IET Information Security

Research Article

ISSN 1751-8709
Received on 9th November 2017
Revised 4th April 2018
Accepted on 4th May 2018
E-First on 13th July 2018
doi: 10.1049/iet-ifs.2017.0555
www.ietdl.org

Detection of double compression in HEVC videos based on TU size and quantised DCT coefficients

Qian Li[1,2] ✉, *Rangding Wang*[2], *Dawen Xu*[3]

[1]*College of Information Engineering, Ningbo Dahongying University, No. 899, Xueyuan Road, Haishu District, Ningbo 315175, People's Republic of China*

[2]*CKC Software Laboratory, Ningbo University, No. 818, Fenghua Road, Jiangbei District, Ningbo 315211, People's Republic of China*

[3]*School of Electronics and Information Engineering, Ningbo University of Technology, No. 201, Fenghua Road, Jiangbei District, Ningbo 315211, People's Republic of China*

✉ *E-mail: liqian_mine@126.com*

Abstract: With the advent of sophisticated and low-cost video editing software, digital videos are highly vulnerable to be tampered. The authenticity and integrity identification of digital videos is an urgent issue. In this study, an effective method to detect double High Efficiency Video Coding (HEVC) video compression with different quantisation parameter (QP) is proposed, which often occurs in the video tampering process. First, the effects of QP on the distributions of Discrete Cosine Transform (DCT) coefficients and Transform Unit (TU) size are analysed. Then a feature set including 17 features is derived from quantised DCT coefficients and TU size. It can characterize the statistical differences between single and double compressed videos. Finally, the Library for Support Vector Machine classifier is exploited to identify whether a given HEVC video has been double compressed or not. Experimental results demonstrate that the authors' detection method has a good comprehensive performance.

1 Introduction

Nowadays, the usage of surveillance systems and video cameras is growing rapidly. Digital videos play an important role in our daily life. However, modifying a digital video without any obvious traces is easily accomplished by using video editing software, such as Adobe Premiere. Consequently, there is an increasing trend in both the quality and quantity of doctored videos, which gives rise to more and more significant technological, legal and ethical issues. Therefore, ascertaining the authenticity and integrity of the digital video information is an urgent problem. Digital video forensics techniques are in full swing currently. Generally, videos are stored and transmitted in a compressed way due to the large amount of video data. However, the tampering process needs to be performed on uncompressed domain, and the tampered video would be re-encoded and re-saved in compressed format to avoid suspicion. So the double compression detection is an important forensic tool for recovering video processing history [1].

In the past few years, most of the studies were focused on double JPEG compression [2–7] and double MPEG compression [8–11]. Wang and Farid [8] detected double MPEG-2 video compression with variable bit rate mode by examining the period artefacts introduced into the Discrete Cosine Transform (DCT) histogram of I frames. In [9, 10], the violation of Benford law for the first digit distribution of non-zero quantised AC coefficients is applied to detect double MPEG-2 compression. Jiang *et al.* [11] proposed a method to detect double MPEG-4 video compression by modelling Markov transition probability matrix based on DCT coefficients. Liao *et al.* [12] used the probabilities of 20 non-zero quantised AC coefficients as features to distinguish double H.264 compressed videos from single H.264 compressed videos. Jia *et al.* [13] proposed a novel method based on the same quantisation parameter (QP) for double High Efficiency Video Coding (HEVC) compression detection. The authors extract the number for the 4×4 PU (Prediction Unit) blocks and calculate the standard deviation of 4×4 PU blocks difference before and after the compression. Finally, they select the appropriate threshold for double compression detection.

As a latest compression coding standard, HEVC is aimed at significantly improving encoding efficiency on the basis of H.264/AVC and especially deal with high-definition video [14]. Currently, some companies of cameras and surveillance systems (such as HIKVISION) are exploiting HEVC coding standard to develop new products. So it will be a prevalent coding standard in the future. However, few works are designed for the detection of double HEVC video compression. Hence, it is very important to investigate the detection methods of double compressed HEVC videos. In our previous algorithm [15], we proposed a method to detect double HEVC videos based on the co-occurrence matrix of DCT coefficients. However, the feature dimension of the algorithm is high. Therefore, in this paper, we present an effective detection method of double compression with low feature dimension, which utilises the combined features constructed by quantised DCT coefficients and Transform Unit (TU) size. Taking into account several aspects (computational complexity, feature dimensions and detection accuracies), we give a comprehensive detection performance of our method compared to reference [11, 12, 15].

The remaining of this paper is organised as follows. Section 2 briefly reviews the process of HEVC intra coding. Section 3 represents in detail the proposed scheme, which mainly analyses the effect of QP on quantised DCT coefficients and TU size during the process of HEVC coding. Section 3 also discusses extraction of the features for detecting double HEVC videos compression. Section 4 shows how to implement the experiments, and gives the detection results. Section 5 summarises this paper.

2 HEVC intra coding

HEVC, as a kind of lossy compression coding, results in that the double compressed videos have some difference with single compressed videos. Incorporating this feature, we attempt to analyse the effect of HEVC compression on coding parameter information. We then select coding parameters to distinguish single compressed videos from double compressed videos. The procedure of HEVC intra-prediction coding will be described in this section.

Compared with the previous coding standard, the conventional hybrid approach is still used in the video coding layer of HEVC.

Table 1 Optional sizes of PU and TU are correspond to current CU

Depth	CU	PU	TU
0	64×64	64×64	$32 \times 32, 16 \times 16$
1	32×32	32×32	$32 \times 32, 16 \times 16, 8 \times 8$
2	16×16	16×16	$16 \times 16, 8 \times 8, 4 \times 4$
3	8×8	$8 \times 8, 4 \times 4$	$8 \times 8, 4 \times 4$

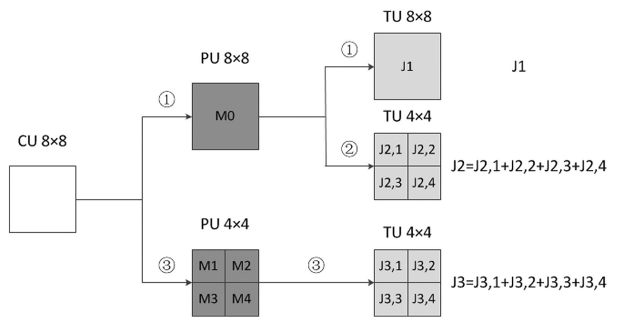


Fig. 1 *Choice of* 8×8 *CU coding mode*

Nonetheless, HEVC has adopted a highly flexible and efficient block partitioning structure by introducing three basic units, i.e. CU (Coding Unit), PU (Prediction Unit) and TU (Transform Unit). For HEVC, videos are encoded frame by frame, and each input frame is divided into several non-overlapped and equal-sized CTUs (Coding Tree Units). HEVC supports a partitioning of CTU into one CU or multiple CUs by using recursive quad-tree coding [16, 17]. Each CU has an associated partitioning into PUs and TUs for prediction and transform, respectively. Table 1 shows that the optional sizes of PU and TU correspond to the size of current CU.

We know that PU and TU partitioning structures have their root at the CU level. PU works as a basic representative block for sharing the prediction information. The best of 35 prediction modes is selected for the current PU as [17] has detailed. TU is a basic unit of residual transform and quantisation. In addition, the actual prediction process operates separately for each TU instead of applying the intra prediction at the PU level. TU structure based on RD cost optimisation is determined by using the selected best prediction mode. All the TUs inside one PU have the same best prediction mode. First, the prediction pixels are achieved by utilising the nearest boundary reference samples from the already decoded neighbouring TUs. Then, as shown in (1), the formed residuals in spatial domain are transformed into frequency coefficients via DCT transform

$$y(x, l) = C(k)C(l)\sum_{m=0}^{N-1}\sum_{n=0}^{N-1} x(m, n)\cos\left(\frac{(2m+1)k\pi}{2N}\right) \times \cos\left(\frac{(2n+1)l\pi}{2N}\right), \quad k, l = 0, 1, \ldots, N-1 \tag{1}$$

where

$$C(k) = C(l) = \begin{cases} \sqrt{(1/N)}, & k, l = 0 \\ \sqrt{(2/N)}, & \text{other} \end{cases},$$

N is the edge length of the current TU, $x(m, n)$ is the residual of TU at the position (m, n), $y(k, l)$ is the DCT coefficient of TU at the position (k, l). Finally, as (2) represents, DCT coefficients are quantised, which vary with given QP

$$q(k, l) = (y(k, l) \times MF[\text{QP mode } 6] + f') \gg (q\,\text{bits} + \text{T\_shift}) \tag{2}$$

where $q\,\text{bits} = 14 + \text{floor}\left(\frac{\text{QP}}{6}\right)$, $MF[x] = \{26214, 23302, 20560, 18396, 16384, 14564\}$, $x = \{0, 1, 2, 3, 4, 5\}$, $\eta = 2^{\text{T\_shift}}$, η represents the scaling factor. $f' = f \ll (q\,\text{bits} + T\_shift)$ indicates the rounding offset, and f is usually equal to $1/3$ for intra frame. $q(k, l)$ denotes the quantised DCT coefficient at the position (k, l), the value of QP is ranged from 0 to 51.

By calculating the Lagrangian RD (Rate Distortion) cost for all possible combined block structures of CU, PU and TU, the one combination with the minimised RD cost that is taken as the optimal coding mode for each CTU. Calculating RD cost should take into account the distortion and the bit requirements for encoding a block. Consider block B. The RD cost J can be achieved by (3), which is related to QP

$$J = \text{SSD}(B|\text{QP}) + \lambda \cdot R(B|\text{QP}) \tag{3}$$

where SSD(B) indicates the sum of squared distortion between original pixels and reconstructed pixels for the block B. λ denotes Lagrangian multiplier. $R(B)$ is the required number of bits to encode block B, which includes quantised DCT coefficients and side information (the coding block partitioning mode, prediction mode, some marks and so on). When B is further divided into N parts, J' represents the total RD cost of N parts as (4) shows. B will be split into N smaller parts if J' is less than J,

$$J' = \sum_{i=1}^{N} (\text{SSD}(B_i|\text{QP}) + \lambda \cdot R(B_i|\text{QP})) \tag{4}$$

In order to vividly illustrate the selection process of code partitioning, the divisions of 8×8 CU are taken for an example like Fig. 1. It has three cases: case 1: when the combination of 8×8 CU, 8×8 PU (the best prediction mode is M_0) and 8×8 TU is selected, its corresponding RD cost is J_1; case 2: when the combination of 8×8 CU, 8×8 PU (the best prediction mode is M_0) and four 4×4 TUs is selected, its corresponding RD cost is $J_2 = J_{2,1} + J_{2,2} + J_{2,3} + J_{2,4}$; case 3: when the combination of 8×8 CU, four 4×4 PUs (the best prediction mode is, respectively, M_1, M_2, M_3, M_4) and four 4×4 TUs is selected, its corresponding RD cost is $J_3 = J_{3,1} + J_{3,2} + J_{3,3} + J_{3,4}$. The combination with $\min(J_1, J_2, J_3)$ is finally selected.

From the above descriptions, it is clear that the block size layout and quantised DCT coefficients are influenced by QP. Therefore, it is necessary to study the differences of block size and quantised DCT coefficients under different QP both theoretically and experimentally in the following section. It helps us to construct some statistics which can distinguish single compressed videos from double compressed videos.

3 Proposed scheme

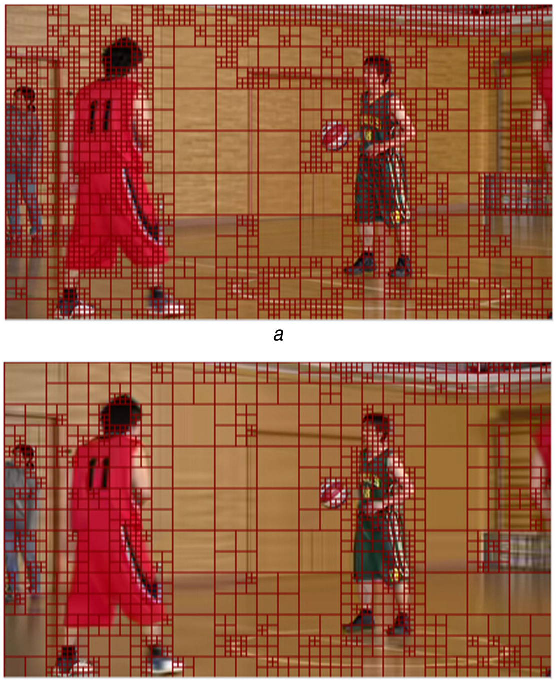
In this section, we propose a double compression detection scheme with different QP. First, we analysed the effect of QP on TU size, and the statistical differences in the distributions of TU size under different QP are shown. However, the selection of TU size is also related to video contents. The distribution of TU size for different videos is volatile, which would affect the classification result. In order to dampen the influence of video content as much as possible, TU size statistic characteristics need to be combined with other valid features. From our previous work [15], we know that the statistical characteristics of quantised DCT coefficients can effectively detect double video compression. Therefore, the TU statistic characteristics and the probability of DCT coefficients constitute the detection features of the algorithm. Finally, we employ SVM (Support Vector Machine) to learn the discriminability from the extracted features.

3.1 Effect of QP on TU size

In this paper, we assume that the double compressed videos have experienced two successive encodings with different QP. Double compressed videos certainly retain part of evidences from the first compression, which inevitably leave detectable artefacts. Analysing statistical differences in the same videos compressed with different QP are significant to propose an algorithm of detecting double HEVC compressed videos.

The irreversible rounding error is produced during the quantisation process of HEVC video compression coding. The magnitude of error is affected by QP, so the distributions of TU size under different QP are different. The influence of different QP on the size of TU will be discussed completely.

Section 2 has described in detail that the optimal block partitioning mode is decided by the principle of minimum RD cost. The calculation of RD cost for encoding a block depends on two factors related to QP: the distortion and the bit requirement consisted of non-zero quantised DCT coefficients and related side



a

b

Fig. 2 *Layout of TU size for Basketballpass*_416 × 240_50

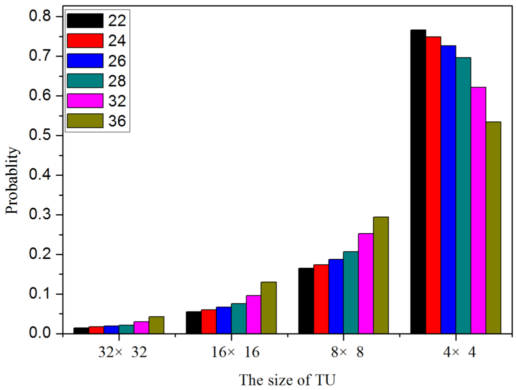


Fig. 3 *Probability distribution of TU size under different QP*

information. It means that the block size layout is affected by QP. We can thus analyse the relationship between TU size and QP for double compression detection.

If the video is compressed by a large QP, the number of non-zero DCT coefficients for a coding block is smaller. That is to say, the code bits for DCT coefficients are little. So the total bits requirement for a block mainly depends on side information. However, there needs a lot of bits for encoding side information if the coding block is further split into several smaller blocks, which may result in a higher RD cost. Therefore, a larger TU size tends to be selected in the case of higher QP. As a demonstration, Figs. 2*a* and *b* show, respectively, the layout of TU size for video compressed with QP = 22 and QP = 36. The results are achieved by using the first frame of the standard YUV sequence which is named Basketballpass_416 × 240_50. The smaller TU in Fig. 2*a* is relative to the distribution of TU size in Fig. 2*b*.

In order to better illustrate the difference in the probability distribution of TU size under different QP, we calculate the probability distribution of TU size for each video, and the mean of 361 videos is shown in Fig. 3. The legend 22, 24, 26, 28, 32, 36 represent the values of the QP. We can conclude from Fig. 3 that the probability of smaller TU size becomes larger as the value of QP increases.

However, according to the coding standard, the selection of TU size is also related to video contents. As Fig. 4 shows, the layouts of TU size for different video frames are different. Therefore, we combine the quantised DCT coefficients in order to dampen the influence of the video content as much as possible.

3.2 Effect of QP on quantised DCT coefficients

From Section 2, the value of quantised DCT coefficients is related to QP as in (2). If QP is higher, there is greater loss of detailed information, and the value range of quantised DCT coefficients is more concentrate. On the contrary, if QP is lower, there is less loss of detailed information, and the value range of quantised DCT coefficients is more disperse. Simultaneously, we plot the distributions of quantised DCT coefficients under different QP from the experiments. Fig. 5 presents only revealed the probabilities of quantised DCT coefficients ranged from −10 to 10, it is the statistical average of 361 videos. In the case of any QP, most of DCT coefficients fall into a small interval and obey the Laplace distribution with mean of 0. However, a lot of non-zero quantised DCT coefficients become zero as QP increased. There is a salient difference in the probability distributions of quantised DCT coefficients under different QP.

3.3 Feature extraction

In this subsection, a set of features constructed by TU size and quantized DCT coefficients is used to characterise the differences between single and double compressed videos. These features are sent to LIBSVM classifier to detect double HEVC compressed videos under different QP.

In order to express the distribution of TUs in one CTU more concise and intuitive, $32 \times 32, 16 \times 16, 8 \times 8, 4 \times 4$ TUs are identified separately by the matrix $\boldsymbol{T}_1, \boldsymbol{T}_2, \boldsymbol{T}_3, \boldsymbol{T}_4$. The matrix $\boldsymbol{T}_1, \boldsymbol{T}_2, \boldsymbol{T}_3, \boldsymbol{T}_4$ has dimensions $8 \times 8, 4 \times 4, 2 \times 2, 1 \times 1$, all elements

a *b*

Fig. 4 *Layout of TU size for different video frames (QP = 36)*

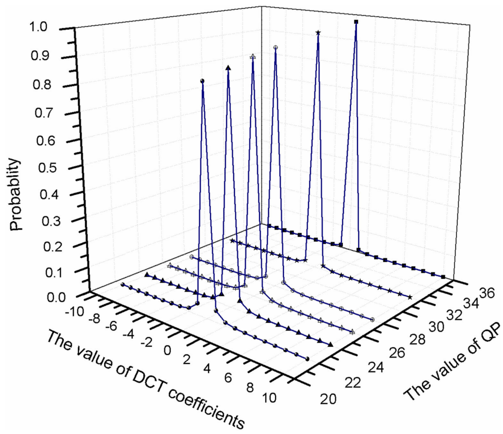


Fig. 5 *Probability distributions of quantised DCT coefficients under different QP*

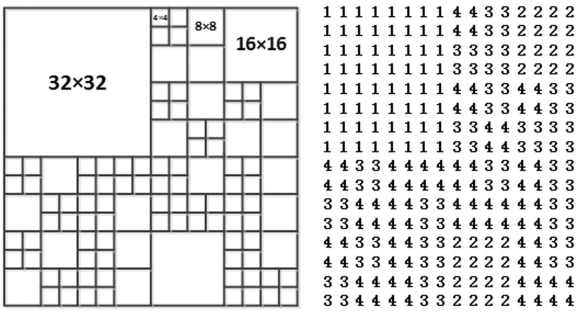


Fig. 6 *Corresponding relationship between TU layout and TU identification matrix for one CTU*

of which are equal to 1, 2, 3 and 4, respectively. In other words, the symbols 1, 2, 3, 4 in Fig. 6 represent $32\times 32, 16\times 16, 8\times 8, 4\times 4$, respectively. For each CTU, the TU layout may be different in the types of TU sizes and the number of TUs. Fig. 6 symbolically depicts the corresponding relationship between the TU layout and the TU identification matrix for one CTU.

Features extraction is shown as follows. Given a compressed video, 64×64 CTU is the basic processing unit used to specify the decoding process for HEVC. Only considering luminance component, the quantised DCT coefficients submatrix $\boldsymbol{D}_n$ and the TU identification submatrix $\boldsymbol{T}_n$ for each CTU are extracted in order. All of $\boldsymbol{D}_n$ and $\boldsymbol{T}_n$ are arranged in two-dimensional (2D) array $\boldsymbol{D}$ and $\boldsymbol{T}$, respectively. $\boldsymbol{D}$ and $\boldsymbol{T}$ are separately defined as follows:

$$\boldsymbol{D}=\begin{Bmatrix} D_1\\ D_2\\ \cdots\\ D_n \end{Bmatrix},\quad \boldsymbol{D}_n=\begin{Bmatrix} D_{n,1,1} & D_{n,1,2} & \cdots & D_{n,1,64}\\ D_{n,2,1} & D_{n,2,2} & \cdots & D_{n,2,64}\\ \cdots & \cdots & \ddots & \cdots\\ D_{n,64,1} & D_{n,64,2} & \cdots & D_{n,64,64} \end{Bmatrix},\quad n=1,2,\ldots,q \tag{5}$$

$$\boldsymbol{T}=\begin{Bmatrix} T_1\\ T_2\\ \cdots\\ T_n \end{Bmatrix},\quad \boldsymbol{T}_n=\begin{Bmatrix} T_{n,1,1} & T_{n,1,2} & \cdots & T_{n,1,16}\\ T_{n,2,1} & T_{n,2,2} & \cdots & T_{n,2,16}\\ \cdots & \cdots & \ddots & \cdots\\ T_{n,16,1} & T_{n,16,2} & \cdots & T_{n,16,16} \end{Bmatrix},\quad n=1,2,\ldots,q \tag{6}$$

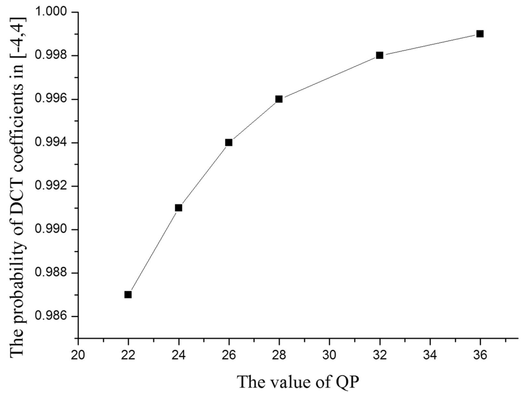


Fig. 7 *Sums of probability for DCT* $\in[-4,4]$ *under different QP*

where $\boldsymbol{D}_n$ and $\boldsymbol{T}_n$, respectively, denote DCT coefficients submatrix and the identification submatrix of TU size in the nth CTU. $D_{n,i,j}$ indicates the value of DCT coefficients in the ith row jth column of $\boldsymbol{D}_n$, and $T_{n,i,j}$ represents the TU identification element in the ith row jth column of $\boldsymbol{T}_n$, q is the number of CTUs, which depends on how many video frames have been encoded and the resolution of the video. Suppose that the resolution of given video is 416×240, the number of CTUs is $q=\left\lceil\frac{416}{64}\right\rceil\times\left\lceil\frac{240}{64}\right\rceil\times N=7\times 4\times N=28N$, in the case of encoding N frames. The result of $\lceil x\rceil$ is an integer which is greater than or equal to x.

A combined feature set $\boldsymbol{F}$ consists of two subsets. The first feature subset is directly derived from $\boldsymbol{D}$. Section 3.2 indicates that the probability distribution of DCT coefficients can be used for double HEVC video compression detection. Besides, the sum of probabilities of DCT coefficients ranged from −4 to 4 is about 99% for different QP, as shown in Fig. 7. So only the probabilities of DCT coefficients in the fixed interval [−4, 4] are treated as the first nine features of $\boldsymbol{F}$. The calculations are given by the following equations:

$$P_{\rm D}(v)=\frac{\sum_{n=1}^{q}\sum_{i=1}^{64}\sum_{j=1}^{64}\delta(D_{n,i,j}=v)}{q\times 64\times 64},\quad v\in[-4,4] \tag{7}$$

$$F(k)=P_D(k-5),\quad k\in[1,9] \tag{8}$$

where $\delta(X=x)=\begin{cases}1, & X=x\\ 0, & \text{else}\end{cases}$, $P_{\rm D}(v)$ denotes the occurrence probability of DCT coefficient v, $F(k)$ is defined as the kth feature of $\boldsymbol{F}$.

Apart from DCT coefficients, the second feature subset constructed by the TU size can also capture statistical differences between single and double compressed videos. The probabilities of $32\times 32, 16\times 16, 8\times 8, 4\times 4$ TUs are denoted by $P_{\rm T}(1), P_{\rm T}(2), P_{\rm T}(3), P_{\rm T}(4)$. According to the identification matrix $\boldsymbol{T}$ of TU size, it can be calculated as follows:

$$P_{\rm T}(x)=\frac{N(x)}{S},\quad x=1,2,3,4 \tag{9}$$

where $N(x)=\frac{\sum_{n=1}^{q}\sum_{i=1}^{16}\sum_{j=1}^{16}\delta(T_{n,i,j}=x)}{w(x)}$, $x=1,2,3,4$, $S=\sum_{x=1}^{4}N(x)$, $w(1)=64$, $w(2)=16$, $w(3)=4$, $w(4)=1$. Simultaneously, the other statistics of TU size are derived, which contains mean, variance, kurtosis and skewness of TU size. They are denoted by M, V, Kur and Skew, and calculated by the following equations:

$$M=\sum_{x=1}^{4}P_{\rm T}(x)\times x \tag{10}$$

$$V = \sum_{x=1}^{4} P_{\mathrm{T}}(x) \times (x - P_{\mathrm{T}}(x))^2 \tag{11}$$

$$\mathrm{Kur} = \frac{S}{S-1} \times \sum_{x=1}^{4} P_{\mathrm{T}}(x) \times \left(\frac{x-M}{\sqrt{V}}\right)^4 \tag{12}$$

$$\mathrm{Skew} = \sum_{x=1}^{4} P_{\mathrm{T}}(x) \times \left(\frac{x-M}{\sqrt{V}}\right)^3 \tag{13}$$

The TU statistic characteristics mentioned above serve as another eight features of $\boldsymbol{F}$. We calculate the probability distributions histograms of DCT coefficients and TU size to describe the differences of single and double compressed videos in Figs. 8 and 9. In Fig. 8, legend 22 represents that the video has undergone single compression, and QP is 22; legend 24-22 represents that the video has undergone double compression, where the QP used for the first compression is 24 and the second compression is 22. The illustrations of other legends in Figs. 8 and 9 are similar. Both the probabilities of occurrence of the DCT coefficients and the distribution of the TU size exhibit are great difference between the single and the double compressed videos.

4 Experiments

4.1 Establishment of sample library

In our experiments, 30 widely known YUV standard sequences, including six different resolutions are selected as source sequences, as shown in Table 2. Moreover, we use video editing package to split each YUV sequence into several non-overlapped video clips to increase sample quantity. We obtain 361 video clips in total, with each clip consisting of 30 frames.

In order to obtain the double compression videos, all of 361 raw video clips are first encoded with QP1, the values of QP1 are 22, 24, 26, 28, 32 and 36. Each HEVC video stream is decoded and recoded with QP2, which different from QP1. For example, the QP2 are 24, 26, 28, 32 and 36 if the QP1 is 22. We choose 361 single compressed videos and 361 double compressed videos as one group, finally, 30 groups of video samples are generated in total. The process of video compression coding is conducted in the HEVC video coding reference software HM12.0.

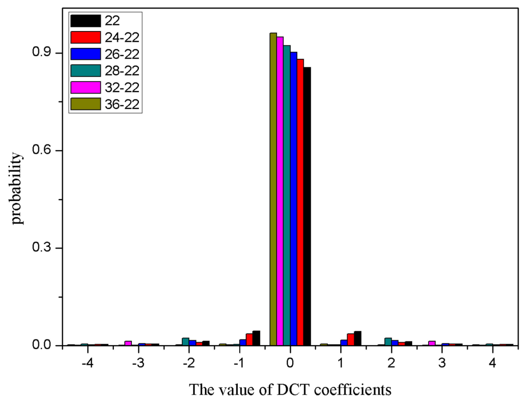


Fig. 8 *Histogram of the probability distribution of DCT* $\in [-4, 4]$

4.2 Experimental results

For each group, single compressed videos are treated as positive-class samples, and double compressed videos are treated as negative-class samples. They are labelled as $+1$, -1, respectively. A feature set is extracted as described in Section 3.3. Then we normalise the value of feature set such that it ranges from -1 to 1. Finally, they are sent to LIBSVM classier [18]. The 50% samples are randomly selected for training a classification model, and the remaining samples are used to test. Detection accuracy is calculated according to (14). For the LIBSVM classifier, radial basis function (RBF) is chosen as the kernel function of svmtrain $(\cdot)$ and 5-fold cross-validation is performed to select the optimal parameters (Gamma and Cost) for the kernel. Each training-testing procedure is repeated over ten times and the average detection accuracy is reported. Table 3 gives the detection accuracies of our method

$$\mathrm{Accuracy} = \frac{\mathrm{TPR} + \mathrm{TNR}}{2} \tag{14}$$

where $\mathrm{TPR} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$, $\mathrm{TNR} = \mathrm{TN}/(\mathrm{TN} + \mathrm{FP})$, TP represents the occurrence that a single compressed video is predicted as single compression, FN denotes the incident that a single compressed video is classified as double compression, TN shows the frequency that a double compressed video is predicted as double compression, FP is the frequency that a double compressed video is classified as single compression.

Here QP1 indicates the first quantisation parameter, QP2 is the second quantisation parameter of double compressed videos. In the case of $\mathrm{QP1} = 22$ and $\mathrm{QP2} = 24$, the detection accuracy of our method is 83.08%. It denotes the detection accuracy of single-compressed videos with 24 and double-compressed videos with 22 followed by 24. It is worth noting that the case of $\mathrm{QP1} = \mathrm{QP2}$ is not considered.

The method can achieve good classification ability in the left bottom of Table 3, but the result of the upper triangle is not ideal. This is probably due to the fact that the first compression coding used higher QP1 will lose more unrecoverable detail information. So it leaves more clear evidence for the second compression

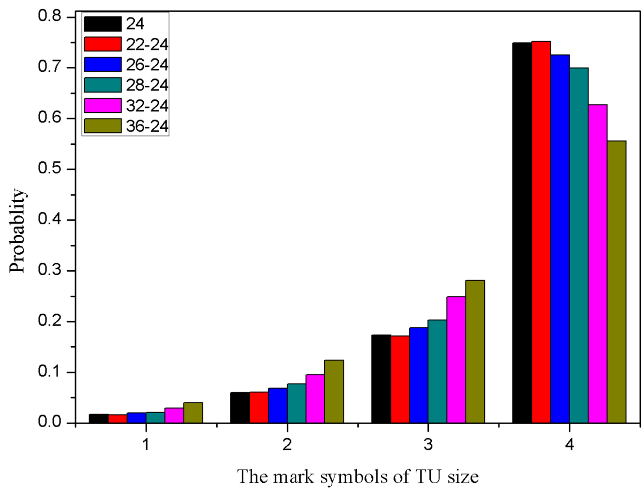


Fig. 9 *Histogram of the probability distribution of TU size*

Table 2 Selection of YUV sequences

Resolution	YUV sequences
416×240	Basketballpass、Blowingbubbles、*BQ*square、Flowervase_416×240、Mobisode2
832×480	Basketballdrill、Basketballdrilltext、*BQ*mall、Flowervase_832×480
1280×720	Johnny、vidyo1、vidyo2、vidyo3
1920×1080	BasketballDrive
2560×1600	Peopleonstreet、Traffic
176×144	akiyo、ridge_close、carphone、container、foreman_part、grandma、
—	hall、highway、miss_america、mother_daughter、news、salseman、silent

Table 3 Detection accuracies of our method (%)

QP1	QP2					
	22	24	26	28	32	36
22	—	83.08	69.72	70.00	51.39	49.44
24	92.22	—	83.33	65.28	58.61	51.39
26	99.44	98.33	—	80.00	72.50	49.17
28	99.72	99.44	95.28	—	67.22	58.89
32	100.00	100.00	100.00	99.44	—	71.39
36	99.72	100.00	99.72	99.72	98.89	—

Table 4 Comparison of three methods comprehensive performance

Algorithm	Our method	[11]	[12]	[15]
the mean of accuracies, %	82.11	80.25	78.96	82.26
feature dimension	17D	162D	20D	136D
computational complexity	$(64 \times 64 + 16 \times 16) \times q$	$(4 \times 2 \times 64 \times 64) \times q$	$(64 \times 64) \times q$	$(4 \times 64 \times 64) \times q$

coding to distinguish double compressed videos from single compressed videos.

4.3 Comparison with other features

We employed the algorithms of double compression detection based on the Markov feature for MPEG-4 videos in [11], DCT histogram feature for H.264/AVC in [12] and co-occurrence matrix feature for HEVC in [15] for comparison. In order to test the effectiveness of the algorithm, we applied the method in [11, 12] to HEVC.

We present a comprehensive performance evaluation for our method compared with the other three methods in Table 4. The average accuracies of our method and in [11, 12, 15] are 82.11%, 80.25%, 78.96% and 82.26%, respectively. In other aspects of feature dimensions and computational complexity, our method calculates once for each element of 64×64 DCT coefficients matrix $\boldsymbol{D}$ and 16×16 TU size identification matrix $\boldsymbol{T}$ and gets 17D features constructed by the distribution of DCT coefficients and TU size. 162D features by calculating Markov transition probabilities of the four directions are obtained in [11]. They repeat the calculation of each value of 64×64 DCT coefficients matrix for eight times. 20D DCT histogram features are extracted in [12], which contains the probabilities of all quantised non-zero AC coefficients ranged from -10 to 10. It only calculates once for each value of 64×64 DCT coefficients matrix $\boldsymbol{D}$. 136D features by calculating the co-occurrence matrix of DCT coefficients of the four directions are obtained in [15]. From Table 4, the algorithm has the higher average detection rate, the lowest feature dimension. The amount of data calculation is greater than [12], but less than [11, 15]. Our method outperforms the other three methods as mentioned above. It can achieve a better balance between detection accuracies, feature dimensions and computational complexity. Compared to our previous work [15], feature dimension and computational complexity in this paper are reduced while maintaining the detection accuracy. In Table 4, the character D represents dimension and q is the number of CTUs be encoded.

5 Conclusion

In this paper, we proposed an effective method to detect double compression in HEVC videos. First, the effect of QP on quantised DCT coefficients and TU size during the process of HEVC coding were analysed. Then, the feature set was extracted by constructing the relative statistics of DCT coefficients and TU size. Finally, LIBSVM classifier was used to distinguish between the single and double compressed videos. It can achieve a better comprehensive performance in consideration of detection accuracies, feature dimensions and computational complexity. Experimental results show that our proposed algorithm can work well when the first quantisation parameter QP1 is higher than the second quantisation parameter QP2.

6 Acknowledgments

This work was supported by the National Natural Science Foundation of China (grant nos. U1736215, 61672302, 61771270), Zhejiang Natural Science Foundation (grant nos. LZ15F020010, Y17F020051, LY17F020013), Ningbo Natural Science Foundation (grant no. 2017A610123), Ningbo University Fund (grant nos. XKXL1509, XKXL1503) and K.C. Wong Magna Fund in Ningbo University. Mobile Network Application Technology Key Laboratory of Zhejiang Province (grant no. F2018001).

7 References

[1] Rui, Y., Weiqi, L., Jiwu, H.: 'Multimedia forensic', *Sci. Chin. Inf. Sci. (Chinese)*, 2013, **43**, pp. 1654–1672

[2] Chen, C., Shi, Y.Q., Su, W.: 'A machine learning based scheme for double JPEG compression detection'. 19th Int. Conf. on Pattern Recognition: Tampa, FL, United States, 2008, pp. 1–4

[3] Bianchi, T., Piva, A.: 'Detection of nonaligned double JPEG compression based on integer periodicity maps', *IEEE Trans. Inf. Forensics Sec.*, 2012, **7**, pp. 842–848

[4] Liu, Q., Cooper, P.A., Chen, L., *et al.*: 'Detection of JPEG double compression and identification of smartphone image source and post-capture manipulation', *Appl. Intell.*, 2013, **39**, pp. 705–726

[5] Yang, J.Q., Xie, J., Zhu, G.P., *et al.*: 'An effective method for detecting double JPEG compression with the same quantization matrix', *IEEE Trans. Inf. Forensics Sec.*, 2014, **9**, pp. 1933–1942

[6] Zhang, R., Wang, R.D.: 'In-camera JPEG compression detection for doubly compressed images', *Multimedia Tools Appl.*, 2015, **74**, pp. 5557–5575

[7] Wang, Q., Zhang, R.: 'Double JPEG compression forensics based on a convolutional neural network', *EURASIP J. Inf. Secur.*, 2016, **23**, pp. 1–12

[8] Wang, W., Farid, H.: 'Exposing digital forgeries in video by detecting double MPEG compression'. Proc. of the Multimedia and Security Workshop, Geneva, Switzerland, 2006, pp. 37–47

[9] Chen, W., Shi, Y.Q.: 'Detection of double MPEG compression based on first digit statistics'. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Busan, Korea, 2009, pp. 16–30

[10] Wan, W., Xinhao, J., Tanfeng, S.: 'Exposing double MPEG compression based on first digit features', *J. Electron. Inf. Technol. (Chinese)*, 2012, **34**, pp. 3046–3050

[11] Jiang, X.H., Wang, W., Sun, T.F., *et al.*: 'Detection of double compression in MPEG-4 videos based on Markov statistics', *IEEE Signal Process. Lett.*, 2013, **20**, pp. 447–450

[12] Liao, D., Yang, R., Liu, H., *et al.*: 'Double H.264/AVC compression detection using quantized nonzero AC coefficients', *Process. SPIE-IS & T/Electron. Imaging*, 2011, **7880**, pp. 78800–78810

[13] Jia, R.S., Li, Z.H., Zhang, Z.Z., *et al.*: 'Double HEVC compression detection with the same QPs based on the PU numbers'. ITM Web of Conf., ITA, Hangzhou, China, 2016, p. 02010

[14] High efficiency video coding, ITU-T Recommendation H.265, 2013

[15] Huang, M., Wang, R., Xu, J., *et al.*: 'Detection of double compression for HEVC videos based on the co-occurrence matrix of DCT coefficients'. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Tokyo, Japan, 2016, pp. 61–71

[16] Sullivan, G.J., Ohm, J., Han, W.J., *et al.*: 'Overview of the high efficiency video coding (HEVC) standard', *IEEE Trans. Circuits Syst. Video Technol.*, 2012, **22**, pp. 1649–1668

[17] Shen, L., Zhang, Z., An, P.: 'Fast CU size decision and mode decision algorithm for HEVC intra coding', *IEEE Trans. Consum. Electron.*, 2013, **59**, pp. 207–213

[18] Chang, C.C., Lin, C.J.: LIBSVM: A Library for Support Vector Machines, 2001