

面向流媒体的 MP3 码流无损数字音频水印*

邹婷婷 王朔中[†] 冯国瑞 王露贤

(上海大学通信与信息工程学院 上海 200072)

摘要 根据 MP3 文件的结构特点, 提出一种无损的数字音频水印算法。该算法在 MP3 码流中嵌入水印, 并不修改 MP3 文件的音频数据, 仅利用文件中帧头的私有位作为标记, 标示该帧主数据中的特定位置是否为水印信息。该方法不损伤音频质量, 具有优良的不可感知性。接收端不需要无水印的原始音频信号, 可顺序进行盲提取, 而且计算量小, 适合于流媒体应用。

关键词 MP3, 音频水印, 私有位, 流媒体

中图分类号: TP391; TP309

文献标识码: A

文章编号: 1000-310X(2012)04-0277-05

Lossless audio watermarking in MP3 files suitable for streaming media applications

ZOU Tingting WANG Shuozhong FENG Guorui WANG Luxian

(School of Communication and Information Engineering, Shanghai University, Shanghai 200072)

Abstract This paper presents a digital watermarking scheme for MP3 files. Secret information is hidden in the MP3 stream without actually changing the audio data based on a detailed analysis of the MP3 file structure. The private bit in each frame head is used to indicate the presence of secret data so that blind extraction of the watermark is straightforward. Since the audio data are unchanged, the watermark does not introduce any distortion to the audio signal. Both embedding and extraction of the watermark are performed sequentially with low computational complexity, and therefore suitable for streaming media applications.

Key words MP3, Audio watermarking, Private bit, Streaming media

1 引言

随着数字多媒体技术的发展及互联网的广泛使用, 数字音像制品的版权保护问题越来越受到人们的关注。数字音频水印是版权保护的有效方法。音乐界及有关人士于 1999 年发起 SDMI (Secure Digital Music Initiative), 建议在 Internet 上发行的

CD 和音乐数据都应加入数字水印。在过去十几年中, 人们提出一系列针对原始音频信号音频水印算法: 回声掩蔽法、变换域法、量化法、分块处理法等。

MP3 是 Internet 上最流行的音乐格式, 针对 MP3 格式的水印是数字音频水印的重要研究课题, 主要有两类方法: 在压缩过程进行嵌入的方法和 MP3 码流方法。第一类方法根据 MP3 编码的特点,

2011-12-31 收稿; 2012-03-22 定稿

*国家自然科学基金项目资助(60832010)

作者简介: 邹婷婷 (1989-), 女, 江西抚州人, 硕士研究生, 研究方向: 数字音频水印。

王朔中 (1943-), 男, 教授, 博士生导师。

冯国瑞 (1976-), 男, 副研究员。

王露贤 (1988-), 女, 硕士研究生。

[†]通讯作者: 王朔中, E-mail: shuowang@shu.edu.cn

在压缩编码过程中寻找适合的位置将数据嵌入。Petitcolas 提出的 MP3stego^[1]在压缩过程中通过调整量化误差改变当前音频帧的编码长度,将水印信息隐藏进 MP3 文件。这种方法计算复杂度高,嵌入较耗时,难以满足网络传播的实时性要求。此外研究者提出了在 MP3 压缩过程中嵌入水印的算法^[2-3],通过改变 MDCT 系数同时实现 MP3 编码和水印嵌入。

第二类方法即 MP3 码流法是根据 MP3 文件格式的特点,在码流中嵌入水印而不涉及编码过程^[5-6]。可用两种方法嵌入水印:一种是直接嵌入在 MP3 码流中,根据水印数据使边信息中的比例因子增加 1 或减小 1,不会被人耳感知;另一种是将水印嵌入到 MP3 码流的主信息中,将编码样本数据加 1 或减 1 来承载水印信息。为了减少失真,在若干个编码样本中随机选择 1 到 2 个样本来嵌入水印。此外还有基于 Huffman 编码的 MP3 信息隐藏方法^[7],通过改变部分 Huffman 码字达到嵌入信息的目的,并确保在听觉上与原始音频无差异,该方法的优点是嵌入信息量大。

本文提出一种新的音频水印嵌入方法,属于 MP3 码流法,这种方法并不修改 MP3 文件中的音频数据,而是将帧头中的私有位作为水印标识,用来指示该帧主数据中特定的 n 位是否为水印数据。除了在 MP3 文件中嵌入水印的常规应用外,由于无需介入 MP3 编码过程,该方法适合于以流媒体方式通过网络传输 MP3 音乐,在发送端实时嵌入水印,

接收端边播放边实时提取水印。

2 MP3 文件格式简述

MP3 全称 MPEG Audio Layer III^[8],它所产生的码流是以帧为单位编排的。对于单声道音频,一帧分为两个粒度,共包含 1152 个音频样本,每个样本有 16 比特。MP3 编码时,音频样本首先由分析子带滤波器转换成 32 个等带宽的子带信号,通过修正的离散余弦变换(MDCT)将每个子带再细分为 18 个次频带。根据心理学模型对每个次频带信号分配比特并量化,再从低频至高频依次进行 Huffman 编码,编码后数据按照规定的比特流格式输出。

MP3 比特流格式如图 1 所示,每一帧包含帧头、帧校验(CRC 校验)、边信息、主信息四个主要部分。帧头的长度固定为 32 比特,包含帧同步、比特率、采样率等位,其中第 24 位为私有位,可供用户使用,我们将利用该位作为水印标记。帧校验的长度为 16 比特,有无可选,由帧头中的保护比特位指示。边信息长度受信道模式控制,单声道时边信息长度为 17 字节,立体声、联合立体声和双声道时边信息长度为 32 字节。边信息含有主音频数据开始位置、Huffman 码表选择、各声道音频数据长度等重要信息。主信息即音频信号的内容,包括尺度因子编码和 576 条谱线的 Huffman 编码两部分,长度由传输率决定。

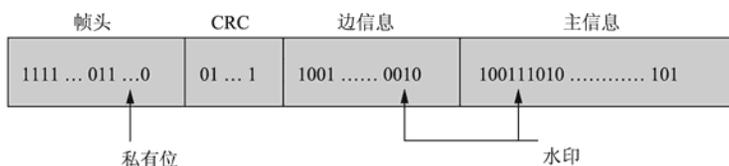


图 1 MP3 码流的帧结构

3 MP3 码流无损音频水印

3.1 水印的嵌入和提取

欲嵌入到数字音频中的水印可以是任何形式的数据,如 ASCII 字符串或图标,我们先将它们转换为二进制序列。为避免未被授权者获取水印信息,根据密钥指定主信息和边信息中的 n 比特代表水印数据,这些比特的位置可以作伪随机轮换。依次扫描 MP3 各帧和二进制水印序列,若当前的 n

比特水印数据与当前帧被指定的 n 比特主信息/边信息数据吻合,则在该帧的私有位写入 1,否则保持为 0。由于音频数据本身并未发生变化,水印的嵌入不会引起任何失真,因此这样“嵌入”的是一种无损水印。

音频和水印序列通常服从均匀分布,则 n 比特音频数据与 n 比特水印数据吻合的概率为 $1/2^n$,即每 $1/2^n$ 帧可承载 n 比特水印数据,因此 MP3 序列的最大水印嵌入率为:

$$c = n / 2^n. \quad (\text{比特/帧}) \quad (1)$$

显然, 当 $n=1$ 和 2 时水印容量达到最大值(每帧 0.5 比特)。图 1 给出了 $n=2$ 的一个例子, 其中私有位为 1, 说明该帧包含两比特水印数据。

提取水印的方法很简单。逐帧进行检测, 只要私有位为 1, 就取出由密钥指定的 n 比特数据; 如果私有位为 0, 则跳过该帧。将取出的二进制数拼接, 解码还原为 ASCII 字符串或图标, 完成水印的提取。

3.2 无损音频水印的流媒体应用

本文提出的水印嵌入方法是对 MP3 码流各帧逐一进行判别, 依次找到与待嵌入水印匹配的数据位, 利用私有位进行标记将它们确认为水印。这种方法计算量很小, 且具有顺序处理的特点, 是一种适合流媒体应用的方案, 接收端可在播放流媒体同时将水印提取出来。

流媒体是采用流式传输的方式在互联网或内网播放的媒体格式, 流媒体系统包含服务器和客户端。音频文件经压缩后被分成一系列数据包, 由服务器向客户端以单播(unicast)或多播(multicast)方式连续、实时地发送; 客户端收到足够的初始数据后即可边接收数据边进行播放, 无需等待下载全部文件, 也不必在计算机内保存副本。

在服务器端用本文方法嵌入水印, 生成含有水印的 MP3 文件, 在文件目录中显示该文件名, 并将 MP3 文件转换成若干个压缩包, 等待请求指令。客户端通过 IP 地址与服务器连接, 获取服务器端的文件目录, 点击文件名即可启动文件传输, 在客户端实时播放 MP3 音乐并提取水印。

用以上方法嵌入的水印并不损伤音质, 具有完全的透明性, 同时由于水印所在位置由密钥控制, 故有良好的安全性。但在流媒体传输中由于信道流量起伏或存在干扰等情况, 部分水印数据可能丢失, 导致最终无法解码, 使得鲁棒性不能满足要求, 故必须采取措施防止误码无限传递, 尽快恢复水印的正确解码。

为使水印具有抗干扰性, 可在 MP3 码流中添加同步标志, 选取某个特定的随机序列用于同步。例如当水印为字符串时, 以两个连续的“~”字符为同步标志, 每隔若干字符插入一个同步标志。在发

生数据丢失时只要找到下一个字符串“~”就可以恢复正常解码。当水印信息为图标时也可选用特定的伪随机序列为同步标志。另外还可采用奇偶校验、循环编码、交织编码等技术引入纠错码, 以降低水印信息的误码率。

4 实验结果与讨论

实验中以古典、乡村、爵士三类音乐为对象, 取 $n=2$, 在 MP3 文件中嵌入和提取水印, 在流媒体实验平台上发送含水印的 MP3 音乐, 客户端在播放音乐的同时提取并显示水印。以下给出实验结果。

4.1 不可感知性

本文提出的水印算法只是改变帧头中的私有位, 没有改变 MP3 音乐每帧的主信息, 具有完全的不可感知性。为了进行验证, 用 PEAQ 算法^[9]对多种类型音乐在水印嵌入前后的音质进行比较, 按五个等级客观评价音频质量。根据该算法计算结果, 得到测试信号和参考信号的客观差异等级 (Objective Difference Grade, ODG), 范围为 $[-4, 0]$, 分数越接近 0 表示音频可察觉的损伤越小。当 ODG 大于 0 时, 说明测试信号与参考信号不能区分。

对采样率为 44.1 kHz, 16 bit 编码的 PCM 音频信号进行测试, 嵌入时先将 PCM 格式的原始音频信号压缩为 MP3, 采用本文的方法嵌入水印。再将含有水印的 MP3 文件解压为 PCM 格式, 以原始的 PCM 音频信号为参考信号, 嵌入后解压的 PCM 信号为测试信号, 用 PEAQ 算法计算水印嵌入前后音频的 ODG 差值, 计算得到的 ODG 全部大于 0, 说明不能区分水印嵌入前后的音频信号, 音频质量没有改变。

4.2 水印容量

当 $n=2$ 时, 水印嵌入量的理论值为 0.5 比特/帧 (包括同步标识)。用不同时间的 MP3 音乐为实验样本, 测试水印的最大嵌入量, 结果如表 1 所示, 与理论分析结果一致。

实际应用中, 为了提高水印可靠性和稳定性, 往往要牺牲水印的嵌入量。比如说添加同步标志控制误码率, 嵌入量则会低于 0.5 比特/帧。相比于差错控制, 实际应用对嵌入量的要求比较低。

表1 水印嵌入量

音乐格式	序号	时长(s)	帧数	嵌入水印量(bits)	平均每帧嵌入量(比特/帧)
古典音乐	1	47	1795	935	0.521
	2	110	4216	2123	0.503
	3	132	5072	2455	0.484
乡村音乐	1	70	2699	1347	0.499
	2	90	3472	1761	0.507
	3	125	4760	2332	0.490
爵士音乐	1	83	3204	1616	0.504
	2	101	3849	1902	0.494
	3	136	5214	2565	0.492
每帧嵌入量的均值					0.499

4.3 流媒体应用实验结果

图2是自行开发的流媒体实验平台用户界面。在服务器端向MP3数据流中嵌入字符或图标,客户端都能正确播放并提取水印。

进行了部分数据丢失影响水印提取的实验。采用添加同步标志的方式,分别以字符串和图标为水印信息。如原始水印字符为“abcdefghijklmn”,每隔4个字符添加同步标志“~”。在客户端,随机选择一个位置,丢弃少量接收到的压缩包,模拟含水印的MP3流在传输中数据丢失。提取嵌入数据后

去除同步标志,解码还原得到“□□□□efghijklmn”。

以二值图标为水印的实验结果见图3,其中图3(a)是嵌入的图标,图3(b)是无干扰情况下正确提取的水印。对传输的流进行干扰,图3(c)是未添加同步码时提取的水印,可见错误一直扩散到整个水印。由密钥控制生成10比特伪随机序列作为同步标志,每3列图标数据添加一个同步标志,提取结果见图3(d),检测到同步信息后恢复了正常的水印提取,见解码的图标右侧“华夏”二字。



(a) 服务器



(b) 客户端

图2 流媒体音频水印实验平台用户界面

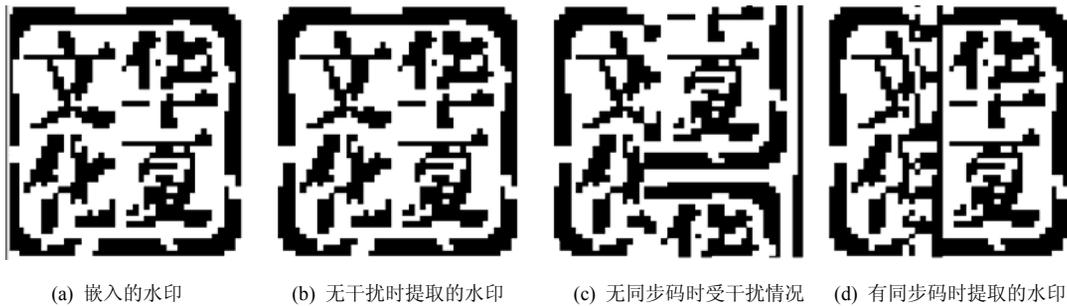


图 3 有干扰情况下二值图标水印流媒体的传输实验

5 结束语

本文根据 MP3 文件的结构特点提出了一种可用于 MP3 版权保护的水印算法,利用私有位作为水印标记,并不修改 MP3 音乐的主要数据。相对于其他基于 MP3 码流的音频水印算法,该方法具有优良的不可感知性和运算简单的特点,适合于在流媒体发送和接收过程中实时完成嵌入和提取操作。水印的安全性由密钥保证,未经授权者无法提取。针对流媒体传输中部分水印信息可能丢失的情况,添加同步标志和纠错编码,并可采用重复嵌入等措施减小网络异常对水印提取的影响。

参 考 文 献

- [1] PETITCOLAS F[EB/OL]. [1999]. <http://www.cl.cam.ac.uk/~fapp2/steganography/mp3stego>.
- [2] WANG C T, CHEN T S, CHAO W H. A New Audio Watermarking Based on Modified Discrete Cosine Transform of MPEG/Audio Layer[C]. IEEE International Conference on Networking, Sensing & Control, 2004: 984-989.
- [3] CHEN B, ZHAO J Y. An Adaptive and Audio Watermarking Algorithm for MP3 Compressed Audio Signal[C]. International Instrumentation and Measurement Technology Conference, 2008: 1057-1060.
- [4] QIAO L T, NAHRSTEDTY N. Non-invertible Watermarking Methods for MPEG Encode Audio[C]. SPIE Proceeding on Security and Watermarking of Multimedia Contents, 1999: 194-202.
- [5] 高海英. 基于 Huffman 编码的 MP3 隐写算法[J]. 中山大学学报: 自然科学版, 2007, 46(4): 32-35.
- [6] International Organization for Standardization. ISO/IEC 11172-3-1993, Information Technology Coding of Moving Pictures and Associated Audio for Digital Storage Media at up to About 1.5 Mb/s, Part 3: Audio[S], 1993.
- [7] International Telecommunication Union. ITU-R BS.1387-1-1998, Method for Objective Measurements of Perceived Audio Quality[S], 1998.
- [8] 刘伟, 王朔中, 张新鹏. 一种基于部分 MP3 编码原理的音频水印[J]. 中山大学学报: 自然科学版, 2004, 43(2): 26-29.
- [9] 晁婷婷, 王新房, 蒋存云. 基于压缩域的 MP3 音频数字水印算法[J]. 计算机工程, 2011, 37(10): 204-206.