A novel two-dimensional histogram modification
for reversible data embedding into stereo H.264 video

Juan Zhao · Zhi-Tang Li · Bing Feng

Abstract Histogram modification (HM) is an efficient technique for reversible data embedding (RDE) into stereo H.264 video. Nevertheless, in the traditional HM, the value of a coefficient is singly modified for embedding data, i.e., the relationships among coefficients are not considered. In this paper, to make full use of the correlation between two coefficients, we present a novel two-dimensional (2D) HM strategy for stereo H.264 video. Firstly, two quantized discrete cosine transform (QDCT) alternating current (AC) coefficients are randomly selected from each embeddable 4×4 luminance block. The values of coefficient pairs are classified into nonoverlapping sets. According to the sets of coefficient pairs, the generated 2D histogram is modified to embed data. When the value of one QDCT AC coefficient is modified by adding or subtracting 1, only one data bit at most could be hidden by using the traditional HM, whereas up to three bits of information can be simultaneously embedded by employing the proposed scheme. The better capacity-distortion performance of the proposed strategy is proved by experiments.

1 Introduction

Reversible data embedding (RDE) algorithm is a promising technology which can be used to hide information into a digital medium in a reversible manner. The original medium can be completely recovered after the hidden data has been extracted [1]. Therefore, RDE methods can be applied in the legal, the military, and the medical fields, where any permanent distortion due to hiding information is intolerable [2]. H.264 is a state-of-the-art video compression standard and has become the most widely deployed video codec. At the H.264 encoder, a residual pixel block is obtained by subtracting a prediction block from its original pixel block in YC_bC_r (YUV) video. After lossy compression (discrete cosine transformation (DCT) and quantization), the residual block becomes a quantized DCT (QDCT) block, which has one direct current coefficient and

J. Zhao · Z.-T. Li (✉) · B. Feng
Huazhong University of Science and Technology, Wuhan 430074, China
e-mail: lzt@hust.edu.cn

some alternating current (AC) coefficients. After entropy encoding (lossless compression) of each QDCT macro block (MB), YUV video is encoded into an H.264 video. Therefore, the information hidden by changing QDCT coefficients can be fully extracted after entropy decoding. It is the most common to embed data into QDCT coefficients, but the distortion caused by hiding data will spread and accumulate (called distortion drift). When RDE techniques are used for video, we can save and enjoy important videos without information and distortion caused by hiding data. Consequently, there will not be too many network videos with secret information so that it is difficult for others to find stego cover. In addition, RDE methods can also be used for video error concealment [3–5].

It is necessary that RDE algorithms assure large capacity and low distortion, which are two competitive factors. A higher embedding capacity will generally result in a higher degree of distortion. Recently, in order to hide more capacity of information with lower distortion of carrier, many researchers have presented various RDE tactics such as lossless compress based methods [6–8], integer transform based methods [9–12], difference expansion based methods [3, 13–22], and histogram modification (HM, also called histogram shifting) based methods [4, 5, 23–35]. Among these schemes, the difference expansion based methods and the HM based methods are two leading techniques which could be used for video RDE [3–5, 17, 20, 22].

In the difference expansion based RDE algorithm presented by Tian [13], the original image is split into pixel pairs, and the difference between two neighboring pixel values is doubled to hide one data bit. In order to solve the overflow and the underflow problems, all expandable locations are recorded by a location map, which is used by most RDE methods. On this basis, Alattar [14] generalized the difference expansion method to three or four pixels. In addition, Kamstra and Heijmans [15] used Low-pass image to predict expandable locations in order that the location map can be compressed.

In typical HM based RDE algorithm, data is embedded into the peak points of the histogram of a medium [4, 23]. In order to embed one data bit, the value of each pixel is changed at most by adding or subtracting 1, which limits the distortion induced by embedding information. The quantity of embeddable data bits depends on the pixel number of the peak point in the histogram. Li et al. [31] proposed a general framework of HM based RDE. A shifting function and an embedding function could be easily designed for a RDE method.

However, for RDE algorithms based on conventional HM, when 1 is added to or subtracted from the carrier, only one information bit could be hidden at most. Additionally, different from image RDE, the distortion drift caused by embedded information will vastly affect the video quality [36, 37]. In order to solve this limiting problem of the capacity-distortion performance and the embedding efficiency, an efficient two-dimensional (2D) HM strategy uniting a pair of QDCT AC coefficients is proposed for stereo H.264 video in this paper. The main contributions of this work are as follows. First, when 1 is added to or subtracted from the value of one QDCT AC coefficient, three data bits at best could be hidden by using the presented method. Second, we embed data into 4×4 blocks which meet our conditions to avoid intra-frame distortion drift. Third, random selection of embeddable blocks and coefficients reduces the statistical distortion.

The remainder of the paper is organized as follows. Section 2 introduces the main idea of HM. In Section 3, the proposed scheme and its implementation details are described. Experimental results validating the effectiveness of our scheme are shown in Section 4. Last, Section 5 draws conclusions.

2 2D HM based RDE scheme

The distribution of QDCT coefficients is shown in Table 1. The probability of zero coefficients is denoted as p_0. For the nine test videos, p_0 is between 82.87 and 99.39 %. The value of overwhelming majority of QDCT coefficients is zero, which indicates that the QDCT coefficient histogram is very sharp. Accordingly, good embedding performance can be achieved by using HM to hide information into QDCT coefficients.

Figure 1 shows a 4×4 QDCT luminance block with zigzag scan, which is a specific sequential ordering of transform coefficient levels from the lowest spatial frequency to the highest. In the 4× 4 QDCT block, Y_0 is a direct current coefficient, Y_1, Y_2, …, and Y_5 are AC coefficients in the low frequency area. Most energy of a block is stored in Y_0 and low frequency area. Coefficients in high frequency such as Y_{13}, Y_{14}, and Y_{15} have little energy. Embedding data into high frequency area causes less impact on the video quality compared with hiding data into low frequency area. However, less statistical distortion can be gained by hiding information into the low frequency area compared with embedding data into the high frequency region. Users can choose different embeddable coefficients according to different application environments. For example, low frequency region could be utilized by covert communication for strong capability of avoiding detection; high frequency region could be used to hide motion vectors for error concealment.

2.1 Usual HM

According to Ni et al.'s HM algorithm [23], data could be embedded into QDCT coefficient through shifting one-dimensional (1D) coefficient histogram. The embedding procedure is shown by Fig. 2 and (1), where Y_{s1} is the $s1$th QDCT coefficient in a 4×4 QDCT luminance block, Y'_{s1} is the marked QDCT coefficient, and $m_i \in \{0,1\}$ is a to-be-hidden data bit. When the value of QDCT coefficient Y_{s1} is 0 or −1, a data bit can be hidden.

$$Y'_{s1}= \begin{cases} Y_{s1}, & \text{if } (Y_{s1}=0 \text{ or } -1) \wedge (m_i=0) \\ 1, & \text{if } (Y_{s1}=0) \wedge (m_i=1) \\ -2, & \text{if } (Y_{s1}=-1) \wedge (m_i=1) \\ Y_{s1}+1, & \text{if } Y_{s1}>0 \\ Y_{s1}-1, & \text{if } Y_{s1}<-1 \end{cases} \tag{1}$$

Table 1 Average percent (%) of different QDCT coefficient values in each P_0 frame

| Sequences	The value of QDCT coefficient												
	−6	−5	−4	−3	−2	−1	0	1	2	3	4	5	6
Akko&Kayo	0	0	0	0.01	0.07	1.53	96.86	1.41	0.1	0.01	0	0	0
Ballroom	0	0.01	0.02	0.07	0.3	2.87	93.27	2.94	0.38	0.1	0.03	0.01	0
Crowd	0	0.01	0.04	0.11	0.45	4.05	90.64	3.97	0.52	0.14	0.05	0.02	0.01
Exit	0	0	0.02	0.06	0.24	2.1	95.1	2.15	0.24	0.05	0.03	0.01	0
Flamenco	0	0	0.01	0.04	0.2	2.44	94.52	2.51	0.24	0.04	0.01	0	0
Objects	0	0.05	0.14	0.34	1.12	6.57	82.87	7.21	1.11	0.32	0.18	0.06	0.02
Race	0.01	0.02	0.05	0.11	0.5	4.25	89.62	4.76	0.52	0.11	0.04	0.02	0.01
Rena	0	0	0	0	0	0.35	99.39	0.25	0	0	0	0	0
Vassar	0	0	0.01	0.03	0.23	3.15	93.13	3.14	0.26	0.04	0.01	0	0
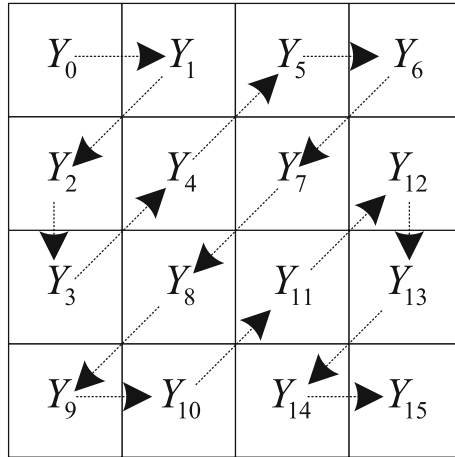
Fig. 1 A 4×4 QDCT luma block with zigzag scan

Accordingly, the information could be extracted from the marked QDCT coefficient Y'_{s1} based on the extracting procedure:

1. if $Y'_{s1}=0$ or -1, the original coefficient $Y_{s1}=Y'_{s1}$, and the extracted data bit $m_i=0$.
2. if $Y'_{s1}=1$ or -2, the original coefficient $Y_{s1}=Y'_{s1}-1$, and the extracted data bit $m_i=1$.
3. if $Y'_{s1}>1$, there will be no hidden data bit, and the original coefficient $Y_{s1}=Y'_{s1}-1$.
4. if $Y'_{s1}<-2$, there will be no hidden data bit, and the original coefficient $Y_{s1}=Y'_{s1}+1$.

In this scheme, the 1D coefficient histogram is defined by

$$h(r_1) = \#\left\{ Y_{s1} \middle| Y_{s1} = r_1 \right\}, \tag{2}$$

where # is the cardinal number of a set, r_1 is an integer.
The embedding capacity denoted as EC is

$$EC = h(0) + h(-1). \tag{3}$$

For QDCT coefficients, the embedding distortion denoted as ED in terms of l^2-error can be formulated as

$$ED = \frac{1}{2}h(0) + \frac{1}{2}h(-1) + \sum_{r_1=-\infty}^{-2} h(r_1) + \sum_{r_1=1}^{+\infty} h(r_1). \tag{4}$$

The usual 1D HM [23] could be employed in 2D HM, where 2D histogram is defined as

$$y(r_1,r_2) = \#\left\{ (Y_{s1}, Y_{s2}) \middle| Y_{s1} = r_1, Y_{s2} = r_2 \right\}, \tag{5}$$

where Y_{s2} is the s2th QDCT coefficient in a 4×4 QDCT luminance block, r_2 is an integer. Selecting two QDCT coefficients Y_{s1} and Y_{s2}, the sender could embed a data bit into each coefficient as 1D HM shown in Fig. 2. Then the mapping of coefficient pairs could be demonstrated by Fig. 3, where data is singly embedded into each coefficient all the same. For example, when the value of the QDCT coefficient pair (Y_{s1}, Y_{s2}) is $(0,1)$, in the 1D HM, the
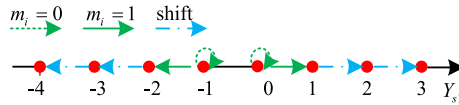
Fig. 2 Conversion of QDCT coefficients in usual 1D HM

value of QDCT coefficient Y_{s1} is expanded to 0 or 1 for hiding a data bit $m_i \in \{0,1\}$, and the value of QDCT coefficient Y_{s2} is shifted to 2. Accordingly, in the 2D HM, the value (0, 1) of QDCT coefficient pair (Y_{s1}, Y_{s2}) will be expanded to (0, 2) if the information bit m_i is 0, and (1, 2) when m_i is 1.

2.2 Presented 2D HM

2.2.1 Data embedding scheme

In the customary 2D HM, each QDCT coefficient is separately modified for hiding a data bit, which is not efficient enough for embedding information. In order to improve the capacity-distortion performance and the embedding efficiency, we present a novel 2D HM illustrated in Fig. 4. The set of all the points shown in Fig. 4 is denoted as A, which is divided into nineteen subsets denoted as A_1, A_2, …, and A_{19} for comparing the embedding capacity and distortion in Section 2.2.2 (in fact, the set A just needs to be divided into thirteen subsets for hiding data). These sets are defined as follows:

$A_1 = \{(0, 0)\}$
$A_2 = \{(Y_{s1}, 0) \mid Y_{s1} > 0\}$
$A_3 = \{(-1, 0)\}$
$A_4 = \{(Y_{s1}, 0) \mid Y_{s1} < -1\}$
$A_5 = \{(0, 1)\}$
$A_6 = \{(0, -1)\}$
$A_7 = \{(0, 2)\}$
$A_8 = \{(0, -2)\}$
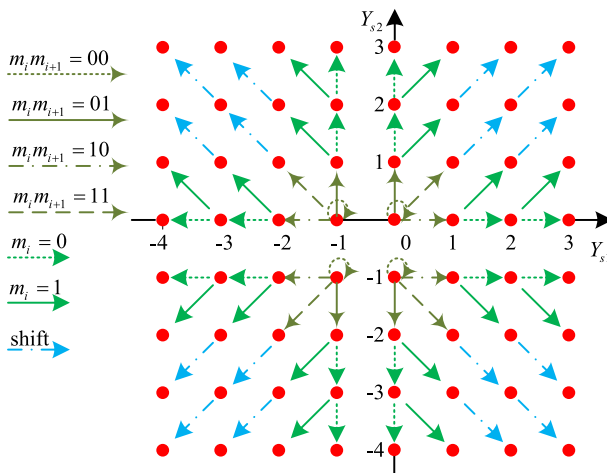

Fig. 3 Conversion of QDCT coefficient pairs in customary 2D HM

$A_9 = \{(0, Y_{s2}) \mid Y_{s2} > 2\}$
$A_{10} = \{(0, Y_{s2}) \mid Y_{s2} < -2\}$
$A_{11} = \{(Y_{s1}, Y_{s2}) \mid Y_{s1} > 0, Y_{s2} > 0\}$
$A_{12} = \{(Y_{s1}, -1) \mid Y_{s1} > 0\}$
$A_{13} = \{(Y_{s1}, Y_{s2}) \mid Y_{s1} > 0, Y_{s2} < -1\}$
$A_{14} = \{(-1, Y_{s2}) \mid Y_{s2} > 0\}$
$A_{15} = \{(Y_{s1}, Y_{s2}) \mid Y_{s1} < -1, Y_{s2} > 0\}$
$A_{16} = \{(-1, -1)\}$
$A_{17} = \{(Y_{s1}, -1) \mid Y_{s1} < -1\}$
$A_{18} = \{(-1, Y_{s2}) \mid Y_{s2} < -1\}$
$A_{19} = \{(Y_{s1}, Y_{s2}) \mid Y_{s1} < -1, Y_{s2} < -1\}$

According to the set which the value of the QDCT coefficient pair (Y_{s1}, Y_{s2}) belongs to, we aim to embed information bits as many as possible at the same time. The embedding procedure could be described as follows:

1. if the QDCT coefficient pair $(Y_{s1}, Y_{s2}) \in A_1$, the marked coefficient pair denoted by (Y'_{s1}, Y'_{s2}) will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}), & \text{if } m_i m_{i+1} = 01 \\ (Y_{s1} + 1, Y_{s2}), & \text{if } m_i m_{i+1} = 10 \\ (Y_{s1}, Y_{s2} + 1), & \text{if } m_i m_{i+1} = 11 \\ (Y_{s1} - 1, Y_{s2}), & \text{if } m_i m_{i+1} m_{i+2} = 000 \\ (Y_{s1}, Y_{s2} - 1), & \text{if } m_i m_{i+1} m_{i+2} = 001 \end{cases} \quad (6)$$
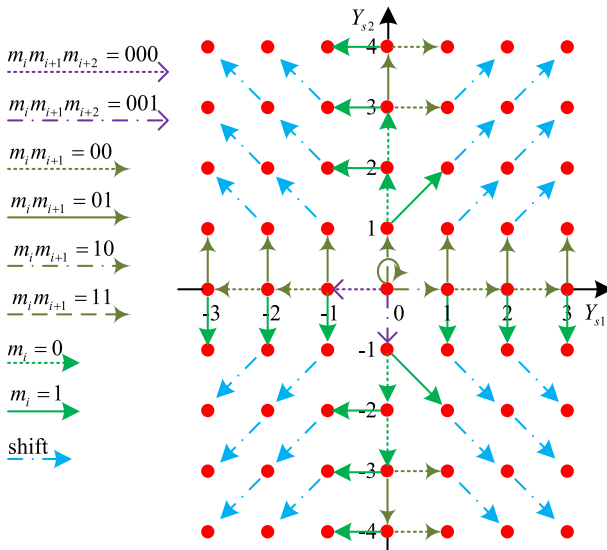

Fig. 4 Conversion of QDCT coefficient pairs in the proposed 2D HM

2. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_2$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}-1), & \text{if } m_i = 1 \\ (Y_{s1}+1, Y_{s2}), & \text{if } m_i m_{i+1} = 00 \\ (Y_{s1}, Y_{s2}+1), & \text{if } m_i m_{i+1} = 01 \end{cases} \qquad (7)$$

3. if the coefficient pair (Y_{s1}, Y_{s2}) belongs to A_3 or A_4, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}-1), & \text{if } m_i = 1 \\ (Y_{s1}-1, Y_{s2}), & \text{if } m_i m_{i+1} = 00 \\ (Y_{s1}, Y_{s2}+1), & \text{if } m_i m_{i+1} = 01 \end{cases} \qquad (8)$$

4. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_5$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}+1), & \text{if } m_i = 0 \\ (Y_{s1}+1, Y_{s2}+1), & \text{if } m_i = 1 \end{cases} \qquad (9)$$

5. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_6$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}-1), & \text{if } m_i = 0 \\ (Y_{s1}+1, Y_{s2}-1), & \text{if } m_i = 1 \end{cases} \qquad (10)$$

6. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_7$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}+1), & \text{if } m_i = 0 \\ (Y_{s1}-1, Y_{s2}), & \text{if } m_i = 1 \end{cases} \qquad (11)$$

7. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_8$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}, Y_{s2}-1), & \text{if } m_i = 0 \\ (Y_{s1}-1, Y_{s2}), & \text{if } m_i = 1 \end{cases} \qquad (12)$$

8. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_9$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}-1, Y_{s2}), & \text{if } m_i = 1 \\ (Y_{s1}+1, Y_{s2}), & \text{if } m_i m_{i+1} = 00 \\ (Y_{s1}, Y_{s2}+1), & \text{if } m_i m_{i+1} = 01 \end{cases} \qquad (13)$$

9. if the coefficient pair $(Y_{s1}, Y_{s2}) \in A_{10}$, the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}-1, Y_{s2}), & \text{if } m_i = 1 \\ (Y_{s1}+1, Y_{s2}), & \text{if } m_i m_{i+1} = 00 \\ (Y_{s1}, Y_{s2}-1), & \text{if } m_i m_{i+1} = 01 \end{cases} \qquad (14)$$

10. if the coefficient pair (Y_{s1}, Y_{s2}) belongs to $A_{11}, A_{12}, A_{13}, A_{14}, A_{15}, A_{16}, A_{17}, A_{18}$ or A_{19}, no message could be hidden, and the marked coefficient pair will be

$$\left(Y'_{s1}, Y'_{s2}\right) = \begin{cases} (Y_{s1}+1, Y_{s2}+1), & \text{if } (Y_{s1}, Y_{s2}) \in A_{11} \\ (Y_{s1}+1, Y_{s2}-1), & \text{if } (Y_{s1}, Y_{s2}) \in A_{12} \cup A_{13} \\ (Y_{s1}-1, Y_{s2}+1), & \text{if } (Y_{s1}, Y_{s2}) \in A_{14} \cup A_{15} \\ (Y_{s1}-1, Y_{s2}-1), & \text{if } (Y_{s1}, Y_{s2}) \in A_{16} \cup A_{17} \cup A_{18} \cup A_{19} \end{cases} \qquad (15)$$

The value of the marked coefficient pair may belong to a new set. The receiver could extract the information and fully restore the video according to the reverse process of Fig. 4.

2.2.2 Embedding capacity and distortion

The embedding capacities of the presented 2D HM and the customary 2D HM, denoted as EC_{pre} and EC_{cus}, can be computed by

$$EC_{pre} = \frac{9}{4} \sum_{(Y_{s1}, Y_{s2}) \in A_1} y(Y_{s1}, Y_{s2}) + \frac{3}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_2 \cup A_3 \cup A_4 \cup A_9 \cup A_{10}} y(Y_{s1}, Y_{s2}) + \sum_{(Y_{s1}, Y_{s2}) \in A_5 \cup A_6 \cup A_7 \cup A_8} y(Y_{s1}, Y_{s2}),$$

and
$$\tag{16}$$

$$EC_{cus} = 2 \sum_{(Y_{s1}, Y_{s2}) \in A_1 \cup A_3 \cup A_6 \cup A_{16}} y(Y_{s1}, Y_{s2}) + \sum_{(Y_{s1}, Y_{s2}) \in A_2 \cup A_4 \cup A_5 \cup A_7 \cup A_8 \cup A_9 \cup A_{10} \cup A_{12} \cup A_{14} \cup A_{17} \cup A_{18}} y(Y_{s1}, Y_{s2}). \tag{17}$$

According to (16) and (17), we can infer the difference of embedding capacity between the presented 2D HM and the conventional 2D HM by

$$
\begin{aligned}
EC_{pre} &- EC_{cus} \\
&= \frac{1}{4} \sum_{(Y_{s1}, Y_{s2}) \in A_1} y(Y_{s1}, Y_{s2}) + \frac{1}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_2 \cup A_4 \cup A_9 \cup A_{10}} y(Y_{s1}, Y_{s2}) - \frac{1}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_3} y(Y_{s1}, Y_{s2}) \\
&\quad - 2 \sum_{(Y_{s1}, Y_{s2}) \in A_{16}} y(Y_{s1}, Y_{s2}) - \sum_{(Y_{s1}, Y_{s2}) \in A_6 \cup A_{12} \cup A_{14} \cup A_{17} \cup A_{18}} y(Y_{s1}, Y_{s2}) \\
&\approx \frac{1}{2} \sum_{Y_{s1} = -\infty}^{+\infty} y(Y_{s1}, 0) + \frac{1}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_9 \cup A_{10}} y(Y_{s1}, Y_{s2}) - \frac{1}{4} \sum_{(Y_{s1}, Y_{s2}) \in A_1} y(Y_{s1}, Y_{s2}) - 2 \sum_{Y_{s1} = -\infty}^{+\infty} y(Y_{s1}, -1).
\end{aligned}
\tag{18}
$$

For QDCT coefficients, the distortion in terms of l^2-error of the presented 2D HM and the customary 2D HM, denoted as ED_{pre} and ED_{cus}, can be formulated as

$$
\begin{aligned}
ED_{pre} &= \frac{3}{4} \sum_{(Y_{s1}, Y_{s2}) \in A_1} y(Y_{s1}, Y_{s2}) + \sum_{(Y_{s1}, Y_{s2}) \in A_2 \cup A_3 \cup A_4 \cup A_7 \cup A_8 \cup A_9 \cup A_{10}} y(Y_{s1}, Y_{s2}) + \frac{3}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_5 \cup A_6} y(Y_{s1}, Y_{s2}), \\
&\quad + 2 \sum_{(Y_{s1}, Y_{s2}) \in A_{11} \cup A_{12} \cup A_{13} \cup A_{14} \cup A_{15} \cup A_{16} \cup A_{17} \cup A_{18} \cup A_{19}} y(Y_{s1}, Y_{s2})
\end{aligned}
\tag{19}
$$

and

$$
\begin{aligned}
ED_{cus} &= \sum_{(Y_{s1}, Y_{s2}) \in A_1 \cup A_3 \cup A_6 \cup A_{16}} y(Y_{s1}, Y_{s2}) + \frac{3}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_2 \cup A_4 \cup A_5 \cup A_7 \cup A_8 \cup A_9 \cup A_{10} \cup A_{12} \cup A_{14} \cup A_{17} \cup A_{18}} y(Y_{s1}, Y_{s2}) \\
&\quad + 2 \sum_{(Y_{s1}, Y_{s2}) \in A_{11} \cup A_{13} \cup A_{15} \cup A_{19}} y(Y_{s1}, Y_{s2}).
\end{aligned}
\tag{20}
$$

According to (19) and (20), we can infer the difference of embedding distortion between the presented 2D HM and the conventional 2D HM by

$$
\begin{aligned}
ED_{cus} &- ED_{pre} \\
&= \frac{1}{4} \sum_{(Y_{s1}, Y_{s2}) \in A_1} y(Y_{s1}, Y_{s2}) + \frac{1}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_2 \cup A_4 \cup A_7 \cup A_8 \cup A_9 \cup A_{10}} y(Y_{s1}, Y_{s2}) \\
&\quad - \frac{1}{2} \sum_{(Y_{s1}, Y_{s2}) \in A_6 \cup A_{12} \cup A_{14} \cup A_{17} \cup A_{18}} y(Y_{s1}, Y_{s2}) - \sum_{(Y_{s1}, Y_{s2}) \in A_{16}} y(Y_{s1}, Y_{s2}) \\
&\approx \frac{1}{2} \sum_{Y_{s1} = -\infty}^{+\infty} y(Y_{s1}, 0) + \sum_{(Y_{s1}, Y_{s2}) \in A_4} y(Y_{s1}, Y_{s2}) - \frac{1}{4} \sum_{(Y_{s1}, Y_{s2}) \in A_1} y(Y_{s1}, Y_{s2}) - \sum_{Y_{s1} = -\infty}^{+\infty} y(Y_{s1}, -1).
\end{aligned}
\tag{21}
$$

The probabilities of QDCT coefficients with values 1 and −1 are denoted by p_1 and p_{-1}. For the test video Rena, $p_0 \approx 99.39$ %, $p_1 \approx 0.25$ %, and $p_{-1} \approx 0.35$ %. Then $EC_{pre} - EC_{cus} \approx 0.243\#A$ and $ED_{cus} - ED_{pre} \approx 0.246\#A$, where $\#A$ is the cardinal number of the set A. For the test video Objects, $p_0 \approx 82.87$ %, $p_1 \approx 7.21$ %, and $p_{-1} \approx 6.57$ %. So $EC_{pre} - EC_{cus}$ is about $0.116\#A$ and $ED_{cus} - ED_{pre}$ is about $0.191\#A$. The results of the two videos verify that larger capacity and less distortion can be achieved by the presented 2D HM compared with the conventional 2D HM. Furthermore, if p_0 is larger, the differences of capacity and distortion between the two methods will be larger.

In Table 2, a pair of coefficients (Y_2, Y_5) in the low frequency region is used for embedding data. The capacity of a video sequence is the average number of bits embedded into one P_0 frame of all the P_0 frames in that sequence. The peak signal-to-noise ratio (PSNR) value obtained through comparing the marked YUV video with the original YUV video is the average of all the frames. It is obvious that the proposed 2D HM is superior to the traditional 2D HM by improving capacity with 44 bits and PSNR with 0.03 dB at least.

2.2.3 Detailed example

Three examples are given to depict the presented hiding scheme clearly.

1. For the QDCT coefficient pair $(Y_{s1}, Y_{s2}) = (0, 0) \in A_1$, the value $(0, 0)$ of QDCT coefficient pair (Y_{s1}, Y_{s2}) is expanded to $(0, 0)$, $(1, 0)$, and $(0, 1)$ when two to-be-embedded data bits (m_i, m_{i+1}) are $(0, 1)$, $(1, 0)$, and $(1, 1)$, severally. Otherwise, the value $(0, 0)$ of QDCT coefficient pair (Y_{s1}, Y_{s2}) will be expanded to $(-1, 0)$ if three to-be-embedded data bits (m_i, m_{i+1}, m_{i+2}) are $(0, 0, 0)$, and $(0, -1)$ if three data bits (m_i, m_{i+1}, m_{i+2}) are $(0, 0, 1)$. Three data bits (m_i, m_{i+1}, m_{i+2}) or two data bits (m_i, m_{i+1}) could be hidden at the same time, and the cost is 1 at most. It can be inferred that the proposed scheme is better than the conventional 2D HM, whose cost will be 0, 1, 1, and 2 if two data bits (m_i, m_{i+1}) are $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$, severally. Furthermore, the histogram of QDCT coefficient is so steep that most chosen pairs could be $(0, 0)$. As a result, the distortion will be significantly reduced by this new mapping.
2. For the QDCT coefficient pair $(Y_{s1}, Y_{s2}) = (1, 0) \in A_2$, it will be expanded to $(1, -1)$ if a data bit m_i is 1. Otherwise, it will be expanded to $(2, 0)$ if two data bits (m_i, m_{i+1}) are $(0, 0)$, and $(1, 1)$ if two data bits (m_i, m_{i+1}) are $(0, 1)$. Two data bits (m_i, m_{i+1}) or one data bit m_i can be embedded at the cost of modifying one coefficient by adding or subtracting 1. By contrast, in the conventional 2D HM, the cost will be 1 if a data bit m_i is 0, and 2 if m_i is 1. Hence, when the value of QDCT coefficient pair (Y_{s1}, Y_{s2}) belongs to the set A_2, compared with the conventional 2D HM, better payload-distortion performance can be obtained by using the proposed 2D HM.
3. For the QDCT coefficient pair $(Y_{s1}, Y_{s2}) = (0, 1) \in A_5$, it will be expanded to $(0, 2)$ if a data bit m_i is 0, and $(1, 2)$ if m_i is 1. It can be seen that the proposed scheme can be used to achieve the same payload-distortion performance as the customary 2D HM when the value of the selected coefficient pair belongs to the set A_5.

To sum up, when the value of QDCT coefficient pair (Y_{s1}, Y_{s2}) belongs to the set A_1, A_2, A_3, A_4, A_9 or A_{10}, our method is better than the customary 2D HM. When the value of the selected pair (Y_{s1}, Y_{s2}) belongs to the other sets, the same embedding performance can be achieved by the two methods. On the whole, video RDE algorithms could be realized through the presented 2D HM, by which good capacity-distortion performance can be acquired. Furthermore, the sharper the histogram is, the better hiding performance can be gained. In addition, when the

Table 2 Capacities and PSNR of the customary 2D HM and the presented 2D HM for hiding data into one P_0 frame

Sequences		Akko&Kayo	Ballroom	Crowd	Exit	Flamenco	Objects	Race	Rena	Vassar
Customary 2D HM	Capacity (bits)	576	1585	4080	507	565	642	2883	554	1106
	PSNR (dB)	39.64	36.62	35.24	38.21	39.39	37.12	37.32	41.46	36.18
Presented 2D HM	Capacity (bits)	643	1728	4257	555	620	686	3125	621	1207
	PSNR (dB)	39.67	36.78	35.47	38.30	39.42	37.27	37.40	41.51	36.40

proposed method is applied in some media, especially a gray-scale image with 8 storage bits [24], the overflow or the underflow problem has to be treated. However, this problem needs no considering when the data is hidden into QDCT coefficients of H.264 video [3].

3 Actualization of the presented RDE algorithm

The presented RDE algorithm based on 2D HM for stereo H.264 video is depicted in Fig. 5. For the sender, to enhance the security and the robustness of the hidden information, the information denoted by M is encrypted into cipher text, which is encoded by the secret sharing technique (secret is divided into n_s shares, where any t_s shares can be used to recover the secret, n_s and t_s are integers) and BCH (n_B, k_B, t_B) code (n_B is the codeword length, k_B is the code dimension, t_B errors can be corrected [38]) before data hiding. The original stereo H.264 video is entropy decoded to get QDCT MBs, where some MBs are chosen for hiding data, and the unselected MBs will be entropy encoded directly. Several data bits are embedded into each two QDCT coefficients, which could be randomly selected from every embeddable 4×4 block in a MB. Correspondingly, for the receiver, the hidden information could be extracted from the marked QDCT coefficients after entropy decoding.

3.1 Embeddable blocks

Inter-view prediction, inter-frame prediction, and intra-frame prediction are used to improve the compressing efficiency of stereo H.264. As a result, the inter-view, the inter-frame, and the
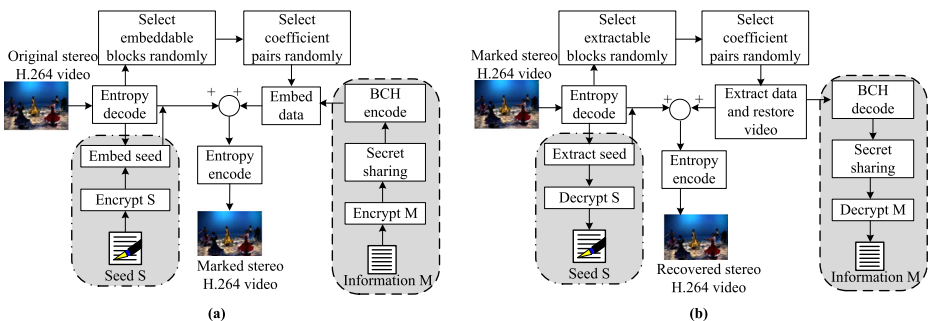

Fig. 5 The flowchart of the presented RDE algorithm **a** Embedding **b** Extraction and recovery

intra-frame distortion drift will be caused by hiding data into stereo H.264 video. So it is necessary to find embeddable blocks for hiding data with limited distortion drift.

Horizontal inter-frame prediction and vertical inter-view prediction of a stereo H.264 video with hierarchical B coding and two views are illustrated in Fig. 6. There are 16 frames in each group of picture, where each view has eight frames. Only intra-frame prediction is used for I_0 frames, the distortion of other frames will not affect I_0 frames. However, hiding data into one I_0 frame will sway all the P_0, B_1, B_2, B_3, and b_4 frames in the two adjacent groups of pictures predicted by this I_0 frame. By contrast, embedding data into P_0 frames or b_4 frames will not lead to inter-view distortion drift, because frames in the right view do not predict frames in the left view. Therefore, compared with embedding data into I_0 frames, better video quality can be achieved through embedding data into P_0 or b_4 frames. Inter-view distortion drift and inter-frame distortion drift could be avoided by embedding data into b_4 frames, but b_4 frames located at the lowest level are easy to be discarded during the process of transmission in the network. Inter-frame distortion drift will be caused by embedding data into P_0 frames. However, P_0 frames are key pictures at the highest level, resulting that they cannot be easily lost during the process of the network transmission. Consequently, stronger robustness can be obtained by hiding data into P_0 frames compared with b_4 frames. In addition, in one group of picture, there are six B_3 frames so that enough random space could be used for hiding data into B_3 frames. Only b_4 frames may be affected by embedding information into B_3 frames. P_0, B_3 and b_4 frames have their virtues and faults, so users could select embeddable frames on demand.

At the encoder, through using inter-view prediction, inter-frame prediction, or intra-frame prediction, the prediction block of a block is achieved to compute its residual block denoted as F^{Ro}, which will be changed into a QDCT block denoted as Y by the 4×4 discrete cosine transform and quantization formulated as

$$Y = round\left[\left(C_f \cdot F^{Ro} \cdot C_f^T\right) \otimes \left(E_f / Q\right)\right], \tag{22}$$

where $C_f = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}$, $E_f = \begin{bmatrix} a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \\ a^2 & ab/2 & a^2 & ab/2 \\ ab/2 & b^2/4 & ab/2 & b^2/4 \end{bmatrix}$, $a = 1/2$, $b = \sqrt{2/5}$, Q is the quantization step size, $\otimes$ is a math operator, in which each element in the former matrix is multiplied by the value at the corresponding position in the latter matrix.

When a data bit is hidden into one QDCT block Y by modifying some QDCT coefficients, the QDCT block after embedding data is denoted by Y^{Emb}. The modification caused by hiding information is denoted by ΔY, which is shown as

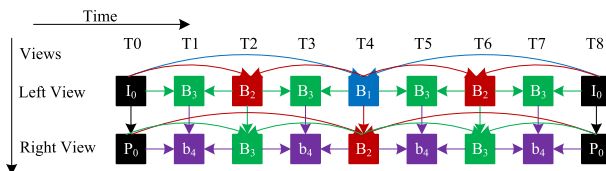$$\Delta Y = Y^{Emb} - Y. \tag{23}$$

Fig. 6 Prediction structure of stereo H.264 video with two views

At the decoder, the residual block acquired by inverse quantization and 4×4 inverse discrete cosine transform is denoted by F^R, which can be computed by

$$F^R = round\left[C_d^T \cdot (Y \otimes E_d) \cdot C_d\right], \tag{24}$$

where $C_d = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1/2 & -1/2 & -1 \\ 1 & -1 & -1 & 1 \\ 1/2 & -1 & 1 & -1/2 \end{bmatrix}$, $E_d = \begin{bmatrix} a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \\ a^2 & ab & a^2 & ab \\ ab & b^2 & ab & b^2 \end{bmatrix}$.

When a data bit is embedded through modifying some QDCT coefficients of one block, the residual block after embedding data is denoted by F^{REmb}. Denote the variation of the residual block between before and after hiding information as ΔF^R, which can be calculated as

$$\Delta F^R = F^{REmb} - F^R = round\left[C_d^T \cdot (\Delta Y \cdot Q \otimes E_d) \cdot C_d\right]. \tag{25}$$

Take the QDCT coefficient Y_2 as an example to explain the distortion caused by embedding data. Assume an integer denoted by t is added to the value of Y_2, the modification of the QDCT block for hiding data is $\Delta Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ t & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$. Then the alteration of the corresponding block in YUV video is

$$\Delta F^R = \frac{1}{2} abtQ \begin{bmatrix} 2 & 2 & 2 & 2 \\ 1 & 1 & 1 & 1 \\ -1 & -1 & -1 & -1 \\ -2 & -2 & -2 & -2 \end{bmatrix}. \tag{26}$$

It can be seen that changing only one QDCT coefficient in a 4×4 block causes the variation of the whole block in the corresponding YUV video. In the same way, for an 8×8 block, modifying one QDCT coefficient will alter the whole 8×8 block, whose affected region is bigger than that of the 4×4 block. In addition, only two kinds of transformations, 4×4 transformation and 8×8 transformation, are used in stereo H.264 standard. Hence, the 4×4 transform block is selected to embed information in this paper. It can be inferred that the
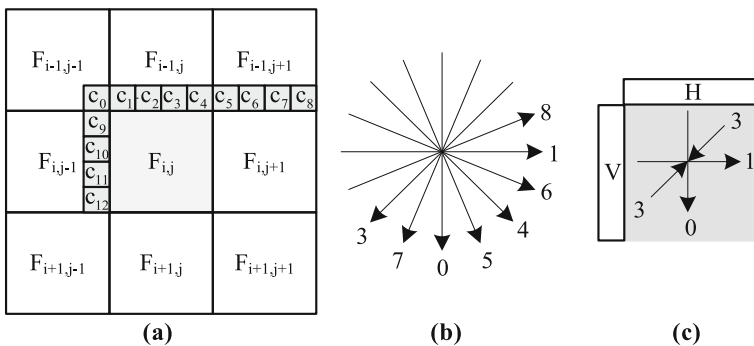

Fig. 7 Intra-frame prediction mode **a** Block position. **b** The predictive direction of 4×4 and 8×8 luma block. **c** The predictive direction of 16×16 luma block. (In mode 2 not shown in the figure, all elements are predicted with the average of upper pixels denoted by H and left pixels denoted by V, i.e., the average Mean (H+V))

boundary pixels denoted by c_0 … c_{12} (shown in Fig. 7) may be changed by hiding data into some QDCT coefficients of the blocks $F_{i,j-1}$ (the position of a block is expressed by (i, j)), $F_{i-1,j-1}$, $F_{i-1,j}$, and $F_{i-1,j+1}$. In addition, if intra-frame prediction is used by the current block $F_{i,j}$, its prediction block denoted by $F^P_{i,j}$ will be reckoned by the pixels c_0 … c_{12}. Therefore, the hiding induced deviation of the blocks $F_{i,j-1}$, $F_{i-1,j-1}$, $F_{i-1,j}$, and $F_{i-1,j+1}$ will propagate to the block $F_{i,j}$. Otherwise, if inter-view prediction or inter-frame prediction is employed to compute the prediction block $F^P_{i,j}$, the block $F_{i,j}$ will not be affected by the change of the pixels c_0 … c_{12}, because the prediction block $F^P_{i,j}$ is calculated by referring other frames.

A 16×16 MB using inter-frame prediction or inter-view prediction is denoted by inter-MB. As depicted in Fig. 8, the 4×4 blocks numbered by 0…8, which are not located at the bottom or the rightmost column of the inter-MB, are selected as embeddable blocks so that intra-frame distortion drift can be avoided. If the current block $F_{i,j}$ shown in Fig. 7a is an embeddable block, the blocks $F_{i,j+1}$, $F_{i+1,j+1}$, and $F_{i+1,j}$, which adjoin $F_{i,j}$, will be in the current inter-MB. Furthermore, the block $F_{i+1,j-1}$ maybe stays in the current inter-MB or one of the blocks numbered by 9, 10, and 11 in the encoded MB at the encoder (or the decoded MB at the decoder). Thus the adjacent blocks $F_{i,j+1}$, $F_{i+1,j+1}$, $F_{i+1,j}$ or $F_{i+1,j-1}$ will not be affected by the current block $F_{i,j}$, resulting that intra-frame distortion drift will not be caused by embedding data into the block $F_{i,j}$. It can be seen that one inter-MB has nine 4×4 embeddable blocks, which could be applied for embedding data without causing intra-frame distortion drift.

3.2 Data embedding

The embedding program is shown in Fig. 5a. There are five main steps to complete the data-hiding process: processing the to-be-hidden information, embedding the seed denoted by S, choosing embeddable blocks, selecting coefficient pairs, and data embedding.

1. *Processing the to-be-hidden information* The original information is encrypted into cypher text with a RSA public key. Then the secret sharing method is used to divide the encrypted information into several groups, where each group of information is encoded by BCH code before data hiding.

2. *Embedding the seed S* To improve the security, random function is used to select embeddable blocks and coefficients. We encrypt the seed S with a RSA public key and hide the seed S into some QDCT MBs of I_0 frames with a public RDE algorithm.

3. *Choosing embeddable blocks* A random threshold denoted by U is generated so that we can randomly select 4×4 embeddable blocks (not at the bottom row or the rightmost column, as shown in Fig. 8) from QDCT inter-MBs with 4×4 luma transformation
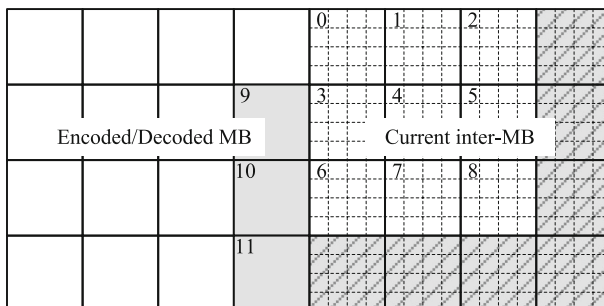

Fig. 8 Embeddable 4×4 QDCT blocks without intra-frame distortion drift in an inter-MB

according to $|Y_0| \geq U$. High threshold U will constrain the quantity of embeddable blocks, so the distortion region will be limited. Consequently, embeddable blocks could be randomly chosen for hiding data.

4. *Selecting coefficient pairs* Two coefficients are chosen from each embeddable blocks randomly as shown in Algorithm1. In this way, the application of arbitrary embedding positions including blocks and coefficients can be used to reduce the modification of statistical histogram and improve the undetectability of RDE algorithm.

5. *Data embedding* Some data bits are hidden into each coefficient pairs by following Fig. 4.

Algorithm 1 depicts the method of choosing two QDCT coefficients from an embeddable block. Random function is used to select two QDCT coefficients (Y_{s1}, Y_{s2}) from 15 AC coefficients (or 5 AC coefficients in the low frequency region) in a 4×4 QDCT luminance block with zigzag scan shown in Fig. 1 (Several pairs of QDCT coefficients could be chosen for hiding data, only one pair is considered in this paper). It can be known that there are 15 ways to choose Y_{s1} from 15 QDCT AC coefficients and 14 ways to select Y_{s2} from the rest 14 QDCT AC coefficients. Thus the optional number of selecting (Y_{s1}, Y_{s2}) is 15×14=210. If a marked block is found by the third party, the probability for computing the embedded data bit at once will be 1/210≈0.00476. When the pair (Y_{s1}, Y_{s2}) is selected from the low frequency area, there are five ways to choose Y_{s1} from 5 AC coefficients in the low frequency region and four ways to select Y_{s2} from the rest 4 AC coefficients. Accordingly, the optional number of selecting (Y_{s1}, Y_{s2}) is 5×4=20. Suppose a marked block is found by others, the probability for directly reckoning the hidden data bit is 1/20≈0.05. Therefore, compared with the small random space, the large random space can prevent the third party from guessing the hidden data bit more efficiently.

Algorithm 1 Random selection of two QDCT coefficients

```
Input: Seed S
Output: Two QDCT coefficients (Ys1, Ys2)
Initialize the random seed with srand(S);
Define the number of AC coefficients N=15 or 5;
For i=1 to N do
 Assign a value to the ith element of the integer array X[] with X[i]=i;
End for
For i=1 to 2 do
 Generate the ith random number with si=rand()%(N+1−i)+i;
 If si!=i
   Swap the positions of X[si] and X[i];
 End if
End for
 Two QDCT coefficients (Ys1, Ys2) are chosen as embedding
coefficients.
```

3.3 Data extraction and video recovery

The procedure of information extraction and video recovery is shown in Fig. 5b. After the entropy decoding of the stereo H.264 video, we extract the seed S from I_0 frame with the public RDE algorithm and decrypt the seed S with the RSA private key. Since the random seed is in one-to-one correspondence with the random sequence, the embeddable blocks and coefficients used by the sender will be in one-to-one correspondence with the extractable blocks and coefficients used by the receiver if the receiver and the sender use the same random seed S.
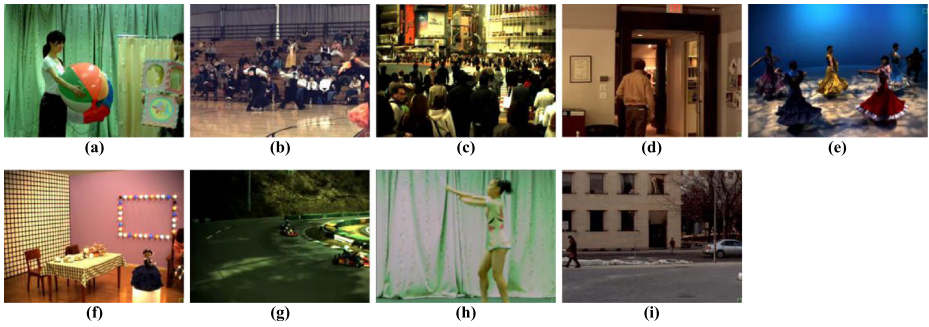
Fig. 9 Test videos (the size of each frame is 640×480) **a** Akko&Kayo. **b** Ballroom. **c** Crowd. **d** Exit. **e** Flamenco. **f** Objects. **g** Race. **h** Rena. **i** Vassar

According to the reverse process of Fig. 4, the hidden information could be extracted, and the stereo H.264 video is recovered completely. At last, we use BCH code and secret sharing techniques to recover the information and get the hidden data.

4 Experimental results and discussions

JM18.4 [39] is utilized to do experiments with nine test videos [40] shown in Fig. 9. Two YUV files are encoded into a stereo H.264 video with 233 frames including 30 P_0 frames. The parameter intra-period is set as 8. Comparing the marked frames with the original frames of the corresponding YUV files, we compute the structural similarity (SSIM), which is the average of all the frames. The embedding efficiency denoted by e is defined as

$$e = L_{emb} \bigg/ \sum_{i=1}^{i=N_{cha}} Lcha_i, \tag{27}$$

where L_{emb} is the number of embedded bits, and $Lcha_i$ is the changed size of a modified coefficient, N_{cha} is the quantity of changed coefficients.

The average embedding capacity with five different thresholds (the threshold U=0, 1, 2, 3, 4) is shown in Fig. 10a. It can be seen that the capacity always depends on the video content, i.e., different embedding capacities can be achieved by hiding data into different videos.
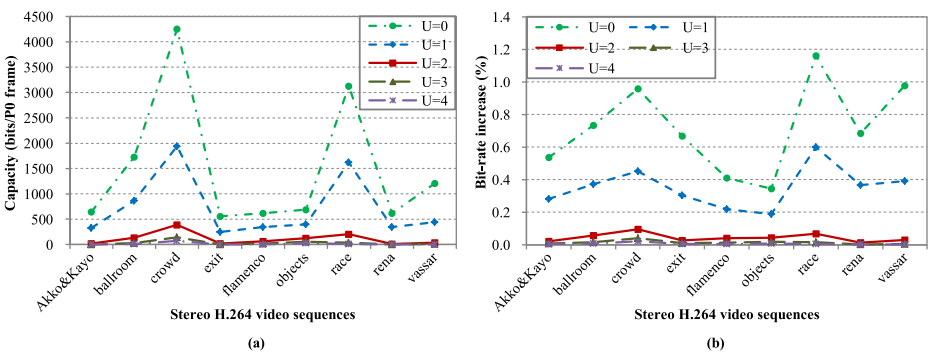

Fig. 10 The capacities and the bit-rate increase of the presented scheme with different thresholds (U=0, 1, 2, 3, 4)

Table 3 Hiding performance of three methods for hiding 2600 bits into one frame of Crowd and Race

Sequences		Ma	P_0 _drift	P_0 _interMB
Crowd	PSNR (dB)	35.32	35.45	35.54
	SSIM	0.9669	0.9671	0.9680
	e	1.17	1.79	1.85
Race	PSNR (dB)	37.22	37.39	37.49
	SSIM	0.9597	0.9606	0.9609
	e	1.16	2.31	2.30

Additionally, the higher the threshold U is, the lower the embedding capacity will be obtained, and the less bit-rate increase will be caused, as shown in Fig. 10b.
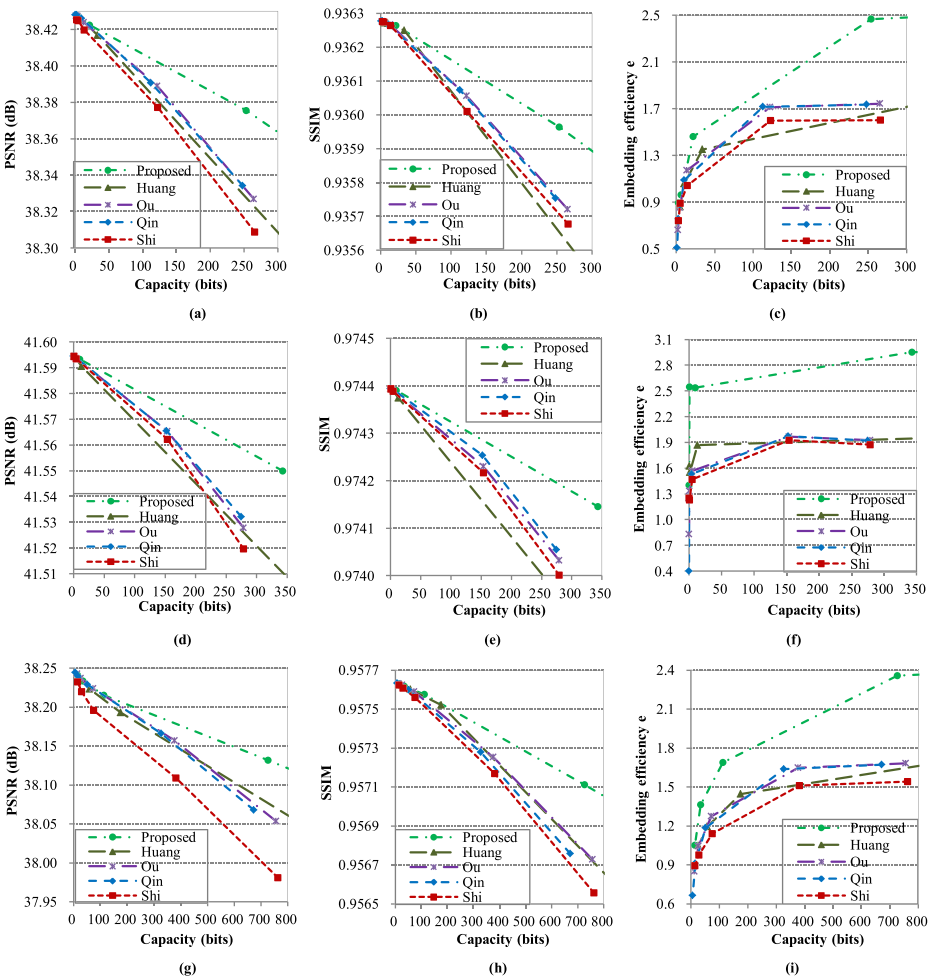

Fig. 11 Performance comparison with other RDE algorithms on stereo H.264 videos **a–c** Exit. **d–f** Rena. **g–i** Average of nine videos

Table 3 shows the embedding performance of three methods. Apart from Ma et al.'s scheme [36] denoted as Ma, where data is hidden into I_0 frame, the proposed 2D-HM-based RDE method is used to embed data into P_0 frame in the other schemes. The scheme embedding information with intra-frame distortion drift is denoted by P_0 _drift. Compared with Ma et al.'s scheme [36] without intra-frame distortion drift, P_0_drift increases PSNR, SSIM, and embedding efficiency e with 0.13 dB, 0.0002, and 0.63 at least for hiding 2600 bits of data into one frame, respectively. The scheme embedding information without intra-frame distortion drift is denoted by P_0_interMB. Compared with P_0_drift, PSNR and SSIM can be enhanced with 0.09 dB and 0.0003 at least by preventing intra-frame distortion drift, severally. It experimentally demonstrates that the way to avoid distortion drift is efficient.

In order to compare the presented 2D HM based RDE algorithm with other RDE methods in the same environment, embeddable blocks (shown in Fig. 8), which could be used for hiding information without causing inter-view distortion drift and intra-frame distortion drift, are selected from inter-MBs of P_0 frames. Huang and Chang's method [28] is denoted by Huang, where data is embedded into three QDCT coefficients Y_2, Y_5, and Y_3. The QDCT coefficient pair (Y_2, Y_5) is utilized by our algorithm, Shi et al.'s method [20] denoted by Shi,

Table 4 Performance comparison for hiding information into the first P_0 frame

	Amount of hidden data (bits)	Sequences	Huang	Ou	Qin	Shi	Ours
Akko&Kayo	860	PSNR (dB)	40.411	40.465	40.466	40.423	40.511
		SSIM	0.97087	0.97114	0.97110	0.97102	0.97145
		e	1.99	1.88	1.88	1.83	2.84
Ballroom	2500	PSNR (dB)	37.276	37.350	37.344	37.218	37.424
		SSIM	0.94599	0.94706	0.94683	0.94645	0.94739
		e	1.93	1.74	1.73	1.57	2.64
Crowd	2400	PSNR (dB)	37.191	37.168	37.171	36.819	37.283
		SSIM	0.97653	0.97670	0.97656	0.97618	0.97685
		e	1.74	1.47	1.44	1.28	2.10
Exit	410	PSNR (dB)	39.071	39.093	39.094	39.062	39.107
		SSIM	0.94302	0.94316	0.94309	0.94305	0.94318
		e	1.72	1.69	1.70	1.54	2.52
Flamenco	170	PSNR (dB)	41.288	41.301	41.306	41.282	41.319
		SSIM	0.97870	0.97880	0.97881	0.97874	0.97888
		e	1.75	1.60	1.62	1.49	2.69
Objects	2200	PSNR (dB)	37.322	37.262	37.251	36.747	37.377
		SSIM	0.97352	0.97369	0.97349	0.97330	0.97372
		e	1.48	1.25	1.22	1.09	1.65
Race	740	PSNR (dB)	39.882	39.910	39.901	39.865	39.960
		SSIM	0.96596	0.96670	0.96669	0.96666	0.96671
		e	2.01	1.79	1.77	1.63	2.82
Rena	980	PSNR (dB)	42.494	42.641	42.650	42.617	42.694
		SSIM	0.97516	0.97562	0.97572	0.97551	0.97599
		e	2.03	2.02	2.02	1.97	2.89
Vassar	690	PSNR (dB)	37.208	37.215	37.219	37.165	37.237
		SSIM	0.92164	0.92182	0.92170	0.92160	0.92188
		e	1.77	1.59	1.51	1.37	2.26

and Ou et al.'s method [29] denoted by Ou. In addition, Y_2 is used by Qin et al.'s method [32] denoted by Qin. In Fig. 11, the points of each line from left to right represent the embedding cases whose threshold U is 4, 3, 2, 1, and 0, respectively. It can be seen that the best SSIM, PSNR and embedding efficiency can be gained by our method when the same amount of information is hidden. The best SSIM and PSNR mean that the best video quality could be obtained by using the proposed algorithm. In addition, the embedding efficiency e of our scheme is much higher compared with that of other methods. It indicates that when the same amounts of carrier bits are modified, more data bits can be hidden through the proposed method compared with other algorithms.

To clarify this further, the same amount of information is hidden by using each RDE method. As we can see in Fig. 10, different embedding capacities will be obtained by hiding data into different videos. So different amount of data is hidden into different videos. In Table 4, the threshold U is 0, and some bits of data are hidden into the first P_0 frame of each video sequence. Compared with the state-of-the-art methods, our RDE algorithm is superior by enhancing PSNR, SSIM, and embedding efficiency e with 0.063 dB, 0.0014 and 1.2 at most (0.013 dB, 0.00001 and 0.17 at least), respectively.

Correspondingly, to clearly display visual results, Fig. 12 shows parts of the marked frames of Flamenco, Objects, and Race. (a), (b), and (c) are parts of the marked frames achieved by using the method Huang [28] to hide data. (d), (e), and (f) are parts of the marked frames obtained by utilizing our method to embed information. Their original frames are shown in Fig. 9. It can be seen that obvious distortion is produced in the frames (a), (b), and (c). Some distortion could be found near the persons' heads and clothes in the frame (a). On the top left of frame (b), the distortion is on the wall and the ceiling. And the distortion could be observed on the top middle of frame (c). In contrast, there is little distortion in the frames (d), (e), and (f). It can be concluded from the experimental results that better visual quality can be acquired by embedding data with the proposed algorithm.

Denote the frame number as n, and the information length as w. The computational efficiency of the presented scheme is related to the data length w and the frame number n. Therefore, the computational complexity of the proposed algorithm can be denoted by $O(n \times w)$.

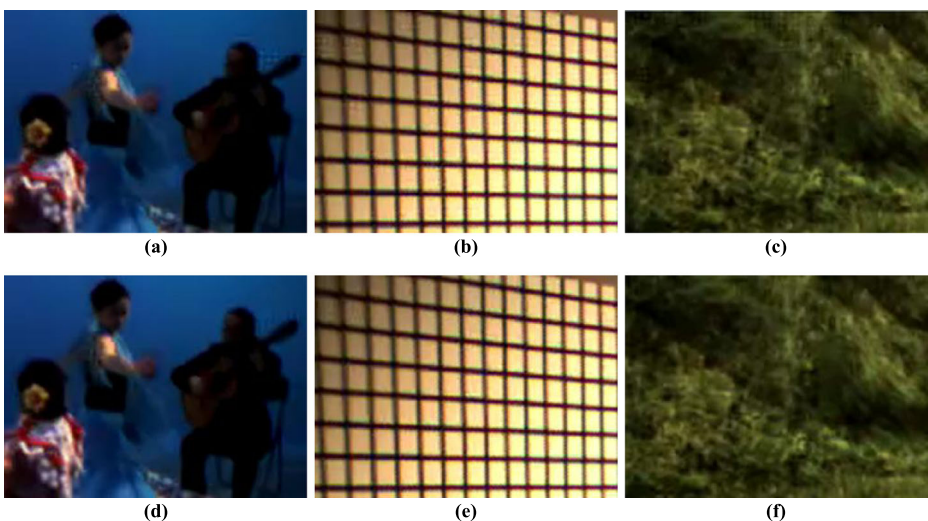|		
(a)	(b)	(c)
(d)	(e)	(f)

Fig. 12 Parts of the marked frames of Flamenco, Objects, and Race with Huang and the presented algorithm

All the experiments were run on a machine with 24 processors (2.5 GHz) and 32 GB of RAM. It took 163 ms on average to run the data hiding procedure for each frame of the test videos. The video decoding procedure without data hiding for each frame of the test videos cost 161 ms on average. The duration caused by hiding data for each frame is 2 ms on average. Therefore, the presented method with low complexity can be used for real-time application.

Furthermore, in our paper [38], we have tested that BCH(7,4,1) is the most powerful code in terms of error correction capability if 1 random error bit is corrected. Otherwise, BCH (63,7,15) is the most powerful code in terms of error correction capability, where the hidden data bits can be recovered 100 % when the frame loss rate is no more than 15 %. So high robustness could be achieved by using the proposed scheme shown in Fig. 5.

5 Conclusions

An efficient RDE algorithm based on 2D HM is presented in this paper. This algorithm is a novel reversible method uniting two QDCT coefficients together for embedding several data bits into one block, where only one coefficient needs modifying by adding or subtracting 1 at most in most instances. Firstly, embeddable blocks limiting distortion drift are chosen at random. Then two QDCT coefficients are randomly selected from an embeddable block for embedding data with 2D HM. Better capacity-distortion performance can be obtained by this presented 2D HM method compared with the state-of-the-art RDE algorithms. The proposed algorithm will be generalized to hide over three data bits with at most one modification in our future work.

References

1. Alattar AM (2004) Reversible watermark using the difference expansion of a generalized integer transform. IEEE Trans Image Process 13(8):1147–1156
2. Al-Qershi OM, Khoo BE (2013) Two-dimensional difference expansion (2D-DE) scheme with a characteristics-based threshold. Signal Process 93(1):154–162
3. An LL, Gao XB, Li XL, Tao DC, Deng C, Li J (2012) Robust reversible watermarking via clustering and enhanced pixel-wise masking. IEEE Trans Image Process 21(8):3598–3611
4. Celik MU, Sharma G, Tekalp AM, Saber E (2005) Lossless generalized-LSB data embedding. IEEE Trans Image Process 14(2):253–266
5. Celik MU, Sharma G, Tekalp AM (2006) Lossless watermarking for image authentication: a new framework and an implementation. IEEE Trans Image Process 15(4):1042–1049
6. Chung KL, Huang YH, Chang PC, Liao HYM (2010) Reversible data hiding-based approach for intra-frame error concealment in H.264/AVC. IEEE Trans Circuits Syst Video Technol 20(11):1643–1647
7. Coltuc D (2012) Low distortion transform for reversible watermarking. IEEE Trans Image Process 21(1):412–417
8. Coltuc D, Chassery JM (2007) Very fast watermarking by reversible contrast mapping. IEEE Signal Process Lett 14(4):255–258
9. De Vleeschouwer C, Delaigle JF, Macq B (2003) Circular interpretation of bijective transformations in lossless watermarking for media asset management. IEEE Trans Multimedia 5(1):97–105
10. Hsu FH, Wu MH, Wang SJ (2013) Reversible data hiding using side-match predictions on steganographic images. Multimedia Tools Appl 67(3):571–591

11. Huang HC, Chang FC (2013) Hierarchy-based reversible data hiding. Expert Syst Appl 40(1):34–43
12. Jawad K, Khan A (2013) Genetic algorithm and difference expansion based reversible watermarking for relational databases. J Syst Softw 86(11):2742–2753
13. Jeni M, Srinivasan S (2013) Reversible data hiding in videos using low distortion transform. IEEE, New York
14. Kamstra L, Heijmans H (2005) Reversible data embedding into images using wavelet techniques and sorting. IEEE Trans Image Process 14(12):2082–2090
15. Li XL, Li B, Yang B, Zeng TY (2013) General framework to histogram-shifting-based reversible data hiding. IEEE Trans Image Process 22(6):2181–2191
16. Lin SFD, Su YL, Huang JY (2007) Error resilience using a reversible data embedding technique in H.264/AVC. 7th WSEAS International Conference on Mulitmedia Systems & Signal Processing (MUSP '07), Hangzhou, China, 112–117
17. Lin C-C, Tai W-L, Chang C-C (2008) Multilevel reversible data hiding based on histogram modification of difference images. Pattern Recogn 41(12):3582–3591
18. Liu Y, Li Z, Ma X, Liu J (2014) A robust without intra-frame distortion drift data hiding algorithm based on H.264/AVC. Multimedia Tools Appl 72(1):613–636
19. Liu YC, Wu HC, Yu SS (2011) Adaptive DE-based reversible steganographic technique using bilinear interpolation and simplified location map. Multimedia Tools Appl 52(2–3):263–276
20. Ma XJ, Li ZT, Tu H, Zhang BC (2010) A data hiding algorithm for H.264/AVC video streams without intra-frame distortion drift. IEEE Trans Circuits Syst Video Technol 20(10):1320–1330
21. Ni ZC, Shi YQ, Ansari N, Su W (2006) Reversible data hiding. IEEE Trans Circuits Syst Video Technol 16(3):354–362
22. Ou B, Li XL, Zhao Y, Ni RR, Shi YQ (2013) Pairwise prediction-error expansion for efficient reversible data hiding. IEEE Trans Image Process 22(12):5010–5021
23. Peng F, Li XL, Yang B (2012) Adaptive reversible data hiding scheme based on integer transform. Signal Process 92(1):54–62
24. Qin C, Chang CC, Huang YH, Liao LT (2013) An inpainting-assisted reversible steganographic scheme using a histogram shifting mechanism. IEEE Trans Circuits Syst Video Technol 23(7):1109–1118
25. Sachnev V, Kim HJ, Nam J, Suresh S, Shi YQ (2009) Reversible watermarking algorithm using sorting and prediction. IEEE Trans Circuits Syst Video Technol 19(7):989–999
26. Shi YJ, Qi M, Yi YG, Zhang M, Kong J (2013) Object based dual watermarking for video authentication. Optik 124(19):3827–3834
27. Sühring K "H.264/AVC software coordination," Aug, 2012; http://iphome.hhi.de/suehring/tml
28. Thabit R, Khoo BE (2014) Robust reversible watermarking scheme using Slantlet transform matrix. J Syst Softw 88:74–86
29. Thodi DM, Rodriguez JJ (2007) Expansion embedding techniques for reversible watermarking. IEEE Trans Image Process 16(3):721–730
30. Tian J (2003) Reversible data embedding using a difference expansion. IEEE Trans Circuits Syst Video Technol 13(8):890–896
31. "Video test sequences," Feb, 2013; http://blog.csdn.net/do2jiang/article/details/5499464
32. Wang ZH, Lee CF, Chang CY (2013) Histogram-shifting-imitated reversible data hiding. J Syst Softw 86(2):315–323
33. Wang X, Li XL, Yang B, Guo ZM (2010) Efficient generalized integer transform for reversible watermarking. IEEE Signal Process Lett 17(6):567–570
34. Wang CT, Yu HF (2012) High-capacity reversible data hiding based on multi-histogram modification. Multimedia Tools Appl 61(2):299–319
35. Xu DW, Wang RD, Shi YQ (2014) An improved reversible data hiding-based approach for intra-frame error concealment in H.264/AVC. J Vis Commun Image Represent 25(2):410–422
36. Zeng XA, Chen ZY, Chen M, Xong Z (2011) Reversible video watermarking using motion estimation and prediction error expansion. J Inf Sci Eng 27(2):465–479
37. Zeng X, Chen Z, Xiong Z (2011) Issues and solution on distortion drift in reversible video data hiding. Multimedia Tools Appl 52(2–3):465–484
38. Zhang XP (2013) Reversible data hiding with optimal value transfer. IEEE Trans Multimedia 15(2):316–325

39. Zhang WM, Chen B, Yu NH (2012) Improving various reversible data hiding schemes via optimal codes for binary covers. IEEE Trans Image Process 21(6):2991–3003
40. Zhang WM, Hu XC, Li XL, Yu NH (2013) Recursive histogram modification: establishing equivalency between reversible data hiding and lossless data compression. IEEE Trans Image Process 22(7):2775–2785

Juan Zhao received her B.S. and M.E. degrees from Henan Normal University, Xinxiang, China, in 2007 and 2010. She is currently a PhD student in the Department of Computer Science and Technology at Huazhong University of Science and Technology. Her research interests include network security and multimedia security.

Zhitang Li received his M.E. degree in Computer Architecture from Huazhong University of Science and Technology, Wuhan, China, 1987, and PhD degree in Computer Architecture from Huazhong University of Science and Technology, Wuhan, China, 1992. His research interests include computer architecture, network security, and P2P networks. He was the director of China Education and Research Network (CERNET) in Central China. He was a vice president of Department of Computer Science and Technology, Huazhong University of Science and Technology, China. He has published more than one hundred papers in the areas of network security, computer architecture, and P2P networks.

Bing Feng received his B.S. degree from Naval Aeronautical Engineering Institute, Yantai, China, 2008. He is currently a master degree candidate in Huazhong University of Science and Technology. His research interest is data hiding.