Hindawi
Security and Communication Networks
Volume 2017, Article ID 8051389, 12 pages
https://doi.org/10.1155/2017/8051389

Research Article

Segmentation Based Video Steganalysis to Detect Motion Vector Modification

Peipei Wang,[1,2] Yun Cao,[1,2] and Xianfeng Zhao[1,2]

[1] *State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China*
[2] *School of Cyber Security, University of Chinese Academy of Sciences, Beijing 100093, China*

Correspondence should be addressed to Xianfeng Zhao; zhaoxianfeng@iie.ac.cn

Received 13 April 2017; Revised 23 July 2017; Accepted 7 August 2017; Published 16 October 2017

Academic Editor: Qingzhong Liu

This paper presents a steganalytic approach against video steganography which modifies motion vector (MV) in content adaptive manner. Current video steganalytic schemes extract features from fixed-length frames of the whole video and do not take advantage of the content diversity. Consequently, the effectiveness of the steganalytic feature is influenced by video content and the problem of cover source mismatch also affects the steganalytic performance. The goal of this paper is to propose a steganalytic method which can suppress the differences of statistical characteristics caused by video content. The given video is segmented to subsequences according to block's motion in every frame. The steganalytic features extracted from each category of subsequences with close motion intensity are used to build one classifier. The final steganalytic result can be obtained by fusing the results of weighted classifiers. The experimental results have demonstrated that our method can effectively improve the performance of video steganalysis, especially for videos of low bitrate and low embedding ratio.

1. Introduction

Steganography, as the art and science of data hiding, realizes covert communication under the camouflage of innocent-looking cover media. It will not arouse eavesdroppers' suspicion because the perceptual and statistical characteristic of embedded file is similar to that of original unaltered counterpart. Facilitated by advanced video compression and computer network technology, digital video has become one of the most influential media. And the boom of highly interactive multimedia applications has created an urgent need to explore the steganography of hiding data into digital videos.

Up to date, video steganographic methods are usually integrated into the video compression process. Such approaches hide information by modifying certain output coefficients during compression procedure, such as MVs [1–6], interprediction modes [7], quantized DCT coefficients [8–10], and variable length codes [11, 12]. In this paper, we focus on attacking MV-based steganography. There are two advantages making MV-based steganography superior to others. First, rich motion information in compressed video streams guarantees sufficient embedding capacity. Secondly, the modification applied to MV will not affect the coding performance much.

In this paper, we focus on attacking MV-based steganography. Many MV-based steganographic methods have been proposed recently. Jordan et al. [1] embedded message bits by modifying the LSBs of nonzero MVs' horizontal and vertical components. Xu et al. [2] suggested modifying the MVs whose magnitudes are above a given threshold. Aly [3] chose the candidate MVs according to their associated prediction errors. By applying mature coding techniques such as wet paper codes (WPCs) [13] and syndrome-trellis codes (STCs) [14, 15] to video steganography, adaptive steganographic schemes have been presented in recent years. In Yao et al.'s work [4], an adaptive MV-based steganography was proposed by considering the statistical distribution change and the prediction error change. And two-layered STCs [15] are used to minimize distortion for embedding process. In order to resist the steganalytic schemes based on MV's local

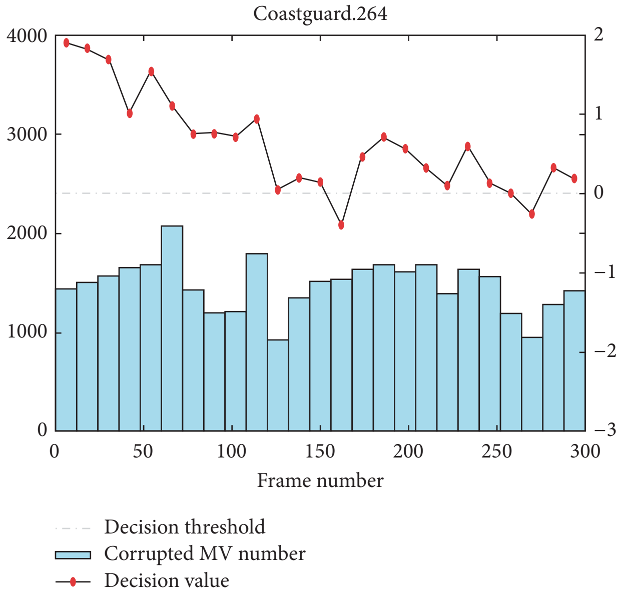


Figure 1: Corrupted MV number and corresponding decision value of every detection interval.

optimality [16], several approaches were proposed. In [5], Cao et al. exploited the opportunity to optimize the ME perturbation using the loss caused by video compression process. The data embedding was implemented using a double-layered (first channel: STCs [15]; second channel: WPCs [13]) coding structure with distortion scale calculated on optimal neighbors. In Wang et al.'s work [6], data was embedded based on the distortion defined by considering motion characteristic of video content, MVs local optimality, and statistical distribution.

In order to reveal the existence of hidden message, current video steganalysis divides the video into detection intervals (DI) with fixed-length and then extracts feature from every DI. The calibration-based approach is a typical steganalytic method. In Wang et al.'s work [17], the calibration-based steganalysis is further improved by matching the parameters between the first and second compression process. Wang et al. [16] extracted features based on the difference between the actual the sum of absolute difference (SAD) and locally optimal SAD after the adding-or-subtracting-one operation on MVs. Recently Zhang et al. [18] suggested checking the local optimality of MVs by considering both distortion and bit estimation associated with MVs. And near-perfect estimation for local optimality is utilized to detect MV-based steganography.

Although various steganalytic approaches have been presented, there are still many challenges in the field of video MVs targeted steganalysis. Just as modifications are implemented in textured regions in image adaptive steganography [19, 20], adaptive approaches [4–6] in video steganography can also constrain their embedding changes to those parts that are difficult to model, such as rich motion frames. As a consequence, the adaptive steganography has become the research focus due to its high embedding capacity and enhanced security. The basic principle of the current video steganalysis is to analyze the embedding perturbation and statistical changes within the fixed-length DIs. However, the embedding changes are not only correlated with the steganographic methods, but also with the video content.

Consequently the restrictions of current video steganalysis can be concluded as follows. Firstly, compared with the processing of embedding, the video content makes a more significant impact on the differences of video statistical characteristics. Moreover, the detection accuracy of steganalytic method relies on the performance of classifiers. Therefore, if the contents of training and testing videos are different a lot in motion intensity, the result of classification will be affected obviously.

In order to solve this problem, a steganalytic method using motion based segmentation is proposed in this paper. The main contributions of this paper include segmenting the whole videos to subsequences according to the block's motion and extracting steganalytic features from categories of subsequences with close motion intensity; building the model of multiple subclassifiers; and fusing the results of weighted subclassifiers to obtain the final steganalytic performance.

The organization of the rest paper is as follows. In Section 2, our steganalytic approach including video segmentation and decision fusion is proposed. Section 3 shows the experimental results and the conclusions and future works are given in Section 4.

2. Motivation

In current video steganography [4–6], the data is adaptively embedded by modifying the MVs according to video content. Because the changes of statistical characteristics are different in frames with different motion intensity, the features extracted from fixed-length DIs are not effective, which can not cope with adaptive steganography well.

In order to demonstrate this phenomenon, we utilize Cao's method [5] to embed information into the video Coastguard. The NPELO features [18] are extracted from

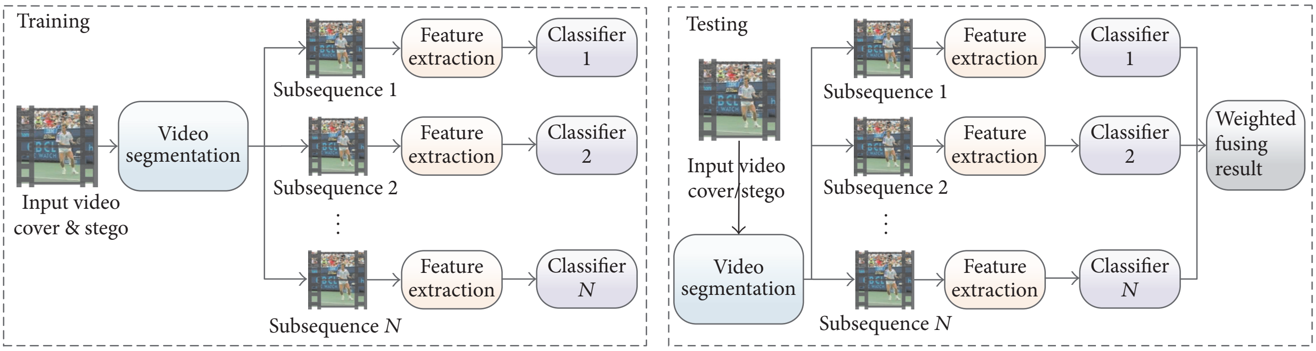


Figure 2: The segmentation based steganalysis.

every 12 frames. Then the features are subjected to a trained classifier. Figure 1 shows the corrupted MV number and the corresponding decision value. It is observed that the feature's effectiveness is greatly affected by the motion intensity in video content. And the features drawn from frames with high motion intensity are usually more effective than that from other frames.

In image steganalysis, many frameworks [21–24] considering image content have been proposed, which guide the proposal of our approach. If we can extract features from subsequences of different motion intensity, the classifiers with features of different effectiveness can be trained independently. By assigning high weight values to effective classifiers, the overall steganalytic performance is expected to be improved.

3. The Proposed Steganalysis

In this paper, the video is segmented to several subsequences which are sorted by the motion intensity. The effectiveness of features extracted from the categories with rich motion is to be improved, which make it easier to distinguish the stego videos from the cover ones.

The schematic diagram of segmentation based steganalysis is shown in Figure 2. Compared with directly extracting features from fixed-length DIs of the whole video in traditional video steganalysis, the input videos are firstly segmented to subsequences both in the training and in testing processes. Then the features are extracted from DIs in category of subsequences with different motion intensity. In the training process, the cover and stego video pairs in the training set are segmented according to video content and then the features of the subsequences with similar motion intensity are subjected to train one classifier. Consequently, N classifiers are trained after this process. In the testing process, the features of different categories of subsequences are fed to different classifiers. The final result is obtained by fusing the results of classifiers assigned with different weight values.

3.1. Video Segmentation. Inspired by the motion continuity in video content, we first segment the video into subsequences by linking the blocks among adjacent frames. Subsequently, based on the characteristics of the linking, the subsequences are sorted to categories of different motion intensity.

3.1.1. MV Flow Based Segmentation. As the integral part of existing video coding standards, Motion Estimation (ME) is designed to reduce the temporal redundancy between video frames. This is achieved by allowing blocks of pixels from currently coded frame to be matched with those from reference frame(s). As a result of ME, MV represents the spatial displacement offset between a block and its prediction.

Therefore, MVs' values are greatly determined by the ME performance. MVs obtained by different ME methods vary a lot. It is accepted that a moving object is usually very different from the static background. Therefore, if MVs do not reflect the real motion, the corresponding prediction residuals are relatively large. Based on this principle, we define "credible MV" as follows.

Definition 1 (credible MV). A certain MV is deemed to be credible if it satisfies

$$\left|\mathrm{SAD}-\frac{\mathrm{SAD}_1+\cdots+\mathrm{SAD}_i+\cdots+\mathrm{SAD}_{n^2-1}}{n^2-1}\right|<\delta \qquad i\in\left\{1,2,\ldots,n^2-1\right\}. \tag{1}$$

Here, SAD is the sum of absolute difference between macroblock (MB) and its prediction and δ is the standard deviation of SAD. This formula computes the difference between the current SAD and the average SAD of its neighbors.

If the difference does not exceed the preset threshold, the SADs' variety in this $n \times n$ area is equable and there is no singularity at the center. It means that compared with its neighbors, the SAD of this macroblock is not distinctly large. Therefore, its MV can indicate the direction and magnitude of real motion, which is called a credible MV.

Inspired by the proverbial concept optical flow [25] in pixel-domain action recognition, we bring forward "MV Flow." In pixel domain, the dense optical flow $\omega_t = (u_t, v_t)$ of frame F_t is computed by tracking pixel point in the next frame F_{t+1}, where u_t and v_t are the horizontal and vertical components. In spite of different implementation domain and details, this process is similar to ME in video compression. And as the outcome of locating macroblock in its reference frame, MV has some same characteristics with ω_t. Thus in our method, the definition of "MV Flow" is given as below.

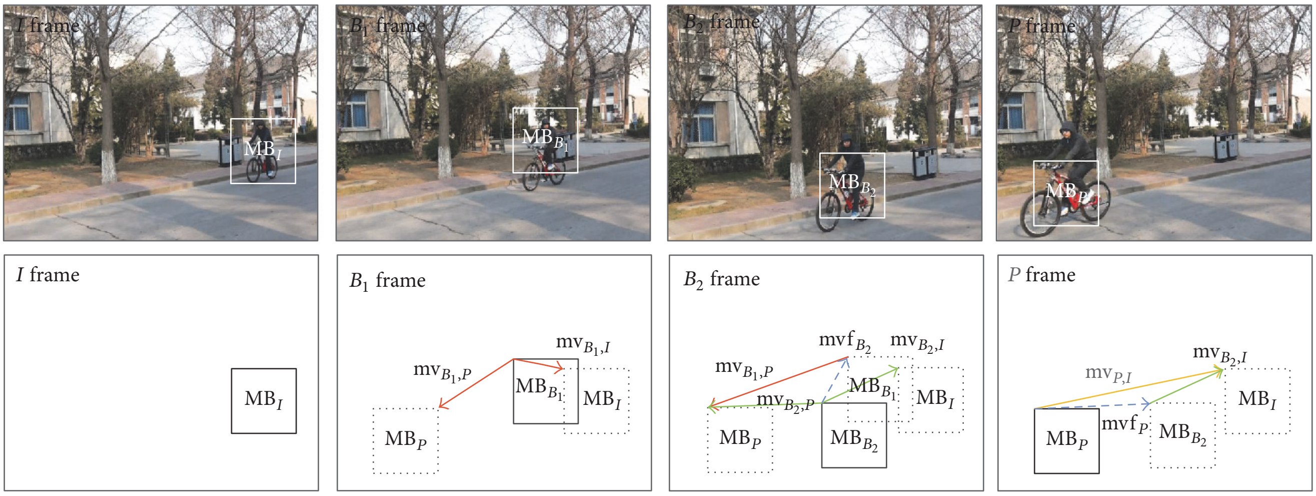


FIGURE 3: The computation of MV Flow.

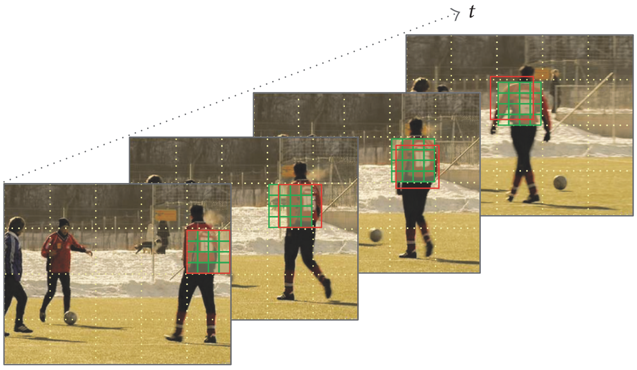


FIGURE 4: Segmentation by linking blocks.

Definition 2 (MV Flow, MVF). The MV Flow of frame F_t is defined as

$$\mathrm{MVF}_t = \left\{\mathbf{mvf}_t^i,\ i = 1 \cdots m\right\}, \quad (2)$$

where m is the number of macroblocks and $\mathbf{mvf}_t$ is the credible MV matched by referring MBs in frame F_{t-1}.

As shown in Figure 3, in every frame, the MVF can be computed from their original credible MVs. In frame F_{B_1}, $\mathbf{mvf}_{B_1}$ equals $\mathbf{mv}_{B_1,I}$, which is the MV obtained by referring to the corresponding MB_I in frame I. In frame F_{B_2}, according to the operation rule of vector (the Triangle Rule), we can get $\mathbf{mvf}_{B_2} = \mathbf{mv}_{B_2,I} - \mathbf{mv}_{B_1,I}$. And similarly from analyzing the constraints of the credible MVs, we can get $\mathbf{mvf}_P = \mathbf{mv}_{P,I} - \mathbf{mv}_{B_2,I}$ in frame F_P.

In MVF field, the macroblocks' shifts between two adjacent frames can be interpolated using credible MVs. As shown in Figure 4, given a compressed video sequence $\{F_t\}_{t=1}^N$, every frame's MVF $\{\mathbf{mvf}_t^i\}_{i=1}^m$ can be obtained by computation shown in Figure 3. The calculated MV $\mathbf{mvf}_t^i$ points to the location of corresponding macroblock in frame F_{t-1}. If the location crosses macroblock boundary, the one with largest overlapping area is selected as the best matching macroblock. In order to link the associated blocks with same size, the block of larger size is subblocks. The principle of this manipulation is based on the motion continuity of video content. If there is motion change or shot switching in this video, the whole video will be segmented into several subsequences by linking the blocks.

3.1.2. Classification of Subsequences. In this subsection, the classifying strategy of subsequences is proposed. First we define the distance to measure the motion in this subsequence and then we proceed to classify the subsequences to different categories with different motion intensity.

As discussed in Section 3.1.1, similar blocks can be linked from MVF. Derived from the distance measurement [26] in pixel domain, we propose the "Linking Distance" to measure motion of block linking during this subsequence. The similar blocks are linked by $\mathbf{mvf}_t = (x_t, y_t)$ $(t = 1, \ldots, T)$, which are vectors on x-y plane along the discrete time axis. Thus the linking can be denoted by L_i, where $L_i = \{\mathbf{mvf}_{i,t}\}_{t=1}^T = ((x_{i,1}, y_{i,1}), \ldots, (x_{i,T}, y_{i,T}))$. The definition of "Linking Distance" of the sequence is given as follows.

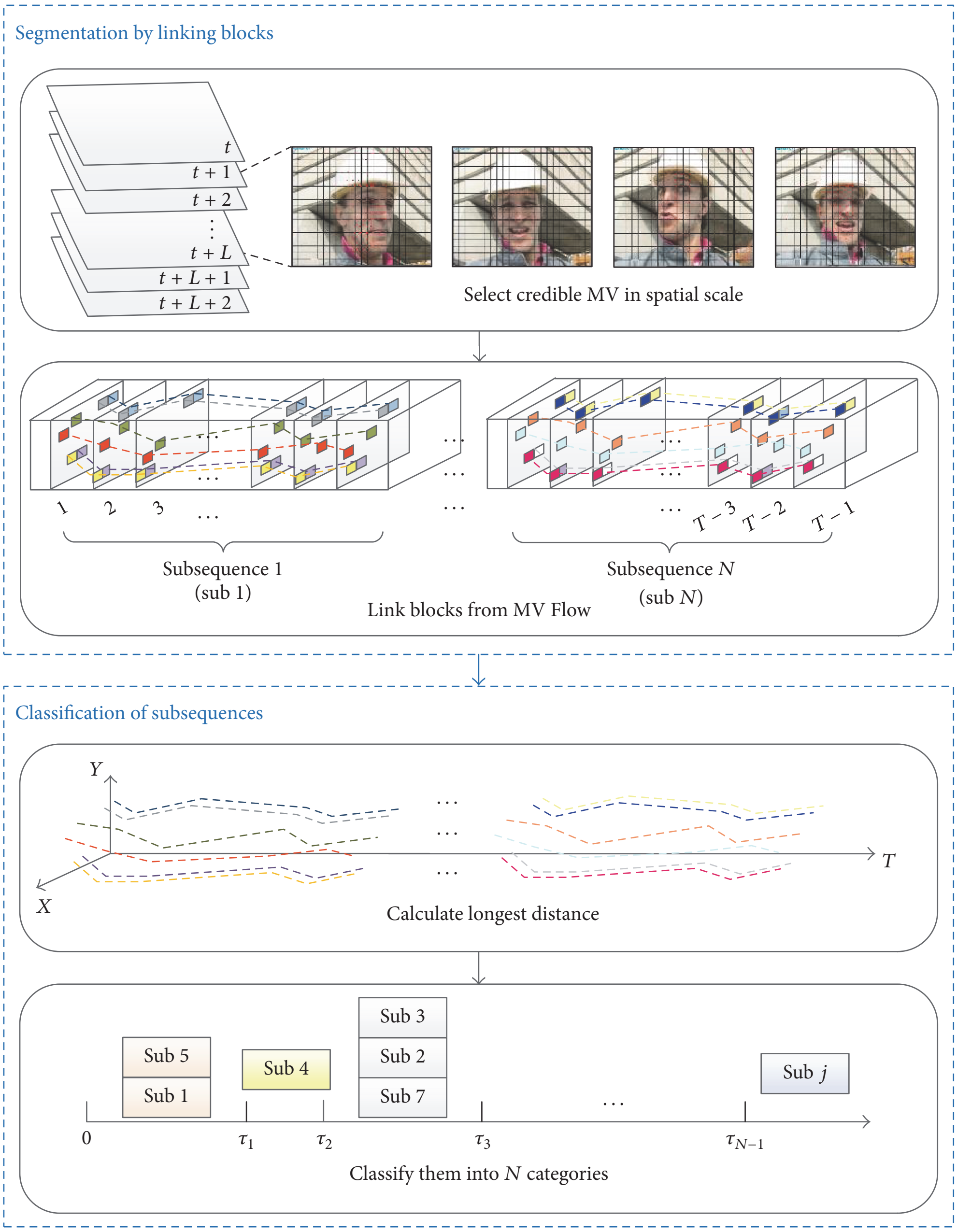


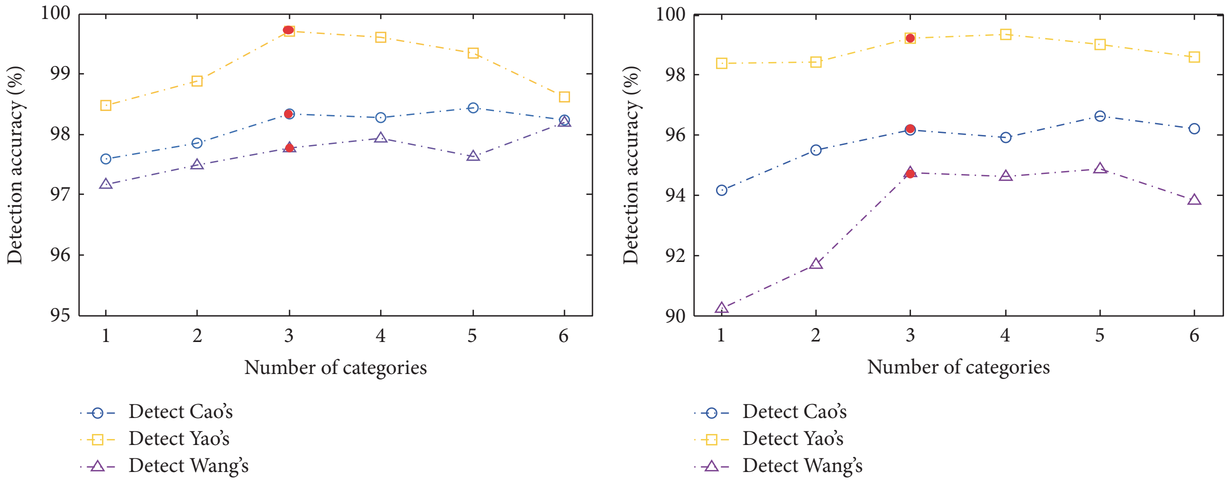
FIGURE 5: The process of video segmentation.

Definition 3 (Linking Distance). The Linking Distance (LD) of this subsequence is defined as

$$\mathrm{LD}_j = \frac{\sum_{i=1}^{M} \sum_{t=1}^{T} \sqrt{\left(x_t - x_{t+1}\right)^2 + \left(y_t - y_{t+1}\right)^2}}{M \times T}, \qquad (3)$$

where LD_j $(1 \le j \le V)$ is the LD of the jth subsequence of the video, V is the number of subsequences, and M is the number of linking in this subsequence. The motion intensity of the subsequence can be measured by calculating the LD of all of the blocks' linking. By setting several threshold values of LD, the jth subsequences Sub_j are classified into categories with different motion intensity, which is formulated as follows:

$$\begin{aligned} \mathrm{Sub}_j &\in \mathrm{Cat}_1, \quad \text{if } 0 \le \mathrm{LD}_j \le \tau_1, \\ \mathrm{Sub}_j &\in \mathrm{Cat}_2, \quad \text{if } \tau_1 \le \mathrm{LD}_j \le \tau_2, \\ &\vdots \\ \mathrm{Sub}_j &\in \mathrm{Cat}_N, \quad \text{if } \tau_{N-1} \le \mathrm{LD}_j, \end{aligned} \qquad (4)$$



(a)

(b)

Figure 6: Steganalytic performances (%) of (a) improved NPELO and (b) improved MVRBR using different number of categories. The red objects refer to the number of categories which is used for further experiments.

where $\{\tau_1, \tau_2, \ldots, \tau_{N-1}\}$ is the set of thresholds to classify the subsequences and N is the number of categories, which is corresponding to N classifiers in the training and testing processes.

As described in Figure 4, the segmentation of compressed video can be realized by linking the similar blocks and classifying the subsequences. The implementation procedure is illustrated in Figure 5. First, when inspecting the MV field, credible MVs are selected to guarantee the representability of macroblocks' real motion. Based on the credible MVs' relevance between frames, MVF can be obtained. Subsequently, under the maximum-overlapping principle, the similar blocks are linked between adjacent frames, which results in dividing the whole video into several subsequences. Then LD of every subsequence is computed and will be further used for classifying the subsequence to one of the categories with settled thresholds. As a consequence, all of the subsequences are sorted to N categories, from which the N kinds of features are extracted and then subjected to the corresponding N classifiers.

The proposed video segmentation method is based on the ME process in video compression. In order to dispose the intracodec blocks' influence, we test several 720P sequences with QP as 28. It is found that only 5%-6% macroblocks are intracoded in P and B frame. Thus the effect of intra-MB can be negligible in our method.

3.2. Decision Fusion. After segmenting the videos into subsequences sorted by motion intensity, the steganalytic features are extracted from each category of subsequences, which are utilized for training or testing. As shown in Figure 2, the features extracted from N different categories are used to train N classifiers in the training process. And in order to test the video, the features of N categories are input to the corresponding trained classifiers, which output N detection results.

Inspired by the fusing methods in image steganalysis [21–24], we assign weight value to each classifier. The weight value of the jth classifier is defined by

$$w_j = \frac{\left(\mathrm{DA}_j - 0.5\right)}{\sum_{j=1}^{N}\left(\mathrm{DA}_j - 0.5\right)},$$
$$\mathrm{DA}_j = \frac{\mathrm{TP}_j + \mathrm{TN}_j}{2}, \tag{5}$$

where DA_j $(0 \le j \le N)$ is the detection accuracy of jth classifier and TP_j and TN_j are the rates of true positive and true negative respectively. If the detection accuracy of a specific category of subsequences is relatively higher, it means that more subsequences are correctly detected as stego. Thus a bigger weight value should be assigned to this type of classifier, and vice versa.

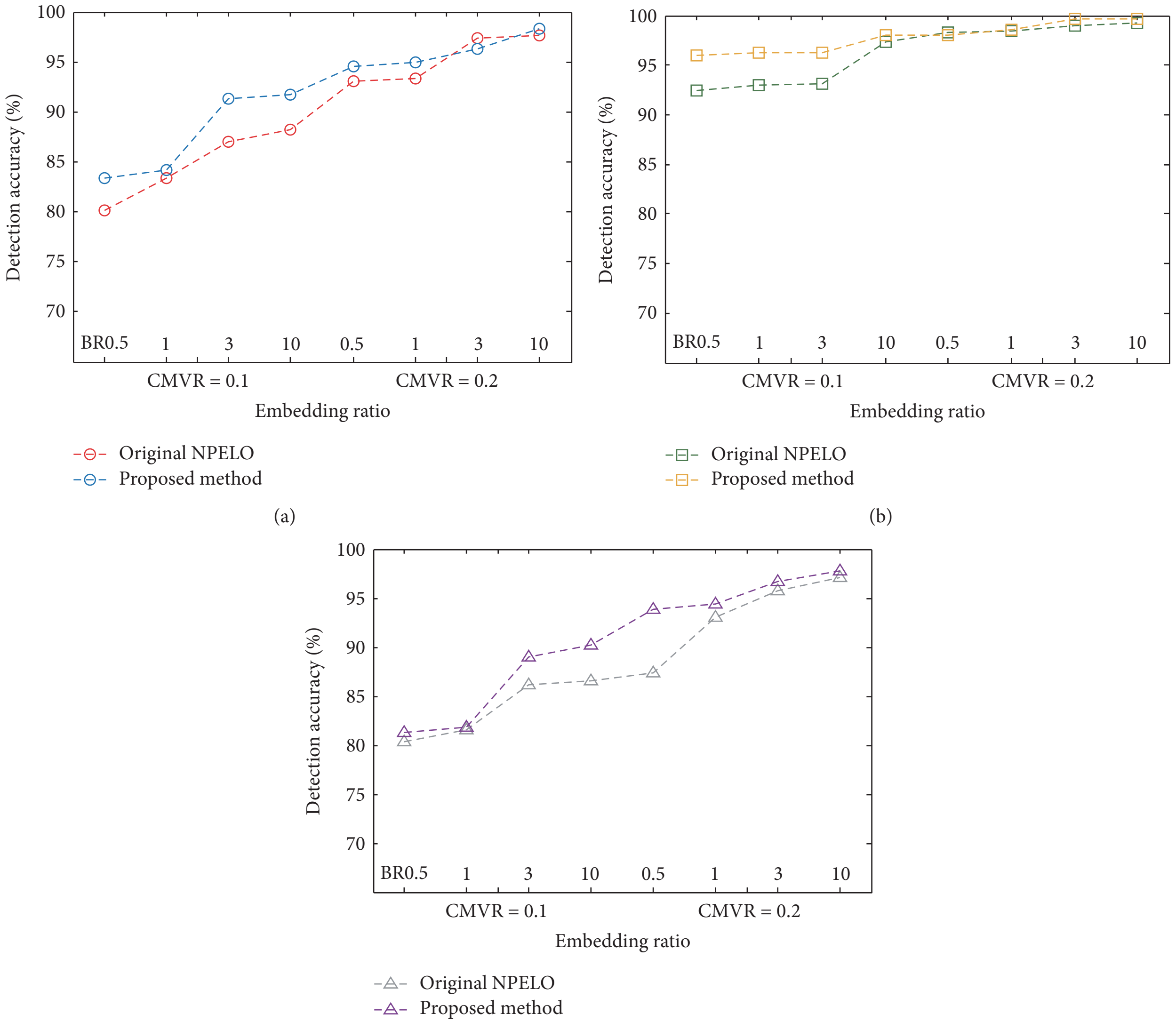
In the voting process, the decision values p_i of subsequences are set in the changing range of 0 to 1, p_i of cover subsequence is 0 and p_i of stego one equals 1. The final decision value is obtained by voting as follows:

$$P = \sum_{j=1}^{N} p_i \dot{w}_i. \tag{6}$$

The input video is detected as a stego video if $P > 0.5$ and cover video otherwise. The final detection result of every single video can be obtained by this fusing manipulation. And the detection performance of the proposed method is evaluated by the detection accuracy of the weighted classifier set, which is the average value of TP and TN of testing video set.

4. Experiments

4.1. Experimental Setup. Our proposed steganalytic scheme is implemented on a well-known H.264/AVC codec named x264 [27]. The video database is composed of 100 standard



(a)

(b)

(c)

Figure 7: Steganalytic performances (%) of NPELO and proposed method against (a) Cao's, (b) Yao's, and (c) Wang's algorithms.

4 : 2 : 0 YUV sequences in CIF format. The raw sequences vary from 150 to 300 frames in length and are coded with 30 fps frame rate.

In order to evaluate our adaptive steganalytic strategy against existing MV-based steganography, Cao's [5], Yao's [4], and Wang's [6] methods are implemented to generate the class of stego videos. The embedding ratio is denoted by corrupted MV ratio (CMVR), which represents the ratio of corrupted MVs' number to the total number of MVs in each frame. Various bitrates (BR) including 0.5 Mbps, 1 Mbps, 3 Mbps, and 10 Mbps are considered with the achieved embedding ratio (ER) of CMVR = 0.1 and CMVR = 0.2, respectively.

In our experiments, the current best steganalytic features NPELO [18] and MVRBR [17] are leveraged to extract features from cover and stego samples. We randomly select 50 percent pairs of videos for training and the remaining ones for testing. Each training and testing is repeated several times and average detection accuracy is used to evaluate the final performance. Moreover, Chang and Lin's support vector machine (SVM) [28] with Gaussian kernel is utilized as classifier.

4.2. Results and Discussion. First, in order to investigate the relationship between categories' number and steganalytic performance, we test the proposed scheme under the conditions of different categories' number. The stego samples are generated by Cao's, Yao's, and Wang's steganographic methods at the CMVR of 0.2. And the bitrate is set at 10 Mbps. Both of improved NPELO and improved MVRBR are utilized for feature extraction. By assigning different values to τ_i in Formula (4), the video subsequences can be segmented to several categories. And the credible MV is defined by the threshold $\delta = 100$ in Formula (1).

Table 1 records the detection accuracies with the corresponding values of τ_i. As illustrated in Figure 6, the steganalytic performance improves with the increase of categories'

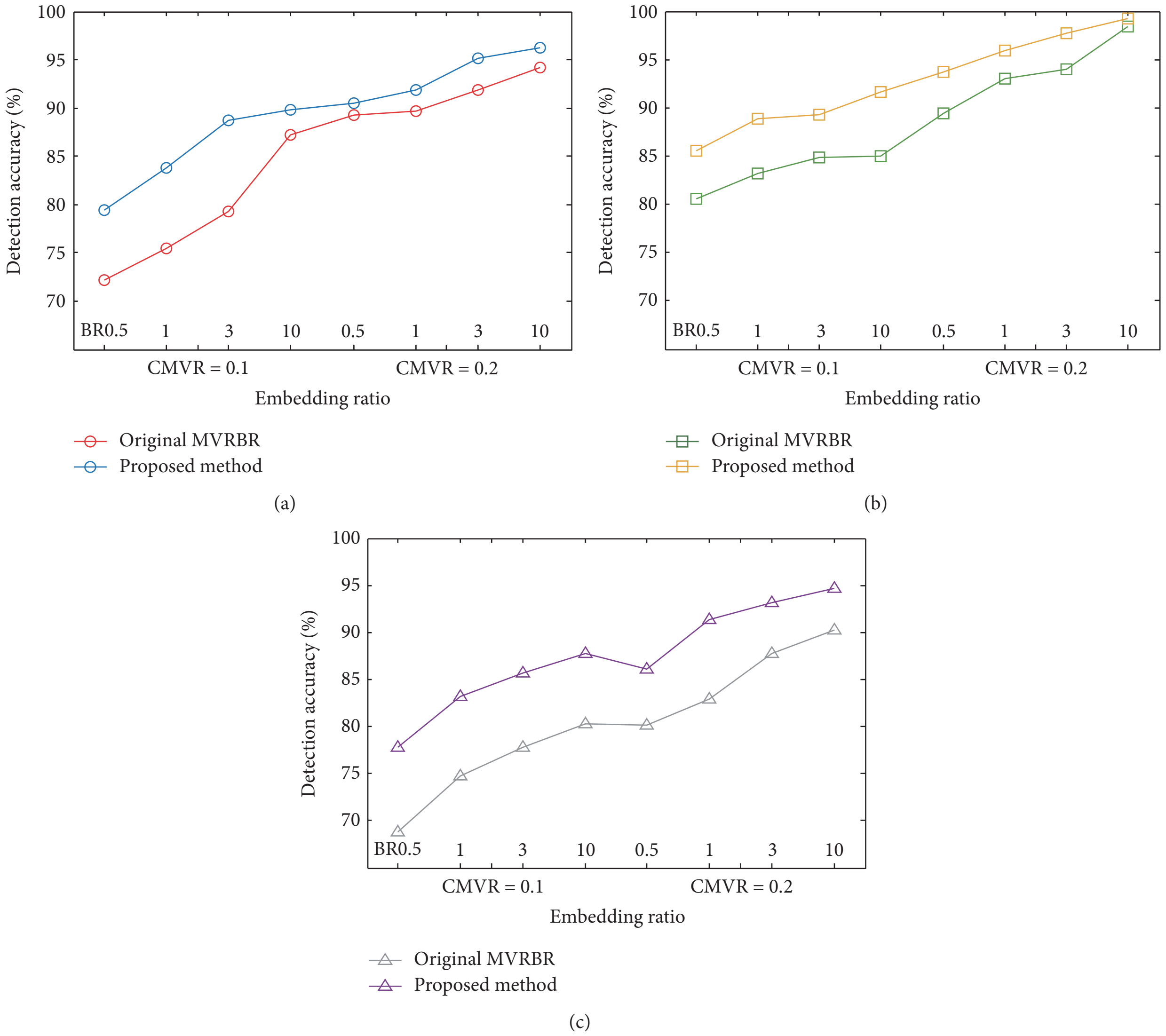


Figure 8: Steganalytic performances (%) of MVRBR and proposed method against (a) Cao's, (b) Yao's, and (c) Wang's algorithms.

number and stabilizes at the number 3 to 5. With the purpose of reducing computation and time complexity, the video subsequences are segmented to three categories of low, middle, and high motion intensity with the parameters $\tau_1 = 2$ and $\tau_2 = 6$ in the following experiments.

Because the statistical characteristics of MVs could be significantly influenced by the variations in coding parameters and embedding ratios, the steganalytic performances against Cao's, Yao's, and Wang's steganography are further evaluated under the various configurations of bitrates and CMVRs.

The detection accuracies of the original NPELO and the proposed methods against current MV-based steganographic algorithms are recorded in Table 2. Figure 7 depicts the comparison of their performances in detecting Cao's, Yao's, and Wang's methods, respectively. It can be seen that the proposed approach performs better than NPELO method when detecting all the three steganographic schemes. Our method achieves the detection accuracy of 99.72% when detecting Yao's scheme with CMVR = 0.2 and 3 Mbps. And the maximum detection accuracies against Cao's and Wang's methods are up to 98.34% and 97.77%.

The performances of original and improved MVRBR are illustrated in Figure 8. It is shown that the proposed approach can effectively improve the accuracy in detecting above three steganographic methods. From Table 3, we can observe that the maximum accuracies are achieved when CMVR is 0.2 and bitrate is 10 Mbps. And the maximum accuracies are equal to 96.19%, 99.24%, and 94.73% against Cao's, Yao's, and Wang's steganographic methods. The steganalytic performances of these methods meliorate with the increase of QP. It is because less losses are induced in higher quality videos, and the features extracted from these subsequences are more effective.

The average improvement percentage of NPELO and MVRBR is 2.02% and 6.36%, respectively. Consequently our approach can effectively improve the performance of

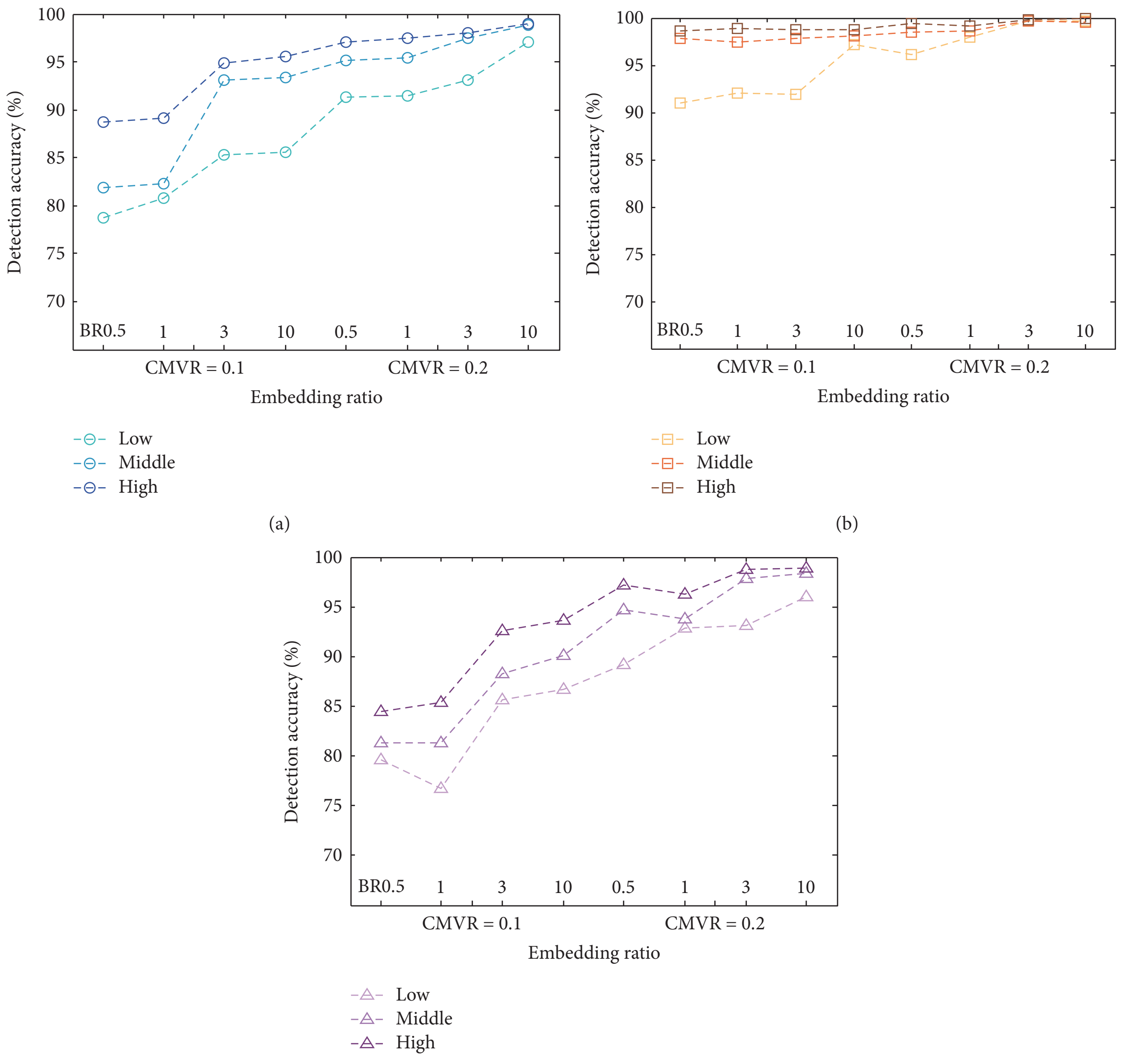


Figure 9: Three categories' steganalytic performances (%) of improved NPELO against (a) Cao's, (b) Yao's, and (c) Wang's algorithms.

steganalytic feature, especially in the cases of low bitrates and low embedding ratios.

Moreover, in order to analyze the impacts of three categories on overall result, steganalytic performances of high, middle, and low motion intensity are tested. As a result, the detection accuracies are recorded in Table 4. As shown in Figures 9 and 10, the performance of high motion intensity is best among the three categories, whereas the low motion intensity performs worst. Correspondingly, the weight value of third classifier is largest and small weight value is assigned to the first classifier. Therefore, the category of high motion intensity makes greatest contribution to the improvement of steganalytic performance, followed by the one of middle motion intensity and the low motion intensity's contribution is least.

5. Conclusion and Future Works

In this paper, a segmentation based steganalytic scheme aimed at MV-based video steganography is proposed. In order to reduce the difference of statistical characteristics caused by diverse video content, the input videos are segmented to subsequences according to block's motion in each frame. Then features are, respectively, extracted from every category of subsequences with close motion intensity, which are used to train one classifier independently. In the testing process, the steganalytic features of each category are sent to the corresponding classifier and the final decision is made through a weighted fusing process. The results of the experiments have shown that the proposed algorithm can effectively improve the performance of video steganalysis, especially for

Table 1: Detection accuracies (%) of improved NPELO and improved MVRBR using different number of categories.

Number of categories	Parameter value $\tau_i \in$	Improved NPELO			Improved MVRBR		
		Cao's	Yao's	Wang's	Cao's	Yao's	Wang's
1	$\{\emptyset\}$	97.60	98.47	97.16	94.15	98.39	90.21
2	$\{4\}$	97.86	98.88	97.50	95.50	98.41	91.68
3	$\{2, 6\}$	98.34	99.70	97.77	96.19	99.24	94.73
4	$\{2, 4, 6\}$	98.28	99.61	97.93	95.93	99.35	94.64
5	$\{1.5, 3, 4.5, 6\}$	98.44	99.35	97.62	96.62	99.01	94.89
6	$\{1.2, 2.4, 3.6, 4.8, 6\}$	98.24	98.61	98.20	96.20	98.61	93.84

Table 2: Detection accuracies (%) of NPELO and proposed method against Cao's, Yao's, and Wang's steganographic algorithms.

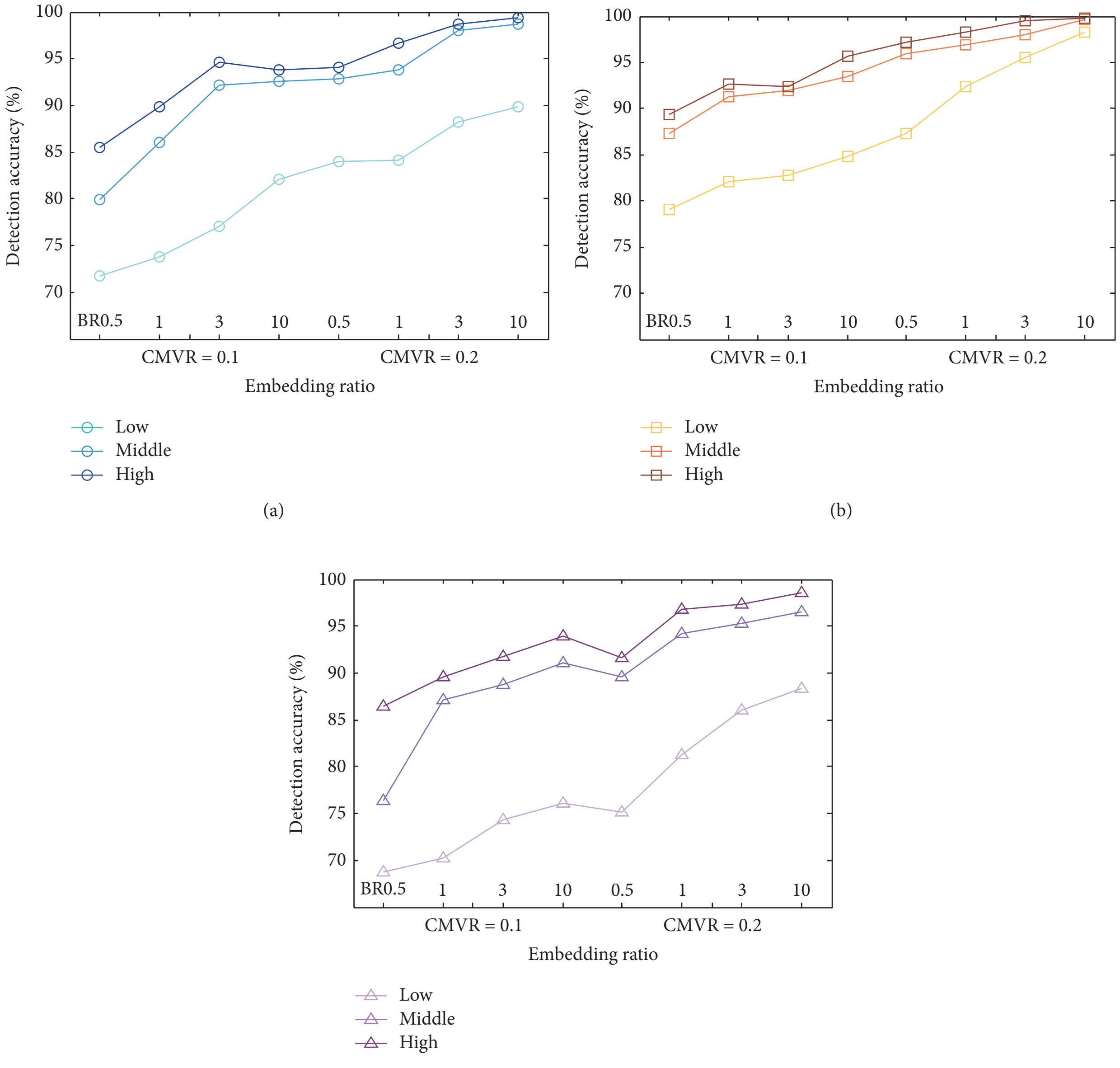
Steganography	BR (Mbps)	0.5		1		3		10	
	ER	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2
Cao's	NPELO	80.11	93.09	83.38	93.30	87.00	97.39	88.17	97.60
	Proposed	83.33	94.60	84.22	94.89	91.26	96.26	91.69	**98.34**
Yao's	NPELO	92.43	98.34	93.00	99.01	93.19	99.19	97.33	98.47
	Proposed	95.96	98.02	96.24	98.59	96.24	**99.72**	97.98	99.70
Wang's	NPELO	80.30	87.32	81.51	93.11	86.11	95.73	86.57	97.16
	Proposed	81.83	93.80	81.27	94.34	88.93	96.65	90.20	**97.77**

Table 3: Detection accuracies (%) of MVRBR and proposed method against Cao's, Yao's, and Wang's steganographic algorithms.

Steganography	BR (Mbps)	0.5		1		3		10	
	ER	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2
Cao's	MVRBR	72.23	89.33	75.41	89.69	79.32	91.91	87.26	94.15
	Proposed	79.40	90.52	83.77	91.83	88.67	95.21	89.80	**96.19**
Yao's	MVRBR	80.63	89.43	83.23	93.01	84.79	96.92	85.04	98.39
	Proposed	85.48	93.69	88.90	95.89	89.27	97.72	91.57	**99.24**
Wang's	MVRBR	68.83	80.18	74.79	82.91	77.82	87.78	80.32	90.21
	Proposed	77.86	86.04	83.22	91.30	85.62	93.14	87.73	**94.73**

Table 4: Three categories' detection accuracies (%) of improved NPELO and improved MVRBR.

Steganalysis	Steganography	BR (Mbps)	0.5		1		3		10	
		ER	0.1	0.2	0.1	0.2	0.1	0.2	0.1	0.2
Improved NPELO	Cao's	Slow	78.78	91.32	80.78	91.51	85.29	93.11	85.57	97.10
		Middle	81.92	95.19	82.28	95.47	93.03	97.53	93.38	98.88
		High	88.67	97.01	89.14	97.47	94.90	98.00	95.53	99.01
	Yao's	Slow	91.01	96.10	92.05	97.96	91.94	99.66	97.13	99.60
		Middle	97.84	98.49	97.50	98.64	97.86	99.68	98.09	99.61
		High	98.66	99.41	98.91	99.17	98.69	99.82	98.72	99.90
	Wang's	Slow	79.65	89.13	76.77	92.85	85.61	93.16	86.74	96.07
		Middle	81.28	94.68	81.26	93.79	88.30	97.86	90.11	98.31
		High	84.43	97.22	85.34	96.32	92.60	98.74	93.59	98.89
Improved MVRBR	Cao's	Slow	71.70	83.99	73.83	84.16	77.03	88.21	82.07	89.88
		Middle	79.87	92.79	85.98	93.81	92.21	98.03	92.60	98.73
		High	85.44	94.11	89.83	96.60	94.68	98.66	93.80	99.38
	Yao's	Slow	79.04	87.24	82.05	92.37	82.74	95.58	84.78	98.31
		Middle	87.36	95.99	91.22	96.85	92.03	97.96	93.53	99.59
		High	89.34	97.20	92.66	98.25	92.39	99.54	95.68	99.81
	Wang's	Slow	68.81	75.14	70.28	81.27	74.31	86.04	76.09	88.36
		Middle	76.36	89.53	87.08	94.21	88.78	95.29	91.12	96.59
		High	86.40	91.56	89.62	96.78	91.73	97.30	93.88	98.62



(a)

(b)

(c)

Figure 10: Three categories' steganalytic performances (%) of improved MVRBR against (a) Cao's, (b) Yao's, and (c) Wang's algorithms.

low bitrate videos which are embedded at low embedding ratio.

In our future work, a larger database will be used for evaluation of our method and the videos are to be classified into more categories with different motion intensity. And the implementations on other video coding standards such as H.265/HEVC will be further considered.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the NSFC under U1636102 and U1536105 and National Key Technology R&D Program under 2014BAH41B01, 2016YFB0801003, and 2016QY15Z2500.

References

[1] F. Jordan, M. Kutter, and T. Ebrahimi, "Proposal of a watermarking technique for hiding data in compressed and decompressed video," *Multiresolution Watermarking for Images; Video A Unified Approach*, 1997.

[2] C. Xu, X. Ping, and T. Zhang, "Steganography in compressed video stream," in *Proceedings of the 1st International Conference*

on Innovative Computing, Information and Control 2006, ICICIC'06, pp. 269–272, September 2006.

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

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

[5] Y. Cao, H. Zhang, X. Zhao, and H. Yu, "Video steganography based on optimized motion estimation perturbation," in *Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security, IHMMSec '15*, pp. 25–31, New York, NY, USA, 2015.

[6] P. Wang, H. Zhang, Y. Cao, and X. Zhao, "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, IHMMSec '16*, pp. 127–137, New York, NY, USA, 2016.

[7] S. K. Kapotas and A. N. Skodras, "A new data hiding scheme for scene change detection in H.264 encoded video sequences," in *Proceedings of the 2008 IEEE International Conference on Multimedia and Expo, ICME 2008*, pp. 277–280, June 2008.

[8] X. Ma, Z. Li, J. Lv, and W. Wang, "Data hiding in H.264/AVC streams with limited intra-frame distortion drift," in *Proceedings of the 1st International Symposium on Computer Network and Multimedia Technology, CNMT 2009*, pp. 1–5, December 2009.

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

[10] T.-J. Lin, K.-L. Chung, P.-C. Chang, Y.-H. Huang, H.-Y. M. Liao, and C.-Y. Fang, "An improved DCT-based perturbation scheme for high capacity data hiding in H.264/AVC intra frames," *Journal of Systems and Software*, vol. 86, no. 3, pp. 604–614, 2013.

[11] S. Kim, S. Kim, Y. Hong, and C. Won, "Data Hiding on H.264/AVC Compressed Video," pp. 698–707, Springer Berlin Heidelberg, Berlin, Germany, 2007.

[12] K. Liao, S. Lian, Z. Guo, and J. Wang, "Efficient information hiding in H.264/AVC video coding," *Telecommunication Systems*, vol. 49, no. 2, pp. 261–269, 2012.

[13] J. Fridrich, M. Goljan, P. Lisonek, and D. Soukal, "Writing on wet paper," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3923–3935, 2005.

[14] T. Filler, J. Judas, and J. Fridrich, "Minimizing embedding impact in steganography using trellis-coded quantization," in *Proceedings of the Media Forensics and Security II*, USA, January 2010.

[15] T. Filler, J. Judas, and J. Fridrich, "Minimizing additive distortion in steganography using syndrome-trellis codes," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 920–935, 2011.

[16] K. Wang, H. Zhao, and H. Wang, "Video steganalysis against motion vector-based steganography by adding or subtracting one motion vector value," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 5, pp. 741–751, 2014.

[17] P. Wang, Y. Cao, X. Zhao, and B. Wu, "Motion vector reversion-based steganalysis revisited," in *Proceedings of the IEEE China Summit and International Conference on Signal and Information Processing, China SIP 2015*, pp. 463–467, China, 2015.

[18] H. Zhang, Y. Cao, and X. Zhao, "A steganalytic approach to detect motion vector modification using near-perfect estimation for local optimality," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 2, pp. 465–478, 2017.

[19] W. Luo, F. Huang, and J. Huang, "Edge adaptive image steganography based on lsb matching revisited," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 201–214, 2010.

[20] T. Pevný, T. Filler, and P. Bas, "Using high-dimensional image models to perform highly undetectable steganography," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 6387, pp. 161–177, 2010.

[21] G. Xiong, X. Ping, T. Zhang, and X. Hou, "Image textural features for steganalysis of spatial domain steganography," *Journal of Electronic Imaging*, vol. 21, article 033015, no. 3, 2012.

[22] H. Amirkhani and M. Rahmati, "New framework for using image contents in blind steganalysis systems," *Journal of Electronic Imaging*, vol. 20, article 013016, no. 1, 2011.

[23] S. Cho, B.-H. Cha, J. Wang, and C.-C. J. Kuo, "Block-based image steganalysis: Algorithm and performance evaluation," in *Proceedings of the 2010 IEEE International Symposium on Circuits and Systems: Nano-Bio Circuit Fabrics and Systems, ISCAS 2010*, pp. 1679–1682, France, June 2010.

[24] R. Wang, X. Ping, S. Niu, and T. Zhang, "Segmentation Based Steganalysis of Spatial Images Using Local Linear Transform," in *Digital Forensics and Watermarking*, vol. 10082 of *Lecture Notes in Computer Science*, pp. 533–549, Springer International Publishing, Cham, 2017.

[25] H. Wang, A. Klaser, C. Schmid, and C. L. Liu, "Dense trajectories and motion boundary descriptors for action recognition," *International Journal of Computer Vision*, vol. 103, no. 1, pp. 60–79, 2013.

[26] M. Vlachos, G. Kollios, and D. Gunopulos, "Discovering similar multidimensional trajectories," in *Proceedings of the 18th International Conference on Data Engineering*, pp. 673–684, San Jose, Calif, USA, March 2002.

[27] VideoLan, http://www.videolan.org/developers/x264.html.

[28] C. Chang and C. Lin, "LIBSVM: A Library for Support Vector Machines 2001," http://www.csie.ntu.edu.tw/~cjlin/libsvm/.

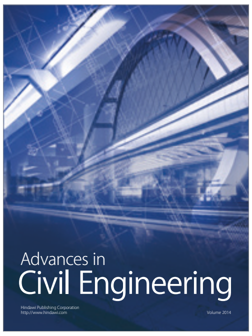
Journal of
Engineering

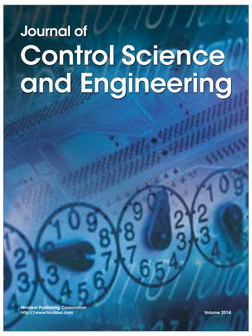
The Scientific World Journal

International Journal of
Rotating Machinery

Journal of
Sensors

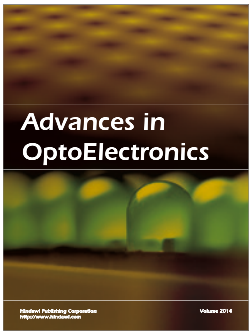
International Journal of
Distributed Sensor Networks


Advances in
Civil Engineering


Journal of
Control Science and Engineering

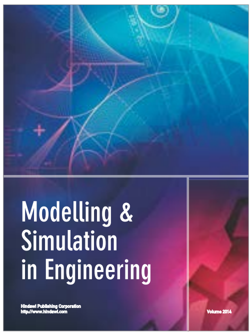
Journal of
Robotics

Journal of
Electrical and Computer Engineering

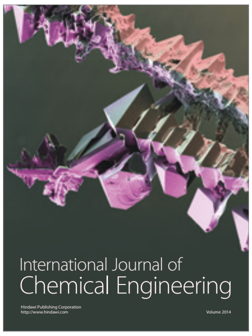

Advances in
OptoElectronics

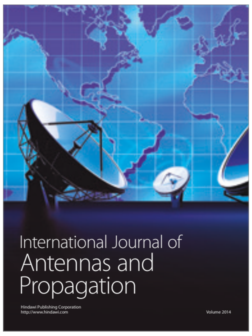
VLSI Design

International Journal of
Navigation and Observation

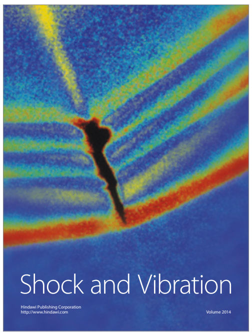

Modelling & Simulation in Engineering

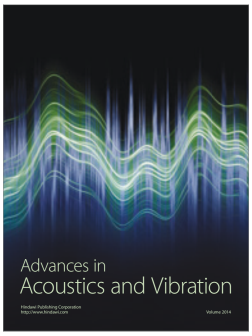
International Journal of
Aerospace Engineering


International Journal of
Chemical Engineering


International Journal of
Antennas and Propagation

Active and Passive
Electronic Components


Shock and Vibration


Advances in
Acoustics and Vibration