

计算机视觉-Final Project2

任务目标

给定若干张名片图，实现名片的图像校正和名片信息的抽取。

任务可以做如下分解：

1. **名片校正**：主要是把名片从背景图像中切割出来，只保留名片本身的完整图像信息(不含背景图等)。
2. **名片主要部分切割**：对于名片图像，可以分为如下几个主要的部分：姓名，Title，单位信息，电话号码等。
3. **字符切割与识别**：完成单字符的切割，并识别，最后输出(姓名，Title，单位信息，地址，手机，座机，传真等号码)等重要信息。

任务二：字符切割和数字识别

任务要求

给定 100 张名片图，附件已经给出 20 张(其他需要大家收集)。完成下面两个功能。

1. 字符切割：完整切割出单个汉字和数字字符，思路可以使用作业 5 和作业 6 的思路，但是需要保证汉字字符的完整性。
2. 字符识别：只识别名片中的手机号码；即只需要训练数字的分类器，并完成识别名片中的手机号码识别。
3. 复杂任务：识别名片中所有信息(包括数字和汉字)；即只需要训练数字的分类器，并完成识别名片中的手机号码识别。(可用 Tesseract 库)

字符切割

首先在以前的代码上做了一点改进，将连通块用一个结构体存放，结构体中包含四个int表示xmin，xmax，ymin，ymax，就不用每次遍历来找顶点了。

```
struct Conn {  
    vector<Point> points;  
    vector<int> xy;  
};
```

然后处理连通块，太小的不要，太大的不要，离的近的合在一起。处理之后放在一个新的Vector connect2里。

具体过程是先对所有连通块按ymin（ymax也可以）排序，再按xmin排序，然后就可以按顺序遍历，合并重叠的和靠近的连通块。

```
void Segmentation::process_connection() {  
    sort(connect.begin(), connect.end(), mycmpy);  
    int tmpy = 136;
```

```

        vector<Conn> tmp_con_y;
        for (int i = 0; i < connect.size(); i++) {
            if (connect[i].points.size() < 200 || connect[i].points.size() >
10000) continue;
            if (abs(connect[i].xy[2] - tmpy) > 55) {
                conn_sort.push_back(tmp_con_y);
                tmpy = connect[i].xy[2];
                tmp_con_y.clear();
            }
            tmp_con_y.push_back(connect[i]);
        }
        conn_sort.push_back(tmp_con_y);

        for (int i = 0; i < conn_sort.size(); i++) {
            sort(conn_sort[i].begin(), conn_sort[i].end(), mycmpx);
            for (int j = 0; j < conn_sort[i].size(); ++j) {
                Conn tmp = conn_sort[i][j];
                if (j == conn_sort[i].size() - 1) {
                    connect2.push_back(tmp);
                    break;
                }
                for (int k = j + 1; k < conn_sort[i].size(); ++k) {
                    if (conn_sort[i][k].xy[0] - tmp.xy[1] < 4) {
                        if (conn_sort[i][k].xy[0] < tmp.xy[0])
tmp.xy[0] = conn_sort[i][k].xy[0];
                        if (conn_sort[i][k].xy[1] > tmp.xy[1])
tmp.xy[1] = conn_sort[i][k].xy[1];
                        if (conn_sort[i][k].xy[2] < tmp.xy[2])
tmp.xy[2] = conn_sort[i][k].xy[2];
                        if (conn_sort[i][k].xy[3] > tmp.xy[3])
tmp.xy[3] = conn_sort[i][k].xy[3];
                        if (k == conn_sort[i].size() - 1) {
                            connect2.push_back(tmp);
                        }
                    }
                    else {
                        connect2.push_back(tmp);
                        j = k - 1;
                        break;
                    }
                }
            }
        }
    }
}

```

最后直接把connect2全部画在图上即可。

```

CImg<unsigned char> Segmentation::draw_single(CImg<unsigned char> input) {
    CImg<unsigned char> result = input;
    const unsigned char red[] = { 255,0,0 };
    for (int i = 0; i < connect2.size(); i++) {
        int x0 = connect2[i].xy[0];

```

```

        int x1 = connect2[i].xy[1];
        int y0 = connect2[i].xy[2];
        int y1 = connect2[i].xy[3];
        result.draw_line(x0, y0, x1, y0, red);
        result.draw_line(x0, y1, x1, y1, red);
        result.draw_line(x0, y0, x0, y1, red);
        result.draw_line(x1, y0, x1, y1, red);
    }
    return result;
}

```

效果如下：



字符识别

使用Tesseract-OCR识别名片“郭飞”结果如下：

```

$70. BB "g-fifé'b 'kfifiFR' o ' 1%]
jtfi'ig-fiffimfif 'fik BB 99 '

```

```

"5' 13 1

```

```

$111: 13898145123

```

```

Milt: %BB$H$EE$E$77%:&101%

```

```

$13; 024—23524589 fig: 024—23524589

```

```

$101: 13898145123

```

只有数字是正确的。

使用百度AIP识别，效果如下：

```
D:\TH\CV\Dataset\IMG_20191204_001325.jpg
{'log_id': 3888580043076795285, 'direction': 0, 'words_result_num': 8, 'words_result':
[{'words': '沈阳普特仪表有限公司'}, {'words': '北京普特仪表厂(沈阳办)'}, {'words': '郭飞'},
{'words': '经理'}, {'words': '手机:13898145123'}, {'words':
'地址:沈阳市和平区民主路77号二楼101号'}, {'words': '电话:024-23524589传真:024-23524589'},
{'words': '手机:13898145123'}]}
[Finished in 5.4s]
```

其他结果见result.txt。