Hybrid Error Concealment Algorithm for MPEG-4 Video Decoding

Hak-Sop SONG[1] Hiroyuki OKADA[1,2] Gen FUJITA[1] Takao ONOYE[1] Isao SHIRAKAWA[1]

[1]Dept. Information Systems Engineering, Osaka University
2-1 Yamada-Oka, Suita, Osaka, 565-0871 Japan
Tel.: +81-6-6879-7808 Fax: +81-6-6875-5902
E-mail: {shs,h_okada,fujita,onoe, sirakawa}@ise.eng.osaka-u.ac.jp

[2]Digital Information Appliances Development Center, Sharp Corporation
1-9-2 Nakase, Mihama-Ku, Chiba-Shi, Chiba, 261-8520 Japan
Tel.: +81-43-299-8537 Fax: +81-43-299-8509
E-mail: okada_hiroyuki@sharp.co.jp

Abstract: In this paper, a novel error concealment algorithm is proposed for the MPEG-4 video decoding. Apart from existing algorithms which fail to exhibit stable performance over various video sequences and error patterns, the proposed algorithm adopts a new hybrid scheme, which can achieve a consistent performance with reduced computational complexity. This algorithm is implemented on the basis of the MPEG-4 decoder, and the experimental results demonstrate that the new approach provides acceptable performance both subjectively and objectively at various bit error rates and video sequences.

1 Introduction

With the recent increase in the demand for the real-time video transmission over mobile communication networks, transmission errors have become a serious issue. Typically, to combat transmission errors, different error resilience and recovery schemes are employed in video transmission systems [1]. Since these schemes can not correct all errors, it is necessary to handle the residual errors.

The error concealment (EC) is to conceal the lost parts of images which are not decoded due to the residual errors. The subjective quality of a decoded video sequence is influenced mainly by the concealment algorithm to be adopted. So far, a number of EC algorithms [2], [3] have been proposed by means of the motion estimation. Although these algorithms have attempted to improve the effectiveness of concealment, there still remains the technical issue of intensive computational complexity.

Thus, a new EC algorithm based on a motion estimation method is devised [4], which can attain both a low computational cost and a stable performance. Nevertheless, the algorithm can not find a correct motion vector (MV) when the correlation between the current and previous frames is very low. To solve this technical issue, in this paper a new EC algorithm is devised, which adopts a hybrid scheme of utilizing the temporal and spatial redundancies in video data.

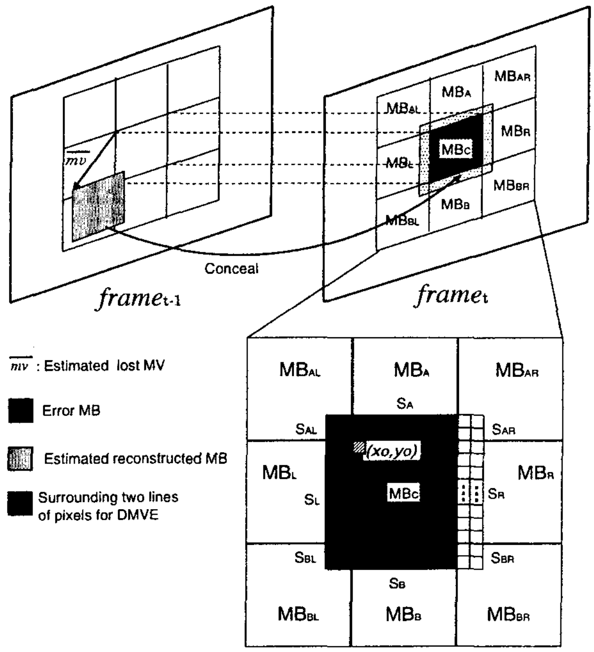


Fig. 1 Example for TEC

2 Conventional Error Concealment Algorithms

Conventionally, three kinds of EC algorithms have been proposed: the temporal error concealment (TEC), the spatial error concealment (SEC), and the hybrid error concealment (HEC) algorithms.

2.1 Temporal Error Concealment

TEC algorithms, such as Block Copy [1], BMA [2], and DMVE [3], utilize the temporal redundancy in a sequence and are intended to be applied for the pictures which have high temporal correlation with previous pictures.

The DMVE estimates the lost MV using the neighboring macroblocks (MBs), then replaces the error MB with the pixels obtained from the estimated MV (Fig. 1). Let $f_t(x, y)$ denote the pixel values of the reconstructed frame at time t, where (x, y) is the spatial coordinate. Assume that the MB_C at $frame_t$ of Fig. 1 is degraded by transmission errors. We denote the estimate of the lost MV by $\overline{mv}$, and its $x-$ and $y-$coordinates by $\overline{mv_x}$ and $\overline{mv_y}$ respectively. The lost

MV is estimated as follows.

$$\overline{mv} = argmin_{mv \in Cmv}\{BME(mv)\} \quad (1)$$

where, Cmv represents a set of candidate MVs and BME represents boundary matching error [3].

$$BME(mv) = \sum_{j \in \rho} D_j(mv), \rho = \{A, L, B, AL, BL\} \quad (2)$$

where, $D()$ represents the squared difference for the surrounding lines $(S_A, S_L, S_B, S_{AL}, S_{BL})$ between of MB_C at $frame_t$ and of the candidate MB at $frame_{t-1}$. Thus, the estimated MB ($\overline{MB}$) using an estimated MV is then

$$\overline{MB} = \sum_{i=0}^{15}\sum_{j=0}^{15}[f_{t-1}(x_0 + i + \overline{mv_x}, y_0 + j + \overline{mv_y})] \quad (3)$$

where the upper-left coordinate of the MB is (x_0, y_0) (Fig. 1). The search range of this algorithm is ±16 pixels both horizontally and vertically. This algorithm uses the set of all possible MVs within the search range, which makes its computational cost high.

2.2 Spatial Error Concealment

SEC algorithms [5]–[7] exploit the spatial redundancy within a frame. If it is given only two available MBs, above and bottom MBs, the $\overline{MB}$ by linear interpolation [7] is then

$$\begin{aligned}\overline{MB} &= \sum_{i=0}^{15}\sum_{j=1}^{15}\frac{1}{16}[(15-j)f_t(x_0+i, y_0-1) \\ &+ (j+1)f_t(x_0+i, y_0+16)] \quad (4)\end{aligned}$$

The SEC algorithms have a comparatively low computational cost. However, those algorithms do not recover high spatial details.

2.3 Hybrid Error Concealment

Since a single technique is not sufficient for an arbitrary video sequence, various HEC algorithms [6], [7] have been proposed to devise a better technique. However, the conventional HEC algorithms show poor performance for video sequences with scene changes or complex motion due to false coding mode estimation. Moreover, although the coding mode estimation is correct, the algorithms need to improve because of the low performance of employing the TEC or the SEC algorithm.

3 Proposed Algorithm

The proposed algorithm performs, alternatively, the Temporal Correlation Phase (TC Phase) as a TEC or the Spatial Correlation Phase (SC Phase) as a SEC to conceal each error MB on the basis of estimating the coding error MB types (Fig. 2).

TC Phase is characterized by low computational cost as compared with conventional methods [2], [3], which are based on the motion estimation method. On the other hand, SC Phase tends towards the low distortion of high frequency components in comparison with typical linear interpolation methods [7].

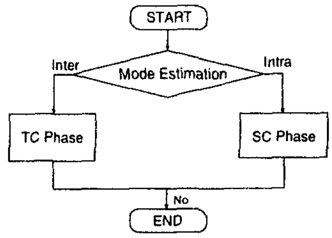


Fig. 2 Flowchart of proposed algorithm

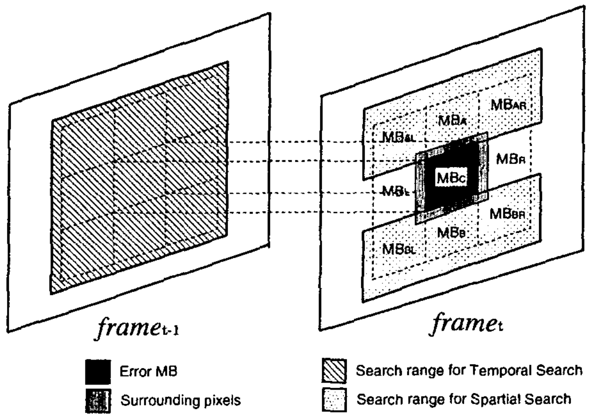


Fig. 3 Proposed algorithm

3.1 Mode Estimation

Mode estimation means to presume the coding type of an error MB to be concealed. The lost MB's type is presumed the coding type to be held a majority in the available neighboring MBs. The proposed mode estimation method has better performance than the other algorithm [6] because the proposed method uses more neighboring MBs.

3.2 TC Phase

The TC Phase utilizes the strong association between the error MV, the MV of the neighboring MB, and the previous frame MV to decrease the computational complexity by limiting the number of candidate MV to 25. This phase estimates the lost MV from the candidate MVs by reference to SAD, the sum of the absolute differences between the surrounding lines of the error MB in the current frame and those of the MB in the previous frame. Then, the error MBs are replaced by means of (3).

[MV estimation of TC Phase]

1) **Generation of candidate MVs**
MV is divided into x and y components. Let $mv[MB_\rho]_{t-1}$ be the MV of MB_ρ at $frame_{t-1}$. Using the temporal and spatial correlation, the 13 candidate MVs ($Cmv_{i,(i=0\sim12)}$) are determined as follows.

$$\begin{aligned}Cmv_{i,(i=0\sim3)} &= mv[MB_\rho]_t, \quad \rho \in \{AL, A, L, AR\} \\ Cmv_{i,(i=4\sim8)} &= mv[MB_\rho]_{t-1}, \rho \in \{AL, L, B, R, C\} \\ Cmv_{i,(i=9\sim12)} &= mv[MB_C]_{t-1} + Dmv[MB_\rho], \\ &\rho \in \{A, L, B, R\}, \quad (5)\end{aligned}$$

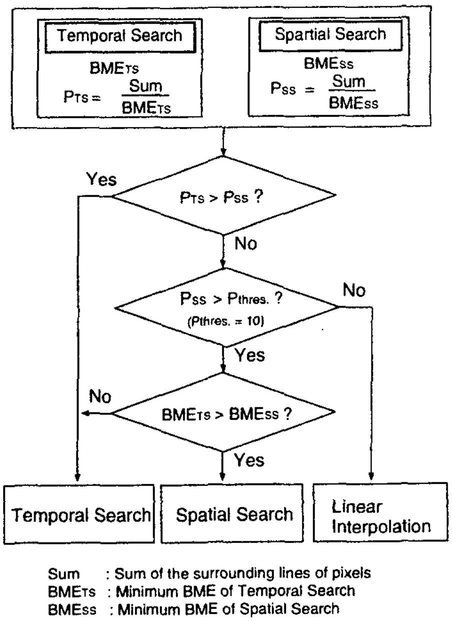


Fig. 4 Flowchart of Y component for SC phase

Table 1 Search of SC Phase

	Temporal Search	Spatial Search
Search	previous frame	current frame
Search range	−16 ∼ +16 pixels both horizontally and vertically	−16 ∼ +16 pixels horizontally, −18, +18 vertically

where,

$$Dmv[MB_\rho] = mv[MB_\rho]_t - mv[MB_\rho]_{t-1}. \quad (6)$$

Then, 5 different mv_x and mv_y are respectively selected among the 13 candidate MVs. Finally, the new 25 candidate MVs are generated by a combination of x and y components to be selected.

2) **Calculation of BME**
The lost MV is estimated using the candidate MVs by step 1). The BME for each candidate MVs is defined as follows.

$$BME(mv) = \sum_{j \in \rho} D_j(mv),$$
$$\rho = \{A, L, B, R, AL, AR, BL, BR\} \quad (7)$$

where, $D()$ represents the SAD for the surrounding lines (available S_A, S_L, S_B, S_R, S_{AL}, S_{AR}, S_{BL}, S_{BR}) between of MB_C at $frame_t$ and of the estimated MB at $frame_{t-1}$ (Fig. 1).

3.3 SC Phase

SC Phase is designed for concealing the errors on scene change or frame refresh. Since the background of the image is fairly consistent, it may be feasible to achieve an even more detailed concealment by searching the pixels around the error MB rather than just applying an interpolation method. In addition, when the refresh frame has errors, it is possible to increase concealment performance by using the pixels of the previous frame rather than by using an interpolation.

SC Phase is designed on the basis of the property of image, and is divided into the luminance (Y component) concealment and the chrominance (UV components) concealment to have robust performance. The EC procedure of the Y component is executed as shown in Fig. 4, also the details of both temporal and spatial search are shown in Table 1 and illustrated in Fig. 3. Whereas the EC of the UV components is employed by a simple linear interpolation of (4). In Fig. 4, the Y component of error MB can be accurately concealed by means of the pixels to be gained from the current or the previous frame or interpolation method.

Table 2 Simulation Features

Video sequence	Akiyo, Foreman, Silent, Merge
Video coding	MPEG-4 Simple profile@Level 1 [8]
Image format	QCIF(176 × 144)
Frame rates	10.0 fps
Bit rates	48 kbps
Bit error rate (BER)	10^{-3}, 10^{-4}, 10^{-5}

Table 3 Comparison of CPU time for EC algorithms

Frame Type	INTER			INTRA		
Number of Error MB	10	20	30	10	20	30
Block Copy [1]	1.00	1.05	1.07	1.00	1.00	1.00
BMA [2]	1.63	2.10	2.59	1.43	1.79	2.13
DMVE [4]	2.24	3.22	4.22	1.91	2.43	3.38
Interpolation [7]	1.07	1.10	1.12	1.05	1.05	1.11
Proposed	1.07	1.10	1.12	1.26	1.50	1.75

4 Simulation Results

Computer simulations have been performed to evaluate the performance of the proposed algorithm. Table 2 shows the simulation features, "Merge" sequence, which is to evaluate the EC performance of scene change, consists of a "Foreman" sequence, a "Akiyo" sequence, and a "Silent" sequence.

Table 3 represents the normalized CPU time which is the ratio of the decoding time required for each algorithm to the decoding time for an undamaged frame. The result indicates that the proposed algorithm has comparably low computational complexity.

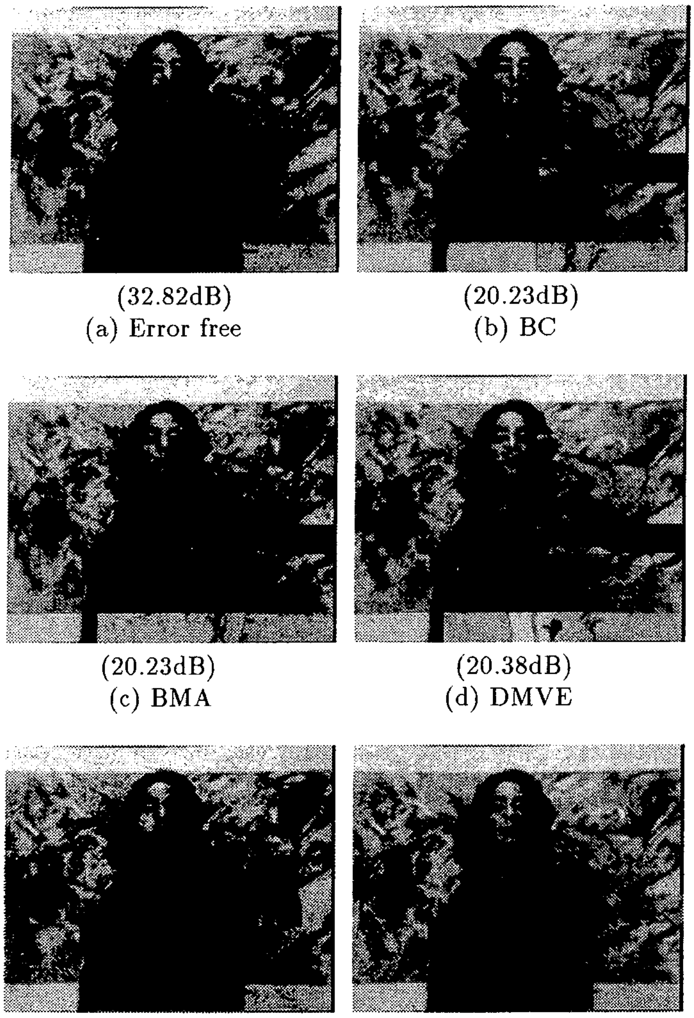
Tables 4 and 5 show the average of the Y component PSNR (peak signal-to-noise ratio) in the decoded images at various bit error rate (BER). In these Tables, the proposed algorithm provides better performance objectively than the conventional ones for all the test sequence. For example, the proposed algorithm provides about 5.0 dB better PSNR than the conventional algorithms on the "Merge" sequence at a BER of 10^{-4}.

Table 4 Comparison of decoded sequence quality for EC algorithms (unit : dB)

Sequence	Akiyo			Foreman		
Error free	39.05			31.24		
BER	10^{-3}	10^{-4}	10^{-5}	10^{-3}	10^{-4}	10^{-5}
Block Copy [1]	33.39	36.41	38.77	24.32	27.23	30.66
BMA [2]	28.78	31.56	36.75	23.82	27.24	30.53
DMVE [4]	30.14	34.15	37.92	25.99	27.91	30.78
Interpolation [7]	26.00	27.56	35.43	19.27	23.22	29.59
Proposed	33.51	36.42	38.77	26.75	28.58	30.94

Table 5 Comparison of decoded sequence quality for EC algorithms (unit : dB)

Sequence	Silent			Merge		
Error free	32.48			35.06		
BER	10^{-3}	10^{-4}	10^{-5}	10^{-3}	10^{-4}	10^{-5}
Block Copy [1]	26.56	30.29	32.31	23.18	24.90	32.51
BMA [2]	23.06	28.22	31.90	22.93	24.29	32.52
DMVE [4]	24.10	28.99	32.07	24.03	24.91	33.23
Interpolation [7]	21.80	26.99	31.56	20.80	23.46	31.44
Proposed	26.44	30.94	32.33	25.93	29.25	34.09



(32.82dB) (a) Error free

(20.23dB) (b) BC

(20.23dB) (c) BMA

(20.38dB) (d) DMVE

(27.15dB) (e) Linear interpolation

(28.88dB) (f) Proposed

Fig. 5 Reconstructed image of the "Merge" sequence (50th frame, scene change from "Akiyo" sequence)

Fig. 5 (a) shows the error free image, and Fig. 5 (b)~(f) show the restored images obtained by applying each of the algorithms respectively. These figures indicate that the proposed algorithm can improve the quality of the decoded image significantly. Consequently, the proposed algorithm has low computational complexity and stable performance for various video sequences.

5 Conclusion

A new hybrid error concealment algorithm that has low computational complexity and stable performance has been presented in this paper. In our approach, the computational cost is reduced by means of the MV correlation for concealing the inter-coded MBs, in addition, the PSNR performance is improved by the use of the modified linear interpolation for concealing the intra-coded MBs. In the simulation, the proposed algorithm has consistently stable performance and low computational cost for various video sequences. Eventually, the new approach can be combined with efficient error detection techniques [9] to further improve decoded video quality.

References

[1] Y. Wang and Q. F. Zhu: "Error control and concealment for video communication: A review," *Proc. IEEE*, Vol. 86, pp. 974–997, May 1998.

[2] W. M. Lam, A. R. Reilbman, and B. Liu: "Recovery of lost or erroneously received motion vector," in *Proc. ICASSP*, Vol. 5, pp. 417–420, Apr. 1993.

[3] J. Zhang, J. F. Arnold, and M. R. Frater: "A cell-loss concealment technique for MPEG-2 coded video," *IEEE Trans. Circuits Syst. Video Technol.*, Vol. 10, No. 4, pp. 659–665, Jun. 2000.

[4] H. S. Song, H. Okada, G. Fujita, T. Onoye, and I. Shirakawa: "Error concealment algorithm by motion estimation method for MPEG-4 video decoder," in *Proc. ITC-CSCC2001*, Vol. 1, pp. 104–107, Jul. 2001.

[5] J. W. Suh and Y. S. Ho: "Error concealment based on directional interpolation," *IEEE Trans. Consumer Electron.*, Vol. 43, No. 3, pp. 295–302, Oct. 1997.

[6] H. Sun, K. Challapali, and J. Zdepski: "Error concealment in digital simulcast AD-HDTV decoder," *IEEE Trans. Consumer Electron.*, Vol. 38, No. 3, pp. 108–116, Aug. 1992.

[7] M. H. Jo and W. J. Song: "Error concealment for MPEG-2 video decoders with enhanced coding mode estimation,"*IEEE Trans. Consumer Electron.*, Vol. 46, No. 4, pp. 962–967, Nov. 2000.

[8] ISO/IEC 14496-2: "Information technology - generic coding of audio-visual objects," *International Standard*, 1999.

[9] H. Okada, H. S. Song, G. Fujita, T. Onoye, and I. Shirakawa: "Error detection based on check marker embedding for MPEG-4 video coding," in *Proc. ITC-CSCC2001*, Vol. 1, pp. 96–99, Jul. 2001.