

MUSICODER——基于深度学习的 轻量化音乐质感转换系统

说明文档



开源代码



项目页面



介绍视频



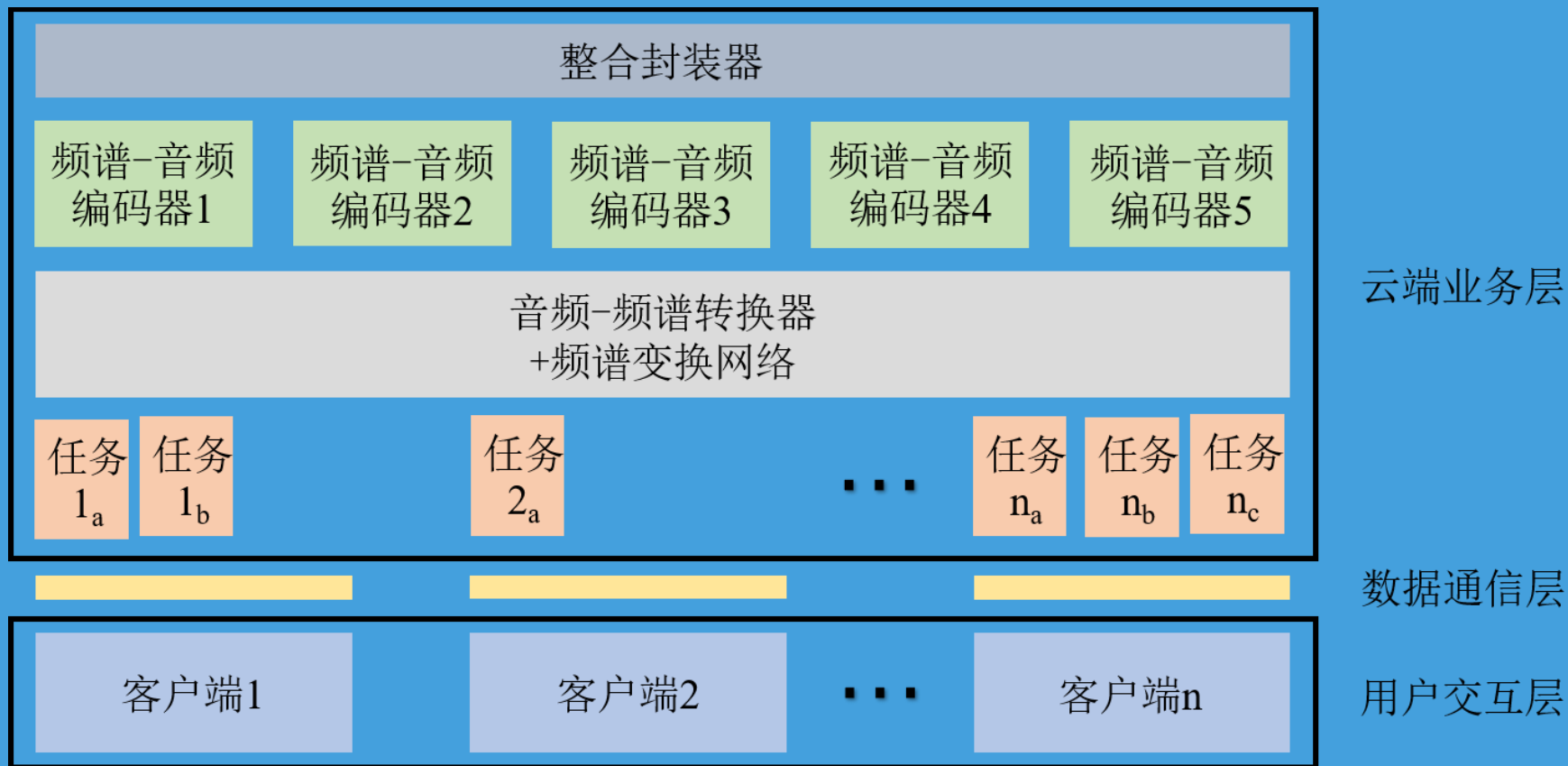
索引

- 项目信息 ----- i
- 整体架构 ----- 1
- 客户端 ----- 2
 - 配置要求 ----- 3
 - 安装部署 ----- 3
 - 最佳实践 ----- 4
 - 上传音乐 ----- 4
 - 播放与剪辑 ----- 5
 - 目标质感选择 --- 6
 - 深度质感转换 --- 7
 - 输出音乐 ----- 8
 - 其它说明 ----- 9
 - 网络连接问题 --- 9
 - 播放音量问题 --- 9
 - 音乐加载问题 --- 9
- 服务器端 ----- 10
 - 配置要求 ----- 11
 - 硬件及系统 ----- 11
 - 软件及环境 ----- 11
 - 预训练模型 ----- 11
 - 安装部署 ----- 11
 - 其它说明 ----- 12
 - 运行时意外 ----- 12
 - 临时文件管理 ----- 12
- 版权信息 ----- 13
 - 学术作品 ----- 13
 - 软件著作权 ----- 13
 - 开源协议 ----- 13



整体架构

MusiCoder系统主要包含两个模块。一个是部署在服务器后端的音乐质感转换网络，一个是供不同用户进行并发交互的客户端程序。二者通过网络连接单元实现数据通信。系统的整体架构见下图





客户端



MusiCoder主界面

我们面向Windows系统开发了供功能试用的64位客户端，并按照功能与软件包体量分为了`std`版和`lite`版。

软件主界面包括输入音乐面板、输入音乐播放器、输出音乐面板、输出音乐播放器、剪辑按钮、质感选择按钮、清晰度选择按钮、保存按钮等交互元素。它们的具体作用请参考[最佳实践](#)。界面上的数据传输状态单元主要解决连接服务器相关的问题，具体信息请参考[网络连接问题](#)。除此之外，界面上提供帮助按钮，可通过它引导至项目页面或查看在线帮助文档。



客户端

配置要求

硬件

- 不少于**4G**内存
- 不少于**100M**空余存储空间

软件

在较新版本的Windows操作系统测试中，尚未出现依赖异常

安装部署

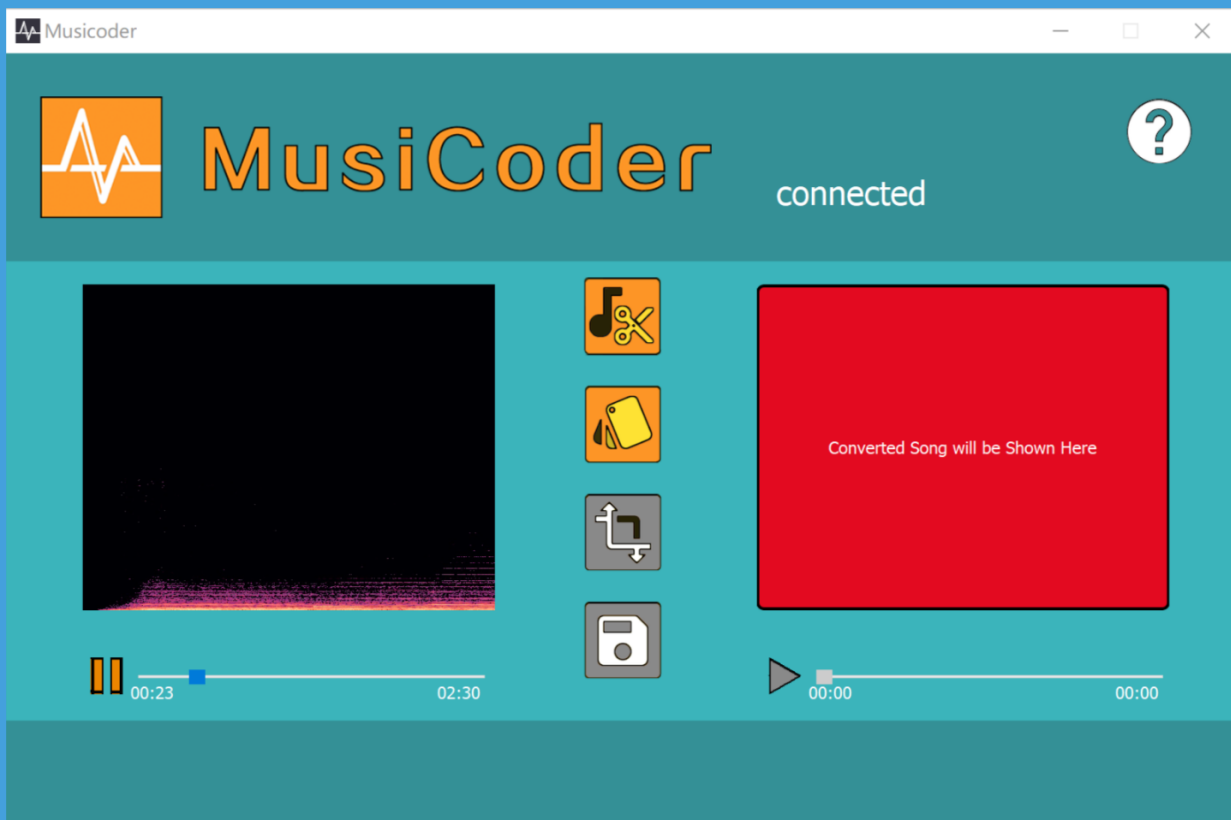
0. **安装前请确认已经关闭了全部安全软件!**
1. 点击安装包，按照提示执行各步骤。默认安装在当前目录下
2. 成功安装后，目录中应包含Style文件夹（放置质感实例音频）、ffmpeg可执行程序（提供音频编解码）及MusiCoder可执行程序
3. 检查网络连接后，点击MusiCoder，即可打开客户端



最佳实践-上传音乐

我们支持 **.mp4** 与 **.wav** 格式的音频文件。仅需简单地将音乐拖拽到输入音乐面板，您就能将音频文件加载进入软件。加载成功后，便会自动建立与服务器的通信，并返回曲目（默认前10秒）的波形图。

在使用全程，您均可通过点击输入音乐面板右上角，删除当前音频文件从而重新上传加载。

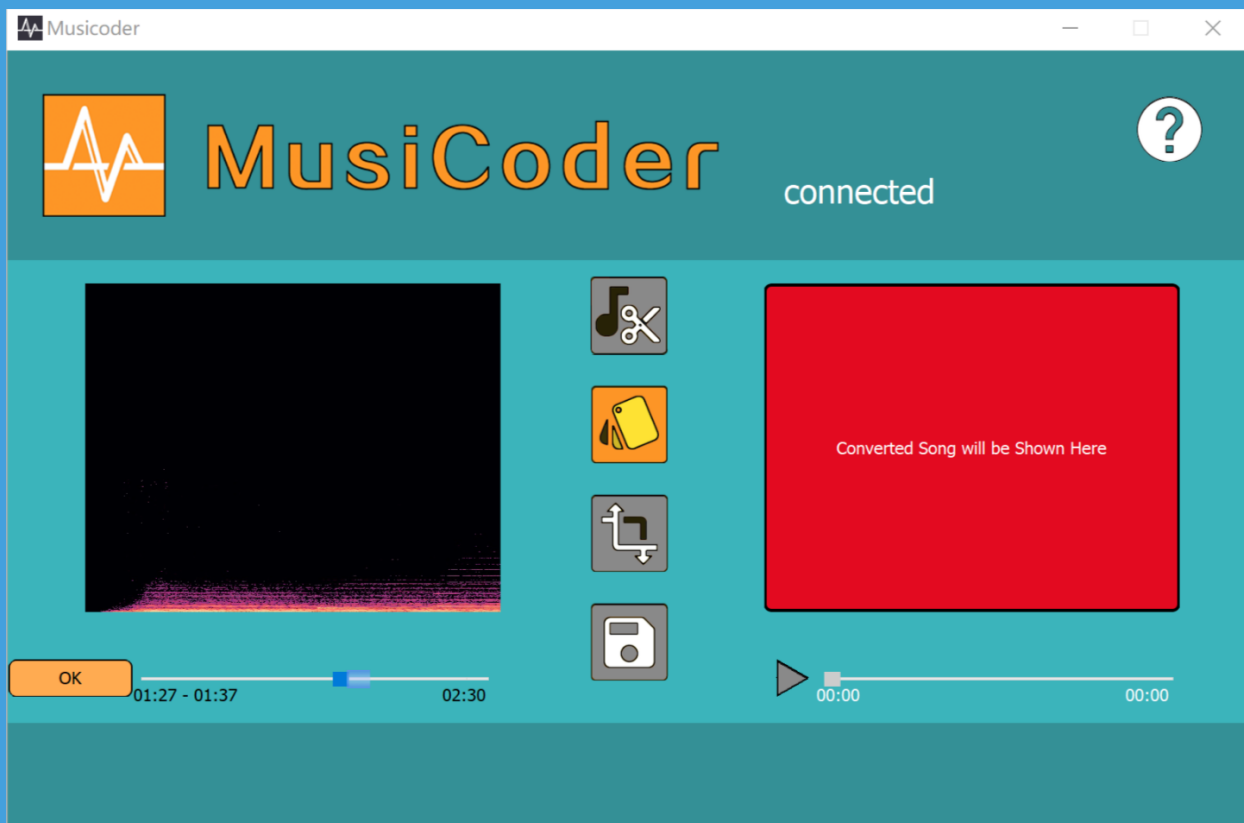


上传音乐并播放



最佳实践-播放与剪辑

点击输入音乐播放器，您可以欣赏导入成功的音乐。试用阶段的MusiCoder支持转换时长为**10秒**的音乐。当剪辑按钮由暗变亮时，您可以进行点击，移动进度条以指定待转换区间，并点击OK进行剪辑。剪辑完成的音乐会实时上传服务器，并更新对应波形图。您可以随时试听剪辑获得的片段。

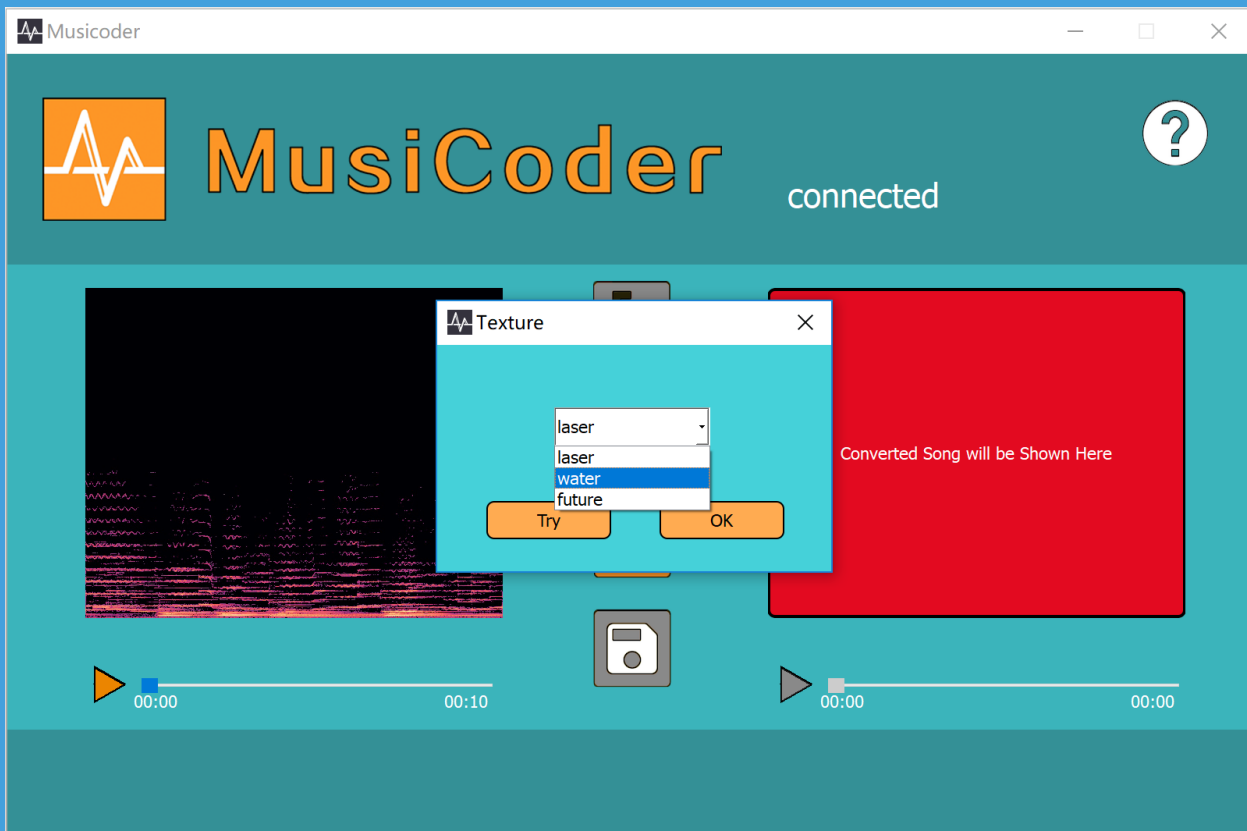


剪辑音乐

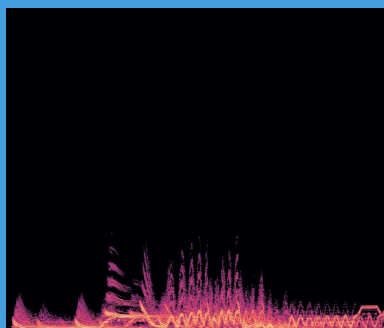


最佳实践-目标质感选择

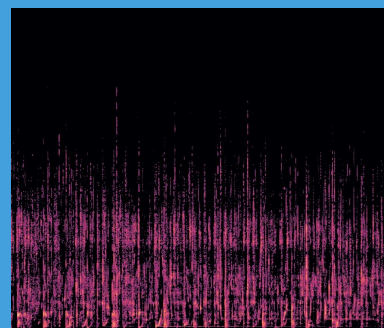
点击质感选择按钮，您可以点击 Try 试听目标质感，并点击 OK 进行选择。我们目前准备了 **Laser** (镭射音效)、**Water** (水流音效) 和 **Future** (未来感音效) 三种备选质感，其对应频谱图如下。



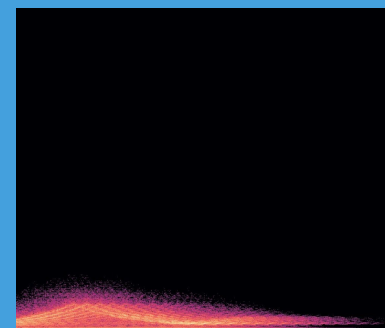
选择质感



Laser



Water



Future

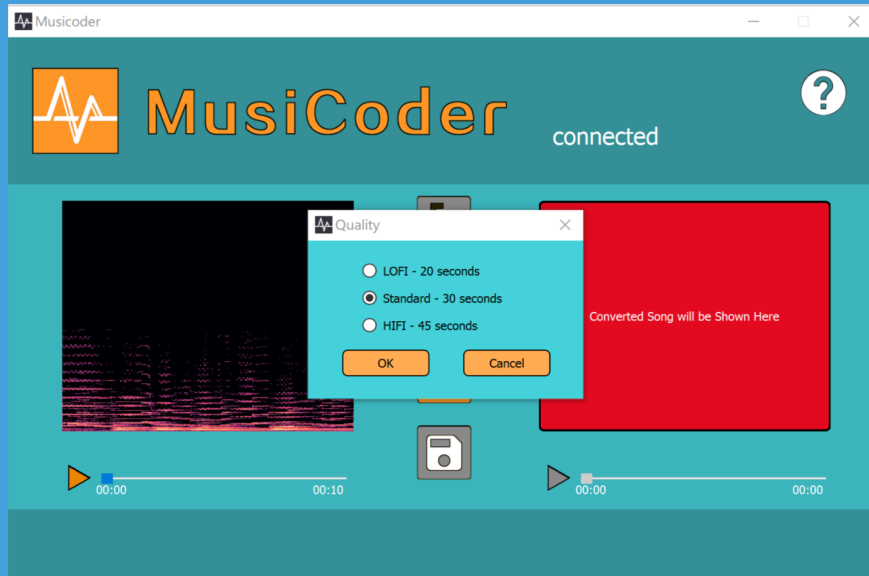


最佳实践-深度质感转换

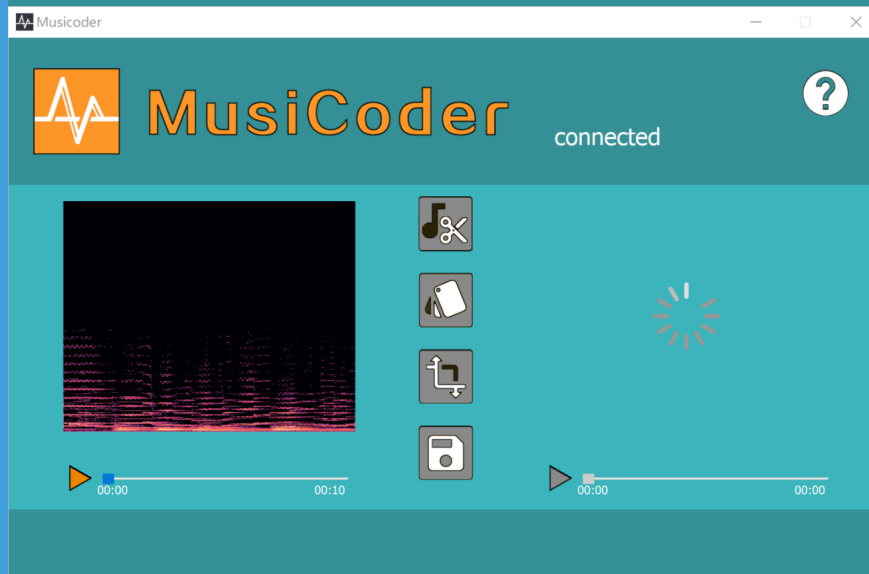
确定目标质感后，在清晰度选择按钮变亮时，请点击。我们提供了三个不同的输出音频清晰等级**LOFI**、**Standard**和**HIFI**，分别代表低质量、中等质量和高质量。在网络质量正常的情况下，分别需要20秒、30秒和45秒。这里清晰度的区别主要来自于云端神经网络通过Griffin-Lim算法重建音乐输出时，迭代次数的不同。完成选择后，点击OK，即可启动服务器端任务。

在转换过程中，输出音乐面板将会出现进度画面。此时请您稍作等待。

选择输出质量



进行转换中



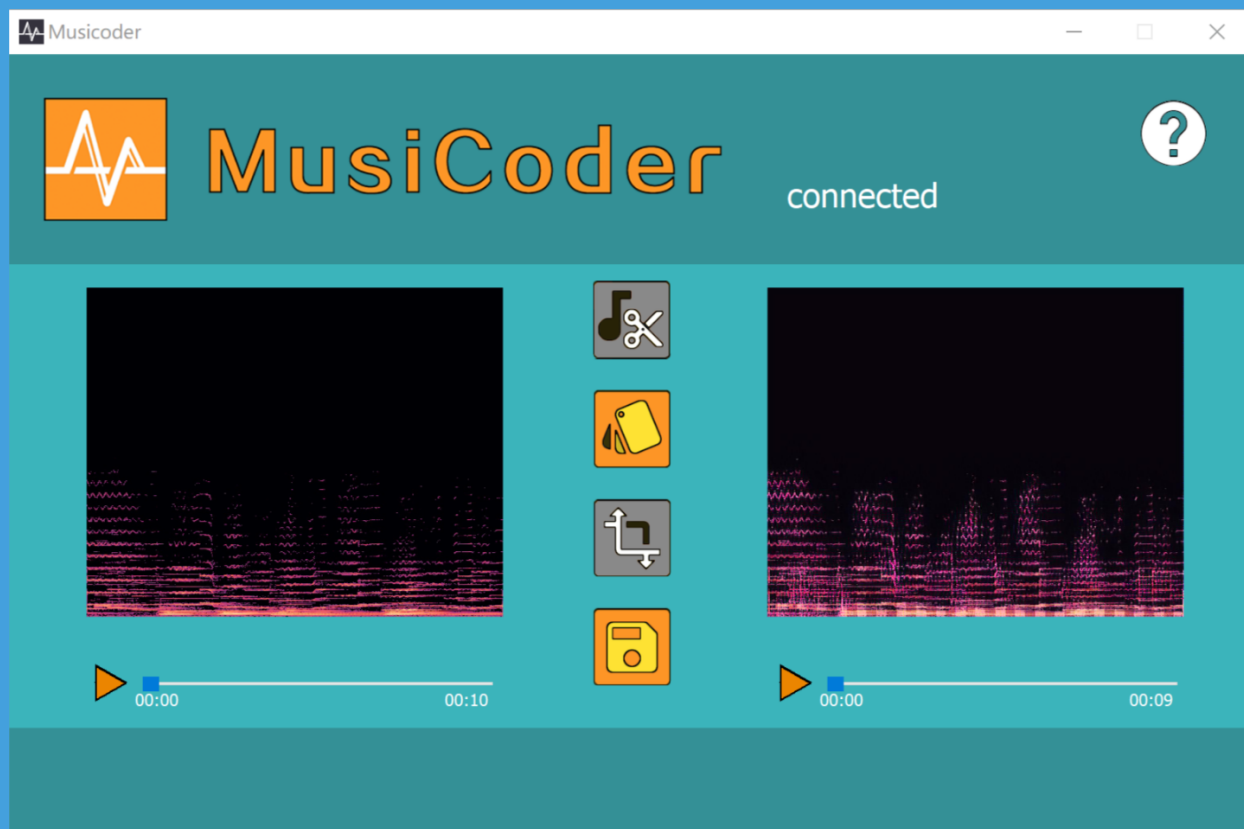


最佳实践-输出音乐

服务器端软件在完成转换后，会自动将生成的具备新质感的音乐及其频谱图发送回客户端。输出音乐播放器将被激活，而输出音乐面板上也会显示频谱图。您可以立刻进行欣赏。

如果您喜欢输出的音乐，在变亮后，可以点击保存按钮，将它保存在本地任意路径。

如果您希望尝试更多质感，或者改变清晰度，无需重新上传剪辑音乐，只需重复**目标质感选择**与**深度质感转换**即可。



完成转换



客户端-其它说明

网络连接问题

作为一款基于云服务的软件，MusiCoder可能会遭遇网络通信异常等情况。我们在数据传输状态单元提供了针对软件状态的显示。**connected**表示连接正常；**connecting**表示正在连接；而**not connected**表示通讯异常，此时请点击**retry**恢复数据传输。

如果您的网络异常出现在完成音乐剪辑后，我们建议您重新上传剪辑音乐，以便获得最佳使用体验。

播放音量问题

本软件的std版本集成了基于**音频强度正则化技术**的平衡模块，确保任何时候播放音乐均获得一致的音量。而lite版本则去除了这一模块，需要您手动调整音量至合适值。

音乐加载问题

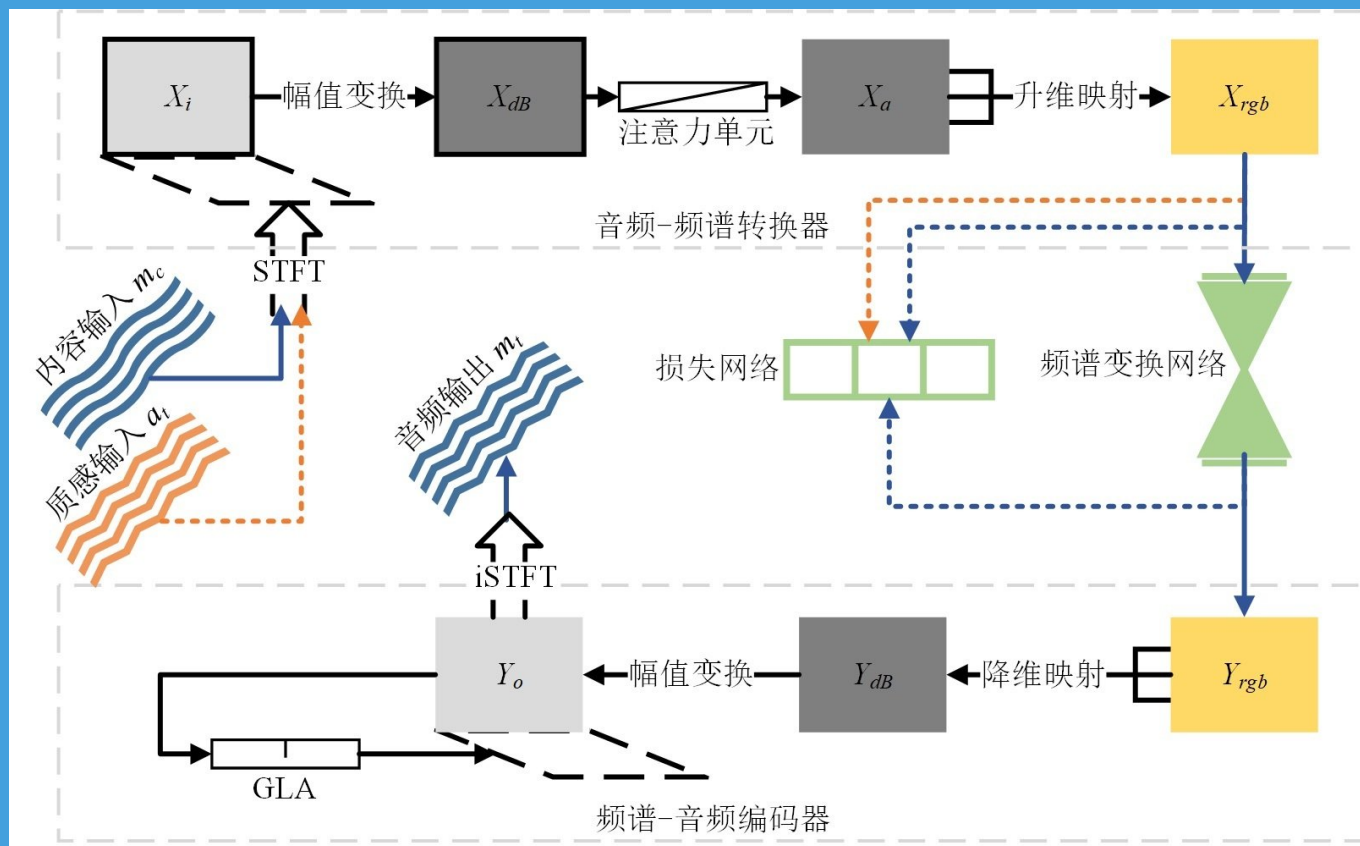
由于音量平衡模块的存在，本软件的std版本在初次上传加载音乐时耗时较长，**期间程序将不提供响应**，请您保持耐心。lite版本则可以实现快速加载。

请按需选择std版本或lite版本。



服务器端

我们在云端服务器上部署了基于深度学习的后台及数据交互/会话管理工具。其中质感转换网络的架构和前馈神经网络简图如下图所示



音乐质感转换网络



频谱变换网络(前馈)



服务器端

硬件及系统

- 不少于**4G**内存
- 运行Ubuntu 12.04 LTS及以上版本操作系统

软件及环境

- Python3环境
- 科学计算与数据可视化库：
matplotlib、scipy、numpy、pylab
- 图形图像依赖库:PIL.Image
- 音频处理库及工具:librosa、pydub、ffmpeg(需设置为环境变量)

预训练模型

请下载下列训练好的模型文件

- <https://www.dropbox.com/s/6xhg6ipsn0fq7yy/future.ckpt.zip>
- <https://www.dropbox.com/s/y2rstqwq21xph99/water.ckpt.zip>
- <https://www.dropbox.com/s/wln82c3c6ibhbfx/laser.ckpt.zip>

安装部署

1. 完成环境配置
2. 开启指定端口[port]
3. 维持原有目录结构部署程序，并将下载的模型解压放置在converter/models下
4. 执行命令，运行程序（后台静默模式）

```
nohup python3 server.py [port] &
```



服务器端-其它说明

运行时意外

在遇到通信异常等非服务器端内部问题时，会话管理程序会**自动回收**离线客户端所占用的资源，确保服务正常运行。

临时文件管理

客户端上传的临时文件存放在 *receive* 目录下，后台质感转换产生的中间文件则存放在 *temp* 目录下。正常情况下，服务器程序会自动化地进行清理。但是，如果客户端意外下线，则可能导致临时文件残留。因此，**我们建议您定时查看对应路径，必要时进行清理即可。**



版权信息

学术作品

我们在开发本系统的过程中，分别从算法原理和技术实现角度取得了一些学术成果，目前已整理成论文：

1. 为大赛匿名. Mtn: End-to-end Music Texture Transfer via Reconstructive Spectral Representation. 2018. (ISMIR 2018在投)
2. 为大赛匿名. MusiCoder—基于深度学习的轻量化音乐质感转换系统. 2018.

软件著作权

本系统已经申请了软件著作权《MusiCoder音乐质感转换系统》，正在审核。流水号为2018R11L350294。

开源协议

我们的开源项目采用Affero通用公共许可证 (AGPL)，保留一切相关权利。

祝您使用愉快