

机器翻译微信小程序

开发文档



2023-4-21

参与人员：谭玲玲 陈焯

目录

[一、项目介绍 2](#_Toc132999127)

[二、项目负责人 2](#_Toc132999128)

[三、项目需求 2](#_Toc132999129)

[四、用户需求 2](#_Toc132999130)

[用户功能需求： 2](#_Toc132999131)

[用户非功能需求 2](#_Toc132999132)

[五、系统需求 3](#_Toc132999133)

[对用户需求一 3](#_Toc132999134)

[对用户需求二： 4](#_Toc132999135)

[对用户需求三： 5](#_Toc132999136)

[对用户需求四： 6](#_Toc132999137)

[对用户需求五： 6](#_Toc132999138)

[对用户需求六： 7](#_Toc132999139)

[六、系统架构 8](#_Toc132999140)

[1、逻辑架构： 8](#_Toc132999141)

[2、开发架构 8](#_Toc132999142)

[3、物理架构 8](#_Toc132999143)

# 一、项目介绍

项目名称：TC机器翻译

# 二、项目负责人

谭玲玲

陈焯

# 三、项目需求

通过微信小程序使用后台机器翻译服务。

1、翻译服务配置

2、翻译服务集成

3、翻译

4、第三方OCR、语音识别和语音合成集合

# 四、用户需求

## 用户功能需求：

1、能够翻译各种语言的文字，例如中文、英语、法语、德语等。

2、能够对用户输入的语言进行语音识别，支持多种语言，并实现现代语语音转为文字的功能。

3、能够将用用户输入的文本或翻译后的文本，使用语言语音成功进行语言播放报。

4、能够提供高质量的OCR识别服务，支持多种文件格式，包括但不限于JPG、PNG等文件格式。

5、能够支持多种语言的OCR识别服务，包括但不限于中文、英语、法语、德语等。

6、应具有易于使用、直观和简洁的界面。

## 用户非功能需求

1、当用户输入需要翻译的文本，需要快速返回翻译后的文本，量化时间（5s以内）

2、用户输入需要翻译的文本，得到翻译后的文本，翻译结果需要保持一定的准确性。

3、当用户没有选择需要翻译的语言选择项时，需要提供一个“自动”选择项，如一般翻译软件默认将中文翻译为英文。

4、OCR功能应该允许用户使用手机相册里的照片，同时也可以运行用户自行拍摄照片来进行翻译。

5、OCR功能应能使用户自行决定照片翻译的位置范例。

6、小程序应记录用户的翻译历史，而用户可以在小程序中查看之前翻译的内容和结果。

7、语言合成功能应该允许用户听到他们所需要的语言的翻译结果。

8、用户可以在应用中设置语言偏好、字体大小等个性化选项。

# 五、系统需求

初期假设

用户进入翻译小程序，程序已经在系统中创建了一条记录，并收录了用户的流量，微信账号等信息。数据和历史信息并输出翻译结果。

正常状态

后台程序先分析用户的输入语言。

接着后台程序按照百度翻译API接口提供的翻译方法将用户输入内容翻译为用户所选择的语言并输出。

后台程序选择菜单选项来添加翻译历史信息。

有哪些会出错

用户信息不存在或找不到。系统应创建一条新记录并记录用户信息。

翻译接口中没有合适的用户期望翻译的语言。程序可以改用其他接口以提供更多的翻译语言。

输出窗口一直都没有显示内容。系统应限制输出程序反应时间，如果翻译内容不明或翻译不出，输出翻译错误。

其他活动

当用户正确在输入信息时，其他用户也可以输入信息并输出该用户的翻译结果。

完成的用户状态

用户登录。将包含历史记录在内的用户信息输入到数据库。记录被添加到系统日志中，系统日志显示和延时的延时间以及相关的用户。

对用户需求一：

* 系统应能够支持多种语言，包括但不限于中文、英语、法语、德语等；
* 用户在输入待翻译文本时，系统应能够进行自动识别语种，并选择对应的翻译模型；
* 系统应具备文本翻译的功能，能够将输入的文本自动翻译成目标语言；
* 系统应能够提供基础的翻译服务，包括但不限于单词、短语、句子等级别的翻译，且翻译质量应尽可能达到自然流畅的程度；
* 系统应支持用户输入多个待翻译文本，且能够在一次请求中同时翻译多个文本；
* 系统应支持用户自定义目标语言，能够将文本翻译成用户选择的目标语言；
* 系统应具有反馈机制，能够接收用户反馈信息并根据反馈信息来优化翻译质量；
* 在网络条件良好的情况下，系统应保持较高的响应速度，翻译时间应尽可能短；
* 系统应提供稳定的服务质量和数据安全保障，能够保证用户隐私数据的安全。

## **对用户需求二**：

**1、需求名称：**

* 语音识别和转换为文字

**2、需求描述：**

* 系统能够对用户输入的语音进行语音识别，并将语音转换为文字形式。系统要求支持多种语言的识别和转换。系统还需要能够在较短的时间内完成识别和转换。

**3、功能/特性：**

* 支持多种语言的语音识别和转换
* 快速、高效的语音转换为文字
* 可以识别不同的语音类型，例如正常语速、快速语速、口音等
* 提供可定制的设置，例如识别的精度、输出的格式等
* 支持多种输入方式，例如麦克风输入、文件输入等

**4、功能细节：**

* 系统支持多种语言的语音识别和转换为文字的功能。
* 用户可输入语音文件或者通过录音设备进行实时语音录入。
* 系统对输入的语音进行识别，并通过语音识别技术将语音转化为相应的文本。
* 系统将文本以可编辑的格式呈现给用户，方便进行后续的文本处理。
* 在语音识别和转化为文字的过程中，系统需要保证语音和文本的准确性和可靠性。

**5、输入：**

* 用户输入的语音

**6、输出：**

* 转换后的文字

**7、前提条件：**

* 用户有语音输入设备，例如麦克风
* 用户提供的语音清晰且符合识别的要求

**8、假设：**

* 系统使用了一些语音识别和转换的工具/库

**9、操作/步骤：**

* 用户通过语音输入设备输入语音
* 系统接收到用户输入的语音
* 系统对语音进行识别和转换，输出对应的文字结果
* 系统将结果返回给用户

**10、异常情况：**

* 用户输入的语音清晰度较差，导致无法识别和转换
* 用户输入了系统不支持的语音语言
* 系统出现错误/异常导致无法识别和转换

## 对用户需求三：

**1、需求名称：**

* 语音合成功能

**2、需求描述：**

* 用户希望系统能够将文本转换为语音，并进行语音播报。需要支持用户输入的文本和翻译后的文本两种情况。系统应该提供语音合成接口来完成这一功能。

**3、需求细节：**

* 语音合成应该支持多种语言，包括但不限于中文、英文等；
* 需要支持通过API调用第三方语音合成服务进行转换；
* 用户需要能够控制语音的音量、语速和语调，以满足不同场景的需求；
* 系统需要支持多种音频格式的输出，以便于在不同设备和场景下进行播放；
* 需要有错误处理机制，防止语音合成出现异常或错误情况；
* 应该提供用户可配置的参数，以满足不同用户的需求；
* 需要满足信息安全和隐私保护的要求。

**4、具体标准：**

* 系统能够准确地将用户输入的文本转换为对应的语音，并且能够语音播报；
* 用户可以自由控制语音的音量、语速和语调，以满足不同场景的需求；
* 系统支持多种主流语音合成服务提供商，并能够根据不同场景自动选择适合的服务商；
* 语音合成功能应该满足较高的精度和流畅度要求，用户能够准确地理解所播报的语音内容；
* 系统应该能够支持多种音频格式的输出，并保证音频质量；
* 系统应该提供灵活的配置文件，以便于管理员可根据实际情况进行调整；
* 系统应该能够安全保护用户的语音合成请求和相关数据，不泄露用户的隐私信息。

## 对用户需求四：

* 支持多种文件格式，包括但不限于JPG、PNG等文件格式。
* 实现高质量OCR识别服务，确保可靠性和准确性。
* 能够根据用户需求进行定制化设置，包括识别速度、精度等方面的调整。
* 能够处理大量文件，保证系统的并发处理能力。
* 提供易于使用的API接口，方便用户进行集成使用。
* 提供安全可靠的数据存储和保护机制，确保用户的隐私和数据安全。
* 能够实现可扩展性，以满足未来用户需求的变化。

## 对用户需求五：

**1、需求描述：**

* 系统应该提供支持多种语言的OCR识别服务，包括但不限于中文、英语、法语、德语等。用户应该能够灵活选择需要识别的语言。

**2、功能要求：**

* 系统应该支持中文、英语、法语、德语等多种语言的OCR识别。
* 用户可以选择需要识别的语言。
* 系统应该能够自动检测文本所使用的语言，进行准确的OCR识别。

**3、性能要求：**

* OCR识别速度应该快速响应，保证用户使用的流畅性。
* 识别准确率应该高，保证用户数据的准确性和完整性。
* 系统应该支持高并发，保证多用户同时使用时的性能表现。

**4、可用性要求：**

* 界面应该友好易用，用户可以轻松进行操作。
* 系统应该提供说明文档和操作指南，方便用户使用和了解系统。
* 系统应该具有良好的容错和错误处理能力，防止因系统故障或用户操作失误导致数据丢失或损坏。

**5、实现约束：**

* 系统应该使用成熟的OCR识别技术和算法，确保准确性和性能。
* 系统应该具有足够的存储空间，为多种语言的OCR识别提供充分的支持。
* 系统应该支持不同格式的文件输入和输出，方便用户使用和数据交互。

**6、具体标准：**

* 用户可以成功选择需要识别的语言，并获得准确的OCR识别结果。
* OCR识别速度快，准确率高，符合上述性能要求。
* 系统界面友好易用，容错处理能力好，符合上述可用性要求。
* 系统满足上述实现约束。

## 对用户需求六：

**1、需求描述：**

* 系统界面应该易于使用、直观和简洁，以满足用户需求。

**2、系统需求：**

* 系统应该具有直观的用户界面，使用户能够在不需要额外培训的情况下快速上手。
* 系统应该提供简洁的界面，避免过多的复杂功能或内容，以减少用户的操作难度和犯错率。
* 系统界面应该友好，对于特定群体的用户也应该尽可能考虑到其习惯或需要，以提高用户满意度。
* 系统应该提供易于导航的界面，使用户能够快速找到需要的功能或信息。
* 系统应该提供清晰、准确的文本提示和帮助信息，以帮助用户理解和操作系统。
* 系统应该提供符合易用性标准的图标和控件，以提高用户的操作效率和准确性。
* 系统应该与操作系统和硬件配置相适应，以保证系统的稳定性和性能。
* 系统需要进行用户体验测试，以确保界面设计和功能操作满足用户需求和期望。

# 六、系统架构

本套小程序系统由数据处理模块和微信小程序发布这两大模块构成。数据处理模块部署一台 web 服务器，web 服务器负责与小程序端和 API 进行信息传输。在微信小程序开发结束发布以后，代码上传至微信后台服务器，客户通过在微信中扫描或者搜索程序名即可从微信服务器中加载程序，然后通过微信小程序实现翻译查询。

**（一）系统功能设计**

数据处理模块主要是处理微信前端发送过来的数据和返回语音识别结果和翻译结果。微信小程序发布模块主要为用户提供一个输入需翻译内容的界面并且和数据处理模块实时通信。用户获得界面是从微信服务器，而翻译结果是经过数据处理模块返回获得。

**（二）系统功能实现**

1、数据处理模块

数据的处理主要分为3个方面，前端 MP3 文件的获取以及转码，请求语音识别和翻译数据，数据的处理和返回。

（1）前端 MP3 文件的获取及转码

在服务器接收到微信小程序的前端发送过来的 MP3 文件后，使用 UUID 进行文件的唯一重命名。根据 BaiduAPI 提供的文档，百度语音识别 API 接 口能够接受 MP3、PCM 等音频格式文件，但是最终 Baidu 后 台 处 理 的 还 是 PCM 文 件。 当 MP3 文 件 转 PCM 文件后，文件的体积会减少，这样可以减少数 据发送的时间。在考虑服务器带宽的情况下，本项目使用 PCM 文件格式进行上传，其中 MP3 文件转 PCM 文件使用 ffmpeg 进行转码。

（2）语音识别和翻译数据的请求将上一个步骤中转码成功的数据进行请求发送，同时对请求结果进行处理，当发现是用户使用错误时，没有识别结果，此时将信息返回，进行适当的 提 示。如果是百度Token验证过期，则重新拉取 token 验证，如果是因为 QPS 达到上限，则在 2~3 秒内进行随机等待，重新请求数据。

（3）数据的处理和返回根据微信小程序的前端需求，返回错误信息进行提示或者当数据正确的时候，返回语音识别和翻译后的信息进行前端显示。

2、微信小程序发布

（1）微信小程序开发环境的搭建

微信小程序开发要先在微信公众平台官网注册账号，完善小程序信息，获取唯一标识场景定制翻译小程序的AppID,然后在官网下载专用的开发工具。微信小程序的开发工具界面，主要包括微信小程序的代码编辑、代码调试和微信小程序模拟器功能。除此之外，在调用网络 API 的部分，需事先设置好通讯域名用于连接我们的数据处理模块的服务器。需要注意的是，微信小程序开发文档中说明了小程序不支持 IP 和端口设置，只支持 https和wss 协议。

（2）微信小程序界面设计

翻译小程序的界面设计主要是实现翻译语言选择，文本输入查询，语音输入查询，图片识别查询，翻译结果展示和历史记录的功能。本项目页面设计使用官方提供的 MINA框架。 MINA 框架提供了自己的视图层描述语言WXML和WXSS，以及基于JavaScript的逻辑层框架，并在视图层与逻辑层间提供了数据传输和事件系统，让开发者可以方便地聚焦于数据与逻辑上。它也提供了一套基础的组件，这些组件自带微信风格的样式以及特殊的逻辑，开发者可以通过组合基础组件，创建出强大的微信小程序。

整体框架为：

