

Research  
Robotics—Article

## 一种基于共点映射的无人车可行驶区域检测方法

刘子熠<sup>a,b</sup>, 余思雨<sup>a,b</sup>, 郑南宁<sup>a,b,\*</sup><sup>a</sup> Institute of Artificial Intelligence and Robotics, Xi'an Jiaotong University, Xi'an 710049, China<sup>b</sup> National Engineering Laboratory for Visual Information Processing and Applications, Xi'an Jiaotong University, Xi'an 710049, China

## ARTICLE INFO

## Article history:

Received 6 April 2017

Revised 13 November 2017

Accepted 28 December 2017

Available online 18 July 2018

## 关键词

可行驶区域

自动驾驶

数据融合

共点映射

## 摘要

城市交通场景的随机性和复杂性使得无人车的可行驶区域检测成为难题。受人类驾驶行为的启发, 本文提出了一种无人车的可行驶区域检测的新方法, 该方法利用了单目相机得到的像素信息和激光传感器得到的空间点云信息的融合, 与共线中的双射类似, 该方法引入了“共点映射”的新概念, 其定义为: 将来自激光雷达的点映射到图像分割边缘上的点的双射。该方法通过将障碍物与超像素融合得到初始可行驶区域, 并基于该初始可行驶区域, 利用自学习模型定位候选可行驶区域。此外, 为了提升算法的鲁棒性, 本文融合了四种特征, 特别提出了一种称为可行驶程度 (drivable degree, DD) 的特征, 该特征定义了激光点的可行驶程度。经过自学习四种特征得到初始可行驶区域之后, 利用贝叶斯框架建立可行驶区域的最终概率图模型。本文的方法没有引入强假设条件, 也不需要训练过程, 但在 ROAD-KITTI benchmark 测试中获得了最佳的性能。实验结果表明, 该方法是一种泛化性强且有效的可行驶区域的检测方法。

© 2018 THE AUTHORS. Published by Elsevier LTD on behalf of Chinese Academy of Engineering and Higher Education Press Limited Company This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

## 1. 引言

一直以来, 道路检测都被认为是无人车领域的关键技术, 并吸引了广大研究者的关注。迄今为止, 道路检测取得了显著进展[1–3]。然而, 仅基于道路检测的自动驾驶决策可能仍无法处理一些紧急情况, 例如, 由于突然转向的车辆或行人使得所检测到的道路变得不可行这种情况。事实上, 在驾驶汽车时, 人类驾驶员通过将障碍物与非障碍物分类来理解情景, 而不仅仅是识别道路。因此, 在紧急情况下, 出于安全原因, 人类驾驶员可以选择在通常不被视为道路的平坦区域上行驶。对于

无人车而言, 检测这种“平坦区域”而不是检测道路区域, 可以为决策过程提供更全面的知识, 使无人车更像人类驾驶员。

虽然大多数现有的道路检测方法通过训练在标签清晰的道路数据集上实现了不错的效果, 但由于场景布局、光照和天气条件的高度可变性, 如何检测在城市和乡村环境中道路标识不明确场景下的道路和车道仍然十分困难。迄今为止, 该问题还没有可靠的解决方案, 迫切需要提出一种鲁棒而有效的解决方法。

在图像分割中, 物体的边界通常出现在深度不连续处。因此, 图像分割应该与深度不连续性检测相融合。

\* Corresponding author.

E-mail address: [nnzheng@mail.xjtu.edu.cn](mailto:nnzheng@mail.xjtu.edu.cn) (N. Zheng).

在投影几何中,类似形是指投影空间的同构,它是一种线到线的双射关系,即所谓的共线。在此,我们引入一个类似于共线中双射的新概念:共点映射。共点映射是一种点到点的双射,它将来自激光传感器的点映射到图像分割边缘的点上。在给定空间域下,将投影空间用一组点的法向矢量表示。由于光照变化,道路的不平坦以及二维(two-dimensional, 2D)图像中的阴影,一些共点不是类似形。为了克服这些问题,本文使用点云数据的法向矢量而不是原始点云数据,后文中有详细描述。在这种情况下,可以清楚地表征像素深度数据的融合,并将像素深度的类似形定义为共点映射。

因此,本文提出了一种自适应的可行驶区域的检测方法,该方法利用共点映射将单目相机获得的像素信息与激光传感器获得的空间信息融合,如图1所示。其中,通过将每个激光点的空间位置与图像像素的坐标结合,建立Delaunay三角剖分图[4],以此得到激光点之间的空间关系,并利用三角剖分三角形的法向矢量完成激光点的障碍分类;接下来,通过自学习模型将障碍物分类结果与图像超像素融合来定位初始可行驶区域;然后,在不同特征空间中计算候选可行驶区域,这些特征包括:可行驶程度(drivable degree, DD)特征、法向量(normal vector, NV)特征、颜色特征和强度特征;最后,利用贝叶斯框架融合候选区域以获得最终的可驱动区域。本文使用ROAD-KITTI benchmark [5]进行实验测试。实验结果表明,与其他融合方法相比,本文所提出的方法无需形状或高度信息的假设条件和训练过程,且取得了最优越的性能;该结果验证了我们的方法具有鲁棒性和高泛化能力。

本文的主要贡献如下:

(1) 提出了一种基于数据融合的无监督检测方法,该方法无需强假设,确保了在多变的城市交通场景下的泛化能力。

(2) 引入了共点映射这一新概念,它描述了激光传感器和相机之间进行数据融合的一种新型约束。

(3) 设计了一个名为DD的新特征,以描述激光点的可行驶程度。

## 2. 相关工作

鲁棒的道路区域检测方法是无人车的关键组成部分。在过去的几十年中,出现了许多针对该问题的解决方法。根据采集数据的传感器不同将这些方法进行

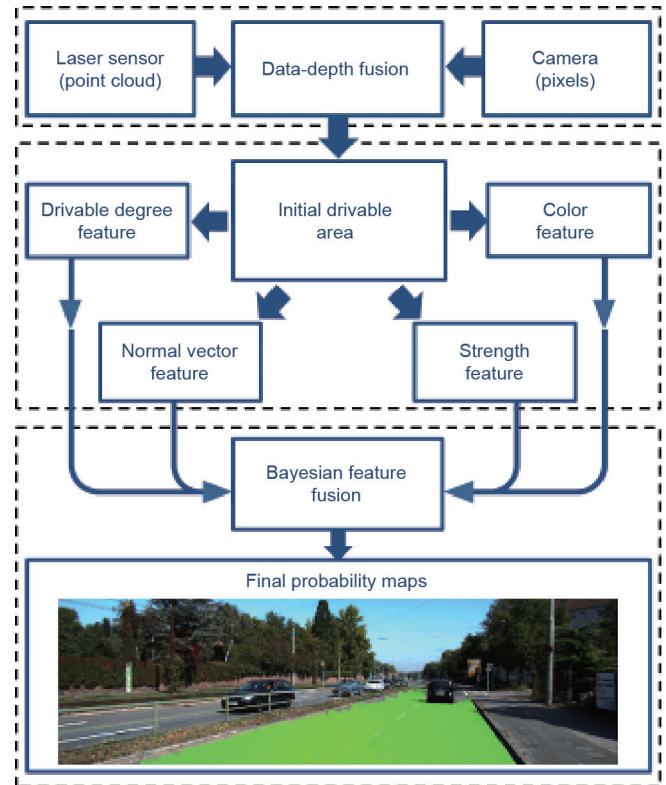


图1. 本文所提出的检测方法的框架。该方法大致可分为三个步骤:数据融合、特征提取和特征融合。

分类:单目相机、立体视觉、激光传感器和多传感器的融合。

基于单目视觉的方法已广泛用于道路检测。与其他传感器相比,视觉传感器体积小、成本低、易于安装。此外,视觉传感器可以提供丰富的视觉信息,具有广阔的检测范围。最重要的是,视觉传感器的原理和结构类似于人类感知组织的原理和结构。通常利用视觉场景的2D信息检测道路区域,如颜色、角点、纹理、边缘和形状。对于颜色信息,通常利用RGB颜色空间[6]、HSI颜色空间[7]或其他颜色空间进行图像分割。Jau等[8]比较了不同光照条件下的RGB和HSI颜色空间分割效果。Finlayson等[9]提出了基于物理的光照不变颜色空间原理,实现了无阴影影响的图像表征,本文也使用了该颜色空间模型。此外,通过利用用于捕获原始彩色图像的相机光谱特性,Maddern等[10]提出了另一种光照不变颜色空间,减少了由太阳光引起的光照变化的影响。

另一个热门研究课题是基于卷积神经网络(convolutional neural network, CNN)的方法,该方法取得了巨大成功[11,12]。最初,CNN用于解决分类问题[13,14];然而,随着近期出现的一系列工作[15,16]基于CNN的语义

分割方法进入了高潮时期。

然而，道路的概念与其他视觉概念不同，因为视觉中的像素外观不是检测道路区域的唯一标准，像平坦度之类的物理属性对道路的概念判定影响更大。这表明仅依靠单目视觉的方法不够可靠。尽管基于CNN的方法已经可以实现良好的性能，但它们严重依赖于训练过程，可能无法处理崭新的道路场景，并且可能具有过拟合问题。与场景分类[17]或相似性学习[18]等问题不同，道路区域检测是一种病态条件问题，需要使用2D信息来解决三维（three-dimensional, 3D）真实世界场景下的任务。虽然可以利用许多如地平线和消失点之类的3D线索来缓解这个问题[19–21]，然而，这些3D线索的检测本身就是一个未解决的问题[22,23]；一些几何假设可能导致检测失灵或降低泛化能力，如图2所示。

近年来，激光传感器的出现激励了许多基于激光传感器的道路检测方法，这些方法可以提供对真实世界3D场景的深度增补测度。这些方法使用激光点的空间位置分析场景，并将平坦区域识别为道路，可以做如下分类。

(1) 基于网格的方法。由于点云数据量庞大，通常使用基于2D网格的方法[24–26]来减采样，同时计算网格内点的统计信息以表征每个网格，可能包括平均高度和最大高度差等统计。虽然这些方法简单直接，对噪声鲁棒且高效，但难以确定合适的阈值。

(2) 基于平面拟合的方法。这些方法的基本假设是，道路是平坦且平滑的，可以通过多个参数进行平面拟合[27,28]。典型的平面估计方法已经很成熟，如随机抽样一致算法（random sample consensus, RANSAC）[29]。然而，由于缺乏真正的地面激光点，这些方法在复杂交通场景下可能失效。

(3) 基于相邻点之间空间关系的方法。这些方法[30,31]利用相邻点之间的空间关系来提取特征（如法向

量）或建立概率模型，以便估计地面激光点或障碍物激光点（如路缘激光点）。

以上所有基于激光传感器的方法都受到点云数据稀疏性的影响，这使得单纯利用激光点是难以重建场景细节的。

检测道路区域可视为两类标签问题，条件随机场（conditional random field, CRF）框架在该领域得到广泛应用[32,33]。根据各个观测，CRF框架将该问题定义为计算数据集整体标签的最大后验概率问题。这是一个通用框架，通过设计能量函数和势函数的不同的通道来定义不同的观测结果。因此，为了平衡不同来源的数据并获得最佳融合结果，基于CRF的方法被广泛用于融合[34,35]。然而，基于CRF的方法会带来很大的计算消耗和存储消耗，并且需要手动标记的数据。

为了克服上述缺点，本文提出了一种基于共点映射的自适应的检测方法，该方法通过将激光传感器的数据与单目相机的数据融合以检测可行驶区域。首先，该方法包含几个预处理步骤：利用超像素分割获得后续步骤的最小图像处理单元；通过交叉校准和共点映射将激光点投射到RGB图像上；利用三角剖分对激光点进行预处理，以建立激光点之间的空间关系。接下来，利用预处理结果进行像素深度数据融合，并利用数据融合结果完成障碍物分类。然后，将超像素与障碍物分类结果相结合，得到初始可行驶区域。之后是特征提取，本文采用DD特征、NV特征、颜色特征和强度特征，所有这些特征都能够以自学方式计算出可行驶概率。最后，在超像素层面，利用贝叶斯框架获得可行驶区域的联合概率分布图。

本文所提出的方法在三个主要方面区别于其他方法：

- (1) 不需要强假设，不需要标记数据进行训练；
- (2) 通过利用共点视角和融合激光传感器与单目相机，对光照变化鲁棒，并且可以应对复杂场景；
- (3) 采用超像素分割，用超像素替换像素作为最小

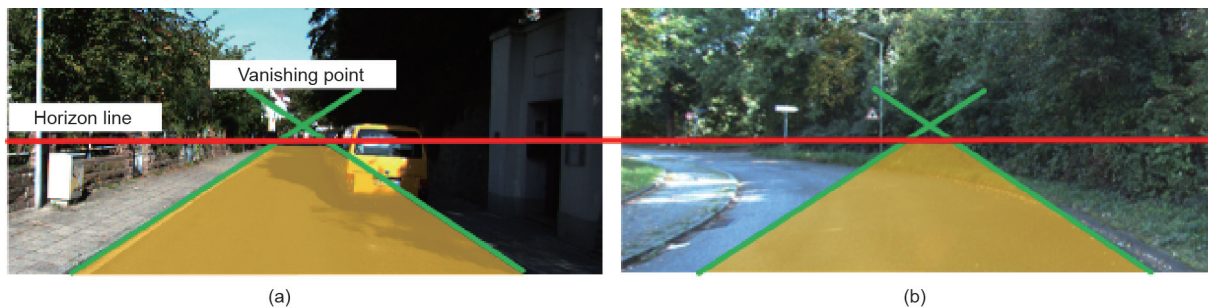


图2. 比较在3D线索可能起作用或失败下的不同道路场景。(a) 3D线索运作良好；(b) 由于消失点的移动和道路形状的弯曲，3D线索可能失败。

处理元素，这种处理的优点将在后面详述。

不同于一些缺乏特征层面融合的方法[36,37]，本文所提出的方法将单目视觉与激光传感器相结合，获得了数据层面和特征层面的丰富信息。此方法自主学习地提取和融合特征。此外，共点视角和超像素表征使得该方法更具有鲁棒性和有效性。实验结果表明，本文的方法比其他方法具有更高的检测精度。因此，我们认为，本文所提出的方法是一种泛化能力强、实用性高且自适应的无人车的可行驶区域的检测方法。

### 3. 基础理论

本章叙述了新方法所需要的基础理论，包括图像的超像素表征、激光点在RGB图像上的投影过程以及对数色度空间的建立过程。

#### 3.1. 超像素表征

超像素算法最早由Ren和Malik [38]提出。一个超像素代表在颜色或纹理上相近的一组像素，这种表示方式保留了原始图像的大部分结构信息。

本文利用超像素代替像素作为图像处理步骤中的最小处理单元，以辅助寻找候选可行驶区域。由于超像素方法性能的提升，用超像素代替像素降低了计算和存储成本，且不会带来较大的精度牺牲。此外，超像素将颜色信息考虑在内，使得算法在处理复杂照明情况下仍可以获得鲁棒的结果。

为了更好地分割原始图像，超像素方法应该满足两个要求：首先，超像素的生成速度应该快；其次，生成的超像素应“附着”在图像边缘上。

根据参考文献[39,40]，一种名为“边缘附着”的超像素满足上述要求，故本文采用边缘附着超像素的生成方法。该方法是在简单线接口计算（simple line interface calculation, SLIC）方法[41]的基础上添加了对边缘的考虑，利用这个增加的边缘项，生成的超像素可以更好地附着到物体边缘，从而保留更多的图像结构信息，同时也产生更好的物体边缘分割。

#### 3.2. 激光点投影过程

如图3所示，本文参照文献[42]，将激光传感器的坐标系投影到相机坐标系下。激光坐标系下任一3D点  $p_{laser} = (x_l, y_l, z_l, 1)^T$  投影到相机坐标系下对应点  $p_{camera} = (x_c, y_c, z_c, 1)^T$  的计算公式如下：

$$p_{camera} = \mathbf{R}_{rect}^0 \mathbf{T}_{velo}^{cam} p_{laser} \quad (1)$$

式中， $\mathbf{R}_{rect}^0$ 表示旋转矩阵，本文通过附加全为零的行和列将 $\mathbf{R}_{rect}^0$ 转化为一个 $4 \times 4$ 的矩阵，并设置 $\mathbf{R}_{rect}^0(4,4) = 1$ 。 $\mathbf{T}_{velo}^{cam}$ 表示转换矩阵，计算公式如下：

$$\mathbf{T}_{velo}^{cam} = \begin{pmatrix} \mathbf{R}_{velo}^{cam} & \mathbf{t}_{velo}^{cam} \\ \mathbf{0} & 1 \end{pmatrix} \quad (2)$$

式中， $\mathbf{R}_{velo}^{cam}$ 和 $\mathbf{t}_{velo}^{cam}$ 分别表示旋转矩阵和平移向量，如文献[42]所示：

$$(u, v, 1)^T = \mathbf{P}_{rect}^{(i)} (x_c, y_c, z_c, 1)^T \quad (3)$$

式中， $\mathbf{P}_{rect}^{(i)}$ 表示所有点 $p_c$ 在第 $i$ 个相机平台下的投影矩阵（本文使用第二个相机），则相机坐标系下的位置信息 $(u, v)$ 可由上述公式求得。

经过投影和矫正就可以得到对齐后的激光点集 $\mathbb{P} = \{\mathbf{P}_i\}_{i=1}^N$ ，其中 $\mathbf{P}_i = (x_i, y_i, z_i, u_i, v_i)$ 。

#### 3.3. 对数色度空间

根据文献[36]，为了获得对阴影和光照条件鲁棒的颜色特征，本文将RGB彩色图像（ $I$ ）变换为对数色度空间（ $I_{log}$ ），得到光照不变图像 $I_{l-c}$ 。每个对数色度空间中的 $\{\log(R/G), \log(B/G)\}$ 像素值对应于原始RGB图像中的 $\{R, G, B\}$ 像素值。如图4所示，本文通过将 $\{\log(R/G), \log(B/G)\}$ 像素值投影到角度为 $\theta$ 的正交轴上，从而获得灰度图像。角度 $\theta$ 定义了与光照变化线正交的不变方向，其数值取决于设备本身，可以进行校准。本文按照参考文献[43]将 $\theta$ 设置为 $45^\circ$ ， $I_{log}$ 可以根据如下公式计算：

$$I_{log}(u, v) = \frac{\log[I_R(u, v)/I_G(u, v)] + \log[I_B(u, v)/I_G(u, v)] \cdot \tan \theta}{1 + \tan^2 \theta} \quad (4)$$

## 4. 像素 - 深度数据融合

本文没有利用可以直接探测障碍的传感器[44]，而

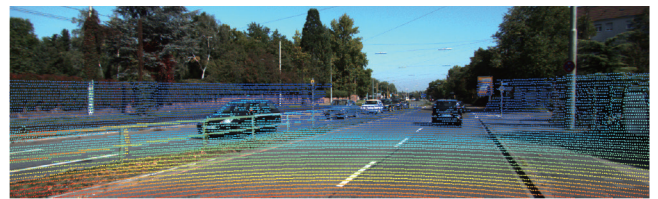


图3. 激光点和图像像素点的对齐示例。这里，每个激光点的深度用不同颜色展示，以便更好地查看对齐效果。

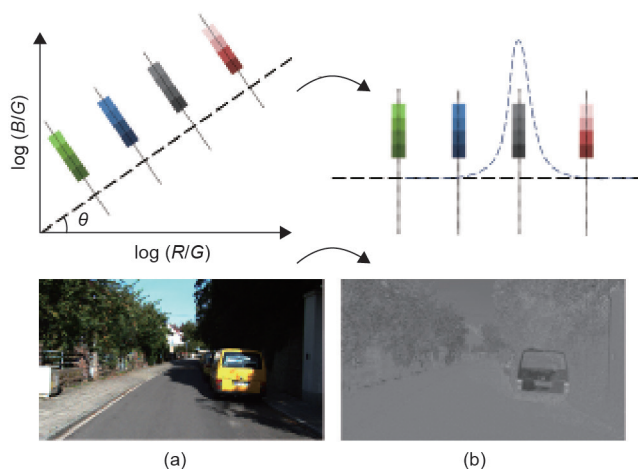


图4. 从(a)图表示的RGB空间到(b)图表示的对数色度空间的转换示例图。从图中可以看出, 经过转换后, 阴影被成功地去除了, 同时, 道路和其他物体(即黄色汽车和绿色植物)之间的色差保留较完整。

是通过将激光传感器提供的空间位置信息与单目相机提供的外观信息相融合以寻找障碍物。本章将介绍像素和深度数据融合的过程。简单而言: 第一步, 将超像素与激光点融合, 以提高方法的效率和鲁棒性; 第二步, 本文称之为“共点映射”过程, 包括将激光传感器得到的深度点与图像分割边缘上的像素点融合, 以消除平坦区域内的激光点; 最后一步是空间信息和图像坐标之间的融合, 此过程通过生成无向图来建模激光点和障碍物之间的空间关系。

#### 4.1. 基于超像素的图像预处理

本文采用3.1节中描述的超像素生成方法。除了上述优点之外, 本文发现, 超像素和激光点是互补的: 第一, 激光点包含无法从单目相机获得的空间位置信息; 第二, 超像素是稠密的, 因此像素和激光点融合后可以得到两者的鲁棒局部统计。此外, 超像素还包含激光传感器无法捕获的颜色信息。因此, 本文利用超像素代替像素, 作为图像处理步骤中的最小单位。另外, 第5节叙述了使用超像素分割获得初始可行区域特征的算法。

#### 4.2. 基于共点映射的边缘信息双射

由于激光点和图像像素本质上属于在相同场景相同时间下的不同观测, 所以它们反映的是相同的结构信息。因此, 二者之间的投影应该满足某些特定约束。受投影几何中的共线概念的启发, 本文引入了一种新的概念, 即“共点映射”, 来描述这种约束。类似于共线的双射, 共点映射定义了将激光传感器得到的点映射到图

像分割边缘上的像素点的双射。

使用共点映射可以提升对齐性能并消除不需要的激光点。具体而言, 边缘可以被视为影响像素的外观和激光点结构的基本元素。像素和激光点之间的边缘信息的双射是数据融合的关键。对齐像素和激光点时, 两者的边缘信息必须正确对齐。使用交叉校准可以获得初始对齐结果。接下来, 本文通过共点映射来改善对齐效果。类似的, 边缘信息的双射可以消除位于平坦区域中的冗余激光点, 加速计算过程。而且, 因为平坦区域内可能存在噪声激光点, 这种消除也可以提高算法的鲁棒性。

如图5所示, 本文仅使用共点映射来消除平坦区域内的激光点。首先, 使用所有超像素的边缘获取整个图像的边缘池。显然, 此过程存在许多冗余边缘信息。但是, 考虑到边缘池保留了所有真实边缘, 具有边缘信息完整性, 这是一种合理的做法。实际上, 本文采用图像扩张来减少对齐误差。接下来, 保留位于边缘池中的激光点, 丢弃其他点。实验结果表明, 27%的激光点被消除了, 且这种消除没有牺牲算法精度, 这一发现表明, 共点映射是可行且有效的。

#### 4.3. 障碍物分类

障碍物分类过程可以看作寻找映射函数 $ob(\mathbf{P}_i)$ 的过程, 其中:

$$ob(\mathbf{P}_i) = \begin{cases} 1, & \text{if } \mathbf{P}_i \text{ is an obstacle point} \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

分类结果如图6所示。

我们认为 $ob(\mathbf{P}_i)$ 的值仅取决于点 $\mathbf{P}_i$ 周围表面的平坦程度, 现实世界中这种平坦程度信息蕴藏在激光点云

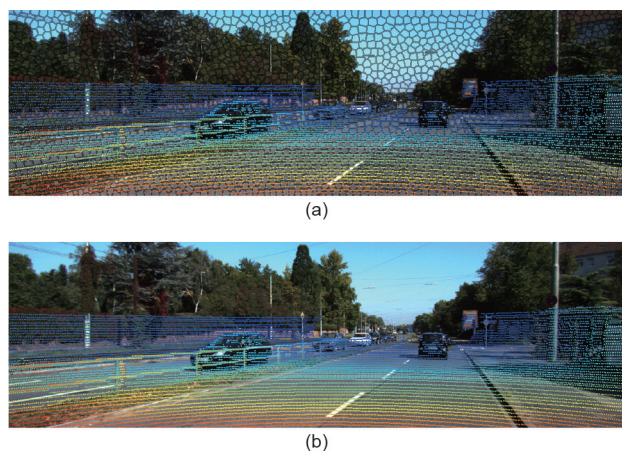


图5. 共点映射的示例图。与(b)图不同, (a)图仅保留靠近超像素边缘的激光点。虽然丢弃了许多点, 但也成功地保留了激光点结构信息。

数据中。因此，障碍物分类问题被分解为两个子问题，即如何找到 $\mathbf{P}_i$ 的周围表面和如何确定这些周围表面是否平坦。

本文通过利用Delaunay三角剖分来解决第一个子问题[4]。对于平面中的一组点集的Delaunay三角剖分，该集合中的任何点都不在该三角剖分中的任一三角形的外接圆内。此属性使得空间的每个点平均具有6个相邻三角形，其最近邻图是Delaunay三角剖分的子图，因此Delaunay三角剖分可用于在激光点云数据之间建立空间关系。对于每一点 $\mathbf{P}_i = (x_i, y_i, z_i, u_i, v_i)$ ，本文对其像素坐标系 $(u_i, v_i)$ 运用三角剖分构建无向图 $G = \{\mathbb{P}, E\}$ ，其中 $E$ 表示 $\mathbf{P}_i$ 之间的具有空间关系的有效边的集合。 $\mathbf{P}_i$ 的周围表面由公式 $\{(u_j, v_j) | j = i \text{ or } \mathbf{P}_j \in Nb(\mathbf{P}_i)\}$ 定义的表面三角形组成，其中， $Nb(\mathbf{P}_i)$ 表示与 $\mathbf{P}_i$ 连接的点的集合。这里需要剔除欧式距离不满足如下公式的边 $(\mathbf{P}_i, \mathbf{P}_j)$ ：

$$\|(\mathbf{P}_i - \mathbf{P}_j)\| < \varepsilon \quad (6)$$

式中， $\|\mathbf{P}_i - \mathbf{P}_j\|$ 表示 $(x_i, y_i, z_i)$ 和 $(x_j, y_j, z_j)$ 的欧氏距离； $\varepsilon$ 表示最大边长阈值。

对于第二个子问题， $\mathbf{P}_i$ 周围表面的平坦程度可以通过其相邻三角形的法向量度量，将 $\mathbf{P}_i$ 的所有相邻三角形的法向量的均值作为 $\mathbf{P}_i$ 的法向量，表示为 $N(\mathbf{P}_i) = (x_i^n, y_i^n, z_i^n)$ 。由此，可以根据如下公式得到 $ob(\mathbf{P}_i)$ ：

$$ob(\mathbf{P}_i) = \begin{cases} 1, & \text{if } \arcsin(z_i^n / \|N(\mathbf{P}_i)\|) > c \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

式中， $c$ 为手动设置的参数； $Nb(\mathbf{P}_i)$ 表示与 $\mathbf{P}_i$ 连接的点的集合， $\mathbf{P}_i$ 的法向量就是 $N(\mathbf{P}_i) = (x_i^n, y_i^n, z_i^n)$ 。可以这样理解该公式：如果 $\mathbf{P}_i$ 的法向量与地面夹角大于 $c$ ，则 $\mathbf{P}_i$ 被分类

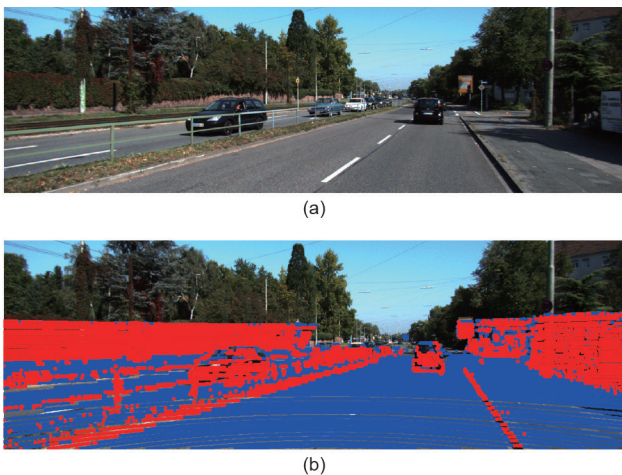


图6. 障碍物分类的结果。该结果表明本文提出的方法不受障碍物像素外观的影响，可以很好地检测出障碍物（即路缘石、墙壁和汽车）。

为障碍物点；反之，为非障碍物点。 $c$ 定义了最大角度阈值，实验中取 $60^\circ$ 。计算法向量以分类障碍物的过程如图7所示。

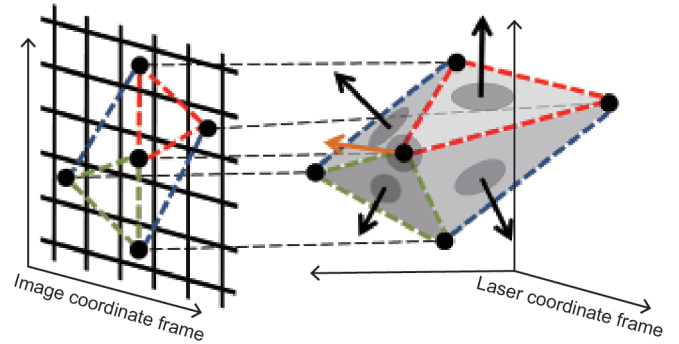


图7. 障碍物分类的图示。每个黑点代表一个激光点。点之间的空间关系由平面Delaunay三角剖分得到，在激光传感器坐标系下计算三角形的法向量（黑色箭头：相邻三角形的法线向量；橙色箭头：中心点的法向量，通过计算周围的黑色箭头的均值得到）。

## 5. 可行驶区域检测

本节介绍了该方法的关键步骤——可行驶区域检测。首先，从障碍物分类结果获得方向射线图 $I_{\text{DRM}}$ 。接下来，通过将 $I_{\text{DRM}}$ 与超像素结合起来获得初始可行驶区域，本文采用不同观测的不同特征表征初始可行驶区域；然后，自学习地计算在初始可行驶区域内的每个超像素属于可行驶区域的概率；最后，利用贝叶斯框架融合由各个特征得到最终可行驶区域的概率图。

### 5.1. 计算初始可行驶区域

#### 5.1.1. 射线检测

按照算法1计算 $I_{\text{DRM}}$ 。

#### 算法1 计算方向射线图

##### Input:

The set of laser points  $\{\mathbb{P}^{(h)}\}_{h=1}^H$ ;

##### Output:

Direction ray map,  $I_{\text{DRM}}$

1: Initial  $I_{\text{DRM}}$  with the size of  $I$  and zeros elements

2: **for**  $h = 1$  to  $H$  **do**

3: Find obstacle set  $\mathbb{O} = \{\mathbf{P}_i | ob(\mathbf{P}_i) = 1 \ \& \ \mathbf{P}_i \in \mathbb{P}^{(h)}\}$

4: **if**  $\mathbb{O} = \emptyset$  **then**

5:  $\mathbf{P}_{\text{bin}}^{(h)} = \arg \max_{\mathbb{P}^{(h)}} \text{dist}(\mathbf{P}_i^{(h)}, \mathbf{P}_{\text{base}})$

6: **else**

7:  $\mathbf{P}_{\text{bin}}^{(h)} = \arg \min_{\mathbf{P}_i^{(h)} \in \mathbb{O}} \text{dist}(\mathbf{P}_i^{(h)}, \mathbf{P}_{\text{base}})$

8: **end if**

9: Line point  $\mathbf{P}_{\text{bin}}^{(h)}$  with point  $\mathbf{P}_{\text{base}}$  in  $I_{\text{DRM}}$

10: **end for**

首先，本文将任一 $\mathbf{P}_i$ 都转换到极坐标系下，以便更好

地表示可行驶区域。也就是说,  $(u_i, v_i)$  被转换为极坐标, 坐标系原点是图像的最中间底部的像素 (记为  $\mathbf{P}_{\text{base}}$ )。因此, 用  $\mathbb{P}$  表示  $\{\mathbb{P}^{(h)}\}_{h=1}^H$ , 其中  $\mathbb{P}^{(h)} = \{\mathbf{P}_i^{(h)}\}_{i=1}^{N^{(h)}}$ ,  $\mathbf{P}_i^{(h)}$  表示属于第  $h$  角度范围内的点。 $\mathbb{P}$  的稀疏性带来了两个影响精度的新问题: 第一, 如何克服连线的“泄漏”问题, 如图8所示; 第二, 如何从稀疏射线中获得稠密的像素区域。

为了解决第一个问题, 本文在图像坐标系下对射线长度进行滤波, 如图8 (d) 所示。由于汽车的宽度不可忽略, 因此一条射线所代表的区域是否可行驶还取决于该区域的宽度。也就是说, 如果射线辐射的区域太窄以至于不能满足车身通过, 无论该区域多么平坦, 都不是可行驶的。因此, 采用最小滤波对射线长度进行处理, 解决了“泄漏”问题, 如图8 (c) 所示。

对于第二个问题, 最直接的解决方案是增大  $H$ 。但是, 这样做会加剧第一个问题。因此, 本文将  $I_{\text{DRM}}$  与超像素融合以获得初始可行驶区域。如4.1节所述, 此解决方案有两个优点: 首先, 它通过用超像素替换像素, 大大减少了数据量; 第二, 它融合了深度信息和颜色信息。

在  $I_{\text{DRM}}$  与超像素融合之后, 本文用一组超像素的点集表示初始可行驶区域, 该点集定义为  $S_{\text{int}} = \{S_i | S_i \cup I_{\text{DRM}} \neq \emptyset\}$ ,  $\mathbb{P}_{S_i}$  表示投影在超像素  $S_i$  区域内的激光点的集合。由此, 以下所有特征的计算均是基于超

像素而不是像素, 既引入鲁棒的局部统计, 又加速了整个算法。

### 5.1.2. 计算 DD 特征

检测可行驶区域的算法需要精选的合理的表征特征。本文提出了DD特征, 如算法2所示。

#### 算法2 计算DD特征

##### Input:

The set of laser points  $\{\mathbb{P}^{(h)}\}_{h=1}^H$ ;

##### Output:

DD feature for every point  $D(\mathbf{P}_i^{(h)})$

- 1: Initial  $D(\mathbf{P}_i^{(h)})$  with the zero ( $h = 1 \rightarrow H, i = 1 \rightarrow N^{(h)}$ )
- 2: **for**  $h = 1$  to  $H$  **do**
- 3:   **for**  $i = 1$  to  $N^{(h)}$  **do**
- 4:     **if**  $ob(\mathbf{P}_i^{(h)}) = 1$  **then**
- 5:       **for**  $j = 1$  to  $N^{(h)}$  **do**
- 6:          $D(\mathbf{P}_i^{(h)}) = D(\mathbf{P}_i^{(h)}) + \text{abs}(z_{\mathbf{P}_i^{(h)}} - z_{\mathbf{P}_j^{(h)}})$
- 7:       **end for**
- 8:     **end if**
- 9:   **end for**
- 10: **end for**

将  $\mathbb{P}^{(h)}$  中的点按与  $\mathbf{P}_{\text{base}}$  在图像坐标系下的距离进行处理, 使得  $\mathbb{P}^{(h)}$  中的点均满足:

$$\text{dist}(\mathbf{P}_i^{(h)}, \mathbf{P}_{\text{base}}) > \text{dist}(\mathbf{P}_{i-1}^{(h)}, \mathbf{P}_{\text{base}}), i = 2 \rightarrow N^{(h)} \quad (8)$$

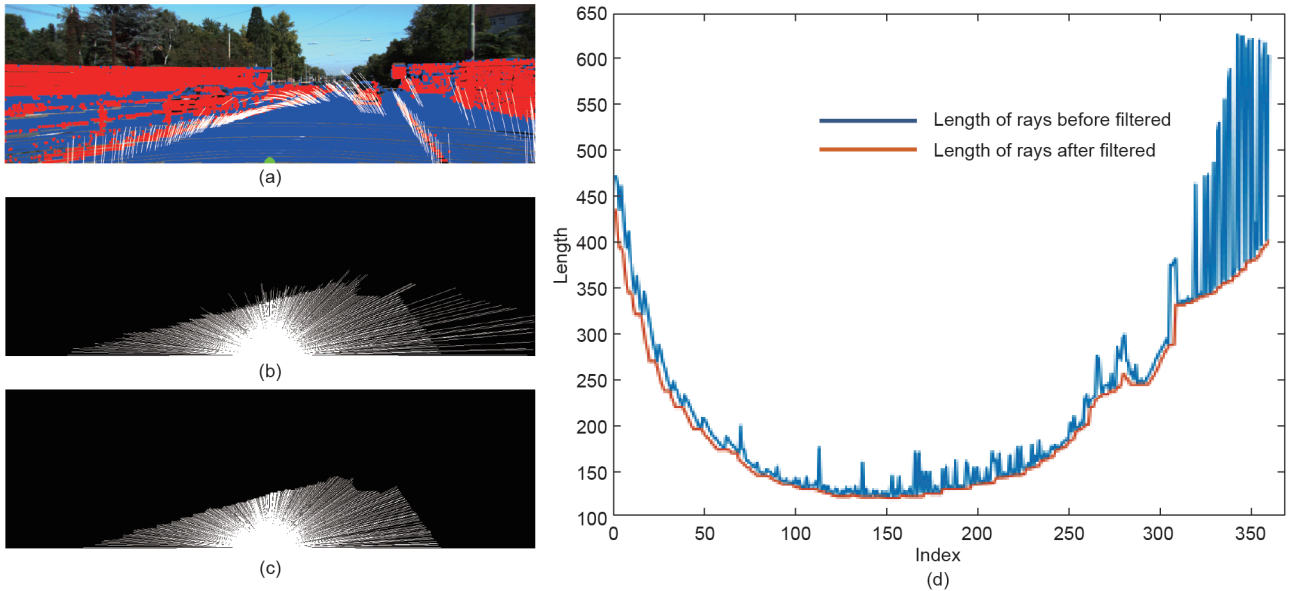


图8.“泄漏”问题的示意图。(a)为“泄漏”问题示意图: 绿点是  $\mathbf{P}_{\text{base}}$ ; 在 (b) 中, 检测射线由白线表示; (c) 表示最小滤波的结果示意图; (d) 展示了最小滤波处理之前 (蓝色曲线) 和之后 (橙色) 的每个射线长度的对照图。

图9展示了DD特征值计算的示意图。图10(c)显示了引入DD特征后的效果[45]。

## 5.2. 自适应的特征模型

基于所获得的初始可行驶区域，可以从不同概率空

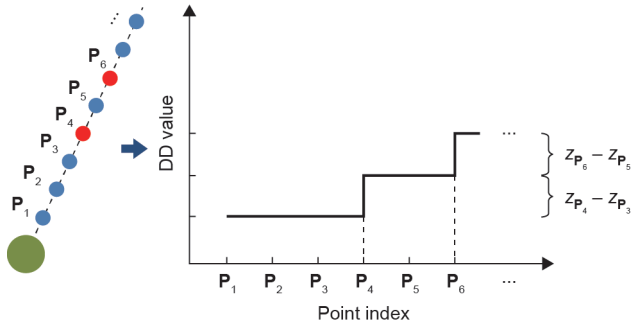


图9. DD计算的示意图。左侧的黑色虚线表示 $I_{DRM}$ 中的射线。蓝点表示非障碍物激光点，红点表示障碍物激光点，大绿点表示 $\mathbf{P}_{base}$ 。在该图中， $\mathbf{P}_4^{(h)}$ 被分类为障碍点；接下来，将 $\mathbf{P}_4^{(h)}$ 和 $\mathbf{P}_3^{(h)}$ 之间的高度差加到 $\mathbf{P}_4^{(h)}$ 之后的所有点的DD特征值上，如 $\mathbf{P}_6^{(h)}$ 。

间中的4个特征自适应地学习候选可行驶区域，此4个特征为：DD、NV、颜色和强度特征。

### 5.2.1. DD 特征

每一超像素 $D(S_i)$ 的DD特征值由如下公式计算：

$$D(S_i) = \frac{1}{\|P_{S_i}\|} \sum_{P_i} D(P_i) \quad (9)$$

由于 $D(P_i)$ 与点的高度差相关，可知， $D(P_i)$ 越小， $P_i$ 所对应的区域的可行驶程度越高。为了得到候选可行驶区域，构造一个类高斯模型用于表征DD特征的自适应概率空间，计算如下：

$$D_{prob}(S_i) = \begin{cases} \exp\{-[D(S_i) - \mu_D]^2 / 2\sigma_D^2\}, & \text{if } D(S_i) \geq \mu_D \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

式中， $\mu_D$ 和 $\sigma_D^2$ 为此类高斯分布的参数，可由基于 $S_{int}$ 的概

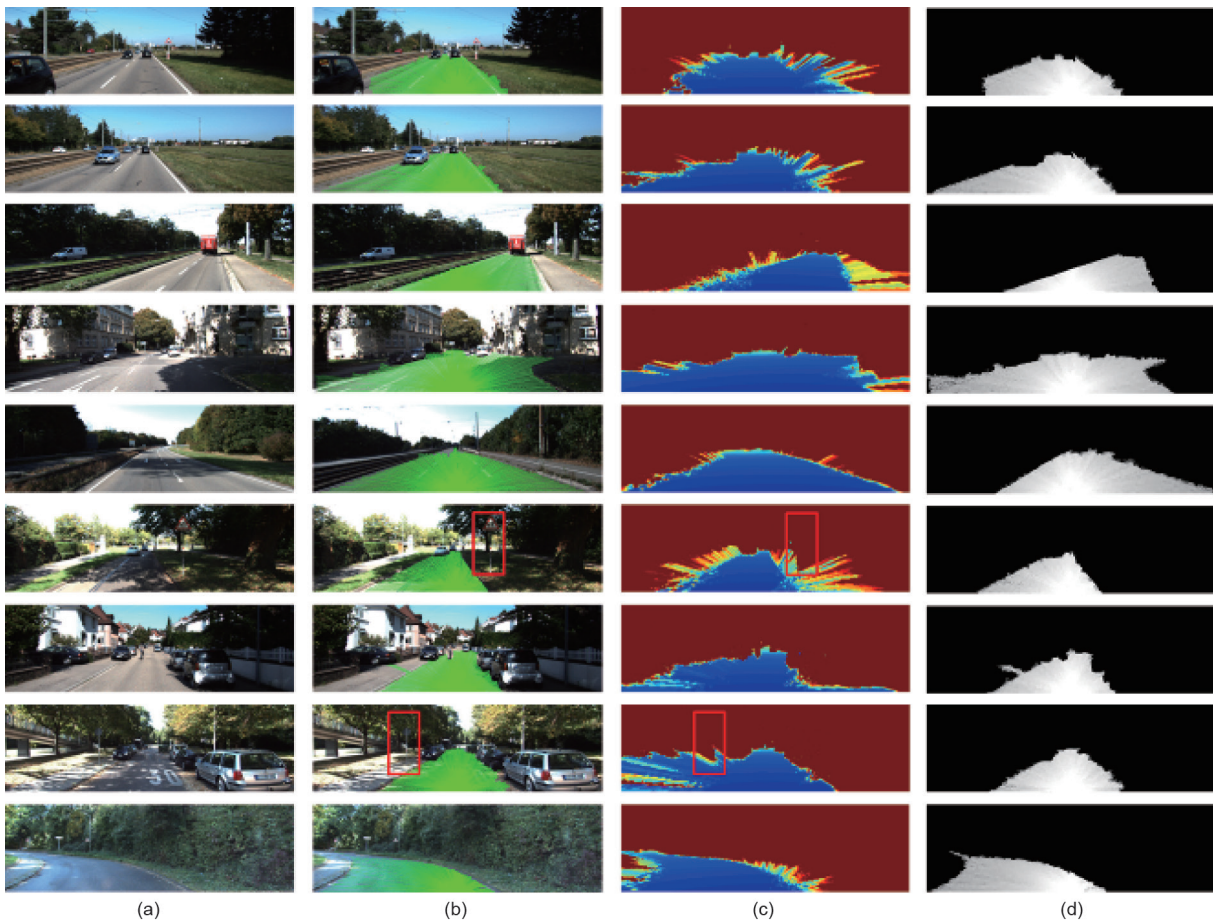


图10.  $I_{DRM}$ 、DD特征计算结果和最终概率图的展示图。(a) 显示原始图像，其中包含不同的场景，如汽车遮挡、重影和道路消失点的漂移。(b) 显示了相应的 $I_{DRM}$ ，覆盖了足够的可行驶区域，以生成合理的 $S_{int}$ 。(c) 显示了DD特征[44]，区域越蓝，可行驶程度越大。图中可见，真实道路区域非常蓝，人行道介于蓝色和红色之间，而汽车和障碍物等障碍物则为红色。(d) 显示最终可行驶区域的概率图，证明本文所提出的方法在不同情况下表现良好。第六行和第八行中的红色框说明了我们的方法对小目标（如第二列中的框内的灯杆）的处理效果良好。灯杆前面的点的DD特征值很高，但是在这些灯杆后面的区域中的点DD特征值突然下降，表明这些小目标已被成功检测到。



率统计，不需要手动设置或者训练； $D_{\text{prob}}(S_i)$ 为在给定NV特征度量下， $S_i$ 属于可行驶区域的概率。

### 5.2.2. NV 特征

根据4.3节，法向量在 $z_i^n$ 方向上的值越大，表面可行驶程度越大。因此，本文定义了每个超像素点的NV特征，用 $N(S_i)$ 表示。通过计算 $\mathbb{P}_{S_i}$ 中 $z_i^n$ 的最小值得到 $N(S_i)$ 。类似于上述DD特征计算，构造包含参数 $\mu_n$ 和 $\sigma_n^2$ 的类高斯模型表征自适应概率特征空间：

$$N_{\text{prob}}(S_i) = \begin{cases} \exp\{-[N(S_i) - \mu_n]^2/2\sigma_n^2\}, & \text{if } N(S_i) \geq \mu_n \\ 1, & \text{otherwise} \end{cases} \quad (11)$$

式中， $N_{\text{prob}}(S_i)$ 表示在给定NV特征度量下， $S_i$ 属于可行驶区域的概率。估计参数 $\mu_n$ 和 $\sigma_n^2$ 的过程与5.2.1节中相同。因此，该模型也是自适应的，不涉及手动设置。

### 5.2.3. 颜色特征

如3.3节所述，利用光照不变图像来获得 $S_{\text{int}}$ 的颜色特征。与 $D(S_i)$ 和 $N(S_i)$ 类似，使用高斯参数 $\mu_c$ 和 $\sigma_c^2$ 建立参数化的概率模型：

$$C_{\text{prob}}(S_i) = \exp\left\{-\frac{[I_{\log}(S_i) - \mu_c]^2}{2\sigma_c^2}\right\} \quad (12)$$

式中， $C_{\text{prob}}(S_i)$ 表示在给定颜色特征下， $S_i$ 属于可行驶区域的概率； $I_{\log}(S_i)$ 表示 $S_i$ 在颜色空间转化后得到的像素值。

### 5.2.4. 强度特征

$S_i$ 的强度特征 $Sg(S_i)$ 度量了超像素的平坦程度，是 $I_{\text{DRM}}$ 与每个超像素的重叠部分。在强度特征下， $S_i$ 是属于可行驶区域的概率建模如下：

$$Sg_{\text{prob}}(S_i) = \frac{Sg(S_i) \text{dist}(S_i, \mathbf{P}_{\text{base}})}{A(S_i)} \quad (13)$$

式中， $A(S_i)$ 表示 $S_i$ 所占图像坐标系中的区域； $\text{dist}(S_i, \mathbf{P}_{\text{base}})$ 表示图像坐标系下 $S_i$ 与 $\mathbf{P}_{\text{base}}$ 的距离。

## 5.3. 贝叶斯框架

本文利用贝叶斯框架进行特征融合。根据相机和激光传感器的观测结果，找出每一超像素属于可行驶区域的后验概率，表示为 $p(S_i = R|\mathbf{Obs})$ ，其中， $\mathbf{Obs}$ 代表上述的所有观测：

$$\mathbf{Obs} = \{D(S_i), C(S_i), N(S_i), Sg(S_i)|S_i \in \mathbb{S}_{\text{int}}\} \quad (14)$$

接下来，将从上述4个特征获得的概率图作为先验条件概率。假设各个超像素是条件独立存在的，则每个超像素属于可行驶区域的后验概率如下：

$$p(S_i = R|\mathbf{Obs}) \propto \prod_{j=1}^4 p(S_i = R|\mathbf{Obs}_j)p(S_i = R) \quad (15)$$

式中， $p(S_i = R)$ 表示超像素 $S_i$ 属于可行驶区域的概率，通过对整个图像集的 $\mathbb{S}_{\text{int}}$ 求平均来得到。

## 6. 实验结果和讨论

为了测试新方法，本文在ROAD-KITTI benchmark上进行实验，使用了289个训练图像和290个测试图像[5]。在鸟瞰图（bird's eye view, BEV）下，使用以下6个指标对实验结果进行评估：最大 $F$ 测量（max  $F$ -measure, MaxF）、平均精度（average precision, AP）、精确度（precision, PRE）、召回率（recall, REC）、误报率（false positive rate, FPR）和假负利率（false negative rate, FNR）。实验使用了3个数据集：标记的城市数据集（urban marked, UM）、多标记城市数据集（urban multiple marked, UMM）和未标记的城市数据集（urban unmarked, UU）。为了证明本文所提出方法的有效性（下文表中使用“**Our method with co-point**”表示），本文将与其在ROAD-KITTI benchmark上的利用激光传感器的前三最优的方法（HybridCRF、MixedCRF和LidarHisto）进行比较。因为本文提出的方法采用的数据融合与参考文献[36]中使用的方法（RES3D-Velo）类似，本文也列出了与RES3D-Velo方法的对照实验结果。为了显示通过共点映射消除激光点所产生的影响，本文比较了带有和未带有共点映射的方法。此外，本文在训练集上展示了带有共点映射方法的性能，以证明该方法是无监督的。

如表1至表4所示，我们的方法在UMM和UU数据集中，在PRE和FPR度量上取得最佳效果，表明其很好地覆盖了道路区域。就表4来看，我们的方法在AP、PRE和FPR度量上均取得最优性能，表明此方法在不同场景中具有鲁棒性。与没有共点映射的方法相比，我们的共点映射方法消除了大约30%的激光数据（表5），但仍然得到了类似的优越性能。这一发现表明，共点映射可以成功地保留图像的语义结构。最重要的是，虽然我们的

方法是无监督的，但与监督方法相比，此方法具有很大竞争力和优势。

此外，为了验证特征融合过程对算法性能的改善程度，本文将特征融合的结果与每个特征得到的概率图、 $S_{int}$ 以及基准进行了比较，如表6至表9所示。同样使用上述数据集的训练集进行对照实验。从表6至表9可以看

出，特征融合对MaxF和AP度量改善效果显著。

与使用全部激光点的方法（见下文表中“Fusion without co-point”）相比，具有共点映射的方法在训练集中平均减少了约30%的激光点（表5），但仍然取得了类似的性能。

$S_{int}$ 在REC和FNR度量上表现出优异的性能，与“基

表1 与目前最优的三种方法在UM (BEV)数据集上的对照实验结果

Method	MaxF	AP	PRE	REC	FPR	FNR
HybridCRF	90.99%	85.26%	90.65%	91.33%	4.29%	8.67%
MixedCRF	90.83%	83.84%	89.09%	92.64%	5.17%	7.36%
LidarHisto	89.87%	83.03%	91.28%	88.49%	3.85%	11.51%
RES3D-Velo	83.81%	73.95%	78.56%	89.80%	11.16%	10.20%
Our method with co-point	83.94%	84.83%	79.11%	89.40%	10.76%	10.60%
Our method without co-point	84.82%	85.37%	78.61%	92.09%	11.42%	7.91%
Our method on training set	85.87%	87.85%	81.56%	90.67%	9.43%	9.33%

表2 与目前最优的三种方法在UMM (BEV)数据集上的对照实验结果

Method	MaxF	AP	PRE	REC	FPR	FNR
HybridCRF	91.95%	86.44%	94.01%	89.98%	6.30%	10.02%
MixedCRF	92.29%	90.06%	93.83%	90.80%	6.56%	9.20%
LidarHisto	93.32%	93.19%	95.39%	91.34%	4.85%	8.66%
RES3D-Velo	90.60%	85.38%	85.96%	95.78%	17.20%	4.22%
Our method with co-point	91.80%	92.09%	91.31%	92.29%	9.65%	7.71%
Our method without co-point	92.41%	92.31%	91.61%	93.22%	9.38%	6.78%
Our method on training set	92.15%	93.56%	92.84%	91.48%	7.70%	8.52%

表3 与目前最优的三种方法在UU (BEV)数据集上的对照实验结果

Method	MaxF	AP	PRE	REC	FPR	FNR
HybridCRF	88.53%	80.79%	86.41%	90.76%	4.65%	9.24%
MixedCRF	82.79%	69.11%	79.01%	86.96%	7.53%	13.04%
LidarHisto	86.55%	81.13%	90.71%	82.75%	2.76%	17.25%
RES3D-Velo	83.63%	72.58%	77.38%	90.97%	8.67%	9.03%
Our method with co-point	82.57%	81.49%	75.54%	91.04%	9.61%	8.96%
Our method without co-point	83.09%	83.25%	75.52%	92.36%	9.76%	7.64%
Our method on training set	82.03%	82.82%	74.63%	91.05%	11.01%	8.95%

表4 与目前最优的三种方法在URBAN (BEV)数据集上对照实验结果，URBAN数据集结合了UM、UMM和UU数据集，此结果为以上三种数据集上取得结果的平均值

Method	MaxF	AP	PRE	REC	FPR	FNR
HybridCRF	90.81%	86.01%	91.05%	90.57%	4.90%	9.43%
MixedCRF	89.46%	83.70%	88.52%	90.42%	6.46%	9.59%
LidarHisto	90.67%	84.79%	93.06%	88.41%	3.63%	11.59%
RES3D-Velo	86.58%	78.34%	82.63%	90.92%	10.53%	9.08%
Our method with co-point	87.02%	86.55%	83.49%	90.87%	9.90%	9.13%
Our method without co-point	87.70%	87.31%	83.50%	92.36%	10.06%	7.64%
Our method on training set	86.68%	89.71%	87.23%	91.78%	42.22%	8.44%

表5 本文方法所使用的激光点数量

	Our method with co-point	Our method without co-point	Eliminating rate
Training set	13 194	18 873	30.09%
Testing set	12 902	18 663	30.87%

表6 在UM训练集 (BEV) 上的融合性能分析

Item compared	MaxF	AP	PRE	REC	FPR	FNR
Baseline	81.34%	86.42%	77.09%	86.09%	11.76%	13.91%
Initial drivable area	78.61%	64.00%	67.25%	94.58%	21.18%	5.42%
Color feature	83.97%	78.65%	77.57%	91.52%	12.17%	8.48%
Strength feature	81.39%	84.44%	80.00%	82.82%	9.52%	17.18%
DD feature	84.56%	74.42%	78.71%	91.34%	11.36%	8.66%
NV feature	85.38%	77.10%	81.81%	89.28%	9.13%	10.72%
Fusion with co-point	85.87%	87.85%	81.56%	90.67%	9.43%	9.33%
Fusion without co-point	86.55%	88.10%	81.90%	91.76%	9.33%	8.24%

表7 在UMM训练集 (BEV) 上的融合性能分析

Item compared	MaxF	AP	PRE	REC	FPR	FNR
Baseline	79.48%	84.04%	72.83%	87.47%	35.59%	12.53%
Initial drivable area	88.06%	79.48%	82.21%	94.81%	22.38%	5.19%
Color feature	91.41%	89.92%	89.96%	92.90%	11.31%	7.10%
Strength feature	88.15%	90.71%	83.72%	93.08%	19.74%	6.92%
DD feature	90.75%	86.23%	89.64%	91.89%	11.58%	8.11%
NV feature	91.83%	89.33%	93.05%	90.64%	7.38%	9.36%
Fusion with co-point	92.15%	93.56%	92.84%	91.48%	7.70%	8.52%
Fusion without co-point	92.99%	93.81%	93.11%	92.87%	7.49%	7.13%

表8 在UU训练集 (BEV) 上的融合性能分析

Item compared	MaxF	AP	PRE	REC	FPR	FNR
Baseline	73.02%	76.96%	68.57%	78.10%	12.74%	21.90%
Initial drivable area	74.26%	57.49%	60.61%	95.85%	22.16%	4.15%
Color feature	78.37%	69.23%	68.64%	91.31%	14.84%	8.69%
Strength feature	78.62%	81.54%	74.39%	83.37%	10.21%	16.63%
DD feature	81.00%	68.05%	72.23%	92.19%	12.61%	7.81%
NV feature	81.23%	69.53%	73.85%	90.24%	11.37%	9.76%
Fusion with co-point	82.03%	82.82%	74.63%	91.05%	11.01%	8.95%
Fusion without co-point	82.99%	83.36%	75.80%	91.69%	10.41%	8.31%

准”有近似的FPR，因此， $S_{int}$ 用于估计参数是合理可行的，详见第5节。

如图11所示，在ROAD-KITTI数据集中，由红线围住的图像区域不是被标定为道路真实值的区域。然而，鉴于这些区域大都平坦可供行驶，并且检测可行区域是本研究主要关注的问题，本文倾向于将这些区域识别为前景。在图11所示的情况下，本文的方法

取得的FPR值更高。实际上，这些有界区域具有模糊的语义，包括人行道和道路之间的过渡区以及停车场的入口和车辆开到人行道之前的斜坡。必要时，这些区域中大都可用于车辆行驶。在现实生活中，我们期望自动驾驶汽车会在紧急情况下选择这些平坦区域作为候选道路（如避开突然转弯的车辆）。因此，规划算法应该充分考虑这些区域。

表9 在URBAN训练集（BEV）上的融合性能分析

Item compared	MaxF	AP	PRE	REC	FPR	FNR
Baseline	77.95%	82.47%	72.83%	83.88%	20.03%	16.11%
Initial drivable area	80.31%	66.99%	70.02%	95.08%	21.91%	4.92%
Color feature	84.58%	79.27%	78.72%	91.91%	12.77%	8.09%
Strength feature	82.72%	85.56%	79.37%	86.42%	13.16%	13.58%
DD feature	85.44%	76.23%	80.19%	91.81%	11.85%	8.19%
NV feature	86.15%	78.65%	82.90%	90.05%	9.29%	9.95%
Fusion with co-point	86.68%	88.08%	83.01%	91.07%	9.38%	8.93%
Fusion without co-point	87.51%	88.42%	83.60%	92.11%	9.08%	7.89%



图11. 具有模糊语义的区域的示例图。由红线围住的图像区域在狭义上不是“道路”，但是这些区域中的许多区域是为车辆行驶而设计的。

## 7. 结论和工作展望

本文提出了一种基于共点映射的自适应可行驶区域检测方法，该方法融合了像素信息和激光点的空间信息，计算了4个特征（DD、NV、颜色和强度特征），并将各个特征在贝叶斯框架中进行融合。基于数据融合，此方法克服了使用单个传感器在处理高度随机和复杂的城市交通场景时的缺点。此方法不需要强假设、训练过程或标记数据。此外，本文使用ROAD-KITTI benchmark 进行实验，证明了此方法的效率和鲁棒性。关于未来的工作，首要任务是将道路划分为可行驶区域和在紧急驾驶情况下可供行驶的区域。接下来，需要一个能够更好地处理模糊语义问题的可行驶区域数据集。最

后，需要使用现场可编程门阵列（field-programmable gate array, FPGA）实现此方法，以便在自动驾驶汽车上进行实时应用。

## 致谢

本研究得到了国家自然科学基金（No. 61773312）、国家重点研发计划（No. 2017YFC0803905）和学科人才引进计划（No. B13043）支持。

## Compliance with ethics guidelines

Ziyi Liu, Siyu Yu, and Nanning Zheng declare that

they have no conflict of interest or financial conflicts to disclose.

## References

- [1] Bar Hillel A, Lerner R, Levi D, Raz G. Recent progress in road and lane detection: a survey. *Mach Vis Appl* 2014;25(3):727–45.
- [2] Zhang G, Zheng N, Cui D, Yang G. An efficient road detection method in noisy urban environment. In: 2009 IEEE Intelligent Vehicles Symposium Proceedings; 2009. p. 556–561.
- [3] Cong Y, Peng JJ, Sun J, Zhu LL, Tang YD. V-disparity based UGV obstacle detection in rough outdoor terrain. *Acta Autom Sin* 2010;36(5):667–73.
- [4] Lee DT, Schachter BJ. Two algorithms for constructing a delaunay triangulation. *Int J Comput Inf Sci* 1980;9(3):219–42.
- [5] Fritsch J, Kuehnl T, Geiger A. A new performance measure and evaluation benchmark for road detection algorithms. In: Proceedings of International Conference on Intelligent Transportation Systems (ITSC); 2013 Oct 6–9; The Hague, the Netherlands. New York: IEEE; 2013.
- [6] Tan C, Hong T, Chang T, Shneier M. Color model-based real-time learning for road following. In: Proceedings of Intelligent Transportation Systems Conference. New York: IEEE; 2006. p. 939–44.
- [7] Rotaru C, Graf T, Zhang J. Color image segmentation in HSI space for automotive applications. *J Real-Time Image Process* 2008;3(4):311–22.
- [8] Jau UL, Teh CS, Ng GW. A comparison of RGB and HSI colour segmentation in real-time video images: a preliminary study on road sign detection. In: Proceedings of the 2008 International Symposium on Information Technology; 2008 Aug 26–28; Kuala Lumpur, Malaysia. New York: IEEE; 2008.
- [9] Finlayson GD, Hordley SD, Lu C, Drew MS. On the removal of shadows from images. *IEEE Trans Pattern Anal Mach Intell* 2006;28(1):59–68.
- [10] Maddern W, Stewart A, McManus C, Upcroft B, Churchill W, Newman P. Illumination invariant imaging: applications in robust vision-based localisation, mapping and classification for autonomous vehicles. Proceedings of the Visual Place Recognition in Changing Environments Workshop; 2014 Dec 5–10; Hong Kong, China. New York: IEEE; 2014.
- [11] Alvarez JM, Gevers T, LeCun Y, Lopez AM. Road scene segmentation from a single image. In: Proceedings of the 12th European Conference on Computer Vision: Volume Part VII; 2012 Oct 7–13; Florence, Italy. Heidelberg: Springer-Verlag Berlin; 2012. p. 376–89.
- [12] Teichmann M, Weber M, Zoellner M, Cipolla B, Urtasun R. MultiNet: real-time joint semantic reasoning for autonomous driving. 2016. arXiv:1612.07695.
- [13] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 2017;64(6):64–90.
- [14] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014. arXiv:1409.1556.
- [15] Badrinarayanan V, Handa A, Cipolla R. Segnet: a deep convolutional encoder-decoder architecture for robust semantic pixel-wise labelling. 2015. arXiv:1505.07293.
- [16] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; 2015 Jun 7–12; Boston, MA, USA. New York: IEEE; 2015. p. 3431–40.
- [17] Cong Y, Liu J, Yuan J, Luo J. Self-supervised online metric learning with low rank constraint for scene categorization. *IEEE Trans Image Process* 2013;22(8):3179–91.
- [18] Cong Y, Liu J, Fan B, Zeng P, Yu H, Luo J. Online similarity learning for big data with overfitting. *IEEE Trans Big Data* 2017;4(1):78–89.
- [19] Alvarez JM, Gevers T, Lopez AM. 3D scene priors for road detection. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2010; 2010 Jun 13–18; San Francisco, CA, USA. New York; 2010. p. 57–64.
- [20] Nan Z, Wei P, Xu L, Zheng N. Efficient lane boundary detection with spatial-temporal knowledge filtering. *Sensors* 2016;16(8):1276.
- [21] Hoiem D, Efron AA, Hebert M. Recovering surface layout from an image. *Int J Comput Vis* 2007;75(1):151–72.
- [22] Kong H, Audibert J, Ponce J. Vanishing point detection for road detection. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition; 2009 Jun 20–25; Miami, FL, USA. New York: IEEE; 2009.
- [23] Sivic J, Kaneva B, Torralba A, Avidan S, Freeman WT. Creating and exploring a large photorealistic virtual space. In: Proceedings of 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops; 2008 Jun 23–28; Anchorage, AK, USA. New York: IEEE; 2008.
- [24] Montemerlo M, Becker J, Bhat S, Dahlkamp H, Dolgov D, Ettinger S, et al. Junior: the Stanford entry in the urban challenge. *J Field Robot* 2008;25(9):569–97.
- [25] Thrun S, Montemerlo M, Dahlkamp H, Stavens D, Aron A, Diebel J, et al. Stanley: the robot that won the DARPA grand challenge. *J Field Robot* 2006;23(9):661–92.
- [26] Urmson C, Anhalt J, Bagnell D, Baker C, Bittner R, Clark MN, et al. Autonomous driving in urban environments: boss and the urban challenge. *J Field Robot* 2008;25(8):425–66.
- [27] Neidhart H, Sester M. Extraction of building ground plans from LiDAR data. *Int Arch Photogramm Remote Sens Spat Inf Sci* 2008;37(Pt 2):405–10.
- [28] Hu X, Rodriguez FSA, Geppert A. A multi-modal system for road detection and segmentation. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings; 2014 Jun 8–11; Dearborn, MI, USA. New York: IEEE; 2014. p. 1365–70.
- [29] Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 1981;24(6):381–95.
- [30] Wellington C, Courville A, Stentz AT. A generative model of terrain for autonomous navigation in vegetation. *Int J Robot Res* 2006;25(12):1287–304.
- [31] Klasing K, Wollherr D, Buss M. Realtime segmentation of range data using continuous nearest neighbors. In: Proceedings of 2009 IEEE International Conference on Robotics and Automation; 2008 May 13–19; Kobe, Japan. New York: IEEE; 2009. p. 2431–6.
- [32] Diebel J, Thrun S. An application of Markov random fields to range sensing. *Adv Neural Inf Process Syst* 2005;18:291–8.
- [33] Shotton J, Winn J, Rother C, Criminisi A. Textonboost: joint appearance, shape and context modeling for multi-class object recognition and segmentation. In: Proceedings of the 9th European Conference on Computer Vision: Volume Part I; 2006 May 7–13; Graz, Austria. Berlin: Springer; 2006. p. 1–15.
- [34] Xiao L, Dai B, Liu D, Tingbo H, Tao W. CRF based road detection with multi-sensor fusion. In: Proceedings of 2015 IEEE Intelligent Vehicles Symposium (IV); 2015 Jun 28–Jul 1; Seoul, Korea. New York: IEEE; 2015. p. 192–8.
- [35] Huang W, Gong X, Yang MY. Joint object segmentation and depth upsampling. *IEEE Signal Process Lett* 2015;22(2):192–6.
- [36] Alvarez AJM, Lopez AM. Road detection based on illuminant invariance. *IEEE Trans Intell Transp Syst* 2011;12(1):184–93.
- [37] Shinzato PY, Wolf DF, Stiller C. Road terrain detection: avoiding common obstacle detection assumptions using sensor fusion. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings; 2014 Jun 8–11; Dearborn, MI, USA. New York: IEEE; 2014. p. 687–92.
- [38] Ren X, Malik J. Learning a classification model for segmentation. In: Proceedings of the 9th IEEE International Conference on Computer Vision; 2003 Oct 13–16; Nice, France. New York: IEEE; 2003. p. 10–7.
- [39] Dollár P, Zitnick CL. Structured forests for fast edge detection. In: Proceedings of 2013 IEEE International Conference on Computer Vision; 2013 Dec 1–8; Sydney, NSW, Australia. New York: IEEE; 2013. p. 1841–8.
- [40] Zitnick CL, Dollár P. Edge boxes: locating object proposals from edges. In: Fleet D, Pajdla T, Schiele B, Tuytelaars T, editors. *Computer Vision—ECCV 2014*; 2014 Sep 6–12; Zurich, Switzerland. Cham: Springer; 2014. p. 391–405.
- [41] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, Süsstrunk S. SLIC superpixels compared to state-of-the-art superpixel methods. *IEEE Trans Pattern Anal Mach Intell* 2012;34(11):2274–82.
- [42] Geiger A, Lenz P, Stiller C, Urtasun R. Vision meets robotics: the KITTI dataset. *Int J Robot Res* 2013;32(11):1231–7.
- [43] Hu X, Rodriguez FSA, Geppert A. A multi-modal system for road detection and segmentation. In: 2014 IEEE Intelligent Vehicles Symposium Proceedings. New York: IEEE; 2014. p. 1365–70.
- [44] Wang T, Zheng N, Xin J, Ma Z. Integrating millimeter wave radar with a monocular vision sensor for on-road obstacle detection applications. *Sensors* 2011;11(9):8992–9008.
- [45] Liu J, Gong X. Guided depth enhancement via anisotropic diffusion. In: Proceedings of the 14th Pacific-Rim Conference on Multimedia; 2013 Dec 13–16; Nanjing, China. Berlin: Springer; 2013. p. 408–17.