

基于机器感知的智能车导航

王 宏 杨 明

清华大学计算机科学与技术系 智能技术与系统国家重点实验室 北京 100084

e-mail: wanghong@mail.tsinghua.edu.cn

摘要: 基于多传感器的机器感知是智能车辆导航的核心问题。它主要包括道路跟踪和障碍物检测等技术。本文结合大量的现有系统和我们的研究工作,对道路跟踪和障碍物检测的方法进行了深入的探讨,然后介绍了三个代表性系统,并对其最新实验结果进行了分析。在此基础上,对该领域今后的发展方向进行了展望。

关键词: 机器感知 道路跟踪 障碍物检测 计算机视觉 智能车辆导航

Machine Perception based Intelligent Vehicle Navigation

Wang Hong Yang Ming

State Key Lab. of Intelligent Technology and Systems, Tsinghua University, Beijing 100084

e-mail: wanghong@mail.tsinghua.edu.cn

Abstract: Multi-sensor based computer perception is a key problem of Intelligent Vehicle Navigation (IVN). It mainly includes road following (RF) and obstacle detection (OD). In this paper, some approaches for RF and OD are surveyed reviewing the existing systems and our research work. Then three typical systems are introduced with their latest experimental results. Some trends of IVN are presented according to the discussion above in the end.

Keywords: Machine perception, road following, obstacle detection, computer vision, intelligent vehicle navigation

1. 引言

智能车又称轮式移动机器人,是一个集机器感知、规划决策、智能驾驶等多种功能于一体的综合系统。除特殊潜在的军事价值外,还因其在公路交通运输中广阔的应用前景受到西方国家的普遍关注。几个工业发达国家已相继将智能车的研究纳入 20 世纪 90 年代和 21 世纪重点研究开发的智能交通系统和智能车路系统的重要组成部分。

智能车系统的机器感知必须同时具备实时性、鲁棒性和实用性这 3 个技术特点。实时性要求系统的传感器感知和数据处理必须与车体的高速行驶同步进行;鲁棒性则要求智能车对不同的道路环境、复杂的路面状况以及多变的气候条件均具有良好的适应性,这是智能车系统研究的难点所在;实用性要求系统在体积与成本等方面能够为普通用户所接受。上述技术特点构成了智能车系统研究所涉及的传感器信息处理、智能导航等主要关键技术。

在智能车导航多传感器感知的复杂任务中,最受重视的是基于视觉的道路跟踪和障碍物检测。目前的方法一般首先从图像中检测出道路信息和障碍物信息,即道路形状和与车辆的相对关系;然后,控制车辆在道路中安全行驶并躲避障碍物。本文首先分别讨论道路检测问题和障碍物检测问题,然后介绍三个代表性系统,并对其最新实验结果进行分析。在此基础上,对智能车辆导航的发展方向进行展望。

2. 道路检测

由于现实生活中的道路千差万别,因此道路检测是一个非常复杂的模式识别问题,目前仍不存在通用的道路检测系统。但是,适当地引入一些关于道路的基本假设将有助于问题的简化。下面是一些在道路检测中常用的基本假设。

道路形状假设: 简化道路模型的一种有效方法就是使用道路形状假设。如 VaMP 系统和戴姆勒-克莱斯勒公司的 UTA 系统都使用回旋曲线作为道路模型。同时一些研究小组使用多项式曲线或更一般化的道路模型。使用道路形状假设,避免了道路几何复杂的重构问题,简化了车体控制,提高了系统对阴影和遮挡的鲁棒性。但该假设往往需要复杂的公式匹配,而且当道路不符合模型假设时会失效。因此道路形状假设的选择至关重要。

道路宽度假设: 如果可以假设道路的宽度固定或变化缓慢,则道路检测可以简化成平

行的道路特征检测问题。韩国 Pohang 理工大学的 PVR III 系统和意大利帕尔马大学的 GOLD 系统都使用路面上的平行车道线来检测道路。美国 CMU 的 RALPH 系统进一步扩展了该方法, 使用任何平行于道路的特征来检测道路, 如道路边界、车辙等^[1]。使用道路宽度假设, 系统对阴影和车辆遮挡的鲁棒性明显提高, 但该假设不适于宽度变换频繁的道路。

路面特征一致假设: 通常, 图像中对应路面的区域具有一致的不同于非道路区域的特征, 例如灰度特征、颜色特征、纹理特征等, 因此可以采用聚类的方法检测出道路区域。如 FMC 公司的 Kuan 计算图像一维彩色特征的直方图, 通过阈值化将像素分类为“道路”像素和“非道路”像素, CMU 的 Thorpe 结合彩色特征和纹理特征对像素进行分类。使用路面特征一致假设, 检测困难道路的能力大大增强。但由于使用了聚类方法, 计算量较大。

兴趣区域假设: 尽管道路跟踪的实时性处理的运算量非常大, 但根据物理约束和连续性约束, 可以通过相邻帧之间的时间相关性加以简化: 只需在某些感兴趣的区域中进行分析 and 寻找期望特征, 而无需分析整幅图像。如德国 Autonomes Fahren 项目使用道路和车体动力学模型来确定车道线是否在图像中的兴趣区域内。类似地, GOLD 系统将待处理图像限制在上一次检测结果的邻域。使用兴趣区域假设可以显著加快道路检测的速度, 降低对硬件的要求。兴趣区域的选择是该假设的关键, 但兴趣区域假设不适合尺度较大的特征。

道路平坦假设: 出于车体控制的需要, 系统在获得图像特征之后, 需要将特征从图像坐标系转换到车辆坐标系。如果假设车辆前方道路是平坦的, 那么就可以利用已知的摄像机标定信息进行变换。例如 GOLD 系统和 RALPH 系统都使用了该假设: 首先根据道路平坦假设, 使用逆透视变换去除图像中的透视效果, 然后在得到的道路准俯视图上进行道路检测。使用道路平坦假设简化了图像坐标到车辆坐标的变换, 但在实际中道路平坦假设只能近似地满足, 或借助于摄像机的动态标定。

实际中的道路往往可分为结构化和非结构化道路两类, 在实际中需采取不同的方法。

结构化的道路具有清晰的车道线和道路边界。在这种情况下, 道路检测问题可以简化为车道线或道路边界检测问题。前面提到的 VaMP 系统、GOLD 系统、RALPH 系统、MOSFET 系统都是针对结构化道路的。它们通常具有较快的处理速度, 但只能在结构化道路上使用。而且这类系统中常用的梯度算子是典型的高通滤波器, 对噪声很敏感, 在处理退化的车道线和道路边界时会遇到困难。此外, 道路两旁的树木和建筑物、桥或其它车辆会在路面上产生阴影, 形成道路纹理。有时, 阴影边界的强度要远高于道路边界的强度。

非结构化道路一般指结构化程度较低的道路, 例如城市交通道路等。针对这类道路的系统除了前面提到的 PVR III 系统外, 还包括美国 CMU 的 SCARF 系统和 UNSCARF 系统, 美国 Martin Marietta 公司的 VITS 系统, 清华大学的 TURF 系统 (Tsinghua University Road Following)^[2]等。由于这类道路没有清晰的车道线和道路边界, 为避免由噪声或模糊的道路边界带来的问题, 这类系统一般使用路面特征一致假设, 因此也可用来检测结构化道路。非结构化道路同样存在阴影问题, 为此 VITS 系统和 SCARF 系统扩展了直方图阈值化方法, 使用多维分类的方法。但是它们需要使用前一帧图像的处理结果进行跟踪。

3. 障碍物检测

障碍物检测对于保证车辆的安全至关重要。道路上常见的障碍物包括车辆、摩托车、自行车、行人等。障碍物检测的方法很大程度上取决于障碍物的定义, 目前对障碍物尚无统一的定义。一些系统将障碍物限定为其周围的车辆, 而多数系统采用广义的定义: 障碍物是车辆行驶道路上任何具有一定高度的物体。一般地, 障碍物的检测方法可分为三类:

基于单幅图像的方法: 这类方法通常只用来检测特定的障碍物, 如周围的车辆。一般根据障碍物的特点, 使用适当的假设和简化模型来进行检测。由于仅使用单幅图像, 算法求解比较简单, 处理速度也较快, 且检测过程不受车辆运动的影响。但由于仅使用单幅图像, 失去了场景的深度信息。当模型和假设过于严格时, 可能会造成障碍物的漏检; 而当过于宽松时, 又可能造成障碍物的误检测。因此障碍物模型和假设的选择是该方法的关键。

基于光流的方法: 这类方法使用来自同一摄像机的一组图像序列。通常图像中的光流与运动场保持一致, 因此可用来检测运动的障碍物。先在图像中计算表征相邻帧间像素运动的光流场, 然后根据光流场估计出摄像机的主运动, 最后通过分析光流场中与主运动不一致的光流检测出障碍物。使用光流估计补偿主运动的计算量较大, 故一些系统直接使用从定位系统获得的车辆运动信息。基于光流的方法无需摄像机标定, 可计算出车辆的主运动和障碍

物的相对运动。但这类方法计算量较大，且当障碍物静止或运动速度很慢时失效。

基于立体视觉的方法：与基于光流的方法不同，这类方法使用来自不同视角的多个摄像机在同一时刻的多幅图像。使用立体视觉的最大优点是摄像机之间的相对关系确定，因此无需类似光流方法的主运动估计。传统的立体视觉使用基于三维重构的方法，但由于通常需要使用复杂的区域匹配，即在几幅图像中搜索像素对应关系，因此计算量也非常大，不适合实时应用。为解决这一问题，一些研究小组使用基于特征匹配的方法，典型代表是 UTA 实验车的智能 Stop&Go 系统和 MIT 的视觉系统。然而，更可靠的方法是使用道路平坦假设，称之为地平面立体视觉方法^[8]。由于道路约束的引入，立体图像的匹配可利用立体摄像机之间的相对关系获得，因此计算复杂度大大降低，非常适合实时的障碍物检测。基于立体视觉的方法能够检测广义障碍物，而且可进行三维重构。但除非使用相关的假设进行约束，一般来说其计算量较大，而且对车辆运动和摄像机标定漂移十分敏感，因此，摄像机的动态标定和障碍物的真伪判别是这类方法的重点。

地平面立体视觉方法：基于道路平坦假设，可使用如下方法检测障碍物：首先，根据两个摄像机之间位置关系（可预先通过摄像机的弱标定获得），将左右两幅立体图像投影到同一坐标系下，例如地平面坐标系或其中任何一幅图像坐标系。如果道路平坦且没有任何障碍物，则在此坐标系下，左右投影图像中同一位置上的像素将对应于路面上的同一区域，因此应该具有类似的灰度。而如果道路上存在障碍物，由于障碍物具有一定高度，因此障碍物上同一点在左右投影图像中的对应像素位置不同，由此产生了平面视差。所以对两幅投影图像取差并阈值化，得到的非零区域将意味着可能存在障碍物。

根据检测障碍物所处坐标系的不同可分为两种方法：第一种方法直接在其中一幅图像的坐标系中检测障碍物。如加州的 PATH 系统，东芝公司的 Hattori 和清华大学的朱志刚采用了类似的方法。该方法的优点是只需知道两个摄像机之间的相对关系。第二种方法是在地平面坐标系中检测障碍物。首先根据已知的左右摄像机相对地平面的关系，使用逆透视变换去除原始图像中的透视效果，将左右图像投影到地平面上得到准俯视图，然后对左右准俯视图取差并阈值化得到可能障碍物区域。GOLD 系统、TURF 系统^[3]都采用了这种方法。由于该方法需要使用左右摄像机与地平面之间的相对关系，因而对车辆的颠簸比较敏感，可能会扩大噪声的影响。但方法的最大优点是坐标系物理意义明确，可以直观地进行障碍物检测，并方便地使用障碍物的位置和大小等信息判断障碍物的危险度，而第一种方法往往需要间接地通过估计碰撞时间来进行判断。

4. 代表性系统和实验

一些研究小组已陆续将最新的解决方案和算法集成到智能原型车辆上，并在各种道路、天气情况下进行大范围测试。下面简要介绍其中最具代表性的三个系统及其最新实验结果。

德国慕尼黑联邦国防大学的 VaMP 实验车：由一辆豪华型奔驰 500 SEL 改装而成。视觉系统主要包括道路检测与跟踪 RDT 模块和障碍物检测与跟踪 ODT 模块。RDT 模块使用回旋曲线模型检测车道线，并通过 Kalman 滤波器跟踪车辆运动状态向量和道路的动态模型。ODT 模块用于检测周围的车辆，包括使用车辆对称性的单目标检测与跟踪和使用车辆阴影的多目标检测与跟踪。VaMP 系统能够实现车道保持、纵向控制、躲避障碍和换道操作等功能。自动驾驶速度设定值由驾驶员根据交通信号、交通情况和预期目标给出。此外，由于系统只使用前向双焦视觉系统，因此驾驶员还负责启动自动换道操作和监视后方车辆。

1995 年，VaMP 实验车进行了从德国慕尼黑到丹麦欧登塞的长距离实验。总行程约 1600 公里，其中 95% 是自动驾驶，执行了 400 多次换道操作。实验中的主要问题出现在道路的施工区域，因车道线十分模糊，道路检测失败。系统的主要不足是摄像机的工作范围有限，尤其在太阳直射的情况下。此外，单目标检测模块存在误报警问题，所使用的假设不仅车辆能够满足，有时其它目标也能满足，例如路面上的交通标志和偶尔出现的桥梁阴影。

美国 CMU 的 NabLab 5 实验车：由 Potiac 运动跑车改装而成，装备有便携计算机、摄像头、GPS 全球定位系统、雷达等设备。系统运行在 PANS 硬件平台之上，该平台为系统提供了计算平台、IO 接口、位置估计、方向盘控制和安全监视等功能。NabLab 5 的核心是基于视觉的 RALPH 系统^[1]。该系统中道路检测使用道路平坦假设和平行于道路的特征，例如车道线、道路边界、车辙等。因此 RALPH 系统的最大优点是能够在车道线欠佳的情况下进行道路跟踪。障碍物检测主要依靠来自激光雷达的距离信息。根据 RALPH 系统输出道路曲

率和障碍物信息，系统控制方向盘使车辆保持在车道内，而油门和刹车由人来控制。

1995 年 6 月，NabLab 5 进行了横穿美国的实验 NHAA (No Hands Across America)，从宾州的匹斯堡到加州的圣地亚哥。总行程 4587 公里，其中自动驾驶占 98.2%，最长连续自动驾驶距离为 111 公里，全程平均速度为每小时 102.72 公里。实验证明该系统对实际道路具有极强的鲁棒性。实验中遇到的主要问题包括雨天、太阳位置较低时路面上形成反射、天桥的阴影、施工区域、退化的道路边界和车道线等。

意大利帕尔马大学的 ARGO 实验车：由一辆兰西亚 Thema 客车改装而成，装备有摄像机、霍尔效应传感器、IO 接口板和奔腾 200MMX 的 PC 机，使用 Linux 操作系统。ARGO 的核心是 GOLD 视觉系统，它基于道路平坦假设，使用单目视觉检测平行的车道线，双目立体视觉检测障碍物。根据 GOLD 系统的输出，ARGO 实验车能够自动控制方向，实现道路跟踪。此外还能够实现人触发的换道操作。GOLD 的处理结果也通过车载显示器和基于 LED 的控制面板传递给驾驶员，当检测到危险时还可通过语音设备向驾驶员报警。

ARGO 实验车于 1998 年 6 月进行了 2000 公里的长距离实验。路线包括大量的高架桥和隧道。受交通堵塞、收费站、交叉路口和施工道路影响，平均时速为 90 公里，最高时速 123 公里。尽管在通过收费站和繁忙交通地段时系统必须关闭，ARGO 系统仍自主驾驶了大约 94%，最长连续自动驾驶距离为 54.3 公里，显示出较高的性能。遇到的主要问题是太阳光在挡风玻璃上的反射对安装在玻璃后的摄像机影响较大，造成采集的图像过饱和。此外在隧道的出入口处摄像机无法适应环境光照的快速变化，从而造成图像质量较差。

5. 结论和展望

从以上鼓舞人心的实验结果中可以得出结论：所有系统都在自动驾驶的绝大部分时间取得了成功，这说明基于视觉的智能车辆导航在技术上是完全可行的。在实验中所暴露出的一些问题也正是今后研究的发展方向^[3]。

受益于计算机硬件的飞速发展，除 VaMP 系统外，其它系统均采用了商业化的硬件平台。但需要强调的是，专用硬件的研制并非没有意义。目前的趋势是使用混合式结构，即使用专用硬件来辅助强大的通用处理器。

另外，目前的传感器仍无法完全满足视觉导航的需要。首先传感器的动态范围较低。常规摄像机的动态范围为 500:1，而人眼的动态范围为 $10^5:1$ ，如果考虑人眼的光线自适应能力，则可达到 $10^9:1$ 。而视觉导航要求单幅图像至少具有 $10^5:1$ 的动态范围。因此，必须提高视觉传感器的动态范围。一种解决办法是使用具有对数响应的 CMOS 传感器。CMOS 传感器没有过度曝光的光晕问题，动态范围较大，鲁棒性强，便于实现芯片级集成，是理想的全天候传感器，但这类传感器固有的低速问题仍有待解决。其次，视觉传感器的分辨率和视角无法同时满足视觉导航的要求。常用摄像机的视角仅为 30° ，在车辆转弯时明显不足。而鱼眼摄像头虽然视角可以达到 180° ，但其图像为非线性，且分辨率较低。

从道路检测的方法来看，虽然新方法的探索仍在继续，但每种方法都存在一些不足^[4]，因此融合将是解决问题的出路所在，包括硬件层次上的融合，即传感器数据融合，由其它视觉传感器获得的深度图像能弥补目前摄像机某些方面的不足；同时也包括软件层次上的融合，即方法和算法上的融合。目前真正能同时集成道路检测和障碍物检测的系统并不多，这种集成的解决方案必能同时提高道路和障碍物检测的质量，从而加快处理速度和效率。

目前，结构化道路的跟踪问题方法已趋向成熟。研究人员已将注意力转向低速复杂的城市交通环境。例如，欧洲最新的 CARSENSE 计划 (Sensing of Car Environment at Low Speed Driving) 强调其创新的重点之一就是“为辅助驾驶系统引入复杂的场景，开辟通向密集城市交通自动驾驶的道路”。因此，诸如交叉口检测、行人检测等问题将成为新的热点和难点。

参考文献

- [1] D. Pomerleau, T. Jochem, Rapidly Adapting Machine Vision for Automated Vehicle Steering, IEEE Expert, April 1996, 11(2): Page(S): 19-27
- [2] 杨明, 陆建业, 王宏, 张钹. 基于视觉的道路跟踪. 《模式识别与人工智能》, Vol. 14, No. 2, 2001, 186-193.
- [3] 王宏, 何克忠, 张钹. 智能车辆的自动驾驶与辅助导航. 《机器人》 Vol. 19, No. 2, 1997, 155-160.
- [4] M. Bertozzi, A. Broggi, A. Fascioli, Vision-Based Intelligent Vehicles: State of the Art and Perspective, Robotics and Autonomous Systems, 2000, 32: 1-16.