DETECTING DOUBLE MPEG COMPRESSION WITH THE SAME QUANTISER SCALE BASED ON MBM FEATURE

Jieyuan Chen[⋆†] *Xinghao Jiang*[⋆†] *Tanfeng Sun*[⋆†] *Peisong He*[⋆†] *Shilin Wang*[⋆†]

[⋆] School of Electronic Information and Electrical Engineering, Shanghai Jiao Tong University
[†] National Engineering Lab for Information Content Analysis Techniques, GT036001, Shanghai, China
Email: cjy_sjtu@outlook.com, {xhjiang, tfsun, gokeyhps, wsl}@sjtu.edu.cn

ABSTRACT

Detecting double MPEG compression is of prime significance in video forensics. However, existing methods are effective only when the primary compression and the secondary compression have different quantiser scales (QS). There is a lack of effective methods dealing with double MPEG compression with the same QS. In this paper, a novel method based on the statistical feature of macroblock mode (MBM) which consists of macroblock type and motion vector in P-frames is proposed to detect double MPEG compression with the same QS. The MBM statistical feature is extracted during multiple decoding procedures when the video is repeatedly compressed with the same QS for several times. Finally, the proposed feature is combined with the support vector machine (SVM) to classify the single MPEG compression and double MPEG compression. Experiments have demonstrated the effectiveness of the proposed method and the robustness to a wide range of QSs and different encoders.

Index Terms— Digital forensics, double MPEG compression, same quantiser scale, MBM statistical feature

1. INTRODUCTION

In recent years, MPEG based codecs have been widely utilized in various applications such as surveillance systems and video cameras. However, with sophisticated video editing software, attackers can maliciously delete or insert video shots to change semantics of original video content. Since most of the tampered videos must undergo the recompression process, double compression detection can be regarded as one of the most important methods in video forensics[1].

Double compression can be categorized into two classes for variable bit rate (VBR) encoding depending on whether the primary QS and second QS are same or not. For detecting double MPEG compression with different QSs, there are several successful methods. By modeling the distribution of DCT coefficients, Wang and Farid provide some statistical features of DCT coefficients to detect double compression in [2, 3]. Benford's law of DCT coefficients is another useful detection tool[4, 5, 6]. Markov statistics is one of state-of-the-art methods to detect double MPEG compression and achieves satisfactory performance in [7, 8].

However, all methods mentioned above fail to detect double MPEG compression with the same QS due to the slight difference between singly compressed and doubly compressed videos. Please note these methods focus on double compression with the matched Group of Pictures (GOP) which does not leave temporal fingerprints, as well as our method. To detect double compression with different GOP structures, methods based on the temporal features are proposed[9, 10]. Similar to double MPEG compression detection with the same QS, in image forensics, detecting double JPEG compression with the same quality factor(QF) is still a challenging problem. Huang *et al.* first addressed this issue based on the JPEG coefficients convergence[11]. Yang *et al.* used the statistical difference of the error blocks, improving the detection accuracy and the detection range of QF[12]. Different from JPEG compression, MPEG compression considers the temporal correlation between frames which is more complicated. Thus detecting double MPEG compression with the same QS is more difficult.

In this paper, we propose a novel method to detect double MPEG compression with the same QS based on macroblock mode (MBM). Inspired by the JPEG coefficients convergence to detect double JPEG compression with the same QF in [11], we define the concept of MBM and analyze the statistical characteristics of MBM. To extract MBM statistical feature, a video is repeatedly re-compressed with the same QS and then the number of marcoblocks with different MBMs is calculated between two sequential compressions. Finally, the MBM statistical feature is fed to the support vector machine to detect double MPEG compression with the same QS. According to the experimental results, MBM statistical feature have been demonstrated to be effective to detect MPEG double compression with the same QS. Besides, the proposed method is robust to a wide range of QSs and different encoders conforming to both MPEG-2 and MPEG-4 standards.

The rest of this paper is organized as follows. Section 2 first introduces the preliminary of MBM statistical model and then the MBM statistical model is illustrated in detail.

Section 3 presents the proposed double MPEG compression detection scheme. Section 4 reports the experimental results and conclusions are drawn in section 5.

2. MBM STATISTICAL FEATURE

2.1. Analysis of Quantised DCT Coefficients of I-frames in Multiple Compression With the Same QS

In MPEG compression, there are three basic types of frames: intra-coded frame (I-frame), predictive-coded frame (P-frame), bi-directionally predictive-coded frame (B-frame). There are also three basic types of macroblocks: intra-coded macroblock(I-MB), inter-coded marcoblock(P-MB) and skipped macroblock(S-MB). All macroblocks in I-frames must be intra coded like JPEG compression. Therefore, when a video is MPEG compressed with the same QS repeatedly, the quantized DCT coefficients of I-frames will have the similar characteristic to JPEG coeffcients shown in [11]. We extend the repeated compression process of JPEG in [12] to MPEG compressed with the same QS. The n-times compressed 8×8 quantised DCT coefficients of an I-MB is denoted as $\boldsymbol{D_n}$, the generation of $\boldsymbol{D_{n+1}}$ can be expressed as:

$$\boldsymbol{D_{n+1}} = [DCT(RT(IDCT(\boldsymbol{D_n} \times \boldsymbol{Q} \times QS)))/(\boldsymbol{Q} \times QS)], \quad (1)$$

where the subscript *n* denotes the times of MPEG compression a video has undergone, $DCT(\cdot)$ and $IDCT(\cdot)$ denote 8×8 discrete cosine transform and inverse discrete cosine transform separately, $\boldsymbol{Q}$ denotes the fixed quantisation matrix, $RT(\cdot)$ denotes rounding and truncating a real number to integer in the range of [0, 255], $[\cdot]$ denotes rounding operation and all the basic arithmetic operation is component-wise. And the error of rounding and truncating the IDCT coefficients ($\boldsymbol{E_n}$) can be calculated as:

$$\boldsymbol{E_n} = RT(IDCT(\boldsymbol{D_n} \times \boldsymbol{Q} \times QS)) - IDCT(\boldsymbol{D_n} \times \boldsymbol{Q} \times QS). \quad (2)$$

According to (1) and (2), we can obtain the recursion formula:

$$\begin{aligned} \boldsymbol{D_{n+1}} &= [DCT(IDCT(\boldsymbol{D_n} \times \boldsymbol{Q} \times QS) + \boldsymbol{E_n})/(\boldsymbol{Q} \times QS)] \\ &= \boldsymbol{D_n} + [DCT(\boldsymbol{E_n})/(\boldsymbol{Q} \times QS)] \\ &= \boldsymbol{D_n} + \boldsymbol{R_n}, \end{aligned} \quad (3)$$

where $\boldsymbol{R_n}$ denotes the quantised DCT coefficients of $\boldsymbol{E_n}$.

Fig. 1. shows the average number of different quantized DCT coefficients per I-frame by multiple compression with the same QS for YUV sequences in CIF format downloaded from [13]. Each curve in Fig. 1. corresponds to a specific QS. The number of different coefficients is presented as F_n, where n represents the times of compression. It is observed that the value of F_n decreases with the compression times increasing. It can be found that the value of F_n is smaller corresponding to a larger QS. According to (3), the larger QS is, the elements of $\boldsymbol{R_n}$ are more likely to be equal to 0, which results in the faster convergence.

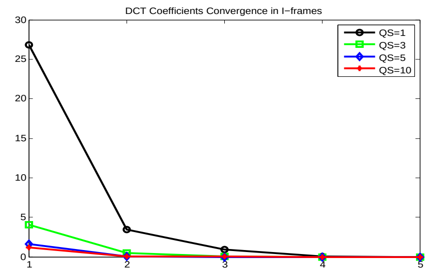


Fig. 1. Variation of the average number of different quantised DCT coefficients per I-frame corresponding to different QS, where the horizontal axis represents the times of compression and the vertical axis represents the average number of different coefficients.

2.2. MBM Statistical Feature

When the quantisation is strong, the difficulty of detecting double JPEG compression with the same QF has been shown in [11, 12] because of the slight difference of JPEG coefficients. However, the elements of quantisation matrix in MPEG compression are larger than these in JPEG compression which means the quantisation in MPEG is much stronger. It infers that it is difficult to discriminate between singly MPEG compression and doubly MPEG compression with the number of different quantised DCT coefficients of I-frames directly. In the inter-coding process, both macroblock type decision and motion estimation are based on the quantised DCT coefficients of I-frames and are more insensitive to the strong quantisation. In this section, we propose the definition of macroblock mode (MBM) and analyze the MBM statistical feature. MBM of each macroblock is defined as a set composed of 2 properties, i.e. macroblock type and motion vector, which can be expressed as:

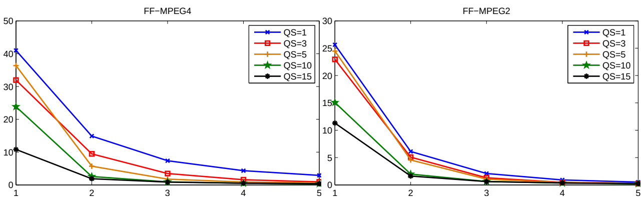
$$MBM(M) = \{M_{type}, M_{mv}\}, \quad (4)$$
$$M_{type} \in \{\text{I-MB,P-MB,S-MB}\}, M_{mv} = \{(u,v)|u,v \in \mathbb{Z}\},$$

where M denotes a macroblock, M_{type} denotes the macroblock type, M_{mv} denotes the motion vectors of M. If M_{type} is I-MB, its M_{mv} is set to $\{(0,0)\}$. Two macroblocks belong to the same MBM if and only if they have the same macroblock type and motion vector at the same time. When the video is compressed for several times, if a macroblock at the same position in n-times compressed video and $(n+1)$-times compressed video belongs to the same MBM, this macroblock is considered as temporarily stable in $(n+1)$-times compression, otherwise, the macroblock is unstable.

The total number of unstable macroblocks of a video between (n)-times compression and $(n+1)$-times compression is denoted as T_n, and the average number of unstable macroblocks per P-frame is formulated as:

$$C_n = \frac{1}{N} \sum_{i,x,y} I(M_n(i,x,y), M_{n+1}(i,x,y)), \quad (5)$$

where N denotes the total number of P-frames, $M_n(i,x,y)$



(a) (b)

Fig. 2. Variation of the average value of C_n for 116 video sequence corresponding to different QS, where the horizontal axis represents the index n of C_n and the vertical axis represents the value of C_n. (a) Videos encoded by FF-MPEG4; (b) Videos encoded by FF-MPEG2.

denotes the macroblock of the n-times compressed video, located at (x, y) of the i-th P-frame. And the indicator function $I(M_1, M_2)$ is defined as:

$$I(M_1, M_2) = \begin{cases} 1 & MBM(M_1) \neq MBM(M_2) \\ 0 & MBM(M_1) = MBM(M_2), \end{cases} \tag{6}$$

where M_1,M_2 are two macroblocks.

Fig. 2. shows the average value of C_n for 116 YUV sequences[1] encoded by FF-MPEG4 and FF-MPEG2 with different QSs. FF-MPEG4 and FF-MPEG2 denote the *mpeg4* codec and the *mpeg2video* codec in FFmpeg separately. It is observed that no matter which QS or encoder are used, the value of C_n is monotonically decreasing and the speed of decrease becomes slower with the increment of n. Furthermore, it is observed that the values of C_n are bigger corresponding to smaller QSs.

Comparing Fig. 2 with Fig. 1., it is found that the variation tendency of MBM (C_n) is similar to the variation tendency of quantised DCT coefficients in I-frames (F_n). It can be calculated that the minimal proportion of unstable macroblocks per P-frame is 2.7% while the maximum proportion of different quantised DCT coefficients per I-frame is only 0.26‰. Thus the convergence speed of C_n is slower than F_n. The reason is that the stability of macroblocks in P-frames is frame-delayed. P-frames locate in the later part of GOP can only refer to its previous P-frame, not I-frame directly. The $(K+1)$-th P-frame cannot be completely stable until the K-th P-frame is stable where $K = 1, 2, \cdots$, presenting the order of P-frames in a GOP. Therefore, the macroblocks in P-frames located in the later part of GOP can still be unstable after several times compression. The slower convergence speed of C_n infers that C_n is more discriminative to detect double MPEG compression with the same QS, especially with large QS.

3. PROPOSED DETECTION SCHEME

As illustrated in Fig. 2., the curves of MBMs have discriminatively different convergence between singly MPEG compressed videos and doubly MPEG compressed videos. To further enhance the robust of MBM feature, we propose a detection method based on machine learning framework. The proposed method is consisted of four steps which are described as follows:

1. Decode the input video V to YUV file Y_1 in order not to introduce the error when changing YUV to RGB and extract its macroblock mode information M_1 of all macroblocks in P-frames synchronously.

2. Compress the YUV file Y_n to MPEG video V_n with the same QS, $n = 1, 2, \cdots$. Then decode V_n to YUV file Y_{n+1} and extract its macroblock mode information M_{n+1} at the same time.

3. Repeat Step 2) for K times. Then calculate C_n using M_n and M_{n+1} according to (5).

4. Construct the vector $\boldsymbol{C} = (C_1, C_2, \cdots, C_K)$ as the feature of this video, then the support vector machine(SVM) is employed to detect doubly MPEG compressed with the same QS.

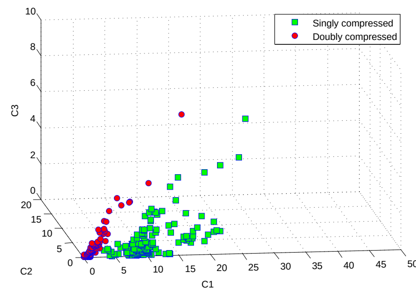


Fig. 3. Feature space of C when $K = 3$ and the samples contain 116 singly compressed videos and 116 doubly compressed videos with the QS equals 15.

Fig. 3. shows the feature space of $\boldsymbol{C}$ when $K = 3$ and intuitively demonstrates the capability of $\boldsymbol{C}$ to differentiate the singly compressed and doubly compressed videos. According to our experimental results, $K = 5$ should be a proper value to construct the feature vector because a larger K only increases a little precision in detection but needs to repeat step 2) for more times which is computationally expensive.

4. EXPERIMENTS

To generate our dataset, a group of 32 well-known YUV sequences has been downloaded with CIF resolution (352×288 pixels)[13]. Each sequence has been split into several non-overlapping subsequences with 100 frames long. Finally, 116 new YUV sequences are obtained. Similar to the samples generated in [11], each YUV sequence is first MPEG compressed with a specific QS as an element of negative class,

[1] The details about 116 YUV sequences are illustated in Section 4.

denoted by $V_{1,QS}$, and then $V_{1,QS}$ is re-compressed with the same QS to construct positive class, denoted by $V_{2,QS}$. The accuracy rate(AR) is calculated as AR=(TPR+TNR)/2, where TPR and TNR means true positive rate and true negative rate separately. For the SVM classifier, rbf kernel is selected and the parameters is determined by a grid-search. All the results by the proposed method are computed by 5-fold cross-validation. In order to demonstrate the robustness of our methods to different encoders, three encoders based on different MPEG standards are selected to test separately. For simplicity, the size of GOP is set as 10 and no B-frame is encoded, other parameters are set as default value. To our best knowledge, there is lack of methods for dealing with such situation, so we choose one of the state-of-the-art method in [7] using Markov statistics to compare with our method.

We first evaluate the performance on the subset S_{qs} which only contains the videos compressed with a specific QS=qs ranging from 1 to 15 and then evaluate the performance on the whole set S_h containing hybrid QSs, defined in (7).

$$S_{qs} = \{V_{i,qs}^k | i = 1, 2, k = 1, 2, ..., 116\}, S_h = \bigcup_{qs=1}^{15} S_{qs}, \quad (7)$$

where $V_{1,qs}^k$ denotes the k-th singly compressed video with QS=qs, $V_{2,qs}^k$ denotes the k-th doubly compressed video with QS=qs. When the QS is equal to 15, the quality of the video is extremely poor, so the largest QS is selected up to 15.

Results are given in Table 1 for different encoders and QSs, where the average ARs of all the QSs ranging from 1 to 15 are given in the "Average" line. FF-MPEG2 and FF-MPEG4 denote the default MPEG2 and MPEG4 encoders in FFmpeg separately, and XVID denotes the common used Xvid encoder. "MBM" denotes the results of the proposed method and "Markov" denotes the results of the state-of-the-art method in [7] using Markov statistics. It is observed that the average ARs of MBM are 96.71%, 96.27%, 92.67% for FF-MPEG2, FF-MPEG4 and XVID separately. The accuracy rates are lower corresponding to XVID due to the more sophisticated macroblock type decision and motion estimation algorithms. The results of "Markov" show that it fails to detect double MPEG compression with the same QS and degrades to random guess. On the other hand, when QS becomes larger, the accuracy rates do not tend to decrease. The results can even be 100% with a large QS compressed by FF-MPEG2 and the AR is up to 99.70% and 95.78% for FF-MPEG4 and XVID separately. There is a minimal value of the accuracy rates when QS is small. The reason is that when QS is small, the value of C_n defined in (5) will be still large when $n > 1$. So there are some videos' value of C_{n+1} is larger than other videos' value of C_n, which makes it difficult to differentiate. When the QS is large, the variance is much smaller resulting in the better performance. Therefore, the proposed method based on MBM statistical feature can be applicable to a wide range of QS.

Table 1. Detection accuracy rates (%) for different encoders and QSs. (Compared with Markov statistics[7])

QS	FF-MPEG2		FF-MPEG4		XVID	
	Markov	MBM	Markov	MBM	Markov	MBM
1	49.09	**93.98**	49.69	**93.67**	57.79	**91.87**
2	49.09	**92.17**	49.91	**93.98**	49.26	**94.58**
3	49.63	**87.94**	49.86	**93.37**	49.74	**91.87**
4	49.91	**90.34**	49.89	**94.58**	49.63	**91.27**
5	49.83	**92.77**	49.69	**92.47**	49.83	**89.16**
6	50.00	**98.80**	49.91	**93.98**	49.69	**88.86**
7	50.20	**98.19**	49.80	**97.89**	49.80	**90.96**
8	50.06	**99.10**	49.60	**95.78**	49.86	**92.17**
9	49.78	**99.10**	50.06	**96.39**	49.66	**95.48**
10	49.80	**99.40**	49.86	**98.49**	49.97	**93.37**
11	49.69	**99.70**	49.86	**96.99**	49.86	**95.18**
12	49.80	**99.40**	49.86	**98.19**	49.71	**89.46**
13	49.51	**100**	49.80	**99.10**	49.88	**95.78**
14	49.34	**100**	49.74	**99.70**	49.83	**95.78**
15	49.26	**99.70**	49.74	**99.40**	49.71	**94.28**
Average	49.67	**96.71**	49.82	**96.27**	50.28	**92.67**
Hybrid	–	**94.28**	–	**94.12**	–	**90.90**

The "Hybrid" line of Table 1 shows results on the whole videos compressed with different QSs in set S_h. The results on XVID is 90.90% and the results on FF-MPEG2 and FF-MPEG4 are all higher than 94%. Markov statistics need to train a unique model for each QS, so we do not evaluate the performance for it on hybrid set. The results illustrate that in fact one hyperplane in SVM is enough to differentiate the doubly and singly compressed videos corresponding to different QSs and shows that MBM statistical feature is universal to different QSs. The results in Table 1 demonstrate the good discriminative capability of detecting double MPEG compression with the same QS and the robustness to different QSs and encoders of proposed method.

5. CONCLUSION

Detecting double MPEG compression with the same QS is a challenging work due to the unobvious difference between singly compressed and doubly compressed MPEG videos. To the best of our knowledge, there are few effective methods to solve such problem in the literature. In this paper, the statistical characteristic of MBMs in P-frames is analyzed and the MBM statistical feature is extracted to detect the double MPEG compression with the same QS. Experimental results demonstrate that our proposed method is capable of detecting double MPEG compression when the primary compression and the secondary compression have the same QS.

Acknowledgment: This work was supported by the National Natural Science Foundation of China (No. 61272249, 61272439, 61572320, 61572321), the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20120073110053). Corresponding author is Xinghao Jiang, any comments should be addressed to xhjiang@sjtu.edu.cn.

6. REFERENCES

[1] Simone Milani, Marco Fontani, Paolo Bestagini, Mauro Barni, Alessandro Piva, Marco Ta-gliasacchi, and Stefano Tubaro, "An overview on video forensics," *APSIPA Transactions on Signal and Information Processing*, vol. 1, pp. e2, 2012.

[2] Weihong Wang and Hany Farid, "Exposing digital forgeries in video by detecting double MPEG compression," in *Proceedings of the 8th workshop on Multimedia and security*. ACM, 2006, pp. 37–47.

[3] Wang Weihong and Farid Hany, "Exposing digital forgeries in video by detecting double quantization," in *Proceedings of the 11th ACM workshop on Multimedia and security*. ACM, 2009, pp. 39–48.

[4] Wen Chen and Yun Q Shi, "Detection of double MPEG compression based on first digit statistics," in *Digital Watermarking*, pp. 16–30. Springer, 2009.

[5] Tanfeng Sun, Wan Wang, and Xinghao Jiang, "Exposing video forgeries by detecting MPEG double compression," in *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*. IEEE, 2012, pp. 1389–1392.

[6] Simone Milani, Paolo Bestagini, Marco Tagliasacchi, and Stefano Tubaro, "Multiple compression detection for video sequences," in *Multimedia Signal Processing (MMSP), 2012 IEEE 14th International Workshop on*. IEEE, 2012, pp. 112–117.

[7] Xinghao Jiang, Wan Wang, Tanfeng Sun, Yun Q Shi, and Shilin Wang, "Detection of double compression in MPEG-4 videos based on Markov statistics," *Signal Processing Letters, IEEE*, vol. 20, no. 5, pp. 447–450, 2013.

[8] Hareesh Ravi, AV Subramanyam, Gaurav Gupta, and B Avinash Kumar, "Compression noise based video forgery detection," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5352–5356.

[9] D Vázquez-Padın, M Fontani, T Bianchi, P Comesana, A Piva, and M Barni, "Detection of video double encoding with GOP size estimation," in *IEEE International Workshop on Information Forensics and Security (WIFS)*, 2012.

[10] Matthew C Stamm, W Sabrina Lin, and KJ Liu, "Temporal forensics and anti-forensics for motion compensated video," *Information Forensics and Security, IEEE Transactions on*, vol. 7, no. 4, pp. 1315–1329, 2012.

[11] Fangjun Huang, Jiwu Huang, and Yun Qing Shi, "Detecting double JPEG compression with the same quantization matrix," *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 4, pp. 848–856, 2010.

[12] Jianquan Yang, Jin Xie, Guopu Zhu, Sam Kwong, and Yun-Qing Shi, "An Effective Method for Detecting Double JPEG Compression With the Same Quantization Matrix," *Information Forensics and Security, IEEE Transactions on*, vol. 9, no. 11, pp. 1933–1942, Nov 2014.

[13] C Montgomery et al., "Xiph. org video test media (derf's collection), the xiph open source community, 1994," Online, https://media.xiph.org/video/derf/.