

学术写作与学术规范

学位论文撰写

苏小红

哈尔滨工业大学 计算机学院



学位论文ABC



学位论文的重要性及其类型

- # 是对研究生研究工作的全面总结
- # 是描述其研究成果、代表其研究水平的重要学术文献资料
- # 是申请和授予相应学位的基本依据
- # 学位论文撰写是研究生培养过程的基本训练之一

基础理论型

- 需要提出成体系的理论

工程应用型

- 需对有价值的重要工程应用中反映出的具体问题进行深入分析



论文研究工作的步骤

文献检索
确定选题

确定方案
准备开题

数据采集
实验研究

1

2

3

论文研究工作 的步骤

中期检查
进度把控

4

7

6

5

论文送审
准备答辩

撰写论文
反复修改

补充实验
整理分析

撰写学位论文的三板斧

1
• 收集资料、
确定题目、
方案论证、
实验验证

2
• 编写提纲、
交给导师审
核提纲、完
成论文初稿

3
• 导师审核初
稿、学生反
复修改论文、
最终定稿





指导学生论文是一种怎样的感受？





指导学生论文是一种怎样的感受？





学位论文的写作过程



1.准备工作

1

- 计算和列表
 - 重新检查并核对实验结果

2

- 绘图
 - 实验结果尽可能用图可视化

3

- 结论
 - 分析结果，作出解释，得出结论

4

- 补充实验
 - 收集更多数据，验证实验结论

5

- 修正结论
 - 分析适用性和例外情况，获得新发现

6

- 整理和记录
 - 去除无可比性数据
 - 记录结果



2. 拟定目录和提纲

拟定目录和提纲

导师审核目录和提纲

结构微调

- 各章节的排列次序是否合理?
- 各章节的内容之间的联系是否紧密?
- 各章节的内容安排是否合乎逻辑?
- 各章节的标题命名是否清晰合理?



3. 完成初稿

根据审核后的提纲撰写初稿

- 一气呵成
- 莫追求完美
- 初稿写完后，再通篇阅读，思考如何修改



4. 补充、修改、完善

充实实验

- 是否有足够的实验数据来支持自己的论点？
- 实验设计是否回答了自己想要验证的所有问题？
- 实验的评价是否公平、客观、合理？
- 是否有与最新的文献方法做实验对比？
- 是否有充分的实验结果分析？
- 从实验结果中是否有新的发现？



4. 补充、修改、完善

精炼内容

- 内容方面是否有不必要的重复内容？
- 重要论点的论述是否充分，内容是否有遗漏？
- 各章内容的篇幅是否平衡？
- 是否符合撰写规范要求？



4. 补充、修改、完善

题目推敲

- 确切概括论文的性质与内容
- 简短、易读、易懂，但又不过于简单，空泛
- 文字中包括被研究对象和被观察的因素，不包括细节和结论
- 字句的次序适当，重要的字句应尽可能排在前边
- 合乎语法



4. 补充、修改、完善

段落调整，逻辑梳理

- 是否某些段落太短或太长？
- 段落之间的逻辑是否连贯？

文句斟酌

- 是否有逻辑不连贯的、语法不通顺的句子？
- 是否有语义模糊、存在歧义的句子？
- 文字叙述是否简明扼要？是否有冗长、含义重复的句子？



4. 补充、修改、完善

论文不厌百回改

结构微调

逻辑梳理

文句斟酌

文献查证

数据核对

1

2

3

4

5





学位论文的构成



学位论文的构成

1. 题目 (封面)
2. 摘要 (中/英文)
3. 目录
4. 正文 (绪论, 主体, 结论)
5. 致谢
6. 参考文献
7. 附录

This is..
套路



学位论文的撰写要求

硕士和博士学位论文

- 不同的要求：
 - 字数
 - 理论研究的深度
 - 创造性成果等
- 相同的要求：
 - 撰写规范基本一致



題 目



学位论文的题目要求

- # 恰当、准确地反映和概括本课题的研究内容
- # 中文题目应**不超过25字**，且不设副标题

A woman in a dark blazer is standing at a whiteboard, pointing at it with her right hand. She has long dark hair and is looking towards the audience. In the foreground, a man's profile is visible on the left, and another woman's profile is on the right, both appearing to be listening or taking notes. The background shows a large window with a view of a city skyline.

摘要





摘要的字数要求

- 博士学位论文
 - 一般为1000 ~ 2000字 (2页)
- 硕士学位论文
 - 一般为500 ~ 1000字 (1页)



摘要的撰写范式

背景、目的和意义

- 1) 研究背景与意义
- 2) 相关领域研究现状
- 3) 论文针对的关键科学问题（包含问题陈述和动机，突出你要解决的问题的重要性）
- 4) 使读者把握论文选题的必要性和重要性

主要研究内容

- 1) 针对论文所要解决的核心问题开展的主要研究工作
- 2) 研究方法或手段，得到的结果和重要结论
- 3) 使读者了解论文的研究思路、研究方案、研究方法或手段的合理性与先进性

简单概括总结

- 1) 对主要结果和结论的意义和价值进行概括
- 2) 使读者快速获得论文成果的学术价值



摘要的研究工作部分

基础理论型

第1项研究内容、
方法、结果

第2项研究内容、
方法、结果

第3项研究内容、
方法、结果

工程应用型

总体设计

核心算法/主要模
块/实现平台、语
言、工具

测试或应用效果



摘要的撰写要求

- # **论文内容的高度概括和简要陈述**
- # **具有独立性和自含性**
 - 独立性：不依赖文章的内容，就能清楚论文的研究结果和结论
 - 自含性：不阅读全文，就能获得论文的主要信息
- # **内容完整、客观、准确，不遗漏、不拔高、不添加**
- # **文字要精练，段落衔接要流畅**
- # **一般使用第三人称**
 - “分析了.....原因”
 - “认为.....”
 - “对.....进行了探讨”



撰写摘要时的注意事项

- # (1) 不要写成流水账，或目录式的内容介绍
- # (2) 不要使用公式、图表和非公知公用的符号与术语
- # (3) 不要标注引用文献编号
- # (4) 不要包含作者主观性的评价
- # (5) 对背景、目的、意义、概念和一般性（常识性）理论叙述不要过多
- # (6) 不宜使用“我”、“你”、“他”等人称代词
- # (7) 名词术语要规范
 - 对个别新术语或无中文译文的术语，可用外文，或在中文译文后加括号注明外文



不好的摘要示例

摘要

在型号研制的工程实践中，需要进行各阶段的测试和试验来检验设计的正确性、安全性、有效性。在综合测试阶段，通过型号部件进行单元测试，来验证型号的状态；通过系统测试来考核型号系统工作流程和性能是否正常，另外，考核型号控制分系统各种功能性能是否正确，需要进行初级的半实物仿真验证试验。目前本单位现有的测试设备中，只能是针对部件的简单测试，能够证明部件的简单特性，偏重于定性检验，缺少测试系统设备，能够集自动测试、验证仿真、目标相对运动为一体。而且现有设备的测试结果无法保存，缺乏对后续分析的参考作用。因此，为了检验型号及型号各系统工作的协调性和正确性，同时能够进行综合测试及半实物仿真，需要设计一种用于对型号各部件进行单元和系统测试的综合测试系统。

本文以航天器的测试仿真需求为牵引，设计了用于型号的自动化综合测试系统，进行了航天器等效器的系统方案设计、PXI 总线技术研究、等效器基于 PXI 的系统硬件、软件的设计和实现、等效器基于 PXI 的系统集成和分析、以及与型号的联调测试。

该系统在工作中已经成为航天器测试中的有利助手，保证研制质量并提高了测试效率，为航天事业的蓬勃发展奠定了坚实的基础。

关键词：等效器；测试仿真；PXI 总线；自动化

研究背景和意义写得太多
介绍本文工作的字数太少



不好的摘要示例

摘要

微信是近年来最常用的一种社交方式之一，微信企业号平台利用微信软件的特点将日常社交、网络和企业资源有机的结合起来。将企业办公系统从有原来的专用客户端或者应用APP转换成只要企业员工的手机装有微信客户端就能完成日常办公事务处理，这样不仅最大程度利用了企业信息化资源，也能使员工通过日常使用的社交软件，随时随地得处理办公事务，达到更简便的效果。通过微信社交软件的广泛性，能够使企业无需对员工的姓名、身份信息、个人照片等个人信息进行重新录入，减少了企业人力资源管理部门人员数据录入的工作量。微信企业号平台上的移动办公系统调用微信上用户的身份认证信息，使员工在登录微信后不需要再次登录移动办公系统，即减轻企业对于员工身份登录与验证系统开发、服务器并发压力与数据录入，也减少企业员工登录各种系统的登录次数。通过这种方式完成日常办公，使企业、个人的工作效率能得到显著的提高。

本项目来源于本人在实习期间参与的基于微信企业平台的移动办公系统的实际开发项目。该项目目标主要是构建和实现在微信企业号平台上的移动办公系统，使企业员工能够通过此系统进行办公、管理操作，完成企业员工的基本邮件、事物审批、签到考勤和微信公众号管理与发布等内容。目的是为了提高企业办公效率，使企业员工更方便的处理日常办公需求，确保员工通过社交软件上的集成功能提高工作技能，更好的服务企业。本人负责该项目的设计方案和具体实现，即基于微信企业平台的移动办公系统。本项目主要包括以下内容：

基于微信企业平台的移动办公系统的原理和技术分析。对平台实现技术的分析。包括平台架构的设计、开发语言的选择、业务服务器的构建以及数据库服务选型。

基于微信企业平台的移动办公系统的原理和技术分析。对平台实现技术的分析。包括平台架构的设计、开发语言的选择、业务服务器的构建以及数据库服务选型。

以上基于微信企业平台的移动办公系统设计与实现的设计与实现是本人在硕士实习期间于哈尔滨工业大学软件工程股份有限公司完成的项目，该项目来源真实的企业单位委托项目，意义重大。最终，本系统得到了顺利实现并上线正常使用，本人也得到了公司的认可。

关键词：微信企业号平台，移动办公系统，考勤签到，MVC

第1段研究背景和意义写得太多

基于微信企业平台的移动办公系统的原理和技术分析。对平台实现技术的分析。包括平台架构的设计、开发语言的选择、业务服务器的构建以及数据库服务选型。

基于微信企业平台的移动办公系统的设计与开发。本文首先对用户企业的需求进行分析，然后设计相应功能模块，其包括：我的办公桌功能、工作流程功能、消息群发功能、微信考勤功能、企业号管理功能和系统管理功能，本文设计和实现了所有功能模块。

语句不通顺

第2段和第5段不适合写在摘要中



不好的摘要示例

摘要

本文主要是通过图像处理技术和BP神经网络技术对青椒果实进行识别，研究的主要内容如下：

1. 从素材图库中分析青椒果实和背景的颜色特点，利用色度学知识分别提取出果实和背景颜色的范围和特点，找到理想的参数对原始图像进行灰度化处理，再结合果实和背景的灰度分布特点，计算出果实和背景的分割阈值，最终将图像分割成二值图像。

2. 利用图像处理技术，对图像进行前期处理，通过去噪、平滑、腐蚀膨胀以及边缘检测，完成青椒果实和背景图像的分离。实验中，主要对四种滤波方法和三种边缘检测方法进行了详细对比，选出最适合处理青椒图像的中值滤波算法和Canny算子边缘检测算法作为最终选定的设计方案。

3. 通过分析青椒果实的形状特征，提取出五个归一化的特征值，作为BP神经网络的输入。实验中提取了青椒果实的五个具有同类特征的特征值，加快了BP神经网络的识别速率。

4. 本文构造BP神经网络图像识别系统，以默认参数建立BP神经网络模型，通过青椒果实的五个归一化特征值进行训练识别，同时不断改变神经网络参数，确定一组收敛快、识别率高的参数组，建立稳定的BP神经网络。

5. 测试样本，最终对识别结果进行统计、分析。

关键词：图像识别，阈值，形状特征，神经网络，青椒

**缺研究背景和意义
语句不通顺
关键词不当**



不好的摘要示例

摘要

据不完全统计，中国每年有 20 万名失踪儿童，只有 0.1% 被找回。防止儿童走失，已经成为社会性问题。采用高科技防止儿童走失，有一个非常有前景的市场。

本文采用智能手表硬件终端与手机 Android 端结合的方式，实现基于位置服务的儿童安全监控管理系统。根据智能手表硬件终端生产商提供的通信协议，服务器端采用 Netty 框架，实现与智能手表硬件终端的通信。通过分析手表硬件终端上传的数据，可以监控佩戴者的位置状态等信息。开发手机应用程序，绑定手表终端，并监控手表数据。手机应用使用 Http 的 get、post 请求，从服务器端获取手表终端数据，手表终端通过 Tcp 协议，上传数据到服务器端。同时，通过百度云推送，可将数据主动发送到手机 Android 端。基于 AMap SDK，实现了基于位置的服务来监控儿童安全。实现了位置跟踪、区域监控、危险场所告警、路线偏差、SOS 语音呼救、轨迹回放、群组管理等功能，能够对儿童进行实时监护。同时采用 K 中心算法等，对历史轨迹的显示进行优化，使家长能够清晰的掌握儿童历史动态。

关键词： 位置服务，儿童安全，精准定位，高并发。

缺少实验结论或系统实现的效果或测试结果

关键词没有很好地体现本文采用的关键技术



不好的摘要示例

摘要

电子商务系统随着现代计算机技术的发展而呈现蓬勃发展之势并已逐渐成熟，然而，与之对应的物流技术却在信息化发展的道路上停滞不前。而随着电子商务商户及其客户、物流、及时性、透明性的要求的逐渐提高，现有的物流管理的方法及技术越来越无法适应时代的发展，同时物流管理居高不下的成本也成为阻滞物流管理技术发展的瓶颈，在我国物流发展的现阶段存在着很多中小型物流企业，这些企业对物流管理信息化的建设依然认识浅薄，以致我国物流信息化的发展停滞不前。并且在已有的物流信息化管理系统中，由于物流信息化管理标准的缺失，这些物流信息化管理系统的通用性和扩展性受到了极大地限制。

基于我国物流信息化管理发展的现状，本文将搭建一个具有良好的通用性和扩展性的物流信息化管理平台，同时优化系统的吞吐量，使得系统平台能更好的适应大量用户并行访问的需求。并且在系统中为相应的电子商务、国家相关部门等系统预留相应的接口，以期该系统和其他相关系统能实现有效的对接，最大限度的实现物流管理和电子商务及其他相关业务和政策协作的高效性，最大限度的实现物流管理系统和电子商务系统及其他供应链系统的信息共享和优化。

另外，在物流管理的过程中，物流管理的成本居高不下的一个主要原因既是在物流配送过程中物流车辆的空载率问题，故本文基于物流车辆路径优化问题的研究现状，给出了物流车辆路径优化的数学模型及优化目标，然后结合粒子群算法，对物流车辆路径规划问题设计了编码和解码方法及其目标约束函数，并依据仿真实验的实际结果对算法的有效性进行了验证。

**大段文字在写研究意义
和背景**

**需增加技术相关的关键词
如粒子群算法、物流车
辆配送路径规划等**



不好的摘要示例

摘要

随着当前计算机技术的广泛应用，传统的大型装备纸质技术使用手册数字化程度低，导致在编制、运输、使用、存储等过程中缺乏高效性和延续性，已不能较好的满足用户需求。产品交互式电子手册（IETM，Interactive Electronic Technical Manual）由美国国防部率先提出，并逐渐受到各国国防部门重视。利用该手册，技术人员对装备的使用、人机交互、信息检索、设备维护、故障诊断、维修保障等方面将更为容易实现。按照国际通用标准规范要求，使用模块化管理形式对装备信息进行统一管理，提高工作自动化程度。但是，受到复杂产品内容过多的影响，手册中的技术文件数量逐渐膨胀庞大，导致普通检索方法产生结果时效率不高、冗余度大、用户易用性下降，在故障排查方面也缺乏一定的诊断能力。

本文首先从现在装备需求出发，以 GJB6600 标准作为基础，结合实际工作项目，建立起一套包括技术文件数据编制、数据管理、交互浏览的系统模型，为使用者提供模型中装备安全警示类、技术说明类、使用操作类等资料文件的编制显示。采用全文资料检索技术来解决技术资料的快速定位相应问题，通过故障诊断技术构建知识库从而达到面向设备故障诊断的能力，应用交互式动态浏览技术提高技术资料交互显示水平，从而大幅提高装备使用维护效率，显著提升装备的维修保障能力。

工程硕士学位论文

产品交互式电子手册 IETM 的
设计与实现

IETM与前面的“产品
交互式电子手册”是同
一个含义，建议题目
改成产品交互式电子手
册数据管理系统

摘要介绍研究意义和
背景的篇幅较多，介
绍论文工作的篇幅偏
少，而且没有介绍系
统实现的效果。



不好的摘要示例

摘要

最近的研究表明在大规模程序中含有许多重复代码，它们中的大部分是由拷贝-粘贴行为造成的。这些重复代码不仅降低了系统软件的可维护性，而且会给软件引入大量的 bug。然而，现有的重复代码查找工具或因消耗的时间和内存过多而无法面向大规模软件，或无法找出那些经过重命名、插入或删除修改的重复代码。

通过对比各种重复代码检测技术的优缺点，本文选用基于 token 串的分析方法，引入了数据挖掘技术，实现了一个重复代码检测模型。首先，分析源代码创建一个序列数据库，将重复代码检测问题转化成为一个序列模式挖掘问题。然后，使用改进的 CloSpan 算法查找支持度至少为 2 的频繁子序列，这些子序列对应于程序中的重复代码段。最后，改进检测结果，去除不太可能是重复代码的结果，并合并相邻的重复模块，组成较大的重复代码段。

与其它方法相比，基于 token 串的方法和 CloSpan 算法的算法复杂度较低，时间和内存开销较低，因此适合于分析大规模程序；模型在词法分析之后忽略标识符的真实名字，将类型相似的标识符都映射为相同值，从而可以检测标识符重命名的拷贝-粘贴片段；由于频繁子序列在其支持序列中是可以中断的，因此在设置 gap 阈值之后可以准确地识别出进行过插入删除操作的拷贝-粘贴片段。

实验结果表明，本模型能够在 40 分钟内从 Httpd 2.2.2 (280,000 行) 中检测出约 3000 组重复代码，其中包括经过各种（重命名、插入和删除）修改后的重复代码。

摘要中给出实验结论即可，无需给出实验结果。



不好的摘要示例

摘要

随着生产技术的不断提高，使汽车钣金件不仅在产量方面做到了逐年提高，也使其产品的质量越来越高。这就使得检测钣金件产品是否存在问题是
情况成为了一种庞大数据量中的小概率事件。而传统的人工检测方式不仅浪
费人力、消耗大量的精力和金钱，而且也很难对这种产量较大的全部钣金件
做出细致而全面的检查。

随着计算机视觉的不断发展，为工业自动化检测方法提供了全新的思路。本文利用卷积神经网络技术对汽车钣金件的螺栓螺帽检测进行了研究。具体研究内容如下：

首先，对钣金件的图片数据进行标注，标注后利用企业所提供的少量数据进行数据采样扩大数据样本。本文采用针对前景的采样和针对背景的采样两次采样实现数据扩充。为了进一步扩大样本，本文采样所得到的数据基础上还进行了一些图像变换进行进一步的样本扩容，为后续训练模型做数据准备。

其次，针对图片本身较大而检测目标较小可能会导致较大计算量和检测不准确的问题。本文设计并实现了初步判定图像是否为前景图片的分类器，该分类器的作用是将钣金件前景和背景进行区分。用它来进行初步粗略的识别出螺栓螺帽所在的前景区域，这样做相当于为后续的目标检测模型添加一层注意力机制，方便后续利用目标检测模型检测出准确的前景位置坐标。

最后，本文利用前景区域的初步识别的判别结果，将其判别结果为前景的图片输入到目标检测模型中。对前景的所在坐标进行定位，检测出前景坐标在图片中的具体位置。利用预测得到的坐标位置与标准钣金件上的螺栓螺帽所在位置的坐标进行对比，得出最终结论。

关键词：钣金件缺陷检测，螺栓螺帽缺失检测，计算机视觉，卷积神经网络

**为什么研究没解释清楚
语言不通顺
段落之间内容不连贯
研究得出什么结论没说清楚**



摘要

对话是指单轮或者多轮的对话，一般是以一问一答的方式进行。对话系统的实现可以有规则匹配，基于检索的对话系统，即通过相似度计算找到正确的回复，和基于生成的对话系统，即依赖给出的情景得到回复的对话系统。

目前的检索对话系统和生成对话系统各有突出点，所以本文希望探索合并两者的突出点以提高对话系统的性能的方式，主要从以下三个方向进行研究：

基于检索的对话系统的研究。在 **baseline** 方面，探索了基于 **tf-idf** 的传统检索模型，探索了利用循环神经网络的检索模型，即通过循环神经网络得到句子的表示，之后再利用点积等相似度计算的方法得到问题和候选回复的相似度以选取分数最高的回复。最后探索了利用自注意力机制对句子进行编码得到更深的向量表示对模型效果的影响。

基于生成的对话系统研究。在 **baseline** 方面，探索了序列到序列的生成对话模型，使用 **LSTM** 与 **BiLSTM** 两种网络作为编解码模型结构，之后，探索了带入注意力机制的序列生成对话模型，最后本文探索了外部知识对模型生成结果的影响，分别探索了利用外部信息对模型结果的影响和利用情景信息以及两者结合对模型结果的影响，最后也探索了 Teacher forcing 等训练技巧是如何提升模型效果，一些预测技巧如何对实验效果产生影响。

检索与生成相结合的对话系统研究。在融合方式上，本文探索了几种不同的检索与生成结合的方式。第一种是通过设置阈值，即以阈值作为分界值，判断检索答案得分与阈值的大小，决定采用答案的方式。之后又探索了如何将生成模型的信息融合入检索模型当中，和如何将检索模型的信息融合入生成模型，最后探索如何将这两种结合方式融合在一起。、

在基于以上三个模块的探索之后，本文对比了这三种对话系统的各个优缺点，及本文得到的一些结论与心得。

**研究背景、目的和意义没解释清楚
语言不通顺
研究得出什么结论没说清楚**



不好的摘要示例

摘要

随着我国在司法机关大力推行案件信息公开，人民群众对公检法等话题相关事件的关注日益增长。由于此类事件多伴有争议性，往往能引起巨大的舆论反响。网络舆情虽能推动司法改革和进步，却往往泥沙俱下，尤其是恶性舆论向司法机关施加的压力将干扰司法权的独立。然而，司法独立与尊重民意并不矛盾。舆情监测技术的支持，可以极大缓解司法机关的公关压力，促使司法机关更公正公平地行使权力。

舆情监测是一项交叉了社会科学和数据科学、融合了多种技术的复杂系统。本课题立足于实际需求，结合了对舆情监测的技术研究和系统实现两方面的工作。本文提供了关于构建稳定运行的舆情监测系统的指南，并基于系统采集的海量真实数据，进行了多项相关技术的研究。

我们探索了多种机器学习算法与图论知识在结构化舆情数据上的应用。包括对不同类型舆情事件的传播规律进行定量和定性的分析；基于初期传播状态对最终传播层级的预测；基于当前传播状态对次日传播趋势的预测；以及对传播网络中的意见领袖进行挖掘等。此外，我们还使用了统计学检验的方法评估了多种特征对预测的贡献。

我们也探索了多种深度学习与自然语言处理技术在文本类型的非结构化数据上的应用。包括在热点话题中智能识别公检法等话题相关的舆情事件；在海量微博评论文本中检测用户的情感；以及在事件相关新闻报道中挖掘事件相关要素等。

系统在各项任务指标上均取得了较好的表现。同时系统得到的海量数据也能为其他领域如社会计算、自然语言处理等研究提供支持。

关键词：舆情监测；司法机关；社会计算；自然语言处理；深度学习

**研究背景、目的和意义没解释清楚
研究得出什么结论没说清楚**



优秀摘要示

摘要

软件 **bug** 管理是开发和维护软件系统的核心工作之一。为提升软件 **bug** 的管理效率，软件公司通常采用 **bug** 仓库统一记录和处理软件 **bug**。结合 **bug** 仓库，已有研究尝试将 **bug** 处理流程自动化，如自动生成测试脚本检测 **bug**、自动分派和定位 **bug** 报告。尽管如此，在软件 **bug** 的报告生成与理解方面仍有许多问题亟待解决，主要包括如何自动分析软件测试失败脚本以撰写 **bug** 报告，以及如何理解 **bug** 报告内容以修复软件 **bug**。围绕上述问题，本文研究三方面内容，分别是 **bug** 报告的生成（系统集成测试 **bug** 形成原因分类），**bug** 报告内容理解（基于众包的有监督 **bug** 报告摘要、基于深度学习的无监督 **bug** 报告摘要），**bug** 报告关联资源理解（基于词嵌入的 **bug** 报告关联资源推荐）。

(1) 提出系统集成测试 **bug** 分类算法。在软件测试过程中，测试人员通常需分析执行失败的测试脚本以找到 **bug** 产生的原因，进而准确提交 **bug** 报告。为提升测试人员撰写 **bug** 报告的效率，本部分研究利用测试脚本所关联的测试日志，自动分类测试脚本执行失败的原因。实验表明算法可应用于工业生产环境并提升测试脚本分析的准确率。

(2) 提出基于众包的有监督 **bug** 报告摘要算法。**Bug** 报告摘要是帮助开发人员高效理解 **bug** 报告内容的重要手段，基于摘要结果，开发人员可以在减少阅读量的同时快速理解 **bug** 报告的主旨。本部分研究利用众包的方式招募志愿者人工完成 **bug** 报告摘要任务，通过分析志愿者选择摘要句的原因，并结合启发式构建规则构造领域特征。实践表明，利用上述方法构建的特征可直接提升 **bug** 报告摘要任务的效果。

(3) 提出基于深度学习的无监督 **bug** 报告摘要算法。监督学习算法的效果往往受标注数据的影响，为了缓解该影响，本部分研究结合 **bug** 报告的领域特征，设计步进式自编码网络，以无监督的方式对 **bug** 报告进行摘要。实验对典型 **bug** 仓库中的 **bug** 报告进行摘要，结果显示该算法显著超过对比算法，可解决无标注数据下的 **bug** 报告摘要问题。

(4) 提出基于词嵌入的 **bug** 报告关联资源推荐算法。开发人员根据 **bug** 报告所描述的需求编写软件代码，为了更好的理解 **bug** 维护需求，本部分研究根据 **bug** 的维护需求推荐相关的软件资源，包括 API 和 API 文档，进而帮助开发人员从代码的角度理解 **bug** 报告。该算法利用词嵌入技术将自然语言和代码中 API 进行统一向量表示，根据向量检索与 **bug** 维护需求语义相关的软件资源，利用这些资源辅助开发人员理解 **bug** 报告。

本文对软件 **bug** 的报告生成与理解进行分析，以提升提交软件 **bug** 报告和理解软件 **bug** 的效率。同时本文提出的基于众包的特征构建方法可扩展解决软件工程其它问题。



优秀抗

摘要

随着软件规模的不断扩大，软件故障不可避免。寻找并修复这些故障需要耗费大量的人力和时间。软件调试的主要任务就是先找到出软件错误的位置，进而修复软件错误。因此，软件错误定位是软件调试阶段的首要任务，如果能够实现快速准确的错误定位，那么软件调试过程消耗的时间和精力就会被大大降低。

然而，目前大部分已有的自动化软件错误定位方法的定位效果都不够准确和稳定，也并未考虑巧合正确性测试用例的影响。基于变异分析的错误定位方法定位精度较高且定位结果不受巧合正确性测试用例的影响，但定位的精度依赖于测试用例的质量和变异算子的选择，只能定位与变异算子相关的错误类型。如果变异算子类型较少或选择不当，其错误定位的精度将受到影响；此外，由于程序变异产生的大量变异体会降低变异效率，进而增加错误定位的开销。

针对以上的这些问题，本文主要完成了以下工作：

首先，分析待测程序的词法和语法，获取抽象语法树，在语法树上进行语句级别的程序插桩，执行插桩后的程序以捕捉测试用例的执行路径信息，进而获取变异目标语句。

其次，设计变异算子，在已有的变异测试常用的变异算子的基础上，分析开源的错误程序代码，提出了四类新的变异算子，使用这些变异算子，在待测程序的抽象语法树上对目标语句进行变异处理，生成一系列变异体。

然后，提出了冗余成功测试用例的过滤方法，实现了变异体和测试用例的选择和执行优化，降低了错误定位的开销。接着根据变异体的执行信息计算语句的可疑度，获得程序语句的错误可疑度列表，并给出了结合变异分析的错误定位方法的一个实例。

最后，为了验证方法的有效性，使用 C# 语言实现了一个基于变异分析的 C 语言程序的错误定位原型工具，以 Siemens Suite 作为测试数据集对其进行了测试，通过实例分析验证了本文新提出的变异算子能够在一定程度上提高错误定位的精度，实验结果表明，本文的错误定位方法相比于其他错误定位方法，错误定位的精度较高。此外，进行实验分析测试程序优化执行后提速的比例，结果是本文改进的优化执行策略提高了变异执行的效率。

关键词：变异分析；错误定位；变异算子；变异体；执行优化



优秀

摘要

目前，计算机软件已广泛应用于多个领域中，其规模也日趋庞大和复杂，人们投入大量精力来保证软件的质量，并提出了很多自动化软件错误定位方法，用来帮助开发者快速定位到软件中的错误。

但是，大部分错误定位方法的结果都是一个语句的可疑值序列，其中按照可疑值降序的顺序列出了不同语句的可疑值。但是，仅仅单独检查给出的一条可疑语句，开发者很难理解错误是如何产生的，往往还需要更多与错误相关的上下文信息。

针对这一问题，本文完成了以下工作：

首先，基于抽象语法树对待测程序进行语句级别和逻辑表达式级别两种不同粒度的插桩，并得到相应的程序执行路径。由于目前缺乏简单易用的 C 语言词法语法分析工具，因此本文使用 C# 语言实现了一个 C 语言词法语法分析工具 **CParser**。

其次，为提高错误定位精度，提出了加权软件行为图，将语句执行频率作为权重，可以更好的处理循环和递归等结构。根据成功执行和失败执行的程序执行路径，分别构造出相应的加权软件行为图。

然后，利用基于分支限界搜索的加权软件行为图挖掘算法，识别成功和失败的加权软件行为图之间最有差异的子图作为错误签名，最后得到的 Top-K 错误签名作为错误定位和错误理解的依据，并给出错误签名用于错误理解的实例。

最后，使用 **Siemens Suit** 对本文方法进行了测试，分别分析本文方法的错误定位精度，对不同类型的错误的定位精度，以及逻辑表达式级别插桩对错误定位精度的改进。结果表明，本文方法的错误定位精度较高，对冗余代码、缺失代码和变量替换错误，以及会直接改变执行路径的错误的定位精度明显更高一些，而且逻辑表达式级别插桩能够在一定程度上提升多种错误定位方法的错误定位精度。

关键词：错误定位；软件行为图；图挖掘；错误签名；分支限界搜索



优秀

摘要

医学中的大部分疾病都和细胞行为有关。细胞图像分割与细胞跟踪是研究细胞行为的重要手段。传统细胞行为的研究往往是基于人工的方式，不但需要大量的人力，而且在跟踪细胞过程中涉及细胞染色等化学操作，影响细胞的生理运动，从而影响跟踪结果。因此借助计算机视觉相关理论完成细胞图像分割与细胞跟踪对于研究细胞行为具有重要的意义。

由于细胞场景中存在细胞聚集、消失与分裂等复杂状态，使得一些经典的图像分割方法和目标跟踪方法不能直接应用于细胞场景中。本文正是基于此对细胞图像分割和细胞跟踪方法进行研究。

本文主要完成了以下工作：

1) 针对黏连细胞问题提出了一种基于多尺度 log-filter (Laplace of Gaussian filter) 的四阶段细胞图像分割方法。该方法主要涉及阈值法提取细胞区域；基于多尺度 log-filter 检测细胞中心；基于城市距离定位黏连细胞连接处，之后通过移除低灰度像素或者最近邻方法得到粗分割结果；最后，通过增加梯度约束项改进主动轮廓模型调整分割结果，使得分割结果更加接近细胞的真实边界。该方法在 MCF-10A 数据集上进行验证，能够得到较好的分割结果。

2) 为应对细胞分裂、消失、新生等复杂情况，将细胞跟踪问题转化为分类问题，分别实现了基于有监督学习和主动学习的细胞跟踪。该方法通过挖掘细胞特征，基于 GBDT 有监督学习训练，可以对未知样本取得很好的跟踪效果。针对目前细胞视频标记样本缺乏问题，采用主动学习的思想解决，通过制定筛选样本策略使得标记很少的样本就可以得到很高的正确率。

3) PHD (Probability Hypothesis Density) 滤波方法是基于随机有限集统计的滤波方法，能够跟踪数目变化的目标，因此适合应用于细胞场景中。本文通过分析显微视频图像中细胞的多个特征，提取细胞中心位置、灰度均值和局部场景图像熵三种特征，在状态空间中表征细胞的运动学特征和图像特征，用于自动化细胞跟踪。实验结果表明本文方法可以有效的同时跟踪多个细胞，与采用单独动力学特征的细胞跟踪相比具有更高的跟踪精度。

关键字：细胞分割；细胞检测；细胞跟踪；PHD 滤波；主动学习；GBDT



关键词



关键词的撰写

- # 供检索用的主题词条
- # 应集中体现论文特色，反映研究成果的内涵，具有语义性，在论文中有明确的出处
- # 用能覆盖论文主要内容的专业技术词条(参照相应的技术术语标准)
- # 按词条的外延层次从大到小排列
- # 中文关键字之间用分号；隔开，英文摘要用逗号隔开，最后一词之后不加标点符号





关键词撰写中的常见问题

- # **关键词过于随意**
- # **关键词不是实词**
- # **关键词由过多单词组成**
- # **关键词中包含过多缩写**
- # **关键词在前言和题目中没有出现过**



如何确定关键词？

- # 把题目、摘要或论文中出现频率较高的重要实词挑选出来即可
- # 至少在论文题目中出现的关键词应列入关键词中
- # 不需要太多，一般情况下5~6个



章节结构





章节结构与标题

- # 章节不宜过少/过多，**目录最多三级**
- # 章节标题要突出重点、简明扼要，要能反映论文的工作内容
- # 标题中尽量不采用英文缩写词，必须采用时应使用本行业的通用缩写词
- # 字数一般应在15字以内，标题中不加标点符号

目 录

摘要	I
ABSTRACT	III
第1章 绪 论	1
1.1 课题来源	1
1.2 课题的目的和意义	1
1.3 国内外的研究现状及分析	4
1.3.1 研究概况	4
1.3.2 多机协同目标分配方法研究现状	6
1.3.3 多机协同航迹规划方法研究现状	9
1.3.4 多机协同航迹重规划方法研究现状	11
1.3.5 多无人机多任务的协同方法概述	12
1.4 存在问题分析	14
1.5 主要研究内容	15

第4章 基于空间模糊表示的多机协同航迹规划方法	
4.1 引言	
4.2 基于任务属性的空间模糊表示	
4.2.1 空间点的模糊集表示	
4.2.2 模糊隶属度函数设计	
4.2.3 截集优选关键路径点	
4.3 面向 MUAS 协同航迹规划的文化算法研究	
4.3.1 空间模糊文化算法的航迹规划框架	
4.3.2 基于模糊表示的信念集构成和更新	
4.3.3 种群集上基于 MODE 的协同航迹规划	

各章题序及标题用小4号黑体，其余用小4号宋体

**摘要及目录的页码标号为
罗马数字。如： I II**

目 录

摘要	I
Abstract	II
第1章 绪论	1
1.1 研究目的与意义	1
1.2 国内外研究现状	2
1.2.1 软件错误定位的国外研究现状	3
1.2.2 软件错误定位的国内研究现状	5
1.2.3 软件错误定位的国内外研究现状分析	6
1.3 课题研究的主要内容及章节安排	
第2章 程序插桩技术	
2.1 程序插桩流程	
2.2 抽象语法树	
2.3 C 语言词法语法分析工具 CParser 的设计与实现	
2.4 程序插桩及改进	
2.4.1 语句级别的程序插桩	
2.4.2 逻辑表达式级别的程序插桩	
2.5 本章小结	
第3章 基于加权软件行为图挖掘的错误定位	
3.1 方法基本框架	
3.2 获取程序执行路径	
3.3 构造加权软件行为图	
3.3.1 软件行为图	
3.3.2 加权软件行为图	
3.4 识别错误签名	
3.4.1 剪枝策略	28
3.4.2 子图的选择和去冗余	29
3.5 基于分支限界搜索的加权软件行为图挖掘	31
3.6 错误签名实例	32
3.7 本章小结	34
第4章 验证系统的实现与测试	35

除绪论外，各章后应有一节
“本章小结”

4.1 实验环境	35
4.2 系统设计与实现	35
4.3 系统的功能测试	37
4.4 错误定位的实验结果与分析	42
4.4.1 实验数据	42
4.4.2 评价指标	42
4.4.3 实验结果总体分析	43
4.4.4 针对不同错误类型的定位结果分析	46
4.4.5 语句级别插桩与逻辑表达式级别插桩的对比结果分析	49
4.5 本章小结	51
结 论	53
参考文献	54
攻读硕士学位期间发表的论文及其它成果	58
哈尔滨工业大学学位论文原创性声明及使用授权说明	59
致 谢	60

硕士论文：4~5章
博士论文：5章

目 录

摘要	31
Abstract.....	32
第1章 绪 论	33
1.1 研究的目的与意义.....	34
1.2 国内外研究现状.....	35
1.2.1 错误定位的国内外研究现状.....	35
1.2.2 变异分析的国内外研究现状.....	36
1.3 课题研究的主要内容及章节安排.....	37
第2章 变异算子设计和程序变异处理	38
2.1 引言	38
2.2 变异分析	39
2.3 获取目标语句	40
2.3.1 抽象语法树	40
2.3.2 程序插装	41
2.3.3 获取执行路径	42
2.4 变异算子的设计	43
2.4.1 变异测试中的变异算子	43
2.4.2 新的变异算子	44
2.5 变异处理生成变异体	45
2.6 本章小结	46
第3章 变异执行优化	47
3.1 引言	47
3.2 变异体执行优化	48
3.3 测试用例的选择和执行优化	49
3.3.1 测试用例的随机选择	49
3.3.2 过滤冗余的成功测试用例	50
3.3.3 测试用例执行优化	51
3.4 本章小结	52
第4章 结合变异执行优化的错误定位方法	53
4.1 引言	53
4.2 结合变异执行优化的错误定位方法的基本原理	54
4.3 结合变异执行优化的错误定位方法的流程	55
4.4 变异体执行信息统计分析	56
4.5 计算语句的可疑度	57
4.6 错误定位实例	58
4.7 本章小结	59
第5章 结合变异执行优化的错误定位原型系统的设计与测试	60
5.1 实验环境	60
5.2 系统设计	61
5.3 系统的功能测试	62
5.4 错误定位的实验结果与分析	63
5.4.1 实验数据	63
5.4.2 评价指标	64
5.4.3 新变异算子对错误定位结果的影响	65
5.4.4 与其他错误定位方法的对比分析	66
5.4.5 变异执行优化对执行效率的影响	67
5.5 本章小结	68
结 论	69
参考文献	70
攻读硕士学位期间发表的论文及其它成果	74
哈尔滨工业大学学位论文原创性声明及使用授权说明	75
致 谢	76

硕士论文：4~5章（最后1章
原型系统设计）



有问题的章节示例

3.4 基于主动学习的细胞跟踪	36
3.5 实验结果与分析	37
3.5.1 实验数据	37
3.5.2 有监督学习实验结果及分析	38
3.5.3 主动学习实验结果及分析	42
3.6 本章小结	43
第4章 基于多特征 PHD 滤波的细胞跟踪	44
4.1 引言	44
4.2 基于 PHD 滤波的多目标跟踪方法	45
4.2.1 PHD 滤波	45
4.2.2 PHD 滤波的粒子滤波实现	46
4.2.3 基于 PHD 滤波的目标状态关联	49
4.3 结合灰度信息的细胞 PHD 滤波跟踪	51
4.3.1 细胞状态向量的建立	51
4.3.2 结合灰度特征的 PHD 滤波细胞跟踪算法	53
4.4 实验结果分析	55
4.4.1 实验 1: MCF-10A 数据集	55
4.4.2 实验 2: DCcellIQ 数据集	57
4.5 本章小结	58
结 论	59
参考文献	60
攻读硕士学位期间发表的论文及其它成果	64
哈尔滨工业大学学位论文原创性声明及使用授权说明	65
致 谢	66

标题中不加标点符号



有问题的章节示例

3.6 配置管理模块的设计	27
3.6.1 模块功能描述	27
3.6.2 数据结构设计	27
3.6.3 错误处理设计	28
3.7 数据传输模块的设计	28
3.7.1 数据路由的定义	28
3.8 远端数据管理模块的设计	29
3.8.1 远端数据管理模块接口设计	29
3.9 调度模块的设计	30

如目录中的子项为唯一项
则不必在目录中单列成子项
即只有一小节不能独立成节



有问题的章节示例

摘要	I
Abstract	II
第1章 绪论	5
1.1 课题的背景及研究目的	5
1.1.1 课题的来源	5
1.1.2 课题的目的	5
1.1.3 课题的意义	5
1.2 国内外研究现状	6
1.2.1 国外研究状况	6
1.2.2 国内研究状况	9
1.3 课题主要研究内容	10
第2章 终端手机短信过滤技术	12
2.1 MAPI 简介	12
2.2 MAPI 原理	13
2.2.1 MAPI 规则客户端描述	13
2.2.2 向收件箱注册 MailRuleClient 对象	13
2.2.3 注销 MailRuleClient 对象	14
2.2.4 客户端应用程序与 MAPI 规则客户端之间的通信	14
2.2.5 用以捕获 SMS 消息的客户端进程	15
2.3 测试 mapirule 程序代码	15
2.4 本章小结	17

第2章不要写论文涉及的关键技术，论文涉及的关键技术穿插到后面的设计与实现中去写

各章标题中尽量不出现英文缩写词，对必须采用者，应使用本行业的通用缩写词

系统涉及的关键技术——通常不单独做为一章，不要泛泛的大篇幅抄书，关键技术融入到设计和实现中介绍



有问题的章节示例

目 录

摘要	1.a
Abstract	II.a
第 1 章 绪论	1.a
1.1 课题背景	1.a
1.1.1 课题来源	1.a
1.1.2 课题研究的目的与意义	1.a
1.2 本课题相关的国内外研究状况	2.a
1.2.1 网格计算研究现状	2.a
1.2.2 网格计算中的作业调度	3.a
1.2.3 多集群中的作业调度	4.a
1.3 本课题研究的主要内容	5.a
1.4 本文组织结构	6.a
第 2 章 系统需求分析	8.a
2.1 系统业务需求分析	8.a
2.2 数据感知的定义	8.a
2.3 系统功能需求	9.a
2.3.1 透明的数据访问需求	9.a
2.3.2 数据缓存需求	10.a
2.3.3 智能的数据传输需求	11.a
2.3.4 数据集信息和存储信息查询操作支持需求	11.a
2.3.5 作业调度需求	12.a
2.4 系统非功能性需求	12.a
2.5 本章小结	13.a
第 3 章 系统设计	14.a
3.1 系统架构设计	14.a
3.2 系统模块划分	15.a
3.3 数据集管理模块的设计	17.a
3.3.1 数据集定义	17.a
3.3.2 数据结构设计	18.a
3.3.3 数据集管理模块接口设计	18.a

标题命名不明确



有问题的章节示例

目 录

摘要	I
ABSTRACT	II
第1章 绪论	1
1.1 课题背景及研究的目的和意义	1
1.2 相关工作的研究概况	2
1.2.1 舆情信息传播与预警相关技术的研究现状	3
1.2.2 舆情文本分类与挖掘相关技术的研究现状	5
1.3 本文的主要研究内容	7
1.4 论文的主要组织结构	8
第2章 舆情消息传播与预警	10
2.1 引言	10
2.2 传播网络分析	10
2.2.1 研究背景	10
2.2.2 算法原理	12
2.2.3 实验数据	13
2.2.4 评价指标	14
2.2.5 结果分析	15
2.3 意见领袖挖掘	16
2.3.1 研究背景	16
2.3.2 算法原理	17
2.3.3 实验数据	18
2.3.4 评价指标	18
2.3.5 结果分析	22
2.4 传播趋势预测	23
2.4.1 研究背景	23
2.4.2 算法原理	23
2.4.3 实验数据	25
2.4.4 结果分析	25

标题命名不明确

第3章 舆情文本分类与挖掘	27
3.1 引言	27
3.2 相关事件识别	27
3.2.1 研究背景	27
3.2.2 算法原理	28
3.2.3 实验数据	32
3.2.4 结果分析	33
3.3 评论情感分析	34
3.3.1 研究背景	34
3.3.2 算法原理	35
3.3.3 实验数据	40
3.3.4 评价指标	41
3.3.5 结果分析	41
3.4 新闻文本挖掘	42
3.4.1 研究背景	42
3.4.2 新闻文本摘要	42
3.4.3 新闻事件图谱	44
第4章 面向检务公开的舆情监测系统的实现	46
4.1 引言	46
4.2 网络爬虫的设计	47
4.3 数据存储的设计	48
4.4 运行逻辑的设计	50
4.5 前端界面的设计	53
结 论	55
参考文献	56
附录 A 舆情监测系统演示图	61
A.1 系统部分功能演示图	61
攻读硕士学位期间发表的论文及其他成果	63
哈尔滨工业大学学位论文原创性声明和使用权限	64
致 谢	65

A blurred background image of a business meeting in a conference room. Several people are seated around a large conference table, looking towards a presentation screen at the front of the room. A woman in a dark suit is standing by the screen, which displays a presentation slide. The room has large windows with a view of a city skyline.

正文部分



正文的撰写

1

绪论

课题背景、研究的目的和意义
国内外研究现状
研究内容和组织结构（章节安排）



2

主体

分章叙述
博士论文：4章
硕士论文：3-4章
本科论文：3章



3

结论

工作成果的总结
博士论文：2-3页
硕士论文：1-2页
本科论文：1页



绪 论





绪论的撰写

本课题的来源、背景、目的和意义

- 研究的理由、目的、背景，热点、前沿、共识问题

国内外研究现状

- 涉及多个子领域，则分类综述
- 只涉及一个子领域，则按国内和国外综述
- 相关领域的研究进展及成果
- 存在的不足或待深入研究的问题（前人的工作和知识空白）
- 引用文献最多的章节

主要研究内容

- 提出问题，将问题的解决划分为若干阶段，从内容上拆分为若干章节

章节安排

字数：3000 ~ 5000字

1.3 国内外的研究现状及分析.....
1.3.1 研究概况.....
1.3.2 多机协同目标分配方法研究现状.....
1.3.3 多机协同航迹规划方法研究现状.....
1.3.4 多机协同航迹重规划方法研究现状.....
1.3.5 多无人机多任务的协同方法概述.....
1.4 存在问题分析.....



绪论的撰写

第1章 绪 论

1.1 课题来源

本论文的研究内容源自“复杂环境下数目未知时^[6]探测、城市监控、航拍摄影、矿产探测、应急检测、气象预报、社会安全防范研究”(国家自然科学基金资助项目, 批准号: 611750^[6])等领域均有长足的发展。此外, 近年来随着无人技术的广泛推广, 无人系统扩展目标跟踪方法研究”(国家自然科学基金资助项^[7](Unmanned System)在各领域获得普遍关注, 除了无人飞行器系统(Unmanned Aircraft System, UAS)被概念化以外, 无人地面系统(Unmanned Ground System, UGS)^[7]和无人海洋系统(Unmanned Maritime System, UMS)^[8]均获得普遍关注。

1.2 课题的目的和意义

由此可见, 未来无人技术发展的趋势必然是陆地、天空、海洋以及外太空多种无人机(Unmanned Aerial Vehicle, UAV)是指不载无人系统一体化的综合应用^[3]。无人系统技术的全面推广, 将以更全面复杂的气动力进行飞行; 能够自主或远程领航控制; 可扩展方式影响人类的未来。

的无人飞行平台^[1]。与有飞行员驾驶的载人飞行器;

无人机技术和 MUAS 经过了几十年的发展, 提出了大量的理论和方法, 在不同的领域都取得了丰硕的成果。其支撑理论涉及到多个学科和技术领域的综合和交叉, 如自动控制、人工智能、学习模式、无线通信、分布式和最优化理论等, 目前还存在着大量亟待解决的问题, 在理论和实践上均具有较高的研究价值。同时, 由于执行的任务越来越复杂; 环境和约束的变化越来越宽泛; 协同和合作的目标越来越具体; 完成任务的时间和打击效果越来越精细, 这就要求 MUAS 必须具备更优良的技术和性能。因此, 只有对现有的技术进行改进和创新, 不断提出新的理论和方法, 才能够保证多无人机系统有效协同执行复杂多约束条件下的多个任务, 这也是本课题的研究目标。因此, 本文从统一目标分配建模、协同目标分配方法优化、协同航迹规划和在线航迹重规划几个方面展开研究, 重点研究三维环境下 MUAS 自主协同规划的若干问题。



国内外研究现状

1.2.2 软件错误定位的国内研究现状

Hao 等人^[33]提出的基于语句覆盖的错误定位方法 SAFL(Similarity-Aware Fault Localization)，通过统计语句的覆盖信息建立覆盖信息矩阵，来计算每条语句对测试结果的贡献度。如果一条语句对失败执行的贡献度越大，对成功执行的贡献度越小，那么它就更加可疑。

赵磊等人^[34]建立了基于覆盖向量的轨迹分析模型，认为高可疑代码与错误代码之间的关联关系能够使用覆盖向量来表示，并利用频繁集来挖掘与高可疑代码相关的错误代码。

谭德贵等人^[35]改进了基于频谱的错误定位算法，引入了失败测试用例的边际权重，失败测试用例对某条语句的覆盖次数越多，该测试用例的权重也越大。该方法可以一定程度上提高错误定位的效率。

许高阳等人^[36]提出了一种基于层次切片的错误定位方法，该方法利用层次切片模型和逐步求精算法，分别在包层、类层和方法层将通过测试的包、类和方法删去，以缩小错误存在的范围。该方法比较适合面向对象语言，而且适合于方法调用较少的程序。

Wen 等人^[37]提出了 PSS-SFL (Program Slicing Spectrumbased Software Fault Localization) 技术，该技术首先分析程序实体间的依赖关系，并删除与失败输出无关的实体。然后建立程序切片频谱模型，计算每个程序实体的可疑度。

Zhang 等人^[38]提出了 CP (Capture Propagation) 方法，认为一条语句的错误可能会影响到程序状态，而随着程序的继续执行，被影响的程序状态可能会被进一

不要流水账



国内外研究现状

1.2.3 软件错误定位的国内外研究现状分析

目前，很多技术已被应用于软件错误定位，包括程序切片技术、程序依赖分析、统计分析、概率图模型和图挖掘算法等，并提出了很多优秀的软件错误定位方法。它们大致可以分为基于程序状态修改的方法、基于程序行为特征对比的方法和基于程序依赖关系的方法这三类。

基于程序状态修改的方法，对小规模程序效果一般比较理想，但由于其算法复杂度都比较大，处理复杂程序时效率会大大降低，而且实现比较困难。除此之外，此类方法需要对程序状态进行强制修改，这样很容易导致程序发生运行异常，或者产生无法判断的程序结果。

基于程序行为特征对比的方法，一般思想比较简单，容易实现，也能够得到不错的错误定位结果。但这类方法对测试用例的质量和数量要求较高，而且往往只考虑了语句或谓词的覆盖信息，缺少对程序结构和执行信息的分析，会影响错误定位精度，而且难以给出与错误相关的上下文信息，开发者仍然需要做许多工作才能够理解并修复错误。

基于程序依赖关系的方法实现难度不大，能够充分考虑程序的依赖信息，需要使用的测试用例较少，而且也可以提供错误的相关行为描述。但其定位的精度仍有待提高，而且提供的上下文信息往往包含过多的冗余信息，会造成难以找到与错误相关的信息。

目前大多数方法在错误理解方面做的还不够，或者只提供可疑语句的列表，缺少可以用于理解错误产生的上下文信息，或者提供了过多的信息，开发者难以从中找到与错误相关的信息，这些都不利于开发者理解错误是如何产生的。因此，如何在给出具有较高精度的错误定位结果的同时，提供有利于开发者理解错误的上下文信息，是非常值得研究的。

1.3 课题研究的主要内容及章节安排

引用他人研究成果或采用他人成果时，应注明出处，不得将其与本人提出的理论分析混淆在一起



研究现状中的引文标注规范

引文标注遵照GB/T7714-2005

- 按顺序标注
- 正文中引用文献的标示应置于所引内容最后一个字的**右上角**，所引文献编号用阿拉伯数字置于方括号“[]”中
- 引用单篇文献时，如“**二次铣削^[1]**”。
- 同一处引用多篇文献时，各篇文献的序号在方括号内全部列出，各序号间用逗号，如遇连续序号，可标注起止序号
 - ...形成了多种数学模型^[7,9,11-13]...
- 多次引用同一文献时，在文献序号的[]后标注引文页码
 - ...间质细胞CAMP含量测定^{[3]101-108}...。 ...含量测定方法规定^{[3]92}...
- 当提及的参考文献需要在文中直接说明时，则用小4号字与正文排齐，如“由文献^[8,10-13]可知”
- 不得将引用文献标示置于各级标题处



研究现状中引用参考文献的要求

- # 所有被引用文献均要列入参考文献中，必须按顺序标注，但同一篇文章只用一个序号
- # 尽量引用原始文献
- # 当不能引用原始文献时，要将二次引用文献、原始文献同时标注



主要研究内容和章节结构

1.5 主要研究内容

本文从多机协同目标分配建模，目标分配求解，协同航迹规划和协同航迹重规划四个角度重点研究了三维环境中 MUAS 协同执行多个复杂目标的多任务协同规划问题，此研究可应用于军事行动中多无人机协同侦查、情报获取、空防一体化和打击敌方地面军事建筑设施等方面。在任务建模阶段，基于最优

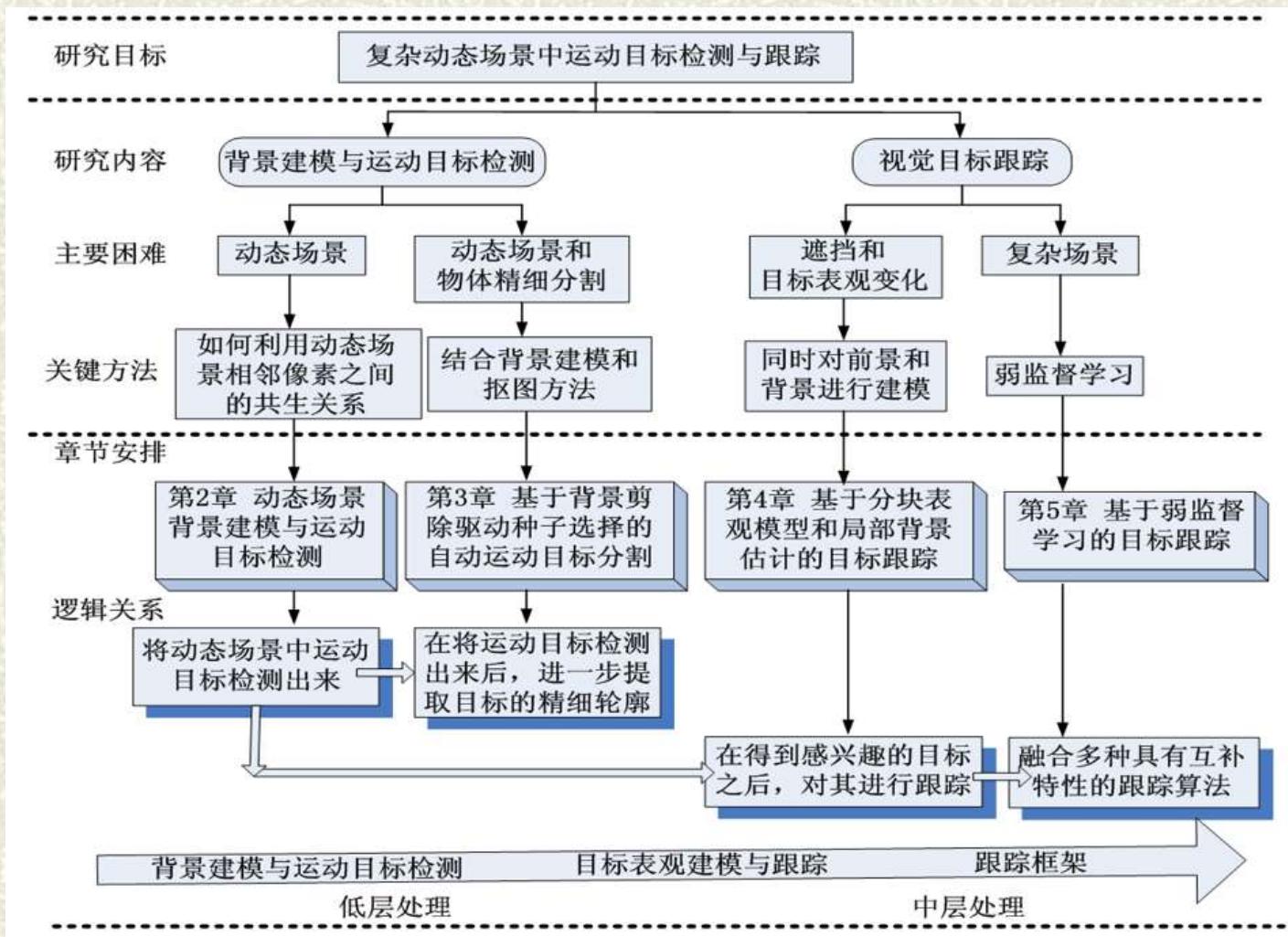
第 3 章研究 MUAS 多目标分配问题的表示方法，并改进离散差分进化算法的框架和策略以提高分配效率。首先基于统一分配模型和航程代价矩阵设计一致的基因编码规则，将分配关系与具有实际物理意义的航程代价关联在一起；利用航程代价映射分配关系，在连续实值空间进行差分操作，然后再将操作获得的临时解通过反射规则映射回离散的分配关系空间，保证获得有效的子代个体，并且提高了分配的合理性；进化过程中还动态混合差分进化策略，以维持算法的开发性与探索性的平衡，使进化过程既可以覆盖更广阔的空间，又可以搜索到更深入的层面，使算法始终向着全局最优的方向前进。

第 4 章在获得了有效的 MUAS 目标分配结果之后，重点研究这些目标执行的全局航迹规划问题。根据三维空间栅格点与任务代价和威胁区的关系设计隶



章节内容关系图示

论文各章之间应该前后关联，构成一个有机的整体





章节内容关系图示

软件
Bug
的报
告生
成与
理解

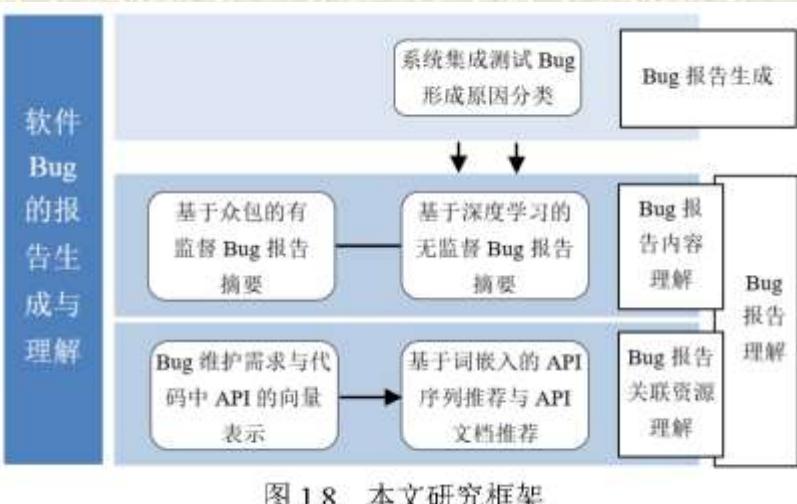


图 1.8 本文研究框架

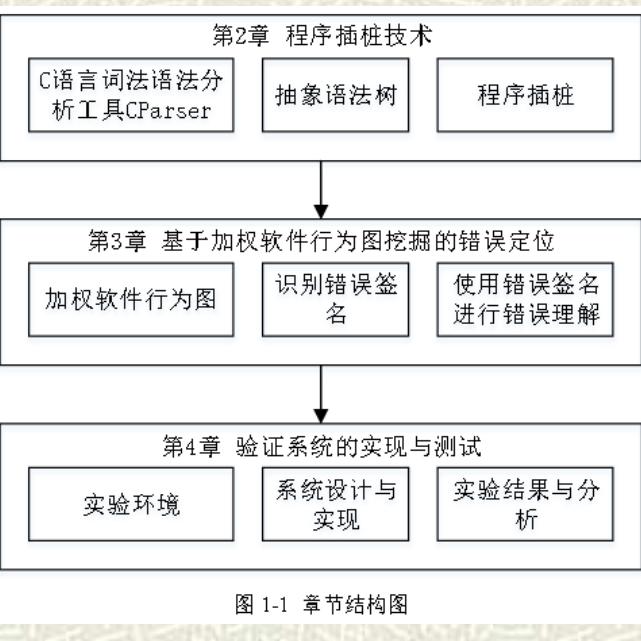


图 1-1 章节结构图



图 1-7 论文主要研究内容及其关系示意图



图 1-2 本文主要工作示意图。



正文





正文字字体、字号

- # 各级标题用**黑体**
- # 正文**小4号宋体字**

- # 中文字体为**宋体**
- # 数字、英文为**新罗马字体**

第4章 基于空间模糊表示的多机协同航迹规划方法

4.1 引言

小3号黑体

多机协同航迹规划是 MUAS 任务规划的重要环节，是实现无人机自主飞行与任务执行的重要阶段。协同航迹规划问题是指在已知、部分已知或未知信息的环境中，预先规划出从各个 UAV 起始点到对应目标点，可以绕过途中各种威胁区和障碍物，安全、可靠、相互无碰撞，且同时满足 UAV 自身约束条件和协同约束限制的多条可行飞行航迹。该问题是一个复杂的 NP 完全问题，因此，求解难度大，是任务规划系统研究的难点之一。

在三维空间进行多机协同航迹规划，需要综合考虑各无人机的运动特性和避障需求，以及它们之间的通信和数据交换。由于三维空间的复杂度，才导致计算量不足，是导致计算量其反映规划问题的关键因素。因此，如何在保证规划质量和效率的前提下，优化航迹规划的策略，是优化航迹规划的研究重点。

1.3 国内外的研究现状及分析

1.3.1 研究概况

4号黑体

近年来，无人机的研究和应用在多数发达国家呈几何级数的趋势增长，这些国家投入了大量的人力和财力组织研发，涌现出一批性能优良的新型无人机

- 小2号黑体
- 小4号宋体
- 小4号新罗马



正文的撰写规范

正文格式

- 各层次题序及标题不置于页面的最后一行（孤行）
- 章序及章名居中排版，章序用阿拉伯数字
- 各节、条题序顶格书写，与标题间空一格，下面阐述内容另起一段
- 排版时每一页下方不要留太多空白，可把后面文字调整上来

第 4 章 基于空间模糊表示的多机协同航迹规划方法

4.1 引言

多机协同航迹规划是 MUAS 任务规划的重要环节，是实现无人机自主飞行与任务执行的重要阶段。协同航迹规划问题是指在已知、部分已知或未知信息的环境中，预先规划出从各个 UAV 起始点到对应目标点，可以绕过途中各种威胁区和障碍物，安全、可靠、相互无碰撞，且同时满足 UAV 自身约束条件和协同约束限制的多条可行飞行航迹。该问题是一个复杂的 NP 完全问题，因此，求解难度大，是任务规划系统研究的难点之一。

在三维空间进行多机协同航迹规划时，要采用有效的方法降低三维空间的复杂度，才能使规划得以顺利进行。通常在三维栅格上进行航迹规划，能够较快处理三维规划的问题。然而，三维栅格上各个栅格点的信息量不足，是导致计算量增加的主要原因。实际上，在空间中不同位置的点，其反映规划问题的关键程度是不同的，因此考虑三维空间中栅格点的表示问题，是优化航迹规划的重要前提。



正文的撰写内容

示例

- 第1章：绪论
- 第2章：基础理论
- 第3章：xx方法研究
- 第4章：xx方法研究
- 第5章：xx原型系统的实现和测试
- 结论



正文的撰写要求

论证的逻辑

- ◆ 层次感，而不是平面感
- ◆ 缝密性，而不是一盘散沙
- ◆ 科学性，而不是宣传性
- ◆ 学理性，而不是口语化
- ◆ 严谨性，而不是随意性
- ◆ 围绕核心问题展开论证



正文字数要求

博士论文

一般为110页左右
(6-10万字)

双面打印

页眉分别为章标题和“
哈尔滨工业大学工学博
士学位论文”

硕士论文

一般为60页左右
(3-5万字)

单面打印

页眉全部为“哈尔滨工
业大学工学硕士学位论
文” / “哈尔滨工业大
学工程硕士学位论文”

本科论文

一般为40页左右

单面打印

页眉为“哈尔滨工业大
学本科毕业设计（论文
）”



论文主体部分的撰写内容和要求

提出问题

文献综述 (第二手数据)
对问题系统论述和分析

设计方案

设计研究框架
设计并论证模型或方案
模型或算法的实现
突出本文创新或改进

实验验证

案例论证或实证分析 (第一手数据)
结果分析或建议、改进措施等

内容充实，阐述全面客观

论证有力，观点明确，理论分析深入

论据充分，数据真实可信

重点写本文的工作，少引文献或参考书籍中的内容



论文主体中的本章小结

- # 各章小结是对各章研究内容、方法与成果的简洁准确的总结与概括，也是论文最后结论的依据
- # 本章小结中不要出现图表和文献引用

3.7 本章小结

为高效求解三维环境下的 MUAS 协同目标分配问题，本章提出了一种基于统一分配模型的协同多目标分配方法。首先，基于 UAVs、目标点序列和它们对应的航程代价，提出了一种统一的基因编码策略，该编码策略在一致的框架下处理各种目标分配模型，并将航程代价加入基因编码中，方便了从离散空间到连续空间的转换。然后，提出了适用于 DE 算法的映射和反映射规则，以解决离散的组合优化问题。改进的算法将无人机和目标对应的离散空间转换为连续的飞行航程代价空间，并用代价值进行差分操作。接下来，为了获得合理的子代个体，又根据反映射规则将差分操作的临时结果转换回无人机和目标的对应关系空间。在此方法中，采用距离信息指导搜索的方向，易于获取实际问题的最优解。最后，在进化过程中引入动态交叉率和混合进化策略，以平衡算法的开发性和探测性；同时，扩展了代间变异机制，提高了搜索种群的覆盖范围，确保算法能够有效跳出局部最优的陷阱。理论分析和实验结果验证了该算法优于同类方法，能够快速获得可行的最优解。



文字表述问题示例

• 2.3 Word2Vec 词向量模型的优化。

经过以上背景分析本文在计算大知识库中问句相似度计算过程中，本章采用常用基于词向量的问句相似度计算方式进行问句匹配。考虑到所有基于词向量的相似度计算方案，均依赖词向量的好坏^[27]，词向量的好坏直接决定了问句相似度的正确率。而根据分析调查，传统词向量主要是指 Word2Vec 生成的词向量，针对该模型生成的词向量词向量仍存在很多问题，其中包括近反义词难以区分，未登录词处理不当等^[28]，并且原有的基于词向量的问句相似度计算方案中因而本章将针对这些问题给出解决方案，在目前的词向量生成模型的基础上优化词向量，从而能够提高基于词向量模型的问句相似度计算的整体效果。

缺少标点，文字不通顺，主语变化，表述不清、逻辑断裂



文字表述问题示例

基于深度学习的小样本学习希望深度网络在含有大量标注的基本类样本上预训练后，使得网络在遇到只有少量标注的新类别样本时，能够实现对这些新类别的快速泛化识别。当前实现小样本学习的一个主流方法是元学习（Meta-Learning）。然而，基于元学习的方法不仅需要大量的训练时间和计算资源，而且性能也无法令人满意。近期，基于传统监督表示学习的小样本学习方法被证明能够取得和元学习方法相媲美的结果，且拥有模型结构简单、训练时间短的特点，显示出了惊人的潜力。为了进一步提升小样本学习模型的性能，本文基于传统的监督表示学习策略，提出了多种改进方法。本文的主要

文字不通顺，主谓宾结构错误

