

Received January 25, 2021, accepted February 8, 2021, date of publication February 16, 2021, date of current version April 16, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3059654

A Video Steganography Method Based on Transform Block Decision for H.265/HEVC

HONGGUO ZHAO^{1,2}, YUNXIA LIU¹, YONGHAO WANG², SI LIU¹, AND CONG FENG¹

¹College of Information Science and Technology, Zhengzhou Normal University, Zhengzhou 450044, China

²School of Computing and Digital Technology, Birmingham City University, Birmingham B5 5JU, U.K.

Corresponding author: Yunxia Liu (liuyunxia0110@hust.edu.cn)

This work was supported by the National Natural Science Foundation of China (NSFC) under Grant 61572447.

ABSTRACT High definition video application has drawn a lot of interest both from academy and industry. The relevant latest video coding technology, H.265/HEVC has been a promising area for video steganography. In this paper, we present a novel and efficient video steganography method based on transform block decision for H.265. In order to improve the visual quality of carrier video, we analyze the embedding error of data hiding with modifying partitioning parameters of CB, PB and TB, and modify the transform block decision to embed secret message and update corresponding residuals synchronously. In order to limit embedding error, we utilize an efficient embedding mapping rule which can embed N ($N > 1$) bits message and at most modify one bit transform partitioning flag. Our experimental results show that the proposed method can achieve better visual quality, larger embedding capacity and less bit-rate increase than state-of-the-art researches.

INDEX TERMS H.265/HEVC, transform block decision, transform partitioning flag, video steganography.

I. INTRODUCTION

THE latest video coding standard H.265/high efficiency video coding (HEVC), published by the ITU-T Video Coding Experts Group (VCEG) and the ISO/IEC Moving Picture Experts Group (MPEG), has drawn a lot of research attractions due to its high compression performance, increased video resolution and abundant application scenarios. As declared in [1]–[3], H.265/HEVC standardization could improve compression performance with 50% bit-rate reduction for equal perceptual video quality when compared to the preceding standard H.264/AVC. Along with increasing diversity of H.265/HEVC based applications, the need for video security is also becoming stronger, especially for high definition (HD) and beyond-HD ($4k \times 2k$, $8k \times 4k$, etc.) format video [1]. Video steganography technology provides an efficient solution to protect video copyright, video traceability, even video based covert communication, etc.

Video steganography is a technique that utilizes human visual redundancy on digital signals to embed secret message [4]–[6]. It might be also presented as watermarking or steganography derived from different applications [6]. Watermarking focuses on copyright protection related areas,

such as declaration of owner rights, authentication of media contents, trace back of illegally spread, etc. However, steganography puts more emphasis on covert communication due to its large embedding capacity. There are several attributes that determine whether video steganography research performs successfully or not, including perceptual video quality (imperceptibility) and embedding capacity (payload). Based on these attributes, video steganography methods can be mainly divided into three categories derived from characteristics of coding structure. They are discrete cosine/sine transform (DCT/DST) coefficients (residuals) [7]–[12], prediction modes [13], [14], [21]/motion vectors [15]–[18], entropy coding [19], [20], [27] and coding block structure decision [21]–[24], etc. Practically in the earlier researches, there are several methods that embed secret message prior to video coding [25]. However, these methods make the embedded secret message easy to be lost after compressing, especially when video carrier is transmitted on the limited network [6]. Thus, combining video steganography and characteristics of coding structure are more favored in the researches of more efficient video steganography methods.

DCT/DST coefficients are active embedding area for video steganography due to its majority in the compressed video bitstream. This category method usually utilizes alternating

The associate editor coordinating the review of this manuscript and approving it for publication was Gangyi Jiang.

current (AC) coefficients to embed secret message after integer transform and quantization. In order to improve perceptual video quality, other coding characteristics are integrated with DCT coefficients, such as diagonal directions [7], intrapicture (intra) prediction modes [8]–[10], histogram shift [11], etc. DCT/DST based video steganography methods usually chose non-zero coefficients to embed secret message [8]–[10], [26], since a full zero coefficient block will not be transmitted in the compressed bitstream. However, the majority of block residuals are full zero. This restricts embedding capacity of DCT/DST based video steganography methods in covert communication scenarios. Intra prediction modes have also been used for video steganography researches [13], [14]. This category method usually embeds secret message by modifying prediction modes. However, prediction modes only exist in intra prediction process, which cannot be applied to inter prediction (i.e. motion vectors). Motion vectors are the syntax elements after motion estimation process and have been embedding positions for many video steganography methods [15]–[18]. These methods mainly modify motion vectors or adjust motion vector search process. As the distortion introduced by video steganography will propagate to subsequent prediction blocks with the motion estimation and compensation process, the embedding error will accumulate, which makes motion vector based video steganography methods generate large embedding error. Entropy coding based video steganography methods [19], [20] mainly embedding secret message by modifying syntax elements of CAVLC or CABAC. However, these methods might not achieve high perceptual video quality and even lead to decoding failure in some cases [6]. Coding block structure decision based video steganography methods are promising research areas. Especially when the traditional macroblock in H.264/AVC is replaced by adaptive coding tree block (CTB) structure in H.265/HEVC, the partitioning of CTB provides sufficient space for embedding secret message.

Actually there have been several related works on coding block structure decision. The main idea of these methods is that the coding block structure is classified or directly modified, according to pre-defined map rules between secret message and characteristics of block partitioning. Yang *et al.* [21] utilizes block types and modes of intra-coded blocks to embed secret message for H.264/AVC. In this method, the intra-4 × 4 coded blocks (I4-blocks) are divided into two groups, then the modes of I4-block are modified according to the map between watermark and intra-prediction modes. In this work, I4-block structures have been chosen as thresholds for the selection of embedding position, however, modification is only manipulated on intra-prediction modes. Tew and Wong [22] utilizes the partitioning of prediction blocks (PB) in H.265/HEVC to embed secret message. The partitioning of PBs are divided into two groups: $2N \times 2N$, $N \times 2N$, $nR \times 2N$, $2N \times nD$ and $N \times N$, $2N \times N$, $nL \times 2N$, $2N \times nU$, which are mapping to secret message 1 and 0, respectively. Shanableh [23] proposed a method that embeds secret

message by modifying splitting decisions of 32×32 and 16×16 coding blocks (CB). First, a model weights between split decisions of coding block and its feature variables (e.g. mean and variance of motion vector, depth of CB) are established, then the secret message is embedded with a function between the predicted (computed by model weights) and true split decisions per CB. When the embedded message is '1', the true split decision is constrained to be identical to predicted split decision. In addition, when the embedded message is '0', the true split decision is constrained to be different from the predicted one. In addition, Shanableh [24] proposed a more effective method which represents the partitioning of 16×16 sub CBs as a sequence of binary flags. Then, 6 or 4 secret message bits can be embedded with maximum 2 partitions modified in single CB. These previous researches mainly utilize the CB or PB partitioning structure to embed secret message, which performed effectively in embedding capacity and perceptual video quality, and the results reveal that video steganography based on coding block structure could be a promising area in H.265/HEVC.

In the former block structure based researches [21]–[24], secret message is always embedded with the modification on partitioning decisions of CB or PB. However, as CB may be directly or recursively inherited from the partitioning of CTB, and both of PB and transform block (TB) have their root at CB level. Modification on CB may introduce extra embedding error on subsequent PBs and TBs. Another challenge is that partitioning of PBs only manipulates once from its root CB level, modification or restrictions imposed on PB may affect the subsequent prediction process. Moreover, the residuals which will be processed later in transformation and quantization process are also affected, which has also been proven in [22]. Based on above consideration, in this paper, we chose TB (the minimum processing unit in the processes of intra/inter prediction, transformation and quantization) as the embedding carriers, which is targeted at less embedding error for carrier video, sufficient embedding capacity and less increment for bitstreams. The contribution and novelty about this paper are highlighted as follows:

(i) *Adaptive Modification on TB Partitioning*: With the goal of minimizing the embedding error and improve the perceptual video quality, this paper utilizes TB as embedding carrier, which has guaranteed that the embedding error is only restricted to TB region and other external process and syntax elements (e.g., prediction modes choosing, motion estimation, partitioning of CTB and PB partitioning structure) are not affected. Additionally, the subsequent residual coefficients are updated synchronously with new TB partitioning structure and a fast traversal algorithm is provided for the effective implementation of video steganography on TB block decision.

(ii) *A Hybrid Effective Embedding Scheme*: With the goal of minimizing the modification on TB partitioning along with embedding maximum secret message, a hybrid effective embedding scheme is proposed in this paper. The main advantage is that in most cases, multiple ($N > 1$) secret message

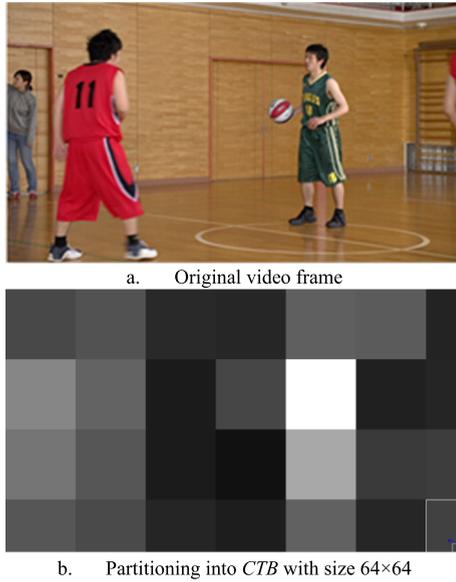


FIGURE 1. The partitioning structure of CTB in BasketballPass.

could be embedded while the partitioning of TB is modified at most once. Additionally, the embedding scheme could be alternative with different needs between embedding capacity and perceptual video quality.

The remainder of this paper is organized as follows. Section II introduces relative key characteristics of HEVC coding block structure decision. Then embedding error analysis among CB, PB and TB, the proposed method of modification on TB partitioning and hybrid embedding scheme are provided in section III. The experimental results are presented in section IV and conclusion is provided in Section V.

II. PARTITIONING STRUCTURE IN H.265/HEVC

H.265/HEVC employs coding tree unit (CTU: an integration of CTB and coding syntax) as the basic processing unit instead of fixed size macroblock in prior video coding standard. One advantage is that partitioning picture into larger CTB (more than 16×16) is beneficial in compression efficiency when encoding high-resolution video [1]. The CTB size is specified by the sequence parameter set (SPS), with optional size of 16, 32 or 64, and 64×64 is mostly used as default size. Fig. 1 gives an example of partitioning one picture into 64×64 size blocks (*BasketballPass* 416×240 resolution, partitioned into 4×7 dimension CTB and every CTB covers 64×64 square region with raster-scan order as shown in Fig. 1 except picture boundary).

Partitioning decision for CB is initiated from CTB. The square region of CTB can be directly as one CB or further partitioned into multiple CBs with the loop of quadtree iteration. The quadtree partitioning will be iterated until the size of CB has reached the allowed minimum CB size specified in the SPS. In most case, minimum CB size is restricted to 8×8 in luminance (luma) component. As a result, the final decision for CB partitioning is chosen from all possible iteration partitioning and the most optimized partitioning mode

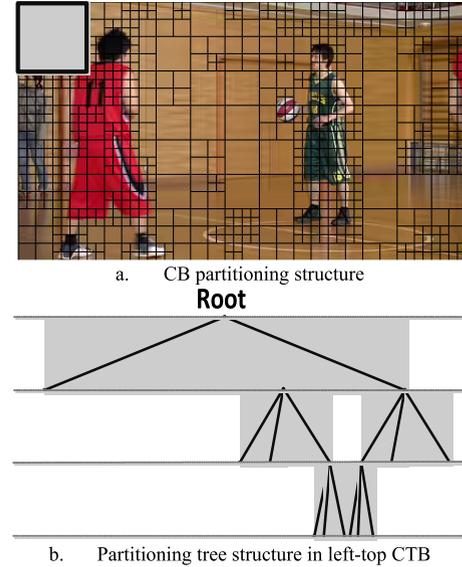


FIGURE 2. The partitioning structure of CBs in BasketballPass.

is determined by calculating and comparing the cost of distortion (weighted by sum of squared differences-SSD) and estimated number of encoded bits. The calculating model includes subsequent prediction, transform and quantization processes in this CB region. The calculating model can be represented as following [3]:

$$\min J_{CB}, J_{CB} = \sum D(i, j) + \lambda \sum R(i, j) \quad (1)$$

where J_{CB} depicts the search process for the cost of distortion and number of bits with one partitioning structure of CBs rooted at CTB, $D(i, j)$, $R(i, j)$ depicts distortion and number of encoded bits when attempting to compress i_{th} CB with j_{th} partitioning structure. One typical example is shown in Fig. 2, where the tree structure has been quadtree-structure partitioned three times at most, and the minimum CB size is 8×8 derived from the top-left CTB.

PB is based on the partitioning of CB, and can be divided into intra and inter prediction. Different prediction will generate different partitioning structure on PB. Generally speaking, in intra prediction, PB size is identical to CB size except when CB size meets the minimum value, where PB will has its root at CB and might be partitioned into four quartering sub PBs. In inter prediction, there are eight types of partitioning structure. It is noted that PB size of 4×4 is not allowed in inter prediction process, and when CB is partitioned into two PBs with the asymmetric size of $N \times N/4$, $N \times 3N/4$ and $N/4 \times N$, $3N/4 \times N$, the CB size of N is restricted to be 16 or higher. Fig.3 depicts these partitioning structures including intra and inter prediction.

After prediction process, TB is the basic process unit for residual transformation. Likewise, TB has its root at CB and employs equivalent quadtree-like partitioning structure. That means a CB might be recursively split into quadrants or just one TB directly, which are determined by calculating the cost of distortion and estimated number of encoded

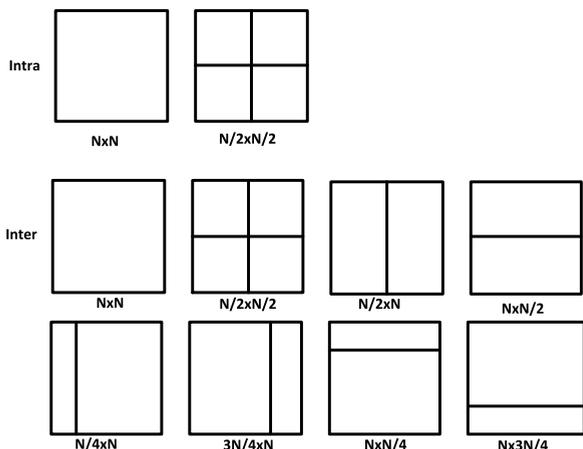


FIGURE 3. Partitioning structure of PB with intra and inter prediction.

bits with current TB partitioning structure. The calculating model is the same as CB in form and can be presented as follows:

$$\min J_{TB}, J_{TB} = \sum D(i, j) + \lambda \sum R(i, j) \quad (2)$$

where J_{TB} presents the search process for the minimum cost of distortion and estimated number of bits rooted at CB, and $D(i, j)$ indicates sum of SSD with i_{th} TB attempting j partitioning decision. Likewise, $R(i, j)$ indicates the estimated number of encoded bits of quantized residual coefficients, partitioning syntax elements, etc., which are used for the representing of this specified TB in the decoder or reconstruction process. It is noted that partitioning CB into TBs is also restricted by the maximum and minimum TB size specified in SPS. The enforcement is that when CB size is identical to the maximum, CB is enforced to be split into quadrants and when CB size reaches the minimum value, CB is enforced to be non-split. Moreover, if the prediction mode is intra and partitioning decision of PB is $N/2 \times N/2$, CB is also implicit to be split into quadrants. And in inter prediction mode, a TB might cover several PB regions for the maximum coding efficiency. Fig. 4 provides one TB partitioning structure marked as red lines in *BasketballPass* on the top-left CTB.

III. PROPOSED TRANSFORM DECISION BASED VIDEO STEGANOGRAPHY METHOD

In this section, we elaborate the scheme of transform block decision based video steganography method for H.265/HEVC. The architecture of the proposed scheme is depicted in Fig. 5. The proposed video steganography scheme consists of three components, including embedding, carrier bitstreams transmitted, and extraction components. First, prior to embedding, the partitioning structures of CB, PB and TB need to be obtained from coding decision of H.265/HEVC, where the partitioning structures are the results, which could achieve the minimum cost of distortion and number of bits with adopting current partitioning structures. And then, selective hybrid embedding modules

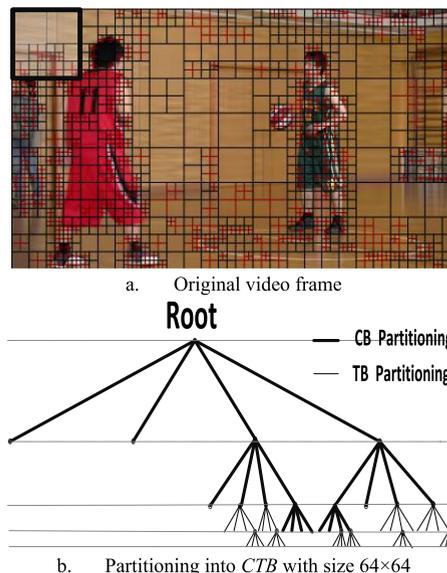


FIGURE 4. The partitioning structure of CTB in *BasketballPass*.

are established according to different applications. If we need large volume of secret message to be embedded, *single-embedding* module can meet this demand. If we need high security and perceptual video quality, *collaborative-embedding* module can be applied. Second, when the secret message has been embedded into carrier video, there are two alternative processes. One is ultimately encoded the carrier video to bitstream by entropy coding and enjoyed by the audiences via network transmission. Another is the reconstructed digital video which is constructed for the future prediction process of subsequent PB, with the similar process of the decoder. Finally, when the carrier video bitstream has reached to receiver, the partitioning structures of CB, PB and TB are obtained and analyzed for the search of targeted embedded TBs after entropy decoding. Selective extraction module is confirmed from two alternative extraction modules of *Single- and Collaborative-extraction*. As the result of extraction process, the embedded secret message depended on TB block decision is extracted.

A. PROPOSED HIDING DATA INTO TRANSFORM BLOCK DECISION

Transform Block Decision indicates embedding secret message into the partitioning structure of 8×8 TBs. The partitioning on TB is not only dependent on the search for the minimum cost of distortion and number of bits representing the residual samples in the decoder, but also associated with values of secret message which will be embedded. Weighting a tradeoff between the cost and secret message is one targeted goal in this work. To clearly depict the mechanism of the proposed method in maintaining high perceptual video quality and low bitstream increment for carrier videos, both of embedding error (distortion) analysis on TB block decision and key technical improvements of the proposed method are illustrated as follows.

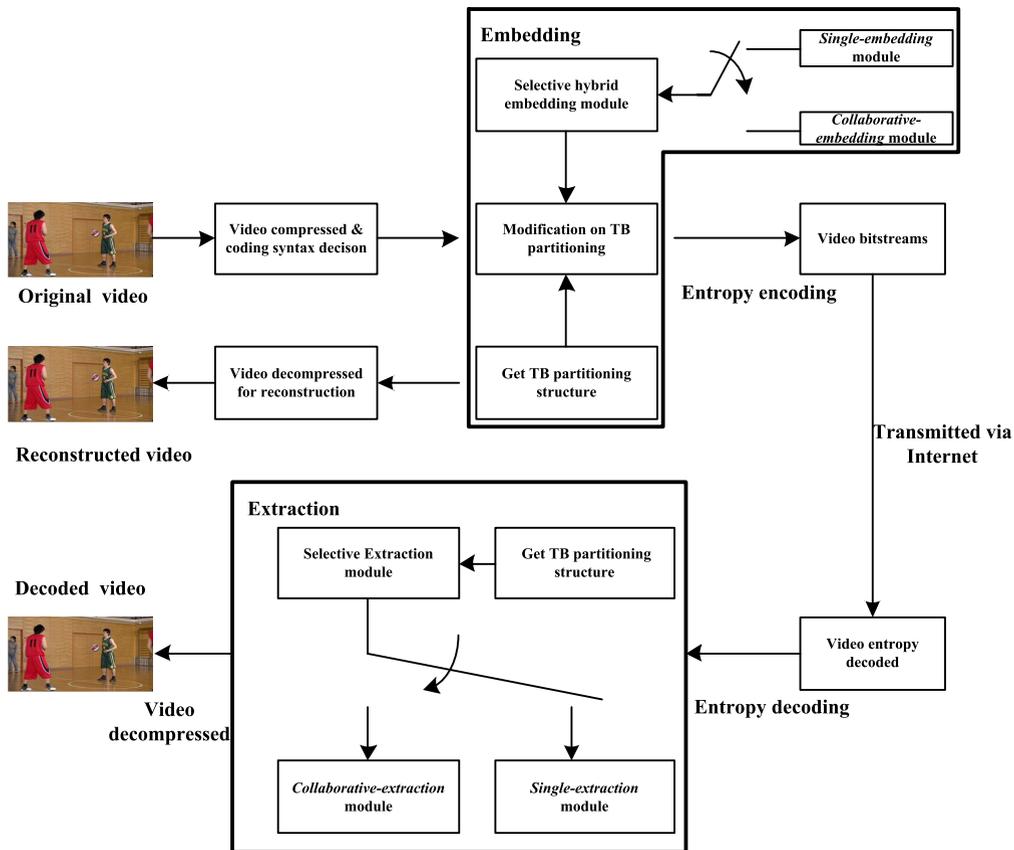


FIGURE 5. Architecture of the proposed transform block decision based data hiding method in H.265/HEVC.

1) EMBEDDING DISTORTION ANALYSIS ON TB BLOCK DECISION

Embedding distortion emerges during the process of embedding secret message into carrier videos, which can decrease the perceptual video quality and increase the volume of compressed bitstream compared to videos without embedding (original video). Embedding distortion on TB block decision would be accumulated with adoption of the modified transform partitioning structure. Embedding secret message on TB block decision might change the optimal transform partitioning structures, according to the binary sequence of secret message and mapping rules of embedding module. The transformation and quantization processes, would be processed depended on the new TB block size. However, compared to embedding secret message on CB or PB block decision, embedding on TB block makes less distortion because of smaller embedding regions in block size and less influenced processes. The smaller embedding regions indicate that embedding is only restricted in TB block which is the smallest process unit in H.265/HEVC encoder. As depicted in Fig. 6, the selected 16×16 CB is located at z-scan address 160 in the left-top CTB of *Basketballpass*. In the worst cases of embedding, embedding on CB and PB block decision will fill the whole 16×16 block as embedding regions marked as yellow regions. However, the embedding region of TB

will be restricted to 8×8 block region (blue regions) while embedding the same secret message. In addition, embedding on CB block will bring an impact to the subsequent processes, including CB partitioning, prediction, transformation and quantization in the whole CB region. Likewise, embedding on PB would also bring an impact to the prediction, transformation and quantization processes in the PB region. However, modification on TB block only affects the transformation and quantization processes, and the embedding distortion is only filled in the specified TB region.

In order to further decrease the embedding distortion and bitstream increment, the size of embedded TB is confined to be 8×8 , which is the minimum transform block that can be used to embed secret message. Since 8×8 TB can be further partitioned into four quadrants of 4×4 blocks or just one 8×8 block, it can adaptively map to the binary sequence of secret message. Fig. 7 provides the distortion performance of mean squared error (MSE) and bitstream increment ratio while encoding *Basketballpass* into 50 frames with group structure of *IPPP*. The average discrepancy of MSE is 0.166, and increment ratio of bitstream is 2.9%, which are referred to the comparison between H.265/HEVC compressed video and carrier video with the proposed method. The results also indicates that the distortion embedding on TB block decision is restricted to a small scope, and in most cases the distortion

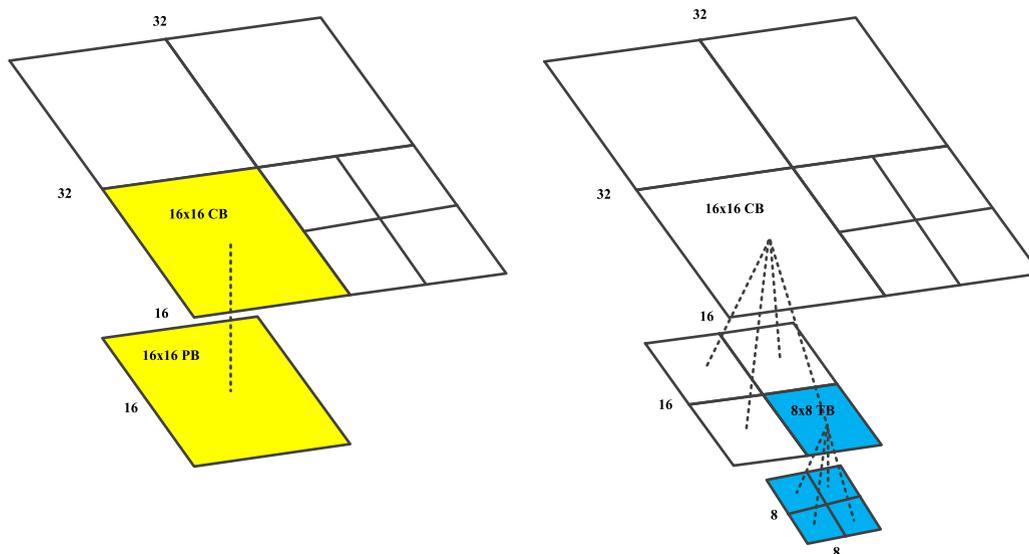


FIGURE 6. Embedding influenced regions among CB, PB and TBs.

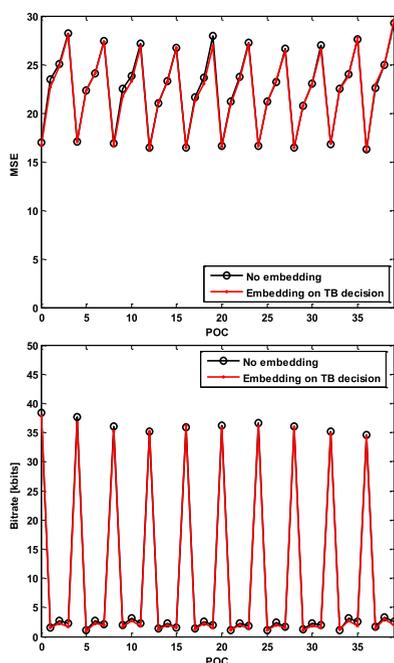


FIGURE 7. Distortion of MSE and bitrate on TB block decision.

introduced by embedding is approximately approaching to the distortion by standard H.265/HEVC compressing.

2) SELECTION CONSTRAINTS FOR TB BLOCK

The embedded TB block is confined to be 8×8 size block. There are two categories of 8×8 TB blocks that can be selected as the candidate blocks for embedding. One category is directly 8×8 size CB, where the subsequent partitioning structure of PB is restricted to only $N \times N$ mode in intra prediction, and the coding block flag(cbf) of three components: *cbf_luma*, *cbf_cb* and *cbf_cr* are restricted to be not all zero in inter prediction. When the 8×8 CB block

meets the $N/2 \times N/2$ partitioning of PBs in intra prediction, the subsequent TB blocks are enforced into quadrants, which cannot adjust the transform partitioning structure according to the embedding module. If all *cbfs* of three components are zero, the syntax elements of transform partitioning will not be entropy encoded into bitstream, which would make the secret message loss at the decoder. Another category is 8×8 or 4×4 size TBs, which are sub-regions of larger CB block ($> 8 \times 8$). That is, rooted at CB level, there exists 8×8 or 4×4 TB blocks determined by H.265/HEVC encoder. Then these 8×8 or 4×4 TB blocks can be selected as candidate embedded carriers. The constraint on this category block is that only in inter prediction, three components *cbfs* (including *cbf_luma*, *cbf_cb*, and *cbf_cr*) are confined to be not all zero, which is similar to the first category. The typical example referred to these two categories of candidate embedded TB blocks is shown in Fig. 8. It is noted that 8×8 CB located at Z-scan address 96 cannot be selected as candidate TB block due to its $N/2 \times N/2$ PB partitioning structure in intra prediction. In addition, 21 candidate embedded TB blocks in one CTB provide sufficient embedding capacity for embedding.

3) SYNCHRONIZATION OF RESIDUALS AND CBF

Since different transform block sizes employ different transformation matrixes [1], the corresponding residuals, which will be probably mismatched to the new transform block partitioning structure, and bring an extra distortion to the carrier videos. In order to eliminate the mismatch error and improve the visual quality, the prediction residuals, are recomputed by calling the transformation, scaling and quantization processes with new transform block partitioning. Moreover, as the QDCT/QDST residuals may be zero after re-transforming and quantizing in some worst cases, the corresponding transform partitioning with embedding secret message, will not be encoded into carrier bitstreams by

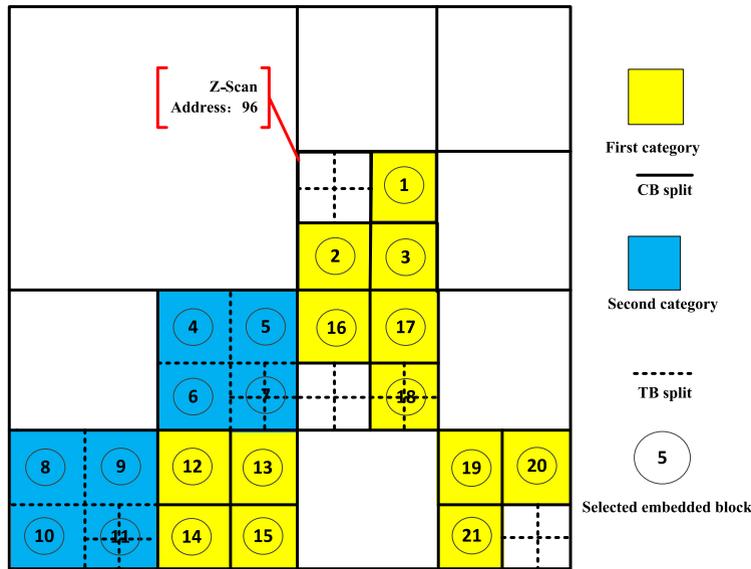


FIGURE 8. Typical example of two categories candidate embedded transform blocks.

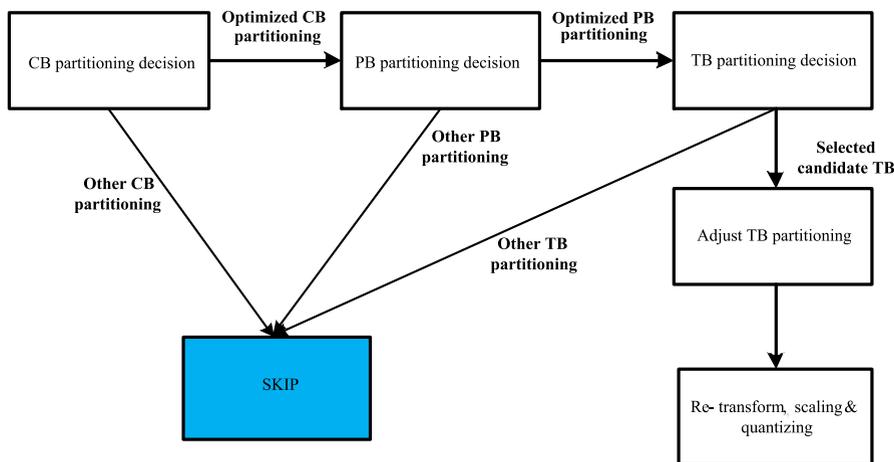


FIGURE 9. Proposed fast traversal algorithm for locating selected candidate TBs.

entropy encode, especially for inter prediction cases. Thus, each component of *cbf_luma*, *cbf_cb*, or *cbf_cr* should be constrained to be synchronous with the former *cbfs* determined by the encoder with H.265/HEVC.

4) FAST TRAVERSAL OF QUADTREE-STRUCTURED Cbs AND TBs

To guarantee the transform partitioning carriers are encoded into bitstream during multi-level iterations for the optimized partitioning structures for CB, PB and TB, we provide a new fast traversal algorithm for locating the most optimized transform partitioning structures. As depicted in Fig. 9, when the encoder begins the iterations for CB partitioning, only the most optimized CB partitioning needs to be traversed and other partitioning attempts for current CB coding are set to be skipped. During the most optimized CB partitioning decision

process, only the most optimized PB partitioning and prediction modes/motion vectors need to be traversed, and other PB partitioning or prediction decision are set to be skipped. After prediction process, only the most optimized TB partitioning decision needs to be continuously traversed and other TB partitioning attempts are set to be skipped. During the process of most optimized TB partitioning, the current most optimized TB partitioning needs to be adjusted according to the mapping rules as illustrated in section 3.2 for candidate embedded TB, and for other non-candidate TB, the current most optimized TB partitioning remains unchanged. For embedded candidate blocks, the residuals and *cbfs* also need to be re-transformed, scaling and quantizing with new adjusted TB partitioning structures.

The main advantage of this fast traversal algorithm is that only the most optimized Cbs, PBs and TBs partitioning loops

are fast traversed, which will greatly reduce the computing resources and time consuming for embedding secret message in the encoder.

B. PROPOSED HYBRID EMBEDDING & EXTRACTION SCHEME

The proposed hybrid embedding scheme embeds secret message into the partition structures of selected candidate TB. That is, after scanning the most optimized partitioning structure of CB and PB, the practical embedding is manipulated on the process of TB partitioning decision. According to the mapping rules of hybrid embedding scheme in this section, the secret message can be embedded with the adjustment of the most optimized transform partitioning of embedded candidate 8×8 TB. The main advantage of hybrid embedding scheme is that multiple bits of secret message can be embedded while at most one transformation partitioning of 8×8 TB needs to be changed. The proposed hybrid extraction scheme is an inverse process of embedding scheme.

1) PROPOSED HYBRID EMBEDDING SCHEME

The proposed hybrid embedding scheme mainly includes *single-* and *collaborative-embedding* modules. The embedding scheme is determined by the count number N of selected candidate 8×8 TBs in one CTB, where CTB is a basic process unit for the proposed embedding scheme. In addition, the embedding threshold Γ is also pre-defined to determine the embedding scheme type. If $N < \Gamma$, *single-embedding* module is selected. If $N \geq \Gamma$, *collaborative-embedding* module is selected. The value of pre-defined threshold Γ is determined by the tradeoff between the embedding capacity and visual quality. High embedding capacity can be achieved with large Γ , and high visual quality can be achieved with small Γ .

The *single-embedding* module is dependent on the most optimized transformation partitioning characteristic determined by the encoder, where each embedded candidate 8×8 TB embeds one bit secret message. The characteristic is a sequence that depicts whether each scheme candidate scheme TB is split into quadrants not. The mapping rule is that secret message bit '0' indicates that the scheme candidate scheme 8×8 TB is not partitioned. Otherwise, secret message bit '1' indicates that the 8×8 TB is enforced to be partitioned into four quadrants of 4×4 TBs.

The *collaborative-embedding* module is developed from matrix encoding, and needs the collaboration of all candidate embedded 8×8 TBs. The collaboration means that each transformation partitioning of candidate embedded TBs is a contributor to jointly embed secret message. In *collaborative-embedding* module, the characteristic sequence of candidate embedded 8×8 TB blocks $P_k = (P_1, P_2, \dots, P_N)$ is established, where each element $p \in P_k$ indicates whether the candidate embedded 8×8 TB block is further partitioned into four 4×4 TB blocks or not. That is, signal '1' presents further partitioning is employed and signal '0' presents further partitioning is discarded. Each signal in sequence P_k is organized

as z-scan order. Another sequence is the embedded secret message $S_k = (S_1, S_2, \dots, S_l)$, where l is the count number of secret message in one CTB, which is determined by the number of selected candidate 8×8 TBs N . The relationship between them is illustrated as

$$l = \text{floor}(\log_2(N + 1)) \quad (3)$$

after we have acquired sequences P_k and S_k , the final embedded position E_P is computed according to successive bitwise operations with each element in P_k and S_k , which is formulated as

$$E_P = \left(\bigoplus_{k=1}^N p \times k \right) \oplus \left(\sum_{k=1}^l s \times 2^{k-1} \right), \quad p \in P_k \text{ and } s \in S_k \quad (4)$$

where E_P indicates the position of 8×8 TB that is needed to be modified in the sequence P_k . It is noted that in the best case, the value of E_P will be zero, which indicates that the current partitioning structure of candidate 8×8 TB will be maintained as the same as the original partitioning structure determined by the encoder. However, in other cases, the transformation partitioning of candidate 8×8 TB indexed by E_P should be changed to the opposite partitioning structure. That is, if the partitioning characteristic indexed by E_P is 1, then this candidate 8×8 TB should be confined to be non-split into four 4×4 TB blocks. Otherwise, if the partitioning characteristic is 0, the corresponding 8×8 TB is confined to be split into quadrants.

The *single-* and *collaborative-embedding* modules are usually exclusive with each other in single CTB, which are determined by the pre-defined threshold Γ and the count number of selected candidate 8×8 TBs N . After the hybrid embedding scheme, the following QDCT/QDST residuals and CBF should be synchronized with the new transformation partitioning structures. In the end, the new embedded partitioning syntax referred to candidate 8×8 TBs will be encoded into bitstream by entropy encoding and the following CTB will be processed with similar embedding loop. The proposed hybrid embedding scheme within one CTB is depicted with algorithm 1 as follows.

2) PROPOSED HYBRID EXTRACTION SCHEME

The proposed hybrid extraction scheme is an inverse process compared to embedding, which is manipulated at decoder. The extraction module is also divided into two categories, *single-* and *collaborative-extraction*, which is determined by the pre-defined threshold Γ and the count number of extracted candidate 8×8 TBs N . In addition, the constraints on the selected candidate extraction 8×8 TB blocks are the same as the embedding, including *intra* or *inter* prediction cases. However, fast traversal process are canceled, because the decoder only reconstruct video samples according to the received bitstream, which is the same to the synchronization of QDST/QDCT residuals and CBF.

Algorithm 1 Proposed Hybrid Embedding module

```

Input: Transformation partitioning sequence of selected candidate  $8 \times 8$  TB blocks  $P_k$ 
         Sequence of embedded secret message  $S_k$ 
Output: Embedded transformation partitioning sequence of candidate  $8 \times 8$  TB blocks  $\hat{P}_k$ 
BEGIN
Confirm Pre-defined embedding threshold  $\Gamma$ 
Count the number of selected candidate  $8 \times 8$  TB blocks  $N$ 
IF  $N < \Gamma$  THEN
    Single-embedding module is chosen
    FOREACH  $P, s$  in the sequence  $P_k, S_k$  DO           //  $p \in P_k, s \in S_k$ 
        IF  $s == 1$  THEN
             $P- > 1$ 
            Enforce current candidate  $8 \times 8$  TB block partitioned into quadrants
            QDCT/QDST residuals and CBF are synchronized
        IF  $s == 0$  THEN
             $p- > 0$ 
            Enforce current candidate  $8 \times 8$  TB block non-partitioned
            QDCT/QDST residuals and CBF are synchronized
    ELSE
        Collaborative-embedding module is chosen
        Confirm the count number of embedded secret message  $l : l = \text{floor}(\log_2(N + 1))$ 
        FOREACH  $p$  in the sequence  $P_k$  DO
             $f_i = (p \times i) \oplus f_{i-1}$            //  $i_{th}$  iteration,  $f_0 = 1$ 
            Calculate the decimal value of Truncated Secret message  $S_T = \sum_{i=1}^l, s \times 2^{i-l} s \in (s_1, s_2, \dots, s_l)$ 
            Located embedded position  $E\_P = f_N \oplus S_T$ 
            FOREACH  $p$  in the sequence  $P_k$  DO
                IF  $p == E\_P$  THEN
                    Change current partitioning structure to opposite partitioning
                    QDCT/QDST residuals and CBF are synchronized
                ELSE
                    Maintain unchanged to the current partitioning syntax element
            CONTINUE
END

```

Fast traversal

Fast traversal

If the count number N is less than threshold Γ , the *signal-extraction* will be applied in the current CTB. Let $P'_k = (p'_1, p'_2, \dots, p'_N)$ be the partitioning characteristic sequence of candidate extracted 8×8 TBs, and $S'_k = (s'_1, s'_2, \dots, s'_N)$ be the extracted binary secret message sequence. Then the corresponding mapping rules referred to *signal-extraction* can be depicted as follows:

$$\begin{aligned} \text{if } p == 1 \rightarrow s' = 1 \\ \text{if } p == 0 \rightarrow s' = 0 \quad p \in P'_k, s' \in S'_k \end{aligned} \quad (5)$$

where $p \in P'_k$ presents each partitioning structure in the sequence of candidate 8×8 TBs.

If the count number N is more than threshold Γ , the *collaborative-extraction* module will be applied. Similar to the process of embedding, the length of truncated secret message sequence l should be confirmed according to the count number of extracted candidate 8×8 TBs with formula (3). Then the extracted secret message sequence can be initially created as $S'_k = (s'_1, s'_2, \dots, s'_l)$ According to the partitioning

characteristic sequence $P'_k = (p'_1, p'_2, \dots, p'_N)$, the decimal secret message S'_T can be formulated as

$$S'_T = \bigoplus_{k=1}^N p \times k, \quad p \in P'_k \quad (6)$$

where $p \in P'_k$ presents each partitioning structure in the sequence of candidate 8×8 TB blocks. In the end, the binary sequence S'_k can be filled by the decimal value of S'_T in current CTB. The proposed hybrid extraction module is depicted in algorithm 2.

C. PRACTICE FOR DATA HIDING METHOD BASED ON TRANSFORM BLOCK DECISION

In this section, one typical example is provided, which depicts a whole process of the proposed selection and constraints for candidate 8×8 TBs, residuals and *CBF*'s synchronization, fast traversal and hybrid embedding /extraction scheme.

The left-top CTB (raster-scan address 1) of the first frame in *Basketballpass* is selected, where the partitioning

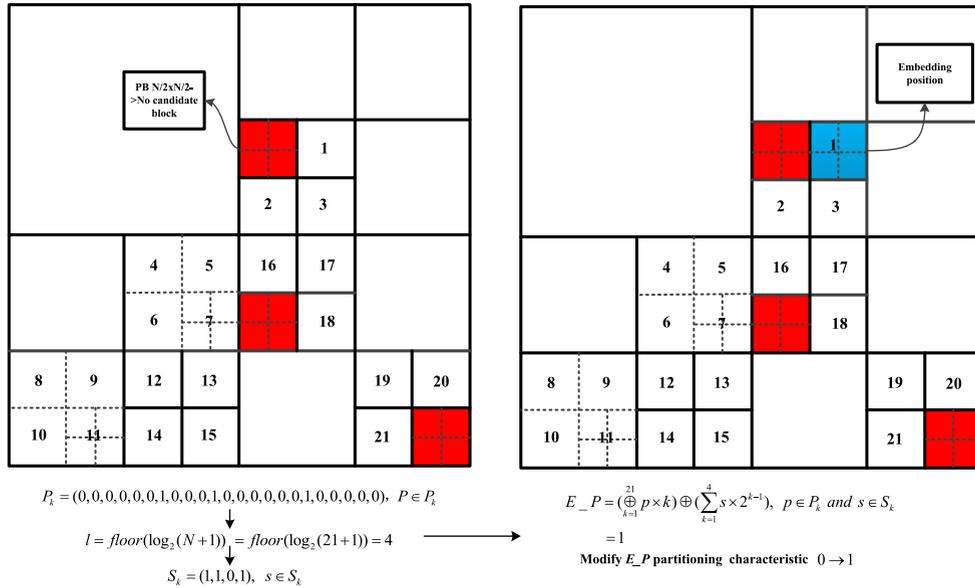


FIGURE 10. Embedding practice for the proposed method.

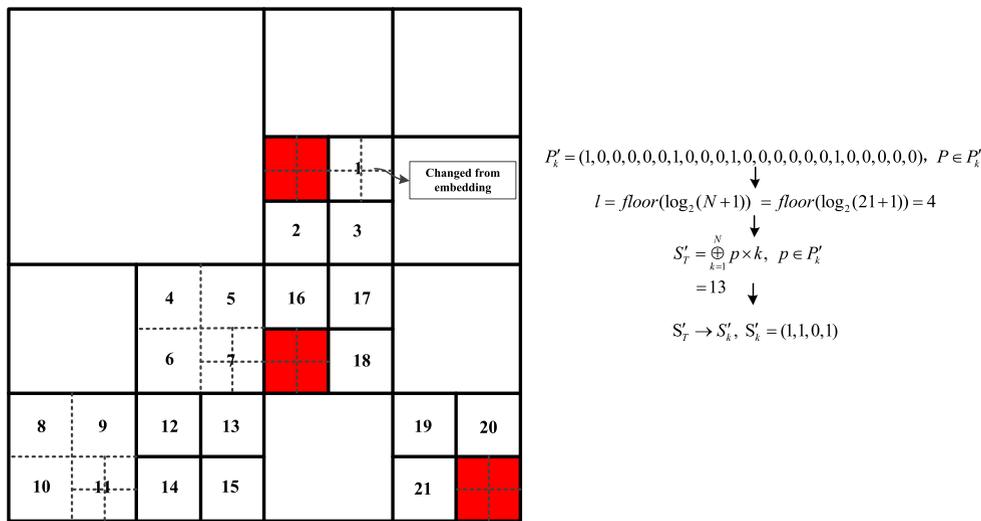


FIGURE 11. Extraction practice for the proposed method.

of picture (GOP) size and period of intra frames are both set to 4 in inter prediction cases, and the quantization parameter (QP) is set to 32, except for embedding capacity evaluation in a range of 32 to 50. A series of typical video sequences, in which the resolution is varied in the range of 416×240 to 1920×1080 , are used as tested samples and are encoded into 30 frames at 30 frames/s.

A. VISUAL QUALITY PERFORMANCE

In order to evaluate the visual quality performance, we use three GOP structures of all intra (IIII), P frames (IPPP), B frames (IBBB), where the pre-defined threshold Γ is set to be 2. Fig.12 depicts the visual quality performance of three video samples, *BasketballPass* (416×240), *KristenAndSara* (1280×720), *ParkScene* (1920×1080) with

different POC frames, where POC presents the display order. It can be seen from Fig. 12 that the proposed method can achieve a high visual quality performance, where the distortion introduced by embedding (represented as the PSNR discrepancy between H.265 and proposed method) are 0.02dB, 0.03dB and 0.01dB in *BasketballPass*, 0.48dB, 0.32dB and 0.33dB in *KristenAndSara*, 0.16dB, 0.03dB and 0.2dB in *ParkScene*, respectively. In some cases, the PSNR with the proposed method is a little higher than H.265/HEVC encoder. The reason is that the cost of transformation partitioning is determined by the distortion and bit-rate as declared in formula (2), higher visual quality always needs larger volume bitstream. Thus, as a result, the final cost of modification on TB with embedding is getting worse than without embedding.

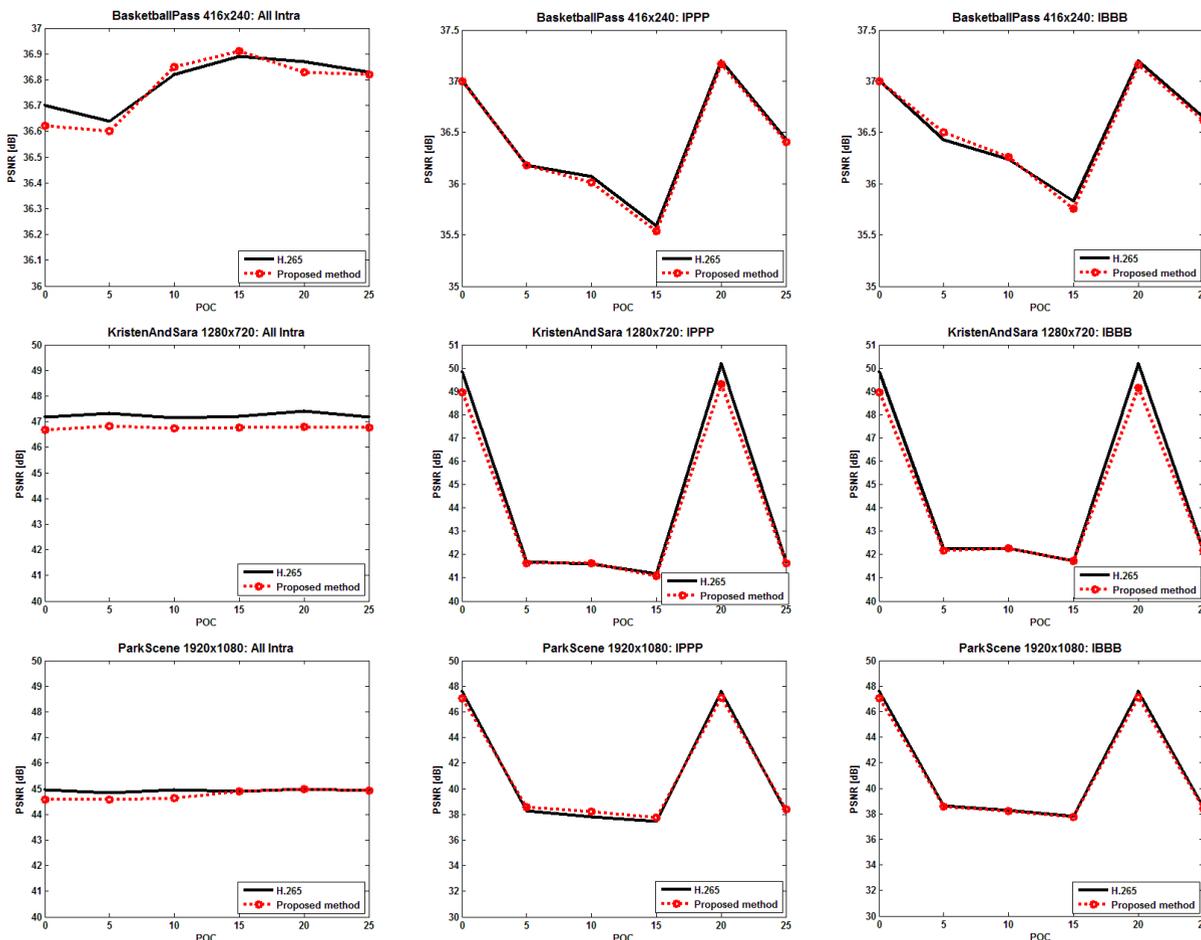


FIGURE 12. Visual quality performance of the proposed method with different GOP structures.

Subjective quality is also shown in Fig. 13, where original video samples, compressed samples with H.265 and carrier samples with embedding are all compared. In addition, different video sequences meet different POCs, where *BasketballPass*, *KristenAndSara* and *ParkScene* meet POC of 0, 5 and 10, respectively. It can be seen that the discrepancy on subjective quality is too slight to be observed, and the discrepancy of PSNR between compressed and carrier samples are 0.08dB, 0.32dB and 0.32dB, respectively.

The average visual quality performance of the proposed method (average value in 30 frames) is depicted in Table 1, where ΔMSE , $\Delta PSNR$ and $\Delta SSIM$ are computed by the video samples without embedding and carrier samples with embedding. It can be seen from Table 1 that the proposed method can achieve a high visual quality performance in three different GOP structures, where the maximums of ΔMSE , $\Delta PSNR$, $\Delta SSIM$ are 0.46, 0.48dB, 0.001, and the minimums are 0.07, 0.02dB, 0.0001, respectively.

B. EMBEDDING PERFORMANCE

We use different thresholds $\Gamma (2 \sim 64)$ to test the embedding performance of the proposed method in Table 2. In addition,

TABLE 1. Average visual quality performance of the proposed method.

Sequence		All Intra	IPPP	IBBB
<i>BasketballPass</i>	ΔMSE	0.11	0.16	0.1
	$\Delta PSNR$	0.03	0.04	0.02
	$\Delta SSIM$	0.0005	0.0001	0.0001
<i>BlowingBubbles</i>	ΔMSE	0.4	0.46	0.31
	$\Delta PSNR$	0.37	0.09	0.06
	$\Delta SSIM$	0.001	0.0006	0.0008
<i>KristenAndSara</i>	ΔMSE	0.14	0.07	0.07
	$\Delta PSNR$	0.48	0.08	0.08
	$\Delta SSIM$	0.0007	0.0001	0.0003
<i>ParkScene</i>	ΔMSE	0.18	0.2	0.19
	$\Delta PSNR$	0.37	0.08	0.08
	$\Delta SSIM$	0.001	0.0008	0.0009

we also provide a comparison of embedding performance in Table 3. In order to evaluate the influence of different QPs to embedding capacity in the proposed method, Fig. 14 provides the experimental results in the end.

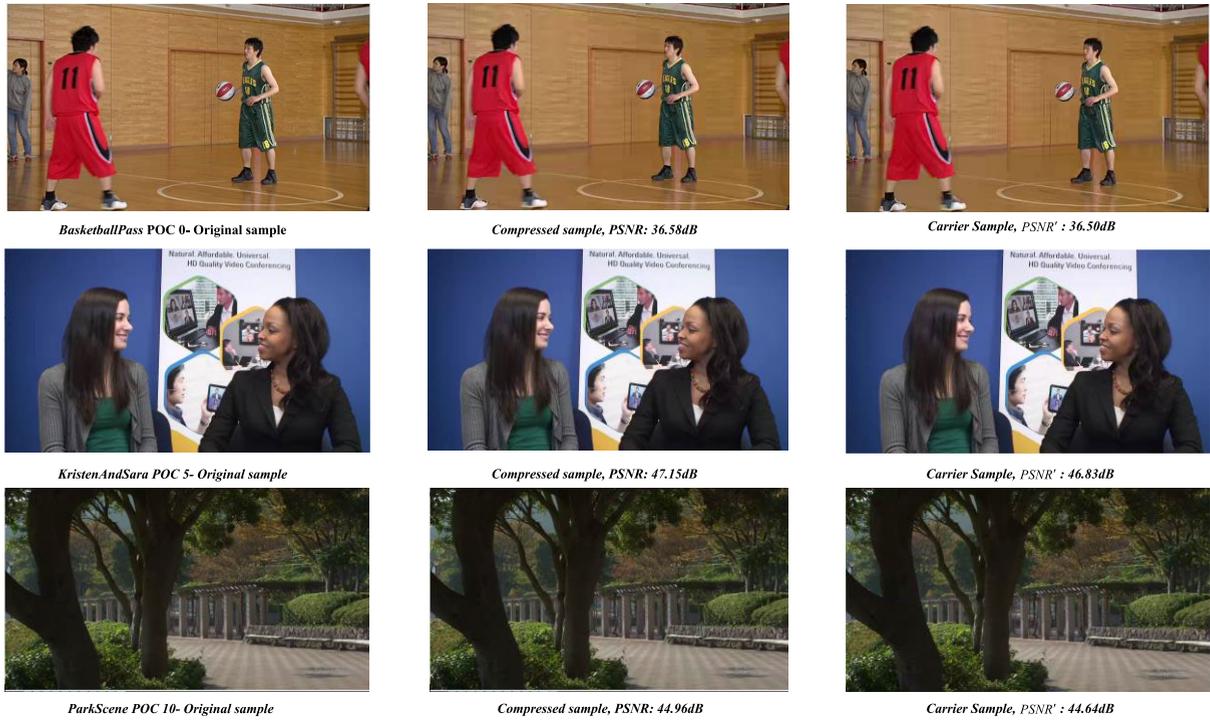


FIGURE 13. Subjective performance of the proposed method.

TABLE 2. Embedding performance of the proposed method.

Sequence	Threshold	$\Delta PSNR$ (dB)	Capacity(bits)	Bit-rate increase(%)
BasketballPass	2	0.03	2784	1.21
	4	0.03	2857	1.26
	8	0.03	3254	1.47
	16	0.06	4777	2.08
	32	0.13	10431	3.77
BlowingBubbles	2	0.37	3502	0.08
	4	0.35	3498	0.17
	8	0.34	3574	0.21
	16	0.54	4655	0.21
	32	2.05	15422	0.31
KristenAndSara	2	0.48	13311	0.33
	4	0.55	14557	0.44
	8	0.61	16616	0.58
	16	1.06	25477	0.89
	32	1.55	50048	1.88
ParkScene	2	2.41	63752	2.49
	4	0.37	45264	0.19
	8	0.4	46198	0.17
	16	0.46	52057	0.20
	32	0.79	77940	0.25
	64	1.92	189718	0.25
	64	2.31	241808	0.28

We use the average values of $\Delta PSNR$ for visual quality measurement, and bit-rate increase for bitstream increment of embedding. Table 2 provides the embedding performance of the proposed method when different thresholds are employed. The average discrepancy $\Delta PSNR$ is 0.82dB, embedding capacity is 38.55kbts/s, and the bit-rate increase is 0.98%. It can be discovered that better visual quality can be achieved with a smaller threshold, where collaborative-embedding

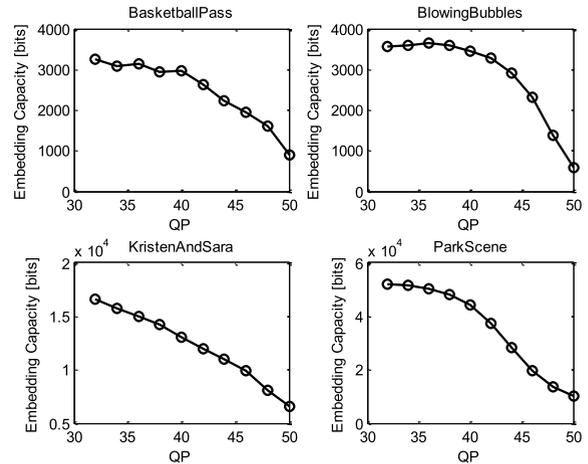


FIGURE 14. Embedding capacities of the proposed method in different QPs.

module is mostly employed. However, larger volume of embedding capacity can be achieved with a higher threshold, where single-embedding is more likely employed. In addition, the bit-rate increase is also increasing with the increment of threshold Γ . As a result, the value of threshold Γ can be adaptively updated according to different practical demands. For example, the scenario of stricter demands on high visual quality can adopt the minimum threshold 2, and the maximum threshold 64 can be applied to the scenarios of large volume secret message transmitted.

Table 3 describes the comparison of embedding performance among the proposed method and researches [4], [26].

TABLE 3. Comparison of embedding performance among the proposed method and [4], [26].

Sequence	Proposed method	[4]	[27]
BasketballPass	ΔPSNR(dB)	0.03	0.22
	Capacity(bits)	3254	3264
	Bit-rate increase(%)	1.47	1.56
BlowingBubbles	ΔPSNR(dB)	0.34	1.78
	Capacity(bit)	3574	8520
	Bit-rate increase(%)	0.21	2.43
KristenAndSara	ΔPSNR(dB)	0.61	0.99
	Capacity(bits)	16616	15000
	Bit-rate increase(%)	0.58	2.45
ParkScene	ΔPSNR(dB)	0.46	0.73
	Capacity(bits)	52057	30696
	Bit-rate increase(%)	0.2	1.3

During the experiment process, threshold Γ is set to be 8 in the proposed method. For the visual quality, the average $\Delta PSNR$ of proposed method are 0.36dB, However, the corresponding $\Delta PSNR$ of [4] and [26] are 1.07dB and 0.93dB. It can be seen that the proposed method can achieve a higher visual quality performance when compared to the researches [4] and [26]. It is noted that for video sequence *BlowingBubbles*, which is sensitive to embedding, the proposed method outperform better than [4] and [26]. For embedding capacity, the average embedding capacity of the proposed method is 18.875kbits/s, while [4] is 8.211kbits/s and [26] is 14.370kbits/s, respectively. Compared to [4] and [26], the proposed method can achieve a higher embedding capacity, and less distortion on visual quality at the same time. For bit-rate increase, the average bit-rate increase of the proposed method is 0.61%. However, the average bit-rate increase of [4] is 1.50%, and [26] reaches 1.94%. It can be seen that the proposed method can achieve the least bit-rate increase when compared to [4] and [26]. In addition, in order to further evaluate the embedding performance of the proposed method, a summary of comparative analysis of the proposed method, [4], [7] and [28]–[30] has been shown in Table 4. Compared to [4] and [7], The proposed method achieved a higher capacity, and due to the embedding location of the transform partitioning syntax, embedding scope of the proposed method could be applied to both of intra and inter process rather than only intra in [4] and [7]. Compared to [28] to [30], which emphasized on the robustness of the video steganography, the proposed method puts more attention on the high visual quality of the carrier videos, which enhance the imperceptibility (high PSNR approaching to original video) and capacity. Moreover, the proposed method is more compatible for compressed video bitstream distributed on the internet scene. The robustness consideration of the proposed method would be our future work as depicted in conclusion section.

Fig. 14 provides the experimental results about the embedding capacities of the proposed method with different large QPs ($30 < QP < 51$), where the threshold Γ is set to be 8. For *BasketballPass*, the embedding capacity varies from 3254 bits to 904 bits when QP varies from 32 to 50. For *BlowingBubbles*, the embedding capacity varies from 3574 bits to

TABLE 4. Summary of the comparison analysis of the proposed method.

Evaluation	Proposed Method	[4]	[27]	[29]	[30]	[31]
Visual Quality	High	High	High	High	High	High
Capacity	High	Low	Medium	High	Medium	High
Robustness	N/A	Yes	N/A	Yes	Yes	Yes
Embedding Location	Coding Syntax	DST	DST	CW T	WC	CWT
Scope	Intra&Inter	Intra	Intra	Chroma	Intra/Inter	Chroma

TABLE 5. Proportions of candidate TB blocks with different GOP structures.

Sequence		All Intra	IPPP	IBBB
BasketballPass	Proportion(%)	63.45	51.93	51.63
	All Blocks Number	15107	8642	8544
BlowingBubbles	Proportion(%)	58.67	49.21	49.85
	All Blocks Number	36066	12240	12000
KristenAndSara	Proportion(%)	58.3	56.18	56.49
	All Blocks Number	111754	30750	31006
ParkScene	Proportion(%)	63.74	53.51	54.28
	All Blocks Number	382272	124423	122142

578 bits. For *KristenAndSara*, the embedding capacity varies from 16616 bits to 6566 bits and for *ParkScene*, the embedding capacity varies from 52057 bits to 10015 bits. It can be discovered that the embedding capacity is getting smaller when QP is getting larger. However, for these four test video samples, when QP is set to be a large number (e.g., 50), the proposed method can still embed secret message into 8×8 TBs. In this worst case, more video frames are needed to embed large volume secret message.

C. STATISTIC CHARACTERISTICS

Table 5 provides the proportions of selected candidate embedded TBs, which are occupied in the whole 8×8 TB blocks when meeting different GOP structures. The average proportion of candidate TBs with *all intra* is 61.04%, with IPPP is 52.70%, and with IBBB is 53.06%, respectively. It can be seen that the proportion of candidate TBs with *all Intra* is larger than others, as there are more all zero residuals ($CBF = 0$) of TBs with larger size (more than 8×8) in *inter* prediction. The number of all 8×8 TBs in one specified video sequence (presented as All Blocks Num) is also provided in Table 5. The experimental result shows the sufficiency of the candidate embedded TBs, which can provide a strong support to embed large volume of secret message with the proposed method.

The proportions of *single-* and *collaborative-* embedding/ extraction modules with different threshold Γ are described in Fig. 15. It can be seen that the proportion of *collaborative*

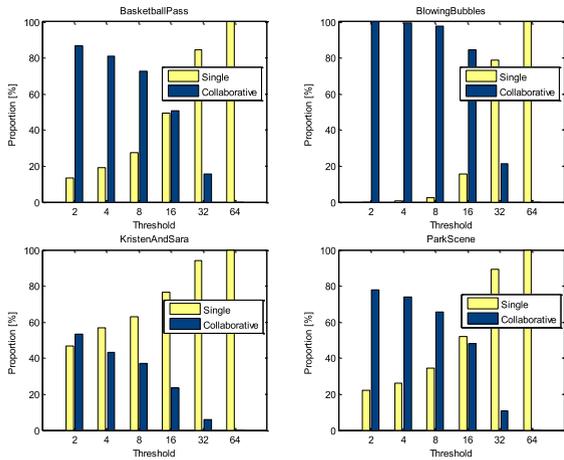


FIGURE 15. Proportions of Single and collaborative with different threshold Γ .

module is getting smaller along with the increase of threshold Γ , however, the proportion of *single* module is becoming bigger, which indicates that the embedding capacity is increasing and the visual quality is dropping at the same time. When the threshold Γ is set to be 2, the proportion of *collaborative* module reaches to the maximum value, and the visual quality of carrier video outperforms best than other cases. However, when the threshold Γ is set to be 64, the proportion of *collaborative* module is becoming to be zero, and the embedding capacity could achieve the maximum values. As a result, a simple way to the tradeoff between visual quality and embedding capacity of the proposed method can be conducted with adjusting threshold Γ under various needs.

V. CONCLUSION

In this paper, we propose a novel and efficient video steganography method utilizing transform block decision for H.265/HEVC. To improve the embedding performance, we employ one adaptive modification mechanism on TB partitioning and one hybrid embedding/extraction scheme. To guarantee the superiority of the proposed method, we employ several technical improvements, including selection constraints for TB Block, synchronization of residuals and CBF, fast traversal for the quadtree-structured CB and TB.

As shown in this work, the modification of TB partitioning can be a desirable technique to decrease the embedding error, and hybrid embedding/extraction scheme is a powerful technique to improve the embedding performance. The experimental results also show that the proposed method can achieve better visual quality, lower bit-rate increase and higher embedding capacity when compared to other relevant state-of-the-art methods. However, since the video steganography is vulnerable to various attacks, such as compression, upscaling, rotation, and cropping, more consideration for the robustness of the proposed method will be our future work.

REFERENCES

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, Dec. 2012.

[2] Y.-H. Chen and V. Sze, "A deeply pipelined CABAC decoder for HEVC supporting level 6.2 high-tier applications," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 25, no. 5, pp. 856–868, May 2015.

[3] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. Keng Tan, and T. Wiegand, "Comparison of the coding efficiency of video coding standards—Including high efficiency video coding (HEVC)," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1669–1684, Dec. 2012.

[4] Y. X. Liu, H. Zhao, S. Liu, C. Feng, S. Liu, "A robust and improved visual quality data hiding method for HEVC," *IEEE Access*, vol. 6, pp. 53984–53987, 2018.

[5] X. Zhang, "Reversible data hiding with optimal value transfer," *IEEE Trans. Multimedia*, vol. 15, no. 2, pp. 316–325, Feb. 2013.

[6] J. Yang and S. Li, "An efficient information hiding method based on motion vector space encoding for HEVC," *Multimedia Tools Appl.*, vol. 77, no. 10, pp. 11979–12001, May 2018, doi: 10.1007/s11042-017-4844-1.

[7] G. Qiu, P. Marziliano, A. T. S. Ho, D. He, and Q. Sun, "A hybrid watermarking scheme for H.264/AVC video," in *Proc. 17th Int. Conf. Pattern Recognit. (ICPR)*, 2004, pp. 865–869.

[8] X. Ma, Z. Li, J. Lv, and W. Wang, "Data hiding in H.264/AVC streams with limited intra-frame distortion drift," in *Proc. Int. Symp. Comput. Netw. Multimedia Technol.*, Dec. 2009, pp. 1–5.

[9] X. Ma, Z. Li, H. Tu, and B. Zhang, "A data hiding algorithm for H.264/AVC video streams without intra-frame distortion drift," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 20, no. 10, pp. 1320–1330, Oct. 2010.

[10] Y. Liu, L. Chen, M. Hu, Z. Jia, S. Jia, and H. Zhao, "A reversible data hiding method for H.264 with Shamir's (t, n)-threshold secret sharing," *Neurocomputing*, vol. 188, pp. 63–70, May 2016.

[11] J. Zhao, Z.-T. Li, and B. Feng, "A novel two-dimensional histogram modification for reversible data embedding into stereo H.264 video," *Multimedia Tools Appl.*, vol. 75, no. 10, pp. 5959–5980, May 2016.

[12] Y. Yao, W. Zhang, and N. Yu, "Inter-frame distortion drift analysis for reversible data hiding in encrypted H.264/AVC video bitstreams," *Signal Process.*, vol. 128, pp. 531–545, Nov. 2016.

[13] C.-H. Liu and T.-C. O. Chen, "Data hiding in inter and intra prediction modes of H.264/AVC," in *Proc. IEEE Int. Symp. Circuits Syst.*, Jun. 2008, pp. 3025–3028, doi: 10.1109/ISCAS.2008.4542095.

[14] D. Xu, R. Wang, and J. Wang, "Prediction mode modulated data-hiding algorithm for H.264/AVC," *J. Real-Time Image Process.*, vol. 7, no. 4, pp. 205–214, Dec. 2012.

[15] D. Li, Y. Zhang, X. Li, K. Niu, X. Yang, and Y. Sun, "Two-dimensional histogram modification based reversible data hiding using motion vector for H.264," *Multimedia Tools Appl.*, vol. 78, no. 7, pp. 8167–8181, Apr. 2019.

[16] Y. Yao, W. Zhang, N. Yu, and X. Zhao, "Defining embedding distortion for motion vector-based video steganography," *Multimedia Tools Appl.*, vol. 74, no. 24, pp. 11163–11186, Dec. 2015.

[17] H. Zhang, Y. Cao, and X. Zhao, "Motion vector-based video steganography with preserved local optimality," *Multimedia Tools Appl.*, vol. 75, no. 21, pp. 13503–13519, Nov. 2016.

[18] H. A. Aly, "Data hiding in motion vectors of compressed video based on their associated prediction error," *IEEE Trans. Inf. Forensics Security*, vol. 6, no. 1, pp. 14–18, Mar. 2011.

[19] D. Xu and R. Wang, "Context adaptive binary arithmetic coding-based data hiding in partially encrypted H.264/AVC videos," *J. Electron. Imag.*, vol. 24, no. 3, Jun. 2015, Art. no. 033028.

[20] D. Xu, R. Wang, and Y. Q. Shi, "Data hiding in encrypted H.264/AVC video streams by codeword substitution," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 4, pp. 596–606, Apr. 2014.

[21] G. Yang, J. Li, Y. He, and Z. Kang, "An information hiding algorithm based on intra-prediction modes and matrix coding for H.264/AVC video stream," *AEU-Int. J. Electron. Commun.*, vol. 65, no. 4, pp. 331–337, Apr. 2011.

[22] Y. Q. Tew and K. Wong, "Information hiding in HEVC standard using adaptive coding block size decision," in *Proc. IEEE Int. Conf. Image Processing*, Oct. 2015, pp. 5502–5506.

[23] T. Shanableh, "Altering split decisions of coding units for message embedding in HEVC," *Multimedia Tools Appl.*, vol. 77, no. 7, pp. 8939–8953, Apr. 2018, doi: 10.1007/s11042-017-4787-6.

[24] T. Shanableh, "Data embedding in high efficiency video coding (HEVC) videos by modifying the partitioning of coding units," *IET Image Process.*, vol. 13, no. 11, pp. 1909–1913, Sep. 2019, doi: 10.1049/iet-ipr.2018.5782.

[25] R. Zhang, V. Sachnev, M. B. Botnan, H. Joong Kim, and J. Heo, "An efficient embedder for BCH coding for steganography," *IEEE Trans. Inf. Theory*, vol. 58, no. 12, pp. 7272–7279, Dec. 2012.

[26] Y. Liu, S. Liu, H. Zhao, and S. Liu, "A new data hiding method for H.265/HEVC video streams without intra-frame distortion drift," *Multimedia Tools Appl.*, vol. 78, no. 6, pp. 6459–6486, Mar. 2019.

[27] Z. Qian, X. Zhang, and S. Wang, "Reversible data hiding in encrypted JPEG bitstream," *IEEE Trans. Multimedia*, vol. 16, no. 5, pp. 1486–1491, Aug. 2014.

[28] M. Asikuzzaman, M. J. Alam, A. J. Lambert, and M. R. Pickering, "Imperceptible and robust blind video watermarking using chrominance embedding: A set of approaches in the DT CWT domain," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 9, pp. 1502–1517, Sep. 2014.

[29] H. Khalilian and I. V. Bajic, "Video watermarking with empirical PCA-based decoding," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4825–4840, Dec. 2013.

[30] M. Asikuzzaman, M. J. Alam, A. J. Lambert, and M. R. Pickering, "Robust DT CWT-based DIBR 3D video watermarking using chrominance embedding," *IEEE Trans. Multimedia*, vol. 18, no. 9, pp. 1733–1748, Sep. 2016.



YONGHAO WANG was born in Jiangsu, China, in 1977. He received the B.E. degree from the Nanjing University of Science and Technology, China, in 1999, and the M.Sc. degree from the University of Central England, Birmingham, U.K., in 2002. He is currently a Senior Lecturer with the DMT in Laboratory, School of Computing and Digital Technology, Birmingham City University. His research interests include digital signal processing, multimedia network distribution, and information security.



HONGGUO ZHAO was born in Shanxi, China, in 1990. He received the master's degree from the Huazhong University of Science and Technology, Wuhan, China, in 2014. He is currently pursuing the Ph.D. degree with the School of Computing and Digital Technology, Birmingham City University, U.K. He joined the College of Information Science and Technology, Zhengzhou Normal University, in 2016. His research interests include video steganography and network security.



SI LIU was born in Xinyang, China, in 1985. He received the master's degree from the Zhengzhou University of Light Industry, Zhengzhou, China, in 2011. He is currently a Lecturer with the College of Information Science and Technology, Zhengzhou Normal University. His research interests include information hiding and artificial intelligence.



YUNXIA LIU was born in Zhoukou, China, in 1972. She received the B.E. degree from Henan Normal University, Xinxiang, China, in 1996, and the Ph.D. degree from the Huazhong University of Science and Technology, Wuhan, China, in 2013. She is currently a Professor with the College of Information Science and Technology, Zhengzhou Normal University. Her research interests include cryptography, network security, and multimedia security.



CONG FENG was born in Shangqiu, China, in 1987. She received the master's degree from Guangxi Normal University, Guangxi, China, in 2013. She joined the College of Information Science and Technology, Zhengzhou Normal University, in 2013. Her research interests include information hiding and artificial intelligence.

...