doi:10.3969/j.issn.1000-436x.2014.z1.008

# 基于振幅值修改的 wav 音频隐写算法

邹明光[1,2]，李芝棠[1,2,3]
（1. 华中科技大学 计算机科学与技术学院，湖北 武汉 430074；2.下一代互联网接入系统国家工程实验室，湖北 武汉 430074；
3. 华中科技大学 网络与计算中心，湖北 武汉 430074）

**摘 要**：在对不同有效域的算法进行总结和分析的基础上，提出了一种基于振幅值修改的 wav 音频隐写算法。算法将秘密音频通过随机数生成器置乱，通过比较原始音频每个采样点分组中振幅值之间的关系，结合待嵌入的秘密信息位进行嵌入修改，嵌入强度依据密钥进行调节。实验结果表明，该算法具有较大的嵌入容量，不可感知性良好，具有一定的抗隐写分析能力，并可实现盲提取。
**关键词**：振幅值修改；音频隐写；嵌入容量；不可感知性
**中图分类号**：TP 309.2 **文献标识码**：A **文章编号**：1000-436X(2014)Z1-0036-05

# Wav-audio steganography algorithm based on amplitude modifying

ZOU Ming-guang[1,2]，LI Zhi-tang[1,2,3]
(1. School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074，China；
2. National Engineering Laboratory for Next Generation Internet Access System，Wuhan 430074，China;
3. Network and Computing Center, Huazhong University of Science and Technology，Wuhan 430074，China)

**Abstract:** Algorithms in different effective domains are analysed, and a wav-audio steganography algorithm based on amplitude modifying is proposed. The secret audio is scrambled through a random number generator. Amplitude values of each sampling point group are compared in original audio.While amplitude values are modified, the secret information bits are embedded. Embedding strength can be adjusted according to the key. Simulation results demonstrate that hiding capacity of this algorithm is large and good invisibility can be achieved. The anti-steganalysis ability of this algorithm is good and blind extraction can be realized.
**Key words:** amplitude modifying; audio steganography; hiding capacity; invisibility

## 1 引言

隐写（steganography）是指利用人的视听觉系统的掩蔽效应，把秘密信息嵌入到载体中，以实现不为人知的隐蔽通信的信息隐藏技术，又称隐写术。隐写的载体通常选择信息冗余较大的图像、音频和视频等大众媒体。由于隐写术利用公共信道作为传输通道，将秘密信息通过一定的算法嵌入到大众媒体中，不仅隐藏了秘密信息，同时还隐藏了通信的行为，可以实现较为安全可靠的隐蔽通信。

随着隐写技术的发展，音频隐写逐渐成为近年来的热门研究领域和发展方向，该技术通过向音频文件中嵌入秘密信息以达到隐蔽通信的目的。根据算法的有效域不同，音频隐写分为时间域算法、变换域算法、压缩域算法、相位算法、回声算法等。时间域算法在音频的时间域上进行替换，这种方法隐写容量较大，但顽健性较差[1]；变换域算法在傅立叶变换域(FFT)、离散余弦变换域(DCT)、离散小波变换域(DWT)等变换域嵌入数据，这种方法计算相对复杂，但顽健性较好[2~4]；压缩域算法在压缩域嵌入数据，这种方法难度高、隐写容量小[5]；相位算法在音频的相位进行替代嵌入，这种方法隐写容量小[6]；回声算法通过引入回声来嵌入数据，这种方法提取复杂，隐写容量小[7]。算法设计时需要

**收稿日期**：2014-10-14
**基金项目**：国家自然科学基金资助项目（61272407）
**Foundation Item:** The National Natural Science Foundation of China (61272407)

根据适用条件的不同在隐藏容量、不可感知性、顽健性等相互矛盾的条件下寻求合理优化选择。

Wav 格式的音频使用脉冲编码调制（PCM）无压缩编码，采用 Little-Endian 字节顺序存储采样点的振幅值，该特性为音频隐写算法提供了良好的条件。本文以 wav 音频为载体进行分析，提出一种基于振幅值修改的大容量隐写算法。该算法根据每 3 个采样点振幅值之间的关系，动态调整中间采样点的振幅值用于隐藏秘密信息。在增大嵌入容量的同时，能有效控制嵌入信息对音频质量产生的影响，取得了较好的实验效果。

## 2　算法描述

### 2.1　嵌入过程

基于振幅值修改的wav音频算法嵌入过程如图 1 所示。

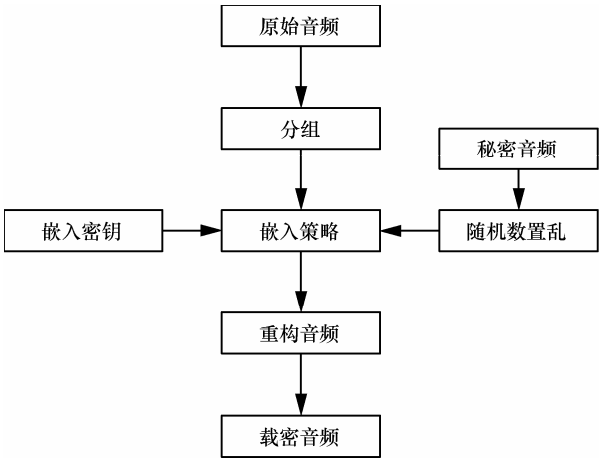


图 1　振幅值修改 wav 隐写算法嵌入过程

具体嵌入步骤如下。

1) 秘密音频随机数置乱

选择整数 $n$ 作为随机数生成器的种子，将秘密音频经过随机数生成器的处理进行置乱。这样不但能给秘密音频加密，增强安全性，而且采用打乱音频采样点排列顺序，嵌入位置随机，有利于增强抗隐写分析能力。

2) 分组计算原始音频的振幅值

以原始音频每连续的 3 个采样点作为分组，分别计算振幅值 $f1$、$f2$、$f3$。

3) 嵌入策略

在原始音频的每一采样点分组嵌入一个比特秘密信息。通过式(1)对每一组中的 $f2$ 和 $\frac{f1+f3}{2}$ 进行比较，根据差值比较的结果，结合嵌入秘密信息是 0 还是 1，采取修改 $f2$ 的方法实现在时域嵌入秘密信息。

$$\Delta = f2 - \frac{f1+f3}{2} \tag{1}$$

选择整数 $\alpha$ 作为嵌入密钥，密钥与嵌入强度成正比，与不可感知性成反比。密钥越大，嵌入到原始音频的信息越多，即嵌入强度越大，此时载密音频的信噪比就越低。密钥用于控制秘密信息的嵌入容量，以保证嵌入后载体的不可感知性。只有当满足条件式(2)时，算法才进行嵌入操作。

$$0 < |\Delta| \leqslant \alpha \tag{2}$$

当嵌入信息 1 时，如果 $\Delta > 0$，则不改变 $f2$ 的值，如果 $\Delta < 0$，则把 $f2$ 修改为 $\frac{f1+f3}{2}$ 的上限减 1，也就是比 $\frac{f1+f3}{2}$ 小的最大整数。$\Delta = 0$，该分组不做嵌入，即

$$f2 = \begin{cases} f2, \Delta > 0 \\ \left|\frac{f1+f3}{2}\right| - 1, \Delta < 0 \end{cases} \tag{3}$$

当嵌入信息 0 时，如果 $\Delta < 0$，则不改变 $f2$ 的值，如果 $\Delta > 0$，则把 $f2$ 修改为 $\frac{f1+f3}{2}$ 的下限加 1，也就是比 $\frac{f1+f3}{2}$ 大的最小整数。$\Delta = 0$，该分组不做嵌入，即

$$f2 = \begin{cases} \left|\frac{f1+f3}{2}\right| - 1, \Delta > 0 \\ f2, \Delta < 0 \end{cases} \tag{4}$$

式(3)和(4)用于控制载密音频的质量，使得秘密信息的嵌入对原始音频的影响最小。

4) 音频重构

将修改后的振幅值写回文件，得到载密音频信号。

### 2.2　提取过程

秘密音频的提取步骤如下。

1) 分组计算载密音频的振幅值

以载密音频每连续的 3 个采样点作为分组，分别计算载密音频分组的振幅值 $f1$、$f2$、$f3$。

2) 提取策略

对每一组中的 $f2$ 和 $\frac{f1+f3}{2}$ 进行比较，计算差

值，即

$$\Delta = f2 - \frac{f1 + f3}{2} \tag{5}$$

$\alpha$ 是嵌入密钥，只有当满足条件(6)时，算法才进行提取操作。

$$0 < |\Delta| \leqslant \alpha \tag{6}$$

根据 $\Delta$ 的值来提取秘密信息，如果 $\Delta > 0$，则提取秘密信息为 1；如果 $\Delta < 0$，则提取秘密信息为 0。

$$X = \begin{cases} 1, \Delta > 0 \\ 0, \Delta < 0 \end{cases} \tag{7}$$

3) 秘密音频重构

$n$ 是随机数生成器的种子，将秘密音频按照随机数生成器的生成顺序进行重组并写入文件，得到秘密音频信号。

可见，采用该算法嵌入的秘密音频的提取算法简单，并且在提取秘密信息时不需要提供原始音频数据，实现了秘密音频的盲提取。

## 3　实验及结果

### 3.1　实验环境

本算法的实验是在 Eclipse4.3 Java 开发平台和 MATLAB R2012b 上完成，软件环境是 Windows 7。原始音频文件选用采样频率为 44.1 kHz，量化精度为 16 bit，时长为 1 min 2 s，大小为 10.5 MB（载体 1）以及时长为 32.391 s，大小为 5.45 MB（载体 2）的 2 个双声道 wav 波形文件。秘密音频文件选用采样频率为 8 kHz，时长为 1.94 s，大小为 3.03 KB 的 amr 音频文件。随机数种子 $n$ 取 3。

### 3.2　实验结果

图 2 给出了嵌入密钥对嵌入容量影响的对比。嵌入密钥 $\alpha$ 与嵌入容量成正比，其他条件相同时，$\alpha$ 越大，符合嵌入条件的嵌入位置就越多，从而能有更大的嵌入容量。在 $\alpha = 50$ 时，时长为 1 min 2 s 的载体 1 能够嵌入约 44 KB 的秘密信息，时长为 32.391 s 的载体 2 能够嵌入约 24 KB 的秘密信息。实验数据可以看出，本算法的嵌入容量较大。

通常用信噪比（SNR）来对算法不可感知性进行评估，信噪比越大表明算法的不可感知性越好。原始音频 $M$ 的采样点个数为 $N$，嵌入秘密音频后的载密音频为 $M'$，则信噪比为

$$SNR = 10\lg \frac{\sum_{i=1}^{N} M(i)^2}{\sum_{i=1}^{N} \left[ M(i) - M'(i) \right]^2} \tag{8}$$

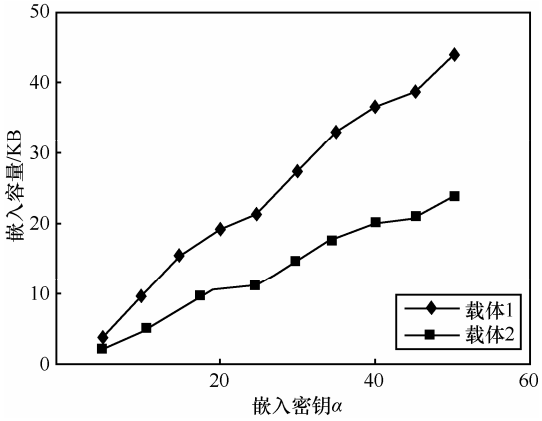


图 2　密钥与嵌入容量的关系

图 3 给出了实验中载体在不同嵌入容量下信噪比的变化情况。在同等的信噪比条件下，载体 2 的嵌入容量约为载体 1 的一半，这是因为载体 2 的原始文件大小也约为载体 1 的一半。载体 1 即使在嵌入 40 KB 的秘密信息时，嵌入后的音频信噪比基本都在 50 dB 以上，保持了较高的信噪比。从实验中可以看出，本算法的不可感知性较好。

位错误率（BER）也常被用来对算法不可感知性进行评估，位错误率描述的是嵌入秘密信息后的载密音频与原始载体对比，修改的比特数与原始载体比特数的比值。原始音频 $M$ 的比特数为 $N$，嵌入秘密音频后的载密音频为 $M'$，则位错误率为

$$BER = \frac{1}{N} \sum_{i=0}^{N-1} \begin{cases} 1, M'(i) \neq M(i) \\ 0, M'(i) = M(i) \end{cases} \tag{9}$$

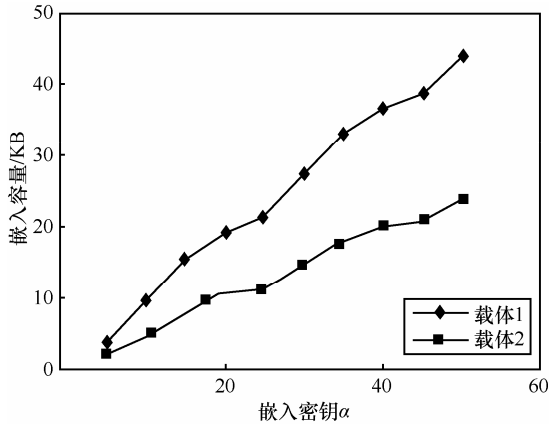


图 3　嵌入容量与信噪比 *SNR* 的关系

表 1 给出了实验中载体 1 和载体 2 在不同嵌入密钥 $\alpha$ 下位错误率的变化情况。使用本算法对音频载体进行嵌入，嵌入后的载密音频与原始音频相比，位错误率保持在一个较低的数值。实验统计的数据表明，本算法能有效控制嵌入信息对音频质量产生的影响。

**表 1　不同嵌入密钥 $\alpha$ 下的位错误率**

| 嵌入密钥 | 载体 1 *BER* | 载体 2 *BER* |
|---|---|---|
| 5 | 0.0372 | 0.0397 |
| 10 | 0.1171 | 0.1212 |
| 20 | 0.2944 | 0.308 |
| 35 | 0.5641 | 0.5899 |
| 50 | 0.7854 | 0.8235 |

实验中，当载体 1 的嵌入密钥 $\alpha$ 取值为 5 时，嵌入容量为 3.10 KB，刚好能把大小为 3.03 KB 的 amr 秘密音频文件完整嵌入，计算嵌入后的信噪比为 85.397 9 dB，嵌入秘密音频前后的音频信号时域波形如图所示，其中，图 4(a)为原始音频信号时域波形，图 4(b)为采用本算法嵌入秘密音频后的音频信号时域波形。从图中也可直观得看出，嵌入秘密音频后对音频文件的改变不大。

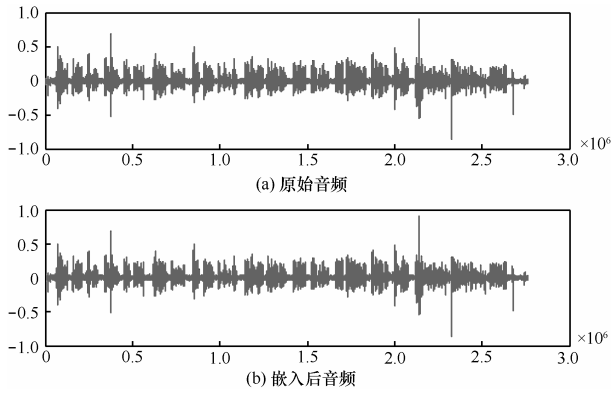


图 4　音频嵌入前后时域波形

对使用本算法处理后的音频进行提取秘密信息操作，将提取出的秘密信息与原始秘密信息逐位对比计算提取正确率。使用 20 段不同种类的音频载体在不同的嵌入率下进行嵌入提取实验，统计得到的提取正确率如表 2 所示。实验表明，使用本算法对音频载体进行嵌入时，无论嵌入率为多少，只要秘密信息不超过最大嵌入容量，算法都可以把秘密信息正确提取出来。

**表 2　秘密信息正确提取率**

| 嵌入率/% | 正确提取率/% |
|---|---|
| 5 | 100 |
| 10 | 100 |
| 30 | 100 |
| 50 | 100 |
| 100 | 100 |

隐写分析方法主要是分析载体是否嵌入秘密信息，估计秘密信息的嵌入量和嵌入位置，甚至秘密信息是什么。本文结合通用音频隐写分析、$\chi^2$ 分析和 RS 分析的特点，对本算法进行抗隐写分析性能分析。

通用音频隐写分析的一般思路是选取检测对象的特征统计量，利用分类器从大量数据中构造检测模型，对检测模型进行训练寻找最佳阈值。然而这类隐写分析需要对分类器进行大量训练，实用性不强；通用特征的选取和阈值大小的选取也较困难，造成检测的复杂度高，准确性低，而且本算法随机嵌入的特点使得这种隐写分析基本无效。

针对时域隐写常用的隐写分析方法有：$\chi^2$ 分析和 RS 分析。

使用 $\chi^2$ 分析时，如果载体使用 LSB 算法进行满嵌入，则相邻采样点直方图 $x_{2i}$ 和 $x_{2i+1}$ 的值会比较接近，根据这个特点可以计算出原始载体嵌入秘密信息的可能性。本算法通过嵌入密钥 $\alpha$ 来控制嵌入强度，使得隐写后载体采样点直方图与原始载体的直方图差别不大。图 5 给出了实验中载体隐写前后音频采样点的直方图，从图中也可以看出，隐写前后直方图基本保持不变。因此 $\chi^2$ 分析不能检测出是否嵌入了秘密信息。

使用 RS 分析时，在采用本算法的情况下，$F_1$ 翻转和 $F_{-1}$ 翻转对音频的影响程度基本一样，有 $R_M \approx R_{-M}$，$S_M \approx S_{-M}$。因此用 RS 分析也不能检测出是否嵌入了秘密信息。从实验和分析可以得知，本算法具有一定的抗隐写分析能力。

## 4　结束语

本文提出了一种基于振幅值修改的时域 wav 音频隐写算法，在嵌入秘密时可根据需要选取不同的密钥以嵌入不同大小的秘密音频。该算法运算复杂度低，嵌入和提取秘密音频速度快，且可实现秘密音频的盲提取。实验结果表明，该算法具有较大的

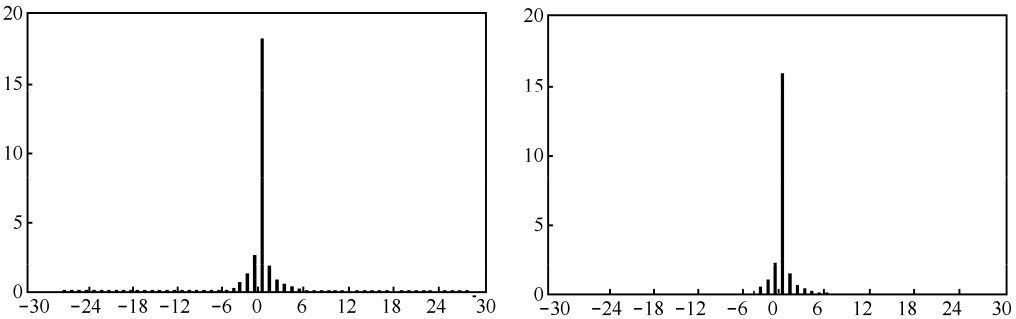


图 5　隐写前后音频采样点直方图的对比

嵌入容量，不可感知性良好，且具有一定的抗隐写分析能力。下一步的研究方向是加强算法对抗各种攻击的顽健性，利用纠错编码、多秘密共享等技术对待嵌入秘密音频进行结构化预处理，进行重复冗余嵌入，以提高秘密音频信息在传输过程中经受各种攻击后的存活率。

## 参考文献：

[1] 黄仿元.基于 LSB 的数字水印算法及 MATLAB 实现[J].计算机应用，2008(2):67-73.
HUANG F Y, Digital watermarking based on LSB and implementation of MATLAB[J]. Journal of Computer Application, 2008(2):67-73.

[2] BHAT K V, SENGUPTA I, DAS A. An adaptive audio watermarking based on the singular valuedecomposition in the wavelet domain[J]. A Review Journal of Digital Signal Processing, 2010, 20(6):1547-1558.

[3] 彭宏，王珣，王卫星.基于音频特征的多小波域水印算法[J].计算机研究与发展，2010,47(2):216-222.
PENG H, WANG X, WANG W X, Audio watermarking approach based on audio features in multiwavelet domain[J]. Journal of Computer Research and Development, 2010,47(2):216-222.

[4] 黄雄华,王宏霞.一种 DCT 域自适应音频水印算法[J].计算机应用研究，2009, 26(8): 2989-2991.
HUANG X H, WANG H X, Adaptive audio watermarking algorithm on DCT domain[J]. Application Research of Computer, 2009, 26(8): 2989-2991.

[5] YAN D Q. Quantization step parity-based steganography for MP3 Audio[J]. FudamentaInformaticae, 2009, 97(1/2):1-14.

[6] BENDER W, GRUHL D, MORIMOTO N. Techniques for data hiding[J]. IBM System Journal, 1996,35(3):313-336.

[7] 张一帆，蒋天发.基于时域扩展回声隐藏的数字音频水印研究[J].计算机与工程，2008,44(31):119-124.
ZHANG Y F, JIANG T F, Research on time spread echo hiding for audio watermarking[J]. Computer Engineering, 2008,44(31):119-124.

## 作者简介：

**邹明光**（1986-），男，江西抚州人，华中科技大学硕士生，主要研究方向为网络与信息安全。

**李芝棠**（1951-），男，湖北监利人，华中科技大学教授、博士生导师，主要研究方向为计算机系统结构、网络与信息安全、P2P 网络。