

Article

# Efficient Video Steganalytic Feature Design by Exploiting Local Optimality and Lagrangian Cost Quotient

Ying Liu <sup>1</sup>, Jiangqun Ni <sup>2,3,\*</sup> and Wenkang Su <sup>2</sup><sup>1</sup> School of Electronics and Information Engineering, Sun Yat-sen University, Guangzhou 510006, China<sup>2</sup> School of Computer Science and Engineering, Sun Yat-sen University, Guangzhou 510006, China<sup>3</sup> Cyberspace Security Research Center, Peng Cheng Laboratory, Shenzhen 518000, China

\* Correspondence: issjqni@mail.sysu.edu.cn

**Abstract:** As the opponent of motion vector (MV)-based video steganography, the corresponding symmetric steganalysis has also developed a lot in recent years, among which the logic-based steganalytic schemes, e.g., AoSO, NPELO and MVC, are the most prevailing. Although currently achieving the best detection performance, these steganalytic schemes are less effective in detecting some logic-maintaining steganographic schemes. In view of the fact that the distributions of covers' local Lagrangian cost quotients are normally more concentrated in the small value ranges than those of stegos and "spread" to the large values ranges after modifying the motion vector, the local Lagrangian cost quotient would thus be an efficient indicator to reflect the difference between cover videos and stego ones. In this regard, combining the logic-based (Lg) and local Lagrangian cost quotient (LLCQ)-based feature, we finally proposed a more effective and general steganalysis feature, i.e., Lg-LLCQ, which is composed of diverse subfeatures and performs much better than the corresponding single-type feature. Extensive experimental results show that the proposed method exhibits detection performance superior to other state-of-the-art schemes and even works well under cover sources and steganographic scheme mismatch scenes, which indicates our proposed feature is more conducive to real-world applications.

**Keywords:** video steganalysis; video steganography; motion vector; local Lagrangian cost quotient



**Citation:** Liu, Y.; Ni, J.; Su, W. Efficient Video Steganalytic Feature Design by Exploiting Local Optimality and Lagrangian Cost Quotient. *Symmetry* **2023**, *15*, 520. <https://doi.org/10.3390/sym15020520>

Academic Editor: Vyacheslav Yukalov

Received: 6 January 2023

Revised: 29 January 2023

Accepted: 31 January 2023

Published: 15 February 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Steganography is the science and art of covert communication by slightly modifying the data in digital media, such as image, audio and video, without drawing any suspicion. In the past decade, many efforts have been dedicated to improving the security of image steganography [1–6], while less attention has been attached to video steganography. On the other hand, with the rapid development of electronic and multimedia technology, video is gradually being widely used in our daily lives, especially with the emergence of the H.264/AVC video format. Compared with other traditional cover medias, H.264/AVC videos have rich compression pipelines in terms of cover elements. Additionally, according to the difference on the selection of cover elements, currently, video steganography can be divided into the following categories: motion vector (MV)-based [7–17], quantized DCT coefficients based [18–26], partition modes based [27–32] and quantization-parameters-based [33–35], among which, motion vector (MV)-based H.264/AVC video steganography is currently the most prevalent.

Generally speaking, MV-based steganographic schemes can be further categorized into two popular types: heuristic schemes [7–9] and content-adaptive ones [10–17]. For heuristic MV-based video steganography, secret messages are usually embedded into the candidate MVs through some predefined selection rules, but these rules tend to increase its potential security risks. Recently, with the emergence of the framework of minimal distortion embedding, content-adaptive MV-based video steganographic schemes have

been gradually proposed, which mainly focuses on the design of efficient distortion cost. For example, Cao et al. [11] assigned a cost to each MV by exploiting its probabilities of satisfying the optimal criteria, and the secret messages were embedded into MVs by using WPCs (Wet Paper Codes) and STCs (Syndrome-Trellis Codes) [36]. More recently, Zhang et al. [12] proposed another video steganographic scheme called MVMPLO (Motion Vector Modification with Preserved Local Optimality) to preserve the local optimality in SAD (Sum of Absolute Differences) sense. Moreover, in order to maintain the statistical distribution of MV after embedding modifications, Yao et al. [13] suggested modifying the MVs with slight changes in MV distributions and prediction errors.

To cope with the abuse of MV-based steganography, considerable progress has been made in developing symmetric MV-based steganalytic features [37–49], including statistic-based features, calibration-based features and logic-based features. With regard to the statistic-based features, they are usually constructed based on the distribution of noise residuals calculated from adjacent MVs, with the assumption that the embedding modifications are additive independent. For instance, Su et al. [37] utilized the statistical characteristics of neighboring MV differences to construct steganalytic features. Moreover, Tasdemir et al. [38] proposed a spatio-temporal rich model-based steganalytic feature, which was built from the MV residuals filtered by a series of high-pass filters.

As for calibration-based features, they are usually constructed by the histogram of the difference between original MVs and corresponding recompressed ones. Following this way, Cao et al. [39] proposed a calibrated feature by recompressing the H.264/AVC video to improve detection performance. Moreover, in response to the mismatching of motion estimation of cover and stego video after recompression, Wang et al. [40] recently proposed an improved version by predicting the motion estimation before calibration.

It is well known that the MVs of cover videos are mostly locally optimal and inconsistent, but the steganographic embedding modification usually breaks these characteristics. Motivated by this defect, a kind of logic-based feature constructed from the logical probabilities of MVs subsequently appeared. Typically, Wang et al. [41] constructed an AoSO (Add-or-Subtract-One) feature by checking whether an MV is locally optimal in SAD sense, which is the first logic-based feature. Based on this, Zhang et al. [42] further utilized the Lagrangian cost function to check the local optimality and then proposed an enhanced feature called NPELO (Near-Perfect Estimation for Local Optimality). Moreover, in response to the case that the originally different MVs of the sub-blocks in the same macroblock tend to be consistent after embedding modifications, Zhai et al. [43] proposed a more powerful feature called MVC (Motion Vector Consistency), which exhibits better detection performance in most cases.

It should be noted that statistic-based features require the block size to be fixed, but the fixed block size option has already been abandoned in the current practical video codings, e.g., H.264/AVC and H.265/HEVC, and uses variable block size instead; therefore, they can not be applied in the detection of variable block-size-based video steganography. As for calibration-based features, both the way of motion estimation and the coding parameters are required to be kept the same before and after calibration, otherwise, the detection performance will be degraded. Although they perform much more effectively than the statistic-based and calibration-based features, logic-based features still have the risk of being defeated by some target steganographic schemes [11,12,16]. Given all this, it is desirable to construct an effective and general steganalytic feature for MV-based H.264/AVC video steganography. In this paper, we propose a novel steganalytic feature composed of diverse subfeatures for existing MV-based steganographic schemes. This work is mainly motivated by the numerical anomaly of local Lagrangian cost quotient introduced by the embedding modifications and, based on which, a local Lagrangian cost quotient (LLCQ)-based feature is proposed. Moreover, for further improving its detection performance, the previous efficient logic-based feature (Lg) is then introduced and integrated into the LLCQ, thereby forming a more powerful steganalytic feature, i.e., Lg-LLCQ. Experimental results show

that our proposed Lg-LLCQ is much more effective and general compared with the existing steganalytic schemes. The main contributions of this paper are summarized as follows.

(1) The effects of the MV modifications on the statistic characteristics of local Lagrangian cost quotient are evaluated, and the steganalytic performances of LLCQ are provided, all of which indicate the effectiveness of the proposed LLCQ feature.

(2) The logic-based (Lg) feature and local Lagrangian cost quotient (LLCQ)-based features are merged for steganalysis, which provides better results than the corresponding single-type feature.

The rest of the paper is structured as follows. In Section 2, Lagrangian cost functions in motion estimation and local optimality for video steganalysis of H.264 video are briefly described. The construction of our proposed steganalytic feature is presented in Section 3, followed by the experimental results and analysis in Section 4.1 for its verification of reasonability and feasibility. Finally, the paper is concluded in Section 5.

## 2. Preliminaries

### 2.1. Several Key Notions in H.264/AVC

**Motion estimation:** Motion estimation is to search for the best matching block in a previously coded reference frame for the current block based on a certain matching criterion within a given search area. **Motion vector:** Motion vector is the relative displacement between the current block and the matching block. **Motion compensation:** **Motion compensation is achieved by subtracting the matching block from the current block to obtain the residual block. Obviously, the motion vector is the product of motion estimation. Additionally, the motion vector and the best matching block can be used in motion compensation to reduce the interframe redundancy.**

### 2.2. Lagrangian Cost Function in Motion Estimation

Motion estimation is one of the most important coding options of video encoders, which can efficiently remove interframe redundancies. During the process of motion estimation, the bitrate and distortion of each macroblock in a search window is calculated, by which the selection of MV can be subsequently determined based on minimal distortion. It should be noted that the distortion is inversely proportional to the bitrate, i.e., the smaller the resultant distortion, the larger the corresponding bitrate. To achieve the trade-off between distortion and bitrate, some effective Lagrangian optimization techniques for rate-distortion optimization are adopted. Specifically, motion estimation is achieved by searching for the best matching block in the previously coded reference frame based on a rate-distortion criterion. The criterion is usually conducted by minimizing the following Lagrangian cost function [50]:

$$\mathbf{mv}_i^* = \arg \min_{\mathbf{mv}_i \in \Omega} \{ \Psi(S_{\text{org}}, S_{\mathbf{mv}_i}) + \lambda_{\text{motion}} \cdot R_{\text{motion}}(\mathbf{mv}_i, \text{ref\_idx}) \}, \quad (1)$$

where  $\mathbf{mv}_i^*$  represents the best matching MV in the search space  $\Omega$ ,  $\Psi(S_{\text{org}}, S_{\mathbf{mv}_i})$  is the distortion obtained by calculating the prediction error between the original block  $S_{\text{org}}$  and corresponding prediction block  $S_{\mathbf{mv}_i}$  indexed by  $i$ th  $\mathbf{mv}_i$  of the cover video and  $R_{\text{motion}}(\mathbf{mv}_i, \text{ref\_idx})$  represents the total number of bits required for coding  $\mathbf{mv}_i$ .  $\text{ref\_idx}$  stands for the index of reference frame utilized in multiple reference frames motion estimation.  $\lambda_{\text{motion}}$  is a weighting factor, which can be obtained by calculating an empirical formula [50] with a given QP (Quantization Parameter).

$$\lambda_{\text{motion}} = \sqrt{0.85 \times 2^{(QP-12)/3}}, \quad (QP \in [0, 51]). \quad (2)$$

As for the choice of distortion measures between the original frame and the corresponding predicted frame, there are two commonly used methods, i.e., SAD (Sum of Absolute Differences) and SATD (Sum of Absolute Transformed Differences), wherein SAD is employed as a pixelwise distortion measure, while SATD is employed as a sub-pixel-

wise one in rate-distortion optimization. Specifically, the pixelwise distortion SAD [42] is given by

$$SAD = \sum_m \sum_n |S_{org}(m, n) - S_{mv_i}(m, n)|, \tag{3}$$

where  $S_{org}(m, n)$  and  $S_{mv_i}(m, n)$  denote the  $(m, n)$  th pixel sample in original coding unit  $S_{org}$  and prediction unit  $S_{mv_i}$ , respectively. Additionally, the sub-pixel-wise distortion SATD [42] is given by

$$SATD = \sum_m \sum_n |T_{4 \times 4}(m, n)|, \tag{4}$$

where  $T_{4 \times 4}$  is the  $4 \times 4$  block obtained by applying Hadamard transform on the prediction error between coding unit  $S_{org}$  and  $S_{mv_i}$ .

In order to quickly calculate the number of bits used to encode and transmit the MVs, i.e.,  $R_{motion}(\mathbf{mv}_i)$ , the Exp-Golomb coding is employed in H.264/AVC to encode the MV difference. Specifically, for a given  $\mathbf{mv}_i$ , we first obtain the corresponding MV difference and denote it as  $\mathbf{mvd} = (mvd_x, mvd_y)$ , then calculate the Exp-Golomb indices for each MV difference component according to the following mapping rule [51]:

$$CodeNum_d = \begin{cases} 2|d|, & d \leq 0 \\ 2|d| - 1, & d > 0 \end{cases} \tag{5}$$

where  $d \in \{mvd_x, mvd_y\}$ . According to the Exp-Golomb indices, the number of bits required to encode and transmit  $\mathbf{mv}_i$  is finally defined as [51]

$$R_{motion}(\mathbf{mv}_i) = 2 \lfloor \log_2(CodeNum_{mvd_x} + 1) \rfloor + 2 \lfloor \log_2(CodeNum_{mvd_y} + 1) \rfloor + 2, \tag{6}$$

where  $CodeNum_{mvd_x}$  and  $CodeNum_{mvd_y}$  denote the Exp-Golomb indices of  $mvd_x$  and  $mvd_y$ , respectively.

### 2.3. Local Optimality for Video Steganalysis

For H.264/AVC cover videos, most of the MVs still remain locally optimal after compression. Specifically, an MV  $\mathbf{mv}_i = (x, y)$  is said to have local optimality with respect to its adjacent MVs, i.e.,  $\Omega(\mathbf{mv}_i) = \{\mathbf{mv}'_i = (x + \Delta x, y + \Delta y) | \Delta x, \Delta y \in \{-1, 0, 1\}\}$ , if, for any  $\mathbf{mv}'_i \in \Omega(\mathbf{mv}_i)$ , there is [42]

$$J_{motion}^{\Psi}(\mathbf{mv}_i) \leq J_{motion}^{\Psi}(\mathbf{mv}'_i), \tag{7}$$

with the Lagrangian cost function given by

$$J_{motion}^{\Psi}(\mathbf{mv}'_i) = \Psi(S_{rec}, S_{mv'_i}) + \lambda_{motion} \cdot R_{motion}(\mathbf{mv}'_i), \tag{8}$$

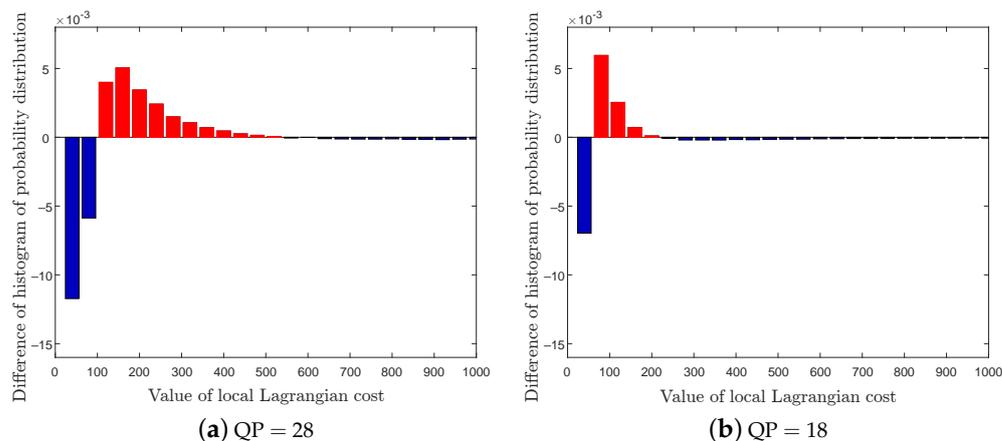
where  $\Psi(S_{rec}, S_{mv'_i})$  represents the prediction error between the reconstructed block  $S_{rec}$  and the corresponding prediction block  $S_{mv'_i}$  measured as SAD or SATD.

It should be noted that due to the influence of quantization in the encoding process, the local optimality of the MV in the encoder may not always be fully maintained in the decoder, by which the newly proposed MV-based steganographic schemes [11,12] can well evade the detection of the previous local optimality feature AoSO [41]. In view of this defect, an improved local optimality feature called NPELO is proposed in [42]. In NPELO, the Lagrangian cost-based criterion is used to check whether an MV is locally optimal. In addition, to more accurately distinguish the local optimal MVs in cover videos from the ones in stego videos, an exponentially magnified relative difference between current Lagrangian cost and minimum Lagrangian cost is further integrated into the NPELO feature, thereby forming a final 36D steganalytic feature. Although the enhanced local optimality feature NPELO can compensate for this defect, the risk of being targeted attacked still exists.

### 3. The Proposed Steganalytic Feature for H.264/AVC Video

#### 3.1. The Statistic Characteristics of Local Lagrangian Cost

As mentioned in Section 2.3, most MVs are locally optimal in cover videos, i.e.,  $J_{motion}^{\Psi}(\mathbf{mv}_i) \leq J_{motion}^{\Psi}(\mathbf{mv}'_i)$ , but the local optimality of MVs in stego videos is likely to be destroyed by embedding modifications, thereby causing numerical anomaly of local Lagrangian cost. To verify this, we performed an analytical experiment to compare local Lagrangian cost for cover and stego videos. In this experiment, the video sequences with resolution  $352 \times 288$  from the database provided in [52] are the original cover videos, and they are all encoded by JM 19.0 reference software [53]. Then, we use the steganographic scheme MVMPLO to generate the corresponding stego videos with relative payload 0.4 bp-nsmv (bits per nonskip motion vector). For ease of intuitively observing the changes in statistical characteristics after embedding modifications, we first separately count the histogram of the probability distribution of local Lagrangian cost on cover and stego video datasets under QP = 28 and QP = 18 (different QPs represent different compression efficiency, and the larger the QP, the lower the corresponding bitrate), and then calculate the histogram differences between the cover and the corresponding stego, as shown in Figure 1. From Figure 1a,b, we can find that the embedding modifications on the video with QP = 28 will arise more and richer histogram differences, also known as stronger steganographic traces, than those of QP = 18. Specifically, the apparent histogram difference at QP = 28 lies within the local Lagrangian cost of [0, 500], while the ones at QP = 18 lie within the range of [0, 200]. Therefore, we can assert that the local Lagrangian cost can effectively differentiate cover videos from stego videos.



**Figure 1.** The difference of histogram of probability distribution for local Lagrangian cost between cover videos and stego videos. (a,b) are the difference of the histogram of probability distribution for local Lagrangian cost under QP = 28 and QP = 18, respectively. Histogram bins with values greater than 0 are shown in red, while those less than 0 are shown in blue.

#### 3.2. The Proposed Feature

The statistical properties of local Lagrangian cost are influenced by video content and movement (e.g., larger values occur in fast-moving and rough regions, while smaller values are for slow-moving and smooth regions). In order to decrease these influences and improve the stability of steganalytic features, we proposed a new indicator, i.e., the local Lagrangian cost quotient (LLCQ), to reflect the difference between cover videos and stego ones. The advantage of LLCQ as compared with the local Lagrangian cost is that the video content is largely suppressed, which has a much narrower dynamic range, thus allowing a more stable statistical description. Moreover, the previous efficient logic-based feature (Lg) is then integrated into the LLCQ, thereby forming a more powerful steganalytic feature, i.e., a logic- and LLCQ-based feature (Lg-LLCQ), which is composed of five types of subfeatures. For ease of description, the specific procedure of feature extraction is sketched

in Figure 2. In a H.264/AVC video with a GOP size of  $K$ , assuming that each GOP contains  $N$  MVs, for each MV  $\mathbf{mv}_i$  ( $i \in [1, N]$ ) in a given GOP, we first obtain the neighborhood set  $\Omega(\mathbf{mv}_i)$ , then calculate  $J_{motion}^{SAD}(\mathbf{mv}_i^k)$  and  $J_{motion}^{SATD}(\mathbf{mv}_i^k)$  associated with the  $k$ th ( $k \in [1, 9]$ ) MV in  $\Omega(\mathbf{mv}_i)$ , according to Figure 3. Thereafter, the five types of subfeatures in Lg-LLCQ are separately constructed as follows.

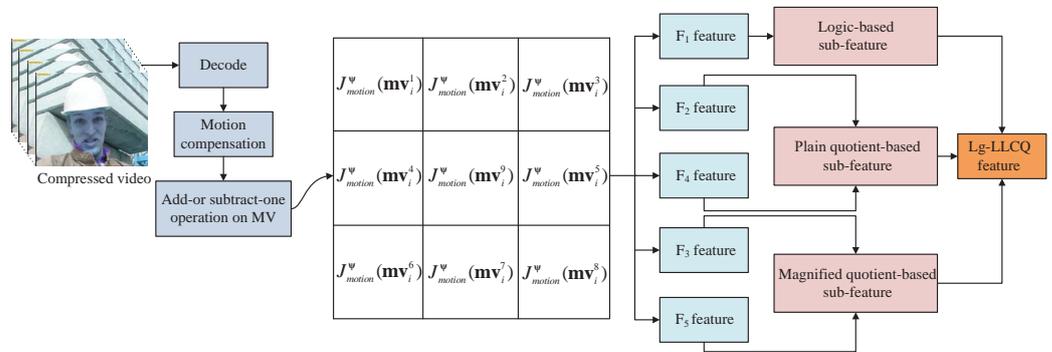


Figure 2. The extraction procedure diagram of the proposed Lg-LLCQ feature.

$\mathbf{mv}_i^1$ ( $x-1, y-1$ )	$\mathbf{mv}_i^2$ ( $x, y-1$ )	$\mathbf{mv}_i^3$ ( $x+1, y-1$ )
$\mathbf{mv}_i^4$ ( $x-1, y$ )	$\mathbf{mv}_i^9$ ( $x, y$ )	$\mathbf{mv}_i^5$ ( $x+1, y$ )
$\mathbf{mv}_i^6$ ( $x-1, y+1$ )	$\mathbf{mv}_i^7$ ( $x, y+1$ )	$\mathbf{mv}_i^8$ ( $x+1, y+1$ )

Figure 3. The specific relationship of neighborhood position of MVs.

### 3.2.1. Feature of Type 1

The feature of type 1, denoted as  $F_1$ , is the NPELO feature [42], which consists of two subfeatures, i.e., SAD-based and SATD-based subfeature. Additionally, the SAD-based subfeature is obtained by checking the local optimality of MVs by computing  $J_{motion}^{SAD}(\mathbf{mv}_i^k)$ , i.e.,

$$\begin{aligned}
 F_1(k) &= P(J_{motion}^{SAD}(\mathbf{mv}_i^k) = J_{\min}^{SAD}(\Omega(\mathbf{mv}_i))) \\
 &= \frac{1}{N} \sum_{i=1}^N \delta(J_{motion}^{SAD}(\mathbf{mv}_i^k), J_{\min}^{SAD}(\Omega(\mathbf{mv}_i))), k \in [1, 9],
 \end{aligned} \tag{9}$$

$$\begin{aligned}
 F_1(k+9) &= \frac{1}{\beta} \sum_{i=1}^N \exp\left\{ \frac{|J_{\min}^{SAD}(\Omega(\mathbf{mv}_i)) - J_{motion}^{SAD}(\mathbf{mv}_i^k)|}{J_{motion}^{SAD}(\mathbf{mv}_i^k)} \right\} \\
 &\quad \cdot \delta(J_{motion}^{SAD}(\mathbf{mv}_i^k), J_{\min}^{SAD}(\Omega(\mathbf{mv}_i))), k \in [1, 9],
 \end{aligned} \tag{10}$$

where  $\beta$  denotes the normalization factor and is given by

$$\beta = \sum_{k=1}^9 \sum_{i=1}^N \exp \left\{ \frac{|J_{\min}^{\text{SAD}}(\Omega(\mathbf{mv}_i)) - J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i)|}{J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i)} \right\} \cdot \delta(J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^k), J_{\min}^{\text{SAD}}(\Omega(\mathbf{mv}_i))), k \in [1, 9].$$

Additionally,  $J_{\min}^{\text{SAD}}(\Omega(\mathbf{mv}_i))$  denotes the minimum value of an element in the set  $\{J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^k) | \mathbf{mv}_i^k \in \Omega(\mathbf{mv}_i)\}$ ,  $\delta(a, b)$  is the Kronecker delta function, with  $\delta$  being 1 if  $a$  equals  $b$  and 0 otherwise, and  $N$  is the MV number in the GOP. Since the construction of the SATD-based subfeature is similar to the SAD-based subfeature, for simplicity, we do not give further details on the construction of the SATD-based subfeature (see [42] if needed).

### 3.2.2. Feature of Type 2

The feature of type 2 is associated with the plain quotient between  $J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^9)$  and  $J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^k)$ , with a given relative spatial position  $k$ , i.e.,

$$F_2(k) = \frac{1}{N} \sum_{i=1}^N \frac{J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^9)}{J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^k) + \alpha}, k \in [1, 8], \tag{11}$$

where  $\alpha$  is a constant introduced to avoid dividing by zero and set as 1.0 in our implementation.

### 3.2.3. Feature of Type 3

The feature of type 3 corresponds to the exponentially magnified quotient between  $J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^9)$  and  $J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^k)$ , with a given relative spatial position  $k$ , i.e.,

$$F_3(k) = \frac{1}{N} \sum_{i=1}^N \left( \frac{J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^9)}{J_{\text{motion}}^{\text{SAD}}(\mathbf{mv}_i^k) + \alpha} \right)^p, k \in [1, 8], \tag{12}$$

where  $p$  is a positive integer not less than 1.

### 3.2.4. Feature of Type 4

The feature of type 4 is associated with the plain quotient between  $J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^9)$  and  $J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^k)$ , with a given relative spatial position  $k$ , i.e.,

$$F_4(k) = \frac{1}{N} \sum_{i=1}^N \frac{J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^9)}{J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^k) + \alpha}, k \in [1, 8]. \tag{13}$$

### 3.2.5. Feature of Type 5

The feature of type 5 corresponds to the exponentially magnified quotient between  $J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^9)$  and  $J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^k)$ , with a given relative spatial position  $k$ , i.e.,

$$F_5(k) = \frac{1}{N} \sum_{i=1}^N \left( \frac{J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^9)}{J_{\text{motion}}^{\text{SATD}}(\mathbf{mv}_i^k) + \alpha} \right)^p, k \in [1, 8]. \tag{14}$$

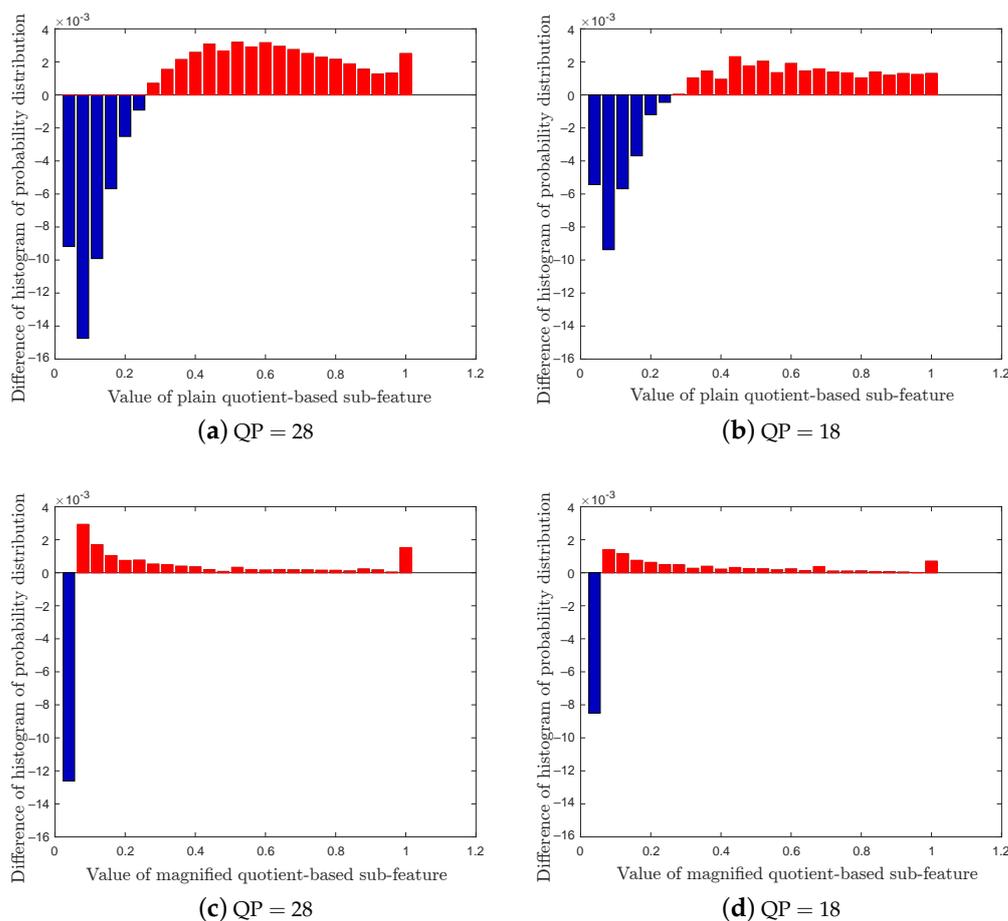
## 3.3. The Final Joint Feature

Combining all these subfeatures together, a final 68-dimensional-feature Lg-LLCQ is defined as

$$F(k) = \begin{cases} F_1(k), & k \in [1, 36] \\ F_2(k - 36), & k \in [37, 44] \\ F_3(k - 44), & k \in [45, 52], \\ F_4(k - 52), & k \in [53, 60] \\ F_5(k - 60), & k \in [61, 68] \end{cases} \tag{15}$$

where  $F_1$  is the logic-based subfeature, and  $F_2 \sim F_5$  are the LLCQ-based subfeatures. For ease of expression in the following, the  $F_1$  feature is denoted as “logic-based subfeature”, the  $F_2/F_4$  features are denoted as “plain quotient-based subfeature” and  $F_3/F_5$  features are denoted as “magnified quotient-based subfeature”.

Figure 4 shows the LLCQ feature of cover and stego videos under  $QP = 28$  and  $QP = 18$ , and the stego videos are produced by steganographic scheme MVMPLO with relative payload 0.4 bpnsmv. Referring to the results in Figure 4a,b, the difference of the histogram of probability distribution for the plain quotient-based subfeature shows that the plain quotient-based subfeature of cover videos are normally more concentrated in the value ranges less than 0.3; it will be “spread” from the small value ranges to the large value ranges in stego videos after embedding modifications. Figure 4c,d shows magnified quotient-based subfeatures of cover videos are normally more concentrated in the value ranges close to 0, and the embedding modifications will magnify some subtle distinctions between cover videos and stego ones. In a word, Figure 4 indicates that the two kinds of subfeatures in LLCQ are significantly different between cover videos and stego ones. Given these apparent differences in steganalytic feature statistics between cover videos and stego ones, we can assert that the proposed LLCQ feature will be feasible for detecting the MV-based steganography.



**Figure 4.** The difference of the histogram of probability distribution for plain quotient-based and magnified quotient-based subfeatures between cover videos and stego videos. (a,b) are the difference of the histogram of probability distribution for plain quotient-based subfeatures under  $QP = 28$  and  $QP = 18$ , respectively. (c,d) are the difference of the histogram of probability distribution for magnified quotient-based subfeatures under  $QP = 28$  and  $QP = 18$ , respectively. Histogram bins with values greater than 0 are shown in red, while those less than 0 are shown in blue.

## 4. Experiments

In this section, extensive experiments are conducted to verify the effectiveness, generalization and stability of the proposed scheme. We compare the detection performance of our proposed Lg-LLCQ feature with three other state-of-the-art steganalytic features, i.e., AoSO, NPELO and MVC, under various application scenes. Specifically, seven kinds of setups are carefully constructed in the following. Setup 1 is the basic test that runs to determine the constant exponential parameter  $p$ , and Setup 2 is first performed to test the effectiveness of the LLCQ feature and then is constructed to further verify the effectiveness of the proposed combined Lg-LLCQ feature. The stability of the Lg-LLCQ feature is assessed under Setup 3, and the independent effects of the subfeatures in Lg-LLCQ, i.e., logic-based subfeature, plain quotient-based subfeature and magnified quotient-based subfeature, are carefully investigated and compared under Setup 4. The experiment under Setup 5 is mainly to show our method can be directly applied to the previous video coding standards. Setup 6 is carried out to test whether the proposed feature is still effective under cover sources and steganographic scheme mismatch scenes. Lastly, and most importantly, a well-designed experiment is conducted to evaluate the steganalytic performance of the proposed Lg-LLCQ feature under different video resolutions. Additionally, bold in the tables indicates the best performance under the given settings.

### 4.1. Experiment Setups

#### 4.1.1. Datasets

In order to comprehensively compare the detection performance of different steganalytic features under different scenes, three public video datasets are introduced for experiments, i.e., DB1 [54], DB2 [52] and DB3 [54]. DB1, DB2 and DB3 all consist of 100 video sequences. Each sequence in DB1 and DB3 includes 100 frames on average, while the sequence in DB2 includes 220 frames on average. Moreover, the video sequences in DB1 and DB2 are stored in the 4:2:0 color sampling format with the same resolution, i.e.,  $352 \times 288$ , and the video sequences in DB3 are stored in the 4:2:0 color sampling format with resolution, i.e.,  $176 \times 144$ .

#### 4.1.2. Steganographic Schemes

To evaluate the effectiveness, generalization and stability of our proposed steganalytic feature, three advanced MV-based steganographic schemes are introduced for experiments. One of them is the conventional scheme proposed by Aly [9] (denoted as Tar1), the other two are content-adaptive embedding schemes, i.e., MVMPL0 [12] (denoted as Tar2) and Cao et al. [11] (denoted as Tar3). For fair comparison, the payload is depicted by a relative indicator, i.e., bpnsmv, also called relative payload, which is the ratio of the number of embedded bits to the total number of nonskip MVs, and all the involved steganographic schemes are implemented using JM 19.0 reference software [53]. It should be noted that the payload is ranging from  $0.1 \text{ bpnsmv} + M$  to  $0.5 \text{ bpnsmv} + M$  in Tar3, where  $M$  is an additional  $M$ -bit in the second channel of Tar3.

#### 4.1.3. Setups for Performance Evaluation

To better test the feasibility of the proposed steganalytic feature in practical application, six kinds of setups are elaborately constructed in this paper.

**Setup 1:** In this setup, the steganalytic performance and the average time consumption of feature extraction are evaluated under different parameter  $p \in \{5, 10, 15, 20, 25\}$ . All these experiments are conducted on DB1 by using Tar2 at payload 0.4 bpnsmv under QP = 28 and QP = 18, and the experiment environment is Visual Studio 2013 on a 3.0 GHz Intel Core E5-2653 CPU with 128 GB memory.

**Setup 2:** Two kinds of experiments are conducted in this setup. One aims at testing the effectiveness of the proposed LLCQ feature in detecting Tar2 on DB1 and DB2 under QP = 28 and QP = 18. The other is performed to further evaluate the steganalytic performance of the combined feature Lg-LLCQ against involved steganographic schemes with

payloads 0.1~0.5 bpnsmv under QP = 28 and QP = 18. Additionally, each video sample in these experiments is coded with *BaselineProfile*, under which the prediction structure is IPPPPPPPPPIPP . . . , i.e., one I-slice followed by nine P slices with one reference frame.

**Setup 3:** In practical applications of steganalysis, the stego videos may contain various payloads and QPs. In order to evaluate the stability of our proposed features against this circumstance, we then mix the stego videos obtained under setup 2 for experiments.

**Setup 4:** For evaluating the independent effects of various subfeatures of the proposed Lg-LLCQ, the logic-based subfeature, plain quotient-based subfeature and magnified quotient-based subfeatures are separately employed to detect Tar2 and Tar3, wherein the stego videos are obtained under setup 2 at payloads 0.2~0.4 bpnsmv.

**Setup 5:** To evaluate the applicability of the proposed Lg-LLCQ under fixed block size and different QPs, the size of all coding units is taken as  $16 \times 16$ , and the QPs are also set as 28 and 18. The proposed Lg-LLCQ and other involved steganalytic features are all employed to detect Tar2 at payloads 0.2 and 0.4 bpnsmv under the same coding configurations in setup 2.

**Setup 6:** The mismatch of cover sources and steganographic schemes is generally considered to be the most influential factor limiting the application of steganalysis in the real world. In this setup, to simulate the real detection scene, an experiment is performed to evaluate the applicability of our proposed feature by training and testing on different cover sources and steganographic schemes. To differentiate various steganalytic and steganographic schemes, in this part, we use the syntax of names following the convention:  $name = \{steganalytic\ scheme\} - \{steganographic\ scheme\}$ , where  $\{steganalytic\ scheme\}$  indicates steganalytic scheme, i.e., AoSO, NPELO, MVC and Lg-LLCQ, and  $\{steganographic\ scheme\}$  is the steganographic scheme, i.e., Tar2 and Tar3.

**Setup 7:** To evaluate the steganalytic performance of the proposed Lg-LLCQ on DB3, the proposed Lg-LLCQ and other involved steganalytic features are all employed to detect Tar2~3 at payloads 0.1 and 0.3 bpnsmv under the same coding configurations in setup 2.

#### 4.1.4. Training and Classification

In our experiments, the features are extracted from each video sample. Throughout all the steganalytic experiments, sixty percent cover–stego pairs are randomly selected for training, while the remaining are randomly shuffled and sent into the SVM classifier one by one for testing. The penalty factor  $C$  and kernel factor  $\gamma$  in Gaussian-kernel SVM are optimized by a five-fold cross-validation on the grid space  $\langle (C, \gamma) | C = 2^{10}, 2^{11}, \dots, 2^{15}, \gamma = 2^{-15}, 2^{-14}, \dots, 2^3 \rangle$ . The detection performance will be quantified as the average accuracy  $\bar{P}_{ACC}$ , which is defined as

$$\bar{P}_{ACC} = 1 - \frac{1}{2}(P_{FA} + P_{MD}), \quad (16)$$

where  $\bar{P}_{ACC}$  is averaged over 50 iterations on each steganalytic experiment and  $P_{FA}$  and  $P_{MD}$  represent the probability of false alarm and missed detection, respectively.

#### 4.2. Performance Evaluation

##### 4.2.1. Evaluation of Computational Complexity and Steganalytic Performance of the Proposed Features under Different Parameter $p$

In this part, we evaluate the computational complexity and steganalytic performance of our proposed Lg-LLCQ in terms of average computation time (Ave-Time) and average accuracy ( $\bar{P}_{ACC}$ ) under different exponential parameter  $p$ , and the corresponding results are listed in Table 1. As can be seen from Table 1, an increasing  $p$  does not always significantly improve detection performance, but negatively affects Ave-Time. In this regard, for the trade-off between Ave-Time and steganalytic performance, we set  $p$  as 10 in this paper.

### 4.2.2. Steganalytic Performance Comparison

With the determination of parameter  $p$ , we can then compare the detection performance of the proposed LLCQ with other state-of-the-art MV-based steganalytic schemes, i.e., AoSO, NPELO and MVC, in detecting Tar2, and the corresponding results are shown in Table 2. As we can see from Table 2, the proposed LLCQ achieves remarkable performance improvement over NPELO, which proves the conclusion drawn in Section 3.3 that the LLCQ feature will be feasible for detecting the MV-based steganography. We then systematically compare the detection performance of the proposed combined feature, i.e., Lg-LLCQ, with involved steganalytic schemes at different payloads on DB1 and DB2, the corresponding results are summarized in Tables 3 and 4, respectively. To further evaluate the detection performance, Figures 5 and 6 give the ROC curves for the detection of the involved steganographic schemes on DB1 and DB2 at 0.3 bpnsmv. For Tar1~3, our proposed Lg-LLCQ exhibits excellent detection performance at various payloads under QP = 28 and QP = 18 and outperforms other competing methods, as shown in Tables 3 and 4 and Figures 5 and 6, indicating the effectiveness of the proposed Lg-LLCQ in detecting the involved steganographic schemes.

**Table 1.** The average time consumption (in seconds) and average accuracy  $\bar{P}_{ACC}$  (in %) of our proposed Lg-LLCQ under parameter  $p$ .

QP	Evaluating Scheme	Parameter $p$				
		5	10	15	20	25
28	Ave_Time	483	491	503	522	541
	$\bar{P}_{ACC}$	93.28	94.36	94.61	94.79	94.58
18	Ave_Time	561	578	593	609	621
	$\bar{P}_{ACC}$	93.61	95.45	95.67	95.84	95.67

**Table 2.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed LLCQ against Tar2 under various payloads (in bpnsmv) and QPs (28 and 18) on DB1 and DB2.

Dataset	QP	Feature	Payload (in Bpnsmv)				
			0.1	0.2	0.3	0.4	0.5
DB1	28	AoSO	50.48	51.49	52.18	53.04	56.38
		NPELO	61.02	72.12	79.28	85.21	88.34
		MVC	57.69	69.71	78.59	85.68	88.76
		LLCQ	<b>61.28</b>	<b>75.82</b>	<b>86.34</b>	<b>88.65</b>	<b>92.24</b>
	18	AoSO	51.87	54.02	60.01	64.13	72.31
		NPELO	63.21	71.47	82.18	88.24	91.45
		MVC	68.56	81.74	89.39	<b>94.71</b>	<b>96.46</b>
		LLCQ	<b>70.28</b>	<b>85.64</b>	<b>91.68</b>	94.15	96.32
DB2	28	AoSO	50.51	50.89	52.23	54.61	58.84
		NPELO	66.54	80.36	86.69	91.26	93.15
		MVC	65.34	79.61	87.26	92.48	94.23
		LLCQ	<b>68.25</b>	<b>85.66</b>	<b>91.32</b>	<b>93.47</b>	<b>95.28</b>
	18	AoSO	50.39	53.62	59.37	65.54	71.62
		NPELO	64.29	78.81	87.69	92.37	94.61
		MVC	<b>69.12</b>	<b>87.41</b>	<b>94.12</b>	<b>96.49</b>	<b>97.45</b>
		LLCQ	68.34	85.16	92.62	95.41	97.08

**Table 3.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed Lg-LLCQ against the steganographic schemes Tar1~3 under various payloads (in bpsnmv) and QPs (28 and 18) on DB1.

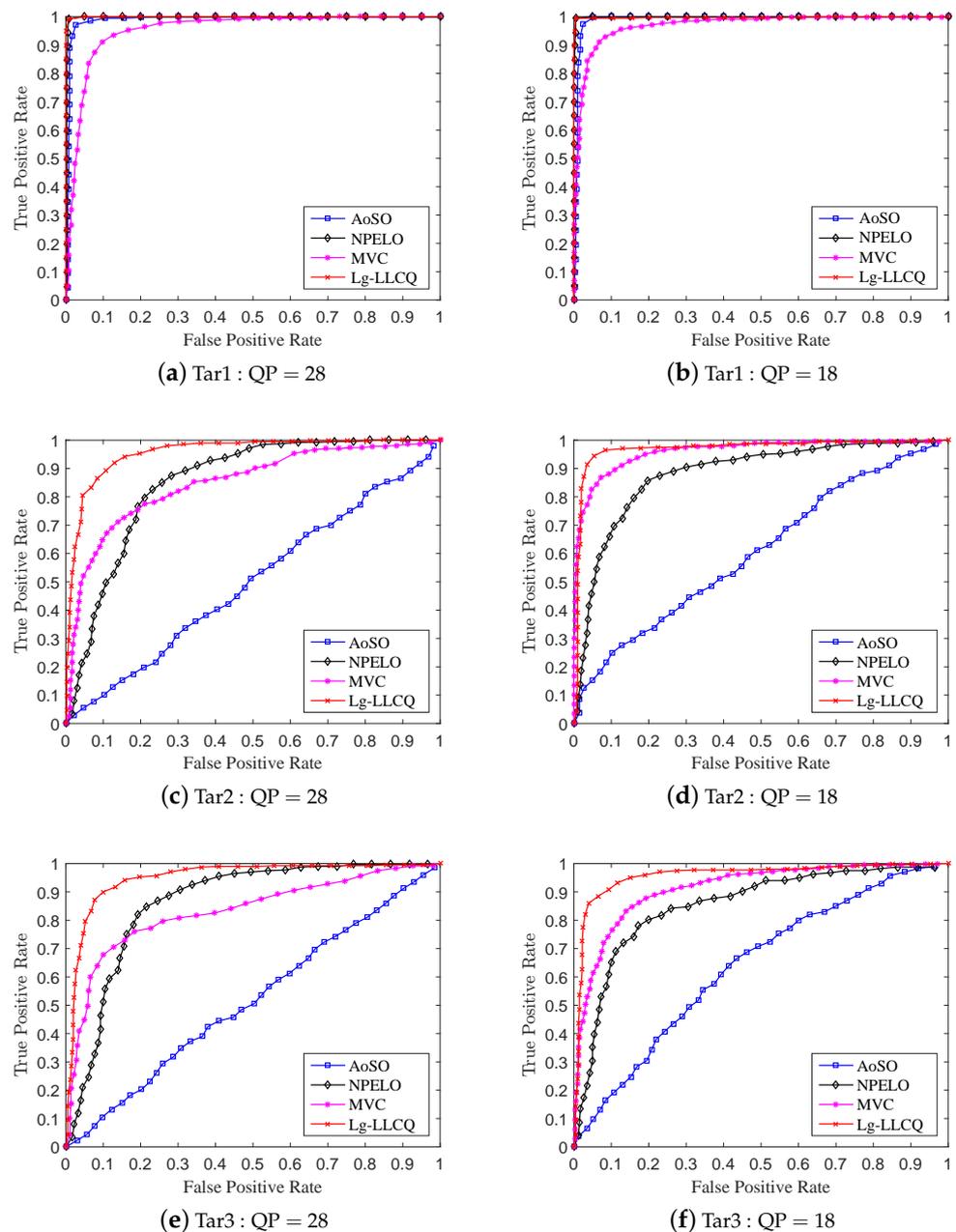
Scheme	Feature	Payload (QP = 28)					Payload (QP = 18)				
		0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
Tar1	AoSO	92.13	95.47	96.49	97.12	97.89	93.19	96.21	97.12	97.68	98.32
	NPELO	93.52	97.43	99.12	99.36	99.71	95.18	98.19	99.28	99.39	99.51
	MVC	78.46	83.89	89.65	92.15	94.48	82.69	89.17	93.78	95.66	97.18
	Lg-LLCQ	<b>95.54</b>	<b>98.34</b>	<b>99.32</b>	<b>99.52</b>	<b>99.76</b>	<b>97.29</b>	<b>98.86</b>	<b>99.33</b>	<b>99.45</b>	<b>99.82</b>
Tar2	AoSO	50.48	51.49	52.18	53.04	56.38	51.87	54.02	60.01	64.13	72.31
	NPELO	61.02	72.12	79.28	85.21	88.34	63.21	71.47	82.18	88.24	91.45
	MVC	57.69	69.71	78.59	85.68	88.76	68.56	81.74	89.39	94.71	96.46
	Lg-LLCQ	<b>62.14</b>	<b>78.96</b>	<b>89.34</b>	<b>94.36</b>	<b>95.87</b>	<b>71.87</b>	<b>89.18</b>	<b>94.21</b>	<b>95.45</b>	<b>97.39</b>
Tar3	AoSO	50.08	50.69	51.54	53.51	55.24	50.04	52.01	57.15	66.59	74.29
	NPELO	65.24	75.62	80.34	84.11	87.26	57.62	68.42	78.12	85.22	89.56
	MVC	58.14	72.16	79.23	84.87	88.34	61.66	76.53	85.33	91.36	95.21
	Lg-LLCQ	<b>69.14</b>	<b>80.77</b>	<b>88.72</b>	<b>92.23</b>	<b>94.34</b>	<b>69.03</b>	<b>84.76</b>	<b>89.36</b>	<b>95.21</b>	<b>96.81</b>

**Table 4.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed Lg-LLCQ against the steganographic schemes Tar1~3 under various payloads (in bpsnmv) and QPs (28 and 18) on DB2.

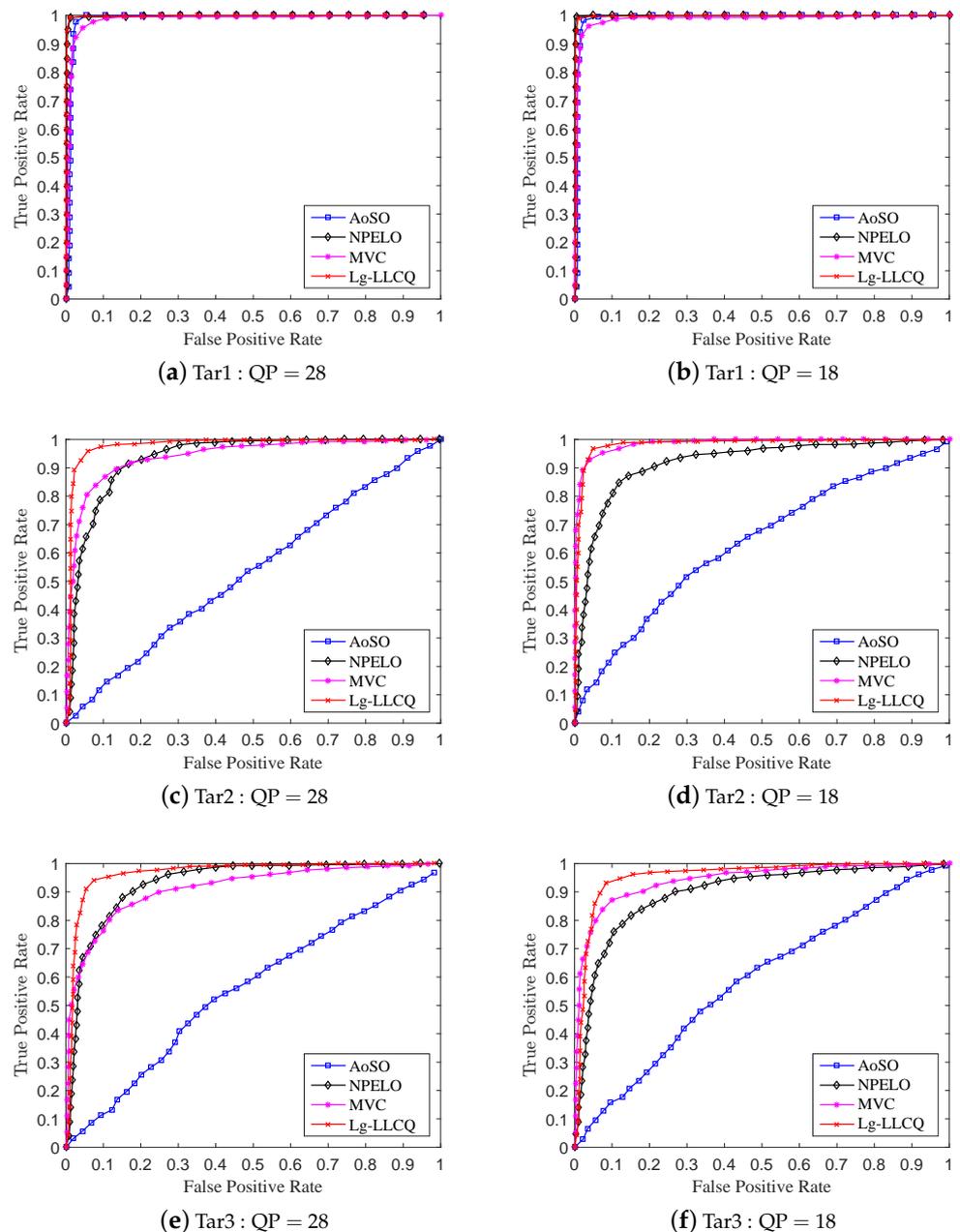
Scheme	Feature	Payload (QP = 28)					Payload (QP = 18)				
		0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
Tar1	AoSO	93.49	96.14	97.09	97.71	98.26	94.31	96.89	97.89	98.14	98.86
	NPELO	94.68	98.12	99.18	99.38	99.54	95.69	98.35	99.18	99.42	99.66
	MVC	85.63	89.14	94.36	95.15	97.69	88.79	93.88	95.79	97.16	98.84
	Lg-LLCQ	<b>95.79</b>	<b>98.61</b>	<b>99.22</b>	<b>99.46</b>	<b>99.61</b>	<b>97.42</b>	<b>98.63</b>	<b>99.36</b>	<b>99.53</b>	<b>99.72</b>
Tar2	AoSO	50.51	50.89	52.23	54.61	58.84	50.39	53.62	59.37	65.54	71.62
	NPELO	66.54	80.36	86.69	91.26	93.15	64.29	78.81	87.69	92.37	94.61
	MVC	65.34	79.61	87.26	92.48	94.23	69.12	87.41	94.12	96.49	97.45
	Lg-LLCQ	<b>69.71</b>	<b>88.03</b>	<b>94.43</b>	<b>97.36</b>	<b>98.27</b>	<b>70.38</b>	<b>89.13</b>	<b>96.02</b>	<b>97.61</b>	<b>98.86</b>
Tar3	AoSO	50.08	51.29	53.18	55.08	57.29	50.39	54.02	57.71	61.69	65.83
	NPELO	69.58	79.31	85.12	87.37	89.37	62.25	74.31	82.34	86.74	88.17
	MVC	63.49	76.21	83.67	88.21	91.69	65.86	79.71	88.37	92.08	93.68
	Lg-LLCQ	<b>75.39</b>	<b>86.41</b>	<b>90.39</b>	<b>92.48</b>	<b>94.26</b>	<b>72.43</b>	<b>85.68</b>	<b>91.51</b>	<b>93.73</b>	<b>95.27</b>

Referring to the results, we find that the tested steganalytic features show excellent performance in detecting Tar1. As for the detection on Tar2~3, the AoSO performs the worst. This is because Tar2~3 maintain the local optimality of MVs in SAD sense, which would have a dramatically negative influence on AoSO's steganalytic performance. The NPELO shows comparable performance under QP = 28 but broadly inferior performance under QP = 18 in detecting Tar2~3 as compared with MVC. The reason that contributed to the detection performance degradation under QP = 18 may be the way of MVs modifications. Specifically, the NPELO is built by checking the local optimality of MVs in rate-distortion sense. However, according to SAD-based modification criterion in Tar2~3, the stego MVs have certain probabilities to preserve the local optimality in rate-distortion sense. Thus, they lead the detection performance of NPELO degradation under QP = 18 as compared with MVC. Although the MVC feature shows overall better performance in detecting Tar2~3, it is still inferior to NPELO in some cases, i.e., some small payloads under QP = 28. The accepted internal mechanism is as follows. MVC is a steganalytic feature built by checking the consistency of MVs in small blocks. However, according to the principles of video encoding, when setting high QP in *BaselineProfile*, the encoder tends to choose larger blocks to encode videos. Additionally, the small blocks in the encoded videos will accordingly decrease, adversely affecting the performance of MVC. The NPELO is designed based on the observation that the embedding modifications will break the local optimality

of MVs. However, it is less effective in detecting some logic-maintaining steganographic schemes, e.g., Tar2~3. Unlike NPELO, in which only the logic-based feature is utilized to detect the MV modifications, our proposed Lg-LLCQ takes into account both the logic-based feature and local Lagrangian cost quotient (LLCQ)-based feature for steganalysis, which may contribute to its substantial performance improvement over NPELO and overall performance improvement as compared with MVC in detecting Tar1~3 under QP = 28 and QP = 18.



**Figure 5.** ROC curves of AoSO, NPELO, MVC and Lg-LLCQ on DB1 at 0.3 bpns/mv. (a) Tar1 under QP = 28; (b) Tar1 under QP = 18; (c) Tar2 under QP = 28; (d) Tar2 under QP = 18; (e) Tar3 under QP = 28; and (f) Tar3 under QP = 18.



**Figure 6.** ROC curves of AoSO, NPELO, MVC and Lg-LLCQ on DB2 at 0.3 bpns/mv. (a) Tar1 under QP = 28; (b) Tar1 under QP = 18; (c) Tar2 under QP = 28; (d) Tar2 under QP = 18; (e) Tar3 under QP = 28; and (f) Tar3 under QP = 18.

#### 4.2.3. Stability Performance Evaluation

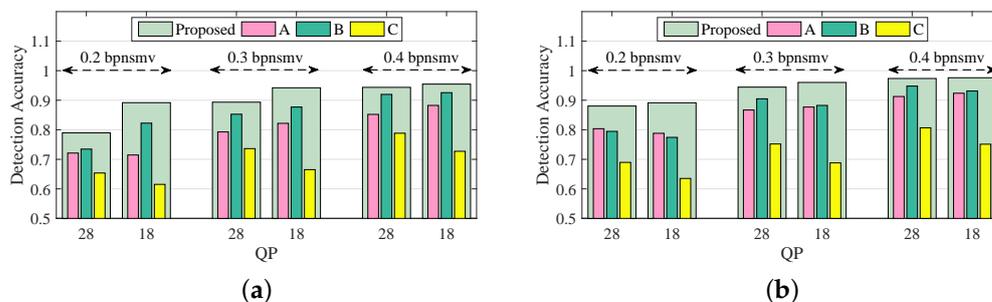
To evaluate the stability of our proposed Lg-LLCQ, the stego video samples with different embedding payloads and QPs under the same steganographic scheme will be grouped together for experiments. Under this circumstance, we compare the detection performance of our proposed Lg-LLCQ with AoSO, NPELO and MVC, and the corresponding experimental results are reported in Table 5. It can be noted from Table 5 that our scheme outperforms all competitors consistently, which demonstrates that our proposed Lg-LLCQ exhibits more stable steganalytic performance against the involved MV-based steganographic schemes.

**Table 5.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed Lg-LLCQ under mixed payloads (0.1~0.5 bpnsmv) and QPs (28 and 18) on DB1 and DB2.

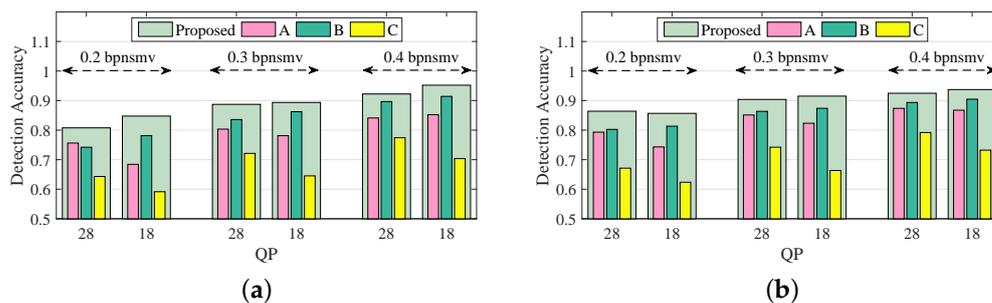
Scheme	DB1				DB2			
	AoSO	NPELO	MVC	Lg-LLCQ	AoSO	NPELO	MVC	Lg-LLCQ
Tar1	95.68	97.82	92.79	<b>98.31</b>	96.26	98.38	94.32	<b>99.12</b>
Tar2	58.31	84.57	86.72	<b>93.62</b>	60.19	86.41	88.26	<b>94.68</b>
Tar3	57.72	83.16	85.04	<b>92.16</b>	57.46	85.67	86.34	<b>93.46</b>

4.2.4. Evaluation of Steganalytic Performance of Subfeatures

Setup 4 aims to evaluate the independent effects of subfeatures of the proposed Lg-LLCQ against Tar2~3, and the results are shown in Figures 7 and 8. It is observed that the proposed feature, built on a combination of all subfeatures, achieves outstanding gains in detecting Tar2~3. It is also noted that (1) the logic-based subfeature and plain quotient-based subfeature provide satisfactory detection accuracies in most cases. (2) The magnified quotient-based subfeature enlarged some subtle distinctions between cover and stego videos to further boost the detection performance.



**Figure 7.** Average accuracy  $\bar{P}_{ACC}$  (in %) of subfeatures against Tar2. (a) DB1 and (b) DB2. A, B and C indicate the logic-based subfeature, plain quotient-based subfeature and magnified quotientbased subfeature, respectively.



**Figure 8.** Average accuracy  $\bar{P}_{ACC}$  (in %) of subfeatures against Tar3. (a) DB1 and (b) DB2. A, B and C indicate the logic-based subfeature, plain quotient-based subfeature and magnified quotientbased subfeature, respectively.

4.2.5. Applicability Performance Evaluation

We then compare the applicability of the proposed Lg-LLCQ with other steganalytic schemes, i.e., AoSO, NPELO and MVC at payloads 0.2 and 0.4 bpnsmv on DB1 and DB2. To accomplish this, we take Tar2 with fixed block size for experiments, and the corresponding results are summarized in Table 6. They show that our proposed Lg-LLCQ still performs the best. On the other hand, they also show the MVC is totally ineffective in detecting Tar2, which can be attributed to the reason that MVC is defined for the sub-blocks within the same macroblock, while the encoder does not exist sub-blocks under fixed-size option in *BaselineProfile*. Although our proposed Lg-LLCQ is originally designed for the detection

of H.264/AVC video steganography, owing to the general applicability of Lg-LLCQ, it can be easily extended to previous video coding standards, i.e., MPEG-2 and MPEG-4.

**Table 6.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed Lg-LLCQ against Tar2 under fixed block size.

Scheme	Feature	QP = 28		QP = 18	
		0.2	0.4	0.2	0.4
Tar2	AoSO	55.71	58.31	67.45	71.78
	NPELO	72.61	77.86	71.46	76.58
	MVC	50.00	50.00	50.00	50.00
	Lg-LLCQ	<b>83.26</b>	<b>89.43</b>	<b>81.27</b>	<b>88.24</b>

#### 4.2.6. Evaluation of Steganalytic Performance under Cover Sources and Steganographic Schemes Mismatch

Since all the involved features achieve excellent performance in detecting Tar1, Tar1 is not considered in this setup. We then proceed to compare the detection performance of our proposed Lg-LLCQ with AoSO, NPELO and MVC under cover sources and steganographic scheme mismatch scenes, which are summarized in Table 7. For Tar2~3, our proposed Lg-LLCQ shows nearly perfect detection performance at various payloads under QP = 18 and QP = 28, indicating the effectiveness of the proposed Lg-LLCQ in the detection of Tar2~3. Note that low QP value, e.g., QP = 18, will weaken the detection model of our proposed Lg-LLCQ in cover sources and steganographic scheme mismatch scenes due to small distortion. This could explain why our proposed Lg-LLCQ is slightly inferior to the MVC feature in detecting Tar2~3 under QP = 18.

**Table 7.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed Lg-LLCQ under cover sources and steganographic scheme mismatch scenes.

Scheme	Trained Model	Payload (QP = 28)					Payload (QP = 18)				
		0.1	0.2	0.3	0.4	0.5	0.1	0.2	0.3	0.4	0.5
Tar2	AoSO-Tar3	50.05	50.41	50.88	51.06	51.49	50.06	51.76	55.89	62.45	65.78
	NPELO-Tar3	63.11	74.55	82.39	86.32	88.61	56.58	69.92	77.17	82.69	84.86
	MVC-Tar3	59.22	74.95	84.81	90.34	93.42	<b>66.78</b>	<b>82.17</b>	<b>89.32</b>	<b>91.66</b>	<b>94.18</b>
	Lg-LLCQ-Tar3	<b>66.73</b>	<b>82.34</b>	<b>90.36</b>	<b>93.65</b>	<b>96.74</b>	64.36	78.67	85.61	88.47	90.26
Tar3	AoSO-Tar2	50.38	50.69	51.24	52.18	53.86	50.19	51.04	53.89	58.23	61.78
	NPELO-Tar2	64.32	72.14	77.39	82.32	84.61	58.65	68.27	73.66	76.89	78.24
	MVC-Tar2	60.27	68.07	78.69	85.24	87.36	62.66	<b>76.12</b>	<b>83.49</b>	<b>85.17</b>	<b>86.38</b>
	Lg-LLCQ-Tar2	<b>68.13</b>	<b>79.84</b>	<b>86.34</b>	<b>91.15</b>	<b>92.08</b>	<b>63.85</b>	73.81	79.63	82.51	83.82

#### 4.2.7. Evaluation of Steganalytic Performance under Different Video Resolutions

We then proceed to evaluate the detection performance of our proposed scheme on DB3. Three SOTA steganalytic schemes, i.e., AoSO, NPELO and MVC, are still included for performance comparison, and the corresponding experimental results are finally listed in Table 8. It is observed that our proposed Lg-LLCQ can still perform the best among involved steganalytic schemes in detecting Tar2~3 under the other resolution database. Table 8 also shows an interesting result that all the test steganalytic schemes perform worse under higher QP value (i.e., QP = 28). This could be due to the fact that a high QP value leads to a decline in the number of MVs. In our paper, the payload is depicted by a relative indicator, i.e.,  $bpnsmv$ , which is the ratio of the number of embedded bits to the total number of nonskip MVs. Therefore, as the number of MVs decreases, the corresponding embedding payload will also reduce, thereby leading to relatively fewer embedding traces, which finally makes it more difficult for steganalyzers to detect.

**Table 8.** Average accuracy  $\bar{P}_{ACC}$  (in %) of AoSO, NPELO, MVC and our proposed Lg-LLCQ against Tar2~Tar3 at relative payloads 0.1 and 0.3 bpsnmv on DB3.

Scheme	Feature	Payload (QP = 28)		Payload (QP = 18)	
		0.1	0.3	0.1	0.3
Tar2	AoSO	50.32	50.59	50.68	53.78
	NPELO	51.42	66.21	52.14	74.52
	MVC	50.58	66.73	61.64	83.16
	Lg-LLCQ	<b>52.81</b>	<b>77.15</b>	<b>63.81</b>	<b>89.22</b>
Tar3	AoSO	50.13	50.41	50.37	51.69
	NPELO	53.54	68.27	51.71	72.18
	MVC	51.82	67.55	55.86	81.58
	Lg-LLCQ	<b>55.63</b>	<b>76.49</b>	<b>59.78</b>	<b>85.62</b>

## 5. Conclusions

In this paper, we proposed an effective and general steganalytic feature named Lg-LLCQ to detect MV-based H.264/AVC steganography. The proposed Lg-LLCQ is composed of two types of subfeatures, i.e., the logic-based (Lg) subfeature and the local Lagrangian cost quotient (LLCQ)-based subfeature, wherein the Lg subfeature is inherited from the previous art NPELO, and the LLCQ subfeature is newly designed according to the statistic of local Lagrangian cost quotient. As for the characterization of this statistic, we referred to the fact that the embedding modifications usually cause the numerical anomaly of the local Lagrangian cost quotient and proposed a local Lagrangian cost quotient(LLCQ)-based indicator, the validity of which was also validated. Moreover, with the introduction of the LLCQ-based subfeature, the defect in the Lg-based subfeature that it is less effective in the detection of some logic-maintaining steganographic schemes can be well-compensated, and the resulting Lg-LLCQ feature shows the best performance in detecting MV-based H.264/AVC steganography. By the way, we will extend the proposed Lg-LLCQ to MV-based H.265/HEVC steganography to further enhance its generality in the future.

**Author Contributions:** Y.L.: conceptualization, methodology, investigation, software and writing—original draft. J.N.: conceptualization, supervision, validation and writing—review editing. W.S.: conceptualization, methodology, investigation, validation and writing—review editing. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was partially supported by the National Natural Science Foundation of China (Grants No. 62202507 and U1936212), Natural Science Foundation of Guangdong Province, China (Grant No. 2022A1515011209), and China Postdoctoral Science Foundation (Grant No. 2021M703767).

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Khalifa, A.; Guzman, A. Imperceptible Image Steganography Using Symmetry-Adapted Deep Learning Techniques. *Symmetry* **2022**, *14*, 1325. [CrossRef]
- Lin, J.; Tsai, C.W.; Yang, C.W.; Liu, K.H. A Large Payload Data Hiding Scheme Using Scalable Secret Reference Matrix. *Symmetry* **2022**, *14*, 828. [CrossRef]
- Lin, J.; Chang, C.C.; Horng, J.H. Asymmetric Data Hiding for Compressed Images with High Payload and Reversibility. *Symmetry* **2021**, *13*, 2355. [CrossRef]
- Su, W.; Ni, J.; Hu, X.; Fridrich, J. Image steganography with symmetric embedding using Gaussian Markov random field model. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *31*, 1001–1015. [CrossRef]
- Liu, J.F.; Tian, Y.G.; Han, T.; Yang, C.F.; Liu, W.B. LSB steganographic payload location for JPEG-decompressed images. *Digit. Signal Process.* **2015**, *38*, 66–76. [CrossRef]
- Hu, X.; Ni, J.; Shi, Y.Q. Efficient JPEG steganography using domain transformation of embedding entropy. *IEEE Signal Process. Lett.* **2018**, *25*, 773–777. [CrossRef]
- Xu, C.; Ping, X.; Zhang, T. Steganography in compressed video stream. In Proceedings of the IEEE First International Conference on Innovative Computing, Information and Control-Volume I (ICIC'06), Beijing, China, 30 August–1 September 2006; Volume 1, pp. 269–272. [CrossRef]

8. Fang, D.Y.; Chang, L.W. Data hiding for digital video with phase of motion vector. In Proceedings of the 2006 IEEE International Symposium on Circuits and Systems, Kos, Greece, 21–24 May 2006. [\[CrossRef\]](#)
9. Aly, H. Data hiding in motion vectors of compressed video based on their associated prediction error. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 14–18. [\[CrossRef\]](#)
10. Cao, Y.; Zhao, X.; Feng, D.; Sheng, R. Video steganography with perturbed motion estimation. In *International Workshop on Information Hiding*; Springer: Berlin/Heidelberg, Germany, 2011; pp. 193–207. [\[CrossRef\]](#)
11. Cao, Y.; Zhang, H.; Zhao, X.; Yu, H. Video steganography based on optimized motion estimation perturbation. In Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security, Portland, OR, USA, 17–19 June 2015; pp. 25–31. [\[CrossRef\]](#)
12. Zhang, H.; Cao, Y.; Zhao, X. Motion vector-based video steganography with preserved local optimality. *Multimed. Tools Appl.* **2016**, *75*, 13503–13519. [\[CrossRef\]](#)
13. Yao, Y.; Zhang, W.; Yu, N.; Zhao, X. Defining embedding distortion for motion vector-based video steganography. *Multimed. Tools Appl.* **2015**, *74*, 11163–11186. [\[CrossRef\]](#)
14. Wang, P.; Zhang, H.; Cao, Y.; Zhao, X. A novel embedding distortion for motion vector-based steganography considering motion characteristic, local optimality and statistical distribution. In Proceedings of the 4th ACM Workshop on Information Hiding and Multimedia Security, Vigo Galicia, Spain, 20–22 June 2016; pp. 127–137. [\[CrossRef\]](#)
15. Zhu, B.; Ni, J. Uniform embedding for efficient steganography of H. 264 video. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 1678–1682. [\[CrossRef\]](#)
16. Ghamsarian, N.; Khademi, M. Undetectable video steganography by considering spatio-temporal steganalytic features in the embedding cost function. *Multimed. Tools Appl.* **2020**, *79*, 18909–18939. [\[CrossRef\]](#)
17. Yao, Y.; Yu, N. Motion vector modification distortion analysis-based payload allocation for video steganography. *J. Vis. Commun. Image Represent.* **2021**, *74*, 102986. [\[CrossRef\]](#)
18. Ma, X.; Li, Z.; Tu, H.; Zhang, B. A Data Hiding Algorithm for H.264/AVC Video Streams Without Intra-Frame Distortion Drift. *IEEE Trans. Circuits Syst. Video Technol.* **2010**, *20*, 1320–1330. [\[CrossRef\]](#)
19. Esen, E.; Alatan, A.A. Robust Video Data Hiding Using Forbidden Zone Data Hiding and Selective Embedding. *IEEE Trans. Circuits Syst. Video Technol.* **2011**, *21*, 1130–1138. [\[CrossRef\]](#)
20. Cao, Y.; Wang, Y.; Zhao, X.; Zhu, M.; Xu, Z. Cover block decoupling for content-adaptive H. 264 steganography. In Proceedings of the 6th ACM Workshop on Information Hiding and Multimedia Security, Innsbruck, Austria, 20–22 June 2018; pp. 23–30. [\[CrossRef\]](#)
21. Ghasempour, M.; Ghanbari, M. A Low Complexity System for Multiple Data Embedding Into H.264 Coded Video Bit-Stream. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *30*, 4009–4019. [\[CrossRef\]](#)
22. Wang, Y.; Cao, Y.; Zhao, X. Minimizing Embedding Impact for H. 264 Steganography by Progressive Trellis Coding. *IEEE Trans. Inf. Forensics Secur.* **2020**, *16*, 333–345. [\[CrossRef\]](#)
23. Wang, Y.; Cao, Y.; Zhao, X. CEC: Cluster Embedding Coding for H. 264 Steganography. *IEEE Signal Process. Lett.* **2020**, *27*, 955–959. [\[CrossRef\]](#)
24. Liu, S.; Liu, Y.; Feng, C.; Zhao, H.; Huang, Y. A HEVC Steganography Method Based on QDCT Coefficient. In *International Conference on Intelligent Computing*; Springer: Cham, Switzerland, 2020; pp. 624–632. [\[CrossRef\]](#)
25. Chen, Y.; Wang, H.; Wu, H.Z.; Wu, Z.; Li, T.; Malik, A. Adaptive video data hiding through cost assignment and STCs. *IEEE Trans. Dependable Secur. Comput.* **2019**, *19*, 955–959. [\[CrossRef\]](#)
26. Chen, Y.; Wang, H.; Choo, K.K.R.; He, P.; Salcic, Z.; Kaafar, M.A.; Zhang, X. DDCA: A Distortion Drift-Based Cost Assignment Method for Adaptive Video Steganography in the Transform Domain. *IEEE Trans. Dependable Secur. Comput.* **2021**, *19*, 2405–2420. [\[CrossRef\]](#)
27. Yang, X.Y.; Zhao, L.Y.; Niu, K. An efficient video steganography algorithm based on sub-macroblock partition for H. 264/AVC. *Adv. Mater. Res.* **2012**, *433*, 5384–5389. [\[CrossRef\]](#)
28. Zhang, H.; Cao, Y.; Zhao, X.; Zhang, W.; Yu, N. Video steganography with perturbed macroblock partition. In Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security, Salzburg, Austria, 11–13 June 2014; pp. 115–122. [\[CrossRef\]](#)
29. Nie, Q.; Xu, X.; Feng, B.; Zhang, L.Y. Defining embedding distortion for intra prediction mode-based video steganography. *Comput. Mater. Contin.* **2018**, *55*, 59–70. [\[CrossRef\]](#)
30. Jia, X.; Wang, J.; Liu, Y.; Kang, X.; Shi, Y. A Layered Embedding-Based Scheme to Cope with Intra-Frame Distortion Drift In IPM-Based HEVC Steganography. In Proceedings of the ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, ON, Canada, 6–11 June 2021; pp. 2720–2724. [\[CrossRef\]](#)
31. Zhao, H.; Liu, Y.; Wang, Y.; Liu, S.; Feng, C. A Video Steganography Method Based on Transform Block Decision for H. 265/HEVC. *IEEE Access* **2021**, *9*, 55506–55521. [\[CrossRef\]](#)
32. Liu, J.; Li, Z.; Jiang, X.; Zhang, Z. A high-performance CNN-applied HEVC steganography based on diamond-coded PU partition modes. *IEEE Trans. Multimed.* **2021**, *24*, 2084–2097. [\[CrossRef\]](#)
33. Wong, K.; Tanaka, K. A data hiding method using Mquant in MPEG domain. *J. Inst. Image Electron. Eng. Jpn.* **2008**, *37*, 256–267. [\[CrossRef\]](#)

34. Wong, K.; Tanaka, K.; Takagi, K.; Nakajima, Y. Complete video quality-preserving data hiding. *IEEE Trans. Circuits Syst. Video Technol.* **2009**, *19*, 1499–1512. [CrossRef]
35. Shanableh, T. Data hiding in MPEG video files using multivariate regression and flexible macroblock ordering. *IEEE Trans. Inf. Forensics Secur.* **2011**, *7*, 455–464. [CrossRef]
36. Filler, T.; Judas, J.; Fridrich, J. Minimizing Additive Distortion in Steganography Using Syndrome-Trellis Codes. *IEEE Trans. Inf. Forensics Secur.* **2011**, *6*, 920–935. [CrossRef]
37. Su, Y.; Zhang, C.; Zhang, C. A video steganalytic algorithm against motion-vector-based steganography. *Signal Process.* **2011**, *91*, 1901–1909. [CrossRef]
38. Tasdemir, K.; Kurugollu, F.; Sezer, S. Spatio-temporal rich model-based video steganalysis on cross sections of motion vector planes. *IEEE Trans. Image Process.* **2016**, *25*, 3316–3328. [CrossRef]
39. Cao, Y.; Zhao, X.; Feng, D. Video steganalysis exploiting motion vector reversion-based features. *IEEE Signal Process. Lett.* **2011**, *19*, 35–38. [CrossRef]
40. Wang, P.; Cao, Y.; Zhao, X.; Wu, B. Motion vector reversion-based steganalysis revisited. In Proceedings of the 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), Chengdu, China, 12–15 July 2015; pp. 463–467. [CrossRef]
41. Wang, K.; Zhao, H.; Wang, H. Video Steganalysis Against Motion Vector-Based Steganography by Adding or Subtracting One Motion Vector Value. *IEEE Trans. Inf. Forensics Secur.* **2014**, *9*, 741–751. [CrossRef]
42. Zhang, H.; Cao, Y.; Zhao, X. A Steganalytic Approach to Detect Motion Vector Modification Using Near-Perfect Estimation for Local Optimality. *IEEE Trans. Inf. Forensics Secur.* **2017**, *12*, 465–478. [CrossRef]
43. Zhai, L.; Wang, L.; Ren, Y. Universal Detection of Video Steganography in Multiple Domains Based on the Consistency of Motion Vectors. *IEEE Trans. Inf. Forensics Secur.* **2020**, *15*, 1762–1777. [CrossRef]
44. Deng, Y.; Wu, Y.; Zhou, L. Digital video steganalysis using motion vector recovery-based features. *Appl. Opt.* **2012**, *51*, 4667–4677. [CrossRef]
45. Wu, H.T.; Liu, Y.; Huang, J.; Yang, X.Y. Improved steganalysis algorithm against motion vector based video steganography. In Proceedings of the 2014 IEEE International Conference on Image Processing (ICIP), Paris, France, 27–30 October 2014; pp. 5512–5516. [CrossRef]
46. Zhai, L.; Wang, L.; Ren, Y. Combined and calibrated features for steganalysis of motion vector-based steganography in H. 264/AVC. In Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security, Philadelphia, PA, USA, 20–22 June 2017; pp. 135–146. [CrossRef]
47. Ren, Y.; Zhai, L.; Wang, L.; Zhu, T. Video steganalysis based on subtractive probability of optimal matching feature. In Proceedings of the 2nd ACM Workshop on Information Hiding and Multimedia Security, Salzburg, Austria, 11–13 June 2014; pp. 83–90. [CrossRef]
48. Ghamsarian, N.; Schoeffmann, K.; Khademi, M. Blind MV-based video steganalysis based on joint inter-frame and intra-frame statistics. *Multimed. Tools Appl.* **2021**, *80*, 9137–9159. [CrossRef]
49. Liu, S.; Hu, Y.; Liu, B.; Li, C.T. An HEVC steganalytic approach against motion vector modification using local optimality in candidate list. *Pattern Recognit. Lett.* **2021**, *146*, 23–30. [CrossRef]
50. Wiegand, T.; Schwarz, H.; Joch, A.; Kossentini, F.; Sullivan, G.J. Rate-constrained coder control and comparison of video coding standards. *IEEE Trans. Circuits Syst. Video Technol.* **2003**, *13*, 688–703. [CrossRef]
51. Puri, A.; Chen, X.; Luthra, A. Video coding using the H. 264/MPEG-4 AVC compression standard. *Signal Process. Image Commun.* **2004**, *19*, 793–849. [CrossRef]
52. Cao, Y.; Zhang, H.; Zhao, X.; He, X. Steganalysis of H. 264/AVC Videos Exploiting Subtractive Prediction Error Blocks. *IEEE Trans. Inf. Forensics Secur.* **2021**, *16*, 3326–3338. [CrossRef]
53. Suehring, K. H.264/AVC JM Reference Software. Available online: <http://iphome.hhi.de/suehring/tml/> (accessed on 15 October 2021).
54. Video Trace Library, YUV Sequences. Available online: <http://trace.eas.asu.edu/yuv/index.html> (accessed on 15 October 2021).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.