



ARTICLE

A HEVC Video Steganalysis Method Using the Optimality of Motion Vector Prediction

Jun Li^{1,2}, Mingqing Zhang^{1,2,*}, Ke Niu¹, Yingnan Zhang¹ and Xiaoyuan Yang^{1,2}

¹College of Cryptography Engineering, Engineering University of the Chinese People's Armed Police Force, Xi'an, 710086, China

²Key Laboratory of Network and Information Security of the Chinese People's Armed Police Force, Xi'an, 710086, China

*Corresponding Author: Mingqing Zhang. Email: api_zmq@126.com

Received: 27 November 2023 Accepted: 18 March 2024 Published: 15 May 2024

ABSTRACT

Among steganalysis techniques, detection against MV (motion vector) domain-based video steganography in the HEVC (High Efficiency Video Coding) standard remains a challenging issue. For the purpose of improving the detection performance, this paper proposes a steganalysis method that can perfectly detect MV-based steganography in HEVC. Firstly, we define the local optimality of MVP (Motion Vector Prediction) based on the technology of AMVP (Advanced Motion Vector Prediction). Secondly, we analyze that in HEVC video, message embedding either using MVP index or MVD (Motion Vector Difference) may destroy the above optimality of MVP. And then, we define the optimal rate of MVP as a steganalysis feature. Finally, we conduct steganalysis detection experiments on two general datasets for three popular steganography methods and compare the performance with four state-of-the-art steganalysis methods. The experimental results demonstrate the effectiveness of the proposed feature set. Furthermore, our method stands out for its practical applicability, requiring no model training and exhibiting low computational complexity, making it a viable solution for real-world scenarios.

KEYWORDS

Video steganography; video steganalysis; motion vector prediction; motion vector difference; advanced motion vector prediction; local optimality

1 Introduction

Steganography aims to embed secret messages in multimedia such as pictures, audio, and video without arousing suspicion, thus enabling covert communication. On the other hand, its adversary, steganalysis, aims to detect the presence of embedded secret messages in ordinary media. Video is the ideal cover for steganography, and there are different steganography methods according to the embedding location [1] in the video, mainly intra-frame prediction modes [2–4], frame selection [5], inter-frame prediction modes [6–10], MVs (Motion Vectors) [11–14], transformation coefficients [15–17], etc. Since many MVs are available for message embedding in video coding, more methods are based on the MV domain. Thus, the MV-based steganalysis technique is a current research hotspot.

With the gradual popularization and application of the HEVC (High Efficiency Video Coding) standard [18], the research of MV-based video steganography and steganalysis techniques based on



the HEVC is particularly important. Yang et al. [19] proposed a steganography method based on MV space coding for HEVC. They provided the construction and coding process of MV space. They defined the mapping relationship between the set of MVs and the points in the space, which can achieve the effect of embedding a $2N + 1$ binary number by changing at most one component among N MV components. Guo et al. [12] first counted the motion trend of each frame and established a MTB (Motion Trend Based) mapping strategy between the MV and the binary bitstream, and then used the SATD (Sum of Transform Difference) difference before and after the MV modification as steganographic distortion for message embedding. Hu et al. [20] proposed a new steganography method, named SAMVP (Steganography by Advanced Motion Vector Prediction), which using the AMVP (Advanced Motion Vector Prediction) technique in HEVC. SAMVP uses the MVP (Motion Vector Prediction) index in the AMVP technique of inter-frame prediction as the embedding cover, which has a sizeable embedding capacity and is lossless. Liu et al. [21] proposed the Adaptive-SAMVP (A-SAMVP) based on SAMVP by defining the cost function and combining it with STC (Syndrome Trellis Code) [22]. Since AMVP encodes MVs by MVP index values and MVDs (Motion Vector Differences), A-SAMVP embeds the information in the index values of the candidate list and uses the code rate difference between two candidate MVPs to define the cost function. The overall performance of the algorithm is improved.

The MV-based steganography algorithm is a modification of the MV and its associated information, which inevitably destroys the optimality of specific parameters in the video coding process, so some traditional H.264/AVC-based steganalysis methods are still effective to some extent in HEVC, such as AoSO (Adding or Subtracting One) [23], NPELO (Near Perfect Estimation for Local Optimality) [24], MVC (Motion Vector Consistency) [25]. Nevertheless, researchers have attempted to design steganalysis features that integrate the unique characteristics of HEVC to improve the detection efficiency of steganalysis algorithms. Shanableh [26] extended the idea of the MVC approach from H.264/AVC to HEVC. They redefined the concept of block group based on the coding depth according to the characteristics of HEVC standard and proposed the feature sets based on MV non-consistency. Huang et al. [27] introduced the convolutional neural network to the MV domain video steganalysis based on the HEVC standard and proposed the VSRNet (Video Steganalysis Residual Network) structure. The method constructs independent VSRNet sub-networks for different embedding rates and finally connects all sub-network structures to form a quantitative steganalysis convolutional neural network. Based on VSRNet, they further introduce information such as Selection-Channel-Aware [28] and MVD [29] to improve the performance of steganalysis. In the new type of MV modification strategy [20,21] based on the HEVC standard, it is possible to modify only the MVP index without changing the MV itself. So, the traditional MV-based steganalysis features are ineffective for this new type of steganography algorithm. However, if the MVP index is modified, the local optimality of the MVP in the candidate list may be destroyed. Based on this observation, Liu et al. [30] constructed steganalysis features based on local optimality on the MVP candidate list and MV, and they proposed the LOCL (Local Optimality in Candidate List) method, which effectively improves the detection performance in HEVC.

However, existing MV-based video steganalysis methods still have some significant shortcomings. Firstly, the current methods ignore the disturbance caused by MV steganography to the local optimality of MVP in HEVC, which leads to low detection effectiveness in current video steganalysis. Secondly, existing methods are based on machine learning models that require a significant amount of training to achieve an ideal detection model. However, these trained steganalysis models often have low robustness, as they tend to exhibit noticeable performance degradation in the presence of cover or algorithm mismatches.

Based on the above analysis, this paper focuses on the local optimality of the MVP candidate list in HEVC and fully explores the statistical differences before and after message embedding to design the steganalysis feature. First, either the traditional steganography of modifying MVDs or the new steganography of modifying MVP indexes may have perturbations on the local optimality of MVP. Second, we propose a steganalysis feature with a dimension of only one based on the local optimality of the MVP, which is defined as the optimality rate of the MVP in HEVC codestreams. The optimality rate of the MVP is 100% in all cover videos and below 100% in all stego videos. Based on this feature, we can accurately determine whether or not the video has been modified by steganography.

The main contributions of this paper can be summarized as follows:

1. We conducted an in-depth analysis on how information embedding, whether based on the MVP index or MVD, can potentially disrupt the optimality of MVP in the AMVP technique. As a result, we introduced a steganalysis feature defined as the optimal rate of MVP with a dimension of one. This feature boasts the lowest dimension compared to existing steganalysis features in the MV domain.
2. Unlike previous steganalysis schemes in the MV domain for HEVC, our proposed method consistently achieves 100% accuracy in distinguishing between cover and stego videos.
3. The proposed scheme eliminates the need for redundant model training during execution, ensuring low computational complexity and making it highly suitable for practical application scenarios.

The rest of the paper are organized as follows. The second part introduces the basics knowledge of AMVP technology. The third part analyzes the effect of on the MVP by message embedding based on MVP index and MVD, and defines the optimal rate of MVP as a feature for steganalysis. Then it is proved theoretically that the optimal rate is 100% in the cover video and below 100% in the stego video. The experimental results and analysis are given in the fourth part. Finally, the paper is concluded.

2 Preliminaries

2.1 The Technology of AMVP (Advanced Motion Vector Prediction)

AMVP is an MV prediction technique for inter-frame encoding proposed in HEVC standard. AMVP uses the correlation of MVs in the spatial and temporal domains to build a list of candidate MVPs (including mvp_0 and mvp_1 for a PU (Prediction Unit)). The optimal MVP mvp_{idx} , $idx \in \{0, 1\}$ is selected from the candidate list, and the final optimal MV mv is obtained by whole-pixel and sub-pixel motion estimation starting from the mvp_{idx} . Then the MVD mvd is obtained by:

$$mvd = mv - mvp. \quad (1)$$

mvd is finally encoded using 0-th order Exp-Golomb codes [18]. The decoder recovers the mv of the current PU by building the same list of candidate MVPs, and only needs the index value idx of mvp in the candidate list and the mvd . So that the recovered $mv = mvd + mvp_{idx}$.

2.2 The Local Optimality of the MVP

HEVC adopts the Lagrangian optimization algorithm to achieve encoding control in selecting the optimal MVP from the candidate list. The definition of Lagrangian rate-distortion is as follows:

$$J_{motion}(mv) = D + \lambda * R, \quad (2)$$

where D represents the pixel distortion caused by encoding using the current mv . The distortion D is usually calculated using the SAD (Sum of Absolute Difference) or SATD (Hadamard Sum of Absolute Transformed Difference). λ is a Lagrangian parameter that controls the balance between bit rate and distortion. R represents the number of bits required to encode the current mv , which is actually the number of bits required to encode mvd and MVP index idx :

$$R = \text{Bits}(mvd) + \text{Bits}(idx) = \text{Bits}(mvd) + 1, \quad (3)$$

where $\text{Bits}(idx) = 1$ is the number of bits required to encode the idx , and $\text{Bits}(mvd)$ is the number of bits required to encode mvd using the 0-th order Exp-Golomb codes.

According to the Lagrangian optimization model, without loss of generality, assuming that the optimal MVP selected by the encoder in the candidate list is mvp_{idx} , then mvp_{idx} must meet **the local optimality of the MVP**:

$$J_{motion}(mvp_{idx}) \leq J_{motion}(mvp_{\overline{idx}}), \quad (4)$$

where \overline{idx} represents the values in the set $\{0, 1\}$ that is different from idx . Formula (4) means the rate-distortion of the MVP corresponding to index idx must be the smallest in the candidate list. Due to the fact that the optimal MVP has been determined by the encoder during the final confirmation of the MVP, which means the reference block is determined. Therefore, the distortion D of the two candidate MVPs is the same, so the local optimality of the MVP in formula (4) can be simplified as:

$$R(mvp_{idx}) \leq R(mvp_{\overline{idx}}). \quad (5)$$

That is to say the number of bits encoding the optimal MVP mvp_{idx} is lower than that of another candidate $mvp_{\overline{idx}}$.

3 The Proposed Steganalysis Method

In this section, we first analyze the security risk of the HEVC steganography method using MVP index and MVD, i.e., both of them can perturb the local optimality of the MVP. Then a steganalysis feature is designed based on the optimality of the MVP in AMVP.

3.1 Motion Vector Domain Based Steganography in HEVC

Based on the analysis in the previous section, the Lagrangian rate-distortion optimization model first finds the optimal mvp from the candidate list of MVPs and then finds the optimal mv by motion estimation. Thus the selected mvp is optimal in the sense of rate-distortion in the candidate list. MV-based steganography in HEVC can use the MVP index idx or MVD mvd as embedding cover. The effects of these two embedding methods on the optimality of the MVP are analyzed below.

3.1.1 Using the Index of MVP for Message Embedding

Each PU encoded with the AMVP technique has an MVP index $idx \in \{0, 1\}$. SAMVP [20] and A-SAMVP [21] are the new type of steganographic approaches with idx as the cover. In the method of SAMVP, when the secret message is the same as idx , the information of the corresponding PU block does not have to be modified. When the secret message differs from idx , the value of idx must be modified to \overline{idx} , and then the corresponding mvp is also modified. According to $mvd = mv - mvp$, since mvp is changed while mv remains unchanged, mvd correspondingly needs to be modified. According to the analysis of formula (2), firstly, as mv is unchanged, then the corresponding best matching block is not changed, so there is no change in pixel distortion D , i.e., no visual distortion for this message

embedding; Secondly, the bit rate has changed, mainly due to the change in the number of bits required for *mvd* encoding using 0-th order Exp-Golomb codes. Therefore, although SAMVP is lossless in visual quality, it may increase the bit rate. To reduce the impact of embedding operations on the bit rate, A-SAMVP constructs an adaptive steganographic method using STC. The scheme improves the performance of the steganography method by taking the differences in bit rate before and after message embedding as the cost function.

Although the schemes in SAMVP and A-SAMVP can be lossless in visual quality, an obvious security risk exists. According to the Lagrangian optimization model, the encoder must satisfy the local optimality in formula (4) after selecting the optimal mvp_{idx} from the MVP candidate list. Therefore, if the mvp_{idx} is artificially modified to $mvp_{\bar{idx}}$, there will be evident modification traces from the decoder. Figs. 1a, 1b show the scenarios where the local optimality of the MVP is corrupted due to message embedding using *idx*. Fig. 1a shows the normal case before message embedding. The optimal *mv* by motion estimation is (3, 9), the two candidates mvp_0 and mvp_1 in the MVP candidate list are (3, 8) and (3, 9), respectively, and the corresponding *mvd* are (0, 1) and (0, 0), respectively. Calculated according to formula (3), the number of encoding bits corresponding to the two MVPs is 4 and 3, respectively. So the optimal MVP index is $idx = 1$, i.e., mvp_1 will be selected as the optimal MVP. Fig. 1b shows the situation after message embedding. Assuming that *idx* changes from 1 to 0, which means the mvp_0 is selected for AMVP. And the number of bits needed to encode the two candidate MVPs at the decoding side is still 4 and 3, respectively. As a result, the optimal MVP index *idx* should be 1 and mvp_1 should be selected as the optimal MVP in theory. However, in practice, the *idx* obtained by decoder is 0, thus destroying the optimality of the MVP.

mv	mvp idx	mvp	mvd	Number of bits <i>R</i>	Optimal in Theoretically
(3,9)	0	(3,8)	(0,1)	4	-
	1	(3,9)	(0,0)	3	Yes

(a)

mv	mvp idx	mvp	mvd	Number of bits <i>R</i>	Optimal in Theoretically
(3,9)	0	(3,8)	(0,1)	4	-
	1	(3,9)	(0,0)	3	Yes

(b)

mv	mvp idx	mvp	mvd	Number of bits <i>R</i>	Optimal in Theoretically
(3,8)	0	(3,8)	(0,0)	3	Yes
	1	(3,9)	(0,-1)	5	-

(c)

Figure 1: An example for the local optimality of MVP in HEVC. The gray background represents the actual situation of MVP observed by the decoding end. (a) Cover video, the MVP satisfies the local optimality. (b) Stego video by modifying the MVP index *idx*, the MVP do not satisfies the local optimality. (c) Stego video by modifying the *mvd*, the MVP do not satisfies the local optimality

Based on the above analysis, using the MVP index idx as the message embedding cover could destroy the MVP's local optimality.

3.1.2 Using the MVD for Message Embedding

Using the MVD as a cover for message embedding is a traditional steganography method [11, 19] in H.264/AVC and HEVC. In HEVC, according to $mvd = mv - mvp$, since mvd is modified to mvd' but mvp remains unchanged, which means the mv needs to be modified:

$$mv' = mvd' + mvp. \quad (6)$$

That is to say, the optimal matching block corresponding to the current PU has changed. Although these steganography methods do not directly modify the MVP index idx , they may still destroy the MVP's local optimality. This is because the rate-distortion of the corresponding two candidate MVPs will be changed after the mvd modification. Still, Fig. 1a shows the PU of cover video before embedding, and Fig. 1c shows the case after the mvd of this PU block is modified. Suppose the mvd changes from $(0, 0)$ to $(0, -1)$ after message embedding (usually, only one component is modified, and the modification amplitude is one). At this time, $idx = 1$ remains unchanged, and according to formula (6), the corresponding mv at the decoding side will be changed to $(3, 8)$. From the decoding end, the mvd at $idx = 0$ becomes $(0, 0)$. According to formula (3), the number of bits required to encode mvp_0 and mvp_1 is 3 and 5, respectively. So the optimal MVP index idx should be 0 theoretically, but it is actually 1, thus destroying the optimality of the MVP. From the above analysis, the steganography methods using the MVD as the cover may also destroy the MVP's optimality.

3.2 The Proposed One-Dimensional Steganalysis Feature Based on the Local Optimality of the MVP

According to the analysis in Section 3.1, in the HEVC standard, both the traditional steganography method using MVD as cover and the new steganography method using MVP indexes as cover may perturb the local optimality of the MVP. Specifically, if the video to be detected is a normal cover video, then all the MVPs of PUs encoded by AMVP in its inter-frame coding frames will satisfy the local optimality (formula (4)); if the video to be detected is a stego video containing secret information, then some of the MVPs of PU encoded by AMVP in its inter-frame coding frames will not satisfy the local optimality. To formalize the degree of disturbance to the local optimality of the MVPs, the **optimal rate of the MVP** for the video is defined as follows:

$$Optimal(mvp) = \frac{\sum_{i=1}^N \delta(J_{motion}(mvp_{idx_i}), J_{min})}{N} \times 100\%, \quad (7)$$

where N is the total number of all PUs encoded with the AMVP technique in a video sequence. δ is a check function, $\delta(x, y) = 1$ when x equals y , and $\delta(x, y) = 0$ otherwise. J is the Lagrangian rate-distortion calculated according to formula (2), and $J_{min} = \min\{J_{motion}(mvp_{idx_i}), J_{motion}(mvp_{\overline{idx_i}})\}$, $idx_i, \overline{idx_i} \in \{0, 1\}$. As can be seen from the above definition, $Optimal(mvp)$ essentially describes the proportion of MVPs that satisfy the local optimality among all MVPs in a video sequence. Steganalysis aims to judge whether a video is a cover or a stego video, so the value of $Optimal(mvp)$ can be used as the basis (feature) for steganalysis. If $Optimal(mvp) = 100\%$, the video is considered a normal cover; if $Optimal(mvp) < 100\%$, the video is viewed as a stego. To prove that $Optimal(mvp)$ can be used for steganalysis, two properties and a corollary and its proof are given below.

Property 1: The optimal rate of MVP in the cover video is 100%.

Proof: According to the HEVC standard, AMVP select the one with the minimum Lagrangian rate-distortion in the MVP candidates list $\{mvp_0, mvp_1\}$ as the optimal MVP for the current PU. Without loss of generality, assuming that the optimal MVP in PU is mvp_0 , then according to the rate-distortion minimization rule, there must be:

$$J_{motion}(mvp_0) \leq J_{motion}(mvp_1), \quad (8)$$

and then there are:

$$J_{\min} = \min \{J_{motion}(mvp_0), J_{motion}(mvp_1)\} = J_{motion}(mvp_0), \quad (9)$$

so that $\delta(J_{motion}(mvp_0), J_{\min}) = 1$, and then:

$$Optimal(mvp) = \frac{\sum_{i=1}^N \delta(J_{motion}(mvp_0), J_{\min})}{N} \times 100\% = \frac{\sum_{i=1}^N 1}{N} \times 100\% = 100\%. \quad (10)$$

Finally $Optimal(mvp) = 100\%$, and the proof is completed.

Property 2: If the local optimality of MVP of some PUs in the stego video is broken, the optimal rate of the MVP in the stego video is less than 100%.

Proof: Without loss of generality, it is assumed that AMVP chooses mvp_0 as the optimal MVP before message embedding. According to the analysis in [Section 3.1](#), message embedding using either the MVP index or the MVD may destroy the local optimality of the MVP.

Case (1). For the steganography method of using the MVP index as cover. If the local optimality of the MVP of some PUs is corrupted after embedding, i.e., the selected optimal MVP in the encoder becomes mvp_1 after embedding (see [Section 3.1.1](#)), and $J_{motion}(mvp_0) \leq J_{motion}(mvp_1)$. So that $J_{\min} = \min \{J_{motion}(mvp_0), J_{motion}(mvp_1)\} = J_{motion}(mvp_0)$, and $\delta(J_{motion}(mvp_1), J_{\min}) = 0 < 1$. So the optimal rate of MVP in the decoder is:

$$Optimal(mvp) = \frac{\sum_{i=1}^N \delta(J_{motion}(mvp_1), J_{\min})}{N} \times 100\% < \frac{\sum_{i=1}^N 1}{N} \times 100\% = 100\%. \quad (11)$$

Case (2). For the steganography method of using the MVD as cover. If the local optimality of the MVP of some PUs is corrupted after embedding, the optimal MVPs selected by these PUs remain unchanged, but according to the analysis in [Section 3.1.2](#), the MVDs have changed. Therefore, $J_{motion}(mvp_1) < J_{motion}(mvp_0)$, and $J_{\min} = \min \{J_{motion}(mvp_0), J_{motion}(mvp_1)\} = J_{motion}(mvp_1)$. And then there is $\delta(J_{motion}(mvp_0), J_{\min}) = 0$, so the optimal rate of MVP in the Decoder is:

$$Optimal(mvp) = \frac{\sum_{i=1}^N \delta(J_{motion}(mvp_0), J_{\min})}{N} \times 100\% < \frac{\sum_{i=1}^N 1}{N} \times 100\% = 100\%. \quad (12)$$

Combining [formulas \(11\) and \(12\)](#), the proof of Property 2 is completed.

Corollary 1: Given a video sequence, if its optimal rate of MVP is less than 100%, the sequence is a stego video.

Proof: According to Property 1, if this video is a cover video, its optimal rate of MVP must be equal to 100%. On the contrary, if its optimal rate of MVP is lower than 100%, it means that the optimality of the MVP of some PUs is perturbed, which is an abnormal phenomenon. This perturbation comes from the message embedding, so that the video can be judged as stego.

3.3 The Process of the Proposed Steganalysis Method

Through the analysis of Properties 1, 2 and Corollary 1, we can use the optimal rate of MVP $Optimal(mvp)$ as the steganalysis feature for determining whether the video sequence of HEVC has been modified by message embedding. The proposed steganalysis process is shown in Fig. 2, and the specific steps are as follows:

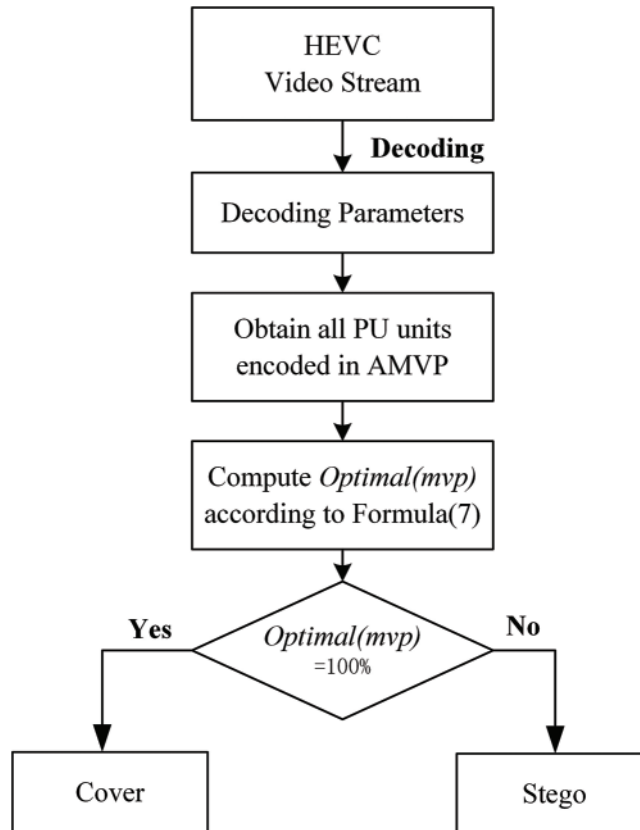


Figure 2: The steganalysis process of the proposed scheme

Step 1: The HEVC compressed video sequence is decoded to obtain the decoding parameters.

Step 2: All PU units encoded using the AMVP technique and their corresponding parameters (MVs, MVP candidate lists, etc.) are collected.

Step 3: The optimal rate of MVP $Optimal(mvp)$ for the video sequence is calculated according to formula (7).

Step 4: Finally, the value of $Optimal(mvp)$ is used for judgment. If $Optimal(mvp) = 100\%$, it indicates that the optimal MVPs of all PU units encoded using the AMVP technique are intact, so the video sequence is judged as a normal cover video. If $Optimal(mvp) < 100\%$, it indicates that the optimal MVPs of some PU units have been damaged, and the video sequence is judged as a stego video.

In the above steganalysis process, the value needed to be calculated for each detected video sequence is only the $Optimal(mvp)$, so there is no model training and classification. Therefore, the computational complexity of the proposed method is low, and the execution efficiency is high.

4 Experiments and Analysis

In this section, some different setups are presented to evaluate the performance of the proposed scheme.

4.1 Experiments Setup

4.1.1 Video Databases

Table 1 shows two video databases used for the experiments. The database DB1 contains 34 well-known standard test sequences [31] with a resolution of Common Intermediate Format (CIF, 352×288), and each video sequence is cut into a fixed length by selecting its first 240 frames (so the total number of frames for experiments is 8160). The DB1 database is widely used in video steganography and steganalysis. It is characterized by low resolution, and all videos have the same resolution. In addition, the video types in DB1 are very rich, including both slow-motion and fast-motion videos, both dynamic and static backgrounds, and both human and landscape videos. Another database, DB2, contains 80 standard test sequences with different resolutions (from 416×240 to 2560×1600), which are downloaded from the internet, and each sequence is cut into a fixed length by selecting its first 100 frames (so the number of total frames is 8000). The DB2 dataset is also a commonly used standard dataset in video steganography and steganalysis. It is characterized by different resolutions and more high-definition videos. In addition, the video types in DB2 are also very rich. All the video sequences in DB1 and DB2 are stored in uncompressed file format, with YUV 4:2:0 color space.

Table 1: Databases with YUV format of test video sequences

Database	Examples	Number of total frames	Resolution	Number of frames in each sequence	Number of sequence
DB1	Foreman, Akiyo, Bridge_close, Bridge_far, City, News, Football...	8160	352×288	240	34
DB2	Basketball, BQMall, Cactus, ChinaSpeed, Johnny, RaceHorses, SlideShow...	8000	2560×1600 1920×1080 1280×720 832×480 416×240	100 100 100 100 100	10 10 20 20 20

4.1.2 Steganography Methods

Three state-of-the-art typical MV-based steganography methods for HEVC are used for message embedding to evaluate the detectability of video steganalysis in the MV domain. The first one is Yang's method [19] (denoted as Tar1), the second one is Hu's method [20] (denoted as Tar2), and the last one is Liu's method [21] (denoted as Tar3). Due to the different design principles of the above three methods, their embedding capacities are evaluated differently. The embedding strength e in Tar1 is a decimal whose range is 0 to 1, representing the probability of whether the secret information is embedded in CTU, which shall be set at 0.1, 0.2, 0.3, 0.4, and 0.5. The embedding threshold T in Tar2 is defined

as $T = abs(abs(H_1) - abs(H_0)) + abs(abs(V_1) - abs(V_0))$, where H_0, V_0 represent the horizontal and vertical components of mvp_0 , H_1, V_1 represent the horizontal and vertical components of mvp_1 , and $abs(x)$ is the absolute value of x . The T will be set at 0, 1, 5, 20, and 1000 for experiments. The embedding capacity for Tar3 is bpap (Bits Per AMVP PU), which shall be set at 0.1, 0.2, 0.3, 0.4, and 0.5. All the steganography methods are implemented based on the official test model HM16.9 [32]. It is worth noting that, since the proposed scheme in this paper is a steganalysis algorithm based on MV domain, we only use the steganography method based on MV domain for message embedding in the experiment. However, according to the existing research on steganalysis algorithm based on multi-domain [25], the message embedding in other domains (such as inter-frame coding mode) may also lead to abnormal statistical characteristics of MV domain. Therefore, in future research, we will study the applicability of the proposed method in other embedding domains.

4.1.3 Competitor Steganalysis Methods

There are two types of competitor steganalysis methods used for the experiments. The first type is the parallel porting of classical methods proposed for H.264/AVC to HEVC, including AoSO [23] and NPELO [24]. The other type is the MV-based steganalysis methods proposed based on HEVC, including the neural network-based VSRNet method [27] and the local optimality in the candidate list-based LOCL method [30]. All the steganalysis feature sets are extracted based on the official test model HM16.9.

4.1.4 Training and Classification

It is worth noting that the steganalysis scheme proposed in this article does not require the use of machine learning methods for training and classification, as we can determine whether there exist message embedding based on the proposed optimal rate of MVP. To implement training and classification for various competitive steganalysis approaches, we use a Gaussian-kernel SVM (Support Vector Machine) [33], whose penalty factor C and kernel factor γ are established via a five-fold cross-validation. Additionally, the accuracy rate-which is calculated as the proportion of correctly identified samples to all samples-is used to gauge the effectiveness of the detection process. Ten randomly selected database splits are used to get the final accuracy rate. Each iteration uses 40% of the cover-stego video pairings for testing, and 60% are randomly selected for training. A desktop computer with a 3.1 GHz Intel Core i9 CPU and 64 GB RAM is used to conduct all of the tests. The evaluation metric for steganalysis detection is the average correct detection rate, which is the average of the true positive rate and false positive rate.

4.2 The Optimal Rate of MVP for Cover video

This experiment verifies the applicability of Property 1 of Section 3.2 under different conditions, i.e., the optimal rate of MVP is calculated according to formula (7) for cover video in DB1 and DB2 databases. To detect the impact of different encoders, the cover videos in this experiment are compressed using two different encoders.

The first encoder is HM16.9, and the cover videos are encoded with QP (Quality Parameters) of 20, 25, and 30. The GOP (Group Of Picture) structure used for HM is IPPPPPPPPPI..., and the experimental results are shown in Table 2. Two metrics are counted in the table, the first is the average of the optimal rate of MVP of all videos in the database, and the second is the proportion of videos with a 100% optimal rate to the whole videos. The experimental data shows that both DB1 and DB2, *Optimal (mvp)* have a mean value of 100% at different QPs. The proportion of videos with a 100%

optimal rate to the whole videos is also 100%. The experimental data indicate that among the cover videos encoded with HM, all the PU encoded with AMVP meet the local optimality of MVP, i.e., they satisfy the Property 1.

Table 2: The optimal rate of MVP for cover videos encoded by HM

Database	QP	<i>Optimal (mvp)</i>	Proportion of videos with 100% optimal rate
DB1	20	100	100
	25	100	100
	30	100	100
DB2	20	100	100
	25	100	100
	30	100	100

The second encoder is the efficient x265 [34]. The parameters used for x265 are the same as HM above. The experimental results are shown in Table 3, and it can be seen that both in DB1 and DB2, *Optimal (mvp)* have a mean value of 100% at different QPs. The proportion of videos with a 100% optimal rate to the whole videos is also 100%. The experimental result indicates that the cover video compressed by x265 also satisfies the Property 1.

Table 3: The optimal rate of MVP for cover videos encoded by x265

Database	QP	<i>Optimal (mvp)</i>	Proportion of videos with 100% optimal rate
DB1	20	100	100
	25	100	100
	30	100	100
DB2	20	100	100
	25	100	100
	30	100	100

The results of the above two experiments show that both the official reference software HM and the optimized high-performance encoder x265 have an optimal rate of 100% of MVP for cover videos under different encoding parameters, which follows the Property 1.

4.3 The Optimal Rate of MVP for Stego Video

This experiment analyzes the detection performance of the proposed method on stego video. We use different steganography methods for message embedding on DB1 and DB2 databases, with an encoder of HM16.9 and a GOP structure of IPPPPPPPI.

The experimental results for the steganography algorithm Tar1 [19] are shown in Table 4. Taking the embedding strength $e = 0.1$ at QP = 20 in the DB1 database as an example, the average value of *Optimal (mvp)* is 99.70%, which means the optimality of the MVP of 0.3% PUs is corrupted. The proportion of videos with *Optimal (mvp)* equal to 100% of the total videos is 0%, which indicates that no video fully satisfies the MVP optimality. As a whole from Table 4, the average value of *Optimal*

(mvp) is lower than 100% for different databases and QPs, and the proportion of the number of videos with *Optimal* (mvp) equal to 100% is 0%, which indicates that the optimality of MVP of all videos is perturbed. Based on the above analysis, it can be determined that all video sequences in this experiment are stego videos, and the reason is that the embedding operation performed by Tar1 on the MVDs destroys the local optimality of the MVP, which is consistent with the theoretical analysis in Section 3.1.2 and Property 2 in Section 3.2.

Table 4: Statistical data of the optimal rate of MVP for steganography method Tar1 [19]. The first value is the average (%) of the optimal rate of all videos, and the second value is the proportion (%) of the videos with a 100% optimal rate to the whole videos

DB	QP	Embedding strength e									
		0.1	0.2	0.3	0.4	0.5					
DB1	20	99.70	0	99.60	0	99.35	0	99.29	0	99.14	0
	25	99.85	0	99.70	0	99.52	0	99.41	0	99.24	0
	30	99.87	0	99.73	0	99.58	0	99.47	0	99.33	0
DB2	20	99.81	0	99.70	0	99.57	0	99.47	0	99.38	0
	25	99.83	0	99.70	0	99.54	0	99.40	0	99.28	0
	30	99.85	0	99.72	0	99.59	0	99.45	0	99.34	0

The experimental results for the steganography algorithm Tar2 [20] are shown in Table 5. With different databases and QPs, the average value of *Optimal*(mvp) is 100% when the embedding threshold $T = 0$ and the proportion of videos with *Optimal*(mvp) equal to 100% of the total videos is 0%. According to the definition of T in Section 4.1.2, $T = 0$ means that mvp_0 is the same as mvp_l , so modifying the MVP index does not change the optimality of the MVP. That means when $T = 0$, the scheme of this paper is invalid. In fact, at $T = 0$, the Tar2 algorithm only selects those PUs whose mvp_0 is the same as mvp_l for message embedding, and its embedding capacity is smaller. When $T \neq 0$, the first indicator in the experimental results are not 100%, and the second indicator are 0%, which indicates that the optimality of the MVP of all videos is perturbed. We can determine these videos as stego based on these two statistical indicators. The reason for the above experimental results is that the embedding operation performed by Tar2 on the MVP index destroys the MVP's local optimality, which is consistent with the theoretical analysis in Section 3.1.1 and Property 2 in Section 3.2.

Table 5: Statistical data of the optimal rate of MVP for steganography method Tar2 [20]. The first value is the average (%) of the optimal rate of all videos, and the second value is the proportion (%) of the videos with a 100% optimal rate to the whole videos

DB	QP	Embedding threshold T									
		0	1	5	20	1000					
DB1	20	100	100	99.90	0	82.27	0	72.84	0	69.17	0
	25	100	100	93.27	0	81.92	0	71.30	0	66.80	0
	30	100	100	94.04	0	82.97	0	71.55	0	66.36	0

(Continued)

Table 5 (continued)

DB	QP	Embedding threshold T									
		0		1		5		20		1000	
DB2	20	100	100	95.17	0	84.85	0	75.01	0	68.57	0
	25	100	100	94.65	0	83.34	0	72.31	0	65.77	0
	30	100	100	95.02	0	83.69	0	71.52	0	64.07	0

The experimental results for the steganography algorithm Tar3 [21] are shown in Table 6. The experimental results are similar to those of Tar2, but the Tar3 method is an adaptive embedding in MVP index with STC. The experimental results show that the proposed scheme can ideally detect the damage to the local optimality of the MVP. The underlying reason is that the embedding operation performed by Tar3 on the MVP index destroys the local optimality of the MVP, which is also consistent with the theoretical analysis in Section 3.1.1 and Property 2 in Section 3.2.

Table 6: Statistical data of the optimal rate of MVP for steganography method Tar3 [21]. The first value is the average (%) of the optimal rate of all videos, and the second value is the proportion (%) of the videos with a 100% optimal rate to the whole videos

DB	QP	Embedding capacity (in bpap)									
		0.1		0.2		0.3		0.4		0.5	
DB1	20	98.88	0	98.99	0	98.08	0	96.46	0	94.40	0
	25	98.17	0	98.49	0	97.69	0	95.98	0	93.74	0
	30	94.46	0	96.24	0	95.53	0	94.17	0	92.03	0
DB2	20	99.46	0	99.26	0	98.38	0	96.78	0	94.59	0
	25	99.49	0	99.17	0	98.06	0	96.15	0	93.64	0
	30	99.08	0	98.91	0	97.56	0	95.32	0	92.56	0

Through the above analysis, for the three state-of-the-art steganography methods, the proposed steganalysis feature in this paper is invalid only in the case of $T = 0$ for the Tar2. In all other conditions, we can accurately distinguish the cover video from the stego video by whether the optimal rate of MVP is equal to 100%. The above experimental findings also verify the correctness of Property 2 and Corollary 1 in Section 3.2.

4.4 Comparison with Other Traditional Machine Learning-Based Steganalysis Methods

To compare the detection performance of existing steganalysis methods against the above three steganography algorithms, this section uses four traditional existing state-of-the-art steganalysis methods to detect the stego videos in Section 4.3. Due to limitations in the length of the paper, we only list some experimental data on database DB1 in Table 7. The feature set of AoSO [23] uses the SAD to describe the MV's local optimal. From the experimental data, AoSO has some detection effect on Tar1 because Tar1 is a steganography method that directly modifies the MV and destroys the local optimality of the MV. AoSO is ineffective on Tar2 and Tar3 because these two steganography

methods embed messages in the MVP index, and the original MV remains unchanged. NPELO [24] is a steganalysis feature set based on the local optimality of MV rate-distortion. The experimental data shows that NPELO performs better in detecting Tar1 than Tar2 and Tar3 for reasons similar to AoSO. The overall performance of NPELO is better than that of AoSO because NPELO considers rate-distortion (including pixel distortion and code bit), making it more reasonable. VSRNet [27] is a neural network-based steganalysis method, which is effective in detecting Tar1 but ineffective for Tar2 and Tar3. The experimental results indicate that VSRNet cannot yet capture the perturbations caused by the steganography method, which embeds messages in the index of MVP. LOCL [30] is a feature set that combines the NPELO and the optimal MVP candidate list, and its performance is better than AoSO, NPELO and VSRNet overall. However, LOCL considers the MVP's optimality together with the MV's optimality when designing the features, but the optimality of the MVP still needs to be fully exploited.

Table 7: Detector accuracy (%) of competitor four traditional steganalysis methods against three steganography methods in database DB1

Steganalysis methods	QP	Steganography method & Embedding parameter	Tar1 [19] Embedding strength e		Tar2 [20] Embedding threshold T		Tar3 [21] Embedding capacity (in bpap)	
			0.1	0.5	0	1000	0.1	0.5
AoSO [23]	20		58.29	67.80	49.17	48.17	48.45	50.04
	30		54.92	63.90	49.20	48.80	48.90	48.15
NPELO [24]	20		71.93	74.39	63.47	73.89	63.22	60.99
	30		68.31	70.44	60.86	71.64	45.40	48.88
VSRNet [27]	20		68.11	70.19	49.34	56.12	49.34	52.42
	30		65.24	67.83	50.23	54.90	49.08	51.98
LOCL [30]	20		72.36	76.80	59.96	72.82	69.55	68.21
	30		65.77	67.56	46.21	87.32	57.74	64.62

From the above analysis, the detection accuracy of traditional steganalysis features for the three steganography algorithms is low (mostly below 80%). In contrast, according to the analysis in Sections 4.2 and 4.3, the proposed optimal rate of MVP can perfectly distinguish cover videos from stego videos in most cases. The reasons can be described from two aspects. First, the proposed method is designed based on an apparent vulnerability of the existing HEVC steganography algorithm in the MV domain, which has not been fully exploited in previous methods. The proposed steganalysis algorithm designed by the vulnerability can effectively attack the MV-based HEVC steganography algorithm. Second, the mechanism adopted by the proposed method is different from that adopted by the above comparison schemes. The value of the optimal rate of the MVP is used as the basis for judging whether it is a stego video. The other comparison schemes adopt the mechanism based on machine learning, which requires model training of cover and stego videos, and the overall detection performance is not high.

4.5 Applicability in B-Frames

The proposed method is based on the AMVP technique and can be applied to all videos encoded with the AMVP technique. Therefore, although the inter-frame coding frames used in the previous experiments are P-Frames, they are also theoretically applicable to B-Frames. If there are two reference

lists for the PU block encoded by the AMVP technique in the B-Frame, the MVP on both reference lists satisfies the properties of [Section 3.2](#).

To verify the applicability of the proposed scheme on B-Frames, we use the GOP structure of IBBBBBBBBBI... for this experiment. The experiment is performed on the DB1 database with the steganography method as Tar2 (it is worth noting that Tar1 and Tar3 are also applicable and are not listed in this paper due to space limitation), and other parameters are the same as in [Section 4.3](#). [Table 8](#) shows the statistical results of the optimal rate of MVP for the cover video and stego video. As can be seen from the data in the table, the average value of *Optimal (mvp)* for the cover videos at different QPs is 100%, and the videos with 100% optimal rate account for 100% of the whole dataset, indicating that the experimental results at B-Frame also satisfy the Property 1 in [Section 3.2](#).

Table 8: Statistical results for the cover and stego video with GOP structure of IBBBBBBBBBI... The steganography algorithm is Tar2 [20], and the database is DB1. The first value is the average (%) of the optimal rate of all videos, and the second value is the proportion (%) of the videos with a 100% optimal rate to the whole videos

QP	Cover		Embedding threshold T									
			0		1		5		20		1000	
20	100	100	100	100	95.80	0	87.44	0	79.63	0	75.97	0
25	100	100	100	100	96.16	0	87.52	0	77.94	0	72.45	0
30	100	100	100	100	96.70	0	88.05	0	77.59	0	71.28	0

The results at the embedding threshold $T = 0$ are consistent with those at $T = 0$ in [Table 5](#), again because only the PUs whose $mvp_0 = mvp_1$ are used for message embedding, so the optimality of the MVP is not destroyed. In contrast, when $T \neq 0$, *Optimal (mvp)* in the experimental results is not 100%, and the proportion of the videos with a 100% optimal rate to the whole videos is 0%, which indicates that the optimality of the MVP of all videos is perturbed. In summary, for the HEVC videos compressed with B-Frames, except for $T = 0$, the proposed optimal rate of MVP can still effectively distinguish between cover and stego videos.

4.6 The Complexity Analysis of the Proposed Feature

In order to analyze the computational complexity of the proposed scheme, this subsection compares the time required for feature extraction with different QPs. [Table 9](#) shows the dimensionality of the four steganalysis features and the average time needed to extract a video sequence (see [Section 4.3](#) for parameter settings, CIF format, 240 frames, DB1). The experiments are run on a desktop computer with a 3.1 GHz Intel Core i9 CPU and 64 GB RAM. The data in the table shows that the feature dimension of the proposed scheme in this paper is only 1, which is the lowest among all methods. Regarding computational complexity, both AoSO and NPELO need to compute the 1-neighbourhood optimality of the MV, and the computational complexity is close. The highest complexity is LOCL because it has to calculate not only the optimality of the MV but also the optimality of the MVP's candidate list. The extraction time of the proposed scheme is only about 1/2 of the other algorithms, because the proposed scheme does not need to calculate the 1-neighborhood optimality of the MV, but only the rate-distortion of the two MVPs. In addition, the smaller the QP, the larger the running time of all algorithms. This is because the smaller the QP, the finer the division of coding blocks, the more MVs in the code stream, and the more data to be processed.

Table 9: Average computational time (in second) of feature extraction for a single CIF sequence with 240 frames (in DB1)

Steganalysis methods	Dimension	QP		
		20	25	30
AoSO [23]	18	4.15	3.30	2.64
NPELO [24]	36	4.45	3.55	2.86
LOCL [30]	37	4.70	3.85	2.95
Proposed	1	2.40	1.81	1.39

Overall, the computational complexity of the scheme in this paper is minimal because the feature dimension is only one, and the rate-distortion of only two MVPs needs to be calculated. More importantly, the proposed scheme does not require extensive machine learning based training. Therefore, the proposed method is very efficient and can be applied to practical scenarios.

5 Conclusion

Problems in information security are always the mutual game between attackers and defenders [35,36], so are video steganography and steganalysis. The development of video coding standards is always aimed at reducing redundancy and improving compressibility to ensure visual quality [18,37]. The primary starting point of steganography is to use data redundancy to embed secret information. Therefore, although the new coding standard provides more coding elements (such as the MVP index in AMVP technology) for information embedding, it also poses a challenge to steganography due to reduced redundancy. For example, with the wide application of HEVC, video information hiding based on HEVC has received more and more attention. The AMVP technology introduced in HEVC provides more embedding space for steganography, which can be used for message embedding by MVD or MVP index. However, the existing steganography of the MV domain based on HEVC also exposes security risks. Firstly, we analyzed in this paper that information embedding using the MVP index or the MVD may lead to perturbations in the optimality of the MVP. Secondly, we design the optimal rate of MVP as a steganalysis feature. Finally, through theoretical analysis and experimental verification, the proposed scheme not only can accurately distinguish the cover video from the stego video, but also does not require traditional model training and classification, and has the characteristics of low complexity.

The future work is primarily structured around two focal points. Firstly, we intend to persist in examining whether novel techniques emerging in alternative standards, such as HEVC and VVC, introduce fresh vulnerabilities to the prevalent steganography algorithms. Secondly, we aim to investigate the utilization of AMVP technology's advantages to seamlessly integrate information into HEVC without compromising the optimality of MVP. This approach aims to elevate the embedding capacity and bolster resistance against steganalysis attacks.

Acknowledgement: We would like to thanks all those who made a contribution to the completion of this paper.

Funding Statement: This work is supported by the National Natural Science Foundation of China (Grant Nos. 62272478, 62202496, 61872384).

Author Contributions: The authors confirm contribution to the paper as follows: Jun Li: Conceptualization, Methodology, Software, Investigation, Writing-Original Draft; Minqing Zhang: Conceptualization, Funding Acquisition, Writing-Review & Editing, Supervision; Ke Niu: Software; Yingnan Zhang: Investigation, Validation; Xiaoyuan Yang: Writing-Review & Editing, Validation.

Availability of Data and Materials: The executable program related to this paper can be downloaded from: <https://github.com/lijun9250lj/HEVC-MVPO-Video-steganalysis>.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] M. Bouzegza, A. Belatreche, A. Bouridane, and M. Tounsi, "A comprehensive review of video steganalysis," *IET Imag. Proce.*, vol. 16, no. 13, pp. 3407–3425, 2022. doi: [10.1049/ipr2.12573](https://doi.org/10.1049/ipr2.12573).
- [2] Y. Dong, X. Jiang, Z. Li, T. Sun, and Z. Zhang, "Multi-channel HEVC steganography by minimizing IPM steganographic distortions," *IEEE Trans. Multimed.*, vol. 25, pp. 2698–2709, 2023. doi: [10.1109/TMM.2022.3150180](https://doi.org/10.1109/TMM.2022.3150180).
- [3] Y. Dong, X. Jiang, Z. Li, T. Sun, and P. He, "Adaptive HEVC steganography based on steganographic compression efficiency degradation model," *IEEE Trans. Depend. Secur. Comput.*, vol. 20, no. 1, pp. 769–783, 2023. doi: [10.1109/TDSC.2022.3144139](https://doi.org/10.1109/TDSC.2022.3144139).
- [4] H. Zhang, Y. Cao, X. Zhao, W. Zhang, and N. Yu, "Video steganography with perturbed macroblock partition," in *Pro. 2nd ACM Workshop Inf. Hid. Multime. secu.*, New York, USA, Jun. 11–13, 2014, pp. 115–122.
- [5] A. Sajjad, H. Ashraf, N. Jhanjhi, M. Humayun, M. Masud and M. A. AlZain, "Improved video steganography with dual cover medium, DNA and complex frames," *Comput. Mater. Contin.*, vol. 74, no. 2, pp. 3881–3898, 2023. doi: [10.32604/cmc.2023.030197](https://doi.org/10.32604/cmc.2023.030197).
- [6] Z. Li, X. Jiang, Y. Dong, L. Meng, and T. Sun, "An anti-steganalysis HEVC video steganography with high performance based on CNN and PU partition modes," *IEEE Trans. Depend. Secure. Comput.*, vol. 20, no. 1, pp. 606–619, 2023. doi: [10.1109/TDSC.2022.3140899](https://doi.org/10.1109/TDSC.2022.3140899).
- [7] J. Liu, Z. Li, X. Jiang, and Z. Zhang, "A high-performance CNN-applied HEVC steganography based on diamond-coded PU partition modes," *IEEE Trans. Multimed.*, vol. 24, pp. 2084–2097, 2022. doi: [10.1109/TMM.2021.3075858](https://doi.org/10.1109/TMM.2021.3075858).
- [8] J. Wang, X. Jia, X. Kang, and Y. Shi, "A cover selection HEVC video steganography based on intra prediction mode," *IEEE Access*, vol. 7, pp. 119393–119402, 2019. doi: [10.1109/ACCESS.2019.2936614](https://doi.org/10.1109/ACCESS.2019.2936614).
- [9] S. He, D. Xu, L. Yang, and Y. Liu, "HEVC video information hiding scheme based on adaptive double-layer embedding strategy," *J. Vis. Commun. Image Rep.*, vol. 87, no. June, pp. 103549, 2022. doi: [10.1016/j.jvcir.2022.103549](https://doi.org/10.1016/j.jvcir.2022.103549).
- [10] R. Li, J. Qin, Y. Tan, and N. N. Xiong, "Coverless video steganography based on frame sequence perceptual distance mapping," *Comput. Mater. Contin.*, vol. 73, no. 1, pp. 1571–1583, 2022. doi: [10.32604/cmc.2022.029378](https://doi.org/10.32604/cmc.2022.029378).
- [11] J. Li, M. Zhang, K. Niu, and X. Yang, "Investigation on principles for cost assignment in motion vector-based video steganography," *J. Inf. Secur. Appl.*, vol. 73, pp. 103439, 2023. doi: [10.1016/j.jisa.2023.103439](https://doi.org/10.1016/j.jisa.2023.103439).
- [12] M. Guo, T. Sun, X. Jiang, Y. Dong, and K. Xu, "A motion vector-based steganographic algorithm for HEVC with MTB mapping strategy," in *The 2019 Int. Workshop on Digit. Watermarking*, Chengdu, China, Springer, Nov. 2–4, 2020, pp. 293–306. doi: [10.1007/978-3-030-43575-2_25](https://doi.org/10.1007/978-3-030-43575-2_25).
- [13] Y. Liu, J. Ni, W. Zhang, and J. Huang, "A novel video steganographic scheme incorporating the consistency degree of motion vectors," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 32, no. 7, pp. 4905–4910, 2022. doi: [10.1109/TCSVT.2021.3135384](https://doi.org/10.1109/TCSVT.2021.3135384).

- [14] J. Li, M. Zhang, K. Niu, Y. Zhang, and X. Yang, "Motion vector-domain video steganalysis exploiting skipped macroblocks," *IET Image Process.*, Dec. 2023. doi: [10.1049/ipr2.13014](https://doi.org/10.1049/ipr2.13014).
- [15] Y. Chen, H. Wang, H. Wu, Z. Wu, T. Li and A. Malik, "DDCA: A distortion drift-based cost assignment method for adaptive video steganography in the transform domain," *IEEE Trans. Depend. Secure. Comput.*, vol. 19, no. 4, pp. 2405–2420, 2022. doi: [10.1109/TDSC.2021.3058134](https://doi.org/10.1109/TDSC.2021.3058134).
- [16] Y. Wang, Y. Cao, and X. Zhao, "CEC: Cluster embedding coding for H.264 steganography," *IEEE Signal Process. Lett.*, vol. 27, pp. 955–959, 2020. doi: [10.1109/LSP.2020.2998360](https://doi.org/10.1109/LSP.2020.2998360).
- [17] L. Yang, D. Xu, R. Wang, and S. He, "Adaptive HEVC video steganography based on distortion compensation optimization," *J. Inf. Secur. Appl.*, vol. 73, pp. 103442, Mar. 2023. doi: [10.1016/j.jisa.2023.103442](https://doi.org/10.1016/j.jisa.2023.103442).
- [18] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the high efficiency video coding (HEVC) standard," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 22, no. 12, pp. 1649–1668, 2012. doi: [10.1109/TCSVT.2012.2221191](https://doi.org/10.1109/TCSVT.2012.2221191).
- [19] J. Yang and S. Li, "An efficient information hiding method based on motion vector space encoding for HEVC," *Multimed. Tools Appl.*, vol. 77, no. 10, pp. 11979–12001, 2018. doi: [10.1007/s11042-017-4844-1](https://doi.org/10.1007/s11042-017-4844-1).
- [20] Y. Hu, W. Gong, F. Liu, L. Liu, and M. Zhu, "Large-capacity lossless HEVC information hiding based on index parameter modification," *J. South China Univ. Technol. (Nat. Sci. Ed.)*, vol. 46, no. 5, pp. 1–8, 2018.
- [21] S. Liu, B. Liu, Y. Hu, and X. Zhao, "Non-degraded adaptive HEVC steganography by advanced motion vector prediction," *IEEE Signal Process. Lett.*, vol. 28, pp. 1843–1847, 2021. doi: [10.1109/LSP.2021.3111565](https://doi.org/10.1109/LSP.2021.3111565).
- [22] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Trans. Inf. Forensic. Secur.*, vol. 6, no. 3, pp. 920–935, 2011. doi: [10.1109/TIFS.2011.2134094](https://doi.org/10.1109/TIFS.2011.2134094).
- [23] K. Wang, H. Zhao, and H. Wang, "Video steganalysis against motion vector-based steganography by adding or subtracting one motion vector value," *IEEE Trans. Inf. Forensic. Secur.*, vol. 9, no. 5, pp. 741–751, 2014. doi: [10.1109/TIFS.2014.2308633](https://doi.org/10.1109/TIFS.2014.2308633).
- [24] H. Zhang, Y. Cao, and X. Zhao, "A steganalytic approach to detect motion vector modification using near-perfect estimation for local optimality," *IEEE Trans. Inf. Forensic. Secur.*, vol. 12, no. 2, pp. 465–478, 2017. doi: [10.1109/TIFS.2016.2623587](https://doi.org/10.1109/TIFS.2016.2623587).
- [25] L. Zhai, L. Wang, and Y. Ren, "Universal detection of video steganography in multiple domains based on the consistency of motion vectors," *IEEE Trans. Inf. Forensic. Secur.*, vol. 15, pp. 1762–1777, 2020. doi: [10.1109/TIFS.2019.2949428](https://doi.org/10.1109/TIFS.2019.2949428).
- [26] T. Shanableh, "Feature extraction and machine learning solutions for detecting motion vector data embedding in HEVC videos," *Multimed. Tools Appl.*, vol. 80, no. 18, pp. 27047–27066, 2021. doi: [10.1007/s11042-020-09826-1](https://doi.org/10.1007/s11042-020-09826-1).
- [27] X. Huang, Y. Hu, Y. Wang, B. Liu, and S. Liu, "Deep learning-based quantitative steganalysis to detect motion vector embedding of HEVC videos," in *2020 IEEE Fifth Int. Conf. Data Sci. Cybersp.*, Hong Kong, China, Jul. 27–29, 2020, pp. 150–155.
- [28] X. Huang, Y. Hu, Y. Wang, B. Liu, and S. Liu, "Selection-channel-aware deep neural network to detect motion vector embedding of HEVC videos," in *2020 IEEE Int. Conf. Sig. Process., Commun. Comput.*, Macau, China, Aug. 21–24, 2020, pp. 1–6.
- [29] Y. Hu, X. Huang, Y. Wang, B. Liu, and S. Liu, "Improving deep learning-based video steganalysis by using MVD detection space," *J. Image Graph.*, vol. 3, no. 28, pp. 702–715, 2022.
- [30] S. Liu, Y. Hu, B. Liu, and C. T. Li, "An HEVC steganalytic approach against motion vector modification using local optimality in candidate list," *Pattern Recognit. Lett.*, vol. 146, pp. 23–30, 2021. doi: [10.1016/j.patrec.2021.02.018](https://doi.org/10.1016/j.patrec.2021.02.018).
- [31] YUV Video Sequences. Accessed: Feb. 10, 2023. [Online]. Available: <https://media.xiph.org/video/derf>
- [32] HEVC HM Reference Software. Accessed: Feb. 05, 2023. [Online]. Available: <http://www.hevc.info>
- [33] C. C. Chang, and C. J. Lin, "LIBSVM: A library for support vector machines," Accessed: Feb. 20, 2023. [Online]. Available: <https://www.csie.ntu.edu.tw/~#x007E; cjlin/libsvm/>
- [34] VideoLAN, x265. Accessed: Feb. 01, 2022. [Online]. Available: <https://www.videolan.org/developers/x265.html>

- [35] Z. Wang, W. Wang, Y. Yang, Z. Han, D. Xu and C. Su, "CNN and GAN-based classification of malicious code families: A code visualization approach," *Int. J. Intell. Syst.*, vol. 37, no. 12, pp. 12472–12489, 2022. doi: [10.1002/int.23094](https://doi.org/10.1002/int.23094).
- [36] L. Zhang, J. Wang, W. Wang, Z. Jin, and C. Zhao, "A novel smart contract vulnerability detection method based on information graph and ensemble learning," *Sens.*, vol. 22, no. 9, pp. 3581–3606, 2022. doi: [10.3390/s22093581](https://doi.org/10.3390/s22093581).
- [37] B. Bross *et al.*, "Overview of the versatile video coding (VVC) standard and its applications," *IEEE Trans. Circ. Syst. Video Technol.*, vol. 31, no. 10, pp. 3736–3764, 2021. doi: [10.1109/TCSVT.2021.3101953](https://doi.org/10.1109/TCSVT.2021.3101953).