2009 International Conference on Multimedia Information Networking and Security

An Efficient Method to Hide Information in MPEG Video Sequences

Peng Zheng Bo Zhao Min-zhong Liu
The Key Laboratory of Aerospace Information Security and Trust Computing，
Ministry of Education, Hubei, 430072, P.R.China
School of Computer, Wuhan University, Hubei, 430072, P.R.China
pzheng51@163.com

Abstract

In the paper, we present an efficient algorithm to hide information in MPEG compressed video. In order to utilize the large amount of video streams, and to speed up the process of hiding information, we provide an algorithm that makes use of compressed coefficients to hide information directly. The advantage of the method is that we can make use of DCT coefficients directly without compressing and decompressing. We have experimented with the algorithm, and the result of experiment indicates that the performance of the algorithm is efficient.

1. Introduction

The rapid growth in the demand and consumption of the digital video data in recent years has led some issues that we need to face, such as content security, authenticity, and digital rights management, etc. Macq[1] summarizes the related research in protecting the ownership rights of digital media. In a relatively short span, the use of digital data hiding for covert communication has made a notable progress. Chandramouli[2] gives a review of steganalysis, and Fridrich[3] shows many effective methods.

In practical video storage and distribution system, the video sequences are stored and transmitted in compressed format. MPEG compressed video is the most popular, because of its efficient compressed ratio and quality of picture.

In the paper we will discus the hiding information methods based on MPEG compressed video. The rest of the paper is organized as follows. In section 2, we introduce related works. In section 3, we propose the idea of hiding information and the position of hiding information. In section 4, we propose the algorithm. Some experimental results are given in section 5. Finally, we give the conclusions and future work in section 6.

2. The Related Work

We focus talking about the methods in the compressed domains. Hartung[4] presents techniques to embed a spread-spectrum watermark into MPEG-2 compressed videos. The basic ideas are generating a watermark signal for each frame of the video sequence exactly at the same manner, and arranging the watermark signal into a two-dimensional signal as the video frames. The capacity of hiding information is not large by the method.

Langelarr[5] presents a compressed domain watermark technique called Differential Energy Watermark(DEW) in which the video is partitioned into groups of blocks each of which is further divided into two sets of equal size as determined by the watermark embedding key. The choice of the DEW standard is not easy.

Yang[6] presents an algorithm dividing the 4*4 block into sub-blocks and modify only the coefficients within possibly as few as only one sub-block. He hides information within low-frequency coefficients. The method is just suitable for H.264.

Langelaar[7] embeds information based on DCT coefficients, but each four MBs is embedded one bit, so the capacity of hiding is not large. And the method need insert stuffing bits in each MB in order to keep its original size.

Wang[8] proposes the method as follows. First, the candidate I frame for data hiding is extracted. Then the Variable Length Code (VLC) of the intracoded block is obtained. The VLC of AC components of the selected block is decoded to get the quantized values which are integer numbers.

This work is supported by Wuhan Scientific Problems Tackling Project(200810321134)

978-0-7695-3843-3/09 $26.00 © 2009 Crown Copyright
DOI 10.1109/MINES.2009.123

Xie[9] proposes the algorithm by adopting process quantify to dither modulate of double polarity parameter of DCT coefficient. When adopting quantifying to embed information, the selection of quantifying parameter is the linchpin, if the parameter is too small, then robustness is bad; if the parameter is too big, then imperceptibility is not good.

3. The Proposed Method

There are three opportunities to hide information in video stream. The first opportunity is to hide secret information before encoded. Such method will encode and decode video sequences while hiding information. Thus, the hiding information is easy lost, and it is not convenient for extracting and detecting the hiding information. The second opportunity is to hide secret information while video encoder and decoder are processing. Facing a large amount of video streams, the cost of such method is very high. The third opportunity is to embed secret information into the compressed video directly. The advantage of the method is needless of encoding and decoding, but compressed ratio confine the capacity of hiding information and the designer of algorithm is more complex. So based on MPEG video structure and research of [4]~[9], considering practical situation, we present a method that makes use of compressed DCT coefficients to hide information in video sequences.

3.1. The Typical Structure of MPEG

An MPEG sequence is composed of group of pictures (GOP). A GOP is composed of frames. There are three types of frame: intra-coded frame (I-frame), prediction-coded frame (P-frame) and bidirectional interpolated frame (B-frame). Each frame is divided into blocks of 16×16 pixels called macroblocks (MBs). There are four types of MBs: intra-coded I-MB, forward-predicated P-MB, backward-predicated B-MB, bidirectionally-coded D-MB. Every I-frame only consists of I-MBs. Every P-frame consists of I-MBs and P-MBs. In the B-frames, there are all of above four kinds of MBs. Every MB is composed of 4 8×8 blocks, and block is the basic unit of compressed code. MPEG doesn't prescribe strictly the types of frames in a GOP. They can be organized flexibly. Examples as following.

Example 1: IIIIIIIIIIII

Example 2: IPPPPPPPPPPP

Example 3: IBPBPBPBPBPB

Example 4: IBBPBBPBBPBB

Among the examples, Example 4 is widespread used due to its good compressed ratio and image quality. The typical proportion of I:P:B is 1:3:8. There are about two GOPs in one second.

The red, green and blue (RGB) signals coming from a color television camera can be equivalently expressed as luminance (Y) and chrominance (UV) components. The chrominance bandwidth may be reduced relative to the luminance without significantly affecting the picture quality. YUV video signals can be sampled and digitized to form discrete pixels. The terms 4:2:2 and 4:2:0 are often used to describe the sampling structure of the digital picture. 4:2:2 means the chrominance is horizontally subsampled by a factor of two relative to the luminance; 4:2:0 means the chrominance is horizontally and vertically subsampled by a factor of two relative to the luminance.

Up to now, there are many standards have been made by MPEG. Among them, MPEG-2 is a generic video coding system supporting a diverse range of applications. Y:U:V of MPEG-2 is 4:2:0.

A two-dimensional DCT is performed on small blocks (8 pixels by 8 lines) of each component of the picture to produce blocks of DCT coefficients. The magnitude of each DCT coefficient indicates the contribution of a particular combination of horizontal and vertical spatial frequencies to the original picture block. The coefficient corresponding to zero horizontal and vertical frequency is called the DC coefficient. The others are called AC coefficients.

3.2. The Positions of Hiding Information

We hide information not in every frame, just in I-frames and P-frames. I-frames have the highest channel capacity for hiding information due to I-frames just including I-MBs. I-frames are the preferred frames for hiding information. In order to increase capacity of embedded we also choose P-frames as embedded frames because when shot transition occurs P-frames include many I-MBs. Figure 1 illustrates such situation. In order to illustrate clearly, we just take 160×128 video picture as an example. In Figure 1, P-frame is very different from the former I-frame, so there are many I-MBs in the P-frame. Between the I-frame and P-frame, there are 2 B-frames that are more similar to the I-frame. B-frames have the lowest channel capacity for hiding information. The number of I-MBs in B-frames is usually few, so we don't consider B-frames.

Mankind's eyes are more sensitive to luminance than to chrominance. So there are 4 groups of luminance DCT coefficients, 2 groups of chrominance DCT coefficients in each MB. We call the 6 groups as Y1, Y2, Y3, Y4, U, and V. The DCT transform tends to concentrate the energy into the low-frequency coefficients and many of the other coefficients are near-zero. The low-frequency coefficients are more sensitive to our visual system than high-frequency coefficients.

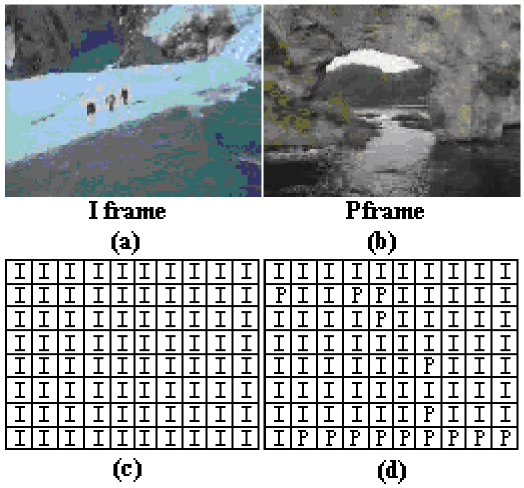


Figure 1. MBs in I-frame and P-frame

If we alter the high frequency, it will reduce the efficiency of run-level encoding. If we alter the low frequency, it will distort the quality of picture. So the proposed hiding information scheme just alters the 8 quantized medium frequency AC coefficients of each luminance block. The other parameters will not be altered. We use AC(0)~AC(63) to represent AC coefficients. We illustrate the positions where the hiding information in figure 2. The 8 little white squares represent AC coefficients that we will use.

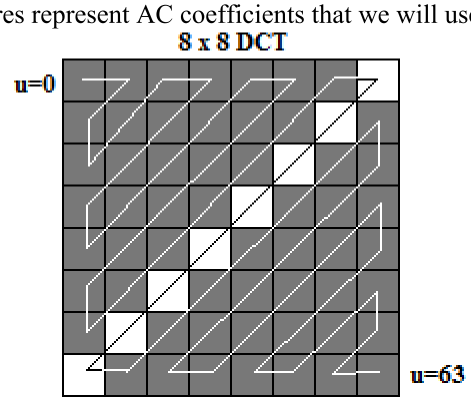


Figure 2. 64DCT Coefficients

In such method, we can restrict the noise to a unperceivable level, and acquire a fast speed.

4. Hiding and Extracting Information

In order to keep the quality of a picture, in the proposed algorithm, we just hide a byte in Yi, (i=1,2,3,4) per I-MB.

Suppose H represents the information that need to be hidden. We define H(k,j) as follows.

H(k,j)= {h (k, j), k=0,1,…, N-1; j=0,1,…,7}

k represents the byte k in H, j represents the bit j in the byte.

The strategy of hiding information: If H(k,j)=0, we replace the AC(28+j) of Yi with 0. If H(k,j)=1 and AC(28+j)=0, we add 1 to AC(28+j), else if H(k,j)=1 and AC(28+j) ≠0, we do nothing.

4.1. The Hiding Information Algorithm

The algorithm of hiding information is described as following:

(1) Locating the first I-frame or P-frame;

(2) Getting the first, second, third, fourth byte that will be hidden, and saving them into H1, H2, H3, H4;

(3) Locating the first I-MB in the selected frame;

(4) Scanning the 8*8 DCT block in the selected I MB in zig-zag fashion;

(5) Replacing the AC(28)~AC(35) of Y1, Y2, Y3, Y4 with H1, H2, H3, H4, based the aforesaid strategy;

(6) If there is information that still not be hidden, getting the next 4 bytes that will be hidden, and saving them into H1, H2, H3, H4, else goto (9);

(7) In the frame if there is any I-MB not being embedded, locating the next I-MB, goto (4);

(8) If there is a I-frame or P-frame that not be embedded, locating the next I-frame or P-frame, goto (3), else report the result, and return;

(9) Putting a mark, return.

4.2. The Extracting Information Algorithm

The algorithm of extracting information is described as following:

(1) Locating the first I-frame or P-frame;

(2) Locating the first I MB in the selected frame;

(3) Scanning the 8*8 DCT block in the selected I MB in zig-zag fashion;

(4) Getting the AC(28)~AC(35) of Y1, Y2, Y3, Y4, if the value is 0, extracting 0,else extracting 1;

(5) If encountering the mark, return;

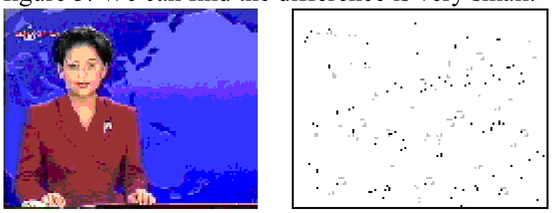
(6) In the frame if there is any I MB not being extracted, locating the next I MB, goto (3), else goto (7);

(7) If there is a I-frame or P-frame that not be extracted, locating the next I-frame or P-frame, goto (2), else return;

5. Experiments Results

There are several sizes of picture in MPEG-2 standard, such as 352×288, 720×576, etc. The pixels of a picture in testing video are 352 × 288. The structure of GOP is IBBPBBPBBPBB. The testing video sequences are news video sequences.

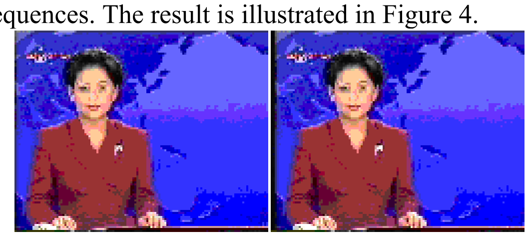
We have implemented the proposed algorithm with Visual Studio 2005. We illustrate the difference of I-frame between the original frame and embedded frame in figure 3. We can find the difference is very small.



(a) Unhided I-frame (b)Frame difference after hided

Figure 3. Frame Difference

The average PSNR between the MPEG-compressed original and the uncompressed original is 37dB[7]. In order to value the quality of embedded picture, we have tested the PSNR of I-frames in testing video sequences. The result is illustrated in Figure 4.



(a) (b)

Figure 4. (a) Un-embedded I-frame, PSNR=35.56dB;
(b) Embedded I-frame, PSNR=34.73dB,the number of embedded is 1584 bytes.

6. Analysis and Conclusions

There are 25 frames in 1 second, so there are about 2 GOPs in 1 second. Two I-frames will emerge in 1 second at least.

If the pixels of a picture is 720×576, the blocks of an I-frame is 90×72. Based on the algorithm, one byte will be hidden in a block, so 6480 bytes can be hidden in an I-frame. Every GOP includes at least one I-frame, so at least 12K information can be hidden in 1 second video sequences. Adding up the capacity of P-frames in a GOP, more information can be hidden.

In this paper, we have presented an efficient information hiding algorithm for digital video content under MPEG-2 compression.

The proposed algorithm embeds information by dividing the 8*8 DCT coefficients block into sub-blocks and hiding the data within the medium-frequency sub-blocks.

The proposed algorithm has gained a very high level of robustness against MPEG-2 compression while causing no visual distortion. It is also shown that hidden information within different types of frames have different levels of robustness against compression due to the different coding strategies of I-frame, P-frame, and B-frame. How to increase the robust of hiding information against the MPEG compressed attacks is the issue that we will deal with for the future.

The proposed algorithm is going to be very useful in practical application such as secret information delivery, captioning, video-in-video, and etc. The method can hide information real time. In the future, it will be applied to video streaming applications for video-on-demand, and thus it will be tested under a packet-switch network transmission environment.

7. References

[1] B. Macq, J. Dittmann, and E. J. Delp, “Benchmarking of image watermarking algorithms for digital rights management”, Proc. IEEE, vol. 92, no. 6, Jun. 2004, pp. 971–983.

[2] R. Chandramouli, M. Kharrazi, and N. Memon, “Image steganography and steganalysis: concepts and practices”, in Proc. 2nd Int. Workshop Digital Watermarking, Seoul, Korea, Jan. 2003, pp. 35–49.

[3] J. Fridrich, M. Goljan, “Practical steganalysis of digital images—state of the art”, in Proc. IST/SPIE 14th Annu. Symp. Electronic Imaging Science Technology, 2002, vol. 4675.

[4]F. Hartung, B. Girod, Watermarking of Uncompressed and Compressed Video, Signal Processing, vol. 66, no. 3 (Special issue on Watermarking) , May 1998, pp. 283-301.

[5]G. C. Langelarr, R. L. Lagendijk, “Optimal Differential Energy Watermarking of DCT Encoded Images And Video”, IEEE Transactions on Image Processing 10, 2001, 148-158

[6]M. Yang, N. Boubrakis, “A High Bitrate Information Hiding Algorithm for Digital Video Content under H.264/AVC Compression”, Circuits and Systems, 48th Midwest Symposium , 2005, pp:935-938.

[7]G.C. Langelaar, R.L.Lagendijk,J.Biemond, “Real-time Labeling of MPEG-2 Compressed Video”, In :Image Representation, vol. 9, no. 4, 1998, pp.256-270.

[8]Y.Wang, E. Izquierdo, “High-Capacity Data Hiding in MPEG-2 Compressed Video”, Proceeding of the 9th International Workshop on Systems, Signals and Image Processing, Manchester, U.K., Nov 2002, pp.212-218.

[9]J. Q. Xie, C.H.Yang, D.Z.Huang, “High Capacity Information Hiding Algorithm for DCT Domain of Image”, Proceedings of the 2008 International Conference on Intelligent Information Hiding and Multimedia Signal Processing, Volume 00, Publisher: IEEE Computer Society, 2008,pp.269-272.