

WAV 波形文件的结构及其应用实践^①

郭兴吉

(绵阳师范学院计算机科学与工程系 绵阳 621000)

摘 要 本文从音频数字化的角度, 介绍计算机处理音频信息的几种方法, 并以此为基础, 详细分析了目前计算机应用领域广泛使用的 WAV 波形文件结构, 并提出了基于 WAV 的几种典型应用模式。

关键词 音频 数字化 声道 采样频率 文档结构

中图法分类号 TP391.42 文献标识码 A

Structure and Application pf WAV File

GUO Xing-ji

(Dept.of Computer science and Technology,Mianyang Normal University,MianYang, 621000)

Abstract As regards audio digitization, the researcher introduced several computer process means of audio information, and then, presented application patterns based on WAV file after analyzing thoroughly the structure of the widely used WAV File in computer application field.

Key words Audio Frequency/digitization/sound Track/Acquisition Frequency/File Structure

一、音频数字化技术

音频数字化, 就是用数字的形式来记录音频信息的一种方式。最常见的数字化方式是采用音频“脉冲编码调制”法, 即 PCM(Pulse Code Modulation)。其数字化原理如下: 首先我们考虑声音(一种典型的音频信息)经过麦克风, 转换成一连串随电压变化的信号, 如图 1 所示。图中的横坐标单位为时间秒, 纵坐标为电压大小。在进行 PCM 编码时, 对该波形以等时距分割(假设用每间隔 0.01 秒进行波形分割), 如图 2 所示。

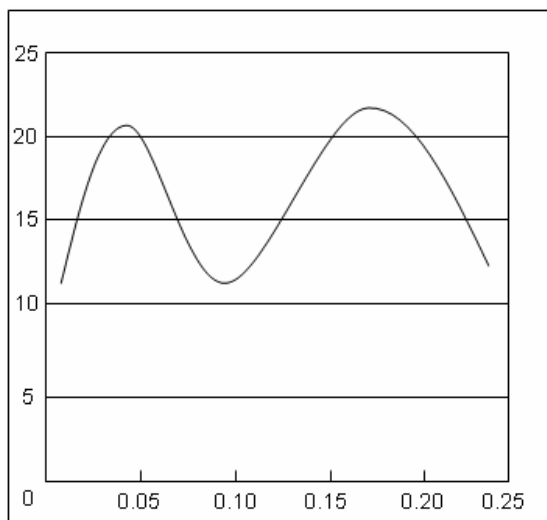


图 1 音频波形示意图

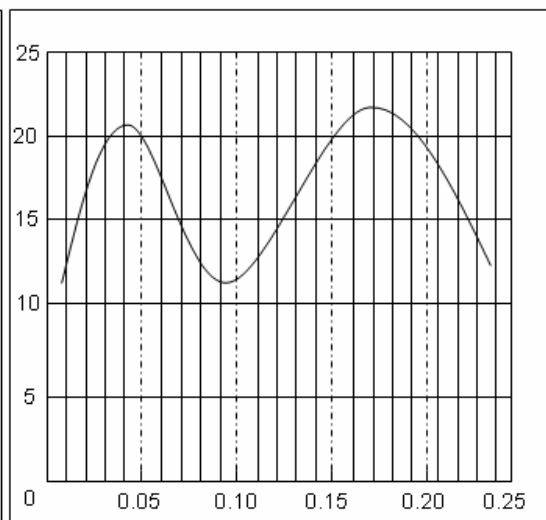


图 2 音频波形数字化示意图

在图 2 中, 把时间分割线与信号波形图的交叉点的坐标位置记录下来, 即可得到如下数字序列(以“时间: 值”形式表示): (0.01,12.00)、(0.02,16.78) (0.21,18.11)、(0.22,16.00)。到此, 我们已经把这个波形以数字形式记录下来了。由于我们已经知道采集数字信息时使用的的时间间隔是固定的 0.01 秒, 因此在记录生成的数字序列中可以去掉时间间隔值, 而只记录纵坐标的值, 得到新的数字序列是: 12.00 16.7818.11 16.00, 这一数字序列就是将以

^① 四川省教育厅科研项目(《智能电视广告监播系统》)经费资助

上波形信号数字化的结果(在以上范例中,我们的采样频率是 100Hz(1/0.01 秒)。事实上计算机常用的 WAV 文件的内容与这个例子非常类似。

从这个例子中可以看出,要想使采集得到的数据尽可能逼近真实波形,只有减小采集的时间间隔(即增大采样频率),然而,增大采样频率又意味着增加采集得到的数据量。同时,无论怎样减小采集的时间间隔,采集得到的数字跟实际的信号波形总是有误差,因此音频信息数字化无法完全记录原始信号,我们称这个数字化造成的波形失真为量化失真。

在目前,计算机中的音频处理广泛采用的就是 PCM 编码。

二、WAV 格式分析

如前所述, WAV 音频格式就是采用 PCM 编码体系对音频信息进行编码存储的。

1. 影响 WAV 文档的因素

采样频率: 每秒钟采集音频数据的次数。采样频率越高, 音频保真度越高。目前计算机广泛配置的 16 位声卡, 使用的采样频率通常包括 11025Hz、22050Hz、44100Hz 和 48000Hz 四种, 其中, 采用 11025Hz 采样的声音效果相当于电话声音的效果; 采用 22050Hz 采样的声音效果相当于 FM 调频广播的效果; 采用 44100Hz 采样的声音效果相当于 CD 声音的效果。

采样位数(振幅采样精度): 即采样值或取样值, 是用来衡量声音波动变化的一个参数, 也是声卡的分辨率。它的数值越大, 分辨率也就越高, 发出声音的能力越强。目前计算机中配置的 16 位声卡的采样位数包括 8 位和 16 位两种。

声道数: 有单声道和立体声之分, 单声道的声音只能使用一个喇叭发声(有的声卡也将单声道信息处理成两个喇叭同时输出), 立体声的 WAV 可以使两个喇叭都发声(一般左右声道有分工), 这样更能感受到音频信息的空间效果。显然, 双声道数据还原特性更接近人们的听力习惯, 但采集得到的数据量会增加 1 倍。

2. WAV 文档格式

弄清 WAV 文档格式, 对于灵活应用 WAV 文档具有非常重要的作用。

图 3 是 WAV 文档的格式。从该图中可以看出, WAV 文档主要由“WAV 头”和“WAV 数据表”两部分构成, 下面我们分别就 WAV 文档头和 WAV 数据表结构进行分析。

(1) WAV 头结构

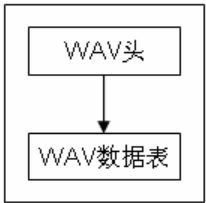


图 3 WAV 文档结构

WAV头结构(44B)			
偏移地址	长度(字节)	数据类型	值或含义
00H	4	Char	格式标记“RIFF” ^①
04H	4	Longint	文件长度 ^②
08H	4	Char	“WAVE”
0CH	4	Char	“fmt”
10H	4	Longint	不定
14H	2	Word	WAV格式类别
16H	2	Word	声道数1或2
18H	3	Longint	采样频率
1CH	3	Longint	数据传输率 ^③
20H	2	Word	块调整值 ^④
22H	2	Word	采样位数
24H	4	Char	标记“data”
28H	4	Longint	音频数据长度 ^⑤

注①: WAVE文档的保存格式标准是RIFF

注②: 文件长度=文件实际长度-8

注③: 数据传输率=声道数×采样频率×采样位数/8

注④: 块调整值=声道数×采样位数/8

注⑤: 音频数据长度=文件长度-头长度

图 4 WAV(44B)头结构示意图

在 Windows 系统中, WAV 头结构有两种, 一种头结构长度为 44B, 另一种头结构长度为 58B, 这两种头结构的 WAV 文档都可以利用播放器(或程序中利用播放器插件)正常播放。图 4 和图 5 分别是这两种 WAV 头的结构示意图。

(2) WAV 音频数据表结构

WAV 音频数据表结构采用“左→右”声道数据交替的方式进行存储。其数据量的大小可利用公式计算得到: 音频数据量=采样频率*声道数*采样秒数*采样位数/8 (字节)

特别地: 音频数据以 128(8 位)或 32768(16 位)两个数值为音频波形中线值, 这两个值表示没有声音, 距离该值越远, 表示波形振幅越大, 音量也越大。

3. 音调与音频数据的关系

在音频数据中, 距离音频中线值越远(即音频数据与中线值差的绝对值越大)则音量越宏伟, 在音频数据中, 音调又是如何体现的呢?

仔细分析 WAV 音频数据以及 PCM 编码原理可以发现, 在 WAV 音频数据中, 音调是通过音频数据的变化率来体现的。由此可以推出: 音频数据变化率越大, 则音调越高, 反之越低。

三、WAV 应用实践

弄清楚了 WAV 文档结构, 即可编程实现对 WAV 音频数据进行多种形式的特效处理了。

1. 增加(减小)音量

要增加(减小)音量, 只需要对每一个音频数据进行单独处理后再保存即可。

```
procedure AddandAbate(RATIO:word); //参数 RATIO: 增加或减小音量的百分比
begin
    while not eof(ff1) do begin //ff1: 源文件的变量实例
        read(ff1, x);
        if x>=MIDLINE then x:=x*(1+RATIO) //MIDLINE: 波形中线值
        else x:=x*(1-RATIO);
        write(ff2, x); //ff2: 目标件的变量实例
    end;
end;
```

2. 加快(减慢)音频播放速度

要减慢或加快音频信息的播放速度, 只需要在音频数据序列中的适当位置插入或删除

WAV头结构(58B)			
偏移地址	长度(字节)	数据类型	值或含义
00H	4	Char	格式标记"RIFF"
04H	4	Longint	文件长度
08H	4	Char	"WAVE"
0CH	4	Char	"fmt"
10H	4	Longint	不定
14H	2	Word	WAV格式类别
16H	2	Word	声道数1或2
18H	3	Longint	采样频率
1CH	3	Longint	数据传输率
20H	2	Word	块调整值
22H	2	Word	采样位数
24H	2	Word	暂时未用, 值为0
26H	4	Char	暂时未用, 值为"fact"
2aH	4	Longint	WAV标志, 值为4
2eH	4	Char	暂时未用
24H	4	Char	标记"data"
28H	4	Longint	音频数据长度

图 5 WAV(58B)头结构示意图

- 3 飞思科技产品研发中心 《Delphi6 灵感编程》[M]北京：电子工业出版社,2002.
- 4 大富翁论坛 .Delphi- 多媒体 [EB/OL]. <http://www.delphibbs.com/delphibbs/listq.asp?page=2&type=1&room=16&sort=1>

作者简介：

郭兴吉 男(汉族)，1964 年 7 月生，绵阳师范学院计算机科学与工程系副教授，研究方向：计算机网络技术、多媒体技术。

联系方式：

电话：0816-2201122

E-Mail: xingji_g@yahoo.com.cn

通信地址：四川，绵阳师范学院计算机科学与工程系

邮编：621000