

doi: 103969/j.issn.0490-6756.2017.03.013

适用于广播信道的时频调制音频隐藏算法

李丹, 高勇

(四川大学电子信息学院, 成都 610065)

摘要: 广播信道中包含大量音频信息, 给音频信息隐藏应用提供了广阔空间, 但在广播信道上缺乏行之有效的音频隐藏算法. 针对上述问题提出了一种时频调制音频信息隐藏算法, 时频调制技术具有较强的抗信道干扰特性, 在广播信道上具有很强的适用性. 该算法在音频载体中嵌入一定幅度的时频调制信息序列, 用时频调制信息序列在不同时隙和频点的排列组合来表示秘密信息, 通过控制时频调制信息序列的幅度将秘密信息隐藏到载体音频中, 不被人耳所察觉. 实验结果表明, 该算法可应用在不同的调频 FM 广播信道上, 并可保持 96% 以上的信息恢复率.

关键词: 音频信息隐藏; 时频调制; 信息隐藏技术; FM 广播信道

中图分类号: TP309.7 **文献标识码:** A **文章编号:** 0490-6756(2017)02-0298-05

Audio information hiding algorithm based on time-frequency modulation method applied in broadcast channel

LI Dan, GAO Yong

(College of Electronics and Information Engineering, Sichuan University, Chengdu 610065, China)

Abstract: A lot of audio information is contained in Broadcast channel, which provides a broad space for information hiding application, but effective audio hiding algorithm rarely shows up. In view of the above problems, a kind of Time-Frequency modulated audio information hiding algorithm is put forward. Time-Frequency modulation technology possesses strong interference-robusted characteristics, which results in its strong applicability in the broadcast channel. In the algorithm, certain range of Time-Frequency modulation information sequence is embedded in the audio carrier, secret information is represented by the permutation and combination of different time slots and frequency points of the Time-Frequency modulation information sequence. By controlling the amplitude of the Time-Frequency modulation information sequence, the secret information is hidden into the audio carrier, which is not detected by the human ear. The experimental results show that the proposed algorithm can be applied to different FM broadcast channels, and can maintain the information recovery rate of over 96%.

Keywords: Audio information hiding; Time-Frequency modulation; Information hiding; FM broadcast channel

1 引言

信息隐藏技术^[1,2]是近些年研究的热门课题, 在诸多领域有着重要应用, 尤其在保密通信、版权保护以及个人信息保护方面得到越来越多的关注.

信息隐藏技术需要信号处理、密码学、统计学以及生物科学作为基础, 随着新技术的不断应用, 这门学科也蓬勃发展起来. 音频信息隐藏技术作为信息隐藏技术的一个重要组成部分, 其原理是将秘密信息隐藏于音频载体中, 不被第三方觉察, 从而达到

收稿日期: 2016-05-17

基金项目: 四川大学科研项目(JG2016013)

作者简介: 李丹(1990-), 男, 安徽阜阳人, 硕士生, 研究方向为语音信号处理. E-mail: scdxdzld@163.com

隐蔽通信的目的. 现有音频信息隐藏算法有很多种, 具有代表性的有: 最不重要位 (LSB) 隐藏算法^[3]、回声隐藏算法^[4,5]、扩频隐藏算法^[6,7]、基于离散余弦变换域 (DCT) 隐藏算法^[8,9]、基于离散小波变换域 (DWT) 隐藏算法^[10,11] 等. 这些算法各有其优缺点^[12]: 最不重要位隐藏算法具有很高的信息隐藏容量, 但是其鲁棒性太差; 回声隐藏算法相比于最不重要位隐藏算法, 具有较好的鲁棒性和透明性, 它的缺点是隐藏容量较小; 扩频隐藏算法鲁棒性较好, 但是在透明性与隐藏容量上需要取舍; 基于离散余弦变换域和基于离散小波变换域的隐藏方法其透明性和鲁棒性不能兼顾, 且操作复杂, 实现难度较高.

广播是被广泛采用的通信手段, 具有传播范围广、传播速度快、穿透能力强等特点, 相对于电视、报纸和互联网, 广播的成本要低的多. 广播信道中包含大量的音频信号, 这给音频信息隐藏提供了广阔的空间. 在广播信道上音频信息隐藏的应用范围除了隐蔽通信外, 还有一个很重要的应用——广播的安全播出.

如今, 广播安全播出的重要性日益突出, 可以通过音频信息隐藏在电台信号中隐藏一段“防伪标识”, 依靠这个标识来鉴定广播信号的来源可靠性.

广播信道上因存在多径衰落效应以及其它干扰, 现有的音频信息隐藏算法在其上应用效果不佳, 现阶段关于广播信道上的语音隐藏问题的研究较少. 针对上述问题, 本文提出了一种采用时频调制技术的音频信息隐藏算法, 时频调制技术利用不同时隙发送不同频率的信号, 使用时隙和频率的排列组合来表示信息^[13], 其在抗衰落和抗多径方面表现较好, 在广播信道上具有较强的适用性.

2 人耳听觉感知特性

音频信号是一种复合音, 它由很多频率成分的谐波组成. 对于频率不同的声音, 人耳具有不同的听辨灵敏度. 用“听阈”来表示人耳刚刚能够听到的声音的强度, 人耳在安静环境下能够感觉到的听阈与频率成非线性关系, 它们的关系式^[14]如式(1)所示.

$$T(f) = 3.64(f/1000)^{-0.8} - 6.5e^{-0.6(f/1000-3.3)^2} + 10^{-3}(f/1000)^4 \text{ (dB SPL)} \quad (1)$$

在一个较强的声音附近, 相对较弱的声音将不被人耳所察觉, 即被强音所掩蔽, 这种现象被称为人耳的“掩蔽效应”. 较强音被称为掩蔽者, 较弱音被称为被掩蔽者. 掩蔽效应分为两种, 即同时掩蔽 (Sim-

ultaneous Masking) 和异时掩蔽 (Non-simultaneous Masking). 掩蔽现象发生在掩蔽者和被掩蔽者同时存在时, 称作同时掩蔽, 也称作频域掩蔽, 声音能否被听到取决于它的频率和强度. 异时掩蔽效应发生在掩蔽者和被掩蔽者不同时存在时, 也称作时域掩蔽. 若掩蔽效应发生在掩蔽者之前的某段时间, 称作前掩蔽; 若掩蔽效应发生在掩蔽者结束之后的某段时间, 则称作后掩蔽.

3 时频调制音频隐藏算法

时频调制音频隐藏算法原理是基于人耳的听觉感知特性, 在音频信号中嵌入时频调制序列, 根据序列嵌入位置音频信号的强度自适应调节时频调制序列的幅度, 从而不被人耳所察觉. 时频调制序列以时隙和频率的组合来表示信息, 该算法可以表示为

$$x(k) = s(k) + m(k) + q(k) \quad (2)$$

式(2)中, $x(k)$ 为嵌入了秘密信息的音频信号; $s(k)$ 为预处理后的音频载体; $m(k)$ 为时频调制信息序列; $q(k)$ 为同步标志信号. 对时频调制音频隐藏算法分为四个部分进行论述: 时频调制信息序列在音频载体中嵌入位置的选择、时频调制信息序列的设计、音频载体预处理、同步标志信号的设计.

3.1 时频调制信息序列在音频载体中嵌入位置的选择

本文选用了三种常用的音频采样率 8 KHz、16 KHz 以及 44.1 KHz 对秘密信息的嵌入频率范围进行论述. 对于这三种不同采样率的音频载体, 选取不同的嵌入频率范围.

对于采样率为 8 KHz 的音频信号, 音频中只能嵌入频率为 0~4 KHz 的时频调制信息序列, 若嵌入音频信号的低频部分, 音频载体的低频部分能量分布较强, 对接收端检测影响较大, 信息恢复率较低, 而在高频部分能量分布较少, 对检测影响较小, 因此对于采样率为 8 KHz 的音频载体, 隐藏信息的嵌入范围可选择在 3.4~4 KHz.

对于采样率为 16 KHz 的音频载体, 可嵌入的隐藏信息频率范围为 0~8 KHz. 音频信号在 6~8 KHz 范围能量分布较弱, 对秘密信息的恢复影响小, 且人耳对这一频率范围“听阈”较高, 隐蔽效果好, 故对于采样率为 16 KHz 音频载体, 隐藏信息的嵌入范围可选择在 6~8 KHz.

对于采样率为 44.1 KHz 的音频信号, 可嵌入的隐藏信息频率范围为 0~22.05 KHz. 一般情况下

广播电台 FM(Frequency Modulation)调频广播发射设备能对 80~15 KHz 频率范围的信号进行有效发射,为了确保含密音频信号能够在 FM 信道有效传播,故对于采样率为 44.1 KHz 的音频载体,隐藏信息的嵌入范围可选择在 10~15 KHz,此频率范围音频信号能量分布较少,人耳的“听阈”较高。

3.2 时频调制信息序列的设计

时频调制信息序列包含 L 种不同频率的正弦信号,即在音频载体中选用了 L 个嵌入频点. 把时频调制信息序列的每 N 个数据点作为一个时隙,把五个时隙作为一组. 每一种频率的正弦信号在一组时隙中的两个时隙存在,这样设计可使该序列具有一定的检错能力,按照上述设定,一种频率的正弦信号在一组时隙中的位置将会有十个不同的排列组合,用这十种不同的位置来表示一个 0~9 的十进制数,具体的定义如表 1 所示,表中的位置 (α, β) 表示该频率正弦信号在一组时隙的第 α 和第 β 时隙存在. 一组时隙的一个嵌入频点可以表示一个一位非负的十进制数, L 个嵌入频点就可以表示一个 L 位的非负的十进制数。

表 1 一组时隙的单个频点上正弦信号的位置与其对应的十进制数

Tab. 1 The position and corresponding decimal number of sinusoidal signal at single frequency point of a set of time slots

位置	(1,2)	(1,3)	(1,4)	(1,5)	(2,3)
代表数字	0	1	2	3	4
位置	(2,4)	(2,5)	(3,4)	(3,5)	(4,5)
代表数字	5	6	7	8	9

时频调制信息序列 $m(k)$ 可用式(3)来表示

$$m(k) = \sum_{l=1}^L T_l(k) \quad (3)$$

式(3)中, $T_l(k)$ 表示第 l 个频点上的时频调制信息序列, $T_l(k)$ 表示为

$$T_l(k) = \sum_{m=1}^M \sum_{i=1}^5 \Delta(l, m, i) \Phi_l(m, i, k) \quad (4)$$

式(4)中, M 为时频调制信息序列的组数,即十进制秘密信息的个数; $\Delta(l, m, i)$ 有两个取值:“0”和“1”,若 $\Delta(l, m, i) = 1$,则表示频率为 f_l 的正弦信号在时频调制信息序列的 m 组第 i 时隙存在,若 $\Delta(l, m, i) = 0$,则表示不存在,具体取值由该组中隐藏的数值来决定. $\Phi_l(m, i, k)$ 表示为

$$\Phi_l(m, i, k) = \sum_{k=1}^N a(l, m, i) \sin(2\pi f_l k) \quad (5)$$

式(5)中, $a(l, m, i)$ 为频率为 f_l 的正弦信号在时频调制信息序列的 m 组第 i 个时隙的信号幅度,其值根据该正弦信号的频率与其在音频载体中嵌入位置的能量来确定,经多次实验 $a(l, m, i)$ 按照式(6)进行选取,掩蔽效果较强,不易被第三方察觉。

$$a(l, m, i) = \begin{cases} 1.2 \times 10^{-5} (T(f_l) - T_m) + 3 \times 10^{-4}, & E(m, i) > E_0 \\ 1.1 \times 10^{-5} (T(f_l) - T_m) + 3 \times 10^{-4}, & E_0 \leq E(m, i) < 0.5E_0 \\ 10^{-5} (T(f_l) - T_m) + 3 \times 10^{-4}, & E(m, i) \leq 0.5E_0 \end{cases} \quad (6)$$

式(6)中, T_m 为(1)式中的 $T(f)$ 最小值; $E(m, i)$ 为时频调制信息序列对应嵌入时隙内音频载体的平均能量; E_0 为音频载体总的平均能量。

3.3 音频载体预处理

为了增加时频调制信息序列在接收端的检测准确性,在时频调制信息序列嵌入之前对音频载体进行带阻滤波处理,采用的滤波器为 512 阶最小均方带阻滤波器,阻带为时频调制信息列嵌入频点 $f_l (l=1, 2, \dots, L)$ 为中心的 10 Hz 频率范围,因为音频载体中该频率范围的能量分布较弱,故带阻滤波对载体音频影响很小,人耳对滤波前后的音频载体很难察觉到不同。

3.4 同步标志信号的设计

同步标志对于含密音频中隐藏信息的准确提取至关重要,如果不能检测到同步标志,秘密信息提取就无法实现. 基于人耳的异时掩蔽效应,我们设计了一个长度为 400 个数据点的同步标志信号,其由频率依次相差为 50 Hz、位于 2.1 节中所讨论的嵌入频率范围的五个正弦信号构成. 利用人耳的后掩蔽效应,将同步标志信号嵌入在一段能量大于 E_0 的音频后. 同步标志信号嵌入到音频载体后,记录该位置与时频调制信息序列嵌入起始点的距离 R , R 为正表示其位于秘密信息嵌入起始点之前,为负则表示其在秘密信息嵌入起始点之后. 在检测时只要在相应的五个频点同时检测到正弦信号的存在,就可以确定该位置就是同步标志信号所在的位置。

4 提取算法

接收方提取秘密信息需要的密钥有:同步标志信号的嵌入频点、持续长度;同步标志信号与时频调制信息序列嵌入起始点的距离 R ; 时频调制信息序列的嵌入频点;时频调制信息序列的组数 M ;

一个时隙包含的数据点数 N .

提取过程采用信号的相关性进行检测,相关函数反映了两个信号的相似程度,定义如下.

$$r_{yz}(m) = \sum_{k=-\infty}^{+\infty} y(k)z(k+m) \quad (7)$$

式(7)中, $r_{yz}(m)$ 表示 $y(k)$ 与 $z(k)$ 的相关函数在时刻 m 时的值,等于将 $z(k)$ 左移 m 个数据点之后再与 $y(k)$ 相乘再累加.

设接收方接收到含密音频为 $x(k)'$, 提取算法的步骤如下.

Step 1 让 $x(k)'$ 经过一个 512 阶的最小均方带通滤波器,滤除同步标志信号所在频率范围以外的音频成分. 设滤波后得到的信号为 $x_1(k)'$, 分别求出 $x_1(k)'$ 与同步标志信号中的五个正弦信号的相关函数,将求得的五个相关函数相加,得到和函数最大值所在的位置,该位置就是同步标志信号的起始位置.

Step 2 根据检测出来同步标志信号的位置以及其与时频调制信息序列嵌入起始点的距离 R , 可以确定时频调制信息序列的嵌入起始位置,并在其后截取长度为 $5M \times N$ 个数据点的音频信号 $x_2(k)'$. 让截取的音频信号经过 512 阶的最小均方带通滤波器,通带范围为时频调制信息序列嵌入频率范围,将滤波后得到的信号 $x_3(k)'$ 均分为 $5M$ 个时隙,这样每个时隙都与时频调制信息序列中的时隙一一对应. $x_3(k)'$ 的每个时隙的信号分别与频率为 $f_1(l = 1, 2 \dots L)$ 、长度为 N 个数据点的正弦信号求相关函数,并将每个相关函数的最大值依次记录在一个 $L \times 5M$ 数组 ψ 的第 l 行中.

Step 3 对 ψ 每一行的 $5M$ 个数据等分成 M 组,记下每一组的最大的两个数在组中位置,对照表 1 译出该组隐藏的数字. 结合该行所对应的权重,就可以得到隐藏的 M 个 L 位非负的十进制秘密信息.

5 仿真结果及对比实验分析

实验选取用的时频调制信息序列包含的频率个数 $L = 3$, L 取值过大会对隐藏的透明性和信息恢复率造成影响,过小会降低信息的隐藏容量. 每个时隙的采样点个数 $N = 3000$, N 的取值过大会降低信息的隐藏容量,过小则会对信息的恢复率有所影响.

5.1 透明性分析

采用分辨率为 16 bit、采样率为 8 KHz 的音频作为载体,载体音频信号、含密音频信号的时域波形分别如图 1 和图 2 所示,通过对比和试听可以直

观发现,两者有很高的相似性.

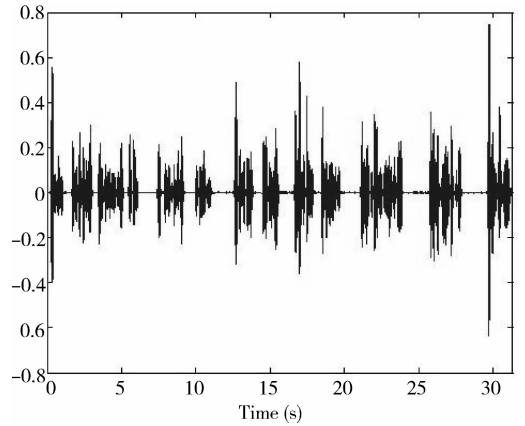


图 1 音频载体

Fig. 1 Audio carrier

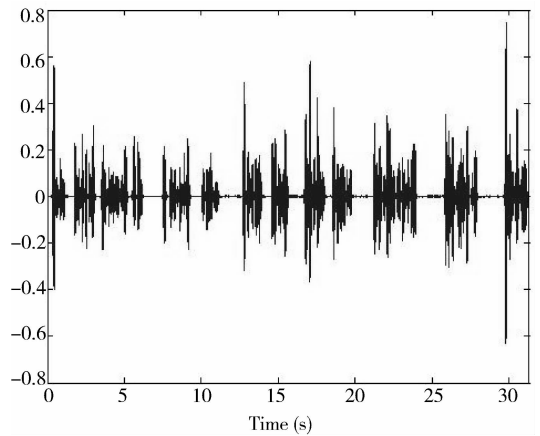


图 2 含密音频

Fig. 2 Stego audio

图 1 和图 2 直观地说明了算法的透明性,在客观上,用信噪比(SNR)来衡量嵌入时频调制信息序列后载体的失真,定义为

$$SNR = 10 \log_{10} \frac{\sum_{d=1}^D h^2(d)}{\sum_{d=1}^D (h'(d) - h(d))^2} \quad (8)$$

其中, $h(d)$ 表示音频载体; $h'(d)$ 表示含密音频信号; D 为音频载体的长度; SNR 的值越大,说明算法的透明性越好. 由该算法计算出的载体信息的 SNR 为 30.49 dB,客观上说明该算法具有很好的透明性.

5.2 隐藏容量

不同采样率条件下,由于音频载体每秒包含的采样点数也不同,其信息隐藏容量也不相同,表 2 给出了 8 KHz、16 KHz、44.1 KHz 三种采样率在 $L = 3$ 、 $N = 3000$ 条件下音频载体的隐藏容量.

表 2 不同采样率下的隐藏容量

Tab. 2 Hiding capacity under different sampling rates

采样率	8 KHz	16 KHz	44.1 KHz
隐藏容量	5.315bit/s	10.630bit/s	29.299bit/s

隐藏容量还与频点的个数 L 、每个时隙包含的采样的数据点 N 有关,增加 L 以及减少 N 都可使隐藏容量增加,但是接收端的信息恢复率也会有所下降。

5.3 对比试验

将时频调制隐藏算法分别与前后向双核回声隐藏算法、基于离散余弦变换(DCT)域的扩频隐藏算法进行了信息恢复率对比实验,实验选取分辨率为 16 bit,采样率为 8 KHz 的 50 组音频信号作为载体。FM 广播发射频点选取的三个空闲频点,分别为 87.6 MHz、94.9 MHz、106.6 MHz。

时频调制算法参数选取为嵌入频点个数 $L = 3$ 、每个时隙包含的采样点数 $N = 3000$ 。嵌入的信息为长度为 100,取值范围为 0~999 的十进制数。

前后向双核回声隐藏算法参数选取为回声核的衰减因子为 0.5,延时 10 ms 表示比特信息“0”,延时 15 ms 表示比特信息“1”,嵌入的信息为长度为 100 的二进制比特流。

基于离散变换域的扩频隐藏算法参数选取为每一帧长度为 256 位,伪随机序列采用长度为 127 的 m 序列。每一帧的离散余弦变换会有 256 个离散余弦系数,选取其中的高 127 位进行隐匿信息嵌入,嵌入因子为 0.05,嵌入的信息为长度为 100 的二进制比特流。

将上述三种算法分别经 FM 广播信道后进行秘密信息提取,得到信息回复率实验结果如表 3 所示。从表 3 可见,时频调制隐藏算法在 FM 广播信道上具有良好的适用性。

表 3 不同 FM 广播发射频点下的信息恢复率

Tab. 3 Information recovery rate under different FM broadcast transmission frequency

发射频点 (MHz)	87.6	94.9	106.6
时频调制隐藏算法	96.45%	96.51%	96.46%
前后向双核回声隐藏算法	85.62%	85.43%	85.54%
基于 DCT 域扩频隐藏算法	74.84%	74.87%	74.93%

6 结 论

本文提出了一种时频调制音频隐藏算法,该算法具有较强的抗信道干扰能力,经对比实验表明该隐藏方法经 FM 广播信道后能保持较高的信息恢复率,具有一定的实用价值。

参考文献:

- [1] Sodipo E. The art of security and information hiding [M]. [s. l.]: Lulu. com, 2008.
- [2] 孔维国,王宏霞,王科人,等. 基于转移概率矩阵的 H.264/AVC 视频帧内预测模式信息隐藏检测算法[J]. 四川大学学报:自然科学版,2014,51(6): 1183.
- [3] Nasef M M. An efficient method to audio steganography based on modification of least significant bit technique using random keys[J]. Int J Comput Netw Inf Secur, 2015, 7: 24.
- [4] Xiang Y, Natgunanathan I, Peng D, et al. A dual-channel time-spread echo method for audio watermarking[J]. IEEE Trans Inf Foren Sec, 2012, 7(2): 383.
- [5] Kim H J, Choi Y H. A novel echo-hiding scheme with backward and forward kernels[J]. IEEE Trans Circuits Syst Vid Tech, 2003, 13(8): 885.
- [6] Shokri S, Ismail M, Zainal N, et al. BER performance of audio watermarking using spread spectrum technique [J]. Procedia Tech, 2013, 11(1): 107.
- [7] 同鸣,郝重阳,刘晓军. 基于听觉掩蔽效应特性的扩频信息隐藏方法[J]. 计算机工程,2006,32(10): 137.
- [8] Hu H T, Hsu L Y. Robust, transparent and high-capacity audio watermarking in DCT domain[J]. Signal Process, 2015, 109(C): 226.
- [9] Charfeddine M, El'Arbi M, Amar C B. A new DCT audio watermarking scheme based on preliminary MP3 study[J]. Multim Tool Appl, 2014, 70(3): 1521.
- [10] Hu H T, Hsu L Y, Chou H H. Variable-dimensional vector modulation for perceptual-based DWT blind audio watermarking with adjustable payload capacity[J]. Digit Sig Process, 2014, 31(5): 115.
- [11] Shirali-Shahreza S, Manzuri-Shalmani M T. High capacity error free wavelet domain speech steganography[C]//2008 IEEE International Conference on Acoustics, Speech and Signal Processing. [s. l.]: IEEE, 2008.
- [12] 谭良,吴波,刘震,等. 一种基于混沌和小波变换的大容量音频信息隐藏算法[J]. 电子学报,2010,38(8): 1812.
- [13] Jiang H J, Tan X, Shen Y H. Performance of space-time time-frequency-shift keying (ST-TFSK) modulation systems [J]. Acta Electron Sinica, 2007, 35(3): 488.
- [14] Painter T, Spanias A. Perceptual coding of digital audio[J]. P IEEE, 2000, 88(4): 451.