**手机验证及OCR识别自动化需求说明**

1. **现状说明**

目前的使用方式及流程：

1. 数据导入：人工将号码数据整理好，分别导入每台手机，并在手机设定好参数，执行手机验证程序开始进行验证工作。
2. 图片导出：在手机任务执行完毕后，将手机中的图片目录进行压缩打包（如果不压缩打包，电脑连接手机无法显示全部图片），并将打包好的图片导入到电脑中。
3. 图片导入服务器：根据服务器的OCR程序运行情况（人工判别哪台服务器有空闲），将电脑中的图片导入到可执行OCR程序的服务器中，并在服务器上进行解压缩操作。
4. 运行OCR程序：配置好OCR程序的运行脚本后，即可开始执行OCR程序，持续监控程序的运行状态，能看到程序是否执行完毕。
5. 整理输出OCR验证结果：将OCR程序的验证结果进行预处理格式化，并导出成目标数据文件，后续可以将该文件导入现场的号码标识库。
6. **总体需求说明**

目前的操作基本上是纯人工操作，后续根据领导的指示，需要减少人工工作量，将手机验证及OCR识别做成自动化流程，避免人工参与。



公共空间的目录结构：

PUB\_SPACE（公共空间）

|——NUM\_DATA（号码数据）

|——WAIT\_DATA（待处理）

|——RUN\_DATA（运行中）

|——END\_DATA（处理完毕）

|——PIC\_DATA（图片数据）

|——WAIT\_DATA（待处理）

|——RUN\_DATA（运行中）

|——END\_DATA（处理完毕）

|——RES\_DATA（结果数据）

|——RES\_OUT（结果输出）

需求流程说明：

1. 公共空间设定：在公司的服务器上设置一个公共空间，用来存放“号码数据”和“图片数据”，操作者在公共空间上放置号码数据后，程序就可以自动获取该数据，在手机验证完后，手机端自动将图片数据上传至公共空间中存储。
2. 自动获取号码数据：手机验证程序能自动去公共空间的“号码数据-待处理”目录下搜索未执行的号码数据文件（根据文件名升序排列获取第一个文件，每个文件等于一个任务），如果找到该文件就将文件放置在空闲的手机上进行任务执行，并将数据文件移至“号码数据-运行中”目录。手机验证程序可以管理多台手机，并判断手机任务空闲的程度，将合理的号码数据任务分配给空闲的手机进行运行。
3. 自动上传图片数据：手机验证程序在完成验证任务后，将数据文件从“号码数据-运行中”目录移至“号码数据-处理完毕”目录，并将截图存储目录压缩打包，上传至“图片数据-待处理”目录中。
4. 自动处理图片文件：OCR程序实时监控“图片数据-待处理”目录，判断相关文件是否已上传完毕（根据文件大小、时间等），在文件上传完毕后，OCR程序将图片数据文件移至“图片数据-运行中”目录，并将图片数据文件传给空闲的OCR节点（自动判断OCR节点的负载均衡），进行图片文件处理解压缩，并执行节点上的OCR程序对图片进行识别。
5. 自动上传结果文件：在OCR节点完成图片验证任务后，将图片数据文件从“图片数据-运行中”目录移至“图片数据-处理完毕”目录，并将OCR节点中生成的验证结果文件上传至“结果数据-结果输出”目录下。
6. **手机验证接口及流程说明**

**1、目前界面参数优化**



**2、接口参数说明（所有参数均可在手机端进行配置修改）**

设备编号：MXX（XX=01~99，例如：M01）

接口类型：ftp/sftp（默认为sftp）

服务器IP：192.168.10.191

用户名：shandong1

密码：shandong1

检索间隔时间：默认30秒

拨号间隔时间：默认3000ms

**3、目录说明：**

1. 获取数据文件目录：/home/shandong1/PUB\_SPACE/NUM\_DATA/WAIT\_DATA
2. 数据文件执行中目录：/home/shandong1/PUB\_SPACE/NUM\_DATA/RUN\_DATA
3. 数据文件执行完毕目录：/home/shandong1/PUB\_SPACE/NUM\_DATA/END\_DATA
4. 图片打包上传目录：/home/shandong1/PUB\_SPACE/PIC\_DATA/WAIT\_DATA

**4、数据文件格式：**

文本格式：一行一个号码

**5、数据文件命名规范：**

文件名格式：设备编号\_序号（6位数字）.txt

样例：M01\_000001.txt

**6、图片目录及文件名命名规范：**

图片目录命名：设备编号\_序号（6位数字）\_开始日期（YYMMDD）\_开始时间（hhmiss）

图片目录样例：M01\_000001\_191213\_093040

打包类型：整个目录ZIP压缩

压缩包命名：

1. 上传中：M01\_000001\_191213\_093040.zip
2. 上传完毕：M01\_000001\_191213\_093040**\_done**.zip （上传完毕后，将文件名加上\_done标识）

**7、基本流程步骤：**

1. 检索数据文件：在手机端点击“启动”按钮后，手机端自动去接口服务器上检索对应设备编号的数据文件（例如：M01\_000001.txt），如果找到对应的数据文件（如果找到多个文件，默认按序号升序排列，获取第一个文件），就下载到手机端，开始执行任务。如果没有找到数据文件，手机端每隔30秒（可配置）去检索一次，直到找到数据文件或人工退出程序为止。
2. 手机端运行任务：手机端获取数据文件后，将所获取的数据文件从“NUM\_DATA/WAIT\_DATA”目录移至“NUM\_DATA/RUN\_DATA”目录下。并根据手机端的预设参数，自动开始执行任务。
3. 图片打包上传：任务全部执行完毕后，将数据文件从“PUB\_SPACE/NUM\_DATA/RUN\_DATA”目录移至“PUB\_SPACE/NUM\_DATA/END\_DATA”目录下。并将图片目录打包为zip文件，自动上传至“PUB\_SPACE/PIC\_DATA/WAIT\_DATA”目录下，上传完毕后，在原zip文件名后面加上“\_done”的完成标识（例如：M01\_000001\_191213\_093040**\_done**.zip）。
4. 继续检索数据文件：完成以上1-3步骤后，手机端继续开启检索程序，获取或等待新的数据文件，并继续执行以上步骤。
5. **OCR识别程序流程说明**

**1、主要目录说明：**

1. 图片文件获取目录：/home/shandong1/PUB\_SPACE/PIC\_DATA/WAIT\_DATA
2. OCR图片识别中目录：/home/shandong1/PUB\_SPACE/PIC\_DATA/RUN\_DATA
3. OCR图片识别完毕目录：/home/shandong1/PUB\_SPACE/PIC\_DATA/END\_DATA
4. 结果输出数据上传目录：/home/shandong1/PUB\_SPACE/RES\_DATA/RES\_OUT

**2、结果文件命名规范**

格式：图片目录命名\_res.txt

样例：M01\_000001\_191213\_093040\_res.txt

**3、基本流程步骤：**

1. 目录检索：自动检索（每隔30秒）图片上传目录“PUB\_SPACE/PIC\_DATA/WAIT\_DATA”，找到文件名后面带有“\_done”标识的zip文件。
2. 自动处理图片：将上传完毕的zip文件移至“PUB\_SPACE/PIC\_DATA/RUN\_DATA”目录下，开始解压缩zip包，完成解压缩后开始调用OCR程序找合适的OCR服务器进行图片识别处理。图片处理完毕后，将解压缩后的图片目录删除，并将原先的zip文件移至“PUB\_SPACE/PIC\_DATA/END\_DATA”目录下。
3. 结果文件上传：将生成的结果文件上传至“PUB\_SPACE/RES\_DATA/RES\_OUT”目录下。
4. 继续检索图片文件：检索图片文件目录，找到上传完毕的图片文件后，继续按照以上步骤对图片数据文件进行处理。