol. 11, pp. 399-417, 1963.

[11] A. Ortega and K. Ramchandran, ”Forward-adaptive quantization with otpimal overhead cost for image and video coding with applications to MPEG video coders,” in *Proc. IS&T SPIE, Digital Video Compression: Algorithms and Technologies*, vol. 2419, pp. 129-138, Feburary 1995.

[12] Joint Video Team, “JVT Reference Software, version JM 10.2,” http://bs.hhi.de/seuhring/tml/download.