

中國民航大學

《软件工程课程设计》

软件需求规格说明书



手语识别

组长学号姓名: 200340070 宋美杰

成员学号姓名: 200340071 宋荣泉

200340059 李坤壕

200340012 李星宇

200340069 石涵夫

200340018 孟琮昊

二〇二二年十一月

目录

第 1 章 引言	1
1.1 软件项目概述	1
(1) 项目意义	1
(2) 软件目标用户	1
1.2 软件功能概述	2
1.3 软件实现难点及特色分析	2
(1) 项目实施难点	2
(2) 项目特色	2
1.4 软件产品与其环境之间的关系	2
1.5 参考资料	3
第 2 章 软件项目需求描述	4
2.1 软件需求的用例模型	4
2.2 软件需求的分析模型	5
2.2.1 用例名：输入自然语言	5
2.2.2 用例名：获取手语视频	6
2.2.3 用例名：翻译自然语言为动画	7
2.2.4 将视频转换为自然语言	8
2.2.5 显示动画	9
2.2.6 显示自然语言	10
2.2.7 分析类图	11
第 3 章 其他需求描述	12
3.1 性能要求	12
3.2 交付要求	12
3.3 验收要求	12
第 4 章 软件原型	13

图目录

图 2-1 手语翻译用例模型	4
图 2-2-1 手语翻译系统“输入自然语言”用例的顺序图	6
图 2-2-2 手语翻译系统“获取手语视频”用例的顺序图	6
图 2-2-3 手语翻译系统“翻译自然语言为动画”用例的顺序图	8
图 2-2-4 手语翻译系统“将视频转化为自然语言”用例的顺序图	9
图 2-2-5 手语翻译系统“显示动画”用例的顺序图	10
图 2-2-6 手语翻译系统“显示自然语言”用例的顺序图	10
图 2-3 手语翻译系统分析类图	11
图 4-1 手语翻译系统 Android 端软件手语翻译界面	13
图 4-2 手语翻译系统 Android 端文字生成动画界面	13

表目录

表 1-1 软件功能描述表	2
表 1-2 软件与外界环境的交互关系	2

第 1 章 引言

1.1 软件项目概述

(1) 项目意义

软件项目的背景及主要功能：

2021 年中国有 8500 万残障人士，占人口比例的 6.2%。但是我们在生活中却很少看到残障人士。其中一个很大的原因是因为对残障人士的保障不足，很多残障人士“不敢”出门。而其中，听力障碍人士达到 2780 万人，占残障人士数量的 30%。我国 cctv 新闻频道设置专门的手语播送，为听力障碍人士提供了很大帮助。然而在生活中听力障碍人士有时会很难和非专业人士很难沟通。如果可以制造廉价的手语翻译工具，将会给众多听力障碍人士带来便利。

现有的手语翻译工具，要么过于简单，只能输入简单地文字，将文字拆分为词语，检索回对应的图片；要么造价过于昂贵，需要佩戴精密的外设，并且佩戴外设又为听力障碍人士带来新的不便。如果可以开发出新的手语翻译工具，可以将连贯的手语翻译成语句，或者将语句翻译成手语，那么可以方便残障人士的沟通。如果可以将手语翻译工具的便捷性以及廉价性变得友好，则可以使很多爱心或公益人士进行推广。如此以来，将大大提高社会对听力障碍人士的友好度，帮助构建和谐社会。

(2) 软件目标用户

手语翻译软件是以帮助聋哑人及公益人士翻译手语为主题，面向的用户主要为语言交流障碍人士，公益人士。A 类用户可以通过手语翻译与其他人进行交流；B 类用户可以与言语交流障碍人士进行沟通。

1.2 软件功能概述

表 1-1 软件功能描述表

功能名称	功能描述
手语翻译自然语言	通过对手语的分析，转化为自然语言
自然语言翻译手语	通过输入自然语言，转化为手语动画

1.3 软件实现难点及特色分析

(1) 项目实施难点

为方便任何人使用，需要多平台（安卓，苹果；PC 端有简单示例）进行开发；

应使软件尽可能资源消耗较小而出色的完成任务

减少用户的负担（如应尽量较少申请权限，软件体积不应过大）

(2) 项目特色

经调研分析，其他公益项目仅可对文字进行搜索，返回单个字（词）对应的图片，且仅支持网页端。本项目既有 app 可供离线使用，又有网页端方便用户临时使用。且 app “小而美”，网页端支持响应式布局。

1.4 软件产品与其环境之间的关系

本项目主要借助于手机，帮助用户实现手语与文字的相互转换。它所运行的外部环境如下表：

表 1-2 软件与外界环境的交互关系

系统组成部分	外部环境
云端服务器	Ubuntu 操作系统
手机客户端	Android 操作系统

网页端	任意设备
-----	------

1.5 参考资料

[Hands - mediapipe \(google.github.io\)](https://github.com/google/mediapipe): 谷歌开源 mediapipe, 内涵多种视觉算法, 网络小巧, 但文档资料较少。

[Android 开发者 | Android Developers \(google.cn\)](https://developer.android.com/): 安卓开发官方网站, 内涵安卓开发的大部分解决方案, 但部分文档错误, 落后。

[Three.js - JavaScript 3D Library \(threejs.org\)](https://threejs.org/): 基于 WebGL 的 3d 开发库。

第 2 章 软件项目需求描述

2.1 软件需求的用例模型

用户通过手语翻译系统可以将手语翻译为文字，也可以用文字生成手语动画。其中，将翻译自然语言为动画包含用户输入自然语言与将动画显示两个行为；而翻译视频为自然语言则包含从用户获取手语视频并显示自然语言给用户两个行为。

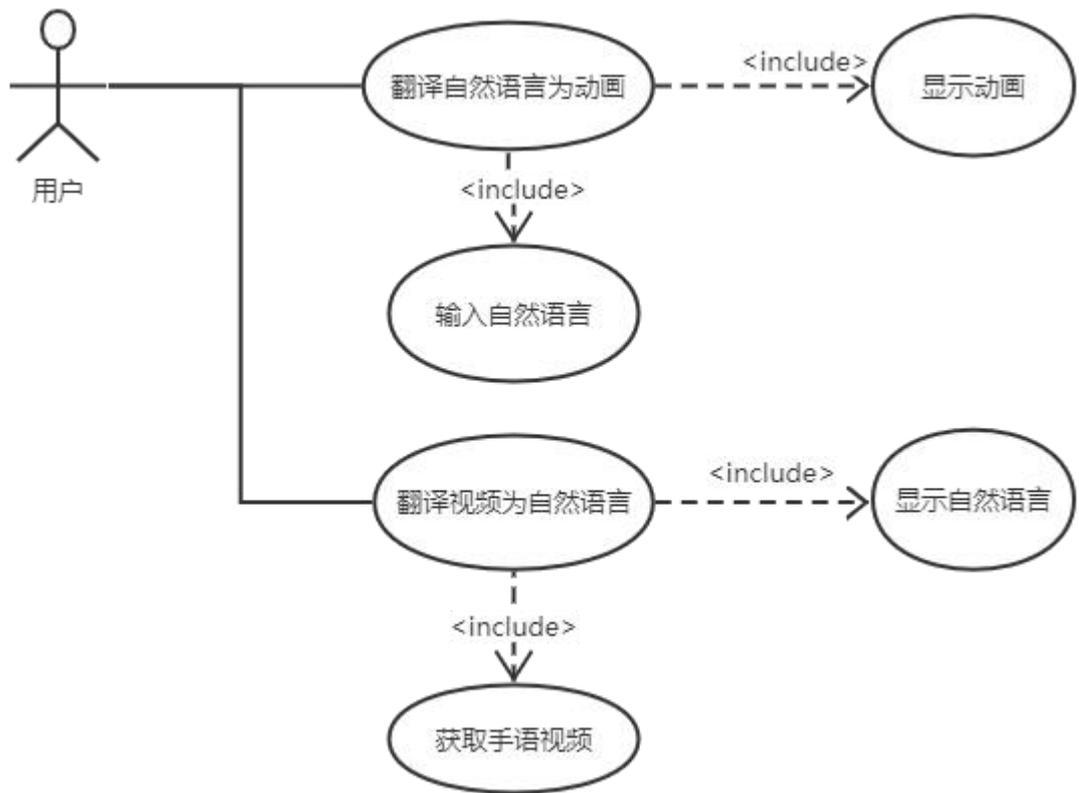


图 2-1 手语翻译系统用例模型

2.2 软件需求的分析模型

2.2.1 用例名：输入自然语言

(a) 输入自然语言

业务目标：用户输入需要翻译的自然语言

执行者：用户

前置条件：app 版本大于 4.0

基本交互动作：

- 1.用户打开手语翻译 app
2. app 跳转到生成手语动画页面
- 3.用户点击输入框，输入自然语言
- 4.app 接收自然语言

扩展交互动作：

1a: 修改输入内容

1a1:直接在输入框进行文本编辑

后置条件：输入数据无损失

可靠性需求： 数据安全性无限接近 100%

(b) 输入自然语言用例交互图

用户进行输入时，在输入框进行编辑，若有错误直接更改即可；完毕点击提交按钮，系统将获取输入内容，并与词库进行比对分析处理。

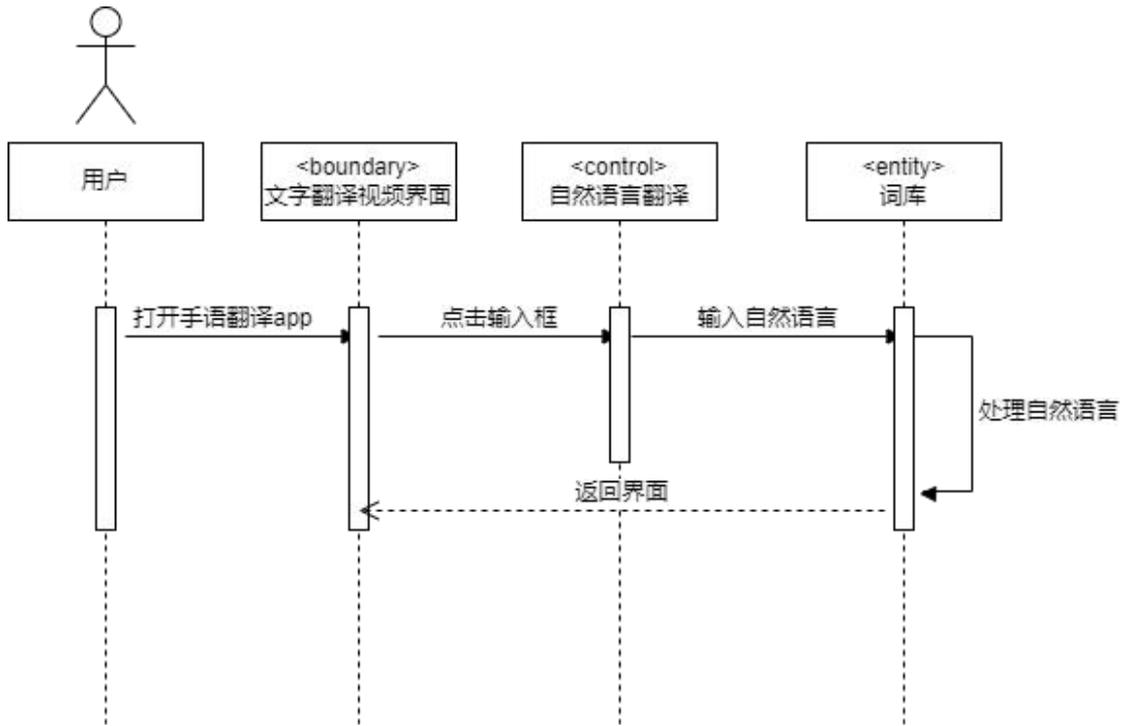


图 2-2-1 手语翻译系统“输入自然语言”用例的顺序图

2.2.2 用例名：获取手语视频

(a) 获取手语视频

业务目标：用户通过摄像头上传视频到 app

执行者：用户

前置条件：android 版本大于 4.0 且拥有摄像头

基本交互动作：

- 1.app 跳转到主页面（手语翻译文字界面）
- 2.app 请求获取摄像头权限
- 3.用户允许 app 开启摄像头
- 4.app 获取摄像头视频流
- 5.对视频进行初步处理，以便后续使用

后置条件：输入数据无损失

可靠性需求：视频解码无损

(b) 获取手语视频用例交互图

用户使用 app 手语翻译文字功能时, 需要从摄像头获取视频。系统取得对捕获到的视频进行拆帧等处理, 等待后续备用。

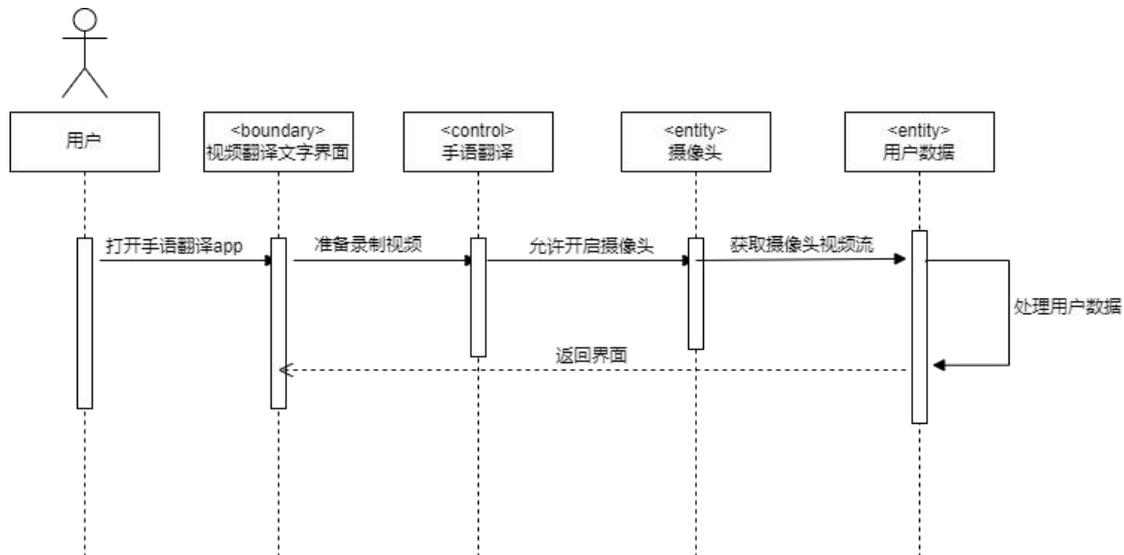


图 2-2-2 手语翻译系统“获取手语视频”用例的顺序图

2.2.3 用例名：翻译自然语言为动画

(a) 翻译自然语言为动画

业务目标：用户输入需要翻译的自然语言

执行者： 用户

前置条件： 用户已输入自然语言

基本交互动作：

1. 用户点击“Submit”按钮
2. app 对获取自然语言进行翻译
3. 生成动画，进行播放

扩展交互动作：

后置条件： 在词库内有对应数据时，可生成完整动画

可靠性需求： 响应事件小于 200ms

(b) 翻译自然语言为动画用例交互图

系统将之前获取到的文字进行拆分，查询，找到数据库中对应的动画信息。

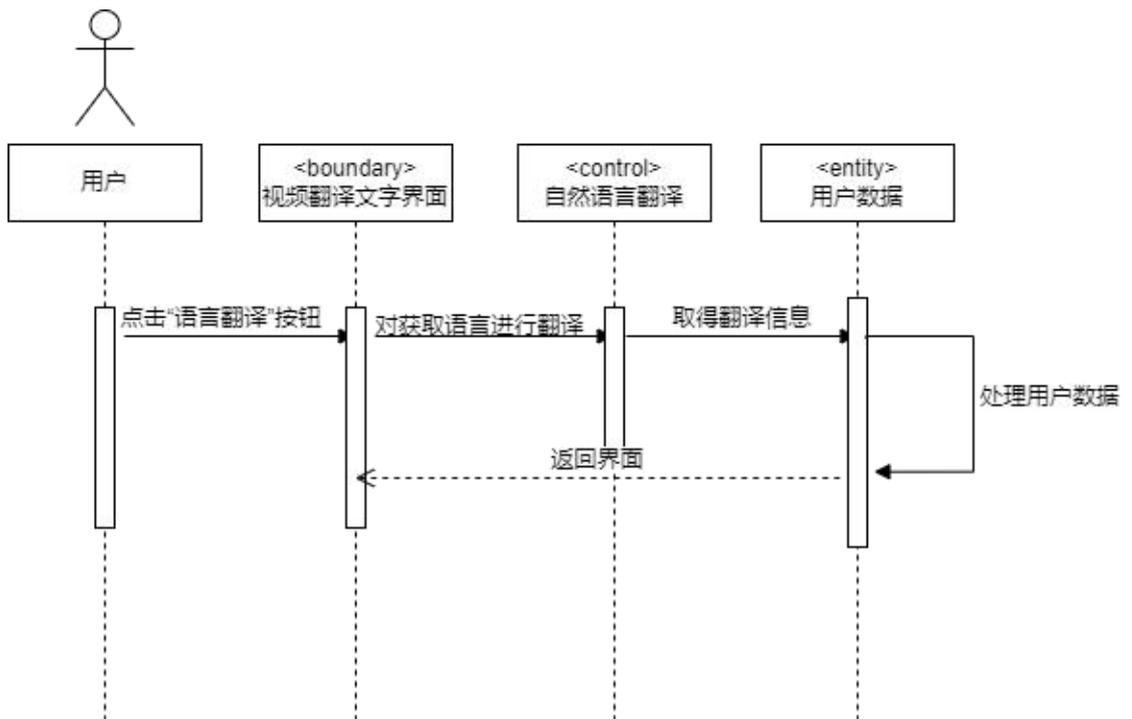


图 2-2-3 手语翻译系统“翻译自然语言为动画”用例的顺序图

2.2.4 用例名：将视频转换为自然语言

(a) 将视频转换为自然语言

业务目标：用户输入需要翻译的自然语言

执行者：用户

前置条件：获取到有信息的视频

基本交互动作：

1. 获取到有效视频帧
2. 对视频帧进行信息提取
3. 数据库进行查询
4. 显示相应文字

扩展交互动作：

后置条件：当视频内容有效时，根据视频可生成对应自然语言

可靠性需求：每秒处理速度达到 30fps

(b) 视频转化为自然语言用例交互图

进入视频翻译界面，对获取到的视频帧进行信息提取，数据库查询获得文字

信息。

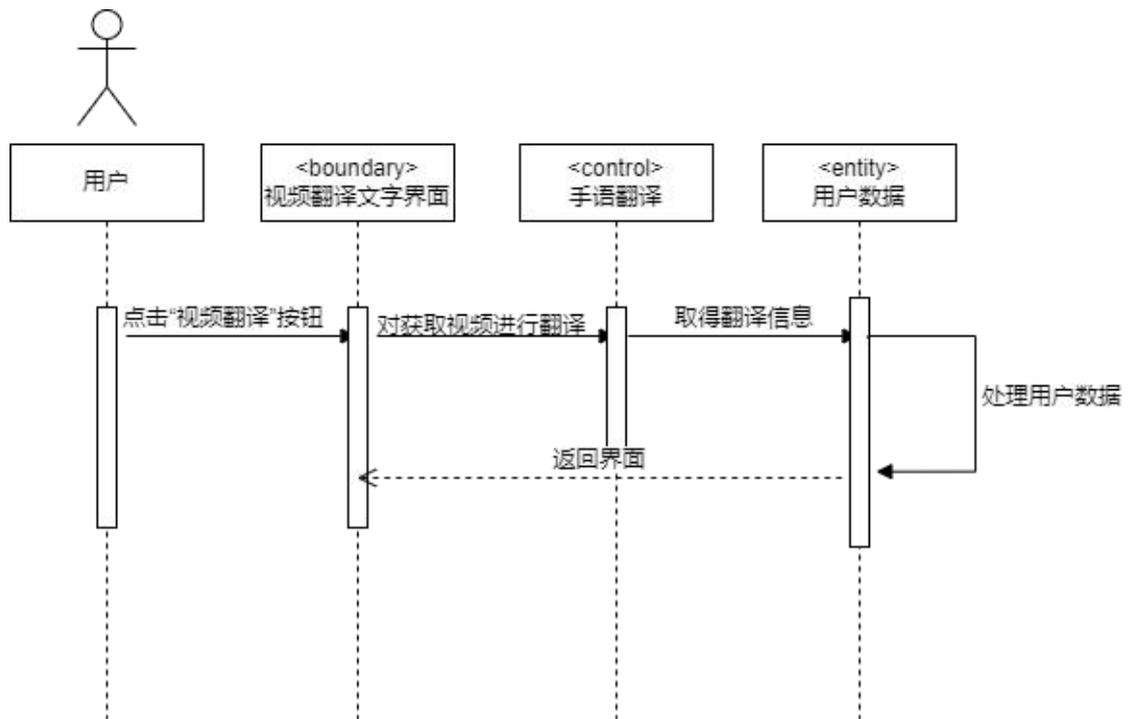


图 2-2-4 手语翻译系统“将视频转化为自然语言”用例的顺序图

2.2.5 用例名：显示动画

(a) 显示动画

业务目标：在 app 显示数据转换后的 3D 动画

执行者：用户

前置条件：app 已经获取 3D 动画数据

基本交互动作：

1. 系统获取到动画信息
2. 根据动画信息生成动画

扩展交互动作：

后置条件：当动画生成后，在屏幕指定位置展示动画

可靠性需求：动作混淆错位等错误发生率为 0%

(b) 显示动画用例交互图

根据动画信息，生成动画。

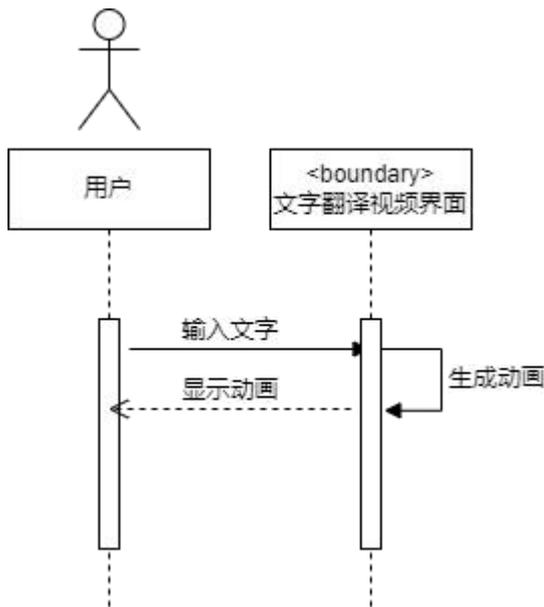


图 2-2-5 手语翻译系统“显示动画”用例的顺序图

2.2.6 用例名：显示自然语言

(a) 显示自然语言

业务目标：在 app 显示数据转换后的自然语言

执行者：用户

前置条件：app 已经获取处理后的自然语言

基本交互动作：

- 1.系统获取到文字信息
- 2.将文字信息展示在文本框

扩展交互动作：

后置条件：当收到自然语言时，在屏幕指定位置生成文字

可靠性需求：保证数据准确度为 100%

(b) 显示自然语言

获取文字信息，显示文字在文本框。

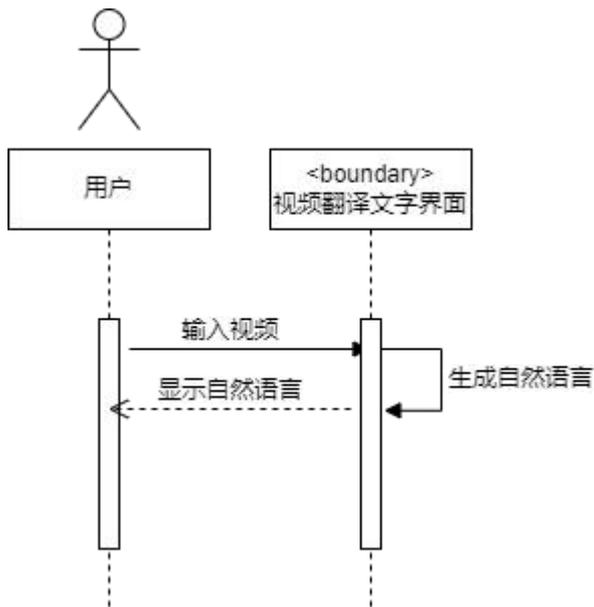


图 2-2-6 手语翻译系统“显示自然语言”用例的顺序图

2.2.7 分析类图

手语翻译类需要从视频翻译文字界面的 `meida_input()` 获取视频信息，进行信息提取后再通过词库获取文字信息，最后通过 `word_output()` 进行文字显示；自然语言翻译类需要从文字翻译视频界面获取文字，通过用户数据进行动画信息查询，再通过 `hands_3d` 进行动画生成。

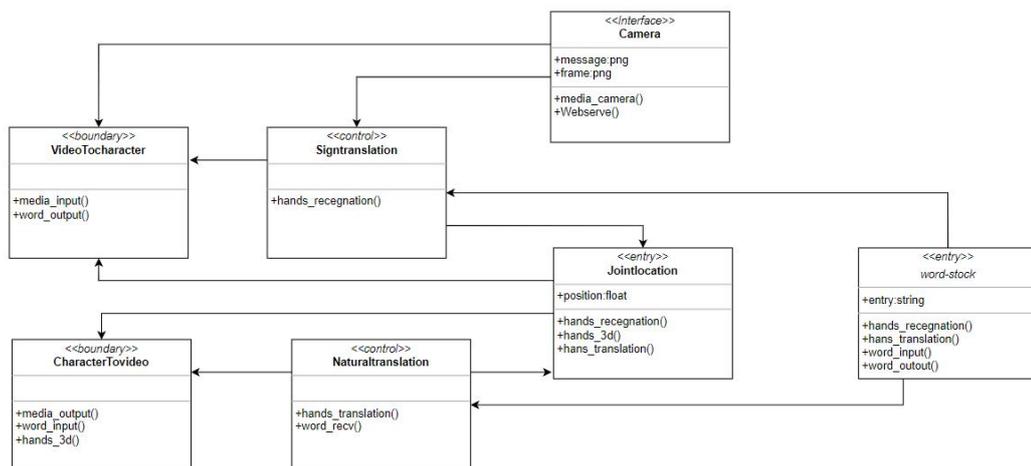


图 2-3 手语翻译系统分析类图

第 3 章 其他需求描述

3.1 性能要求

1).分析时间要求:为保证分析的流畅性,每帧视频的翻译时间应小于 30ms,即达到每秒 30fps 的流畅度;

2).数据回传时间:对于云端部分,任何信息应在 1s 内完场请求响应以及分析全过程,确保及时性;

3).可靠性要求:将视频翻译为文字应该准确率在 80%以上,文字生成动画应完全准确(词库有对应数据时)。

3.2 交付要求

交付内容:

1) .web 服务端源码及软件、app 端源码、网页端源码。

2) .软件设计规格说明书的电子文档

3) .使用说明书

3.3 验收要求

1).要求整个系统正常运行过程中无 Bug,能在用户非正常操作的情况下报告错误但不至于崩溃。

2).要求在光线明亮时可以准确识别手语并翻译。

3).要求整个系统各部分优化完毕,不存在编程过程中遗留的调试代码等影响用户体验的部分。

第 4 章 软件原型

1) 本界面为手语翻译系统的手语翻译界面，通过对手语视频解析生成对应文字。

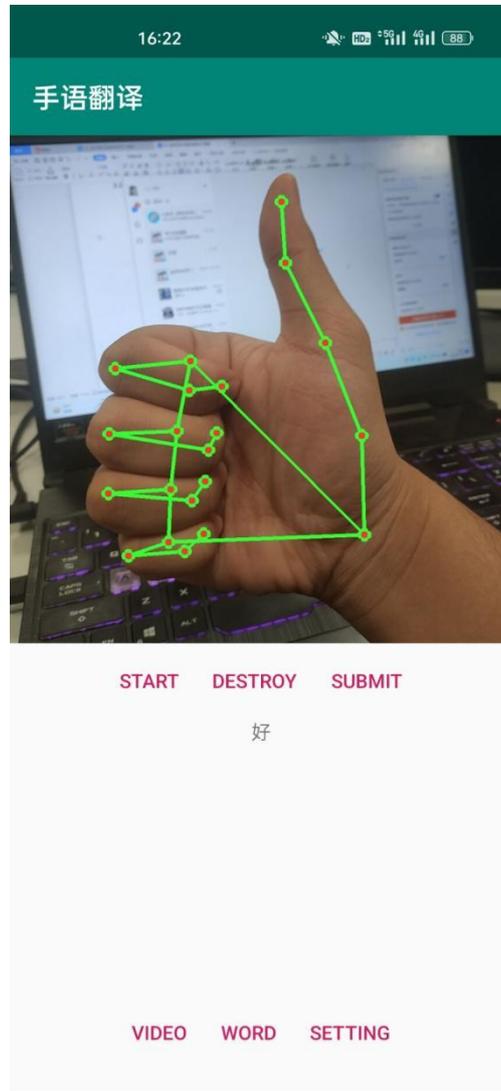


图 4-1 手语翻译系统 Android 端软件手语翻译界面

2) 本界面为手语翻译系统的文字生成动画界面，通过输入文字生成手语的相应动画。

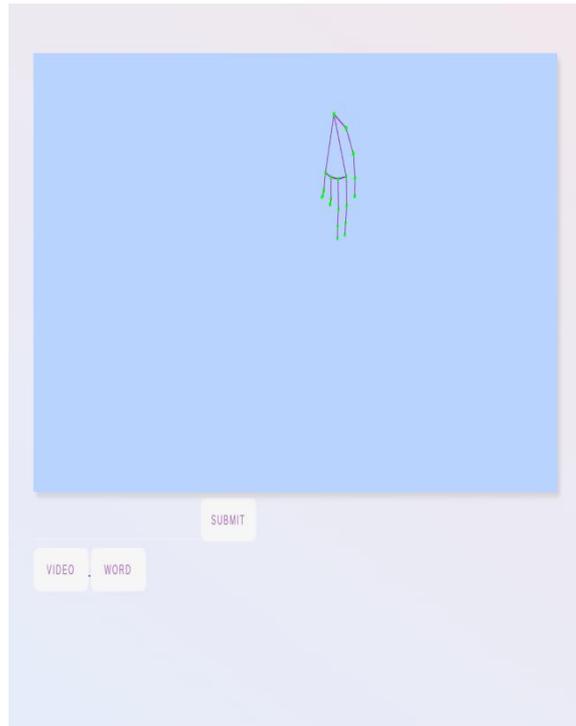


图 4-2 手语翻译系统 Android 端文字生成动画界面