

DOI:10.3969/j.issn.1001-4551.2018.01.019

基于抓拍数据时空判别的车牌检错纠错研究^{*}

刘晴辉¹, 徐 巍², 董红召^{1*}

(1. 浙江工业大学 智能交通联合研究所, 浙江 杭州 310014; 2. 杭州市交警支队, 浙江 杭州 310014)

摘要:针对车牌识别设备应用场合背景复杂, 车牌识别率易受到光线、车牌污损与遮挡、车辆行驶速度等因素的影响, 车辆抓拍识别数据中存在错误车牌号码这一问题, 分析了车辆抓拍识别系统提高识别准确率的方法, 提出了一种以车牌识别设备采集到的源数据为基础进行车牌号码检错与纠错的方法。首先基于抓拍数据计算路段平均行驶速度, 通过对抓拍数据时空判别进行了车牌检错, 然后对错误车牌进行了字符替换并作时空判别进行了车牌纠错。最后在杭州市主城区大量的抓拍数据与机动车登记信息数据中进行了实验, 并以车辆抓拍识别设备采集的视频图像数据进行了校验。研究结果表明: 该方法可以较好地检测出错误识别的车牌号码, 进一步提高车辆抓拍数据的准确率。

关键词:智能交通; 车牌识别技术; 字符替换; 行程时间

中图分类号: TP301; U491

文献标志码: A

文章编号: 1001-4551(2018)01-0101-05

Error detection and correction based on time-space judgement of data captured by VLPR system

LIU Qing-hui¹, XU Wei², DONG Hong-zhao¹

(1. ITS Joint Research Institute Zhejiang University of Technology, Hangzhou 310014, China;

2. Traffic Police Detachment of Hangzhou, Hangzhou 310014, China)

Abstract: Aiming at the problem that complex application background of license plate recognition equipment, the recognition rate is vulnerable to light, license plate fouling and blocking, vehicle speed and other factors, there are wrong license plates information in the collected data, the method of improving the recognition accuracy of VLPR was analyzed. Based on the data collected from license plate recognition devices, a method of error detection and correction was proposed for license plate recognition. First the average travel speed of road based on the captured data was calculated and wrong license plate by time-space judgement was detected. Then character of wrong license plate was replaced and the wrong license plate by time-space judgement was corrected. Finally this method was validated in the big data collected from license plate recognition devices of Hangzhou and motor vehicle registration information data of Hangzhou. Besides, the experimental results were verified by video image data collected from license plate recognition devices. The results indicate that this method can detect the wrong license plate, improve the accuracy of data collected by VLPR system.

Key words: intelligent transportation; vehicle license plate recognition (VLPR); character replacement; travel time

0 引言

车牌识别技术是指能够检测到受监控路面的车辆并自动提取车辆牌照信息进行处理的技术。在城市道路的交叉路口, 安装有车牌号码抓拍识别设备, 记录车

辆的经过时间、位置、车牌号码、车辆类型等信息(一条过车记录)。车牌识别是智能交通的重要技术之一, 在交通管理、治安防控和嫌疑车辆追踪定位等领域发挥着重要的作用, 具有广泛的应用前景^[1-2]。通过对大量的过车记录数据以及抓拍视频图像进行统计分

收稿日期: 2017-03-07

基金项目: 浙江省自然科学基金资助项目(LY17F030017); 杭州市社会发展科研自主申报项目(20160533B82)

作者简介: 刘晴辉(1990-), 男, 河南信阳人, 硕士研究生, 主要从事智能交通方面的研究。E-mail: 727564723@qq.com

通信联系人: 董红召, 男, 教授, 博士生导师。E-mail: its@zjut.edu.cn

析,车辆抓拍识别设备的总体识别率并不高,这是因为车牌识别设备对光线比较敏感,某些车牌字符的相似性,车牌字符磨损、遮挡以及车辆行驶速度等都会影响车牌字符识别的准确率,这些因素导致过车记录数据中含有错误车牌号码信息。

国内外对车牌识别技术研究主要集中在车牌定位、字符分割及字符识别等方面。车牌定位算法方面,如基于灰度图像的算法、基于彩色图像的算法、融合定位算法等^[3-6];字符分割算法方面,如垂直投影与模板匹配结合法、区域标记法、形态学法^[7-10]等;字符识别算法方面主要有模板匹配法、特征分析匹配法、机器学习法^[11-12]等;这些方法都依赖于抓拍获取的较高质量图像,即抓拍识别设备的好坏、拍摄角度、外界环境、车辆等因素都会影响算法得到的结果。

本研究提出一种对过车记录数据作时空判别进行车牌号码检错与纠错的方法。

1 车牌号码检错纠错原理

首先对常用的符号进行定义:

$$\mathbf{A}_T^{k_A} = \begin{Bmatrix} k_A & t_1 & h_1 & \cdots \\ \vdots & \vdots & \vdots & \\ k_A & t_i & h_i & \cdots \\ \vdots & \vdots & \vdots & \\ k_A & t_n & h_n & \cdots \end{Bmatrix} \quad (1)$$

式中: $\mathbf{A}_T^{k_A}$ —抓拍点位 A 在时间段 T 内的过车记录数据集用集合, k_A —抓拍点位 A 的特征编号; t_i, h_i —过车记录中的抓拍时间和识别出的车牌号码。

$$\mathbf{B}_T^{hphm} = \{h_1, h_2, \cdots, h_n\}^T \quad (2)$$

$$\mathbf{B} = \{h_1, h_2, \cdots, h_m\}^T$$

式中: $\mathbf{B}_T^{hphm}$ —集合 $\mathbf{A}_T^{k_A}$ 中所有车辆车牌号码组成的集合;集合 $\mathbf{B}$ —本地所有已发放车牌号码数据集。

1.1 路段平均行驶速度计算方法

本研究基于过车记录数据计算得到的路段平均行驶速度是检错纠错方法中时空判别的依据。抓拍点位 A 与 B 之间的路段 T 时段的平均速度计算方法如下:

抓拍点位 A 与 B 在时段 T 的过车记录集合分别用 $\mathbf{A}_T^{k_A}$ 和 $\mathbf{A}_T^{k_B}$ 表示:

$$\mathbf{A}_T^{k_A} = \begin{Bmatrix} k_A & t_1 & h_1 & \cdots \\ \vdots & \vdots & \vdots & \\ k_A & t_i & h_i & \cdots \\ \vdots & \vdots & \vdots & \\ k_A & t_u & h_u & \cdots \end{Bmatrix}$$

$$\mathbf{A}_T^{k_B} = \begin{Bmatrix} k_B & t_1 & h_1 & \cdots \\ \vdots & \vdots & \vdots & \\ k_B & t_j & h_j & \cdots \\ \vdots & \vdots & \vdots & \\ k_B & t_v & h_v & \cdots \end{Bmatrix} \quad (3)$$

式中:集合 $\mathbf{A}_T^{k_A k_B}$ —所有在时段 T 内既经过抓拍点位 A 又经过抓拍点位 B 的车辆过车记录集合; t_i, t'_i —车辆 h_i 在 A, B 的抓拍时间。

$$\mathbf{A}_T^{k_A k_B} = \begin{Bmatrix} h_1 & t_1 & t'_1 & k_A k_B \\ \vdots & \vdots & \vdots & \\ h_i & t_i & t'_i & k_A k_B \\ \vdots & \vdots & \vdots & \\ h_w & t_w & t'_w & k_A k_B \end{Bmatrix} \quad (4)$$

集合 $\mathbf{A}_T^{k_A k_B}$ 的抓拍时间作差取绝对值之后的过车记录集合用 $\mathbf{C}_T^{k_A k_B}$ 表示:

$$\mathbf{C}_T^{k_A k_B} = \begin{Bmatrix} h_1 & \Delta t_1 & k_A k_B \\ \vdots & \vdots & \vdots \\ h_i & \Delta t_i & k_A k_B \\ \vdots & \vdots & \vdots \\ h_w & \Delta t_w & k_A k_B \end{Bmatrix} \quad (5)$$

式中: $\Delta t_i = |t_i - t'_i|$ 。

集合 $\mathbf{A}_T^{k_A k_B}$ 代表所有在时段 T 内既经过抓拍点位 A 又经过抓拍点位 B 的过车记录集合,但并不代表车辆一次性通过 A 与 B 之间的路段,通过设置一个路段行程时间筛选时间窗(下文简称时间窗)来去除 $\mathbf{C}_T^{k_A k_B}$ 中的冗余数据。定义一个速度区间 $[v_1, v_2]$, $L(k_A k_B)$ 表示抓拍点位 A 与 B 之间路段的长度,则对应的时间窗为 $[L(k_A k_B)/v_2, L(k_A k_B)/v_1]$ 。去除集合 $\mathbf{C}_T^{k_A k_B}$ 中 Δt_i 不处于时间窗内的过车记录数据,通过时间窗筛选后的数据集集合为有效数据,用 $\overline{\mathbf{C}_T^{k_A k_B}}$ 表示:

$$\overline{\mathbf{C}_T^{k_A k_B}} = \begin{Bmatrix} h_1 & \Delta t_1 & k_A k_B \\ \vdots & \vdots & \vdots \\ h_j & \Delta t_j & k_A k_B \\ \vdots & \vdots & \vdots \\ h_s & \Delta t_s & k_A k_B \end{Bmatrix} (s \leq w) \quad (6)$$

则路段 $k_A k_B$ 在时段 T 内车辆的平均行驶速度为:

$$\overline{V_{k_A k_B}^T} = \frac{L(k_A k_B)}{\frac{1}{s} \sum_{j=1}^s \Delta t_j} \quad (7)$$

1.2 车牌号码检错原理步骤

(1) 集合 $\mathbf{B}_T^{hphm}$ 与集合 $\mathbf{B}$ 匹配, $\mathbf{B}_T^{hphm}$ 与 $\mathbf{B}$ 的差集为 $\mathbf{B}_T^{hphm}$ 中未发放车牌号码组成的集合,记为集合 $\mathbf{B}_T^1$ ($\mathbf{B}_T^1 =$

$B_T^{hphm} - B = \{h_1, h_2, \dots, h_p\}^T$, B_T^1 中的车牌号码必定是错误识别的车牌号码,初步得到抓拍点位 A 在 T 时段的错误识别车牌号码集合; B_T^{hphm} 与 B 的交集表示 B_T^{hphm} 中已经发放车牌号码组成的集合,记为集合 B_T^2 ,对 B_T^2 中的车牌号码作如下时空判别($B_T^2 = B_T^{hphm} \cap B = \{h_1, h_2, \dots, h_q\}^T$)。

(2) 遍历集合 B_T^2 中的车牌号码,对 $\forall h_i \in B_T^2, h_i$ 对应的过车记录的经过时间为 t_i 。设置一个时间半径参数 t_r ,车牌号码 h_i 在时间段 $[t_i - t_r, t_i + t_r]$ 的所有过车记录集合记为 $A_{h_i}^{t_i}$ (按时间先后排列):

$$A_{h_i}^{t_i} = \left\{ \begin{matrix} h_i & k_1 & t_1 \\ \vdots & \vdots & \vdots \\ h_i & k_j & t_j \\ \vdots & \vdots & \vdots \\ h_i & k_s & t_s \end{matrix} \right\} (t_i - t_r < t_j < t_i + t_r) \quad (8)$$

(3) 集合 $A_{h_i}^{t_i}$ 中过车记录两两时间作差,得到集合 $\overline{C_{h_i}^{t_i}}$ 。

$$\overline{C_{h_i}^{t_i}} = \left\{ \begin{matrix} h_i & k_1 k_2 & \Delta t_1 \\ \vdots & \vdots & \vdots \\ h_i & k_j k_{j+1} & \Delta t_j \\ \vdots & \vdots & \vdots \\ h_i & k_{s-1} k_s & \Delta t_{s-1} \end{matrix} \right\} \quad (9)$$

式中: Δt_j —抓拍路口 k_j 到 k_{j+1} 之间的路段 $k_j k_{j+1}$ 车辆 h_i 耗费的行程时间。

(4) 如果满足时空判别条件:

$$\forall \frac{L(k_j k_{j+1})}{\Delta t_j} \leq \lambda \cdot \overline{V_{k_j k_{j+1}}^{t_j}} \quad (10)$$

式中: $\overline{V_{k_j k_{j+1}}^{t_j}}$ —路段 $k_j k_{j+1}$ 在时间段 $[t_j, t_{j+1}]$ 的平均行驶速度,可以通过计算路段平均行驶速度方法得出; λ —平均速度的修正系数,是一个大于 1 的无量纲参数。

则 h_i 为正确识别的车牌号码,否则 h_i 识别错误。遍历集合 B_T^2 ,所有不满足时空判别条件的车牌号码认为是错误识别的车牌号码,将该车牌号码添加入集合 B_T^1 中,得到新的集合 $\overline{B_T^1}$ 。则最终集合 $\overline{B_T^1}$ 为检错得出的错误识别车牌号码集合。检错主要步骤流程图如图 1 所示。

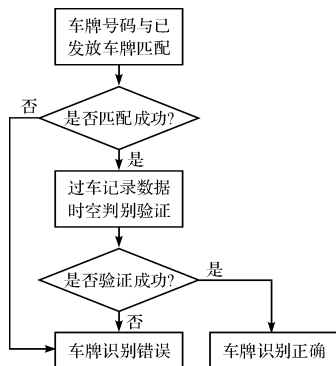


图1 车牌号码检错步骤流程图

1.3 车牌号码纠错原理步骤

(1) 对 $\forall h_i \in \overline{B_T^1}$, 设 h_i 的一条过车记录为 $\{h_i, t_0, k_0, \dots\}$, 首先对 h_i 进行字符替换。由于过车记录中主要为本地车辆,同时受到实验条件所限(缺少全国其他地区已发放车牌数据),车牌号码的前两位字符不做替换。考虑到替换车牌号码与原车牌号码要有较高的相似性,只对 h_i 后 5 位字符的其中一位进行替换,替换字符为 24 个英文字符(不包含字母 O 和 I) 与 10 个数字字符,产生新的替换车牌号码集 R_1 。

$$R_1 = \{h_{re}^1, h_{re}^2, \dots, h_{re}^n\}^T \quad (11)$$

(2) 集合 R_1 中有很多车牌号码不符合编排规则,而且很多符合编排规则但实际并未发放。将集合 R_1 与 B 取交集,得到符合编排规则且已经发放的替换车牌号码集合,记为 R_2 :

$$R_2 = R_1 \cap B = \{h_{re}^1, h_{re}^2, \dots, h_{re}^m\}^T \quad (12)$$

(3) 遍历集合 R_2 ,对 $\forall h_{re}^i \in R_2$, 车辆 h_{re}^i 在时间段 $[t_0 - t_r, t_0 + t_r]$ 的过车记录用集合 $A_{h_{re}^i}^{t_0}$ 表示:

$$A_{h_{re}^i}^{t_0} = \left\{ \begin{matrix} h_{re}^i & k_1 & t_1 \\ \vdots & \vdots & \vdots \\ h_{re}^i & k_j & t_j \\ \vdots & \vdots & \vdots \\ h_{re}^i & k_s & t_s \end{matrix} \right\} (t_0 - t_r < t_j < t_0 + t_r) \quad (13)$$

车辆 h_i 在时间段 $[t_0 - t_r, t_0 + t_r]$ 的过车记录表示为:

$$A_{h_i}^{t_0} = \left\{ \begin{matrix} h_i & k_1 & t_1 \\ \vdots & \vdots & \vdots \\ h_i & k_i & t_i \\ \vdots & \vdots & \vdots \\ h_i & k_x & t_x \end{matrix} \right\} (t_0 - t_r < t_i < t_0 + t_r) \quad (14)$$

(4) 集合 $A_{h_i}^{t_0}$ 的过车记录数据中 h_i 是错误识别的数据,但对应的抓拍时间 t_i 和抓拍点位标识 k_i 是正确信息。若替换车牌号码 h_{re}^i 是 h_i 的正确车牌号码,则过车记录集合 $A_{h_{re}^i}^{t_0}$ 和 $A_{h_i}^{t_0}$ 的并集为车辆完整的过车记录集合,记为集合 $A_h^{t_0}$ (按时间排序)。这里为了方便集合表示,用 h 代替 h_i 和 h_{re}^i 。

$$A_h^{t_0} = \left\{ \begin{matrix} h & k_1 & t_1 \\ \vdots & \vdots & \vdots \\ h & k_j & t_j \\ \vdots & \vdots & \vdots \\ h k_{s+x} & t_{s+x} \end{matrix} \right\} \quad (15)$$

(5) 循环 1.2 节的(3 ~ 4) 步骤, 如果满足时空判别条件, 则 h_{re}^i 是 h_i 的正确替换车牌号码, 否则, 继续遍历集合 R_2 , 直到遍历完集合 R_2 中所有车牌号码, 纠错主要步骤流程图如图 2 所示。

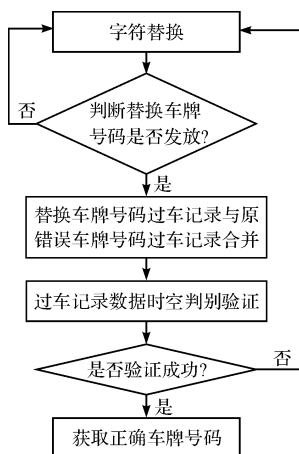


图 2 车牌号码纠错步骤流程图

2 参数的确定

车牌号码检错纠错方法涉及到的参数有 5 个, 分别是路段行驶速度区间参数 v_1 、 v_2 , 时间半径 t_r , 路段平均行驶速度修正系数 λ 。

(1) 路段行驶速度区间参数 v_1 、 v_2 的确定是为了利用过车记录数据计算路段平均行驶速度时去除冗余数据。 v_1 、 v_2 需要符合不同等级路段的设计时速, 同时考虑交叉路口信号周期、驾驶行为、车辆类型、交通拥堵等因素对平均行驶速度的影响。如果速度区间 $[v_1, v_2]$ 区间长度过小, 则筛选后的样本数据量小, 不利于平均速度的计算; 如果 v_1 选取过小, 则时间窗上限值 $L(k_A k_B)/v_1$ 过大, 不能很好地剔除冗余数据, 从而导致计算得到的路段平均行程时间值偏大, 路段平均行驶速度值偏小。通过大量的数据分析, 选取 $v_1 = 5 \text{ km/h}$, $v_2 = 100 \text{ km/h}$ 。

以杭州市绍兴路(香积寺路与大关路之间路段, 路段长度为 0.93 km) 一段时间的过车记录数据为例, 时间窗的大小为 $[34, 670]$, 单位为 s。筛选之前的路段行程时间分布图如图 3 所示, 样本量为 179。由图 3 可以看出, 有多辆车经过这个路段的行程时间过大, 这些显然是冗余数据, 在计算路段平均行程时间以及路段平均行驶速度时必须剔除。

经过时间窗筛选之后的路段行程时间分布如图 4 所示, 样本量为 120, 剔除了 59 个样本数据, 并且这 59 个数据中大部分数值较大, 为冗余数据, 利用筛选后的行程时间, 计算得出的该路段在这一时段的平均行程

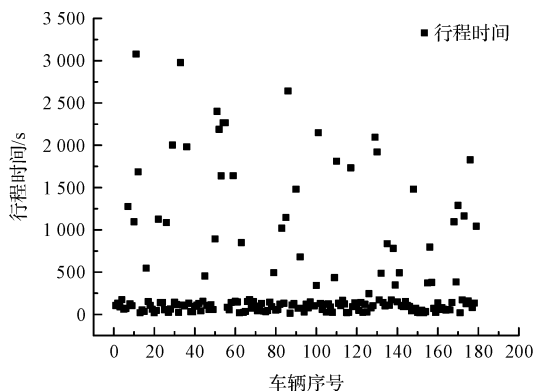


图 3 筛选前路段行程时间分布图

时间为 134.16 s, 平均行驶速度为 24.99 km/h。因此参数 v_1 、 v_2 的值选取合理, 设置的时间窗能够较好地剔除冗余数据。

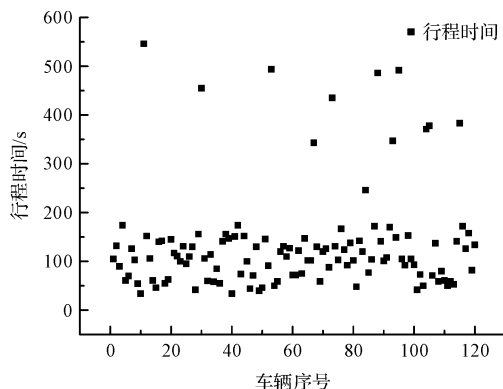


图 4 筛选后路段行程时间分布图

(2) 参数 t_r 可以减小车牌号码检错与纠错步骤计算的数据量, 提高检错和纠错方法的运行效率。 t_r 值的选择要符合人们日常开车出行的时间规律, 同时要保证覆盖车辆一次出行的时间范围。通过对大量的过车记录数据分析以及日常经验, t_r 参数的值选择 12 h 比较合适。

(3) 路段平均行驶速度修正系数 λ 是一个大于 1 的无量纲参数。路段平均行驶速度的计算与 v_1 、 v_2 参数的选择有关, 因此参数 λ 是对计算得出的路段平均行驶速度的修正, 同时参数 λ 也限定了车牌检错纠错方法中时空判定条件的宽松程度。为了避免将正确采集的车牌号码错误的检测出来, 车牌号码检错步骤中选择较大的 λ 值, 而较小的 λ 值更有利于车牌号码的纠错。因此车牌号码检错与纠错步骤 λ 值分别取 3 和 1.5。

3 实 验

本研究选取了杭州市主城区车辆抓拍识别设备采集的过车记录数据以及杭州市车辆登记信息数据进行实验, 并以车牌识别设备采集的视频图像数据进行验

证。过车记录数据结构如表1所示。

表1 过车记录数据结构表

字段名	字段类型	描述
XH	int	序号
KKID	bigint	抓拍点位编号
JGSJ	datetime	抓拍时间
HPHM	varchar(50)	车牌号码

实验选取了杭州市绍兴路与香积寺路交叉口南抓拍点位3个时段(每个时段15 min)共计511张车辆抓拍识别设备采集的图像数据进行对比实验。各时段本地车、外地车、未识别车流量数据如表2所示。

表2 实验数据流量表

类别	时段1	时段2	时段3
未识别流量	29	22	5
外地车流量	15	30	20
本地车流量	114	97	179
总计	158	149	204

本研究通过检错算法对3个时段共计390辆本地车车牌号码进行检错实验,分别检测出5、9、3个错误识别车牌号码。对比视频图像数据进行验证,实际分别有7、10、6个车牌号码识别错误。平均检错正确率(检错出的车牌号码中实际错误车牌号码个数与检错出的车牌号码总数的比值)为100%。平均检错检出率(检错出的车牌号码中实际错误车牌号码个数与实际错误识别车牌号码总数的比值)为74%。

本研究对检错步骤检测出的17个车牌号码进行纠错,并与视频图像数据进行比对,共有5个错误识别的车牌号码纠错后给出了唯一并与实际相符的正确车牌号码,纠错成功率(纠错后给出唯一正确车牌号码数与纠错车牌号码总数的比值)为29.4%。其他12个错误车牌号码纠错后得到的正确车牌号码大于一个或者未能得到正确的车牌号码。3个时段实际错误识别车牌号码量、检错成功量与纠错成功量如表3所示。

表3 检错纠错实验结果表

类别	时段1	时段2	时段3
实际错误识别量	7	10	6
检错检测错误识别量	5	9	3
检错正确检测出的数量	5	9	3
成功纠错量	2	2	1

4 结束语

本研究提出了基于过车记录数据时空判别的车牌号码检错纠错方法,同时对该方法进行了实际测试,验证了该方法对错误识别车牌号码具有较高的检错成功率,并能够纠正部分错误识别的车牌号码,进一步提高车牌号码识别数据的准确率。

该方法依赖于较高的车辆抓拍点位密度以及实时更新的已发放车牌数据。目前城市抓拍识别设备主要分布在主城区的交叉路口,限制了该方法的应用范围;同时,由于抓拍点位密度不足,也在一定程度上对检错纠错时空判别的有效性产生影响。

参考文献(References):

- [1] 文洁玲,上官明灿,孙文侠,等.浅谈车牌识别技术应用现状[J].公路交通科技:应用技术版,2006(7):160-161.
- [2] 徐海铭.一种车牌识别系统的设计与实现[J].机电工程技术,2015,44(1):66-68.
- [3] 张浩鹏,王宗义.基于灰度方差和边缘密度的车牌定位算法[J].仪器仪表学报,2011,32(5):1095-1102.
- [4] 吴惟希,赵刚.一种基于色彩的车牌图像定位算法[J].信息安全研究,2016,2(1):58-65.
- [5] 杨硕,张波,张志杰.多特征融合的车牌定位算法[J].计算机应用,2016,36(6):1730-1734.
- [6] 王枚,苏光大,王国宏.复杂环境下的车牌定位及目标真实性验证[J].光学精密工程,2009,17(4):886-894.
- [7] 卜英家.基于计算机视觉的车牌识别系统的算法研究[D].成都:电子科技大学电子工程学院,2015.
- [8] SHAN B. License plate character segmentation and recognition based on RBF neural network[C]. Second International Workshop on Education Technology and Computer Science, Wuhan: IEEE,2010.
- [9] 崔文学,崔义川,王朝晖,等.基于模板匹配和垂直投影的车牌字符分割算法[J].齐齐哈尔大学学报:自然科学版,2015,31(6):12-16.
- [10] 甘玲,林小晶.基于连通域提取的车牌字符分割算法[J].计算机仿真,2011,28(4):336-339.
- [11] 陈学保.车牌字符识别算法的研究[D].重庆:重庆大学光电工程学院,2013.
- [12] 董玲娇.车牌自动识别中的字符特征提取[J].机电工程,2008,25(9):106-108.

[编辑:李辉]

本文引用格式:

刘晴辉,徐巍,董红召.基于抓拍数据时空判别的车牌检错纠错研究[J].机电工程,2017,35(1):101-105.

LIU Qing-hui, XU Wei, DONG Hong-zhao. Errordetection and correction based on time-space judgement of data captured by VLPR system[J]. Journal of Mechanical & Electrical Engineering, 2017,35(1):101-105.

《机电工程》杂志:http://www.meem.com.cn