

# 一种名片图像的文字区块分割方法

胡景锋, 吕 岳

(华东师范大学计算机科学与技术系, 上海 200062)

**摘 要:** 针对以手机摄像头作为图像采集器获得的名片图像, 介绍一种文字区块的分割方法。对用手机摄像头获取的彩色名片图像, 利用4个方向的 sobel 算子进行边缘检测, 并对边缘图进行二值化处理, 求出连通元, 通过连通元的合并和去噪声处理找到名片图像上的文字区块。实验结果表明, 该方法适用于复杂背景的名片图像的字符区块分割问题, 可适应文字在语种、字体、尺度、色彩等方面的变化, 对文字区域的定位准确率较高。

**关键词:** 边缘检测; 字符块分割; 名片识别; 连通元

## Text Areas Segmentation Approach for Business Card Images

HU Jing-feng, LV Yue

(Department of Computer Science and Technology, East China Normal University, Shanghai 200062)

**【Abstract】** The algorithm is presented for detecting text areas in business card from color image captured by mobilephone camera. It is carried out by three steps: edge detection using sobel arithmetic operators in four directions, connected component analysis, text area candidate detection and decision. Experimental results demonstrate that the proposed approach has good performance for segmenting text areas in mobilephone camera captured images with varying font-sizes, font-colors, background and languages.

**【Key words】** edge detection; character block segmentation; business card recognition; connected component

### 1 概述

名片是现代商务活动中一种重要的信息载体。大量名片给人们带来方便的同时也带来日常管理的麻烦, 很多手机和PDA虽然可用于名片管理, 但其中能实现名片自动输入的并不多, 或效果不甚理想。随着文本图像分析技术的发展尤其是文字识别技术的日趋成熟, 印刷的简体汉英混排字符的识别正确率已超过 97%<sup>[1]</sup>。手机的图像采集系统(Camera)、处理器(MPU)系统和内存系统(RAM)的日益完善也为手机识别名片提供了必要的硬件基础, 基于摄像头的文本图像信息的识别技术正在飞速发展<sup>[2]</sup>。

提取名片图像中的信息主要包括以下4个部分:

- (1) 获取名片图像。
- (2) 从名片图像中提取出有意义的文字区块。
- (3) 对字符块进行识别。
- (4) 对识别出字符串提取信息并进行理解。

名片图像可以直接通过手机摄像头获取, 字符识别技术已十分完善。与算法比较成熟的文字识别相比, 图像中的文字定位还须深入研究。在整个名片图像的识别系统中, 文字区域分割成为一个关键步骤, 直接关系到整个系统性能的优劣。

近年来, 学者对彩色文档、图像和视频帧图像中有无字符块及字符块的位置、分割等问题进行了大量的研究<sup>[2]</sup>, 取得较好的效果, 但大多是对扫描仪或摄像机获得的高质量图像进行处理, 且各有自己侧重的研究重点和研究对象。用于对摄像头拍摄的文字图像信息的处理和识别的有效方法多是针对文档对象进行处理<sup>[3]</sup>, 面对背景复杂、字体、语言多样的名片图像, 其效果并不理想。

### 2 名片图像识别算法简介

传统的名片图像识别是先对彩色图像进行灰度化、二值化、平滑、去噪等预处理, 再对得到的二值图像进行版面分析、分割和识别。文献[4]对版面复杂的名片图像进行了版面分析和研究, 但只是针对扫描仪获取的名片图像。而在手机拍摄图像的过程中, 由于光照、背景和拍摄本身的原因, 可能会使名片图像产生几何变形, 名片的矩形边框在拍摄获得的图像中不再是矩形, 变成了不规则的四边形, 有的甚至产生扭曲。另外, 名片图像中背景颜色不一, 图形、文字、线条交错, 使得直接对名片图像进行二值化处理的效果并不理想, 甚至无法处理, 既可能丢失一些有用信息, 也给分割和识别工作带来很多困难。

通过分析样本名片图像可见, 由于无法保证摄像头与名片平面相平行, 使名片图像发生形变, 其中名片的外围矩形边框发生形变最明显, 变成不规则的四边形, 有的边框甚至不再是直线, 但从局部看, 名片图像中的字符块图像和图形(Logo 等)块等关键信息局部图像发生的形变却不太明显, 许多识别算法可以进行有效处理。本文针对文字区块的分割问题介绍一种方法, 在手机摄像头获取的彩色名片图像的基础上, 利用4个方向的 sobel 算子进行边缘检测, 得到对应的边缘图像, 然后根据连通元的特征提取出文字区块和图形

**基金项目:** 国家自然科学基金资助项目(60475006); 上海市曙光人才计划基金资助项目(05SG29); 教育部新世纪优秀人才计划基金资助项目(NCET-05-0430)

**作者简介:** 胡景锋(1984 - ), 男, 硕士研究生, 主研方向: 图像处理和模式识别; 吕 岳, 教授

**收稿日期:** 2007-12-03 **E-mail:** jf\_hu@tom.com

区块图像。算法流程如图 1 所示。

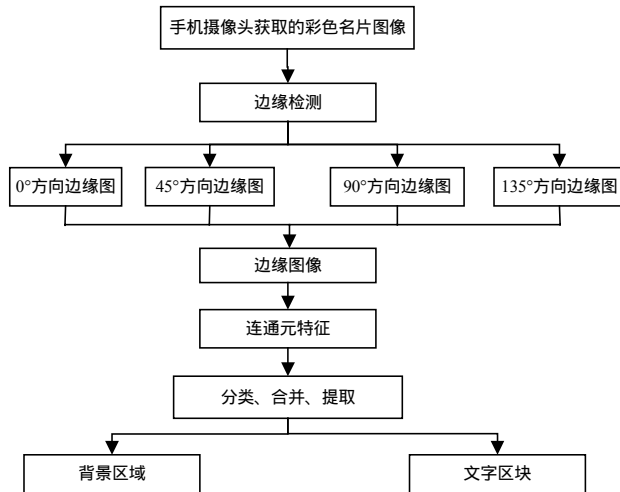


图 1 本文算法流程

### 3 文字块提取

字符是由笔画组成的,包括竖笔画、横笔画、斜笔画,其中斜笔画有 2 种,45°角方向和 135°角方向,综合起来就是 0°,45°,90°和 135°这 4 种方向的笔画。和名片图像中的背景相比,字符块图像的边缘特征非常明显,因此从名片图像中找出富含边缘的区域,即可能的文字区块。最后,对这些可能的文字区块分析,经过去除噪声,与区块合并等处理,便可得到文字区块图片。

#### 3.1 边缘检测

用于检测垂直、水平方向的边缘一般选用sobel算子,本文加入 45°和 135°方向的两个算子<sup>[2]</sup>,如图 2 所示。

-1	-2	-1	-2	-1	0	-1	0	1	0	-1	-2
0	0	0	-1	0	1	-2	0	2	1	0	-1
1	2	1	0	1	2	-1	0	1	2	1	0

图 2 4 个方向的 sobel 算子

通过选用图 2 中 4 个方向的sobel算子,对手机摄像头采集到的彩色名片图像RGB每个分量进行处理,得到 4 个方向的边缘图。用 $I_\theta$ 表示对  $\theta$ 角方向得到的边缘图,取 4 个方向上的最大值作为图像的边缘值,即

$$I = \max_{\theta \in C} I_\theta \quad C = \{0, \pi/4, \pi/2, 3\pi/4\}$$

上述处理是针对彩色图像的一个分量,笔者对RGB 3 个分量作同样的处理,并用 $I_R, I_G, I_B$ 分别表示 3 个分量,最后得到的彩色边缘图像如图 3 所示。

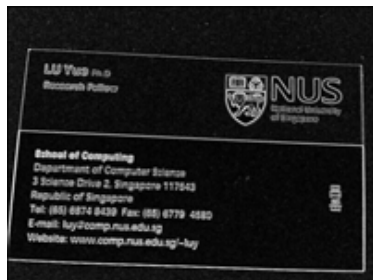


图 3 边缘检测后的边缘图像

#### 3.2 特征提取

求出名片图像的边缘图像以后,边缘图中亮度较大的区域为可能的字符区块,为找到这些区域,先对边缘图二值化,

再选择合适的特征区分出背景区域和字符区域。

##### 3.2.1 灰度化

对上述所求得到的彩色边缘图进行灰度化,彩色图像是用 RGB 值表示一个像素点,而灰度图像就是一个灰度级。这是一个三维向量到一维量的一个投影。通常用一个线性变换来完成灰度化的过程,如:

$$P(x, y) = k \cdot I_R(x, y) + l \cdot I_G(x, y) + m \cdot I_B(x, y)$$

其中, $I_R(x, y), I_G(x, y), I_B(x, y)$ 分别表示像素点 $(x, y)$ 处红、绿、蓝分量的值; $P(x, y)$ 表示对应像素点的灰度值; $k, l, m$ 是预先确定的参数,并且 $k+l+m=1$ 。

##### 3.2.2 二值化

对灰度边缘图进行二值化处理,把多灰度级图像转化为仅包含黑白两级的图像。本文采用迭代法选取图像最佳分割阈值的算法,设定最大迭代次数为  $t$ ,进行二值化。

##### 3.2.3 特征提取

从二值图像的内容可以看出,边缘像素点相互连通。在二值图像中,基于二值信息的几何特征是区分文本,图形和图像的重要特征之一,常用的几何特征包括连通元素的高度、宽度、宽高(或高宽)比、周长、像素点密度等<sup>[5]</sup>。本文选择连通元素的位置 $p(x, y)$ 、高度 $h$ 、宽度 $w$ 、宽高(或高宽)比 $r$ 这些几何特征区分文字区块和背景。

对二值图像求连通元(8-连通)<sup>[6]</sup>,求出包含每个连通元的最小矩形框,并计算每个连通元矩形框的在整幅图像中的坐标 $p(x, y)$ 、宽度 $w$ 和高度 $h$ , $r$ 值定义为

$$r = \min(h/w, w/h)$$

图 4 显示的矩形框标识了每一个求得的连通元。

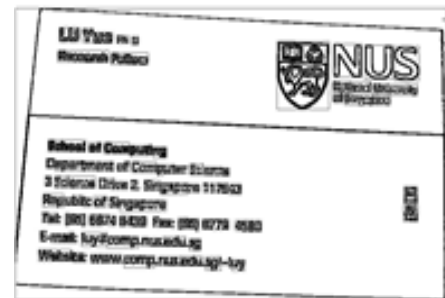


图 4 二值边缘图的连通元

#### 3.3 特征选择与分类

设二值边缘图中求得的连通元素结果为 $B = \{B_1, B_2, B_3, \dots, B_n\}$ , $B_i$ 表示第 $i$ 个连通元素,连通元素总个数为 $n, i = 1, \dots, n, B_i = (p_i, (x_i, y_i), w_i, h_i, r_i)$ 分别记录每个连通元素的左下角坐标、宽度、高度和宽高比。名片图像的宽度和高度分别记为 $W$ 和 $H$ 。

名片图片中有边框,即图像中名片区域和背景区域的交界,表现在已求出连通元素中面积较大的那个连通元即是边框对应的边缘图像,记此连通元为 $B_{\max}$ 。考虑到拍摄名片图像时,背景和名片底色可能是一致的,因此,在名片图像中就不存在边框,边缘图像中也不存在 $B_{\max}$ 。

判断有无 $B_{\max}$ 的条件为

$$w(B_{\max})/W > r_1, \quad h(B_{\max})/H > r_1$$

实验中只需是名片图像中的信息,因此提取出坐标落在 $B_{\max}$ 中的连通元素即可。如果不存在 $B_{\max}$ ,则选择所有的 $B_i$ 。

另外,一些狭长的线条予以去除,它们需要满足的条件如下:

$$\min(w(B_i), h(B_i)) < r_2 \quad \text{and} \quad r < r_3$$

### 3.4 字符区块提取

本文采用连通元合并的办法来分割出名片图像中的字符区块(block), 用一个矩形(box)对所有的连通元素进行扫描, 如果有两个或两个以上的连通元素与此 box 相交, 则认为其相邻, 合并这两个(或多个)连通元为一个大的区块。

以两个连通元为例说明, 设选取的box大小为 $w_b \times h_b$ , 待判断的两个连通元的坐标分别为 $p_i(x_i, y_i)$ ,  $p_j(x_j, y_j)$ , 宽度和高度分别为 $(w_i, h_i)$ ,  $(w_j, h_j)$ , 那么如果它们满足下面条件, 即可认为它们是相邻的。

$$|x_i - x_j + \frac{w_i - w_j}{2}| \leq w_b + \frac{w_i + w_j}{2}$$

$$|y_i - y_j + \frac{h_i - h_j}{2}| \leq h_b + \frac{h_i + h_j}{2}$$

将两个邻近的连通元合并, 合并后的连通坐标记为 $p_m(x_m, y_m)$ , 宽度和高度记为 $w_m, h_m$ 则有

$$x_m = \min(x_i, x_j), y_m = \min(y_i, y_j),$$

$$w_m = |x_i - x_j + \frac{w_i - w_j}{2}| + \frac{w_i + w_j}{2},$$

$$h_m = |y_i - y_j + \frac{h_i - h_j}{2}| + \frac{h_i + h_j}{2}$$

对所有的连通元素进行处理后, 须去除孤立的噪声连通元, 其判断标准为

$$w < w_b \text{ or } h < h_b$$

如果进行合并处理后的连通区块中满足条件, 则认为是噪声连通元, 予以去除。合并并去除噪声后的连通元素就是所要求的字符区块对应的位置。

最后对得到的每个区块, 在原彩色名片图像中, 按各区块对应的位置和大小分割提取出彩色字符区块图像。

## 4 实验结果

本实验选用MOTOA1200 手机摄像头拍摄的 208 张图片进行处理, 包括中文名片、英文名片、中英文混合名片的图像, 获取的名片图像的分辨率为 200 万像素, 图片格式为jpg。实验采用测试参数为 $k=l=m=1/3$ ,  $t=50$ ,  $r_1=2/3$ ,  $r_2=10$ ,  $r_3=0.04$ ,  $w_b=h_b=16$ 。结果表明, 99.5%的图片成功分割出字符区块, 尽管有些名片的背景比较复杂, 而且手机在拍摄过程中由于摄像头本身的原因, 名片图像会产生变形, 本文方法仍能取得比较理想的分割效果。

图 5 为原图中对应分割出的文字区块, 图 6 为分割出的各字符区块, 图 7 给出更多的实验结果。

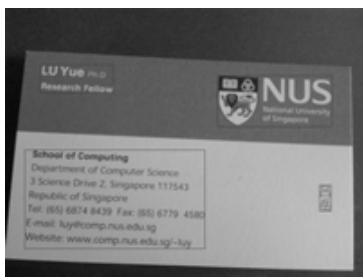


图 5 合并后的文字区块



图 6 分割出的各字符区块

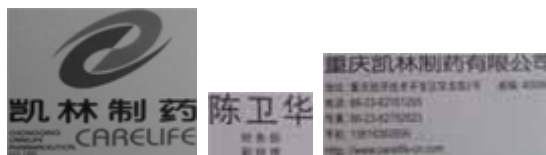


图 7 其他文字区域分割实例

## 5 结束语

本文提出一种新的分割方法, 可在多背景的情况下, 从名片图像中提取出文字区块, 包括了不同颜色、不同字体或者不同种语言的文字。该方法完整保留了文字的颜色信息, 避免了信息丢失。实验结果表明, 该算法在多种类型的名片图像中都有很好的工作效率。今后将加强对文字的自动识别、名片内容的识别和理解, 以及自动建立通讯录等技术的研究, 完善整个系统。

## 参考文献

- [1] 丁晓青, 郭繁夏. 中文OCR技术最新进展(续)[J]. 电子出版, 1996, (1): 7-9.
- [2] Liu Chunmei, Wang Chunheng, Dai Ruwei. Text Detection in Images Based on Unsupervised Classification of Edge-based Features[C]//Proceedings of the 8th International Conference on Document Analysis and Recognition. Seoul, Korea: [s. n.], 2005: 610-614.
- [3] Donaldson K, Myers G K. Bayesian Super-Resolution of Text in Video with a Text-Specific Bimodal Prior[C]//Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, San Diego. CA, USA: [s. n.], 2005: 1188-1195.
- [4] Lin Lin, Chew Lim Tan. Text Extraction from Name Cards with Complex design[C]//Proceedings of the 8th International Conference on Document Analysis and Recognition. Seoul, Korea: [s. n.], 2005: 977-980.
- [5] 张利, 朱颖, 吴国威. 版面分割中游程平滑后的图文特征分类[J]. 中国图象图形学报, 1998, 3(10): 827-831.
- [6] 徐正光, 鲍东来, 张利欣. 基于递归的二值图像连通域像素标记算法[J]. 计算机工程, 2006, 32(24): 186-188, 225.