

# 基于单目视觉的车辆下边沿估计和逆透视变换的车距测量<sup>\*</sup>

王永森, 刘宏哲

(北京联合大学北京市信息工程重点实验室, 北京 100101)

**摘 要:** 前方车辆测距在自动驾驶汽车技术领域中起着至关重要的作用。针对目前基于单目视觉的车辆测距技术忽略了车辆与地面相接的下边沿问题, 提出一种基于车辆下边沿估计和逆透视变换的单目视觉测距方法, 实现了对前方车辆进行横向和纵向的高精度车距测量。该方法首先通过对车辆关键点估计和几何关系模型完成对车辆下边沿的估计, 然后从中计算测距关键点, 再利用基于点的逆透视变换测距模型进行距离计算。实验结果表明, 与其他基于单目视觉的车辆测距方法相比, 该方法提高了测距的精度和稳定性。

**关键词:** 单目视觉; 车辆测距; 视觉测距; 逆透视变换; 自动驾驶汽车

**中图分类号:** TP391.4

**文献标志码:** A

**doi:** 10.3969/j.issn.1007-130X.2020.07.012

## Vehicle distance measurement with vehicle lower edge estimation and inverse perspective mapping based on monocular vision

WANG Yong-sen, LIU Hong-zhe

(Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, China)

**Abstract:** Front vehicle ranging plays a vital role in the field of autonomous vehicle technology. Aiming at the problem that the current vehicle ranging technology based on monocular vision neglects the lower edge of the vehicle connected to the ground, this paper proposes a monocular vision ranging model based on vehicle lower edge estimation and inverse perspective transformation, which realizes high-precision distance measurement of lateral and longitudinal directions of vehicles in front. Firstly, the vehicle's key point estimation and geometric relationship model are used to estimate the lower edge of the vehicle, then the distance measurement key points are calculated, and the point-based inverse perspective transformation distance measurement model is used to calculate the distance. The experimental results show that, compared with other monocular vision vehicle ranging algorithms, this method improves the accuracy and stability of ranging.

**Key words:** monocular vision; vehicle distance measurement; visual ranging; inverse perspective mapping; autonomous vehicle

### 1 引言

随着经济的快速发展, 我国的机动车产量和保有量已经连年高居世界第一, 机动车虽然给人们的

生活带来了方便, 但也使交通事故频发, 其中由于驾驶员操作不当而引发的道路交通事故占到了总事故的 70% 以上。自动驾驶汽车的发展能够有效地解决这一难题, 保障交通系统的运行安全并且提高运输效率, 对促进国民经济的发展具有重要的意

<sup>\*</sup> 收稿日期: 2019-12-05; 修回日期: 2020-02-27

基金项目: 国家自然科学基金(61871039); 北京市属高校高水平教师队伍建设支持计划(IDHT20170511); 北京联合大学领军人才项目(BPHR2019AZ01); 北京联合大学项目(202011417004, 202011417005, WZ10201903); 北京联合大学研究生科研创新资助项目

通信作者: 刘宏哲(liuhongzhe@bnu.edu.cn)

通信地址: 100101 北京市北京联合大学北京市信息工程重点实验室

Address: Beijing Key Laboratory of Information Service Engineering, Beijing Union University, Beijing 100101, P. R. China

义和经济价值。作为自动驾驶汽车重要组成部分,驾驶场景视觉感知与分析已经成为自动驾驶领域中的重点研究对象,其中前方车辆检测是场景感知工作中至关重要的一部分,并且准确测量当前车辆与前方车辆之间的距离可以对车辆进行准确定位,对自动驾驶场景分析与决策系统起着非常重要的作用,对安全行车意义重大<sup>[1-6]</sup>。

目前,驾驶环境下前方车辆测距方法根据传感器不同分为基于雷达和基于视觉2类。基于毫米波雷达和激光雷达等主动式传感器的测距方法<sup>[7-9]</sup>虽然测距精度较高,但价格昂贵,同时易受外界环境干扰。基于视觉传感器的测距方法应用范围广泛,所获得的信息丰富,价格低廉,并且目前已经出现了很多优秀的目标检测算法<sup>[10-16]</sup>,为视觉测距提供了很好的目标定位基础。基于视觉的测距方法又分为基于单目视觉和基于多目视觉2类。基于多目视觉的测距方法<sup>[17-19]</sup>需要融合多个摄像头的信息,算法复杂,计算量大,相机标定困难,成本相对较高。基于单目视觉的测距方法实现原理简单,成本低廉,实时性更好。

近年来,国内外对单目视觉测距方法的研究取得了一些进展。文献[20]提出了一种面向未来智能交通的自适应前方车辆单目视觉测距方法,通过计算图像坐标系下车道标志线、交通标志、车辆尺度信息,依据相机成像数学模型计算前方目标车辆距离,但其需要特定场景环境,需要获取多个道路相关信息,不具普适性。文献[21]为了克服传统单目视觉测距中时变的摄像机外参数对测量结果的影响,利用道路车道标识线平行约束条件,提出了一种基于车道线消隐点的单目视觉自适应车辆测距方法,但其需要对道路车道线信息进行提取,测距精度受车道线提取准确度的影响,只适用于特定结构化道路。文献[22]提出了一种使用单目相机估算车辆在弯道时前方车辆距离的算法,通过分析车辆在弯道行驶时的特征对前方车辆进行检测,在此基础上实现车距测量,该方法受限于车辆在弯道行驶时的特征信息,只适用于特定场景。文献[23]利用光流法来计算前方车辆距离,从单目相机测得的光流和车辆控制端的信息来估计车辆的距离和速度,但该方法依赖控制信息,且受光变环境影响较大,不具有普适性。

文献[24-26]所研究的测距方法是基于相机成像横向模型算法来计算目标车辆的距离,该方法需要已知目标车辆的实际物理宽度,车辆宽度大约在

1.4~2.6 m,根据相似三角形原理计算车辆距离,如果预估的车辆宽度和实际车辆宽度有较大差异则会产生很大的测距误差。文献[27]利用纵向测距模型来计算车辆距离,基本原理也是根据相机成像原理,但是与利用横向测距模型方法不同的是,他们的方法是用自身的相机高度来代替车辆的物理宽度参数,这样相当于已知一个较为准确的物理值,不需要估算车辆的宽度,提高了测距的精度。但是,该方法忽略了车辆下边沿问题,无法确定车辆与地面相接的准确位置信息,且受相机安装角度的精度影响,导致测距误差的产生。文献[28]提出了一种基于水平方向和垂直方向的相似三角形测距方法,减少了由于车辆自身在远距离上的俯仰角而导致的误差,但该方法没有考虑到车辆与地面相接的下边沿位置问题,造成测量误差。文献[29]提出了一种基于逆透视变换的测距方法,利用逆透视变换得到道路俯视图,根据线性关系进行距离计算,该方法虽然简单易行,但忽略了车辆测距的关键点选取问题,只能估计车身整体大致位置,无法精确定位,且对相邻车道内的车辆测距误差较大。

通过对上述研究成果进行分析,目前基于单目视觉的车辆测距方法忽略了车辆下边沿对测距产生的影响,导致测距关键点选取不够精准,造成测量误差的产生。为了提高驾驶环境中前方车辆单目视觉测距的精度和稳定性,本文提出了基于车辆下边沿估计和逆透视变换的车距测量方法,利用车辆关键点估计车辆三维信息,提取车辆与地面相接的下边沿位置坐标,计算测距关键点,通过基于点的逆透视变换测距模型,实现对车辆距离的高精度测量。

## 2 方法总体设计

如图1所示,为了实现车辆下边沿估计,本文方法借助基于深度学习的目标检测算法完成驾驶场景中车辆目标以及车辆关键组件的检测。车辆下边沿估计模型通过分析车辆视角,对车辆所处位姿进行分类,参照车辆组件位置,计算车辆关键点坐标,同时依据相机成像原理和几何先验模型进行车辆三维信息估计,得到车辆下边沿信息。逆透视变换模型根据相机成像原理对相机进行标定,得到相机内外参数,从车辆下边沿信息估计测距关键点坐标,根据逆透视过程,对测距关键点进行距离计算,完成车距测量。

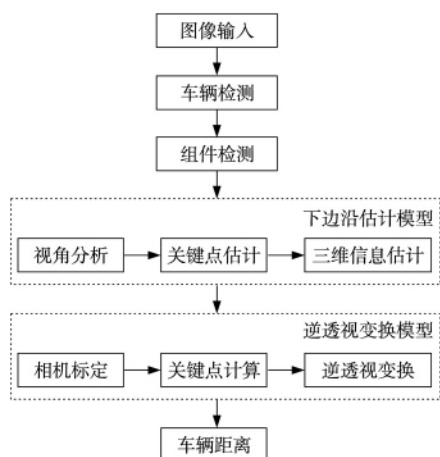


Figure 1 Overall design of the method

图1 方法总体设计

### 3 车辆下边沿估计模型

车辆下边沿是车辆在道路上时车尾的垂直截面与地面相接的位置,如图 2a 所示,而常规的车辆检测结果如图 2b 所示。传统的测距模型建立在车辆矩形检测框的基础上,测距点的选择基于矩形框的下边,如图 2c 所示,通过对比,真实的车辆下边沿和检测得到的矩形框的下边存在差异,导致测距误差的产生。特别地,当车辆下边沿估计不准时,将对横向距离估计造成非常大的影响。为了从根本上提高测距精度,找到更加准确的车辆下边沿非常重要。本文建立车辆下边沿估计模型,对车辆下边沿进行估计。

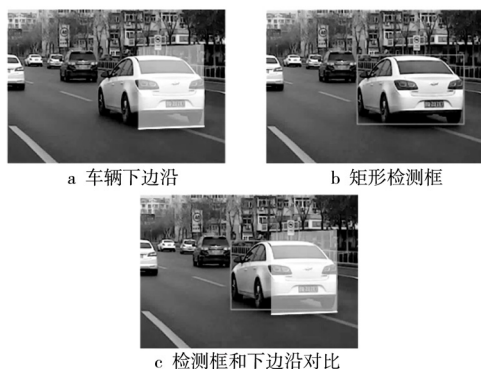


Figure 2 Comparison of the lower edge of the vehicle and the detection rectangle

图2 车辆下边沿与检测矩形框对比

#### 3.1 车辆视角分析

车辆在行驶过程中,由于所在位置相对于当前车辆会时刻发生变化,相机所捕捉到的车辆在图像中会呈现不同的位姿角度,相当于观察车辆的视角不同,从而增加了车辆下边沿估计的难度。如图 3

所示,对车辆在行驶过程中的位置变化进行分析,观察车辆的视角变化的同时,车辆的下边沿相对于车辆矩形检测框的位置也发生变化。



Figure 3 Comparison of different vehicle perspectives

图3 不同车辆视角对比

根据车辆位姿和观察视角的不同,将车辆关键点的选择分为 2 种情况:

(1) 可以观察到车辆侧面的情况。

如图 3 所示,针对前 6 幅图像中可以确定车辆侧面的情况,即能看到车轮所在的侧面,如果不对车辆下边沿进行估计,而直接采取车辆矩形检测框进行距离测量,将会造成测距误差的产生,尤其对横向距离的估计影响较大。

如图 4 所示,为实现车辆下边沿的估计,本文分别选取车辆侧面的 2 个车轮和地面相接的地方作为 2 个关键点  $w_0(w_{0(x)}, w_{0(y)})$ ,  $w_1(w_{1(x)}, w_{1(y)})$ , 除该 2 点之外,选取车辆车尾或者车头与相机中可以看到侧面相交的棱,定义为直线  $L(x = p_0)$ , 其中  $p_0$  为图像坐标横轴的坐标值,即需要的 5 个参数为  $N(w_{0(x)}, w_{0(y)}, w_{1(x)}, w_{1(y)}, p_0)$ 。



Figure 4 Vehicle side view

图4 车辆侧方视角

(2) 无法观察到车辆侧面的情况。

如图 5 所示,图像中的目标车辆接近于正前方车辆的观察视角,这种情况下车辆的下边沿和矩形检测框非常接近,使用矩形检测框作为车辆的下边沿进行测距所产生的误差相对较小。所以,对于该种情况的车辆不进行车辆下边沿估计,使用矩形检

测框的下边作为车辆下边沿进行测距。

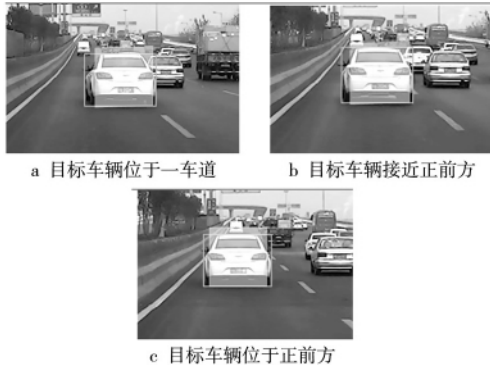


Figure 5 Perspective from the rear of the vehicle

图 5 车辆正后方视角

在实际应用过程中,为了确定车辆所处的位姿以进行车辆关键点估计,定义变量  $\lambda$ ,  $\lambda$  表示检测到的车尾矩形框中心点横坐标与车辆外接矩形框宽度的比值(车辆外接矩形框的左上角顶点位置为坐标原点),根据  $\lambda$  的大小判定车辆所处位姿和观察视角,若  $\lambda \notin [0.42, 0.58]$ ,属于车辆侧方视角;若  $\lambda \in [0.42, 0.58]$ ,属于车辆正后方视角。

### 3.2 车辆关键点估计

车辆关键点估计,首先利用目标检测算法<sup>[11,13]</sup>对图像中的车辆进行检测,得到车辆外接矩形框坐标位置  $B(x_1, y_1, x_2, y_2)$ ,截取车辆最小外接矩形框包围的像素信息。分析 3.1 节情况(1)中关键点位置信息,可通过对相应组件检测得到关键点位置坐标。如图 6 所示,利用 2 个侧方车轮和车辆尾部的位置信息,估计出车辆关键点。



Figure 6 Vehicle key components detection

图 6 车辆关键组件检测

使用 SSD(Single Short multi box Detector)目标检测框架<sup>[13]</sup>对截取的矩形框包围的车辆图像进行检测,如图 7 所示,得到车轮的矩形框位置



Figure 7 Key points estimation

图 7 关键点估计

$W_0(x_1, y_1, x_2, y_2)$ ,  $W_1(x_1, y_1, x_2, y_2)$  和车尾的矩形框  $T_0(x_1, y_1, x_2, y_2)$ 。

关键点计算过程中,前侧车轮和后侧车轮组件矩形框分别使用  $wb_0$  和  $wb_1$  表示,它们可由式(1)和式(2)得到:

$$wb_0 = \begin{cases} W_0, & \text{if } W_0(y_1) < W_1(y_1) \\ W_1, & \text{else} \end{cases} \quad (1)$$

$$wb_1 = \begin{cases} W_0, & \text{if } W_0(y_1) > W_1(y_1) \\ W_1, & \text{else} \end{cases} \quad (2)$$

变量  $p_0$  由式(3)得到:

$$p_0 = \begin{cases} T_0(x_1), & \text{if } T_0(x_1) - W_0(x_1) < \\ T_0(x_2) - W_0(x_1) \\ T_0(x_2), & \text{else} \end{cases} \quad (3)$$

2 个关键点  $w_0$  和  $w_1$  的坐标由式(4)和式(5)得到,即  $G = (w_0, w_1, p_0)$ 。

$$w_0 = \left( wb_0(x_1) + \frac{wb_0(x_2) - wb_0(x_1)}{2}, wb_0(y_2) \right) \quad (4)$$

$$w_1 = \left( wb_1(x_1) + \frac{wb_1(x_2) - wb_1(x_1)}{2}, wb_1(y_2) \right) \quad (5)$$

### 3.3 车辆三维信息估计

如图 8 所示,车辆目标的三维信息在图像中使用顶点坐标表示为  $C = \{a, b, c, d, e, f, g, h\}$ ,其中,  $\{a, b, c, d\}$  表示目标车辆靠近相机侧垂直切面上从左上角按照顺时针方向的 4 个顶点在二维图像上的像素坐标点,  $\{e, f, g, h\}$  表示目标车辆远离相机侧垂直切面上从左上角按照顺时针方向的 4 个顶点在二维图像上的像素坐标点。



Figure 8 Vehicle 3D information

图 8 车辆三维信息

由  $G$  得到  $N = (w_{0(x)}, w_{0(y)}, w_{1(x)}, w_{1(y)}, p_0)$ ,

依据几何模型计算车辆三维信息中的顶点坐标。计算过程如下:

如图 9 所示,由式(6)计算经过  $w_0(w_{0(x)}, w_{0(y)})$  和  $w_1(w_{1(x)}, w_{1(y)})$  2 个关键点的直线方程  $F(x)$ :

$$(w_0, w_1) \Rightarrow F(x) \quad (6)$$

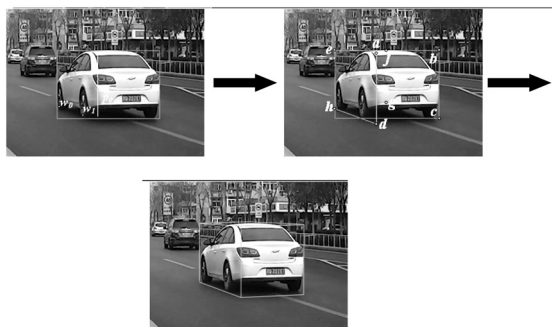


Figure 9 Vehicle 3D information estimation

图 9 车辆三维信息估计

由式(7)计算直线方程  $F(x)$  与直线  $L(x = p_0)$  相交的点,记为构建 3D 立体框  $C = \{a, b, c, d, e, f, g, h\}$  中的顶点  $d$ :

$$\left. \begin{array}{l} F(x) \\ L(x = p_0) \end{array} \right\} \Rightarrow d \quad (7)$$

由式(8)可得直线方程  $F(x)$  与车辆矩形框左侧边  $K(x = k_0)$  相交的点,记为  $C = \{a, b, c, d, e, f, g, h\}$  中的顶点  $h$ :

$$\left. \begin{array}{l} F(x) \\ K(x = k_0) \end{array} \right\} \Rightarrow h \quad (8)$$

然后取车辆矩形框右侧边下顶点为  $c$  点,取车辆矩形框左上角顶点为  $e$  点,由几何原理,依次平移线段  $eh$  到  $d$  点和  $c$  点,可得  $a$  点和  $b$  点,然后依次平移线段  $hd$  使  $d$  点分别和  $c$  点、 $b$  点重合,得到  $g$  点、 $f$  点,即可计算出车辆三维信息  $C = \{a, b, c, d, e, f, g, h\}$ 。

由车辆三维信息  $C = \{a, b, c, d, e, f, g, h\}$ ,一般情况下可直接由线段  $cd$  得到车辆下边沿,但关键点  $G$  会因组件检测的偏差导致误差增大,该情况下根据几何模型计算出的顶点坐标也将出现误差,随之估计得到的车辆三维结构将出现畸变,影响车辆下边沿估计的准确性。所以,为了降低误差,对  $C$  进行滤波校验。

在像素坐标系下,计算经过  $c$  和  $d$  2 点直线的斜率  $k_{cd}$ ,  $k_{cd}$  决定车辆下边沿在像素坐标系中的倾斜程度。根据驾驶场景中车辆的角度位姿先验,定义滤波函数  $Fsen(x)$  对其进行滤波,如式(9)所示:

$$Fsen(x) = \begin{cases} 0, & x > 0 \\ x, & -0.1 < x < 0 \\ -0.1, & x < -0.1 \end{cases} \quad (9)$$

通过式(9)完成对斜率  $k_{cd}$  的修正,其中  $x$  表示斜率。利用相机投射投影成像存在消失点的原理<sup>[30-33]</sup>,对  $f$ 、 $g$  和  $e$  进行修正,公式表示如式(10)~式(12)所示:

$$f(x, y) = \left( f(x), \frac{e^2}{8} \cdot e(y) \right) \quad (10)$$

$$g(x, y) = \left( \left(1 + \frac{8 - e^2}{8}\right) \cdot g(x), g(y) \right) \quad (11)$$

$$e(x, y) = \left( e(x), \left(1 + \frac{8 - e^2}{8}\right) \cdot e(y) \right) \quad (12)$$

完成上述过程之后,校验几何模型是否满足式(13)所示的图像中像素坐标关系:

$$\begin{cases} e(y) < h(y) < d(y) \\ b(y) < c(y) \leq d(y) \\ e(x) < d(x) < b(x) \\ e(x) < g(x) < b(x) \\ e(y) < g(y) \leq h(y) \end{cases} \quad (13)$$

令三维信息估计结果标志变量为  $t$ ,如果满足此条件,则  $t = 1$ ; 不满足,则  $t = 0$ 。最后通过估计到的车辆三维信息,根据式(14)提取车辆下边沿  $R(x_1, y_1, x_2, y_2)$ :

$$R = \begin{cases} (c(x), c(y), d(x), d(y)), & t = 1 \\ (B(x_1), B(y_1), B(x_2), B(y_2)), & t = 0 \end{cases} \quad (14)$$

## 4 逆透视变换模型

相机逆透视变换,目的是将图像像素坐标对应还原回物理世界坐标系中,得到图像的俯视图,如图 10 所示。相机成像的基本原理是将世界坐标系中的物理坐标点映射到像素坐标系中,过程如图 11a 所示,逆透视变换依据该原理,将像素坐标系中的像素坐标转换到世界坐标系中,为相机成像的逆过程,如图 11b 所示。

本文基于逆透视变换的测距模型的创新之处在于,利用前方车辆目标处在道路平面这一先验,结合标定原理,计算尺度缩放因子,然后进行基于点的逆透视变换实现测距,不同于利用全部像素逆透视变换之后的线性关系进行测距的方法。文献[29]采用的逆透视变换测距方法是将图像中道路信息还原成物理道路平面俯视图,利用该平面与真实道路平面具有的线性比例关系,确定车辆的距离。但是,该方法忽略了车辆的视角姿态,在具有线性比例的关系图中难以找到车辆的下边沿位置坐标,只能近似估计车辆的整体车身距离,并且在整体路面进行逆透视变换时需要缺失的像



素信息进行估计,从而增大了误差的产生。故本文方法先找到车辆下边沿,计算测距关键点,结合缩放因子,对该关键点进行逆透视变换,计算车辆距离,从而减小测距误差。

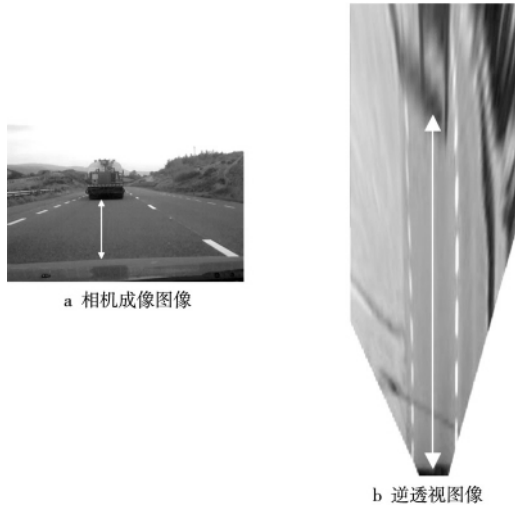


Figure 10 Inverse perspective mapping

图 10 逆透视变换

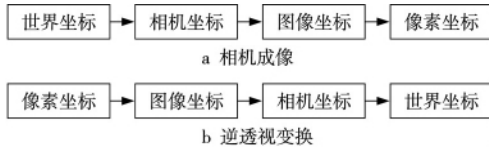


Figure 11 Coordinate transformation process

图 11 坐标变换过程

单目 RGB 相机在成像过程中,像素的深度距离信息会在相机坐标到图像坐标映射过程中丢失,如图 12 所示。在完成相机标定之后,得到相机外参旋转矩阵  $R_{3 \times 3}$  和平移矩阵  $T_{3 \times 1}$ 。外参的作用是将世界坐标系和相机坐标系对齐,经过分析发现,外参包含一个隐藏的物理参数值,即世界坐标系原点和相机坐标系原点之间的位移,利用该物理参数值,以及目标车辆和当前车辆处于同一道路水平面先验,可求出道路平面像素点所对应的尺度缩放因子,从而实现基于点的逆透视测距。

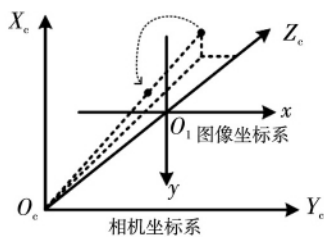


Figure 12 Camera coordinate system

and image coordinate system

图 12 相机坐标系和图像坐标系

逆透视测距详细过程如下所示:

令  $K_{3 \times 3}$  为相机内参矩阵(包含 5 个内参数),表示为:

$$K_{3 \times 3} = \begin{bmatrix} f/d_x & r & u_0 \\ 0 & f/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (15)$$

其中,  $f$  为相机焦距,  $d_x$  和  $d_y$  分别是相机横向和纵向单位像素的物理尺寸,  $(u_0, v_0)$  是图像坐标系原点  $O_i$  在像素坐标系中的像素位置,  $r$  是内参矩阵因子。

设世界坐标系中的坐标点  $(X_w, Y_w, Z_w)$ , 对应像素坐标系中的坐标点  $(u, v)$ , 由式(16)所示的成像矩阵方程可得到式(17):

$$Z_c \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f/d_x & r & u_0 \\ 0 & f/d_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ O & 1 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (16)$$

其中,  $O$  为  $1 \times 3$  的零矩阵。

$$Z_c \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_{3 \times 3} \cdot \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ O & 1 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (17)$$

令  $P_{3 \times 4} = K_{3 \times 3} \cdot \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ O & 1 \end{bmatrix}$ , 可得:

$$Z_c \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = P_{3 \times 4} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (18)$$

通过 4 组对应点和单应性矩阵原理,可得相机参数  $P_{3 \times 4}$ 。

由车辆下边沿信息  $R(x_1, y_1, x_2, y_2)$  和式(19)可得测距关键点像素坐标  $E(x, y)$ :

$$\begin{cases} x = \frac{R(x_1) + R(x_2)}{2} \\ y = \frac{R(y_1) + R(y_2)}{2} \end{cases} \quad (19)$$

令  $Z_c = s$ ,  $s$  即尺度缩放因子,由式(16)可得:

$$s \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_{3 \times 3} \cdot \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ O & 1 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (20)$$

$$s \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K_{3 \times 3} \cdot \begin{bmatrix} R_{3 \times 3} & T_{3 \times 1} \\ O & 1 \end{bmatrix} \cdot \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (21)$$

由车辆下边沿位置先验,得  $Z_w = 0$ , 将  $Z_w = 0, (u, v) = E(x, y)$  代入式(21), 可得  $s = s_0$ , 又由式(21)可得:

$$\begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = s \cdot R_{3 \times 3}^{-1} \cdot K_{3 \times 3}^{-1} \cdot \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} - R_{3 \times 3}^{-1} \cdot T_{3 \times 1} \quad (22)$$

将  $Z_w = 0, (u, v) = E(x, y), s = s_0$  代入式(22), 可得矩阵  $\begin{bmatrix} X_w \\ Y_w \\ 0 \end{bmatrix}$ ,  $(X_w, Y_w)$  即为车辆纵向和横向物理距离。

## 5 实验

为了验证本文的测距方法,在具有雷达标注的距离信息的 KITTI 数据集<sup>[34]</sup>上进行实验。为了从多个角度对本文的测距方法进行验证,本文进行2组对比实验。第1组是将本文方法与未采用车辆下边沿估计直接进行逆透视变换的测距方法进行对比,目的是验证车辆下边沿对测距本身产生的影响;第2组是将本文方法与其他测距方法对比,目的是验证本文方法的测量精度。

本文从 KITTI 官方的训练数据集中去除用于训练检测网络的 5 000 幅图像,剩余的 2 481 幅图像用于实验验证。本文只对 KITTI 数据集中标签为“Car”类别的目标进行验证,并且对像素宽度小于 100 的车辆目标不进行下边沿估计操作,直接进行距离计算。KITTI 数据集的标签格式如图 13 所示,每个 TXT 格式的标签文件代表对应文件名的图像中所有目标的信息,如图 13 中,每行数据代表一个实例目标,其中每行中的多个值使用空格分开。对本次实验中需要用到的标签数据进行如下说明:首先,每行中的第 1 个值为字符串类型,表示该目标的类别;第 5~8 位为 4 个浮点数,代表目标的 2D 边界框坐标;第 9~11 位为 3 个浮点数,代表目标的尺寸,分别是高、宽、长(单位为 m);第 12~14 位为 3 个浮点数,代表目标中心点所处的位

置坐标  $(x, y, z)$ , 该坐标建立在相机坐标系下(单位为 m),坐标轴以相机中心点为坐标原点,正右方为  $x$  轴正向,正下方为  $y$  轴正向,正前方为  $z$  轴正向。本次实验使用  $x$  和  $z$  作为实验验证的真值,分别代表目标车辆的横向和纵向距离。

本文实验使用的操作系统是 Ubuntu 16.04, 内存为 32 GB, GPU 为 NVIDIA GeForce GTX 1080Ti, 编程环境为 Python2.7, 深度学习框架为 PyTorch 0.4.0 版本。使用 PyTorch 实现的 Faster R-CNN 目标检测算法<sup>[11]</sup>进行前期的车辆检测,使用 SSD 目标检测算法<sup>[13]</sup>进行车辆关键组件的检测,并将检测阶段和测距阶段进行了融合。

针对驾驶场景中处于不同状态的车辆进行测距分析,如前方车辆转弯、上坡和遮挡。当前方车辆转弯时,本文方法计算的车距是前方车辆相对于当前车辆的纵向和横向的距离,如图 14a 所示,纵向距离  $x$ , 横向距离  $y$ , 该种情况和非转弯的车辆测距情况相同,不影响测距结果;当前方车辆在上坡而当前车辆在水平路面时,如图 14b 所示,上坡路面角度为  $\alpha$ , 此时,横向测距结果不受影响,纵向测距结果将发生变化,变为测量值的  $\sec \alpha$  倍,根据《公路工程技术标准》(JTJ B01-2003)规定:各级公路的最大纵坡不应大于  $3\% \sim 9\%$ , 即当最大纵坡取  $9\%$  时,  $\sec \alpha = 1.0040$ , 即误差变化比例为 0.0040, 车距为 100 m 时,误差增加 0.40 m;当前方车辆和当前车辆同时处于坡上时,如图 14b 所示,两者处于同一水平面,测距结果不发生变化;当前方车辆受到遮挡时,如图 14c 所示,车辆检测阶段会检测不到完整车辆,在车辆组件检测时,如果车辆组件检测结果不完整,则采用车辆矩形框下边中心点作为测距关键点进行车距测量,即相当于采用不进行车辆下边沿估计的方式进行测距,测距误差见 5.1 节。

本文旨在通过深度学习技术来提高车辆视觉测距的精度,所使用的车辆检测方法为 Faster R-CNN<sup>[11]</sup>, 该方法为通用的目标检测方法,同其他车距测量方法中需要首先获取车辆外接矩形框的目的相同,本文的创新之处在于基于车辆外接矩形框

```
Cyclist 0.89 0 -2.22 1056.16 190.50 1241.00 374.00 1.59 0.53 1.89 2.75 1.68 3.14 -1.55
Car 0.00 0 1.90 359.43 179.30 516.30 270.97 1.44 1.64 3.78 -3.03 1.57 13.30 1.68
Car 0.00 1 1.82 444.13 163.92 559.20 247.04 1.77 1.71 4.28 -2.52 1.61 17.60 1.68
Van 0.00 1 -0.47 870.59 115.22 1101.06 213.85 2.71 2.12 5.66 10.90 1.13 21.30 -0.01
Car 0.00 0 -1.72 724.49 171.09 790.50 215.13 1.45 1.71 3.93 5.18 1.42 25.97 -1.53
Car 0.00 1 1.74 534.25 166.20 596.08 216.76 1.73 1.65 3.96 -1.62 1.53 26.83 1.67
Car 0.00 2 -1.69 709.81 162.43 765.70 205.68 1.78 1.89 4.37 5.57 1.38 32.13 -1.52
Car 0.00 2 1.69 562.02 166.42 605.57 206.89 1.68 1.51 3.29 -1.14 1.43 31.78 1.66
```

Figure 13 KITTI dataset annotation information

图 13 KITTI 数据集标注信息

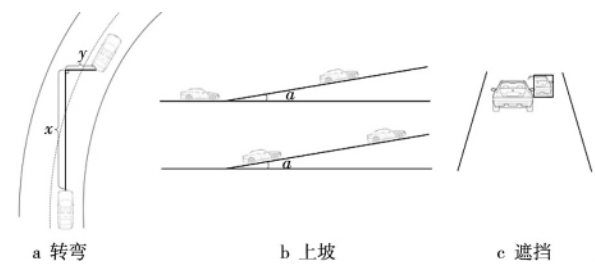


Figure 14 Vehicle distance measurement and analysis in different states

图 14 不同状态下车距测量分析

的检测结果,借助轻量快速的目标检测框架 SSD<sup>[13]</sup> 完成对车辆下边沿的估计。测距各阶段用时和算法处理帧率如表 1 所示。

Table 1 Processing time and processing frame rate for each stage

表 1 各阶段处理用时和处理帧率

	车辆检测 /ms	下边沿估计 /ms	测距计算 /ms	用时 /ms	帧率
其他方法	50	0	2	52	19.2
本文方法	50	18	1	69	14.5

5.1 车辆下边沿对测距误差影响验证

本文对不同纵向物理距离的车辆目标进行分阶段统计测试,总共分为 4 组进行实验,分别为 0~20 m,20~40 m,40~60 m,60~80 m。本次实验的计算规则是分别统计使用下边沿估计和未使用下边沿估计车距测量值与真值之间的差值,即绝对误差。

在实验各阶段,本文提取了符合测量标准的车辆实例,各距离阶段内车辆数量不同,取各阶段测距误差的平均值,实验结果见表 2。数量表示各阶段符合测试条件的车辆数,平均误差表示所有距离阶段的车辆测距误差的平均值。从实验结果中可看出,采用下边沿估计的测距精度明显高于未采用下边沿估计的测距精度,其次,纵向测距误差相对于横向测距更大,尤其在未采用车辆下边沿估计的测距方式中更为明显,而本文方法在横向测距的精度以及稳定性上表现更好。通过实验结果还可以发现,目标纵向距离越大测距误差就会越大,且横向距离和纵向距离的测量误差变化程度会有所不同。如图 15 所示,在车辆目标的纵向距离逐渐增大的过程中,横向测量误差逐渐增大,且变化非常明显,而纵向测量误差虽然也在增大,但相对稳定,增幅较小。从这里更可以看出,本文方法相对于未使用下边沿估计的方法具有更高的测距精度和稳定性。

Table 2 Influence comparison of ranging error in the lower edge estimation

表 2 下边沿估计测距误差影响对比

距离/m	数量	无下边沿/m		有下边沿/m	
		纵向	横向	纵向	横向
0~20	3 744	1.069 0	0.669 1	<b>0.184 3</b>	<b>0.186 1</b>
20~40	3 742	2.692 2	1.095 9	<b>0.555 8</b>	<b>0.200 9</b>
40~60	1 541	3.057 2	1.568 1	<b>0.800 8</b>	<b>0.233 6</b>
60~80	352	5.748 5	2.235 8	<b>1.128 4</b>	<b>0.285 3</b>
平均	9 379	2.218 9	1.045 9	<b>0.469 2</b>	<b>0.203 5</b>

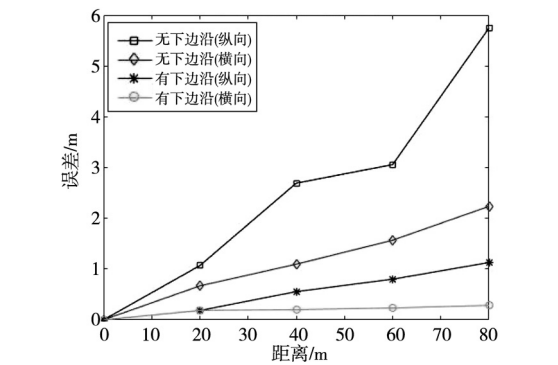


Figure 15 Variation of ranging error

图 15 测距误差变化

5.2 本文方法与其他测距方法对比验证

将本文方法和其他测距方法进行对比实验。选取文献[27,28,35,36]中的方法进行实验,同样分成 4 个不同的距离阶段,分别为 0~20 m,20~40 m,40~60 m,60~80 m,根据不同距离阶段内的车辆目标,分别计算测距误差的平均值,包括绝对误差和相对误差,实验结果如表 3 所示。文献[35]的测距方法只用于对车辆在纵向上的车距测量,该方法在近距离阶段(40 m 之内)测距精度较高,在远距离的 60~80 m 阶段,测距误差变大。文献[27,28,36]的方法可用于纵向和横向上的车距测量。从表 3 可以看出,随着车辆纵向距离的增大,这几种方法的测距误差也有较大程度的增大,而本文方法随着纵向距离的增大,测距误差并没有出现较大的变化,说明本文方法稳定性更好。无论是分阶段的对比还是平均误差的对比,本文方法明显优于其他几种方法,且测距的精度更高,尤其在横向距离测量上误差更小,表明本文方法针对相邻车道内的车辆目标更加有效。

6 结束语

本文针对基于单目视觉的车辆测距技术中忽



**Table 3 Comparison between this method and other ranging methods**

**表3 本文方法与其他测距方法对比**

方法	距离/m	绝对误差/m		相对误差/%	
		纵向	横向	纵向	横向
VDEU <sup>[27]</sup>	0~20	0.37	0.25	1.85	1.25
	20~40	0.28	0.29	2.77	2.92
	40~60	2.24	2.40	3.74	4.00
	60~80	3.56	3.81	4.45	4.76
	平均	1.61	1.69	3.20	3.23
RTMR <sup>[28]</sup>	0~20	0.42	0.25	2.10	1.25
	20~40	0.33	0.38	3.27	3.75
	40~60	2.69	2.80	4.48	4.67
	60~80	4.76	4.46	5.95	5.58
	平均	2.05	1.97	3.95	3.81
IPMD <sup>[35]</sup>	0~20	0.19	-	0.95	-
	20~40	0.11	-	1.05	-
	40~60	2.72	-	4.53	-
	60~80	5.07	-	6.34	-
	平均	2.02	-	3.22	-
VDMI <sup>[36]</sup>	0~20	0.31	0.08	1.57	0.42
	20~40	0.25	0.13	2.50	1.25
	40~60	1.67	1.00	2.79	1.67
	60~80	2.33	2.10	2.91	2.62
	平均	1.14	0.83	2.44	1.49
Ours	0~20	0.18	0.19	0.92	0.93
	20~40	0.14	0.05	1.39	0.50
	40~60	0.80	0.23	1.33	0.39
	60~80	1.13	0.29	1.41	0.36
	平均	0.56	0.19	1.26	0.54

略了车辆与地面相接的下边沿问题,提出了一种基于车辆下边沿估计和逆透视变换的单目视觉测距方法,实现了对前方车辆进行横向和纵向的高精度车距测量。相对于其他测距方法,本文通过下边沿估计模型实现对车辆下边沿的定位,确定更加准确的测距关键点,利用基于点的逆透视变换测距模型进行距离计算。实验结果表明,本文方法的测距精度更高,稳定性更好。在未来的工作中,我们将进一步探索新的车辆下边沿估计模型,在进行车辆下边沿估计时,降低因车辆组件过小而产生的影响,提高算法的鲁棒性,进一步提高车距测量的精度。

#### 参考文献:

[1] Sun Zhe-hang, Bebis G, Miller R. On-road vehicle detection: A review[J]. IEEE Transactions on Pattern Analysis and

Machine Intelligence, 2006, 28(5): 694-711.

- [2] Park K Y, Hwang S Y. Robust range estimation with a monocular camera for vision-based forward collision warning system[J]. The Scientific World Journal, 2014, 2014: 1-9.
- [3] Zhu Mei-xia. Research of collision avoidance mechanism for autonomous vehicles at the intersection[J]. Computer Engineering & Science, 2013, 35(9): 181-186. (in Chinese)
- [4] Ding Peng-heng, Zhang Jin-dong, Zhou Yang-hao, et al. Anti-collision warning algorithm based on visual perception in front of vehicle[C] // Proc of the 2017 International Conference on Mechanical, Electronic, Control and Automation Engineering (MECAE 2017), 2017: 51-55.
- [5] Lin H Y, Chen L Q, Lin Y H, et al. Lane departure and front collision warning using a single camera[C] // Proc of 2012 International Symposium on Intelligent Signal Processing and Communications Systems, 2012: 64-69.
- [6] Wu C F, Lin C J, Lin H Y, et al. Adjacent lane detection and lateral vehicle distance measurement using vision-based neuro-fuzzy approaches[J]. Journal of Applied Research and Technology, 2013, 11(2): 251-258.
- [7] Mao Xue-song, Inoue D, Matsubara H. Demonstration of in-car Doppler laser radar at 1.55  $\mu\text{m}$  for range and speed measurement[J]. IEEE Transactions on Intelligent Transportation Systems, 2013, 14(2): 599-607.
- [8] Folster F, Rohling H. Data association and tracking for automotive radar networks[J]. IEEE Transactions on Intelligent Transportation Systems, 2005, 6(4): 370-377.
- [9] Mao Xue-song, Inoue D, Kato S, et al. Amplitude-modulated laser radar for range and speed measurement in car applications[J]. IEEE Transactions on Intelligent Transportation Systems, 2012, 13(1): 408-413.
- [10] Girshick R. Fast R-CNN[C] // Proc of 2015 IEEE International Conference on Computer Vision (ICCV), 2015: 1440-1448.
- [11] Ren Shao-qing, He Ke-ming, Girshick R, et al. Faster R-CNN: Towards real-time object detection with region proposal net-works[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.
- [12] Redmon J, Divvala S, Girshick R, et al. You only look once: Unified, real-time object detection[C] // Proc of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016: 779-788.
- [13] Liu Wei, Anguelov D, Erhan D, et al. SSD: Single shot multi-box detector[C] // Proc of European Conference on Computer Vision (ECCV 2016), 2016: 21-37.
- [14] Lin T Y, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C] // Proc of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 2117-2125.
- [15] Geng Lei, Peng Xiao-shuai, Xiao Zhi-tao, et al. Preceding vehicles detection based on integrated learning and constraint of positions information[J]. Computer Engineering & Science, 2018, 40(10): 1844-1850. (in Chinese)

- [16] Li Xi-ying, Zhou Zhi-hao, Lv Shuo. A vehicle face detection algorithm based on selective search[J]. Computer Engineering & Science, 2018, 40(10): 1829-1836. (in Chinese)
- [17] Lim B, Woo T, Kim H. Integration of vehicle detection and distance estimation using stereo vision for real-time AEB System[C] // Proc of the 3rd International Conference on Vehicle Technology and Intelligent Transport Systems, 2017: 211-216.
- [18] Lv Xue-zhi, Wang Mei-ting, Qi Yong-feng, et al. Research on ranging method based on binocular stereo vision[J]. Advanced Materials Research, 2014, 945-949: 2075-2081.
- [19] Long Hui-min, Guo Hai-yan, Liang Feng, et al. Distance measurement algorithm based on binocular stereo vision[J]. Applied Mechanics and Materials, 2014, 635-637: 948-952.
- [20] Zou Bin, Yuan Yu-xiang. High precision distance measurement based on monocular vision for intelligent traffic[J]. Journal of Transportation Systems Engineering and Information Technology, 2018, 18(4): 46-53. (in Chinese)
- [21] Guan Chuang, Wei Lang, Qiao Jie, et al. A vehicle distance measurement method with monocular vision based on vanishing point[J]. Electronic Measurement Technology, 2018, 41(11): 83-87. (in Chinese)
- [22] Yu Gui-zhen, Zhou Bin, Wang Yun-peng, et al. Measuring algorithm for the distance to a preceding vehicle on curve road using on-board monocular camera[J]. International Journal of Bifurcation and Chaos, 2015, 25(14): 1540038.
- [23] Ho H W, de Croon G C, Chu Qi-ping. Distance and velocity estimation using optical flow from a monocular camera[J]. International Journal of Micro Air Vehicles, 2017, 9(3): 198-208.
- [24] Wu Bing-fei, Chen Ying-han, Kao C C, et al. A vision-based collision warning system by surrounding vehicles detection[J]. KSII Transactions on Internet and Information Systems, 2012, 6(4): 1203-1222.
- [25] Wu Jun, Li Wen-jie, Geng Lei, et al. Preceding vehicle detection and ranging based on monocular vision[J]. Computer Engineering, 2017, 43(2): 26-32.
- [26] Bao Dong-shang, Wang Pei-kang. Vehicle distance detection based on monocular vision[C] // Proc of 2016 International Conference on Progress in Informatics and Computing (PIC), 2016: 187-191.
- [27] Han J, Heo O, Park M, et al. Vehicle distance estimation using a mono-camera for FCW/AEB systems[J]. International Journal of Automotive Technology, 2016, 17(3): 483-491.
- [28] Nakamura K, Ishigaki K, Ogata T, et al. Real-time monocular ranging by Bayesian triangulation[C] // Proc of 2013 Intelligent Vehicles Symposium, 2013: 1368-1373.
- [29] Tuohy S, O'Cualain D, Jones E, et al. Distance determination for an automobile environment using inverse perspective mapping in OpenCV[C] // Proc of IET Irish Signals and Systems Conference, 2010: 100-105.
- [30] Moghadam P, Starzyk J A, Wijesoma W S. Fast vanishing-point detection in unstructured environments[J]. IEEE Transactions on Image Processing, 2012, 21(1): 425-430.
- [31] Kong Hui, Audibert J Y, Ponce J. Vanishing point detection for road detection[C] // Proc of 2009 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009: 96-103.
- [32] Nguyen L, Phung S L, Bouzerdoum A. Enhanced pixel-wise voting for image vanishing point detection in road scenes[C] // Proc of 2017 IEEE International Conference on Acoustics, Speech and Signal Processing, 2017: 1852-1856.
- [33] Shi Jin-jing, Wang Jin-xiang, Fu Fang-fa. Fast and robust vanishing point detection for unstructured road following[J]. IEEE Transactions on Intelligent Transportation Systems, 2016, 17(4): 970-979.
- [34] Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite[C] // Proc of 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012: 3354-3361.
- [35] Adamshuk R, Carvalho D, Neme J H Z, et al. On the applicability of inverse perspective mapping for the forward distance estimation based on the HSV colormap[C] // Proc of 2017 IEEE International Conference on Industrial Technology (ICIT), 2017: 1036-1041.
- [36] Liu Jun, Hou Shi-hao, Zhang Kai, et al. Vehicle distance measurement with implementation of vehicle attitude angle estimation and inverse perspective mapping based on monocular vision[J]. Transactions of the Chinese Society of Agricultural Engineering, 2018, 34(13): 70-76. (in Chinese)

#### 附中参考文献:

- [3] 朱梅霞. 无人驾驶车辆的路口防撞机制研究[J]. 计算机工程与科学, 2013, 35(9): 181-186.
- [15] 耿磊, 彭晓帅, 肖志涛, 等. 基于集成学习与位置信息约束的前方车辆检测[J]. 计算机工程与科学, 2018, 40(10): 1844-1850.
- [16] 李熙莹, 周智豪, 吕硕. 基于选择性搜索算法的车脸部件检测[J]. 计算机工程与科学, 2018, 40(10): 1829-1836.
- [20] 邹斌, 袁宇翔. 面向智能交通的单目视觉测距方法研究[J]. 交通运输系统工程与信息, 2018, 18(4): 46-53.
- [21] 关闯, 魏朗, 乔洁, 等. 一种基于消隐点的单目视觉车辆测距方法[J]. 电子测量技术, 2018, 41(11): 83-87.
- [25] 吴骏, 李文杰, 耿磊, 等. 基于单目视觉的前方车辆检测与测距[J]. 计算机工程, 2017, 43(2): 26-32.
- [36] 刘军, 后士浩, 张凯, 等. 基于单目视觉车辆姿态角估计和逆透视变换的车距测量[J]. 农业工程学报, 2018, 34(13): 70-76.

#### 作者简介:



王永森(1994-), 男, 河南郸城人, 硕士生, 研究方向为图像处理和深度学习。  
E-mail: 861935973@qq.com

WANG Yong-sen, born in 1994, MS candidate, his research interests include image processing, and deep learning.