Data Hiding in Encoded Video Sequences based on H.264

Xiaoni Li[1], Hexin Chen[1], Dazhong Wang[2], Tian Liu[1], Gang Hou[1]
1. School of Communication Engineering
Jilin University
Changchun, China
e-mail: lxn0612@163.com
2. Department of Electronic and Electrical Engineering
University of Sheffield
Sheffield, United Kingdom
e-mail: dazhongwang.student@gmail.com

Abstract—**Different data hiding methods would be required in different operating environments. To improve the quality of videos embedded by data, this paper will apply the differences of matching blocks in the H.264 encoder during the inter prediction stage and enforce the size of the matching blocks to implement data hiding. Bit rate increases after being embedded by data, to resolve the problem, this paper will mainly make use of the method of modifying the trailing coefficients sign-bit of fixed-length coding to control the bit rate in Context Adaptive Variable Length Coding (CAVLC). To equipoise the quality and bit rate of videos, this paper proposes an algorithm that embeds data separately in both of these two coding blocks.**

Keywords- H.264; Inter coding; CAVLC

I. Preface

For the past few years, illegal using of digitalized multimedia and problems of information security have been more serious. From the end of the 20^{th} century, data hiding technique got quick development which is helpful for resolving and preventing the digitalized data from being falsified, counterfeited, pirated and attacked. However, data hiding also received quick development in other application such as the data with caption, the video index, a role as a additional channel that transmit specific data in communicational system and so on. Currently, in the researching and application of the multi-media data hiding technique, state of digital video as a host media is widely used. However, the existence form of the digital video and compression could not be dissociable. To follow after formulating H.261, H.263, H.263 + and other compressed standards, ITU-T and ISO/IEC established a new video compressed coding standard, H.264, jointly. Moreover, the H.264 is the main content of the 10^{th} part of MPEG-4. H.264 is the latest video compressing coding standard[1] which is a mixed coding technique of DPCM and converting coding.

Basing on the above techniques, H.264 applies inter prediction, multiple reference frames, inter prediction with variable size, integer converting and adaptive entropy coding technique. Compared with other video coding standards, H.264 is equipped with much higher coding efficiency and network friendliness. All the above outstanding performances will provide H.264 standard important function in digital telecast, video real-time communication, network video-media transmitting and multimedia short messages.

As H.264 standard employs complicated coding and high compressing efficiency which causes H.264 standard very impressionable with tiny changes. This gives the data hiding great difficulties and challenges in which retrogression of video quality and increasing of bit rate after data embedding are the main problems. In traditional methods, normally, data hiding could be achieved by modifying the coefficients of the compressing domain. In reference [2], data is embedded through one-frame AC coefficient. In reference [3], divide the 15 AC coefficients in 4×4 coding block to two groups and modify the AC coefficient using weight function to hide data. In reference [4], [5] and [6], it also employs modifying the coefficients of the compressing domain to achieve embedding data.

The first problem that modifying the coefficients bring in is video quality degradation, the second one is bit rate increase as the modified coefficient coded with entropy coding[7]. To solve the problem of video quality degradation embedded by data, this paper adequately utilizes the size differences of matching blocks while H.264 encoder is inter prediction, through enforcing selecting the block size to hide data which could hide data as well as preserving the video quality. For the problem of increasing of bit rate after data embedding, this paper mainly applies trailing coefficients sign-bit in CAVLC to control bit rate. According to experiments, the method proposed in this paper is a feasible solution both in retrogression of video quality and increasing of bit rate after data embedding.

Plan of this paper is list below: in the second section, the basic framework of H.264 encoder will be introduced and the inter coding and the CAVLC will be introduced mainly. In the third section, data hiding applying inter coding in order to improve video quality will be introduced. In the fourth section, data hiding applying the CAVLC in order to control bit rate will be introduced. In the fifth section, the above two embedding algorithms will be integrated which

978-1-4244-5540-9/10/$26.00 ©2010 IEEE

provide better balance between video quality and bit rate. In the sixth section, experimental results and analysis will be represented.

II. H.264 ENCODER

H.264 encoder employs the joint coding method of converting and predicting. In Figure 1, frame/current F_n inserted is processed by encoder united by macro block. Firstly, operate the process with intra coding or inter coding method. The functional components of H.264 encoder is shown in figure 1:

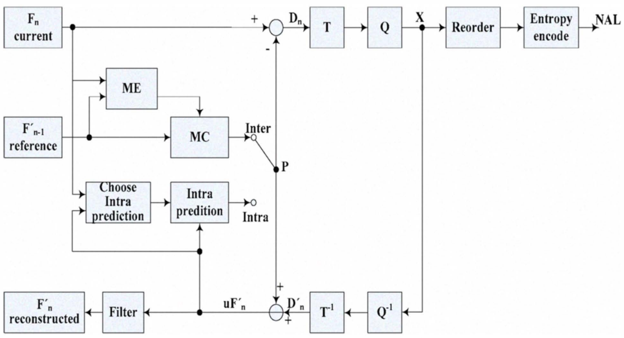

Figure 1. The H.264 encoder.

If the inter coding is applied, the predicting value PRED(indicated by P in the figure) could be obtained by the reference coded image with Motional Compensation(MC) in current image, where F'_{n-1} is the reference image. To enhance predicting efficient and compressing rate, the practical image could be selected in frames that have been coded, decoded, reconstructed and filtered in the past or future (which shows in sequences).

Conduct subtracting operation of the predicting value and current block to generate a residual block D_n. After being converted and quantified, a group of converted quantified coefficients X will be generated which will be supplied to transmission and memorizing after entropy coding, constituting a compressed code stream with a few edge data(e.g. quantitative parameters of predicting model and motional vectors) needed while coding and NAL(Network Adaptive Layer).

To provide farther reference images for prediction, the encoder should have the function of reconstructing images. Therefore, the block match image should become D'_n by inverse quantifying and inverse converting. And add the predicted value P to obtain uF'_n (unfiltered frame). In order to remove the noise caused in coding and decoding ring circuit and improve the imaging quality of the reference frame and the compressed image performance, a ring circuit filter should be set to generate an output F'_n which is the reconstructed image. And this image could be used as the reference image.

III. EMBEDDING DATA IN INTER CODING

The inter prediction is a predicting model applying coded video frame/field and motion compensation based on blocks. H.264 brings in several advanced techniques to realize optimization including variable-size block matching, motional vectors with high precision, multi-reference frame prediction and de-blocking filtering. Among all the modified inter-frame motional predicting techniques, H.264 changes the traditional models employing 16×16 or 8×8 simply in motional predicting matching units, but applies the method of macro-block partition and subsplit to generate multiple matching units(16×16, 16×8, 8×16, 8×8, 4×8, 8×4, 4×4) with different sizes and shapes[8][9]. In order to choose the best block size for a macro-block, the H.264 reference code makes use of the computationally intensive Lagrangian Rate-Distortion(RD) optimization, the general equation of which is:

$$J_{mode} = SSD + \lambda_{mode} \times R \quad (1)$$

Where λ_{mode} is the Lagrange multiplier used in mode decision, R reflects the number of bits associated with choosing the mode and macro-block quantizer value Qp, including the bits for the macro-block header, the motion vector(s) and all the DCT residue blocks. SSD is the sum of the squared differences between the original and the reconstructed block.

In this paper, while selecting the matching blocks, it is not operating selecting according to the Lagrangian optimized algorithm, but the embedded data factitiously. Firstly, code the seven kinds of unit matching blocks. While coding, according to different requests, different coding methods could be chosen such as selecting binary coding for simplicity. Also, according to different applying times of different matching blocks, the Hofmann coding could be used. After above processes, transform the data needed to be embedded to a group of binary numbers and according to the coding standard, match the blocks coded and data needed to be embedded. In this way, the matching blocks selected would be the blocks employed in inter coding.

For simplicity, in this paper, binary coding is employed, and only four matching units distributed with 2 bits respectively are selected to be coded. Then, transform the data needed to be embedded to a group of binary numbers and separate these numbers by 2 bits mutually. Finally, match every group of numbers along with the matching blocks one by one. Table 1 shows binary coding of four matching blocks in models. Figure 2 represents the matching blocks selecting process. Figure 3 shows the proceeding of data hiding.

TABLE I. BINARY CODES OF THE BLCOK TYPE

Block type	Binary code
16×16	00
16×8	01
8×16	10
8×8	11

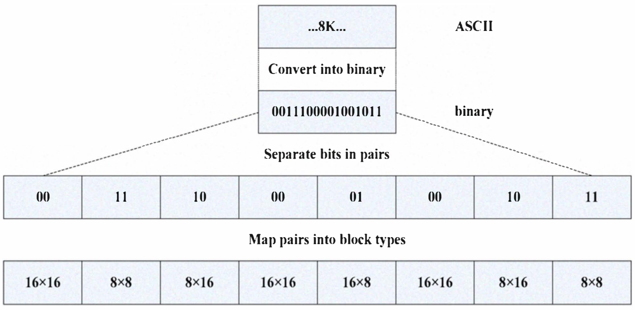

Figure 2. The message mapping into block types.

While the block is selected, the message is hided, the proceeding of data hiding is shown in figure3.

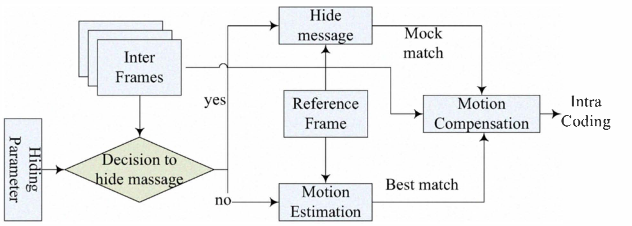

Figure 3. The proceeding of data hiding.

The proposed scheme was integrated within version 11.0 of the reference JVT software. Table 2 represents configuration parameters of the encoder. Table 3 shows PSNR and Bit Rate Variations.

TABLE II. CONFIGURATION PARAMETERS OF THE ENCODER

Profile	Baseline
Frames	200
Frame Rate	30 fps
Number of reference frames	2
8×8	11
ME Algorithm	Full Search
RD Optimization	Disabled
Rate Control	Disabled

TABLE III. PSNRADN BIT RATE VARIABLES

Sequence	PSNR variation(dB)	Bit Rate variation(%)
highway	0.00	-0.02
grandma	0.01	0.04
Mother and daughter	-0.01	0.20
Bridge-close	0.00	0.35
	Average: 0.00 dB	Average: 0.14 %

It could be made out that enforcing the size of matching blocks to achieve data hiding in inter prediction nearly has no effects on video quality, but increase embedding quantity greatly.

IV. EMBEDDING DATA IN CAVLC

CAVLC is applied in residual data coding of luminance and chrominance[10]. The residuals could express below characters after being converted and quantified: the non-zero coefficients of the 4×4 data being predicted, converted and quantified mainly center on low frequency domain and most high frequency coefficients are zeros; zig-zag scanning of the quantified data, non-zero coefficients around the DC coefficients are larger than others and most non-zero coefficients on high frequency are +1 and -1; the adjacent numbers of the 4×4 non-zero coefficients are correlative. CAVLC make adequate use of the data characters of residuals being integer converted and quantified to reduce the redundant information in data further.

CAVLC process is shown below: firstly, code the numbers of the non-zero coefficients(TC) and the trailing coefficients(T1); secondly, code every sign-bit of the trailing coefficients; thirdly, code the levels(Levels) of the non-zero coefficients except the trailing coefficients; fourthly, code the number of zeros(TotalZeros) generated before the last non-zero coefficient, finally, code the numbers of zeros(RunBefore) generated before every non-zero coefficient. The frequency components and the bit-length of every elements are represented in Table 4:

TABLE IV. THE BIT-LENGTH AND FREQUENCY COMPONENTS OF THE PARAMETERS

Coding Parameter	Values to be coded	Bit-length	Frequency components
Coeff token	TC and T1	Variable	All
Trailing coefficients sign-bit	Signs of T1	Fixed	Highest
Level prefix	Non-zero coefficients excluding T1	Variable	Medium to low
Level suffix	Non-zero coefficients excluding T1	Variable	Medium to low
Total zeros	Total number of zeros before the last non-zero coefficient	Variable	NA
Run before	Number of zeros before every non-zero coefficient	Variable	NA

While coding the sigh-bit of the trailing coefficients, generally, one bit could express the sigh-bit as 0 represents +1 and 1 represents -1. The coding sequence is according to the inverse scanning sequence which is starting from the high frequency data and number of the trailing coefficients is no bigger than 3. In the coding elements, only the trailing coefficients sigh-bit is the fixed-length coding and it is the highest frequency components of the QDCT coefficients. Therefore, operating data hiding in the coding of the trailing coefficients sigh-bit could both guarantee the video quality and control the increasing of bit rate perfectly. In this paper, it is operating data hiding in the coding of the trailing coefficients sigh-bit to achieve the aim that the bit rate embedded by data is unchanged.

In the experiments, S_{i1}, S_{i2}, S_{i3} represent the trailing coefficients sigh-bit of i^{th} 4×4 QDCT inverse-scanned, where S_{i1} is the sigh-bit of the highers frequency. As the maximum number of the trailing coefficients is 3, the maximum sigh-bit is 3. Assume B_i is a binary number that will be embedded into i^{th} 4×4 QDCT and the number of the trailing coefficients is not smaller than 1. the embedding process is shown below:

$$S'_{i1} = \begin{cases} \overline{S}_{i1}, & if(B_i \neq S_{i1}) \\ S_{i1} & otherwise \end{cases} \tag{2}$$

Where $\overline{S}_{i1}$ represents the complement of S_{i1}, S'_{i1} is the sigh-bit embedded by data and it is the same with the value of B_i. Therefore, while extracting data, the extracted embedded data $\hat{B}_i$ would be the sigh-bit of the first trailing coefficient of i^{th} QDCT block at the receiving terminal.

In the experiments, for simplicity, only one frame 176×144 image is selected to operate data embedding. Table 5 gives the changes of video quality and bit rate embedded by data.

TABLE V. THE CHANGES OF VIDEO QUALITY AND BIT-RATE

Video sequence	Number of embedded bits	Bit-rate increase	PSNR
Claire	930	0 %	43.2 dB
Carphone	870	0 %	43.4 dB
Mother	765	0 %	45.2 dB
Tempete	350	0 %	48.3 dB

It can be obtained from Table 5, embedding data in CAVLC trailing coefficient sigh-bit coding could hold the bit flows embedded by data and given the condition of plentiful embeddings, the PSNR value of videos is around 45dB. In this paper, if every macro-block is embedded by one bit information, the video quality is not affected basically. Therefore, it could be according to different applying environment, balance the embedding quantity and video quality.

V. DATA HIDING IN BOTH MODELS SYNCHRONOUSLY

Hiding data in inter-frame prediction could guarantee the video quality and hiding data in CAVLC could control the video bit rate that will not increase after data embedding. To balance the video quality and bit rate better, these two embedding algorithms could be combined and hide data in these two coding models synchronously. In the above method, according to the steps in the third section, operate data hiding in inter coding and according to the steps in the fourth section, operate embedding data in CAVLC. The proportion of the embedding data in each model could be different according to different applying environment. When it needs high video quality, embedding data in inter coding is mainly applied. When it needs strict requests of the bit rate, embedding data in CAVLC is mainly applied.

VI. CONCLUSION

According to different requests of video embedded by data in this paper, two embedding plans are proposed. To eliminate the effects of embedding data to video quality, the macro-block partition and the subsplit of H.264 standard is adequately applied in inter coding to form multiple matching units with different sizes and shapes. In this paper, being different from traditional selection of optimized matching blocks based on Lagrangian optimized algorithm, the matching blocks is selected according to corresponding relationship between the embedding data and the matching blocks, which means the embedding data selected the size of matching blocks. Therefore, there is no additional data embedded in actual video sequences, the video quality will not be affected and it will realize balance between the embedding quantity and the bit rate. To hold the video bit rate with data embedded, data embedding in CAVLC is proposed as the second embedding plan in the paper. As the trailing coefficient sigh-bit represents the maximum frequency part of DCT, changing of high-frequency component will not affect the video quality significantly. Moreover, as the trailing coefficient sigh-bit is coded with fixed length, modifying the sigh-bit will not increase the video bit rate. Based on the above reasons and given the condition that embedding data in sigh-bit could guarantee the video quality, non-increasing of bit rate could be achieved. To achieve better operation, the two embedding algorithms could be combined. According to different applying environment, proportion of embedding quantity of the two models is different. In this way, the video quality and the bit rate could be balanced better and data could be hidden preferably.

ACKNOWLEDGMENT

This work is supported by the National Science Foundation of China (Grant No.60832002, 60702036), Projects of International Cooperation and Exchanges NSFC (Grant No. 609111301281), and Basic Scientific Research Operation Expenses of Jilin University (Grant No. 421033135410).

REFERENCES

[1] C. Gomila and P. Yin, "New features and applications of the H.264 video coding standard," in Proc. of Intl. Conf. on Info. Tech.: Research and Education, pp. 6-10, 2003.

[2] Noorkami M, Mersereau R M. Compressed-domain video watermarking for H.264. Proceedings of ICIP. Genova, Italy, pp. 890-893, 2005.

[3] Jun Yan Hsieh, Long Wen Chang. A high capacity watermarking for 1.264/AVG based on frequency weighting. 2006 International

Symposium on Intelligent Signal Processing and Communication Systems.Yonago Convention Center, Tottori, Japan, pp. 899-902, 2006

[4] Golieri A, Nasiopoulos P , Wang Z J. An improved salar quantization-based digital video watermarking scheme for H.264/AVC. IEEE International Symposium on Circuits and Systems. Island of Kos, Greece, pp. 1434-1438, 2006

[5] Sharmeen Shahabuddin, Razib Iqbal, Shervin Shirmohammadi1,, Jiying Zhao. Compressed-domain Temporal adaptation-resilient watermarking for H.264 video authentication. ICME 2009. pp. 1752-1755, 2009

[6] Shinfeng D. Lin, Chih-Yao Chuang, and Hsiang-Cheng Meng. A Video Watermarking in H.264/AVC Encoder. 2009 Fifth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. pp. 340-343, 2009

[7] Maneli Noorkami,, and Russell M. Mersereau. Digital Video Watermarking in P-Frames with Controlled Video Bit-Rate Increase. IEEE Transactions on Information forensics and Security. Vol. 3, No.3. pp. 441-455, 2008

[8] Spyridon K. Kapotas and Athanassios N. Skodras. A New Data Hiding Scheme for Scene Change Detection in H.264 Encoded Video Sequences. ICME 2008. pp. 277-280, 2008

[9] Spyridon K. Kapotas, Eleni E. Varsaki , Athanassios N. Skodras. Data Hiding in H.264 Encoded Video Sequences. IEEE International Workshop on Multimedia Signal Processing. pp. 373-376, 2007

[10] Sung Min Kim, Sang Beom Kim, Youpyo Hong, and Chee Sun Won. Data Hiding on H.264/AVC Compressed Video. ICIAR 2007, LNCS 4633, pp. 698–707, 2007.