# 课程设计报告

王屹翔 2014040201028

魏涛 2014020905028

1. **摘要**

音高识别在音乐识别、乐器调音、自动记谱、表演辅助等方面有着重要的作用和应用，此次设计旨在根据现有的一些技术手段和处理方法，开发出一种有效的、可行的音高识别系统。在设计中主要通过分隔端点、区分音区以及确定基音频率（基音频率可确定音高）三部分进行音高的识别，利用自相关法和谐波分析法进行基音频率确定，并结合直接法对比识别精度。测试结果显示该识别系统能较为精确的识别出音高（及音符），并在此基础上可进行对一段乐音的自动记谱。

**【关键词】** 音高识别 端点分隔 自相关法 谐波分析法

1. **前言**

音乐学习过程中往往需要知道一段音乐的乐谱，对于只有音频信号的乐音而言乐谱往往不容易得到。因此，本次设计希望得到一个能对音频信号进行较为精确的识别的系统，以便自动记录音乐的乐谱。

十二平均律中每个音都对应一个频率，该识别系统就是根据提取出的频率识别出一段音乐中的每个音。

1. **正文**

**1.设计要求**

识别出某段音乐中的每个音。

**2.系统组成**

音区区分

端点分割

识别基音频率

频率对比

**3.设计原理**

**1）音高**是指各种不同高低的声音，即音的高度，主要由频率（基频）决定。在现实情况里，声音会带有很多谐波，且基频的能量可能比谐波小，这就要求我们不能直接进行FFT找出频率峰值，而是通过其他手段找出信号的基频，并由此来确定音高。

**2）端点分隔**：对信号进行分帧处理后，每一帧内信号频率可近似看为稳定，我们对每帧信号的能量进行判断，分为乐音区、过渡区和静音区，判断标准为阈值H。

**3）基音频率识别：**

**①自相关法**

自相关函数在基音周期的整数倍位置上出现峰值，因此检测峰值的位置就可提取基音周期值。由于短时自相关函数中保留的幅度太多，有许多峰值，可能被检测出来的峰值就会偏离原来峰值的真实位置。因此需要采取中心削波对音乐信号进行预处理来改善自相关函数的性能计算。

**②谐波峰值法**

谐波峰值法是基于快速傅里叶变换(FFT)的分析法，将信号通过FFT变换得到离散的频率谱，最大峰值对应于基音频率。但在实际中，基音峰值的幅度不一定是最大的，因此需要提取峰值频率的最大公因子作为基频。

由于音乐信号的频宽较大，对于音高跨度较大的乐曲，如果乐器谐波比较丰富，采用谐波峰值法就很有可能把二次甚至三次谐波误定为音高。

**4.设计内容**

**1）端点分隔**

将音乐片段信号进行分帧，设置帧长为m，帧移为 n，然后计算每帧的短时能量。假定音乐信号开始前N帧信号为纯背景噪声，初始门限值H 可以定为该 N 帧信号短时能量值的几何平均值的倍，取1至1.5之间的数。如果当前帧判定为噪声，那么门限的更新采用下述方法：

式中的是一个经验值取 0.9，为当前噪声区能量几何平均值的a倍。如果在乐音区发生频率突变，则也是为新的音符。

初始门限设定

是否静音区

门限更新

结束

Yes

No

图1.端点分隔

分帧

计算短时能量

开始

**2）音区区分：**

①计算前1帧的短时能量的平均值得到初始门限值。

②若当前帧的短时能量超过了前面计算出的门限值，且此后连续3帧都超过了该

门限， 即认为该位置为音符的起始点，否则认为处于过渡带。

③若当前帧的短时能量低于门限值，且此后连续3帧都低 于该门限，即认为该

位置为音符的终止点，否则认为处于乐音区。

④由式（\*）重新计算门限。

⑤回到第②步，直到整个音乐信号结束。

结束

No

No

Yes

当前帧能量>门限

门限更新

连续3帧>门限

Yes

开始

计算前1帧短时能量均值

初始门限设定

Yes

过渡带

No

连续3帧<门限

Yes

图2.音区区分

起始点

**3）基音频率识别：**

**①自相关法**

自相关函数要进行乘法运算，其运算量还是比较大的，为进一步减少运算量可对中心削波函数进行修正，采用三电平中心削波的方法，其输入输出函数为：



式中为削波电平，的经验公式为



开始

三电平削波

自相关处理

三电平削波

自相关处理

低通滤波

结束

图3.自相关法

FFT

式中，为频谱中前个大的峰值的幅度，为比例系数。

这使得大多数次要的峰被滤除掉了，而只保留了明显显示周期性的峰，且仅保留了0、1、-1，大大减少了运算复杂度。

自相关函数公式为



一般当为最大值时的值就是基音周期，但由于谐波峰值的干扰很有可能为最大值时的值为谐波周期，这样就会将谐波频率误认为基音频率。为了防止这种基音误取情况的发生，可对自相关系数再进行一次三电平削波、自相关处理，此时得到的自相关系数为最大值时的值即为基音周期。这是因为第一次自相关处理得到的也是以基音周期为周期的序列，在基音及各谐波位置上的幅值比较大，对做一次自相关处理可进一步排除谐波的干扰，得到准确的基音周期。基音频率估值可由计算，为采样频率。

当得到基音频率估值后，就需要以为参数来设计数字低通滤波器。设计滤波器的目的是为了在排除高频谐波分量的干扰和减小音乐信号的频宽的基础上用FFT来准确定位基音频率，因此可以为通带频率、为阻带频率、通带衰减3 dB、阻带衰减20 dB为技术指标采用双线性变换法设计数字低通滤波器。音乐信号经滤波后，通过快速傅里叶变换FFT后就可得到无谐波干扰的理想频谱

**②谐波峰值法**

由于基音的幅度不一定最大，但幅度大的分量一定是基音的某次谐波(一般不大于5)或基音本身，因此可对最大峰值频率求出1到5倍因子作为候选基音，然后对每个候选基音的n个谐波的幅度求和，和值最大的候选基音的置信度最大，继而是基音的可能性最大。公式为：



其中为候选基音，为最大峰值频率，为置信度，为某次谐波的幅度，为谐波的个数。

**5.实验、调试及测试结果与分析**

开始

FFT

最大峰频率

候选基音

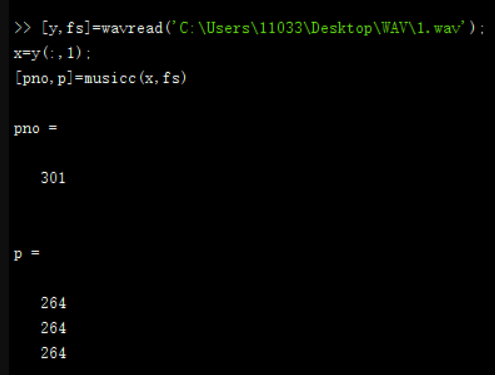
置信度

基音频率

结束

图3.谐波分析法

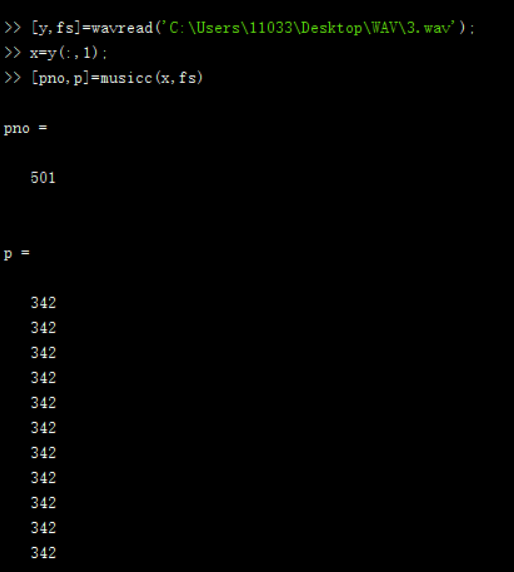
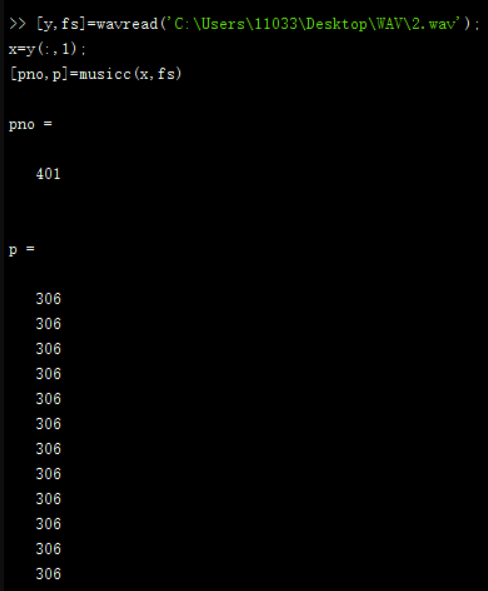
1）用自相关法对得到的音源中的音do进行测试



当帧的步长为250ms时do的检验（精度为4Hz）

音do的规定频率为261.63，可见音源的精确度较高。

2）用自相关法对其他音进行测试

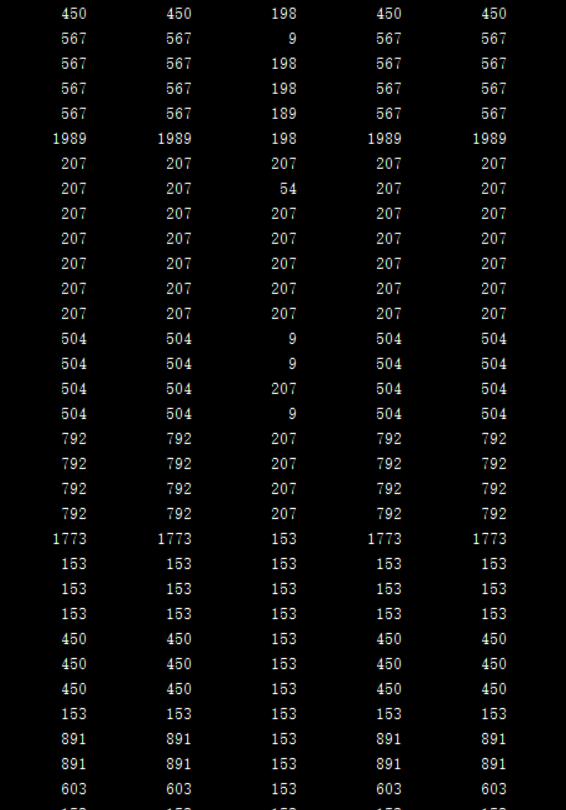
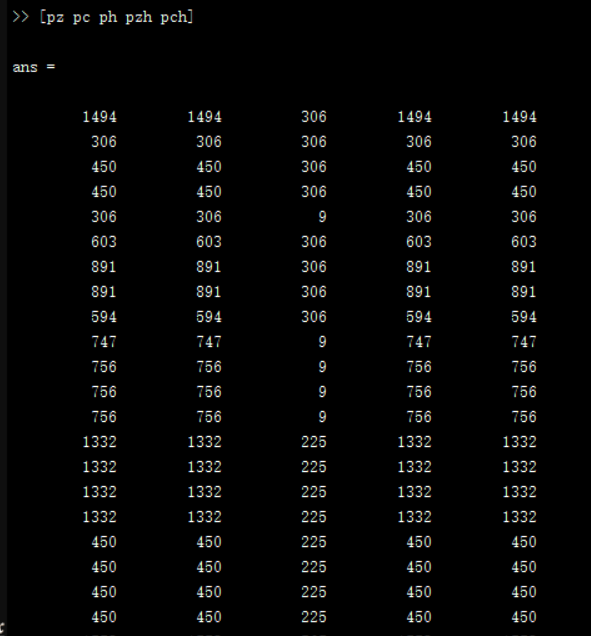


re

mi

通过对照十二平均律音高和频率对照表可知该设计系统能识别出音高。

3）利用八音盒版卡农的音频信号对比自相关法、谐波分析法、直接法（不进行处理直接提取基因频率）、直接和谐波混合法以及自相关和谐波混合法的识别结果。（混合法指在一定条件下分别使用其中两种方法中的其中一种，可利用逻辑判断进行处理）



开头部分

中间部分

结果显示五种方法中谐波分析法与其他四种方法差异较大，而进一步的实验测试表明只有当基音的谐波分量较多时才可直接使用谐波分析法，否则只能和其他方法混合使用。

1. **结论**

本次设计的系统能对音频信号进行较为精确的音高识别，在其音高识别的部分提供了五种可供选择的识别方式，但谐波分析法具有不能识别基音谐波较少音频的缺陷。同时其他方法具有高度的一致性，充分说明了该设计系统的准确性与精确性。

1. **参考文献**

[1]徐鹏进.基于音高与端点联合检测的音符识别算法.计算机应用,2011,Vol.31.Suppl.2.

[2]翟景瞳.改进的音高识别算法.计算工程与应用,2009,45(20).

1. **附录**

