



基于比例因子转移概率的 AAC 音频压缩历史检测算法

黄其娟, 王让定, 严迪群

(宁波大学信息科学与工程学院, 浙江 宁波 315211)

摘要: 音频压缩历史的检测是音频取证的重要组成部分, 对判断音频是否经过篡改和伪造有着十分重要的意义。通过研究发现, AAC 音频比例因子的值会随着压缩次数的增加而逐渐减小。基于此, 提出了一种基于比例因子转移概率差值统计特性的 AAC 音频压缩历史检测算法。实验结果表明, 该算法能对多次压缩的 AAC 音频进行准确分类, 其低码率转高码率间 AAC 音频平均分类准确率达到了 99.75%, 同码率间准确率达 97.28%。另外, 对比实验也证明了本文算法的性能优于现有算法。

关键词: AAC 音频; 压缩历史; 比例因子; 转移概率

中图分类号: TP391

文献标识码: A

doi: 10.11959/j.issn.1000-0801.2017277

AAC compression detection based on scaling factor transition probability

HUANG Qijuan, WANG Rangding, YAN Diqun

College of Information Science and Engineering, Ningbo University, Ningbo 315211, China

Abstract: Audio compression history detection is an important part of audio forensics, which is important to detect whether audio has been tampered or forged. An algorithm of AAC audio compression history detection was presented by using the transition probability differences of scale factors as the discriminative feature. Experimental results demonstrate that the proposed method can distinguish the single, double and triple compressed AAC audios effectively, and from the low-bit-rate to high-bit-rate, the average classification accuracy achieves 99.75%, the same-bit-rate detection accuracy achieves 97.28%. In addition, the results of comparison experiments show that the proposed algorithm outperforms the state-of-the-art algorithm.

Key words: AAC audio, compression history, scale factor, transfer probability

收稿日期: 2017-07-13; 修回日期: 2017-09-26

通信作者: 王让定, wangrangding@nbu.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61672302, No.61300055); 浙江省自然科学基金资助项目 (No.LZ15F020010, No.Y17F020051); 宁波大学科研基金资助项目 (No.XKXL1405, No.XKXL1420, No.XKXL1509, No.XKXL1503); 宁波大学王宽诚幸福基金资助项目

Foundation Items: The National Natural Science Foundation of China (No. 61672302, No.61300055), Natural Science Foundation of Zhejiang Province of China (No.LZ15F020010, No.Y17F020051), The Scientific Research Foundation of Ningbo University (No.XKXL1405, No.XKXL1420, No.XKXL1509, No. XKXL1503), K.C. Wong Magna Fund in Ningbo University



1 引言

随着多媒体和互联网的不断发展,人们可便捷地借助音频、图像、视频设备来分享、获取各种各样的重要信息。这些多媒体信息无时无刻不萦绕在人们的生活中,在给人们生活带来便利的同时,也存在着很多信息安全隐患。信息化的快速发展使得信息安全形势愈加严峻^[1,2],由此,多媒体取证技术应运而生,并得到了快速发展。作为多媒体取证技术的重要分支之一,音频取证也得到迅速发展。

音频压缩历史检测是音频取证技术中的热点问题之一,在对音频进行篡改、插入等操作时,都会对音频进行压缩,导致音频压缩历史的变化。音频压缩历史的检测是判断音频是否被篡改或伪造的一个必要条件。

压缩域音频历史的检测主要集中于双压缩音频的检测。近年来,国内外研究者对 MP3 压缩域音频的压缩历史的检测取得了一些研究成果。D'alessandro 等人^[3]通过分析不同比特率下音频的功率谱,发现其在 16~20 kHz 频带内呈现不同分布,从而提出通过频谱分析检测 MP3 音频真实比特率的方法。Liu 等人^[4]通过修正离散余弦变换(modified discrete cosine transform, MDCT)系数绝对值高于设定阈值的比例,实现单压缩音频和双压缩音频分类,但该阈值的设定基于作者经验值,不具备普适性。Qiao 等人^[5]使用小值 MDCT 系数数量、相邻帧间过渡连续性及系数离散数值分布等特征进行双压缩音频检测。在非压缩域音频压缩历史检测方面,Luo 等人^[6]通过提取 MDCT 系数作为其压缩历史检测的有效特征,在检测是否经过有损编码器重压缩的基础上检测其压缩采用的比特率。Luo 等人^[7]通过增加梅尔倒谱系数(Mel frequency cepstrum coefficient, MFCC)作为检测特征,提高了算法检测准确率,但特征维数较高,增加了计算量。Ren 等人^[8]去除待测音频

的前 200 个采样点,重新压缩得到校准音频,并利用待测音频与校准音频量化 MDCT 系数的不同,进行 MP3 双压缩音频检测和压缩音频原始码率估计。

在 AAC 音频压缩历史检测方面,Seichter 等人^[9]利用卷积神经网络对 AAC 压缩音频进行分类并估计压缩音频的原始码率。Jin 等人^[10]使用 Huffman 码表索引出现的概率以及其 Markov 单步转移概率作为特征,利用 LIBSVM 对单压缩和双压缩 AAC 音频进行分类。虽然该方法在低码率转高码率的压缩音频取得了较好的检测效果,但在同码率的压缩音频方面,识别率较低。李晗等人^[11]利用 AAC 在重压缩过程中频率成分的变化、能量值的不同,使用堆栈自动编码器(SAE)进行双压缩历史的检测,但其并没有对具体码率的单双压缩的音频进行检测,且检测准确率较低。

本文的研究对象仅为经过三次及三次以下的 AAC 压缩音频。可以发现,在音频压缩过程中,比例因子随压缩次数而发生变化,随着压缩次数的增加,比例因子会逐渐减小。根据音频压缩前后比例因子统计特性的差异,提出了一种基于比例因子统计特性的 AAC 音频压缩历史检测算法。该算法将比例因子转移概率差值作为检测特征,利用 LIBSVM 对多压缩的 AAC 音频进行分类。实验结果表明,本文算法检测准确率较高,能有效检测 AAC 音频的压缩历史。

2 AAC 音频比例因子特性变化分析

2.1 AAC 中的比例因子

AAC 是高级音频编码(advanced audio coding)的简称,是 ISO/IEC MPEG-2 和 MPEG-4 标准的重要组成部分^[12]。MPEG-2 AAC 于 1997 年发布,因其在相同码率,尤其是低码率的情况下,有着比 MP3 音频更好的音质^[13],因此被看好能取代 MP3 成为新一代的音频编码标准。AAC 编码框架如图 1 所示,主要由心理声学模型、增益控

制、滤波器组、时域噪声整形、强度立体声耦合、预测、M/S 编码、量化和无噪声编码等部分组成。

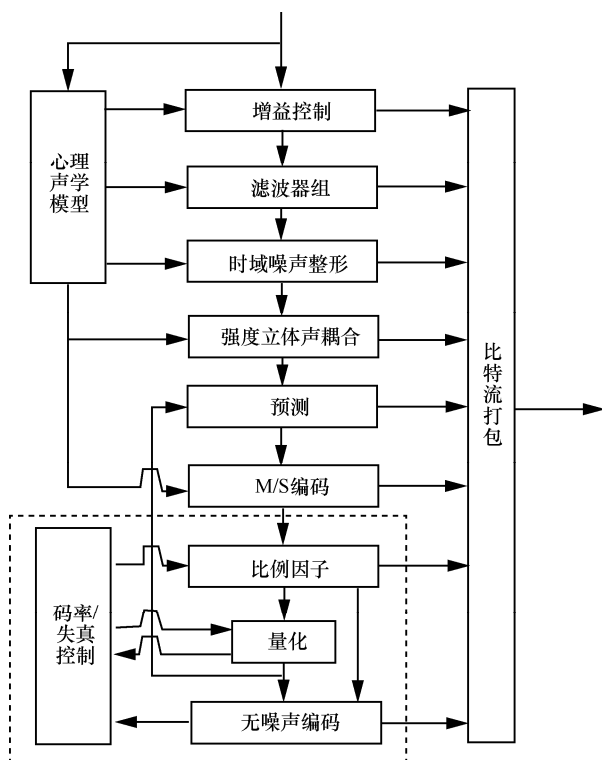


图 1 AAC 编码框架

AAC 进行 MDCT 时使用长窗块、短窗块、开始块和结束块这 4 种变换块类型：瞬时特性变化较为平稳的块使用长窗，变化较为剧烈的块使用短窗，另外两种块则用于长窗块和短窗块的过渡。使用长窗块时，AAC 具有较高的频域分辨率和较高的编码效率。而当音频信号变化剧烈时，AAC 通过使用短窗块来提高信号的时域分辨率。本文则主要关注长窗块情况下的编码。

AAC 标准将 1 024 个频域系数分成若干个连续的组分别进行量化，这样的组就称为比例因子带。比例因子带依据人耳听觉系统的临界带划分，其个数和长度取决于当前块类型和原始音频的采样率。如对原始采样率为 44.1 kHz 的音频，AAC 将其长块和短块分别划分为 49 个和 14 个比例因子带。每个比例因子对应一个比例因子带，AAC 中比例因子采用差分编码和 Huffman 编码的方式，比例因子的最大值可取到 255。

AAC 量化模块利用心理声学模型计算所得掩蔽阈值，求出各比例因子带的最佳比特分配方案。量化由帧循环、外循环和内循环 3 个迭代循环完成，其中帧循环负责初始化变量、计算可用比特数，外循环将量化环节的失真限制在可控范围内，内循环则实施对频谱系数的具体量化。AAC 内循环中的量化函数式为：

$$x_{\text{quant}} = \left\lfloor \left((|mdct_{\text{line}}| \times 2^{\frac{1}{4}(sf_{\text{decoder}} - 100)})^{\frac{3}{4}} + 0.4054 \right) \right\rfloor \quad (1)$$

其中， $\lfloor \cdot \rfloor$ 是向下取整函数， sf_{decoder} 表示量化步长， $mdct_{\text{line}}$ 、 x_{quant} 分别是量化前后的频谱系数。通过调整量化步长调节比例因子表示范围，将每个比例因子带的量化失真控制在最大可允许失真范围内，增加量化步长等价于间接减小比例因子^[14]。由于信号中高频分量的能量值较小，使得需要用较小的量化步长来保留高频信号的精度。在二次压缩之后，很多的高频信息会被量化成零值^[3]，相比一次音频压缩其量化步长增大，比例因子值减小。根据一次压缩和二次压缩时的统计信息，量化步长和比例因子是 AAC 二次压缩检测的重要特征。

2.2 比例因子特性变化分析

AAC 压缩音频（多次压缩）分为以下 3 类：低码率转高码率的 AAC 音频，这类压缩音频通常被称为假音质音频，由较差音质（低码率）的音频转码形成的虚假高品质（高码率）音频；相同码率的 AAC 音频，该类音频的产生多是由于原始压缩音频被篡改后，为了不让格式的改变引起他人怀疑而重新以相同码率重新压缩回篡改前的音频格式；高码率转低码率的 AAC 音频，这种情况虽也有可能发生，但目前在取证场景下的意义仍未被提出。因此，本文后续的压缩音频默认为上述第一类和第二类音频。

为验证 AAC 多次压缩对比例因子的影响，随机选取 100 首 10 s 的 WAV 音频，包括乡村、蓝调、流行、爵士等风格。将这 100 首 WAV 音频压



缩成 60 kbit/s、120 kbit/s 的 AAC 音频, 然后继续进行压缩操作, 分别得到 60 kbit/s→60 kbit/s、60 kbit/s→120 kbit/s 的二次压缩音频, 60 kbit/s→60 kbit/s→60 kbit/s、60 kbit/s→120 kbit/s→120 kbit/s 的 3 次压缩音频。提取每种码率 AAC 音频的比例因子, 并对其出现概率进行统计。AAC 比例因子的范围为[0,255], 如图 2 所示, 其比例因子值的分布近似服从拉普拉斯分布, 且主要值集中于[140,200], 为保证研究的普适性, 本算法仅对这个范围内的比例因子出现概率进行统计。

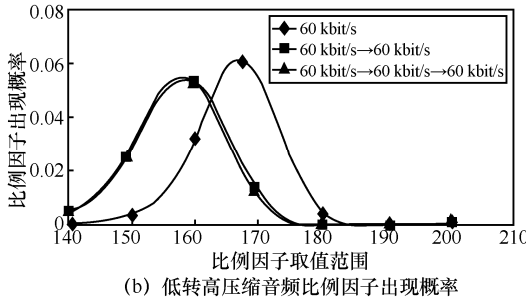
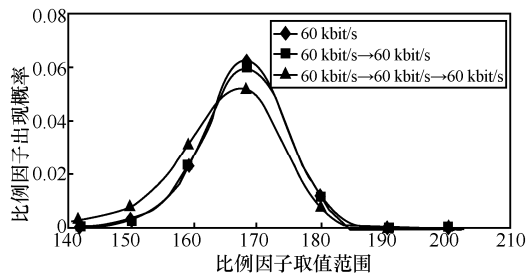


图 2 100 首 AAC 音频比例因子出现概率

由图 2 可以看出, 随着压缩次数的增加, 比例因子值相对减小。其中, 同码率音频压缩比例因子变化较小, 由低码率转高码率的音频比例因子变化相对较大。由此认为可通过分析比例因子的统计特性来判断 AAC 的压缩历史。由图 2 还可以看出, AAC 音频压缩使比例因子发生相应变化, 但压缩音频与原始音频比例因子较为接近, 难以区分音频的压缩历史。且随着压缩次数的增加, 比例因子最大值即曲线峰值的出现概率逐渐减小, 而其他比例因子的出现概率也会发生变换, 这说明压缩会使得部分比

例因子发生转化, 即可能会发生比例因子的相应转移。

3 特征构造与提取

假设一次压缩 AAC 音频为 $M1$, 对其进行解压再压缩, 得到新的 AAC 音频 $M2$, 然后继续进行三次压缩, 得到 AAC 音频 $M3$, 如图 3 所示。在本文的研究中, 假设二次压缩和三次压缩得到的 $M2$ 、 $M3$ 音频码率相等。

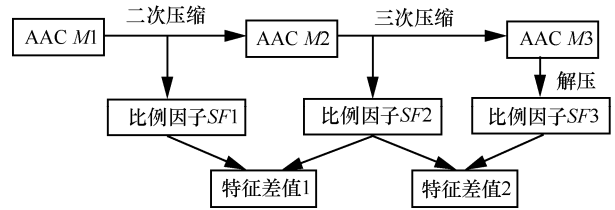


图 3 特征提取流程

图 3 中 $SF1$ 、 $SF2$ 、 $SF3$ 表示将一次压缩 AAC 音频 $M1$ 、二次压缩 AAC 音频 $M2$ 和三次压缩 AAC 音频 $M3$ 解压得到的比例因子矩阵, 可将其定义为:

$$SF1 = \begin{pmatrix} sf_{1,1} & \cdots & sf_{1,n} \\ \vdots & \ddots & \vdots \\ sf_{m,1} & \cdots & sf_{m,n} \end{pmatrix} \quad (2)$$

$$SF2 = \begin{pmatrix} sf'_{1,1} & \cdots & sf'_{1,n} \\ \vdots & \ddots & \vdots \\ sf'_{m,1} & \cdots & sf'_{m,n} \end{pmatrix} \quad (3)$$

$$SF3 = \begin{pmatrix} sf''_{1,1} & \cdots & sf''_{1,n} \\ \vdots & \ddots & \vdots \\ sf''_{m,1} & \cdots & sf''_{m,n} \end{pmatrix} \quad (4)$$

其中, n 代表比例因子带的个数, 若为单声道, 则 $m = 2I$; 若为双声道, 则 $m = 4I$ 。 I 代表待测音频长窗帧的总数。

本文统计了 140~200 内的比例因子, 但由于 140~150、180~200 内的比例因子出现概率过小, 在统计转移概率时, 仅选取在 150~180 内的比例因子进行统计。文中定义了两种转移概率的情况, 分别是块内转移概率 $p\{sf_{i,j+1}=n | sf_{i,j}=m\}$

和块间转移概率 $p\{sf_{i+1,j} = n | sf_{i,j} = m\}$, 即:

$$p\{sf_{i+1,j} = n | sf_{i,j} = m\} = \frac{\sum_{i=1}^{31} \sum_{j=1}^{30} \delta(sf_{i,j} = m, sf_{i+1,j} = n)}{\sum_{i=1}^{31} \sum_{j=1}^{30} \delta(sf_{i,j} = m)} \quad (5)$$

$$p\{sf_{i+1,j} = n | sf_{i,j} = m\} = \frac{\sum_{i=1}^{30} \sum_{j=1}^{31} \delta(sf_{i,j} = m, sf_{i+1,j} = n)}{\sum_{i=1}^{30} \sum_{j=1}^{31} \delta(sf_{i,j} = m)} \quad (6)$$

其中, $m, n \in \{150, 151, \dots, 180\}$, $\delta(X=i, Y=j) = \begin{cases} 1, X=i, Y=j \\ 0, \text{其他} \end{cases}$, $sf_{i,j}$ 是比例因子值。式 (7)~式 (9)

所示为得到的 $M1$ 、 $M2$ 、 $M3$ 的转移概率矩阵 $P1$ 、 $P2$ 和 $P3$, 将 31×31 组元素重组得到 961 维特征:

$$P1 = \begin{pmatrix} p_{1,1} & \cdots & p_{1,31} \\ \vdots & \ddots & \vdots \\ p_{31,1} & \cdots & p_{31,31} \end{pmatrix} \quad (7)$$

$$P2 = \begin{pmatrix} p'_{1,1} & \cdots & p'_{1,31} \\ \vdots & \ddots & \vdots \\ p'_{31,1} & \cdots & p'_{31,31} \end{pmatrix} \quad (8)$$

$$P3 = \begin{pmatrix} p''_{1,1} & \cdots & p''_{1,31} \\ \vdots & \ddots & \vdots \\ p''_{31,1} & \cdots & p''_{31,31} \end{pmatrix} \quad (9)$$

选取 100 首 $M1$ 、 $M2$ 和 $M3$ 音频, 提取比例因子块内和块间转移概率, 取其均值。图 4 所示为 $M1$ 、 $M2$ 和 $M3$ 块内和块间转移概率。图 4 中 (1)、(2)、(3) 分别代表 $M1$ 、 $M2$ 和 $M3$, 横坐标共 961 维, 代表比例因子由 M 转变为 N , M 、 N 如式 (10) 和式 (11) 所示, 纵坐标代表比例因子的转移概率:

$$M = x / 31 + 150 \quad (10)$$

$$N = x \% 31 + 149 \quad (11)$$

其中, $x = 1, 2, \dots, 961$ 。

从图 4(a) 和图 4(b) 中可以看出, $M1$ 、 $M2$ 和 $M3$ 三次压缩的比例因子转移概率差别较小, 无法利用其进行压缩音频的分类。

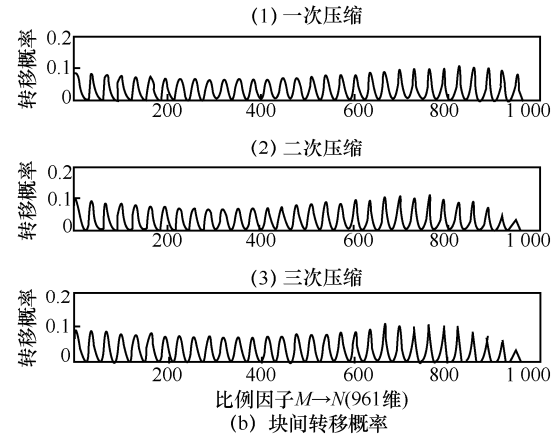
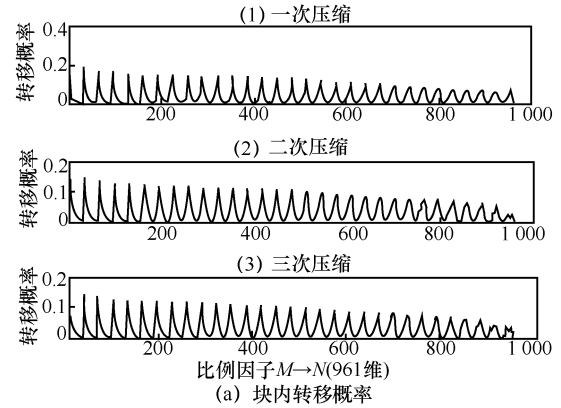


图 4 AAC 压缩音频的转移概率

利用得到的转移概率矩阵 $P1$ 、 $P2$ 和 $P3$, 计算转移概率的差值矩阵 $\Delta P1$ 、 $\Delta P2$, 如式 (12) 和式 (13) 所示:

$$\Delta P1 = P2 - P1 = \begin{pmatrix} \Delta p_{1,1} & \cdots & \Delta p_{1,31} \\ \vdots & \ddots & \vdots \\ \Delta p_{31,1} & \cdots & \Delta p_{31,31} \end{pmatrix} \quad (12)$$

$$\Delta P2 = P3 - P2 = \begin{pmatrix} \Delta p'_{1,1} & \cdots & \Delta p'_{1,31} \\ \vdots & \ddots & \vdots \\ \Delta p'_{31,1} & \cdots & \Delta p'_{31,31} \end{pmatrix} \quad (13)$$

由上述 100 首 $M1$ 、 $M2$ 和 $M3$ 音频计算得到的比例因子转移概率差值散点图如图 5 所示, 图 5(a) 为比例因子块内转移概率差值, 图 5(b) 为比例因子块间转移概率差值, 图中“+”代表 $M1$ 、 $M2$ 差值, “o”代表 $M2$ 、 $M3$ 差值。从图 5 中可以看出, $M1$ 、 $M2$ 转移概率差值大于 $M2$ 、 $M3$ 的差值, 且区别较为明显。

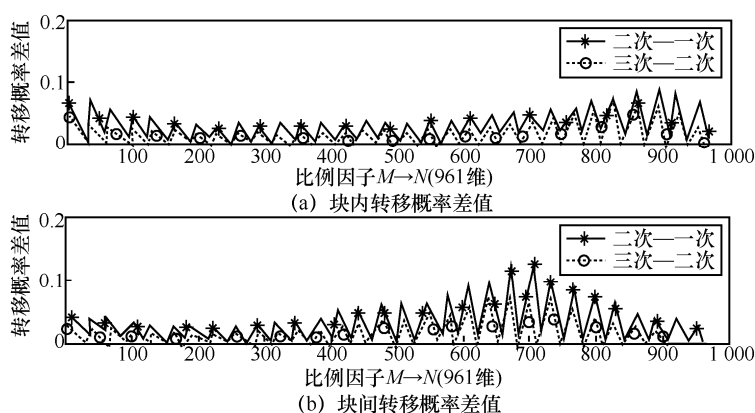


图5 压缩 AAC 音频的转移概率差值

4 实验结果与分析

4.1 实验设置

本文采用的数据库是随机选取的 800 首 10 s 长的 WAV 音频。这 800 首 WAV 音频包含各种不同的风格,如蓝调、流行、古典、乡村和民谣等。实验中使用的 AAC 编解码器为使用最为广泛的开源软件 FAAC-1.28 和 FAAD2-2.7。对上述数据库的 800 首 WAV 音频使用 FAAC-1.28 进行压缩获得单压缩的 AAC 音频,共 7 种比特率,60 kbit/s、75 kbit/s、90 kbit/s、120 kbit/s、135 kbit/s 和 150 kbit/s 各 800 首。而双压缩音频是单压音频在经 FAAD2-2.7 解码后再次使用 FAAC-1.28 编码得到的音频,共得到 28 类二次压缩音频。再对该二次压缩 AAC 音频样本解压重压缩,得到 28 类与二次压缩码率相同的三次压缩 AAC 音频样本。

本文选择 LIBSVM 作为分类器,并从上述一次、二次和三次压缩样本中随机选取 70%用于训练模型,而剩下的 30%作为测试模型。为了使测试的结果更加准确,表 1~表 4 中的检测准确率均是训练和测试重复 10 次结果的平均值。

为表述方便,将比例因子块内转移概率差值记为特征集 $SFTPD_{intra}$,块间转移概率差值记为特征集 $SFTPD_{inter}$ 。

4.2 检测结果

在本实验中,计算比例因子转移概率时,共

有 961 维特征,特征维数过高使得在用 SVM 进行分类时,其效果并不是很理想。因此,使用 SVM-RFE (recursive feature elimination based on SVM, 基于 SVM 的递归特征消除)^[15]算法对选出的特征按照从优到劣进行排序。SVM-RFE 算法于 2002 年由 Guyon 等人提出。SVM-RFE 是在 RFE 的过程中使用 SVM 作为分类器,它从全集出发,根据所训练分类器的各项参数来制定特征排序的标准,逐个剔除最不相关的特征,并通过迭代直到剩下最后一个特征,从而实现特征的排序。

(1) 块内转移概率检测结果

表 1 是特征集 $SFTPD_{intra}$ 对 AAC 音频压缩历史的检测结果,其中 BR1、BR2 和 BR3 表示第一次、第二次和第三次压缩时所设置的码率,本文中仅考虑 BR2=BR3 的情况,而表 1 中的值则代表具体的检测准确率。以表中第 6 行第 8 列的值 (99.93%) 为例,该值表示本方法对先后使用 120 kbit/s 和 150 kbit/s 码率压缩 (FAAC/FAAD2) 的双压缩 AAC 音频的综合检测率达到了 99.93%。

由表 1 可知,低码率转高码率音频检测的平均正确率达到 99.88%,同码率间检测的平均正确率为 97%。该特征集在低码率转高码率的情况下取得了较好的检测效果,均超过了 99%。但针对相同码率的压缩音频,检测率相对偏低约 2.88 个百分点,这是由于在音频压缩码率相同时,其比例因子变化较小,转移概率差值变化也相对较小。

表 1 特征集 $SFTPD_{intra}$ 的检测准确率

BR1	BR2=BR3						
	60	75	90	105	120	135	150
60	100%	100%	100%	100%	100%	100%	100%
75		94.5%	100%	100%	100%	100%	100%
90			97.5%	99.5%	100%	100%	100%
105				96.5%	100%	100%	100%
120					95.83%	100%	99.93%
135						98%	99.17%
150							96.67%

(2) 块间转移概率检测结果

表 2 是特征集 $SFTPD_{inter}$ 对 AAC 音频压缩历史的检测结果。其低码率转高码率音频检测的平均正确率达 99.67%，同码率间检测的平均正确率为 96.17%。由检测结果可知，该方法对先后经过 120 kbit/s 和 150 kbit/s 的双压缩 AAC 音频检测准确率仅为 96.83%，比特征集 $SFTPD_{intra}$ 低约 3.1 个百分点，且与之前规律类似，利用特征集 $SFTPD_{intra}$ 的检测准确率在低码率转高码率上较高，在同码率间检测结果上表现一般。

表 2 特征集 $SFTPD_{inter}$ 的检测准确率

BR1	BR2=BR3						
	60	75	90	105	120	135	150
60	100%	100%	99.67%	100%	100%	100%	99.67%
75		96%	100%	99.83%	100%	99.83%	100%
90			96%	98.5%	99.83%	100%	100%
105				94.17%	100%	100%	100%
120					96.33%	99.83%	96.83%
135						95.5%	99%
150							95.17%

以上结果表明，特征集 $SFTPD_{intra}$ 与特征集 $SFTPD_{inter}$ 的检测准确率几乎相同，两个特征集对 AAC 压缩音频的检测准确率都较为理想。但是，由于在统计比例因子转移概率时，仅统计了在 150~180 内比例因子的转移情况，其特征集 $SFTPD$ 统计的范围并不是最为全面的，这些因素会使 AAC 压缩音频检测的准确率降低。

(3) 融合特征检测结果

本文选取块内、块间转移概率差值作为检测 AAC 压缩音频的特征，而块内、块间转移概率的差值特征均属于单一特征。为证明实验的可靠性，现将块内、块间的转移概率差值特征融合，即将块内、块间转移概率差值相加后，使用 SVM-RFE 对特征进行排序并选择，最终得到融合特征。

表 3 是融合特征对 AAC 音频压缩历史的检测结果，其低码率转高码率音频检测的平均正确率达 99.69%，同码率间检测的平均正确率为 97.67%。与特征集 $SFTPD_{intra}$ 、 $SFTPD_{inter}$ 对比发现，其检测准确率在低码率转高码率上相差不大，但在同码率间检测上有所提升，融合特征检测准确率比 $SFTPD_{intra}$ 高约 0.67 个百分点，比 $SFTPD_{inter}$ 高约 1.5 个百分点。总体而言，融合特征在 AAC 压缩音频分类准确率上表现较好。

表 3 融合特征的检测准确率

BR1	BR2=BR3						
	60	75	90	105	120	135	150
60	100%	99.83%	100%	100%	100%	100%	99.5%
75		96.83%	100%	100%	99.67%	99.83%	100%
90			98.67%	99.5%	100%	100%	97%
105				97.17%	100%	99.67%	100%
120					96.67%	99.83%	98.67%
135						96.67%	100%
150							97.67%

4.3 对比试验

为了更加全面地评估本文提出的 AAC 压缩音频的检测方法，重构了 Jin 等人^[10]方法的特征，并与其进行了对比。在参考文献[10]中，具体分类特征构建的方法是：Huffman 码表在 AAC 音频一次压缩和二次压缩的使用情况存在差异，将 Huffman 码表索引出现的概率作为第一特征，将

表 4 参考文献[10]与文中特征集 $SFTPD_{intra}$ 特征的双压缩检测结果比较

BR1	BR2=BR3						
	60	75	90	105	120	135	150
60	+22.83%	0	0	0	0	0	0
75		+15.92%	0	0	0	0	0
90			+21.75%	-0.33%	0	0	0
105				+22.87%	+1.5%	0	0
120					+16.66%	+2.33%	+0.03%
135						+15.92%	-0.26%
150							+11.17%

其 Markov 单步转移概率作为第二特征,将第一特征与第二特征融合得到第三特征,并用 LIBSVM 对单压缩和双压缩 AAC 音频进行分类,其低码率转高码率检测的平均正确率达 99.72%,同码率间检测的平均正确率为 77.38%。表 4 是参考文献[8]实验检测率与本文特征集 $SFTPD_{intra}$ 检测率对比,“+”代表本文算法结果高于参考文献[10]检测率的百分点,“-”代表本文算法结果低于参考文献[8]检测率的百分点。

上述结果表明, Jin 等人^[10]方法的特征对低码率转高码率的 AAC 音频其检测率较高,与本文算法中的比例因子转移概率差值的检测结果相差不大,但是本文算法解决了参考文献[10]中同码率间的检测效果较差的问题,将检测准确率提高了约 18 个百分点。

5 结束语

本文针对多次压缩的 AAC 音频,提出了一种基于比例因子转移概率差值统计特性的 AAC 音频压缩历史检测算法。通过研究压缩过程中比例因子的变化,利用其转移概率差值特征对多压缩 AAC 音频进行分类。实验结果表明,低码率转高码率间 AAC 音频检测准确率达 99.75%,而同码率间检测准确率也达 97.28%。虽然该算法在检测 AAC 音频的压缩历史时准确率较高,但也存在着一定的局限性。如未考虑噪声等其他因素攻击情况下算法的顽健性;仅针对一种编解码算法,未

涉及相同编码标准的不同编码算法的压缩历史检测算法。因此在今后的工作中将对上述问题进行更深入的研究。

参考文献:

- [1] 郑志彬. 信息网络安全威胁及技术发展趋势[J]. 电信科学, 2009, 25(2): 28-34.
ZHENG Z B. Overview of mobile communication services security[J]. Telecommunications Science, 2009, 25(2): 28-34.
- [2] 王帅, 汪来富, 金华敏, 等. 网络安全分析中的大数据技术应用[J]. 电信科学, 2015, 31(7): 145-150.
WANG S, WANG L F, JIN H M, et al. Big data application in network security analysis[J]. Telecommunications Science, 2015, 31(7): 145-150.
- [3] D'ALESSANDRO B, SHI Y Q. MP3 bit rate quality detection through frequency spectrum analysis[C]//ACM Workshop on Multimedia and Security, September 7-8, 2009, Princeton, USA. New York: ACM Press, 2009: 57-62.
- [4] LIU Q, SUNG A H, QIAO M. Detection of double MP3 compression[J]. Cognitive Computation, 2010, 2(4): 291-296.
- [5] QIAO M, SUNG A H, LIU Q. Improved detection of MP3 double compression using content-independent features[C]//IEEE International Conference on Signal Processing, Communication and Computing, Aug 5-8, 2013, Kunming, China. New Jersey: IEEE Press, 2013: 1-4.
- [6] LUO D, LUO W Q, YANG R, et al. Compression history identification for digital audio signal[C]//2012 IEEE International Conference on Acoustics, Speech and Signal Processing, March 25-30, 2012, Kyoto, Japan. New Jersey: IEEE Press, 2012: 1733-1736.
- [7] LUO D, LUO W Q, YANG R, et al. Identifying compression history of wave audio and its applications[J]. ACM Transactions on Multimedia Computing, Communications, and Applications, 2014, 10(3): 1-19.
- [8] REN Y, FAN M, YE D, et al. Detection of double MP3 com-

- pression based on difference of calibration histogram[J]. Multimedia Tools & Applications, 2016, 75(21): 13855-13870.
- [9] SEICHTER D, CUCCOVILLO L, AICHROTH P. AAC encoding detection and bitrate estimation using a convolutional neural network[C]//IEEE International Conference on Acoustics, Speech and Signal Processing, March 20-25, 2016, Shanghai, China. New Jersey: IEEE Press, 2016: 2069-2073.
- [10] JIN C, WANG R, YAN D, et al. An efficient algorithm for double compressed AAC audio detection[J]. Multimedia Tools & Applications, 2016, 75(8): 4815-4832.
- [11] 李晗. AMR 和 AAC 音频双压缩检测研究[D]. 广州: 华南理工大学, 2015.
- LI H. Research on AMR and AAC audio dual compression detection[D]. Guangzhou: South China University of Technology, 2015.
- [12] INSTITUTION B S. ISO/IEC 13838-7/FPDAM 1. Information technology. Generic coding of moving pictures and associated audio information. Part 7: advanced audio coding (AAC). Amendment 1: signalling of bandwidth extension[S].
- [13] Wikipedia. Advanced audio coding[EB/OL]. (2011-06-15) [2017-07-13]. https://de.wikipedia.org/wiki/Advanced_Audio_Coding.
- [14] GAO Z H, WEI G. The core technology of the broadband MP3 audio compression[J]. Electro Acoustic Technology, 2000, 9(5): 9-13.
- [15] WESTON J, GUYON I. Support vector machine-recursive feature elimination (SVM-RFE): US, US 8095483 B2[P]. 2012.

[作者简介]



黄其娟 (1993-), 女, 宁波大学信息科学与工程学院硕士生, 主要研究方向为多媒体通信与信息安全等。



王让定 (1962-), 男, 博士, 宁波大学高等技术研究院教授、博士生导师, 主要研究方向为多媒体通信与取证、信息隐藏与隐写分析、智能抄表及传感网络技术。



严迪群 (1979-), 男, 博士, 宁波大学信息科学与工程学院副教授、硕士生导师, 主要研究方向为多媒体通信、信息安全、基于深度学习的数字语音取证等。