Motion Vector Recovery for Error Concealment Based on Macroblock Distortion Modeling

Jae-Won Suh and Yo-Sung Ho

Kwangju Institute of Science and Technology,
1 Oryong-Dong Puk-Gu, Kwangju, 500-712, Korea
{won, hoyo}@kjist.ac.kr
http://vclab.kjist.ac.kr/index.html

Abstract. If channel errors are introduced during the video transmission, we can apply error concealment techniques to reduce the transmission error effects by exploiting spatial and temporal redundancies of the video signal. Recently, motion vector recovery and motion compensation with the estimated motion vector has been one of the critical issues for the temporal-domain error concealment. In this paper, we prove that it is reasonable to use the estimated motion vector for concealing the lost macroblock by providing new macroblock distortion models. Based on the proposed distortion models, we develop several motion vector recovery algorithms.

1 Introduction

Fast growth of digital transmission services has generated a great deal of interest in digital transmission of video signals over a bandlimited channel. The bandwidth limitation necessitates the efficient coding schemes to compress an enormous amount of digitized video data, such as H.261, H.263, MPEG-1, and MPEG-2. Applications of digital broadcasting require about 50:1 or greater compression ratios. The MPEG-2 video coding standard [1] successfully satisfies the requirement of compression ratio using a hybrid algorithm of motion compensation (MC) and discrete cosine transform (DCT).

Due to the complex coding structure of the MPEG-2 video compression algorithm, compressed bitstreams are very sensitive to channel disturbances. Even one bit error can degrade not only the current frame but also succeeding frames. Therefore, we need error resilient coding techniques both to protect the transmission data and to reduce the transmission error effects. If we consider one layer bitstream without the network information, there are mainly two different approaches. First, we can apply channel coding techniques at the transmission level, such as forward error correction (FEC) codes. However, these approaches increase transmission bit rates because of the parity bits for error detection and correction. Therefore, if we consider the limited channel bandwidth, FEC may not always be the best solution.

As an alternative approach, we can apply error concealment techniques to hide the lost macroblock (MB) data. Error concealment techniques can be classified into two main groups by exploited redundancies in the video sequence:

M.H. Shafazand and A M. Tjoa (Eds.): EurAsia-ICT 2002, LNCS 2510, pp. 180–187, 2002.

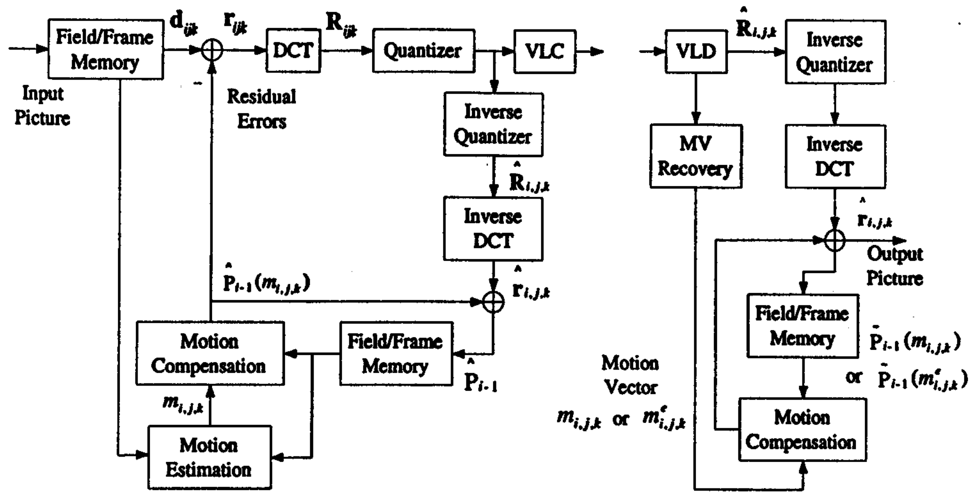


Fig. 1. Modified MPEG-2 Video Codec

spatial-domain error concealment techniques and temporal-domain error concealment techniques. Spatial-domain error concealment techniques interpolate the lost MB data using adjacent luminance and color values of the corrupted MB [2][3]. Temporal-domain error concealment techniques estimate lost motion vector (MV) of the corrupted MB and compensate it by the estimated MV under the assumption that MVs of spatially adjacent MBs are highly correlated [4]-[7]. Although we cannot recover the lost data perfectly, we can get similar MB data without increasing transmission bit rate.

In this paper, we propose new MB distortion models for the predictive frame (P-frame) to effectively conceal the corrupted MB data and increase the performance of error concealment. Based on the proposed MB distortion models, we prove that MC with the estimated MV is reasonable for temporal-domain error concealment techniques. There is a similar analysis for overall distortion of the decoded frame for H.263 [8]. In order to evaluate the proposed MV recovery algorithms, we compare and analyze simulation results with the previous works.

2 MPEG-2 Video Codec

Fig. 1 shows a block diagram of the MPEG-2 video codec. For intra frame (I-frame) coding, the first frame of group of pictures (GOP) is partitioned into nonoverlapping blocks of 8×8 pixels. After two-dimensional DCT is applied to each block independently, DCT coefficients are quantized and encoded using variable length coding (VLC). In the succeeding frames, temporal redundancies are removed by motion estimation (ME) and MC. The residual errors are the difference between the MB to be encoded of the current frame and the best matching block in the reference frame. The residual errors and MV information are encoded and transmitted. The decoding procedure is just opposite of the encoding procedure. In our modified MPEG-2 video decoder, if we receive a corrupted packet, we estimate MV of the lost MB in the MV recovery block and compensate the lost MB using the estimated MV.

The next problem is how to detect the transmission error because error concealment is largely dependent on the capability of error detection. We solve the error detection problem by checking the MB address (MBA). MBA is the unique address of the absolute position of the MB. Whenever we decode correctly received packets, we store the last decoded MBA. If we cannot decode the received packet data and resynchronization is established in the next slice header, we assume that (stored MBA + 1) is the beginning position of the erroneous MBs.

3 MB Distortion Models for P-frame

The MPEG-2 video codec is shown in Fig. 1, where $\{i,j,k\}$ indicates the k-th MB in the j-th slice of the i-th frame and bold character means MB data. $\mathbf{r}_{i,j,k}$ denotes residual errors, and $\mathbf{R}_{i,j,k}$ is DCT transformed data of the residual errors. $\hat{\mathbf{R}}_{i,j,k}$ means quantized data of $\mathbf{R}_{i,j,k}$ and $m_{i,j,k}$ represents a MV. $\hat{\mathbf{P}}_{i-1}(m_{i,j,k})$ and $\tilde{\mathbf{P}}_{i-1}$ indicate reference frames that are used for ME and MC at the encoder and the decoder, respectively, where i is the coding order. They should be differently notified because some parts of the reference frame at the decoder might be concealed.

First of all, we define the slice error probability by Eq. (1). Let $S_{i,j}$ be the event that the j-th slice of the i-th frame is corrupted.

$$\begin{aligned} Pr(S_{i,j}) &= \sum_{n=1}^{N_{i,j}} {}_nC_k P_b^n (1-P_b)^{N_{i,j}-n} \\ i &= 1,2,\cdots,N, \quad k \le n, \quad and \quad j = 1,2,\cdots,J \end{aligned} \tag{1}$$

where P_b is bit error probability, $N_{i,j}$ is the number of bits of the j-th slice of the i-th frame, J is the number of the slices in the frame, and N is the number of frames.

The distortion $\mathbf{e}^I_{i,j,k}$ for the intra MB of the P-frame can be represented by

$$\mathbf{e}^I_{i,j,k} = E\big[(\mathbf{d}_{i,j,k} - \hat{\mathbf{d}}_{i,j,k})^2\big](1 - Pr(S_{i,j})) + E\big[(\mathbf{d}_{i,j,k} - \tilde{\mathbf{d}}_{i,j,k})^2\big]Pr(S_{i,j}). \tag{2}$$

In Eq. (2), if we receive the corrupted MB data, concealed data $\tilde{\mathbf{d}}_{i,j,k}$ replaces quantized MB data $\hat{\mathbf{d}}_{i,j,k}$ by error concealment operation. The first term represents the quantization error and the second term means the concealment error. If we assume that the quantization error and the concealment noise are uncorrelated, the second term in Eq. (2) can be decomposed into

$$E\big[(\mathbf{d}_{i,j,k} - \tilde{\mathbf{d}}_{i,j,k})^2\big] = E\big[(\mathbf{d}_{i,j,k} - \hat{\mathbf{d}}_{i,j,k})^2\big] + E\big[(\hat{\mathbf{d}}_{i,j,k} - \tilde{\mathbf{d}}_{i,j,k})^2\big]. \tag{3}$$

By Eq. (3), we can rewrite Eq. (2) as

$$\mathbf{e}^I_{i,j,k} = E\big[(\mathbf{d}_{i,j,k} - \hat{\mathbf{d}}_{i,j,k})^2\big] + E\big[(\hat{\mathbf{d}}_{i,j,k} - \tilde{\mathbf{d}}_{i,j,k})^2\big]Pr(S_{i,j}). \tag{4}$$

$\tilde{\mathbf{d}}_{i,j,k}$ can be obtained by $\tilde{\mathbf{P}}_{i-1}(m^e_{i,j,k})$, that is, we can employ ME of the lost MB and MC algorithms to conceal the the lost MB.

The distortion $\mathbf{e}^P_{i,j,k}$ for the inter MB of the P-frame can be expressed as

$$\begin{aligned}\mathbf{e}^P_{i,j,k} &= E\big[\{(\mathbf{r}_{i,j,k} + \hat{\mathbf{P}}_{i-1}(m_{i,j,k})) - \tilde{\mathbf{P}}_{i-1}(m^e_{i,j,k})\}^2\big]Pr(S_{i,j}) \\ &+ E\big[\{(\mathbf{r}_{i,j,k} + \hat{\mathbf{P}}_{i-1}(m_{i,j,k})) - (\hat{\mathbf{r}}_{i,j,k} + \tilde{\mathbf{P}}_{i-1}(m_{i,j,k}))\}^2\big](1 - Pr(S_{i,j})).\end{aligned} \quad (5)$$

If we receive corrupted MB data, we estimate the MV of the lost MB and compensate with the estimated MV to conceal the lost MB, as in the first term of Eq. (5). In order to separate meaningful terms, if we assume that the quantization noise of the residual errors and the concealment noise are uncorrelated, the first term of Eq. (5) can be rearranged as

$$\big(E\big[(\mathbf{r}_{i,j,k} - \hat{\mathbf{r}}_{i,j,k})^2\big] + E\big[\{\hat{\mathbf{r}}_{i,j,k} + \hat{\mathbf{P}}_{i-1}(m_{i,j,k}) - \tilde{\mathbf{P}}_{i-1}(m^e_{i,j,k})\}^2\big]\big)Pr(S_{i,j}). \quad (6)$$

If we use another assumption that the quantization noise of the residual errors and the mismatch between reference frames are uncorrelated, the second term of Eq. (5) can be represented as

$$\big(E\big[(\mathbf{r}_{i,j,k} - \hat{\mathbf{r}}_{i,j,k})^2\big] + E\big[\{\hat{\mathbf{P}}_{i-1}(m_{i,j,k}) - \tilde{\mathbf{P}}_{i-1}(m_{i,j,k})\}^2\big]\big)(1 - Pr(S_{i,j})). \quad (7)$$

By Eq. (6) and Eq. (7), we can rewrite Eq. (5) as

$$\begin{aligned}\mathbf{e}^P_{i,j,k} &= E\big[(\mathbf{r}_{i,j,k} - \hat{\mathbf{r}}_{i,j,k})^2\big] \\ &+ E\big[\{\hat{\mathbf{P}}_{i-1}(m_{i,j,k}) - \tilde{\mathbf{P}}_{i-1}(m_{i,j,k})\}^2\big](1 - Pr(S_{i,j})) \\ &+ E\big[\{\hat{\mathbf{r}}_{i,j,k} + \hat{\mathbf{P}}_{i-1}(m_{i,j,k}) - \tilde{\mathbf{P}}_{i-1}(m^e_{i,j,k})\}^2\big]Pr(S_{i,j}).\end{aligned} \quad (8)$$

In Eq. (8), we cannot treat the first term to improve the performance of error concealment since it is the quantization noise of the residual errors. If there is no error in the reference frame, we can remove the second term. From Eq. (4) and Eq. (8), it seems reasonable to conclude that we estimate the lost MV from $\tilde{\mathbf{P}}_{i-1}$ and compensate for the lost MB using the estimated MV $m^e_{i,j,k}$ for error concealment.

4 Error Concealment Techniques

In order to conceal the lost MB, we use upper and lower MB data of the lost MB. A simple estimate value for the lost MV is zero with an assumption that no motion has occurred between the previous reference frame and the current frame. Use of the zero MV produces a reasonably good approximation in the static scene. However, we cannot expect good result in the dynamic scene. Therefore, we should consider motion compensated error concealment algorithms.

In order to estimate MV, we can exploit neighboring MVs of the lost MB. MV of the lost MB can be obtained by taking the average value of MVs of the vertically adjacent MBs (AVG) [4]. In this scheme, if vertically neighboring

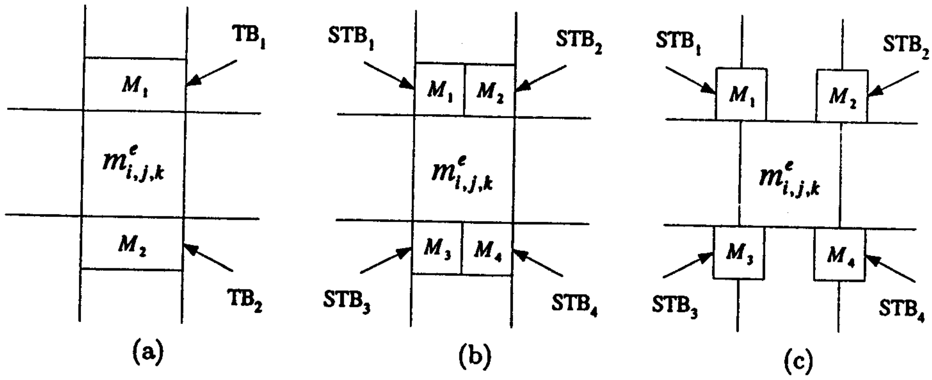
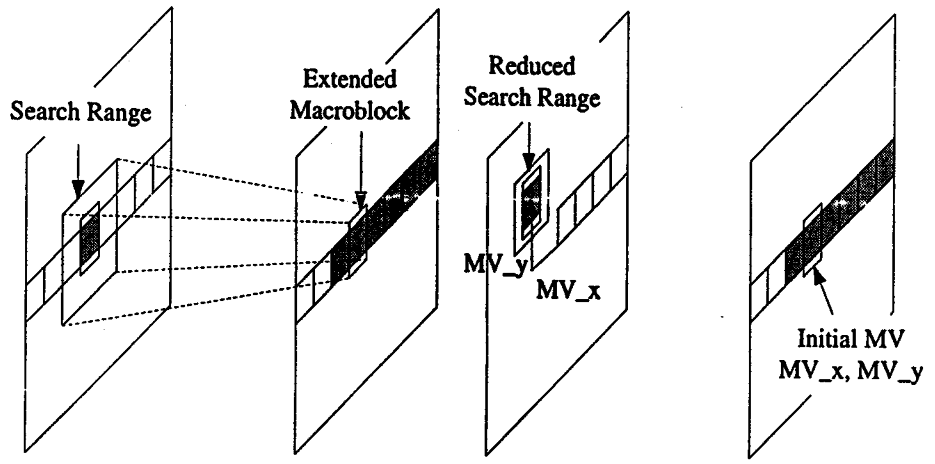


Fig. 2. Modified Average Algorithm

MBs have MVs, we can obtain reasonably good reconstruction quality for the lost MB. However, if only one or none of the vertical neighbors has a valid MV, quality of the reconstructed image is not satisfactory. For this reason, we devise the following modified average (MAVG) schemes for recovering the lost MV [6].

As shown in Fig. 2(a), if vertically adjacent MBs are coded by intra mode, we define 16×8 target blocks (TBs), TB_1 and TB_2, above and below the lost MB. The MV of each TB is computed by the block matching algorithm at the decoder. We substitute the lost MV by the average of the MVs of the two TBs. In order to obtain more accurate MV for the lost MB, we separate each 16×8 TB into two 8×8 small target blocks (STBs), as shown in Fig. 2(b). We can take average of the MVs of the four STBs for the lost MV. Although these two methods produce good performance, they entail a considerable amount of processing complexity at the decoder. In order to reduce the processing time, we define alternative STBs, as shown in Fig. 2(c). Because estimated MVs of STB_2 and STB_4, M_2 and M_4, can be used as M_1 and M_3 in the next MV recovery process, we can reduce the computational complexity by half.

Other MV recovery algorithms use boundary pixels of the lost MB to estimate the lost MV. Within the given search range (SR), the boundary matching algorithm (BMA) [5] calculates the squared sum of differences (SSD) between outer one pixel boundary line of the above, below, and left sides of the lost MB in the current frame and inner one pixel boundary line of the target block in the previous reference frame. BMA replaces the lost MB with the target block data that has the smallest total SSD. The decoder motion vector estimation algorithm (DMVE) [7] is very similar to BMA. Different thing is that DMVE uses outer several pixel boundary lines (two to eight) of the lost MB in the current frame and the previous reference frame to calculate the SSD. However, these algorithms, BAM and DMVE, have significant limitation: left outer boundary pixels of the lost MB are not available when successive MBs are lost. Nevertheless, if left outer boundary pixels are used in computing the SSD, it means that the MV estimation process includes error concealment mismatch from the second MB of the corrupted MBs. In order to resolve this problem, we propose an extension matching algorithm (EMA).



(a) Extension Matching Algorithm (b) EMA with Initial Motion Vector

Fig. 3. EMA with Initial Motion Vector

As shown in Fig. 3(a), we form an extended MB by including pixels above and below the lost MB. The MV is then searched to minimize the SSD between extended pixels in the current and the previous frame. In this algorithm, the extended width (EW) and the SR should be selected appropriately. Since the best matching block should be searched for within the given SR, this technique may entail a considerable amount of processing complexity at the decoder. In order to reduce the computational complexity of EMA, we modify the algorithm with an initial estimate of the MV, as shown in Fig. 3(b). We first set the initial MV for the lost MB by AVG. The initial MV establishes a starting point of the SR and enables to reduce the SR. If none of the vertical neighbors has a valid MV, we use normal SR. As a result, we can effectively reduce the processing time and get good performances.

5 Simulation Results

In order to evaluate the performance of the error concealment algorithms, we perform computer simulations on various test video sequences: FOOTBALL, BICYCLE, and BALLET. They have the 4:2:0 format of 720 × 480 pixels. These are encoded using the pattern of IBBPBBPBBPBB at a frame rate of 30 frames/sec. We consider TS packet transmission system for considering noisy channel. We assume that the first P-frame has lost one TS packet.

We compare the performance of conventional algorithms including AVG [4], BMA [5], and DMVE [7] and proposed algorithms, MAVG [6], EMA, and IEMA. In order to estimate MV of the lost MB, BMA, DMVE, and EMA use [-25, 24] SR. While BMA takes one pixel boundary line, DMVE, EMA, and IEMA exploit variable pixel boundary lines from one to eight. From the simulation, we observe that DMVE produces best results when it takes two pixel boundary lines. EMA and IEMA generate the best performance if they use one pixel boundary line. As

Table 1. Concealed PSNR for Various MV Recovery Algorithms

	FOOTBALL	BICYCLE	BALLET
Original	32.59	26.57	29.12
AVG	30.62	25.15	28.41
BMA	31.25	23.37	28.53
DMVE	31.23	23.30	28.52
MAVG	31.50	24.16	28.76
EMA	31.31	23.76	28.70
IEMA	32.21	25.42	29.01

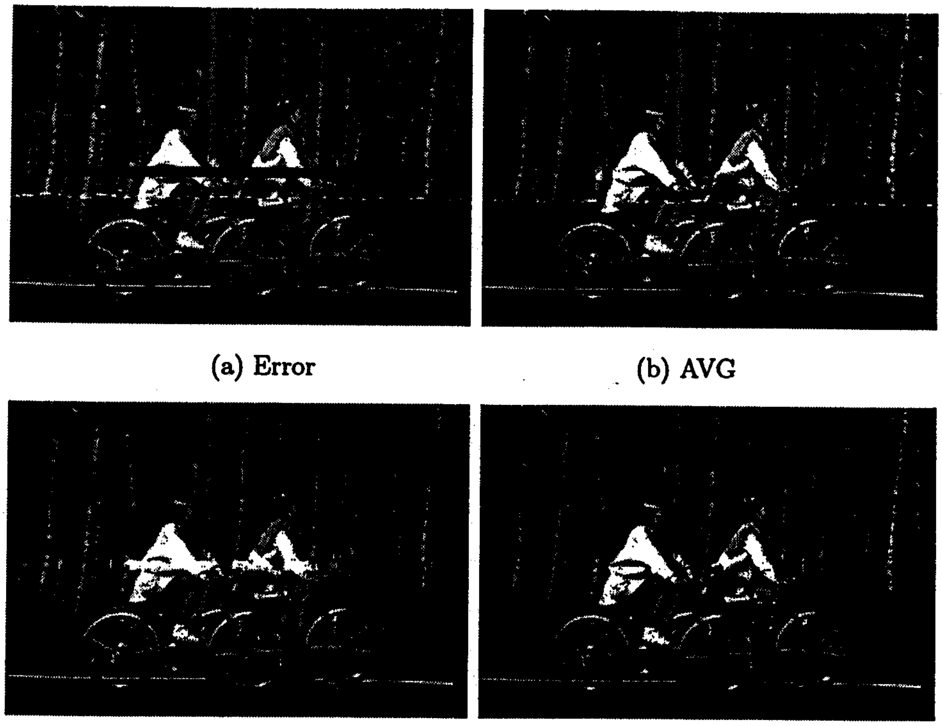
described earlier, IEMA has [-5, 4] reduced SR and [-25, 24] normal SR. MAVG has alternative STB structure to reduce computation time and uses [25, 24] SR for ME of the STB.

Table 1 summarizes the peak-signal-to-noise ratio (PSNR) for the test sequences. BMA, DMVE, and the proposed EMA produce very similar PSNR results. If objects have dynamic motion, such as FOOTBALL and BALLET, BMA, DMVE, and EMA generate good PSNR results. However, if objects have static motion, such as BICYCLE, AVG and MAVG generate good performance. Because the proposed IEMA takes the advantages of each algorithm, it provides the best PSNR results. The subjective quality test of reconstructed frames is very important for the performance comparison of error concealment. As shown in Fig. 4, IEMA shows a good subjective video quality.

6 Conclusion

In this paper, we have shown that motion vector estimation and motion compensation with the estimated motion vector is reasonable for concealing the lost macroblock in temporal-domain error concealment based on proposed MB distortion models. We have reviewed merits and demerits of conventional motion vector recovery algorithms and proposed motion vector recovery algorithms to improve the performance of error concealment. Experimental results of PSNR and subjective quality lead us to the conclusion that the proposed extension matching algorithm with an initial motion vector is one of the best solution for error concealment.

Acknowledgement.
This work was supported in part by the Korea Science and Engineering Foundation (KOSEF) through the Ultra-Fast Fiber-Optic Networks (UFON) Research Center at Kwangju Institute of Science and Technology (K-JIST), and in part by the Ministry of Education (MOE) through the Brain Korea 21 (BK21) project.



(a) Error (b) AVG

(c) BMA (d) IEMA

Fig. 4. Concealed Frames of BICYCLE

References

1. ISO/IEC IS 13818-2 (MPEG-2 Video): Information Technology - Generic Coding of Moving Pictures and Associated Audio Information, April 1996.
2. Aign, S. and Fazel, K.: Temporal and Spatial Error Concealment Techniques for Hierarchical MPEG-2 Video Codec, IEEE International Conference on Communication, Vol. 3, June (1995) 1778-1783.
3. Suh, J.W. and Ho, Y.S.: Error Concealment Based on Directional Interpolation, IEEE Transactions on Consumer Electronics, Vol. 43, No. 3, Aug. (1997) 295-302.
4. Sun, H., Challapali, K., and Zdepski, J.: Error Concealment in Digital Simulcast AD-HDTV Decoder, IEEE Transactions on Consumer Electronics, Vol. 38, No. 3, Aug. (1992) 108-116.
5. Lam, W.M., Reilbman, A.R., and Liu, B.: Recovery of lost or erroneously received motion vectors, Proc. ICASSP, April (1993) V417-V420.
6. Suh, J.W. and Ho, Y.S.: Recovery of Motion Vectors for Error Concealment, Proc. IEEE Region 10 TENCON, Sep. (1999) 750-753.
7. Zhang, J., Arnold, J.F., and Frater, M.R.: A Cell-Loss Concealment Technique for MPEG-2 Coded Video, IEEE Transactions on Circuits and Systems for Video Technology, Vol. 10, No. 4, June (2000) 659-665.
8. Zhang, R., Regunathan, S.L., and Rose, K.: Video Coding with Optimal Inter/Intra-mode switching for packet losss resilience, IEEE Journal on Selected Areas in Communications, Vol. 18, No. 6, June (2000) 952-956.