

最稳定极值区域与边缘增强的船名定位方法

钱江 张桂荣 姚江 季建中 何平 顾宋华

(江苏省泰州引江河管理处 江苏 泰州 225300)

摘要 船名自动定位技术是船名身份自动识别的第一步,对推动船舶自动化管理有着重要意义。提出一种结合最稳定极值区域和边缘增强的船名定位方法。该方法在MSER(Maximally Stable Extremal Regions)区域检测Canny边缘,通过先验知识分析连通域最终实现定位船名。实验表明,该方法能适应昼间、夜间、强光和弱光多种光照环境,昼间数字定位准确率高达90.9%,昼间船名定位准确率达78.5%。

关键词 最稳定极值区域 边缘增强 船名定位 连通域分析

中图分类号 U614 TP3

文献标识码 A

DOI:10.3969/j.issn.1000-386x.2019.02.047

VESSEL NAME LOCATION METHOD BASED ON MAXIMALLY STABLE EXTREMAL REGIONS AND EDGE ENHANCEMENT

Qian Jiang Zhang Guirong Yao Jiang Ji Jianzhong He Ping Gu Songhua

(Taizhou Yinjiang Canal Administration of Jiangsu Province, Taizhou 225300, Jiangsu, China)

Abstract The automatic location technology of vessel name is the first step of automatic identification of ship name identity, which is of great significance for promoting ship automation management. In this paper, a vessel name location method combining maximally stable extremal regions and edge enhancement was proposed. Canny edge in the MSER region was detected, and the vessel name was located by connected component analysis with prior knowledge. Experiments show that the method can be applied to various illumination environments, such as daytime, night, strong light and weak light. The accuracy of digital positioning in daytime is as high as 90.9%, and the accuracy of ship name positioning in daytime is 78.5%.

Keywords Maximally stable extremal regions (MSER) Edge enhancement Vessel name location Connected component analysis

0 引言

船名是船舶的身份标识,完整、清晰及有效的船名识别有利于船舶安全通航监管。在当前的航务管理中,船名的识别主要依靠工作人员肉眼判断,随着船舶数量的大量增加人工方式极易出现误判问题,利用机器的自动识别方法亟待提出。船名定位是利用计算机视觉实现船舶身份自动识别的第一步,但由于该技术尚处于起步阶段,较少学者对其进行研究,因此在船名定位的研究阶段可参考车牌识别方法。

在车牌识别领域,王永杰等^[1]通过根据车牌字符

的分布信息精确定位车牌,结合边缘密度、模板匹配提出了多信息融合的快速车牌定位方法,但该方法忽视了颜色对于分布信息的影响。钟伟钊等^[2]基于车牌字符颜色与车牌底的搭配固定的纹理特征,提出基于颜色属性与纹理结构相结合的方法。文献[3]直接用每个像素值作为特征,用支持向量机训练分类器,之后在概率图上用均值平移法(MeanShift)来寻找文字区域,该方法在背景简单的图像或视频中效果较好,但难以适应自然场景的复杂背景。魏娜等^[4]在融合滤波的基础上,提出一种基于模糊算子的彩色图像边缘检测方法以克服复杂背景的影响。魏金光等^[5]基于均值差分边缘检测,结合连通分量分析技术和单链聚类算法提

出了一种快速的车牌定位方法。文献[6]结合纹理和连通域的方法,从多尺度的概率图中提取候选区。邓泓等^[7]提出了一种像素连接的车牌定位方法,结合 Sobel 算子与非极大值抑制法进行边缘检测,达到理想检测效果。王晓群等^[8]提出一种基于自适应数学形态学的位方法,在预处理阶段适应图像增强算法进行对比度增强,粗定位阶段采用改进行列自适应数学形态学结构元素的方法进行定位,精定位步骤结合区域标记法和投影法获取精准车牌位置。综上,车牌定位技术现已较为成熟,算法主要基于车牌的颜色、纹理、边缘和连通域等特征。

不同于车牌的形状规矩、字体统一,船名背景复杂、字体大小不统一、船名位置不确定、字符众多,因此车牌识别算法不能很好地应用于船名定位。本文提出了船名定位的新思路,首先使用 Retinex 算法增强图像特征,然后搜索最稳定极值区域缩小目标范围,利用先验知识筛选字符区域,最后合并单个字符得到船名区域。本文算法的提出,在一定程度上弥补了船名定位在计算机视觉方面的空白。

1 船名字符定位方法

1.1 定位流程

船名字符定位流程如图1所示:1) 输入图像并使用 Retinex 算法进行对比度增强预处理;2) 搜索图像的最稳定极值区域(MSER),并将 MSER 区域作为船名字符候选区域;3) 检测图像边缘,得到符合边缘条件的候选区域;4) 利用几何形状、尺寸等先验知识对候选区域进行筛选过滤,得到符合先验知识的字符区域;5) 合并单个字符得到最终船名区域。

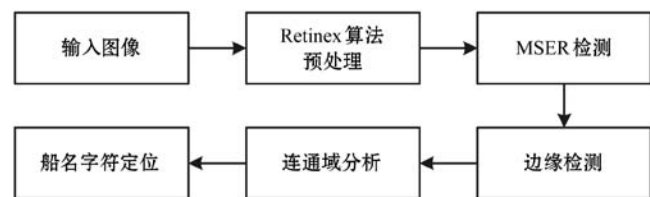


图1 船名字符定位流程图

1.2 图像预处理

由于本文所采用的图像都是在室外获取的,因此光照、天气和环境变化多样,拍摄出来的船舶图像容易到外界因素干扰产生噪声。如若不对获取到的船舶图像进行预处理,会大大影响后续的字符定位精度。本文采用 Retinex 算法进行对船舶图像进行对比度增强处理,Retinex 算法针对雾天、逆光场景能很好地还原图像。

Retinex 理论始于 Land 和 McCann 于 20 世纪 60 年代作出的一系列贡献,其基本思想是人感知到某点的颜色和亮度并不仅仅取决于该点进入人眼的绝对光线,还和其周围的颜色和亮度有关。Retinex 理论认为原始图像 $S(x,y)$ 可以看成是光照图像 $L(x,y)$ 和反射率图像 $R(x,y)$ 的乘积,即 $S(x,y) = R(x,y) \times L(x,y)$ 。基于 Retinex 的图像增强的核心就是从原始图像 $S(x,y)$ 中估计出光照 $L(x,y)$,从而分解出 $R(x,y)$,消除光照不均的影响。首先将图像变换到对数域,图像变换过程如式(1)。Retinex 理论假设图像的噪声来源于图像各个位置反射率不同,因此去除反射图像的噪声便可还原图像,反射分量噪声可通过归一化去除,归一化过程如式(2)。将去噪后的反射分量求指数,反变换到实数域即得到增强后的图像。

$$\begin{cases} s(x,y) = \log S(x,y) \\ l(x,y) = \log L(x,y) \\ r(x,y) = \log R(x,y) \\ s = r + l \end{cases} \quad (1)$$

$$r'(x,y) = \frac{r(x,y)}{\max(r(x,y))} \quad (2)$$

图2为船舶图像预处理前后对比图,(a)为原图像,(b)为使用 Retinex 算法预处理后的图像。从图中可以看出,预处理后图像对比度得到提高,船名特征得到增强。



(a) 预处理前



(b) 预处理后

图2 船舶图像预处理前后对比图

1.3 检测 MSER 区域

自然场景中的文字检测主流方法有两大类:基于纹理特征的方法和基于连通域的方法。基于连通域的方法首先提取候选区(颜色聚类和极值区域提取等方法),利用一些人为设定的规则或者自动分类器过滤一些非文字区域。由于连通域数量相对较小,此类方

法相对高效,对旋转、尺度变化和字体差异不敏感,是自然场景文字识别的主流方法。船名文字一般是一个可靠的连通域,且连通域内像素差异较小,因此本文采用基于连通域的方法进行文字检测,提取最稳定极值区域作为文字候选区域。

最稳定极值区域 MSER 算法对图像具有较强的仿射变换不变性。最稳定极值区域的提取方法如下:对一幅灰度图像(灰度值为 0~255)取阈值进行二值化处理,阈值从 0 到 255 依次递增,在得到的所有二值图像中,图像中的某些连通区域变化很小甚至没有变化,则该区域就被称为最大稳定极值区域,连通区域变化的数学定义为 $q(i) = |Q_{i+\Delta} - Q_{i-\Delta}| / |Q_i|$ 。

图 3(a)为船舶图像的 MSER 区域示意图,图中每一种灰度代表一个 MSER 区域,图 3(b)为 MSER 区域二值图,图中白色区域为 MSER 区域。从图中可以看出,船名区域都为最稳定极值区域,此外还有一些细长条船身边缘及部分船身区域也被检测为最稳定极值区域。



(a) MSER 区域图



(b) MSER 区域二值图

图 3 MSER 区域示意图

1.4 边缘检测

图像的边缘指的是某像素点周围像素值不连续且像素灰度有阶跃性变化,是图像局部强度变化最明显的位置,表征一个特征区域的终结和另一特征区域的开始。船名中无论是中文字符还是数字字符,都具有明显的边缘特征,因此边缘可作为特征把船名与背景区分开来。

Canny 算子检测边缘法与一般边缘算法相比,采用非极大抑制法去除弱边缘,且采用双阈值进行边缘连接,既不丢失强边缘又去除弱边缘。因此本文选用 Canny 算子检测边缘。Canny 边缘计算过程如下:首先

对图像进行高斯模糊;其次计算垂直方向和水平方向梯度,根据梯度计算边缘幅值和角度,方向梯度计算如式(3),边缘幅值和角度计算如式(4);之后采用非极大抑制法舍弃弱边缘,比较相邻像素边缘点值,若中心像素值角度区域小于相邻像素值角度区域,则舍弃该点边缘;最后设置双阈值,对大于高阈值的边缘点直接保留,小于低阈值的边缘点直接舍弃,介于两阈值之间的若能连接到边缘则保留否则舍弃。

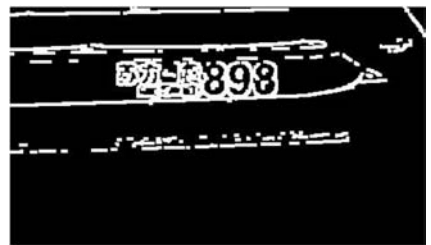
$$\begin{cases} G_x(x,y) \approx [S(x,y+1) - S(x,y) + S(x+1,y+1) - S(x+1,y)]/2 \\ G_y(x,y) \approx [S(x,y) - S(x+1,y) + S(x,y+1) - S(x+1,y+1)]/2 \end{cases} \quad (3)$$

$$\begin{cases} G(x,y) = \sqrt{G_x^2(x,y) + G_y^2(x,y)} \\ \theta(x,y) = \arctan(G_y(x,y)/G_x(x,y)) \end{cases} \quad (4)$$

虽然 Canny 算子检测的边缘为强边缘,但并非所有边缘都是本算法关注的边缘。船名文字既有最稳定极值的特征,又有边缘特征,因此在最稳定极值区域内的边缘才是船名定位的重点关注区域。首先取原始 Canny 边缘与 MSER 区域交集,得到 MSER 区域内边缘,之后使用边缘增长法增强边缘。所谓边缘增长就是设定一个边缘增长值,边缘沿着梯度方向的垂直方向增长。图 4(a)为 Canny 边缘检测结果,图中字符边缘明显,但也检测到船身边缘和水面等边缘,图 4(b)为结合 MSER 区域后的边缘增长图,从图中可以看出边缘变“粗”且过滤了一些背景边缘。



(a) Canny 边缘图



(b) 边缘增长图

图 4 Canny 边缘示意图

图 4(b)的边缘增长图展示了文字定位中的重要边缘信息,将 MSER 区域与边缘增长图取交集,即可得到有强边缘的 MSER 区域。图 5 为 MSER 二值图在边缘增长前后的对比图,(a)为边缘增长前的二值图,图中的字符连通域与船边缘连通域存在粘连,(b)为结

合边缘增长后的二值图,图中连通域相对独立但却又不丢失字符特征。



(a) 边缘增长前的二值图



(b) 连通域分析后的二值图

图 5 边缘增长前后的二值图对比

1.5 连通域分析

基于 MSER 和边缘检测后,基本得到字符候选区域,但候选区域中存在大量非字符区域,因此去除非字符区域是本节重点考虑问题。由于字符区域都具有一定的几何特性,如大小、面积和宽高比等。因此通过先验知识合理设定阈值可以达到对非字符区域去除的目的。

本文采用的先验知识主要包括:区域面积、区域矩形度、外接矩形的宽高比和外接矩形高度等。区域是一个点集,包含该区域内每个像素点坐标,外接矩形就是对该区域取最小外接矩形;区域面积即为该区域内像素点个数;区域矩形度即为区域面积与外接矩形面积之比,该值越接近 1 则该区域外形越接近矩形;外接矩形宽高比即外接矩形宽度与高度之比;对以上先验知识设定阈值,可以得到符合先验知识的文字候选区域。

图 6 为连通域分析前后的二值图对比,从图中可看出,细长条的边缘连通域和不规则的背景连通域都得到很好地去除,最后仅留下字符区域,至此完成船名字符定位。



(a) 连通域分析前的二值图



(b) 连通域分析后的二值图

图 6 连通域分析前后的二值图对比

2 实验及分析

本文算法在 Visual Studio 2015 环境中基于 C++ 开发语言及 OpenCV3.1.0 开源机器视觉库进行。

船名定位测试实验中,测试图像共计 272 张,其中昼间图像 158 张,夜间图像 114 张,共计 2 000 个船名字符,其中汉字字符 749 个,数字 1 251 个。船名字符定位测试结果如表 1 所示,昼间数字定位精度高达 90.9%,汉字定位精度 60% 左右,汉字 + 数字昼间精度为 80.1%。字符定位结果显示数字定位精度远远高于汉字字符,这是因为数字个数种类较少,结构简单,而船名中常用汉字种类达上百种且汉字结构复杂,几何形状各异,导致在候选字符筛选过程中错误去除字符,最后导致汉字定位不准确。

表 1 船名字符定位测试结果

	%		
	汉字字符	数字字符	汉字 + 数字
昼间	62.2	90.9	80.1
夜间	59.2	84.8	75.3

船名定位是在船名字符定位基础上进行的,将相邻的字符区域组合成最终船名位置。船名定位测试结果如表 2 所示,表中定位完全正确表示最终定位的船名区域包含船名所有字符,定位部分字符表示最终定位的船名区域仅包含船名部分字符,定位错误表示最终定位的船名区域并非实际船名区域。从表中可看出,昼间定位准确率高达 78.5%,定位错误率仅 3.1%,夜间定位准确率达 72.8%,定位错误率仅 6.1%。由于汉字字符定位准确率不高,造成船名中很多汉字漏定位,因此 20% 的船名仅能正确定位部分字符。部分图像船名定位结果如图 7 所示。

表 2 船名定位测试结果

	%		
	定位完全正确	定位部分字符	定位错误
昼间	78.5	18.4	3.1
夜间	72.8	21.1	6.1



图7 部分图像定位结果图

3 结 语

本文提出一种结合最稳定极值区域和边缘增强的船名定位方法。该方法首先采用 Retinex 算法增强待处理图像,然后检测 MSER 区域,之后采用 Canny 算子检测边缘,得到边缘增长后的文字候选区域,最后分析连通域属性,设定先验知识阈值,得到最终文字区域。该算法通过实验和分析,验证了该方法在昼间和夜间均能实现船名精准定位。实验结果表明,昼间数字定位准确率高达 90.9%,昼间船名定位准确率达 78.5%。

本文对于船名字符定位算法的研究在一定程度上弥补了船名定位在计算机视觉方面的空白,另一方面该研究也推进了基于计算机视觉的船舶身份自动识别进程。本文所提出的算法仍有需改进之处,如针对船名中的汉字字符的定位准确率相对较低,提高字符定位准确率尤其是中文字符准确率将是今后研究需要突破的难点。

参 考 文 献

[1] 王永杰,裴明涛,贾云得. 多信息融合的快速车牌定位[J]. 中国图象图形学报, 2014, 19(3): 471-475.

[2] 钟伟钊,杜志发,徐小红,等. 基于字符边缘点提取的车牌定位方法[J]. 计算机工程与设计, 2017, 38(3): 795-800.

[3] Kim K I, Jung K, Kim J H. Texture-based approach for text detection in images using support vector machines and continuously adaptive mean shift algorithm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(12):1631-1639.

[4] 魏娜,王振臣,张聪,等. 一种新型车牌定位算法的研究[J]. 激光与红外, 2012, 42(8): 936-939.

[5] 魏金光,侯杰,张帆. 基于均值差分边缘检测和单链聚类的实时车牌定位算法[J]. 电视技术, 2017, 41(9): 176-181.

[6] Pan Y F, Hou X, Liu C L. A hybrid approach to detect and localize texts in natural scene images[J]. IEEE Transactions on Image Processing, 2011, 20(3): 800-813.

[7] 邓泓,李水泉,彭莹琼. 像素连接车牌定位技术的研究[J]. 江西师范大学学报(自然科学版), 2017(6): 651-655.

[8] 王晓群,刘宏志. 基于自适应数学形态学的车牌定位研究[J]. 图学学报, 2017, 38(6): 843-850.

(上接第 244 页)

[4] Sariyanidi E, Gunes H, Cavallaro A. Automatic analysis of facial affect: A survey of registration, representation, and recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(6): 1113-1133.

[5] Ding C, Choi J, Tao D, et al. Multi-directional multi-level dual-cross patterns for robust face recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2016, 38(3): 518-531.

[6] Lu J, Liong V E, Zhou X, et al. Learning compact binary face descriptor for face recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(10): 2041-2056.

[7] Majumder A, Behera L, Subramanian V K. Emotion recognition from geometric facial features using self-organizing map[J]. Pattern Recognition, 2014, 47(3): 1282-1293.

[8] Ghimire D, Lee J, Li Z N, et al. Recognition of facial expressions based on salient geometric features and support vector machines[J]. Multimedia Tools and Applications, 2017, 76(6): 7921-7946.

[9] 李忠健,杨宜民,梅平. 基于改进的隐马尔科夫模型的人脸识别方法[J]. 工业控制计算机, 2014, 27(11):29-31.

[10] 朱黎,胡涛,罗锋. 基于自适应 LBP 人脸识别的身份验证[J]. 计算机工程与科学, 2014, 36(11):2217-2222.

[11] 张辉,刘新. 基于改进 LBP 的单样本人脸识别算法[J]. 计算机应用与软件, 2017,34(12): 220-223, 282.

[12] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. Las Vegas, NV, United States: IEEE, 2016: 770-778.

[13] Wen Y, Zhang K, Li Z, et al. A discriminative feature learning approach for deep face recognition[C]//European Conference on Computer Vision. Springer, Cham, 2016: 499-515.

[14] Ding C, Tao D. Robust face recognition via multimodal deep face representation[J]. IEEE Transactions on Multimedia, 2015, 17(11): 2049-2058.