



当前位置：首页 » PHP技术

音频特征提取——pyAudioAnalysis工具包

2017-05-05 00:01 本站整理 浏览(8)

作者：桂。

时间：2017-05-04 18:31:09

链接：<http://www.cnblogs.com/xingshansi/p/6806637.html>



前言

语音识别等应用离不开音频特征的提取，最近在看音频特征提取的内容，用到一个python下的工具包——**pyAudioAnalysis**: An Open-Source Python Library for Audio Signal Analysis，该工具包的说明文档可以[点击这里](#)下载，对应的github链接[点击这里](#)。

一、常用工具包简介

阅读榜

- 1 [Cookie与Session的区别-总结很好的文章](#)
- 2 [《大型网站架构技术》系列技术文章整理收藏](#)
- 3 [《PHP框架Yii学习》系列技术文章整理收藏](#)
- 4 [《CURL技术知识教程》系列技术文章整理收藏](#)
- 5 [《Yaf零基础学习总结》系列技术文章整理收藏](#)
- 6 [《PHP内核探索系列文章》系列技术文章整理收藏](#)
- 7 [localhost是什么意思_127.0.0.1 localhost介绍](#)
- 8 [vim设置默认字体、窗口大小和配色方案](#)
- 9 [《细说PHP》PDF电子版下载地址](#)
- 10 [notepad++设置字体和字体大小](#)
- 11 [推荐几本好的PHP书籍](#)
- 12 [PHP实例实现剩余时间倒计时显示](#)
- 13 [Kindeditor漏洞 编辑代码内容被执行](#)
- 14 [PHP取数组第一个元素的方法](#)
- 15 [PHP框架MVC原理和模板smarty原理](#)

目前针对音频信号，C/C++、Python、MATLAB等常用的工具包有：



二、pyAudioAnalysis工具包简介

pyAudioAnalysis是一个音频处理工具包，主要功能如图：



其中Feature Extraction包括（顺序有先后）：



补充说明一下：

1-Zero Crossing Rate：短时平均过零率，即每帧信号内，信号过零点的次数，体现的是频率特性

2-Energy：短时能量，即每帧信号的平方和，体现的是信号能量的强弱

3-Entropy of Energy：能量熵，跟频谱的谱熵（Spectral Entropy）有点类似，不过它描述的是信号的时域分布情况，体现的是连续性

4-Spectral Centroid：频谱中心又称为频谱一阶矩，频谱中心的值越小，表明越多的频谱能量集中在低频范围内，如：voice与music相比，通常spectral centroid较低

5-Spectral Spread：频谱延展度，又称为频谱二阶中心矩，它描述了信号在频谱中心周围的分布状况

6-Spectral Entropy：谱熵，根据熵的特性可以知道，分布越均匀，熵越大，能量熵反应了每一帧信号的均匀程度，如说话人频谱由于共振峰存在显得不均匀，而白噪声的频谱就更加均匀，借此进行VAD便是应用之一

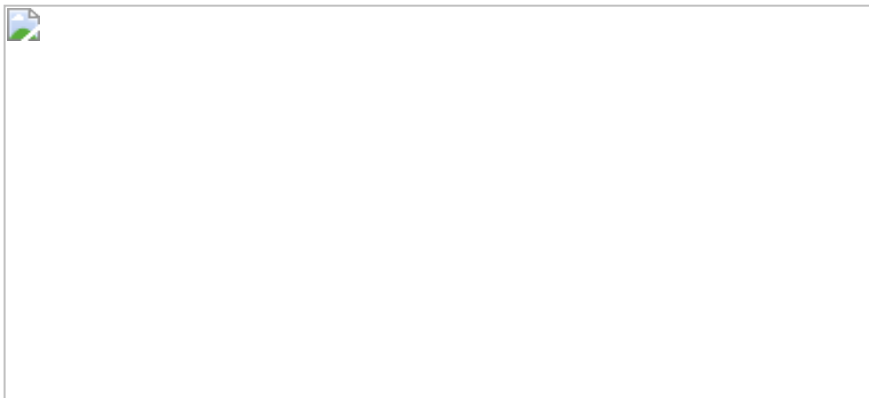
7-Spectral Flux：频谱通量，描述的是相邻帧频谱的变化情况

```
function [vsf] = FeatureSpectralFlux (X, f_s)
```

```
% difference spectrum (set first diff to zero)
afDeltaX = diff([X(:,1), X],1,2);

% flux
vsf = sqrt(sum(afDeltaX.^2))/size(X,1);
end
```

8-Spectral Rolloff：频谱滚降点，给出定义：



9~21-MFCCs：就是大名鼎鼎的梅尔倒谱系数，这个网上资料非常多，也是非常重要的音频特征。

22~33-Chroma Vector：这个有12个参数，对应就是12级音阶，还是看原文解释：A 12-element representation of the spectral energy where the bins represent the 12 equal-tempered pitch classes of western-type music (semitone spacing).

34-Chroma Deviation：这个就是Chroma Vector的标准方差。

这个在音乐声里可能用的比较多，目前没有接触这类特征：



什么是Chroma特征呢？给出一个示意图



code示例：

```
from pyAudioAnalysis import audioBasicIO
from pyAudioAnalysis import audioFeatureExtraction
import matplotlib.pyplot as plt
[Fs, x] = audioBasicIO.readAudioFile("sample.wav");
F = audioFeatureExtraction.stFeatureExtraction(x, Fs, 0.050*Fs, 0.025*Fs);
plt.subplot(2,1,1); plt.plot(F[0,:]); plt.xlabel('Frame no'); plt.ylabel('ZCR');
plt.subplot(2,1,2); plt.plot(F[1,:]); plt.xlabel('Frame no'); plt.ylabel('Energy'); plt.show()
```

如果希望了解更多的音频特征，这里给出一个链接，[点击这里](#)，包含的特征有：

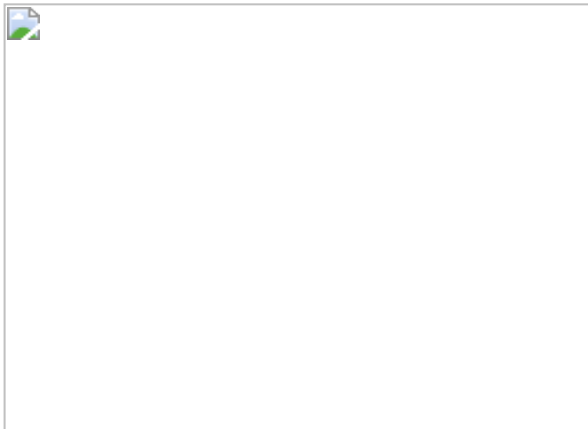


对应都有graph、sound可以点击，sound是对应的音频，graph对应的是特征的效果图，比如打开zeroCross:



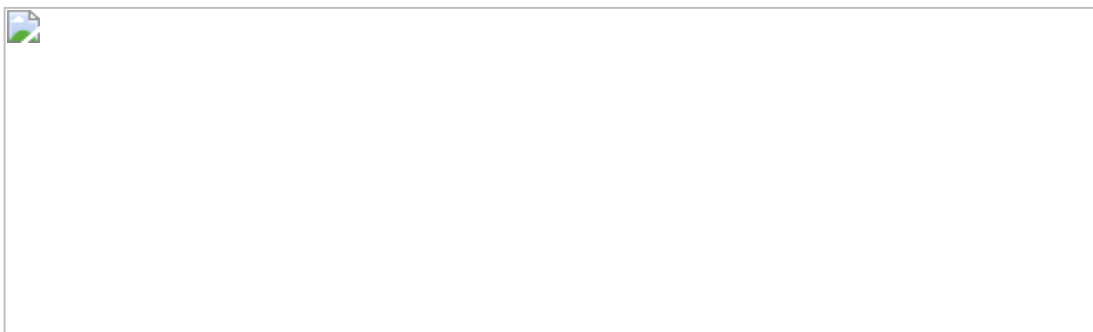
三、pyAudioAnalysis工具包安装

pyAudioAnalysis对应链接[点击这里](#)。安装这个工具包需要依赖：



A-hmmlearn安装

hmmlearn的链接[点击这里](#)。安装hmmlearn有几个前提：



下载之后，我把hmmlearn-master放在python-3.5.2-0\Lib\目录，cmd窗口下cd进去，输入：

```
pip install -U --user hmmlearn
```

即可安装成功：



B-Simplejson工具包安装：

Simplejson是Python的JSON编码和解码器，它具有简单、快速、完整、正确和易于扩展的特点，对应的链接[点击这里](#)。Simplejson工具包直接conda安装即可：



C-eyed3安装：

eyed3 : A tool for working with audio files, specifically MP3 files containing ID3 metadata. 它提供了读写 ID3 标签(v1.x 和 v2.3/v2.4)的功能。同时可检测 MP3 文件的头信息，包括比特率、采样频率和播放时间等。eyed3直接conda install没有成功，对应的链接[点击这里](#)。选择了这个版本：



放在了python库的Lib文件夹下：C:\Users\Nobleding\Anaconda3\pkgs\python-3.5.2-0\Lib，cd到对应目录下，pip install 文件名.whl，即可完成安装



D-pydub安装：

pydub是音频处理常用的工具包，例如：

打开一个wav格式文件：

```
from pydub import AudioSegment
song = AudioSegment.from_wav("never_gonna_give_you_up.wav")
```

打开一个mp3格式文件：

```
song = AudioSegment.from_mp3("never_gonna_give_you_up.mp3")
```

或者其他音频、视频格式：

```
ogg_version = AudioSegment.from_ogg("never_gonna_give_you_up.ogg")
flv_version = AudioSegment.from_flv("never_gonna_give_you_up.flv")

mp4_version = AudioSegment.from_file("never_gonna_give_you_up.mp4", "mp4")
wma_version = AudioSegment.from_file("never_gonna_give_you_up.wma", "wma")
aac_version = AudioSegment.from_file("never_gonna_give_you_up.aiff", "aac")
```

更多细节信息可以访问[其主页](#)。我在[github](#)上下载对应的工具包，里边有对应的安装说明。



如果处理wav文件，没有其他要求，如果音频是其他格式它要求电脑安装 [ffmpeg](#) or [libav](#)。如果没有安装，运行会有提示：



[ffmpeg](#)下载，选择版本



解压并添加环境变量，并利用ffplay测试一下打开一个mp4文件：



ffmpeg安装成功。这个时候import pydub,不再有warning信息



E-pyAudioAnalysis安装

[Github](#)给出的是linux下的安装思路，这里下载之后将pyAudioAnalysis放在了
\\Anaconda3\\Lib\\site-packages文件夹下，输入指令：



成功调用，原数据是支持Python2的，很多细节要修改，给出一个简单读取wav的测试：

```
from pyAudioAnalysis import audioBasicIO
import numpy as np
import matplotlib.pyplot as plt
[Fs, x] = audioBasicIO.readAudioFile("count2.wav");
time = np.arange(0,len(x))*1.0/Fs
plt.plot(time,x)
```

效果图：



猜你在找

[特征提取，特征选择](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取之——Haar特征](#)

[特征提取与特征选择](#)

[特征提取之Haar特征](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[特征提取](#)

[SIFT特征提取，特征匹配](#)

[LBP特征：图像特征提取](#)

[特征选择和特征提取](#)

[特征工程 vs. 特征提取](#)

Copyright © 2011-2015 爱程序网 沪ICP备14001972号-1 业务咨询：

