TechOcr程序进展情况汇报

按照原先计划，TechOcr程序V0.2版本已经完成对“营业执照”的识别；提交Java开发，并完成联调。

一．上周（6/13-6/16）在V0.2版本上进行的工作

完成预计的外部输出、JNI封装工作，并花了两天时间检测内存泄漏。从原来的每次运行（检测一张图片，并输出结果）增加1M+内存，改善到每次约增加20K，达到可容忍范围。

二．上周（6/17-6/19）开始的V0.3版本

在前两个版本识别“营业执照”的开发过程中，遇到的一个问题就是：样本数量太少。V0.3版本的主要工作，是识别样本数量众多的一种版式——身份证。

在网上随机下载了20张身份证图片（分辨率大于1024×768），通过V0.2版本的规则定义了版式，其中有三张可以识别。

三．身份证识别的难点

1．身份证图片有底纹，需去除底纹干扰。而在判定底纹时，还需要先去除光线的影响，此部分属于图像预处理阶段的工作。

此部分工作需要图像学基础知识，特别是色彩空间的知识。上周末对RGB、YUV、HSL等知识进行了学习，并完成了色彩空间转换，像素通道拆分、合并等功能的程序。

2．通过网络图片才知道，还有少数民族格式的身份证；与一般的身份证相比，多了少数民族文字。

转换为版式定义需求，就是：版式特征字的位置，不能靠绝对位置定位，只能根据相对位置来定位。

3．很多身份证图像，缺失部分边缘。也就是没有完整的轮廓矩形，不符合V0.2程序的识别要求。预计依靠文字的相对位置，来检测倾斜、旋转。

四．V0.3版本的主要功能

不依赖OCR程序，自行完成对汉字的定位；将Tesseract OCR程序，弱化为对单个汉字的识别。并在定位过程中，实现对版式的模糊识别。