文件编号：从容应队-SWC2020-从容应队

受控状态：■受控 □非受控

保密级别：□公司级 □部门级 ■项目级 □普通级

采纳标准：CMMI DEV V1.2



[项目LOGO]

素音

**Divoice**

项目开发文档

**Version 1**

2019.12.19

**Written by 从容应队**

[Team LOGO]

**All Rights Reserved**

目录

[1 项目概述 1](#_Toc22846313)

[1.1 项目背景 1](#_Toc22846314)

[1.2 项目定位 1](#_Toc22846315)

[1.2.1 应用场景 1](#_Toc22846316)

[1.2.2 目标人群 1](#_Toc22846317)

[1.3 项目方案 1](#_Toc22846318)

[1.4 项目目标 1](#_Toc22846319)

[1.5 项目价值 1](#_Toc22846320)

[2 开发计划 1](#_Toc22846321)

[2.1 最终呈现形式 1](#_Toc22846322)

[2.2 主要功能描述 1](#_Toc22846323)

[2.3 运行环境 1](#_Toc22846324)

[2.4 验收标准 1](#_Toc22846325)

[2.5 关键问题 1](#_Toc22846326)

[2.6 进度安排 1](#_Toc22846327)

[2.7 开发预算 1](#_Toc22846328)

[3 可行性分析 2](#_Toc22846329)

[3.1 技术可行性分析 2](#_Toc22846330)

[3.2 资源可行性分析 2](#_Toc22846331)

[3.3 市场可行性分析 2](#_Toc22846332)

[4 需求分析 2](#_Toc22846333)

[4.1 数据需求 2](#_Toc22846334)

[4.1.1 静态数据 2](#_Toc22846335)

[4.1.2 动态数据 2](#_Toc22846336)

[4.1.3 数据词典 2](#_Toc22846337)

[4.1.4 数据采集 2](#_Toc22846338)

[4.2 功能需求 2](#_Toc22846339)

[4.2.1 \*\*功能模块 2](#_Toc22846340)

[4.3 性能需求 4](#_Toc22846341)

[4.3.1 时间特性 4](#_Toc22846342)

[4.3.2 适应性 4](#_Toc22846343)

[4.4 界面需求 4](#_Toc22846344)

[4.5 接口需求 4](#_Toc22846345)

[4.5.1 硬件接口 4](#_Toc22846346)

[4.5.2 软件接口 4](#_Toc22846347)

[4.6 其他需求 4](#_Toc22846348)

[5 概要设计 4](#_Toc22846349)

[5.1 处理流程 4](#_Toc22846350)

[5.2 总体结构设计 4](#_Toc22846351)

[5.3 功能设计 4](#_Toc22846352)

[5.4 用户界面设计 4](#_Toc22846353)

[5.5 数据结构设计 4](#_Toc22846354)

[5.6 接口设计 4](#_Toc22846355)

[5.6.1 外部接口 4](#_Toc22846356)

[5.6.2 内部接口 4](#_Toc22846357)

[5.7 错误/异常处理设计 5](#_Toc22846358)

[5.7.1 错误/异常输出信息 5](#_Toc22846359)

[5.7.2 错误/异常处理对策 5](#_Toc22846360)

[5.8 系统配置策略 5](#_Toc22846361)

[5.9 系统部署方案 5](#_Toc22846362)

[5.10 其他相关技术与方案 5](#_Toc22846363)

[6 数据库设计 5](#_Toc22846364)

[7 详细设计 5](#_Toc22846365)

[7.1 \*\*功能模块 5](#_Toc22846366)

[7.1.1 功能描述 5](#_Toc22846367)

[7.1.2 性能描述 5](#_Toc22846368)

[7.1.3 输入 5](#_Toc22846369)

[7.1.4 输出 5](#_Toc22846370)

[7.1.5 程序逻辑 5](#_Toc22846371)

[7.1.6 限制条件 5](#_Toc22846372)

记录更改历史

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **序号** | **更改原因** | **版本** | **作者** | **更改日期** | **备 注** |
| 1 | 创建 | 1 | 张澳 | 2019/12/19 | 创建项目开发文档 |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |

# 项目概述

## 项目背景

随着互联网的不断普及，中国互联网和移动互联网的发展也愈发成熟，人们手中的手机也愈发智能化，网络成为了我们接触世界的一扇极其重要的门扉，而手机则是人们叩开这扇大门的钥匙——相比电脑更加方便、灵巧。同时，在互联网这座庞大的信息海洋里，为了更好地展示自己的才能以及不同的生活场景，使更多的人能够足不出户便遨游世界，网络主播、up主等新型职业应运而生，并不断发展壮大，网络直播、录播、小视频、vlog等新型娱乐方式也早已走入人们的生活之中。

然而，在观看这些大多由个人制作的产品时，我们却发现了一个问题，大多数主播、up主受制于设备不专业的问题，且并没有掌握专业的音频编辑技能，从而导致产出音频、视频作品往往难以如人意，在吸引观众、粉丝上的竞争力也就难免逊色于同行了。这一问题也就将大多数欲闯入主播、up主这一行业的人拒之门外了。

普遍来说，解决音频降噪、混音等等音频问题的方法总结起来有两个：

1. 买更好的硬件，例如声卡、话筒等，这类设备依据性能也分为高中低三个档次，价格 从几百元到几万元不等，而价格偏低的产品也往往性能堪忧。而高价的产品也只是依据通用的算法来处理普遍的问题，对一些特殊的情况也会出现无法满足用户需求的情况，例如背景噪声实在太过嘈杂，甚至大于人（主）声的情况。
2. 寻找专业音频处理师或者调音师，为自己提供更加精细的、更加优质的音频处理服务，这样的情况下音频处理会更加多选以及符合用户的特定要求，但随着服务而来的是价格也愈发昂贵，而大多数个人主播、up主难以承受这样庞大的支出。

而除了上述情况外，主播要求不高的话，还可以通过找寻一个安静度极高的环境来降低环境噪音，但此举也只是满足了降噪的需求，无法实现对音频的多元化修饰，这样产出的作品吸引力也往往较低而无法达到预期。一些up主也会自学音频编辑处理的技能，但还是免不了每一次音频处理后长时间的编辑处理。

为解决用户降噪难，音频处理难的问题，结合如今正大热的机器学习技术以及新兴的快应用技术，本团队构想出Divoice，期望能够依据用户提供的音频，标记并分离出不同的声音，已方便用户实现降噪、主音凸显的功能，并提供一系列主流音频处理功能：混音，增强选定波段声音，添加音轨效果等辅助音频编辑功能。满足主播、up主以及普通用户对音频处理的需求，节省出80%以上音频编辑时间，降低音频、视频录制的门槛，使更多的普通大众也能享受到创作出优良的作品，给自媒体行业加速。

除此之外，Divoice也会随着用户提供个人专属音频素材以及反馈继续学习，更加精确识别与提取用户声音，为用户提供更加优质的音频处理服务。

## 项目定位

### 应用场景

无论是主播、up主还是普通用户，当用户需要在嘈杂的环境中录制音频时，当用户想要获取一段音频的伴奏或是纯净人声时，当用户苦于电脑处理音频太过繁琐时，当用户渴望听清嘈杂环境中的特定声音是，Divoice都可以为用户提供专业、优质、个性化的音频分离服务和强大的音频处理服务，用户仅需要提供自己录制的音频，并选择性的提供单一音色的声音素材，Divoice便可以依据强大的计算能力，智能分离当前音频中的不同声音并生成不同的音轨存储这些分离出的声音，再依据用户的选择对当前所有音频进行处理，快速、便捷地生成用户所需音频文件，极大地节省用户的时间。

### 目标人群

此项目适用于有音频处理需求的人，主要包括：

1. 主播、up主，帮助他们能够仅凭手机就能实现对自己录制文件的降噪、混音处理， 节省时间、降低成本，大大提高工作效率。激励普通大众也生成属于自己的特色音频/ 视频文件。
2. 企业的宣传部门与音频工作有关的工作人员，打造优质的企业之声。
3. 专业音频处理人员，为其提供纯净的声音素材，大大降低工作难度，灵活工作时间 与办公地点。

## 项目方案

## 项目目标

## 项目价值

# 开发计划

## 最终呈现形式

## 主要功能描述

## 运行环境

## 验收标准

## 关键问题

## 进度安排

## 开发预算

# 可行性分析

## 技术可行性分析

## 资源可行性分析

人力资源包括队员四人，包括深入学习过深度学习技术、掌握软件构架搭建、PPT制作以及视频制作能力突出的技术人员，也有组织能力突出的领导人员。

模型训练数据集：\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

Devoice需要一台服务器，由学校提供。\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

## 市场可行性分析

在当前市场上，手机端的音频处理软件并不多，并且大多以音频剪辑、合成以及语音识别为主，并没有能够达到音频分离功能的软件，这是市场的一大空白。在PC端的音频处理软件中，Adobe Audition、Audacity等部分软件能达到简单的音频分离，但分离效果较差，会造成音质的损坏，并且此类软件普遍操作复杂，入门困难，使用难度较大。

Divoice作为以音频分离为主的手机端快应用软件，主要有以下几点优势：

1. 以手机快应用的形式，无需下载，使用便捷，用户能够随时随地使用，而不需要求助于电脑等笨重的设备。
2. 以人手一把的手机为载体，注定了潜在用户数量众多，用户群体也不会局限于专业的音频处理人员，普通用户也能够在观看视频、直播时，分离出所需要的音频，获得更好的体验。
3. Divoice的分离方法是基于深度学习技术，用户只需输入原始音频文件，分离功能由软件自动实现，不需要用户自身额外的操作，使用方法简单便捷，入门容易。
4. 系统在运营过程中会不断地收集用户的反馈和评价，进而动态地更新模型，实现语音分离的高效率、高质量，为用户提供更加智能便捷的服务。

# 需求分析

## 数据需求

### 静态数据

### 动态数据

### 数据词典

### 数据采集

## 功能需求

### \*\*功能模块

表 核心功能模块描述

|  |  |  |  |
| --- | --- | --- | --- |
| **功能模块** | **功能** | **功能描述** | **优先级** |
|  |  |  |  |
|  |  |  |
|  |  |  |

表 \*\*\*\*用例规约

|  |  |
| --- | --- |
| **用例名称** |  |
| **功能简述** |  |
| **用例编号** |  |
| **执行者** |  |
| **前置条件** |  |
| **后置条件** |  |
| **涉众利益** |  |
| **基本路径** |  |
| **扩展路径** |  |
| **字段列表** |  |
| **设计规则** |  |
| **未解决的问题** |  |
| **备注** |  |

## 性能需求

### 时间特性

### 适应性

## 界面需求

## 接口需求

### 硬件接口

### 软件接口

## 其他需求

# 概要设计

## 处理流程

## 总体结构设计

## 功能设计

## 用户界面设计

## 数据结构设计

## 接口设计

### 外部接口

### 内部接口

## 错误/异常处理设计

### 错误/异常输出信息

### 错误/异常处理对策

## 系统配置策略

## 系统部署方案

## 其他相关技术与方案

# 数据库设计

# 详细设计

## \*\*功能模块

### 功能描述

### 性能描述

### 输入

### 输出

### 程序逻辑

### 限制条件