

# 基于透视变换的手机身份证字符分割算法

李志杰<sup>1</sup>, 袁鹏泰<sup>1</sup>, 李博涵<sup>2</sup>

(1. 南京航空航天大学 计算机科学与技术学院, 江苏 南京 211106;

2. 江苏易图地理信息科技股份有限公司, 江苏 扬州 225009)

**摘要:**在 Android 手机下,提出一种改进的从复杂背景中分割出无畸变身份证字符图像的技术。对于手机扫描得到的满足要求的图像进行预处理,包括:灰度化、缩放、中值滤波、自适应阈值 Canny 边缘检测;在此基础上使用概率 Hough 变换得到边缘检测图中的直线。为了避免同一条直线出现断裂的情况,需要对 Hough 变换后的图像进行同一条直线的连接处理。在连接后的直线中,找出身份证外圈的四条轮廓线并求出四条线的交点,利用这四个交点和透视变换分割出无畸变、无背景的目标身份证图像。根据所需信息区域在身份证上位置固定,利用比例坐标,在之前分割出的无背景身份证图像上,分割出含有所需信息的矩形区域图像,最终用投影法做行分割和单字符分割。实验结果表明,该方法具有较高的分割速度与正确率。

**关键词:**Android 手机;复杂背景;身份证字符分割;Canny 边缘检测;Hough 变换;透视变换;投影法

中图分类号:TP301.6

文献标识码:A

文章编号:1673-629X(2018)07-0058-05

doi:10.3969/j.issn.1673-629X.2018.07.013

## A Segmentation Algorithm of Identification Card Characters of Mobile Phone Based on Perspective Transform

LI Zhi-jie<sup>1</sup>, YUAN Peng-tai<sup>1</sup>, LI Bo-han<sup>2</sup>

(1. School of Computer Science and Technology, Nanjing University of  
Aeronautics and Astronautics, Nanjing 211106, China;

2. Jiangsu Province Easy Map Geographic Information Technology Company Limited by Shares,  
Yangzhou 225009, China)

**Abstract:** We put forward an improved technique to segment the image of the undistorted ID card characters from the complex background in Android mobile phone. The image obtained by scanning and satisfying special conditions is preprocessed firstly including graying, scaling, median filtering and self-adaptive threshold Canny edge detection. Based on previous steps, the probabilistic Hough transform is used for line detection in edge detection image. To avoid a fracture in the same straight line, it is necessary to connect the same line of the image after the Hough transformation. In the straight line after the connection, we find out the outer four contours of the ID card and find the intersection point of the four lines. Using the four intersection points and the perspective transformation, we can segment the target ID card image without distortion and background. Because the required information is fixed on the ID card, the rectangular image of the area containing the required information can be segmented on the ID card image without background which is previously segmented by using the proportional coordinate. Finally, the projection method is used for row segmentation and single character segmentation. Experiment shows that this method has higher segmentation speed and correct rate.

**Key words:** Android mobile phone; complex background; segmentation of ID card characters; Canny edge detection; Hough transform; perspective transform; projection method

## 0 引言

我国的居民身份证记载了每位公民的个人基本信息,包括姓名、性别、民族、出生年月日、住址和唯一的

身份证编号。不仅是基于社会治安的需要和个人权益的保障,而且在人才招聘等场合也需要输入公民身份证信息。

收稿日期:2017-05-17

修回日期:2017-09-26

网络出版时间:2018-02-24

基金项目:国家自然科学基金((61375021))

作者简介:李志杰(1992-),男,硕士,研究方向为数字图像处理和模式识别;李博涵,副教授,研究方向为地理信息系统和图像处理。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20180224.1510.002.html>

随着携带高清摄像头的安卓手机越来越普及,研究采用基于带摄像头的安卓手机拍摄图像的身份证识别系统,可让使用成本更低、运用更加灵活。

数字图像处理和模式识别技术在安卓手机上应用广泛,其中就包括身份证的自动识别。这使得身份证的信息获取从人工识别转化为自动识别。目前身份证信息大多需要人工录入和手工填写,效率十分低下,录入和填写信息时可能发生错误,而且长时间的识别也会使人眼疲劳。所以,该方法已经不适应当今智能手机等领域飞速发展的现状。研究安卓手机如何自动读取身份证信息是十分必要的,不仅可以有效克服人工识别的局限性,而且具有识别效率高、识别准确度高的优点。

在身份证识别系统<sup>[1]</sup>中,采集到的含有身份证的图像很可能会同时存在以下一种或几种问题:除了身份证之外的背景与身份证的边缘不明显;身份证不是一个宽和高比例固定的矩形;身份证上有反光等光照影响;采集图像上有噪声。从复杂背景中分割出无畸变的身份证字符图像<sup>[2-4]</sup>是关键环节,决定了身份证识别系统的识别速度与正确率。对此,文中提出了一套完整的身份证图像分割和字符分割方法。

## 1 预处理

通过安卓手机扫描采集到的图像是像素  $1\ 024 * 768$  的三通道图像,每个像素的  $R$  值、 $G$  值、 $B$  值相等。其中, $R$  值、 $G$  值、 $B$  值分别为三通道图像中像素的红、绿、蓝分量值。扫描要求:身份证四条外围线的位置分别在图像上  $1/5$  内、下  $1/5$  内、左  $1/5$  内、右  $1/5$  内。

首先,对采集图像按式 1 计算各像素的灰度值 GrayValue,目标是将采集图像灰度化:

$$\text{GrayValue} = (306 * R + 601 * G + 117 * B) \gg 10 \quad (1)$$

其次,为了加快处理速度,采用线性插值将灰度图像缩放成像素  $341 * 256$  的图像,即缩放比例为  $3 : 1$ 。然后为了减小噪声的干扰,对缩放后的灰度图像进行中值滤波<sup>[5-6]</sup>。最后对图像进行自适应阈值 Canny 边缘检测<sup>[7-8]</sup>。

为了降低硬件资源的消耗,该设计采用线性插值法进行图像缩放。线性插值法就是根据待插值位置周围最近的几个像素来计算插值,对于平面图像来说,共有四点。其中插值计算公式为:

$$\begin{aligned} f(i+u, i+v) = & (1-u)(1-v)f(i, j) + \\ & (1-u)vf(i, j+1) + \\ & u(1-v)f(i+1, j) + \\ & uvf(i+1, j+1) \end{aligned} \quad (2)$$

其中, $u, v$  均为非负整数,分别表示像素的横纵坐

标位置; $u, v \in [0, 1]$  分别表示插值位置与  $(i, j)$  的相对位置; $f(i, j)$  表示图像  $(i, j)$  位置的像素值。

中值滤波 (median filter) 是一种基于排序统计理论的能有效抑制噪声的非线性信号处理技术。它是一种常用的非线性平滑滤波和邻域运算,类似于卷积,但不是加权求和计算,基本原理是把邻域中的像素按灰度等级进行排序,然后选择该组的中间值作为输出像素值,让周围的像素值不变,从而消除孤立的噪声点。邻域通常为  $3 * 3, 5 * 5$  区域,可以是不同的形状,如线状、圆形、十字形、圆环形等。

中值滤波对脉冲噪声有良好的滤除作用,特别是在滤除噪声的同时,能够保护信号的边缘,使之不被模糊。中值滤波的算法比较简单,也易于用硬件实现。这些优良特性是线性滤波方法不具有的。中值滤波在图像处理中常用于保护边缘信息,是经典的平滑噪声的方法。

在众多经典边缘检测<sup>[9]</sup>方法中选择 Canny 边缘检测算子,是因为相比传统的微分算子,如 Roberts、Prewitt、Sobel、Laplace、Log 等,基于最优化算法的 Canny 边缘检测算子具有较好的信噪比和检测精度,对传统单阈值法提出了改进,根据图像的梯度直方图分别选取高低两阈值。但传统 Canny 算子由于使用基于全图的固定阈值选取方法,容易丢失对比度较低的弱边缘。因此有相关学者提出了一种自适应阈值 Canny 边缘检测方法。

该方法<sup>[10]</sup>先利用图像的梯度方差作为判据对图像进行分块,图像上点  $(i, j)$  处的梯度幅值为:

$$A(i, j) = \sqrt{E_x^2(i, j) + E_y^2(i, j)} \quad (3)$$

其中:

$$\begin{cases} E_x = \frac{\partial G}{\partial x} * f(x, y) \\ E_y = \frac{\partial G}{\partial y} * f(x, y) \end{cases} \quad (4)$$

设所选的高斯函数  $G$  为:

$$G(x, y) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (5)$$

其中, $\sigma$  为高斯滤波器参数; $f(x, y)$  为原图像函数。

继而对各子分块采用最大类间方差法<sup>[11]</sup>分别求得最佳阈值。假设一幅图像灰度值分为  $0 \sim (l-1)$  级,灰度值  $i$  的像素数为  $n_i$ ,则总像素数  $N$  及各像素值概率  $P_i$  为:

$$\begin{cases} N = \sum_{i=0}^{l-1} n_i \\ P_i = n_i / N \end{cases} \quad (6)$$

然后用一整数  $t$  将图像中的像素按灰度级划分成

两类:  $C_0 = \{0, 1, \dots, t\}$  和  $C_1 = \{t+1, t+2, \dots, l-1\}$ 。

则  $C_0$  产生的概率  $\omega_0$  及均值  $\mu_0$  为:

$$\omega_0 = \sum_{i=0}^t p_i = \omega(t) \quad (7)$$

$$\mu_0 = \sum_{i=0}^t ip_i / \omega_0 = \mu(t) / \omega(t) \quad (8)$$

$C_1$  产生的概率  $\omega_1$  及均值  $\mu_1$  为:

$$\omega_1 = \sum_{i=t+1}^{l-1} p_i = 1 - \omega(t) \quad (9)$$

$$\mu_1 = \sum_{i=t+1}^{l-1} ip_i / \omega_1 = [\mu - \mu(t)] / [1 - \omega(t)] \quad (10)$$

其中,  $\mu$  为整体图像灰度的统计均值。

$$\mu = \sum_{i=0}^{l-1} ip_i = \omega_0 \mu_0 + \omega_1 \mu_1 \quad (11)$$

由上可得类间方差  $\eta$  及最佳阈值  $t^*$ 。

$$\eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (12)$$

$$t^* = \text{Argmax}_{t \in \{0, 1, \dots, l-1\}} \sigma_B^2 \quad (13)$$

其中:

$$\sigma_B^2 = \omega_0 (\mu_0 - \mu)^2 + \omega_1 (\mu_1 - \mu)^2 \quad (14)$$

最后采用插值算法得到最后的自适应阈值分割矩阵。效果见图 1。

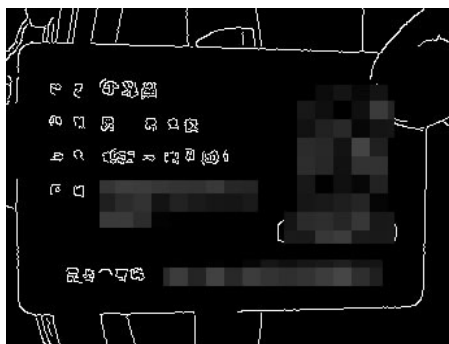


图 1 Canny 边缘检测图

## 2 求身份证四个顶点坐标

### 2.1 Hough 变换找直线

对 Canny 边缘检测后的图像使用概率 Hough 变换<sup>[12-13]</sup>。在 Hough 变换中, 图形边缘  $M$  每个像素点  $(x_i, y_i)$  被映射到一个代表正弦曲线投票模式的  $N_\theta \times N_\rho$  的累加器矩阵, 它描绘了一个标准的  $(\theta, \rho)$  参数平面, 代表所有可能通过该像素点的直线, 如式 15:

$$\rho = x \cos \theta + y \sin \theta \quad (15)$$

其中,  $\rho$  为离原点的距离;  $\theta$  为直线的方向。

对应于同一条直线上的所有像素点的正弦曲线在参数空间中交于理想的一个点, 这在累加器  $A(\rho, \theta)$  矩阵中会达到一个峰值。

概率 Hough 变换采用空间映射和直线检测交替进行的方式, 其基本思想是在整个图像空间中随机取

点映射, 每次映射后检测累加器  $A(\rho, \theta)$  对应点的数值, 当检测到超过特定阈值的直线时, 将位于该直线上的所有候选边缘点删除, 从而有效地缩小了搜索和计算空间<sup>[14]</sup>。

对 Hough 变换得到的直线集合, 因为身份证最外围的四条边会出现同一条直线产生断裂的情况, 为了接下来找出这四条线的准确性, 需要分别对位置在图像上 1/5 内、下 1/5 内、左 1/5 内、右 1/5 内的直线进行同一条直线的连接处理, 具体为:

(1) 当一条直线  $L_1$  在图像右 1/5 处时, 若一条直线  $L_2$  满足:  $L_2$  倾斜角度与  $L_1$  夹角在 5 度以内; 两条直线列间距在 3 个像素点以内; 两条直线行间距在 30 个像素点以内; 两条直线的长度均大于 7 个像素点, 则将两条直线连在一起。

(2) 当一条直线  $L_3$  在图像左 1/5 处时, 若一条直线  $L_4$  满足:  $L_4$  倾斜角度与  $L_3$  夹角在 5 度以内; 两条直线列间距在 3 个像素点以内; 两条直线行间距在 30 个像素点以内; 两条直线的长度均大于 7 个像素点, 则将两条直线连在一起。

(3) 当一条直线  $L_5$  在图像上 1/5 处时, 若一条直线  $L_6$  满足:  $L_5$  倾斜角度与  $L_6$  夹角在 5 度以内; 两条直线行间距在 3 个像素点以内; 两条直线列间距在 30 个像素点以内; 两条直线的长度均大于 7 个像素点, 则将两条直线连在一起。

(4) 当一条直线  $L_7$  在图像下 1/5 处时, 若一条直线  $L_8$  满足:  $L_7$  倾斜角度与  $L_8$  夹角在 5 度以内; 两条直线行间距在 3 个像素点以内; 两条直线列间距在 30 个像素点以内; 两条直线的长度均大于 7 个像素点, 则将两条直线连在一起。

其中, 直线两个端点的坐标已知, 分别为  $(x_1, y_1)$  和  $(x_2, y_2)$ , 由直线公式  $y = kx + b$  求得直线的斜率  $k$ :

$$k = (y_2 - y_1) / (x_2 - x_1) \quad (16)$$

然后由反正切函数公式和直线斜率的绝对值求直线的倾斜角度  $\theta$ :

$$\theta = \arctan(|k|) \quad (17)$$

最后直线长度  $\text{dist}$  由式 17 可得:

$$\text{dist} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (18)$$

### 2.2 求身份证四条外围边及其交点

在连接后的直线集合中, 需要找到身份证外围的四条直线。由找到的四条直线求出在预处理后图像中身份证的四个顶点坐标, 由缩放比例可得在采集图像中身份证四个顶点的坐标, 过程如下:

(1) 分别在预处理后图像的上 1/5 和下 1/5 内, 找到与水平的夹角在 0 度至 20 度区间的最长线;

(2) 分别在预处理后图像的左 1/5 和右 1/5 内,找到与竖直线相交角度在 0 度到 20 度区间的最长线;

(3) 对求得的四条直线求交点。假设两条直线方程分别为  $y = k_1x + b_1$  和  $y = k_2x + b_2$ , 则交点坐标为:  $((b_2 - b_1)/(k_1 - k_2), (k_1b_2 - k_2b_1)/(k_1 - k_2))$ 。

### 3 透视变换

以矩形区域的四个角点为例,经过透视变换后  $A' \rightarrow A, B' \rightarrow B, C' \rightarrow C, D' \rightarrow D$ , 矩形的形状发生改变。透视投影实质上就是  $P'$  平面上的每一点在视角的作用下投影到  $P$  平面的过程。若将  $P$  平面定义为物体的正视图平面,那么透视变换就是透视投影平面上的每一个像素点变换到正视图上对应的像素点的过程。为了实现透视畸变的校正,只需要找到  $P$  平面上的点和物体正视图上对应点的变换公式。

二维图像透视变化公式可以表示为:

$$u = \frac{ax + by + c}{gx + hy + 1} \quad (19)$$

$$v = \frac{dx + ey + f}{gx + hy + 1} \quad (20)$$

其中,  $(x, y)$  为投影图像的坐标;  $(u, v)$  为正视图上的坐标;  $a, b, c, d, e, f, g, h$  为畸变参数。

由透视变换<sup>[15-16]</sup>和采集图像中身份证四个顶点的坐标,在采集图像中分割出宽和高比例在 1.49 ~ 1.69 区间的无背景身份证图像,并且解决图像因采集而产生的图像畸变问题。

### 4 目标信息区域分割

根据姓名、性别、身份证号三个信息区域在身份证上位置固定,利用比例坐标,在以上步骤分割出的无背景身份证图像上,分割出含有以上三个信息的矩形区域图像,分割后的结果见图 2~4。

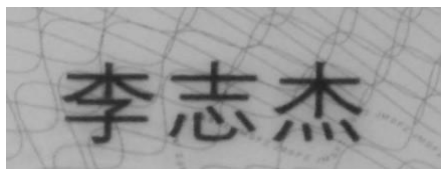


图2 姓名信息区域



图3 性别信息区域



图4 身份证号信息区域

### 5 字符分割

以姓名信息区域的字符分割为例,其他部分的字符分割<sup>[17-18]</sup>同理可得。

对分割出来的采集图像上的姓名区域,先灰度化,再大津阈值二值化。

对二值化后的姓名区域图像进行投影法字符分割,将各个字符分割开,具体过程如下:

(1) 先从上到下统计每行像素值为 0 的像素点个数,即统计每行黑色像素点个数,找到个数为 0 的行的行号(这一行在图像中的第几行),即找到一行全是白色像素点的行号,如果此时出现连续个数为 0 的行,则找出连续行号的中间值,由此得到分割行的行号,即得到下一步图像水平分割的分割点;

(2) 在姓名区域彩色图像上进行行分割;

(3) 对行分割后的二值图,从左到右统计每列像素值为 0 的像素点个数,找到个数为 0 的列,如果此时出现连续个数为 0 的列的列号(这一列为图像中的第几列),则找出连续列号的中间值,由此得到分割列的列号;

(4) 在行分割后的姓名区域彩色图像上进行列分割,这样就可以得到姓名区域分割出来的彩色字符图像。分割后的字符图像见图 5。

值得注意的是,对于类似“刘”、“小”、“川”这样有内部间隔即容易被分割开的汉字,可以利用字内间距比字间间距小和汉字字体大小固定设定阈值进行分割。

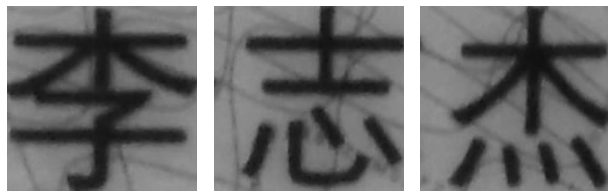


图5 姓名区域分割出的字符图像

### 6 实验结果与分析

实验是在带有 1 600 万像素摄像的 Android 手机上进行。使用 JNI 机制,采用 C 语言进行算法的程序化,将写好的 C 源文件用 ndk 编译成 so 库以供 Java 调用。

实验使用扫描身份证的方法进行识别,并且在 C 源文件中统计身份证字符分割部分所耗费的时间,显示在 AndroidStudio 中。

一共进行了 50 组实验。因为汉字分割是字符分割的重难点,所以用手机扫描了 50 组不一样的身份证,50 组姓名中包含了各类汉字,比如汉字是上下结构、左右结构等种类。统计数据得:完成一次身份证字符分割,平均耗费时间为 0.4 s,可以看出基于透视变



换的手机身份证字符分割速度很快;实验中字符分割基本正确,正确率达到 99%。

### 7 结束语

针对安卓手机上复杂背景中的身份证字符分割,设计了一种基于透视变换的算法,并设计了一种目标区域内连接同一条直线的方法。实验结果表明,该方法具有较为良好的分割能力。另一方面,该方法对全白背景中的身份证字符分割效果较差,还需要进一步的探索。

#### 参考文献:

[1] 严 曲. 身份证识别系统的原理及算法研究[D]. 长沙:中南大学,2005.

[2] 许新征,丁世飞,史忠植,等. 图像分割的新理论和新方法[J]. 电子学报,2010,38(2A):76-82.

[3] PAL N R, PAL S K. A review on image segmentation techniques[J]. Pattern Recognition, 1993, 26(9):1277-1294.

[4] 李 静,卢凯旋. 二代身份证的自动分割方法研究[J]. 计算机工程与应用,2015,51(14):165-169.

[5] ALVAREZ L, LIONS P L, MOREL J M. Image selective smoothing and edge detection by nonlinear diffusion. II[J]. SIAM Journal on Numerical Analysis, 1992, 29(3):845-866.

[6] 赵高长,张 磊,武风波. 改进的中值滤波算法在图像去噪中的应用[J]. 应用光学,2011,32(4):678-682.

[7] 许宏科,秦严严,陈会茹. 一种基于改进 Canny 的边缘检测

算法[J]. 红外技术,2014,36(3):210-214.

[8] 金 刚. 自适应 Canny 算法研究及其在图像边缘检测中的应用[D]. 杭州:浙江大学,2009.

[9] 段瑞玲,李庆祥,李玉和. 图像边缘检测方法研究综述[J]. 光学技术,2005,31(3):415-419.

[10] 张 帆,彭中伟,蒙水金. 基于自适应阈值的改进 Canny 边缘检测方法[J]. 计算机应用,2012,32(8):2296-2298.

[11] 杜 奇,向健勇,袁胜春. 一种改进的最大类间方差法[J]. 红外技术,2003,25(5):33-36.

[12] 段汝娇,赵 伟,黄松岭,等. 一种基于改进 Hough 变换的直线快速检测算法[J]. 仪器仪表学报,2010,31(12):2774-2780.

[13] ILLINGWORTH J, KITTLER J. The adaptive Hough transform[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 1987, 9(5):690-698.

[14] 王燕清,辛柯俊,陈德运,等. 基于启发式概率 Hough 变换的道路边缘检测方法[J]. 计算机科学,2013,40(9):279-283.

[15] 牛 彦. 关于透视变换的研究[J]. 计算机辅助设计与图形学学报,2001,13(6):549-551.

[16] 代 勤,王延杰,韩广良. 基于改进 Hough 变换和透视变换的透视图像矫正[J]. 液晶与显示,2012,27(4):552-556.

[17] 肖西华,江志兴,梁 旭,等. 移动平台下的身份证图像字符分割方法研究[J]. 计算机工程与应用,2015,51(24):201-204.

[18] 苗红霞,张 龙,徐文杰,等. 一种身份证图像字符分割的改进方法[J]. 微处理机,2016,37(3):51-55.

(上接第 57 页)

[9] HINTON G E. Learning distributed representations of concepts[C]//Proceedings of the eighth annual conference of the cognitive science. [s. l.]:[s. n.], 1986:1-12.

[10] BENGIO Y, DUCHARME R, VINCENT P, et al. A neural probabilistic language model[J]. Journal of Machine Learning Research, 2003, 3:1137-1155.

[11] MNIH A, HINTON G. Three new graphical models for statistical language modelling[C]//Proceedings of the twenty-fourth international conference on machine learning. Corvallis, Oregon, USA: ACM, 2007:641-648.

[12] MIKOLOV T, KARAFIAT M, BURGET L, et al. Recurrent neural network based language model[C]//11th annual conference of the international speech communication association. Makuhari, Chiba, Japan:[s. n.], 2010:131-138.

[13] COLLOBERT R, WESTON J, BOTTON L, et al. Natural language processing (almost) from scratch[J]. Journal of Machine Learning Research, 2011, 12(1):2493-2537.

[14] MIKOLOV T, CHEN Kai, CORRADO G, et al. Efficient estimation of word representations in vector space[C]//International conference on learning representations workshop track. [s. l.]:[s. n.], 2013:1301-1313.

[15] KIM Y. Convolutional neural networks for sentence classification[C]//Proceedings of the 2014 conference on empirical methods in natural language processing. [s. l.]:[s. n.], 2014:1746-1751.

[16] PENNINGTON J, SOCHER R, MANNING C D. Glove: global vectors for word representation[C]//Proceedings of the empirical methods in natural language processing. [s. l.]:[s. n.], 2014:1532-1543.