

湖南大学博士研究生毕业（学位）论文开题报告

姓名	田峥	学号	B0910T0015	已修学分	8
所属学院	计算机与通信学院			一级学科	计算机科学与技术
二级学科	计算机应用技术				
指导教师	李仁发 教授			开题时间	2010年 7月
研究方向	计算机视觉、模式识别				
论文题目	无人驾驶汽车上视觉感知计算的软硬件关键技术研究				

一、文献综述

无人驾驶汽车作为智能交通技术的综合集成，在军事和民用领域具有广阔的应用前景，受到发达国家的高度重视。视觉感知技术相当于无人驾驶汽车的眼睛，在导航及目标检测等方面具有无可替代的作用，成为各科研机构研究的热点。

无人驾驶汽车领域的研究工作可追溯到 20 世纪 70 年代初期的移动机器人研究。20 多年来，世界上许多知名科研院所和企业无人驾驶汽车的软硬件领域展开了大量的基础性研究，从相关研究的过程来看，无人驾驶技术的发展可以分为以下两个阶段：

在第一阶段，由于受到当时硬件计算能力的制约，而图像处理的计算量非常大，因此必须使用昂贵的专用硬件处理单元才能满足实时计算的要求。如德国慕尼黑联邦国防军大学的 VaMoRs-P 系统使用基于晶片机（Transputer）结构的并行处理板^[18]，美国卡内基-梅隆大学的 NavLab I 系统使用 16k 的 MP-2 并行机^[19]，意大利帕尔玛大学的 GOLD（General Obstacle and Lane Detection）系统使用大规模并行结构 PAPERICA^[20]。这一阶段对视觉感知算法的研究还只是停留在结构化道路的检测与跟踪上，而研究者们把更多的精力放在了硬件计算平台的设计、测试和验证上。

在第二阶段，随着计算机硬件水平的飞速发展，处理器和存储器价格不断下降的同时，民用计算机的体积不断变小，能耗不断降低，性能也有了显著的提升。而且新一代处理器的指令集中已经包括了 SIMD（Single Instruction Multiple Data）指令，这些指令特别适合多媒体数据的并行处理^[21]。这些都使得基于图像处理的视觉感知算法在通用的标准 PC 机上进行计算成为了可能。于是，几乎所有的研究小组都将他们的系统移植到了商业化的硬件平台上。如 VaMoRs-P 系统使用 3 台双 Pentium II 计算机代替了原有的 Transputer 并行处理系统^[22]，GOLD 系统使用 1 台 Pentium 200M 的 PC 机代替了原来的 16*16 的并行结构 PAPERICA^[23]。这一阶段研究者的主要精力都从对硬件平台的研究转移到了对视觉算法的研究上，算法的研究对象也逐渐扩大，从最初的道路识别、障碍物识别延伸到了行人检测、交通标志、信号灯识别等领域。

根据不同时间段所从事研究工作的侧重点不同，可以将无人驾驶汽车领域的研究工作分为两个主要的方向：其一是对视觉感知算法的研究，属于模式识别和计算机视觉的范畴，研究内容主要包括道路识别、障碍物检测、车辆和行人检测等算法，以提高算法的准确性和可靠性为目的；另一个方向主要是对无人驾驶汽车的硬件计算平台的研究，属于计算机硬件技术和系统结构的范畴，研究的主要内容包包括高性能计算、并行与分布式处理、流水线技术等，以提高无人驾驶汽车验证平台的计算能力，降低平台计算功耗为目的。下面分别对无人驾驶汽车领域两个研究方向的发展现状做详细介绍。

1) 视觉感知算法

基于视觉的道路识别是无人驾驶汽车导航的核心问题。自无人驾驶汽车诞生以来，基于视觉的道路识别算法的研究就从未间断过，该领域的研究现状在一定程度上代表了整个无人驾驶汽车上视觉感知计算的发展水平。

由于现实生活的道路情况千差万别，因此基于视觉的道路识别问题是一个非常复杂的模式识别问题，目前仍然没有出现一个比较通用的道路检测系统。根据道路本身的结构特征，可以将道路分为结构化道路（Structured Road）和非结构化道路（Unstructured Road）。结构化道路通常是指日常生活中的高速公路这一类的道路，其具有明显的结构化特征：道路标志线比较明显，道路基本是直的（转弯处的弯度较小）等，这些特征受周围环境的影响较小。非结构化道路的结构化程度要显著低于结构化道路，而且其路面特征与周围环境差别较小甚至几乎没有，通常是指野外、沙滩、森林等一类的自然环境。

为了简化问题，各研究机构都引入了基本的条件假设来提高识别的准确率。即使是信息最为丰富、识别难度最低的结构化道路也都如此。这些基本假设包括：道路模型假设、路面特征一致假设、道路平坦假设、道路宽度假设、兴趣区域假设等。其中路面特征一致假设和道路平坦假设为两个最为基本的假设，几乎所有的应用系统^[24-26]都默认了这两个基本假设。在道路模型方面，各研究机构根据自己的需要引入了不同的模型，这些模型主要包括双曲线模型，平行线模型，三角形模型，梯形模型等。这些几何模型的提出给道路的识别以及结果的验证提供了一个评价的标准。LOIS 系统^[27]和 LANA 系统^[28]使用回旋曲线作为道路模型，由于回旋曲线上任意点的曲率与到原点的距离成正比，所以只需要两个参数就能确定道路的形状。另外，意大利 Parma 大学的 MOB-LAB 试验车则使用了更一般的三角模型^[29]。

帕尔玛大学的 GOLD 系统^[24]和卡内基-梅隆大学的 YARF 系统^[25]是结构化道路识别中两个比较典型的系统。在 GOLD 系统中，算法主要通过检测车道线来做导航。算法基于道路是等宽的假设，将道路的车道线建模为一组平行线。因此，算法首先通过摄像机标定参数将拍摄到的透视图片做逆透视变换，转化为鸟瞰图，以保证车道线在图片中也是平行的，这一转化过程隐式的引入了道路是平坦的这一假设，因为，如若道路是不平坦的，那么逆透视之后，图像中的车道线也不会是平行的。类似的方法也被 RALPH^[26]系统所使用。此后，算法使用 Black-White-Black 算法来提取车道线特征。该方法认为，车道线上的任意一像素点的亮度都会比其一定距离的左、右区域内的像素点值要高。该方法充分使用了结构化道路的几何特征和路面特征；因此，在结构化道路环境下，该方法具有较高的鲁棒性。但是，当车道线不十分明显，或者道路的宽度发生变化时，使用该方法将会导致检测的失效。YARF 系统则更多的关注了道路的几何特征，认为道路为地平面上垂直于“脊椎曲线”的横截面。据此，算法建立了一个线性的道路模型，并通过统计的方法来得到道路的参数。

总的来说，结构化道路识别的方法都以准确提取出车道线特征为前提；不能准确提取出车道线将会直接导致道路识别的失败。因此，这类方法失效于车道线特征不明显的非结构化道路。

目前，非结构化道路主流的检测算法大致有两种：基于图像分割的方法和基于模板匹配的方法。在基于图像分割的非结构化道路识别算法中，最大类间方差分割（OTSU）方法备受青睐^[30]。OTSU 是一个以概率统计为基础的算法，它通过引入对图像分割效果的评判参数，计算最优化的分割阈值 k ，从而实现对灰度图像二值化处理。该方法的计算简单，效率很高，因此许多非结构化道路识别的算法都是用了该方法或者其变体。OTSU 方法的缺点是只能将图像分为两类，在复杂环境下，其分割结果往往差强人意。

卡内基-梅隆大学的 SCARF 系统^[31]和 UNSCARF 系统^[32]则向我们展示了另一种非结构化道路的方法：基于模板匹配的方法。UNSCARF 系统使用 ISODATA（Iterative Self-Organizing Data Analysis Technique）分类算法将相似颜色和位置的像素归为一类，然后再匹配道路形状模型来确定道路模型。它可以检测没有车道线、道路表面腐蚀的道路，且不需要任何有关道路的先验知识。SCARF 是一个彩色的视觉系统，它使用了贝叶斯分类算法进行分类，在此基础之上，使用经典的模式识别方法来确定道路表面的相似度，从而来检测道路。它不仅仅可以检测一般的道路，也是第一个可以检

测到十字路口的系统，开辟了城市自主导航的先河。算法首先根据颜色信息分别对道路区域和非道路区域进行高斯建模，然后利用贝叶斯定理计算出一系列的最有可能的道路候选区域，最后通过模板匹配的方法从这些候选区域中选择相似度最大的区域作为道路区域。

总的来说，基于图像分割的方法计算量较小，实时性较好，但是算法对阴影、路面腐蚀等情况的适应性较小。相对而言，基于模板匹配的方法鲁棒性较好，但是其计算量较大，对计算平台处理能力的要求较高。从道路识别算法发展的时间轴上来看，目前的研究主要集中在三个方向：1、研究利用新的方法进行道路识别的算法，如 Xu 等人在他们的作品中使用了 CNN (Cellular Neural Network) 技术来提高道路识别算法的效率^[33]；Lee 等人使用了 SVM 技术，算法先使用道路表面的颜色来训练 SVM，再使用其来检测道路^[34]；2、利用更有效的图像分割技术和概率统计方法来提高基于道路识别算法的鲁棒性；3、研究降低基于模板匹配道路算法的计算复杂度的方法，或者针对算法的计算特性研究专用的计算结构，以满足实时性的要求。而本文将重点从第三个方向入手展开对无人驾驶汽车上视觉感知算法的研究。

2) 硬件计算平台

纵观无人驾驶汽车上硬件体系结构的发展，大致经历了一个从专用到通用的发展过程。在无人驾驶汽车发展的最初阶段，由于图像处理的计算量非常大，而当时的单个处理器单元的计算能力非常有限，因此不得不采用专用的硬件方案才能满足视觉算法的计算需求。研究者们把大量的精力都放在了多处理器单元的体系结构、并行计算和测试验证的研究上。比较有代表性的工作是德国慕尼黑联邦国防军大学的 VaMoRs-P 系统^[35]和帕尔玛大学的 GOLD 系统^[36]。

VaMoRs-P 系统的计算平台由基于 Transputer 的并行处理单元和一台 PC 机组成。Transputer 并行处理单元由大约 60 个 T-222 型 (16 位，主要负责图像处理和数据通信) 和 T-800 型 (32 位，主要负责复杂数字运算和图像理解) Transputer 构成，PC486 作为整个 Transputer 系统的控制单元和人机交互接口，通过特殊的总线接口连接到并行处理单元。后来，随着通用处理器的发展，Dickmanns 等人将 Transputer 系统中的部分处理单元用新一代的微处理器 Motorola PowerPC 601 代替，让 MPC 专门负责视觉计算的部分，而 Transputer 单元只需要负责数据的通信，从而大大简化了 VaMoRs-P 系统的硬件结构^[37]。而随着 Intel Pentium 系列处理器的诞生，VaMoRs-P 被移植到了 3 台基于 Dual-Pentium Pro 的计算机上^[22]。

GOLD 系统最初也是采用专用的硬件结构，其计算平台由一个被称为 PAPERICA 的包含 16*16 个 PE (Processing Element) 的 SIMD 并行计算单元和一台 Sparc 的工作站主机构成。PAPERICA 集成了图像采集、显示接口、形态学处理、图像重映射等多个功能单元，主要负责道路和障碍物检测算法的底层图像处理部分，而中上层的算法处理则交由主机工作站来完成。PAPERICA 和主机工作站之间采用流水线的机制协同工作，PAPERICA 将部分处理后的中间结果交给工作站继续处理，从而实现算法的并行执行。整个系统的运行效率取决于 PAPERICA 部件的处理速度，由于摄像头采集的速度是 50 帧每秒 (fps)，因此每个时间片的长度为 20ms，PAPERICA 可以在 5 个时间片内处理完 1 帧的数据，因此整个系统的处理速度可以达到 10fps。

同样，随着通用处理器的指令集中对 SIMD 指令的支持，GOLD 被移植到了一台具备 MMX 指令集的 Pentium 200MHz 处理器上，只是由于当时的编译器还不支持基于 MMX 的指令集，因此只能通过直接写汇编语言来保证算法的并行性。但即使整个系统没有实现最优化，其执行速度也至少比以前快了一个数量级^[23]。

由此可见，随着通用处理器性能的提升，价格的降低，无人驾驶汽车的硬件结构从专用平台发展到通用平台是一个必然趋势。无人驾驶汽车领域的研究也逐渐从对硬件结构的研究转移到了对视觉算法的研究。但是，随着无人驾驶汽车上的传感器设备越来越多，所使用的算法越来越复杂，为了满足计算需求，保证计算的实时性，研究者们开始将大型服务器机组、千兆级网络交换设备搬到了无人驾驶汽车上，而计算功耗也随之大幅上涨，专用的车载发电机和空气冷却装置也派上了用场。

清华大学的 THMR-V 系统装备有彩色摄像头、GPS、磁罗盘光码盘定位系统、激光测距仪等设

备。计算平台采用两台 Pentium II 计算机，其中一台进行视觉信息处理，另一台完成信息融合、路径规划、行为控制、决策控制等功能。四台 IPC 工控机分别完成激光测距信息处理、定位信息处理、通讯管理、驾驶控制等功能^[38]。国防科技大学的 CITAVT-IV 系统使用了一台 Pentium III 533MHz 计算机来做视觉计算，使用 Motorola 的 Delta 机作为主控平台，包括一块 VME-167 进行系统的开发和调试；三块 VME-162 来完成数据处理任务；另外还用了一台笔记本电脑来实时监控系统的状态，并完成和操作人员的必要交互工作^[14]。整个系统的总功率接近两千瓦，采用一套专用的供电组合对系统进行供电，包括一个三千瓦的车载直流发电机，一个蓄电池组和两个 220V 逆变器。

2007 年 DARPA Urban Challenge 比赛的冠军，卡内基-梅隆大学的 BOSS 系统，配备了 11 个激光雷达、5 个微波雷达、2 个摄像头和 1 个差分 GPS，计算设备采用了 10 台 2.16G 的酷睿双核服务器阵列；整个系统仅仅是计算部分的功耗就超过了 500W^[39]。类似的，在 2009 年第一届中国“智能车未来挑战”赛中，湖南大学的参赛车队使用了包括摄像头、激光雷达、GPS 等多种类型的传感器，采用 4 台酷睿双核的笔记本作为计算平台，而西安交通大学和武汉大学的参赛车队则采用刀片服务器作为计算平台。这些系统都配备了专用的供电装置。

在现阶段，无人驾驶汽车领域的研究者们大都只追求算法的效果，而几乎不去考虑硬件计算平台的成本和功耗。这也就导致了计算设备的规模越来越大，功耗越来越高。但是，无人驾驶技术在实际应用中通常对计算平台有成本、体积和功耗的要求，一旦软硬件的系统原型都已经构建成熟，使用最优化的嵌入式计算引擎就成为了必要。以色列耶路撒冷希伯来大学是较早的针对视觉感知算法研究嵌入式专用计算平台的研究机构，他们在 2004 年与 MobilEye 公司合作研制出了一款用于道路、车辆和行人检测的专用 ASIC 芯片 EyeQ，芯片包含一个通用的 ARM 核和 4 个专用的图像处理单元^[12]；并在 2008 年研制出了 EyeQ 的升级版 EyeQ2；帕尔玛大学的 VisLab 实验室也在 2008 年研制出了一款基于 TI DM642 的 DSP 的 VIDA 视频处理模块^[17]。

另外，嵌入式环境下的算法开发并不像通用的计算平台那样有完整的开发工具链和可视化的调试环境，随着无人驾驶汽车上用于视觉感知计算的专用部件的产品化、实用化程度越来越高，开发人员对算法开发的高效性和便捷性的要求也会越来越高。因此，在专用计算平台之上研究和设计一套可为实时视觉应用的快速开发提供完整工具集的软件基础框架变得势在必行。帕尔玛大学的 VisLab 实验室就研制出了一套用于无人驾驶汽车上视觉应用开发的完整框架^[40]。

结合国内外相关的研究现状可以看到，目前无人驾驶汽车上计算平台规模大、能耗高的问题已经成为了制约无人驾驶汽车领域的研究从实验室走向实际应用的一个关键问题。随着嵌入式计算能力的不断提升，视觉感知算法在专用的嵌入式平台上进行计算成为了可能。有理由相信，无人驾驶汽车上硬件体系结构的发展将会经历一个从通用计算平台到高性能专用计算平台的发展过程。而且在无人驾驶汽车的研究领域，面向视觉感知算法的专用嵌入式计算引擎及其相关的软件开发基础框架的研究也将会越来越受到各研究机构的重视。目前这方面的相关研究还处于探索阶段，国外的少数研究机构已经取得了一些初步成果，而在国内的相关研究并不多见。因此，在目前视觉感知算法逐渐发展成熟，无人驾驶技术开始走向实际应用的情况下，开展面向视觉算法的专用嵌入式计算引擎及其软件开发基础框架的相关研究是必要而且迫切的。

综上所述，本文拟针对无人驾驶汽车上视觉感知算法的特定应用，以提高视觉感知算法的计算性能、降低无人驾驶汽车计算平台的功耗，并提升算法开发人员的设计效率为目标，研究算法的计算特性，设计专用的嵌入式计算引擎，研究实时视觉感知算法的设计流程和方法，提出一个面向视觉感知算法的软件开发基础框架结构。可以设想，如能取得预期效果，本文的开展将对促进无人驾驶汽车上视觉感知算法研究的发展，推广无人驾驶汽车验证平台的实际应用发挥一定的作用。

参考文献

- [1] Ichiro Masaki, "Machine-Vision Systems for Intelligent Transportation Systems", IEEE Intelligent

- Systems & Their Applications magazine, 1998, 13(6):24-31
- [2] Bishop R. A Survey of Intelligent Vehicle Application Worldwide. In: Proc. of IEEE Intelligent Vehicle Symposium, Dearborn, USA, 2000:25-30
- [3] <http://www.cs.cmu.edu/afs/cs/project/alv/www/index.html>
- [4] http://en.wikipedia.org/wiki/EUREKA_Prometheus_Project
- [5] M.Bertozzi, A.Broggi, M.Cellario, et al. Artificial vision in road vehicles. In: Proc. of IEEE Special Issue on Technology and Tools for Visual Perception, vol.90, no.7, pp.1258–1271, Jul.2002
- [6] U. Franke, D. Gavrila; S. Gorzig, et al. Autonomous Driving Goes Downtown, IEEE Intelligent Systems, 1998, 13(6): 40– 48
- [7] M. M. Meinecke and M. A. Obojski, Potentials and limitations of precrash systems for pedestrian protection, in Proc. 2nd Int. Workshop Intell. Transp. Syst., 2005.
- [8] Fengliang Xu, Xia Liu, Fujimura K. Pedestrian detection and tracking with night vision. IEEE Transactions on Intelligent Transportation Systems(S1524-9050), 2005, 6(1):63-71
- [9] Jochem, T., Pomerlean, D., Kumar, D. and Armstrong, J., PANS: a portable navigation platform. In: Proceedings of the IEEE Symposium on Intelligent Vehicles, 1995:107-112
- [10] S. Thrun, M. Montemerlo, H. Dahlkamp, et al. Stanley: The robot that won the darpa grand challenge: Research articles. J. Robot. Syst., vol. 23, no. 9, pp. 661-692, 2006
- [11] Massimo Bertozzi, Alberto Broggi, and Alessandra Fascioli, Vision-based Intelligent Vehicles: state of the art and perspectives, Journal of Robotics and Autonomous Systems, 32(1):1-16, June 2000
- [12] Stein, G.P., Rushinek, E., Hayun, G., et al. A Computer Vision System on a Chip: A case study from the automotive domain. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005:130-130
- [13] 张鹏飞, 何克忠, 欧阳正柱等. 多功能室外智能移动机器人实验平台——THMR-V. 机器人. 2002, 24(2): 97-101
- [14] 孙振平, 安向京, 贺汉根. CITAVT-IV——视觉导航的自主车. 机器人, 2002, 24(2): 115-120
- [15] Hong Cheng; Nanning Zheng; Xuetao Zhang; et al. Interactive Road Situation Analysis for Driver Assistance and Safety Warning Systems: Framework and Algorithms. IEEE Transactions on Intelligent Transportation Systems, March 2007, 8(1):157-167
- [16] <http://ccvai.xjtu.edu.cn/mes.do?method=list>
- [17] VisLab, VisLab Embedded Lane Detector Technical Information, July 2009, http://vislab.it/resources/Products/veld/VELD_Specs.pdf
- [18] Maurer M., Behringer R., Dickmanns D., et al. VaMoRs-P: An advanced platform for visual autonomous road vehicle guidance. In Proc. of SPIE, Boston, November 1994:239-248,
- [19] Todd Jochem and Shumeet Baluja, A Massively Parallel Road Follower. In: Proceedings of Computer Architectures for Machine Perception, December, 1993:2-12
- [20] Broggi A, Conte G, Gregoretti F, et al. Design and Implementation of the PAPRICA Parallel Architecture, The Journal of VLSI Signal Processing, 1998, 19 (1):5-18
- [21] 杨明, 陆建业, 王宏等, 基于视觉的道路跟踪, 模式识别与人工智能, 2001, 14(2): 186-193
- [22] Dickmanns E.D., Vehicles capable of dynamic vision: a new breed of technical beings?, AI(103), No. 1-2, August 1998, pp. 49-76
- [23] M. Bertozzi, A. Broggi, A.Fascioli. Obstacle and lane detection on ARGO autonomous vehicle, In: Proc. of IEEE Intelligent Transportation Systems Conf.'97, Boston, MA.
- [24] Massimo Bertozzi and Alberto Broggi, Vision-based Vehicle Guidance, IEEE Computer, July 1997, 30(7):49-55,

- [25] Kluge, K. and C. Thorpe. Representation and recovery of road geometry in YARF. in Intelligent Vehicles '92 Symposium., Proceedings of the. 1992.
- [26] Pomerleau, D. RALPH: rapidly adapting lateral position handler. in Intelligent Vehicles '95 Symposium., Proceedings of the. 1995
- [27] Kluge K., S. Lakshmanan. A deformable-template approach to lane detection. in Proc. IEEE Intelligent Vehicles Sym., Detroit, MI, Sep.25-26, 1995, pp.54-59
- [28] C. Kreucher, S. Lakshmanan, LANA: a lane extraction algorithm that uses frequency domain features. Robotics and Automation, IEEE Transactions on, vol. 15, pp. 343-350, 1995
- [29] Alberto Broggi, Simona Berte, Vision-Based Road Detection in Automotive Systems: a Real-Time Expectation-Driven Approach, Journal of Artificial Intelligence Research, 3:325-348, December 1995
- [30] M. Sezgin, B. Sankur. Survey over image thresholding techniques and quantitative performance evaluation. Journal of Electronic Imaging 13(1):146-165, 2004
- [31] Crisman, J.D., C.E. Thorpe, SCARF: a color vision system that tracks roads and intersections. Robotics and Automation, IEEE Transactions on, 1993. 9(1): p. 49-58.
- [32] Crisman, J.D., C.E. Thorpe. UNSCARF - a color vision system for the detection of unstructured roads. in Proceedings of IEEE International Conference on Robotics and Automation. 1991
- [33] Guobao Xu, Yixin Yin, Lu Yin, et al. Visual information processing using cellular neural networks for mobile robot. In: Proc. of 2007 IEEE International Conference on Grey Systems and Intelligent Services, November 18, 2009
- [34] Brooks, C.A., K. Iagnemma. Visual detection of novel terrain via two-class classification. In: Proceedings of the 2009 ACM symposium on Applied Computing. March 2009:1145-1150
- [35] Dickmanns E.D., Behringer R., Dickmanns D., et al. The Seeing Passenger Car 'VaMoRs-P'. In: Proc. of IVS, 1994:68-73
- [36] M. Bertozzi, A. Broggi. GOLD: a Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection. IEEE Transactions on Image Processing, 1997
- [37] MAURER M, BEHRINGER R, THOMANEK F, et al. A Compact Vision System For Road Vehicle Guidance, In: Proc. of the 13th International Conference on Pattern Recognition, 1996
- [38] 李华, 丁冬花, 何克忠. THMR-V 导航控制算法的研究. 机器人, 2001, 23(6):525-530
- [39] C. Urmson, J. Anhalt, D. Bagnell, et al. Autonomous driving in urban environments: Boss and the urban challenge. Journal of Field Robotics, vol. 25, no. 8, pp. 425-466, 2008
- [40] Massimo Bertozzi, Luca Bombini, Alberto Broggi, et al. GOLD: A Complete Framework for Developing Artificial Vision Applications for Intelligent Vehicles, IEEE Intelligent Systems, 23(1):69-71, January-February 2008

(本表可附页)

二、选题背景及意义

智能交通系统是现代科技手段在交通领域的具体应用，是以缓和道路堵塞、减少交通事故、提高道路使用者的方便舒适为目的，利用信息通讯技术实现的交通系统的总称，是未来道路交通发展的方向。无人驾驶汽车（Unmanned Ground Vehicles, UGVs）是智能交通系统中一个非常重要的研究领域，融合了传感器、自动控制、计算机视觉，计算机体系结构等多项技术，是计算机科学、人工智能和智能控制等学科高度发展的产物^[1]。无人驾驶技术在军事国防、汽车制造、城市交通、公共安全等领域具有广阔的应用前景和潜在的实用价值，因此得到了世界许多国家的高度重视^[2]。发达国家从 20 世纪 70 年代开始进行无人驾驶汽车研究，美国国防部在 20 世纪 80 年代初期就开始大规模资助自主陆地车辆 ALV（Autonomous Land Vehicle）的研究^[3]；欧洲也启动了世界上最早大规模研究智能地面车辆的 Eureka-PROMETHEUS 项目^[4]，欧洲几乎所有的轿车生产厂家和大部分的著名研究院所、大学都参与了这一项目，目前在可行性和实用化方面都取得了突破性进展。

视觉感知技术相当于无人驾驶汽车的眼睛，它在行车导航及障碍物检测等方面具有无可替代的作用，是被公认的最具发展潜力的感知技术。据统计，无人驾驶汽车中 90% 的环境信息是通过视觉来感知的^[5]。无人驾驶汽车上的视觉感知技术不同于以往传统的静态场景下的视觉分析技术，它需要面对各种复杂多变的外部环境，需要从高速运动的视觉传感器中获取大量的动态信息，因此无人驾驶汽车的实际应用场景对视觉感知算法的稳定性和适应性要求非常高，其算法具有高复杂度、高计算量的特点。正因为视觉感知技术在无人驾驶汽车上的重要性和其算法的复杂性，它成为了计算机领域的一个研究热点。

近年来，一些企业和科研院所围绕道路识别、障碍物检测、行人车辆检测、多摄像头数据融合等方面开展了深入研究，在 IEEE（Institute of Electrical and Electronics Engineers）和 ACM（Association of Computer Machinery）相关会议发表了一些重要的研究成果。无人驾驶汽车领域的顶级会议 IVS（IEEE Intelligent Vehicles Symposium）每年都会跟踪无人驾驶汽车上视觉感知技术的最新进展，人工智能与模式识别领域的权威期刊如 TPAMI（IEEE Transactions on Pattern Analysis and Machine Intelligence）、IJCV（International Journal of Computer Vision）等和重要的国际会议如 ICCV（International Conference on Computer Vision）、CVPR（IEEE Computer Society Conference on Computer Vision and Pattern Recognition）等也都将无人驾驶汽车上视觉感知技术的研究作为主题内容之一。此外，戴姆勒-奔驰^[6]、大众^[7]、本田^[8]等汽车生产厂商也已投身到视觉感知技术的研发当中，并投入巨资建立了各自庞大的研究计划。卡内基-梅隆大学^[9]、美国斯坦福大学^[10]等著名学府也一直在关注相关的研究工作，启动了一系列的研究项目；意大利的帕尔玛大学^[11]、以色列耶路撒冷希伯来大学^[12]等研究机构也发表了相关研究成果；国内清华大学^[13]、国防科技大学^[14]、西安交通大学^[15]等科研小组也在持续进行相关研究。

为了提高计算机对视听觉信息的理解能力与处理效率、推动信息服务及相关产业发展、确保国家安全与公共安全，我国国家自然科学基金委员会于 2008 年将“视听觉信息的认知计算”列为重大研究计划，其中研制具有自然环境感知与智能行为决策能力的无人驾驶车辆验证平台是“视听觉信息的认知计算”的总体目标之一。为扩大重大研究计划的影响，国家自然科学基金委于 2009 年 6 月在西安主办了首届“智能车未来挑战”赛，旨在通过赛事交流，推动和促进视听觉信息认知计算模型、关键技术与验证平台研究的创新与发展，该赛事的成功举行对促进我国非结构化视听觉信息认知技术的发展，推动无人驾驶车辆验证平台的研究与创新具有重大意义^[16]。

目前我国无人驾驶汽车验证平台的研究尚处于从实验室走向实际应用的探索阶段。而阻碍这一发展过程的一个非常重要的因素就是：现有的无人驾驶汽车所搭载的计算平台和运行环境尚无法满足实际应用场景对其功耗和实时性的要求。目前我国无人驾驶汽车上的主流计算单元仍然是基于 X86 架构的通用 PC 机或服务器，而为了满足算法的计算需求，保证算法的实时性，越来越多的计算设备被搬到了无人驾驶汽车上，这也就导致了计算平台的规模越来越大，功耗越来越高。这样，

如果需要进行长距离或长时间的测试任务或面对实际应用环境，计算功耗的需求就无法满足。例如在首届“智能车未来挑战”赛中，绝大部分的参赛车队都采用了多台通用 PC 机或刀片服务器作为其计算平台，都面临着系统供电不足的问题。

因此，针对无人驾驶汽车上视觉感知算法的特殊应用，研究高性能、低功耗的计算平台和运行环境，成为了促进我国无人驾驶汽车验证平台的研究从实验室走向实际应用的一个关键技术。而随着主流嵌入式系统计算能力的不断提升，特别是 DSP (Digital Signal Processor)、FPGA (Field Programmable Gate Array) 和异构多核高性能嵌入式计算平台的发展，使得为无人驾驶汽车上视觉感知算法设计专用的嵌入式计算引擎成为了可能。事实上，已经有相关的研究机构展开了这方面的工作。具有代表性的工作是帕尔玛大学的 VisLab 实验室在 2008 年研制出了一款基于 TI (Texas Instruments) 公司 DSP 处理器 DM642 的 VELD (VisLab Embedded Lane Detector) 视频处理平台^[17]，专门负责处理无人驾驶汽车上的视觉感知计算。

但是，在将高性能嵌入式计算平台引入到无人驾驶汽车上的同时，又给视觉感知算法的开发人员带来了新的问题，他们在专注于算法实现和性能提升的过程中，还需要同时兼顾硬件上的问题。由于采用的是专用的嵌入式计算平台而非通用处理器，算法开发人员不得不从底层硬件开始搭建整个算法系统，同时还需要考虑不同处理器之间编程模型的差异，这无疑加大了算法开发的难度，同时增加了整个系统调试和维护的复杂度。因此，为了屏蔽底层硬件的差异，使算法开发人员专注于特定应用中的视觉问题而不需要考虑其他细节，展开面向视觉感知算法开发的软件基础框架的研究是非常必要的。如能在嵌入式计算平台的基础上对底层相关操作进行封装，结合可视化的开发和调试环境，为算法研发人员提供完整的工具集，将对提高视觉感知算法的开发效率、促进我国无人驾驶汽车验证平台实用性的发展发挥至关重要的作用。

综上所述，无人驾驶汽车上视觉感知算法的软硬件关键技术研究具有重要的现实理论意义和应用价值。如能针对特定的应用，结合实时视觉感知算法的设计流程和方法，研究并设计专用的嵌入式计算引擎，实现算法对实时性和低功耗的要求；同时研究并设计一个面向无人驾驶汽车上视觉感知算法开发的软件基础框架，为算法开发人员提供完整工具集和可视化调试环境，提升算法的开发效率，解决多个处理单元如何高效协作运行的问题，突破无人驾驶汽车嵌入式计算平台在实际应用中所面临的瓶颈，对于推进我国无人驾驶汽车验证平台的研究从实验室走向实际应用具有重要意义。

(本表可附页)

三、研究的主要内容

本文面向无人驾驶汽车上视觉感知技术，以提高视觉感知算法的计算性能、降低无人驾驶汽车计算平台的功耗，提升算法开发人员的设计效率为目标，针对无人驾驶汽车上视觉感知算法的计算特性，设计专用的嵌入式计算引擎，并通过研究实时视觉感知算法的设计流程和方法，提出一个面向视觉感知算法的软件开发基础框架结构，为算法研发人员提供一套完整的工具集。研究内容主要包括如下 5 个方面：

1) 分析常用视觉感知算法的计算特性，研究提高计算性能的方法

当前无人驾驶汽车上的视觉感知算法主要包括：道路识别、障碍物检测、车辆行人检测、交通灯和交通标识识别等，通过在自主研发的无人驾驶汽车实验平台上对这些算法进行测试、比较，深入分析其计算特性，如计算复杂度、资源占用情况、计算瓶颈、可并行程度等。然后尝试建立通用的计算模型，并探索提升其计算性能的关键方法；

2) 研究计算模型与计算平台的对应关系，设计面向实时视觉应用的专用计算引擎

针对目前无人驾驶汽车主流计算平台功耗大、成本高的问题，探索利用高性能嵌入式系统作为专用计算引擎来提高计算效率、降低平台计算功耗的方法。通过分析算法的计算特性与硬件平台的性能、功耗、成本之间的对应关系，设计专用的异构多处理器片上计算系统和专用的数据通路，并采用“数据流”和“控制流”相结合的方式，在尽量保持结构化程序优良可编程性的特点下，充分发挥“数据流”计算结构对于海量数据处理的优势；

3) 分析开发人员的设计需求，研究面向视觉感知算法开发的软件基础框架

为提高专用计算引擎的通用性，使算法开发人员从底层硬件管理的负担中解脱出来，研究面向实时视觉应用的软件开发基础框架设计方法。充分借鉴计算机科学领域的软件设计方法学和操作系统中分层设计思想的成果和优势，研究面向实时视觉应用的核心算法开发方法和流程，设计层次化的软件体系结构，定义清晰的抽象层设计规范以及层与层之间的交互接口；

4) 研究专用计算引擎下的面向实时视觉应用的模块化设计方法

为提升视觉算法开发的便捷性和可扩展性，研究为算法开发提供完整工具集的模块化功能部件的设计方法。通过分析嵌入式环境下的视觉应用算法开发流程，掌握开发人员在算法实现过程中对软硬件资源的各项需求，借鉴模块化的设计思想，设计用户可裁剪和定制的功能单元，如经典视觉算法 API、数据预录和回放等功能。同时研究可视化的开发和调试技术，设计并搭建一个友好的图形化用户接口，使得新的应用程序或算法能以插件的形式应用于该计算平台上；

5) 研究基于算法开发基础框架的参考实现和测试验证方法

针对视觉算法开发基础框架的结构特性，研究层次化和模块化的测试方法。利用基础框架设计一个基于视觉计算的道路识别算法原型，并在自主研发的无人驾驶汽车实验平台上，进行长距离的实际行车测试，对算法的可靠性、实时性，以及软硬件基础框架的功耗性和便捷性进行评估和验证。

(本表可附页)

四、工作的重点与难点，拟采取的解决方案

重点和难点

1) 算法计算特性分析与专用计算结构建模

无人驾驶汽车上的视觉感知算法具有计算数据量大、计算复杂度高特性，利用常规计算结构难以合理高效地进行处理。如何提取出视觉感知算法的通用计算特性，针对此特性搭建专用计算结构模型，研究和设计高性能计算引擎，在保证计算实时、高效的情况下，增加系统设计和应用的灵活性，是本文研究的基础和关键，也是在计算机体系结构领域研究的首要科学问题；

2) 面向视觉感知算法的设计方法学

面向无人驾驶汽车上视觉感知算法开发的软件基础框架研究是一个较新的领域。如何针对无人驾驶汽车上视觉感知算法的计算需求和设计流程，结合上述专用计算结构，根据视觉感知算法的分析和开发流程，充分借鉴计算机科学领域的软件设计方法学和硬件设计自动化方面的成果和优势，研究面向无人驾驶汽车上视觉感知算法的设计方法，提出相应的设计规范，使算法开发流程规范化、标准化，提升算法开发人员的设计效率，是本文研究的关键科学问题。

研究方案

本文拟提出一个视觉算法开发基础框架结构，并在此框架结构的基础上展开包括专用计算引擎、硬件抽象技术、模块化功能设计、算法封装与接口技术以及可视化调试技术等各种面向实时视觉感知计算的软硬件关键技术研究。视觉算法设计框架的基本结构如图 1 所示，它主要包含 5 个抽象层次，从下至上依次是：硬件层、硬件抽象层、基本服务层、高级服务层和应用层。

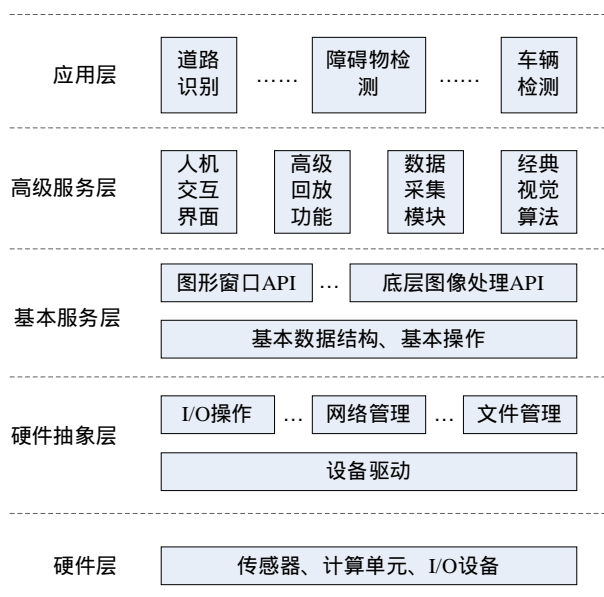


图 1 视觉算法基础框架结构图

图中的每一个抽象层次之间都通过清晰定义的函数接口来进行衔接，且越往上，抽象程度越高，功能越强大。**硬件层**主要包括专用的嵌入式计算引擎和算法所需的常用硬件资源；**硬件抽象层**能屏蔽底层硬件设备（处理器、传感器等）的区别，同时实现底层的数据校准、数据同步等操作，向上提供访问计算平台各种硬件资源的接口；**基本服务层**定义基本数据结构和基本操作，如图像结构体、内存的分配与释放、图像的保存与删除等；并向上封装常用的底层图像处理函数（如二值化、Hough变换等）和图像图形显示接口；**高级服务层**以模块的形式实现视觉算法开发中常用的高级功能，如

道路识别和障碍物检测的经典算法、数据预录和回放等，并为开发人员提供可视化的开发调试界面；以上四层可以为实时视觉应用的快速开发提供一套完整的工具集。开发人员可以根据特定算法的应用需求对高级服务层中的各个功能模块进行裁剪和定制，然后利用下层提供的功能接口，将新的应用程序以插件的形式无缝集成到应用层中。

针对本文所提出的研究目标和研究内容，我们拟采用如图 2 所示的技术路线，主要分四个步骤来完成研究工作，具体步骤如下：

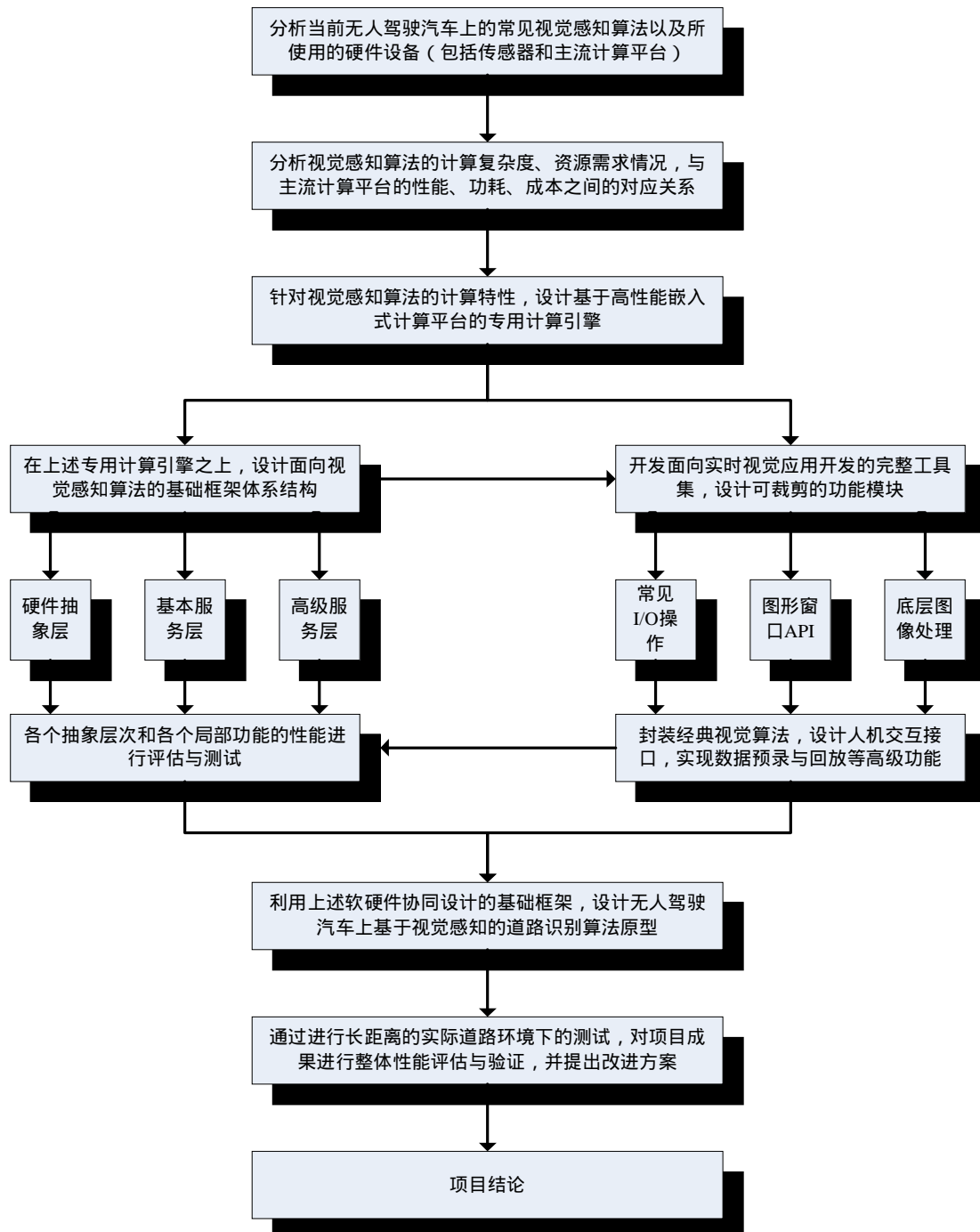


图 2 技术路线图

- 1) 首先分析现有无人驾驶汽车上的常用视觉感知算法，以及所采用的主流计算平台；重点通过对自主研发的无人驾驶汽车实验平台进行各项实际测试，分析算法的计算复杂度和资源需求情况，与计算平台的性能、功耗、成本之间的对应关系；

- 2) 针对当前主流视觉感知算法的计算特性，结合前期分析得出的计算模型与计算平台之间的对应关系，利用自主研发的高性能、低功耗嵌入式计算平台，设计专用的视觉计算引擎。可选的嵌入式系统构建方案包括：基于达芬奇技术的 TMS320DM6446 计算平台、采用 ARM+硬件编解码的 FIC8120 计算平台、TI DM64x+系列的 DSP 以及 Xilinx Virtex-4 系列的 FPGA 计算平台；
- 3) 在专用计算引擎的基础上，深入分析开发人员在算法开发过程中对软件环境的设计需求，结合嵌入式条件下的软件开发和调试过程，搭建一个面向实时视觉算法的软件设计基础框架。设计过程主要注意两点：1、采用分层设计的思想，构建基础框架的体系结构，定义每个抽象层的设计规范，以及层与层之间的交互接口；2、采用模块化的思想，设计为视觉算法的开发提供完整工具集的功能部件，包括 I/O 操作、底层图像处理等，以及录像回放、经典视觉算法等高级功能模块；
- 4) 基础框架的设计过程中，基础框架的各个抽象层次和各个功能部件都将在自主研发的无人驾驶汽车实验平台进行实际道路环境的测试，以验证其可靠性、实时性和实用性等性能指标。整体框架设计完成之后，使用该设计框架开发一个无人驾驶汽车上基于视觉计算的道路识别算法原型，对其便捷性和实用性进行评估；并进行长时间长距离不间断的实况测试，以验证设计框架的整体性能；最后根据测试结果，提出改进方案。

(本表可附页)

五、论文工作量及进度

根据工作内容和目前已有的研究基础，在后续攻读博士学位的两年时间里一共分为四个大的阶段。每个阶段都有其具体的任务安排和要达到的目标，包括需发表的论文数。详细内容如下：

第一阶段：参考相关文献和课题组前期研究经验，分析无人驾驶汽车视觉感知算法中计算特性及其对平台的要求（2010年7月--2010年9月）

- 1) 综合分析目前主流道路识别和障碍物检测实现方案，分析其中实时视觉感知技术的计算瓶颈和性能要求，通过系统工具统计其程序行为特征，为下一步基础框架体系结构和算法工具集的设计和实现提供基础。在此阶段，需要根据应用领域和研究要求，选定几个符合要求的高性能嵌入式计算平台，并确定编程模型。
- 2) 发表与论文相关的有较高水平的学术论文 1~2 篇。

第二阶段：面向视觉感知算法的专用嵌入式计算引擎的研究（2010年10月--2011年4月）

- 1) 跟踪国际上比较知名的高性能嵌入式计算、视觉感知算法研究小组，了解其在硬件平台和算法设计等各方面研究的最新进展；
- 2) 在第一阶段工作的基础上，通过测试评估进一步明确无人驾驶汽车视觉感知算法的计算特性及其相关需求，为计算平台的设计与优化提供理论依据；
- 3) 综合考虑实时和功耗这两个嵌入式应用中的关键因素，实现在满足实时性的条件下，根据各处理器反馈的各自的性能，设计全局功耗最佳的计算拓扑结构。根据各处理单元的负载状况和各任务完成情况，结合服务等级和功耗等参数，设计自适应的低功耗实时任务调度算法；
- 4) 将现有视觉感知算法在专用计算引擎上进行移植和测试，检验理论研究成果；并研究算法实现流程，总结实践经验，为第三阶段的研究提供理论依据。
- 5) 发表与论文相关的有较高水平的学术论文 1~2 篇

第三阶段：面向视觉感知算法的软件开发基础框架设计与实现（2011年5月--2011年12月）

- 1) 研究已有的算法开发框架和基本流程，分析各种框架的优缺点；并在第二阶段的研究基础上，提出基础框架的基本结构；
- 2) 将底层各种不同的硬件操作以及各种设备驱动进行封装，向上提供统一的编程接口，对算法开发人员屏蔽底层硬件的差异；
- 3) 集成可裁剪的图像处理功能模块，使得现有无人驾驶汽车上成熟的视觉感知模块能够以插件的形式无缝集成到框架中；引入高效的图形窗口 API，封装常用 I/O 操作和资源管理功能，以提供便捷直观的算法开发调试环境；
- 4) 结合无人驾驶汽车上某一特定的应用背景，提出其基于特定计算引擎的形式化模型，并在所提出的基础框架上对其算法开发流程进行研究；
- 5) 发表与论文相关的有较高水平的学术论文 2~3 篇

第四阶段：视觉感知算法的参考实现研究（2012年1月--2012年7月）

- 1) 针对第三阶段研究的某一特定应用背景，改进无人驾驶汽车验证平台，集成第二阶段中的专用计算引擎和第三阶段中的算法开发基础框架，并在此基础上进行特定视觉感知算法的设计与开发，将理论研究成果在实际平台上验证；
- 2) 对算法进行测试，分析他们在实际平台上运行时的能耗、平均响应时间、突发响应速度等性能，运用实际平台对算法开发效率进行评价，将现有算法的运行结果与实测结果进行对比分析，根据分析结果对计算引擎和基础框架进行完善；

- 3) 文档准备，撰写博士论文；
- 4) 发表与论文相关的有较高水平的学术论文 1~2 篇

(本表可附页)

六、论文预期成果及创新点

预期研究成果

预期的研究成果主要有以下三部分：

- 1) 构建一个面向视觉感知计算的专用计算引擎，并在此基础上形成一个可扩展的视觉算法开发基础框架，该框架可为实时视觉算法的快速开发提供完整工具集。
- 2) 基于上述专用计算引擎和算法开发基础框架设计无人驾驶汽车上道路识别算法的参考实现。所开发算法的可靠性、实时性不低于其在现有主流计算平台上运行的性能，而且前者的计算功耗和成本低于后者。
- 3) 在国内外重要学术期刊和国际会议上发表不少于 5 篇高质量论文，其中 SCI/EI 检索不少于 5 篇。
- 4) 完成博士毕业论文

论文研究的创新点

本文提出研制面向无人驾驶汽车上视觉感知计算的专用嵌入式计算引擎及相应的软件开发基础框架，研究内容涉及到计算机视觉、体系结构、软件设计方法学等多个领域的原理与方法，具有方法创新性和设计设想独特性。本文的创新点主要有：

1. 提出利用高性能嵌入式计算平台的优势来改善现有无人驾驶汽车上通用计算平台功耗大、复杂度高的问题，利用计算机体系结构领域的原理和方法来为视觉感知算法设计专用的计算结构，在保证算法实时性的前提下，研究降低平台计算功耗的方法。对促进视觉感知技术在嵌入式领域相关研究的发展具有一定的指导意义；
2. 针对嵌入式计算平台底层硬件差异大、编程模型不一致的特点，提出设计基于专用计算引擎的视觉算法开发基础框架，通过借鉴计算机科学领域的软件设计方法学和硬件设计自动化方面的成果和优势，来提升算法研发人员的设计效率。对推广嵌入式计算平台在无人驾驶汽车上的运用具有一定的促进作用。

(本表可附页)

七、完成论文拟阅读的主要文献

为了完成论文，已阅读或者拟阅读以下主要文献：

- | IEEE Transactions on Intelligent Transportation Systems 上近 5 年相关论文
- | IEEE Transactions on Pattern Analysis and Machine Intelligence 上近 5 年相关论文
- | Proceedings of the IEEE Intelligent Vehicles Symposium 近年相关论文
- | Proceedings of the IEEE Conference on Computer Vision 近年相关论文
- | Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 近年相关论文
- | 其他相关嵌入式重要国际会议、期刊和国内权威期刊近年相关论文
- | 相关国内外博士论文

为完成论文工作，预计阅读论文和技术文献 100 - 150 篇

(本表可附页)

指导教师意见	<p>指导教师签名：_____ 二〇〇 年 月 日</p>
评议小组名单	<p>由学院主管领导确定 3~5 名具有高级职称的教师为评议小组成员</p> <p>组长：_____</p> <p>组员：_____、_____、_____、_____、_____</p> <p style="text-align: right;">（公章）</p> <p>主管领导签名：_____ 二〇〇 年 月 日</p>
评议小组意见	<p>1、论文选题： 有理论意义； 有工程背景； 有实用价值； 意义不大。</p> <p>2、论文的难度： 偏高； 适当； 偏低。</p> <p>3、论文的工作量： 偏大； 适当； 偏小。</p> <p>4、研究方案的可行性： 好； 较好； 一般； 不可行。</p> <p>5、研究生对文献资料及课题的了解程度： 好； 较好； 一般； 较差。</p> <p>6、研究生在论文选题报告中反映出的综合能力和表达能力： 好； 较好； 一般； 较差。</p> <p>7、研究生在论文选题报告中反映出的创新能力： 好； 较好； 一般； 较差。</p> <p>8、对论文选题报告的总体评价： 好； 较好； 一般； 较差。</p> <p>（在相应的方块内作记号“ ”）</p>
评议结论	<p>评议小组组长签名：_____ 二〇〇 年 月 日</p>