高级音频编码(AAC)的一种信息隐藏方法

唐步天[1], 郭 立[1], 刘振华[2]
(1. 中国科技大学电子科学与技术系, 合肥 230026; 2. 中国科学院研究生院信息安全国家重点实验室, 北京 100039)

摘要: 信息隐藏是 20 世纪 90 年代逐步兴起的研究课题。针对高级音频编码(Advanced Audio Coding,简称 AAC),简要描述了其编解码过程,说明了 AAC 信息隐藏的主要的概念性方法,然后提出了一种利用频域量化值的统计特性与霍夫曼码书选择间的相关性的信息隐藏的概念性方法,并描述了信息隐藏和信息提取的过程。实验在约 3s 的 AAC 压缩音频中隐藏了 526bit 的信息,证明了该概念性方法的可行性。还分析了目前的信息隐藏的对抗技术——隐秘信息检测技术对其的适应能力。

关键词: 信息隐藏;高级音频编码(AAC);霍夫曼码书选择
中图法分类号: TP309.7 TN912.3 **文献标识码**: A **文章编号**: 1000-3630(2008)-04-0533-06

An information hiding method in advanced audio coding (AAC)

TANG Bu-tian[1], GUO Li[1], LIU Zhen-hua[2]
(1. *Department of Electronics Science and Technology, University of Science and Technology of China, Hefei* 230026, *China;* 2. *Key State Lab of Information Security, Graduate University of Chinese Academy of Science, Beijing* 100039, *China*)

Abstract: The problem of information hiding has been studied since 1990′s. Based on describing the coding-decoding process of AAC and the general ideas of hiding in AAC, a hiding idea of using the relationship between Huffman codebook selection and statistic properties of quantified values in frequency domain is proposed, and the process of information hiding and extracting is also described. The availability of the idea is proved by an experiment of hiding 526 bits in an AAC compressed audio signal of 3 seconds. Some present hiding detection methods are limited to adapt this idea.

Key words: information hiding; Advanced Audio Coding(AAC); Huffman codebook selection

1 引 言

信息隐藏是近十多年逐步开展并兴起的一热门研究领域,它是一门研究如何将秘密信息隐藏于感官媒体的学问,研究人员通常以 1996 年在剑桥召开的第一届信息隐藏国际会议作为信息隐藏研究的起跑点。信息隐藏的兴起得益于多媒体技术和互联网络的发展,它的知识范围涉及电子学与信号处理、计算机技术、生理学和物理学等领域,目前视频图象、音频和文字等的隐藏的研究都得到了开展,研究成果也在版权保护、隐私保护、秘密通讯等领域得到较多的应用, 最典型的就是在数字水印和 DRM 中的应用。

在图象和声音中进行信息隐藏的最简单的办法就是 LSB 方法,随着研究的深入,扩频通信的思想被引入到信息隐藏研究之中,研究者提出了利用修改音频数据的傅立叶系数的相位值的相位编码法、DCT 频域变换法和小波变换的信息隐藏方法,另外还有利用分形的信息隐藏方法, 这些研究涉及到如 bmp, JPEG 图象和 wav,midi,mp3 等媒体格式文件。

音频信号的信息隐藏除了可以利用上述的普适的隐藏思想外,Gruhl 等提出了回声隐藏的方法,另外,还有学者提出了利用倒谱域的隐藏方法[1]。同时,由于多媒体技术的发展,跟随新的应用音频媒体格式的研究也同步地得到了开展。

收稿日期:2007-05-01;修回日期:2007-08-14
基金项目:自然科学基金资助项目(60577039)。
作者简介:唐步天(1973-),男,江苏盐城人,博士研究生, 研究方向为信息隐藏、信息安全。
通讯作者:唐步天,E-mail:tangbutian@abchina.com

高级音频编码（AAC）是运动图象专家组在MPEG-1后推出用于MPEG-2标准的一种音频编码，它是不向后兼容(NBC)的编码方式。2000年MPEG-4音频标准继续沿用了MPEG-2 AAC，同时追加了一些新的编码特性，所以它现在被称为MPEG-4 AAC。

2 AAC编码过程概述

AAC定义的编解码的原理[2,3]如图1、图2所示：

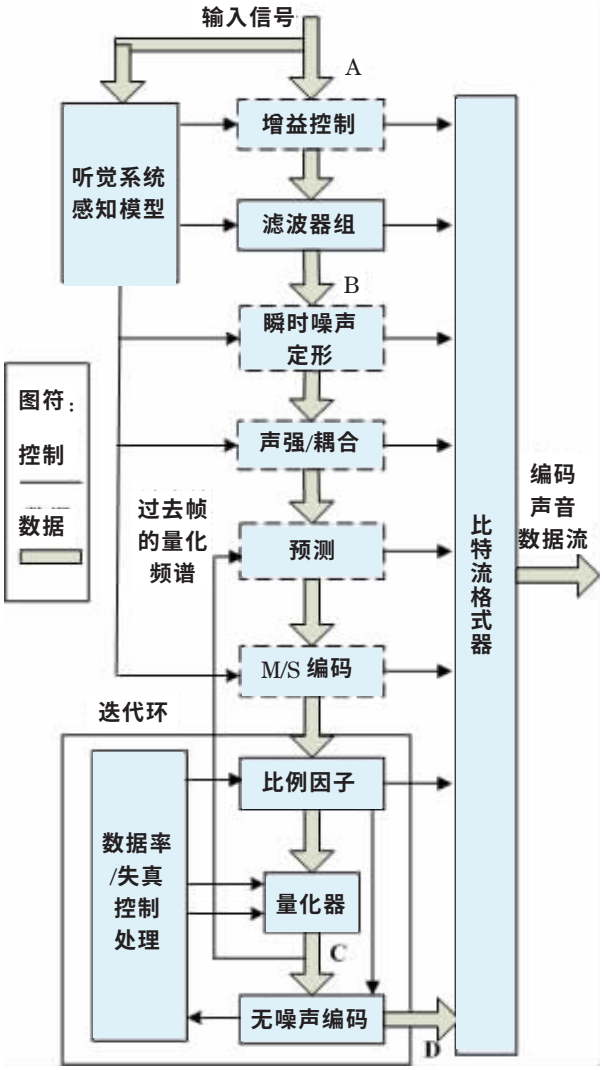

图1 AAC编码原理图
Fig.1 Coding of AAC

AAC标准定义了3种配置：

基本配置。它仅舍弃了增益控制模块，能提供最好的声音质量。

低复杂性配置。它舍弃了预测模块、增益控制，TNS的阶数也受到限制。

可变采样率配置。这种配置增益控制是必需的，不包括预测和声道耦合，TNS的阶数和带宽也受到限制，因此它比基本配置和低复杂性配置更简单。

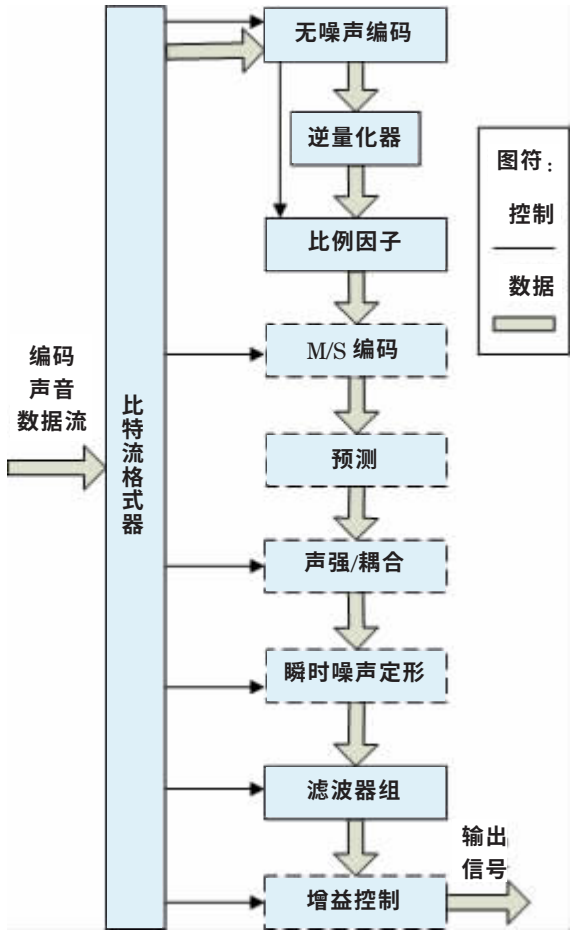

图2 AAC解码原理图
Fig.2 Decoding of AAC

以上配置，滤波器组、比例因子模块、量化器、无噪声编码/解码模块和比特流格式器都是必不可少的。下面说明这些主要模块的功能。

滤波器组：在编码中的功能是将时域音频信号进行MDCT变换得到频谱；在解码器中的作用是将频谱进行IMDCT变换重建时域音频信号。

比例因子模块：在编码中，利用比例因子将归一化的频谱非归一化；在解码中，利用比例因子将非归一化的频谱归一化。

量化器：在编码中，非均匀地将频谱量化成整数值；在解码中，将整数值反变换成非归一化的频谱值。

无噪声编码/解码模块：无噪声编码模块将频谱的量化整数值用霍夫曼编码压缩，并按比例因子记录各生理听觉频谱分区。无噪声解码模块则用根据记录的各生理听觉频谱分区建立比例因子信息，同时用霍夫曼码书解码出各分区的频谱的量化整数值。

比特流格式器：在编码时，将霍夫曼压缩的分区信息、霍夫曼压缩的频谱的量化整数值、M/S判决信息

(可选)、预测器状态信息(可选)、强度立体声/耦合控制信息(可选)、瞬时噪声定形信息(可选)、滤波器控制信息、增益控制信息(可选)打包输出比特流;在解码时,将输入比特流分解成上述的各种信息成分。

3　三种概念性方法

基于 AAC 编码的信息隐藏,目前的研究主要基于如下的三种概念性方法:

(1) 时域幅度修改。这是最普遍适用的信息隐藏的概念性方法,Ryuki Tachibana 的方法在采取频域嵌入手段的同时,也适当地采用了时域幅度修改的方法[4]。这种概念性方法的隐藏点选取在图 1 中的数据流 A 中。

(2) 利用频域实数信号的听觉冗余。AAC 编码中的滤波器组对时域音频信号进行 MDCT 变换得到了频域上的实数信号,在频域上,能以不影响主观听觉效果为前提,在特定的频率区间内以不同的比例系数嵌入特定的信号。此方法常用于对水印信号的嵌入隐藏。C. Neubauer 和 Ryuki Tachibana 均使用这种概念性的方法进行水印信号的嵌入隐藏[4,5]。这种概念性方法的隐藏点选取在图 1 中的数据流 B 中。

(3) 利用频域量化值的的听觉冗余。AAC 编码中频域经 TNS、M/S 处理后,需要转化成整数值存储,这种转化是通过量化完成的。AAC 的量化是非均匀量化,在生理听觉的各个频带内的量化使用不同的比例因子,量化后的整数值的微小差异不会引起的该频带的听觉效果的差异。这种概念性的方法已经在 MP3 的信息隐藏中应用,在 AAC 中同样实用。这种概念性方法的隐藏点选取在图 1 中的数据流 C 中。

C. Neubauer 和 Ryuki Tachibana 的研究主要集中于水印的研究,目的在于对加强对数字权利的认证和追踪,只需从概率上对隐藏的数字标识进行归属或是否的判别。

本文提出的一种概念性的方法利用的是频域量化值的自身统计特性与霍夫曼码书选择间的相关性。频域量化值的自身统计特性决定了其自身被无损压缩的能力和其对霍夫曼码书的适应程度,若利用可选码书压缩数据可以得到相等的最优长度比特,那么信息的隐藏可以以使用的可选码书来表达。利用这种概念性方法,信息被隐藏在图 1 中的数据流 D 中,而隐藏的过程是在无噪声编码模块完成的。这种概念性方法主要在于信息的隐藏与恢复,目的用于秘密通讯和隐私保护,需要完全地恢复信息。

4　利用频域量化值的统计压缩特性的信息隐藏概念性方法

AAC 编码中的无噪声编码和比特流格式器模块会将频域量化值进行压缩试算,选择最佳压缩方案,将频域量化值压缩成比特流输出,其基本过程如下:

(1) 预先根据人类听觉系统的临界频带生理特征,划分出若干个频域区间,各个区间称为比例因子频段(sfb)。

(2) 对每一帧的频域量化数据,对应到各个比例因子频段。

(3) 对每个比例因子频段,按量化值的特点,用固定的 0 到 11 号霍夫曼码书中的几种进行压缩试算,选取压缩后比特数最少的码书作为该比例因子频段的压缩码书。

(4) 对相邻的比例因子频段(或者在要求压缩比特数小于两个区间的压缩比特数的和值时),对使用相同码书者进行区间合并。

(5) 将各比例因子频段起止段、所用的霍夫曼码书、频域量化数据压缩后的比特、各比例因子频段的比例因子及其他信息按比特流组装格式拼装成比特流输出。

AAC 编码中的 0 到 11 号霍夫曼码书分别反映了截然不同的概率分布函数,并设定了所能表示的量化系数的最大值[3] (0 号码书代表压缩数据全为 0)。表 1 列出了各码书的一个压缩分组的数据个数及数据的最大绝对值等条件。

AAC编码中的压缩试算规则如表 2 所示。

表 1　码书适用性表

Table 1　Uses of codebooks

码书序号	数据个数	最大绝对值	带符号值
0		0	
1	4	1	是
2	4	1	是
3	4	2	否
4	4	2	否
5	2	4	是
6	2	4	是
7	2	7	否
8	2	7	否
9	2	12	否
10	2	12	否
11	2	16(溢出)	否

表 2 试算规则表

Table 2 Regulations for test calculation

数据最大绝对值条件	试算规则
情况 1:最大绝对值<=1	用 1,2,3 号码书压缩试算
情况 2:1<最大绝对值<=2	用 3,4,5 号码书压缩试算
情况 3:2<最大绝对值<=4	用 5,6,7 号码书压缩试算
情况 4:4<最大绝对值<=7	用 7,8,9 号码书压缩试算
情况 5:7<最大绝对值<=12	用 9,10 号码书压缩试算
情况 6:最大绝对值>=13	用 11 号码书压缩

比例因子频段中的不同频域量化值经不同的霍夫曼码书编码后的得到的比特数通常会有所差异,这也是 AAC 编码采用多个码书进行压缩试算,然后选优的原因。但如果同一比例因子频段内频域量化值经不同的霍夫曼码书编码后得到相等的最优长度比特的概率是可观的,那么这种概率就为信息隐藏提供了隐蔽空间。

在 $2562890625(15^8)$的计算强度内,通过程序计算,得到了不同试算条件下的相等的最优长度比特的概率,如表 3 所示。

表 3 不同试算条件概率表

Table 3 Possibilities under different conditions

试算规则	实验压缩数据个数	样本空间	得到等长比特且比特数最少的点数	概率
(1,2,3 号码书)	4	80	7	8.75%
(3,4,5 号码书)	4	544	52	9.55%
(5,6,7 号码书)	4	5936	96	1.61%
(7,8,9 号码书)	4	44064	372	0.84%
(9,10 号码书)	4	340000	1168	0.34%
小计		$25^4-1=390624$	1668	0.43%
(1,2,3 号码书)	8	6560	155	2.36%
(3,4,5 号码书)	8	384064	16616	4.32%
(5,6,7 号码书)	8	42656096	417676	0.98%
(7,8,9 号码书)	8	2519843904	3824832	0.15%
小计		$15^8-1=2562890624$	4259279	0.17%

根据表 3 中的概率估算,1s 22.05kHz 采样的音频信号平均可隐藏约 20bit~50bit。

设因子频段内频域量化值序列为 da, 压缩试算所用的霍夫曼码书集合为 $H=\{h1,h2,\cdots,hn\}$,第 i 种霍夫曼码书对 da 的编码函数表示为$f(i,da)$,对编码后的比特求比特数函数为 lenbit, 求最小值函数为 min,则其概念性方法可表示为:

若存在 $i<j$,且 $i,j\in$H,使得

$$\text{lenbit}(f(i,da))=\text{lenbit}(f(j,da))=\min(\text{lenbit}(f(h1,da)),\text{lenbit}(f(h2,da)),\cdots,\text{lenbit}(f(hn,da)))$$

则可用选 i 号码书代表隐藏比特 0,用选 j 号码书代表隐藏比特 1,反之也可。

若更进一步,可以考虑对频域量化值修改,要隐藏 1,则修改频域量化值,使得满足表达式;要隐藏 0,则修改频域量化值,使得不满足表达式。当然,此想法需要更多地研究 0 到 11 号霍夫曼码书的概率特性,目前还未取得此方面进展。

利用概念性方法,成功地进行了信息隐藏的实验,下面说明隐藏过程和提取过程。

5 信息隐藏和信息提取

为信息隐藏基本过程如下:

(1) 将要隐藏的信息转换成比特流 $b(1),b(2),\cdots,b(n)$。通常也把隐藏比特流的长度作为 b 的一部分,以便于信息提取。

(2) 生成与隐藏比特流同等长度的 0、1 比特序列做隐藏密钥。生成函数可以选用密码学中的 m 伪随机序列生成函数,假定生成函数为$f(n)=c,c\in\{0,1\}$。

(3) 设定记数器 $cnt=0$。

(4) 从音源读入 PCM 音频数据,按 AAC 算法分帧、通过滤波器组及其他可选模块,直到量化完成,即将进入无噪声编码模块。此时,可以得到帧的量化整数值。

(5) 在帧内对每个比例因子频段的量化值进行进行压缩试算。

```
while (还有比例因子频段要压缩试算)
{
 if (该比例因子频段用两个码书 i 和 j(i<j)可以得到相同的最小长度的压缩比特)且(cnt<隐藏比特流长度)
        {
         if (b(cnt)⊕f(cnt))
          选用码书 i;
         else
          选用码书 j;
         cnt=cnt+1;
        }
}
```

(6) 通过比特流格式器,输出 AAC 码流。

(7) 保留或发送隐藏密钥,以便信息提取。

信息提取基本过程如下：

(1) 输入作为隐藏密钥的 0、1 比特序列 $f(n)$。

(2) 读入 AAC 码流，通过比特流格式器，分解出每帧的数据及帧中各比例因子频段所用的压缩码书。

(3) 设定记数器 cnt=0。

(4) 对每个比例因子频段的量化值进行压缩试算。

while (还有比例因子频段要压缩试算)

{

if (该比例因子频段用两个码书 i 和 $j(i<j)$可以得到相同的最小长度的压缩比特)且(cnt<隐藏密钥比特序列长度)

{

设定 $flag$=0;

if (码书 i=该比例因子频段所用的压缩码书)

$flag$=1;

else

$flag$=0;

$b(cnt)=flag \oplus f(cnt)$;

$cnt=cnt+1$;

}

}

(5) AAC 无噪声解码，反量化，直到输出音频信号流。

(6) 将提取的比特流 b 转成信息输出。

6　实验结果与分析

实验对 22.05kHz 采样的约 3s 的“中华人民共和国”音频文件，采用 AAC 压缩编码(低复杂性配置，11.025kHz 带宽)，在其中隐藏了 526bit (约 65byte)的信息。信息提取程序正确地进行了信息提取。

表 4 列出了几种不同类型的音频片段的隐藏实验结果。

目前，隐秘信息检测技术主要分三类：感官检测、统计检测和特征检测[6]。

由于本隐藏方法的音频输出质量取决于AAC的编码性能，而 AAC 压缩算法是符合人类生理听觉特点的，对人类的听觉感知影响很小，从表 4 中可以看出听觉感官检测对本隐藏方法不具备有效性。

目前的统计检测方法主要有针对信号的时域或变换域的统计检测方法，时域对应于图 1 中的数据流 A，变换域对应于图 1 中的数据流 B 和 C，由于本隐藏方法未在数据流 A、B 和 C 中调制隐秘信息，因而如果将通常的时域或变换域的统计检测方法用于对本隐藏方法进行检测，只会造成误判。

只有针对本隐藏方法的原理，采用以己之矛攻己之盾的思想，进行 AAC 隐藏点(比例因子频段的码书)的特征性分析的特征检测方法，对隐秘信息的存在与否的分析才具有针对性和实用性。我们的下一步研究目标就是针对这种隐藏方法，试图建立特征分析模型。然而即使确证了隐秘信息的存在，要实现在未知隐藏密钥的情况下的隐藏的隐秘信息的提取，还是很困难的(毕竟在隐藏时，每个可隐藏点可以有隐藏和不隐藏两种选择)，这类似于 LSB 隐藏方法，可以进行概率性的检测分析，很难实施信息的盲提取。表 5 列出了各实验音频片段的隐藏前后各码书的使用频度，可供特征检测方法研究者参考。

本隐藏方法的缺点是无法对抗主动攻击，因为不同的软件对最优压缩码书的次序选择不一样(AAC 编码中并未对最优压缩码书的次序选择进行强制性规定)，采用不同的软件对 AAC 进行重新编码，会使隐藏的信息丢失。

表 4　不同音频片段的实验结果

Table 4　Experiment results of different audio fragment

片段序号	音频种类与内容	时长/s	隐藏比特数	信噪比/dB	听觉效果
1	古典(贝多芬交响曲)	9.77	1843	18.8429	无明显感知影响
2	乐器(钢琴曲)	13.08	4022	20.6707	无明显感知影响
3	托福语音(一半静音)	10.35	265	19.4164	无明显感知影响
4	自然声音(海滩声音)	5.06	758	17.6293	无明显感知影响
5	冲锋号	3.77	505	19.8788	无明显感知影响
6	乐器(二胡)	12.35	3647	21.5715	无明显感知影响
7	流行乐(万水千山总是情)	9.82	1507	18.9457	无明显感知影响

表5 隐藏前后码书使用频度表

Table 5 Frequencies of using codebook under no hiding and hiding condition

		片段序号1	片段序号2	片段序号3	片段序号4	片段序号5	片段序号6	片段序号7
无隐藏的AAC	码书0	0	0	98	1	2	0	0
	码书1	1498	2899	372	1043	443	2934	1185
	码书2	190	573	42	169	205	1119	216
	码书3	811	1461	304	75	268	1154	687
	码书4	1189	1444	106	842	208	1824	1282
	码书5	1194	2021	314	83	186	1586	1264
	码书6	1502	1425	241	1388	156	786	1677
	码书7	696	1173	215	30	233	740	578
	码书8	820	957	243	329	184	373	733
	码书9	99	121	58	3	120	345	39
	码书10	140	170	137	44	83	181	136
	码书11	0	1	33	3	5	2	6
隐藏后的AAC	码书0	0	0	98	1	2	0	0
	码书1	1120	2061	310	715	290	1962	969
	码书2	205	599	42	175	205	1159	223
	码书3	1112	2125	356	391	407	1915	817
	码书4	1164	1374	103	833	200	1747	1258
	码书5	1047	1619	292	91	168	1483	1097
	码书6	1510	1441	247	1389	157	791	1692
	码书7	784	1609	218	28	239	966	734
	码书8	846	1002	247	331	191	387	756
	码书9	202	211	80	9	141	433	113
	码书10	149	203	137	44	88	199	138
	码书11	0	1	33	3	5	2	6

7 结 论

AAC是一种符合人类生理听觉特点的有效的有损音频压缩算法,这种算法将信号在频域划分成若干比例因子频段,为了对频域量化值进行最佳压缩,每个比例因子频段都用了几种霍夫曼码书进行压缩试算,并选取最优者。AAC的压缩试算特点,使得AAC给信息隐藏提供了除时域信号值、频域信号实数值和频域信号量化值之外的第四种隐藏空间——压缩码书。实验成功地利用压缩码书进行了信息的隐藏和提取。更进一步的研究,可考虑将频域量化值修改和压缩码书的选择结合起来进行信息隐藏。

参 考 文 献

[1] LI X, YU H H. Transparent and robust audio data hiding in cepstrum domain[A]. International Conference on Multimedia and Expo 2000[C]. New York. 2000, 1. 397-400.

[2] 张益贞, 刘滔. Visual C++实现MPEG/JPEG编解码技术[M]. 北京: 人民邮电出版社, 2002. 274-295.
ZHANG Yizhen, LIU Tao. Implementing MPEG/JPEG coding and decoding technology by using Visual C++[M]. Beijing: People Posts and Telecommunications Publishing House, 2002. 274-295.

[3] GB/T 17957.7-2002. 信息技术运动图象及其伴音信息的通用编码第七部分: 先进音频编码(AAC)[S]. 北京: 中国标准出版社, 2003.
GB/T 17957.7-2002. Information technology-generic coding of moving pictures and associated audio information Part 7: Advanced audio coding (AAC)[S]. Beijing: Standards Press of China, 2003.

[4] Ryuki Tachibana. Two-dimensional audio watermark for MPEG AAC audio[J]. Proc.SPIE, 2004, **5306**(1): 139-150.

[5] Neubauer C. Herre J. Audio watermarking of MPEG2 AAC bit streams[A]. Proc. of 108th Convention of Audio Engineering Society(AES)[C]. Paris, 2000.

[6] 郭芬红. 几种典型的隐秘信息检测算法[J]. 华南金融电脑. 2006, (2): 89-91
GUO Fenhong. Several typical detection methods for secret information hiding[J]. Financial Computer of South China, 2006, (2): 89-91.