

# 记录小助手

产品需求文档

# 前期讨论

- 刚开始的讨论方案为进行图像识别如图像清晰度增强等功能
- 形成原因：可能因为网络传播，手机拍摄的技术或者手机像素等问题造成图像模糊。
- 否决原因：后来考虑到目前技术的不断发展，手机的质量像素也越来越高，对于该方案的需求可能没有那么的大。受众的人群不广泛。
- 所以我们进行了修改我们的方案，增加了功能的同时，让受众的人群更广阔。

# 背景

- 现在许多的大学生会有这样的情况，在课堂上老师在讲授课程，听与写的能力较难同时进行，当能完全清楚地了解老师讲的内容，笔记的部分缺失；当做了笔记，但老师讲的内容可能没能完全吸收。同时大部分的学生都是拿出手机拍摄了老师的PPT的照片，但如果课后没能及时整理做笔记，导致期末准备考试时出现无从下手的情况。亦或者部分学生习惯了看电子版的内容，而不想用笔记录笔记。针对以上的一些问题，我们设计了这样一款APP帮用户解决这样子的痛点。

# 修改为以下方案

- APP名称：记录小助手
- 受众人群：学生，工作族，日常记录等
- 主要功能：可以通过语音识别进行记录某些短小的信息，可用于日常记账，记录灵感等等  
可以通过文字图像识别将图片上的文字识别出来，用于更好地整理笔记。  
附加功能：图像清晰度增强的功能，可能因为距离，拍摄角度等原因导致图像不清晰，这时可以通过图像清晰度增强使图像更清晰，降低错误率。

# 核心价值：

- 通过运用图像清晰度增强，提取文本以及手写文字识别等功能，为学生在课堂上做笔记，老师在审阅试卷时，以及在日常生活中的记录提供方便，节省用户的时间，提高学习效率。

# API价值宣言

- 微软Azure API中的提取文本API的价值 可以通过拍摄老师课上的PPT，识别PPT中的文字提取文本，节省学生在课上做笔记的时间，能更好地听老师讲课。
- 百度开放API中的手写文字识别的价值
- 可以将我们平时手写书摘、读书笔记、课堂笔记等内容的识别，实现对手写文字内容的扫描及线上存储。
- 老师对学生日常作业及考试试卷中的手写内容进行自动识别，对于一些学生比较潦草的字体能更加方便老师的阅读
- 百度开放API中图像清晰度增强的价值 拍摄时可能因为拍摄距离/手机像素等原因导致图像模糊不清导致文字识别出现错误的情况，这时图像清晰图增强可以使图像变清晰，减少识别错误的情况。
- 百度开放API中的语音识别成文字的价值 当用户在某一时刻想要记录自己的想法/与老师讨论时记录老师的观点时，可以通过语音识别成文字的功能，为用户提供方便。

# 对图像清晰度增强功能的补充

- 因为我们考虑到APP的日常运作的资金问题，所以我们将图像清晰度增强的功能设置为一种VIP的功能，前期进行推广宣传时可以通过用户做任务的方式（如分享给好友等）获得使用该功能的机会。后期用户增多的时候则进行金钱上的交易，如通过包月会员，或者使用一次需要付费多少的价格来进行收取费用。
- 因为这是推广部分的内容，具体还需后期再进行讨论决定。
- 其次，因为图像清晰度增强的功能是针对一些用户使用的图片模糊不清的情况，我们APP的图像文字识别已满足用户的基本需求，所以将其设置为会员制的功能。

# 用户痛点

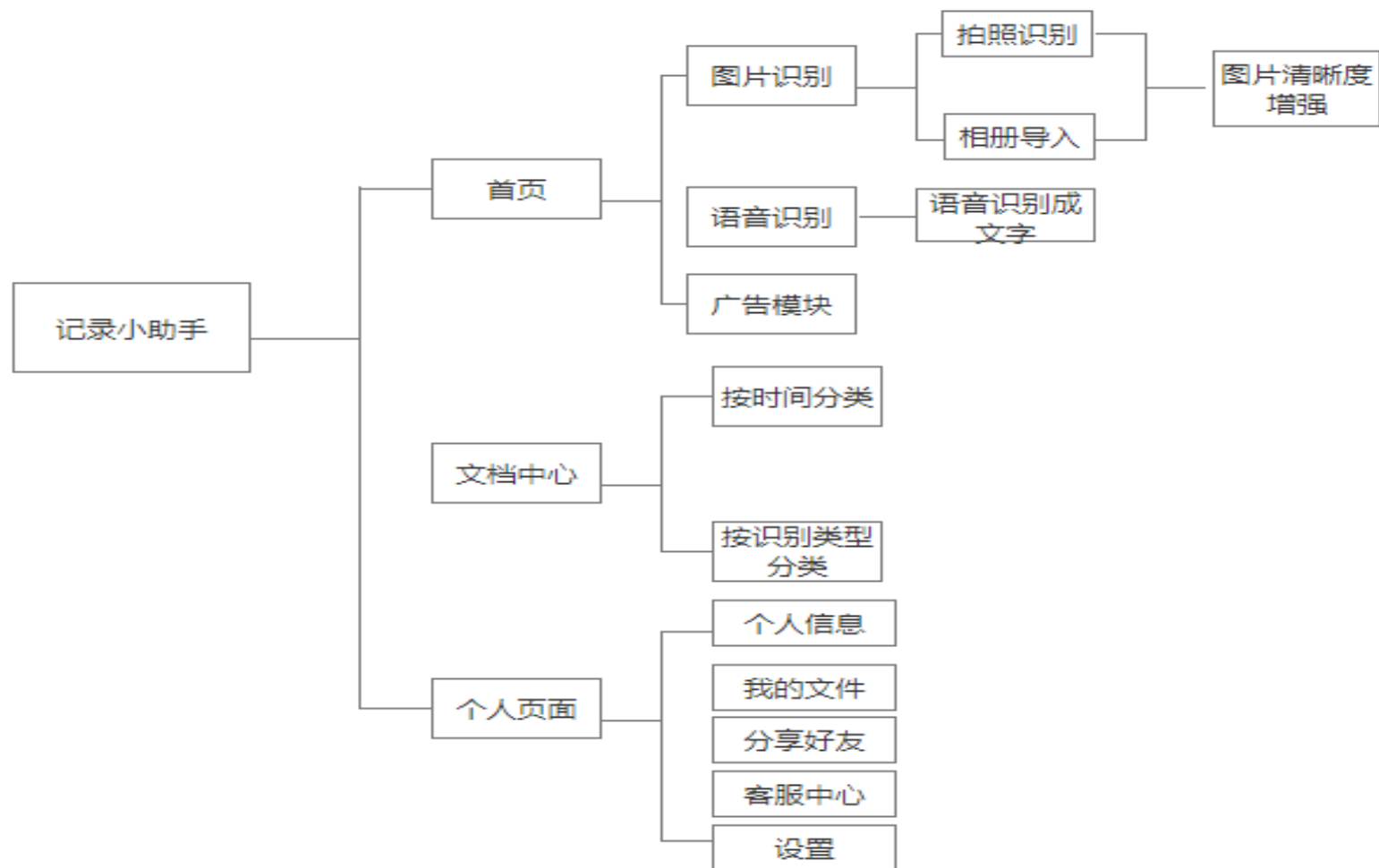
- 用户痛点
- **场景一：**
- 在课堂上听老师讲授知识以及抄写笔记是比较难同时进行的一件事，一心难以二用。
- **场景二：**
- 老师在批改作业时总会出现写字潦草的学生，这会增加其工作的时间与难度
- **场景三：**
- 对于一些会议手写记录需要整理成电子档，需要手动输入到电脑中
- **场景四：**
- 手写的笔记整理，后期查找相应内容较为麻烦
- **场景五**
- 当用户突然灵感一闪，想出了某个问题的解决办法，但身边没有纸笔。



# 需求列表

- 具体应用场景
- 在课堂上，小明正一边听着老师讲课一边写着PPT的笔记，但当记完一段笔记时，发现老师已经讲完那一章节的内容已经翻页了，正打算空出一部分页面来继续记下笔记，这时小明发现了一个新的APP，能够通过拍照识别文字，这时小明拿出手机，打开趣印APP，对着PPT拍摄照片，点击识别的按钮，PPT的文字便自动识别在手机屏幕上。
- 小花老师刚成为老师没多久，还不太适应学生的各种较为潦草的字体，当他在审阅学生的作文时，发现有什么内容并不正确，于是用手机拍照后，用趣印APP识别出来，他很快知道学生的作文内容并给分。同时，因为学生的字体潦草扣了分数。
- 小明因为太晚才到课室，发现靠前的座位已经被占满了，他只能坐在后排，当拿出手机准备拍照记录PPT笔记的时候发现因为距离的原因导致图片有点模糊，当他直扫时发现有部分文字缺失，然后他选择了图像清晰的按钮，使图片变得清晰，缺失的文字也出现了。
- 当小红在记录需要买什么东西时，你本想拿起纸笔记录，后来想起自己有记录小助手这个APP，他拿起手机把需要买的东西语音录入到手机上，等到到了商场，他便能较为方便地对照着手机的清单购物。

# 产品架构图



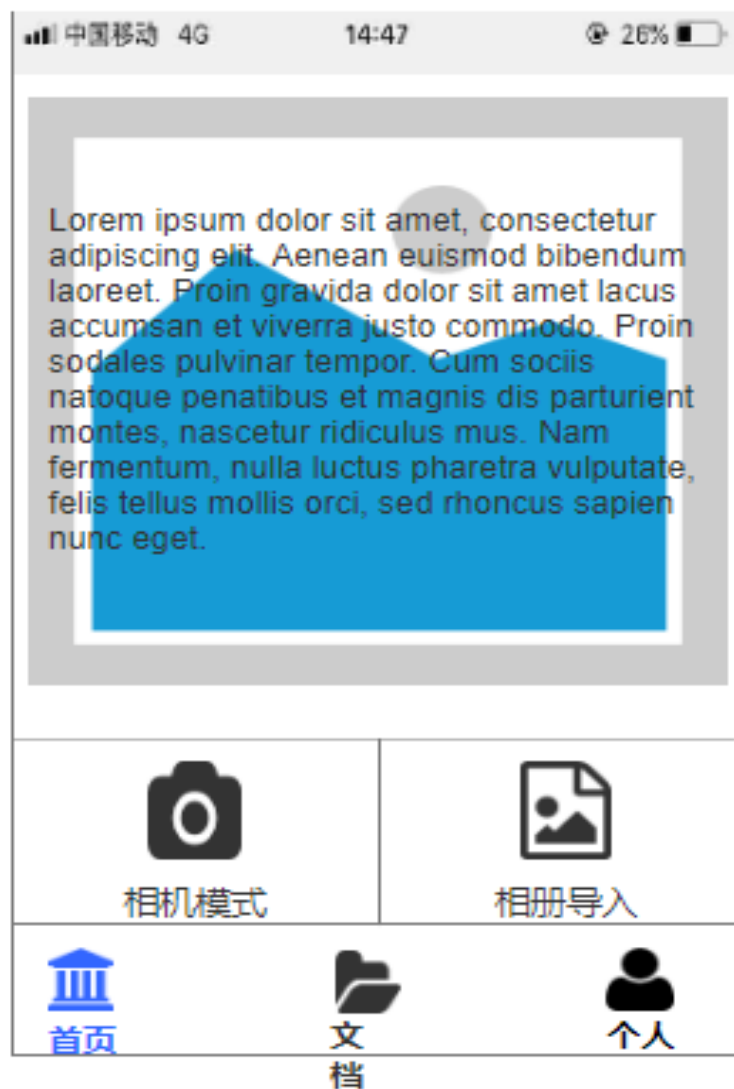
## 产品流程图



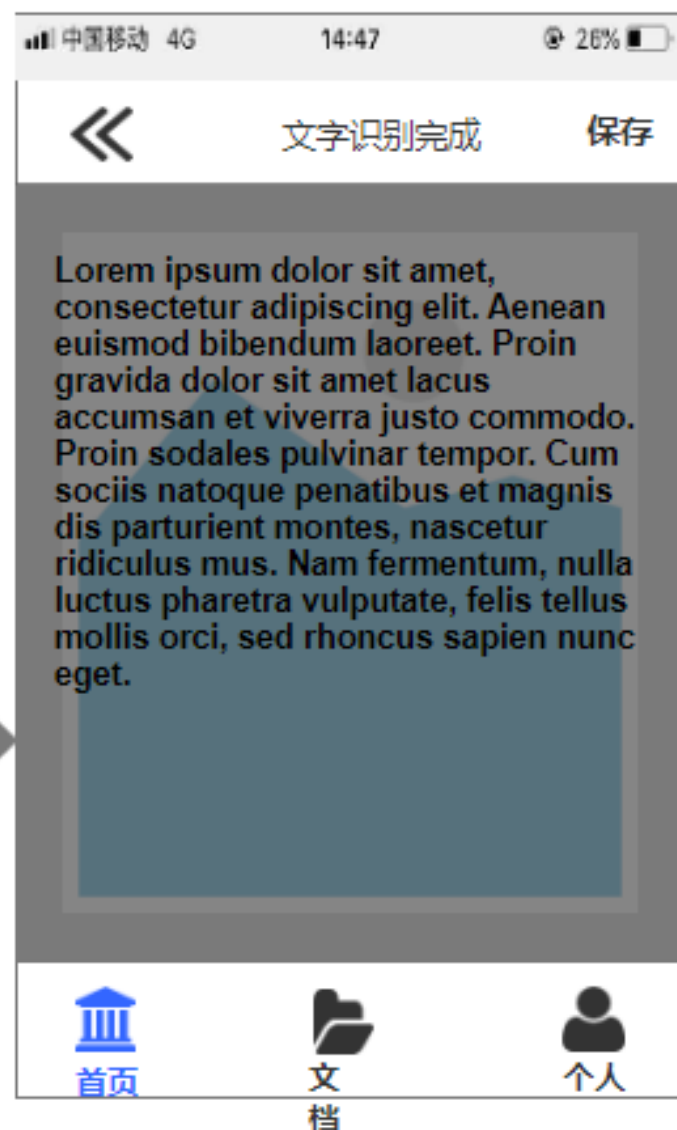
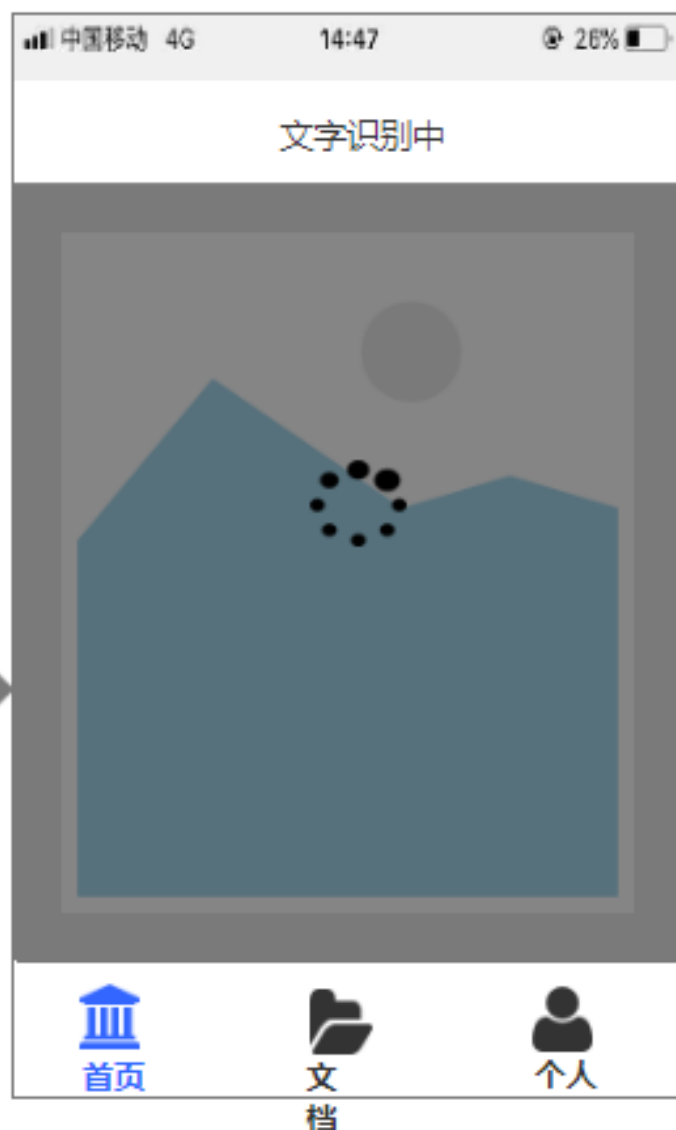
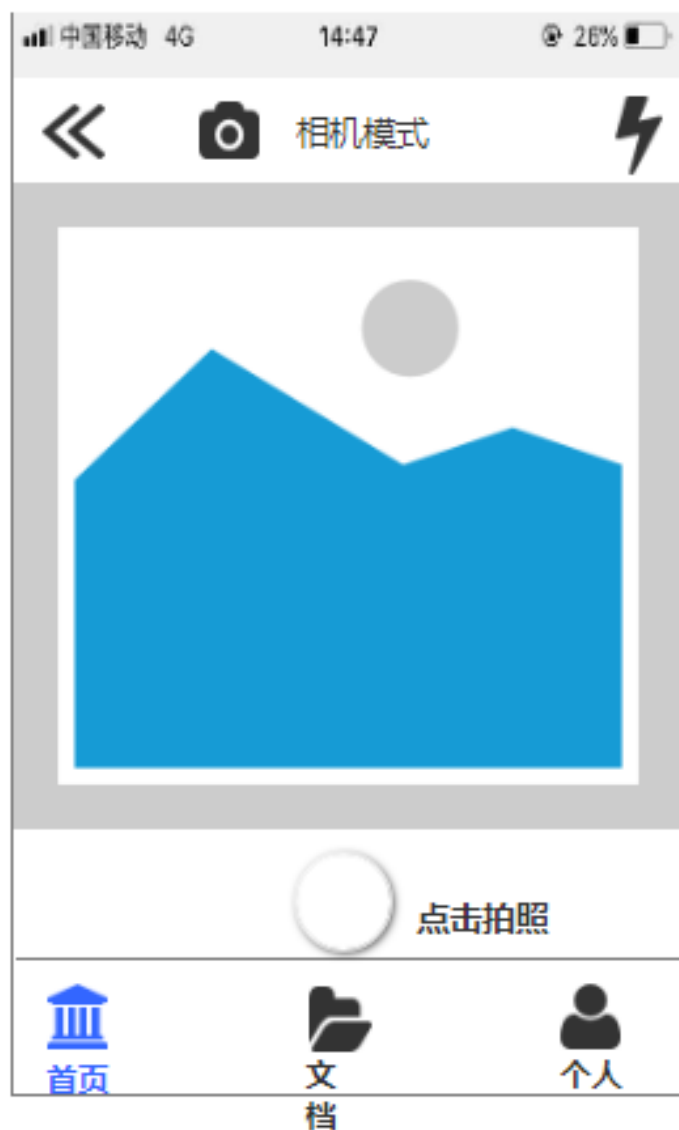
## 登录页面与首页



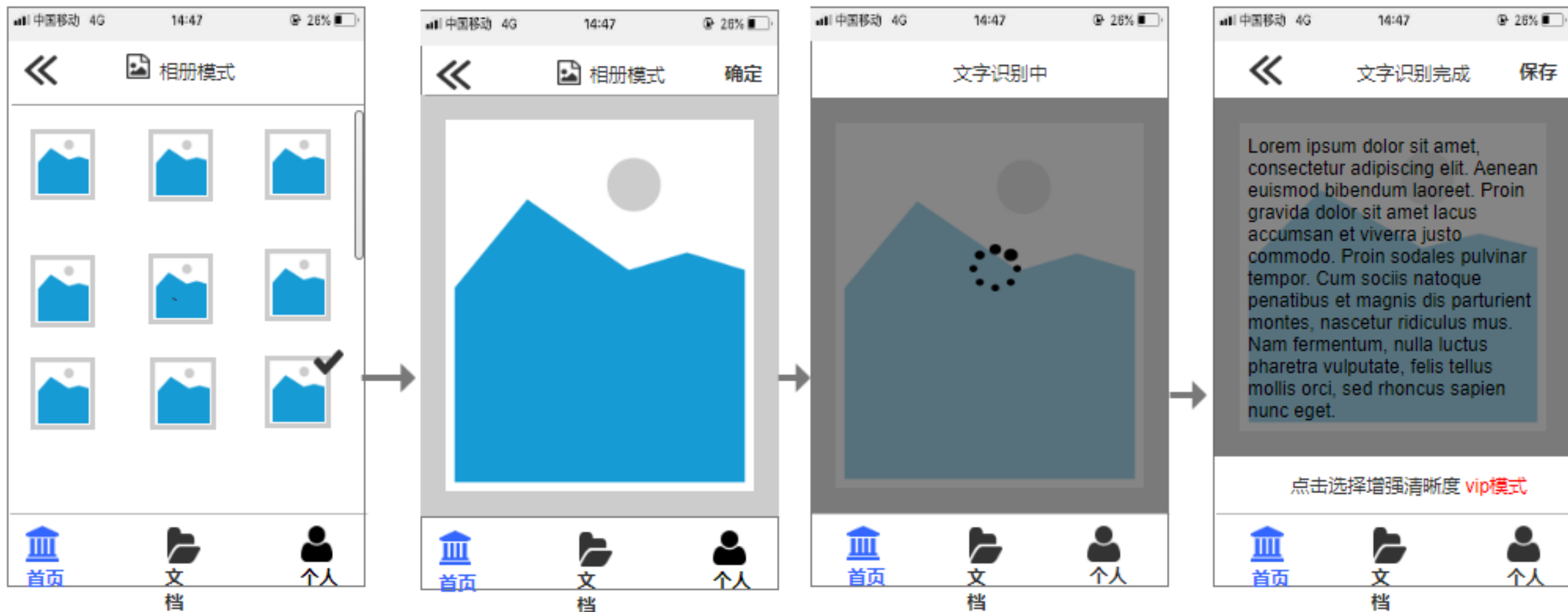
## 图片文字识别与语音识别内容



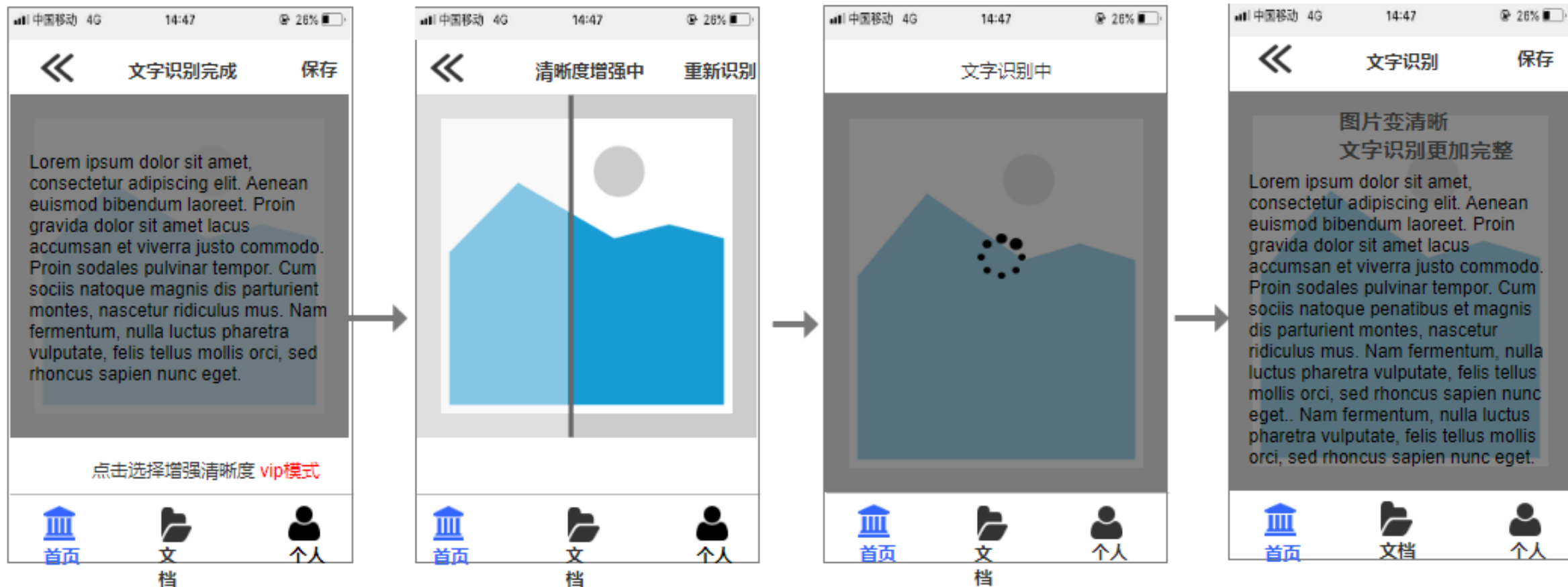
## 文字识别的相机模式



## 图片文字识别的相册模式



## 图像清晰度增强功能



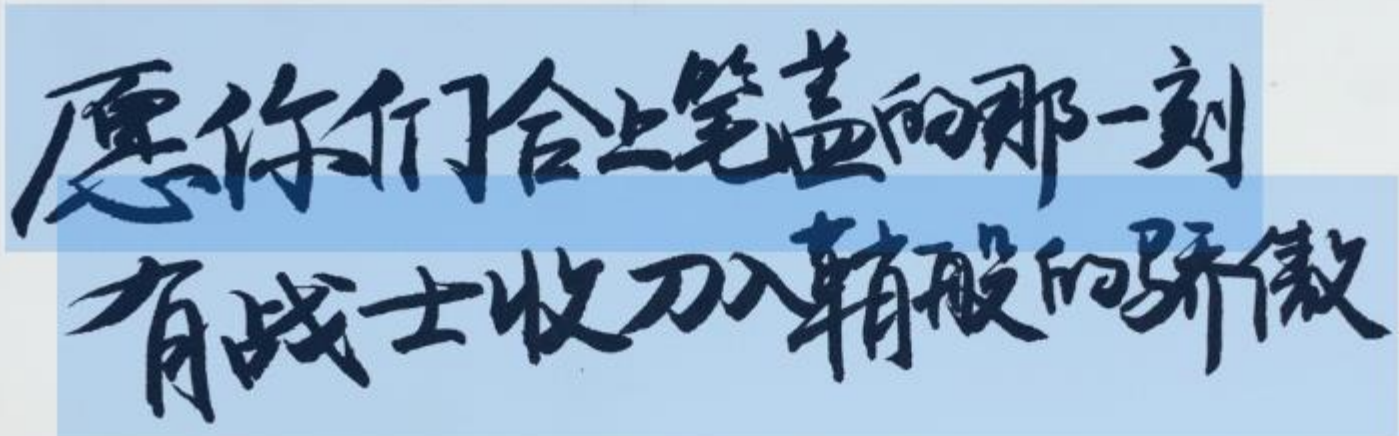


## 语音识别功能



## 各平台API对比分析

### 百度手写文字识别API



#### 识别结果

序号	内容	宽度	高度	左间距	上间距
1	愿你们合上笔盖的那一刻	898	177	140	183
2	有战士收刀入鞘般的骄傲	958	188	177	305

或



愿你们合上笔盖的那一刻  
有战士收刀入鞘般的骄傲

上传本地照片

或

使用网络图片，请粘贴

图片格式仅支持PNG、JPG、JPEG，图片大小不超过3MB

### 识别结果

- 1 愿你们合上笔盖的那一刻
- 2 有战士收刀入鞘般的骄傲

## •百度开放平台的使用后风险报告

### •API市场竞争程度

•在百度搜索引擎中搜索“语音识别”时，第一个搜索出来的是“百度语音识别成文字”的广告，其次是讯飞开放平台的广告。

•在bing搜索引擎中搜索“语音识别”时，第一个是一些较为不出名的开放平台的广告，第一页中有腾讯AI开放平台，而百度AI开放平台的搜索结果并不在第一页。

总结：在百度中，百度AI开放平台的在百度搜索引擎中的比重及广告宣传力度还是较大的，但在bing搜索引擎中就比较难找到。

### •输入输出限制

#### i. 文字识别

百度AI开放平台的文字识别主要是识别中文的电子版与手写板，对于英文的识别仍然是一个缺陷。所以我们主要是使用微软中的的文本识别API。

#### ii. 语音识别

百度AI开放平台的语音开放平台中主要是包含输入法、搜索、粤语、英语、四川话、远场模型，部分的其他方言无法识别出来

#### iii. 图像清晰度增强

这个功能的API目前还处于邀测的状态，具体的输出限制内容还没有发现，但他要求需要处理的图片base64编码后大小不超过4M，最短边至少64px，最长边最大4096px，长宽比3：1以内，像素乘积不超过 1280\*720。

•定价

i. 文字识别

每日 500 次免费调用量，免费额度用尽后按照使用次数收费，月调用量0-5万次数时，为0.005元一次。对比了腾讯云，阿里云等等的开放平台，百度的定价是较为合理的。具体如下图所示。

通用文字识别

每日 50000 次免费调用量，免费额度用尽后按照如下价格进行计费。如需付费使用，可点击 [开通按量后付费](#)。

月调用量（万次）	通用文字识别（元/次）
0<调用次数<=5	0.0050
5<调用次数<=10	0.0045
10<调用次数<=20	0.0040
20<调用次数<=50	0.0035
50<调用次数<=100	0.0030
100<调用次数	0.0025

说明：“调用次数”只包括成功调用，调用失败不计费

ii. 语音识别

短语音识别包含输入法、搜索、粤语、英语、四川话、远场模型。定价为每账号前50000次调用免费，免费额度用尽后开始计费，按阶梯后付价格。具体价格如下图：

短语音识别标准版价格

短语音识别包含输入法、搜索、粤语、英语、四川话、远场模型。可按天/月/年购买QPS。价格如下：

QPS购买方式	单价（元）
1QPS/天	80
1QPS/月	1400
1QPS/年	12000

说明：若原有QPS默认配额为2，则购买10QPS/月后，该月QPS提升至12。不限调用量。

语短音识别极速版价格

短语音识别极速版采用按调用量阶梯后付及次数包两种方式：

分段阶梯后付方式，调用单价按照自然月累积调用量所落阶梯区间而变化。月初，上月累积的调用量清零，重新开始累积本月调用量。用多少付多少。每账号前50000次调用免费，免费额度用尽后开始计费，按阶梯后付价格如下：

月调用量（万次）	语音识别极速版（元/次）
0<调用次数<=600	0.0042
600<调用次数<=3000	0.0036
3000<调用次数<=6000	0.0029
6000<调用次数<=15000	0.0019
15000<调用次数	0.0014

iii. 图像清晰度增强

目前这个功能的API目前还处于邀测的状态，所以没有定价的内容，但我们可以向百度AI平台申请试用以及合作的咨询。