

Information Hiding in H.264 Compressed Video

Sajiyya K., Divya Sreenivasan, Sruthi L., Ansha Beevi S.

Department of Computer Science
College of Engineering, Karunagappally
Kollam, India

Abstract-In today's world the art of sending and displaying hidden information especially in public places, has received more challenges. Information hiding is the art in ways that avert the revelation of secret messages. Video files are generally a collection of images. So most of the presented techniques on images and audio can be applied to video files too. The great advantages of video are the large amount of data that can be hidden inside and the fact that it is a moving stream of images. Here, information hiding methods in the H.264/AVC compressed video domain are surveyed. This concept is illustrated by using various data representation schemes. Venues at which information hiding takes place are then identified. Related information hiding methods at each venue are briefly reviewed. Comparison among the considered information hiding methods is also conducted. Finally, a new technique to hide the data in moving objects is explained. The system uses motion vector as hiding component and third LSB method as hiding scheme. Original message is being hidden within a carrier such that the changes so occurred in the carrier are not observable in the video. This method describes how motion vector can be used as a carrier to hide data. The purpose of data hiding is to secretly transmit the message, thus maintaining the confidentiality of data. The advantage of proposed data hiding in video are: user cannot find the original data, it cannot be easily cracked and it increases the security and size of stored data.

Keywords: Information hiding, Data embedding, Compressed video, Compression, H.264

I. INTRODUCTION

Motion picture, or known as video, has become one of the most important media in the entertainment field. In the past, video was virtually a sequence of frames. Later added more features in video, including instant frame access, high resolution, high frame rate, fast forward, etc. Therefore, MPEG (Motion Picture Expert Group) standard is established in year 1993 and enables the VCD (Video Compact Disc) technology, then MPEG-2 in year 1995, which enables the DVD (Digital Video Disc) and Satellite TV. In the pursuit of higher efficiency in coding of video, H.264 (H.264/Advance Video Coding) is proposed by the Video Coding Experts Group and it has become one of the most commonly practiced video compression formats since year 2003. The design of H.264 deals with an enhanced

compression performance on video representation for various purposes, including broadcast, storage, video telephony and streaming applications. H.264 achieves a significant improvement in rate distortion trade off by offering high video quality for relatively low bit rate as compared to the previous generations of video compression standard.

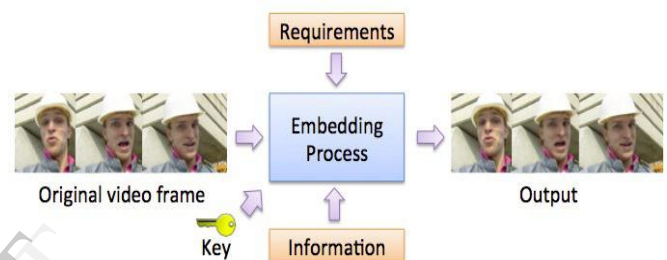


Fig .1. General framework of information hiding

Information hiding refers to the process of inserting secret message into a host to serve specific purposes. The components of information hiding are summarized in Fig 1. In particular, modify the watermarking framework to a general framework of data hiding to emphasize on the information insertion part. Here, the data is external to the content (e.g., ownership information and secret message) or deduced from the content. The data is inserted into the host by means of modifying part(s) of the host based on the representation scheme in use and a key so that the output (i.e. content and inserted information as a single unit) satisfies the imposed properties and requirements. These properties include high secrecy, reversibility, perceptual quality of the inserted information, etc. An information hiding system, in general, is expected to meet three key requirements, namely, accurate recovery of embedded information, large payload (payload is the number of bits that get delivered to the end user at the destination) and imperceptibility of embedding. In a pure information hiding framework, the technique for embedding the message should be unidentified to anyone other than the sender and the receiver.

In the video domain, the application of data embedding can be coarsely categorized as steganography, watermarking, error recovery (resilient) and general data embedding. In the case of, steganography is the art and science of concealing the existence of the secret information inserted into a cover such as video, audio image, text, etc. The existence of the embedded information should be undetectable, that is, the modified content should be

perceptually and statistically similar to its original unaltered counterpart. On the other hand, watermarking is information is inserted into video to visibly or invisibly render the ownership information. The inserted watermark data is retrieved from the content (i.e., when it is alleged to be an illegal copy) to prove ownership during a dispute, to detect any attempt in destroying the inserted watermark, or to identify act of tampering the content.

Information hiding is also utilized to realize error recovery (also referred to as error concealment). Here, the main components of the compressed video (e.g., motion vectors, prediction modes) are embedded into the video itself. When error occurs during transmission, the main components can be extracted to patch the lost or corrupted parts at reasonable perceptual quality. Lastly, general data embedding refers to the principle idea of inserting information into a host for specific purpose. For example, text description of a video is embedded into itself for annotation purpose. Note that most of the current digital video formats enable the inclusion of text through metadata or using a separate file without requiring the use of data hiding. However, text embedding is viable because the data always stays intact with the video, but at the expense of possible video quality degradation due to information hiding.

II. DATA REPRESENTATION SCHEMES

A. Bit Plane Replacement

The general method of bit plane replacement is to embed information in a particular bit plane (i.e., location) agreed upon by both the sender and receiver. In this approach, one bit can be embed into a digital host without causing significant perceptual impact on the host content. This technique is commonly applied to entities such as raw pixel value, motion vector information, audio sample, etc. It achieves low distortion, high payload, and its implementation is relatively straightforward.

Bit plane replacement is widely referred to as LSB (least significant bit), which modifies the right most bit of an entity. The idea of LSB embedding is generalized by Wang and Chan in still image [1]. They propose an optimal bit plane substitution method to hide data into an image. A genetic algorithm is used to reduce the search space to achieve the highest possible image quality based on a given embedding rate. In the video domain, Kapotas et al. apply this technique to coefficients in the luminance and chrominance channels [2]. Their results show that by modifying the last three bit planes of an 8-bit coefficient, the distortion is imperceptible. However, the aforementioned methods are not optimal because LSB is adjusted locally. Yu et al. propose an information hiding method based on the prediction process in image compression. In particular, the LSB of the prediction errors are modulated to embed information. Shahid et.al apply LSB technique for message embedding in the video bit stream [3]. They developed a strategy to embed message in the 1 LSB, 2 LSBs, or 1 and 2 LSBs together. Generally, the LSB scheme is the most straightforward method to encode

secret information, with insignificant perceptual quality degradation when applied on the raw pixel values. However, a direct LSB scheme is irreversible.

B. Spread Spectrum

In still image, spread spectrum based information hiding method is widely used for watermarking purposes. In this class of data representation scheme, hidden information is embedded within a range of acceptable changes and it is imperceptible to the human visual system. Cox et al. introduce three watermarking methods based on spread spectrum [4]. These techniques can be formulated as:

$$v'_i = v_i + aw_i \quad (1)$$

$$v'_i = v_i(1 + aw_i) \quad (2)$$

$$v'_i = v_i(e^{aw_i}) \quad (3)$$

where $w_i \in (0, 1)$ is the information, v_i and v'_i are the input and output values, respectively, and a is a scaling parameter which determines the extent to which w alters v_i . Equation 1 explains that the output is generated by adding information to the host, which is invertible. Outputs of Equation 2 and 3 are generated based on multiplication and exponential operations, and they are invertible only if v_i not equal to zero.

C. Histogram Manipulation

Fig 2 shows the histogram of pixels in a compressed still image. The pixels values are usually associated with another value to embed information. In particular, Ni et al. utilize the zero and peak points of the histogram and slightly modify the (grayscale) pixel values to embed data[5]. The original histogram is preprocessed to zero out the T-th bin (i.e., the bin next to the peak point) for embedding purposes by increasing the value c to $c + 1$ for all c greater than T.

D. Mapping Rules

Mapping rules rely on a set of code words for information inserting and extraction purposes. It depends on the structure or syntax of the content and requires a common mapping rule agreed by the sender and receiver.

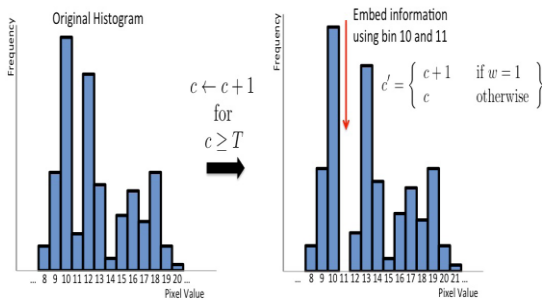


Fig. 2. Hiding information using histogram

Basically code words are associated with pre-defined meaning and they are chosen based on the information to be inserted. As such, an algorithm is forced to operate in a mode determined by the information to be embedded instead of the original decision. During the encoding process, each macroblock in frame is decomposed into different block sizes according to the information to be inserted. Kapotas et al. implement this approach to embed information during the prediction process [6]. The proposed method takes advantage of the different block sizes used by the H.264 encoder during the inter prediction stage in order to hide the data. It is a blind data hiding scheme, i.e. the message can be extracted directly from the encoded stream without the need of the original host video.

E. Divisibility

The divisibility of a value by a specific divisor can be exploited as an essential criteria for reversible information hiding. For example in Wong et al. scale the magnitude of each coefficient in the macroblock by a prime number when 1 is to be embedded, or leave them as they are to embed 0. In order for this method to be viable, divisibility of all coefficients by the chosen prime number needs to be checked to avoid error during extraction. Another method considers a pair of neighboring pixels x and y , and the value n [7]. These values are then transformed. The transformed pixels can be modified by adding an integer. Traditionally, the divisibility scheme is invented for reversible data hiding. This approach maintains high perceptual quality of the embedded video. However it is of high computational complexity.

F. Matrix Encoding (ME)

ME is a general principle that can be applied on top of the data representation schemes to improve their embedding efficiencies, i.e., the number of modifications per embedded bit. Choi et al.'s method embeds a image in selected bit planes of all color channels of a host image using ME. The output image assumes less distortion when compared to the ordinary non-ME based data representation schemes. In a hybrid data representation scheme is proposed by applying ME recursively on selected blocks and modifying only large

coefficients when necessary. Payload is improved while trading off with embedding efficiency.

III. OPPORTUNITIES FOR INFORMATION HIDING IN H.264

A. Prediction Process

Several researchers have manipulated the block prediction process in vector quantization based image compression to hide information. Different coding methods are applied on dedicated blocks, such as truncate coding and side-match vector quantization. In the compressed video domain, similar approaches are taken by block size, exploiting mode, entities, etc. that are related to the prediction process.

Intra-frame prediction

If a macroblock is encoded in intra-mode, the prediction is carried out by using one of the 14 prediction modes while referring to the previously coded and reconstructed blocks, where they themselves could be macroblocks predicted using the intra prediction mode. To exploit mode selection for information hiding, mapping rules are usually considered to increase the payload without causing significant bitrate overhead. Zhang et al. introduce hiding based on intra prediction modes [8].

Inter-frame prediction

In order to improve the coding efficiency in inter-prediction mode, H.264 standard has adopted seven different block sizes and the motion estimation algorithm is invoked for each block size. The block type that results in the minimum number of bits will be selected. In a method it forces the encoder to choose a particular block type according to the information to be inserted. In this technique, each block type is assigned to represent two bits. Then the information is divided into segments and each segment is encoded using block size. These macroblocks are then motion estimated using the forced block size. This technique only affects the visual quality of the video insignificantly. The payload is high and it is proportional to the size of host video.

Motion Vector Displacement

Information hiding can be achieved by utilizing the motion vector attributes, including phase angle, horizontal and vertical magnitudes. Jordan et al. initiate this technique for video watermarking purpose. Then, Zhang et al. and Dai et al. propose enhanced versions of Jordan et al.'s technique by restricting data hiding to specific types of inter-frame [9]. Later, consider to embed information using DCT coefficients in I-frame and magnitude of motion vectors in P-frame to achieve higher payload. There is a different information hiding approach aiming to achieve a minimum prediction error and bitstream size overhead. Instead of using magnitude and phase angle, this technique

exploits the prediction errors caused by the associated motion vector displacement to determine its suitability for information hiding. In particular, the prediction error is compared to an adaptive threshold. This technique causes low distortion in the video and suppresses bitstream size increment.

Motion vector search range

Hierarchical-based motion estimation is adopted in H.264 format to support a range of block sizes and quarter-pixel precision for achieving high compression efficiency. For each macroblock, the motion estimation process starts by searching for the best macroblock in the integer-pixel level, then to the sub-pixel level around the best integer-pixel position, and finally continues searching at quarter-pixel level around the selected sub-pixel position to find the best matching point. The information can be embedded by modifying the search points of the motion estimation process according to the mapping rule. In particular, this technique utilizes two non-overlapping sets of search points to embed information.

B.Transform process

Similar to data hiding in still image, luminance DCT coefficients are commonly utilized as the venue to hide data by using bit plane replacement embedding technique. A method to embed information into the quantized DCT coefficients (luminance) in I-frame. In a work on a similar technique to hide ROI (region of interest) information into the quantized DCT coefficients[10]. ROI information is used to represent significant object in still image and it is constructed based on skin pixel. If the current pixel is marked as a skin pixel, its position, width and height values are embedded into two LSBs of the non-zero DCT coefficients of the current frame. This technique achieves lossless reconstruction, but the results indicate that the frame payload is insufficient to host the entire ROI information.

C.Quantization process

In a technique, quantization scale of each macroblock (if it is coded) is utilized for data hiding. This concept is able to preserve the video bit stream size with low embedding complexity. In another one it maintains the quality of modified video exactly to that of the original host even after data embedding. If '0' is to be embedded, the macroblock is left as it is. Otherwise, the macroblock is manipulated by dividing the quantization scale by a prime number and multiplying each non-zero DCT coefficient by the same prime number.

D.Entropy coding

Two entropy coding methods, namely CAVLC (Context-Adaptive Variable Length Coding) and CABAC (Context-Adaptive Binary Arithmetic Coding) are available in H.264 compression format, and they are also utilized for data embedding purposes. In CAVLC, run-level coding is utilized to compactly represent strings of zeros. In a method embed information in LSB of syntax elements during the binarization process in CABAC, which is a process to concatenate all the syntax elements in binary format.

H.264 Video Compression Standard

In comparison to the previous standards, H.264 deals with new features to further improve video compression efficiency. Notably, these features include multiple frames reference capability, intra-prediction in Intra-frame quarter pixel interpolation, deblocking filtering post-processing, and FMO (flexible macroblock ordering). In general, H.264 divides the sequence of frames into several GOPs (group of pictures). These frames are labeled as I (Intra), P (Predicted) and B (Bi-directionally predicted) frames, depending on the order in which they appear.

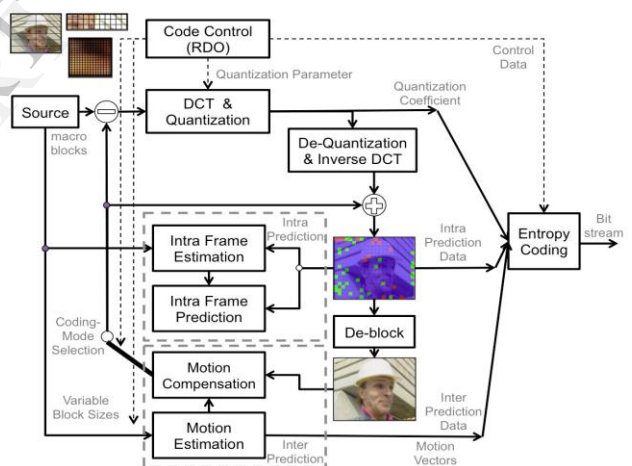


Fig. 3. H.264 hybrid video encoder.

The hybrid encoding process of H.264 video compression standard is shown in Fig. 3. At the Source part, each frame is divided into non-overlapping blocks of equal size called macroblocks. Each macroblock can be further divided into smaller blocks being the smallest possible block size. These macroblocks are subjected to DCT, quantization, and entropy coding.

Comparison of Venues for Information Hiding in H.264

Intuitively, among all the information hiding techniques, intraprediction type selections and block size offer simple ways to encode information by associating the indices with

Hiding Component (Venue)	Hiding Scheme	Related Process	Found In	Payload	Size Overhead	Impact on Video Quality	Computational Complexity
Chroma & luma DCT coefficient	Spread Spectrum/ Divisibility	Transform	I/P/B Frame, Intra prediction	High	Low	Low	Low
Motion estimated residual value	Bit Plane Replacement	Prediction/ Transform	P/B Frame, Inter prediction	High	Low	Low	Low
Motion vector	Mapping Rule	Prediction	P/B Frame, Inter prediction	Moderate	Moderate	High	Moderate
Quantization parameter	Spread Spectrum/ Divisibility	Quantization	I/P/B Frame, Intra/Inter prediction	Low	High	High	High
Block size selection	Mapping Rule	Prediction	I/P/B Frame, Intra/Inter prediction	Low	Moderate	Low	High
Intra-prediction type selection	Mapping Rule	Prediction	I/P/B Frame, Intra prediction	Low	Moderate	High	Moderate
CAVLC codeword *	Mapping Rule	Entropy Coding	I/P/B Frame, Intra/Inter prediction	High	High	High	Low
CABAC binary syntax element *	Bit Plane Replacement	Entropy Coding	I/P/B Frame, Intra/Inter prediction	High	Low	High	Low

* Quality of the embedded video is severely distorted.

Fig.4. Comparison of Venues for Information Hiding H.264

groups of bits. These techniques maintain coding efficiency with insignificant fluctuation in bitrate, but cause video bitstream size increment. On the other hand, hiding data in block size selection provides minimal impact on quality of the video and bitrate. The current approaches merely divide the block size selection into two groups and offer minimal payload. Hence, a straightforward improvement is to extend the selection groups to encode more information. The availability of motion vector in large number leads to maximum payload for information hiding in B and P frames. However, techniques involving motion vector increase the complexity of the encoding process. On the other hand, transform coefficients can provide arguably the maximum payload for information hiding purposes. However, this approach may lead to noticeable degradation in video quality and significant bit-stream size increment when embedding at high rate. Modulating the quantization parameter may cause significant degradation in visual quality. Therefore, matrix encoding technique is usually applied to reduce the number of modifications required at the expense of lower payload and higher complexity.

IV. NEW DEVELOPMENTS

The hiding method can be the LSB scheme, but the compression standard used here is H.264. So there will be chances for distortion of first LSB. So here can use the third LSB as hiding scheme. This technique is commonly applied to entities such as raw pixel value, audio sample, motion vector information, etc. Here utilizes the motion vector as hiding component. This scheme achieves high payload, low distortion and its implementation is relatively straightforward. Also imperceptibility of embedding, accurate recovery of embedded information. Also a motion vector based information hiding method can be used.

CONCLUSION

In this work, surveyed the conventional information hiding methods in the compressed video domain, focusing on the H.264 video compression standard. Commonly considered data representation schemes and the hiding venues were summarized. The general trend of information hiding in the compressed video domain were presented. Then, categorized the existing information hiding methods based on the venues at which they operate and highlighted their strengths and weaknesses. Also a motion vector based information hiding method is explained.

REFERENCES

- [1] C.-K. Chan and L. Cheng, Hiding data in images by simple LSB substitution, *Pattern Recognition*, vol. 37, no. 3, pp. 469474, Mar. 2004.
- [2] S. Kapotas and A. Skodras, Real time data hiding by exploiting the IPCM macroblocks in H.264/AVC streams, *J. of Real-Time Image Process.*, vol. 4, pp. 3341, Oct. 2009.
- [3] Z. Shahid, M. Chaumont, and W. Puech, Considering the reconstruction loop for data hiding of intra- and inter-frames of H.264/AVC, *Signal, Image and Video Process.*, pp. 119, Apr. 2011.
- [4] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoan, Secure spread spectrum watermarking for multimedia, *IEEE Trans. Image Process.*, vol. 6, no. 12, pp. 16731687, Dec. 1997.
- [5] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, Reversible data hiding, *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 3, pp. 354362, Mar. 2006.
- [6] S. Kapotas, E. Varsaki, and A. Skodras, Data hiding in H.264 encoded video sequences, in *IEEE 9th Workshop on Multimedia Signal Process.*, pp. 373376, Oct. 2007.
- [7] D. Coltuc, Improved capacity reversible watermarking, in *IEEE Int. Conference on Image Process.*, vol. 3, pp. 249252, Oct. 2007.
- [8] Y. Hu, C. Zhang, and Y. Su, Information hiding based on intra prediction modes for H.264/AVC, in *IEEE Int. Conference on Multimedia and Expo*, pp. 12311234, Jul. 2007.
- [9] J. Zhang, J. Li, and L. Zhang, Video watermark technique in motion vector, in *Proc. of XIV Brazilian Symp. on Comput. Graph. and Image Process.*, pp. 179182, Oct. 2001.
- [10] P. Meuel, M. Chaumont, and W. Puech, Data hiding in H.264 video for lossless reconstruction of region of interest, in *EUSIPCO 07: 15th European Signal Process. Conference*. Poznan, Poland: EURASIP, pp. 23012305, Sep. 2007.