

中国软件杯 大学生软件设计大赛



“乐闻”视障辅助资讯软件 用户使用说明书

参赛学校： 河海大学

组 名： 菜菜求求奖奖

组 长： 秦骁

组 员： 王子潇、范德儒、顾书宁

指导老师： 陈慧萍

目录

1. 引言.....	3
1.1 编写目的.....	3
1.2 项目背景	3
1.3 参考资料	3
2. 总体设计.....	4
2.1 需求概述	4
2.2 软件结构	5
3. 程序描述.....	6
3.1 注册登录模块.....	6
3.2 新闻发布模块	8
3.2.5.1 新闻分类预测算法.....	10
3.2.5.2 命名实体识别算法.....	10
3.3 新闻搜索模块	12
3.4 新闻推荐模块	14
3.4.5.1 基于用户的协同过滤算法.....	15
3.4.5.2 基于新闻的协同过滤算法.....	16
3.4.5.3 面向视障用户的基于行为的融合推荐算法.....	17
3.5 新闻朗读模块	19
3.6 操作提示模块	21
3.7 拍照识别模块	23

1. 引言

1.1 编写目的

为明确软件需求、安排项目规划与进度、组织软件开发与测试,撰写本文档。
本文档供项目经理、开发人员参考。

1.2 项目背景

近年来,人工智能的迅速发展,这将深刻改变人类社会生活、改变世界。我国注重人工智能产业的发展,致力于建设创新型国家和世界科技强国,人工智能的水平在不断提高,人们的对人工智能产品的期望值不断提升,一些带有身体缺陷的人士也渴望通过人工智能产品提升自身的生活品质。人工智能技术中,图像文字识别、语音文字转换、自然语言处理等任务在近些年来取得关键性的突破,电子产品的普及,为视障人士(半盲和低视力)的无障碍阅读提供了无限可能。

然而,目前多数的资讯软件都是面向普通人群开发的,功能局限于新闻推送、新闻详情显示、收藏新闻等,未考虑到视障人士的使用需求,不附带语音提示、行为识别等功能,需要用户精确点击才能精准操作,这对视障人士来说是十分困难的,即使具备新闻播放的功能,也是面向普通用户收听的需求,视障人士难以精确定位想要阅读的新闻或点击播放按钮。

本产品致力于打造一款能够辅助视障人士获取新闻资讯的软件,借助如上所提到的技术,我们通过移动云提供的 AI 能力接口和本团队训练的深度学习模型,实现了辅助视障人士获取资讯的用户友好型的软件,其包括丰富的系统语音指示、手势识别操作、语音交互等功能。该项目的开发可以明显提高视障人士的新闻阅读体验。

1.3 参考资料

[1] 张海藩编著,《软件工程导论(第5版)》,清华大学出版社

[2] 沈顺天.微信小程序项目开发实战[M].北京:机械工业出版社.

[3] Craig Walls. Spring Boot 实战[M]. 北京: 人民邮电出版社.

[4] KarypisG. Evaluation of item-based top • n recommendation algorithms[C] // Proceedings of the tenth international conference on Information and knowledge management. ACM, 2001: 247-254.

2. 总体设计

2.1 需求概述

“乐闻”视障人士辅助资讯软件提供一下几个业务模块:

(1) 注册登录模块

- ① 登录功能
- ② 注册功能

(2) 新闻发布模块

- ① 新闻类型识别功能
- ② 新闻实体识别功能
- ③ 新闻上传功能

(3) 新闻搜索模块

- ① 按类别搜索功能
- ② 按作者搜索功能
- ③ 按实体搜索功能

(4) 新闻朗读模块

- ① 浏览记录统计功能
- ② 用户偏好统计功能
- ③ 文本转语音功能

(5) 操作提示模块

- ① 预选中定位提示功能
- ② 确认提示功能

(6) 新闻推荐模块

(7) 拍照识别模块

① 图像文字识别功能

② 图像内容识别功能

2.2 软件结构

本系统总体分为七个模块，注册登录模块、新闻上传模块、新闻搜索模块、新闻朗读模块、操作提示模块、新闻推荐模块、拍照识别模块。

注册登录模块：视障用户通过微信小程序登录，进入小程序时向微信服务器发送请求，服务器识别用户唯一的标识码来判断用户是新用户或是老用户。若是新用户，则为其注册，初始化所有用户参数；若是老用户，则允许其直接登录。

新闻发布模块：新闻发布者可以上传新闻标题、新闻简介、新闻内容，并添加新闻的标签，系统对新闻的类别和标题中所含的实体进行识别，自动添加新闻标签。上传成功后，用户可在系统中查看。

新闻搜索模块：用户可以向系统输入一段语音进行查询，系统接收语音后对语音进行实体识别，然后对新闻库进行查询，对新闻的标题、作者、简介内容进行模糊查询，在新闻实体信息中匹配相关的新闻。

新闻朗读模块：用户在新闻概览界面收听新闻标题后可以选择感兴趣的新闻，确认选择进入新闻后，更新用户的阅读偏好信息、浏览记录、新闻信息。并系统播放新闻详情，包括新闻标题、新闻简介、新闻浏览量、新闻发布时间、新闻作者、新闻收藏量、新闻内容。

操作提示模块：用户触击和滑动都会触发相应的语音提示功能。用户单击屏幕，系统会提示目前所处界面的功能和如何触碰将引发何种事件；用户双击屏幕，系统会收到确认信号，进入下一级界面或者开始播放此块按钮的文字语音内容；用户向上下滑动屏幕，系统会根据用户滑动的方向来移动窗口位置；用户向左滑动屏幕，如果存在上级界面，系统会返回上级界面。

新闻推荐模块：对于新用户，其将接收基于新闻热度和用户自定义喜好的冷启动推荐系统的推荐新闻；对于有浏览、收藏记录的老用户，面向视障人士的综合基于用户和新闻的协同过滤算法得到的推荐新闻，当用户和新闻浏览和收藏数据量越庞大，推荐系统的性能越佳。

拍照识别模块：用户可以选择拍照识别文字和物体。拍照识别文字指：用户

拍摄书籍或报纸上印刷体的文字，系统播报识别语音来辅助视障人士阅读实体书。

拍照识别物体指：用户可以拍照识别身边或远处的东西，系统播报识别协助辨别其类别、特征、状态。

本系统的软件结构图如图 2.1 所示。

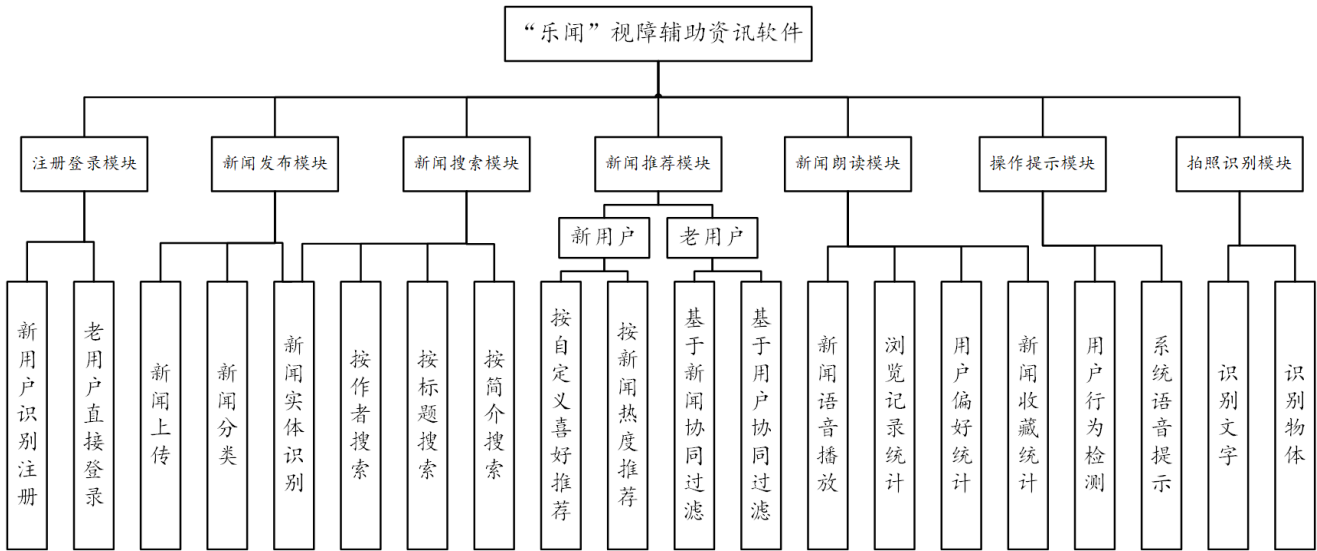


图 2.1 软件结构图

3. 程序描述

3.1 注册登录模块

3.1.1 功能

登录：用户通过微信小程序扫码进行登录，当微信服务器接收到用户的登录请求，识别用户的唯一标识码已经注册过，即允许用户直接登录进入系统。

注册：用户通过微信小程序进入系统，当微信服务器接收到用户的登录请求，发现用户的标识码未在应用软件服务器注册过，则自动为用户注册，初始化用户信息，注册成功后允许用户进入系统。

3.1.2 性能

准确性：用户的唯一标识码必须是有效的

及时性：用户扫码尝试小程序后，系统及时验证，给出登录或者注册反馈

时效性：用户的登录验证需在一定网络延迟内完成

3.1.3 输入项目

登录注册模块输入项如表 3.1 所示：

表 3.1 登陆注册输入项

名称	标识	数据类型	输入方式
用户标识符	User_code	Char	系统自动输入

3.1.4 输出项目

注册登录验证的反馈信息：

- (1) 验证成功：进入系统
- (2) 首次登录：自动注册
- (3) 网络延时过高：用户登录请求过期

3.1.5 算法

用户扫码进入微信小程序时，系统对用户标识符进行识别验证，将用户标识符发送给微信服务器端，得到用于验证的唯一标识符的返回结果。如果唯一标识符已经在数据库中，说明用户曾经注册过，允许其直接登录进入系统；如果唯一标识符不在数据库中，说明用户是首次使用本微信小程序，系统自动为其注册，将用户唯一标识符记录进数据库。

注册登录模块算法流程图如图 3.1 所示：

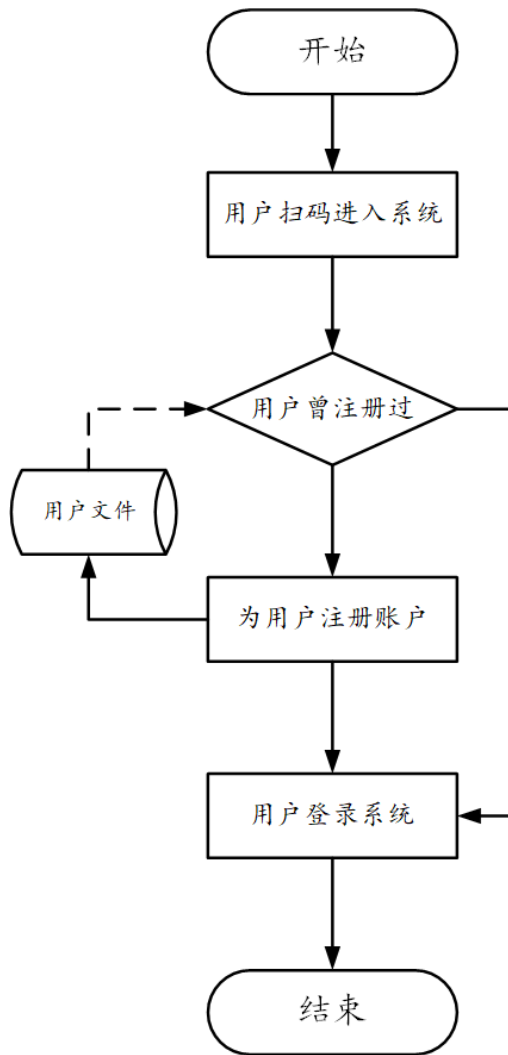


图 3.1 注册登录模块算法流程图

3.1.6 测试要点

- (1) 测试是否正确判断用户为首次登陆
- (2) 测试用户信息是否正确写入数据库
- (3) 测试能否正确得到验证消息反馈

3.2 新闻发布模块

3.2.1 功能

新闻上传：新闻发布者通过服务端发送新闻的标题、简介、内容，并且可以手动添加一些新闻标签，系统自动截取系统时钟作为新闻发布时间。

新闻分类：新闻发布者上传新闻后，新闻内容经过新闻分类模型得到新闻的类别，为新闻添加类别属性。

新闻实体识别：新闻发布者上传新闻后，新闻标题和简介经过实体识别模型，为新闻添加实体标签。

3.2.2 性能

及时性：发布者上传新闻后，新闻经过分类和实体识别要在一定时间内完成，使用户能够及时获取最新新闻。

准确性：新闻发布的内容要与作者上传一致

3.2.3 输入项目

新闻发布模块输入项如表 3.2 所示：

表 3.2 新闻发布模块输入项

名称	标识	数据类型	有效范围	输入方式
新闻编号	Nid	Int	数字字符	系统自动输入
新闻标题	Ntitle	Char	50 字符以内	手动输入
新闻简介	Nabstract	Char	200 字符以内	手动输入
新闻内容	Ncontent	Char	5000 字符以内	手动输入
发布时间	Time	Int	1970 计时法	系统自动输入
实体集	Tags	Char	人名、地点、机构实体	系统自动输入
作者	Writer_id	Int	数字字符	系统自动输入
浏览量	Views	Int	数字字符	系统自动输入
收藏量	Stars	Int	数字字符	系统自动输入

3.2.4 输出项目

新闻发布的反馈信息：发布成功/失败

3.2.5 算法

3.2.5.1 新闻分类预测算法

项目组成员采用将卷积神经网络 CNN 应用到文本分类任务，利用多个不同 size 的 kernel 来提取句子中的关键信息，从而能够更好地捕捉局部相关性。

假设每个句子中的单词可以表示为 k 维词向量，一个含有 n 个词的句子就可以由 $n * k$ 维的向量表示，Text-CNN 利用形状为 $h * k$ 的卷积核对句子的 $n * k$ 向量进行卷积操作，得到向量 $(n - h + 1) * 1$ 维的向量 c ，再对 c 做最大池化，得到 $1 * 1$ 维的特征 \hat{c} 。可以选择多个大小不同 (h 不同) 的卷积核重复如上卷积池化操作，得到由 m 个卷积核抽取的特征 $z = [\hat{c}_1, \hat{c}_2, \dots, \hat{c}_m]$ ，可以通过全连接层得到文本的分类结果。

新闻分类预测的算法流程图如图 3.2.5.1 所示：

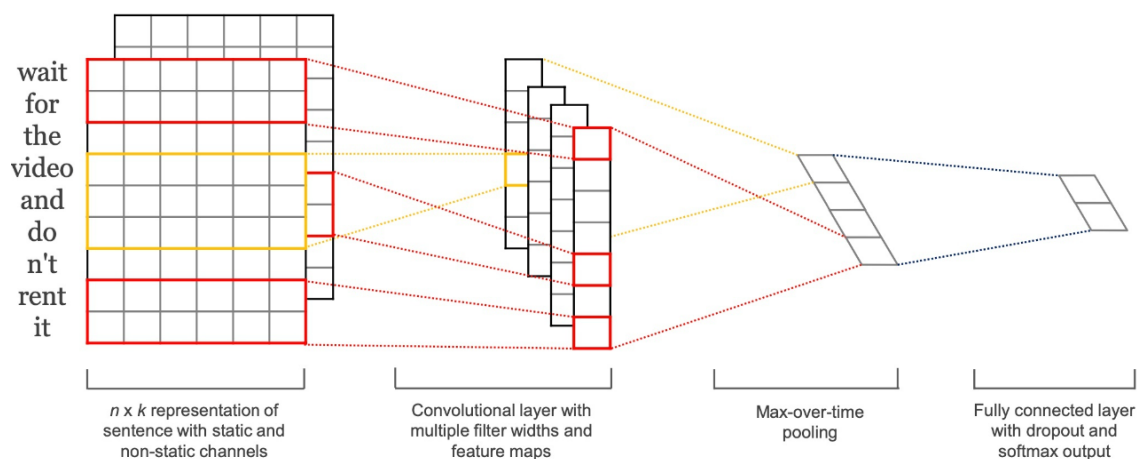


图 3.2.5.1 新闻分类预测的算法流程图

3.2.5.2 命名实体识别算法

项目组成员对新闻数据集进行数据采集和清洗等预处理，并构建 [CLS] 和 [SEP] 标注的新闻样本集，向 Bert 预训练模型输入训练数据集，进行新闻领域的二次训练，通过评价指标迭代优化模型参数，最终得到可以识别语句中的实体的模型。

命名实体识别的算法流程图如图 3.2.5.2 所示：

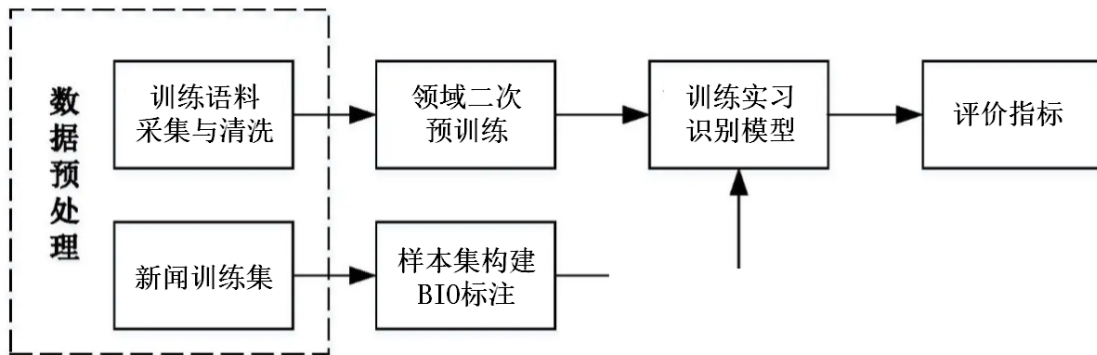


图 3.2.5.2 命名实体识别的算法流程图

3.2.5.3 总体算法流程

新闻发布者输入想要发布的新闻标题、简介、内容，并且可以手动添加相关标签，点击上传，系统会自动进行新闻类型预测和新闻标题的实体提取，从外部时钟自动获取上传时间，将新闻添加到数据库中。

新闻发布模块的总体算法流程图如图 3.2.5.3 所示：

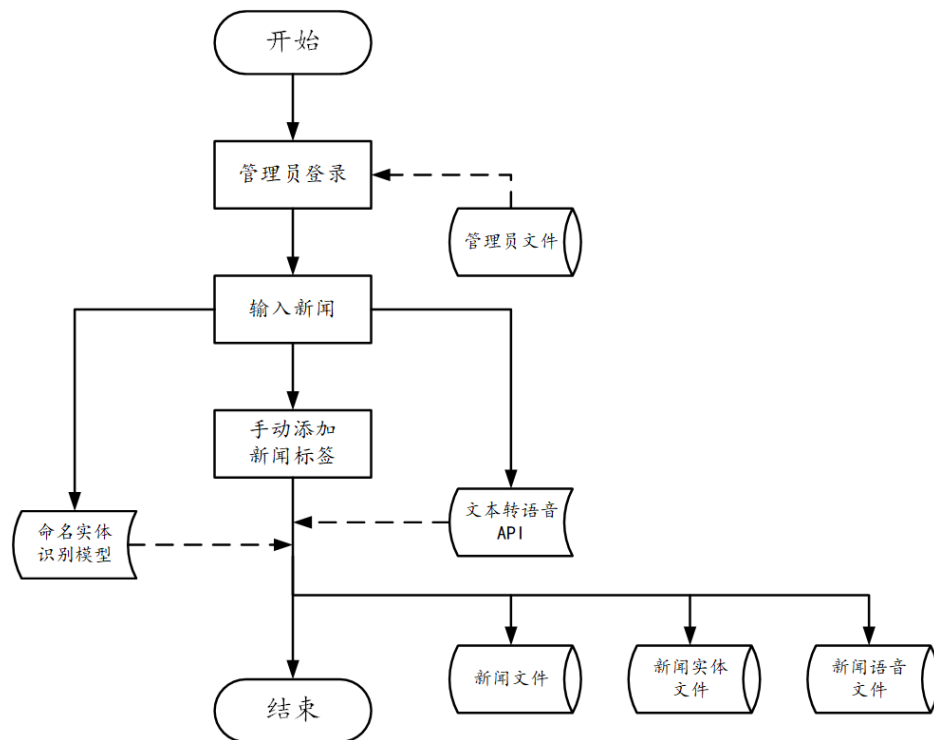


图 3.2.5.3

新闻发布模块的总体算法流程图

3.2.6 测试要点

- (1) 验证用户上传的新闻标题、简介、内容、实体是否有误
- (2) 验证模型识别的新闻类别、实体、上传时间是否有误
- (3) 验证新闻的语音转换结果识别有误

3.3 新闻搜索模块

3.3.1 功能

针对视障用户，系统采用语音输入的搜索方法。用户输入语音，系统先将语音转换成文本，对文本进行实体识别，然后将文本和识别到的实体在数据库中对新闻标题、简介、作者和新闻实体对照表进行查询，将结果按浏览量排序一起返回给用户。

3.3.2 性能

- 准确性：对新闻的搜索结果要求具有较高准确率
- 高效性：对新闻的搜索和结果返回要在短时间内完成

3.3.3 输入项目

新闻搜索模块输入项如表 3.3 所示：

表 3.3 新闻搜索模块输入项目

名称	标识	数据类型	有效范围	输入方式
搜索语音	Query_voice	Wav	语音长度 20s 以内	语音输入

3.3.4 输出项目

新闻搜索的结果：新闻标题搜索结果；新闻简介搜索结果；新闻实体搜索结果；新闻作者搜索结果。

3.3.5 算法

用户上传搜索语音，系统自动调用语音识别 API，将语音转换成搜索文本，将系统对搜索文本进行实体识别，将识别的实体和原搜索文本都加入查询语句，在数据库中对新闻标题、新闻简介、作者、新闻实体表进行模糊查询。将查询结果一起返回给用户。

新闻搜索模块的算法流程图如图 3.3 所示：

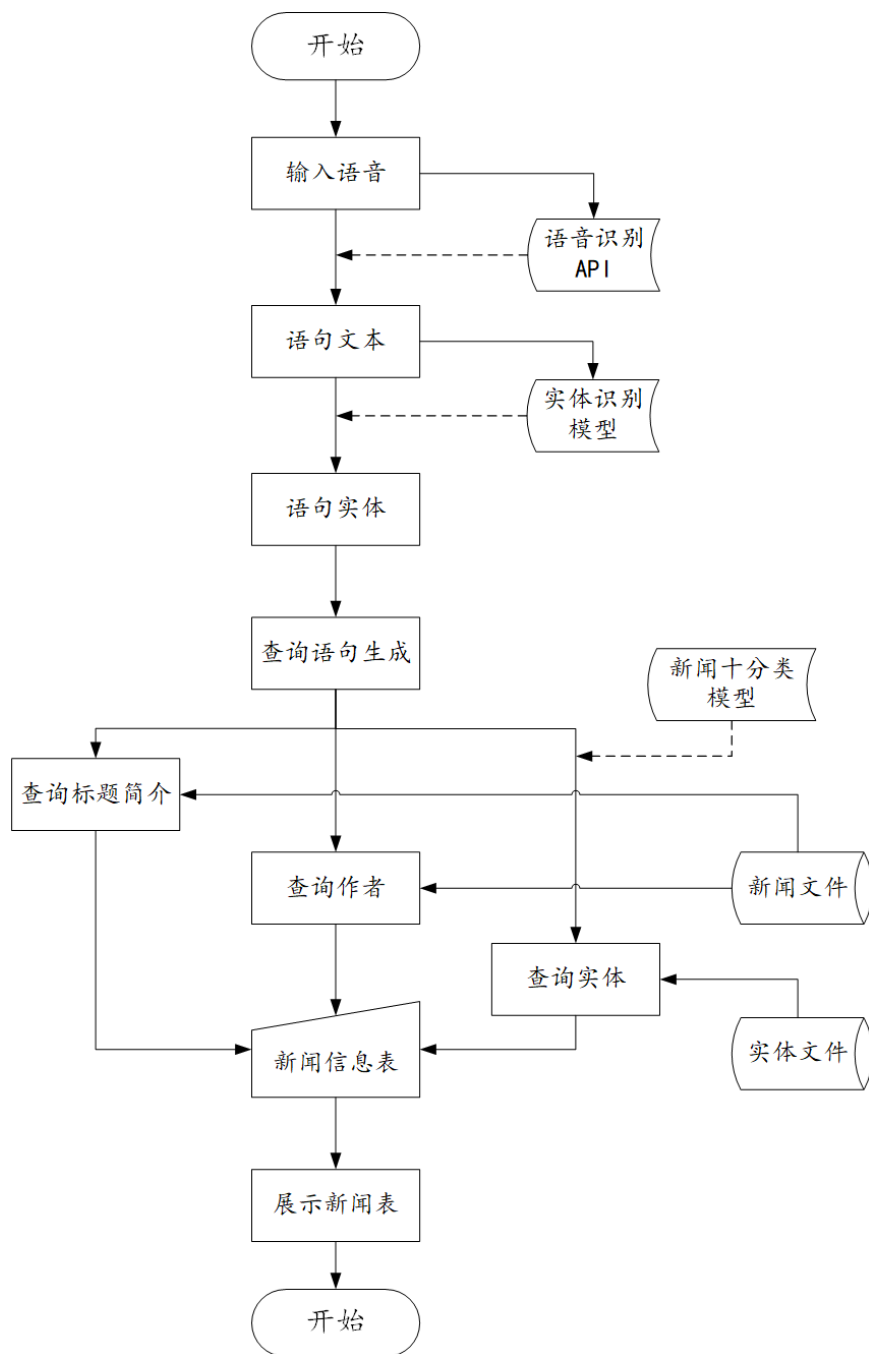


图 3.3 新闻查询算法流程图

3.3.6 测试要点

- (1) 测试是否能够接收用户语音
- (2) 测试是否能够正确调用语音转文本 API 及其准确性
- (3) 测试是否能够正确识别搜索文本的实体及其准确性
- (4) 测试是否能够在数据库中通过查询语句正确查询
- (5) 测试是否能够正确返回新闻的查询结果

3.4 新闻推荐模块

3.4.1 功能

根据用户喜好、浏览情况、收藏情况，针对性地为其推荐新闻。

3.4.2 性能

准确性：对于新闻推荐的结果需要较高准确率，精准推送用户感兴趣的新闻

3.4.3 输入项目

新闻推荐模块输入项如表 3.4 所示：

表 3.4 新闻推荐模块输入项目

名称	标识	数据类型	有效范围	输入方式
用户编号	Uid	Int	数字字符	系统自动输入
新闻推荐量	Rcmd_cnt	Int	数字字符	系统自动输入

3.4.4 输出项目

新闻推荐模块输入项为向用户推荐的新闻预览表。

3.4.5 算法

用户进入系统时，系统识别用户编号，根据给定的新闻推荐量，系统自动调用新闻推荐模块。对于新用户，系统采用冷启动的方式，根据当前所有新闻的热度和类别自动推荐。对于老用户，系统根据其浏览记录、收藏信息，寻找其他相似的用户，进行基于用户的协同过滤推荐，见图 3.4.1；系统根据其浏览、收藏的新闻的信息，寻找其他相似的新闻，进行基于新闻的协同过滤推荐，见图 3.4.2。结合上述两种推荐算法，系统采用一种面向视障用户的基于视障用户行为综合推荐算法，见图 3.4.3。

3.4.5.1 相关定义

定义用户 u 对有阅读记录的新闻 i 的评分为 R_{ui} ，阅读持续时间为 D_{ui} ，根据阅读持续时间得到的评分为 RD_{ui} ，阅读进度为 P_{ui} ，根据阅读进度得到的评分为 RP_{ui} ，是否收藏过该新闻为 C_{ui} ，根据收藏记录得到的评分为足 RC_{ui} 。用户 u 对新闻 i 的阅读记录且 $Record$ ，包括四个部分：新闻编号 $news_ld$ ，阅读持续时间

D_{ui} , 阅读进度 P_{ui} , 是否收藏 C_{ui} 。新闻信息数组 $newsArray$, $newsArray = \{n1, n2, n3 \dots ni\}$, 新闻 i 与新闻 j 的相似度为 W_{ui} , 书籍之间的相似度矩阵为 W , 阅读持续时间是用户最近一次进入新闻的阅读界面的时间与退出该阅读界面之间的时间差值。

3.4.5.2 基于用户的协同过滤

根据 Johns. Breese 在论文中提出的公式计算用户的兴趣相似度, 该公式如公式 1 所示:

$$W_{uv} = (\sum_{i \in N(u) \cap N(v)} \frac{1}{\log(1 + |N(i)|)}) / \sqrt{|N(u)| \cdot |N(v)|} \quad (\text{公式 1})$$

在得到新闻之间的相似度后, 通过公式 2 计算用户 u 对新闻 i 的感兴趣程度, 由于使用收藏与否作为判定, 所以 $r_{ui} = 1$ 对所有 r_{ui} 成立, 图 3.4.5.2 所示, 为基于物品的协同过滤的流程图。

$$P(u, i) = \sum_{v \in S(u, K) \cap N(i)} W_{uv} \cdot r_{vi} \quad (\text{公式 2})$$

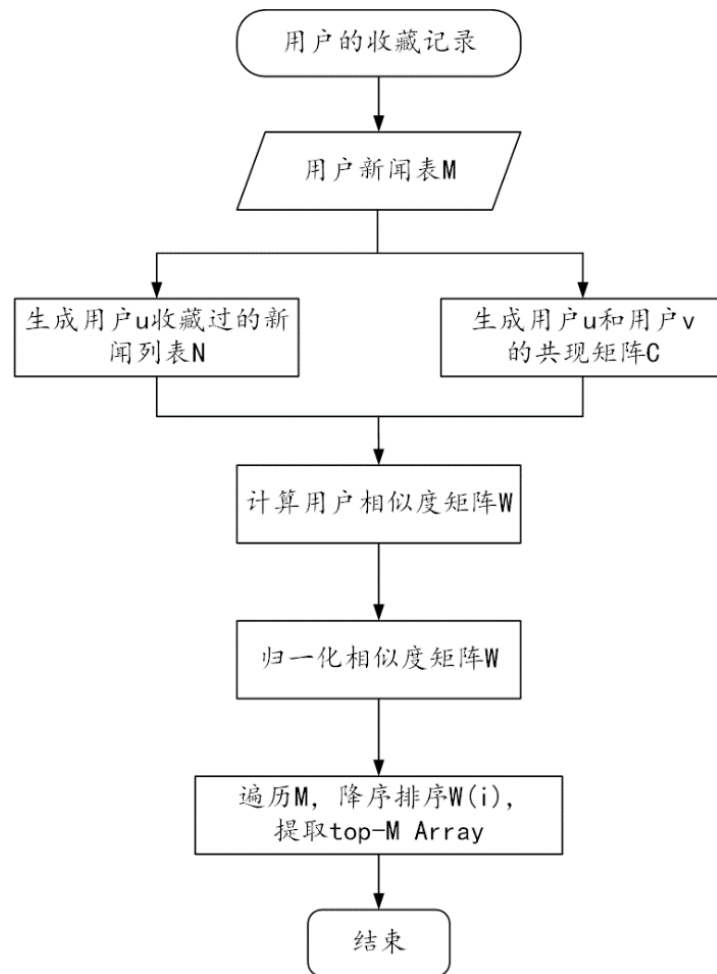


图 3.4.5.2 基于用户的协同过滤推荐算法流程图

3.4.5.3 基于新闻的协同过滤

根据 Johns. Breese 在论文中提出的公式计算物体的兴趣相似度 W_{ij} ，该公式如公式 3 所示：

$$W_{ij} = (\sum_{i \in N(u) \cap N(v)} \frac{1}{\log(1 + |N(i)|)}) / \sqrt{|N(u)| \cdot |N(v)|} \quad (\text{公式 3})$$

在得到新闻之间的相似度后，通过公式 4 计算用户 u 对新闻 j 的感兴趣程度，由于使用收藏与否作为判定，所以 $r_{ui} = 1$ 对所有 r_{ui} 成立，图 3.4.5.3 所示，为基于物品的协同过滤的流程图。

$$P(u, j) = \sum_{i \in S(j, K) \cap N(u)} W_{ij} \cdot r_{ui} \quad (\text{公式 4})$$

在引文《Evaluation of item—based top • n recommendation algorithms》中，可以得知提前将相似度矩阵按最大值进行归一化，可以提高推荐结果的准确率。可以用公式 5 得到归一化的相似度矩阵形，。

$$W'_{ij} = W_{ij} / \max_j W_{ij} \quad (\text{公式 5})$$

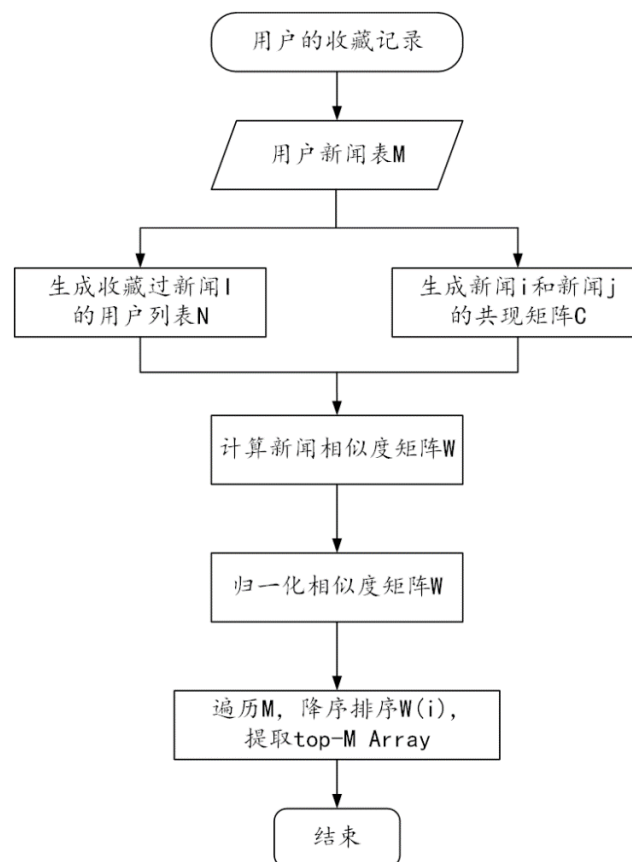


图 3.4.5.3 基于新闻的协同过滤推荐算法流程图

3.4.5.4 面向视障用户的基于行为的融合推荐算法

经过用户反馈发现，现有推荐算法的效果还远远差强人意。视障人士在使用电子书阅读 APP 获取资讯时，由于视力障碍带来的行动不便，很少会去主动关注别人，主动点击 APP 上看过的电子书的喜欢、收藏按钮，在中国盲文社提供的用户的阅读记录的数据中，一共有 223 个用户，但只有 18 个用户有收藏的记录，因此这将导致显式的用户反馈信息不足，如果直接运用一般的推荐算法给视障人群推荐，由于评分矩阵稀疏的问题，推荐效果并不理想。要想得到用户对电子书的有效评分，收集其隐式的反馈信息就显得尤为重要。而用户在 APP 上的操作行为又是很容易收集的，对用户某些操作行为的分析就能得到用户的兴趣取向。例如，对于视障人士而言，他们一般是通过听书的方式进行阅读，要想判断一本书是不是他们感兴趣的类型，他们耗费的时间远远多于普通人，经过了解，一般而言，他们至少得听 1 分钟以上才能进行判断该书是不是自己感兴趣的书籍，普通人可能只需要十几秒不到就能迅速做出判断。所以，我们可以这样假设，用户阅读这本书的时长小于或等于 1 分钟，代表他对这本书不感兴趣，否则，阅读时间越长，则代表他对这本书感兴趣的程度越深。而且，考虑到视障人士很容易产生偶然的点击记录，他们经常会不小心点击进入了某本电子书的播放界面，从而产生错误的阅读记录。所以，用户的阅读时长将是本算法的一个重要的考虑因素。另外，由于用户的兴趣并非总是固定不变的，例如，用户可能过去喜欢阅读保健养生类的小说，现在的兴趣又转移为古典文学类的小说。所以，我们可以这样假设，最新的阅读记录具有的反馈价值总是反映着用户的兴趣取向，所以我们只处理最新上传的阅读记录。用户看过某本电子书的阅读进度也是该用户对这本电子书的兴趣程度的重要指标。所以，电子书的阅读进度信息也是本算法的一个重要考虑因素。当然，上文中提到的收藏列表虽然信息量少，但是显式的反馈信息仍然不能忽视，所以也会作为本算法的一个考虑因素。

本算法根据 D 的取值确定 RD 的取值，根据 P 的取值确定 RP 的取值，根据 C 的取值来确定 RC 的取值， RD ， RP ， RC 的每项取值范围均是 0-1 分。其中 RD ， RP ， RC 的映射值采用引文的设定，见表 3.4.5.4.1、表 3.4.5.4.2、表 3.4.5.4.3。

表 3.4.5.4.1 D 与 RD 映射表

<i>D</i>	<i>RD</i>
小于或等于 1min	0.00
不大于 3min	0.17
不大于 6min	0.33
不大于 29min	0.50
不大于 104min	0.67
其他	1.00

表 3.4.5.4.2 P 与 RP 映射表

<i>P</i>	<i>RP</i>
0.0%	0.00
不大于 2.0%	0.14
不大于 8.0%	0.29
不大于 12.0%	0.43
不大于 20.0%	0.57
不大于 50.0%	0.86

表 3.4.5.4.3 C 与 RC 映射表

<i>C</i>	<i>RC</i>
0	0
1	1

用户 U 对书籍 i 的兴趣度评分且耐的计算方式如公式 6 所示：

$$R_{ui} = \alpha \cdot RD_{ui} + \beta \cdot RP_{ui} + \gamma \cdot RC_{ui} \quad (\text{公式 6})$$

其中， α ， β ， γ 是比重参数， $\alpha + \beta + \gamma = 1$ ，基于中国盲文社提供的用户的阅读记录数据集上，凸优化问题求解得到的系数分别为 $\alpha = 0.74427$ ， $\beta = 0.01015$ ， $\gamma = 0.24556$ 。

类似基于物品的协同过滤算法一样得到用户 u 对新闻 j 的评分 Q_{uj} ，计算方法如公式 7 所示，其中，新闻 $N(u)$ 表示用户 u 感兴趣的书记的集合， $S(j, K)$ 是与新闻 j 最相似的 K 条新闻的集合。

$$Q_{uj} = \sum_{i \in S(j, K) \cap N(u)} W_{ji} \cdot R_{vi} \quad (\text{公式 7})$$

考虑前面提出的面向 APP 的融合推荐算法进行修正，假设由面向 APP 的融合推荐算法得到的用户 u 对新闻 j 的评分 P_{ui} ，则用户 u 对新闻 j 的最终评分 R_{ui} 如公式 8 所示，经过实验发现 α 取值为 0.1 时结果最优。

$$R_{uj} = R_{uj} + \alpha Q_{uj}$$

(公式 8)

算法流程图如图 3.4.5.4 所示：

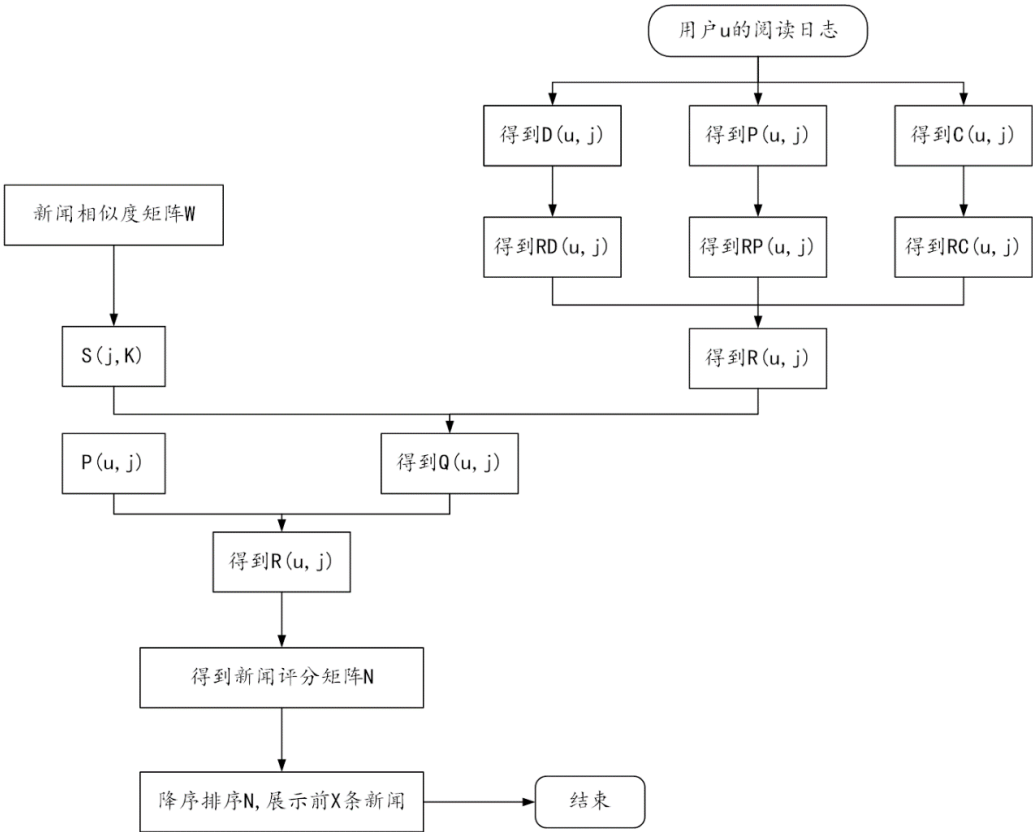


图 3.4.5.4 面向视障用户的基于行为的融合推荐算法流程图

3.4.6 测试要点

- (1) 测试为新用户推荐的新闻是否基于新闻热度
- (2) 测试为老用户推荐的新闻是否基于融合推荐算法及其准确性

3.5 新闻朗读模块

3.5.1 功能

用户选择想要收听的新闻，系统进入新闻展示详细信息并朗读新闻详细内容，更新浏览记录、用户喜好、用户收藏信息。

3.5.2 性能

便捷性：用户选择新闻将自动播放，不需要用户繁杂操作

高效性：用户选择新闻后，朗读新闻的响应时间要足够短

3.5.3 输入项目

新闻朗读模块输入项如表 3.5 所示：

表 3.5.3 新闻朗读模块输入项

名称	标识	数据类型	有效范围	输入方式
新闻编号	Nid	Int	数字字符	手动输入
用户编号	Uid	Int	数字字符	系统自动输入

3.5.4 输出项目

新闻朗读模块输出项如表 3.5.4 所示：

表 3.5.4 新闻朗读模块输出项

名称	数据类型
朗读语音	MP3 或 WAV

3.5.5 算法

用户选择新闻，新闻条目内容经过语音识别 API 转换为对应语音向用户播放，系统更新用户喜好文件和浏览记录文件，若用户选择收藏/取消收藏新闻，则新闻收藏记录更新；否则不更新新闻收藏记录文件。

新闻朗读模块的算法流程图如图 3.5 所示：

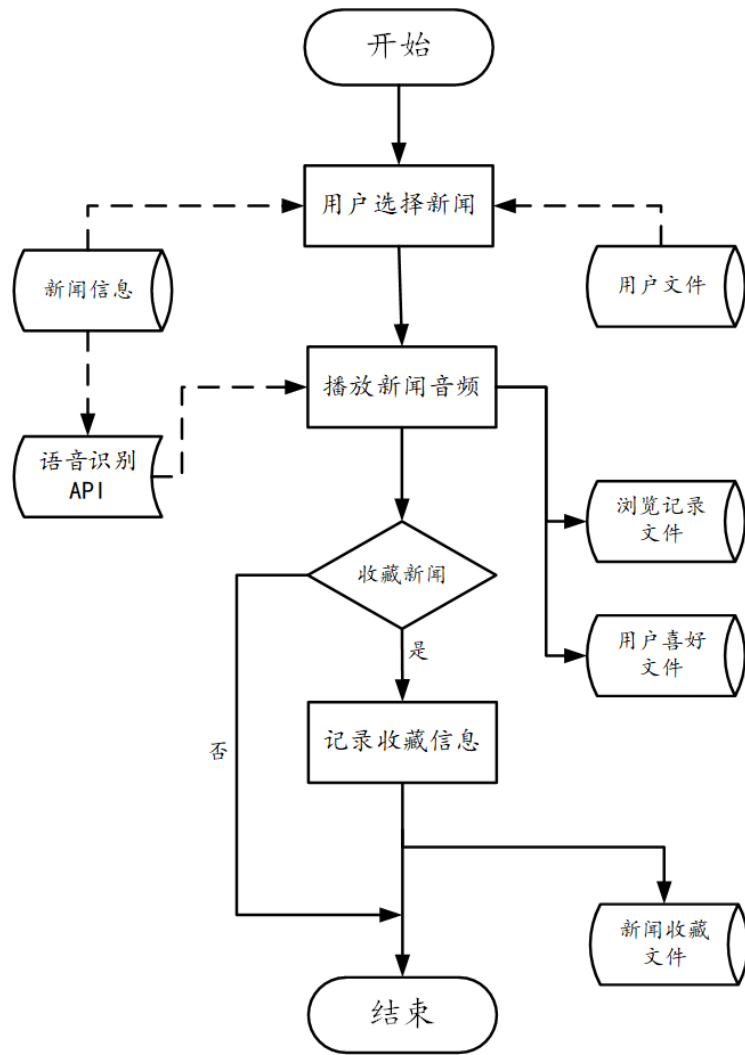


图 3.5 新闻朗读模块的算法流程图

3.5.6 测试要点

- (1) 测试新闻朗读的内容是否与文字内容一致
- (2) 测试浏览记录、新闻收藏、用户喜好文件是否相应更新

3.6 操作提示模块

3.6.1 功能

系统检测用户动作，发出相应的语音提示。

3.6.2 性能

便捷性：系统随时自动给出用户操作提示

3.6.3 输入项目

操作提示模块输入项如表 3.6.3 所示：

表 3.6.3 操作提示模块输入项

名称	标识	数据类型	有效范围	输入方式
用户行为	Tid	触摸事件	触摸开始、触摸移动、 触摸结束、触摸取消	手动输入

3.6.4 输出项目

操作提示模块输出项如表 3.6.4 所示：

表 3.6.4 操作提示模块输出项

名称	数据类型
朗读语音	MP3 或 WAV

3.6.5 算法

用户触摸屏幕激活系统行为检测，系统做出相应反应。当用户向右滑动时，系统判断是否存在上级界面，若存在上级界面，则将当前界面退回上级界面；当用户单击屏幕时，则播放当前预选中按钮作用的语音提示；当用户双击屏幕时，则确定选中当前预选中的按钮，激活当前按钮的功能函数；当用户长按屏幕时，系统判断当前界面是否处于新闻总览界面，若时，则启动语音搜索功能，启动语音搜索模块，否则刷新当前界面；当用户上下滑动屏幕时，系统改变当前预选中按钮位置，相应地进行上下/条目移动。

操作提示模块的算法流程图如图 3.6 所示：

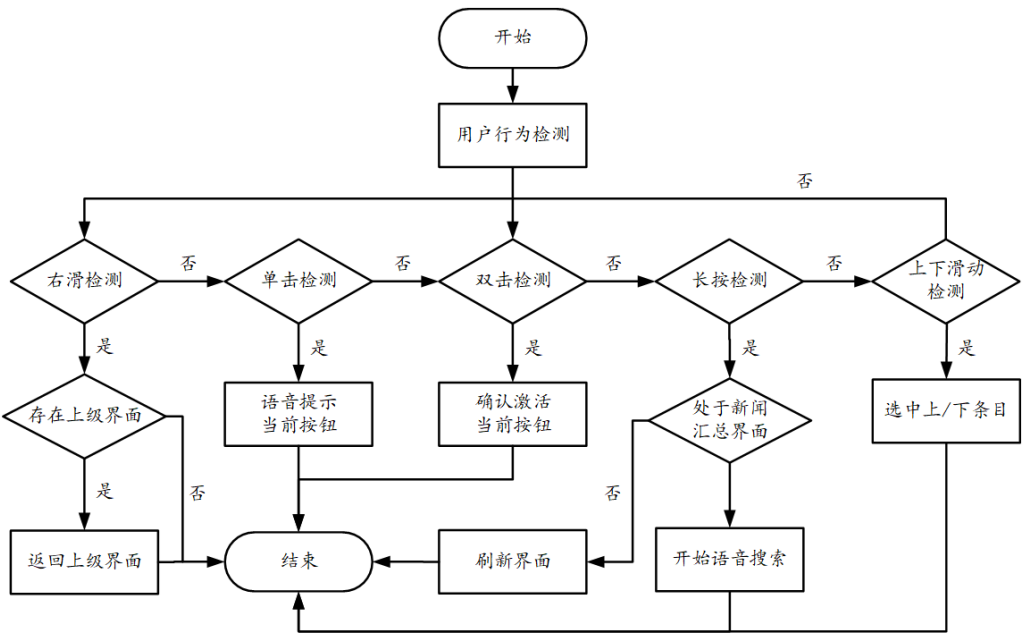


图 3.6 操作提示模块地算法流程图

3.6.6 测试要点

- (1) 测试系统是否精确识别用户行为
- (2) 测试是否根据用户行为做出相应反应

3.7 拍照识别模块

3.7.1 功能

用户在小程序端启动拍照功能，可以进行图片文字识别朗读或者图片内容物体识别。

3.7.2 性能

便捷性：用户可以在界面多次拍照识别

3.7.3 输入项目

拍照识别模块的输入项如表 3.7.3 所示：

表 3.7.3 拍照识别模块输入项

名称	标识	数据类型	有效范围	输入方式
待识别图片	Pic_bs64	Char	Jpg、png 格式 Base64	手动拍摄 自动转换

3.7.4 输出项目

拍照识别模块输出项如表 3.7.3 所示：

表 3.7.3 拍照识别模块输出项

名称	数据类型
朗读语音	MP3 或 WAV

3.7.5 算法

若用户选择拍照识别文字功能，系统接收用户拍摄照片，识别图片中印刷体文字并朗读；若用户选择拍照识物功能，系统接收用户拍摄照片，识别图片中的物体，并朗读其类别、状态、颜色、形状等信息。

拍照识别模块的算法流程图如图 3.7 所示：

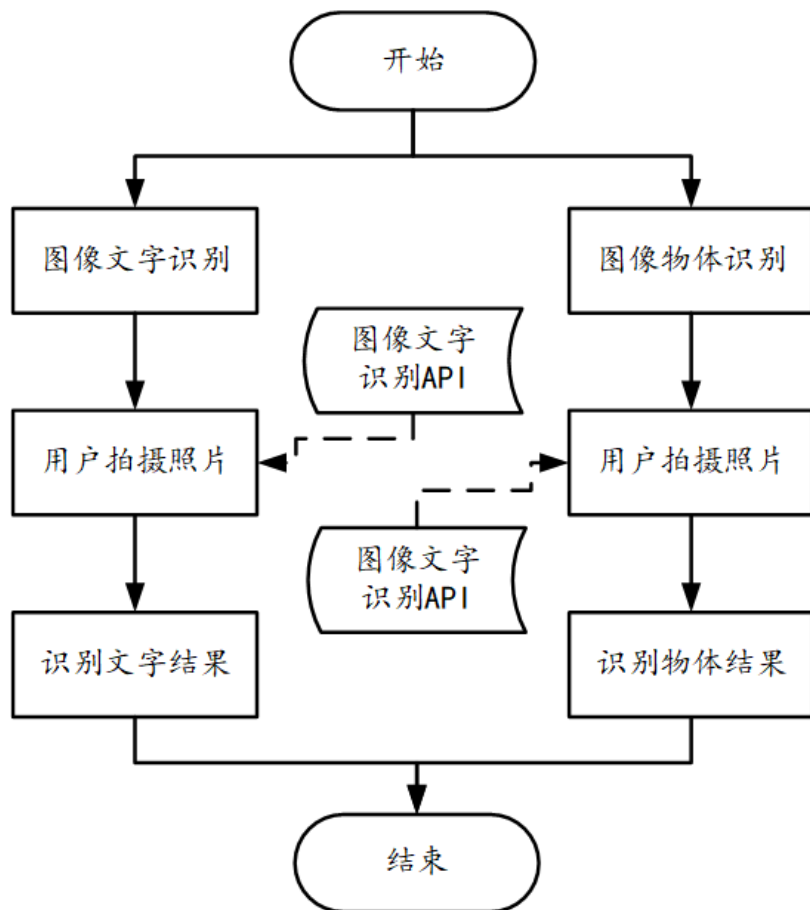


图 3.7 拍照识别模块的算法流程图

3.7.6 测试要点

- (1) 测试拍摄接口是否正常工作
- (2) 测试拍照识别是否正确调用文字识别和物体识别接口