Science in China Series F: Information Sciences

www.scichina.com
info.scichina.com
www.springerlink.com

Error detection based on MB types

FANG Yong[1,2†], JEONG JeChang[2] & WU ChengKe[1]

[1] National Key Lab on ISN, Xidian University, Xi'an 710071, China;
[2] College of Electronics and Communications Engineering, Hanyang University, Seoul 133791, Korea

This paper proposes a method of error detection based on macroblock (MB) types for video transmission. For decoded inter MBs, the absolute values of received residues are accumulated. At the same time, the intra textural complexity of the current MB is estimated by that of the motion compensated reference block. We compare the inter residue with the intra textural complexity. If the inter residue is larger than the intra textural complexity by a predefined threshold, the MB is considered to be erroneous and errors are concealed. For decoded intra MBs, the connective smoothness of the current MB with neighboring MBs is tested to find erroneous MBs. Simulation results show that the new method can remove those seriously-corrupted MBs efficiently. Combined with error concealment, the new method improves the recovered quality at the decoder by about 0.5—1 dB.

MPEG-4, video transmission, error detection, error concealment

Most of current video coding standards (such as MPEG-x and H.26x) are based on the framework of motion estimation (ME), discrete cosine transform (DCT) and variable-length coding (VLC). At the same time of improving coding efficiency, this framework makes video bit stream sensitive to transmission errors. On one hand, the usage of VLC causes spatial propagation of errors; on the other hand, the usage of motion compensation further causes temporal propagation of errors. Although the application of channel codes improves the error robustness of bit stream significantly, random bit errors in bit stream are still unavoidable. Hence, coding-level error control techniques (such as error-resilient coding, error concealment, etc.) are necessary. Error-resilient coding refers to enhancing error resilience of bit stream by introducing redundancy, such as inserting resynchronization markers[1], using reversible VLC (RVLC)[2], partial backward decodable bit steam (PBDBS)[3] and error-resilient entropy coding (EREC)[4], etc. Error concealment refers to conceal errors and alleviate the impact of errors on image quality through some approaches if there is any error in bit stream.

Due to the usage of VLC, when errors are found in bit stream, the decoder usually fails to determine the exact positions of errors because errors may happen anywhere before. Hence, before

Received August 24, 2007; accepted October 14, 2007
doi: 10.1007/s11432-008-0029-x
†Corresponding author (email: yfang79@gmail.com)
Supported by the BK21 Foundation of Korea and the National Natural Science Foundation of China (Grant Nos. 60532060, 60672117 and 60607010)

any error concealment technique can be applied, error positioning is important. It is error detection. The process of error detection can be divided into two steps: grammar-based detection and content-based detection. The former is used to find whether there is any error in bit stream; if any, the latter is used to deduce the positions of errors. Content-based detection can be executed in frequency domain or pixel domain. In refs. [5, 6], pixel differences between two neighboring lines are used to detect transmission errors in pulse code modulation (PCM) and differential PCM (DPCM) coded images. Mitchell and Tabatabai[7] proposed a method which can detect the damage to a single DCT coefficient. However, this method is based on the framework of DCT and fixed-length codes (FLC). Lam and Reibman[8] proposed a frequency-domain method of error detection. However, all these methods are not suited to the present hybrid ME, DCT, and VLC framework. In addition, these methods are of high computational complexity and do not exploit inter-frame correlations in video signal.

In recent standards, such as H.263, MPEG-2 and MPEG-4, error detection is operated usually in transmission level[9,10]. For example, video bit stream is packetized into real-time protocol (RTP) packets, and each RTP packet is stamped with an index. Hence, the receiver can detect errors in video bit stream by extracting packet indexes. However, such approaches need the support from bottom-layers and may be unsuitable to some cases.

This paper proposes a content-based method of error detection which works in pixel domain. The method sums up the absolute values of the residues of decoded inter MBs and uses textural complexities of motion-compensated reference blocks to predict that of current MBs. Inter residues are compared with intra textural complexities to find erroneously-decoded MBs.

1 Main methods for error detection

During the decoding process, the decoder tries to find errors in bit stream by parsing grammars of video bit stream. When any of the following cases happens, the decoder affirms that errors happen in bit stream: 1) Illegal VLCs are received; 2) more than 64 DCT coefficients are decoded in one block; 3) for inter MBs, motion vectors are outside the predefined scope.

When any error is found, the decoder stops decoding and searches for the next resynchronization marker directly. In this case, MBs in the video packet can be divided into three classes: correctly-decodable MBs, erroneously-decodable MBs, and undecodable MBs. There are three classes of approaches to dealing with erroneous packets: all MBs are concealed (disregarding of whether any MB is decodable or not); only undecodable MBs are concealed (disregarding of whether any decoded MB is erroneous or not); correctly-decodable MBs are distinguished from erroneously-decodable MBs, and then erroneously-decodable MBs and undecodable MBs are concealed. Obviously, the third way is the best choice.

Errors may happen anywhere of bit stream, such as MB types, motion vectors, transform coefficients, etc. Due to the usage of data partitioning (DP), header information and motion vectors are usually put in different packets from transform coefficients and provided with higher protection using forward error correct (FEC). For this reason, in this paper, only error detection of transform coefficients is considered, while motion vectors and header information are assumed to be received correctly. As for how to detect and conceal errors in MB types and motion vectors (MVs), please refer to refs. [11, 12].

Usually, errors in alternate-current (AC) components are much more difficult to detect than those in direct-current (DC) components. Hence, our method will be limited to error detection of

AC components. In addition, because inter MBs occupy a major proportion of video bit stream, our method emphasizes on error detection of inter MBs.

2 Error detection based on macroblock types

2.1 Review on encoding process

The original image (P frame) is divided into 16×16 MBs. First, motion estimation is made to find the optimal prediction of each MB so that the sum of absolute differences (SAD) between the current MB and the reference block is minimized,

$$\mathrm{SAD}(x,y)=\sum_{i=0,\,j=0}^{15,15}\left|\mathrm{orig}(i,j)-\mathrm{ref}(i+x,j+y)\right|, \tag{1}$$

where orig(i, j) is the current MB, and ref (i+x, j+y) is the reference block. The (x, y) resulting in the smallest SAD(x, y) is the motion vector of the current MB. The minimum of SAD(x, y) is denoted as $\mathrm{SAD}_{\mathrm{orig}}$. Then, the DC component of the current MB is calculated as follows:

$$\mathrm{MB\_DC}=\sum_{i=0,\,j=0}^{15,15}\mathrm{orig}(i,j)/256. \tag{2}$$

Then, the intra textural complexity of the current MB is calculated as follows:

$$\mathrm{MB\_Comp}=\sum_{i=0,\,j=0}^{15,15}\left|\mathrm{orig}(i,j)-\mathrm{MB\_DC}\right|. \tag{3}$$

Finally, MB_Comp is compared with $\mathrm{SAD}_{\mathrm{orig}}$. If

$$\mathrm{MB\_Comp}<(\mathrm{SAD}_{\mathrm{orig}}-C), \tag{4}$$

where C is a const and set to 512 in MPEG-4 verification model, then intra mode is chosen, otherwise inter type is chosen. For inter MB, the residue between the current MB and the optimal reference block will be transformed, quantized, and coded. The coded residue is denoted as $\mathrm{diff}_{\mathrm{enc}}(i,j)$. We sum up the absolute values of $\mathrm{diff}_{\mathrm{enc}}(i,j)$ and denote the sum as $\mathrm{SAD}_{\mathrm{enc}}$,

$$\mathrm{SAD}_{\mathrm{enc}}=\sum_{i=0,\,j=0}^{15,15}\left|\mathrm{diff}_{\mathrm{enc}}(i,j)\right|. \tag{5}$$

Because it is lossy coding, $\mathrm{SAD}_{\mathrm{enc}}<\mathrm{SAD}_{\mathrm{orig}}$. In addition, $\mathrm{SAD}_{\mathrm{enc}}$ decreases with the increase in quantization level Q_p. Figure 1 shows when the 2nd frame of "Foreman" (QCIF) is encoded, the impact of Q_p on $(\mathrm{SAD}_{\mathrm{orig}}-\mathrm{SAD}_{\mathrm{enc}})$ (Figures 2—4 are similar). In Figure 1, the x-axis is the index of each MB and the y-axis is $(\mathrm{SAD}_{\mathrm{orig}}-\mathrm{SAD}_{\mathrm{enc}})$ of each MB. We find that with the increase in Q_p, $(\mathrm{SAD}_{\mathrm{orig}}-\mathrm{SAD}_{\mathrm{enc}})$ increases significantly.

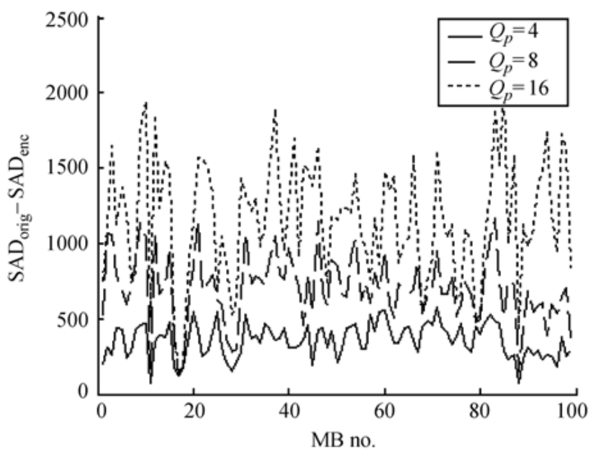

Figure 1 Impact of quantization level on $\mathrm{SAD}_{\mathrm{enc}}$.

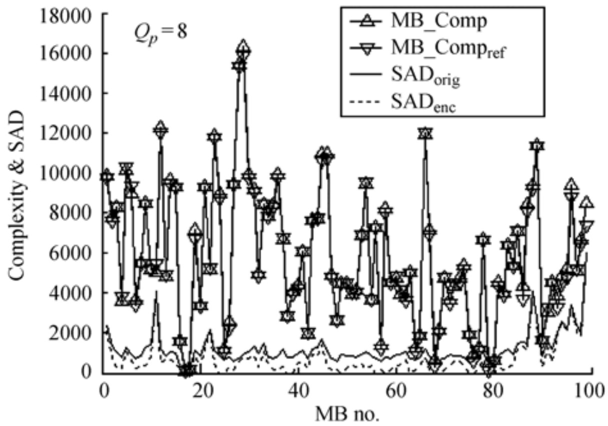

Figure 2 Comparison of MB_Comp_{ref} and MB_Comp.

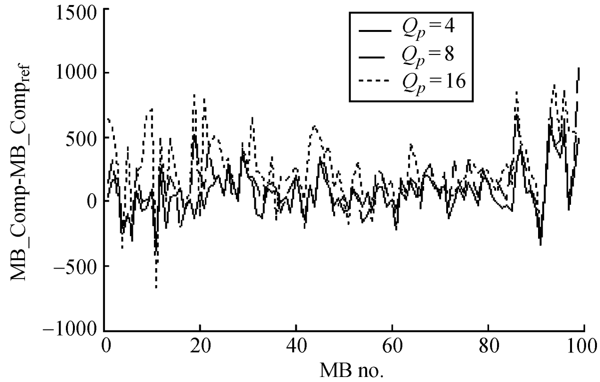

Figure 3 Impact of quantization level on MB_Comp_{ref}.

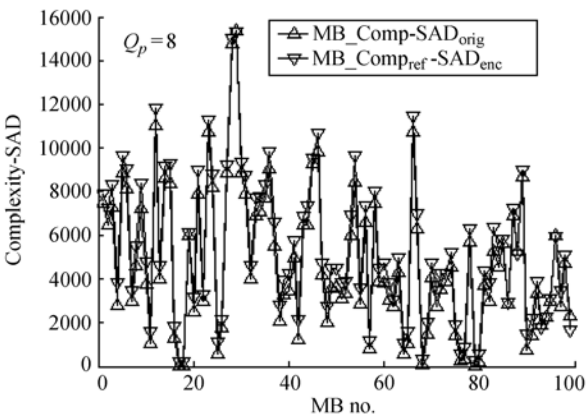

Figure 4 Comparison of predicted difference and actual difference (between complexity and SAD).

2.2 Error detection of MBs

At the decoder, assume the current MB is decodable and in inter type. After inverse DCT (IDCT), we obtain the residue $\mathrm{diff}_{dec}(i, j)$ (= $\mathrm{diff}_{enc}(i, j)+\mathrm{err}(i, j)$, where $\mathrm{err}(i, j)$ is transmission error). We calculate the sum of absolute values of $\mathrm{diff}_{dec}(i, j)$ (denoted as SAD_{dec})

$$\mathrm{SAD}_{\mathrm{dec}} = \sum_{i=0,\,j=0}^{15,15} \left|\mathrm{diff}_{\mathrm{dec}}(i,j)\right|. \tag{6}$$

If the transmission is error-free, then $\mathrm{SAD}_{\mathrm{dec}} = \mathrm{SAD}_{\mathrm{enc}}$. However, due to transmission errors, $\mathrm{SAD}_{\mathrm{dec}}$ may be different from $\mathrm{SAD}_{\mathrm{enc}}$. The difference between $\mathrm{SAD}_{\mathrm{dec}}$ and $\mathrm{SAD}_{\mathrm{enc}}$ is decided by $\mathrm{err}(i, j)$:

1. If high frequency (HF) components of the residue of the current MB are lost, $\mathrm{err}(i, j)$ will make $\mathrm{diff}_{\mathrm{dec}}(i, j)$ flatter, then $\mathrm{SAD}_{\mathrm{dec}} < \mathrm{SAD}_{\mathrm{enc}}$.

2. If extra HF components are decoded, $\mathrm{err}(i, j)$ will make $\mathrm{diff}_{\mathrm{dec}}(i, j)$ more fluctuate, then $\mathrm{SAD}_{\mathrm{dec}} > \mathrm{SAD}_{\mathrm{enc}}$.

These two cases have different impacts on reconstructed images. For case 1, the impact is often trivial and can be ignored because human eyes are not sensitive to HF components. In addition, error concealment works poorly in this case. However, for case 2, reconstructed images may be seriously degraded, especially when there are extra HF components of large magnitudes. For this reason, our interest will be focused only on case 2 hereinafter.

Now, we recall the decision process of MB types. When inequation (4) holds, intra type is chosen. Intuitively, we substitute $\mathrm{SAD}_{\mathrm{dec}}$ in inequation (4). When there is no transmission error, then

$$\mathrm{SAD}_{\mathrm{dec}} < (\mathrm{MB\_Comp} + C). \tag{7}$$

If transmission errors cause extra HF components, $\mathrm{SAD}_{\mathrm{dec}}$ will increase so that inequation (7) may not hold any more. In this case, we can affirm that the MB is in error (because in the case of large inter residue, a wise encoder should have coded the MB in intra type). As for the case that inequation (7) still holds even though there are errors and extra HF components in bit stream, transmission errors are not very serious and can be ignored. Hence, this case will not be discussed hereinafter.

However, because the decoder is unaware of the exact current MB, MB_Comp cannot be calculated. To get around this problem, the motion compensated reference block of the current MB, MC_ Ref (i, j), is used to estimate MB_Comp. We denote the intra textural complexity of MC_ Ref(i, j) as $\mathrm{MB\_Comp}_{\mathrm{ref}}$. Simulation results show that $\mathrm{MB\_Comp}_{\mathrm{ref}}$ is a good approximate to MB_Comp. From Figure 2, we find that $\mathrm{MB\_Comp}_{\mathrm{ref}}$ fits MB_Comp well. Different from $(\mathrm{SAD}_{\mathrm{orig}} - \mathrm{SAD}_{\mathrm{enc}})$, $(\mathrm{MC\_Comp} - \mathrm{MC\_Comp}_{\mathrm{ref}})$ is around zero and departs from zero slightly with the increase in Q_p (Figure 3).

On Substituting $\mathrm{MB\_Comp}_{\mathrm{ref}}$ in inequation (7), we obtain

$$\mathrm{SAD}_{\mathrm{dec}} < (\mathrm{MB\_Comp}_{\mathrm{ref}} + C'), \tag{8}$$

where C' is a const. When inequation (8) does not hold, the MB is considered to be erroneous. The rationality of inequation (8) stems from the assumption that given by $\mathrm{SAD}_{\mathrm{orig}} < (\mathrm{MB\_Comp} + C)$, then $\mathrm{SAD}_{\mathrm{enc}} < (\mathrm{MB\_Comp}_{\mathrm{ref}} + C')$. To verify the assumption and determine C', we give simulations. From Figure 4, we can find that $(\mathrm{MB\_Comp}_{\mathrm{ref}} - \mathrm{SAD}_{\mathrm{enc}})$ fits $(\mathrm{MB\_Comp} - \mathrm{SAD}_{\mathrm{orig}})$ very well. This result justifies the above assumption. At the same time, we set $C' = C$.

It is valuable to point out that three methods dealing with damaged video packets (discard all MBs; remain all decodable MBs; remain only correctly-decodable MBs) can seem as special cases of inequation (8). If we let $C' = -\infty$, inequation (8) will never hold, then all MBs in the video packet are considered to be erroneous, which corresponds to the case that all MBs are discarded in damaged video packets. On the contrary, if we let $C' = \infty$, inequation (8) will hold always, then all decodable MBs in the video packet are considered to be correct, which corresponds to the case that all MBs in damaged video packets remained.

2.3 Implementation of the proposed method

To use our method, a prerequisite is that MC_Ref(i, j) should be received correctly. It means that in all preceding frames (in the same group-of-picture (GOP)), neighboring pixels around the current MB are received correctly. Due to the randomness of errors, the probability that errors happen in neighboring areas of consecutive frames is not high, hence this prerequisite can seem basically effective. For intra MBs, due to the lack of reference blocks, in our simulation, the DC component of the current MB is compared with that of spatially neighboring MBs. If there is a large difference, the current MB is considered to be erroneous. In addition, edge pixels are used to test the connective smoothness of the current MB with neighboring MBs. If there is a large difference between edge pixels of the current MB and neighboring MBs, the current MB also seems to be erroneous. For erroneous intra MBs, intra interpolation is used to conceal errors. As for MB types and MVs, due to their low bandwidth, we use DP plus unequal error protection (UEP) to transmit them.

3 Simulation results and discussions

To achieve better performance, error detection should be integrated with error concealment in practice. Many methods for error concealment have been proposed, such as maximally smooth image recovery (MSR) proposed by Wang[13], spatial interpolation using projections onto convex sets (POCS) proposed by Sun[14], etc. These methods are of high computational complexity. In our simulation, error concealment is simple. For intra MBs, intra interpolation based on weighted pixel averaging is used for error concealment. The weight used for averaging is the inverse of the distance between the source and destination pixels. For inter MBs, the motion compensated reference blocks are copied directly to conceal damaged MBs.

MPEG-4 coding program is used in our simulation. The size of GOP is 50, i.e., after every 50 frames, one I frame is inserted. Each video packet includes one row of MBs. To illustrate the performance of our proposed method better, I frames, header information, and motion vectors are placed in different packets from DCT coefficients of P frames. Three standard QCIF image sequences: “Foreman”, “Coastguard” and “Container” are simulated under 6 kinds of different channel bit-error-rates (BERs) (shown in Table 1).

Table 1 Objective comparison of three classes of methods

BER	Foreman			Coastguard			Container		
	NED	NER	ED	NED	NER	ED	NED	NER	ED
0		34.92			33.76			34.66	
1.35E-3	26.23	26.01	26.71	24.34	23.59	24.63	29.51	29.05	29.68
1.24E-3	26.56	26.32	26.99	24.52	23.80	24.81	29.72	29.18	29.85
9.21E-4	27.47	27.18	27.79	25.11	24.59	25.46	30.36	29.79	30.52
8.77E-5	33.13	33.00	33.21	30.44	30.11	30.67	33.80	34.01	33.89
1.12E-4	32.97	32.87	33.10	29.96	29.76	30.19	33.67	33.38	33.71
8.99E-4	27.52	27.20	27.81	25.14	24.66	25.52	30.40	30.40	30.61
Average	28.98	28.76	29.27	26.59	26.09	26.88	31.24	30.97	31.38

Once a bad video packet is found, the decoder tries three kinds of different schemes. In scheme 1, all MBs in this video packet are discarded and concealed. It is referred to as NED scheme. In scheme 2, all decodable MBs remained and only undecodable MBs are concealed. It is referred to as NER scheme. In scheme 3, the proposed method is executed to find damaged MBs. Only correctly decoded MBs remained, while both undecodable MBs and erroneously decoded MBs are concealed. It is referred to as ED scheme. The objective results are included in Table 1, where it shows the average luminance peak-signal-noise-ratio (PSNR) (in dB). From Table 1, we can find clearly that recovery image quality is improved significantly using our proposed method. Among these methods, method 2 shows the worst performance, meaning that in most of the cases, erroneously-decoded MBs degrade recovery quality even more seriously than undecodable MBs. Compared with method 2, about 0.5—1 dB luminance PSNR improvement can be found using method 3. Compared with method 1, the average PSNR of method 3 is also improved by about 0.3 dB.

4 Conclusions

In this paper, we propose a method of error detection which uses MB types and inter-frame information. Combined with error concealment, this method can remove those MBs with serious errors efficiently and improve recovery quality significantly. Compared with other present methods, our method is of low computational complexity and easily operated in practical.

1 Ferguson T J, Ranowitz J H. Self-synchronizing Huffman codes. IEEE Trans Inf Theory, 1984, 30(4): 687—693[DOI]

2 Takishima Y, Wada M, Murakami H. Reversible variable length codes. IEEE Trans Commun, 1995, 43(2): 158—162[DOI]

3 Gao S S, Tu G F. Robust H.263+ video transmission using partial backward decodable bit stream. IEEE Trans Circ Syst Vid, 2003, 13(2): 182—187[DOI]

4 Redmill D W, Kingsbury N G. The EREC: An error resilient technique for coding variable-length blocks of data. IEEE Trans Image Proc, 1996, 5(4): 565—574[DOI]

5 Ngan K N, Stelle R. Enhancement of PCM and DPCM images corrupted by transmission errors. IEEE Trans Commun, 1982, 30(1): 257—265[DOI]

6 Rose K M, Heiman A. Enhancement of one-dimensional variable-length DPCM images corrupted by transmission errors. IEEE Trans Commun, 1989, 37(4): 373—379[DOI]

7 Mitchell O R, Tabatabai A J. Channel error recovery for transform image coding. IEEE Trans Commun, 1981, 29(12): 1754—1762[DOI]

8 Lam W M, Reibman A. An error concealment algorithm for images subject to channel errors. IEEE Trans Image Process, 1995, 4(5): 533—542[DOI]

9 Fan C, Cui H J, Tang K. Chessboard-interpolation-based multiple description video coding. Sci China Ser F-Inf Sci, 2004, 47(1): 66—74

10 Wang G J, Zhang Q, Zhu W W, et al. An end-to-end robust approach for scalable video over the Internet. Sci China Ser F-Inf Sci, 2004, 47(2): 246—259

11 Hong M C, Kondi L, Scwab H, et al. Video error concealment techniques. Signal Process-Image, 1999, 14(6): 437—492

12 Sun H, Challapali K, Zdepski J. Error concealment in digital simulcast AD-HDTV decoder. IEEE Trans Consum Electr, 1992, 38(3): 108—117[DOI]

13 Wang Y, Zhu Q F, Shaw L. Maximally smooth image recovery in transform coding. IEEE Trans Commun, 1993, 41(10): 1544—1551[DOI]

14 Sun H, Kwok W. Concealment of damaged block transform coded images using projections onto convex sets. IEEE Trans Image Process, 1995, 4(4): 470—477[DOI]