

·信号与信息处理·

## 面向药瓶标签的字符定位方法研究

姜月秋<sup>1</sup>, 仇 维<sup>1</sup>, 高宏伟<sup>2</sup>

(1.沈阳理工大学通信与网络工程中心, 辽宁 沈阳 110159; 2.沈阳理工大学信息科学与工程学院, 辽宁 沈阳 110159)

**摘要:**针对药瓶标签的特点,提出了一种数学形态学和投影特征相结合的药瓶标签字符定位方法。该方法首先对药瓶标签图像进行预处理,减小圆柱面造成的字符畸变程度,对处理得到的图像进行二值化和边缘检测等。其次对图像进行多次数学形态学处理,使图像分割成几个连通区域,根据先验知识以及轮廓跟踪方法得到字符所在的大致区域。最后利用水平投影和图像分割方法将粗定位得到的图像分割成三个单行字符,完成对字符区域的精确定位,便于后续的单字符分割和识别。实验结果表明,该方法能对药瓶标签进行有效的定位。

**关键字:**药瓶标签;字符定位;数学形态学;膨胀;投影特征

**中图分类号:**TP391.4

**文献标识码:**A

**文章编号:**1673-1255(2012)05-0047-05

## Character Positioning Methods of Medicine Bottle Label

JIANG Yue-qiu<sup>1</sup>, QIU Wei<sup>1</sup>, GAO Hong-wei<sup>2</sup>

(1. Communication and Network Engineering Center, Shenyang University of Science and Technology, Shenyang 110159, China;

2. Communication and Network Engineering Center, Shenyang University of Science and Technology, Shenyang 110159, China)

**Abstract:** According to the characteristics of medicine bottle labels, a method of character positioning which combines mathematical morphology with projective characteristics is provided. According to the method, firstly, the images of the medicine bottle labels are preprocessed so as to reduce the distortion extent of characters caused by cylindrical surfaces and the processed images will be performed binary and edge detection etc. Secondly, the images are performed mathematical morphology process many times to be divided into several connecting regions. According to the prior knowledge and contour tracking scheme, the characters area can be got. Finally, the images can be segmented into three separate characters by the level projection and image segmentation method. The accurate positioning of target characters is completed so as to segment and recognize succeeding single character. The experiment results show that the method is effective on the medicine bottle label position.

**Key words:** medicine bottle label; character positioning; mathematical morphology; dilation; projective characteristic

因药品质量关系到人的健康甚至生命安全,所以国家药品管理部门及药厂对药品的检测要求也越来越严格。其中药品外包装的打码效果就是一个重要的检测方面,打码内容包括药品的生产日期、批号以及截止日期,它传递着药品的详细信息。而光学字符识别(optical character recognition, OCR)技术是文字电子化过程中最重要的环节,目前市场已有的OCR产品已经相对成熟,它们已经通过市场检

验并被广大用户大规模使用。因此药瓶标签字符识别作为OCR技术的一种实际应用,有着非常广泛的应用前景。药瓶标签字符的定位是药瓶标签字符识别算法中关键的一步,它定位的准确与否直接影响到下面药瓶标签字符的分割和识别的准确率。

字符定位是图像处理 and 模式识别领域的一项传统的课题。尽管许多学者已经研究了多年,并在某些方面取得了不错的成果,但目前仍有许多学者在

收稿日期:2012-07-02

作者简介:姜月秋(1975-),女,博士,教授,研究方向为图像处理.

研究这项课题。由于具体的条件不同,存在的干扰不同,适用的范围不同,解决的方法也不相同,所以字符定位的研究仍具有理论和实践意义。目前字符定位研究的范围包括车牌定位、卡片(如电话卡,充值卡等)字符定位<sup>[1]</sup>、商品货号定位<sup>[2]</sup>、铁路货车编码定位<sup>[3]</sup>、邮政编码定位、纸币序列号定位等,其中对车牌定位方法的研究比较多。目前常用的车牌定位方法主要是沿着两条技术路线进行:基于灰度图像的车牌定位方法和基于彩色车牌图像的车牌定位方法。其中基于灰度图像的车牌定位方法包括基于边缘特征的定位方法、基于车牌纹理的定位方法、基于数学形态学的定位方法以及基于神经网络的定位方法。以药瓶为对象,结合数学形态学和投影特征对药瓶标签字符进行定位。

## 1 药瓶标签图像的预处理算法

在药瓶标签字符定位方面,与车牌定位有一定的区别:(1)字符的畸变不同,车牌的倾斜是二维的,它的倾斜是线性的,而药瓶标签字符所在表面是圆柱形的,且它变化时不满足线性关系,图像中央的字符畸变程度较小,图像两侧的字符畸变程度较大。传统的车牌倾斜校正的方法不能满足药瓶标签字符的校正。(2)干扰不同,车牌中车牌区域有和背景区域明显区别的颜色特征和字符密集等特征,而药瓶标签字符中,存在的干扰也是一些字符,和目标区域没有明显的区别。(3)车牌和药瓶标签本身固有特征不同,所以在确定分割阈值时所需要的约束条件不同,可根据药瓶标签本身固有的特征进行定位。其中车牌都是单排文字且车牌的宽高比以及矩形面积一定等,而药瓶标签字符是固定的三排,且三排所表示的内容是固定的。

### 1.1 药瓶标签字符区域畸变矫正

由于药瓶是圆柱体,照相机获得的圆柱面表面字符是由三维的曲面投影到二维平面后的图像。由于投影使得三维曲面上的字符在二维平面上有畸变,图像中央的字符畸变程度较小,图像两侧的字符畸变程度较大,越靠近两侧边缘,其字符畸变程度越大,这是一种非线性的畸变。因此,要想准确定位,必须对所获得的透视平面图像进行校正,将柱面展开成平面。

下面给出具体畸变校正算法<sup>[4]</sup>如下:把圆柱体的半径分为 $n$ 等份,每一份长度为 $l$ ,也就是说 $l \times n = R$ 。如图所示,每一份长度在柱面上对应的实际曲线长度分别是 $l_1, l_2, l_3 \dots l_n$ 。

根据推导可知

$$\cos \theta_i \times R = i \times l \quad (1)$$

$$l_i = (\theta_{i-1} - \theta_i) \times R \quad (2)$$

所以,由上述两式可得

$$l_i = (\arccos(i-1) \times l/R - \arccos(i \times l/R)) \times R \quad (3)$$

由式(3)可以得出投影平面与柱面展开平面之间的数学关系,由此式可进行带有畸变字符的非线性校正。



图1 原始图片

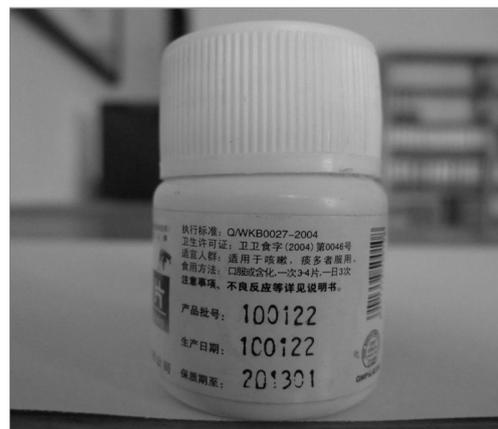


图2 矫正后图像

### 1.2 图像灰度化

在对药瓶标签字符进行定位前,需要对图像进行预处理,首先对图像进行灰度转换,灰度转换采用通用的RGB三原色到灰度的转换式

$$\text{Gray}(i,j) = 0.299R(i,j) + 0.587G(i,j) + 0.114B(i,j) \quad (4)$$

其中,  $\text{Gray}(i,j)$  为转换后的灰度图像在  $(i,j)$  处的灰度值;  $R,G,B$  分别为彩色图像的颜色分量。选择适当的阈值对图像进行二值化。

### 1.3 平滑滤波

噪声一般产生于图像的产生和传输过程中,是一种不可预测的随机信号。对于噪声的去除,在空间域中通常采用均值滤波或中值滤波方式来平滑图像,以达到去除噪声的效果。在此选择中值滤波对图像进行去噪处理。基本思想是选择模板,然后对噪声图像进行处理。对于选择  $3 \times 3$  的模板,假设  $f_1, f_2, f_3, f_4, f_5, f_6, f_7, f_8$  为模板对应像素点的值,重新排序后的序列为:  $f'_1, f'_2, f'_3, f'_4, f'_5, f'_6, f'_7, f'_8$ , 则滤波后此点的像素值为  $f'_5$ 。

### 1.4 边缘检测

在数字图像中,图像灰度产生剧烈变化的区域称之为边缘,它是灰度值不连续的结果。这种不连续可以利用梯度的方法来检测到,简单的梯度算子边缘检测方法自然就成为了边缘检测与提取的主要手段。常用的边缘检测算子有 Sobel 算子、Kirsh 算子、Prewitt 算子、Roberts 算子和 Gauss-Laplace 算子等<sup>[6]</sup>。通过对这几种算子的比较得出, Gauss-Laplace 算子因为卷积模板中带有高斯平滑滤波器而能够比其他一般算子更好地检测到图像的边缘。它首先对原始图像进行平滑滤波,然后再运用拉普拉斯算子检测边缘。由于 Prewitt 算子与 Sobel 算子大致相同,仅仅是权值不同,所以只对 Sobel 算子、Krish 算子、Roberts 算子和 Gauss-Laplace 算子进行了实验,其中采用 Krish 算子时运行需要的时间比较长,采用 Sobel 算子和 Roberts 算子时有些边界线会显得断断续续,只有 Gauss-Laplace 算子能够比较好地检测到边缘。

## 2 药瓶标签字符的定位

### 2.1 药瓶标签字符的粗定位

先对图像进行多次的数学形态学运算,将整个图像分割成几个连通区域,根据轮廓跟踪方法,提取出目标区域,如图3所示,方法如下:

(1) 数学形态学<sup>[7]</sup>是由一组形态学的代数运算子

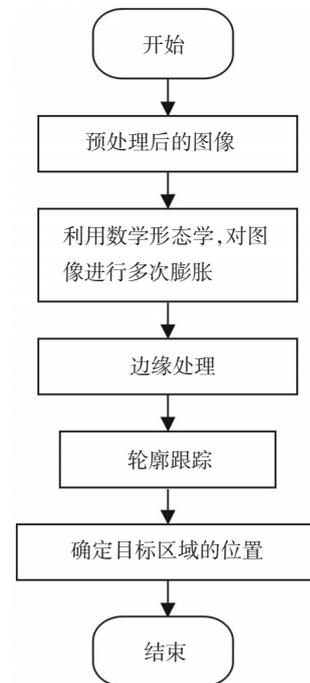


图3 字符粗定位的流程图

组成的。最基本的数学形态学运算子有:腐蚀、膨胀、开和闭运算。其中膨胀运算的作用就是把图像周围的背景合并到物体中。如果两个物体之间距离比较近,那么膨胀运算可能会使这两个物体连通在一起。膨胀运算的定义为  $X \oplus S = \{x | S[x] \cap x \neq \emptyset\}$ 。首先在 ImagePreproc 类中添加 HDilation 函数用于对图像进行水平方向的膨胀运算,结构元素选取为  $\{-1,0\},\{0,0\},\{1,0\}$ ,使整个图像分割成几个连通的区域。

(2) 对图像进行边缘提取,图像经过水平方向膨胀后边缘会形成连通区域,所以需要将边缘提取出来,否则影响后续的轮廓跟踪。边缘提取方法如下:在 DIB 类中添加 EdgeProcess 函数用于边缘提取,初始化一个新内存,将新像素点的灰度值设为 0,设源图像高度为 IHeight,宽度为 IWidth,对源图像进行逐点扫描,经过多次实验得到分界点为 100,当行数小于 100 或行数大于 IHeight-100 时将图像中所在范围内的像素点的灰度值取为 0,这样可以将横向边缘去掉。当列数小于 100 或列数大于 IWidth-100 时将图像中所在范围内的像素点的灰度值取为 0,这样可以将竖向边缘去掉。经过以上两步处理,可以将边缘去掉,便于下一步处理。

(3) 对图像进行轮廓跟踪,提取出字符所在的连通区域。具体方法如下:在 DIB 类中添加 TraceDIB

函数,设两个点为 MaxPoint 和 MinPoint,用来统计每个连通区域中的最高点和最低点所在的位置。将图像扫描的8个方向储存到一个数组 Direction[8][2]中,这八个方向为  $\{-1,1\},\{0,1\},\{1,1\},\{1,0\},\{1,-1\},\{0,-1\},\{-1,-1\},\{-1,0\}$ ,如图4所示。将图像按照从左到右,从下到上的顺序搜索,第一个白色像素点作为第一个边界点即起始点,起始扫描方向为左上方,如果左上方点是边界点,则求出其高度和宽度,和起始点高度和宽度进行比较,若当前点高度和宽度均大于起始点高度和宽度,则将当前点的值赋给 MaxPoint,若当前点高度和宽度均小于起始点高度和宽度,则将当前点的值赋给 MinPoint。将当前点作为新的边界点,并逆时针旋转  $90^\circ$ ,继续用同样的方法搜索下一个边界点,直到返回最初的边界点为止。

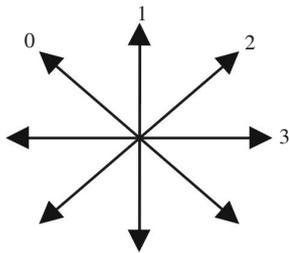


图4 扫描的八个方向

通过对整个图像的扫描,统计得到了每个连通区域的 MaxPoint 的宽度和高度、MinPoint 的宽度和高度。根据实验以及先验知识可以以下面公式作为判断准则,将同时满足以下两个条件的连通区域提取出来并将提取出的区域中的像素值设为 0 和 255 以外的其他值,在此将其设为 128。

$$\text{MaxPoint.Height} - \text{MinPoint.Height} \geq 170 \quad (5)$$

$$\text{MaxPoint.Width} - \text{MinPoint.Width} < 230 \quad (6)$$

最后,当整个图像扫描完以后,再重新将图像从左上方开始扫描,将图像中像素值为 128 的像素点全部赋值为 255,得到的白色区域即为图像字符所在的大致区域,如图5所示。

(4)经过前3个步骤的操作,即可得到字符图像的大致区域,其中可能会有一些干扰区域存在,所以

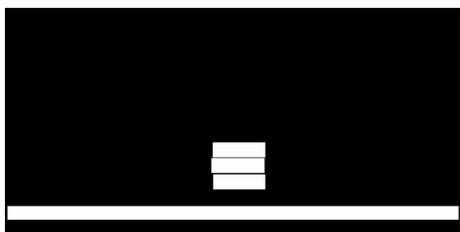


图5 轮廓提取和目标区域提取结果

需要对字符进行精确定位,如图6所示。

100122  
100122  
201321

图6 目标区域提取

## 2.2 药瓶标签字符的精定位

对药瓶标签字符进行粗定位后,由于可能会存在干扰的区域,所以需要对该图像进行精确定位。采用水平扫描的方法对图像进行处理,最终定位出3个单行字符,便于后续的单字符分割和字符的识别。如图7所示,具体的实现方法如下:

在 ImageAppView 类中添加函数 HprojectDIB 对粗定位后的图像进行行扫描,统计每一行的黑色点数,将当前行的黑色点数用 IBlackNumber 表示,前一行的黑色像素点用 B\_IBlackNumber 表示,对图像扫描完后,将图像投影到水平方向,完成对图像的水平投影。添加 Hscan 函数对整个投影图像进行行扫描,由于噪声的干扰。设一个 BOOL 型的变量 is\_BH-WhitePiont,若前一行的黑色像素点个数大于 10,则这个变量的值为 FLASE,否则为 TRUE.根据实验得到的判定准则,将3个字符区域的6个点分别保存在两个数组 Start\_Point[count]和 End\_Point[count]中,如图8所示,最终将单行字符区域提取出来,实现字符的精确定位。

## 3 结 论

根据药瓶标签的固有特征,提出了一种数学形态学和投影特征相结合的药瓶标签字符定位方法。粗定位过程根据数学形态学中的膨胀运算将图像分割成几个连通区域,根据轮廓跟踪和设定的约束条件确定了药瓶标签字符的大致区域,但还存在干扰区域,需要进行精确定位。精确定位过程中根据水平投影图和图像分割方法,利用先验知识将目标区域中的三行字符准确的分割成3个单行字符,保证了定位的准确性,而且有利于后续的单字符分割和识别。并利用 VC++6.0 平台对此方法进行了验证,通过采集的 100 幅图像进行实验,试验结果是:100 幅图像

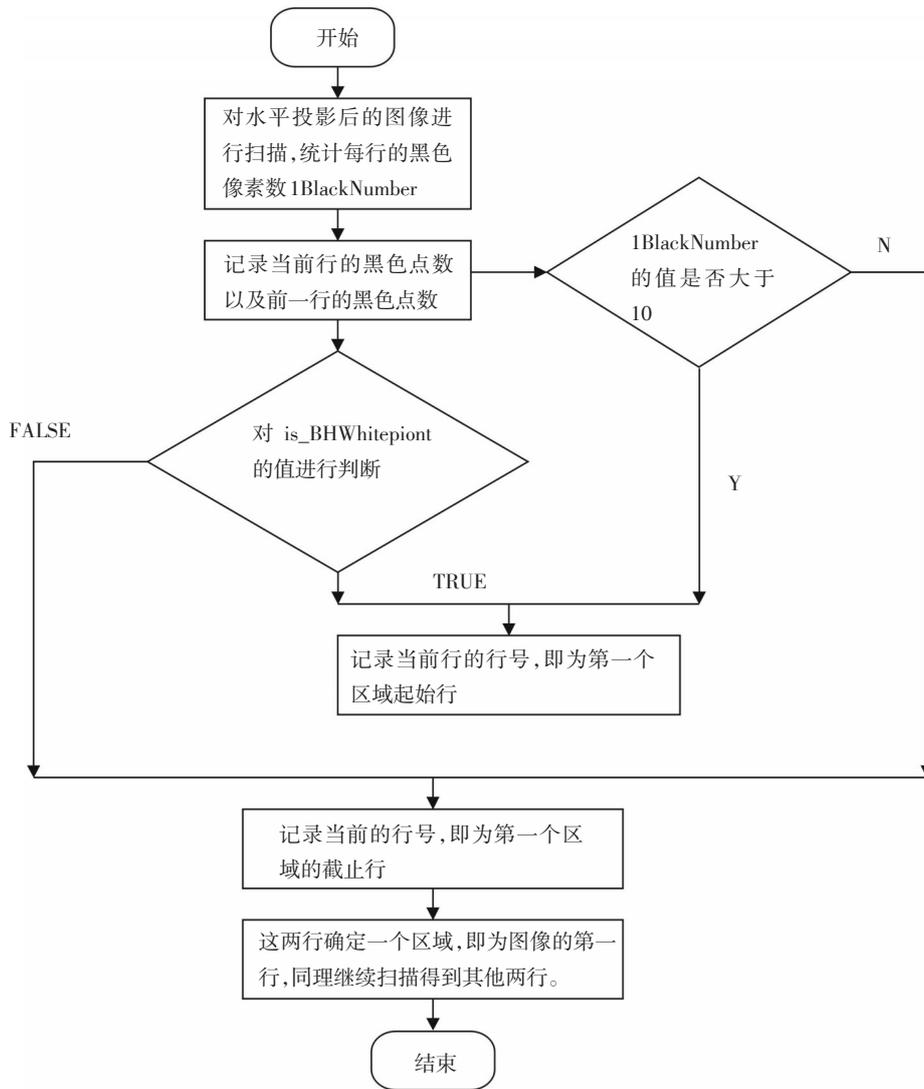


图7 图像精确定位流程图

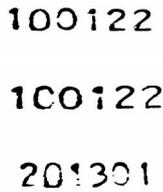


图8 目标区域精确定位

中,药瓶标签定位的准确率达96%,结果表明,此方法能准确地定位出目标字符。但是还是有不足的,这只是针对特定的一种药瓶标签字符的定位。

参考文献

[1] 朱桂方.卡片字符高速识别与校验系统地研究与设计[D].西安:电子科技大学,2010:20-22.

[2] 张惠.同源视频检索与商标货号识别[D].北京:北京邮电大学,2011:34-37.

[3] 常淑英,臧永杰,戴士杰.基于字符间隙和垂直投影特征的铁路货车编码分割算法研究[J].河北工业大学学报,2011,40(2):59-61.

[4] 马亦嘉,葛万成.字符识别中的畸变字符校正算法研究[J].通信技术,2009,42(3):218-220.

[5] Jia Wenjing, Zhang Huai Feng, He Xiangjia, et al. Mean Shift for Accurate License Plate Localization [C]//Proceedings of the 18<sup>th</sup> International IEEE Conference Intelligent Transportation Systems, 2005:566-571.

[6] 文婷,周激流,何坤.基于图像特征的边缘检测[J].计算机工程与应用,2011,47(12):189-192.

[7] 刘翔,周桢,饶昊.基于灰度图像的车牌定位方法研究[J].微计算机信息,2011,27(8):124-126.

[8] 孟晓莉,赵安军,马光思.基于数学形态学的车牌定位研究与实现[J].计算机技术与发展,2010,20(11):84-87.