Proceedings of the 6th WSEAS International Conference on Multimedia Systems & Signal Processing, Hangzhou, China, April 16-18, 2006 (pp7-12)

Combined Copyright Protection and Error Detection Scheme for H.264/AVC

XIAOMING CHEN, YUK YING CHUNG, FANGFEI XU,
AHMED FAWZI OTOOM, *CHANGSEOK BAE
School of Information Technologies, The University of Sydney, NSW 2006, AUSTRALIA
*Post-PC Research Group, ETRI,161 Gajeong Dong, Yuseong Gu, Daejeon, KOREA
xche2902, vchung, faxu2377,

Abstract: - This paper proposes a new copyright protection and video error detection scheme for H.264 video using information embedding technique and turbo coding. There are two categories of information to be embedded into the intra frames of H.264 video at the encoder. These are the turbo coded copyright information and the parity bit of the sums for the non-zero coefficients in 8x8 sub-blocks. These two categories of information are extracted at the H.264 decoder and both can be used to detect transmission errors in intra frames of the video. Finally, the corrupted copyright information can be automatically reconstructed by the turbo decoder. Experimental results show that the proposed scheme can efficiently protect the copyright information and detect more video errors than the H.264 codec.

Key-Words: - video compression, copyright protection, error detection, H.264/AVC, MPEG

1 Introduction

Today, the modern compression techniques make digital video in a compact size so that it is easy to be transmitted over varieties of networks such as Internet and wireless network. However typically Internet and wireless network are error-prone and they offer no QoS guarantee (packet loss and bit errors unavoidable). Also the high compression ratio for the video data introduces more sensitivity to channel noises. This makes developing efficient video error resilient methods critical. On the other hand, the digital video used in the multimedia service is usually exposed to possibilities of illegal copies, and one may not be able to distinguish the illegal copies from the original video. As a result, developing reliable copyright protection approach is needed.

This paper proposes a new hybrid approach that combines error detection and copyright protection for H.264 videos using information embedding techniques and one of the Error Control Code (ECC) – turbo code. H.264 is the denomination of ITU-T's most recent video codec recommendation, which is also known as MPEG-4 Advanced Video Codec (AVC) [1]. H.264 consists of two layers - the Video Coding Layer (VCL) and the Network Abstraction Layer (NAL). The VCL in H.264 has several built-in coding tools that enhance the error detection and resilience for the compressed video stream [1, 2, 3]. However in H.264 the error detection is mainly syntax-based and they are not sufficient to detect all kinds of errors reliably.

Typically common video error detection approaches include header based and FEC based error detection [4], data hiding based approach [5] and video feature characterizing based detection methods [6]. Data hiding or information embedding [5] is also a frequently used approach for copyright protection. By using this approach, some copyright information can be embedded into media and the information is not perceptually visible to the viewer of the media. The copyright information can only be extracted by the owner of the media using their specific algorithms. The copyright information to be embedded into media can be protected by the ECC - Error Control Code, which is an efficient tool that deals with data transmission errors. As it is self-explained, the ECC is a kind of code that has the capability of automatic error detection and correction. It is ideal for reconstructing corrupted data. The turbo code used in this work is one of the most advanced ECC developed.

In this paper a new approach has been proposed and developed for embedding turbo coded copyright information into 8x8 sub-blocks of macroblocks in H.264 video for copyright protection. This approach can detect errors in the transmitted DCT coefficients for intra frame blocks. The error detection is achieved by utilizing the embedded turbo coded copyright information and the sum of non-zero DCT coefficients for 8x8 sub-blocks. First the copyright information is pre-processed by a turbo encoder, and then the turbo coded copyright signal is embedded into the "optimal position" of some 8x8 sub-blocks of macroblocks in intra-frames. At the same time for the rest of 8x8 sub-blocks, the Sum of Non-Zero Coefficients (SNZC) within a sub-block is calculated and the parity of the sum is also embedded into the "optimal position" in that sub-block. At the decoder, the embedded turbo coded information is extracted, but not recovered immediately. The extracted signal is first used to compare with the original turbo coded copyright information, in order to detect and locate transmission errors. For the rest of sub-blocks the error detection is done by recalculating the SNZC for 8x8 sub-blocks and comparing its parity to the extracted bit. After the error detection, the extracted copyright signal is recovered by the turbo decoder, so that the original copyright information can be correctly reconstructed. Fig.1 illustrates the proposed system structure.

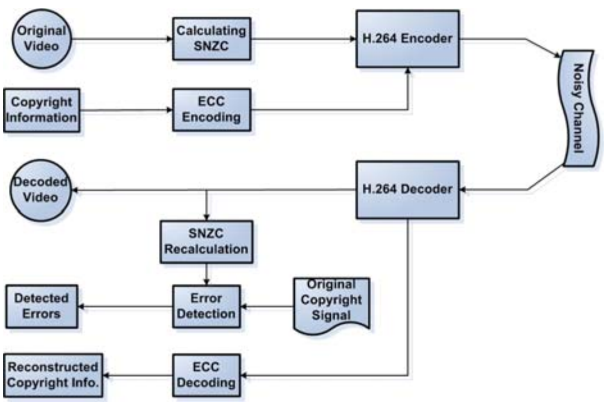

Fig.1 Proposed system structure

In this paper, Section 2 describes the turbo encoder/decoder used; Section 3 introduces the proposed copyright protection and video error detection algorithm; in Section 4, the experimental results are presented; and Section 5 comprises the conclusion.

2 Turbo Coding

Introduced by Berrou, Glavieux and Thitimajshima in 1993, turbo code [7] is a class of error-correcting codes that are widely used in the satellite communications, the third generation cellular and personal communication services. Unlike the traditional error control codes, turbo codes produce a likelihood measure for each bit to be encoded instead of producing a stream of binary digits from the received signals. The turbo codes can be encoded into three sub-blocks of bits, which are the real data to be encoded; the parity data computed using a convolutional code [8]; and the parity data for a known permutation [7] of the payload data that is also computed using a convolutional code.

A typical turbo encoder may consist of two convolutional encoders, separated by an N-bit interleaver, together with an optional puncturing mechanism. The function of the interleaver is to take each incoming block of N data bits and rearrange them into a pseudo-random fashion prior to encoding by the second convolutional encoder. The role of the puncturer is to periodically delete selected bits to reduce coding overhead. Fig.2 describes standard turbo encoder and decoder.

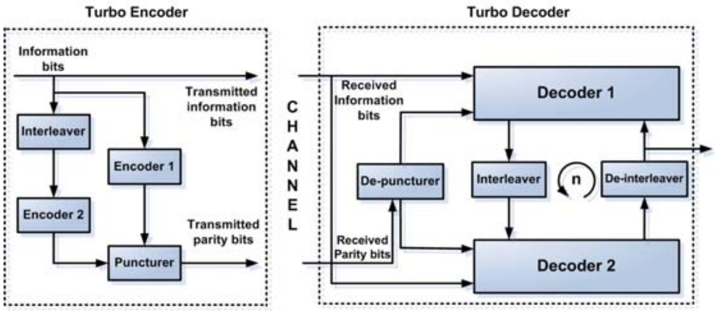

Fig.2 Turbo Encoder / Decoder

The turbo encoder/decoder used in this work is turbo(3,1), which generates ***3l*** bits turbo code for ***l*** bit information.

3 The proposed scheme

In this work, there are two classes of information to be embedded into the intra frames in H.264 video:

- The turbo encoded copyright information; and
- The parity bit calculated from the **S**um of **N**on-**Z**ero **C**oefficients (SNZC) for 8x8 sub-blocks.

For each 8x8 sub-block of a 16x16 macroblock (4x4 luma prediction mode) in H.264 video, the copyright information and the parity bit calculated from the SNZC is interlaced and embedded into one selected 4x4 sub-block of the 8x8 sub-block. The selected 4x4 sub-block will contain the "optimal position" for embedding information. The "optimal position" in an 8x8 sub-block refers to the position of the last non-zero coefficient that has the largest index number (highest frequency) in the zig-zag scanning process among the four 4x4 sub-blocks. It is expected that the embedding of the "optimal position" will not affect the quality of the video. The information embedding is done by modifying the LSB1 (the right most Least Significant Bit) of the

DCT coefficient (luma) at the optimal position. Fig.3 explains the process of choosing "optimal position":

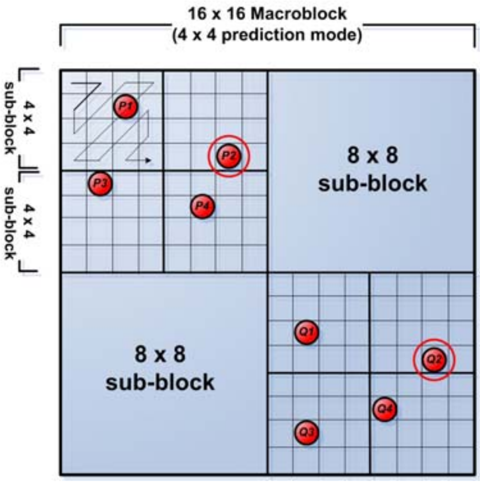

Fig.3 Optimal Position

From Fig.3, in the upper left 8x8 sub-block, the last non-zero coefficients in its four 4x4 sub-blocks are P1, P2, P3 and P4 respectively. It is obvious that the P2 is at the highest frequency position in the zigzag scanning compared with P1, P3 and P4. So in this case P2 will be the "optimal position" for embedding information. Similarly, in the bottom right 8x8 sub-block, the "optimal position" would be Q2.

In order to keep the high quality of the video, a threshold ***t*** should be set to identify the number of first few DCT coefficients that *cannot* be modified for embedding information. The choice of ***t*** values will be discussed later.

The details of the proposed approach can be described as follows:

At the H.264 encoder.

(1) For the current 8x8 sub-block of a 16x16 macroblock (if best prediction mode is 4x4), calculate and select a 4x4 sub-block that contains the "optimal position" for information embedding;

(2) Check if the index of the "optimal position" ***op*** in the zigzag scanning is greater than the threshold ***t***. If yes this position will be ready for embedding the next bit information, otherwise go to the next 8x8 sub-block and return to the step (1);

(3) Interlaced embedding of the turbo coded copyright information and the parity of the sum of non-zero coefficients: (as described in Fig.4) proceeds as follows:

a) For the $\boldsymbol{i}^{th}$ bit to be embedded, if ***i*** is odd, take one bit from ECC (turbo) processed copyright information signal, embed that bit into the "optimal position" (by modifying the LSB1); and

b) If ***i*** is even, calculate the sum of non-zero coefficients of all the four 4x4 sub-blocks within the current 8x8 sub-block, and embed the parity of the sum into the "optimal position"; and

(4) Go to the next 8x8 sub-block or macroblock, then return to step (1).

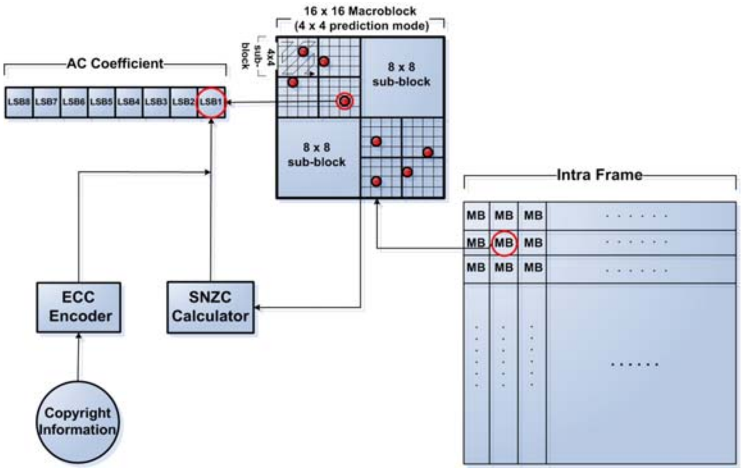

Fig.4 Information embedding process

At the H.264 decoder

(1) For the current 8x8 sub-block of a macroblock, locate the 4x4 sub-block that contains the "optimal position" ***op***. If ***op*** is greater than the threshold ***t***, prepare to read information from that position;

(2) Extracting information from the optimal position (described in Fig.5) proceeds as follows:

a) For the $\boldsymbol{i}^{th}$ bit to be extracted, if ***i*** is odd, read one bit from the current "optimal position", place it into the extracted copyright information buffer, and compare this bit to the original encoded copyright information. If they are different, then a transmission error in the current 8x8 sub-block is detected. Mark this sub-block as "dirty";

b) If ***i*** is even, read one bit from the "optimal position". This bit is the parity of the calculated sum of non-zero coefficients in the current 8x8 sub-block. Then at the decoder the SNZC is recalculated, and compared to the extracted SNZC. If

Proceedings of the 6th WSEAS International Conference on Multimedia Systems & Signal Processing, Hangzhou, China, April 16-18, 2006 (pp7-12)

they are not equal, mark the current 8x8 sub-block as "dirty".

(3) If there are still more 8x8 sub-blocks or macroblocks, go to next block and return to the Step (1); and

(4) Read the complete turbo-processed copyright information signal from the buffer; then pass it to the turbo decoder. The turbo decoder will recover errors and try to reconstruct the most accurate copyright information.

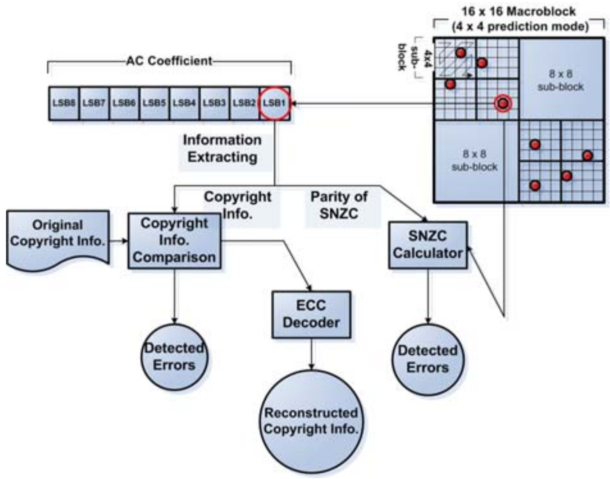

Fig.5 Information extracting / error detection process

In the proposed error detection scheme, the detection is based on 8x8 sub-block. This means once an error is detected in a 4x4 sub-block then the entire 8x8 sub-block containing this 4x4 sub-block is marked as a corrupted block. The purpose of doing this is to cover as many blocks as possible for detection while maintaining the highest video quality. In fact, the choice of value of threshold ***t*** affects the number of corrupted blocks that can be detected. Generally, the value of ***t*** and the number of detectable error blocks is pro rata: the higher value of ***t***, the higher video quality but the less error blocks can be detected. According to the experimental results, to retain high video quality, the threshold ***t*** should be at least set to the value 6. However, choosing the right threshold depends on the requirements of specific applications.

4 Experimental Results

In the following two experiments, the ***trevor*** video sequence with 150 QCIF frames (176 x 144 pixels) is used to encode H.264 sequence while embedding information in intra frames.

4.1 Experiment 1

In the first experiment, the proposed copyright protection scheme using turbo code is evaluated (the threshold ***t*** value is set to 6). The testing copyright information used is 8 text symbols "abcdefgh". The Bit Error Rate (BER) and the retrieved Copyright Information Correct Rate (CICR) are used to measure the copyright protection capability of the proposed approach. And:

$$CICR = (\chi / \tau) * 100\%$$

While χ is the number of correctly retrieved copyright information bits, and τ is the total number of copyright information bits. BER is the average bit error rate for the turbo coded bit stream instead of the original copyright information. There are 6 cases are used to simulate errors for the video stream. Random noises of 5%, 10%, 15%, 20%, 25% and 30% are added into the non-zero coefficients of intra frames of the testing video for case 1 through to case 6. The CICRs and retrieved copyright information for the original H.264 codec and the proposed approach are compared in Table 1 and Fig.6. Table 2 compares the quality of decoded images for the original H.264 codec and the proposed approach.

Comparison of CICR and extracted copyright information					
Noise Ratio	BER (%)	Proposed approach		Original H.264 Codec	
		CICR (%)	Extracted Information	CICR (%)	Extracted Information
Case 1 (5%)	2.6042	93.3125	abcdefgc	98.4375	abcdefwh
Case 2 (10%)	5.7292	93.3125	abcdefgc	92.1875	abct§ggh
Case 3 (15%)	5.1563	93.3125	abcdefgc	90.6250	#`cfEfeh
Case 4 (20%)	10.9375	93.3125	abcdefgc	84.3750	Qs← dmfgx
Case 5 (25%)	11.9792	93.3125	abcdefgc	84.3750	!òBegnox
Case 6 (30%)	17.1875	70.3125	a cÉÿÿc	43.7500	┴î¸àú®ç

Table 1 Comparison of retrieved copyright information

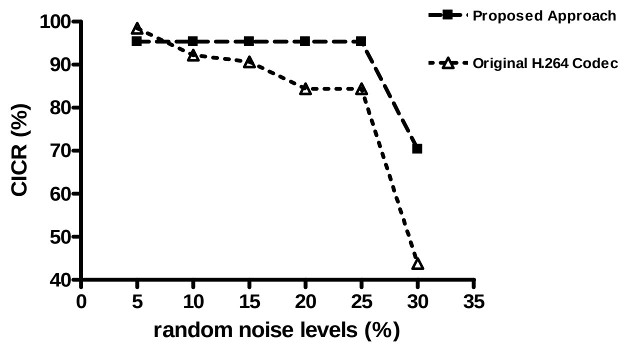

Fig.6 Comparison of CICR

Original H.264 decoded images	Proposed approach decoded images (copyright info. embedded)
	
Frame 0	Frame 0
	
Frame 36	Frame 36
	
Frame 60	Frame 60

Table 2 Comparison of decoded images

From Table1 and Fig.6 above, it is apparent that the proposed new copyright protection approach can achieve higher accuracy when retrieving the copyright information. Typically the copyright information without turbo coding will be unreadable if the BER is over 5.7%. On the other hand, in many cases the turbo encoded copyright information can be automatically recovered and efficiently reconstructed even if the BER is around 12%. From the Table 2, it is apparent that the video quality can be maintained by the proposed approach.

4.2 Experiment 2

In the second experiment, the video error detection capability of the proposed approach is measured. For the intra frames (IntraPeriod = 6) in the testing video sequence, the number of detected 8x8 error sub-blocks which are not detectable by the original H.264 decoder is recorded and illustrated in the Table 3 and Fig.7. The testing results shown in Table 3 and Fig.7 are under different random noise levels (from 5% to 30%) while the threshold ***t*** is set to 6, 7 and 8. The noise is added in the same way as experiment 1.

Number of error blocks detected for intra frames						
Noise levels / Threshold	5%	10%	15%	20%	25%	30%
t = 6	77	103	126	145	158	178
t = 7	47	56	64	92	101	118
t = 8	26	48	47	59	57	47

Table 3 Number of error blocks detected for intra-frames

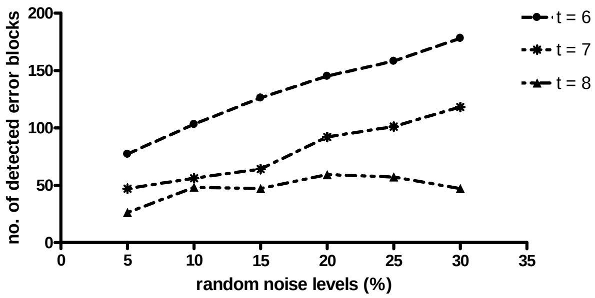

Fig.7 Comparison of error detection capability with different ***t*** values

From the Table 3 and Fig.7 above, it is discovered that the proposed new approach can efficiently detect errors in intra frames that are not detectable by the original H.264 codec. It is also discovered that when the threshold ***t*** is set to 6, the proposed approach can detect maximum number of error blocks. The error detection performance drops noticeably when ***t*** is set to 7 and 8. However the PSNR (luminance) for ***t = 6*** is similar to that of ***t =7*** and ***t = 8***. For the testing video ***travor***, the PSNR for the original H.264 codec is 37.52 db, and the PSNRs for ***t*** = 6, ***t = 7*** and ***t = 8*** for the proposed approach are 36.88 db, 36.98 db and 37.00 db respectively. It shows that ***t=6*** is a good choice for the testing video. On the other hand, setting the ***t*** value to 7 or larger values can preserve even better video quality. The trade off between the video quality and error detection capability depends on the requirement of specific applications.

5 Conclusion and Future Work

This paper proposes a new information embedding scheme for detecting errors in intra frames for H.264 video stream and for claiming ownership of the video. One type of signal to be embedded into the video is obtained by applying copyrights information to a turbo encoder. Another type of information to be embedded is the parity bit of the sum of non-zero coefficients for 8x8 sub-blocks. These two classes of information are interlaced and embedded into the selected “optimal position” in intra frames of the

H.264 video. At the receiver, the extracted copyright information and the recalculated sum of non-zero coefficients are used to detect and locate the transmission errors in the intra frames. Experimental results on a sequence of 150 QCIF images show that the embedded copyright information can be protected and efficiently reconstructed by the turbo codec under a certain amount of noises. The proposed scheme can detect the errors in the transmitted DCT coefficients for intra frames which are not detectable by the original H.264 codec.

The proposed scheme makes use of embedded copyright information and combines this with video error detection. However, this scheme can only detect errors in intra frames. It is also desirable to embed copyright information into the motion vectors so it can detect the errors in P and B frames of H.264 video. Besides, the turbo codec used in this work is possible to be tuned to achieve higher error correcting performance.

References:

[1] Stephan Wenger, H.264 Over IP, *IEEE Transactions on Circuits And Systems For Video Technology*, Vol. 13, No. 7, July 2003, pp. 645-656.

[2] Wiegand, T. et al., Overview of the H.264/AVC Video Coding Standard, *IEEE Transactions on Circuits and Systems For Video Technology*, Vol. 13, No. 7, Jul 2003, pp. 560-576.

[3] J. Ostermann et al, Video Coding with H.264/AVC: Tools, Performance and Complexity, *IEEE Circuits And Systems Magazine*, Vol. 4, No. 1, First Quarter 2004, pp. 7-28.

[4] S. Lin and D.J. Costello, *Error Control Coding: Fundamentals and Applications*, Prentice-Hall, Englewood Cliffs, JF, 1983.

[5] C.S. Bae et al., Error Detection and Copyrights Protection Scheme for MPEG-2 Based on Channel Coded Watermark Signal, *Proceeding of the International Conference on Information Technology and Applications (ICITA05)*, 4^{th} – 7^{th} July, 2005, Sydney, Australia.

[6] Y. Wang and Q.F. Zhu, Error control and concealment for video communication: a review, *IEEE Proceedings*, VOL. 86, NO. 5, pp. 947-997, May 1998.

[7] Wikipedia: Turbo Code, http://en.wikipedia.org/wiki/Turbo_code.

[8] Morelos-Zaragoza, Robert H., *The Art of Error Correcting Coding,* John Wiley & Sons Chichester April 19, 2002.