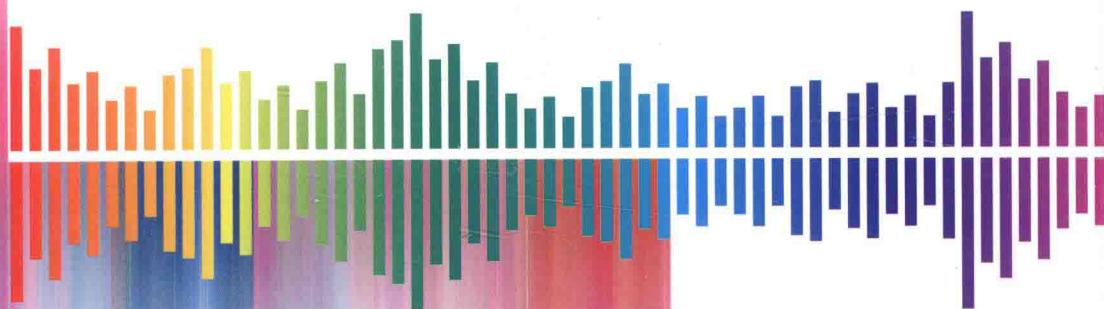


数字音频 水印技术及应用



刘海燕 著

清华大学出版社

数字音频 水印技术及应用

刘海燕 著

清华大学出版社
北京

内 容 简 介

本书主要对音频信号的常用格式、听觉系统的感知特点,数字音频水印的概念及原理、攻击、嵌入算法、检测算法、水印加密算法、纠错编码的使用以及音频水印的应用等方面进行了详细的阐述。

本书可作为信息与通信工程、计算机应用、信号与信息处理等专业的高年级本科生和研究生的教材或参考书,也可以作为信息安全与保密通信、多媒体数字产品保护等领域的相关人员的参考书。

版权所有,侵权必究。侵权举报电话:010-62782989 13701121933

图书在版编目(CIP)数据

数字音频水印技术及应用/刘海燕著. —北京:清华大学出版社,2016

ISBN 978-7-302-43346-0

I. ①数… II. ①刘… III. ①数字音频技术—密码术—应用—信息安全 IV. ①TP309

中国版本图书馆 CIP 数据核字(2016)第 062862 号

责任编辑:柳 萍 赵从棉

封面设计:常雪影

责任校对:赵丽敏

责任印制:刘海龙

出版发行:清华大学出版社

网 址: <http://www.tup.com.cn>, <http://www.wqbook.com>

地 址:北京清华大学学研大厦 A 座 邮 编:100084

社 总 机:010-62770175 邮 购:010-62786544

投稿与读者服务:010-62776969, c-service@tup.tsinghua.edu.cn

质量反馈:010-62772015, zhiliang@tup.tsinghua.edu.cn

印 装 者:三河市金元印装有限公司

经 销:全国新华书店

开 本:165mm×240mm 印 张:7.75 字 数:144千字

版 次:2016年5月第1版

印 次:2016年5月第1次印刷

定 价:59.00元

产品编号:066905-01

随着数字技术的进步,数字产品(如电子出版物、音频、视频、动画、图像产品等)得到了极大的发展。同时,随着计算机网络技术的飞速发展,数字产品容易传播和发布,给人们的生活和工作带来了极大的便利。然而,随之而来的副作用(如在没有被授权的情况下随意篡改、复制和传播有版权的内容),极大地损害了产品所有者的知识产权及商家的利益。如何防止数字产品被侵权、盗版和恶意篡改,已经成为世界各国急需解决的热门课题之一。

现有的版权保护系统多采用加密和数字签名等技术解决信息安全和版权保护问题。基于私用或公共密钥的加密技术可以用来控制数据访问,它将明文消息变换成旁人无法理解的密文消息。加密后的产品是可以访问的,但只有那些具有正确密钥的人才能解密。除此之外,还可以通过设置密码,使得数据在传输时变得不可读,从而可以为从发送到接收过程中的数据提供有效的保护。数字签名是一种由信息发送方以电子形式给消息签名的方法,其实质是一段特殊的字符串,该数字串是除签名者外其他任何人都无法伪造的,具有对签名真实性证明的作用。它通过使用私用密钥,对每个消息进行签名,而公共的检测算法用来检查消息的内容是否符合相应的签名。但是,使用这些技术在版权保护方面存在很大的局限性,主要表现在:①多媒体数据流结构复杂、数据量庞大,对它们进行加密非常困难,而且不利于多媒体数据的传播;②加密数据的不可读性,容易引起监控者的特别注意,并以此为依据展开对密文的破译和对发送者、接收者的攻击;③随着计算机硬件技术的迅速发展,具有并行计算能力的破解技术日益成熟,仅仅通过增加密钥长度来达到增强安全的方法越来越不适应计算和传输的要求;④由于互联网的高度普及,使得加密数据一旦被破解,便可以在广泛的范围内自由传播,一切的密码保护措施都失去意义。因此,需要寻求一种不同于传统技术的更加有效的手段来保障数字信息的安全传

输和保护数字产品的版权。

数字水印技术作为多媒体信息安全研究领域中的重要分支,为解决数字产品版权和认证内容完整性等问题提供了一种有效的方法。数字水印技术一方面弥补了密码技术的缺陷,因为它可以为解密后的数据提供进一步的保护;另一方面,数字水印技术也弥补了数字签名技术的缺陷,这是因为它可以在原始数据中一次性嵌入大量的秘密信息。由于这些特点,数字水印技术已成为多媒体信息安全研究领域发展的热点技术之一,受到国际学术界和企业界的高度关注。

从应用范围上可把水印分为图像水印、视频水印和音频水印等。当前,对图像水印和视频水印技术的研究较多,而对于音频水印的研究较少,这主要是由于:①人类听觉系统比较灵敏,这对数据隐藏的实现造成了很大的困难;②功能强大的音频处理软件在互联网上有很多并且能够免费下载,它们可以极大地改变语音信号的结构并且不会造成音质的损失;③数字音频水印对信号的同步有比较高的要求。然而,随着MP3、MPEG、AC-3等新一代压缩标准的广泛应用,对数字音频作品的知识产权保护显得越来越重要。因此,对数字音频水印技术的研究也变得越来越重要。

本书共8章。第1章是音频基础,简单介绍音频信号的数字化、常用的数字音频格式以及听觉系统感知特点。第2章简要介绍数字水印的概念、原理、关键技术及应用与分类,同时介绍国内外研究现状。第3章介绍常用的图像置乱技术,主要包括幻方变换、仿射变换、Hilbert曲线、正交拉丁方变换、Arnold变换、Fibonacci-Q变换、混沌置乱等。第4、5章重点阐述抗同步攻击的鲁棒性水印的算法,包括数字水印的生成技术、预处理技术、嵌入技术、提取和检测技术。第6章主要研究脆弱性音频水印算法,包括基于MFCC特征的多功能音频水印算法和基于音频内容特征的脆弱性音频水印算法。第7章介绍音频水印的检测算法。第8章主要介绍数字音频水印技术的应用。

本书得到了北京市教育委员会2015年度科技计划项目(批准号:KM201310857001)的支持。

由于作者水平有限,书中难免存在各种疏漏和不当之处,敬请读者批评指正。

编者
2016年3月

第 1 章 音频基础	1
1.1 音频信号的数字化	1
1.2 常用的数字音频格式	1
1.3 听觉系统感知特点	3
1.3.1 掩蔽效应	3
1.3.2 MPEG 心理声学模型 I	4
第 2 章 数字音频水印技术基础	7
2.1 数字水印的定义与性质	7
2.2 数字水印的分类	8
2.3 数字水印生成技术	9
2.4 数字水印系统的基本框架	10
2.5 基于通信系统的数字水印	11
2.6 数字音频水印嵌入算法	13
2.6.1 时间域水印算法	13
2.6.2 频率域水印算法	16
2.6.3 时间域与频率域算法比较	24
2.6.4 压缩域音频水印算法	25
2.7 音频水印的攻击	26
2.7.1 普通攻击类型	26
2.7.2 同步攻击类型	26
2.8 音频水印算法的评价标准	27
2.8.1 不可感知性评价标准	27
2.8.2 水印容量评价标准	29
2.8.3 鲁棒性评价标准	29
2.9 数字音频水印的应用与发展	30

第 3 章 图像置乱技术	33
3.1 幻方变换	33
3.2 仿射变换	34
3.3 Hilbert 曲线	35
3.4 正交拉丁方变换	35
3.5 Fibonacci 变换	36
3.6 Arnold 变换	37
3.7 混沌置乱	38
3.8 Arnold-Logistic 算法设计	40
第 4 章 抗同步攻击的音频水印算法	42
4.1 抗同步攻击算法的分类	42
4.2 同步信号的确定方法	43
4.2.1 采用 m 序列作为同步信号	43
4.2.2 采用巴克码作为同步信号	44
4.3 恒定音频水印算法	44
4.3.1 算法原理	44
4.3.2 嵌入与提取算法	45
4.3.3 频带选择	47
4.3.4 小波基函数选取	48
4.4 实验结果及分析	50
4.4.1 不可感知性测试	53
4.4.2 抗攻击能力测试	53
4.5 本章小结	55
第 5 章 纠错编码在音频水印算法中的有效性研究	57
5.1 纠错编码策略	58
5.1.1 重复码	58
5.1.2 BCH 码	58
5.1.3 级联码	59
5.1.4 卷积码	60
5.2 ECC 与嵌入强度的关系	60
5.3 码率对纠错能力的影响	61
5.4 码长与嵌入强度的关系	62
5.5 纠错编码在水印系统中的性能研究	63
5.6 纠错编码的使用	66

第 6 章 脆弱性水印算法	68
6.1 脆弱性水印概述	68
6.1.1 脆弱性水印的基本特征	68
6.1.2 鲁棒性水印与脆弱性水印的区别	69
6.1.3 脆弱性水印算法的一般框架	70
6.1.4 脆弱性音频水印算法分类	70
6.2 多功能音频水印算法	72
6.2.1 基于 MFCC 特征的相似帧计算	72
6.2.2 算法设计	73
6.2.3 篡改定位与恢复	75
6.2.4 实验结果及分析	76
6.3 基于音频内容的 DWT 域脆弱音频水印算法	83
6.3.1 算法设计	84
6.3.2 篡改定位与恢复	85
6.3.3 算法分析、参数选取与讨论	85
6.3.4 实验结果及分析	88
第 7 章 音频水印检测算法研究	92
7.1 相关检测阈值确定算法	92
7.1.1 相关检测阈值确定	92
7.1.2 信噪比与相关检测的可靠性	95
7.1.3 实验及结果分析	95
7.2 基于 DWT 域乘嵌入音频水印的优化检测算法	96
7.2.1 小波变换系数的统计分布	97
7.2.2 乘嵌入音频水印的优化检测算法	98
7.2.3 实验结果及分析	101
第 8 章 数字音频水印的应用	103
8.1 数字音频水印在广播系统中的应用	103
8.1.1 音频水印在广播节目文件传输系统中的应用	103
8.1.2 音频水印在广播节目监测系统中的应用	104
8.2 数字音频水印在版权保护方面的应用	105
8.3 数字音频水印在保密通信方面的应用	105
参考文献	107

第 1 章 音频基础

1.1 音频信号的数字化

音频信号(audio signal)是带有语音、音乐和音效的有规律的声波的频率、幅度变化信息载体。音频信号的数字化一般经过采样和量化的方法来实现,量化精度和采样频率是音频信号数字化的两个重要参数^[1]。

量化是对模拟音频信号的幅度轴进行数字化,决定了模拟信号数字化的动态范围。通常高质量的数字音频信号最流行的描述样本格式是 16b 线性量化,如 Windows 可视音频格式(Windows audio visual, WAV)和音频交换文件格式(audio interchange file format, AIFF)。另一种较低质量的音频信号格式是 8bit 律或 a 律的非线性量化方式。这些量化方法使信号产生了一些畸变,在 8bit 律中显得更为明显。量化位越高,信号的动态范围越大,数字化后的音频信号就越可能接近原始信号,但所需的存储空间也越大。

采样率是指声音信号从模拟信号转换成数字信号过程中单位时间内采样的次数。音频信号的采样率一般为 8、9.6、10、12、16、22.05、44.1 和 48kHz。采样频率越高,声音的还原就越真实、越自然。在当今的主流采集卡上,采样频率一般共分为 22.05、44.1、48kHz 三个等级,22.05kHz 只能达到 FM 广播的声音品质,44.1kHz 则是理论上的 CD 音质界限,48kHz 则更加精确一些。对于高于 48kHz 的采样频率人耳已无法辨别,所以在计算机上没有多少使用价值。采样率影响数据隐藏量,因为它给出了可用频谱的上限,假设信号的采样频率为 8kHz,则引入的修改分量的频率不会超过 4kHz。

1.2 常用的数字音频格式

数字音频的常见格式有很多种,每种格式都有自己的优点、缺点及使用范围。

1. WAV 格式

WAV 格式是微软公司开发的一种声音文件格式,也称为波形声音文件,是

最早的数字用音频格式,也是最常用的音频文件格式。Windows 平台及其应用程序广泛使用 WAV 音频文件类型,本书中主要采用的也是 WAV 格式的音频文件。WAV 格式支持许多压缩方法,多种音频位数、采样频率和声道;采用 44.1kHz 的采样频率,16 位量化位数,因此 WAV 的音质与 CD 相差无几,但 WAV 格式对存储空间需求太大,不便于交流和传播。

2. CD 格式

CD 存储采用了音轨的形式,是 CD Audio 的缩写,其扩展名为 CDA。采用 CD 格式的音频文件声音质量基本上与原声相同,声音质量纯正,是近似无损的一种音频格式,唱片主要采用这种格式。标准 CD 格式的采样频率为 44.1kHz,16 位量化位数。但采用该种格式的文件太长,所以在存储文件时需要的数据量较大,导致不能进行编辑。

3. MP3 格式

MP3 全称为 Moving Picture Experts Group Audio Layer 3,是 ISO 标准 MPEG-1 和 MPEG-2 的第三层,采样率为 16~48kHz。MP3 能够以高音质、低采样率对数字音频文件进行压缩,即文件数据量较大的音频文件能够在音质丢失很小的情况下把文件压缩得更小。

4. RA 格式

RA 是 Real Audio 的缩写。由于其强大的压缩比,所以能有效地解决带宽资源受限的问题。RA 格式首要考虑的就是容错性与压缩比的问题,其次是音频质量的问题,使用该格式的音频文件的失真率很小。

5. WMA 格式

WMA 的全名是 Windows Media Audio。WMA 一般情况下与 MP3 的音频格式处于同样的地位,但 WMA 比 MP3 格式更有优势,特别是在压缩与解码技术方面,也能得到更好的音频质量,同时在在线播放和网络传输方面也有很大的应用价值。

6. MIDI 格式

MIDI 是由 Musical Instrument Digital Interface 缩写而成,扩展名为“MID”。MIDI 传输的是音符和控制参数等指令,描述的为音乐语言。最初的用意是为了解决电子乐器之间和电子乐器与计算机之间交流的问题。MID 文件本身并不包含波形信息的数据,主要依靠硬件生成,只适合记录音频信息,而不适合处理音频文件。

7. MOD/MTM/FAR/KAR/IT 格式

MOD/MTM/FAR/KAR/IT 是一种模块的格式文件,通过不同的编码就

会有多种不一样的格式。这样既可以保存数字声音信号的采样数据,又可以保存乐曲演奏的命令,它们同时具有数字音频与 MIDI 的共同特性,因此对音频硬件的依赖性较小。

1.3 听觉系统感知特点

数字水印主要利用人类感知系统(视觉和听觉)的冗余实现水印不可感知性,因此数字音频水印技术是依赖人类听觉系统(human auditory system, HAS)的某些特性,来满足嵌入水印的不可感知性的要求。

1.3.1 掩蔽效应

人类听觉系统的一个重要特性是听觉的掩蔽效应。一个较弱的声音(被掩蔽音)的听觉感受被另一个较强的声音(掩蔽音)影响的现象称为人耳的掩蔽效应。掩蔽效应与两个声音的声强、频率、相对方向及延续时间有关。利用掩蔽效应可以用有用的声音信号去除或掩蔽无用的声音信号。图 1-1 所示为人耳听觉掩蔽曲线,由于 A 频带音频信号的能量大于相邻频带的音频信号,因此在掩蔽曲线之下的其他频带信号都被掩蔽起来,即使其能量已超越人耳绝对阈值曲线仍然无法被人耳察觉^[2]。掩蔽效应可分为频域掩蔽(simultaneous masking)和时域掩蔽(temporal masking)两种。

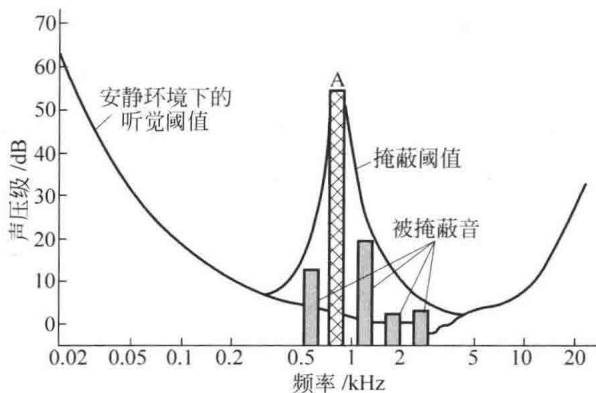


图 1-1 听觉掩蔽曲线

1. 频域掩蔽

所谓频域掩蔽是指掩蔽声与被掩蔽声同时作用时发生掩蔽效应,又称同时掩蔽^[3]。如果一个较弱的声音落在一个较强声音的临界频带中,那么强信号(掩蔽声)就会将弱信号(被掩蔽声)掩蔽掉。通常来说,频域中的一个强音会掩蔽与之同时发声的附近的弱音,弱音离强音越近,越容易被掩蔽;反之,离强音

较远的弱音不容易被掩蔽。

2. 时域掩蔽

所谓时域掩蔽是指掩蔽效应发生在掩蔽声与被掩蔽声不同时出现时,又称异时掩蔽。异时掩蔽又分为超前掩蔽和滞后掩蔽。若掩蔽声音出现之前的一段时间内发生掩蔽效应,则称为前掩蔽(pre-masking);否则称为后掩蔽(post-masking)。产生时域掩蔽的主要原因是人的大脑处理信息需要花费一定的时间,异时掩蔽也随着时间的推移很快会衰减,是一种弱掩蔽效应。一般情况下,超前掩蔽只有5~20ms,而滞后掩蔽却可以持续50~100ms。

图1-2给出了频域掩蔽和时域掩蔽现象。从图中可知,频域掩蔽在掩蔽者持续的时间内一直有效,它是一种较强的掩蔽效应,而时域掩蔽随着时间的推移很快衰减。

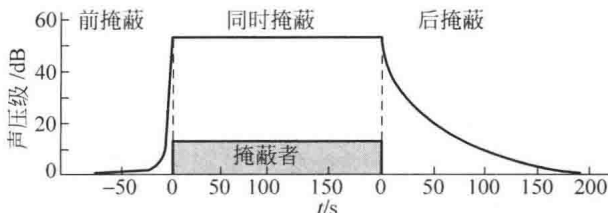


图1-2 三种掩蔽现象的强度以及持续时间

1.3.2 MPEG 心理声学模型 I

因为掩蔽阈值与信号声压级有关,使用心理声学模型对水印整形的过程与感知音频编码的量化噪声控制过程相似,所以在采用心理声学模型的水印系统设计中,经常采用已经比较成熟的感知音频编码的心理声学模型。音频水印中常常使用的感知模型主要有 MPEG-1 心理声学模型 I 和心理声学模型 II,其中心理声学模型 II 采用了感知熵的概念。使用心理声学模型,最终获得一个全局掩蔽阈值。当量化噪声的声压级等于这个掩蔽阈值时,听觉系统刚好能够感知到噪声,所以这个掩蔽阈值又称为临界感知失真(just noticeable distortion, JND)。音频水印使用临界感知失真来限制嵌入的水印最大能量,当水印的能量被限制在临界感知失真阈值以下时,就能保证水印的不可感知性^[4]。本节详细介绍 MPEG-1 中的心理声学模型的实现过程^[5]。

1. 计算原始信号的功率谱

用 FFT 计算音频信号 $s(i)$ 的信号功率谱密度 $X(k)$ 。 $X(k)$ 的计算公式为

$$X(k) = 10 \lg \left| \frac{1}{N} \sum_{i=0}^{N-1} h(i) s(i) \exp(-j2\pi ki/N) \right|^2, \quad k = 0, 1, \dots, N/2 \quad (1-1)$$

式中, $h(i)$ 为 Hanning 窗函数, 用来减少边界效应; N 为每帧样本数。

2. 音调成分和非音调成分的认识

掩蔽者的音调性会影响掩蔽阈值, 因此要根据 $X(k)$ 识别出类音调信号和类窄带噪声信号。一个音调分量是满足如下条件的局部最大值, 其余视为噪声分量。

$$\begin{aligned} X(k) &> X(k+1) \quad \text{且} \quad X(k) \geq X(k-1) \\ X(k) - X(k+m) &\geq 7\text{dB} \\ m \in [-2, +2], & \quad 2 \leq k \leq 63 \\ m \in [-3, -2, +2, +3], & \quad 63 \leq k \leq 127 \\ m \in [-6, \dots, -2, +2, \dots, +6], & \quad 127 \leq k \leq 250 \end{aligned}$$

3. 计算单个掩蔽者的掩蔽阈值

掩蔽者的掩蔽阈值取决于掩蔽者的声压级、自身掩蔽级和掩蔽函数。类音调信号和类窄带噪声信号的自身掩蔽级是不同的, 分别是 $Y_1(z_i)$ 和 $Y_2(z_i)$ (单位: dB):

$$Y_1(z_i) = -1.025 - 0.275z_i - 4.5 \quad (1-2)$$

$$Y_2(z_i) = -1.025 - 0.175z_i - 0.5 \quad (1-3)$$

其中 z_i 是临界频带率。类音调信号和类窄带噪声信号的掩蔽函数 $Y_f(z_i, z_j)$ 相同:

$$Y_f(z_i, z_j) = \begin{cases} 17\Delta z - 0.4X(z_i) + 11, & -3 \leq \Delta z < -1 \\ [0.4X(z_i) + 6]\Delta z, & -1 \leq \Delta z < 0 \\ -17\Delta z, & 0 \leq \Delta z < 1 \\ [0.15X(z_i) - 17]\Delta z - 0.15X(z_i), & 1 \leq \Delta z < 8 \end{cases} \quad (1-4)$$

其中 $X(z_i)$ 是临界频带率为 z_i 的掩蔽者的声压级; $\Delta z = z_j - z_i$, 是被掩蔽者与掩蔽者的距离; $T_1(z_i, z_j)$ 或 $T_2(z_i, z_j)$ 是临界频带率为 z_i 的掩蔽者在临界频带率 z_j 处产生的掩蔽阈值。

类音调掩蔽者为

$$T_1(z_i, z_j) = X(z_i) + Y_1(z_i) + Y_f(z_i, z_j) \quad (1-5)$$

类窄带噪声掩蔽者为

$$T_2(z_i, z_j) = X(z_i) + Y_2(z_i) + Y_f(z_i, z_j) \quad (1-6)$$

式中, Y_1 、 Y_2 分别为类音调和类窄带噪声的掩蔽系数; Y_f 为掩蔽函数。由于临界频带率之差 Δz 增大, 掩蔽作用降低, 所以当 $\Delta z < -3\text{Bark}$ ^① 或 $\Delta z \geq 8\text{Bark}$

① Bark(巴克)是临界频带的单位, 1Bark=一个临界频带宽度。频率小于 500Hz 时, 1Bark 约等于 $\text{freq}/100$; 频率大于 500Hz 时, 1Bark 约为某个纯音中心频率的 20%。

时,不考虑掩蔽,这时设 $T_1(z_i, z_j)$ 和 $T_2(z_i, z_j)$ 为 $-\infty$ 。

4. 计算全局掩蔽阈值

掩蔽是可叠加的,因而 z_j 处的全局掩蔽阈值 $T_g(z_j)$ 为该点的安静阈值 $T_0(z_j)$ 和所有有调、无调成分在该点产生的掩蔽阈值之和,即

$$T_g(z_j) = 10\lg(10^{0.1T_0(z_j)} + \sum_i 10^{0.1T_1(z_i, z_j)} + \sum_i 10^{0.1T_2(z_i, z_j)}) \quad (1-7)$$

5. 决定最低掩蔽阈值

在每个子带中,根据总体掩蔽阈值的情况,决定各个子带的最小掩蔽阈值:

$$T_{\min}(n) = \min[T_g(z_j)] \quad (1-8)$$

音频掩蔽现象及 MPEG 心理声学模型对数字音频水印技术是十分重要的。大多数音频水印技术为了保证嵌入数据的不可感知性,都直接或间接地利用了音频掩蔽现象。嵌入的水印信息要适应并高度依赖于宿主音频信号,其时域和频域分布由宿主音频信号的时域和频域掩蔽特性决定。水印信号强度随宿主音频信号而变化,比如在静音区的水印信号强度应较小,这保证了嵌入水印信息在具有最大能量的同时不可感知,能量的最大化增强了水印抵抗攻击的能力。

第 2 章 数字音频水印技术基础

数字音频水印技术能够有效地保护音频产品的版权以及认证其内容完整性,已经成为学术界的一个研究热点^[6,7]。本章从数字音频水印的定义、性质、框架、分类、嵌入算法、水印攻击、评价标准、应用与发展 9 个方面介绍水印技术。

2.1 数字水印的定义与性质

所谓数字水印,就是在数字媒体信息(如图像、声音、视频等)中添加某些数字信息,保护数字媒体的版权,证明产品的真实可靠性,跟踪盗版或者提供产品的附加信息。

一个成功的数字音频水印系统需要具备以下一些性质^[8,9]:

(1) 水印信息必须隐藏在载体音频数据中,而不能存储于文件头或单独的文件中,否则可以轻易地被去除或者改变;

(2) 水印信息不应对原始音频信号的声音质量产生可听的失真,即应具有不可感知性;

(3) 水印信息在统计上必须具有不可检测性,防止攻击者逆向嵌入过程来除去水印;

(4) 水印算法必须具有一定的鲁棒性,能抵抗添加噪声、压缩、滤波、重采样、重量化、A/D 和 D/A 转换等一般信号处理,未经授权的人很难或者根本无法去掉水印;

(5) 水印应易于提取,嵌入和检测的计算量要低,以方便集成到一般电子产品中;

(6) 水印算法必须具备某种同步机制,以对抗时间域上的同步攻击;

(7) 盲检测,原则上水印的检测不应需要原始音频,即实现盲检测,以便工程上的应用;

(8) 水印算法应该公开,安全性最好依赖于密钥而不是算法的秘密性。

2.2 数字水印的分类

数字水印有很多种分类方法,分类出发点不同导致了分类的不同,它们之间既有联系又有区别。常用的分类方法主要包括以下几种^[10,11]。

1. 按应用要求划分

根据水印的应用要求,可将其划分为鲁棒性水印和脆弱性水印。鲁棒性水印主要用于数字产品的版权保护,它必须保证对原始版权的准确无误的标识,保证在宿主信号可能发生的各种失真变换下,以及各种恶意攻击下都具备很高的抵抗能力。脆弱性水印主要用于数据的真伪鉴别和完整性认证。该水印技术在原始真实信号中嵌入某种标记信息,通过鉴别这些标记信息的改动,达到对原始数据完整性检验的目的。它应对各种攻击敏感,但是脆弱型水印的脆弱性并不是绝对的。对宿主信号的某些必要性操作,如滤波或压缩,脆弱型水印也应体现出一定的鲁棒性,从而将这些不影响主信号最终可信度的操作与那些蓄意破坏操作区分开来。

2. 按水印所附载的媒体划分

按照数字水印所附载的载体数据,分成图像水印、音频水印、视频水印、文本水印以及用于三维网格模型的网格水印等。随着数字技术的发展,会有更多种类的数字媒体出现,同时也会产生相应的水印技术。

3. 按检测过程划分

根据水印检测过程是否需要载体,分成非盲水印和盲水印。非盲水印在检测过程中需要原始数据,而盲水印在检测过程中不需要原始数据。一般来说,非盲水印的鲁棒性比较强,但其应用受到存储成本的限制,也不利于工程上的应用。目前学术界研究的数字水印检测算法大多数是盲水印检测算法。

4. 按水印位置划分

根据数字水印的隐藏位置,分成空间域水印和变换域水印。空间域数字水印是直接信号空间上叠加水印信息,具有较大的信息嵌入量,但其鲁棒性较差。变换域数字水印则分别是在离散余弦变换(discrete cosine transform, DCT)、离散小波变换(discrete wavelet transform, DWT)、离散傅里叶变换(discrete Fourier transform, DFT)等频域隐藏水印。变换域数字水印能较好地利用人类视觉、听觉系统的特性,具有较强的鲁棒性。

5. 按用途划分

不同的应用需求造就了不同的水印技术。按水印的用途,可以将数字水印划分为证件防伪水印、版权标识水印、篡改提示水印等。

6. 按内容划分

水印按内容分成有意义水印和无意义水印。有意义水印是指水印本身也是某个数字图像(如商标图像)或数字音频片段的编码;无意义水印则只对应于一个序列号。有意义水印的优势在于,如果由于受到攻击或其他原因致使解码后的水印破损,人们仍然可以通过视觉观察确认是否有水印。但对于无意义水印来说,如果解码后的水印序列有若干码元错误,则只能通过统计决策来确定信号中是否含有水印。

上述提到的各种分类方法并不是孤立的,它们经常相互交叉^[12]。例如,一个鲁棒性水印算法可以采用盲检测也可以采用非盲检测;鲁棒性水印或者脆弱性水印也可以使用不同的载体^[13,14]。

2.3 数字水印生成技术

数字水印生成是数字水印处理过程的第一个关键步骤,构成水印的序列通常应该具有不可预测的随机性。由于人类视觉系统对亮度具有极高的敏感性,甚至超过对纹理所具有的敏感性,故水印不应该是高亮度的,亦不应该含有纹理,应该具有与噪声相同的特性。

通常意义上说,数字水印生成过程就是在密钥 K 的控制下,由原始身份信息、版权信息、认证信息、保密信息或其他有关信息 m 生成适合于嵌入载体 x 中的待嵌入水印信号 w 的过程,如图 2-1 所示。



图 2-1 数字水印生成算法框图

人们通常把利用原始载体信息 x 生成水印的方法称为自适应生成方法。这类方法的好处是,水印的鲁棒性和透明性能达到较好的折中,有时便于原始载体的自恢复。无原始载体信息 x 参与生成水印的方法称为非自适应生成方法。从各输入项的参与情况看,数字水印生成方式可分为四类:

(1) 原始水印信息 m 参与的自适应水印生成方式,水印生成过程可能需要密钥 K ,也可能不需要,故有 $w=G(m,x,K)$ 或 $w=G(m,x)$;

(2) 无原始水印信息参与的自适应水印生成方式,水印生成过程可能需要密钥 K ,也可能不需要,故有 $w=G(x,K)$ 或 $w=G(x)$;

(3) 原始水印信息参与的非自适应水印生成方式,水印生成过程可能需要密钥 K ,也可能不需要,故有 $w=G(m,K)$ 或 $w=G(m)$;