

基于 Live2D+Mlkit 的社交软件的设计与实现

马文艺, 姜丽莉, 范言文

(南京工业大学浦江学院 计算机与通信工程学院, 江苏 南京 210044)

摘要: 系统是一款采用 Live2D 模型方式、通过 Mlkit 技术捕捉面部特征、达成类似视频式社交互动的即时通信软件。在现有的 IM 软件中, 最为常见的是纯文本、语音、视频、图文并茂这 4 类形式。此系统在现有 IM 软件功能的基础上, 将常见的视频通话方式改为以模型化的方式代替本身样貌进行通话。目前运用 Live2D 技术的领域也在增多, 尤其是快速发展的直播行业, 在目前 3D 技术与硬件条件不匹配的情况下, 使用 2D 技术进行折中, 可以很好地解决硬件条件不足的问题。

关键词: Live2D; Mlmit; 即时通信; 社交

中图分类号: TP311

文献标识码: A

文章编号: 2096-4706 (2022) 11-0037-04

Design and Implementation of Social Software Based on Live2D+Mlkit

MA Wenyi, JIANG Lili, FAN Yanwen

(School of Computer and Communication Engineering, Nanjing Tech University Pujiang Institute, Nanjing 210044, China)

Abstract: This system is an instant messaging software that uses Live2D model method, captures facial features by Mlkit technology, and achieves a video-like social interaction. In the existing IM softwares, the most common four types are plain text, voice, video, pictures and texts. On the basis of the existing IM software functions, this system changes the common video call method to a modeled method instead of their own appearance to make a call. At present, the fields of using Live2D technology is also increasing, especially in the rapid development of the live broadcasting industry. In the current situation that 3D technology does not match the hardware conditions, using 2D technology as a compromise can well solve the problem of insufficient hardware conditions.

Keywords: Live2D; Mlmit; instant messaging; social contact

0 引言

社交软件已经深入人们的生活, 而人们使用社交软件更多的是在以文字、语音、朋友圈等形式在网络上进行交流。据此可见, 人们对于视频社交并不敏感, 仅有极少部分的情况会选择以视频形式进行互动。个人认为不选择使用视频形式有以下两种原因: (1) 对于非亲友或熟识的人, 大部分人不愿意暴露自己真实样貌; (2) 不是所有人都有好看的外表, 有自信去向他人展现。

基于此, 以模型化代替本身样貌进行社交应能成为较优解, 模型化形式可以在保证互动性情况下, 有效避免本身样貌的显露, 有助于线上互动聊天。对于模型, 目前正热门的 Live2D 是一个不错的选择, 目前该技术运用领域也在增多, 尤其是对于目前 3D 技术与硬件技术存在不匹配矛盾的直播行业, 该技术未来还是会有很好的应用前景。

1 系统总体架构

系统架构分为五层, 从下至上分别为: 基础设施层、存储层、业务层、接入层、移动终端。

1.1 业务层

业务层为整个系统的中心, 负责为整个系统提供服务, 响应移动端用户的操作, 将数据返回以及存储到存储层。业务层包括通讯业务、登录业务、账号服务、好友服务、消息服务、聊天室服务, 其中即时通信业务包括音频通信、文字通信、虚拟视频通信业务。

1.2 移动终端

移动终端也就是客户端, 使用 Android 技术开发的 APP, 用户通过 APP 可实现即时通信功能。本系统核心技术 Live2D+Mlkit 运行在移动终端, Mlkit 调用移动终端的相机获取实时视频流, 调用 CPU 对视频流进行处理与计算, 得到实时的人脸坐标数据, 通过自定义的数据转换器转换为 Live2D 配置数据, 完成 Live2D 与人脸动作的实时同步。

1.3 其他层

基础设施层包括操作系统、服务器、存储设备, 主要负责服务的提供与数据的存储; 存储层包括 MySQL、OSS (对象存储), MySQL 是关系型数据库主要存储用户、聊天室、聊天记录等信息, OSS 主要存储系统中的资源文件, 如用户头像等; 接入层用于连接 Android APP 与服务端, 使用 OkHttp3 技术, 按照 Resultful 规则, GET 用于获取各种系统数据, POST 用于新增系统数据, PUT 用于更新系统数据, DELETE 用于删除数据; 用户的每个操作都是向业务层发送请求, 业务层再操作存储层的数据, 为了保证系统易于

收稿日期: 2022-04-19

基金项目: 江苏省高等学校大学生创新创业训练计划项目 (202113905005Y)

维护,在业务层与存储层都需要加入日志服务,用于记录用户操作日志,当系统遇到 Bug、崩溃等问题时可根据日志进行回溯。社交软件需要实现即时通信的功能,本系统使用 WebSocket 技术在客户端与服务端建立持久的连接,保证信息的即时性。

综上所述,系统总体架构如图 1 所示。

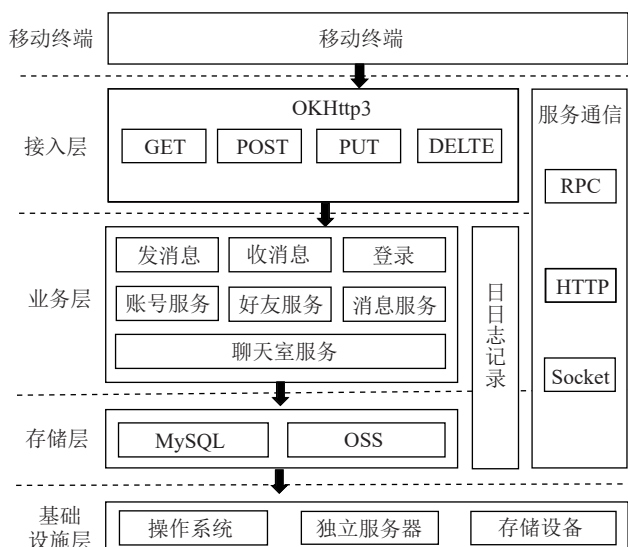


图 1 系统总体架构图

2 系统开发技术与工具

本系统客户端采用 Android 技术,服务端使用 Java 后端技术,通过客户端向服务端、服务端向客户端发送数据进行交互。

2.1 Android 相关技术

Live2D: 日本 Cybernoids 公司开发的一种 2D 绘图渲染技术,多用于电子游戏,通过一系列的连续图像和人物建模来生成类似三维模型的二维图像。

Mlkit: Google 的机器学习套件,是一个移动端 SDK,提供了强大易用的视觉和自然语言 API 来解决应用程序中常见的问题。

Fastjson: Alibaba 的一个可以将 Java 对象转为 JSON 数据、将 JSON 转为 Java 对象的 Java 库。

OkHttp3: 一个主流的网络请求的开源框架。

2.2 服务端技术

Spring Boot: 为快速启动且最小化配置的 Spring 应用而设计,简化了传统 Spring 应用的开发。简化了烦琐的配置项目、提供了内置的 Http 服务器、简化了 Maven 的依赖配置。有助于开发者快速构建架构,进行开发,帮助开发者专注于需求业务的开发。

WebSocket: 一种在单个 TCP 连接上进行全双工通信的协议。

2.3 数据库存储技术

MySQL: 关系型数据库,主要用于存储用户、聊天室等数据。

Redis: Key-Value 型数据库,用于用户 Token 数据的缓存。

3 需求分析

3.1 客户端功能

客户端通过 OkHttp3 和 WebSocket 与服务端通信,为用户提供包括账号登录、账号创建、完善个人信息、文字通信、语音通信、虚拟视频通信、搜索好友、添加好友、修改好友备注、删除好友、创建聊天室、删除聊天室等功能。其中文字通信需要支持颜文字的发送,语音通信拥有关闭麦克风、关闭扬声器功能;虚拟视频通信拥有相机翻转、切换成语音功能。用户功能用例图如图 2 所示。

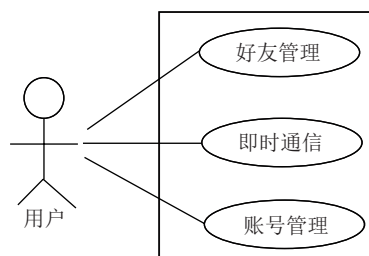


图 2 用户功能用例图

3.2 服务端功能

服务端功能应该包括权限验证、Token 颁发、WebSocket 服务、数据提供服务、数据存储服务等功能。权限验证需要验证用户 Token 是否合法、是否过期等功能;Token 即为用户访问服务端的令牌,每位每次登录都会生成独一无二的 Token;WebSocket 服务需要支持用户文字、语音、视频通信,同时还包括接受用户语音请求、视频请求以及请求结果的消息,实时反馈给用户。输入提供服务提供了调用业务层功能的 API,此服务提供即时通信以外的其他服务;数据存储服务用于随时存储用户产生、更新的各种数据,如聊天室、好友关系、聊天记录等数据。

3.3 管理端功能

管理端为管理系统中的各种数据提供了可视化的管理界面,还提供数据分析、数据统计等功能。管理员功能用例图如图 3 所示。

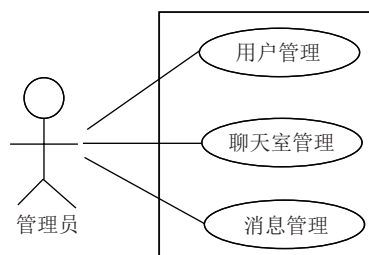


图 3 管理员功能用例图

3.4 系统总体功能框架

综合 3.1 ~ 3.3 所述,本系统的总体功能框架图如图 4 所示。

4 系统核心功能实现

4.1 Android 整合 Live2D 与 Mlkit

整合 Live2D。主流的 Android 应用开发使用 Android

Studio 作为编译器, Gradle 作为包管理器。Live2D 官方版本已达到 4.0 版本, 但官方并未提供 Jar 包, 只提供了 Native 的 SDK, 所以本系统引入了 2.0 版本的 Live2D 的 Jar 包。参考官方 Demo 编写自定义的 Live2D Manager, 通过该管理器可以自定义更换 Live2D 模型和传入自定义动作的配置。

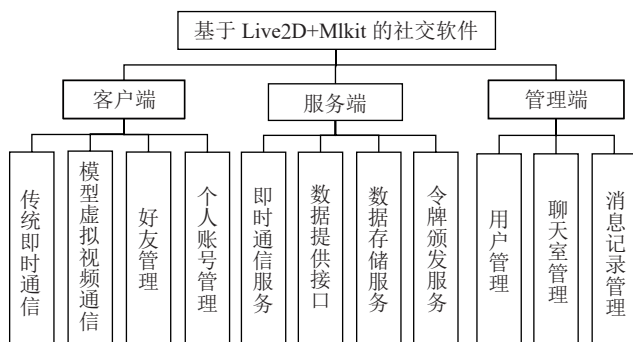


图 4 系统总体功能框架图

整合 Mlkit。Mlkit 是谷歌专门为移动端提供的机器学习套件, 提供了非常强大的人脸和 NPL 技术的 API, 使用 Gradle 可轻松整合。通过 Android 提供的相机服务, 实时获取相机的视频流, 经过 Mlkit 处理, 实时返回人脸坐标数据。

4.2 人脸与 Live2D 模型动作实时同步

4.2.1 处理人脸坐标数据

每个人脸坐标数据都是不同的, 首先本系统需要录入用户的坐标数据作为参考, 通过坐标数据的变换来判断眨眼、张嘴等动作。例如眼部的坐标如图 5 所示



图 5 眼部坐标图示例

以眼部坐标为例, 可以看到眼部一共有 16 个左边点, 这些左边对应着 X 轴、Y 轴, 可以看到坐标 4、5、11、12 处于眼部中间, 理论上只需要将坐标 4 的 Y 减去坐标 12 的 Y 得到的值无限接近 0, 即可知道眼睛的状态。

眼睛单纯的睁闭映射到 Live2D 模型上会显得僵硬和突然, 可通过以下公式计算眼睛的开合度, 会使模型的动作更加逼真。 H 代表某位用户眼睛睁开时 Y 轴相减的高度, Y_u 代表当前用户上眼的 Y 轴坐标, Y_d 代表下眼的 Y 轴坐标, H_r 代表模型最终的开合度, 最终公式为:

$$H_r = (Y_u - Y_d) / H$$

Live2D 眼睛的开合度只支持 1、0, 而嘴巴支持的开合度为 ± 30 , 如果是计算嘴巴的开合度, 公式可改为:

$$H_r = (Y_u - Y_d) / H \times 30$$

4.2.2 将人脸坐标数据转为 Live2D 动作配置数据

Live2D 有一套标准的配置参数表, 通过配置参数即可更改模型的脸部动作, 如图 6 所示。



图 6 模型动作更改示例

通过编写一个数据转换器将人脸数据转换为模型配置, 最终实现人脸与模型动作的实时同步, 最终效果如图 7 所示。



图 7 人脸与模型动作同步示例

需要注意的是, 在虚拟视频通信中, 双方和相机需要保持在一定的距离范围内, 距离误差越大, 实际检测结果误差也就越大。

4.3 文本通信的实现

文本通信属于本系统中实现起来较为容易的功能, 流程为: (1) 用户发送文本消息; (2) 服务端接收并解析, 获取到接收人并分发; (3) 接收端接收到消息后, 通过 Android 广播的方式更新视图并显示。

4.4 音频通信的实现

相比于文本通信, 音频通信需要接收方接收请求才可进行音频通信, 具体流程为: (1) 用户发起音频通信请求; (2) 服务端接收到请求并解析, 获取到接收人并分发; (3) 接收端接收到请求后, 唤起音频接收 Activity, 用户可选择接受或拒绝; (4) 接受端接收或拒绝的响应会发送到服务端, 服务端进行解析, 并将结果分发; (5) 发起端接收到结果, 如果对方接受则开始音频流传输, 如果拒绝则将终止当前 Activity。

4.5 虚拟视频通信的实现

虚拟视频在音频通信的基础上增加了获取用户实时视频流并转换为 Live2D 模型配置, 在接收方接受后, 不仅会传输音频流, 还将 Live2D 实时配置数据传输给接收方, 接收方接收后通过 Live2D Manager 实时更改模型的动作, 也就是将发起方人脸动作同步给接收方模型, 从而实现了所谓的虚拟视频通信。在此期间并没有传输用户的视频流, 只是传输了模型的配置数据。

4.6 虚拟视频通信转为音频通信的实现

用户在虚拟视频通信过程中, 本系统支持转为音频通信, 流程为: (1) 用户 1 点击转为音频通信按钮时会直接终止本地相机, Live2D 配置数据传输也会停止, 与此同时会发送消息给服务端; (2) 服务端接收到消息后, 解析并分发

给用户 2; (3) 用户 2 接收到此消息后, 随即也会终止本地相机的工作和 Live2D 配置数据的传输, 音频数据流传输正常进行。

5 结 论

本系统虽然最终实现了即时通信的功能, 但仍然存在一些问题: (1) 系统使用的是 Live2D 2.0 版本的 SDK, 而 Live2D 版本已经到达 4.0, 新的版本支持的功能更多, 经过多版本的迭代, 性能也会更好, 本系统的 APP 运行占用了大量的系统内存, 对于旧版的手机还是会存在性能问题; (2) Milkit 技术也存在一些问题, 因是实时视频流处理, 且是运行在移动终端, 同样会占用过多的系统内存, 影响用户的体验, 同时对眼睛开合检测的准确率略低; (3) 本系统的即时通信功能基于 WebSocket 进行传输, 对于普通的文字消息用户发送到服务端, 服务端接收后分发给目标用户, 期间的延迟并不明显; 对于音频流使用 WebSocket 传输就会显得吃力, 延迟更高。可以使用 RTC 技术实现客户端点对点通信来解决这个问题。

这些问题分别对应着软件、硬件、技术方面, 在各方面技术高速发展的今天, 或许不久以后就可以得到完美的解决, 当然也有可能今天就可以得到解决, 只是自己掌握的知识不够。

5G 的时代, 传统的 IM 通信已经成为人们生活、工作中不可分割的一部分, 虽然提高了人们的工作效率, 随之而来的是越来越大的工作压力。虚拟人物视频社交或许可以成为人们释放压力、进行娱乐的一种新方式。

参考文献:

- [1] 龙军. Java 多态性浅析 [J]. 科技视界, 2012 (21): 123-124.
 - [2] 陈涵生. Java 语言与 Java 技术 [J]. 计算机工程, 1997 (4): 3-6.
 - [3] 刘云浩, 李沛. 基于 SpringBoot 的后台服务器开发 [J]. 中国科技信息, 2018 (17): 61-62.
 - [4] 张峰. 应用 SpringBoot 改变 web 应用开发模式 [J]. 科技创新与应用, 2017 (23): 193-194.
 - [5] 小阿哥. SpringBoot 编程思想 [M]. 北京: 电子工业出版社, 2019.
 - [6] 陈木鑫. SpringSecurity 实战 [M]. 北京: 电子工业出版社, 2019.
 - [7] 高勇, 马思伟, 宋博闻. 国内虚拟主播产业链发展现状及趋势研究 [J]. 新媒体研究, 2020, 6 (1): 10-14.
 - [8] 陈阔. 三维点与二维图像协同的面部表情识别方法研究 [D]. 北京: 北京工业大学, 2016.
 - [9] 常一孜. 运动捕捉技术工艺流程及数据处理——面部运动捕捉及数据处理 [J]. 现代电影技术, 2016 (1): 39-42.
 - [10] 邱丽梅. 基于人脸特征点和线性回归的 3D 人脸姿态估计方法 [J]. 三明学院学报, 2008, 25 (4): 390-394+460.
 - [11] FETTE L, MELNIKOV A. The WebSocket Protocol [J]. Proceedings of the ACM on Programming Languages, 2011, 6455: 1-71.
- 作者简介: 马文艺 (1998—), 男, 汉族, 江苏灌南人, 本科在读, 研究方向: Web 应用开发; 姜丽莉 (1985—), 女, 汉族, 江苏徐州人, 讲师, 硕士, 研究方向: 内存数据库研发; 范言文 (2001—), 男, 汉族, 江苏灌南人, 本科在读, 研究方向: 大数据应用开发。

(上接 36 页) 有以下特色:

- (1) 平台分为不同级别, 适用不同层次、不同年级的学生需求。
- (2) 循序渐进的提升学生的设计、创新能力。
- (3) 将三维数字化和仿真等技术, 应用到机械类课程教学中, 更能可视化、清晰化, 展现教学内容, 便于讲解和观摩。
- (4) 数字化、虚拟仿真实验室平台, 解决了实验设备跟不上课程要求、设备陈旧、设备老化等问题。
- (5) 在计算上进行模拟设计、仿真, 可充分发挥自己的设计思维, 提高了学生对未来职业的理解, 同时让学生在在校期间就能了解企业对人才能力的要求, 提高学生的动手及设计创新能力。

5 结 论

通过机械创意创新数字化虚拟仿真平台的设计, 实现不同年级、不同层次的学生需求, 让学生所学理论知识与实践动手、企业需求相结合, 提高学生数字化仿真、软件使用等能力, 为全国大学生各类数字化、虚拟仿真等创新设计竞赛奠定基础。同时, 通过不同平台的动手实践操作, 循序渐进的提高大学生的设计创新能力。

参考文献:

- [1] 李平, 毛昌杰, 徐进. 开展国家级虚拟仿真实验教学中心

- 建设提高高校实验教学信息化水平 [J]. 实验室研究与探索, 2013, 32 (11): 5-8.
 - [2] 李红, 陈步云, 徐理军, 等. 构建实验教学质量评价体系提升实验教学质量 [J]. 实验室研究与探索, 2009, 28 (6): 234-236+246.
 - [3] 余建潮. 依托实验教学示范中心加强学生实验能力的培养 [J]. 实验技术与管理, 2013, 30 (3): 117-119+129.
 - [4] 张敬南, 张镠钟. 实验教学中虚拟仿真技术应用的研究 [J]. 实验技术与管理, 2013, 30 (12): 101-104.
 - [5] 邵琰, 刘敏超. 虚拟仿真教学实验平台在应用型大学的搭建及虚拟仿真技术在化工类课程教学中的探索性研究 [J]. 广东化工, 2015, 42 (14): 243-244.
 - [6] 王晓暄. 虚拟仿真实验教学中心建设研究 [J]. 新疆农机化, 2020 (1): 40-42.
 - [7] 罗晓东, 尹立孟, 王青峡, 等. 基于虚拟仿真技术的实验教学平台设计 [J]. 实验室研究与探索, 2016, 35 (4): 104-107.
 - [8] 丁羽, 李传峰, 李平. 基于创新能力培养的三维实体设计课程教学研究 [J]. 新疆农机化, 2018 (4): 41-43.
 - [9] 丁淑辉, 李学艺, 王海霞. 基于应用型创新人才培养的三维设计教学方法研究 [J]. 教育教学论坛, 2015 (33): 259-260.
- 作者简介: 丛日永 (1982.02—), 女, 汉族, 辽宁朝阳人, 高级工程师, 硕士研究生, 研究方向: 机械制造及自动化、机械设计及理论、机械工程方面研究。