

# Detection of Double MP3 Compression Based on fluctuation intensity of Quantized MDCT Coefficients

Guangxin CHEN<sup>1</sup>, Xiangwei KONG<sup>1</sup>, Wei ZHONG<sup>2</sup>, Bo WANG<sup>1</sup>

1. Information Security Research Center, Dalian University of Technology, 116024, Dalian, China

2. Beijing Application Institute of Electronic Technology, 100091, Beijing, China

**Abstract:** MP3 is the most popular audio format in nowadays, and have great application in internet. For some applications of audio tampering, double-compression is often needed. The distribution of the quantized MDCT coefficients for original MP3 files approximates Laplace distribution; the distribution will be varied when the MP3 file is double compressed. In this paper, we conduct the Fourier transform on the quantized MDCT coefficients histogram sequences, and extract the fluctuation intensity with curve fitting as the feature to detect double MP3 compression. Experiments indicate that our method is highly effective for detection of double MP3 compression.

**Keywords:** Audio forensics; Double compression; MP3; MDCT coefficients

## 量化 MDCT 系数波动强度检测双重 MP3 压缩

陈光新<sup>1</sup>, 孔祥维<sup>1</sup>, 钟巍<sup>2</sup>, 王波<sup>1</sup>

1.大连理工大学信息安全研究中心, 大连, 中国, 116024

2.北京电子技术应用研究所, 北京, 中国, 100091

**摘要:** MP3 是目前最流行的音频文件格式, 在网络中得到了广泛应用。在某些音频编辑情况下可能会对 MP3 文件进行双重压缩, 原始 MP3 文件的量化 MDCT 系数近似服从拉普拉斯分布, 当音频文件经过双重 MP3 压缩时会改变原始 MP3 文件的量化 MDCT 系数分布。通过对量化 MDCT 系数直方图序列进行傅里叶变换, 使用直线拟合傅里叶变换后曲线, 获得其波动强度作为特征检测双重 MP3 压缩。实验结果表明本方法能够有效鉴别双重 MP3 压缩。

**关键词:** 音频取证; 重压缩; MP3; MDCT 系数

### 引言

随着多媒体技术的飞速发展并得到广泛应用, 网络上具有多种存储声音信息的数字音频格式, 如 MP3、WAV、MIDI、WMA 等, 在实际中都有不同的应用场合[1]。MP3 作为目前最流行的音频格式之一, 因文件存储容量小和音质高等特点使其在网络下载音乐以及数字录音机中都得到了广泛应用, 网络下载的音乐格式中 MP3 占据了主流地位, 而且大部分数字录音机也都支持 MP3 格式。

在 MP3 文件被广泛应用的同时, 不断发展的数字音频处理和编辑软件的功能也越来越强大, 如 Cool Edit, GoldWave 等, 这些软件使得对音频内容的编辑变得越来越简单, 很容易对 MP3 文件进行修改。带有恶意的音频编辑(如增加或删除音频内容等操作)可以将原始音频内容篡改, 从而达到改变原始音频语义的目的。由于篡改技术的不断增强, 使得人耳根本无法对其进行区分, 这给数字音频的真实性带来了威胁, 从而迫切需要对数字音频文件的真实性和完整性进行鉴定。

双重 MP3 压缩是指 MP3 文件被解压到时域然后

重新压缩的过程。在音频篡改中经常伴随着双重 MP3 压缩。通过检测双重 MP3 压缩, 可以初步确定该音频文件是否经过篡改, 为音频篡改检测提供有力证据。

针对这些问题, Yang<sup>[2]</sup>等人首先提出了一种根据量化 MDCT 系数值为±1 的个数检测伪造高质量 MP3 文件的方法, Brian<sup>[3]</sup>等提出了根据频谱分析确定 MP3 文件真实比特率的方法, 文献[2]、[3]算法仅适用于由低比特率转为高比特率的 MP3 文件; 后来 Yang<sup>[4]</sup>等又提出了一种基于 Benford 法则的检测双重 MP3 压缩算法, 文中根据量化 MDCT 系数首位有效数字是否符合 Benford 法则来检测音频文件是否经过双重 MP3 压缩; Liu<sup>[5]</sup>等人根据高比特率文件相对于低比特率文件含有更多的音频有效信息这一特点提取相关特征进行双重 MP3 压缩检测, 文献[4]、[5]中只对 128kbps 以下比特率的 MP3 文件进行了实验, 而且对于第一次压缩比特率高于第二次压缩比特率时的检测范围较小; Giancarlo<sup>[6]</sup>等提出了使用主观评价法和客观评价法分析 MP3 文件质量的方法, 但该方法只适用于单次压缩再解压缩的文件和前后两次压缩参数相同的 MP3 文件。根据双重压缩给量化 MDCT 系数带来的影响为出发点, 文中给出了一种新的检测双重 MP3 压缩的方法, 有效扩大了第一次压缩比特率高于第二次压缩比特率时的检测范围。

# 1.MP3 压缩和双重 MP3 压缩过程

MP3 (MPEG-1 Layer 3), 是音频压缩标准的一种。目前 MPEG 音频压缩标准共分为三个层次: MPEG-1 Layer 1、MPEG-1 Layer 2、MPEG-1 Layer 3。层数越高, 压缩复杂度就越高, 但由于 MP3 高品质和高压缩比, 使其应用最为广泛, MP3 压缩效率可达到 1:10~1:12。

## 1.1 MP3 压缩过程

MP3 文件在不影响人耳听觉效果的前提下, 通过压缩编码可以大量节省存储空间, MP3 的编码<sup>[7]</sup>是一种基于频域的编码, 其编解码基本流程分别如图 1、图 2 所示:

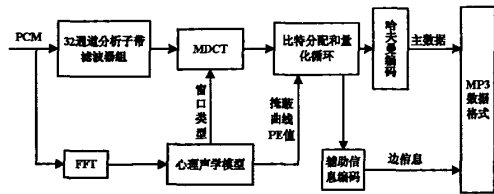


图 1 MP3 编码算法流程图

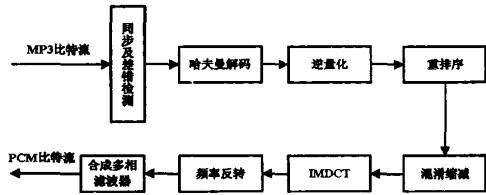


图 2 MP3 解码算法流程图

编码时, 输入的 PCM 信号首先通过 32 子带分析滤波器, 然后通过 MDCT 变换将子带的时域信号变换为 6 或者 18 个频线值 (短窗时 6 个, 长窗时 18 个)。接着通过心理声学模型分析音频内容并且估计掩蔽阈值, 以此得到在不影响音频听觉前提下的最大量化误差; 得到的频线值根据掩蔽阈值进行量化, 再将量化值进行反量化计算量化噪声, 看其是否小于掩蔽阈值, 如果大于掩蔽阈值则再次调整量化因子重新量化, 直到量化噪声最优, 结束量化过程。最后量化后的值通过 Huffman 编码后输出。

MP3 解码过程相对于编码过程较为简单。首先将 MP3 数据进行 Huffman 解码, 然后逆量化到实际的频域幅值, 最后经过 IMDCT 变换和合成多相滤波器将数据变换为时域 PCM 信号。

## 1.2 双重 MP3 压缩过程

对 MP3 文件的修改经常伴随着双重 MP3 压缩过程。图 3 (a) 所示为伪造高质量 MP3 文件过程, 为制作高比特率 MP3 文件需将低比特率文件进行解码然后再以较高的比特率进行二次编码过程。图 3 (b) 所示为篡改 MP3 文件过程, 如需对 MP3 文件进行某种方式的修改, 需先将 MP3 文件解码到时域波形,

修改完成后再进行 MP3 压缩过程。

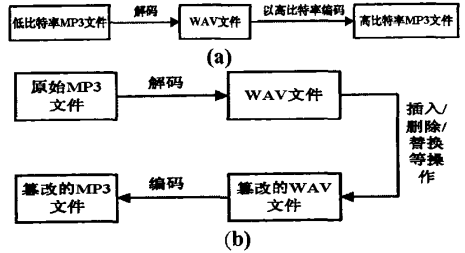


图 3. (a) 制作伪造高质量 MP3, (b) 伪造 MP3 文件

双重 MP3 压缩是指对 MP3 文件进行两次 MP3 压缩, 即要经过两次量化过程。

在重压缩情况下, 原始 MP3 的 MDCT 系数首先被量化因子  $q_1$  量化。当 MP3 文件解码时, 这些数值被截短并进行四舍五入操作。当进行第二次压缩时, MDCT 系数是从这些截短的整数值计算而来, 导致此时的 MDCT 系数会和第一次压缩时的 MDCT 系数不同, 接着这些系数被量化因子  $q_2$  再次量化, 形成重压缩的 MP3 文件。重压缩会直接影响量化 MDCT 系数的变化, 因此可以利用统计特性的变化进行重压缩的检测。

## 2.双重 MP3 压缩检测算法

双重 MP3 压缩会改变原始 MP3 文件量化 MDCT 系数的分布, 通过分析 MP3 量化 MDCT 系数统计分布, 提取量化 MDCT 系数直方图序列傅里叶变换后曲线的波动性水平作为特征, 实现了双重 MP3 压缩的有效检测。

### 2.1 量化 MDCT 系数特征

原始 MP3 文件的量化 MDCT 系数个数近似服从拉普拉斯分布<sup>[8]</sup>, MP3 文件经过双重 MP3 压缩之后, 其量化 MDCT 个数分布会发生变化, 即量化 MDCT 系数直方图分布会背离拉普拉斯分布曲线。

使用 LAME v3.97<sup>[9]</sup>软件将原始 WAV 文件<sup>[10]</sup>经过单次压缩与双重压缩, 分别统计其量化 MDCT 系数的个数并进行傅里叶变换, 观察变化规律, 如图 4 所示, 其中 (a)(c)(e)(g) 为 MP3 文件量化 MDCT 系数直方图 (其中(a)MP3 文件经过一次压缩, 压缩率为 128kbps; (c)MP3 文件经过双重压缩, 压缩率为 192kbps-128kbps (表示经过两次压缩, 第一次压缩的比特率为 192kbps, 第二次压缩的比特率为 128kbps, 下同); (e) 128kbps-128kbps; (g) 96kbps-128kbps); (b)、(d)、(f)、(h) 分别为 100 个 MP3 文件在(a)、(c)、(e)、(g)对应情况下量化 MDCT 系数傅里叶变换之后叠加示意图。

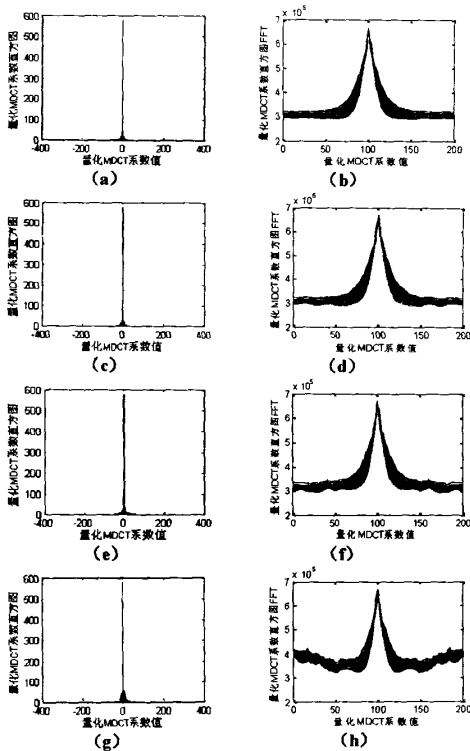


图 4. 单次 MP3 压缩与双重 MP3 压缩的量化 MDCT 系数分布

从图 4(b)可以看出, 经过单次压缩的 MP3 文件其量化 MDCT 系数直方图经傅里叶变换之后旁瓣呈光滑曲线, 近似服从拉普拉斯分布; 而通过图 4(d)(f)(h)我们可以发现经过双重压缩之后其直方图的傅里叶变换旁瓣呈现明显的波动, 与单次压缩的 MP3 文件分布有明显差异, 而且随着第一次压缩比特率的降低, 波动现象越来越明显。

## 2.2. 检测原理

以上述性质作为检测依据, 即经过双重压缩的 MP3 文件其量化 MDCT 系数的直方图序列经傅里叶变换之后旁瓣会出现波动现象, 否则不会。为了表征波动现象的出现从而对 MP3 文件是否经过双重压缩进行区分, 选取该曲线的波动性水平作为取证依据的参数, 定义如下:

$$T = \sum_{i=1}^M |Y(i) - F(i)| \quad (1)$$

式中  $F(i)$  表示量化 MDCT 系数直方图分布序列经傅里叶变换后的值,  $Y(i)$  表示拟合序列  $F$  的直线序列 (以最小二乘法拟合曲线) 值,  $M$  表示拟合点数, 为充分表示旁瓣波动区间以及适应于一定范围内的旁瓣宽度, 不相同比特率文件之间  $M$  会有所差别。算法流程如下: 首先对待测 MP3 文件进行 Huffman

解码获得量化 MDCT 系数, 然后求得系数直方图序列  $N$ , 在此基础上对序列  $N$  进行傅里叶变换得到其频域序列  $F$ , 根据式 (1) 获得波动水平特征值  $T$ , 最后根据训练得到的阈值  $Th$  判断所得傅里叶变换序列  $F$  是否有波动现象, 若有则认为双重压缩的 MP3 文件, 否则认为是原始的 MP3 文件, 双重 MP3 压缩检测算法流程如图 5 所示。

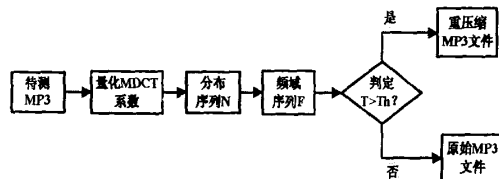


图 5 算法流程图

## 3. 实验结果与分析

我们利用 LAME v3.97 对 240 个 WAV 原始文件<sup>[10]</sup>进行双重压缩。实验中 WAV 文件分别压缩为: 96kbps, 112kbps, 128kbps, 160kbps, 192kbps, 224kbps, 256kbps, 320kbps, 然后分别解压缩再压缩为以上八种比特率, 总共获得  $240 \times 9 \times 8 = 17280$  个样本。

为验证所提算法的有效性, 与文献[4]算法进行了比较测试。对获得的  $240 \times 9 \times 8 = 17280$  个样本进行测试, 由于每种比特率的量化 MDCT 系数分布不同, 即不同比特率文件其旁瓣宽度不同, 阈值  $Th$  的选择会有所差异。为了得到精确的分类结果, 选用简单决策树算法<sup>[11]</sup>进行分类, 因为简单决策树可以获取最优阈值, 并能使平均准确率达到最高。每组 240 个音频中 140 个用来训练, 100 个用来测试, 测试结果如表 1 所示 (BR1 表示第一次压缩比特率, BR2 表示第二次压缩比特率), 结果如表 1 所示。

从实验结果可以看出, 该方法能够有效检测经过双重 MP3 压缩的文件, 特别是对  $BR1 \leq BR2$  时的双重压缩问题具有很好的检测效果, 但是当  $BR1 > BR2$  时, 该方法正确率会随着  $BR1$  的增大而降低, 但相对于文献[4]中的方法, 本算法仅使用一维特征进行检测, 有效扩大了  $BR1 > BR2$  时的检测范围, 而且检测正确率有了明显提高, 实验中发现文献[4]算法在检测时正确率存在异常, 如在 128kbps 时检测正确率为 0%, 文献中检测时也出现了异常现象。因为本算法是根据双重量化过程中所引起的 MDCT 系数的统计特性的变化, 而 MP3 压缩比特率决定其 MDCT 系数分布。当由高比特率转码为低比特率时其 MDCT 系数分布虽会改变, 但由于高比特率含有较多的有效信息, 再次压缩为低比特率时足以满足单次压缩所需的有效信息量, 所以由高比特率转码为低比特率的 MDCT 系数分布与单次压缩的相近, 导致无法区分。

表 1.测试结果对比

		BR1(kbps)								单次压缩		
		96	112	128	160	192	224	256	320			
本文 算法 实验 结果	BR2 (kbps)	96	100%	93%	77%	16%	5%	4%	4%	3%	95%	
		112	100%	100%	100%	100%	76%	21%	3%	4%	97%	
		128	100%	100%	100%	100%	99%	82%	16%	9%	89%	
		160	100%	100%	100%	100%	100%	100%	91%	28%	4%	98%
		192	100%	100%	100%	100%	100%	100%	100%	97%	56%	95%
		224	100%	100%	100%	100%	100%	100%	100%	100%	99%	98%
		256	100%	100%	100%	100%	100%	100%	100%	100%	98%	91%
		320	100%	100%	100%	100%	100%	100%	100%	80%	100%	
参考 文献[4] 实验 结果	BR2 (kbps)	96	100%	100%	100%	18%	3%	4%	0%	1%	99%	
		112	100%	100%	100%	54%	36%	32%	30%	33%	74%	
		128	100%	100%	100%	38%	2%	0%	0%	0%	100%	
		160	100%	100%	100%	77%	69%	45%	26%	24%	84%	
		192	100%	100%	100%	100%	51%	40%	20%	27%	88%	
		224	100%	100%	100%	100%	100%	100%	100%	96%	85%	
		256	100%	100%	100%	100%	100%	100%	100%	100%	88%	
		320	100%	100%	100%	100%	100%	100%	87%	84%		

#### 4.结束语

为了检测双重MP3压缩对MDCT系数进行了研究。该方法根据对文件进行双重MP3压缩时需要MDCT系数进行双重量化,改变了量化MDCT系数分布,从而提取量化MDCT系数直方图傅里叶变换后序列的波动性水平作为特征,并将这一特征应用于检测文件是否经过双重MP3压缩。实验结果表明当 $BR1 \leq BR2$ 时该方法能有效检测经过双重MP3压缩的文件,平均正确率为99.44%,当 $BR1 > BR2$ 时,在一定范围内仍然能有效检测。下一步工作将针对 $BR1 > BR2$ 的情况下进行研究,使得算法的适用性更广。

#### 参考文献

[1] 王建林. 数字音频文件格式探析[J]. 浙江传媒学院学报. 2004, 3:84-86.  
Wang JianLin. A Tentative Analysis of Digital Formats[J]. Journal of Zhejiang and Television College. 2004,3:84-86.

[2] Rui Yang, Yun Q. Shi, Jiwu Huang. Defeating Fake-quality MP3[C]. Proceedings of the 11th ACM workshop on Multimedia and security.2009, 117-124.

[3] Brian D'Alessandro, Yun Q. Shi. Mp3 Bit Rate Quality Detection through Frequency Spectrum Analysis[C].Proceedings of the 11th ACM workshop on Multimedia and security: 2009, 57-61.

[4] Yang R, Yun Q. Shi, Jiwu Huang. Detecting double compression of audio signal [C]. Media Forensics and Security II, San Jose, California, USA, 2010, 7541, 1-10.

[5] Qingzhong Liu, Andrew H.Sung, Mengyu Qiao. Detection of Double MP3 Compression[J]. Cogn Comput.2010, 2(4):291-296.

[6] Giancarlo Vercellesi, Andrea Vitali, Martino Zerbini. MP3 Audio Quality for Single and Multiple

encoding[J].2007 IEEE International Conference on Multimedia and Expo:1279-1282.

[7] 蒋学鑫. MP3 实时编解码系统的研究与开发[D]. [硕士论文]. 电子科技大学,2007.  
Jiang XueXin. The Research and Development of the Real-time MP3 Encoding and Decoding[D]. [Master Thesis]. University of Electronic Science and Technology of China.

[8] R.Yu, X. Lin, S. Rahardia and C.C.Ko. A Statistics Study of the MDCT Coefficients Distribution for audio[C].International Conference on Multimedia and Expo, Taipei, China, 2004,1483-1486.

[9] LAME v3.97. <http://lame.sourceforge.net/>, 2010.11.

[10] Audio database. <http://www.ee.columbia.edu/~dpwe/sounds/musps/scheislan.html>. 2010,11.

[11] Sergios Theodoridis, Konstantinos Koutroumbas. Pattern Recongnition [M]. 3, 机械工业出版社, 2006.9:174-178.

#### 作者简介

孔祥维(1963—), 女, 博士生导师, E-mail: kongxw@dlut.edu.cn