Steganography algorithm based on modified EMD coded PU partition modes for HEVC videos

Zhenzhen Zhang[1], Zhaohong Li [2,*], Jindou Liu [2], Huanma Yan [3]
[1] School of Information Engineering, Beijing Institute of Graphic Communication
Beijing-102600, China
[e-mail: zhangzhenzhen@bigc.edu.cn]
[2] School of Electronic and Information Engineering, Beijing JiaoTong University,
Beijing-100044, China
[e-mail: zhhli2@bjtu.edu.cn; 18120008@bjtu.edu.cn]
[3] Jeejio Technology Co., Ltd, Beijing, 100086, China
[e-mail: yanhuanma@jeejio.com]
*Corresponding author: Zhaohong Li

Abstract

As High Efficiency Video Coding (HEVC) is a worldwide popular video coding standard, the steganography of HEVC videos has gained more and more attention. Prediction Unit (PU) is one of the most important innovative modules of HEVC, thus PU partition mode based steganography is becoming a novel branch of HEVC steganography. However, the embedding capacity of this kind of steganography is limited by the types of PU partition modes. To solve the problem, modified Exploiting Modification Direction (EMD) coded PU partition modes based steganography is proposed in this paper, which can hide a secret digit in a $(2^{n+x}-1)$-ary notational system in a pair of PU partition modes and thus enlarging the capacity. Furthermore, two mapping patterns for PU partition modes are analyzed and the one that performs the better is selected as the final mapping pattern. Firstly, 8×8 and 16×16 sized PU partition modes are recorded according to the optimal mapping pattern in the video encoding process. Then PU partition modes are modified by using the proposed method to satisfy the requirement of secret information. Finally, the stego video can be obtained by re-encoding the video with the modified PU partition modes. Experimental results show that the embedding capacity can be significantly enlarged, and compared with the state-of-the-art work, the proposed method has much larger capacity while keeping high visual quality.

Keywords: steganography, PU partition mode, modified EMD, High Efficiency Video Coding (HEVC)

Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

1. Introduction

Steganography is a hot research issue in the field of information security. It can imperceptibly conceal secret information in the carrier such as texts, images and videos, and has extensive applications in copyright protection, identity authentication, covert communication and so on. Among the various branches of steganography technique, video steganography has attracted more and more attention because video is one of the most widely used digital media and has larger capacity for data embedding compared with other carriers such as texts and audios.

Video steganography is closely linked to video coding standard. High Efficiency Video Coding (HEVC) is the latest video coding standard possessing the best compression efficiency compared with its predecessors such as H.264. HEVC is specifically designed for the increasingly popular high-definition and ultra-high-definition videos [1]. At present, a number of enterprises including Apple and Microsoft have designed products supporting HEVC standard. Therefore, it is worthwhile to develop steganography schemes for HEVC videos.

Modifying DCT/DST coefficients is one common approach to hide data in compressed HEVC videos. Chang et al. [2] classified the blocks in intra-coded frames into 5 categories. Different perturbation patterns were introduced to the luma blocks of each category to avoid intra-frame distortion drift. Similar ideas were taken by Liu et al. to embed data in 4×4 DCT blocks [3] and 4×4 DST blocks [4] of selected frames. Ref. [2-4] were selection-based intra drift free methods. Arige et al. [5] designed the first cancellation-based intra drift free method for HEVC videos which was more general than Ref. [2-4]. Besides 4×4 blocks, 8×8 blocks and 16×16 blocks which satisfies certain coarseness conditions were also employed as carriers in Ref. [6], which made its embedding capacity enlarged.

Intra-prediction mode (IPM)-based steganography is also a hot topic for HEVC videos. Wang et al. [7] proposed an IPM-based stegraphy method by introducing (4, 3) standard array decoding table. The method achieved satisfying embedding capacity while modifying IPMs as few as possible. After that, the authors designed a new mapping table between the angle differences of IPMs and secret information to embed data [8]. The study in [9] explored diamond coding to increase the embedding capacity and could conceal 2-3 bits data by modifying 2 IPMs at most. Wang et al. [10] summarized the existing IPM-based steganography methods and proposed a cover selecting rule by analyzing the probability distribution of 4×4 IPMs.

Besides DCT/DST coefficients and IPM, motion vector and quantization parameter (QP) are another two popular carriers to hide secret information. Van et al. [11] proposed a low complexity steganography method by modifying the motion vector differences and DCT coefficients selected from the frames classified by the inter-prediction dependencies between them. An efficient steganography method based on motion vector space encoding was presented in [12]. Since not every CTU could be used as carrier, the capacity of the method was limited. In the aspect of QP based steganography methods, Zuo et al. [13] combined the just noticeable coding distortion model and exploiting modification direction method to modify the optimal QP values for each coding unit (CU) and embed secret information.

The aforementioned HEVC steganography methods are based on the traditional carriers that are the same with previous video coding standards such as MPEG-4 and H.264. However, HEVC has

developed some innovative technologies including flexible quatree partition architecture (CU, PU and Transform Unit (TU)) and Sample Adaptive Offset. The innovative technologies also offer abundant carriers to embed secret information, thus they opened up a new research field for HEVC stegraphy.

Shanableh [14] utilized the flags indicating the partition of 16×16 sub-CUs as carriers to embed data, where up to an average payload of 32.6Kbit/s can be achieved. Tew et al. [15] divided the prediction block (PB) into two categories. One indicated "0" and the other indicated "1". Data were hidden by modifying the partition mode of PBs. Inspired by Ref. [15], Xie et al. [16] classified the PU partition mode into 3 categories except for "$N\times N$"and "$2N\times 2N$" mode, which enlarged the embedding capacity. By modifying Ref. [16], a multi-level steganography method was designed in Ref. [17], where the embedding capacity can be adjusted according to the actual requirement. Exploiting Modification Direction (EMD) was applied to PU partition mode in Ref. [18] to further enlarge the embedding capacity. These PU partitioning mode based steganography methods have achieved excellent visual quality, but their capacities are relatively low for the limited number of CUs. Even the method of [18], which has improved the capacity a lot compared with [15-17], still required three consecutive CUs to form a group to hide a digit in 7-ary notational system. Stimulated by this, we adopt a modified EMD coding method [19] which only required two elements in each group to pre-code PU partition modes. By modifying the characteristic value function and designing a new codebook, a high-performance HEVC video steganography is proposed in this paper.

CUs with different sizes own different PU partition modes, which have an impact on embedding capacity, thus we investigate the variety of PU partition modes for different sized CUs, and select CUs with sizes of 16×16 and 8×8 to hide information. Furthermore, it is essential to map PU partition modes into integers for PU partition modes based steganography, and two mapping patterns are discussed in this paper. After that, PU partition modes of 16×16 sized CUs and 8×8 sized CUs are forced to be altered according to the modified EMD rules to embed secret data. The experimental results show that the proposed method has much larger capacity than the state-of-the-art works while keeping high visual quality.

The rest of the paper is organized as follows. CU and PU partition modes in HEVC are demonstrated in Section 2. Section 3 discusses the modified EMD method for PU partition modes and the designed mapping patterns. Detailed description of the proposed steganography method is also presented in Section 3. Experimental results are illustrated in Section 4, and Section 5 summarizes the paper.

2. Background of CU and PU partition modes in HEVC

HEVC was jointly released by ITU-T [20] and ISO/IEC [21] and has brought in various new techniques to increase compression efficiency. One of the key techniques is the introduction of quadtree partition structure. In the quadtree partition structure, CU, PU and TU are introduced. CU is the basic coding unit, PU and TU are the basic units for intra/inter prediction and transform/quantization, respectively.

In encoding process, an image is firstly divided into non-overlapping Coding Tree Unit (CTU) whose size is set by the user with 64×64 as default. Then each CTU is recursively divided into

square shapes in a quadtree structure. Every leaf node of a CTU is called a CU which is the basic coding unit in HEVC. There are 4 allowed types of CUs: 64×64, 32×32, 16×16 and 8×8. As shown in **Fig. 1b**, the quadtree structure of the left top CTU surrounded by red square in **Fig. 1a** is displayed. The white lines indicate the boundary of CUs, and the digit denotes the coding order of each CU. We can see that the quadtree partition of CTU is closely related to the image texture. In the region with simple texture such as the tree trunk in **Fig. 1a**, the large CU size is adopted. While in the region with complex texture such as the border between the tree trunk and the road, more detailed CUs are adopted.

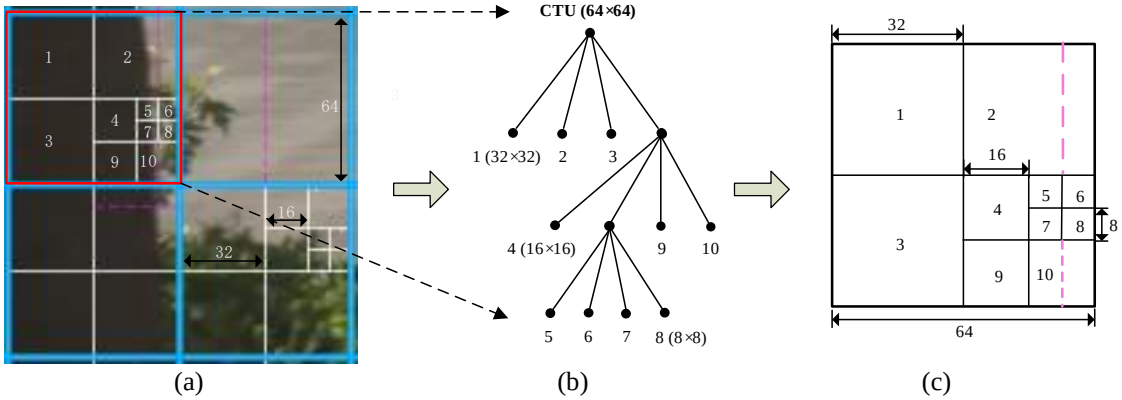


Fig. 1 Example of quadtree partition of CTU into CUs and PUs. (a) CTU partition. (b) CU partition quatree. (c) PU partition.

When performing CTU quadtree tree partition, PU partition mode is determined for each CU in the traversal way according to the rate distortion function. PU partition modes are different under different prediction mode, as presented in **Fig. 2**. In intra prediction mode, one $2N\times2N$ (N=4, 8, 16, 32) CU can be encoded as one single PU or be further split into 4 PUs with equal size ($N\times N$). In inter prediction mode, the CU can be divided into symmetric or asymmetric PUs. Both symmetric and asymmetric PUs are composed of 4 partition modes. **Fig. 1c** shows the PU partition mode of the left top CTU marked red in **Fig. 1a**. The black line implies that the PU has the same size as the corresponding CU, and the purple line means that the CU needs to be subdivided into two PUs in order to get minimum rate distortion.

From **Fig. 2**, we can see that more various PU partition modes are available in inter prediction mode than in intra prediction mode. Since we aim to hide information by modifying the PU partition mode, we select the PU partition mode in inter prediction mode, which means that P picture are adopted to hide information in this paper. Li et al. [18] stated that 8×8 and 16×16 sized CU account for the dominant number of CUs, so we only take 8×8 and 16×16 sized CU into consideration. We can see that 8 PU partition modes for 8×8 sized CU are presented in **Fig. 2**, but in HM 16.15 [22] which is the reference software of HEVC, $N\times N$ PU partition mode is not performed for CUs with size larger than 8×8. Therefore, there are 7 optional PU partition modes for 16×16 CU. **Table 1** and **Table 2** exhibit the supported PU partition modes for 8×8 and 16×16 CU, respectively. Due to the variety of PU partition modes, we determine to use PU partition modes of 8×8 sized CUs and 16×16 sized CUs to hide information in this paper.

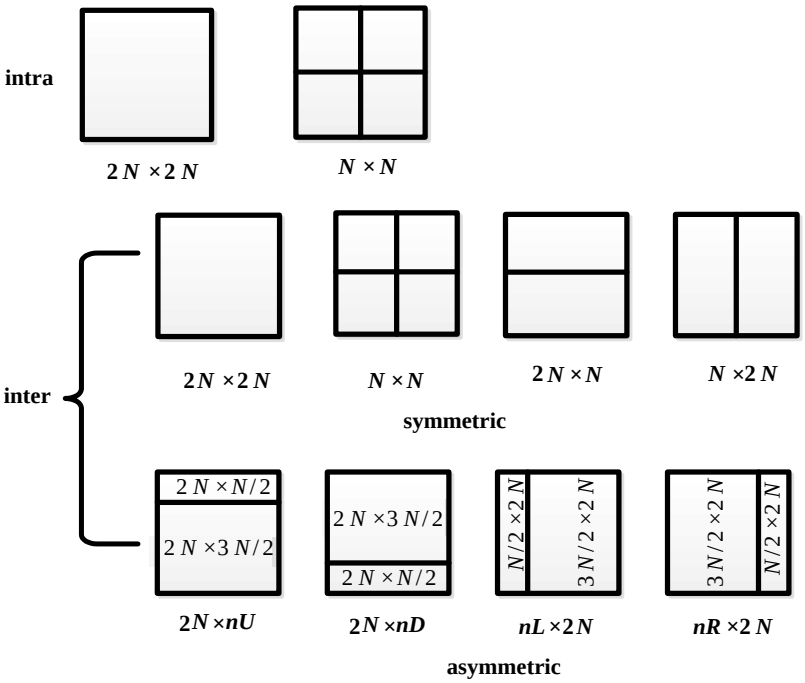


Fig. 2 PU partition modes in different prediction modes

Table 1. PU partition mode for 8×8 CU

Index	1	2	3	4
PU mode	4×4	4×8	8×4	8×8

Table 2. PU partition mode for 16×16 CU

Index	1	2	3	4
PU mode	16×16	16×8	8×16	4×16
Index	5	6	7	-
PU mode	16×12	16×4	12×16	-

3. The proposed steganography method

In the latest EMD-coded PU partition modes based steganography [18], each secret digit in a $(2n+1)$-ary notational system can be concealed by modifying only one PU partition mode in the group of n PU partition modes. However, the embedding capacity can be enlarged by utilizing the modified EMD method which introduces the concept of codebook. In the modified EMD coding, a $(2^{n+x}-1)$ base digit can be concealed in one PU partition mode group. Stimulated by this, the modified EMD method is introduced to pre-code PU partition modes for further improving the

embedding capacity in this paper. The proposed steganography based on the modified EMD is firstly illustrated. Then two mapping patterns for PU partition modes are discussed. After that, the detailed data hiding process which adopts the optimal designed mapping pattern is described.

3.1 Proposed steganography based on PU partition modes

In the proposed steganography, CTU is assigned as the basic processing unit. For a CTU which consists of T 16×16 sized CUs, it can be divided into a series of groups containing n ($n \leqslant T$) consecutive CUs each. Since one CU corresponds to one PU partition mode, each CU group will contain n PU partition modes. After mapping (two mapping patterns have been designed and will be discussed in next section) each PU partition mode into an integer $p_i, i = 1,2,\cdots,n$, a vector of $(p_1, p_2,\cdots, p_n)$ which denotes one PU partition modes group can be achieved. Then the characteristic value F of the group can be calculated as a weighted sum modulo ($2^{n+x}-1$), as shown in Eqs. (1) and (2), where $(b_1, b_2,\cdots, b_n)$ denotes a basis vector which depends on the integer x.

$$F = f(p_1, p_2,\cdots, p_n) = \left[\sum_{i=1}^{n}(p_i \cdot b_i)\right] mod\ (2^{n+x}-1) \tag{1}$$

$$[b_1, b_2,\cdots, b_n] = \begin{cases} 1,6,\cdots,(n-1)\times 6, & if\ (x \geq 3\ and\ n \geq 2) \\ 1,4,\cdots,(n-1)\times 4, & if\ (1 \leq x \leq 3\ and\ n \geq 2) \end{cases} \tag{2}$$

The hiding rule of the proposed steganography with modified EMD is to make the characteristic value F equal to the secret digit s in a ($2^{n+x}-1$)-ary notational system. If $s = F$, no alteration of the current group of PU partition modes is needed. On the contrary, a codebook generated by the modified EMD is designed to change the PU partition mode. The index of the codebook satisfies Eq. (1), and the rule of the codebook is to make minimum alteration of the integers which represent different PU partition modes.

Table 3 illustrates the codebook when n=2 and x=3. If $s \neq F$, the index d is calculated by Eq. (3). Then look up the codebook and find the row with the index d, the vector $(d_1, d_2,\cdots, d_n)$ in the specific row is used to modify the PU partition modes group. Finally the vector of the modified PU partition modes group $(p_1', p_2',\cdots, p_n')$ can be obtained by Eq. (4), and the characteristic value $f(p_1', p_2',\cdots, p_n')$ will be equal to the secret digit s.

$$d = \begin{cases} s - F & if\ (s \geq F) \\ 2^{n+x} - 1 - |s - F| & if\ (s < F) \end{cases} \tag{3}$$

$$p_i' = p_i + d_i \tag{4}$$

When applying modified EMD coded PU partition modes, two important issues should be taken into consideration. One is the selection of n and x. The other one is the overflow or underflow problem.

Table 3. Codebook when n = 2 and x = 3 with $b_1 = 1$ and $b_2 = 6$

Index	d_1	d_2	Index	d_1	d_2
0	0	0	16	-3	-2
1	1	0	17	-2	-2
2	2	0	18	-1	-2
3	3	0	19	0	-2
4	-2	1	20	1	-2
5	-1	1	21	-2	-2
6	0	1	22	-3	-1
7	1	1	23	-2	-1
8	2	1	24	-1	-1
9	3	1	25	0	-1
10	-2	2	26	1	-1
11	-1	2	27	2	-1
12	0	2	28	-3	0
13	1	2	29	-2	0
14	2	2	30	-1	0
15	3	2	-	-	-

As for the first consideration, the setting of n and x is closely related to the embedding capacity. Supposing that the number of 16×16 sized CU in one video picture is denoted by Q, since $\log_2(2^{n+x}-1)$ bits can be hidden in one group containing n PU partition modes, the embedding capacity can be obtained according to Eq. (5), where C means the capacity of one video picture. The partial derivative of C to n in Eq. (6) shows that the capacity will decrease with the increase of n. Therefore, in order to get the maximum capacity, we set n=2 in this paper. Please note that the value of n should be larger than one. The reason is that if n=1, only one PU partition mode is contained in one group, and the minimum alteration of the PU partition mode cannot be guaranteed when the difference between the characteristic value F and secret digit s is large, which will further affect the visual quality of the video.

$$C = Q\log_2(2^{n+x}-1)/n \tag{5}$$

$$\begin{aligned}\frac{\partial C}{\partial n} &= \frac{\partial}{\partial n}\left(Q\log_2(2^{n+x}-1)/n\right) \\ &= Q\frac{1+\left(\frac{1}{2^{n+x}-1}\right)-\log_2\left(2^{n+x}-1\right)}{n^2} < 0, \quad n \ge 2, x \ge 1\end{aligned} \tag{6}$$

Referring to the relationship between C and x, it is obvious that more bits can be hidden with larger x. However, the value of x is limited by the available number of PU partition modes. As n=2, (p_1, p_2) denotes a group of PU partition modes, 49 combinations of (p_1, p_2) can be got because both p_1 and p_2 own 7 available values. One requirement of the modified EMD coding is that the characteristic values F calculated by all the combinations of (p_1, p_2) should cover every value in the $(2^{n+x}-1)$-ary notational system, i.e. from 0 to $2^{n+x}-2$. Through theoretical analysis, it is

found when $x > 3$, the range from 0 to $(2^{n+x} - 2)$ is too large to cover by these 49 combinations of (p_1, p_2), hence it is better to set $x \le 3$. Taking the above two factors together, we set n=2 and x=3 for 16×16 sized PU partition modes in this paper. For 8×8 sized CUs, the hiding rule is the same as 16×16 sized CUs. By analyzing the setting of n and x in a similar way to 16×16 sized CUs, we select n=2 and x=2 for 8×8 sized CUs to get the maximum embedding capacity in this paper. The codebook for n=2 and x =2 is exhibited in **Table 4**.

Table 4. Codebook for n = 2 and x = 2 with $b_1 = 1$ and $b_2 = 4$

Index	d_1	d_2	Index	d_1	d_2
0	0	0	8	1	-2
1	1	0	9	-2	-1
2	2	0	10	-1	-1
3	-1	1	11	0	-1
4	0	1	12	1	-1
5	1	1	13	-2	0
6	2	1	14	-1	0
7	-1	2	-	-	-

For 16×16 sized CUs, from Eqs. (3) and (4) and **Table 3**, we can see that the modified PU partition mode p_i' is obtained by p_i adding or subtracting an integer. The modifying operation may lead to p_i' larger than 7 or smaller than 0, which is called overflow or underflow. Take (p_1, p_2)=(1,2) for instance, the characteristic value F=f (p_1, p_2) is 13 with n=2 and x=3. If the secret digit s = 4, according to the codebook and Eq. (3) and (4), $(p_1', p_2') = (p_1, p_2) + (-3, -1) = (-2, 1)$ can be got, which indicates the case of underflow that p_1' is out of the range from 0 to 6 and the PU partition mode corresponding to -2 cannot be found. To solve the problem, the 49 combinations of (p_1, p_2) are traversed, and the candidate groups, whose f (p_1, p_2) value equals to the secret digit, are selected. Let $\boldsymbol{P} = \{\boldsymbol{p}_l = (p_{l_1}, p_{l_2}), l \in \{1, 2, \cdots, k\}\}$ denotes the set of the candidate groups, where k means the number of candidate groups, the Manhattan Distance md of each group from (p_1, p_2) is calculated as Eq. (7) and the group which owns the minimum Manhattan Distance from (p_1, p_2) is considered as the final group to replace (p_1, p_2). The overflow and underflow question for 8×8 sized CUs are considered in the same way as 16×16 sized CUs.

$$md = |\boldsymbol{p}_l - \boldsymbol{p}| = |p_{l_1} - p_1| + |p_{l_2} - p_2| \tag{7}$$

3.2 Optimal mapping pattern for PU partition modes

As demonstrated in Section 3.1, each PU partition mode should be mapped into an integer for being coded by the modified EMD method. There are 7 and 4 different PU partition modes for one 16×16 sized CU and 8×8 sized CU, respectively, hence 7 integers range from 0 to 6 and 4 integers range from 0 to 3 are adopted to represent PU partition modes for one 16×16 sized CU and 8×8 sized CU, respectively. However, different mapping patterns will affect the visual quality and bitrate of the

video at different levels. Therefore, two mapping patterns for PU partition modes are considered in this paper, and the mapping pattern which causes the less change in the visual quality and bitrate of the video is selected.

The first mapping pattern is determined by the relationship between the distribution of PU partition modes and the probability of modifying other integers to each specific integer in the modified EMD method. We call this mapping pattern "mapping 1" for abbreviation. Besides the consideration of mapping 1, the similarity between adjacent PU partition modes is also considered in the second mapping pattern which is referred to as "mapping 2".

Table 5. The probability of modifying other integers to each specific integer for 16×16 sized CU in the modified EMD coding method

integer	0	1	2	3	4	5	6
probability	0.0984	0.1402	0.1777	0.1801	0.1784	0.1346	0.0905

For mapping 1, we firstly calculate the probability of modifying other integers to each specific integer in the modified EMD method. The situation with n=2 and x=3 as shown in **Table 3** for the proposed steganography is tested, which indicates that a pair of PU partition modes denoted by (p_1, p_2) are taken as cover. All available integers for (p_1, p_2) pairs are traversed in the test. For each pair of (p_1, p_2), every $2^{n+x}-1$ base number which suggests the secret digit, is embedded following the proposed steganography. During the traversal, (p_1, p_2) will be changed to fit the $2^{n+x}-1$ base number, and the probability of modifying (p_1, p_2) to each integer in the set $\{0,1,2,3,4,5,6\}$ is calculated, as shown in **Table 5**. We can see that the integer "3" owns the largest probability among the 7 integers, it means that most PU partition modes would be altered to the PU partition mode which is mapped to integer 3 during the embedding process.

Besides, the distribution of each PU partition mode for videos with different resolutions has been investigated and shown in **Fig. 3**,. Each video contains six P pictures, and the quantization parameter adopted is 32. It is observed that $2N\times 2N$ PU partition mode takes the first place at any resolution for 16×16 sized CU, $N\times 2N$ PU partition mode the second, and $2N\times N$ PU partition mode the third. When designing mapping rule, it is a feasible way to map the PU partition mode to the specific integer according to the occurrence frequency of the PU partition mode and the probability of modifying other integers to the specific integer. For example, $2N\times 2N$ PU partition mode is the dominant one in the distribution of PU partition modes, so we map it to the integer 3 that owns the largest probability to be altered to. In this way, the distribution of each PU partition mode would change the least before and after embedding with modified EMD method. With this consideration, mapping 1 of PU partition modes for 16×16 sized CU can be obtained, as listed in **Table 6**.

The relationship between the distribution of PU partition modes and the probability of modifying other integers to each specific integer for 8×8 sized CU is also considered in the proposed stenography. In order to save the space of the paper, considering the experiments and analysis are the same as 16×16 sized CU, we just present the designed mapping 1 pattern for 8×8 sized CU, as shown in **Table 7**.

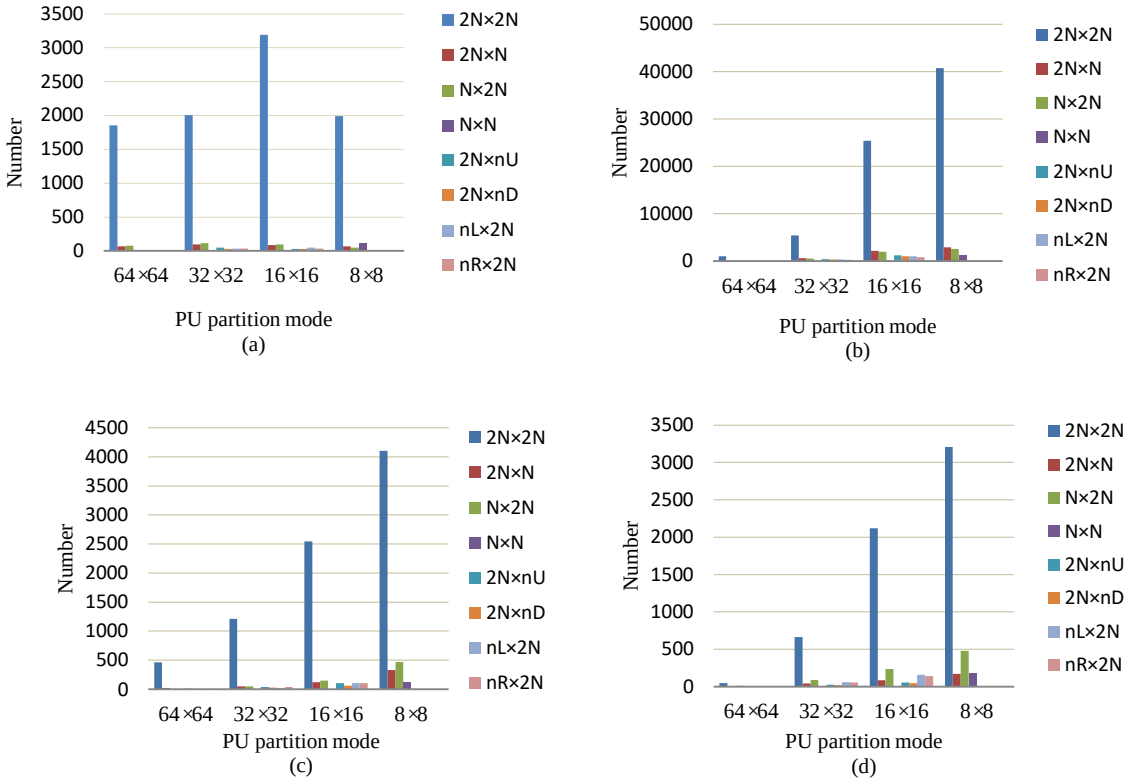


Fig. 3 Distribution of PU partition modes for videos with different resolutions (a) BasketballDrive_1920×1080. (b) PeopleOnStreet_2560×1600. (c) Chinaspeed_1024×768. (d) Race Horses_832×480.

Table 6. Mapping 1 of PU partition modes for 16×16 sized CU

Integer	0	1	2	3	4	5	6
PU partition mode	$2N\times nU$	$nL\times 2N$	$2N\times N$	$2N\times 2N$	$N\times 2N$	$nR\times 2N$	$2N\times nD$

Table 7. Mapping 1 of PU partition modes for 8×8 sized CU

Integer	0	1	2	3
PU partition mode	$2N\times N$	$N\times 2N$	$2N\times 2N$	$N\times N$

The idea of modified EMD method is to change the PU partition mode to its adjacent mode if modification is required. The change of PU partition mode will increase the video's bitrate and decrease its visual quality. In order to reduce the effect of the modification, we also consider the similarity between adjacent PU partition modes on the basis of mapping 1. Since $2N\times2N$, $N\times2N$, and $2N\times N$ PU partition mode take the largest, second largest, and third largest probability of all the optional PU partition modes, respectively, the mapping of them in mapping 2 remains the same as that in mapping 1. The adjacent similarity is considered for the left four PU partition modes and mapping 2 can be obtained. **Table 8** and **Table 9** exhibit the specific mapping pattern of mapping 2 for 16×16 sized CU and 8×8 sized CU, respectively. Take integer "1" in **Table 8** as an example, both its neighbors "0" and "2" indicate the similar PU partition mode with it.

Table 8. Mapping 2 of PU partition modes for 16×16 sized CU

Integer	0	1	2	3	4	5	6
PU partition mode	$2N\times n$D	$2N\times n$U	$2N\times N$	$2N\times 2N$	$N\times 2N$	nL$\times 2N$	nR$\times 2N$

Table 9. Mapping 2 of PU partition modes for 8×8 sized CU

Integer	0	1	2	3
PU partition mode	$2N\times N$	$2N\times 2N$	$N\times 2N$	$N\times N$

In order to take a comparison of the two mapping rules, four YUV videos with different resolutions are adopted to implement data hiding on 16×16 sized CUs using the proposed steganography. Quantization parameter (QP for abbreviation) is one of the most important parameters for HEVC steganography, thus three QPs are tested in the experiments. Since data embedding process influences the video's visual quality and bitrate, ΔPSNR and BRI are used as measurement to indicate the change of visual quality and bitrate, respectively. Eqs. (8) and (9) illustrate the definition of ΔPSNR and BRI, where $PSNR$ and $PSNR'$ denote the PSNR of the cover video without and with data embedding, respectively, BR and BR' represent the bitrate of the cover video without and with data embedding, respectively.

$$\Delta PSNR=PSNR'-PSNR \tag{8}$$

$$BRI=(BR'-BR)/BR \tag{9}$$

Fig. 4 and**Fig. 5** exhibit the average ΔPSNR and BRI of the four videos for the two mapping patterns. From **Fig. 4**, we can see that both the two mapping patterns make a decrease on visual quality at any QPs, but mapping 2 decreases the less. The same phenomenon occurs in **Fig. 5**,

where mapping 2 has the lower BRI comparing with mapping 1. Therefore, mapping 2 is adopted in the proposed method, and the experimental results are based on mapping 2 hereafter.

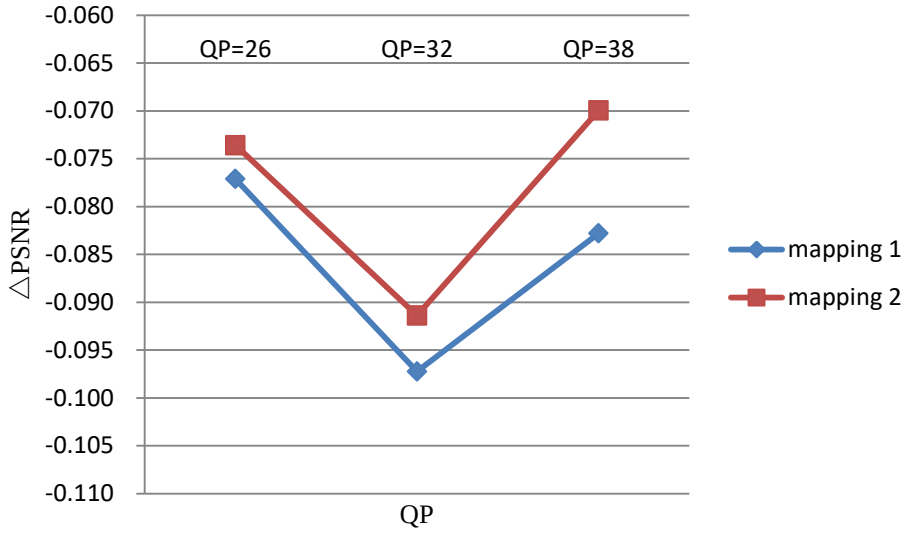


Fig. 4 The average ΔPSNR of the four videos for the two mapping patterns

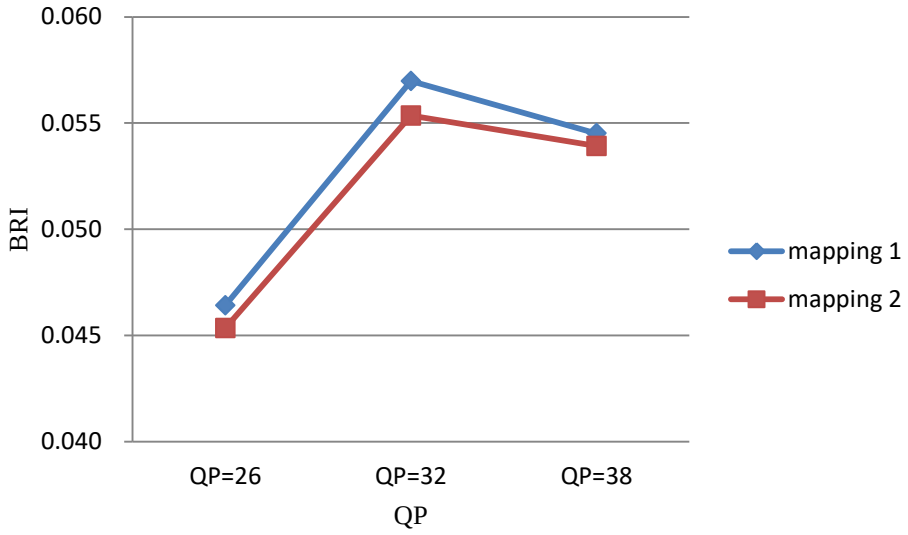


Fig. 5 The average BRI of the four videos for the two mapping patterns

3.3 Embedding and extraction process

Here, we will demonstrate the data embedding and extraction process of the proposed steganography in detail.

Step 1: For one CTU, encode it with HEVC encoders and record the numbers of 16×16 sized CUs, 8×8 sized CUs and the optimal PU partition mode for each CU.

Step 2: For both the 16×16 sized CUs and the 8×8 sized CUs, divide the CUs into multiple groups containing n CUs each. For each group, calculate the characteristic value of the group according to Eqs. (1) and (2). For the given secret digit s in a $(2^{n+x}-1)$-ary notational system, calculate the index d of the codebook by utilizing Eq. (3), and then the modified PU partition modes of the current group can be obtained.

Step 3: Repeat Step 2 to get the modified PU partition modes for each group of the CUs. After that, the CTU is re-encoded by using the modified PU partition modes.

Step 4: Perform Steps 1-3 for all the CTUs in every P-picture of the video.

In order to demonstrate the steps of the proposed method more reproducible, the pseudocode of the embedding process of one PU partition modes group from 16×16 sized CUs is given in **Table 10**, where n=2 and x=3.

Table 10. The pseudocode of the embedding process of one PU partition modes group from 16×16 sized CUs

Input: a group containing two PU partition modes, denoted by $\boldsymbol{p}=(p_1, p_2)$, a basis vector $\boldsymbol{b}=(b_1, b_2)$, and a secret digit s in a $(2^{2+3}-1)$-ary notational system

Output: the modified group of $\boldsymbol{p}$, denoted by $\boldsymbol{p}'$

$F = \boldsymbol{p}\cdot\boldsymbol{b} = p_1\times b_1 + p_2\times b_2$
if $s \geq F$
  $d = s\text{-}F$
else if $s < F$
  $d = 2^{2+3}-1-|s-F|$
end if
$\boldsymbol{p}_{\text{temp}} = (p_1+d_1, p_2+d_2)$
if (any element of $\boldsymbol{p}_{\text{temp}}$ larger than 7 or smaller than 0)
  for i=0:1:7
    for j = 0:1:7
      record all the $\boldsymbol{p}_l = (p_{l_1}, p_{l_2}) = (i, j)$ that makes $i\times b_1 + j\times b_2 == s$, and store them in one set
      $\boldsymbol{P} = \{\boldsymbol{p}_l = (p_{l_1}, p_{l_2}), l \in \{1, 2, \cdots, k\}\}$
    end for
  end for
for l=1:1:k
  calculate the Manhattan Distance md of each element in $\boldsymbol{P}$ from $\boldsymbol{p}$, i.e. $|\boldsymbol{p}_l - \boldsymbol{p}| = |p_{l_1}\text{-}p_1| + |p_{l_2}\text{-}p_2|$
end for
$\boldsymbol{p}' = \arg\min_{\boldsymbol{p}_l}(|p_{l_1} - p_1| + |p_{l_2} - p_2|)$
else if
  $\boldsymbol{p}' = \boldsymbol{p}_{\text{temp}}$
end if

In the stage of extracting secret information from the stego video, we just need to divide the 16×16 sized CUs and 8×8 sized CUs into multiple groups and calculate the characteristic value of

each group, then concatenate the characteristic value of each group to get the secret information as the secret digit equals to the characteristic value of each PU partition modes group in the stego video.

4. Results and Discussion

4.1 Configurations

In this section, the experimental results of the proposed algorithm are presented. Twelve common YUV video sequences with different resolutions are used to conduct the experiments, including 1 sequence with resolution 2560×1600: Traffic; 4 sequences with resolution 1920×1080: Kimono, ParkScene, Cactus, Ducks_take_off; 2 sequences with resolution 1600×1200: Intersection, Mainroad; 3 sequences with resolution 1280×720: City, Crew, Johnny; 2 sequences with resolution 832×480: Keiba, PartyScene. With respect to the video codec platform, HM 16.15 [22] with the main profile is employed. The frame rate, GOP size, and GOP structure are set as 25, 4 and IPPP, respectively. All the videos are tested with three different QPs: 26, 32 and 38.

4.2 Evaluation of subjective visual quality

High subjective visual quality is one of the good behaviors of an excellent video steganography, which means that human eyes cannot perceive the visual distortion of the stego-video. **Fig. 6** displays the first P picture of sequence Kimino before and after data embedding at different QPs. It is shown that the difference between the original picture and the information hidden picture is so small that it is difficult for human eyes to visually perceived, which indicates the high subjective visual quality of the proposed method.





Fig. 6 Subjective visual quality evaluation at different QPs for Kimino 1920×1080. The original pictures are laid on the first column, and pictures with hidden information are placed in the second column. The QP adopted in the first line, second line and third line are 26, 32 and 38, respectively.

4.3 Objective performance

In order to evaluate the performance of the proposed algorithm objectively, three evaluation metrics are adopted in this paper: ΔPSNR, capacity and BRI. ΔPSNR and BRI, as shown in Eqs. (8) and (9), are used to assess the influence of the proposed embedding process on the visual quality and bitrate of the cover video, respectively. The capacity of the proposed algorithm is calculated as the average bits that can be hidden in one P picture.

Table 11 exhibits the results of the proposed algorithm on the three evaluation metrics. From **Table 11**, we can get that the average capacity of all the sequences for QP 26, 32 and 38 are 4190, 1379.83, and 437.75 bits per picture, respectively, which shows the high capacity of the proposed method. Furthermore, it shows a tendency that more bits can be hidden for sequences with lower QP or higher resolution in most cases. For example, the capacity of sequence "Traffic" with QP 26 is 10971 bits, 7474 bits more than that with QP 32, and 7700 bits more than sequence "Kimino" with QP 26 which has lower resolution. This is because more 8×8 and 16×16 sized CUs will be allocated by the encoder in the case of lower QP or higher resolution, leading to more available PU partition modes that can be embedded in.

In **Table 11**, PSNR (clean) means the PSNR of the video sequence without data hiding. ΔPSNR indicates the visual quality distortion caused by the proposed embedding process, the smaller the absolute value of ΔPSNR is, the less distortion it will be. From **Table 11**, we can see that the value of ΔPSNR is within 0.1dB in most cases, even larger than zero in some cases. For example, the value of ΔPSNR for sequence "Intersection" with QP 38 is 0.009dB. The average ΔPSNR value of all the sequences for QP 26, 32, and 38 are -0.0546, -0.0373, and -0.0127, respectively. These results certify that the stego videos after data embedding with the proposed method still keep high visual quality. Furthermore, it is shown that the absolute value of ΔPSNR decreases with the increase of QP as a whole, which means less visual quality distortion for larger QPs. The reason for this is that the numbers of 8×8 and 16×16 sized CUs are smaller for larger QP than lower QP, and less 8×8 and 16×16 PU partition modes are modified for larger QP.

Since the proposed method altered the PU partition modes from optimal to non-optimal, it will cause the increase of the bitrate. BRI is a common measure for bitrate increase and low BRI is pursued. From **Table 11**, we can see that except the sequence "Keiba", the BRI of the proposed method for the other sequences is less than 0.06, and the average BRI of all the sequences for QP 26,

32, and 38 is 0.040, 0.025, and 0.017, respectively, which means that the bitrate increase of the proposed method is acceptable.

Table 11. The ΔPSNR, capacity and BRI of the proposed algorithm with different QPs

Sequence	QP	PSNR (clean)	ΔPSNR	BRI	Capacity
Traffic	26	40.070	-0.071	0.047	10971
2560×1600	32	37.305	-0.036	0.023	3497
	38	34.405	-0.006	0.014	1049
Kimono	26	41.108	-0.047	0.044	3271
1920×1080	32	38.352	-0.030	0.037	1708
	38	35.580	-0.010	0.024	610
ParkScene	26	38.971	-0.075	0.050	7716
1920×1080	32	36.113	-0.040	0.027	2423
	38	33.413	0.002	0.017	608
Cactus	26	37.936	-0.057	0.044	5722
1920×1080	32	35.846	-0.037	0.027	2190
	38	33.378	-0.006	0.018	726
Ducks_take_off	26	35.166	-0.066	0.037	11734
1920×1080	32	32.727	-0.038	0.020	3156
	38	30.446	-0.003	0.013	837
Intersection	26	44.411	-0.012	0.009	532
1600×1200	32	41.271	-0.004	0.006	158
	38	38.316	0.009	0.007	66
Mainroad	26	40.376	-0.028	0.019	3100
1600×1200	32	37.729	-0.009	0.005	242
	38	34.706	0.000	0.005	57
City	26	39.320	-0.035	0.020	1399
1280×720	32	36.480	-0.016	0.008	334
	38	33.210	-0.007	0.003	62
Crew	26	41.100	-0.034	0.052	957
1280×720	32	39.010	-0.016	0.022	191
	38	36.721	-0.008	0.018	58
Johnny	26	42.798	-0.015	0.020	406
1280×720	32	40.683	-0.009	0.010	123
	38	38.011	0.001	0.006	37
Keiba	26	38.908	-0.084	0.081	1746
832×480	32	35.745	-0.094	0.073	976
	38	32.850	-0.063	0.058	419
PartyScene	26	36.561	-0.131	0.057	2726
832×480	32	32.629	-0.119	0.036	1560
	38	28.965	-0.061	0.025	724

4.4 Comparative analysis

To evaluate the performance of the proposed method, two outstanding PU partition modes based steganography algorithms are selected for comparison. The first one, named Yang's method [17], is published in 2018, in which a multilevel information hiding algorithm is proposed and different levels can be chosen according to the requirement of embedding capacity and video visual quality.

The second one, called Li's method [18], is the latest PU partition mode based steganography and has substantial embedding capacity. In this section, the three algorithms are implemented on four different video sequences for comparison, including Kimono and ParkScene with resolution 1920×1080, Keiba and PartyScene with resolution 832×480. Because different algorithms cause different compression bitrates though the same video sequence is tested, in order to demonstrate the relationship between visual quality, embedding capacity and the compression bitrate more intuitively, the curves of bitrate and capacity (RC) and bitrate and distortion (RD) are plotted to compare the results.

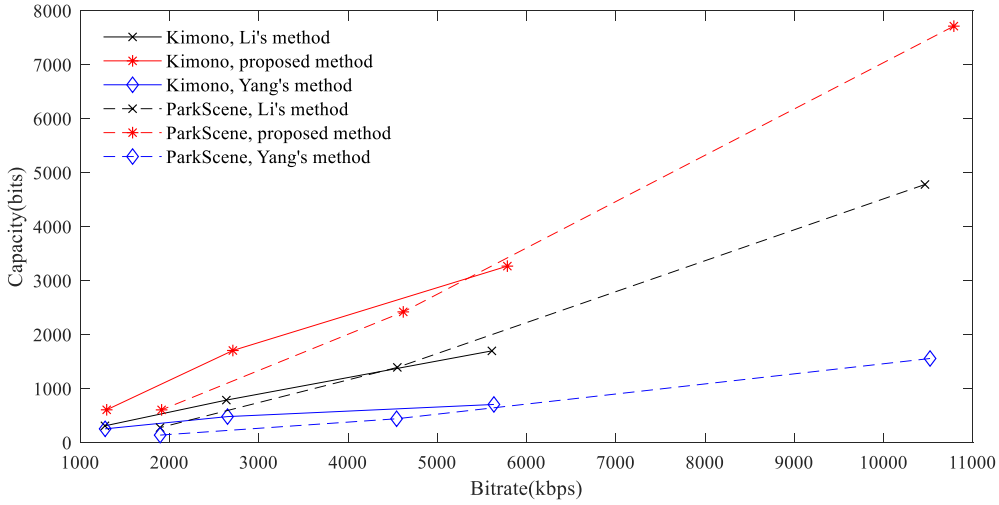


Fig. 7 Comparison of RC with Li's method and Yang's method for sequences Kimino and ParkScene

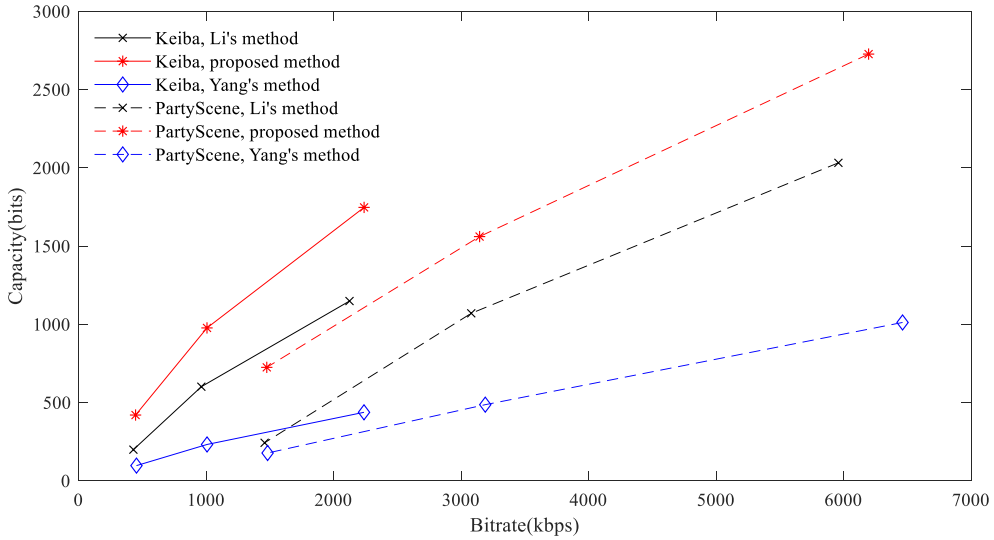


Fig. 8 Comparison of RC with Li's method and Yang's method for sequences Keiba and PartyScene

Fig. 7 and**Fig. 8** describe the RC curves of the three methods. As the capacity of each of the four levels in Yang's method is quite different, and Level 4 has the maximum capacity, we chose Level 4 for Yang's method for comparison. The cross lines, diamond lines, and asterisk lines represent the RC curves of Li's method, Yang's method and the proposed method, respectively. From **Fig. 7** and**Fig. 8**, we can see that the capacity of the proposed method is the largest among the three methods at any bitrate. Averagely, the capacity of the proposed method is 68% higher than Li's method, and it is about 4 times that of Yang's method. Since $2N\times 2N$ PU partition modes which account for the largest proportion are not used to hide information, the lowest embedding capacity is obtained in Yang's method. Furthermore, in Li's method, $\log_2 3$ bits and $\log_2 7/3$ bits can be hidden for modifying one 8×8 PU partition mode and one 16×16 PU partition mode, respectively, while the numerical values are $\log_2 15/2$ bits and $\log_2 31/2$ bits for modifying one 8×8 PU partition modes and one 16×16 PU partition modes, respectively, in the proposed method. This explains the largest capacity of the proposed method theoretically.

RD curves of the three methods are plotted in **Fig. 9** and**Fig. 10**. From **Fig. 9**, we can see that the PSNR of the proposed method keeps nearly the same with Yang's method and Li's method. In **Fig. 10**, the PSNR of the proposed method is second-best among the three methods, with 0.06dB lower than Li's method and 0.23dB higher than Yang's method on average. The reason for the little lower PSNR of the proposed method than Li's method is that more PU partition modes are modified in the proposed method to get a larger embedding capacity. Since the capacity of Li's method is enlarged by 68% by the proposed method, it is acceptable to get a 0.06dB visual quality distortion.

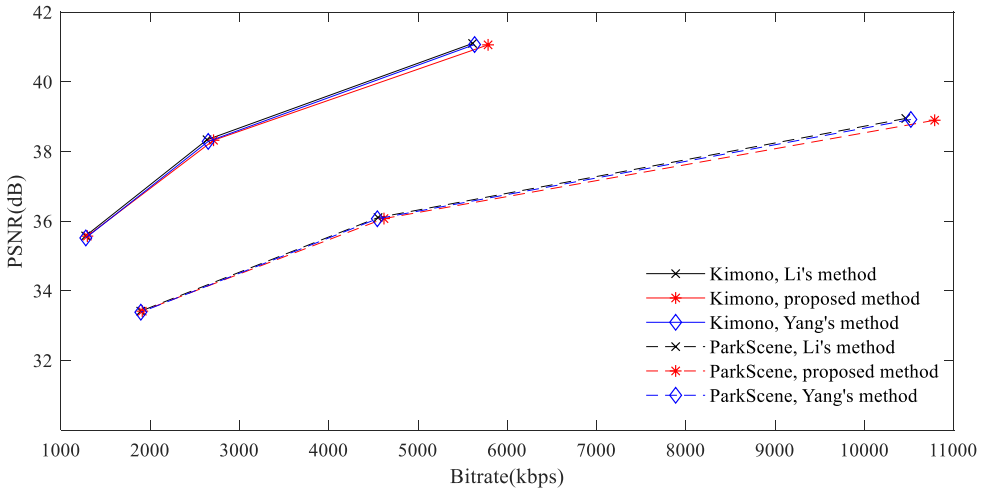


Fig. 9 Comparison of RD with Li's method and Yang's method for sequences Kimino and ParkScene

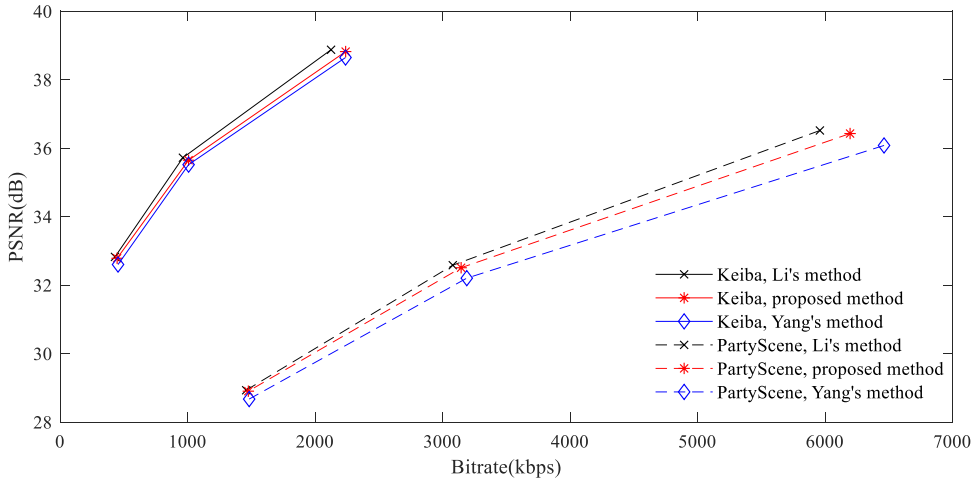


Fig. 10 Comparison of RD with Li's method and Yang's method for sequences Keiba and PartyScene

4.5 Anti-steganalysis analysis

Steganalysis is the counter-technique to steganography, which aims to detect the presence of secret information in digital media. Thus having resistance to steganalysis attacks is an important requirement for a good steganography method. Sheng et al. [23] proposed the first HEVC steganalysis method for PU partition mode based steganography. The amount and proportion change rate of 4×4, 8×8, and 16×16 PU in I-pictures before and after recompression were utilized as the feature set, which can effectively detect the steganography. Thus Sheng's method is reconstructed and used to test the security of the proposed method. 37 YUV sequences with various resolutions are used as the video set, among which 30 sequences are used as training samples and the left are used as testing samples. The training and testing process are repeated 100 times and the mean value of detection rate is listed in **Table 12**. From **Table 12**, we can see that the detection rate of Sheng's method is about 0.5 at any QP, which indicates that Sheng's algorithm is almost inefficient to detect the proposed method, and thus the proposed steganography has acceptable resistance to Sheng's method.

Table 12. The detection rate of Sheng's method to test the proposed algorithm

QP	26	32	38
Detection rate	0.5021	0.5013	0.4986

5. Conclusions

In this paper, an effective HEVC steganography algorithm based on modified EMD coded PU partition modes is proposed. The characteristic value calculation equation is modified and the optimal mapping pattern is investigated in order to enlarge the capacity of the PU partition mode based steganography. Then 8×8 and 16×16 sized PU partition modes are preferred as carriers to

hide information according to the embedding process described in Section 3.3. As a result, the embedding capacity can be enlarged significantly, with nearly 1.68 times and 4 times of Li's method and Yang's method, respectively. At the same time, the proposed method keeps a high visual quality with no more than 0.1 dB distortion less than cover videos in most cases. In future work, we will look for more outstanding coding methods for PU partition modes or other HEVC features to further improve the performance of the proposed method.

Abbreviations

CTU: Coding Tree Unit; CU: coding unit; EMD: Exploiting Modification Direction; HEVC: High Efficiency Video Coding; IPM: Intra-prediction mode; PB: prediction block; PU: Prediction Unit; QP: quantization parameter; TU: Transform Unit

Acknowledgements

The authors would like to thank the editor and the reviewers for their helpful comments and valuable suggestions.

Funding

This work was supported by the National Natural Science Foundation of China (No.61702034), Joint Funding Project of Beijing Municipal Commission of Education and Beijing Natural Science Fund Committee (KZ201710015010), and initial funding for the Doctoral Program of BIGC (27170120003/037)

Availability of data and materials

Please contact the authors for data (in the experiment used videos or coding files with implemented method) requests.

Author Contributions

Zz Zhang and Zh Li conceived the idea of this work. Zz Zhang, Jd Liu and Hm Yan performed the experiments. Zz Zhang and Zh Li drafted the manuscript. Zhli, Jd Liu and Hm Yan refined the manuscript. All authors read and approved the final version of the manuscript.

Ethics approval and consent to participate

Approved.

Consent for publication

Approved.

Competing interests

The authors declare that they have no competing interests.

References

[1] A. Costanzo, M. Barni, in *13th IEEE International Conference on Signal Processing*. Detection of double AVC/HEVC encoding, (Chengdu, China, 2016), pp. 2245-2249

[2] P.C. Chang, K.L Chung, J.J.Chen, C.H. Lin, T.J. Lin, A DCT/DST-based error propagation-free data hiding algorithm for HEVC intra-coded frames. *J. Vis. Commun. Image Represent*. 2014, 25(2), 239-253 (2014)

[3] Y. Liu, S. Liu, H. Zhao, et al. in *International Conference on Intelligent Computing*. A data hiding method for H. 265 without intra-frame distortion drift, (Liverpool, United Kingdom, 2017), pp. 642-650

[4] Y. Liu, S. Liu, H. Zhao, S. Liu, A new data hiding method for H. 265/HEVC video streams without intra-frame distortion drift. *Multimedia Tools and Applications*. 78(6), 6459-6486 (2019)

[5] A. Arige, M. Mitrea, I. Boujelbane, HEVC intra-frame drift cancellation matrix. *Journal of Visual Communication and Image Representation*. 62, 56-67 (2019)

[6] K. Biswas, A robust and high capacity data hiding method for H. 265/HEVC compressed videos with block roughness measure and error correcting techniques. *Symmetry*. 11(11), 1360 (2019)

[7] J. J. Wang, R. D. Wang, W. Li, et al. Video information hiding in intra prediction and block codes for hevc. *Guangdianzi Jiguang/J Optoelectron Laser*. 26(5), 942-950 (2015)

[8] J. J. Wang, R. D. Wang, D. W. Xu, et al. An information hiding algorithm for HEVC based on angle differences of intra prediction mode. *Journal of Software*. 10(2), 213-221 (2015)

[9] Y. Chen, Y. Zhou, X. Wang, et al. Video information hiding algorithm based on diamond coding. *Journal of Computer Applications*. 37(10), 2806-2812 (2017)

[10] J. Wang, X. Jia, X. Kang, et al. A cover selection HEVC video steganography based on intra prediction mode. *IEEE Access*. 7, 119393-119402 (2019)

[11] L. P. Van, J. D. Praeter, G. V. Wallendael, et al. in *IEEE International Conference on Image Processing (ICIP 2015)*. Out-of-the-loop information hiding for HEVC video, (Quebec City, Canada, 2015), pp. 3610-3614

[12] J. Yang, S. Li, An efficient information hiding method based on motion vector space encoding for HEVC. *Multimedia Tools and Applications*. 77(10), 11979-12001 (2018)

[13] L. W. Zuo，T. Luo，G. Y. Jiang，et al．Large-capacity information hiding method for HEVC based on just noticeable coding distortion. *Journal of Image and Graphics*. 22(4), 0443-0451 (2017)

[14] T. Shanableh, Data embedding in high efficiency video coding (HEVC) videos by modifying the partitioning of coding units. *IET Image Processing*. 13(11), 1909-1913 (2019)

[15] Y. Tew, K. S. Wong, in *IEEE International Conference on Image Processing (ICIP 2014)*. Information hiding in HEVC standard using adaptive coding block size decision, (Paris, France, 2014), pp. 5502-5506.

[16] W. Xie, Y. Yang, Z. Li, et al. in *International Conference on Cloud Computing and Security*. An information hiding algorithm for HEVC videos based on PU partitioning modes, (Haikou, China, 2018), pp. 252-264.

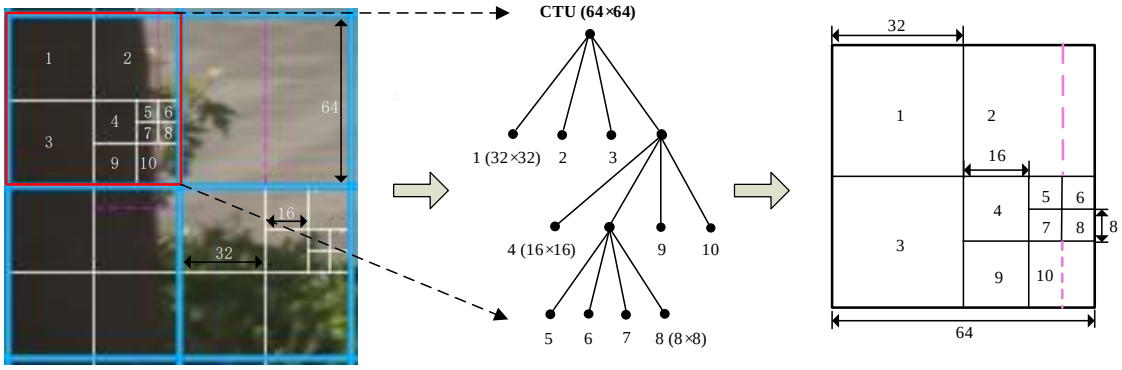
[17] Y. Yang, Z. Li, W. Xie, et al. High capacity and multilevel information hiding algorithm based on pu partition modes for HEVC videos. *Multimedia Tools and Applications*. 78(7), 8423-8446 (2019)

[18] Z. Li, L. Meng, X. Jiang, et al. High capacity HEVC video hiding algorithm based on EMD coded PU partition modes. *Symmetry*. 11(8), 1015 (2019)

[19] C. Kim, Data hiding by an improved exploiting modification direction. *Multimedia Tools and Applications*. 69(3), 569-584 (2014)

[20] Rec. ITU-T H.265: High Efficiency Video Coding. Available online: https://www.itu.int/rec/T-REC-H.265 (accessed on 31 August 2020).
[21] International Organization for Standardization. ISO/IEC 23008-2:2017. Information Technology—High Efficiency Coding and Media Delivery in Heterogeneous Environments—Part 2: High Efficiency Video Coding. Available online: https://www.iso.org/standard/69668.html (accessed on 31 August 2020).
[22] HEVC reference software : HM. Available online: https://hevc.hhi.fraunhofer.de/svn/svn_HEVCSoftware/tags/ (accessed on 31 August 2020).
[23] Q. Sheng, R. Wang, M. Huang, et al. A prediction mode steganalysis detection algorithm for hevc. *J Optoelectron-laser*. 28(4), 433-440 (2017)

Figure title and legend



(a) (b) (c)

Fig. 1 Example of quadtree partition of CTU into CUs and PUs. (a) CTU partition. (b) CU partition quatree. (c) PU partition.

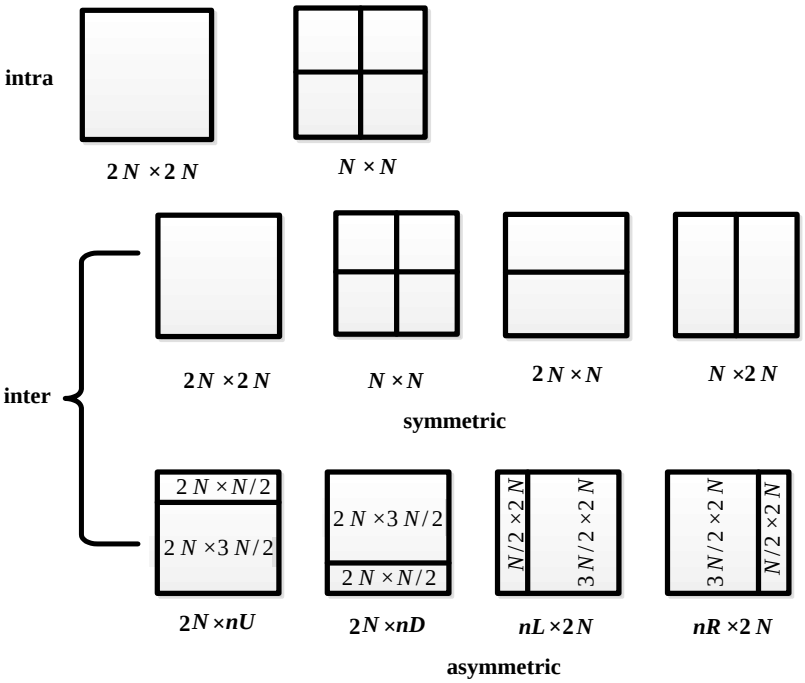


Fig. 2 PU partition modes in different prediction modes

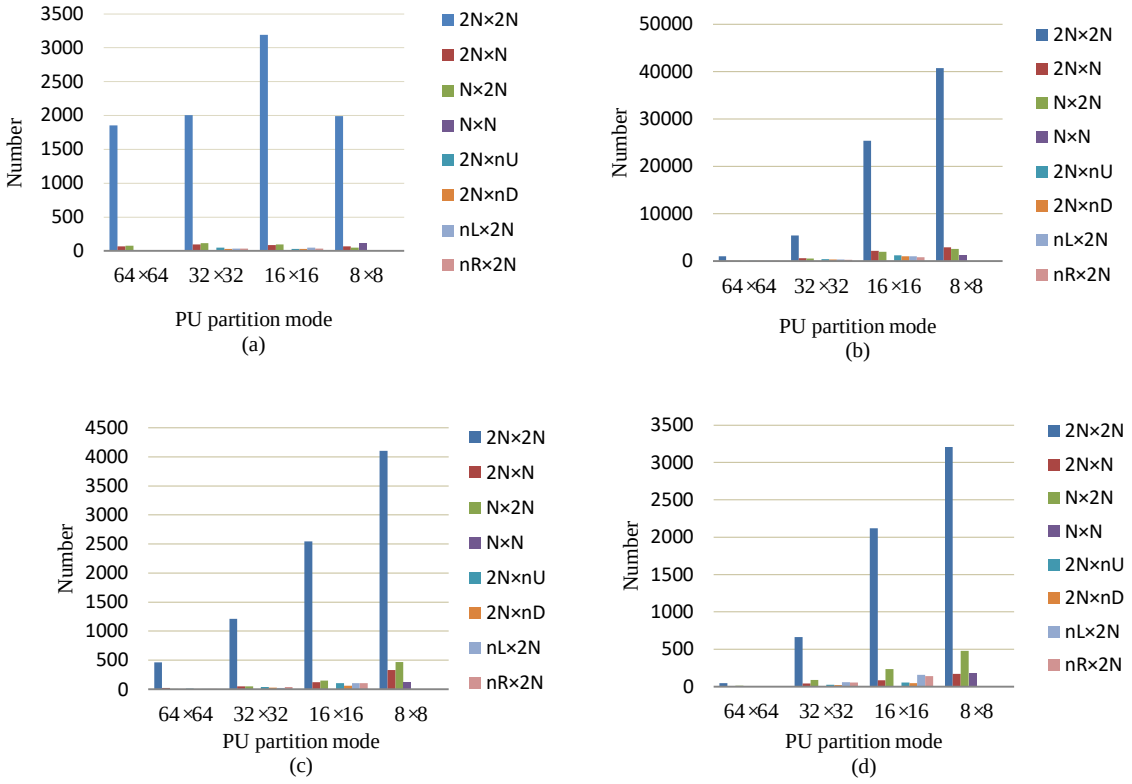


Fig. 3 Distribution of PU partition modes for videos with different resolutions (a) BasketballDrive_1920×1080. (b) PeopleOnStreet_2560×1600. (c) Chinaspeed_1024×768. (d) Race Horses_832×480.

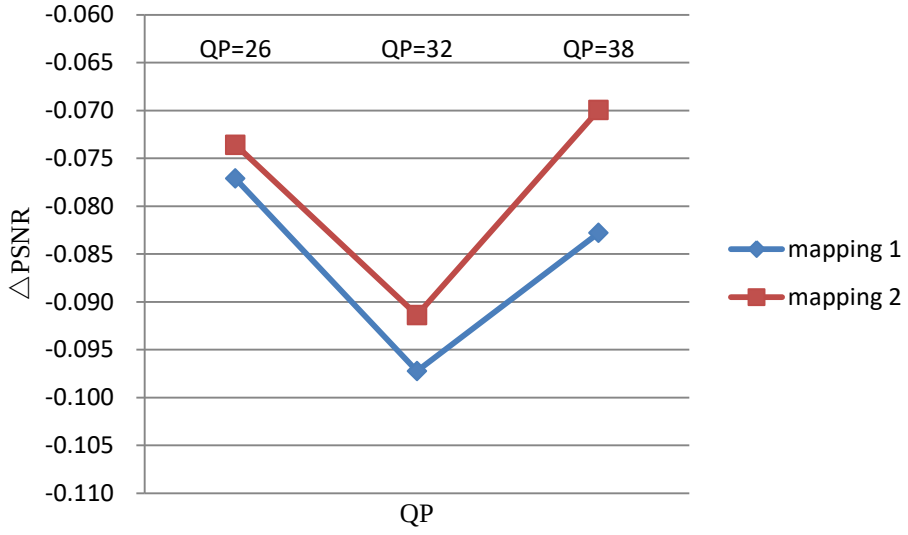


Fig. 4 The average ΔPSNR of the four videos for the two mapping patterns

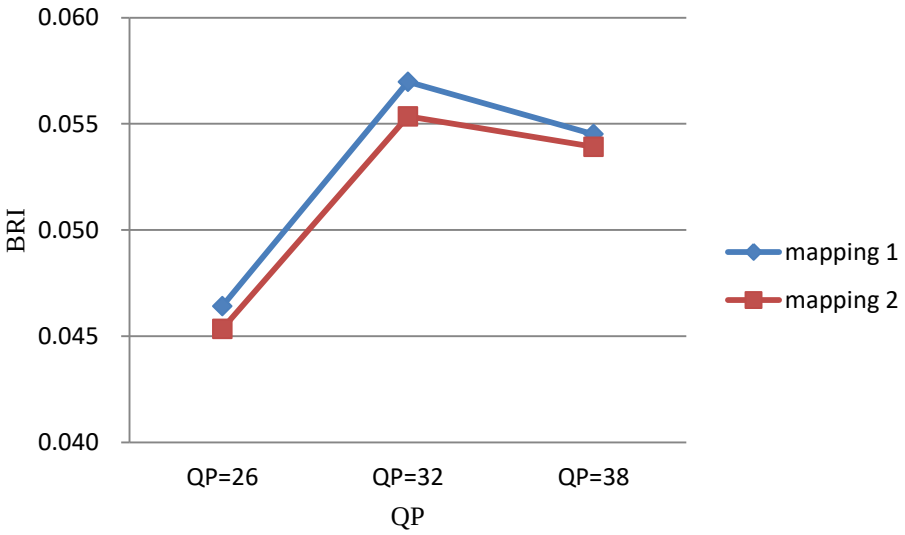


Fig. 5 The average BRI of the four videos for the two mapping patterns



Fig. 6 Subjective visual quality evaluation at different QPs for Kimino 1920×1080. The original pictures are laid on the first column, and pictures with hidden information are placed in the second column. The QP adopted in the first line, second line and third line are 26, 32 and 38, respectively.

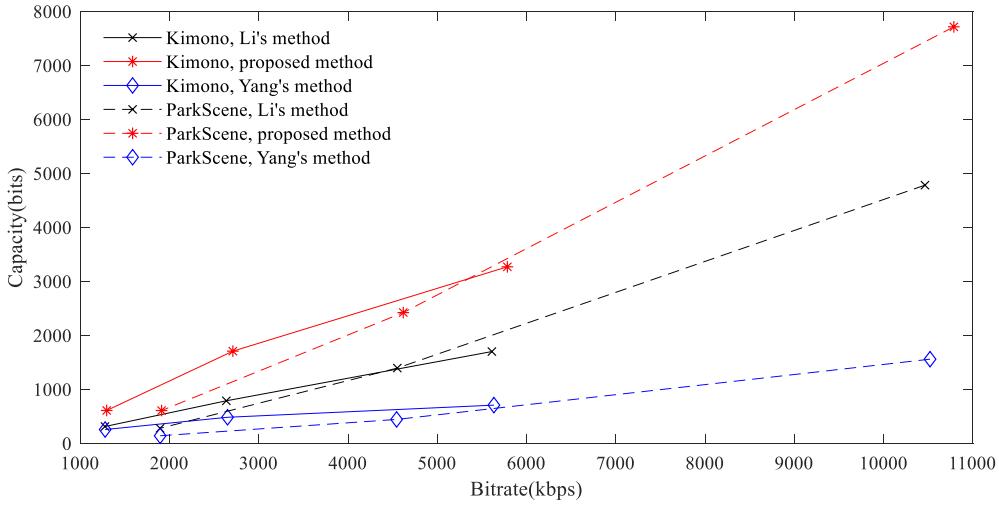


Fig. 7 Comparison of RC with Li's method and Yang's method for sequences Kimino and ParkScene

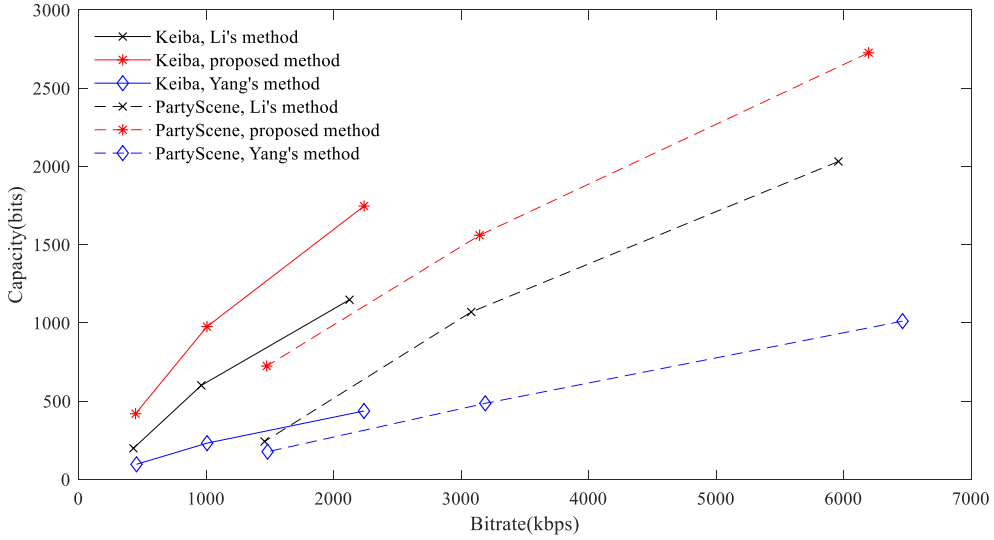


Fig. 8 Comparison of RC with Li's method and Yang's method for sequences Keiba and PartyScene

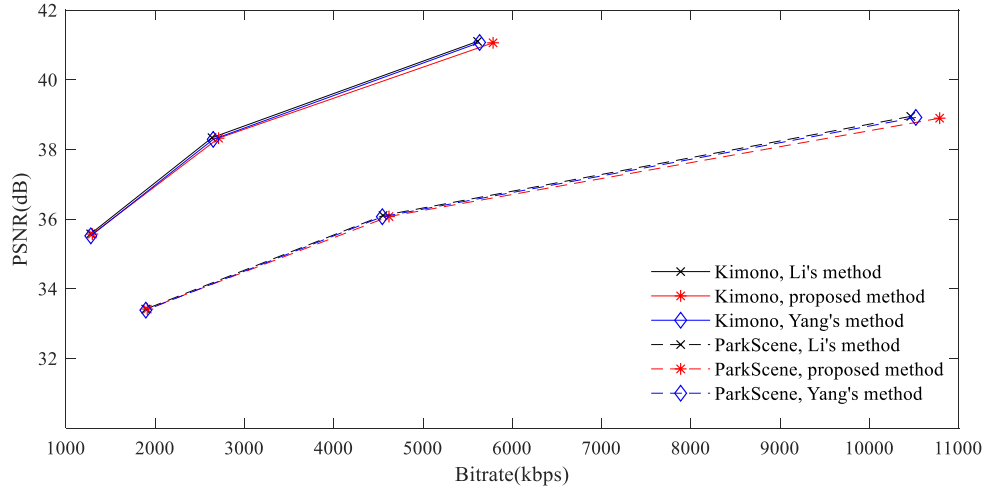


Fig. 9 Comparison of RD with Li's method and Yang's method for sequences Kimino and ParkScene

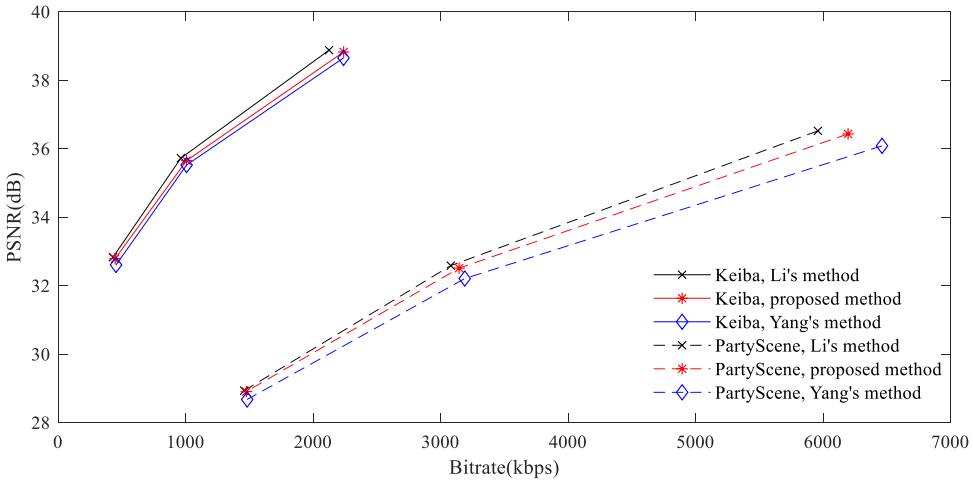


Fig. 10 Comparison of RD with Li's method and Yang's method for sequences Keiba and PartyScene