Detecting Fake-Quality MP3 based on Huffman Table Index

Xianmin Yu, Rangding Wang
College of Information Science and Engineering,Ningbo University,Ningbo 315211,China
Email: mlhappy1016@163.com, wangrangding@nbu.edu.cn

Diqun Yan, Pengfei Ma
College of Information Science and Engineering,Ningbo University,Ningbo 315211,China
Email: yandiqun@nbu.edu.cn, mpfabc@126.com

***Abstract*—MP3 is the most widely used audio format nowadays in our daily life and music on the internet are often of this format. In order to seek commercial benefit, low bit rate MP3 audios are usually transcoded to high bit rate, resulting in fake-quality MP3 audios. Through analyzing the frequency statistics of Huffman table indexes, a method for detecting such fake-quality MP3 audios is presented in this paper. Experimental results show the effectiveness of the proposed method.**

***Index Terms*—MP3; Fake-Quality; Huffman table index**

I. Introduction

With the decline of the traditional music market and the popularity of digital music market, more and more people choose to buy their favorite music in online music store. Often these music are in MP3 format due to the fact that MP3 is the most widely compatible music format and is supported by most media player applications and hand-held music devices. The price of MP3 files varies with different compression bit rates and the higher the MP3 compression bit rate, the more expensive the price of MP3 because high compression bit rate means more bits are used to record the music.

On the one hand, the online music store brings people the convenience of purchasing music. On the other hand, it also may bring about potential property losses to people. When people buy a MP3 audio in online music store, the only way to evaluate the MP3 audio is checking the provided compression bit rate before listening. Certainly, high compression bit rate version is more attractive than low compression bit rate one for people pursuing high-quality music. In order to obtain high profits, some unscrupulous traders make fake-quality MP3 audio. Fake-quality MP3 here means that a MP3 file with a low compression bit rate is decompressed at first and then recompressed at a high bit rate, as illustrated in Fig.1. It is easy to create fake-quality MP3 file by using pervasive audio editing software. When people buy such a fake-quality MP3 file, they will suffer a loss. At the same time, the fake-quality MP3 file will also have a negative impact on the genuine sellers and the prospect of the digital music market.

Figure 1. The creation of fake-quality MP3 audio

Therefore, it is necessary to design reliable method to detect fake-quality MP3 audios. It is drawn from Fig.1 that the creation of fake-quality MP3 audios must involve double MP3 compression. So it is reasonable and feasible to find out fake-quality MP3 files by detecting double MP3 compression. Although there are many literatures dealing with the detection of double compression of images and video [1-8], few works have been done on the MP3 audios [9-12]. Regarding the detection of double MP3 compression, Shi et.al [9] use the number of small-value quantized modified discrete cosine transform(QMDCT) coefficients as features to discriminate fake-quality MP3 from normal MP3. In Shi's another work [10], they discover that for single compressed MP3 audio but not for double compressed MP3 audio, the distribution of the first digits of the QMDCT coefficients can be well modeled as generalized Benford's law. Liu et al [11] utilized the ratio of the number of QMDCT coefficients whose absolute values higher than a certain threshold value to all the number of QMDCT coefficients as feature to identify the double compressed MP3 audios. In Liu's another literature [12], they discover that the small

Corresponding author：Rangding Wang.
Email：wangrangding@nbu.edu.cn

doi:10.4304/jsw.9.4.907-912

QMDCT coefficients are more sensitive to the double MP3 compression. Therefore, they choose the number of zero QMDCT coefficients as feature.

All the existing methods mentioned above of detecting double MP3 compression are extracting the statistical characteristics for analyzing the QMDCT coefficients. In this paper, we propose a method to detect double MP3 compression by extracting features base on the frequency statistics of Huffman table indexes. The double MP3 compression here is only considered from low compression bit rate to high compression bit because our main purpose is to detect fake-quality MP3 file.

The rest of this paper is organized as follows. The second section briefly reviews the MP3 encoding and introduces the Huffman tables in MP3 encoding; our detection scheme is presented in the third section. The next section presents experimental results and discussions, followed by conclusions in the last section.

II. Huffman Tables in MP3 Encoding

A. Review of MP3 Encoding

The lossy compression algorithm of MP3 is designed to greatly reduce the amount of data required to represent the audio informations with a faithful reproduction of the original uncompressed audio for most listeners. Fig.2 illustrates the entire MP3 encoding.

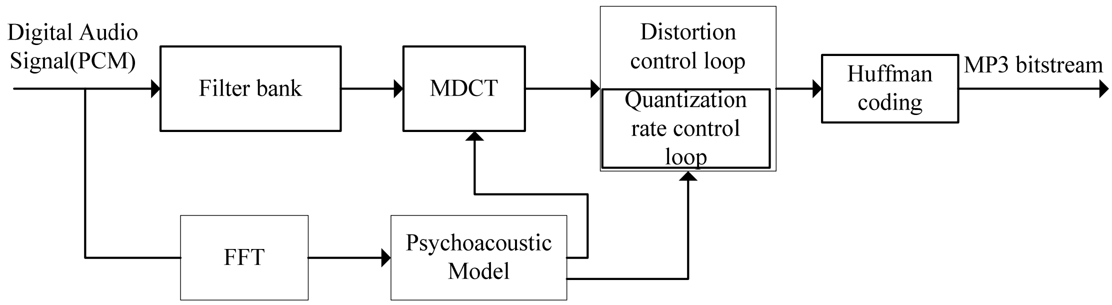


Figure 2. Block diagram of MP3 encoding

Through a polyphase filter analysis, the input PCM signals are first separated into 32 equally spaced frequency subbands, and then the MDCT window further divides each of the subbands into 18(long window) or 6(short window) finer subbands for creating a granule with a total of 576 frequency lines. The psychoacoustic model, providing information on which parts of the audio signals are audible and which parts are not, analyzes the audio content and estimates the masking threshold. This gives the minimum quantization error that can be introduced into the audio signal while still maintaining perceptually unimpaired signal quality. The 576 spectral lines are quantized according to the masking threshold obtained by the psychoacoustic model, which is done iteratively in two nested loops, a distortion control loop and a rate control loop. Finally, the quantized spectral values are encoded by using Huffman code tables to form a bitstream. The MP3 decoding is the inverse process of encoding, which contains the steps in the order of Huffman decoding, inverse quantization, inverse MDCT and aliasing cancellation, and filter-bank synthesis.

B. Huffman Tables in MP3 Encoding

In order to improve the compression efficiency, lossless Huffman coding is adopted in MP3 encoding. Quantized MDCT (QMDCT) coefficients is divided into three zones from high to low frequency: the all-zero region (Rzero), the small value region (Count1) and the large value region (Big_value), shown in Fig.3.

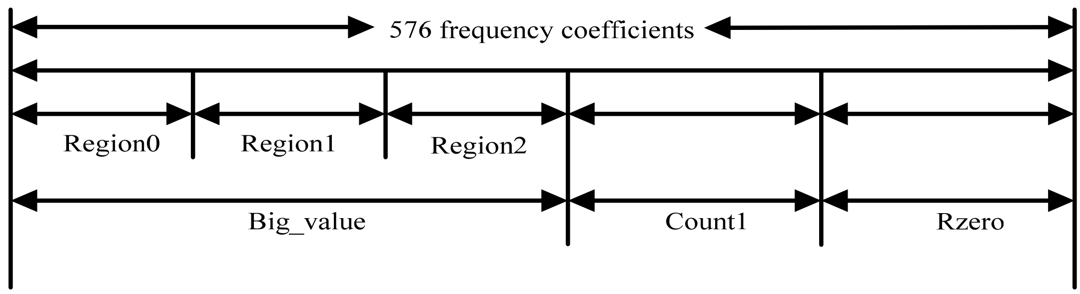


Figure 3. The partition of QMDCT coefficients

The Rzero region is not necessary to be coded because all the values in this region are zero. The other two regions are encoded from a different set of Huffman tables. For MP3, 34 Huffman tables are used to map Huffman codes to symbols. Table 32 and Table 33 are used for the count1 region. Tables 0 through 31 are for the

Big_value region, as shown in Tab.1. The Big_value region is further subdivided into three subregions (region0, 1, 2) and each has its specific Huffman table. The Huffman table used for each subregion is determined by the maximum value of corresponding region. In this paper, only tables for Big_value region are considered.

TABLE 1.

HUFFMAN TABLES IN MP3 STANDARD FOR BIG_VALUE REGION

Table index	Max value	Table index	Max value	Linbits
0	0	16	16	1
1	1	17	19	2
2	2	18	23	3
3	2	19	31	4
4	not used	20	79	6
5	3	21	271	8
6	3	22	1039	10
7	5	23	8207	13
8	5	24	31	4
9	5	25	41	5
10	7	26	79	6
11	7	27	143	7
12	7	28	271	8
13	15	29	527	9
14	not used	30	2016	11
15	15	31	8207	13

III. PROPOSED DETECTION SCHEME

A. The frequency statistics of Huffman tables in different compression bit rates

In order to identify the fake-quality MP3 audios, the first step is to find out the distinction between fake-quality MP3 audios and normal MP3 audios. As fake-quality MP3 audios are transcoded from low compression bit rate to high compression bit rate, some features must be different from the real high compression bit rate ones. It is nature to think that the ideal feature used for detecting fake-quality MP3 audios is sensitive to compression bit rate and hence can reflect the essential characteristic of MP3 audio with different compression bit rates. If such a feature can be found, it is easy to distinguish fake-quality MP3 audios from normal MP3 audios because the feature of fake-quality MP3 files must be different with that of normal MP3 audios.

So the key point is to discover the difference among different compression bit rates and extract the feature which can reflect such difference. As for different compression bit rates, high bit rate means that more bits are used for encoding the QMDCT coefficients. It is drawn from the principle of MP3 encoding that the more the bits used for encoding, the smaller the quantization step. Therefore, the QMDCT coefficients are larger at high bit rate than at low bit rate. From the description in Section 2.2, it is known that the maximum value of QMDCT coefficients determines which Huffman table is used for Huffman coding and the bigger the maximum value of QMDCT coefficients, the bigger the Huffman table index is chosen. Consequently, the number of big Huffman table index (larger than 15) used at high bit rate is larger than that at low bit rate.

The frequency distribution of Huffman table indexes is statistically analyzed. 100 wave audios with each 10s length including different styles are chosen to conduct this experiment. These wave audios are sampled with frequency of 44.1 kHz, quantized with 16 bits. The Lame MP3 codec [13] is used to generate MP3 compressed audios with different bit rates. The mean frequency distribution of Huffman table indexes for all the 500 MP3 audios is illustrated in Fig.4, where the horizontal axis represents the index of Huffman table, and the vertical axis represents the frequency.

It is clearly seen from Fig.4 that the frequency of small Huffman table indexes from 0 to 15 in low bit rate is larger than that in high bit rate and conversely the frequency of big Huffman table indexes from 24 to 31 in low bit rate is smaller than that in high bit rate. So the sum frequency of all small Huffman table indexes or big Huffman table indexes can be extracted as features characterizing different compression bit rates.

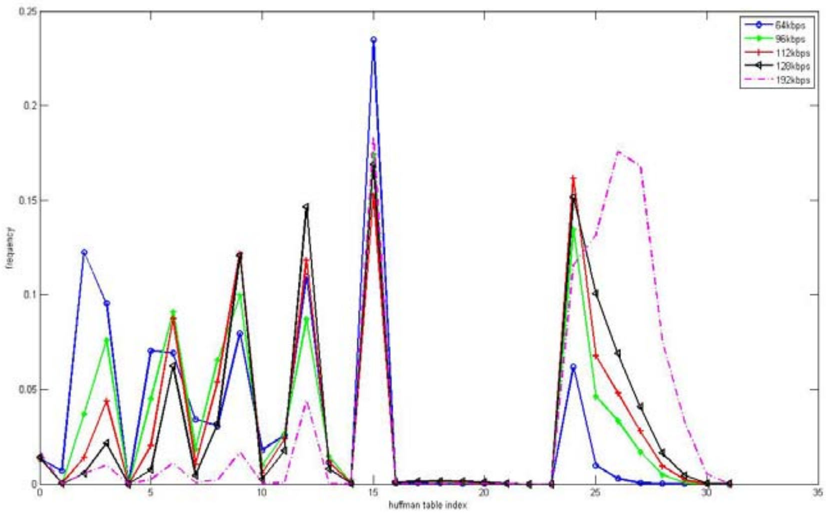


Figure 4. The frequency distribution of Huffman table indexes

B. Feature Extraction

As mentioned above, the creation of fake-quality MP3 involves double MP3 compression. When the low bit rate MP3 is decompressed at first to wav audio, the wav audio obtained after decompression suffers a quality loss because of the lossy MP3 compression. So when such wave audio is recompressed by a higher bit rate, compared with the wav audio that is not undergone MP3 decompression, more bits and lower quantization step are used during MP3 recompression in order to keep the same quality. Therefore, the QMDCT coefficients in fake-quality MP3 are bigger than that in normal MP3 at the same compression bit rate and consequently the sum frequency of big Huffman table index from 24 to 31 in fake-quality MP3 is larger than that in normal MP3. This phenomenon is clearly illustrated in Fig.5, where the horizontal axis represents the index of all 100 MP3 audios, and the vertical axis represents the sum frequency of big Huffman table index from 24 to 31.

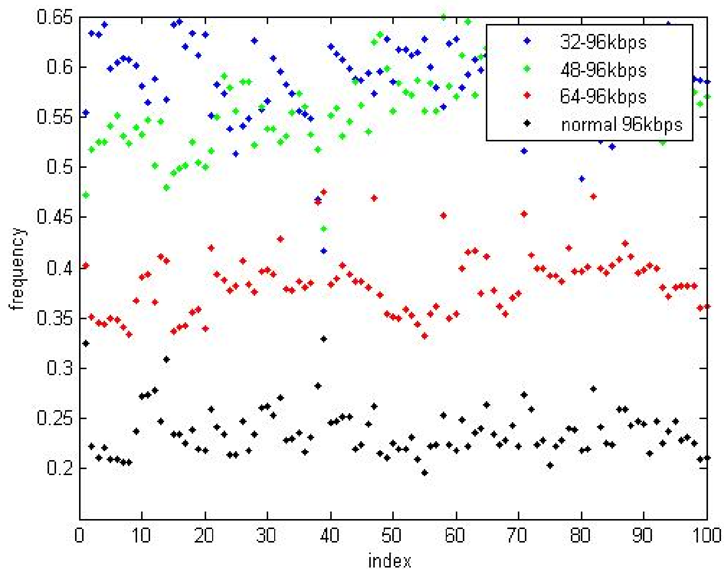


Figure 5. The distribution of sum frequency of big Huffman table indexes from 24 to 31

Suppose the sum frequency of big Huffman table indexes from 24 to 31 is denoted by F. In this paper, F is considered as the feature for detecting fake-quality MP3. It is concluded from Fig.5 that it is easy to find out fake-quality MP3 by using the feature F.

C. Proposed Method

Given a MP3 audio M under examination, our proposed detection method as follows is simple and highly effective.

Step1: Decompress M into time domain and obtain the side information

Step2: For each granule in M, obtain the Huffman table indexes used in Big_value region through side information.

Step3: Repeat Step2 until reach the end. Finally, we can get the sequence of Huffman table indexes.

Step4: Calculate the sum frequency of Huffman table indexes from 24 to 31 and then get the feature F.

Step5: Take F as input for SVM classifier [14]. Apply SVM classifier and determine whether the testing MP3 audio is fake-quality or not.

IV. Experiments

A. Experimental Setup

In the first, we create 100 single compressed MP3 audios of 10 sec each with different bit rates of 32, 48, 64, 96 and 112kbps. These MP3 audios are generated by wav audios through LAME codec which is the latest and most popular MP3 encoder nowadays. The wav audios used are sampled with frequency of 44.1 kHz, quantized with 16 bits, and mono channel is used. Then the MP3 audios with bit rates of 32, 48, 64 and 96kbps are decompressed and compressed again to higher bit rates respectively. These double compressed MP3 audios serve as fake-quality MP3 audios, while the above-mentioned single compressed MP3 audios of bit rate 48, 64, 96 and 112kbps serve as the normal MP3 audios.

The positive and the negative frequency features of the Huffman code table indexes are extracted in the procedures of decoding the single and fake-quality MP3 audio samples respectively. To distinguish the fake-quality MP3 from the single MP3 audios, ten experiments are implemented by the Support Vector Machine (SVM) for classification. The 70% features are chosen randomly for training a model, and the rest of them are used for testing in each of the experiments, finally the average testing accuracy is computed.

B. Experimental Results and Discussions

Table 2 lists the average testing detection accuracy over 10 experiments in each group wherein 'B' denotes the bit rate of the normal MP3 audios and the second-time-compression bit rate of the fake-quality MP3 audios, 'B1' denotes the first-time-compression bit rate of the double MP3 compression audios. Testing results consist of true positive (TP), true negative (TN), false positive and false negative (FN). The testing accuracy is calculated by using $(TPR+TNR)/2$, where $TPR = TP/(TP+FN)$ and $TNR = TN/(TN+FP)$

.

TABLE 2.

Average testing accuracy (%) over ten experiments for each type of detection

B1 (kbps)	B (kbps)			
	48	64	96	112
32	99.17	100	100	100
48		99.83	100	100
64			98.50	99.33
96				99.00

In Table 2, the testing result 99.83%, corresponding to the value of 'B' 64, and the value of 'B1' 48, is the detection accuracy for discriminating between the 64kbps MP3 audios which are actually transcoded from the 48kbps and the normal 64kbps MP3 audios; the testing result 98.50%, corresponding to the value of 'B' 96, and the value of 'B1' 64, is the detection accuracy for distinguishing between the 96kbps MP3 audios that are actually transcoded from the 64kbps and the normal 96kbps MP3 audios, and so on.

The experimental results show that the detection of the fake-quality MP3 audios is very good: in the ten results, five reach 100% and others are all over 98%.

V. Conclusions and Future Work

In this paper, by analyzing the frequency statistics of Huffman tables in different compression bit rates, the sum frequency of big Huffman table indexes from 24 to 31 is extracted as F for detecting fake-quality MP3 audios. Compared with the features extracted through MDCT coefficients, the extraction of our features is simple and rapid because it can be obtained directly from the MP3 bitstream without fully decoding. The experimental results show that the proposed method is highly effective to detect fake-quality MP3 audios.

In future, the main work is to research the transition probability of Huffman table indexes by combining the Hidden Markov Model (HMM) in the hope of finding more refined methods for detecting MP3 audios experienced twice or more MP3 compression.

Acknowledgement

This work is supported by the National Science Foundation of China (61170137, 61300055), Doctoral Fund of Ministry of Education of China (20103305110002), Zhejiang Province Key teaching material Construction project (ZJB2009074), Scientific Research Fund of Zhejiang Provincial Education Department (Y201119434), Zhejiang Scientific and Technical Key Innovation Team of New Generation Mobile Internet Client Software (2010R50009), Natural Science Foundation of Ningbo (2013A610057), Open Fund of Zhejiang Provincial Top Key Discipline of Information and Communication Engineering (XKXL1310).

REFERENCES

[1] J. Fridrich and J. Lukášs. Estimation of primary quantization matrix in double compressed JPEG images[C]. Proceedings of DFRWS 2003, Cleveland, OH, USA, August 5-8 2003.

[2] D. Fu, Y. Q. Shi, and W. Su. A generalized Benford's law for JPEG coefficients and its application in image forensics[C]. Proceedings of the SPIE on Security, Steganography, and Watermarking of Multimedia Contents IX, 2007, 6505: 47-48.

[3] C. Chen, Y. Q. Shi, W. Su. A machine learning based scheme for double JPEG compression detection[C]. IEEE International Conference on pattern Recognition(ICPR2008), Tampa, Floria, USA, December 2008: 1-4.

[4] B. Li, Y. Q. Shi, and J. Huang. Detecting Double Compressed JPEG Images by Using Mode Based First Digit Features[C]. IEEE International Workshop on Multimedia Signal Processing(MMSP), Cairns, Queensland, Australia, October 8-10, 2008: 730-735.

[5] Z. Fan and R. L. de Queiroz.. Identification of bitmap compression history: JPEG detection and quantizer estimation[J]. IEEE Transaction on Image Processing: Feb 2003, 12(2): 30-235.

[6] W. Wang, H. Fraid, and D. College. Exposing digital forgeries in video by detecting double JPEG compression[C]. Proceedings of the 8th workshop on Multimedia and security, New York, 2006:37-47.

[7] W. Chen and Y. Q. Shi. Detection of double MPEG video compression using first digits statistics[C]. International Workshop on Digital Watermarking(IWDW08), Busan, Korea, November 2008: 16-30.

[8] F. Huang, J. Huang, and Y. Q. Shi. Detecting Double JPEG Compression With the Same Quantization Matrix[J]. IEEE Transaction on Information Forensics and Security, 2010, 5(4): 848-856.

[9] R. Yang, Y. Q. Shi and J. Huang. Detecting fake-quality MP3. MM&Sec , 2009: 117-124.

[10] R. Yang, Y. Q. Shi and J. Huang. Detecting double compression of audio signal[C]. Media Forensics and Security II, 2010, 7541(1): 75410.

[11] Q. Liu, A. Sung and M. Qiao. Detection of Double MP3 Compression[J]. Cognitive Computation: May 2010, 2(4): 291-296.

[12] M. Qiao, A. Sung and Q. Liu. Revealing real quality of double compressed MP3 audio[C]. Proceedings of the international conference on ACM Multimedia New York, NY, USA, 2010: 1011-1014.

[13] http://lame.sourceforge.net/.

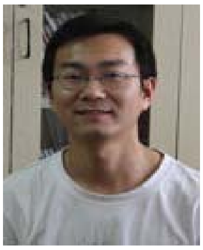
[14] C.C. Chang, C.J. Lin, LIBSVM: A library for support vector machines, Available from <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2001.

[15] Quanjin Liu, Zhimin Zhao, Yingxin Li, Xiaolei Yu, Yong Wang, A Novel Method of Feature Selection based on SVM. Journal of Computers, Aug 2013, 8 (2013): 2144-2149.

[16] Lei Zhang, Yanfei Don, Research on Diagnosis of AC Engine Wear Fault Based on Support Vector Machine and Information Fusion. Journal of Computers, Sep 2012, 9 (2012): 2292-2297.

Xianmin Yu is born in 1988. He is currently the master student in College of Information Science and Engineering, Ningbo University, China. His research interests mainly include Multimedia Information Security, digital watermarking, information hiding, audio steganalysis and audio signal processing.

Rangding Wang is born in 1962. Received his M.S. degree in the Department of Computer Science and Engineering from the Northwest Polytechnic University, Xian in 1987, and received his Ph.D. degree in the School of Electronic and Information Engineering from Tongji University, Shanghai, China, in 2004. He is now a professor at Faculty of Information Science and Engineering, Ningbo University, China. His research interests mainly include multimedia security, digital watermarking for digital rights management, data hiding, and steganography for computer forensics.



Diqun Yan received B.S and M.S. degrees in Circuit and System from Ningbo University, ZhejiangChina, in 2002 and 2008, respectively. He is currently a Ph.D. candidate in College of Information Science and Engineering, Ningbo University. His research interests include digital audio processing and data hiding.

Pengfei Ma is born in 1987. He is currently the master student in College of Information Science and Engineering, Ningbo University, China. His research interests mainly include Multimedia Information Security, double audio compression detection, information hiding, audio forensics and digital signal processing.