

基于信号频谱分析的音乐检索专题研讨

【目的】

- (1) 加深信号频谱的概念，以及对信号频域分析基本原理和方法的理解。
- (2) 培养学生理论联系实际的素质，提高学生的工程实践能力和创新能力。
- (3) 培养学生查阅文献、自主学习的能力。

【知识点】

信号的抽样、频域分析，滤波器等

【背景知识】

音乐检索的主要方法是基于内容的检索，即利用音乐的音符、旋律、节奏、歌曲风格等语义级的特征或者声学层特征从数据库中检索乐曲。本研究专题使用基于信号频谱分析的方法实现音乐检索。下面给出使用信号频谱分析方法实现音乐检索的原理和方案，该方案包含两部分，第一部分为模板库的构建，第二部分为音乐信号的识别，如图 1 所示。

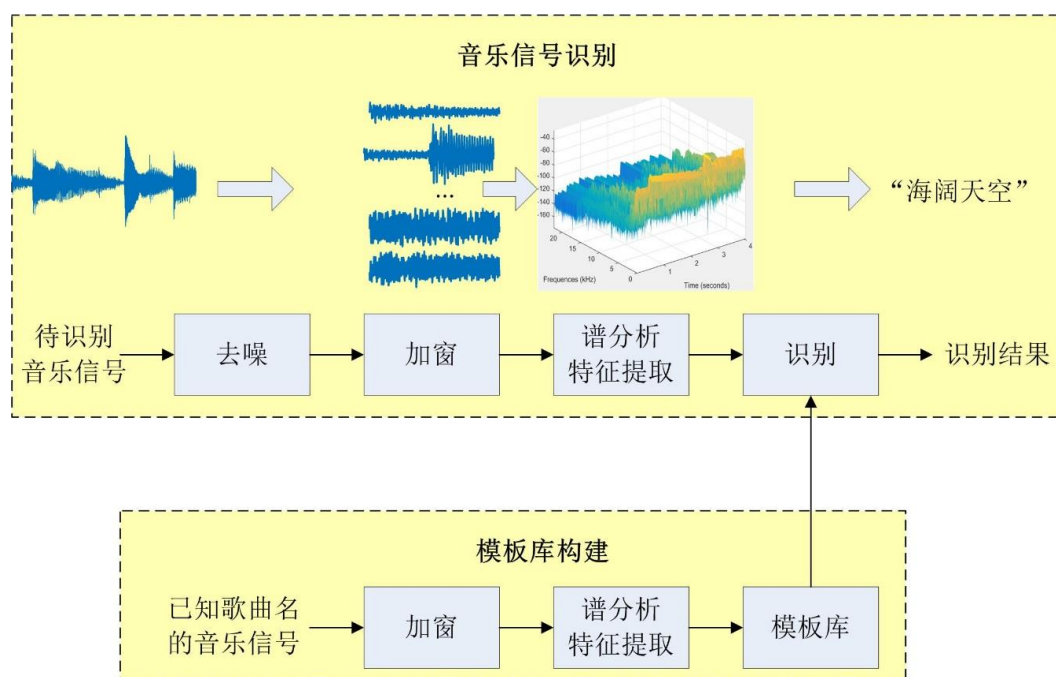


图 1 听音识曲系统

模板库的构建：首先在网上下载多首音乐，然后对每首音乐提取特征，多首音乐的特征即构成模板库。音乐信号并不是一个平稳随机过程，所以不能直接对整个信号进行频谱分析，需要对其进行加窗分帧操作。将音乐信号分为毫秒级的多个音乐片段，每个音乐片段可以看做是平稳随机过程，然后对音乐片段进行频谱分析。为了保证两个音乐片段之间的平滑过渡，需要设置一个偏移量，也就是前后两帧的数据需要共同拥有同一节数据，偏移量一般设置为帧长的二分之一。加窗分帧操作可以通过窗函数实现，每移动一次窗函数，便得到一帧的音乐片段。

音乐信号的识别：打开麦克风对播放音乐进行录制，按照模板库构建过程，对信号进行加窗分帧

和特征提取，最后将待识别信号的特征与模板库进行匹配，得到该首音乐的歌曲名称。

为提高歌曲识别的速度，可以对音乐信号进行降采样处理，需要结合信号抽样和多速率信号处理理论，对音乐信号的频谱进行分析，得到合适的降采样率，以达到在不造成频谱混叠的情况下降低处理的数据量。

【研讨内容】

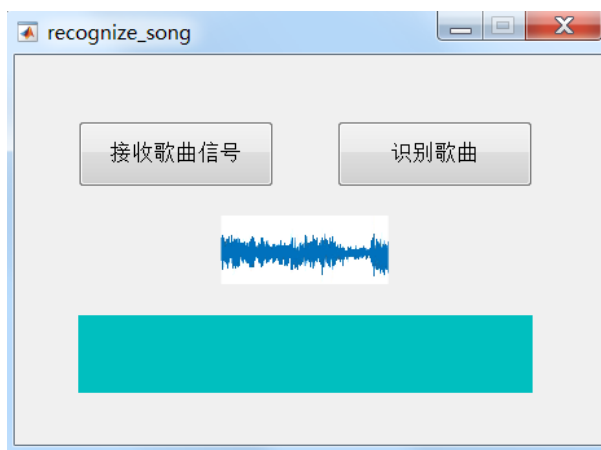
- (1) 建立音乐检索数据库。数据库有若干首歌曲或乐曲构成。格式自定，歌曲数量不少于 3 首。
- (2) 设计并实现基于频谱分析的特征提取算法，如梅尔频率倒谱系数（MFCC）。
- (3) 实现乐曲信号的匹配算法，当特征选用 MFCC 时，匹配算法可采用动态时间规整（DTW）算法。

【温馨提示】

- (1) MFCC 方法参考资料见“Mel Frequency Cepstral Coefficients for Music modelling.pdf”，也可通过上网查阅相关的中文资料。
- (2) 曲目之间不宜过于类似，当采用 MFCC +DTW 时，为保证检索在较短时间内完成，数据库每首音乐的时间不超过 10s，可将歌曲文件中间部分剪裁 10s 作为数据库歌曲。
- (3) 信号的匹配可直接调用现有的算法，DTW 算法的 matlab 程序见“dtw.m”。
- (4) 在实现基于频谱分析的特征提取算法，需要综合频谱分析、滤波器等知识点。
- (5) 采用 MFCC +DTW 实现音乐检索时，其识别正确率可能较低，由于本研学内容重点在于频谱分析与滤波器设计的综合应用，识别正确率不过多强调。
- (6) 若想提高识别正确率，可采用基于 hash 的特征提取与匹配方法，该方法的参考文献见“Hash.pdf”，相比于 MFCC +DTW 方法，基于 hash 的特征提取与匹配方法编程较复杂。

【研讨要求】

- (1) 编写程序，实现音乐检索系统，编程语言不限。
- (2) 鼓励采用界面的方式完成该系统，如下所示：



点击“接收歌曲信号”按钮，系统开始对播放的歌曲录音；点击“识别歌曲”按钮，系统启动

歌曲识别算法，识别结束将识别结果显示在上图青色的方框中。

(3) 完成音乐检索系统的研究报告，报告包括课题背景、方案设计、实验结果与分析以及参考文献。其中关于信号的匹配算法不需要详细的理解，重点说明基于频谱分析的特征提取算法。