

关于音频 onset detection 算法的阅读

Origin:Using Audio Onset Detection Algorithms

本文档只记录了部分的内容,主要以 `aubio` 相关内容为主,并非整个文档的完整内容,记录人:lihaiping1603@aliyun.com

QMUL 算法: 这个基于信号的算法,它结合了能量(观察信号的能量)和相位(观察 FFT 状态偏差)所共同构成复域。它包括一个自适应增白组件,该组件平滑了信号的时间和频率变化,从而通过“将每个频段的幅度引入一个类似的动态范围,使振幅较大的峰值更加明显。该算法跟进复域内的峰值计算每个频域内发生突发事件的可能性,并使用峰值选取算法来标记 onset.

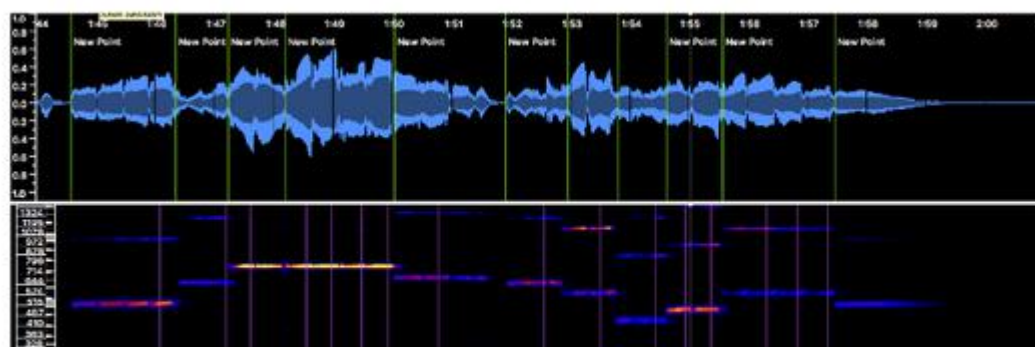


图7:这个例子显示了真音符onsets(绿色)与感知到的QMUL音符onsets(紫色)之间的关系。

Aubio:关于 onset detection 的算法 `aubio` 和 QMUL 算法类似,它改进了 onset detection 中的自动校正功能,通过计算节拍周期, `phase` 相位对齐方法.主要是根据周期,相位,节拍来做预测.这个算法有两个主要的变量参数值: `threshold` 阈值 0.01-0.99(主要用于峰值拾取)和 `onset` 模式(对于 `detection` 功能,包含高频内容,复域,能量和光谱差异)。

例如分析下面的长笛音乐,采用了复域的分析方法.然后系统参数变量调优中 `FFT bin` 大小为 1024,增量大小为 512,峰值阈值 `threshold` 设置为 0.5,寂静阈值设置为 -50dB,以及内部最小的 `onset` 间隔设置为 40ms.由于相位声编码器的存在,窗口大小 `window` 设置为 1024,跳数 `hop` 设置为 512. 通过改变峰值选取算法的阈值,越低或者越高,会导致过多或者过小的 `onset`。

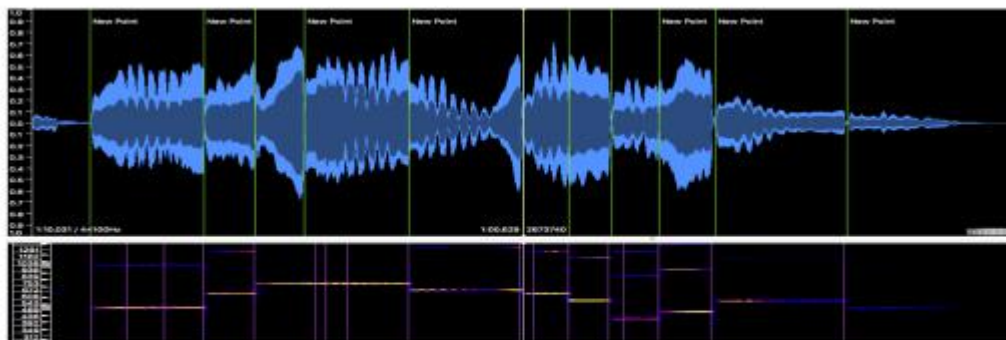


Fig. 8: This example shows how true note onsets (in green) relate to perceived note onsets (in purple) of aubio.

实际情况和 aubio 的 onset 算法分析结果的对比展示如上图。在上图可以看出来，在这个音乐中 11 个真实的 onset 都被正确的发现了。

Pyin 算法：它和上述算法的不同之处在于，他的目的是检测音高，而不是显式的 onset detection，并且是一种基于概率的方法。它提取给定频域范围内的音高。由于该算法的设计在基频估计的基础上附加了一个时间戳，证明了他是一个有效的 onset detection 竞争者。同时这个信息可以用于推断音符的 onset。

使用示例，系统使用 FFT bin 大小为 1024 个样本，增量大小为 512，YIN threshold(一组相关概率的音高候选值)全部设置为 11,抑制低幅度音高估计值设置为 0.1(将振幅抑制在一定值以下),onset 的灵敏度设置为 0.7(相当于峰值拾取)，分析的结果如下图：

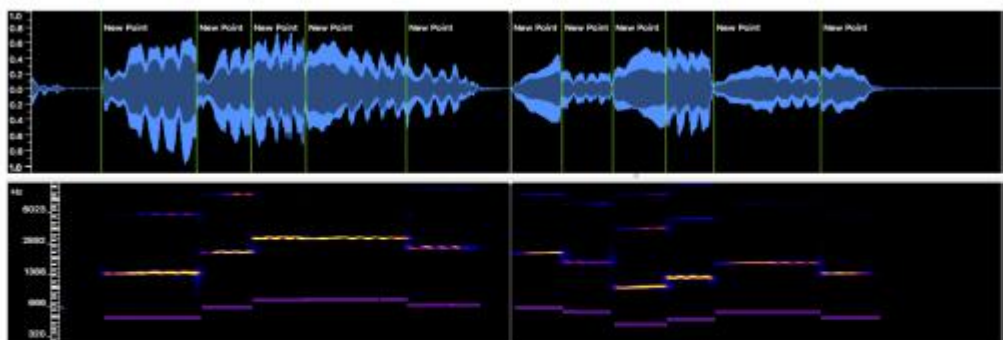


Fig. 9: This example shows how true note onsets (in green) relate to perceived note onsets (in purple) of pyin