TechOcr程序进展情况汇报

按照原先预计的转正、预识别、识别三个步骤，TechOcr程序目前已经完成了核心程序的开发，目前尚余对外部输出、JNI封装等部分的开发尚未完成，另需要些时间检测内存泄漏。预计本周完成V0.2版本的各项工作。

一．V0.2版本已经实现的功能

1．对待识别物件的假设

待识别物件的原件，在最外边，必须有两两平行，且相互垂直的四条边。例如矩形（打印纸张）、含圆角的矩形（身份证）。

如果原件本身不符合上述条件，可以将原件放置在一矩形背景内，且文字与矩形边平行。并在拍摄图像中包含整个矩形背景。

物件的内部可以有其他直线（如各种表格），但这些直线必须与外部的矩形边平行或垂直。

2．对待识别图像的假设

上述两种情形下的矩形边，必须全部包含在图像中。

图像可以任意倾斜、旋转，甚至横置、倒置。物件中平行、垂直的矩形边，在图像中不一定平行或垂直。但上述物件矩形所占面积，必须达到图像总面积的25%。

此条件的另一个隐含限制是，图像中只能有一个待识别物件。

如果以图像的水平、垂直中心线为坐标轴。矩形边两两相交的四个交点，必须分别处于四个象限。

每一个待识别的汉字，在图像中的大小，应不小于50×50点阵。

对于满足上述条件的图像，可以做到精确定位。

3．程序可根据图像，通过特征字位置的判定，自动识别版式；也可根据指定版式进行识别。

4．除识别接口外，程序还提供接口，以定义版式。分为定义版式名称、定义特征字、定义待识别区三个API。可供其他各类程序调用。

5．程序在识别接口中返回以下内容：矩形边的交点、最佳匹配版式名称以及待识别图像与此版式的匹配度、各个待识别区域名称以及识别内容。

二．尚需改进的地方

1．图像预处理部分

在测试中，程序对信用卡的识别不佳（指使用凹凸数字表示卡号的信用卡），程序无法分辨卡面图案与凹凸数字，需对此类图像进行特别的预处理，以消除卡面图像的干扰。

印章对图像的干扰，目前尚未消除。预计可以通过颜色的不同来进行区分消除。

2．矩形识别部分

目前版本中，如果识别出的边线超过4条，在如何选取上，还有很多工作要做。预计使用直线两侧直方图对比的方法完成。

3．预识别部分

目前可以识别横置、倒置的图像，但做了很多重复计算，造成效率的降低。在极端情况下，会执行四次预处理。通过优化可以合并为一次预处理，以提高效率。

4．实际识别部分

目前没有实现对“串行”（指物件印刷部分，与打印部分错位）的分析，尚未想到合适的算法来纠正。

5．其他外围辅助工具

包括定义版式的辅助工具；对识别物件内容的记录、分析、再学习；自学习版式等。

三．项目完成过程中的一些心得

整个过程，与其说是编程序，还不如说是对解析几何、线性代数的再学习。有些算法找来源程序后，看不懂其中的数学含义，只是通过测试验证后就拿来使用，暴露出数学功底的薄弱。

矩阵相乘是不满足交换律的，就是因为两个矩阵相乘的次序，就花了很多时间来调试。

四．此项目的进一步发展方向

1．实现对版式的自学习，实现对名片、铭牌的自适应识别。

2．整个识别项目的移动化，在IOS、Android系统上实现整个项目。