EXPOSING VIDEO FORGERIES BY DETECTING MPEG DOUBLE COMPRESSION

Tanfeng Sun[1, 2], *Wan Wang*[1] *and Xinghao Jiang*[1, 2*]

[1]School of Electronic Information and Electrical Engineering,
Shanghai Jiao Tong University, Shanghai, China
{tfsun, wangwan, xhjiang}@sjtu.edu.cn
[2]Shanghai Information Security Management and Technology Research Key Lab,
Shanghai, 200240, China

ABSTRACT

In this paper, an improved video tampering detection model based on MPEG double compression is proposed. Double compression will import disturbance into Discrete Cosine Transform (DCT) coefficients, reflecting in the violation of the parametric logarithmic law for first digit distribution of quantized Alternating Current (AC) coefficients. A 12-D feature can be extracted from each group of pictures (GOP) and machine learning framework is adopted to enhance the detection accuracy. Furthermore, a novel approach with a serial Support Vector Machine (SVM) architecture to estimate original bit rate scale in doubly compressed video is proposed. Experiments demonstrate higher accuracy and effectiveness.

Index Terms— Video forgery, double compression, first digit distribution, machine learning

1. INTRODUCTION

Nowadays, digital video recording systems are widely used. It is important to reliably authenticate the truism and validity of a given video. Some positive results have been gained in the field of digital forgery detection, as in [1][2]. But a lot of them focus on the detection of image tampering, as in [3].

Digital videos are usually compressed with MPEG-x or H.26x coding standard. The tampering must be operated in uncompressed domain to accomplish frame deletion, frame inserting, etc. Considering size and format, the tampered video must be re-encoded. Thus, the occurrence of double compression may expose digital forgery.

Ref. [4] proposed to use the periodicity of the artifacts introduced into the Discrete Cosine Transform (DCT) coefficient histograms as evidence of double compression. A parametric logarithmic law, i.e. the generalized Benford's law was first modeled to detect JPEG double compression in [5]. Ref. [6] validated the feasibility of this principle in MPEG coded video.

This paper proposed a more effective detection scheme with satisfying accuracy. First, extract the first digit distribution of Alternating Current (AC) coefficients from each intra-coded frame. Then, fit the distribution with parametric logarithmic law and generate a 12-D feature. Finally, use machine learning framework to ensure the reliability and efficiency. In addition, the original bit-rate scale of doubly compressed video can be estimated by a serial architecture of Support Vector Machine (SVM).

The rest of this paper is organized as follows: In Section 2, the MPEG coding principle is described. In Section 3, the application of Benford's law in MPEG video and the algorithm of double compression detection are presented. Section 4 describes the experiments and results analysis for videos in 12 categories. Section 5 draws the conclusion.

2. MPEG CODING PRINCIPLE

Each MPEG video is organized as a hierarchy of layers, including video sequence, group of pictures (GOP), picture, slice, macroblock and block. The algorithm and corresponding experiments in this paper are focus on MPEG-2 Constant Bit Rate (CBR) video, but they are backward compatible to MPEG-1.

2.1. Intra coding principle

The intra coding processes in MPEG and JPEG are very similar, including DCT, quantization and entropy coding.

First, each 8x8 block will perform DCT to achieve energy concentration and signal data de-correlation. The low frequency DCT coefficients, especially Direct Current (DC) coefficient, are of most significance. While high frequency DCT coefficients are almost zero. On account of the energy concentration in DCT, higher frequency DCT coefficients will be more coarsely quantized. Specifically, each coefficient is divided by a corresponding quantization matrix

* Corresponding Author is Dr. Xinghao Jiang.
This paper is supported by the National Natural Science Foundation of China (No.60802057, 61071153) and by the New Century Outstanding Person Plan of China National Education Department (No.NCET-10-0569).

value. Note that each value in the matrix is pre-scaled by multiplying quantizer scale (Q_Scale), which takes effect on a macroblock basis. The quantizer scale can be assigned as an encoder parameter in Variant Bit Rate (VBR) video, while it will change adaptively in CBR video to maintain the output bit rate. Finally, entropy coding, i.e. Variable Length Coding (VLC) will be applied to quantized DCT coefficients.

2.2. Inter coding principle

For macroblocks in P and B frames, motion compensation technique is applied to exploit temporal redundancy. A spatial search is performed in a search range. If a relatively good match can be found, this macroblock will go through inter coding, otherwise it will perform intra coding.

In inter coded macroblock, motion vector is calculated and transmitted to indicate how the reference macroblock should be moved to generate predicted macroblock. Then, the predicted macroblock is subtracted from real macroblock, leaving a less complicated residual error. Note that the residual error is spatially coded as intra coding, except that the quantization matrix is a flat matrix with a constant value. Most of the quantized DCT coefficients are hopefully zero because residual error has less information.

3. DOUBLE COMPRESSION DETECTION ALGORITHM

3.1. First digit distribution in MPEG video

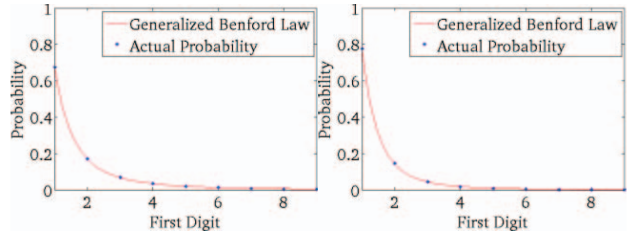
It has been proved that the first digit distribution can be utilized well in JPEG double compression detection. It is reasonable to deduce that the first digit distribution in MPEG has the same characteristic. To verify the application of parametric logarithmic law, the first digit distribution of quantized AC coefficients was extracted from both original video and doubly compressed video. Then, the probabilities were fitted with the parametric logarithmic law, as in

$$y = N \log_{10}(1 + \frac{1}{s + x^{q}}) \,. \tag{1}$$

Where x stands for the first digits, y is the corresponding value in the logarithmic law, and N, q, s are three parameters which are limited to [0.1, 3], [0.1, 3] and [-1, 1].

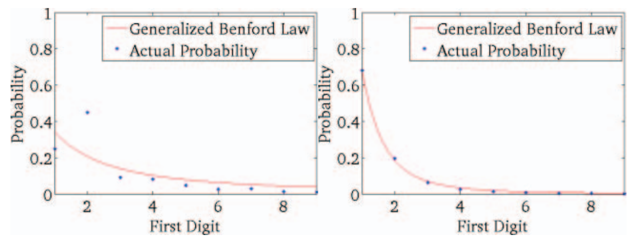
Fig. 1 shows that the first digit distribution in original MPEG video follows the parametric logarithmic law well. But the first digit distribution in non-intra frame differs from that in intra frame. AC coefficients with first digit 1 account for a bigger proportion, while the probabilities of AC coefficients with first digit 7, 8 and 9 are smaller. Thus, the first digit distribution in non-intra frames may not follow the logarithmic law as well as intra frames. This phenomena is caused by inter coded macroblocks in P and B frames.

Fig. 2 shows the fitting results of doubly compressed MPEG video. For the case where target bit rate to doubly compress is larger than original bit rate, the violation of the



(a) (b)

Fig. 1 Fitting results for original MPEG video
(a) Intra (I) frame. (b) Non-intra (P) frame.



(a) (b)

Fig. 2 Fitting results for doubly compressed MPEG video
(a) Target bit rate is larger than original bit rate. (b) Target bit rate is smaller than original bit rate.

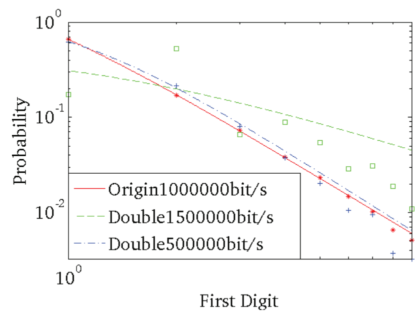


Fig. 3 Fitting results in log-log scale
Solid, dashed and dash dotted line stand for original video, doubly compressed video with larger bit rate and smaller bit rate.

parametric logarithmic law is visually obvious. But if target bit rate is smaller than original bit rate, the first digit distribution has the tendency to follow the logarithmic law. To seek for the abnormality in the latter situation, the fitting results were re-plotted in log-log scale, see Fig. 3. It's clear that the first digit distribution drifts further in doubly compressed MPEG video than original MPEG video.

3.2. MPEG double compression detection algorithm

Due to the sensitiveness of first digit distribution to video content and target bit rate, machine learning framework is adopted to enhance accuracy. Fig. 4 demonstrates the algorithm architecture. The detailed process is as follows:

1) For both query and training video, the first digit distribution of quantized AC coefficients is extracted.

2) Test the first digit distribution with parametric logarithmic law. Three goodness-to-fit statistics are calculated, including squares due to error (SSE), root mean

squared error (RMSE) and R-square. SSE and RMSE closer to zero, R-square closer to one means a good fit.

3) Combine the first digit probabilities and goodness-to-fit statistics to compose a 12-D feature. Only I frames are taken into consideration because the fitting results for intra frames are better than that for non-intra frames.

4) Each GOP with a 12-D feature is treated as a detection unit, so the SVM classifier will judge on a GOP basis. The GOP proportion D is defined as

$$D = M / N. \tag{2}$$

Where M stands for the number of GOPs which are labeled as double compression, and N means the total number of GOPs. If D passes the threshold T, it is extremely possible that the video has gone through double compression. Note that T is adaptive according to the demand of TNR and TPR. Generally T might be set as 0.50.

3.3. Original bit rate estimation algorithm

Taking a deep look into the fitting results of doubly compressed MPEG video, the difference between target bit rate decreasing situation and increasing situation is notable. This is the trigger for a more detailed classification. Fig. 5 shows the serial SVM architecture for this estimation.

It has been verified that the violation of the parametric logarithmic law will be much more obvious if target bit rate is larger than original bit rate. So the bit rate increasing situation can be classified by SVM1. SVM2 focuses on the judgment of bit rate decreasing situation and original video. As a result, the probability can be calculated with

$$p = C / N. \tag{3}$$

Where C stands for the number of GOPs which are labeled as a certain class and N is the total number of GOPs in a video. P means the probability, namely, confidence index.

4. EXPERIMENTS AND ANALYSIS

Source videos for experiments are 12 YUV sequences with resolution 352x288, see Table I. MPEG2Encode and MPEG2Decode are used for video pre-processing. YUV sequences are first encoded to MPEG-2 CBR video with bit rate 500000, 750000, 1000000, 1250000 and 1500000 bps. To simulate double compression, all 1000000 bps videos will be decoded and re-encoded with the other four bit rates.

4.1. MPEG double compression detection experiment

The C# library MPEG2Event and Matlab Curve Fitting Toolbox are utilized to extract 12-D features. LIBSVM with Radial basis function (RBF) kernel is chosen to accomplish machine learning. Video clips in each category are treated as test database in turn, and the remaining 11 categories are used for training. The GOP proportion defined in Section 3.2 is shown in Table II. Video clips with GOP proportion larger than 0.50 are determined as doubly compressed.

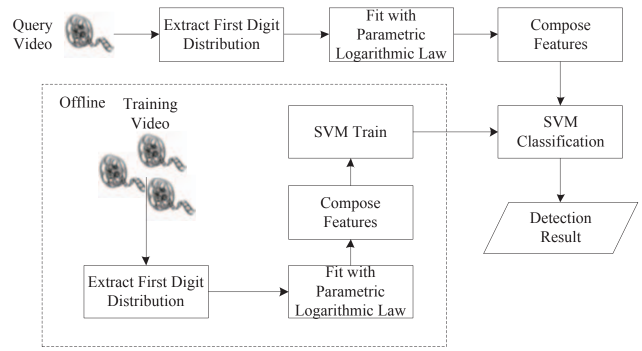


Fig. 4 Double compression detection algorithm architecture

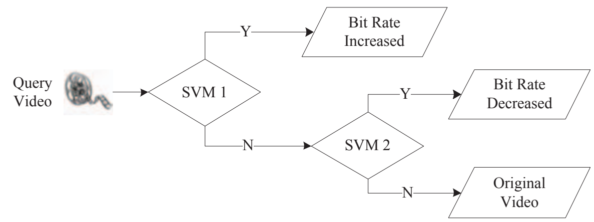


Fig. 5 Serial SVM architecture for original bit rate estimation

Table I Source YUV sequences

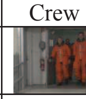
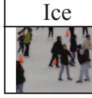
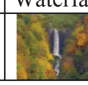
Bus	Container	Crew	Football	Foreman	Hall
					
Ice	Mobile	Mother	News	Tempete	Waterfall
					

Table II GOP proportion for double compression

Bit Rate	500000		750000		1250000		1500000	
(bps)	ori[a]	dbl[b]	ori	dbl	ori	dbl	ori	dbl
bus	0.42	0.58	0.00	1.00	0.00	1.00	0.00	1.00
container	0.00	0.96	0.00	1.00	0.00	1.00	0.00	1.00
crew	0.08	0.56	0.00	0.68	0.00	1.00	0.00	1.00
football	**0.57**	0.71	0.38	0.95	0.19	1.00	0.00	1.00
foreman	0.08	0.64	0.00	0.92	0.00	0.96	0.00	1.00
hall	0.00	0.92	0.00	1.00	0.00	1.00	0.00	1.00
ice	0.00	0.70	0.00	0.95	0.00	0.95	0.00	1.00
mobile	0.00	0.96	0.00	0.96	0.00	1.00	0.00	1.00
mother	0.12	0.92	0.00	1.00	0.00	0.92	0.00	1.00
news	0.12	1.00	0.40	0.68	0.16	1.00	0.12	1.00
tempete	0.05	1.00	0.00	0.95	0.00	1.00	0.00	1.00
waterfall	0.00	1.00	0.00	1.00	0.00	1.00	0.24	1.00

[a] Original MPEG video. [b] Doubly compressed MPEG video.

From Table II, conclusions can be drawn as follows:

1) For the case that target bit rate is larger than original bit rate, all GOPs in most videos can be classified correctly.

2) Only one misjudgment (bold in table) occurred, generating 97.92% TNR and 100.00% TPR. The reason for misjudgment is the sensitiveness of first digit distribution to video content and bit rate. Specifically, videos with lower bit rate are harder to detect because larger quantizer scale leads to a more coarsely quantization process.

Fig. 6 Comparison result of this algorithm and algorithm in Ref. [6] Left and middle column are original accuracy and reproduced accuracy of Ref. [6]. Right column is the accuracy of our algorithm.

Fig. 6 describes the comparison result of our algorithm with the algorithm in [6]. For stronger persuasion concern, the algorithm in [6] was reproduced with same video clips from our experiments. The decrease in accuracy resulted from the change of video content and bit rate. Ref. [6] chose a 36-D feature which contained the average first digit distribution and respecting goodness-to-fit statistics for P and B frames. The features from non-intra frames are eliminated in our algorithm due to the worse fitting results caused by inter coded macroblocks. Experimental results have proved the raise in accuracy with 12-D features.

4.2. Original bit rate estimation experiment

For each video clip, the GOP proportion of original video, target bit rate decreasing and increasing video is obtained. As defined in Section 3.3, the GOP proportion reflects the probability of what class this video clip belongs to. Table III demonstrates the normalized probabilities for all video clips. The straightforward approach is to select class with the highest probability as final determination.

It can be seen from Table III that only 4 (bold in table) in 96 video clips were misjudged, with 95.83% accuracy. It's notable that in majority videos, nearly all the GOPs are classified into the correct class. Conclusion can be drawn that this algorithm can accurately judge whether a video clip has been doubly compressed or not and estimate the original bit rate scale in the doubly compressed video.

5. CONCLUSION

This paper proposed an improved algorithm to detect MPEG double compression. The extracted 12-D feature is only from I frame. Meantime, a novel approach to estimate original bit rate scale using serial SVM architecture is developed. Experiments have demonstrated the accuracy and effectiveness for video tampering detection.

ACKNOWLEDGMENT

We would like to give our special gratitude to the fruitful technical discussions with and selfless help offered by Prof. Y. Q. Shi at NJIT in U.S.A..

REFERENCES

[1] W. H. Wang, H. Farid, "Exposing digital forgeries in interlaced and deinterlaced video" [J]. *IEEE Transactions on Information Forensics and Security*, 2(3), pp. 438-449, 2007.

[2] M. Kobayashi, T. Okabe, and S. Yoichi, "Detecting forgery from static-scene video based on inconsistency in noise level functions" [J]. *IEEE Transactions on Information Forensics and Security*, 5(4), pp. 883-892, 2010.

[3] B. L. Shivakumar, S. S. Baboo, "Detecting copy-move forgery in digital images: a survey and analysis of current methods" [J]. *Global Journal of Computer Science and Technology*, 10(7), pp. 61-65, 2010.

[4] W. H. Wang, H. Farid, "Exposing digital forgeries in video by detecting double MPEG compression" [C]//*Multimedia and Security Workshop 2006*, ACM SIGMM, Geneva, pp. 37-47, 2006.

[5] D. D. Fu, Y. Q. Shi, and W. Su, "A generalized Benford's law for JPEG coefficients and its applications in image forensics" [C]//*Security, Steganography and Watermarking of Multimedia Contents IX*, SPIE, San Jose, pp. 65051L, 2007.

[6] W. Chen, Y. Q. Shi, "Detection of double MPEG compression based on first digit statistics" [C]//*7th International Workshop on Digital Watermarking*, Springer Verlag, Busan, pp. 16-30, 2009.

Table III Classification Probability

Bit Rate	500000 ori			750000 ori			1250000 ori			150000 ori			500000 dbl			750000 dbl			1250000 dbl			150000 dbl		
(bps)	ori	sma[a]	lar[b]	ori	sma	lar	ori	sma	lar	ori	sma	lar	ori	sma	lar	ori	sma	lar	ori	sma	lar	ori	sma	lar
bus	0.83	0.17	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	**0.50**	**0.50**	**0.00**	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00
container	0.92	0.08	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00
crew	0.88	0.12	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.48	0.52	0.00	0.32	0.68	0.00	0.00	0.00	1.00	0.00	0.00	1.00
football	**0.38**	**0.57**	**0.05**	0.62	0.38	0.00	0.86	0.10	0.05	0.90	0.05	0.05	0.29	0.67	0.05	0.10	0.90	0.00	0.00	0.14	0.86	0.00	0.00	1.00
foreman	0.92	0.08	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.28	0.68	0.04	0.04	0.96	0.00	0.04	0.00	0.96	0.00	0.04	0.96
hall	0.96	0.04	0.00	0.96	0.00	0.04	1.00	0.00	0.00	0.96	0.04	0.00	0.04	0.96	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00
ice	0.90	0.10	0.00	0.85	0.15	0.00	1.00	0.00	0.00	0.95	0.05	0.00	0.35	0.65	0.00	**0.05**	**0.30**	**0.65**	0.05	0.00	0.95	0.00	0.00	1.00
mobile	0.96	0.04	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.88	0.12	0.00	0.04	0.92	0.04	0.04	0.96	0.00	0.00	0.00	1.00	0.00	0.00	1.00
mother	0.84	0.16	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.12	0.88	0.00	**0.00**	**0.20**	**0.80**	0.16	0.40	0.44	0.00	0.00	1.00
news	0.84	0.16	0.00	0.88	0.08	0.04	0.80	0.00	0.20	0.88	0.00	0.12	0.00	0.68	0.32	0.40	0.60	0.00	0.00	0.00	1.00	0.00	0.00	1.00
tempete	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.05	0.95	0.00	0.00	0.00	1.00	0.00	0.00	1.00
waterfall	1.00	0.00	0.00	1.00	0.00	0.00	0.81	0.19	0.00	0.52	0.48	0.00	0.00	1.00	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00	0.00	1.00

[a] Target bit rate is smaller than original bit rate. [b] Target bit rate is larger than original bit rate.