

文章编号: 2095-2163(2020)09-0095-06

中图分类号: U216

文献标志码: A

基于显著性检测的接触网标识牌识别研究

蒋丽洁, 柴晓冬, 李立明, 郑树彬

(上海工程技术大学 城市轨道交通学院, 上海 201620)

摘要: 实时、准确的轨道定位是轨道检测及维护的基本条件, 轨道定位的常用方法为接触网标识牌定位方法, 存在识别率低、效率低等问题, 本文提出基于显著性检测的接触网标识牌识别算法弥补这些缺点。首先, 利用马尔科夫链显著性检测算法定位接触网立柱, 获取包含标识牌的感兴趣区域; 其次, 采用形态学与垂直投影的方法对标识牌单个数字进行分割; 最后, 利用多层感知机神经网络(Multi-Layer Perceptron, MLP)分类器完成标识牌数字识别。通过实验验证了该图像处理算法具有极高的可行性, 接触网立柱定位及标识牌识别的成功率可达93%以上, 为轨道定位提供新思路。

关键词: 轨道定位; 显著性检测; 马尔科夫链; 神经网络; 标识牌识别

Research on recognition of contact signs based on significance detection

JIANG Lijie, CHAI Xiaodong, LI Liming, ZHENG Shubin

(College of Urban Rail Transit Institute, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] Real-time and accurate track positioning is the basic condition for track inspection and maintenance. The common method for track zone positioning is the catenary sign positioning method, but there is a problem of low recognition rate. Therefore, a catenary sign recognition algorithm based on saliency detection is proposed to solve this shortcoming. First, the Markov chain saliency detection algorithm is used to locate the catenary column and obtain the area of interest that contains the sign; then the morphology and vertical projection methods is used to segment the single number of the sign; finally the multilayer perceptron neural network (Multi-Layer Perceptron (MLP) classifier is used to complete the number recognition of the sign. Experiments have verified that the image processing algorithm is extremely feasible, and the success rate of catenary column positioning and sign recognition can reach more than 93%, providing new ideas for track zone positioning.

[Key words] track positioning; saliency detection; Markov chain; neural network; sign recognition

0 引言

随着科技的进步, 提出将图像处理与机械化铁路巡检结合起来, 通过研究接触网立柱信息, 完成对接触网立柱或标识牌数字信息的识别, 最终达到轨道区间定位的目的。武翔宇等应用 Hough 直线检测的接触网支柱定位方法, 利用 HOG 特征和 SVM 分类器结合的方式成功完成支柱号码牌识别^[1], 其缺点是需要大量的正负样本来收集特征; 徐可佳将 4C 系统所拍摄接触网支柱号作为研究对象, 设计了基于 Faster R-CNN 的深度学习算法^[2], 但图像分辨率不高; 刘华春等采用视觉显著性检测算法定位车牌^[3]; 陈曦等利用显著性检测算法自动过滤不感兴趣区域的干扰, 高效的计算方式减少计算资源的浪费^[4]; 姜博文等将马尔可夫链与先验知识融合, 摆脱前景区域的影响^[5]。

对图像简单线性迭代 (simple linear iterative cluster, SLIC) 超像素分割后, 本文提出利用吸收马尔可夫链的显著性检测, 对接触网立柱区域初步定位, 将接触网区域进行前景分割后, 排除一切背景的干扰, 利用基于 MLP 分类器的算法对接触网标识牌信息进行识别, 为地铁线路的轨道区间定位提供新思路, 识别过程算法如图 1 所示。

1 基于显著性检测的接触网立柱定位算法

为了构造吸收马尔可夫链的显著性检测模型, 首先利用强健的背景先验的算法来选择吸收节点, 以吸收马尔可夫链中的转移节点到吸收节点的被吸收时间作为转移节点的显著值。

1.1 强健背景先验的吸收节点选择

边界连通性是衡量图像中区域连接到边界程度的标准。由于 SLIC 超像素分割是在 CIELAB 颜色

基金项目: 国家自然科学基金(51975347); 上海市科委重点支撑项目(18030501300)。

作者简介: 蒋丽洁(1995-), 女, 硕士研究生, 主要研究方向: 图像处理; 柴晓冬(1962-), 男, 博士, 教授, 主要研究方向: 通信与信息工程、故障诊断研究; 李立明(1982-), 男, 硕士, 高级实验师, 主要研究方向: 图像处理、计算机视觉; 郑树彬(1979-), 男, 博士, 教授, 主要研究方向: 轨道交通车辆、轨道设备状态检测技术与理论。

通讯作者: 柴晓冬 Email: cxdyj@163.com

收稿日期: 2020-05-20

空间内进行的,节点间的权重 $d_{app}(p,q)$ 以欧氏距离算法来计算。相应的,根据节点间沿最短路径累积边界权重来计算两节点的测地线距离公式(1):

$$d_{geo}(p,q) = \min_{p_1=p, p_2, \dots, p_n=q} \sum_{i=1}^{n-1} d_{app}(p_i, p_{i+1}). \quad (1)$$

其中, n 代表最短路径经过的节点个数,定义该式中 $d_{geo}(p,p) = 0$ 。

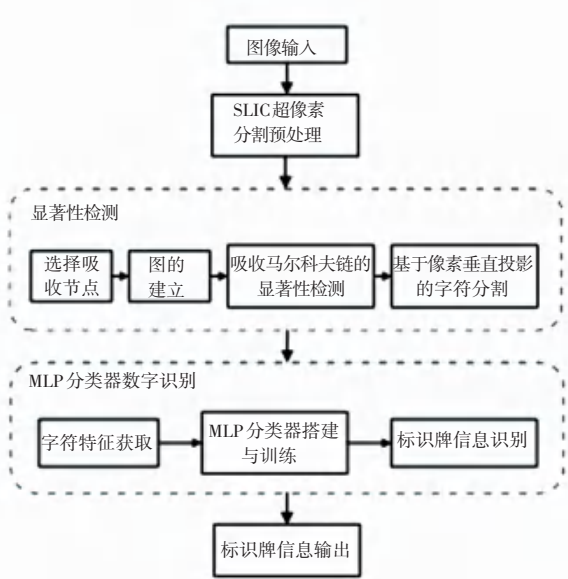


图1 识别过程算法

Fig. 1 Algorithm recognition process

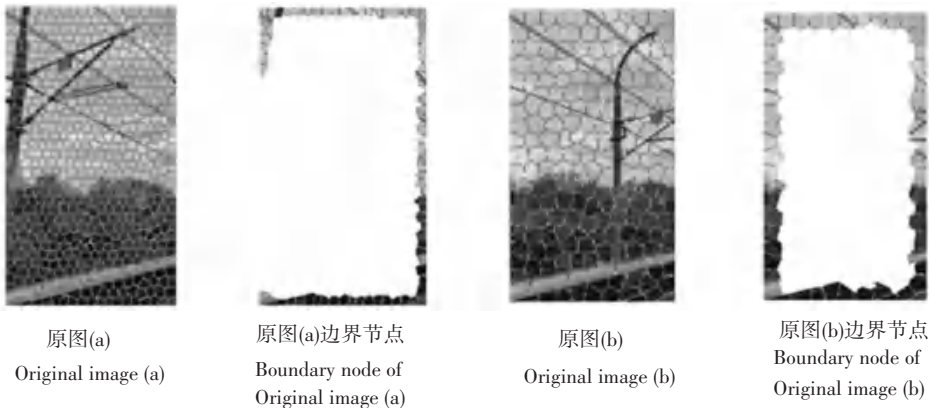


图2 基于强健的背景先验的边界选择

Fig. 2 Boundary selection based on a strong background prior

1.2 吸收马尔可夫链的显著性检测

设吸收马尔可夫链的状态集合 $S = \{s_1, s_2, \dots, s_n\}$ 中,吸收状态与转移状态的数目分别为 k 和 m ,将转移状态放在吸收状态之前,则矩阵 P 的标准形式如式(6)所示:

$$P = \begin{pmatrix} Q & R \\ \mathbf{0} & I \end{pmatrix} \quad (6)$$

其中,转移状态间、转移状态与吸收状态间的概率转移矩阵分别为 $Q \in [0,1]_{m \times m}$, $R \in [0,1]_{m \times k}$ 。该链的基本矩阵 N ,计算公式如式(7)所示:

$$N = (I - Q)^{-1} = I + Q + Q^2 + \dots \quad (7)$$

边界连通性定义公式为式(2)、式(3), N 是扩展的超像素数目,图像边界的像素块 $\delta(\cdot) = 1, \sigma_{cl} = 10$ 。

$$BndCon(R) = \frac{\sum_{i=1}^N S(p, p_i) \times \delta(p_i \in Bnd)}{\sqrt{\sum_{i=1}^N S(p, p_i)}}. \quad (2)$$

$$\sum_{i=1}^N S(p, p_i) = \sum_{i=1}^N \exp\left(-\frac{d_{geo}^2(p, p_i)}{2\sigma_{cl}^2}\right). \quad (3)$$

进一步扩展边界连通性的背景概率 w_i^{bg} , 其计算公式被定义为式(4):

$$w_i^{bg} = 1 - \exp\left(-\frac{BndCon^2(p_i)}{2\sigma_{BndCon}^2}\right). \quad (4)$$

判断边界中像素节点是否为背景的不等式(5),记宽度为 d 个像素的边界中超像素的个数为 K 。本文取 $\beta = 0.8$ 。

$$\frac{w_i^{bg} - \left(\sum_{i=1}^k w_i^{bg} / K\right)}{\sum_{i=1}^k w_i^{bg} / K} < \beta. \quad (5)$$

根据边界节点选择方式,对超像素数目为 625 和 200 图像做边界选择,选出的背景集合,如图 2 所示。在原图(a)中,接触网立柱区域处在图像边界,通过该方法很好的把立柱筛选出来。

元素 n_{ij} ($n_{ij} \in N$) 描述的是起始状态为 s_i , 被吸收之前经历过转移状态 s_j 的期望时间, 行和 $\sum_j n_{ij}$ 揭示了被吸收之前, 从状态 s_i 出发经历所有转移状态的期望时间。每个转移状态的期望吸收时间组成向量为式(8):

$$y = N \times c. \quad (8)$$

其中, c 为 m 维的单位列向量。

为了确保边界节点的显著值能够计算, 复制图像的边界节点作为虚拟吸收节点, 图3为复制边界后建立的包含虚拟吸收节点的图模型 $G = (V, E)$, 无向边构成集合 E , 超像素作为节点 V 。



图3 吸收马尔可夫链的虚拟吸收节点

Fig. 3 Virtual absorbing node absorbing Markov chain

边的权重 $w_{i,j}$ 的计算公式为式(9):

$$w_{i,j} = e^{-\frac{\|c_i - c_j\|}{2\sigma^2}}. \quad (9)$$

其中 c_i, c_j 分别表示节点在 CIELAB 颜色空间的均值, 权重的控制力度由常数 $\sigma = 0.05$ 来把握。

重新排列新的图模型对应的关联矩阵 $A = (a_{ij})$, 使得在矩阵 A 中 t 个转移节点在前, r 个吸收节点在后, a_{ij} 为新图模型中相邻节点之间的相关性, i 相邻节点集合 $N(i)$ 。

$$a_{ij} = \begin{cases} w_{i,j} & j \in N(i) \\ 1 & i = j \\ 0 & \text{others} \end{cases}. \quad (10)$$

对角矩阵 $D = \text{diag}(\sum_j a_{ij})$ 表示连接到每个节点的权重之和。

在该图模型下吸收马尔可夫链的概率转移矩阵 P 的计算公式为式(11):

$$P = D^{-1} \times A. \quad (11)$$

已知转移节点与吸收节点的个数分别为 t 和 r , 依据公式(8)可以提取矩阵 Q , 进一步得到基本矩阵 $N = (I - Q)^{-1}$, 利用公式(10)计算每个节点被不同吸收节点的吸收时间 y , 最后对时间 y 进行归一化处理得到每个节点的显著值, 式(12):

$$S(i) = \overline{y(i)} \quad (i = 1, 2, \dots, t). \quad (12)$$

实验结果显示排除包括树木在内的背景的干扰, 完成立柱定位检测, 如图4所示。

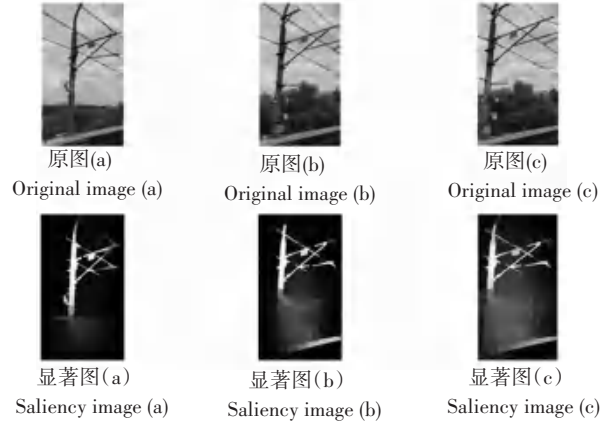


图4 显著性检测结果

Fig. 4 Significance test results

对显著图通过立柱标识牌的几何特征包括区域面积、周长、矩形度和致密程度在内的特征值, 进一步分割得到标识牌区域, 如图5所示。

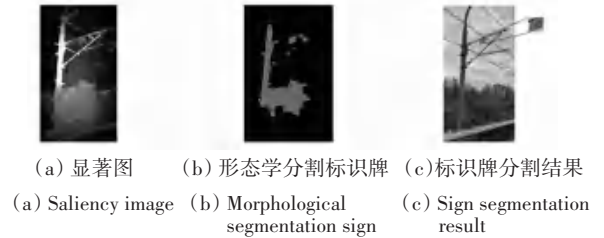


图5 标识牌区域分割

Fig. 5 Segmentation of the identification plate area

1.3 基于像素垂直投影的字符分割

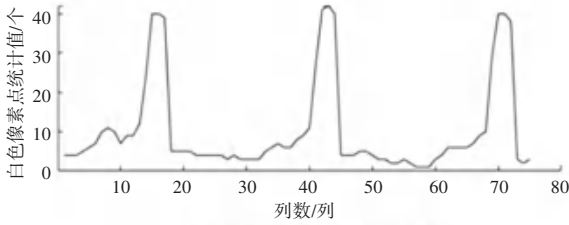
获取标识牌区域之后, 需要对单个字符进行分割, 像素投影法算法简单、速率快, 能够达到本文快速识别算法的要求。首先, 二值化标识牌区域, 使得数字区域转化为白色像素点, 其余区域为黑色。垂直投影是统计二值图中每列的白色像素, 将投影曲线用函数的形式表示出来, 函数图像会出现明显的波峰波谷, 波谷代表白色像素为0的某列, 即数字的分割点。垂直投影得到的函数图像如图6所示。

从图6可以看出, 标识牌中3个数字区域对应像素投影函数图中3个连续的波形区域, 数字间隔对应像素投影函数图的波谷。以垂直投影图中波谷对应的列数对标识牌进行分割, 分割结果如图7所示。

2 基于MLP分类器的标识牌字符识别

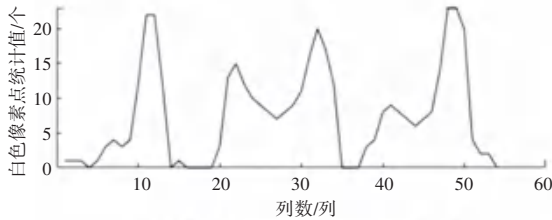
为了规避标识牌区域小、像素值不高的问题, 通过多层感知机神经网络 (Multi-Layer Perceptron,

MLP)分类器,将每个标识牌区域的纹理特征作为训练样本,合理的选择神经网络模型的参数,最终输出识别结果。



(a) 标识牌“111”

(a) Identification plate "111"



(b) 标识牌“194”

(b) Identification plate "194"

图6 白色像素垂直投影及二值图

Fig. 6 Vertical projection and binary image of white pixels



图7 数字分割结果

Fig. 7 Digital segmentation results

2.1 标识牌纹理特征提取

提取图像局部纹理特征,可以用局部二值模式(Local Binary Pattern, LBP)描述子来描述。LBP模式的基本操作单元是小邻域内的图像纹理,具有计算量小、旋转不变性和灰度不变性等优点^[6]。

利用旋转不变模式 $LBP_{(P,R)}^r$ 与统一模式 $LBP_{(P,R)}^{u_2}$ 相结合的旋转不变统一模式来获得标识牌的纹理特征,式(13)和式(14):

$$LBP_{(P,R)}^r = \min_{0 \leq i \leq P-1} \{ROR(LBP_{(P,R)}, i)\}. \quad (13)$$

$$LBP_{(P,R)}^{u_2} = \sum_{i=0}^{P-1} |s(g_{(i+1) \bmod P} - g_c) - s(g_i - g_c)|. \quad (14)$$

旋转不变统一模式的计算公式为式(15):

$$LBP_{(P,R)}^{r, u_2} = \begin{cases} \sum_{i=0}^{P-1} s(g_i - g_c) & LBP_{(P,R)}^{u_2} \leq 2, \\ P + 1 & \text{others.} \end{cases} \quad (15)$$

将提取得到的特征值转化为 LBP 直方图,数字“1”、“9”、“4”的标识牌样本图的 LBP 图谱如图8所示,降维处理之后的特征直方图维数为32。

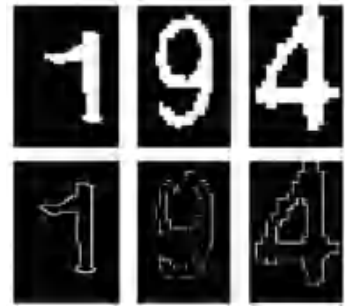


图8 LBP 特征图谱

Fig. 8 LBP feature map

2.2 字符识别算法

多层感知器(MLP)层间连接方式是全连接。输入特征向量 X 、层与层之间的权重 W 与偏置 b 、激励函数是 MLP 神经网络最主要的参数,通过公式(16)可了解基本 MLP 模型的数学关系, $Softmax(\cdot)$ 代表输出层的逻辑回归函数。

$$y(x) = Softmax(b_2 + W_2(f(W_1 X + b_1))). \quad (16)$$

在 MLP 神经网络模型中,为了改变线性规则需要添加激活函数(Activation Function),使得输出结果与特征存在非线性相关,本文采用 Sigmoid (S型)激活函数。该激活函数与其导函数可表示为式(17):

$$\varphi(x) = \frac{1}{1 + e^{-x}}, \varphi'(x) = \frac{e^{-x}}{(1 + e^{-x})^2}. \quad (17)$$

MLP 神经网络的训练过程就是调整边界权重与偏置参数的过程,最终结果是使得总均方误差尽可能小。本文所用 BP 算法是最经典的训练算法,以输入层到隐含层为代表,参数更新如下:

将输入层、隐含层的节点分别定义为 x_i, h_j , 层间节点的权值定义为 w_{x_i, h_j} , 层与层之间权重更新如(18):

$$w_{x_i, h_j} = w_{x_i, h_j} - \eta_{x_i, h_j} \frac{\partial E}{\partial w_{x_i, h_j}} \quad (18)$$

其中, 隐藏层的学习率为 η_{x_i, h_j} 。

将隐含层偏置定义为 b_h , 则输入到隐含层的偏置的更新为公式(19):

$$b_h = b_h - \eta_h \frac{\partial E}{\partial b_h} \quad (19)$$

其中, 该层偏置的学习率为 η_h 。

本文采用均方误差 (Mean Squared Error, MSE) 作为误差估计的方法, 用来测试网络的训练性能, MSE 的数学计算公式如(20)所示:

$$MSE = \frac{1}{2} (h(x) - out_{oi})^2, \quad (20)$$

其中, $h(x)$ 、 out_{oi} 分别为输出层的输出数据与期望输出数据, 则 N 类输出的总均方误差公式为式(21):

$$E = \sum_{i=1}^N MSE. \quad (21)$$

3 实验结果及分析

本实验研究对象为一段基于接触网立柱的短视频, 将数字图像尺寸统一调整到 50×120 的像素大小, 经过对各类数字的纹理特征提取, 输入层特征向量维度为 10×32 , 输出层即为“0-9”十类。

神经网络模型搭建完成后, 对标识牌区域测试集和样本集进行特征提取, 训练得到的特征, 其训练误差曲线如图 9 所示。将误差目标设定为 0.001, 经过 168 次迭代后, 误差小于设置的目标值, MLP 分类器训练截止, 输入测试集对标识牌信息识别。

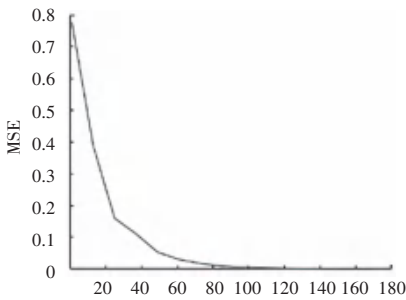


图 9 误差曲线
Fig. 9 Error curve

为了解决图像数据集过少的问题, 标识牌区域旋转合并到数据集中, 最终得到 200 张标识牌图像, 分割数字字符后, 基于 MLP 神经网络的标识牌识别最终结果如图 10 所示。在标识牌区域有接触网线的干扰下, 仍能准确识别数字信息。

本章算法从基于视觉显著性的研究出发, 在将接触网立柱区域前景分割后, 根据标识牌区域的特

征对其进行下一步的分割, 利用像素投影完成单个字符的提取, 最终通过基于 MLP 分类器算法完成标识牌信息识别, 如图 11 所示。该算法最终的识别准确率达到了 93.5%。

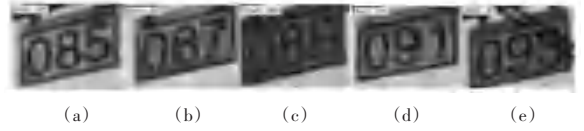


图 10 标识牌识别结果

Fig. 10 Identification result of the sign



(a) 接触网立柱显著图

(a) Significant diagram of catenary column



(b) 标识牌信息识别

(b) Sign information recognition

图 11 接触网立柱标识牌识别结果

Fig. 11 Identification results of the identification plate of the catenary column

4 结束语

由于立柱区域与背景树木颜色及灰度值太过相近, 传统分割效果并不明显。受视觉显著性检测的启发, 本文提出以强健的背景先验与吸收马尔可夫链结合的方式提取立柱显著图, 为了提高检测效率, 以 SLIC 超像素算法对图像进行预处理; 利用标识牌的底层特征对标识牌区域精定位。该算法的提出很好的定位了接触网立柱及标识牌区域。搭建 MLP 神经网络分类器, 输入标识牌纹理特征并训练该模型, 使得该神经网络模型参数达到最优, 对标识牌的识别准确率达到 93.5%。因此该算法基本达到了对接触网标识牌识别定位的要求, 为智能化轨道巡检提供了新思路。
(下转第 108 页)