H.264/AVC Data Hiding Based on Intra Prediction Modes for Real-time Applications

Samira Bouchama, Latifa Hamami, and Hassina Aliane

Abstract— **The existing data hiding methods for the newest video codec H.264/AVC exploit its several modules such as the discrete cosine transform coefficients or the prediction modes. In this paper, a new data hiding approach is presented by exploiting the intra prediction modes for the 4x4 luminance blocks. The objective is to ensure a relatively high embedding capacity and to preserve the encoding and the decoding times in order to satisfy real-time applications. The intra prediction modes are divided into four groups composed of modes of close prediction directions. The data embedding is based on modifying modes of the same group in order to maintain visual quality and limit the number of additional calculation procedures. The increase of embedding capacity relies on the group composed on four modes since it allows the embedding of two bits per mode.**

Index Terms— **Data hiding, H.264/AVC, intra prediction, real time applications, authentication**

I. Introduction

DURING over a decade the field of digital data hiding of multimedia documents has known a growing activity. Already used in various applications such as secret communication, copyright protection and authentication, new applications are expected to appear with new constraints and challenges to meet [1].

Applied to compressed video, several methods have been proposed to the latest video coding standard H.264/AVC known for its coding efficiency compared to the previous video standards. Indeed, the H.264 supports advanced characteristics which play an important role in its performance, such as the 4x4 integer transform, spatial Intra prediction, variable block size motion, etc [2]. Even if the provided high compression ratio seems to limit the possibilities of embedding secret data in this codec, several information hiding methods have been proposed using either the integer discrete cosine transform (DCT) coefficients, the motion vectors, the intra pulse code modulation (IPCM) mode or the Intra/inter prediction modes [3], [4].

The use of intra prediction for data hiding is indeed motivating because it can preserve the video quality perfectly. Moreover, several approaches could be adopted to increase the embedding capacity or to control the increase in bitrate and encoding/decoding time which is demanded in case of real time applications [5]. However, it is difficult to reach a tradeoff between the different application constraints.

Few works of data hiding based on the Intra prediction modes have been proposed. Hu et al.[6] have presented data hiding method based on modifying the intra 4x4 prediction modes by a mapping between the modes and secret data. To improve the embedding capacity and to control the bitrate increase, the mapping rules have been improved by Yang et al [3]. In this scheme, each three 4x4 intra blocks that meet the embedding requirements for the watermark insertion constitute a group. Every two watermark bits are modulated to the prediction modes of these three blocks and only one block is necessary to change its prediction mode for embedding two watermark bits. In the embedding process the optimal mode is changed by the suboptimal one among odd or even modes according to the mapping rules and the modified blocks are then re-encoded.

The security of the secret data is ensured by encryption and scrambling on the one hand, and by the selection of the embedding position template which is controlled by a private key, on the other hand. The scheme promises a high PSNR and slight bitrate increase after watermark embedding. However these methods needs prior knowledge of the prediction modes and cannot be used for real time applications.

Xu et al. [5] used the same idea of data embedding based on changing the mode by the suboptimal one among odd or even modes according to the parity of the bit to be embedded. In this scheme, one mode can carry only one watermark bit. To enhance security, blocks can be randomly selected and the secret information is encrypted by a chaotic sequence. At the contrary of the previous methods, this one can be used for real time application.

In this paper we present a new approach of exploiting the 4x4 intra prediction modes to increase data hiding capacity for real time applications. The idea is based on dividing the intra prediction modes into four groups composed of modes that have close prediction directions. Embedding the secret data is based on applying modification between modes of the same group. Our method is based on the assumption that applying modification directly between modes of close directions could:

--Preserve the encoding and decoding times by reducing the number of additional calculation procedures generated by the search of the suboptimal mode as it is the case for the previous proposed methods [3], [5] and [6].

Manuscript received July 23, 2012; revised August 10, 2012. This work was supported in part by the Research Center on Scientific and Technical Information (CERIST) and the Polytechnic School of Algiers (ENP).

S.Bouchama and H.Aliane are with the Research Center on Scientific and Technical Information, Ben Aknoun, Algiers, phone: 00213 (0)21912136; fax: 00213(0)21912126; (e-mail: {bouchama,haliane}@cerist.dz).

L. Hamami is with the Department of Electronics, Polytechnic School of Algiers, Algeria (e-mail: latifa.hamami@enp.edu.dz).

--Reduce the preprocessing time by avoiding the exploitation of prior knowledge of modes used for the mapping [3], [5].

--Maintain the visual quality of the video sequence.

Furthermore, we have checked experimentally that modifying a mode to the one which has a close direction limits the increase in bitrate comparing to a random modification between modes.

The increase of embedding capacity relies on the group composed on four modes since it allows the embedding of two bits per mode instead of one as explained in the subsection III.A.

This paper is organized as follows: in the following section the Intra prediction is introduced; then, the proposed data hiding approach is detailed in section III. Results and analysis are presented in section IV and finally, conclusions are drawn in section V.

II. THE INTRA PREDICTION

The H.264/AVC standard uses the spatial correlation property through an Intra prediction. A block (or macroblock) of interest can be predicted from blocks generally located above and to the left since they have already been encoded and reconstructed. The subtraction from the current block is then coded and represented with a reduced number of bits compared to the one used for the direct processing of the block itself. Intra coding includes the following modes:

Intra 4x4 mode in which each 4x4 luminance block can be coded using one of nine prediction modes (Fig.2); the Intra 16x16 mode is recommended for image regular areas; the chroma samples prediction; and the prediction mode based on blocks of size 8 by 8 pixels (8x8) which is basically identical to the 4x4 prediction mode but used only in the H.264/AVC High Profile [2], [7].

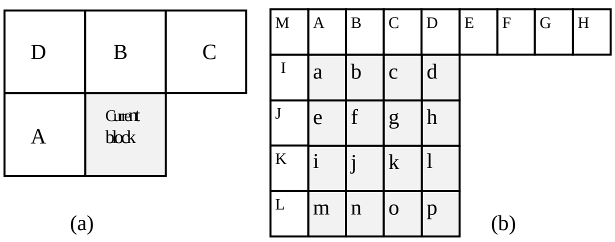

Fig. 1. (a) Neighboring blocks, (b) Neighboring samples

A. The intra prediction modes

The 4x4 intra prediction modes are more convenient for areas of significant details in the picture. The macroblock is divided into sixteen 4×4 luma blocks and for block the sixteen samples (from a to p as illustrated in Fig.1.(b)) are predicted depending on the previously encoded and reconstructed samples (A-M). Nine modes are defined in the intra prediction of the 4x4 luminance block. Fig. 2 represents the nine prediction modes with the corresponding prediction directions. In DC prediction the mean of the left handed and upper samples are used to predict the entire block [2], [7].

In order to compress the prediction mode, the H.264/AVC supports the most probable mode (MPM) which is an estimation of the prediction mode using the above and left adjacent blocks (A and B in Fig.1.(a)) of the current block. The MPM is the minimum of the prediction modes of A and B. If one of these prediction modes is not available the corresponding value is set to 2 (DC mode) [2].

B. The mode decision for the 4x4 Intra prediction modes

Rate distortion optimization (RDO) technique is used to achieve the best coding performance. For Intra 4x4 prediction, all the mode combinations are performed and the optimum mode decision for each of the sixteen 4x4 sub-blocks is obtained by minimizing the following formula [8]:

$$J = D + \lambda_{mode} * R. \qquad (1)$$

D and R represent respectively the distortion and the estimated bitrate for encoding the mode and λ_{mode} denotes the Lagrangian multiplier which depends on quantization parameter QP. λ_{mode} can be calculated with,

$$\lambda_{mode,I,P} = 0.85 \; * \; 2^{(QP-12)/3} \qquad (2)$$

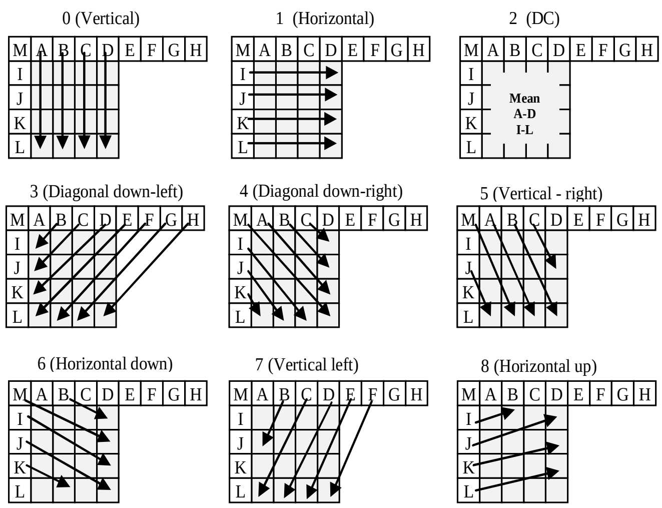

Fig. 2. The 4x4 intra prediction modes

III. PROPOSED APPROACH

A. General presentation

Embedding the secret data is performed in I frames using the 4x4 intra prediction modes. The 16 x 16 intra prediction modes are not worth exploiting because they can carry only few bits since the number of modes is only four and these macroblocks are generally less numerous in common frames. Furthermore, they concern homogeneous areas for which the human visual system (HVS) is more sensitive to small degradations.

Observing the nine 4x4 intra prediction modes (Fig. 2), it is possible to notice that there are prediction modes that have close directions and which use the same neighboring samples to calculate the prediction. In this case, it is possible to avoid any problem of unavailability if the embedding proceeding is based on applying modification between modes of close directions. We have verified experimentally that applying such modification between modes of close

ISBN: 978-988-19251-6-9
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

directions gives better results in terms of bitrate comparing to a random modification of modes.

Based on this idea, groups of modes could be defined as follows:

Group 1: Modes 1 and 8.

Group 2: Modes 3 and 7.

A third group can contain the modes 4, 5 and 6. However, in order to enhance the embedding capacity a fourth mode is added in this group so that it would be possible to embed two bits per mode as explained in the next subsection. The mode that can be added is the DC mode because the availability of the 4th, 5th and 6th implies the availability of the second mode (although, for this mode, the encoder assigns a default value in case of unavailability of neighboring samples), however it is more difficult to modify the DC mode to modes 4, 5 or 6 because the D block (Fig.1.(a)) may not be available. To bypass this problem, a gap is created in the mode 2 so that this case will never occur. Thus, we proceed as follows: To preserve the bitrate, the mode 2 is modified to the most probable mode (MPM) in case MPM is different from 2, otherwise it is modified to the mode 1 or 0 according to the availability.

In order to exploit all the nine modes, it is possible to use the vertical mode to embed two bits by modifying the mode 0 to the mode 2. Thus, we can define the third and the fourth groups as follows:

Group 3: Modes 2, 4, 5 and 6.

Group 4: Modes 0 and 2.

The modification of the modes is validated by the rate distortion optimization (RDO).

B. *Embedding process*

The embedding process is applied to the different modes as follows:

For the group 1, the horizontal mode (1) is unchanged if the watermark bit is equal to 1, and it is modified to (8) if the watermark bit is equal to 0. Similarly, the mode (8) remains unchanged if the embedded bit is 0, and it is modified to 1 if the bit is equal to 0. For each case (mode = 1or mode =8), the embedding is not performed if the MPM is equal to 1 or 8 in order to limit the increase in the bitrate and to ensure a reliable detection.

The same principle is applied to group 2 as illustrated in Table 1.

For the modes 4, 5 and 6, two bits are embedded in the mode and the introduced modification is presented in Table 2. For example, if the best mode is equal to 4 and the MPM is different from 4, 5 or 6, the mode 4 becomes 2 in case two successive watermark bits are equal to “00”, and the mode becomes equal to 6 if the watermark bits are equal to “11”.

Finally, if the best mode is equal to the vertical mode 0 and the MPM is different from 0, the best mode is modified to DC mode only if the watermark bits W_i and W_{i+1} are equal to 0, otherwise, the best mode is not changed. This is justified by the availability of the DC mode as mentioned in the previous subsection, and also by the fact that the vertical mode cannot be changed to the other modes without checking the availability of the modes and without risking an important increase in the bitrate.

C. *Retrieving process*

At the decoding stage and after the entropy decoding, the retrieving process can be done easily from the compressed intra-coded 4x4-blocks since we can deduce the values of the embedded bits according to the modes as presented in Tables I and II. Of course, at this step we maintain the condition that for a given mode, the most probable mode mustn’t be equal to this mode or another one from the same group, except for the mode DC which undergoes a special processing as explained in section III.A.

TABLE I
EMBEDDING PROCESS FOR THE GROUPS 1 AND 2

Watermark bits	Modification of modes: 3 and 7	1 and 8
0	7	8
1	3	1

TABLE II
EMBEDDING PROCESS FOR THE GROUPS 3 AND 4

Watermark bits	Modification of modes: 4, 5 and 6	0
00	2	2
01	4	0 *(not modified)*
10	5	0
11	6	0

D. *Data hiding security and robustness*

In the described embedding process, the embedding position is known and the hidden bits can be easily extracted. Thus, the security of the secret information in based on cryptographic tools that can be performed during a preprocessing step. Moreover, it is possible to change the mapping used to embed the watermark bit (Tables I and II) but it needs to be transmitted as side information. We can also use a random selection of the blocks to reinforce the security aspect but in this case the embedding capacity would be reduced.

Concerning the robustness of the approach, any possible attack of the video stream would cause trouble at the decoding step. However, the most obvious attack which challenges the robustness must be the re-encoding of the video after the stream decoding. For this reason the proposed approach is more suitable for applications such as covert communication or authentication of the compressed video rather than copyright protection.

IV. RESULTS AND ANALYSIS

The data hiding method has been tested on eight QCIF video sequences: Foreman, Silent, Container, Coast Guard, Carphone, Grandma, News and Bridge-close. For each sequence, 199 frames were coded at a frame rate of 30 pictures/s. With the main profile, the group of pictures (GOP) structure was ‘‘IBPBPBPBPB’’.

For tests, we have used a pseudo random binary sequence as secret data which were inserted in I frame according to the method described previously, by integrating the algorithm in the H.264/AVC codec. Results are tabulated in Table III.

The embedding capacity depends on the number of the 4x4 intra coded macroblocks in I frames and also on the existence of modes carrying 2 bits (such as 4, 5 and 6).

TABLE III
RESULTS OBTAINED FOR EIGHT VIDEO SEQUENCES

Video sequences	Embedding Capacity (bits)	Increase in Bitrate (%)	Decrease in PSNR (dB)
Foreman	16291	5.19	0.06
Grandma	11412	5.11	0.06
Container	8331	5.50	0.05
Coast-Guard	9974	2.18	0.05
Carphone	13436	4.03	0.05
Silent	17716	5.72	0.05
News	11645	4.45	0.05
Bridge-close	12007	3.26	0.04

Results show that the lowest embedding capacity is 8331 bits, it is given for Container sequence which is characterized by minimal details implying a relatively important number of 16x16 coded macroblocks that are not used for data embedding. The highest embedding capacity is 17716 obtained for Silent sequence, which contains, on the contrary, significant details. This embedding capacity is spread over the 20 I frames contained in the test sequences of 199 frames.

We have implemented the method proposed in [5] in the same conditions. Results show that the embedding capacity obtained in our approach is higher (Fig.3) and the increase in bitrate obtained by implementing method [5] is close to our results. It varies between 3.4 and 5.6 for the sequences mentioned in Fig. 3.

We have also compared our approach to methods proposed in [3] and [6] in the same conditions, our approach presents the highest payload as shown in Fig.3.

However, according to the results, this method is not convenient for sequences with minimum details such as Grandma and Container, because the provided embedding capacity may not be enough, relatively to the high obtained bitrate, comparing to sequences containing more details such as Silent or Foreman.

Indeed, as shown in Table III, the bitrate is relatively high. It varies between 2.18% and 5.72 % noticed for Silent sequence characterized by a relatively high embedding capacity. Thus this approach may not be suitable for applications demanding in bitrate.

By way of indication, the method proposed in [6] presents results related to the increase in bitrate, that vary between 2.3% and 3.3 % for the sequences mentioned in Fig. 3, whereas the results are less than 2% for the method proposed in [3] for the same sequences.

However, our objective was to avoid using any special mapping that requires prior knowledge of the frames features, or prior encoding of the video, by proposing a method that achieves a simple embedding, easy to implement and suitable for reel time applications as mentioned later. Meanwhile, a relatively high embedding capacity is ensured especially for the highly textured sequences, and using the method based on the suboptimal mode during encoding process (for real time applications), in the same conditions, doesn't achieve this embedding capacity for almost the same increase in bitrate.

The objective assessments of the video measured by the PSNR have shown that the video quality is maintained since the decrease of the PSNR values, comparing to the unwatermarked sequences, are in the range of hundredths (for example, for Container sequence, the PSNR decreased by only 0.05 dB). The subjective quality is perfectly preserved as shown in Fig. 4 which represents the I frames of Container and Bridge-close sequences before and after data embedding.

The increase in the encoding time is negligible as presented in Fig.5. Experiments have been performed on a PC Pentium 4.3 GHz, 1 Go of RAM running under Windows XP.

Finally, the detection step is reliable for all the test sequences without applying any attack. The data extraction can be done easily without a complete decoding of the sequence, after the entropy decoding stage as explained in the subsection III.C. Moreover, results have shown that the watermark doesn't affect the decoding time which is perfectly maintained.

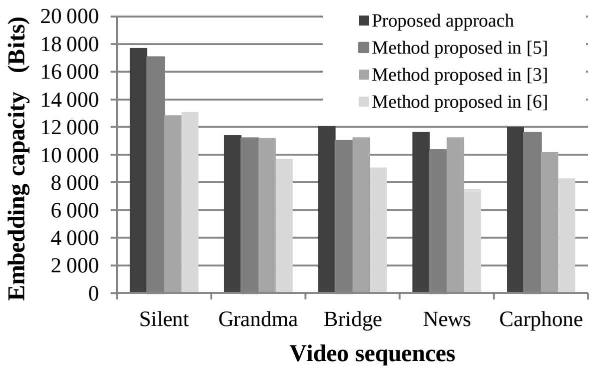

Fig. 3. Embedding capacities of five sequences.

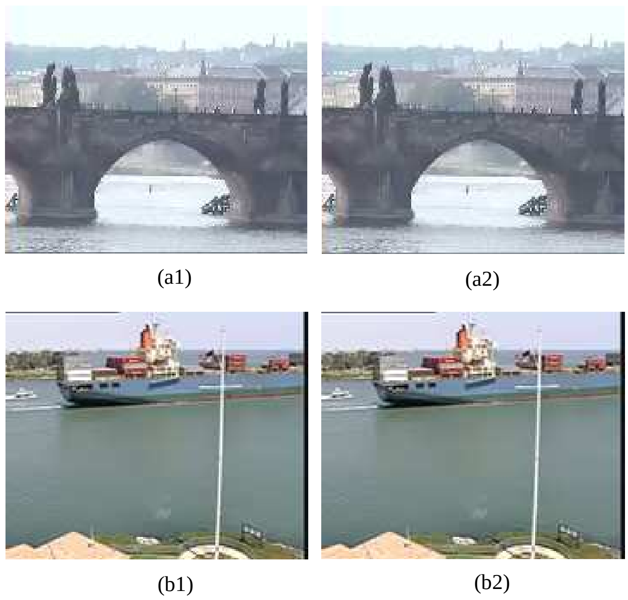

Fig. 4. (a1) and (b1) represent the unwatermarked I frames of the sequences Bridge and Container respectively, and (a2) and (b2) are the related watermarked frames.

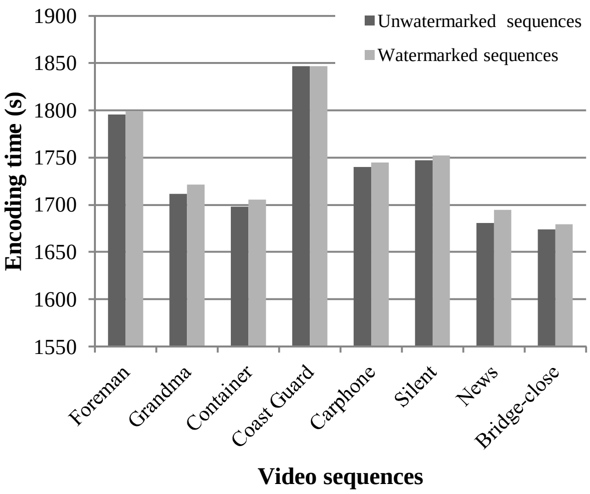

Fig. 5. The encoding time of the unwatermarked and the watermarked test sequences.

V. Conclusion

In this paper, a new approach in exploiting the 4x4 intra prediction modes to embed secret data has been presented. Results have shown that a high embedding capacity can be reached, especially for detailed sequences, while a negligible increase in the encoding time is noticed and the decoding time is not disturbed.

The use of intra prediction to embed secret data in the video codec H.264/AVC allows a better preservation of the video quality comparing to the previous data hiding methods such DCT coefficients based data hiding methods.

Therefore, the use of Intra prediction modes is indeed motivating for all the possibilities it offers in terms of preserving video quality, enhancing embedding capacity and maintaining encoding and decoding times. However, this approach still needs improvements in order to control the increase in bitrate.

References

[1] J.L. Dugelay and G. Doerr, “A guide tour of video watermarking,” *Signal Process Image Comm*, vol.18, pp.263–82, 2003.

[2] I.E.G. Richardson, *H.264 and Mpeg-4 video compression: video coding for next-generation multimedia*, The Robert Gordon University, Aberdeen, UK, Wiley, 2003.

[3] G. Yang, J. Li, Y. He, and Z. Kang, “An information hiding algorithm based on intra-prediction modes and matrix coding for H.264/AVC video stream,” *Int. J. Electron. Commun*, vol.64, pp.331–337, 2010.

[4] S.K. Kapotas and A.N. Skodras, “Real time data hiding by exploiting the IPCM macroblocks in H.264/AVC streams,” *J. Real-Time Image Proc*. vol.4, pp.33–41, 2009.

[5] D. Xu, R.Wang, and J.Wang, “Prediction mode modulated data-hiding algorithm for H.264/AVC,” *J Real-Time Image Proc*. vol.5, 2010.

[6] Y. Hu, C. Zhang, and Y.Su, “Information hiding based on intra prediction modes for H.264/AVC,” in *2007 IEEE Int. Conf. on Multimedia and Expo (ICME)*, pp.1231–1234.

[7] T. Weigand, G. Sullivan, G. Bjontegaard, and A.Luthra, “Overview of the H.264/AVC video coding standard,” *IEEE Trans. on Circuits and System for Video Technology*, vol. 13, pp.560–576, 2003.

[8] P. Wang, Z. Zheng, and L. Li, “ A video watermarking scheme based on motion vectors and mode selection,” in *2008 Proc. of IEEE Int. Conf. on Computer Science and Software Engineering*, pp.233-237.