

一种实用的图像字符定位方法

An Practical Character-Location Algorithm

路嗣恩 朱玉文

(北京理工大学计算机科学与工程系 北京100081)

Abstract Recognizing characters in an image is important for a service robot to understand the scene around it, and this needs locating the characters precisely first. This paper presents a practical character-location algorithm in gray-scale image. Because this approach uses the large size mean filter mask and dynamic threshold, it can precisely get the nature figure of the target area. Hence it not only improves the location precision greatly, but also recognizes the noise effectively.

Keywords Character location, Image segmentation, Service robot

1. 引言

自主服务机器人在完成任务的过程中,需要通过识别指示牌或门牌来识别目的地或实现机器人自身的定位。识别指示牌的首要任务是从获取的场景图像中定位字符。关于字符定位的研究,国内外开展的已经很多,但大部分局限于一些应用范围很窄的场合,如车牌识别等^[5,6]。这些场合一般都有很多的先验知识。如在车牌识别中,车牌的方位、字符的个数、排列关系都可以预先估计出来。本文提出的字符定位方法使用了大均值模板和动态阈值,能够准确地检测出目标区域的自然形状。在整个定位过程中,对结果起主要作用的参数只有一个,因此不需要很多的先验知识,而且避免了人为因素对结果的影响。基于以上的优点,本方法可以用于更复杂的场景,比如图像中字符大小不等,字符的拍摄角度不同,字符的多少不一,字符的排列方式不固定等等。

2. 边缘检测

在一幅含有字符串的灰度图像中,字符串最明显的纹理特征是字符与背景之间明显的明暗交替,即在字符的边缘有着明显的梯度值变化。所以针对这个特征,按下式进行水平方向的差分,对灰度图像进行边缘检测。

$$O(i,j)=\frac{1}{S}\sum_{k=-S}^{k=S}[I(i,j+D+K)-I(i,j-D+K)] \tag{1}$$

式中, $O(i,j)$ 为得到的差分图像, $I(i,j)$ 是输入的灰度图像, S 为求和半径, D 为差分半径。虽然理论上当求和半径 S 、差分半径 D 的取值为字符笔划宽度的一半时,可获得最佳的效果,但对结果影响不大。所以为了减少人为因素的影响,取 $S=D=1$,而且对得到的差分图像不进行阈值化处理。



图1 灰度图

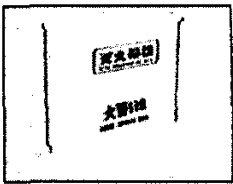


图2 差分图

从图2中可以看出,字符串区域的突出效果很明显:通常是低灰度值最密集的一个区域。当然图像中也存在着一些噪声,这些噪声主要是一些在水平方向上具有很大梯度值的边缘,如物体的垂直轮廓等。

3. 定位算法

为了在差分图中定位字符串区域,通常的做法是先阈值化,再使用数学形态学的方法,先闭后开。但这样需要确定阈值及闭、开中膨胀和腐蚀算子的形状和大小。大量的实验表明,这是引进误差的主要原因,而且这种方法得到的定位图形状和算子的形状密切相关。因此本文的方法针对这两个方面进行改进:1. 避免设置过多的参数,以降低人为因素的影响;2. 获得字符区域的自然形状,这样在定位的同时可初步实现与噪声区域的区分。具体的方法如下:

3.1 预定位

首先,预估图像中字符串区域的面积 S 。因为机器人在移动的过程中只有近距离的明显的标志字符

对它才有意义,而过小或过远的字符应该忽略,所以这个值应为近距离常见字符、标志牌等的平均面积。然后用一个大小为 S 的模板对差分图像进行均值滤波。模板的长宽可灵活选择,因为最后对结果产生影响的主要是它们的乘积,即 S 的大小。

基于速度的原因,在对图像进行滤波时一般采用小模板,如大小为 3×3 、 7×7 像素等。相对于这些模板,本文采用的模板很大,通常长、宽均为几十个像素或更多。采用如此大的模板的原因为:1. 过小的模板不能体现字符区域低灰度值分布比较集中的特点。2. 采用大的模板对图像做滤波肯定会对速度造成影响,但是由于整个定位过程中只做一次这样的滤波,而且采用高效率的卷积算法,因此有效地降低了对速度的影响。所谓高效率的卷积算法是指在卷积的过程中并不重新计算模板中所有像素的均值,而只是计算模板中变化的像素对整个均值带来的影响。这样模板的尺寸虽大,但模板在每次移动的过程中通常变化的只是最右边一列的像素,因此极大地改进了卷积的速度。3. 考虑到服务机器人对于速度的要求,这种采用大模板做滤波的过程是可以接受的。如果要用于其它对速度要求更严格的场合,可采用 Fourier 变换进一步提高速度。



图3 滤波后的差分图

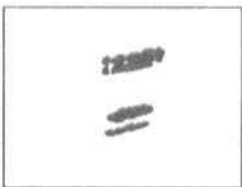


图4 连通域图

从图3中可以看出,滤波对噪声的弱化效果很明显。

然后进行阈值化。阈值的选择条件为:如果用该阈值对图像进行处理,阈值化后图像中最大连通域的面积最接近于 S 。具体算法为:

- ①对滤波后的差分图的灰度级数据进行统计;
- ②寻找一个灰度级 t_1 ,使得 $0 \sim t_1$ 的像素和最接近 S ;
- ③用 t_1 对滤波后的差分图进行阈值化处理,求阈值化后图中最大连通域的面积 s' ;
- ④如果 $s' \geq S$,停止。 t_1 即为所求的阈值。否则,

计算 $\Delta s = S - s'$;

- ⑤寻找一个灰度级 t_2 ,使得 $t_1 \sim t_2$ 的像素和为 Δs 。 $t_1 = t_2$,返回③。

用计算出的阈值对滤波后的差分图进行阈值化,这时图中最大连通域(图4中最上方的连通域)就对应着差分图像中低灰度值最多的区域。从图4中可以看出,这个最大连通域准确地反映了字符区域的形状,因此求它的一个外接多边形并适当扩大,就完成了预定位过程。本文为了不失一般性,只求外接矩形。

在得到外接多边形后必须进行识别,以确定检测出的是所要的字符还是噪声。有时在图像中会存在一些比字符对这种方法更敏感的噪声(这些噪声通常是一些纹理变化更规则或更频繁的图案),因此在这些图像中首先检测出来的肯定是这些噪声。但由于最大连通域比较准确地反映了字符或噪声的形状,因此对得到的外接多边形进行判断,就可以实现字符和噪声的初步区分。如果是噪声,只需简单地将多边形区域从整个图像的差分图中抹去,然后以这时的差分图像为准,重新开始定位过程,就可有效地避免噪声的干扰。

如果连续 N 次检出的都是噪声,这时说明图像中包含字符的概率很低或者图像的拍摄位置不合适。由于机器人处于不断的运动过程,相同的字符会多次出现在不同质量的图像中,因此不需要也没有必要在一幅图像中花费过多的时间,这时可终止整个定位过程。

3.2 精确定位

从整个差分图中分割出外接多边形的差分图,然后将其衰减为一个低精度的二值图像。衰减的比例 R 依据图像的实际尺寸选取(如 4×4 或 8×8 像素),阈值为预定位阶段得到的最大连通域的所有像素在差分图中的灰度均值。根据需要,可对阈值化后的图像做进一步的处理,如去掉像素过少或处于边缘的连通域等。不失一般性,本文没有处理。

最后求阈值化后图像中黑色连通域的外接矩形,就实现了字符的精确定位。

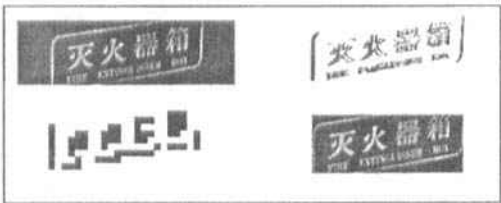


图5 预定位后的灰度图、差分图、阈值化图和精确定位后的灰度图

有时在一幅图像中会有多个字符串,需要将其全部检出。因此在第一次定位完成后,需要按照预定

位阶段中对待噪声的方法一样,将其从整个差分图中抹去,然后开始新的定位过程。具体的定位次数 M 视具体的需要而定,但不宜过多,一般为2、3次,因为通常一幅图像中有用的标志字符不会太多。



图6 图1的第二次定位

4. 实验与结论

本文采集了50幅常见标志字符的图像,图像尺寸大小为 640×480 。最大连通域的面积 S 设为4000个像素,精确定位时衰减比例 R 为 8×8 ,每次预定位过程中噪声允许连续出现的次数 N 为5,每幅图像的定位次数 M 为3。

除了两幅图像由于光线太暗和背景过于复杂的原因未能成功外,其余均取得了很好的定位效果。部分试验结果如下:图中白色矩形分别为第1、2、3次定位得到的结果。

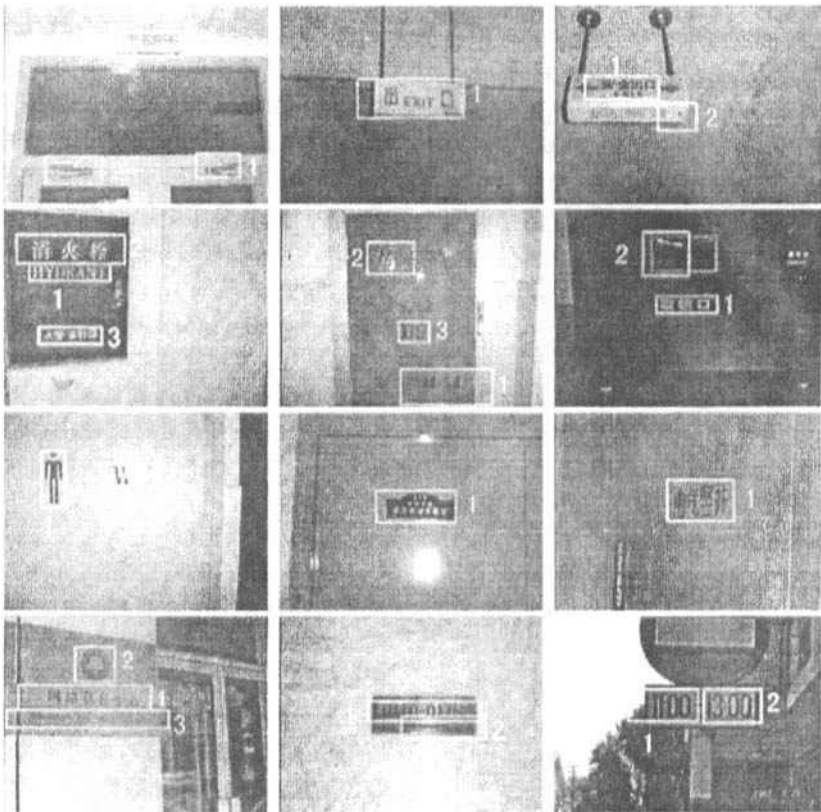


图7 部分实验结果

因此可以看出,本方法确实有效。整个过程唯一一起决定作用的参数是预定位时最大连通域的面积 S 。但它有两个自由度,因此适用的范围很灵活。另外,这个方法可以根据具体的应用场合加很多约束条件,如求最大连通域的外接多边形进行形状的判断等等。本文为了不失一般性,这些工作都没有做。如图7中的一些定位错误的区域,加以一定的约束条件,都可以去除,从而进一步提高最终的定位精度。

参 考 文 献

1 Sonka M, Hlavac V, Boyle R. Image Processing, Analysis, and
• 60 •

Machine Vision. Pacific Grove, CA . 1999
2 章毓晋编著. 图像处理和分折. 北京:清华大学出版社,1999
3 贾云得编著. 机器视觉. 北京:科学出版社,2000
4 崔屹编著. 图像处理与分析:数学形态学方法及应用. 北京:科学出版社,2000
5 曹迪铭,宣国荣. 牌照字符分割中的区域分裂与合并. 计算机工程,2000,26(10):18~20
6 Comelli P, Ferragina P. Optical recognition of motor vehicle license plates[J]. IEEE Transaction on Vehicular Technology, 1995, 11(44): 790~798