

DCT Coefficient-Based Error Detection Technique for Compressed Video Stream.

K. Bhattacharyya, H. S. Jamadagni

Centre for Electronics Design & Technology

Indian Institute of Science, Bangalore, India

E-mail: (bkaushik, hsjam)@cedt.iisc.ernet.in

Abstract: In the process of transmission of digital encoded video bit streams over physical network, the data is corrupted by random bit errors, which has a several effect on the decoded video quality. The success of an error concealment algorithm depends heavily on the detection of the corrupted regions. In this paper we propose a new error detection algorithm for discrete cosine transform based video coding. Our algorithm detects so far undetectable erroneous regions with a very little increase in computational complexity. We show that erroneous blocks in the decoded frames can be detected monitoring the variation of average energy from block to block. To avoid false detection due to the presence of strong edges we employed a very fast DCT coefficient based edge detection algorithm. Computer simulation shows that the quality of the recovered image is significantly improved in comparison to the well-known methods.

1. Introduction

Several international standards for video coding, such as H.261 [1], MPEG-1 [2] and MPEG-2 [3] compress the video data by removing the correlations by a discrete cosine transform (DCT), a motion compensation (MC), and a variable length coder (VLC). However, as the data are highly compressed, the effects of bit errors and cell losses become serious. Several error concealment [EC] schemes have been proposed in the literature to restore the degradation in the decoded video caused by errors. Wang and Zhu [4] provide a survey of existing EC techniques.

EC techniques attempt to fill in the corrupted area by utilizing the redundancies in the received video data. In EC the detection of erroneous blocks is an important problem. Most of the EC techniques were developed and tested in a network environment, such as asynchronous transfer mode, where it is assumed that the ATM adaptation layer (AAL) detects the cell loss. However for applications such as wireless communication where bit error is more critical than cell loss, an error detection algorithm (EDA) is essential.

Bit error in the data stream may generate two types of error. It may change a fixed length synchronization code word or a variable length code word. A few error detection algorithms have been proposed in literature [5]-[7]. There are two layers in these error detection algorithms. The first layer detects syntax violation and synchronization failures. The second layer checks the decoded data for inconsistencies. The

second layer is needed because all errors in a VLC data do not always lead to syntax or synchronization violation. An error can change a valid VLC into another valid but incorrect VLC of a different length. The decoder can not recognize an error at this point. However synchronization with the bit stream is lost and it is likely that VLCs at later stage in the bit stream will be discovered to be illegal. An error also may change a valid VLC to another valid but incorrect VLC of the same length. Decoder can not recognize the error but all the subsequent differentially coded data will be corrupted till the next synchronization code.

In all the proposed algorithms, after a synchronization failure or an invalid code word is detected, the group of blocks [GOB in H.261] or slice [MPEG-1 & MPEG-2] is marked as erroneous, the data stage finally identifies the blocks that are corrupted. In this stage various statistics of the decoded blocks are utilized. Keck [5] suggested to start concealment about two macroblocks to the left of the first detected area. Park *et al.* [6] proposed a boundary vector comparison method to detect incorrectly decoded macroblocks. Chu and Leou [7] used four statistical parameters of a block and its neighborhood to detect a corrupt block.

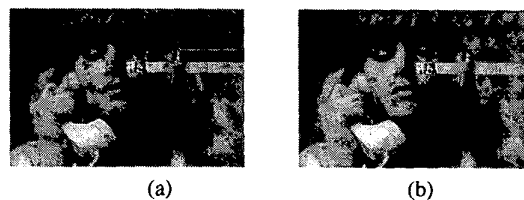


Figure 1: (a) Decoded Intra-coded picture without using error detection. (b) Same frame decoded with primary error detection and concealment.

As shown in Figure 1, the above algorithms fail to detect errors, which do not cause a synchronization failure or generate syntax errors. These errors predominantly occur in the intra-coded frames (I-frames in MPEG-1 & MPEG-2), which cause a visual distortion that can last a few seconds. Also, these algorithms are computationally expensive. In addition, in the presence of strong vertical edges these algorithms mark a correctly decoded block as a corrupted one.

In this paper we propose an algorithm for efficient detection of corrupted blocks. We utilized the properties of the DCT coefficient of each block to calculate the average luminance values, and to detect the presence of edges. We

show that our method of error detection considerably improves the performance of the decoder.

The paper is organized as follows. Section 2 presents some fundamental properties of the DCT coefficients, which are utilized for our work. Section 3 describes proposed error detection algorithm. Section 4 presents the results. Section 5 concludes the paper.

2. Discrete Cosine Transform (DCT)

In DCT based compression schemes the image is divided into 8×8 blocks and DCT is performed on each of these blocks. The DCT operation can be represented as

$$F_{uv} = \frac{C_u C_v}{4} \sum_{i=0}^7 \sum_{j=0}^7 \frac{\cos(2i+1)u\pi}{16} \frac{\cos(2j+1)v\pi}{16} f(i, j) \quad (1)$$

where

$$C_u, C_v = \begin{cases} \frac{1}{\sqrt{2}} & u, v = 0 \\ 1 & \text{otherwise} \end{cases} \quad (2)$$

Each DCT coefficient is a linear combination of all pixel values within the block. There is a relationship between the pixel values and its DCT coefficients, e.g. the DC DCT coefficient represents the average energy of the block. The value of each AC coefficient reflects the variation in gray level values in certain directions at a certain rate [8].

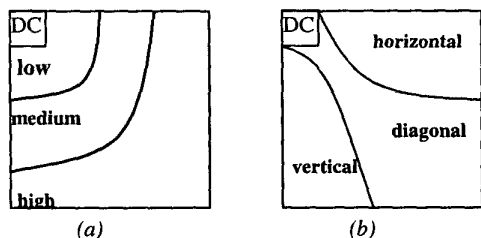


Figure 2: (a) Frequency distribution and (b) block features of DCT coefficients [9].

Shen & Sethi [8] showed that the edge information from an image can be directly extracted by some simple measurements on the AC coefficients of each block in the compressed domain. They used the following relationship to determine the presence of edge in a certain block:

$$A = \sum_{u=0}^7 \sum_{v=0}^7 |F_{u,v}| \quad (u, v) \neq (0,0) \quad (3)$$

The following matrix of DCT coefficients can be used to obtain accurate edge orientation information [8]:

$$\tan \theta = \left(\sum_{v=1}^7 F_{0v} \right) / \left(\sum_{u=1}^7 F_{u0} \right) \quad (4)$$

To determine the edge orientation more coarsely the following equation can be used:

$$\tan \theta = F_{01} / F_{10} \quad (5)$$

3. DCT coefficient Based Error Detection

Our proposed error detection algorithm has three layers, namely the VLC layer, the synchronization layer, and the data layer.

The inner most layer of the decoder is the block level decoder, this layer decodes the variable length code words. This layer consists of VLC codeword and syntax sub-layer, the codewords are decoded and all invalid codeword occurrences are checked. In this layer the validity of the decoded information is also checked. If a codeword is not found in the corresponding codebook or if it is reserved, or a syntax violation is detected in the syntax sub-layer the slice or the GOB is declared to contain an error. Decoding is given up and control is handed over to the synchronization layer.

We observed that in a picture the average energy of a block does not change abruptly. The average energy of the spatially neighboring blocks are strongly correlated except in the presence of a strong edge. In the presence of a strong edge the average energy of the blocks, which falls in the edge direction, are again correlated.

When decoding in the VLC layer has been completed, control is passed to the code synchronization layer, which enforces a search to detect a byte aligned start code. This layer is also activated whenever a byte aligned start code is detected during normal decoding of the block layer. This layer is responsible for synchronization at the slice or GOB layer and initiates the evaluation of the higher layer headers and extensions.

For intra-coded pictures at the end of the decoding of the frame the control is passed to data layer which checks individual blocks of the picture for possible anomalies. In this layer the average energy change from block to block is monitored. We compared the DC DCT coefficient of the luminance part of each block with its immediate spatial neighbors (four connected). If the energy variation is beyond a certain threshold (say T_1) the neighborhood blocks are checked for the presence of edge.

To detect edge activities in the eight-connected neighborhood of the suspected block, we applied the edge detection algorithm of equation (3) in a modified form, where the summation extends only over the first five coefficients. Simulation results show that the first five coefficients contains

enough information to detect coarse edges in a block. One specific goal of our algorithm is to detect straight edges. Therefore we search in the neighboring up and down blocks for high and medium frequency components. If we find the presence of strong edges in the neighborhood we extract the edge direction applying equation (5). From the edge information thus obtained we again compare the energy along the edge direction, if again the energy variation is beyond certain threshold T_2 , the block is detected erroneous and subsequent concealment techniques are applied to it.

We applied the proposed error detection algorithm for I frame only, because we observed from simulation results that the probability of undetectable VLC errors are more for I frame than P or B frames. Also error in an I frame has more visual effect on the whole sequence than the error in P and B frame due to temporal propagation of errors.

In terms of computational complexity our algorithm is more attractive, because to check the average intensity of a block we need only four comparisons, as the DC DCT coefficient for each block is readily available. For detecting edge activity for a block we need only five additions, and to determine edge orientation in a block in the worst case we need fourteen additions and one division (eqn. 4). Whereas, in pixel domain methods we need sixty-four additions to calculate the average intensity of a block. The detection of edge activity is generally done by computing the following measured defined by Moravec [11], which is actually the variance of a window of size $(2a+1)$ by $(2b+1)$ centered at pixel (x, y) in an image $f(x, y)$.

$$V(x, y) = \sqrt{\sum_{k=-a}^a \sum_{l=-b}^b [f(x, y) - f(x-k, y-l)]^2} \quad (6)$$

In pixel domain, detection of edge orientation of a block can be done by the method proposed by Kwok and Sun [12], where they applied a 3×3 Soebl operator on each pixel of a block, with a appropriate voting mechanism to select the most prominent edge orientation. These pixel domain methods are computationally much more expensive than our proposed method.

Our proposed algorithm can also be applied to P and B frames. But it would be more computationally expensive. Due to the use of motion compensation the DCT coefficients are not directly available for P and B blocks. Though we can estimate them by using the motion information [10], the increased computational complexity is not worth the improvement in picture quality.

4. Simulation Results

Our proposed error detection algorithm provides better results in computational complexity, accuracy and picture quality. As the intra-coded frames anchor to several other predicted frames, minimization of errors in these frames stops temporal propagation of error. We tried our algorithm on the MPEG-1 sequence 'Susie', coded at a bit rate of 1.5 Mbps with average PSNR of 37 dB.

By extensive simulations, we observed that a value of 150 for T_1 , 50 for T_2 give optimal performance. To evaluate

the performance of the proposed approach, we have disturbed only the I pictures of the sequence by uniformly distributed random bit errors with different bit error rates (BER).



(a)



(b)

Figure 3: (a) Frame 63 decoded with simple error detection and copy concealment. (b) Same frame decoded with proposed error detection and copy concealment

The peak signal-to-noise-ratio (PSNR) in dB is employed in this study as the objective performance measure for the Y component of the 'Susie' images. In Figure 5 PSNR performance of the proposed error detection algorithm is compared with the simple error detection technique where errors are detected by only syntax and semantics violation without any block by block data consistency check.

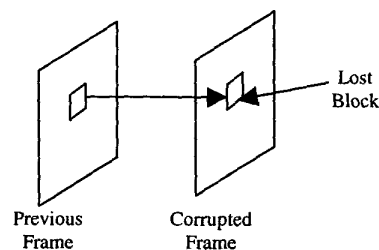


Figure 4: Simple Copy Concealment.

We have used simple temporal copying concealment algorithm to conceal the detected erroneous regions. As shown in Figure 4, in copy concealment technique, the detected erroneous blocks are concealed by copying from the collocated blocks from the temporal neighborhood. In Figure 6, we present the PSNR's of 100 frames of 'Susie', when the BER is 1×10^{-4} . We demonstrate the subjective quality of the decoded images in Figure 3. Figure 3(a) shows a intra-coded frame of 'Susie' sequence, decoded and concealed using usual detection

and concealment algorithm. Figure 3(b) shows the same frame decoded using our proposed detection and copy concealment techniques.

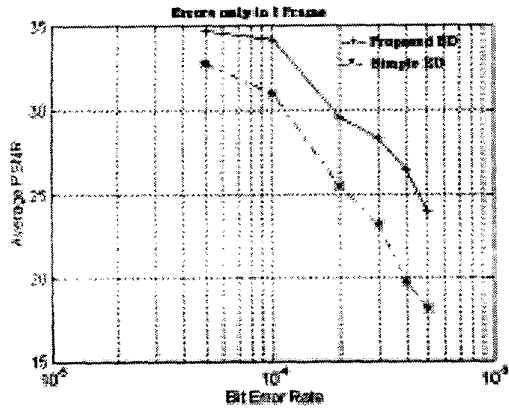


Figure 5: PSNR performance comparison for different error detection technique (Errors only in I frames) and copy concealment.

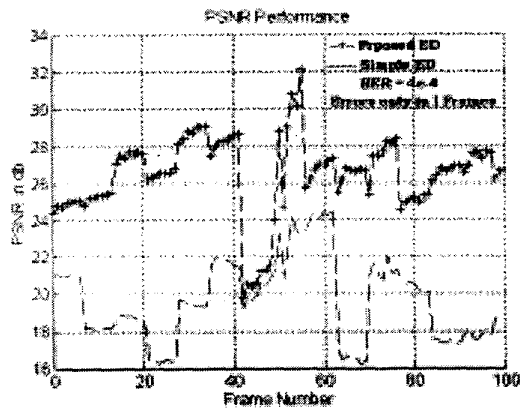


Figure 6: PSNR performance comparison for different error detection techniques (Errors in I frame only).

5. Conclusion

In this paper we presented a new generic error detection technique for DCT based video coding, such as H.261, MPEG-1, MPEG-2. In our approach the location of the corrupted blocks are detected by a multistage EDA, where we have employed a novel way to detect blocks corrupted due to error in the VLC data in intra-coded frames. Our method is computationally less expensive and accurate as it employs some novel properties of the DCT coefficients for calculation of the average energy of the blocks and edge detection. In the simulation we applied the proposed EC algorithm to the MPEG-1 decoder system, and the effects of randomly injected

bit errors in the MPEG-1 bit streams were successfully concealed. Thus we believe that the proposed EC techniques can be employed for any application with DCT based video decoder.

6. REFERENCES

- [1] "ITU-T recommendation H.261: video codec for audiovisual services at $p \times 64$ kbit/s," Tech. Rep., ITU-T, 1990.
- [2] Motion Picture Experts Group, "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbits/s MPEG-1," International Standard 11172-2, ISO/IEC/JTC1 SC29/WG11, 1994.
- [3] Motion Picture Expert Group (JTC1/SC29/WG11) and Experts Group on ATM Video coding (ITU-T SG15), "Generic coding of moving pictures and associated audio MPEG-2" Draft international Standard 13818, ISO/IEC, Nov. 1994.
- [4] Y. Wang, Q. Zhu, "Error control and concealment for video communication: A review," *Proceeding of the IEEE*, vol. 86, no. 5, May 1998, pp. 974-997.
- [5] W. Keck, "A method for robust decoding of erroneous MPEG-2 video bitstreams," *IEEE Transactions on Consumer Electronics*, vol. 42, no. 3, August 1997, pp. 411-421.
- [6] J. W. Park, J. W. Kim, "DCT coefficients recovery based error concealment technique and its application to the MPEG-2 bit stream error," *IEEE Transactions on circuits and systems for video technology*, vol. 7, no. 6, December 1997, pp. 845-853.
- [7] W. Chu, J. Leou, "Detection and concealment of transmission errors in H.261 images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 1, February 1998, pp. 74-83.
- [8] B. Shen, I. K. Sethi, "Direct feature extraction from compressed images," *SPIE vol. 2670, Storage & Retrieval for Image and Video Databases IV*, 1996.
- [9] K. R. Rao and P. Yip, "Discrete Cosine Transform - Algorithms, Advantages, Applications," *Academic Press Inc.*, 1990.
- [10] B.L. Yeo and B. Liu, "On the extraction of DC sequence from MPEG compressed video". In *Proc. of the IEEE International Conference on Image Processing*, volume 2, 1995.
- [11] D. H. Ballard and C. M. Brown, *Computer Vision*, pp. 68-70, *Prentice-Hall*, Englewood Cliffs, New Jersey, 1982.
- [12] W. Kwok, H. Sun, "Multi-Directional Interpolation for spatial error concealment," *IEEE Transactions on Consumer Electronics*, vol. 39, no. 3, August 1993.