A High Bitrate Information Hiding Algorithm for Digital Video Content under H.264/AVC Compression

Ming Yang
Information Technology Research Institute
Wright State University, Dayton, OH 45435, USA
Email: myang@cs.wright.edu

Nikolaos Bourbakis
Information Technology Research Institute
Wright State University, Dayton, OH 45435, USA
Email: nikolaos.bourbakis@wright.edu

Abstract

*With the proliferation of digital multimedia content, such as image, audio, video, and animation, information hiding techniques have attracted more and more research interests. High bitrate information hiding is different from digital watermarking in that it tries to hide relatively large amount of auxiliary information instead of just one or a few verification bits. This research is mainly focused on high bitrate information hiding within digital video content. Channel capacity and the robustness of hidden information against lossy video codec are the two main concerns. In the proposed algorithm, 1 bit is hidden within each 4*4 DCT coefficient block by means of vector quantization. Low-frequency coefficients are chosen for information hiding due to their relatively large amplitudes and the corresponding small step sizes in the quantization matrix. The proposed algorithm is tested under H.264/AVC (Advanced Video Codec), which is the state-of-the-art video coding standard. Experimental results show that the proposed algorithm is very robust and a very high percentage of the hidden information survives H.264/AVC compression. It is also observed that hidden information within different types of frame/slice have different levels of robustness against H.264/AVC compression.*

Keywords

Information Hiding, Video, Channel Capacity, Bitrate, Compression, H.264/AVC.

INTRODUCTION

Information security is a key issue in modern information and communication systems. Traditionally, the security of information is ensured with data encryption, which is a method for protecting information from undesirable attacks by converting it into a non-recognizable form. Nowadays, researchers start to resort to information hiding techniques to protect the host data. Information hiding system is adopted in many application scenarios. With the modern information hiding techniques, it is possible to hide secret message within digital multimedia such as digital audio, digital image, and digital video without degrading the perceptual quality of the host content. Given a cover object I, a message M, the embedding algorithm E, and a key K, the embedding process can be defined as a mapping of the form: $I \times K \times M \times E \rightarrow I'$, where I' is the stego-object. The information retrieval procedure is depicted as: $I' \times K \times E' \rightarrow M'$, where I' is the stego-object, K is the key, E' is the extracting algorithm, and M' is the retrieved message. High bitrate information hiding algorithms can be classified into spatial-domain algorithms and transform-domain algorithms. The most straightforward spatial-domain algorithm for digital image embeds information by modifying the least-significant-bit of each pixel. However, this algorithm is very sensitive to lossy compression. Transform-domain algorithms embed information in the transform domain of the host image/video. Swanson, Zhu, and Tewfik [1] proposed a vector projection based high bitrate information hiding algorithm. Alturki and Mersereau [3] proposed an algorithm that embeds the data by whitening the image and quantizing each DCT coefficients. Chae and Manjunath [4] make use of lattice structure to code the information and place the information in the mid-frequency coefficient locations of the DCT block. In this research, we are trying to design and implement a high bit-rate information hiding algorithm for digital video. Here we have 4 objectives:

(1) High Bit-rate: 1 bit per 4*4 block;
(2) Blind information retrieval;
(3) None or minimum visual degradation;
(4) Robustness to lossy video codec.

The proposed algorithm differs from previous work in that we make use of vector quantization instead of scalar quantization to embed the information, which makes the hidden data less sensitive to noise. Rather than choosing all coefficients for data embedding, we choose only the low-frequency coefficients along the zig-zag scanning path to hide the information and thus avoid degrading the performance of run-length coding.

INFORMATION HIDING STRATEGY

Temporal Location of Hidden Information

Video information hiding is similar to image information hiding since digital video is in essence a sequence of still images. With the Group-Of-Pictures (GOP) structures, there are three different kinds of frames: I-frame, P-frame, and B-frame, each of which is compressed with different strategies. Due to different compression algorithms, it is expected that I-frames have the highest channel capacity for information hiding, while B-frames have the lowest channel capacity for information hiding. Figure 1 illustrates the basic structure for video information hiding.

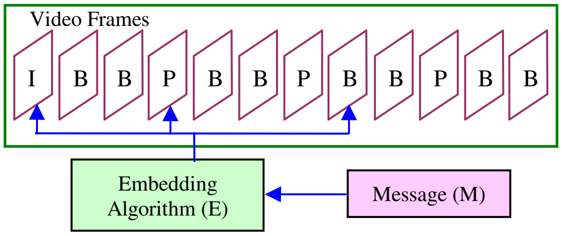

Figure 1 Video information hiding

Spatial Location in YUV Domain

In video compression, the source picture and the decoded picture are both comprised of three sample arrays: 1 luminance (luma) array and two chrominance (chroma) arrays. Since human vision system is more sensitive to luminance than to chrominance, the luma array is kept as it is and the chroma arrays are down-sampled. For the video information hiding, we need to change the luma or chroma sample arrays to embed the message. It is observed in human vision experiments that human eyes are more sensitive to down-sampled chroma components change than to luma components change. As such, in order not to lead to any color distortion, only the luma sample array is chosen as the host of information hiding. The luma array will be modified as a grey-level image to hide information.

Spatial Location in DCT Domain

Here, we need to embed one bit within each 4*4 DCT block. The reasons are: (1) H.264 uses 4*4 integer transformations for decorrelation; (2) integer transformation is a very close approximation of DCT. Existing algorithm proposed to embed information by perturbing the whole block. However, the algorithm has two drawbacks: (1) too much visual distortion due to the modification of every coefficient within the block; (2) the modification of high-frequency coefficients will degrade the performance of run-length coding. In the proposed algorithm, we divide the 4*4 block into sub-blocks and modify only the coefficients within possibly as few as only one sub-block. It is found that low-frequency coefficients have much higher average amplitude compared to high-frequency coefficients. Intuitively, it is better to hide information within low-frequency coefficients. Experimental results also show that the proposed algorithm works best at low-frequency coefficients due to their high amplitudes. As such, the choice of sub-band for information hiding is biased to low-frequency DCT coefficients.

INFORMATION HIDING/RETRIEVAL ALGORITHM

In the proposed algorithm, the message is hidden within Y components only in order not to cause any color distortion. First of all, we extract the Y components from the video trace files. After that, the Y component of each frame, which is essentially a grey-level image, is used for information hiding. The information algorithm proposed for still image information hiding [8] will be applied here. Of course, due to the new features of H.264/AVC, some modifications have to be made in order to make the embedding algorithm robust to H.264/AVC. After information hiding, the Y components will be put back to the original video trace and thus the stego-trace, which holds the message, is generated. Figure 2 illustrates the algorithm.

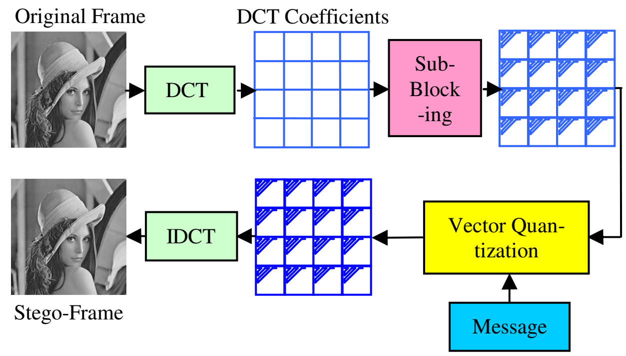

Figure 2 Information Hiding Algorithm

Information Hiding Algorithm

The embedding methodology can be described as following:

(1) The 4*4 DCT block is scanned in zig-zag fashion;
(2) The 8 low-frequency coefficients are converted to an 8-D vector;
(3) Here we have:
V: the 8-D vector $V = (c_0, c_1, c_2, ..., c_6, c_7)$
T: the threshold for vector quantization
$|V|$: the length of vector V
$[l]$: round-off operation

$$l = |V| = \sqrt{\sum_{i=0}^{7} c_i^2}$$

$$l_T = \left[\frac{|V|}{T}\right] = \left[\frac{\sqrt{\sum_{i=0}^{7} c_i^2}}{T}\right]$$

(4) One bit is embedded by modifying l_T: $l_T' = l_T \pm 0.25$ (+0.25 to embed 1, and -0.25 to embed 0);
(5) $l' = l_T' * T$, $V' = \frac{l'}{l} * V$;
(6) Put the vector V' back to its original location in the 4*4 DCT block;
(7) Repeat the same operation for each 4*4 DCT block until all the information bits have been embedded.

Blind Information Retrieval without Original Image

(1) DCT transform the stego-frame (frame with hidden information);
(2) For each 4*4 DCT block, scan the coefficients in zig-zag scanning order;

(3) Pick up the 8 lowest frequency coefficients and convert them to an 8-D vector *V"*;

(4) $l'' = |V''|$

(5) $l_T'' = \frac{l''}{T} = \frac{|V''|}{T}$

(6) $I = l_T'' - [l_T'']$

If $I >= 0$, then 1 is extracted as the information bit;

If $I < 0$, then 0 is extracted as the information bit.

(7) Repeat the same operation to each 4*4 DCT block until all the information bits have been extracted.

H.264 ADVANCED VIDEO CODEC

H.264/AVC, also known as MPEG-4 part 10, is the latest development of video coding standard. This standard was proposed by Joint Video Team (JVT) formed by ITU-T Video Coding Experts Group (VCEG) and ISO Motion Picture Expert Group (MPEG). In H.264/AVC, many new features have been added in order to improve the coding performance.

Intra Prediction

In H.264/AVC, if a block of a macroblock is encoded in intra mode, then a prediction block is formed based on previously encoded and reconstructed blocks. The samples above and to the left of the current block will be used in the encoder and decoder to form a prediction reference. The difference between the prediction block and the current block will be transformed and then entropy-coded.

Inter Prediction

Inter prediction generates a prediction block from one or more than one previously encoded video frames. The prediction is basically formed by motion estimation. The new features added to H.264/AVC are multiple reference frames, quarter-pixel precision motion estimation, and variable block sizes.

Picture/Slice Management

A picture may be split into one or more than one slices in H.264/AVC. A slice can be parsed and decoded independently without using the data from other slices. Slices can be coded with different coding types: I slice, P slice, B slice, SP slice, and SI slice.

Integer transform

After intra- and inter- prediction, each residual macroblock is transformed, quantized, and entropy-coded. H.264 uses integer transform, which is an approximation of DCT transform. After the integer transformation, the DC coefficient is extracted from each of the 16 4*4 luma block, and then a 4*4 DC coefficient block is obtained. The 4*4 DC block will be transformed with Hadamard Transformation to further de-correlate the DC coefficients.

EXPERIMENTAL RESULTS

We have implemented the proposed algorithm with Visual C++ .NET. After information hiding, we compress the stego-video with H.264/AVC under different quantization parameters (QP). The performance of the video information hiding algorithms will be evaluated with the metrics of Bit-Error-Rate (BER), Peak-Signal-to-Noise-Ratio (PSNR), and subjective testing by comparing the original video frames and the stego video frames with hidden information. The robustness of the algorithm to H.264/AVC compression is measured with the metric of (1.0-BER), i.e. how many percents of the hidden information could survive the video codec. A GOP structure of "IBPB" is adopted. The video sequence is IBPBIBPBIBPBIBPB….. and so on. In figure 3, X-axis is the index of the frame, and Y-axis is the survival rate of hidden information in that frame. The periodical spikes reflect the GOP structure we have been using. As expected, information within I-frame has the best chance to survive H.264/AVC. Between the two adjacent spikes, we can see there is a smaller spike, which corresponds to the P-frame between two B-frames. Figure 4 is the comparison between the original video frames and the stego-video frames. As can be seen, there is no perceptual difference between them.

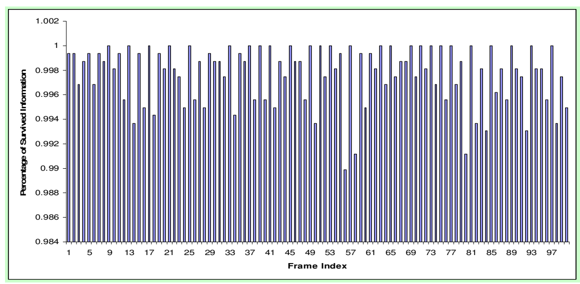

Figure 3 (a) Performance Curve under QP = 10

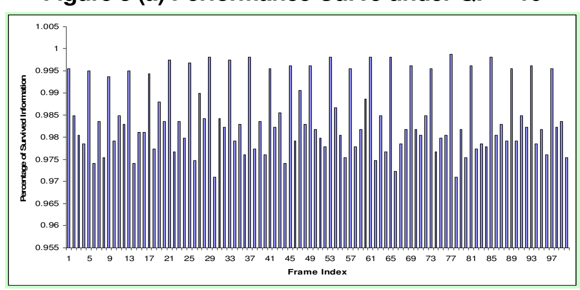

Figure 3 (b) Performance Curve under QP = 15

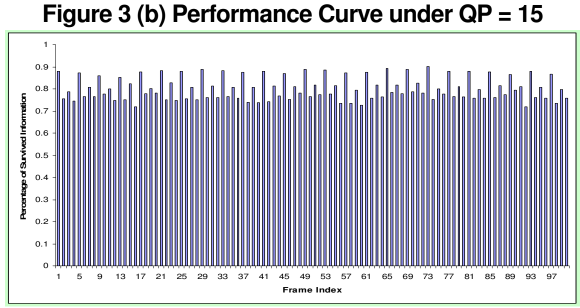

Figure 3 (c) Performance Curve under QP = 20

(a) (b)

Figure 4 (a) Original Frames (b) Stego Frames

CONCLUSIONS AND FUTURE WORK

In this paper, we present a high bitrate information hiding algorithm for digital video content under H.264/AVC compression. The proposed algorithm embeds information by dividing the 4*4 DCT coefficients block into sub-blocks and hiding the data within the low-frequency sub-blocks. The proposed algorithm has gained a very high level of robustness against H.264/AVC 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. In the future, the proposed algorithm will be applied to video streaming applications for video-on-demand, and thus it will be tested under a packet-switch network transmission environment. The proposed algorithm is going to be very useful in practical application such as side information delivery, captioning, video-in-video, speech-in-video, and etc.

REFERENCES

[1] Swanson, M.D., Zhu, B., and Tewfik, A.H., "Data Hiding for Video-in-Video", Proceedings of the 1997 International Conference on Image Processing (ICIP '97), Vol. 2, pp. 676-679, Santa Barbara, California, October, 1997.

[2] Kundur, D., "Implications for High Capacity Data Hiding in the Presence of Lossy Compression", Proceeding of International Conference on Information Technology: Coding and Computing, pp. 16-21, March, 2000.

[3] Alturki, F., and Mersereau, R., "A Novel Approach for Increasing Security and Data Embedding Capacity in Images for Data Hiding Applications", Proceeding of International Conference on Information Technology: Coding and Computing, 2001, pp. 228 –233, April, 2001.

[4] Chae, J.J., and Manjunath, B.S., "Data Hiding in Video", Proceeding of the 6th IEEE International Conference on Image Processing (ICIP'99), Kobe, Japan, vol. 1, pp. 311-315 , Oct. 1999.

[5] Watson, A.B., "DCT Quantization Matrices Visually Optimized for Individual Images", Human Vision, Visual Processing, and Digital Display IV, Bernice E. Rogowitz, Editor, Proc. SPIE 1913-14, (1993).

[6] Wang, Y., and Izquierdo, E., "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.

[7] Maniccam, S.S., and Bourbakis, N., "Lossless Compression and Information Hiding in Images", Pattern Recognition Journal, Vol. 36, 2004.

[8] Yang, M. and Bourbakis, N., "A High Bitrate Multimedia Information Hiding Algorithm in DCT Domain", Proceeding of The 8th World Conference on Integrated Design and Process Technology (IDPT 2005), Beijing, China, June 2005.

[9] Shimizu, S., "Performance Analysis of Information Hiding", in Proceeding of Security and Watermarking of Multimedia, Contents IV, SPIE Vol. 4675, pp. 421-432, San Jose, USA, January 2002.

[10] Lin, E.T., and Delp, E.J., "A Review of Data Hiding in Digital Images", Proceedings of the Image Processing, Image Quality, Image Capture Systems Conference, PICS'99, Savannah, Georgia, pp. 274-278, April, 1999.

[11] Tekalp, A.M., "Digital Video Processing", Book, Prentice Hall, ISBN 7-302-02927-X.

[12] Celik, M., Sharma, G., Tekalp, A.M., and Saber, E., "Lossless Generalized LSB Data Embedding", to Appear in IEEE Transactions on Image Processing.