

基于注意力机制的非平坦路面单目车距估计方法研究

刘永涛, 李怡飞, 高隆鑫, 陈轶嵩, 王泰琪

(长安大学汽车学院, 西安 710064)

摘要: 提出一种基于注意力机制的单目车距估计算法, 以提高非平坦路面下的车距估计精度。通过将通道和空间注意力引入ImVoxelNet神经网络, 增强卷积层对车辆轮廓感知和特征区分能力, 有效减少车辆漏检现象; 基于感兴趣区域角点标定, 剔除逆透视变换时的冗余信息, 改善了图像畸变问题; 针对车辆姿态变化, 提出了考虑姿态干扰的相机外参矩阵, 建立了非平坦路面下的相机坐标转换模型; 利用真实与逆透视图像的比例关系构建车距估计模型, 实现对前车纵、横向距离准确估算。试验表明, 本文方法在非平坦路面条件下, 纵向80 m和横向4 m的间距范围内测距相对误差小于3%, 验证了所提方法的有效性和准确性。

关键词: 3D目标检测; 逆透视变换; 测距; 单目视觉

中图分类号: TP391.4 文献标志码: A DOI: 10.3969/j.issn.2095-1469.2025.03.08

An Attention Mechanism-Based Monocular Vehicle Distance Estimation Method for Uneven Road Surfaces

LIU Yongtao, LI Yifei, GAO Longxing, CHEN Yisong, WANG Taiqi

(School of Automobile, Chang'an University, Xi'an 710064, China)

Abstract: A monocular vehicle distance estimation based on attention mechanisms is proposed to improve estimation accuracy on uneven road surfaces. Channel and spatial attention mechanisms are incorporated into the ImVoxelNet neural network to enhance contour perception and feature discrimination, thereby reducing missed vehicle detections. Redundant information in inverse perspective mapping is eliminated through region-of-interest corner calibration, mitigating image distortion. To address variations in vehicle pose, a pose-interference-aware camera extrinsic parameter matrix is proposed, and a coordinate transformation model for uneven surfaces is established. Finally, the proportional relationship between real-world and inverse perspective images is used to construct a distance estimation model, achieving accurate estimation of the longitudinal and lateral distances. Experimental results show that the proposed method maintains a relative error below 3% within a longitudinal range of 80 meters and a lateral range of 4 meters, validating its effectiveness and accuracy.

收稿日期: 2024-11-01 改稿日期: 2024-12-12 网络首发日期: 2025-02-20

基金项目: 国家重点研发计划项目(2021YFB2501202); 陕西省“两链”融合重点专项揭榜挂帅项目(2023JBGS-13); “车路一体智能交通全国重点实验室”开放基金(2024-B005); 长安大学中央高校基本科研业务费专项资金资助项目(300102223204)

参考文献引用格式:

刘永涛, 李怡飞, 高隆鑫, 等. 基于注意力机制的非平坦路面单目车距估计方法研究[J]. 汽车工程学报, 2025, 15(3): 353-365.

LIU Yongtao, LI Yifei, GAO Longxing, et al. An Attention Mechanism-Based Monocular Vehicle Distance Estimation Method for Uneven Road Surfaces[J]. Chinese Journal of Automotive Engineering, 2025, 15(3): 353-365. (in Chinese)



Keywords: 3D object detection; inversion perspective mapping; distance measurement; monocular vision-based algorithms

自动驾驶汽车在一定程度上能协助驾驶员做出行为决策, 为用户提供更安全的驾驶服务, 减少因人为操作不当带来的风险。视觉感知技术通过获取障碍物、标志、车道线等信息, 对图像特征进行分析和处理, 为自动驾驶汽车的安全、规范行驶提供信息支持。因此, 检测并精确定位前方目标车辆, 将大大提高自动驾驶系统的决策精度, 对提升自动驾驶汽车的行驶安全性具有重要意义。

目前, 基于纯视觉的车辆距离估计方法主要分为单目视觉和双目视觉 2 种。其中, 双目视觉测距方法依赖大量的计算资源, 导致响应时间长, 并且需要摄像头之间的参数同步校准, 而车辆行驶过程中的震动等因素可能会导致校准出现偏差, 影响深度测量的精确性。而单目视觉测距结构简单, 响应速度快, 避免了传感器之间的校准误差, 在自动驾驶辅助系统中有广阔的应用前景^[1]。

近年来, 国内外学者在车辆单目测距方面的研究集中分为 2 类: 基于 2D 目标检测的测距方法和基于 3D 目标检测的测距方法。其中, 2D 目标测距方法又分为基于数据回归建模、基于成像几何模型、基于逆透视变换 3 种方法^[2]。对于数据回归建模的方法, 沈志熙等^[3]应用了从数据推断模型的逆向思路, 提出了用已有距离样本点进行数据回归的方法拟合测距模型, 降低了成像时畸变带来的影响, 但仅考虑平直路面的情况。HUANG Dengyuan 等^[4]利用道路消失点和边缘线对道路区域进行分割, 利用阴影区域特征回归的方法检测出车辆位置, 实现了车辆距离的估计, 但受光照的影响较大。BUI 等^[5]根据从相机光轴上 2 个不同位置图像中提取的信息, 提出了基于位置的距离回归模型, 计算出目标物体的尺寸、距离和方向, 该方法依赖光轴不同位置的图像信息, 在车辆运动场景下无法达到很好的检测效果。张欣等^[6]将车道线在连续帧中的位置变化作为回归变量, 提出了基于该回归变量的障碍物距离估计模型, 该方法依赖车道线的定位, 无

法实现在车道线破损或缺失的道路环境下进行车距估计。尽管数据回归的方法能估计动态物体距离, 但是依赖大量的前期数据, 数据的质量和数量直接影响了模型的检测效果, 模型的泛用性不强。

而基于成像几何模型的方法具有适应性强, 计算量小的优点, 王荣本等^[7]通过 Hough 变换确定车道线, 以车道线为基准选取感兴趣区域, 使用几何关系变换方法估计车距, 但由于基于灰度梯度的方法检测障碍物, 因此光照角度和强度对检测结果有较大影响。MAO Jiafa 等^[8]根据相机成像几何模型推导出目标距离的计算公式, 并通过模数转换的原理, 建立了距离、视场、焦距和分辨率之间的关系, 但该方法只能估计静态物体的距离, 无法适用于车辆测距的动态场景。ZHANG Zhisheng 等^[9]提出了基于投影点与图像对应点映射关系的绝对定位估计方法, 利用基于几何投影点之间的关系, 实现了摄像机与目标之间的绝对距离测定, 该方法依赖目标投影点位置的选择, 对动态物体的测距效果较差。因此, 成像几何模型的方法存在相机标定复杂、不适用于非结构化道路等问题。

基于逆透视变换方法避免了数据回归方法和成像几何模型方法的不足, 具备结构简单并泛用性强的优点。ZHANG Daiming 等^[10]提出了基于道路消失点的逆透视变换方法, 建立透视图与实际道路图像的纵向比例系数标定, 并优化了上下坡路段的测量精度, 但是该方法道路消失点的选取对检测精度有较大影响; 王永森等^[11]提出了利用 2D 目标识别检测框的下边沿和逆透视变换算法, 通过检测框的关键点信息测量车距信息, 但是 2D 检测框下边沿与真实的尾部存在位置差异, 影响距离估计精度。基于单目 2D 检测算法来估计车辆的前后距离能大致反映目标车辆的真实情况。然而, 2D 目标检测主要用于检测缺失或存在信息, 缺少 3D 目标检测所具有的形状特征, 导致车辆的预测位置与实际位置存在较大偏差。目前, 单目 3D 目标测距算法的

研究尚处于发展阶段，ZHE Ting等^[12]通过3D目标检测方法获得车辆俯视图的实际面积和图像中相应的投影面积，建立了面积-距离几何回归模型估计车距。MARTINEZ-DIAZ等^[13]使用基于区域的卷积神经网络结合角点检测器来检测兴趣点，解决了在连续图像中估计移动对象距离时累积误差的问题。虽然以上工作在水平整洁的平坦道路上具有良好的车距估计效果，但在应对丘陵、减速带、坑洼路面等非平坦路面造成的摄像头角度参数变化时，会对真实与逆透视图像间的比例造成干扰，导致车距估计模型误差增大。

综上所述，目前针对非平坦路面下的单目3D测距，存在着逆透视变换畸变大、非平坦路面造成相机参数摄动等问题，导致准确估计前车距离面临挑战。针对以上问题，本文通过引入注意力机制，提出了增强特征感知的ImVoxelNet单目车距检测算法，并建立了非平坦路面下考虑车辆姿态变化的相机外参矩阵和坐标转换模型。通过对逆透视变换区域图像标定和剔除冗余信息，降低了图像畸变程度，提升了车距估计精度。本文主要贡献有如下3点。

1) 通过将通道注意力与空间注意力模块串行，并与SE注意力并行，提出了融合注意力机制的ImVoxelNet神经网络，与现有文献[14]和[15]相比，增强了模型对图像目标的特征感知和区分能力。

2) 对比传统逆透视变换方法^[16-17]，本文通过对逆透视变换感兴趣区域标定，剔除图像变换冗余

信息，建立了相机成像输入与真实世界坐标之间的数学映射关系，有效减少了图像变换过程中的几何畸变。

3) 考虑了车辆姿态角变化引起的相机外参矩阵变化，建立了非平坦路面下相机坐标转换模型，降低了相机外参矩阵摄动对车距估计的影响。

1 基于注意力机制的单目视觉车辆3D目标检测

使用视觉图像识别车距时，需要找到一个基准点进行距离标定。3D目标检测方法能准确获取车辆方向、位置和体积等三维信息，准确体现其外部几何形状特征。传统ImVoxelNet神经网络支持单目图像处理，可以将2D图像信息转换为3D体素，实现对室外场景的车辆目标检测^[15]。然而，传统ImVoxelNet算法对障碍物部分区域的目标识别效果较差，存在车辆漏检、错检和检测框大小并不贴近实际尺寸的问题，对后续车距估计精度产生很大影响。

注意力机制能自适应优化图像特征，提高分类和检测性能，在避免漏检情况下提高检测精度^[18]。因此，本文通过将注意力机制引入ImVoxelNet卷积神经网络中，以增强对目标特征的感知能力，提高车辆检测精度。本文提出的基于注意力机制的ImVoxelNet目标检测算法如图1所示，主要由4个核心模块组成，分别为注意力机制、三维体素构建模块、三维特征提取模块和三维检测头。

通道和空间注意力分别代表图像整体和局部最

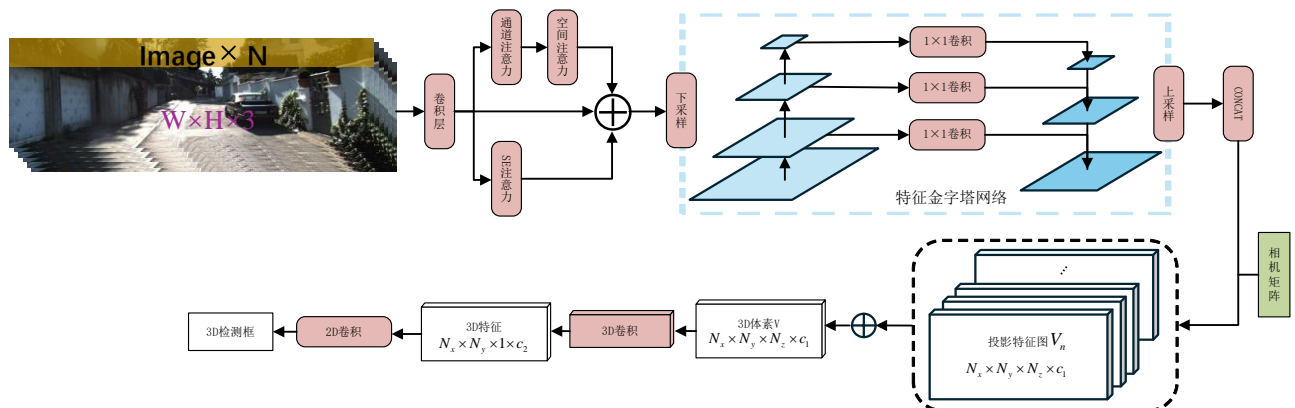


图1 基于注意力机制的3D目标检测 ImVoxelNet算法框架

具有价值的信息，对特征图从通道和空间 2 个角度提取关键信息更能提升检测准确性。图 2a 所示为通道注意力模块，该模块在二维卷积提取特征图后再对特征图提取通道注意力，将特征图分别通过最大池化层和平均池化层以聚合特征图上的空间信息并输入至全连接层得到压缩后的一维向量，该向量包含 2 种池化后特征信息，通过加和能提升目标检测的特征感知能力。将带有通道注意力信息的一维向量输入 Sigmoid 激活函数，如式 (1) 所示。

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \quad (1)$$

式中： F 为特征图；MLP 为全连接感知层； σ 为 Sigmoid 激活函数； M_c 为通道注意力。

图 2b 为空间注意力模块。含有通道注意力的

特征图分别经过平均池化层和最大池化层，将 2 个池化结果拼接为 1 个双通道图像，并进行 7×7 的卷积操作。通过 Sigmoid 激活函数得到空间注意力权重，计算如式 (2) 所示。

$$M_s(F) = \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \quad (2)$$

式中： M_s 为空间注意力； $f^{7 \times 7}$ 为卷积核为 7×7 的卷积操作。

利用 SE (Squeeze and Excitation) 注意力模块的并行运算结构将全局空间信息压缩成 1 个通道向量，捕获了通道特征的全局分布，利用非线性激活函数获取通道特征权重，提高了网络对不同通道特征重要度的敏感性和特征依赖关系。

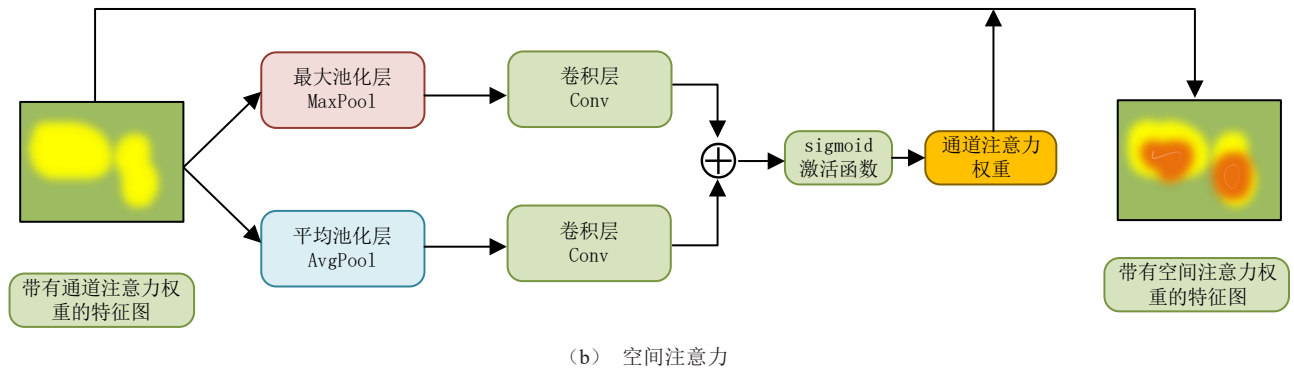
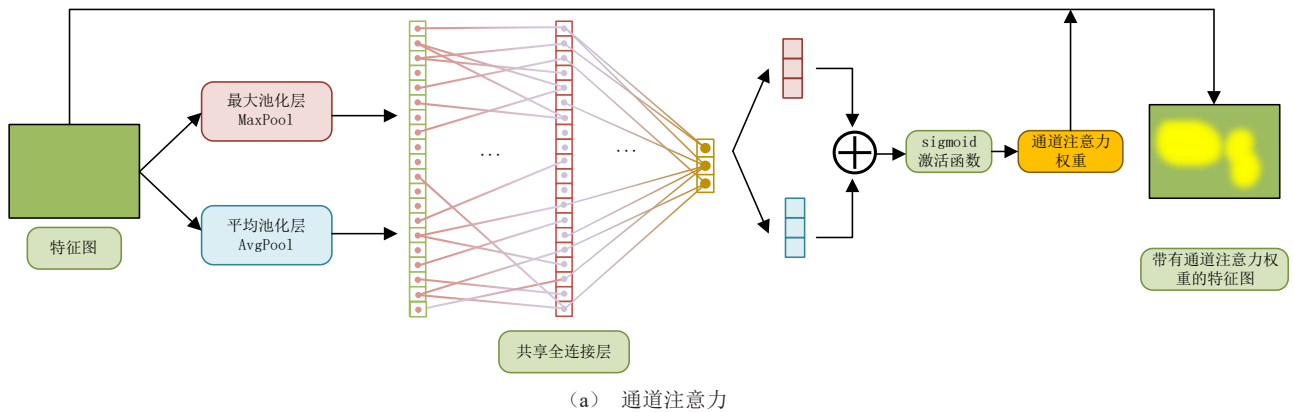


图 2 注意力机制

在三维体素构建阶段，如图 1 所示，首先通过卷积神经网络和特征金字塔网络对单目图像进行二维特征信息提取。然后，将二维特征投影到三维体素空间 V_n 上，并利用针孔相机模型确定二维特征与

三维体素之间的关联关系，如式 (3) 所示。

$$\begin{bmatrix} u \\ v \end{bmatrix} = \Pi \begin{bmatrix} 1/4 & 0 & 0 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{K} \mathbf{R}_n \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \quad (3)$$

式中：二维坐标 (u, v) 为三维空间点 (x, y, z) 在图像平面上的像素坐标； \mathbf{H} 为图像透视映射矩阵； \mathbf{K} 为相机内在矩阵； \mathbf{R}_n 为相机外在矩阵。

式 (3) 构建出平面图像的三维估计体积，将二维目标信息转化为三维信息。三维特征提取通过三维卷积和下采样处理，将得到的三维信息转化为包含三维信息的二维特征图。三维检测头将输入的二维特征张量通过 2D BEV 锚点返回 3D 检测框的 7 个参数，从而实现精确的三维目标检测。其中，位置参数 (x, y, z) 的变化量计算式为：

$$\begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix} = \begin{bmatrix} \frac{x^{\text{gt}} - x^{\text{a}}}{d^{\text{a}}} \\ \frac{y^{\text{gt}} - y^{\text{a}}}{d^{\text{a}}} \\ \frac{z^{\text{gt}} - z^{\text{a}}}{d^{\text{a}}} \end{bmatrix} \quad (4)$$

尺寸参数 (l, w, h) 的变化量计算式为：

$$\begin{bmatrix} \Delta l \\ \Delta w \\ \Delta h \end{bmatrix} = \begin{bmatrix} \log \frac{l^{\text{gt}}}{l^{\text{a}}} \\ \log \frac{w^{\text{gt}}}{w^{\text{a}}} \\ \log \frac{h^{\text{gt}}}{h^{\text{a}}} \end{bmatrix} \quad (5)$$

偏航角 θ 的变化量计算式为：

$$\Delta \theta = \sin(\theta^{\text{gt}} - \theta^{\text{a}}) \quad (6)$$

式中：gt 为真值；a 为锚框。

边界框对角线长度计算式为：

$$d^{\text{a}} = \sqrt{(w^{\text{a}})^2 + (l^{\text{a}})^2} \quad (7)$$

2 基于图像标定的逆透视变换方法

通过逆透视变换，二维图像可以在一定程度上恢复世界坐标系中的长宽比例，为车距估算提供基础。传统逆透视变换方法实现了图像坐标系与世界坐标系的坐标转换，但是存在水平线扭曲以及消失点处的图像畸变问题^[19-20]。这是由于在转换过程中，图像某些部分会被拉伸，其他部分可能会被压缩，导致图像产生畸变。为解决这一问题，本文提出基于图像标定的逆透视变换方法。该方法受文献 [20] 中的图像四边形角点计算方法启发，通过选定路面感兴趣区域四边形角点，剔除图像中的冗余

区域，以改善无关图像信息造成的逆透视变换图像畸变问题。

2.1 逆透视变换标定区域选取

逆透视变换时将图像坐标系对应到三维世界坐标系 xy 平面。记 $W = \{(x, y, z)\}$ 为三维世界坐标系，相机在空间中的位置参数如图 3 所示。

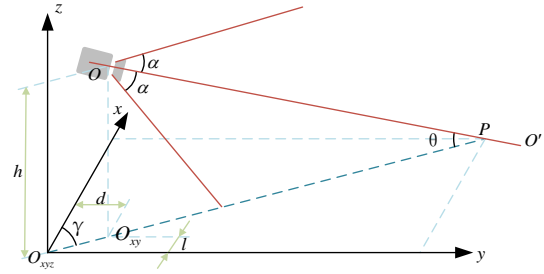


图 3 相机位置参数

图中， O_{xyz} 为空间坐标系 W 的原点， O 为相机所处空间位置， OO' 为相机光轴， O_{xy} 为相机在 xy 平面上的投影， $(l, d, h) \in W$ 为相机位置相对于 x 轴、 y 轴和 z 轴的空间坐标， γ 为相机的航向角， θ 为相机的俯仰角， 2α 为相机孔径。

记 $I_0 = \{(u_0, v_0)\}$ 为相机图像坐标系， $S = \{(x, y, 0)\}$ 为 W 映射到 $z = 0$ 平面的二维图像。根据式 (8)，可将图像坐标系转化为世界坐标系的平面二维形式。

$$\begin{cases} x(u_0, v_0) = h \times \cot \left[(\theta - \alpha) + \frac{2au_0}{m-1} \right] \times \\ \quad \cos \left[(\gamma - \alpha) + \frac{2av_0}{n-1} \right] + l, \\ y(u_0, v_0) = h \times \cot \left[(\theta - \alpha) + \frac{2au_0}{m-1} \right] \times \\ \quad \sin \left[(\gamma - \alpha) + \frac{2av_0}{n-1} \right] + d, \\ z(u, v) = 0. \end{cases} \quad (8)$$

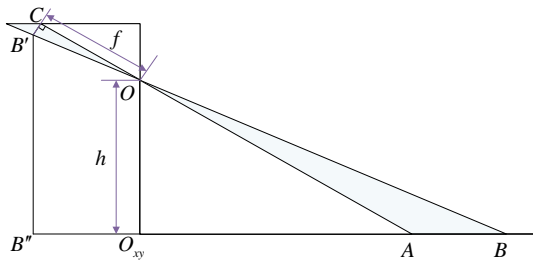
式中：地面位置为 $z = 0$ ； $m \times n$ 为相机的逆透视变换区域分辨率。

进一步，根据式 (9) 和式 (10) 逆变换去除图像的透视效果，转化成正常比例的鸟瞰图图像坐标系 $I_1 = \{(u_1, v_1)\}$ 。

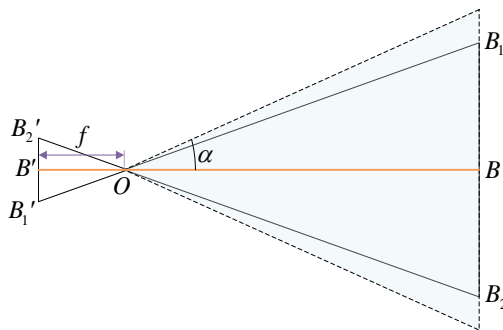
$$u_1(x, y, 0) = \frac{m-1}{2\alpha} \left(\arctan \left\{ h \sin \left(\arctan \left(\frac{y-d}{x-l} \right) / (y-d) \right) \right\} - \frac{m-1}{2\alpha} (\theta - \alpha) \right). \quad (9)$$

$$v_1(x, y, 0) = \frac{n-1}{2\alpha} \left(\arctan \left(\frac{y-d}{x-l} \right) - (\gamma - \alpha) \right). \quad (10)$$

根据路面感兴趣区域四边形角点计算方法^[20], 本文提出基于选取点对的图像变换区域标定方法, 感兴趣区域角点坐标计算原理如图 4a、b 所示。其中, B_1 和 B_2 是要选取的目标点对, B_1' 和 B_2' 是目标点对经过小孔成像后的映射位置, A 点是相机成像视野的下边缘, 虚线区域表示相机孔径范围。其中, OB 的距离计算如式(11)所示。



(a) xz 平面



(b) xy 平面

图 4 感兴趣区域角点计算

$$OB = \sqrt{\left(H \tan \left(\frac{\pi - \theta}{2} - \alpha \right) + L \right)^2 + H^2}. \quad (11)$$

根据几何关系, $\angle AOB$ 可以表示为:

$$\angle AOB = \sin^{-1} \left(\frac{AB \sin(\angle OAB)}{OB} \right) = \sin^{-1} \left(\frac{L \sin \left(\pi - \alpha - \frac{\theta}{2} \right)}{\sqrt{\left(H \tan \left(\frac{\pi - \theta}{2} - \alpha \right) + L \right)^2 + H^2}} \right). \quad (12)$$

$\angle OB'C$ 的值计算式为:

$$\angle OB'C = \frac{\pi + \theta}{2} - \sin^{-1} \left(\frac{L \sin \left(\pi - \alpha - \frac{\theta}{2} \right)}{\sqrt{\left(H \tan \left(\frac{\pi - \theta}{2} - \alpha \right) + L \right)^2 + H^2}} \right). \quad (13)$$

CB' 可以由三角几何关系通过式(14)得出。

$$CB' = \frac{OC \sin(\angle AOB)}{\sin(\angle OB'C)}. \quad (14)$$

CB' 为 B_1 和 B_2 在映射图像上的纵坐标值, 其在图像上的横坐标值 $B_1'B'$ 和 $B_2'B'$ 由相似三角形原理可计算为:

$$B_1'B' = B_2'B' = \frac{OB' \times BB_1}{OB}. \quad (15)$$

通过上述计算, 可得到现实中特定点对 B_1B_2 在图像中映射的对应位置 $B_1'B_2'$ 。类似地, 可以确定世界坐标下两对能组成长方形的特定点对, 并转化为图像上的 4 个角点坐标。

2.2 非平坦路面下相机外部参数校准

非平坦路面下车辆不同自由度姿态变化会造成车载摄像头俯仰角 θ 、偏航角 γ 、侧倾角 φ 和摄像头高度 h 的变化^[21]。如果相机外参矩阵不能实时反映车辆姿态角变化, 逆透视变换时图像会出现畸变, 影响最终距离估计精度。俯仰角 θ 和偏航角 γ 可以通过外部摄像机校准估计得到, 并通过平移得到图像水平线与真实世界水平面的旋转关系。而车辆侧倾角和俯仰角使摄像头高度发生变化, 不仅会导致图像与真实世界角度偏转, 并会使标定区域发生偏移。如图 5a 所示, 相机坐标系 O_{xy} 中 y 轴正方向垂直于道路平面, φ 表示车辆侧倾角, h 表示摄像头在车辆上的安装高度, H_φ 表示摄像头相对地面在 O_{xyz} 坐标系下的高度, W 表示车辆外侧轮距并假设摄像头安装在车辆横向中间位置。根据式(16)可计算得到存在侧倾角时摄像头的离地高度为:

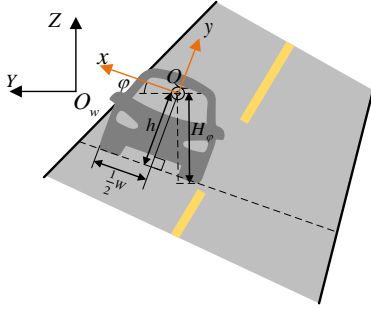
$$H_\varphi = \frac{h}{\cos \varphi} + \left(\frac{1}{2} W - h \tan \varphi \right) \sin \varphi. \quad (16)$$

由侧倾角引起的高度变化量为:

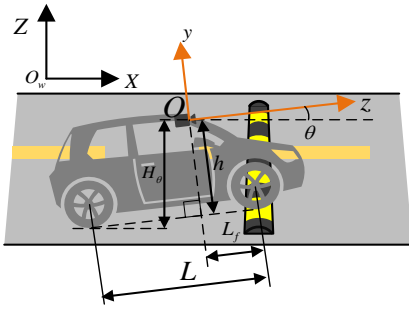
$$\Delta H_\varphi = H_\varphi - h. \quad (17)$$

根据式(18)可计算得到世界坐标系下沿Y轴方向上的位移变化量为:

$$\Delta Y = \frac{1}{2}W(1 - \cos \varphi) \quad (18)$$



(a) 侧倾角对摄像头高度影响



(b) 俯仰角对摄像头高度影响

图5 侧倾角和俯仰角对摄像头高度影响

图5b中,坐标系Oyz中z轴正方向指向摄像头光轴方向,θ表示车辆的俯仰角,H_θ表示摄像头距离地面的实际高度,L表示车辆前后轴距,L_r表示相机到车辆前轮中心的纵向距离。根据式(19)可计算得到存在侧倾角时摄像头的离地高度为:

$$H_{\theta} = \frac{h}{\cos \theta} + [L - (h \tan \theta + L_r)] \sin \theta \quad (19)$$

由俯仰角引起的高度变化量为:

$$\Delta H_{\theta} = H_{\theta} - h \quad (20)$$

根据式(21)计算得到世界坐标系下沿X轴方向上的变化量为:

$$\Delta X = (L - L_r)(\cos \theta - 1) \quad (21)$$

车辆在侧倾角和俯仰角影响下摄像头的总高度变化量如式(22)所示。

$$\Delta Z = \Delta H_{\varphi} + \Delta H_{\theta} \quad (22)$$

通过车辆姿态传感器可获取外部参数θ、γ和φ,对应的旋转矩阵R(θ, γ, φ)如式(23)所示。

$$R(\theta, \gamma, \varphi) = [r_{i,j}]_{3 \times 3} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} \cos \gamma & 0 & \sin \gamma \\ 0 & 1 & 0 \\ -\sin \gamma & 0 & \cos \gamma \end{bmatrix} \begin{bmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (23)$$

变化后位置向量T如式(24)所示。

$$T = [T'_1, T'_2, T'_3]^T = [T_1, T_2, T_3]^T + [\Delta X, \Delta Y, \Delta Z]^T \quad (24)$$

式中:向量[T₁, T₂, T₃]^T为原位置向量。

因此,相机外参矩阵E为:

$$E = \begin{bmatrix} r_{1,1} & r_{1,2} & r_{1,3} & T'_1 \\ r_{2,1} & r_{2,2} & r_{2,3} & T'_2 \\ r_{3,1} & r_{3,2} & r_{3,3} & T'_3 \end{bmatrix} \quad (25)$$

相机内参矩阵I为:

$$I = \begin{bmatrix} f_w/x_s & 0 & w/2 \\ 0 & f_h/y_s & h/2 \\ 0 & 0 & 1 \end{bmatrix} \quad (26)$$

式中:w、h为图像的宽和高;x_s、y_s为对应世界坐标系下x轴和y轴的尺寸参数;f_w和f_h为水平和垂直焦距。

由此,可得到由外部参数估计摄像头与真实世界坐标对应关系的单应性矩阵T为:

$$T = I \times E \times \eta \quad (27)$$

式中:η=[X, Y, Z, 1]^T为相机矩阵。

3 车辆测距模型

假设矩形ABCD为由第2节所提基于选取点对的图像标定区域,如图6a、b所示,其中,矩形abcd为车辆3D检测框的底面,OO'是照相机光轴。图6a中,d_y为AB与车辆检测框cd线的纵向距离。记标定区域ABCD经逆透视变换后的图像区域为A'B'C'D',可近似认为划定区域尺寸与逆透视变换后图像尺寸具有如下比例关系:

$$\frac{A'B'}{AB} = \frac{A'D'}{AD} \approx k \quad (28)$$

因此,d_y所对应的AB到cd真实世界纵向距离d_{real}可计算为:

$$d_{\text{real}} = \frac{d_y}{k} \quad (29)$$

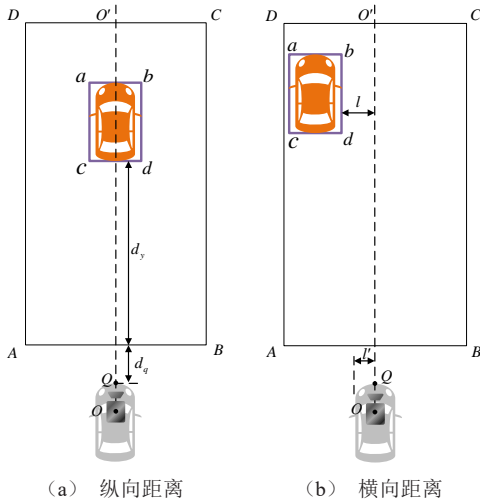


图 6 逆透视变换图像车距估计

进一步,需计算车头位置 Q 距 AB 线段的纵向距离 d_q 。如图 7 所示, AB 是设定的逆透视变换区域, QQ' 是两车之间的实际距离, 图 6 中的 d_y 对应图 7 中逆透视变换 AQ' 距离, QP 由照相机安装位置决定, PA 由感兴趣角点标定位置决定。当照相机俯仰角 $\bar{\theta}$ 、安装位置、逆透视变换区域确定以后, QA 保持不变, 并假设受轴荷转移影响忽略不计。

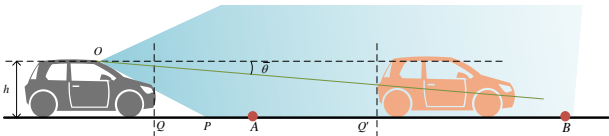


图 7 纵向实际车距估计

所以 d_q 可计算为:

$$d_q = QP + PA \quad (30)$$

进而可得到车辆之间的实际纵向距离 d_{total} 为:

$$d_{\text{total}} = d_{\text{real}} + d_q \quad (31)$$

对于相邻车道的车辆横向距离同样可根据上述方法计算。如图 6b 所示, 目标车辆检测框线 bd 与 OO' 的真实距离 l_{real} 计算为:

$$l_{\text{real}} = \frac{l}{k} \quad (32)$$

所以目标车辆距自车的横向距离 d_x 可计算为:

$$d_x = l_{\text{real}} - l' \quad (33)$$

式中: l' 为照相机光轴与自车侧边之间的横向距离。

4 试验验证与分析

为验证本文所提方法的有效性, 在某型号车辆上使用单目摄像机抓取已知车辆距离位置信息的图片。其中, 相机安装在前挡风玻璃内的内后视镜处, 安装高度为 1.42 m, 安装俯仰角为 5° , 航向角为 0° 。

利用本文所提方法与相关对比文献分别进行车辆距离估算, 并将估算结果与实际测量车辆间距进行对比验证。本节分别从 3D 目标检测算法、逆透视变换方法和道路路面条件 3 个角度对本文所提方法进行试验验证。

4.1 基于注意力机制的 ImVoxelNet 目标检测试验

试验计算机操作系统为 Ubuntu 20.04, 内存为 90 GB, GPU 为 1 张 GeForce RTX 4090 (24 GB), 训练环境为 Python3.8, 深度学习框架为 PyTorch1.11.0, Cuda 版本为 11.3。使用 KITTI 数据集对改进后的 ImVoxelNet 在以上环境配置下训练。试验结果见表 1, 由表可知, 改进后 ImVoxelNet 算法的 3D 检测框在简单、中等、困难场景下的 AP40 指标值分别提升了 3.07%、2.68%、3.11%, 表明本文改进型 ImVoxelNet 算法具有更高的目标检测精度。

相比传统 ImVoxelNet 算法, 本文所提目标检测算法由于考虑了通道特征的注意力分布, 能检测在无遮挡和部分遮挡环境下的车辆。如图 8b 所示, 本文所提改进型 ImVoxelNet 算法有效检测出了图 8a 中未检测到的红圈遮挡车辆, 并且车辆外部三维检测框能准确反映车辆方位和轮廓信息, 有效减少了车辆的漏检。

表 1 3D 目标检测试验结果

模型	3D AP40/%		
	简单	中等	困难
ImVoxelNet ^[15]	87.46	78.40	76.85
改进 ImVoxelNet	90.53	81.12	79.96

图 9 为本文所提改进型 ImVoxelNet 算法在不同天气和光照条件下的试验结果, 由图可知, 在弱光或浓雾环境下, 由于部分车辆轮廓特征在低能见度

环境下明显减弱，所以出现如红圈所示的漏检现象。

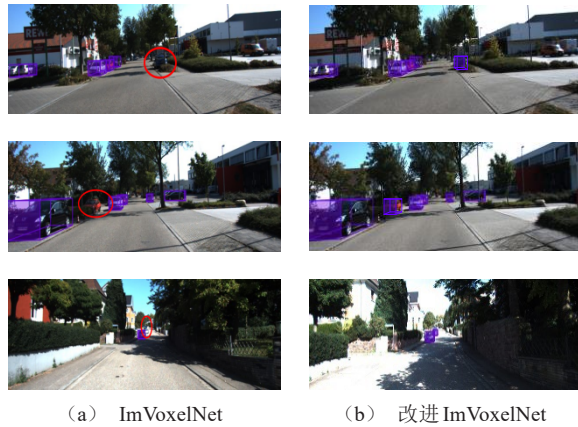


图 8 3D 目标检测试验对比

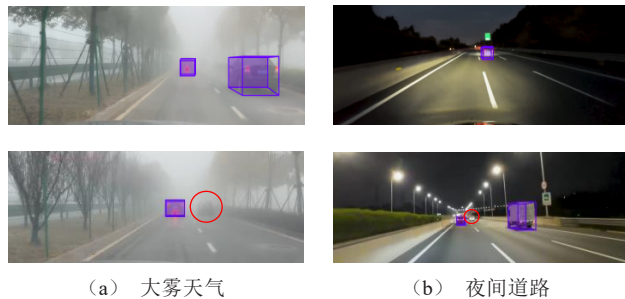


图 9 低能见度条件下的车辆 3D 目标检测效果

4.2 基于图像标定的逆透视变换算法试验

本小节开展基于图像标定的逆透视变换算法有效性验证试验。通过采集两组原始图像数据如图 10a 所示，根据式 (12) 和式 (13) 计算得到标定四边形区域角点坐标 (A, B, C, D) ，进行逆透视变换区域标定。本文仅对该标定区域进行逆透视变换，剔除了原图像中的护栏、障碍物等场景无关信息，其逆透视变换图像在横、纵向没有产生较大畸变，如图 10b 所示，可近似反映真实世界的比例关系。而传统逆透视变换方法由于没有进行图像标定，导致图像与真实场景不具备固定比例关系，由图 10c 的两组试验结果可知，汉字和车道线在横向产生较大畸变，这将造成后续车距估算模型精度下降。

4.3 非平坦路面下单目视觉车距估计试验

该部分验证所提车距估计算法在非平坦路面下的车距估计精确性。试验方法为：首先对前车与自

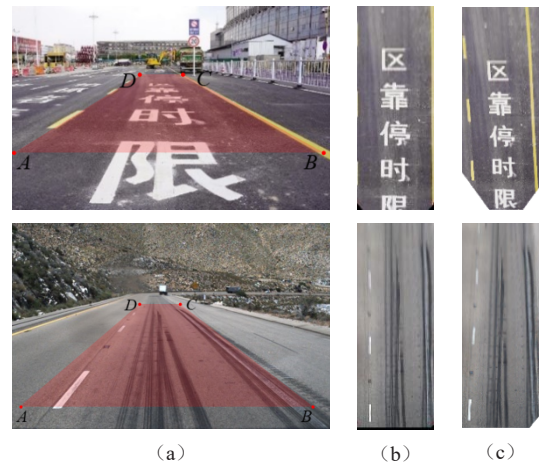


图 10 逆透视变换试验结果

车距离分别各测量 10 次，通过调整前车和自车的真实间距，使实际距离误差在 1% 范围内。按照第 3 节所提车距估计算法分别开展纵向间距为 10、20、40、80 m 的平坦及非平坦路面下的横、纵向车距估计试验，所选取的平坦及非平坦试验道路如图 11 所示。

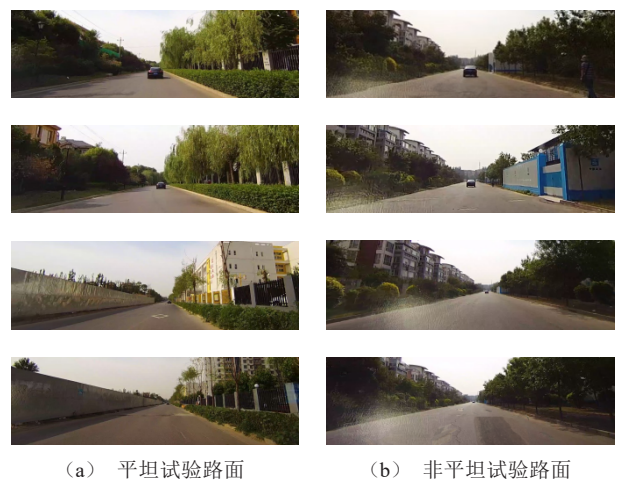


图 11 车距估计试验道路场景

车距估计试验结果见表 2。可以看出，在平坦路面条件下，纵向车距估计平均误差为 0.495 m，相对误差 < 3%，横向车距估计平均误差为 0.117 5 m，相对误差 < 2%；在非平坦路面条件下，纵向车距估计平均误差为 0.76 m，横向车距估计平均误差为 0.62 m，相对误差均 < 2%。

图 12 表明车辆实际间距与车距估计误差呈正相关性, 车辆实际距离在 20 m 以内时, 平坦路面和非平坦路面条件下的车距估计误差均较小; 在实际距离超过 40 m 时, 两种路面条件下的车距估计误差均明显增大。这是由于前车与观测点距离增大, 摄像头对距离感知灵敏度降低而导致的。此

外, 在非平坦路面条件下, 受车辆姿态角变化及相机安装高度动态变化等因素影响, 车距估计误差均比平坦路面的车距估计误差有所增长。综上所述, 试验结果表明, 本文所提车距估计算法在平坦路面及非平坦路面条件下均表现出良好的车距估计效果。

表 2 平坦路面及非平坦路面下的车距估计试验结果

道路条件	实际纵向车距/m	纵向估计车距/m	纵向车距估计误差/m	实际横向车距/m	横向估计车距/m	横向车距估计误差/m
平坦路面	10	9.92	0.08	1	0.98	0.02
	20	19.95	0.05	2	1.96	0.04
	40	39.42	0.58	3	3.11	0.11
	80	78.73	1.27	4	3.69	0.31
非平坦路面	10	9.88	0.12	1	0.92	0.08
	20	19.63	0.37	2	1.78	0.22
	40	40.81	0.81	3	3.52	0.52
	80	81.74	1.74	4	5.66	1.66

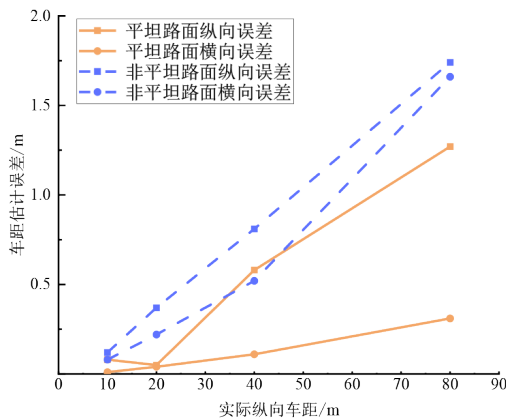


图 12 非平坦路面与平坦路面车距估计对比

为进一步表明所提方法优势, 该部分安排以下对比试验。

1) 文献 [22] 使用车辆宽度的尺度变化估算车距, 需要利用同一车辆在不同时刻检测框信息进行推理估计, 增大了目标检测算法误差带来的影响。而本文方法只需要当前时刻的车辆检测框信息即可估计车距。

2) 文献 [23] 使用贝叶斯框架与水平三角测量方法, 通过估计前车宽度作为先验信息, 利用帧间位置变化跟踪前车并不断更新车距信息, 在倾斜

道路上水平三角测量误差较大, 车距估计效果较差。而本文考虑了非平坦路面下的倾斜问题, 降低了路面坡度带来的估计误差。

3) 文献 [24] 运用了变参数的逆透视变换测距模型, 降低了车辆姿态角变化引起的误差, 但逆透视变换畸变较大。本文通过对逆透视变换进行区域标定, 大大降低了逆透视变换引起的图像畸变, 提高了车距估计精度。

根据表 2 和表 3 的试验结果, 绘制非平坦路面下车辆纵、横向车距估计试验结果对比图, 如图 13 所示。

由图 13 可知, 现有车距估计方法的估计误差与纵向实际距离呈正相关性。具体而言, 在车辆相距较近的 20 m 范围内, 4 种测距方法均展现出了较高的估计精确度。然而, 当车辆间的实际纵向距离扩展至 40 m 以外时, 这 4 种方法在纵向和横向上的误差均有显著上升。通过对 4 种方法在纵、横向误差对比分析可知, 同一测距方法在纵向与横向上的误差差异相对较小, 而不同测距方法之间的误差差异则相对较大, 而本文方法的误差增长趋势更平

表 3 非平坦路面下车辆纵、横向车距估计对比试验结果

方法	实际车距/m	绝对误差/m		相对误差/%	
		纵向	横向	纵向	横向
文献 [22]	10	0.37	0.25	1.85	1.25
	20	0.28	0.29	2.77	2.92
	40	2.24	2.40	3.74	4.00
	80	3.56	3.81	4.45	4.76
	平均值	1.61	1.69	3.20	3.23
文献 [23]	10	0.42	0.25	2.10	1.25
	20	0.33	0.38	3.27	3.75
	40	2.69	2.80	4.48	4.67
	80	4.76	4.46	5.95	5.58
	平均值	2.05	1.97	3.95	3.81
文献 [24]	10	0.31	0.08	1.57	0.42
	20	0.25	0.13	2.50	1.25
	40	1.67	1.02	2.79	1.67
	80	2.33	2.21	2.91	2.62
	平均值	1.14	0.83	2.44	1.49

缓。可以得出，相较于对比方法，本文的车距估计方法在非平坦路面下具有更高的检测精度。

5 结论

本文提出了一种基于注意力机制的单目视觉车距估计方法，可以有效估算非平坦路面条件下的前车间距。通过将注意力机制融入 3D 目标检测算法，

参考文献 (References)

[1] BAO Dongsheng, WANG Peikang. Vehicle Distance Detection Based on Monocular Vision [C]//2016 International Conference on Progress in Informatics and Computing (PIC), Dec. 23-25, 2016, Shanghai, China. Piscataway NJ:IEEE, c2016: 224-228.

[2] 赵江洪,刘茈菱,杨甲,等.几何单目视觉测距研究综述[J].测绘科学,2023,48(9):49-65.
ZHAO Jianghong, LIU Ziling, YANG Jia, et al. Review of Geometric Monocular Visual Ranging Research [J]. Science of Surveying and Mapping. 2023, 48(9): 49-65. (in Chinese)

[3] 沈志熙,黄席樾.基于数据回归建模的单目视觉测距算法[J].计算机工程与应用,2007,43(24):15-18,40.

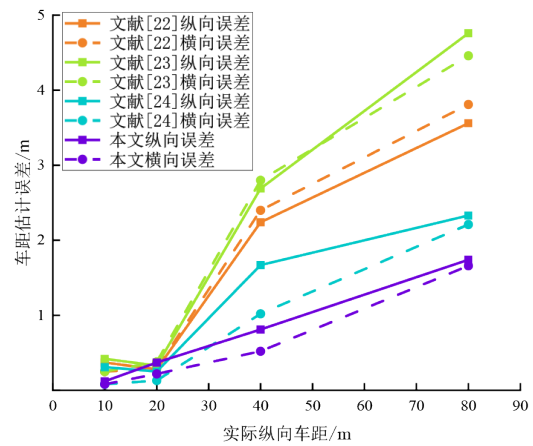


图 13 非平坦路面下车辆纵、横向车距对比试验结果

成功提高了 3D 目标检测精度、降低了车辆漏检率；通过计算和标定图像感兴趣区域，剔除逆透视变换无关信息，有效降低了逆透视图像的畸变程度；考虑如减速带、丘陵、坑洼等非平坦路面条件下车辆姿态角变化对相机外参矩阵影响，非平坦路面下的相机坐标转换模型，降低了相机矩阵参数摄动对车距估计精度的影响。通过对比试验及数据分析，验证了本文利用单目视觉相机估计车距方法的有效性。本研究局限性在于仅考虑了直线道路场景，无法适用于道路转角较大的场景，并且在光照不足环境下只能对特征轮廓明显的车辆进行检测和车距估计。未来工作将进一步探索在复杂行驶环境下的车距估算方法，以进一步提高车距估计方法适用性。

SHEN Zhixi, HUANG Xiyue. Monocular Vision Distance Detection Algorithm Based on Data Regression Modeling [J]. Computer Engineering and Applications, 2007, 43(24): 15-18, 40. (in Chinese)

[4] HUANG D Y, CHEN C H, CHEN T Y, et al. Vehicle Detection and Inter-Vehicle Distance Estimation Using Single-Lens Video Camera on Urban/Suburb Roads [J]. Journal of Visual Communication and Image Representation, 2017, 46: 250-259.

[5] BUI M T, DOSKOCIL R, KRIVANEK V. Distance and Angle Measurement Using Monocular Vision [C]//2018 18th International Conference on Mechatronics-Mechatronika, Dec. 5-7, 2018, Brno, Czech Republic.

- Piscataway NJ:IEEE, c2018: 1-6.
- [6] 张欣,魏志恒,王兵.单目视觉的前车位置识别及安全车距监测[J].河北大学学报(自然科学版),2022,42(3):321-326.
ZHANG Xin, WEI Zhiheng, WANG Bing. Position Recognition of the Vehicle Ahead and Safety Distance Monitoring Based on Monocular Vision [J]. Journal of Hebei University(Natural Science Edition), 2022, 42(3): 321-326.(in Chinese)
- [7] 王荣本,李斌,储江伟,等.公路上基于车载单目机器视觉的前方车距测量方法的研究[J].公路交通科技,2001,18(6):94-98.
WANG Rongben, LI Bin, CHU Jiangwei, et al. Study on the Method of Measuring the Leading Vehicle Distance Based on the On-Board Monocular Camera[J]. Journal of Highway and Transportation Research and Development, 2001, 18(6): 94-98.(in Chinese)
- [8] MAO Jiafa, HUANG Wei, SHENG Weiguo. Target Distance Measurement Method Using Monocular Vision [J]. IET Image Processing, 2020, 14(13): 3181-3187.
- [9] ZHANG Zhisheng, HAN Yanxiang, ZHOU Yifan, et al. A Novel Absolute Localization Estimation of a Target with Monocular Vision[J]. Optik, 2013, 124(12): 1218-1223.
- [10] ZHANG Daiming, FANG Bin, YANG Weibin, et al. Robust Inverse Perspective Mapping Based on Vanishing Point [C]//IEEE International Conference on Security, Pattern Analysis, and Cybernetics, Oct. 18-19, 2014, Wuhan, China. Piscataway NJ: IEEE, c2014: 458-463.
- [11] 王永森,刘宏哲.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[J].计算机工程与科学,2020,42(7):1234-1243.
WANG Yongsen, LIU Hongzhe. Vehicle Distance Measurement with Vehicle Lower Edge Estimation and Inverse Perspective Mapping Based on Monocular Vision [J]. Computer Engineering and Science, 2020, 42(7): 1234-1243.(in Chinese)
- [12] ZHE Ting, HUANG Liqin, WU Qiang, et al. Inter-Vehicle Distance Estimation Method Based on Monocular Vision Using 3D Detection [J]. IEEE Transactions on Vehicular Technology, 2020, 69(5): 4907-4919.
- [13] MARTINEZ-DIAZ S, PADILLA-MEDINA J A. 3D Distance Measurement from a Camera to a Mobile Vehicle, Using Monocular Vision [J]. Journal of Sensors, 2021, 2021: 5526931.1-5526931.8.
- [14] 王飞,何军,黄雄炬,等.基于多视角的轻量化车路协同感知模型[J].无线电通信技术,2024,50(4):688-695.
WANG Fei, HE Jun, HUANG Xiongju, et al. Lightweight Vehicle Road Collaborative Perception Model Based on Multiple Perspectives [J]. Radio Communications Technology, 2024, 50(4): 688-695.(in Chinese)
- [15] RUKHOVICH D, VORONTSOVA A, KONUSHIN A. ImVoxelNet: Image to Voxels Projection for Monocular and Multi-View General-Purpose 3D Object Detection [C]//2022 IEEE Winter Conference on Applications of Computer Vision (WACV), Jan. 4-8, 2022, Waikoloa, HI, USA. Piscataway NJ: IEEE, c2022: 1265-1274.
- [16] ZHURAVLEV D. Towards Real-Time 3D Object Detection Through Inverse Perspective Mapping [J]. Software Engineering Methods in Systems and Network Systems, 2024, 934: 371-385.
- [17] LEE H, JO A, HYUN Y, et al. LiDAR Ground Detection-Based Dynamic Inverse Perspective Mapping of BEV Lanes [J]. IEEE Sensors Journal, 2023, 23(24): 30950-30959.
- [18] YANG Chunling, ZHANG Chunchao, YANG Xuqiang, et al. Performance Study of CBAM Attention Mechanism in Convolutional Neural Networks at Different Depths [C]//2023 IEEE 18th Conference on Industrial Electronics and Applications (ICIEA), Aug. 18-22, 2023, Ningbo, China. Piscataway NJ: IEEE, c2023: 1373-1377.
- [19] BERTOZZ M, BROGGI A, FASCIOLI A. Stereo Inverse Perspective Mapping: Theory and Applications [J]. Image and Vision Computing, 1998, 16(8): 585-590.
- [20] YANG Handuo, MA Tao, TONG Zheng, et al. Deployment Strategies for Lightweight Pavement Defect Detection Using Deep Learning and Inverse Perspective Mapping [J]. Automation in Construction, 2024, 167: 105682.1-105682.17.
- [21] AHMAD M A, MUAD A M. Calibration of Inverse Perspective Mapping from Different Road Surface Images [C]//2021 IEEE 11th International Conference on System Engineering and Technology (ICSET), Nov. 6, 2021, Shah Alam, Malaysia. Piscataway NJ: IEEE, c2021: 139-142.
- [22] 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.
- [23] NAKAMURA K, ISHIGAKI K, OGATA T, et al. Real-Time Monocular Ranging by Bayesian Triangulation [C]//2013 IEEE Intelligent Vehicles Symposium (IV), June 23-26, 2013, Gold Coast, Australia. Piscataway NJ:

IEEE, c2013:1362-1367.

- [24] 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.

LIU Jun, HOU Shihao, 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)

作者简介



刘永涛(1989-), 男, 安徽砀山人, 博士, 副教授, 主要研究方向为人车系统安全、智能车辆控制技术。

E-mail: liuyongtao86@163.com

通信作者



王泰琪(1992-), 男, 陕西咸阳人, 博士, 讲师, 主要研究方向为车辆性能测试。

E-mail: wangtaiqi@chd.edu.cn