

摘 要

数字化和网络化的普及，使数字信号成为人们沟通交流的主要载体。为了进一步高效方便地利用数字信号给工作生活带来便利，随之而来出现了各种对数字多媒体信号进行编辑的工具。这些工具的确丰富了大家的日常生活，并给工作带来了极大的便利。然而，对于不法分子而言，这些工具却成为他们达到非法目的提供了便利条件。比如，音频创作者创作的音频内容，不法分子非法地复制传播来牟利；不法分子恶意地修改敏感音频的内容，达到不可告人的目的等等。

为了解决数字音频信号的版权保护问题，以及安全性问题，文本主要研究了基于保密特征的数字音频双水印技术。

(1) 给出了一种数字音频信号的保密特征，该特征的提取与密钥密切相关，对于非授权用户在没有密钥的前提下，很难提取正确的音频特征。

(2) 分析了保密特征的特性，给出了一种基于该特征的水印嵌入量化方法，并构造了一种基于该特征的数字音频双水印算法。首先将原始信号分帧、提取音频信号的保密特征。然后将由版权信息以及用于篡改定位的标识信息生成的水印，基于要嵌入的水印信息来量化保密特征，并进一步地完成水印的嵌入。本文将水印嵌入在保密特征中，提高了水印系统的安全性。同时算法能够进行版权保护以及对音频内容的取证，提高了水印算法的实用性。最后，实验结果证实了算法的有效性。

关键词：数字音频；数字水印；版权保护；数字取证

Abstract

The popularization of digitization and network has made digital signals as the main carrier of communication. In order to further efficiently and conveniently utilize digital signals to bring convenience to work and life, various tools for editing digital multimedia signals have emerged. These tools have indeed enriched our daily life and brought great convenience to work. However, for the criminals, these tools provide convenience for them to achieve their illegal goals. For example, audio content created by audio creators is illegally copied and distributed by criminals for profit; criminals maliciously modify the content of sensitive audio for ulterior motives, and so on.

In order to solve the copyright protection and security issues of digital audio signals, this paper mainly studies the digital audio double watermarking technology based on confidential features.

(1) We defined the confidential feature of digital audio signal. The extraction of this feature is closely related to the key. It is difficult for the unauthorized users to extract the correct audio feature without the key.

(2) The characteristics of the confidential features are analyzed. The quantization method for watermark embedding based on this feature is presented and a digital audio double watermarking algorithm based on this feature is constructed. Firstly, the security feature of audio signal is extracted. Then, the watermark generated by copyright information and identification information used for tampering location is quantified. Then, the watermark is further embedded. In this paper, the watermark is embedded in the confidential feature to improve the security of the watermarking system. At the same time, the algorithm can protect the copyright and has the ability of tamper localization, which improves the practicability of the watermarking algorithm. Finally, experimental results confirm the effectiveness of the algorithm.

Keywords: digital audio; digital watermark; copyright protection; digital forensics

目 录

第 1 章 引言.....	1
1.1 研究背景和意义.....	1
1.2 国内外研究现状.....	2
1.3 本文创新点.....	5
1.4 本文主要工作和论文组织结构	5
第 2 章 数字音频水印理论基础	6
2.1 数字音频水印的性质.....	6
2.2 数字音频水印的分类.....	7
2.3 数字音频水印的评价方法	8
2.4 数字音频水印常见的攻击方法	11
2.5 本章小结.....	13
第 3 章 基于保密特征的数字音频双水印算法	14
3.1 基础知识.....	14
3.2 水印嵌入和提取.....	17
3.3 实验分析.....	20
3.4 本章小结.....	25
第 4 章 总结与展望	26
4.1 论文总结.....	26
4.2 研究展望.....	26
参考文献.....	28

第1章 引言

1.1 研究背景和意义

数字信号处理技术以及高速网络的发展和普及,使数字信号作为信息交流和传播的载体变得更加便捷。人们可以便捷地从网络上传输分享相关数字作品。为了生活、工作所需,同时也出现了大量的数字媒体编辑工具,满足了工作需要,丰富了大家的生活。然而,引发了一系列的相关问题,比如版权被侵害、信号被修改等。随着大量数字媒体作品的出现,非法盗用、修改别人作品的现象也更加凸显。为了保护内容创作者的合法权益,鼓励创造者的创作积极性,需要一种技术手段来解决目前遇到的一系列问题。

提到数字内容的保护,大家首先想到与此相关的密码学技术。该技术的原理是,在作品完成发行之前,对作品加密成不为人知的密码序列,然后再公开发布。密码技术将已知的明文消息加密成不为人知的密文消息,一定程度上的确是保护了作品的内容。然而,在当今社会,它的局限性也更加明显。因为,一旦拥有密码的授权者将密文消息解密,得到明文消息之后,该消息就会在缺少监管的情况下肆意传播。这给不法分子创造了机会。因为不法分子在得到明文消息之后,存在对明文消息进行编辑、处理甚至是恶意修改的可能。为了弥补加密技术的弊端,学者们研究了一种可以保护明文消息传输的技术——数字水印技术^[1-4]。

多媒体数字水印技术可以简单概括为,将一些特殊用途的信息(版权标识、用于取证的消息)通过一定的方法嵌入到原始多媒体信号中。同时确保嵌入信息之后,不能影响多媒体内容的正常使用,嵌入的信息不能被人们主观地感受到,以及在正常使用的情况下不能丢失嵌入的信息。在需要的时候,从多媒体信号中提取需要的信息,来验证该多媒体内容的拥有者,或者对多媒体信号进行真伪鉴定^[5]。于是,数字水印技术弥补了加密技术的不足,是保护明文数据传输的一种有效手段。不仅具有较强的研究意义,同时也具有良好的社会效益^[6]。

根据嵌入载体的不同,简单地可将多媒体数字水印技术分为三类:数字图像水印技术^[7-10]、数字视频水印技术^[11-14]、数字音频水印技术^[15-20]。数字音频属于数字多媒体内容中的一种,和数字图像、数字视频相比,它可以看作是一维信号,具有数据量小,易传输等优点,成为最常用的传输、记载相关内容的数字载体之一。正

因为如此，音频也成为最容易被盗版侵害，被恶意攻击的信号之一。为了保护音频内容的版权和确保数字音频内容的真实性，本文主要研究以音频为载体的数字音频水印技术。

1.2 国内外研究现状

数字音频和数字图像相比，同属于数字多媒体信号。不同的是，数字图像水印技术起步较早，即便是在今天，也是相关领域的热门研究领域。然而，数字音频水印技术的研究相对晚一些。但是从提出到现在，热度不减。上世纪 90 年代，国际会议“国际信息隐藏学术研讨会”在英国召开，从那个时候开始，数字音频水印技术就成为了诸多学者感兴趣的研究课题。现在，已经有相当多的机构支持开展关于数字水印技术的相关研究，包括日本 Sony 公司、日本 NTT 信息与通信研究中心，IBM 公司等。

在国内，也有相当一部分专家学者对数字音频水印技术有浓厚的研究兴趣。随之很多高校和科研院所也加入到研究队伍当中来。我国 1999 年信息隐藏第一届研讨会召开。第二年 1 月，国家层面大力支持相关技术的研究，如“863”计划等。经过各研究机构和学者的努力，到目前位置，我们已经在音频水印相关技术领域积累的丰硕的研究成果。在某些方面，也已经达到了国际领先水平。

对于音频水印而言，根据应用目的不同，可以分为两类，一类是鲁棒音频水印技术，主要应用在对音频内容的版权保护方面；另一类是脆弱（半脆弱）音频水印技术，主要应用在验证音频内容的真伪方面。对于鲁棒音频水印算法而言，因为主要用来版权保护，所以，嵌入的水印主要是由版权标识等信息生成，而嵌入域往往选择鲁棒的嵌入域来进行嵌入，比如采用频域变换的方法来构造水印嵌入域。对于脆弱音频水印算法而言，由于算法的目的是用来验证该音频的真伪，所以嵌入的水印常常由用来取证的信息生成，而嵌入域则会选择那些在音频被攻击后容易被改变的部分作为嵌入域，比如将水印嵌入在时间域。

（1）基于时间域的音频水印算法

数字音频水印初期，文献[21,22]提出了基于最低有效位（Last Significant Bits, LSB）的方法。最低有效位算法指的是，对于音频载体信号而言，把水印信息通过替换的方法嵌入到样本的最低比特位。该方法易实现，具有较大的嵌入容量。在音频水印早期，该方法是主要的被采用的音频水印算法之一。在日常应用中，音频信号最低有效位几乎是最敏感的位置，也就是说该位置上的信息由于其敏感性而容易

改变。比如，在程度轻微的信号处理后，该位置的信息和原信号相比，会有很大程度的改变。于是，基于 LSB 的音频水印技术，常用来构造脆弱水印算法^[23]。

为了提高基于 LSB 方法的鲁棒性，文献[24]采用了两步的方法。第一步是，首先进行 LSB 编码，目的是将水印嵌入到 LSB 编码的相对高的层；之后，对由水印嵌入而带来的噪声进行降噪处理。该算法对于压缩攻击具有一定的鲁棒性。文献[25]给出了一种基于 LSB 的数字音频水印算法。不像传统的基于 LSB 的方法，作者并没有直接将信息嵌入到样本点的最低有效位，而是将嵌入位置提高到了一定的位置。另外，作者改变相邻的值，并且通过降噪的处理方法来提高水印系统的不可听性。

基于回声的数字音频水印算法属于另一种常见的时间域音频水印技术。研究表明，人耳听觉系统（Human Auditory System, HAS）在收到较强的信号之后，紧接着的 50-200ms 时间之内，再收到较弱的信号将不会被人耳感知，即人耳听觉系统时域掩蔽特性。基于回声的数字音频水印算法正是基于该特性来嵌入水印信号^[26]。实验结果表明，该算法对于有损压缩具有一定的鲁棒性。遗憾的是，水印提取的准确率有待提高。

总体而言，时间域音频水印算法效率高，易实现，再某些特定的场合具有一定的应用。然而，由于其算法鲁棒性不强，影响水印提取的正确率。

（2）基于频率域的音频水印算法

基于离散傅里叶变换（Discrete Fourier Transform, DFT）的水印嵌入方法是一种早期的常见的音频水印算法。文献[27]基于离散傅里叶变换构造了一种频域水印算法。首先将原始音频信号分帧，然后对各帧信号进行 DFT。水印信息生成后，将水印信息嵌入到 DFT 系数的幅值和相位中，实现了水印的盲提取。由于该算法水印嵌入位置和特征是公开的，容易被攻击者针对性地对水印进行破坏。文献[28]提出了一种基于频域音频水印算法，改进了传统的基于 DFT 算法的不足。文中采用基于混沌映射的方法来随机选取嵌入位置，解决传统方法中将水印嵌入在固定频率点带来的安全性不足的问题。

音频信号离散余弦变换（Discrete Cosine Transform, DCT）可以将信号能量主要集中在到中低频系数中。基于此，不少学者构造了性能较好的音频水印算法。文献[29]将音频信号进行离散余弦变换，并选取中频系数将二值图像生成的水印信息嵌入到中频系数中。文献[30]结合离散余弦变换和奇异值分解，提出了一种鲁棒性较强的音频水印算法，但是比较而言，水印嵌入容量有所降低。文献[31]提出了一种自适应的水印嵌入方法。首先对信号进行离散余弦变换，然后依照信噪比来衡量水

印强度，达到自适应的目的。

除了离散余弦变换之外，离散小波变换（Discrete Wavelet Transform, DWT）域也是常被用来进行水印嵌入的频域之一。文献[32]首先对原始信号进行小波变换，然后选取不同的系数来嵌入不同功能的水印，实现了版权保护和内容取证的功能。为了进一步提高算法的鲁棒性，使水印可以抵抗更强的攻击，文献[33]同时结合两种频域变换（DWT 和 DCT），将水印嵌入到频域系数范数中。算法满足了较强的鲁棒性和较好的不可听性要求。

（3）基于 Patchwork 的音频水印算法

Patchwork 的思想最先应用在图像水印中，改变像素点集亮度来将水印信息嵌入。Patchwork 图像水印技术具有很强的鲁棒性。近年来，学者们将 Patchwork 的思想应用在音频水印中，已有部分优秀的研究成果。文献[34]给出了基于 Patchwork 的音频水印算法，进行多次、多层的水印嵌入，使算法实现了优秀的性能。文献[35]提出了一种抗翻录攻击的音频水印算法。结合离散余弦变换和 Patchwork 的思想，构造了一种音频特征，给出了基于该特征的水印嵌入方法。实验结果证实了，该算法能够有效抵抗翻录攻击，为音频内容版权保护积累了技术。

除了上述频率域的音频水印技术之外，倒谱域也是常用来进行水印嵌入的一种技术手段。倒谱技术常用在语音编码、语音合成以及语音识别等领域。后来，有学者将倒谱域用在音频水印算法中。严格来讲，倒谱域音频水印算法可以归类于频域算法的类别，也具有其它常见频域算法的优点，比如鲁棒性较好等。文献[36]较早提出了基于倒谱域的音频水印算法。文中给出了统计均值调制 SMM (statistical mean manipulation) 的水印嵌入思想。分析指出音频信号处理操作对单个复倒谱系数而言影响较大，但是对于统计均值则影响很小。水印嵌入时，可以修改某一部分倒谱系数值，以此来使统计均值满足一定的要求。算法也将安全性考虑在内，构造了较为完善的水印系统。

基于以上分析，可以看出当前数字音频水印技术主要还存在以下几个问题：

（1）应用的深入和攻击的多样化，给数字音频水印系统的安全性带来了威胁。当前部分水印算法存在安全性不足的问题，比如，水印算法是公开的，如果嵌入水印的位置、量化的特征也是容易被获取的话，攻击者则会很容易地获取嵌入的信息，甚至是攻击者对含水印信号重新嵌入属于攻击者的信息，来证明攻击者对音频内容的所有权。

（2）应用场景的多样化，需要满足特定场景需求的音频水印技术。目前的音频水印技术，鲁棒的水印算法应用在版权保护上，而内容取证的功能欠缺；脆弱的

水印算法应用在取证方面，证明音频内容所有权的能力不足。在某些场合，可能会需要验证音频内容的拥有者，同时验证音频内容的真伪。当前，能用在此种场合的音频水印技术的研究成果相对较少。

1.3 本文创新点

本文基于保密特征的音频水印技术研究，主要工作有：

1. 定义了一种数字音频信号的保密特征，分析了该特征的特性，并且给出了基于所提特征的水印量化和嵌入方法。
2. 提出了一种基于保密特征的音频双水印算法，实验验证了该算法不仅能够对音频内容进行版权保护，也能对含水印信号的真伪进行鉴定。仿真实验验证了算法的有效性。

1.4 本文主要工作和论文组织结构

针对现有数字音频水印技术存在的问题，本课题主要研究了基于保密特征的数字音频水印技术。给出了数字音频信号的加密方法，定义了一种加密信号的特征，并且给出了基于所提特征的水印量化和嵌入方法，基于此提出了一种基于保密特征的数字音频水印算法。

第一章：引言。本章主要介绍了本文的研究背景和意义，引出了本文要解决的问题。紧接着，对国内外的研究现状进行了总结和分析。

第二章：数字音频水印理论基础。本章主要介绍了数字音频水印的相关知识，包括数字音频水印的性质和分类，以及常见的数字音频水印算法的评价标准。

第三章：基于保密特征数字音频双水印算法。本章属于本文的主要内容，详细介绍了本文提出的水印算法。对嵌入水印的特征进行加密，将水印嵌入到加密的特征中，提高了整个水印系统的安全性。同时，将帧号映射为水印信息，可以在含水印信号攻击后，通过重构帧号来检测被攻击的内容。对算法进行了实验仿真，验证了算法的有效性。

第四章：总结与展望。本章对全文的研究成果进行了总结，对研究的不足之处进行了思考，并对下一阶段的工作进行展望。

第 2 章 数字音频水印理论基础

数字化和信息化的普及，使得数字音频进一步成为交流沟通的主要载体。为了保护数字音频创作者的合法权益，以及保护敏感数字内容的真实性，数字音频水印技术成了热度不减的研究热点。接下来，本章将介绍有关数字音频水印相关的概念、对含水印信号常见的攻击，以及相关的评价方法。

2.1 数字音频水印的性质

简单来说，数字音频水印技术是一种将某种特殊用途的信息添加到载体信号中的一种技术。依据嵌入的信息，可以用来版权保护，或者对载体信号真实性取证。完整的数字音频水印算法包括水印嵌入过程（如图 2-1）和水印提取过程（如图 2-2）[37,38]。

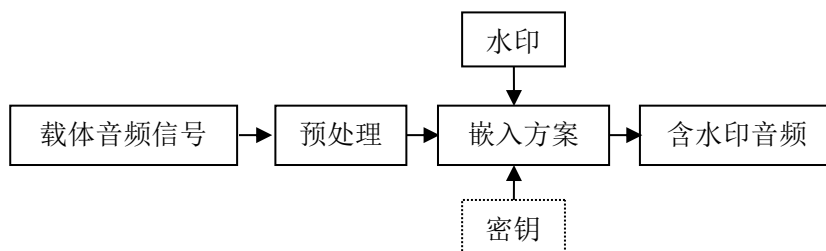


图 2-1 水印嵌入过程

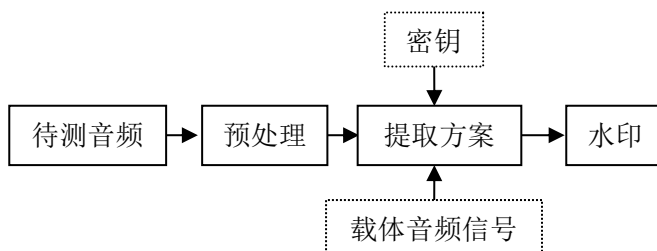


图 2-2 水印提取过程

总的来说，对于音频水印系统而言，常见的性质如下：

(1) 嵌入的信息需要通过一定的技术手段隐藏或者嵌入到载体音频数据中。最好是分布于整段音频内容，而不是特定的一段或者一部分。避免被针对性地抹去

删除。

(2) 载体音频中信息的嵌入不能造成信号的失真，而且嵌入的水印不应该被听觉感知到。

(3) 根据不同的用途，算法应该满足一定的鲁棒性或者是脆弱性。对于鲁棒性要求而言，含水印信号在遭受重采样、压缩以及添加噪声等信号处理后，水印要能准确地提取；对于脆弱性而言，嵌入的水印要具有一定的敏感性，含水印信号在攻击之后，嵌入的水印应该存在较大程度的破坏。

(4) 为了应用的方便，水印提取的时候，应该能够做到不需要原始信号，即算法应该满足盲提取的要求。

(5) 根据要求，水印系统的安全性不应该建立在对算法本身保密的基础上。水印系统的安全主要是靠密钥的安全来进行保密。为了水印系统的安全，密钥空间应该尽可能得大。

2.2 数字音频水印的分类

根据不同的分类方法，数字水印可以分为不同的类别。看似不同，却又有联系。通常的分类方法主要有这么几种。

(1) 按应用场景来划分。

根据用途的不同，可以将数字音频水印技术划分为两类，第一类是鲁棒音频水印技术，第二类是脆弱（半脆弱）音频水印技术^[39]。第一类的鲁棒音频水印技术主要应用在要保护数字音频内容的版权上。对于这一类水印算法而言，嵌入的水印信息通常是版权标识。另外，在经过一些信号处理操作，甚至是某些攻击之后，要尽可能地确保水印信息被准确地提取。如此才能保证音频水印算法的有效性。第二类的脆弱音频水印技术主要应用在需要确保数字音频内容是真实的场合。此类水印技术嵌入的水印对攻击具有较强的敏感性。含水印信号在被攻击的同时，嵌入的水印信息同时也被攻击。验证端正是通过水印提取是否有误来判断音频内容的真伪^[40]。

(2) 按检测过程划分。

依据水印提取的时候是不是需要用到原始信号，将音频水印分为盲提取的音频水印技术和非盲提取的音频水印技术。这两种水印技术根据字面意思，比较容易理解，盲提取的音频水印技术水印提取和检测过程不需要用到原始信号，可以直接从含水印的信号中提取水印信息。非盲提取的音频水印技术在水印提取和检测过程中必须有原始音频信号的配合。这两种技术各有优缺点，非盲水印提取技术鲁棒性较

强，然而需要大的存储空间，因为增加了水印提取所需要的原始音频信号的存储空间。但是该技术由于自身的弊端，在实际推广应用中比较难。为了方便技术的实际应用，音频水印技术都应该做到水印的盲提取^[41]。

（3）按水印嵌入位置划分。

音频水印技术主要是靠将具有特殊用途的水印信息嵌入到载体信号中来达到应用的目的。水印的嵌入位置简单来说，有时间域的嵌入和频率域的嵌入。基于此，将音频水印技术分为时间域音频水印技术和频率域音频水印技术。时间域音频水印技术常见的算法有基于 LSB 的音频水印算法，将水印嵌入到最低有效位中。频率域音频水印技术常见的算法有基于离散余弦变换（DCT）的音频水印技术^[42,43]、基于离散傅里叶变换（DFT）的音频水印技术^[44-46]以及基于离散小波变换（DWT）的音频水印技术^[47-49]。和时间域音频水印技术相比，基于变换域的水印技术更具有较强的鲁棒性，基于变换域的音频水印算法也越来越多。

（4）按生成水印的含义来划分。

数字音频水印算法中，嵌入的水印一般情况下，是由其它的信息来生成的二进制序列，这里的其它信息分为有意义的信息和无意义的信息两种。无意义的信息指的是如伪随机序列等，有意义的信息指的是一副有意义的数字图像（如商标）、一音频片段等。如果嵌入的水印是由无意义的序列来生成，那么在水印检测和提取阶段，一旦提取的水印出现误码的情况，就需要借助于其它的方法来判断提取水印的有效性。而有意义的水印，即使水印提取存在部分错误，仍可以借助于我们主观的观察来判断水印提取是否有效。所以，从这个角度而言，有意义的水印更有优势。

2.3 数字音频水印的评价方法

一般情况下，数字音频水印系统的评价指标有三个，不可听性（不可感知性）、鲁棒性和水印嵌入容量。三个指标相互制约，需要根据实际需求在三个指标中达到平衡。算法设计中，要让这三个指标同时达到最优是很难做到的。比如，对于一个不可听性要求较高的算法，那么嵌入容量一定不会太大。因为嵌入容量越大，表明嵌入的水印信息越多，对原始信号的修改就越大。这样一来，不可听性就会下降。当然，只要不可听性达到一定的标准就好，即满足通常的不可听性要求即可。一个性能卓越的数字音频水印系统要求能够尽可能做到三者的平衡。比如，对于用于取证的数字音频水印系统，鲁棒性不是最核心的需求。于是，对于此类算法而言，可以适当降低鲁棒性，提高算法的不可听性和水印嵌入容量。对于用于版权保护的数

字音频水印系统而言，鲁棒性就显得比较重要。对于此类算法，可以适当降低水印嵌入容量，提高算法的鲁棒性，以达到实际需要的目的。总体来说，音频水印系统三个评价指标的关系，可以由图 2-3 来表明。

(1) 不可听性评价方法。

水印的嵌入，或多或少地都会对原始音频信号进行修改。修改的幅度太大，会引起明显的噪声。不可听性是从某一个方面来衡量水印的嵌入对原始信号的修改程度。通常由主观和客观两种。

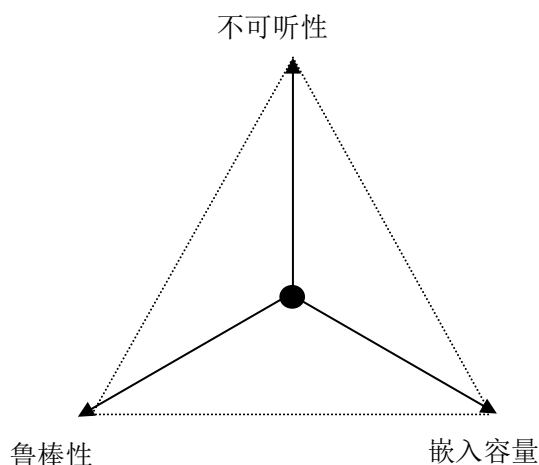


图 2-3 数字音频水印算法不可听性、鲁棒性和嵌入容量之间的关系

主观的方法，将原始音频以及含水印音频播放给听众进行试听，依据主观区分度 SDG (Subjective Difference Grades) 进行打分。对打分结果求平均值，平均值作为最终的评测结果。表 2-1 给出了 SDG 值的打分方法。可以看出，SDG 值越大，水印的不可听性越好^[50]。

表 2-1 SDG 不同分值对应的不可听性描述

SDG	描述
0.0	不可听性最好
-1.0	不可听性较好，但能感觉
-2.0	轻微的不舒适感
-3.0	比较刺耳
-4.0	非常刺耳

客观方法主要是根据信噪比 SNR(Signal to Noise Ratio)，和由 PEAQ(Perceptual Evaluation of Audio Quality)测试工具得出的听觉质量客观区分度 ODG(Objective Difference Grade)，SNR 定义如式 (2-1) 所示。信噪比 SNR 的评价方法实际上是将

嵌入的水印看作给原始信号添加的噪声，而式 (2-1) 则是用来计算原始信号和含水印信号的差异度，用差异度来衡量水印嵌入对原载体信号的改变程度。

$$\text{SNR} = 10 \lg \left(\frac{\sum_{l=1}^N a(l)^2}{\sum_{l=1}^N (a(l) - a'(l))^2} \right) \quad (2-1)$$

其中， $a(l)$ 和 $a'(l)$ 分别表示含载体信号和含水印信号的第 l 个样本点， N 表示信号的样本总数。易得 SNR 值越大，两信号差异就越小，从而水印不可听性越好。大量实验结果表明，SNR 值大于 20 时，水印信号的不可听性较好。表 2-2 给出了 ODG 的评分标准。

表 2-2 ODG 不同取值对应的不可听性描述

ODG	描述	质量等级
0.0	完全不可听	优
-1.0	稍微可感觉，但不影响	良
-2.0	可听，不刺耳	中
-3.0	比较刺耳	差
-4.0	非常刺耳	极差

(2) 鲁棒性评价标准。

鲁棒性指的是信号中嵌入的水印抗攻击的能力，也可以理解为含水印信号在经过信号处理后，水印是否还能够准确提取的能力。常见的测试指标有误码率 (Bit Error Rate, BER) 和相关系数 (Normalized Correlation NC)。

BER 是直接测试提取的水印和原始水印的不同程度，计算方法如式 (2-2)。

$$\text{BER} = \frac{W_r}{W_o} \times 100\% \quad (2-2)$$

其中， W_r 表示水印提取的错误比特数， W_o 表示水印嵌入总的比特数。从 BER 的定义来看，BER 值越小，代表水印提取错误率越低，算法鲁棒性越强。反之，BER 值越大，代表水印提取错误率越高，算法鲁棒性越弱。

相关系数体现的是含水印音频提取的水印和原始水印之间的相似性。对于采用二值图像作为水印的系统，常用相关系数来测试算法的鲁棒性。相关系数 NC 的计算方法如式 (2-3) 所示。

$$\text{NC} = \frac{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W(i, j) W'(i, j)}{\sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W(i, j)^2} \sqrt{\sum_{i=1}^{M_1} \sum_{j=1}^{M_2} W'(i, j)^2}} \quad (2-3)$$

其中, $M_1 \times M_2$ 为二值图像的大小, $w(i, j)$ 为原始二值图像的像素值, $w'(i, j)$ 为从含水印信号中提取的二值图像的像素值。NC 值越大 (越接近 1), 则表示提取的二值图像和嵌入的二值图像越相近, 表明算法的鲁棒性越好。NC 值越小 (越接近 0), 则表示提取的二值图像和嵌入的二值图像相似性越低, 表明算法鲁棒性较弱。

(3) 水印嵌入容量。

水印嵌入容量衡量的是单位时间内载体信号中嵌入的包含水印信息的数据量, 通常采用比特率来描述, 单位为比特/秒 (bps)。根据国际留声机工业联盟的要求, 数字音频水印系统中嵌入容量通常要大于或等于 20bps。水印嵌入容量的计算方法如式 (2-4) 所示。

$$C = \frac{f_s}{L/n} \quad (2-4)$$

其中, f_s 表示载体音频信号采样频率, L 的含义是音频信号的采样点数, n 的含义是水印信息总比特数。

2.4 数字音频水印常见的攻击方法

数字水印算法的抗攻击能力, 是水印技术走向应用的一个关键指标。数字音频信号属于一维信号, 经过部分攻击后, 被攻击内容后面的部分会向前移动, 对水印造成严重的破坏。于是, 根据攻击是否会破坏水印的同步, 将攻击分为两类, 普通攻击和去同步攻击。

2.4.1 普通攻击

普通攻击也被称为常规信号处理, 常见的有添加噪声, 滤波处理, 重采样, 以及压缩等。

(1) 添加噪声。

添加噪声指的是, 在含水印的音频中添加噪声。常见的添加的噪声的类型有高斯噪声, 或者是白噪声。添加噪声之后, 含水印音频的样本幅度会发生改变。因此, 添加噪声处理会影响水印的准确提取。于是, 对于水印提取和音频样本幅度关系密切的算法而言, 添加噪声会较为严重地影响水印的提取。

(2) 滤波处理。

数字音频信号在经过频域变换之后, 会较为明显地看到音频信号的低频部分和

高频部分。滤波处理常见的有低通滤波处理和高通滤波处理，以及带通滤波处理。低通滤波主要涉及到音频信号的低频部分，而高通滤波则主要涉及到音频的高频部分。

(3) 重采样处理。

重采样处理，是一种改变含水印信号采样频率的处理方法。一般的操作过程是，首先改变含水印信号的采样频率，然后再恢复到含水印信号的原采样频率。重采样分为上采样和下采样。上采样指的是，通过对含水印信号进行插值，来提高含水印信号的采样频率，然后再降低采样频率至含水印信号的原采样频率。下采样指的是，通过对含水印信号进行删除样本的操作，来降低含水印信号的采样频率，然后再通过插值的方法来提高采样频率至含水印信号的原采样频率。

(4) 压缩处理。

为了数据存储的方便或者是为了使用不同播放格式的要求，需要对含水印信号进行压缩处理。通常的压缩方法，是根据人耳听觉掩蔽效应，把音频信号中的冗余信息去除，来减少音频信号的数据量和码率。最常见的压缩方法是 MP3 压缩。MP3 压缩能够保证压缩信号的听觉质量，能够以较低的采样频率进行压缩，压缩带来的听觉质量损失几乎可以忽略不计。对于那些基于音频信号的听觉冗余来嵌入水印信息的算法而言，MP3 压缩将会严重影响水印的准确提取。

2.4.2 去同步攻击

对于攻击者而言，去同步攻击的目的是通过改变水印的位置，来影响水印的准确提取。一般而言，去同步攻击后，水印信息较大概率地会依然存在于被攻击信号中，但是水印信息的位置会发生大幅度地改变。严重增加水印检测器检测到水印信息的能力。因此，去同步攻击是音频水印算法较为难以抵抗的一种攻击方法。常见的去同步攻击主要有以下几种：

(1) 变调。

变调处理会让含水印音频的频率发生偏移，对于数字音频水印系统而言，变调攻击将会给水印系统带来严重的威胁。

(2) 抖动攻击。

抖动攻击是，在含水印的音频信号中均匀地在每个固定长度的样本中插入或者删去一个或者几个样本。

(3) 随机剪切或增加样本。

对含水印信号小幅度地随机剪切或者增加部分样本，人耳在听的时候，有可能会听不出来含水印信号的改变。然而，和原始含水印信号相比，攻击后的信号的样本数量和原始信号相比，将会有大的改变。该攻击对于大多数时域数字音频水印算法或者基于扩频技术数字音频水印算法的水印信息检测，造成很大的影响。

(4) 时间标度修改。

时间标度修改攻击是在时域内进行的，指的是在含水印数字音频信号中周期性地添加样本或删除样本，还包括通过复杂的时间标度修改技术来对音调进行维持。这会使得含水印数字音频信号的长度可能会发生增加或减少。

(5) D/A 和 A/D 转换。

一个音频信号属于模拟信号还是数字信号是根据携带这个音频的物质所决定的，比如计算机上的音频信号由声卡输出，而后经过录制存储到磁带中，就需要通过 D/A 转换这一过程。音频信号经过 D/A 或 A/D 转换后，不仅会改变样本的幅值，样本的位置也可能会发生平移，这就会破坏水印信息和音频信号之间的同步结构。

2.5 本章小结

本章主要对数字音频水印算法的相关知识进行了梳理和简单论述，主要包括数字音频水印的性质、水印算法的分类，以及常见的水印系统的评价方法，比如不可听性、鲁棒性和水印嵌入容量的评价标准。本章的相关内容可以用于对第三章提出的音频水印算法性能进行综合评价，衡量算法的性能。

第3章 基于保密特征的数字音频双水印算法

数字信号处理技术的普及，提升了大家对数字多媒体信号处理的能力，方便了大家对数字音频、图像以及视频的处理和编辑。然而，也为不法分子通过对数字媒体的编辑到达某种不可告人的目的提供了便利。比如，侵权行为、对数字媒体内容的恶意攻击行为频发，侵害了合法拥有者的利益，甚至是消弱了大家创新创造多媒体内容的积极性。如何保护媒体所有人正当的合法权益，已成为急需解决的技术难题。

对于数字音频内容而言，数字音频水印技术为音频内容的版权保护和内容取证提供了一种技术手段。该技术将有特殊用途的信息不被听觉感知地隐藏到音频信号中，生成含水印的信号。在遇到版权纠纷或者是需要对内容进行真伪鉴定时，提取隐藏在音频信号中的信息，达到对音频进行保护的目的。数字音频水印技术发展至今，在版权保护应用上已有丰富的成果^[28-34]。此外，在对音频内容真伪鉴定上，也起到了积极的推动作用^[35-36]。不管是版权保护的水印技术或者是取证的水印技术，往往仅关注一个方面的应用。随着应用的深入，某些音频内容的拥有者不仅希望保护作品的所有权，同时也希望对音频内容进行真伪鉴定。对于此种需求，当前可用的研究成果还相对较少。

为了解决当前水印技术的缺陷，同时满足版权保护和内容取证的需求，本文提出了一种基于保密特征的数字音频双水印算法。首先将原始音频信号分帧，并对各帧信号加密，在加密信号的基础上提取特征。然后，根据版权信息和用于取证的信息生成的水印，量化特征。最后生成含水印的信号。实验结果表明，本文算法提高了水印系统的安全性，对信号处理具有一定的鲁棒性。对于恶意攻击，能够检测被攻击的位置。

3.1 基础知识

3.1.1 离散小波变换

离散小波变换（DWT）属于一种频域变换，具有广泛的应用。它打破了傅里叶变换进行信号分析的限制，产生了对信号的时频分析方法，实现了对信号不同分辨

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。如要下载或阅读全文，请访问：<https://d.book118.com/298073045053006023>