A video steganalysis method based on coding cost variation

Jianyi Liu¹, Cong Zhang¹, Ru Zhang¹, Yi Li¹ and Jie Cheng²

Abstract
Aiming at the problems existing in existing steganalysis algorithms, this article proposes Motion Vector Coding Cost Change video steganalysis features based on Improved Motion Vector Reversion-Based features and Subtractive Probability of Coding Cost Optimal Matching features based on Subtractive Probability of Optimal Matching features from the perspective of the change of coding cost. Motion Vector Coding Cost Change features can be well consistent with the coding cost before recoding by analyzing the sub-pixel coding cost of recoding. By counting the sub-pixel coding costs of motion vectors before and after video recoding, the Sum of Absolute Difference values of motion vectors instead of predicted residuals are applied to steganalysis and detection, and the steganographic algorithm based on motion vectors is effectively detected. Experiments show that Motion Vector Coding Cost Change features have higher detection accuracy than Add-or-Subtract-One, Improved Motion Vector Reversion-Based, and other typical features in various steganography methods, and Subtractive Probability of Coding Cost Optimal Matching features have higher detection effect and better robustness than Subtractive Probability of Optimal Matching features.

Keywords
Video steganalysis, motion vector, sub-pixel coding cost, recoding, Subtractive Probability of Coding Cost Optimal Matching

Date received: 22 October 2019; accepted: 4 January 2021

Handling Editor: Svalastog, J

Introduction
Traditional steganography and steganalysis techniques are mostly targeted at image carriers, and the technology is becoming more mature. As one of the most popular media in the Internet, video is composed of a series of spatially and temporally related images. It has a larger amount of data than a single frame image and can accommodate more secret messages. To this end, more researchers have turned their attention to steganography and steganalysis of video carriers.

Video steganography in compressed domain is the main research direction, which can be divided into special type and general type. The special video steganalysis techniques in compressed domain can be divided into four types: steganalysis method based on intra prediction mode, steganalysis method based on inter prediction mode, steganalysis method based on transformation coefficient, steganalysis method based on motion vector, and steganalysis method based on MV (Motion Vector).

For steganalysis method based on intra prediction model, Zhao et al.¹ proposed that the intra prediction mode reverted to the original mode after the video was recompressed. Therefore, designing 13-dimensional IPMC (Intra Prediction Mode Calibration) features by calculating the coding cost change of macroblocks in the intra prediction mode could have better detection results for the steganography method based on the intra prediction mode.

¹Beijing University of Posts and Telecommunications, Beijing, China
²State Grid Information and Telecommunication Branch, Beijing, China

Corresponding author:
Jianyi Liu, Beijing University of Posts and Telecommunications, Xitucheng Road 10, Haidian District, Beijing 100876, China.
Email: liujy@bupt.edu.cn
DCTR (Discrete Cosine Transform Residual) features prediction model, in Zhang et al., the phenomenon is helpful to improve detection accuracy. For steganalysis method based on inter prediction model, in Zhang et al., the phenomenon that the intra prediction mode reverts to the original model after recoding is also used. By calculating the changes of the intra prediction mode before and after recoding, the 40-dimensional features were designed, which had a good detection effect.

Research on video steganalysis based on MV is relatively mature, which can be divided into four types. The first method is to take advantage of the statistical properties of MV. Su et al. constructed a 12-dimensional feature, which comprehensively utilized the central moment of MV histogram from both time domain and space domain. In Zhang and XiJian, the method of temporal and spatial correlation based on pixels is applied to steganographic method detection based on MV. Ye et al. used four-phase scanning to construct the symbiosis matrix to capture the abnormal changes in the space-time correlation of MVs caused by steganography. Tasdemir et al. transferred the SRM (Spatio Rich Model) applied in the field of image steganalysis into the video field, and detected the temporal and spatial characteristic changes of MVs in three dimensions, which has a good effect. However, this method can extract higher characteristic dimensions, and the calculation is more complex and time-consuming. The second is to use the MV calibration method. Cao et al. found through experiments that MV tended to recover to the original when video was re-encoded. A 15-dimensional Motion Vector Reversion-Based (MVRB) feature was designed by utilizing the distance offset of MV before and after video recompression and the change of predicted residual size. On this basis, Wang et al. proposed the IMVRB (Improved Motion Vector Reversion-Based) steganalysis features of video. This method first decoded the video to obtain the available parameters, and then selected the motion prediction method to re-encode the video through the distance offset of MV. This method obtained the available parameters and simulated the unavailable parameters, which had better detection effect than MVRB features. The third method uses the local optimality of MV to determine whether steganography exists. Wang et al. proposed an 18-dimensional steganalysis feature of Add-or-Subtract-One (AoSO) video by taking advantage of the locally optimal property of MV after coding, namely, SAD (Sum of Absolute Difference) minimization. The AoSO feature constructs features by adding or subtracting the SAD value within a range of the MV neighborhood. AoSO is the most effective feature of MV steganography. The fourth method is based on adaptive method to determine whether steganography exists. Wang et al. divided video subsequence by computing the dynamic attitude (Frame Dynamic Degree) of video frames, and carried out feature extraction for subsequence with high dynamic attitude, which is helpful to improve detection accuracy.

The general video steganalysis techniques in compressed domain are still in research stage. In Liu and Li, the convolutional neural network for image steganalysis was improved. The SRM algorithm was used to input the neural network to distinguish the original video frame and the encrypted video frame, which could detect the steganography method based on MV and the steganography method based on intra prediction mode.

In this article, a video steganalysis method based on coding cost change is proposed, which improved the IMVRB feature based on the coding cost change of MV before and after recoding and the SPOM (Subtractive Probability of Optimal Matching) feature based on the coding cost change of local optimal probability, so as to improve detection performance and robustness. This method could also effectively detect the video steganography method which better maintains the local optimalness of the MV’s SAD (Sum of Absolute Difference) value.

**Analysis of the variation of recompressed coding cost value**

Most of the motion-based steganalysis methods use the characteristics of the MV’s SAD value. Due to the existence of quantization errors, although the MV offset of the video is small after recompression, the SAD value of the MV has changed greatly before and after re-encoding, even in the same prediction mode. This section mainly analyzes the coding cost change of the MV before and after re-encoding.

Let \( C_{k,j} \) be the value of the MV’s coding cost of the macroblock with number \( l \) in the prediction residual frame numbered \( k \), and let \( C'_{k,j} \) be the value of the MV’s coding cost of the macroblock with number \( l \) in the prediction residual frame numbered \( k \) after recoding, then

\[
C_{k,j} = \sum_{i\in MBR,l} \sum_{i\in MBR,l} |COST_{i,j,k}|
\]

(1)

\[
C'_{k,j} = \sum_{i\in MBR,l} \sum_{i\in MBR,l} |COST'_{i,j,k}|
\]

(2)

where \( COST \) represents the coding cost value before recoding, \( COST' \) represents the coding cost value after recoding, and \( i \) and \( j \) represent the horizontal and the
vertical sequence numbers of the pixel in macroblock \( MB \), respectively. Similar to the statistical method of the influence of recoding on the SAD value of MV, the influence of recoding on the coding cost value of MV can be measured by the following statistic

\[
e_C = \frac{C - C_0}{C_0}
\]

In the case that coastguard_cif.yuv sequence is selected from the standard video library website, and the macroblock size is set to 16. JM coding software was used for compression coding, and the \( e_C \) variable of the first 60 frames was counted, as shown in Figure 1.

Through the analysis in Figure 1, it can be found that in the case of recoding with the same prediction pattern, despite the quantization error, the coding cost value of the video MV will not change greatly before and after the recoding. Even if it changes, the change is modest. Compared with the change of SAD value, the distribution is relatively concentrated, especially the high proportion that the coding cost value of MV can keep consistent. Although there are quantization errors, the key feature of the improved steganalysis method based on MVs is the consistent proportion of MVs before and after recoding.

**Analysis of the selection of coding cost’s variation features**

**Selection of coding cost’s variation features**

In the process of video recoding, although due to the existence of quantization error, the coding cost value of video MV generally will not change greatly before and after recoding, especially the consistent proportion is very high. Further analysis will be carried out in the following.

First, decode the original video and the re-encoded video, then record the decoded MV of the original video and the prediction residual matrix to calculate the coding cost value, and finally, use the same method to calculate the coding cost value of the re-encoded video for the sub-pixel coding cost calculation. For macroblock \( MB \), the decoding coding cost and the recompression coding cost are \( COST_{MB} \) and \( COST'_{MB} \), respectively, and the coding cost changes before and after re-encoding are defined

\[
\delta(COST_{MB}) = |COST_{MB} - COST'_{MB}|
\]

Let \( \Pr(\delta(COST_{MB}) = 0) \) represent the probability that the coding cost of MV decoding is the same as the coding cost after recompression.

Three classic video sequences are selected from the standard video library website: bus_cif.yuv, mobile_cif.yuv, and coastguard_cif.yuv. By compressing the three video sequences separately, the respective carrier video, the MV larger component least significant bit (LSB) + 1 steganographic video, and the larger component LSB of the MV match embedding steganographic video are obtained. The experimental statistics have the following rules after decoding and re-encoding: the proportion that the same macroblock’s coding cost of the carrier video without steganographic techniques after the decoding and recompression remains consistent is high, that is, the ratio that the value is 0 is very high. Steganography reduces this ratio and the value changes greatly after recompression compared with decoding. The specific experimental results are shown in Table 1.

It can be seen from Table 1 that the probability value of the carrier video \( \Pr(\delta(COST_{MB}) = 0) \) before and after re-encoding is very high, and the corresponding probability value \( \Pr(\delta(COST_{MB}) = 0) \) of the steganographic video embedded by the MV larger component LSB + 1 and embedded by the MV larger component LSB match is relatively low. It can be seen that the video without steganographic operation can better maintain the law that the coding cost remains consistent before and after re-encoding. But the video after steganography will destroy this law. Therefore, designing features from this perspective can well capture whether this law changes, so that the carrier video is distinguished from the steganographic video effectively.

The MV and coding cost information of the video macroblock is used for feature calculation. Through the foregoing analysis, the probability value \( \Pr(\delta(COST_{MB}) = 0) \) before and after re-encoding has a very effective distinguishing property, so it can be set as a one-dimensional feature. Considering the change of MV before and after re-encoding and the change of sub-pixel coding cost, we can refer to the design idea of
IMVRB video steganalysis feature, and the MVCCC (Motion Vector Coding Cost Change) feature design has the following four feature sets

The feature set 1

\[ f^1(k) = f^1_{IMVRB}(k) \]  \hspace{1cm} (5)

This feature set is consistent with the feature set of the IMVRB, reflecting the change of the MV before and after the video recompression.

The feature set 2

\[ f^2(k) = \frac{1}{\gamma} \sum_{i=0}^{n} \delta COST_{X_i} \]  \hspace{1cm} (6)

where

\[ \gamma = \sum_{i=0}^{n} \delta COST_{X_i} \]  \hspace{1cm} (7)

The feature set 2 counts the variation of the sub-pixel coding cost of the MV with an offset distance \( k \), where \( X_i \) represents a macroblock, \( u \) represents the offset distance of the MV, and \( n \) represents the total number of macroblocks.

The feature set 3

\[ (k)f^3 = \frac{1}{\eta} \sum_{i=0}^{n} \left( \|m_{X_i} - m'_{X_i}\| \right) \times \delta COST_{X_i} \]  \hspace{1cm} (8)

where

\[ \eta = \sum_{i=0}^{n} \left( \|m_{X_i} - m'_{X_i}\| \right) \times \delta COST_{X_i} \]  \hspace{1cm} (9)

The feature set 3 counts the comprehensive variation of the sub-pixel coding cost of the MV with an offset distance \( k \) and offset distance.

The MVCCC video steganalysis feature proposed in this article uses the above three feature sets and the probability value \( \Pr(\delta(COST_{MB}) = 0) \) before and after re-encoding, \( F = (f_0, f_1, \ldots, f_{10}) \), a total of 11 dimensions, the specific formula is as follows

\[ F_k = \begin{cases} f^1(k) : k = 0, 1, 2, 3 \\ \Pr(\delta(COST_{MB}) = 0) : k = 5 \\ f^2(k-5) : k = 6, 7, 8 \\ f^3(k-8) : k = 9, 10, 11 \end{cases} \]  \hspace{1cm} (10)

In order to visually observe the discrimination of the above features, a sample carrier video and steganographic video are selected to extract the above 11-dimensional features, as shown in Figure 2.

It can be observed that the probability that the MV offset distance of the first-dimensional feature is 0 is consistent with the sub-pixel coding cost before and after the re-encoding of the fifth-dimensional feature of the carrier video is significantly larger than the steganographic video, but the change in the sub-pixel coding cost of the 6th–11th dimensions and the comprehensive variation of the MV and sub-pixel coding cost of the carrier video are smaller than the steganographic video.

Finally, the support-vector machine (SVM) is used for classification, and the Gaussian kernel function is selected. The penalty factor used includes \( C = \{10^{-3}, 10^{-2}, \ldots, 10^4, 10^5\} \), and the best \( C \) value is searched by traversing.

Selection of the variation features of the probability of coding cost optimal matching

Through the introduction above, when the video is re-encoded, although the video MV SAD value and the coding cost value may change before and after the re-encoding due to the existence of the quantization error, the variation of the SAD value and the probability value of coding cost optimal matching need further analysis. Given a video segment \( V \), where the number of MVs is \( N_{mv} \) and the number of MVs that are locally optimal for SAD values is \( N_{om} \), then the probability of SAD local optimum in \( V \) is

\[ P_V = \frac{N_{om}}{N_{mv}} \]  \hspace{1cm} (11)

| Video sequence   | Carrier video | LSB + 1 steganographic video | LSB matching steganographic video |
|------------------|---------------|------------------------------|----------------------------------|
| bus_cif.yuv      | 81.48         | 62.75                        | 65.41                            |
| mobile_cif.yuv   | 86.67         | 65.40                        | 68.23                            |
| coastguard_cif.yuv | 84.42        | 61.37                        | 64.51                            |

LSB: least significant bit.
Similarly, for a macroblock $MB$, the MV has the least cost $COST_{MB}$ in the range of $mv_{Dh,Dv} = \{mv_{h} + \Delta h, mv_{v} + \Delta v\}$, $\Delta h, \Delta v \in \{-1, 0, 1\}$, which is locally optimal. The number of MVs that meet the local optimum of the coding cost is $N_{om}'$, then the probability of CCPOM (Probability of Coding Cost Optimal Matching) that defines the local optimum of the coding cost in $V$ is

$$P'_v = \frac{N'_{om}}{N_{mv}}$$ (12)

The recompressed video is defined as $V'$. For a given video $V$, the POM (Probability of Optimal Matching) of $V$ is $P_{test}$, the POM of $V'$ is $P_{cab}$, the CCPOM of $V$ is $P'_{test}$, the CCPOM of $V'$ is $P'_{cab}$, and the SPOM is $P_{err}$. The formula is

$$P_{err} = P_{cab} - P_{test}$$ (13)

Similarly, the variation of probability of coding cost optimal matching before and after the video re-encoding is defined as

$$P'_{err} = P'_{cab} - P'_{test}$$ (14)

First, the impact of steganography on video CCPOM is analyzed. Three classic video sequences are selected in the standard video library website which are bus_cif.yuv, soccer_cif.yuv, and coastguard_cif.yuv, respectively. The respective carrier video and MV larger component LSB + 1 steganographic video are obtained by compressing the first 150 frames of the three video sequences, respectively. The experimental steganography operation has the effect on the video CCPOM is as follows: the probability of carrier video’s coding cost optimal matching CCPOM without steganography operation is higher, and the CCPOM value of the steganographic video is lower. The specific experimental results are shown in Table 2.

It can be seen from Table 2 that the CCPOM value of the original carrier video is high, and the steganographic video will reduce the probability of local optimization of the coding cost, that is, the CCPOM value will be much lower.

Then, the effect of the steganographic operation on the variation of the probability $P'_{err}$ of local optimization of the coding cost before and after the video re-encoding is analyzed. With the same three videos as described above, in the case of performing LSB + 1 on the larger component of the modified 0%, the modified 20%, and the modified 40% number of MVs, the change value of the probability of local optimization of the coding cost before and after the re-encoding is recorded. The specific results are shown in Table 3.

It can be seen from Table 3 that the probability $P'_{err}$ of local optimization of the coding cost of the video after steganography before and after recompression will become larger and larger as the embedding rate increases. Therefore, this value is very distinguishable and can be used as a classification feature.

Referring to the design idea of SPOM video steganalysis features, SCCPOM (Subtractive Probability of Coding Cost Optimal Matching) feature design is divided into the following two feature sets:

The feature set 1

$$f^1(k) = P_{err}(k), (k = 0, 1, 2, 3, 4)$$ (15)

The feature set reflects the change of the probability of SAD which is locally optimal before and after video

| Video sequence | Carrier video | LSB + 1 steganographic video |
|----------------|---------------|-----------------------------|
| bus_cif.yuv    | 87.18         | 25.75                       |
| soccer_cif.yuv | 87.32         | 23.40                       |
| coastguard_cif.yuv | 85.75   | 20.37                       |

CCPOM: Coding Cost Optimal Matching; LSB: least significant bit.

| Video sequence | Carrier video | 20% LSB + 1 | 40% LSB + 1 |
|----------------|---------------|-------------|-------------|
| bus_cif.yuv    | 2.51          | 7.75        | 12.3        |
| soccer_cif.yuv | 1.21          | 6.40        | 9.23        |
| coastguard_cif.yuv | 2.35     | 7.23        | 11.6        |

LSB: least significant bit.
recompression, and \( k \) is 0, 1, 2, 3, and 4, respectively, represent the calculated values \( P_{\text{err}} \) at positions \( mv(h,v) \), \( mv(h+1,v) \), \( mv(h-1,v) \), and \( mv(h,v+1) \).

The feature set 2

\[
f^2(k) = P_{\text{err}}(k), \quad (k = 0, 1, 2, 3, 4) \quad (16)
\]

The feature set reflects the variation of probability of coding cost optimal matching before and after video recompression, and \( k \) is 0, 1, 2, 3, and 4, respectively, represent the calculated values \( P_{\text{err}} \) at positions \( mv(h,v) \), \( mv(h+1,v) \), \( mv(h,v-1) \), \( mv(h-1,v) \), and \( mv(h,v+1) \).

The SPCCOM (Subtractive Probability of Coding Cost Optimal Matching) video steganalysis feature proposed in this article uses the above two feature sets, the original position SAD and the design of the probability of coding cost optimal matching, \( F = (f_0, f_1, \ldots, f_5) \), a total of 12 dimensions, the specific formula is as follows:

The SPCCOM feature has a total of 12 dimensions. The specific formula is as follows

\[
F_k = \begin{cases} 
  f^1(k) : k = 0, 1, 2, 3, 4 \\
  f^2(k-5) : k = 5, 6, 7, 8, 9 \\
  P_r : k = 10 \\
  P_s : k = 11 
\end{cases} \quad (17)
\]

The feature includes the SAD value of the video MV at its original position, the position where the vertical component is incremented by one, the position where the horizontal component is decremented by one, and the position where the horizontal component is incremented by one and the position where the vertical component is decremented by one. Therefore, when the two values of different videos are different in size, and the situation which the original MV is close to the local optimum is not the same. Therefore, some videos with low values will be misjudged as steganographic videos. If these videos are not steganized, the SAD values before and after re-encoding and the variation of the probability of coding cost optimal matching will be low. A proposal to combine the two is proposed and classification method is the same as the previous section.

**Design of the steganalysis method**

**Design of the steganalysis method based on the variation of coding cost value**

Through the analysis of the selection of coding cost features in the above section 2.1, it can be found that the proportion of the MV’s coding cost of the carrier video which is consistent before and after re-encoding is high, while the corresponding proportion of the steganographic is low, which has a good distinguishing effect. So this phenomenon can be used to design one-dimensional features. In addition, the IMVRB video steganalysis feature based on the method of restoring MV proposed in the previous section uses the variation of the prediction error SAD before and after re-encoding to design features, but the prediction error is not a key component of video coding motion estimation calculation. So, the change of prediction error is used to design features, although the change after video steganography can be reflected to some extent, it is still not an accurate solution. It can provide better detection results for simple steganographic methods such as Xu et al.\(^{17}\) and Aly,\(^{18}\) while reduce the embedding rate and the distortion of the video after the steganographic operation for the method proposed by Cao et al.\(^{19}\) for STC (syndrome-tellis codes) coding technology\(^{20}\) to a large extent. Therefore, using prediction error to analyze steganographic operations is no longer sufficient. Although this cannot be met, this idea can be used to design features that change the coding cost by reflecting the changes in the coding cost before and after re-encoding.

The SATD (Sum of Absolute Transformed Difference) value of the H.264 video sub-pixel coding cost reflects the size of code stream to a certain extent, so the scheme uses the sub-pixel coding cost. The overall scheme flow is as follows:

1. Prepare the carrier and the original video set.
2. Recode video by the motion search method selected and counted using the similarity.
3. Decode the video and the re-encoded video, record MV and predict residual information, and calculate the coding cost.
4. Calculate feature using MV and coding cost of decoding and re-encoding.
5. Train SVM for classification.

The specific process of MVCCC feature extraction in this article is shown in Figure 3. First, the video is decompressed, the effective compression parameters are collected, and the MV and prediction error (PE) are recorded. When the video is decoded, the video is compressed to obtain residuals and MVs by entropy decoding, inverse quantization, and IDCT transform, and the current macroblock is obtained by macroblock decoding in the reference frame by motion compensation. The available compression parameters can be collected for recompression after the decoding process.

Important available parameters can be broadly classified into three categories according to their role. The first category is related to the basic characteristics of...
the video, including resolution, frame number, and frame rate. The second category is related to video quality, such as compression parameters and bit rate. The last category determines the location of macroblocks, such as image groups and macroblock partitioning. In order to ensure the consistency of the above important parameters in the recompression process, the above video compression parameters are collected in this step.

Then, the video is recompressed, matching the most similar motion prediction methods and recording MVs and the information about prediction error. The process of video recompression is to predict the motion for each macroblock of the current frame, find the reference block in the reference frame, and obtain error and MV. The error is subjected to DCT transform, quantization, and entropy coding processing, and the MV is subjected to entropy coding processing to obtain a compressed video for transmission. Since the current block is still possible as a reference block after encoding, the current block is decoded (inverse quantization and iDCT transform) to obtain a reconstructed macroblock during the compression process.

In recompressed motion prediction, full search, hex search, and small diamond search are traversed. The MV obtained by the above method is used as a candidate for matching. In the matching process, the video offset distance is used to measure whether the MV is similar to that obtained by decoding. The MV which is the most similar to the decoded MV will be used as the MV in recompression. Then, the MV and the information about prediction error of each macroblock are recorded. Finally, the MV and the information about prediction error are obtained before and after the recompression to calculate the sub-pixel coding cost using the H.264 video sub-pixel coding cost formula.

**Design of the steganalysis method based on the variation of probability of coding cost optimal matching**

Through the analysis of the selection of coding cost features in the above section 2.2, we can find that the probability of MV’s coding cost optimal matching of carrier video is very high, and the variation of the probability value before and after re-encoding is very relatively low. And the corresponding proportion of steganographic video will increase, which has a good distinction. So this phenomenon can be used to design features. In addition, the SPOM feature, which also uses the prediction error SAD variation, is less effective in dealing with the steganographic method of Cao et al.,19 and because it has only one-dimensional features, it is less robust and needs to be improved. The scheme first uses the motion estimation matching method in the re-encoding process as in the MVCCC feature in section “Selection of coding cost’s variation features,”10 and then needs to record the value of the MV in the range of its addition and subtraction. Then, after recompression, the same operation is performed to calculate the original MV’s SAD and the probability value of coding cost optimal matching in the range of its addition and subtraction. Finally,
calculate the change of the SAD and the probability value of coding cost optimal matching after decoding and re-encoding and decoding. The procedure of the scheme is as follows:

1. Prepare the carrier and the original video set.
2. Decode the video and the re-encoded video, record MV and predict residual information, and calculate the coding cost.
3. Recode video by the motion search method selected and counted using the similarity, decode the re-encoded video, record MV information, and calculate the probability of coding cost optimal matching.
4. Calculate feature using SAD and the probability of coding cost optimal matching.
5. Train SVM for classification.

**Experimental results and analysis**

**Introduction to the experimental environment**

1. Test sequence: 30 CIF (352*288) video frames form a video library with a pixel format of YUV4:2:0. These videos are downloaded from the standard video library website on the Internet. Since the length of the video is different from 100 to 2000 frames, each video is not cross-cut into 60 frames, and a total of 200 sub-video sets are used for experiments. The video is compressed using the JM10.2 open-source video encoder. The motion estimation fast search algorithm is DIA, HEX, and so on.
2. Steganographic method: in order to ensure the comprehensiveness of the selected steganographic algorithm, the first one we selected is the method based on MV magnitude of Xu et al., and the second is the method based on prediction residual of Aly, and the third is the method based on STC and wet paper coding which could better maintain the local optimality of MVs of Cao et al. The corrupted motion vector ratio (CMVR) is used to represent the steganographic embedding rate, and the CMVR of the carrier video is set to 0, considering the case where the CMVR is 0.2, 0.4, and 0.6.
3. Steganalysis method: in addition to the features we proposed, the IMVRB video steganalysis feature of Wang et al., the SPOM steganalysis feature of Ren et al., and the AoSO video steganalysis feature of Wang et al. were used as a comparison object; these three features are quite effective steganalysis methods for MV domains.
4. Training and detection: in each set of steganalysis experiments, 70% of the carrier video and corresponding steganographic video are randomly selected for training support-vector machines, and the remaining 30% of the sample pairs are used for steganographic classification decisions.

**Analysis of results**

First, verify the validity of the MVCCC features, and perform experiments according to the above experimental configuration. Figures 4–6 show the detection accuracy of the IMVRB video steganalysis feature of Wang et al. and the AoSO video steganalysis feature of Wang et al., and the MVCCC feature based on the coding cost change proposed in this article for the Xu et al. video steganography method, the Aly video
The experimental results show that the MVCCC features proposed in this article are better and more effective than the IMVRB video steganalysis feature and the AoSO video steganalysis feature in the Xu steganography method, the Aly steganography method, and the Cao steganography method. For the Aly steganography method, the detection result of the IMVRB video steganalysis feature is quite close to the detection result of the proposed method, but the detection result of the AoSO video steganalysis feature is slightly lower. For the Cao steganography method, the accuracy of the proposed scheme is much higher than the IMVRB video steganalysis feature and the AoSO video steganalysis feature. The feature of this article first uses the motion prediction matching method to effectively restore the video compression process, avoiding the increase of detection error rate due to the uncertainty of motion estimation mode. Then, considering the estimated coding cost value of the sub-pixel MV motion in re-encoding, the value before re-encoding can be well maintained, and the MV is usually optimized in the sense of distortion rate. The AoSO video steganalysis feature only uses the size of the distortion as the basis for judgment which is not an accurate criterion for the encoding process. The IMVRB video steganalysis feature also uses SAD as a feature, so the effect is not particularly good when it copes with the steganographic algorithm such as the Cao method to maintain the distortion of local optimal features. Therefore, the features of this article have a more effective detection rate than the IMVRB video steganalysis feature and the AoSO video steganalysis feature.

Next, the validity of the SCCPOM feature is verified. The comparison feature is the SPOM feature extracted by Ren and the steganography method is the Cao video steganography, which is performed at different embedding rates, as shown in Figure 7.

It can be seen from Figure 7 that the SPOM feature is not effective in the steganographic algorithm which maintains the local optimality of SAD such as Cao et al. But the SCCPOM feature proposed in this article increases the variation of probability of coding cost optimal matching before and after video re-encoding based on SPOM features, and it performs better when detecting the method of keeping SAD local optimality. The SPOM feature has only one-dimensional features, so the robustness is poor. The SCCPOM feature combines the variation of the local optimal characteristics of the video and the local optimal probability before and after re-encoding, which has better detection effect and robustness.

Conclusion

In this article, two video MV steganalysis methods based on coding cost changes are proposed. Experiments show that MVCCC features have higher detection accuracy than typical features such as AoSO and IMVRB in various steganography methods, and SPCCOM features have higher detection effect and better robustness than SPOM features. This article verifies the validity of the coding cost feature and provides a favorable reference for further research on video steganalysis.
Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported, in part, by the National Key R&D Program of China (2019YFB1406504) and the National Natural Science Foundation of China (U1836108, U1936216, 62002197, and 62001038).

ORCID iD

Jianyi Liu https://orcid.org/0000-0003-3133-4452

References

1. Zhao Y, Zhang H, Cao Y, et al. Video steganalysis based on intra prediction mode calibration. In: Proceedings the 14th international workshop on digital-forensics and watermarking (IWDW’15) (Lecture Notes in Computer Science), Beijing, China, 17–19 September 2016, pp.119–133. New York: Springer.
2. Wang P, Cao Y, Zhao X, et al. A steganalytic algorithm to detect DCT-based Data Hiding Methods for H.264/AVC Videos. In: Proceedings of the 5th ACM workshop on information hiding and multimedia security (Ih&mmsec’17), Philadelphia, PA, 20–21 June 2017, pp.123–133. New York: ACM.
3. Holub V and Fridrich J. Low-complexity features for JPEG steganalysis using undecimated DCT. IEEE Trans Inform Forensic Secur 2015; 10(2): 219–228.
4. Zhang H, Cao B and Zhao X. Video steganalysis method based on inter-frame prediction mode recovery detection. China, 201711021119.7 Patent, 2018.
5. Su Y, Zhang C and Zhang C. A video steganalytic algorithm against motion-vector-based steganography. Signal Pr 2011; 91: 1901–1909.
6. Zhang YX and Xifian P. Video steganalysis based on spatio-temporal correlation. J Image Graph 2010; 15(9): 1331–1337.
7. Ye H, Zhang W, Yao Y, et al. Motion vector-based video steganalysis using spatial-temporal correlation. In: 6th international congress on image and signal processing (CISP’), Hangzhou, China, 16–18 December 2014, pp.148–153. New York: IEEE.
8. Tasdemi K, Kurugollu F and Sezer S. Spatio-temporal rich model for motion vector steganalysis. In: IEEE international conference on acoustics, speech and signal processing (ICASSP’), Brisbane, QLD, Australia, 19–24 April 2015, pp.1717–1721. New York: IEEE.
9. Cao Y, Zhao XF and Feng DG. Video steganalysis exploiting motion vector reversion-based features. Signal Proces Lett 2012; 19(1): 35–38.
10. Wang P, Cao Y, Zhao X, et al. Motion vector reversion-based steganalysis revisited. In: IEEE China summit and international conference on signal and information processing, Chengdu, China, 12–15 July 2015, pp.405–408. New York: IEEE.
11. Wang K, Zhao H and Wang H. Video steganalysis against motion vector-based steganography by adding or subtracting one motion vector value. IEEE Trans Inf Forensic Secur 2014; 9: 741–751.
12. Wang Y, Cao X, Zhao H, et al. An adaptive detecting strategy against motion vector-based steganography. In: Proceedings 2015 IEEE international conference on multimedia and expo (ICME’15), Turin, 29 June–3 July 2015, pp.1–6. New York: IEEE.
13. Liu P and Li S. Universal steganalysis convolutional neural network for video intra-frame inter-frame coding. Newt New Media Tech 2018; 7(6): 13–16.
14. Fridrich J and Kodovsky J. Rich models for steganalysis of digital image. Trans Inform Forensic Secur 2012; 7(3): 862882.
15. Chang CC and Lin CJ. LIBSVM. A library for support vector machines, 2016, http://www.csie.ntu.edu.tw/~cjlin/libsvm (assessed on 22 December 2016)
16. Ren Y, Zhai L, Wang L, et al. Video steganalysis based on subtractive probability of optimal matching feature. In: Proceedings of the 2nd ACM workshop on information hiding and multimedia security (Ih&mmsec’14), New York, 11–13 June 2014, pp.83–90. New York: ACM.
17. Xu C, Ping X and Zhang T. Steganography in compressed video stream. In: First international conference on innovative computing, information and control, Beijing, China, 30 August–1 September, vol. 1, 30 August–1 September 2006, pp.1422–1425. New York: IEEE.
18. Aly HA. Data hiding in motion vectors of compressed video based on their associated prediction error. IEEE Trans Inf Forensics Secur 2011; 6: 14–18.
19. Cao Y, Zhang H, Zhao X, et al. Video steganography based on optimized motion estimation perturbation. In: Proceedings of the 3rd ACM workshop information hiding multimedia security, Portland, OR, 17–19 June 2015, pp.25–31. New York: ACM.
20. Filler T, Judas J and Fridrich J. Minimizing additive distortion in steganography using syndrome-trellis codes. IEEE Trans Inf Forensics Secur 2011; 6: 920–935.