# 

# 创新实验课 课题立项报告

课题名称: 智能音乐助手

课题名称（英文）： AI-MUSIC

课题组长： 冯韵菱

课题组成员： 冯韵菱 张靖鸿 许宏涛

报告时间：2020年 9月 9日

1. 项目创意来历及项目意义（建议500字）

初学吉他必然会遇到一些问题，诸如调音不准，按弦位置错误，或有错音杂音等等。

之所以产生这些问题，是因为没掌握正确的练习方法。当然，即使注意了练习要点，刚开始也有可能出现以上问题，需要及时纠正。因此，在乐器的学习过程中教练显得尤为重要。由于教乐器需要付出更多的精力和时间，而且是针对性的教学，乐器课收费价格昂贵，而且有指导下的训练仅限于课堂，线下个人的练习纠错困难。对于整个行业而言，一方面，优质的老师很稀缺，随着行业竞争的加剧，老师的成本也变得越来越高；另一方面，老师的教学水平参差不齐，流动性强，也给热爱乐器的同学带来了选择上的困难。相比于传统乐器教学，AI+乐器教学的价格优势十分明显，可以随时上课，也不存在教学质量参差不齐的问题，最后，用户的学习数据可以沉淀到教学系统中，通过自适应学习系统，捕捉和回应学生不同的需求和反馈，一定程度实现个性化的教学。因此，乐器的智能教学更加有机会打破现有教育资源供应的瓶颈，并能以普惠的价格提供给所有家庭，使得无论一二线还是三四线城市以及农村家庭都能承受，实现教育公平。

1. 项目研究主要内容（建议1000字左右）

* 乐器的音频特性分析与识别

音频特性分析是识别以及之后其他研究内容的基础。它包括以下内容：乐器的单音特性分析，乐器的连续弹奏特性分析以及和弦的音频特性分析。其中单音特性分析与识别即为对乐器的每一个基本音的采样分析，根据常识，它应该是一段时间内呈现一定周期的声音信号与一些噪声的叠加。即时域存在一定周期性，频域存在频谱峰。我们拟通过这些特点或者更多观察发现完成乐器的单音识别。乐器的连续弹奏特性分析即为对乐器连续弹奏的声音进行采样分析，由于时域上声音的叠加干扰，在音符相接处信号更加复杂，同时识别上我们想借助单音识别的基础，这就要求我们学习更多的音频处理知识以完成连续音到单音的切割、从一个音调中滤除其他音的干扰。和弦的音频特性分析与识别则最为复杂，它是同一个时间中多个频率的叠加，我们拟借助时域与频域的多重分析完成识别。

* 乐器的错误原因学习

完成了音频特性分析与识别之后，我们通过对演奏者演奏的乐曲进行采样，在一定的误差范围内，我们易判断出演奏者的错音。在此之后，我们想进行演奏错误原因的判断。这是一个错误分类问题，我们想从吉他入手，收集演奏错误对应的样本，比如调音问题，按弦不准等等，我们将进行样本的一些预处理，提取出音频特征点以供机器学习。在此之后，我们可以通过演奏者具体的错误类型给以具体的纠正与指导。

* 机器的乐理学习

在以上三个任务均以完成的情况下，我们拟将人工智能运用到作曲领域，创造一个人工智能音乐作曲系统。它要经过乐理知识的学习，能够根据所提供的部分弦乐信息完成整个乐章的谱写，为机器的使用者提供灵感与参考，提升用户的使用体验。

1. 项目创新点与项目特色（建议200字左右）

自动纠错

在传统的乐器教学项目中，往往是教学视频或是弹唱方式的整合或者按弦的教学。这样的学习模式对于乐器自学者最大的问题在于错误无人指出，仅仅是以机械的方式帮助初学者记忆，如果自学时一开始的错误就无人指出，被不断强化记忆，这对乐器的学习是很不好的。而AI-music的纠错功能相比之下显得新颖。

对多种乐器适用

在一个乐器的算法完成之后，只要更改其他乐器的各种采样信息，就可以不断扩大适用的乐器种类。比如完成吉它的音频分析识别后，我们只需简单地采样钢琴每个音调的信息，便可以完成钢琴演奏的纠错功能。

个性化

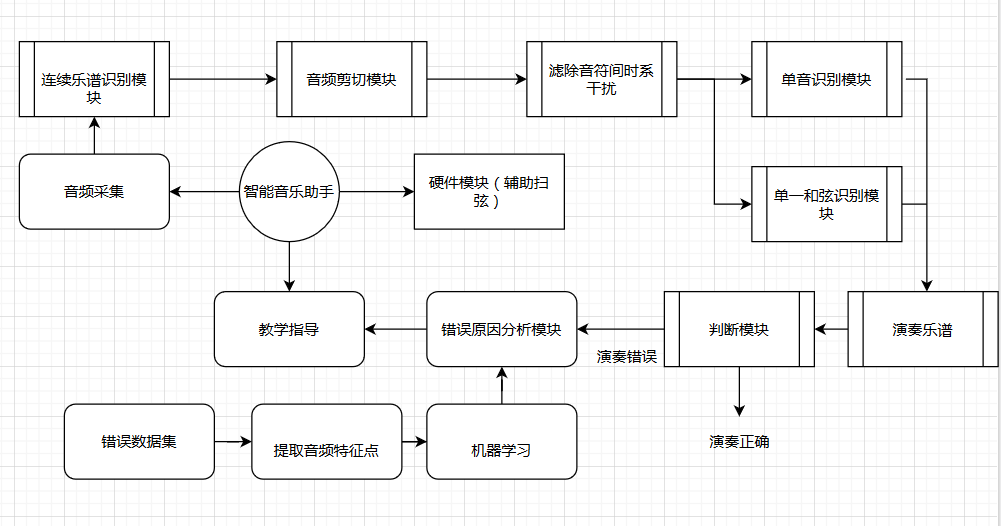
相比于传统的大课堂教学方式，AI-music通过用户的演奏信息收集，更易完成个性化的分析。

1. 系统方案和技术路线（建议1000字左右）

* 关键技术和设计思路
* **音频处理技术**

借鉴传统的数字信号处理方式，利用python进行音频文件的软件处理。

* **机器学习**
* 不同于传统音频处理的时域频域分析方式，借助音频样本的学习使得处理结果更加准确
* **数学物理方法**
* 从最原始的乐器发声方式出发，研究乐器奏乐不同于一般的语音信号的特点
* 系统模块图和必要的说明



* 功能概述（包括子模块的功能）
* **音频采集模块**

进行演奏乐曲的音频信号收集

* **音频剪切模块**

根据音频信号特点，如幅值大小、过零率特性完成一段音乐的分割，得到每个音符的分割时间点

* **过滤模块**

得到分割后的音符并不是完全的单音，其不可避免地存在上一个甚至几个音的残留叠加。过滤模块进行一定的叠加消除，便于最后的识别

* **单音识别模块**

对单个音符的乐谱进行识别，得到其频率，进而判断其声调

* **和弦识别模块**

对单个和弦的乐谱进行识别，得到该和弦的所有子频率，进而判断该和弦

* **判断模块**

将实际音谱与正确的音谱做比较，判断演奏时出错的位置与时长信息

* **特征提取模块**

对错误样本（各种弹奏错误的音频样本）进行音频特征点的提取，以作为机器学习模块的学习对象。

* **机器学习模块**

对音频特征点进行学习，调整各特征点权重参数，经过足够的样本训练后，可以完成对音频错误原因的识别

* **硬件辅助模块**

根据乐谱完成扫弦功能

* 技术亮点（比如：算法改进、算法创新等）
* 机器学习方式的错误原因识别不同与传统的音频时域频域分析，基于样本训练的纠错模型可以给予用户更明确的错误指正
* 使用的软件和开发测试环境

Python3.7.1

Python IDE

* 开发难度

音频分析与识别方面，单音识别难度适中，音符分割难度适中，但要得到非常准确的每个音符的前后节点十分困难。和弦识别较为困难。

机器学习方面，不论是大量错误样本的获取与音频特征点提取或是建立目标函数以及对多层网络的权重调整都具有很大挑战性。

因此，总体来说开发难度较大

1. 项目进度安排和人员分工

* 写明前期安排（比如花多久时间去学知识、熟悉环境等），中期安排（实现功能的进度），后期安排，列出几个时间关键点
* **项目进度安排**
* 首先进行音频知识的再次学习，包括时域频域分析，传统的音频信号处理，之后写算法

在9.11完成音频的处理与分析，在10月中旬完成机器学习的相关知识学习以及错误0型判别。之后进行硬件搭设，在10月底基本完成项目内容。后期安排进行样本的测试与系统算法优化。在10月26日-11月30日间，完成各模块的串联调式。在12月7日-12月21日，完成演示视频录制、实验报告撰写。

* 组内人员分工情况
* 冯韵菱

音频剪切模块，音频的时域频域分析，判断模块，机器学习模块

* 张靖鸿

滤除音符时系干扰模块，音频的时域频域分析，音频特征点提取，机器学习模块

* 许宏涛

单音识别模块，音频错误样本采集，机器学习模块，扫弦硬件搭设

1. 已有基础

* 已取得的相关研究积累和已取得的成绩（包括查阅的文章、调研、网站查询等）
* 单音的识别已经完成，在识别误差为4hz的容许范围下，识别准确率为90%。
* 连续演奏的分割已经完成，但仍存在一些分割错误，比如相邻音符之间的相互残留。
* 通过对分割音谱的演奏分割，正常速度的演奏识别已经完成。
* 通过对时域和频域的双重检测，进一步提高了连续音检测的准确率
* 已具备的研究条件，尚缺少的研究条件及解决方法
* 基本学习完所要用到的音频处理相关知识
* 缺少更精准的分割算法
* 缺少供机器学习中纠错的样本

1. 预期成果形式，达到的技术指标

* 单音识别方面，要达到99%以上的准确率(允许1hz误差)
* 整段乐谱识别要达到98%以上的准确率（允许1hz误差）
* 经过机械学习后的纠错模块能够判断错误原因，并达到90%以上的准确率
* 整个音频分析流程时延不大于20ms（假设音频时长为5min）
* 整个代码占用内存不超过10mb