A High Capacity HEVC Steganography Using Intra Prediction Modes in Multi-sized Prediction Blocks

Yi Dong[1], Tanfeng Sun[1,2] and Xinghao Jiang[1,2*]

[1] School of Electronic Information and Electrical Engineering
Shanghai Jiao Tong University
[2] National Engineering Lab on Information Content Analysis Techniques, GT036001
Shanghai, China
{aa44,tfsun,xhjiang}@sjtu.edu.cn

Abstract. The existing video steganographic schemes based on intra prediction modes for video coding standards H.264/AVC and HEVC all use single-sized blocks to embed the secret payload. Thus, the steganographic properties of HEVC multi-sized tree-structured intra partition still need exploration. In this paper, a novel video steganography algorithm is presented. Base on the fact that visual quality degradation caused by steganography is basically the same for both large-sized Prediction Blocks (PBs) and small-sized PBs, this algorithm tries to exploit intra prediction modes in multi-sized PBs in each Coding Tree Units (CTU). The innovation of this paper includes: 1) Improvement in capacity without introducing great degradation in visual quality. 2) High coding efficiency maintained by defining cost function based on rate distortion. The Experimental results show that this algorithm outperforms the latest intra prediction modes based HEVC steganography algorithm in both capacity and perceptibility while preserving coding efficiency as well.

Keywords: HEVC, intra prediction modes, video steganography.

1 Introduction

With the development of broadband network and mobile Internet technology, the transmission and service based on video media are booming. Video media in HEVC format, because of its high resolution and small file size, are very suitable as carrier of secret communication, with the possibility of large capacity communication provided. On the other hand, unlike image steganography, HEVC video steganography can naturally conceal that the communication is occurring from user behavior [1, 2]. While HEVC video steganography can ensure the rationality of the user's behavior and reduce the risk of exposing the hidden communication.

Many works have been done in both H.264/AVC and HEVC [3, 4]. Hu et al. [5] proposed a steganography algorithm based on intra prediction mode in H.264/AVC. Yang et al. [6] have improved Hu's method by matrix coding. Bouchama [7] divided the intra prediction modes in H.264/AVC into four groups according to their predic-

tion direction, the result shows a better video quality while ensuring high capacity. Zhang et al. [8] analyzed the texture of the video, and proposed a high security adaptive embedding algorithm using STC. Wang et al. [9] proposed intra prediction mode based method for HEVC, a mapping between angle difference and secret message was established to embed data. Dong et al. [10] further proposed the prediction mode steganography technology under the HEVC standard, and made a breakthrough in the capacity limitation of the previous HEVC intra prediction mode based algorithm, while also improving the security.

As far as selection of intra prediction modes is concerned, previous steganography methods in H.264/AVC all choose to use intra prediction modes in 4×4 macroblock to embed the secret message. This selection rule is reasonable in H.264/AVC since the capacity of embedding into 16×16 macroblock is low, and this kind of macroblock usually concern homogeneous areas for which the Human Visual System (HVS) is more sensitive to small degradations. Thus, in the previous HEVC steganography schemes, authors in [9, 10] still use the PB of 4×4 size as the embedding cover. However, this selection rule ignores some objective conditions in HEVC: 1) HEVC partitioning is achieved using tree structures. It supports variable-sized PBs selected according to needs of encoders in terms of video content and resolution. 2) Previous H.264/AVC steganography schemes are all tested under low resolution video dataset. But in high resolution dataset for HEVC, larger size PBs occur more frequently. Using only small size PBs will limit capacity to a great extent. In short, the capacity of previous HEVC steganography schemes is limited since unique techniques in HEVC are not considered sufficiently.

In order to solve the problem mentioned above, and make full use of new features introduced by HEVC, an extension of our previous work in [10] is made. The innovation of this paper includes: 1) Improvement in capacity without introducing great degradation in visual quality. 2) High coding efficiency maintained by defining cost function based on rate distortion.

The rest of this paper is organized as follows. In Section 2, detailed analysis on why large-sized PBs can be modified without introducing great degradation in visual quality is presented. Section 3 describes the proposed HEVC steganography algorithm. In Section 4, experiments and analysis on multi-resolution dataset are presented. Finally, conclusion is drawn in Section 5.

2 Analysis of HEVC Intra Coding Scheme

In this section, the HEVC intra coding scheme will be first described, with which analysis of visual quality degradation in HEVC intra steganography can be thoroughly introduced next.

2.1 HEVC Intra Coding Scheme

The HEVC standard introduces CTU and Coding Tree Block (CTB) structure to intra coding scheme. Each frame in a video is first split into block-shaped CTUs, which each contain luma CTBs and chroma CTBs. The blocks specified as CTBs can either be directly used as Coding Blocks (CBs) or be further partitioned into multiple CBs.

As shown in Fig.1, Partitioning is achieved using tree structures. An intra predicted CB of size *M*×*M* may have one of the two types of PB partitions referred to as PART-2N×2N and PART-N×N, the first of which indicates that the CB is not split and the second indicates that the CB is split into four equal-sized PBs.

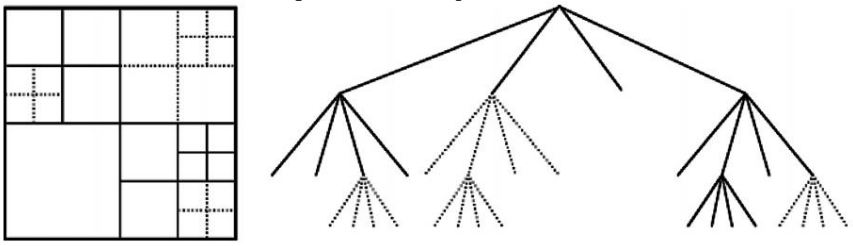

Fig. 1. HEVC tree structured partitioning

The PB size, which is the block size at which the intra prediction mode is established is the same as the CB size except for the smallest CB size (usually 8×8) is allowed in the bitstream. For the latter case, a flag is present that indicates whether the CB is split into four PB quadrants, each PB with their own intra prediction mode. The actual region size at which the intra prediction operates depends on the residual coding partitioning.

For residual coding, a CB can be recursively partitioned into Transform Blocks (TBs). The partitioning is signaled by a residual quadtree. Intra prediction operates based on the TB size, and previously decoded boundary samples from spatially neighboring TBs are used to form the prediction signal. Directional prediction with 33 different directional orientations is defined for (square) TB sizes from 4×4 to 32×32.

2.2 Analysis of Visual Quality Degradation in HEVC Intra Steganography

According to the HEVC intra coding scheme, this subsection will present the analysis of visual quality degradation caused by HEVC steganography, in order to illustrate the reason why large-sized PBs can be modified without introducing significant visual distortion.

Spatial-domain intra prediction has previously been successfully used in H.264/AVC. The intra prediction of HEVC operates similarly in the spatial domain, but is extended significantly—compared to the eight prediction directions of H.264/AVC, HEVC supports a total of 33 angular prediction directions with DC and Planar mode.

The residual signal of the intra prediction, which is the difference between the original block and its prediction, is transformed by a linear spatial transform. The coefficients are then scaled, quantized, entropy coded, and transmitted together with the prediction information. When the prediction mode m_1 of i^{th} N×N size PB is modified to m_2. The original residual of this PB, denoted as $RS^o{}_{i,N}$, can be expressed as:

$$RS^o{}_{i,N} = P^o{}_{i,N} - Pre^o{}_{i,N} \tag{1}$$

Where ${P^o}_{i,N}$ denotes the original pixel value in the i[th] PB, and ${Pre^o}_{i,N}$ denotes the prediction value calculated by original mode m_1.After obtaining the ${RS^o}_{i,N}$, the bits ${B^o}_{i,N}$ used to encode this PB can be expressed as follows:

$$ {B^o}_{i,N} = Ent(RT(\frac{DCT({RS^o}_{i,N})}{Q \times QS})) \tag{2} $$

Where *DCT*(.) denotes the integer discrete cosine transform, *RT*(.) denotes the rounding and truncating operations, *Ent*(.) denotes entropy coding, *Q* denotes the fixed quantization matrix, *QS* denotes the quartier scale. In the decoding process, the reconstruction residual $RSR^o_{i,N}$ can be calculated as:

$$ RSR^o_{i,N} = IDCT\left(IEnt\left({B^o}_{i,N}\right) \times Q \times QS\right) \tag{3} $$

Where *IDCT*(.) denotes the inverse integer discrete cosine transform, *IEnt*(.) denotes the inverse entropy coding, and decoded pixel value $PR^o_{i,N}$can be expressed as:

$$ PR^o_{i,N} = FTR(RSR^o_{i,N} + {Pre^o}_{i,N}) \approx {P^o}_{i,N} \tag{4} $$

Where *FTR*(.) denotes deblocking filter and Sample Adaptive Offset (SAO) operations. Equation (4) shows that the difference between decoded pixel value and original value is mostly depended on quantization. For the next M×M sized PB, its prediction value is determined as:

$$ {Pre^o}_{i+1,M} = SF_N\ (PR^o_{i,N}) \tag{5} $$

$SF_N(.)$ denotes the intra estimation, prediction and smoothing operation applied by HEVC according to the block size. It shows that as long as we keep the size of candidate PB unchanged, the prediction value of next PB will not be affected significantly. After modifying the prediction modes to m_2, the sum of modified prediction value and its residual is still equal to the true pixel value. So, after processing the modified residual value with same parameters as the original, following equation can be obtained:

$$ {PR^m}_{i,N} = FTR(RSR^m_{i,N} + {Pre^m}_{i,N}) \approx {P^o}_{i,N} \tag{6} $$

Where ${PR^m}_{i,N}$ denotes the modified reconstruction value and $RSR^m_{i,N}$ denotes the modified reconstruction residual. Thus, the conclusion can be drawn:

$$ {PR^m}_{i,N} \approx {P^o}_{i,N} \approx PR^o_{i,N} \tag{7} $$

From equation (7), it shows that visual quality of videos generated by this kind of steganography algorithms will not degrade significantly. Another conclusion can be drawn that the difference among values in equation (4-7) is mainly caused by the choice of *Q* and *QS*. Thus, the degradation of visual quality will be mainly caused by the increment of QP, not increment of embedded bits. Intra mode steganography has more potential capacity with multi-sized PBs.

To summarize, the visual quality of generated video file won't be affected by changes in the HEVC intra coding process. Thus, improving capacity by utilizing angular intra prediction modes in multi-sized PBs is viable in theory.

3 The Proposed HEVC Steganography

In this section, based on the above analysis, large-sized PBs can be modified without introducing great degradation in visual quality. The remaining problem is to keep the coding efficiency during the process of embedding secret message. The framework of the proposed algorithm is shown in Fig.2.

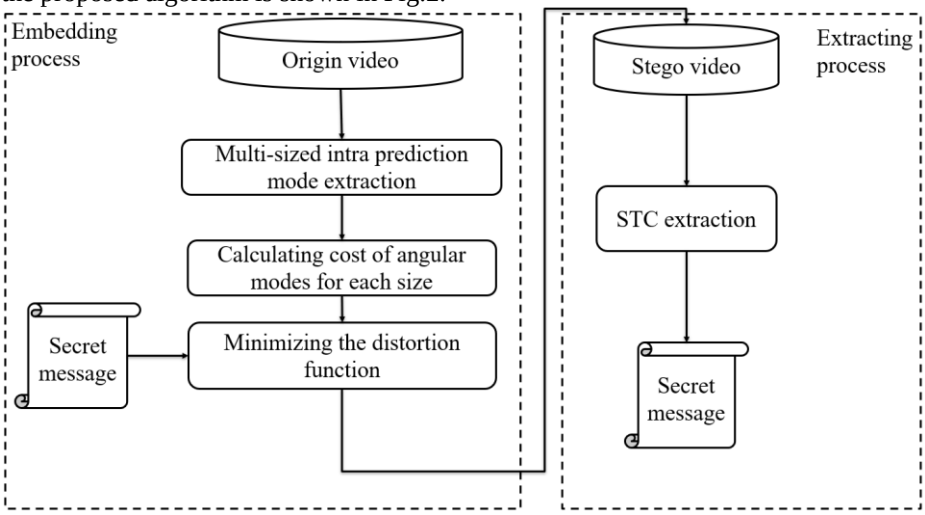

Fig. 2. The framework of the proposed algorithm

3.1 Selection Rule of Intra Prediction Mode

According to the recursive procedure of block partition in HEVC, when QP increases, the number of small blocks decreases. Utilizing large-size PBs will improve the steganography capacity in videos with high QP.

Based the 35 intra prediction directions used in HEVC, it is noticeable that unlike the prediction directions in H.264/AVC, modes of HEVC have a regular pattern. In HEVC, two number-adjacent directions have similar prediction direction. The common way to modify the intra prediction mode is to replace it with a mode that is similar in prediction direction. In this case, these modes are grouped as follows:

$$\{(M_i, M_j) \mid 2 \mid M_i = 2 \mid M_j, \quad \mathrm{i,j} \in (0,34)\} \tag{8}$$

Where the $\mid$ symbol means exact division, M_i means that the current PB prediction mode has the i^{th} prediction direction. One element in the group denotes the bit 0, another denotes the bit 1. According to equation (8), the final grouping is {(0, 1), (2, 3)…}. However, changing mode 0 and mode 1 will significantly affect coding efficiency since they are usually used to encode homogeneous areas. Thus, the first group (0, 1) is removed here. Finally, all the qualified prediction modes of PBs in I frames are extracted, and taken as cover sequence.

3.2 Rate Control Method

However, according to the analysis in the above section, if the best mode is altered, then the coding efficiency will suffer. In HEVC, Rate Distortion Optimization (RDO) technique is used to achieve the best prediction direction:

$$J = D + \lambda R \tag{9}$$

Where J denotes the RD cost, λ denotes the Lagrangian multiplier which depends on quantization parameter QP, D and R represent the distortion and the estimated bitrate of the current PB respectively. The best intra prediction mode is judged by the lowest RD cost. Therefore, if the corresponding RD cost of the current PB is increases, which means residual signal of stego block is larger than original block, the number of bits used to encode this block will increase. Thus, the total coding efficiency will decline.

In order to reduce the proposed algorithm's influence on coding efficiency. The STC method is utilized to embed the secret message into cover:

$$Hx^T = m \tag{10}$$

Where H denotes the parity check matrix generated by STC algorithm, m is the secret message and x is the modified cover sequence. Detail description and implementation of STC can be found in [11]. According to the selection rule of Intra prediction mode, one prediction mode has one candidate mode that can replace it. Thus, this is a binary STC problem. After grouping these prediction modes, the following equation is used to map them into binary sequence:

$$c_i = m_i \bmod 2 \tag{11}$$

Where c_i denotes the binary cover and m_i denotes the original cover.

From equation (9), each RD cost is calculated through estimated bitrate and distortion of each PB. Thus, difference in RD cost can represent the coding efficiency reduction caused by changing the prediction mode of the current PB. The cost of changing one PB is defined as:

$$\varphi_i = \left|J_i - J_j\right|, \ i,j \ \ is\ from\ the\ same\ group \tag{12}$$

Where φ_i is the cost of changing the i^{th} PB and J_i is the RD cost of the prediction mode with the i^{th} prediction direction. The difference of RD cost between two prediction modes is used in the same group as the cost for changing one to another. The total distortion D_c is present as:

$$D_c = \sum_{k=1}^{n} \varphi_k \tag{13}$$

Where n presents the total number of all qualified PBs in the video file. Finally, the secret data can be embedded into a video with little distortion.

4 Experiments and Analysis

Since the proposed algorithm is the first intra prediction modes based algorithm using multi-sized PB designed specifically for HEVC, previous single-sized PB based algorithms in [9,10] are selected for comparison. Performance comparisons between the proposed steganography algorithm and the previous one are made in terms of embedding capacity, SSIM, and bitrate. Moreover, the effect of introducing large-sized PBs on visual quality and effect of different QPs will also being analyzed in this section.

4.1 Experiment Setup

Dataset and Development Environment. The proposed steganography algorithm has been implemented in an open source software X265. HEVC is the state-of-art video codec standard, designed for high definition videos aiming to achieve higher coding efficiency. For this reason, the proposed algorithm was tested on HEVC standard test dataset with multi-resolution. In these experiments, pseudo random binary sequences are generated as secret data, and payload is set to α=0.5, in order to produce the stego sets. The GOP size is 10 and coding structure is IPPP. The video coding platform for HEVC decoding is HM16. The algorithm is developed with Visual C++ 2013. The details of experiment dataset in listed in Table 1.

Table 1. Details of the dataset.

Video name	Resolution	Frame number
Traffic	2560×1600	150
PeopleOnStreet	2560×1600	150
ParkScene	1920×1080	240
BasketballDrive	1920×1080	501
Johnny	1280×720	600
FourPeople	1280×720	600

Indicators. In this section, several indicators are used to measure the performance of the proposed algorithm and other algorithm. Capacity, SSIM, Bit Increase Ratio (BIR) and Bit Increase Ratio with Normalized Capacity (BIR-NC). BIR is defined as:

$$\mathrm{BIR} = \frac{TB_{steg} - TB_{ori}}{TB_{ori}} \times 100\% \tag{14}$$

Where TB_{steg} is the total bits of original video and TB_{ori} is the total bits of modified video.

Because the original cover length is different for these two algorithms, a new indicator is proposed to measure the capacity under different BIR. BIR is normalized with 1 Kbits to show the coding efficiency reduction, named as BIR-NC. The physical

meaning of BIR-NC is the bit increase ratio using secret payload of the same size. This is very common in real application. The definition of BIR-NC is：

$$\text{BIR-NC} = \text{BIR}/\text{Capacity} \tag{15}$$

4.2 Comparison Experiment

In this section, the proposed steganography algorithm will be compared with previous algorithms[9,10] on a different dataset, because algorithm[9] only design a mapping rule between secret message and cover, while proposed an individual dataset . For objective comparison, comparison experiment of [9] is performed under the same setup in [9]. Moreover, for algorithm [10], it also includes the distortion control method as the proposed one, so this algorithm is performed on the above dataset.

Table 2. Comparison results with algorithm [10]

Sequences	Algorithms	QP	Resolution	SSIM	Capacity (Kbits)	BIR	BIR-NC (%/Kbits)
Traffic	**Proposed**	40	2560×1600	**0.9577**	**155.037**	0.0519	**0.0335**
	[**10**]	40	2560×1600	0.9433	60.418	**0.0240**	0.0397
People-OnStreet	**Proposed**	40	2560×1600	**0.9318**	**254.517**	0.0291	**0.0114**
	[**10**]	40	2560×1600	0.9288	122.208	**0.0157**	0.0128
ParkScene	**Proposed**	40	1920×1080	**0.9481**	**815.81**	0.0340	**0.0417**
	[**10**]	40	1920×1080	0.9303	310.06	**0.0168**	0.0542
Basket-ballDrive	**Proposed**	40	1920×1080	**0.9431**	**215.051**	0.0702	**0.0326**
	[**10**]	40	1920×1080	0.9394	732.22	**0.0243**	0.0332
Johnny	**Proposed**	40	1280×720	**0.9784**	**105.433**	0.0911	**0.0864**
	[**10**]	40	1280×720	0.9746	440.05	**0.0396**	0.0899
FourPeople	**Proposed**	40	1280×720	**0.9547**	**173.580**	0.0758	**0.0437**
	[**10**]	40	1280×720	0.9363	798.65	**0.0396**	0.0496

Fig. 3. (a) Content of *Fourpeople* in 720P. (b) Content of *Johnny* in 720P. *Johnny* has more homogeneous areas than *Fourpeople*. Texture-rich video always has a lower BIR-NC.

From Table 2, it shows that our algorithm outperforms the algorithm [10] in capacity, SSIM, but has higher BIR. However, considering capacity and BIR at the same time, our algorithm can achieve smaller BIR at the same capacity, as shown by BIR-NC. For videos in all resolutions, our algorithm has better BIR-NCs, which are 15.6%, 10.9%, 23.1%, 1.8%, 3.9% and 11.9% smaller than algorithm[10]. The average BIR-NCs of the proposed algorithm are 86.7% of the algorithm [10] on 2K videos, 87.6% on 1080P videos and 92.1% on 720P video. It can be also observed that BIR-NC can differ even when the resolutions of video are the same, such as 0.0035 for 2K video *Traffic* or 0.0014 for *PeopleonStreet*. Several tools were used to analyze the difference among these videos. It shows that TBs with complex texture often has a higher residual signal. This kind of TBs can tolerate more bit changes than other blocks. In addition, in a high-resolution video, larger sized PBs also often occurs in texture-rich area. As shown in Table 2 and Fig. 3, the conclusion is drawn that under the same resolution, videos with lower BIR-NC always has more complex texture. Thus, a texture-rich video is more suitable for our algorithm than plain video in terms of preserving coding efficiency.

As far as the embedding payload is concerned, the capacity of our algorithm is 240% larger than algorithm [10] in average. Even if algorithm[10] embed less payload, the SSIM still proves inferior to ours. The reason for this phenomenon may be that in our algorithm, all the block partitioning is exactly the same as in the original video, but in algorithm[10], only the position of 4×4 PBs is preserved. As shown in equation (5), different smoothing filter (usually stronger) may apply to other PBs, which may leads to the degradation in SSIM. This result also proves that utilizing large-sized PBs will not cause severe visual degradation. The effectiveness of utilizing multi-sized PBs is well proven in Table 2.

Next, Wang[9]'s work [9] is compared with our algorithm. Algorithm [9] designs a mapping rule between difference of intra directions and secret message, but do not consider the distortion of HEVC intra prediction mode. Five videos named as video1, video2, video3, video4 and video 5, which are originally used in [9], are tested in this section. The experiment setup is the same as it in their work [9]. Comparison results in capacity, BIR and difference in PSNR is shown in Table 3. PSNR and this dataset are used here solely since they are originally used in [9] to illustrate the performance.

Table 3. Comparison results with algorithm [9]

Sequences	Algorithms	QP	Resolution	ΔPSNR(dB)	Capacity (Kbits)	BIR	BIR-NC (%/Kbits)
Video1	**Proposed**	22	832×480	**-0.02**	**357.688**	0.0354	**0.0099**
	[**9**]	22	832×480	-0.06	161.82	**0.0158**	0.098
Video2	**Proposed**	22	832×480	**-0.02**	**313.863**	0.0255	**0.0081**
	[**9**]	22	832×480	-0.06	110.70	**0.0190**	0.017
Video3	**Proposed**	22	1280×720	**-0.02**	**331.941**	0.0339	**0.010**
	[**9**]	22	1280×720	-0.04	9.534	**0.0076**	0.080
Video4	**Proposed**	22	1280×720	**-0.01**	**322.784**	0.0255	**0.0079**
	[**9**]	22	1280×720	-0.01	8.058	**0.0045**	0.056
Video5	**Proposed**	22	1280×720	**-0.01**	**310.289**	0.0244	**0.0079**
	[**9**]	22	1280×720	-0.03	5.502	**0.0046**	0.084

In Table 3, it shows that even if BIRs of algorithm [9] are lower, our algorithm has better BIR-NCs, which are 89.9%, 52.4%, 87.5%, 85.9%, and 90.6% smaller than algorithm[9]. The average BIR-NCs of the proposed algorithm are 18.7% of the algorithm [9] on their dataset. The main reason is that the proposed algorithm has great advantage of capacity compared to [9]. The average capacity of the proposed algorithm is 327313, which is about 32 times larger than the average capacity of algorithm[9],10069. In theory, algorithm[9] only use a few 4×4 PBs to embed the secret message while our algorithm using all size of PBs. Lower BIR-NC of the proposed algorithm standing for better preserving coding efficiency than algorithm[9].

To summarize, our algorithm outperforms the existing HEVC intra prediction mode algorithm with higher capacity and better coding efficiency. A texture-rich high-resolution video is more suitable for our algorithm in terms of preserving coding efficiency. In next sections, performance of the proposed algorithm will be further discussed by analyzing the influence of different QPs.

4.3 Analysis on Visual Quality with Different QPs.

In the section 4.3 and section 4.4, two algorithms will be performed to prove that utilizing multi-sized PBs for steganography will not cause significant visual degradation. The first one is the proposed algorithm with multi-sized PBs, another is the algorithm exactly the same as the proposed one except for using single-sized PBs. Experiments are conducted on different QPs to discuss the influence of QP setting.

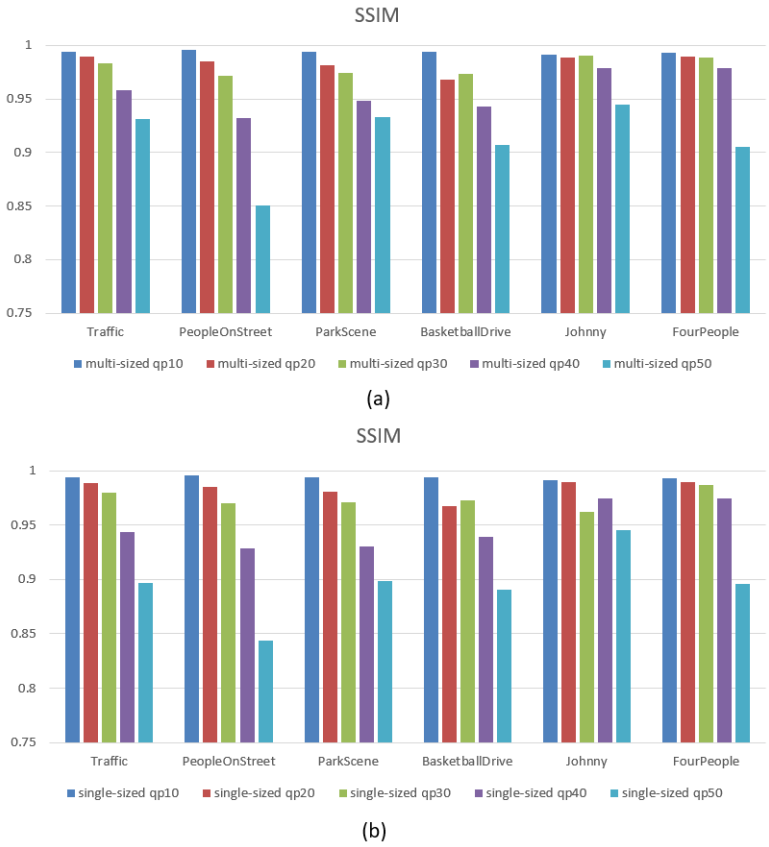

Fig. 4. SSIM on different videos with different QPs. (a) algorithm using multi-sized PBs (b) algorithm using single-sized PBs

As explained in Section 2, video quality of stego video will not dramatically degrade when larger size PBs are used. To prove this, SSIMs between original video and modified video are used to demonstrate the perceptibility and visual quality of the proposed HEVC algorithm. PSNR is not used here because SSIM can present visual quality better. The results are shown in Fig. 4.

It is shown that SSIMs between algorithm using multi-sized PBs and algorithm using single-sized PBs are similar under the same QP. The average SSIM of the proposed algorithm is 1.006 time higher than the algorithm using single-sized PBs. The decreasing of SSIM is mainly caused by the increment of QPs, not by the different sizes of PBs. The overall SSIM value ranges from 0.85 to 0.99 and decreases with the increase of QP. It shows that under the same resolution and QP, videos with more homogeneous areas have higher SSIM than texture-rich videos, such as *Johnny* and *Fourpeople*. The reason may be that although texture-rich areas can tolerate more bit changes, they also bring more pixel changes. Thus, it is totally safe to utilize larger-sized PBs for embedding the secret payload in terms of perceptibility.

4.4 Analysis on Capacity and Coding Efficiency with Different QPs

Table 4, Table 5 and Fig. 5 demonstrate the influence on coding efficiency and capacity of proposed algorithm and the one using single sized PBs. BIR-NC is used here to illustrate the bit changes when embedding with the same capacity with different QPs. Several phenomenon have been discussed in this section.

Table 4. Capacity of algorithm using multi-sized PBs.

Sequences\QP	**10**	**20**	**30**	**40**	**50**
Traffic	975138	557316	329568	155037	40681
PeopleOnStreet	809190	499548	377428	254517	89310
ParkScene	742010	419565	240369	81581	22990
BasketballDrive	690645	522367	445689	215051	81961
Johnny	682386	251338	185027	105433	43872
FourPeople	804557	459767	328896	173580	44538

Table 5. Capacity of algorithm using single-sized PBs.

Sequences\QP	**10**	**20**	**30**	**40**	**50**
Traffic	839925	439749	219407	60418	718
PeopleOnStreet	676915	396140	259114	122208	6646
ParkScene	652061	358280	171898	31006	961
BasketballDrive	278658	145110	287885	73222	5035
Johnny	576710	197728	119581	44005	2298
FourPeople	686454	371155	229741	79865	2371

From Table 4 and Table 5, the average capacity of the proposed algorithm is 783987 for QP equals to 10, 451650 for QP equals to 20, 317829 for QP equals to 30, 164199 for QP equals to 40 and 53892 for QP equals to 50. In average, capacity of the proposed algorithm are 4.9 times larger than algorithm using single-sized PBs. It can be observed that the capacity of algorithms decreases logarithmically with the increment of QPs. When QP increases, more reconstruction pixels will be calculated with the same quantized value, which leads to a smaller RD cost for large partition mode. The sizes of PBs are exponential {4, 8, 16, and 32}, meaning the large-sized PBs exponentially merge the small-sized PBs during this process. This may be the reason for logarithmically decrement in capacity. This also causes another phenomenon--the number of larger size PBs increases with QPs, which leads to a significant degradation in capacity for algorithm using single-sized PBs. For example, the capacity of the *Traffic* video using multi-sized PBs with QP50 is 40681, while reduced to 718 when using single-sized PBs. The bigger the QP, the bigger the advantage of the proposed algorithm using multi-sized PBs.

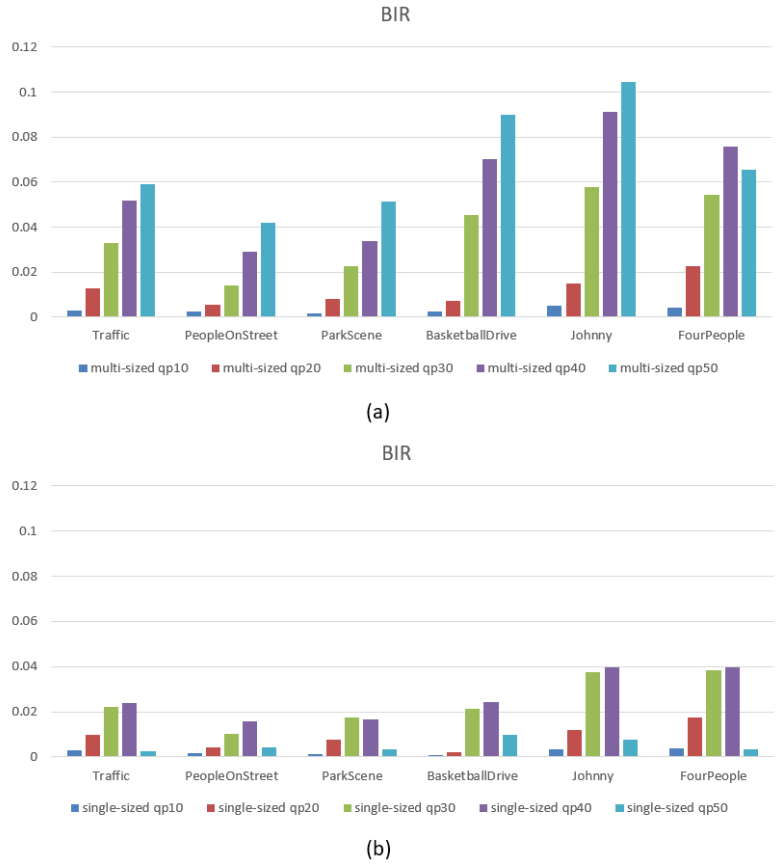

Fig. 5. BIR on different videos with different QPs. (a) algorithm using multi-sized PBs (b) algorithm using single-sized PBs

Fig. 5 shows that algorithm using multi-sized PBs always has higher BIR than algorithm using single-sized PBs. The average BIRs of the proposed algorithm are 0.0032 for QP equals to 10, 0.012 for QP equals to 20, 0.038 for QP equals to 30, 0.059 for QP equals to 40 and 0.069 for QP equals to 50. In average, the BIR of algorithm using multi-sized PBs is 3.9 times larger than algorithm using single-sized PBs. Thus, combining the results on capacity (4.9 times larger), the proposed algorithm still has the advantage when the payload is the same as is shown by BIR-NC in Fig.6. For example, the BIR of the *Traffic* video using multi-sized PB with QP50 is 0.059207, and it reduces to 0.00257 when using single-sized PBs. Nevertheless, the BIR-NC is 0.145 and 0.358 correspondingly. Furthermore, with the increase of QPs, the proposed algorithm achieves better BIR-NC. The averaging BIR-NC of the proposed algorithm when QP equals to 40 is 89.3% of the algorithm using single-sized PBs, but reduces to 63.7% when QP equals to 50. The reason is that with QP increases, more TBs with homogeneous areas appear. This kind of TBs can only tolerate less bit changes than other blocks. Thus, under the same videos, lower BIR-NC can be

achieve by lowering QP. A small QP video is more suitable for our algorithm in terms of preserving coding efficiency.

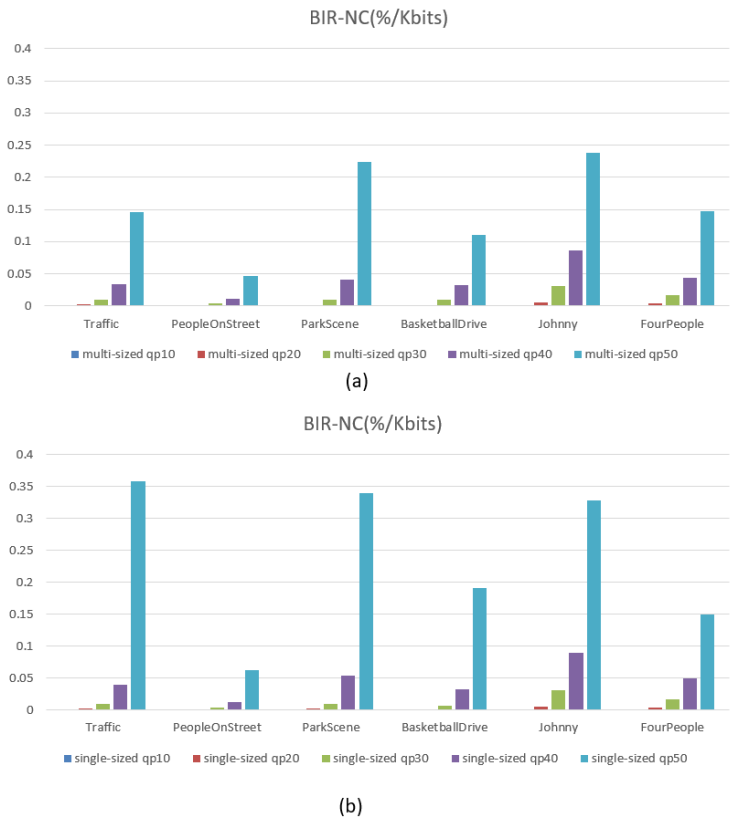

Fig. 6. BIR-NC on different videos with different QPs. (a) algorithm using multi-sized PBs (b) algorithm using single-sized PBs

To summarize, our algorithm outperforms the existing HEVC intra prediction mode algorithm with higher capacity and better perceptibility. Results prove that utilizing multi-sized PBs for steganography will not cause severe degradation in visual quality, while providing great improvement in capacity. A texture-rich high-resolution video is more suitable for our algorithm in terms of preserving coding efficiency. Some interesting phenomenon in the experiment results is discussed, which may be helpful to further improve the performance of the proposed steganography algorithm.

5 Conclusion

The existing video steganographic schemes based on intra prediction modes for H.264/AVC and HEVC all use single-sized blocks to embed the secret payload. In

this paper, a novel video steganography is presented. The innovation of this paper includes: 1) Improvement in capacity without introducing great degradation in visual quality. 2) High coding efficiency maintained by defining cost function based on rate distortion. Detailed experiments have been conducted to prove the effeteness of the proposed algorithm. Our algorithm outperform the latest HEVC intra prediction mode based steganography. The conclusion is drawn that large-sized PBs can be modified without introducing significant visual degradation, and a texture-rich high-resolution video is preferred for our algorithm. Future work can be made in security improvement or adopt the algorithm to an adaptive algorithm.

Acknowledgement. This work is supported by the National Natural Science Foundation of China (No. 61572320, 61572321). Corresponding author is Professor Xinghao Jiang, any comments should be addressed to xhjiang@sjtu.edu.cn.

References

1. Aparna R, Ajish S. A Review on Data Hiding Techniques in Compressed Video[J]. International Journal of Computer Applications, 2016, 134(13): 1-4.
2. Sullivan G J, Ohm J R, Han W J, T Wiegand. Overview of the High Efficiency Video Coding (HEVC) Standard[J]. IEEE Transactions on Circuits & Systems for Video Technology, 2013, 22(12):1649-1668.
3. Chang P C, Chung K L, Chen J, CH Lin, TJ Lin. A DCT/DST-based error propagation-free data hiding algorithm for HEVC intra-coded frames[J]. Journal of Visual Communication and Image Representation, 2014, 25(2): 239-253.
4. Tew Y, Wong K S. Information hiding in HEVC standard using adaptive coding block size decision[C]// IEEE International Conference on Image Processing. IEEE, 2015:5502-5506.
5. Hu Y, Zhang C, Su Y. Information hiding based on intra prediction modes for H. 264/AVC[C]//Multimedia and Expo, 2007 IEEE International Conference on IEEE, 2007: 1231-1234.
6. Yang, G., Li, J., He, Y.: An information hiding algorithm based on intra-prediction modes and matrix coding for H.264/AVC Video stream. International Journal of Electronics and Communications 65(4), 331-337 (2011)
7. Bouchama S, Hamami L, Aliane H. H. 264/AVC data hiding based on intra prediction modes for real-time applications[C]//Proceedings of the World Congress on Engineering and Computer Science. 2012, 2200(1): 655-658.
8. Zhang L, Zhao X. An Adaptive Video Steganography Based on Intra-prediction Mode and Cost Assignment[C]//International Workshop on Digital Watermarking. Springer, Cham, 2016: 518-532.
9. Wang J J, Rang-Ding Wang, Da-Wen Xu, Wei Li. An Information Hiding Algorithm for HEVC Based on Angle Differences of Intra Prediction Mode[M]// Cloud Computing and Security. Springer International Publishing, 2015:1578-1585.
10. Dong Y, Jiang X, Sun T, Xu D. Coding Efficiency Preserving Steganography Based on HEVC Steganographic Channel Model[C]// International Workshop on Digital Watermarking, Magdeburg, Germany, 2017:149-162.
11. Filler T, Judas J, Fridrich J. Minimizing additive distortion in steganography using syndrome-trellis codes[J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3): 920-935.