



燕山大学  
YANSHAN UNIVERSITY

# 硕士学位论文

MASTER'S DISSERTATION

论文题目 移动机器人视觉 SLAM 闭环检测  
算法研究

作者姓名 张朋珍  
学位类别 工学硕士  
指导教师 朱奇光副教授

2017 年 5 月



中图分类号：TP242.6

学校代码：10216

UDC：621.3

密级：公开

## 工学硕士学位论文

# 移动机器人视觉 SLAM 闭环检测算法研究

硕士研究生：张朋珍

导师：朱奇光副教授

申请学位：工学硕士

学科专业：光学工程

所在单位：信息科学与工程学院

答辩日期：2017 年 5 月

授予学位单位：燕山大学



A Dissertation in Optical Engineering

# **CLOSED-LOOP DETECTION BASED VISION SLAM OF MOBILE ROBOT**

by Zhang Pengzhen

Supervisor: Vice Professor Zhu Qiguang

**Yanshan University**

May, 2017



## 燕山大学硕士学位论文原创性声明

本人郑重声明：此处所提交的硕士学位论文《移动机器人视觉 SLAM 闭环检测算法研究》，是本人在导师指导下，在燕山大学攻读硕士学位期间独立进行研究工作所取得的成果。论文中除已注明部分外不包含他人已发表或撰写过的研究成果。对本文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。本声明的法律结果将完全由本人承担。

作者签字：

日期： 年 月 日

## 燕山大学硕士学位论文使用授权书

《移动机器人视觉 SLAM 闭环检测算法研究》系本人在燕山大学攻读硕士学位期间在导师指导下完成的硕士学位论文。本论文的研究成果归燕山大学所有，本论文的研究内容不得以其它单位的名义发表。本人完全了解燕山大学关于保存、使用学位论文的规定，同意学校保留并向有关部门送交论文的复印件和电子版本，允许论文被查阅和借阅。本人授权燕山大学，可以采用影印、缩印或其它复制手段保存论文，可以公布论文的全部或部分内容。

保密☐，在 年解密后适用本授权书。

本学位论文属于

不保密☐。

(请在以上相应方框内打“√”)

作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日





## 摘 要

移动机器人的闭环检测这一环节是机器人同时定位和地图构建(SLAM)中判断自身是否需要更新地图的关键步骤。由于机器人自身视觉传感器等因素引起的误差,不能保证闭环检测的准确度。因此,本文对现有闭环检测算法主要步骤存在的问题展开分析,主要研究内容如下:

首先,概述课题研究意义和背景,对视觉闭环检测算法的研究现状和一些关键问题进行分析,详细分析了 SLAM 的基本方法以及闭环检测中常用的一些特征检测算法,为后续的研究提供理论模型和依据。

其次,针对图像关键点匹配准确性差的问题,采用基于颜色特征结合改进加速鲁棒特征算法(SURF)的图像匹配算法。利用颜色特征进行图像的粗匹配,选取与测试图像最相近的 5 幅作为待匹配对象,利用 Krawtchouk 矩将 SURF 方法中 Hessian 矩阵获取的关键点计算出梯度方向和幅值,得到特征向量,对待匹配图像提取改进 SURF 特征再与测试图像进行精确匹配,进行仿真实验,证明改进算法提高了匹配的正确率。

再次,针对闭环检测中的视觉词袋模型(BoVW)准确率低的问题,设计出一种改进 BoVW 方法,采用基于 Krautchouk 矩的 SURF 算子对基于指导滤波的去雾后的场景图像进行特征提取,利用最大最小距离结合粒子群聚类的方法构建视觉单词本,通过图像分类实验,验证了改进算法的有效性。

最后,针对闭环检测算法的准确度不足问题,设计基于后验处理的闭环检测方法,利用基于 Krautchouk 矩的 SURF 算法提取图像特征,构建金字塔式的分层视觉单词树,通过加权词频-逆向文件频率(TF-IDF)熵得分计算方法进行相似性度量,通过与阈值的比较确定初始闭环,利用时间上的约束和极线几何关系的后验处理得到有效正确的闭环。最后通过实验,验证了算法的可行性。

**关键词:** 机器人; 闭环检测算法; Krautchouk 矩; BoVW 模型; TF-IDF

## Abstract

The closed loop detection of mobile robot Simultaneous Localization And Mapping (SLAM) is an important part of the robot to judge whether it needs to update the map. Due to the error caused by the robot's vision sensor, the accuracy of the closed loop detection cannot be guaranteed. Therefore, in this paper, the existing problems of the important steps in the closed-loop detection algorithm are studied. The major research contents are as follows:

First, the significance and background of the research are introduced. The research status and some key problems of the closed-loop detection algorithm are analyzed. And the basic methods of SLAM and some characteristic detection algorithms in closed-loop detection are analyzed in detail. Which providing the theoretical model and basis is for the following work.

Second, in view of the problems of slow image processing speed and accuracy of feature points matching in mobile robot vision-based SLAM. The paper proposes a novel matching method based on color feature and improved Speeded Up Robust Features (SURF) algorithm. First, image sequences are roughly matched by color characteristics, and make sure five images that is most similar to the test image. Then, Hessian matrix is used to get the feature points that described by Krawtchouk moments, and calculate the feature points' gradient direction and amplitude. And then, the new feature vector has been obtained. The simulation experiment proves that proposed algorithm improves the precision and efficiency of image matching.

Again, an improved Bag Of Visual Words(BoVW) method is designed for the low accuracy of the BoVW model in closed-loop detection. The SURF operator based on Krautchouk moments is used to extract the features of the de-fogged scene based on the guidance filter. Using the maximum and minimum distance function combining particle swarm clustering algorithm to construct visual words, the contrastive experiments of image classification are carried out to prove that the proposed algorithm improves the accuracy rate.

Finally, for the problem of the lack of accuracy in closed-loop detection method, the new closed-loop detection method based on posterior processing is designed. The SURF algorithm based on Krautchouk moment is used to extract the image features, and a pyramid hierarchical visual word tree is constructed. And the similarity measure is calculated by the modified weighted Term Frequency-Inverse Document Frequency (TF-IDF) entropy score calculation method. The initial closed loop is determined by comparing with the threshold value, and the correct closed loop is obtained by posterior processing of time constraint and geometric relation of polar line. The experiments are carried out to verify the feasibility of the proposed algorithm.

**Keywords:** robot; closed loop detection; Krautchouk moment; BoVW model; TF-IDF



## 目 录

摘 要 .....	I
Abstract .....	II
第 1 章 绪 论 .....	1
1.1 课题研究背景和意义 .....	1
1.2 闭环检测的研究现状 .....	2
1.3 闭环检测的关键问题 .....	3
1.4 文章主要内容及结构 .....	3
第 2 章 闭环检测基础理论介绍 .....	5
2.1 SLAM 基础理论 .....	5
2.2 场景图像描述 .....	6
2.2.1 Harris 算子 .....	6
2.2.2 SIFT 算子 .....	7
2.2.3 SURF 算子 .....	10
2.3 本章小结 .....	12
第 3 章 基于多特征分层次的图像匹配算法研究 .....	13
3.1 引言 .....	13
3.2 改进 SURF 算法 .....	13
3.2.1 Krawtchouk 矩不变量的构造 .....	13
3.2.2 特征点描述 .....	16
3.2.3 特征向量匹配 .....	17
3.3 多特征分层次的图像匹配算法 .....	17
3.3.1 颜色特征提取 .....	17
3.3.2 颜色特征匹配 .....	18
3.3.3 多特征分层次匹配算法 .....	19
3.4 实验结果与分析 .....	20
3.4.1 图像特征匹配实验 .....	20
3.4.2 改进 SURF 的性能比较 .....	23
3.5 本章小结 .....	25
第 4 章 基于改进 BoVW 算法的研究 .....	26
4.1 引言 .....	26
4.2 BoVW 模型描述 .....	27
4.3 改进的 BoVW 算法 .....	28

4.3.1 基于指导滤波的 BoVW 模型去雾预处理.....	28
4.3.2 基于改进的聚类算法构建视觉单词树 .....	30
4.3.3 基于信息熵的图像相似性度量 .....	31
4.4 实验研究及分析 .....	33
4.4.1 提取的特征点对比 .....	33
4.4.2 聚类效果对比 .....	34
4.4.3 图像分类的效果对比 .....	35
4.5 本章小结 .....	36
第 5 章 基于后验处理的闭环检测算法研究 .....	38
5.1 引言 .....	38
5.2 闭环检测算法 .....	38
5.3 改进的闭环检测算法 .....	40
5.3.1 视觉单词树的创建和描述 .....	40
5.3.2 基于改进的 TF-IDF 熵的构造 .....	41
5.3.3 基于加权 TF-IDF 熵的得分匹配方法 .....	42
5.3.4 基于后验处理的闭环检测方法 .....	43
5.4 实验研究及分析 .....	45
5.5 本章小结 .....	48
结 论 .....	49
参考文献 .....	50
攻读硕士学位期间承担的科研任务与主要成果 .....	55
致 谢 .....	56

## 第1章 绪论

### 1.1 课题研究背景和意义

从机器人的出现开始,人工智能技术的迅猛发展和不断创新,人们对机器人的热情和研究越发高涨,移动机器人已成功应用于各行各业,在军事、工业、医疗、农业都有巨大的需求。特别是在军事反恐、太空探测、地下和海洋开发、火山灾难等完成高危险性和高难度工作时,移动机器人代替人类作业是必然趋势。近年来,随着世界各国对机器人的投入水平和研究力度的不断提高,对于机器人的研究水平和技术已上升到一个新层次,具有高度智能性和自主性的移动机器人纷纷进入人们的生产和生活中<sup>[1-3]</sup>。近期,AlphaGo机器人与围棋高手的“人机大战”受到了社会广泛关注,最终机器人大比分战胜人类,体现出了机器人的高度智能性与超强学习能力。

对自主移动机器人研究的深度和层次也表明了一个国家的综合实力和科学技术水平。在未知环境条件下,移动机器人通过自身传感器感知的信息进行自主的导航、路径规划等运动决策即自主移动机器人。自主导航技术是自主移动机器人的核心问题,也是完成实际应用的关键。自主导航可总结为三个基本问题:“我在哪?”、“要去何地?”、“怎么去?”对应于移动机器人的定位问题、地图创建问题、路径规划任务。同时定位和地图构建(Simultaneous Localization And Mapping, SLAM)是移动机器人自主性的基础和关键问题<sup>[4,5]</sup>。定位即机器人在整个环境中确定自身的位置、姿态,包括:声纳、里程计、码盘、视觉定位等方法;地图创建即对所处环境的描述与模型构建,创建地图分为三个基本问题:对所处环境信息的获取、地图的表示以及地图的更新,地图的表示方法一般概括为三种:栅格地图、特征地图以及拓扑地图<sup>[6]</sup>。

闭环检测(Closed-Loop Detection, CLD)问题的研究是移动机器人SLAM过程中的数据关联问题,指的是机器人在不停地自主判断自身所处的环境是否为之前遍历的相同位置,以此来决定是否要更新环境地图,对减少移动机器人位姿状态和地图状态变量不确定性十分重要,避免错误和冗余结构引入地图<sup>[7,8]</sup>。因为移动机器人自身装备产生的累积误差,在长时间和大规模环境下的SLAM过程中所引发的地图构

建不准确和定位错误等问题显得尤为突出，闭环检测的目的是实现移动机器人的精确定位和地图构建，完成移动机器人的 SLAM 任务<sup>[9]</sup>。

## 1.2 闭环检测的研究现状

基于视觉传感器获取信息的 SLAM 就是 VSLAM(Vision-based Simultaneous Localization And Mapping, VSLAM)技术，闭环检测问题是其中的难题之一<sup>[10]</sup>。根据现有的研究方法，闭合环路检测算法大致概括为三类。

(1)地图与地图间的匹配方法 Clemente 和 Davison 等在 2007 年提出了一种闭环检测算法，该算法通过获取场景图像的相似度及其特征的相对距离，计算得出一对子图间最大的公共特征集，主要是通过寻找子图之间公共特征的相同之处来实现闭环检测<sup>[11]</sup>。然而，该方法需要利用稀疏地图来实现，信息量不够，闭环检测的精确度不能保证，所以此方法不适合基于视觉的 SLAM。虽然在 2011 年进行了改进<sup>[12]</sup>，但仍需要多次创建地图，在提高闭环检测精确度的同时增加了算法复杂度和系统负担，不适合实时性应用。

(2)图像与地图间的匹配方法 Williams 和 Cummins 等提在 2008 年提出将重定位技术引入了闭环检测过程，此算法利用获取的场景图像与地图上特征点的相同点，通过随机抽样一致性算法和三点位姿算法得到地图上相似点的位姿信息。但此方法需要利用分类用的分类器，其占用内存过大，不适合大规模环境下应用<sup>[13]</sup>。祝继华和周颐等在 2015 年提出通过提取图像的特征进行图像拼接的方式来实现栅格地图的拼接。但由于此方法利用局部收敛的循环最近点算法计算目标函数，使得算法具有不确定性，鲁棒性差。

(3)图像与图像间的匹配方法 Angeli 和 Newman 等人在 2008 年提出一种利用视觉“词袋(bags of words)”来完成图像匹配方法。该算法利用提取场景图像特征来构建词袋模型，通过视觉词典集来表征图像，计算场景图像间的相似程度并计算两者来自同一位置的概率来实现环路闭合检测<sup>[14]</sup>。此方法已投入实际的应用，基于图像与图像间匹配的方法也做出诸多改进。Cummins 和 Newman 等人在 2011 年在原方法的基础上进行了改进，通过加速鲁棒特征算法对场景图像进行特征提取，在保证准确率的同时减少了闭环检测时间，但由于在视觉“单词”集建立时进行了离线训练，给系统带来负担<sup>[15,16]</sup>；梁志伟和陈燕燕等人在 2013 年提出的一种基于图像建模的改进方法，在线构建视觉词典树，改善了 Cummins 方法的缺点，但其利用



了传统的 K-means 算法,使得结果极易生成局部最小<sup>[17]</sup>;李永峰和张国良等人在2015年提出了一种基于历史模型集的闭环检测方法,把连续关键帧的特征点确定为一个历史模型集合,然后利用帧到模型配准模型成功完成了闭环检测。综上所述,虽然对闭环检测的研究已取得不少成果,但在不同环境中仍然存在准确度不够、实时性不强等问题。

将闭环检测问题转化为场景图像间的匹配,也是最适合于实际应用的方式<sup>[18]</sup>。所以,本文进行的闭环检测工作也是采用图像与图像之间的匹配这一方式来进行的。

### 1.3 闭环检测的关键问题

(1)对环境的有效采样 移动机器人在行走时,利用视觉传感器采集视觉场景,获取环境内容的不同表示不同的位姿信息,所以高质量、低重复率、全面的场景图像的获取显得尤为重要,因其直接影响着场景图像匹配的正确率,从而影响闭环检测的准确度。依据机器人行走速度和传感器帧率来设置传感器的采样频率同样重要,从而保证全覆盖、低重复度的场景图像库的建立。

(2)对环境外观的有效建模和表达 对获取的场景外观进行处理是闭环检测的基础环节,主要通过相机参数标定、畸形图像校正、对场景图像底层特征的提取和描述以及场景图像的建模等技术。传统方法中,直接利用图像处理的方法解决SLAM的视觉问题存在缺陷,如:图像连续性和局部相似性等图像本身存在的特殊性,引起广泛关注与研究,并取得积极成果<sup>[19,20]</sup>。

(3)闭环检测的后验处理 由于机器人进行场景图像获取的连续性、场景图像的复杂度、自身传感器的感知误差等原因,会导致初始的闭环响应存在一些错误,所以对闭环响应的后验处理是一个必须环节。现有算法利用对极几何约束进行约束,只有满足条件的闭环才设定为正确闭环,然后更新地图。只利用单一方法进行闭环检测后验处理明显准确度不够,所以,寻找更多有效的正确闭环确定操作,从而得到更加可靠的闭环结果,又是一重要环节。

### 1.4 文章主要内容及结构

本课题研究的主要内容在河北省自然科学基金项目“基于多传感器融合的不确定移动机器人导航及环境建模研究(F2016203245)”下展开。通过对移动机器人

VSLAM 中的闭环检测问题研究，在阅读大量国内外文献的情况下，以提高闭环检测的准确度为中心，本文的主要研究内容有：对获取场景图像的特征提取与匹配、BoVW 模型的研究与改进、视觉字典树相似性得分的优化、基于后验处理的闭环检测算法。

本文的内容结构与章节安排如下：

第1章为绪论，叙述课题的研究背景，并对移动机器人的闭环检测研究现状进行了描述，分析了其中存在的问题和难点。

第2章对 SLAM 中闭环检测的相关理论进行了介绍。主要包括 SLAM 的基本理论以及几种闭环检测常用的特征检测算法。

第3章针对移动机器人 SLAM 中图像关键点匹配准确度不高的问题，提出采用颜色特征和改进 SURF 算法融合的方法。首先，采用颜色特征对图像集进行粗匹配，选取与测试图像最相近的 5 幅作为待匹配对象；其次，改进 SURF 算法，利用 Krawtchouk 矩得到新的特征向量，再与测试图像精确匹配，得到最佳匹配结果，进行仿真实验，验证其性能。

第4章针对闭环检测中的 BoVW 模型准确率低等问题，提出一种改进的 BoVW 方法，首先对场景图像进行基于指导滤波的去雾预处理，利用基于 Krawtchouk 矩的 SURF 算子进行场景图像的特征提取，然后，利用基于最大最小距离改进的粒子群聚类建造视觉单词本，然后通过基于信息熵的空间金字塔的直方图表示场景图像，使得图像间的相似度计算准确简单，最后通过图像分类的对比实验，验证算法的有效性。

第5章针对闭环检测算法的精确度不足问题，提出基于后验处理的闭环检测方法，对移动机器人获取的场景图像，利用基于 Krawtchouk 矩改进的 SURF 算法提取图像的特征，然后，构建金字塔式的分层视觉单词树，通过加权 TF-IDF 熵得分方法求得图像间的相似度，通过与阈值的比较从而确定初始的闭环，而后利用时间上的约束和极线几何关系的后验处理进行错误闭环的剔除，获取有效正确的闭环。

最后，总结本文所做工作的研究成果以及仍存在的问题，对下一步的工作进行展望。

## 第2章 闭环检测基础理论介绍

本章分析了 SLAM 的研究现状和基本方法以及闭环检测中常用的一些特征检测算法。

### 2.1 SLAM 基础理论

机器人进行自主的定位是移动机器人的一个基础且关键的问题。虽然目前可以通过全球定位系统等设备进行机器人的定位和导航，但在一些特殊环境不能奏效，而且由于时间消耗和通讯质量等原因不能保证实时性，同时定位和地图构建可以有效解决这些问题。移动机器人即时定位与地图创建是指，在未知环境，移动机器人在自己的位置不能确定的状况下创建地图，同时利用地图自行的导航和定位。在这过程中，移动机器人需要确定自己在所处环境中的位置，同时也需要进行环境地图的创建，两者是密不可分、相互依存的。

SLAM 的主要问题包括：环境的表达大致概括为栅格表示、特征地图、拓扑地图。栅格地图即把一个整体切分成一系列相等的栅格，判断出每一个有障碍物与否，不考虑环境的其它特征；特征地图也叫几何特征表示，对机器人获取的环境感知信息进行提取和抽象，用线段、边、角等几何信息描述环境；拓扑地图是将环境表示为一张有拓扑意义的图，图中的各个节点对应环境中各个地点，连接各个节点间的弧线相当于各个地点间的路径。不确定信息的处理：在未知环境中，在地图表示时，运动和感知信息的不确定性；在迭代过程，旧信息与新信息关系的处理；依据获得的信息进行的决策具有不确定性。数据的关联即将不同地点和不同时刻的感知信息匹配和联合，从而实现全局环境地图获取和定位。自定位方法是将移动机器人的定位方法分为：基于地图定位方法和无地图定位方法。探索规划方法是为了提高机器人创建地图的效率，在尽量少的时间感知尽可能大的范围。

现有的 SLAM 方法大致包括两种：一种是滤波器相关的，另一种是平滑相关的。基于滤波的方法主要运用迭代贝叶斯原理，通过概率计算来实现。常见的方法有：粒子滤波的方法、稀疏扩展信息滤波的方法和扩展卡尔曼滤波的方法等。基于平滑的方法中，利用所有控制信息和观测来完成移动机器人的运动轨迹获取和地图构建，常见方法有平方根 SAM。

## 2.2 场景图像描述

在视觉 SLAM 中的首要问题即对图像进行准确有效地描述。机器人获取的场景图像包含丰富信息，如何提取有效信息、简化图像的描述尤为重要。因此，选择一种快速且鲁棒的描述子十分关键。首先介绍几种常用描述子，重点介绍 SURF 方法。

### 2.2.1 Harris 算子

C. Harris 和 M.J. Stephens 于 1988 年研究出的一种特征点提取算子，通过微分运算来检测角点的方法，被称作 Harris 算子<sup>[21]</sup>。

Harris 算子检测公式为

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (2-1)$$

式中  $I(x, y)$  ——像素点  $(x, y)$  的灰度；

$I(x+u, y+v)$  ——移动了  $(u, v)$  的灰度；

$w(x, y)$  ——高斯窗函数。

Harris 算子用  $\frac{\partial I}{\partial x}$ ,  $\frac{\partial I}{\partial y}$  表示图像在某个像素点  $(x, y)$  的灰度变化方向，如果两者

的变化足够大，则确定为角点。 $\mathbf{M}$  自相关矩阵的定义如下

$$\mathbf{M} = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 & \frac{\partial I \partial I}{\partial x \partial y} \\ \frac{\partial I \partial I}{\partial x \partial y} & \left(\frac{\partial I}{\partial y}\right)^2 \end{bmatrix} \quad (2-2)$$

为了避免特征值的计算，实际的响应函数为

$$R = \det \mathbf{M} - k (\text{trace} \mathbf{M})^2 \quad (2-3)$$

式中  $k$  ——经验参数；

$\det \mathbf{M}$  ——矩阵行列式；

$\text{trace} \mathbf{M}$  ——矩阵的迹，若  $R$  大于所设定的阈值，则确定为角点。

Harris 算子的基本实现步骤如下：

(1) 对于图像  $I(x, y)$  求出在  $x, y$  方向的梯度  $\mathbf{I}_x, \mathbf{I}_y$ ；

(2) 分别求出在  $x, y$  方向的梯度乘积， $\mathbf{I}_x^2 = \mathbf{I}_x \cdot \mathbf{I}_x$   $\mathbf{I}_y^2 = \mathbf{I}_y \cdot \mathbf{I}_y$   $\mathbf{I}_x \mathbf{I}_y = \mathbf{I}_x \cdot \mathbf{I}_y$ ；

(3)进行高斯加权,  $I_x^2 \otimes w \quad I_y^2 \otimes w \quad I_x I_y \otimes w$ ;

(4)计算  $R$  值,  $R = \left[ I_x^2 \times I_y^2 - (I_x I_y)^2 \right] - k (I_x^2 + I_y^2)^2$ ;

(5)设定合理阈值, 确定角点。

虽然 Harris 角点检测是一种计算简便、方便操作的角点检测算法, 具有旋转、仿射不变性, 但由于其对边缘的敏感性和不是尺度无关的缺点, 应用时具有局限性。

## 2.2.2 SIFT 算子

尺度不变特征变换的方法(Scale-Invariant Feature Transform, SIFT)是由 David G. Lowe 在 1999 年发明、并在 2004 年做了完善, 是用来描述图像局部特征的描述子, 能够保持尺度、回转、平移不变的特性<sup>[22]</sup>。SIFT 算法主要包括以下几个部分。

### 2.2.2.1 检测尺度空间极值点及精确定位

一个多尺度空间函数  $L(x, y, \sigma)$  通过图像  $I(x, y)$  与尺度空间唯一可用的核函数高斯函数  $G(x, y, \sigma)$  卷积得到

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (2-4)$$

$$G(x_i, y_i, \sigma) = \frac{1}{2\pi\sigma^2} \exp \left[ -\frac{(x-x_i)^2 + (y-y_i)^2}{2\sigma^2} \right] \quad (2-5)$$

式中  $(x, y)$ ——在图像中的位置坐标;

$\sigma$ ——尺度因数。

之后进行降阶采样操作, 得到由底到顶的金字塔型图像, 同时对得到的金字塔中各层图像进行不同尺度的高斯滤波, 然后一幅图像产生了多组的(Octaves)图像, 每组都含有多层图像, 如图 2-1 所示。

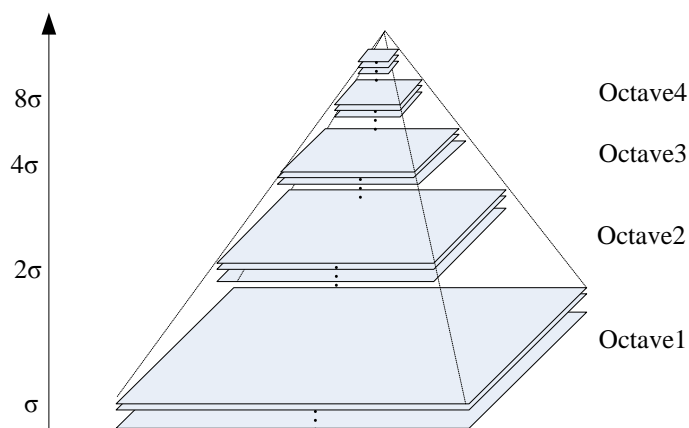


图 2-1 金字塔型图像

如图 2-2 所示, 采用金字塔邻接层相减方式可得到高斯差分图像, 高斯差分图像的计算公式为

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2-6)$$

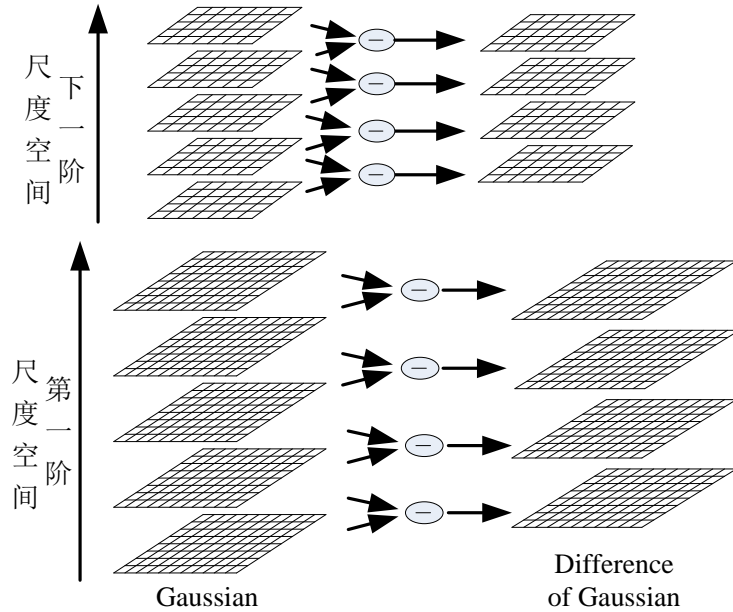


图 2-2 高斯差分图像

生成高斯差分图像, 然后计算极值点。极值点是通过以自身所在当前层的 8 个像素点, 上下邻接层的 9\*2 个像素点的总数 26 个像素点进行对比所得出的。求出极值点后, 要对极值点进行筛选, 需排除图像的边缘点, 还要去除对噪声敏感的不稳定点。

利用高斯差分函数的 Taylor 展开进行曲线拟合

$$D(\mathbf{X}) = D + \frac{\partial D^T}{\partial \mathbf{X}} \mathbf{X} + \frac{1}{2} \mathbf{X}^T \frac{\partial^2 D}{\partial \mathbf{X}^2} \mathbf{X} \quad (2-7)$$

其中,  $\mathbf{X} = (x, y, \sigma)^T$ 。

由公式(2-7), 求  $\mathbf{X}$  偏导, 并令其等于 0 可得极值点的位置

$$\hat{\mathbf{X}} = -\frac{\partial D^T}{\partial \mathbf{X}} \left( \frac{\partial^2 D}{\partial \mathbf{X}^2} \right)^{-1} \quad (2-8)$$

可以通过极值点的函数值  $D(\hat{\mathbf{X}})$  去除低对比度的点。相应极值点的函数值为

$$D(\hat{\mathbf{X}}) = D + \frac{1}{2} \frac{\partial D^T}{\partial \mathbf{X}} \mathbf{X} \quad (2-9)$$

若  $\left|D(\hat{\mathbf{X}})\right| \geq 0.04$ ，则这个极值点予以保留，反之丢弃。

然后，计算极值点的矩阵  $\mathbf{H}$

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (2-10)$$

令  $\alpha, \beta$  是  $\mathbf{H}$  的最大和最小特征值，则

$$\text{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta \quad (2-11)$$

$$\text{Det}(\mathbf{H}) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta \quad (2-12)$$

式中  $\text{Tr}(\mathbf{H})$ ——矩阵  $\mathbf{H}$  对角线元素之和；

$\text{Det}(\mathbf{H})$ ——矩阵  $\mathbf{H}$  的行列式。

设  $\alpha = \gamma\beta$ ，可得

$$\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\gamma + 1)^2}{\gamma} \quad (2-13)$$

当  $\frac{\text{Tr}(\mathbf{H})^2}{\text{Det}(\mathbf{H})} < \frac{(\gamma + 1)^2}{\gamma}$  将极值点保留，否则剔除。在 Lowe 的原文中， $\gamma$  取 10。

#### 2.2.2.2 为关键点进行方向分配

为每个关键点赋予方向，可以获得图像的旋转不变性。通过图像梯度的方法计算每个关键点的梯度，幅值和幅角计算公式为

$$m(x, y) = \sqrt{[L(x+1, y) - L(x-1, y)]^2 + [L(x, y+1) - L(x, y-1)]^2} \quad (2-14)$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (2-15)$$

#### 2.2.2.3 特征点描述

选取以特征点为中心的  $8 \times 8$  的一个窗口，小方格表示像素点，箭头长短和方向表示梯度模值的大小和方向，把窗口的方向统计为  $2 \times 2$  个子区域的幅值和幅角，然后每个子窗口取 8 个方向，用 8 个直方图表示平均方向。

Lowe 原文中，用  $4 \times 4$  个子区域(8 个方向)，得到一个  $16 \times 8 = 128$  维(16 个  $4 \times 4$  子区域)特征向量。由于提取特征点数量较多和维度比较高，增加了 SIFT 算法的计算消耗。

### 2.2.3 SURF 算子

加速鲁棒特征(Speeded Up Robust Features, SURF)算法是由 Herbert Bay 等人在 2006 年提出的一种新的图像局部特征提取的方法,因其良好特性,被广泛应用于图像特征提取等领域<sup>[23]</sup>。SURF 算法主要步骤包括以下几个部分。

#### 2.2.3.1 积分图像

一幅图像中任意一点  $(i, j)$  的积分值  $n(i, j)$  是图像中任意一点  $(i, j)$  该到图像左上角点的对角线块区灰度值的和,计算公式为

$$n(i, j) = \sum_{i' \leq i, j' \leq j} p(i', j') \quad (2-16)$$

式中  $(i', j')$ ——图像的任意点;

$p(i', j')$ —— $(i', j')$  点的灰度值。

$n(i, j)$  可由公式(2-17)和(2-18)计算得到

$$S(i, j) = S(i, j-1) + p(i, j) \quad (2-17)$$

$$n(i, j) = n(i-1, j) + S(i, j) \quad (2-18)$$

式中  $S(i, j)$ ——任意一列的积分, 并且  $S(i, -1) = 0$ ,  $n(-1, j) = 0$ 。

对图像的所有像素点进行扫描可求取积分图像, SURF 算法利用积分图进行计算, 提升了特征提取的速度。

#### 2.2.3.2 Hessian 矩阵

Hessian 矩阵是 SURF 算法的核心, 其表示为

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(X, \sigma) & L_{xy}(X, \sigma) \\ L_{xy}(X, \sigma) & L_{yy}(X, \sigma) \end{bmatrix} \quad (2-19)$$

式中  $L_{xx}(X, \sigma)$  ——高斯二阶微分  $\frac{\partial^2 g(\sigma)}{\partial x^2}$  与在图像  $I$  中  $x$  处的卷积,

$$g(\sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}};$$

$\sigma$ ——尺度空间因子,  $H(x, \sigma)$  中其他元素与  $L_{xx}(X, \sigma)$  类似。



Hessian 矩阵行列式为

$$\det(H) = L_{xx}(X, \sigma) \times L_{yy}(X, \sigma) - L_{xy}^2(X, \sigma) \quad (2-20)$$

当其值为正数时，该点为极值点，反之，则不是。

### 2.2.3.3 SURF 尺度空间构建

与 SIFT 相反，SURF 算法是改变滤波器的大小，图像保持不变来进行尺度空间的构建。SURF 构建尺度空间时，同样是把空间划分为几组，每组是由对同一输入图像进行不断加大滤波器滤波后的一系列响应组成，是定值。

对于  $9 \times 9$  的模板，其响应长度  $l_0$  为 3，是盒子滤波器边长的  $1/3$ 。为保证模板只有一个中心，下一层的  $l_0$  为 5，至少保证一边增加一个像元，则滤波模板大小为  $15 \times 15$ ，由此依次建立的模板尺寸为： $9 \times 9$ 、 $15 \times 15$ 、 $21 \times 21$ 、 $27 \times 27$ ，每组进行相应操作，通过每一层增加尺寸得到下一层的，以此构建出图像的金字塔。如图 2-3 所示，横坐标为尺度，纵坐标为组别。

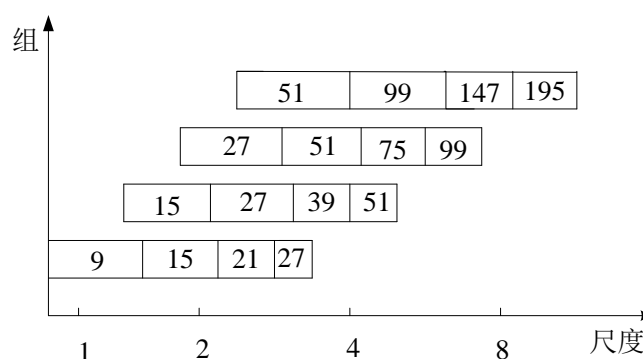


图 2-3 4 组不同滤波器尺寸表示

### 2.2.3.4 关键点检测

在实际计算时为了提高速度，采用方框滤波代替高斯核函数，方框滤波与图像卷积后的值记为  $D_{xx}$ 、 $D_{xy}$ 、 $D_{yy}$ 。Hessian 矩阵的行列式近似计算为

$$\det(H_{approx}) = D_{xx}D_{yy} - (\omega D_{xy})^2 \quad (2-21)$$

式中  $\omega$ ——常数，取 0.9。

由行列式值的正负确定是否为极值点。为了确定关键点，采用  $3 \times 3 \times 3$  的立体邻域进行非最大化抑制，把像素点当前所在层和上、下邻接层的 3 层总数共 26 个点计算的值进行对比，选出极大值的点或者极小值的点作为关键点，同时作为当前尺度的候选点。

### 2.2.3.5 方向的分配

以特征点为圆心， $6s$  为半径的圆形区域内计算 **Harr** 小波响应， $s$  是特征点所在的尺度。求主方向，选定  $\frac{3}{\pi}$  度的扇形区域作为窗口，以  $0.2$  弧度为单位转动此区域，记下滑动过程中窗口内图像的 **Haar** 小波在  $x$  和  $y$  方向的响应值  $dx$  和  $dy$ ，然后求和，可得特征矢量  $(m_w, \theta_w)$  为

$$m_w = \sum_w dx + \sum_w dy \quad (2-22)$$

$$\theta_w = \arctan \left( \frac{\sum_w dx}{\sum_w dy} \right) \quad (2-23)$$

式中  $m_w$ ——特征矢量模值；

$\theta_w$ ——特征矢量方向角度。

响应累加值最大的定为主方向，计算公式为

$$\theta = \theta_w | \max \{ m_w \} \quad (2-24)$$

当有  $m_w$  的峰值相当于主峰值 80% 能量时，则将对应的方向作为该特征点的辅方向，可能存在多个辅方向。

### 2.2.3.6 SURF 特征描述子

生成特征点后，计算每个特征点的 **Harr** 小波响应，在特征点周围构造一个长度  $20s$  的方框， $s$  表示特征点所在尺度。每个子块有 4 维  $(\sum dx, \sum dy, \sum |dx|, \sum |dy|)$  向量组成，共有 16 个  $(4 \times 4)$  小块，所以 **SURF** 特征描述符为 64 维  $(4 \times 4 \times 4)$  特征矢量。**SURF** 特征描述子不仅小波响应本身就有光照不变性，而且其生成的向量还有旋转、尺度和平移不变性，有助于场景图像匹配的准确性。

**SURF** 算子是 **SIFT** 算子的加强版，在匹配结果差别不大的情况下，**SIFT** 算法复杂度高，而且 **SURF** 算法的运算比 **SIFT** 算法的快很多，效率高。

## 2.3 本章小结

本章节阐述了 **SLAM** 的基础理论，并且分析了闭环检测中场景图像特征提取及匹配一些方法的优缺点，下一章节针对 **SURF** 算子中主方向计算不够准确的问题做出优化、改进。

## 第3章 基于多特征分层次的图像匹配算法研究

### 3.1 引言

在移动机器人视觉同时定位与地图创建中,在其行走的同时对周围环境进行图像采集,通过与系统图像库比对,从而完成相应定位和创建地图等任务。所以,图像处理在 VSLAM 过程十分关键<sup>[24]</sup>。

针对图像处理速度慢和特征点匹配准确性差的问题,提出基于颜色特征和改进 SURF 算法融合的图像匹配算法。首先,采用颜色特征对图像序列进行粗匹配操作,选取与测试图像最相近的 5 幅作为待匹配对象;其次,改进 SURF 算法,利用 Krawtchouk 矩描述 Hessian 矩阵获取的关键点,计算出梯度方向和幅值,得到新的特征向量,对待匹配对象提取改进 SURF 特征,再与测试图像进行精确匹配,得到最佳匹配图像。

### 3.2 改进 SURF 算法

#### 3.2.1 Krawtchouk 矩不变量的构造

在 SURF 算法中,采用 Hessian 矩阵获取图像局部最值是十分稳定的<sup>[25]</sup>,但在特征点描述时,求取主方向过度依赖局部区域像素的梯度方向,使得主方向准确性降低,易产生误匹配点对,即使不大的偏差角度也会造成特征匹配的误差放大使得匹配不成功<sup>[26-28]</sup>。为解决该问题并提高 SURF 算法精度,本文引入 Krawtchouk 矩不变量,Krawtchouk 矩是 Yap 在 2003 年提出的一种新型正交矩,计算简单且描述性强,具有在图像任何的感兴趣区域提取局部特征的优势<sup>[29]</sup>。利用 Krawtchouk 矩不变量作为图像的特征向量具有良好的几何不变性,确保了利用 Krawtchouk 矩不变量进行特征描述的鲁棒性<sup>[30]</sup>。因此,在 SURF 算法基础上引入 Krawtchouk 矩不变量对特征点进行描述<sup>[31,32]</sup>。

首先,利用 Hessian 矩阵确定特征点,然后通过计算特征点相邻区域 Krawtchouk 矩不变量来描述特征点,通过 Krawtchouk 矩分别求取特征点在 X 和 Y 方向上的响应,确定主方向。

$n$  阶离散的 Krawtchouk 多项式定义为<sup>[33]</sup>

$$\mathbf{K}_n(x, p, N-1) = \sum_{i=0}^{N-1} a_{i,n,p} x^i = {}_2F_1\left(-n, -x; -(N-1); \frac{1}{p}\right)$$

$$x, n = 0, 1, \dots, N-1, N > 1, p \in (0, 1) \quad (3-1)$$

其中,  ${}_2F_1$  表示 hypergeometric 函数, 其定义如下

$${}_2F_1(a, b; c; k) = \sum_{i=0}^{\infty} \frac{(a)_i (b)_i}{(c)_i} \cdot \frac{k^i}{i!} \quad (3-2)$$

其中,  $(a)_i$  表示 Pochhammer 函数, 其定义如下

$$(a)_i = a(a+1) \cdots (a+i-1) = \frac{\Gamma(a+i)}{\Gamma(a)} \quad (3-3)$$

$$\omega(x, p, N-1) = \binom{N-1}{x} p^x (1-p)^{N-1-x} \quad (3-4)$$

N 个多项式  $\mathbf{K}_n(x, p, N-1)$  经由函数公式(3-4)得到满足公式(3-5)的离散基函数, 并且满足正交性

$$\sum_{x=0}^{N-1} \omega(x, p, N-1) \cdot \mathbf{K}_m(x, p, N-1) \cdot \mathbf{K}_n(x, p, N-1) = \rho(n, p, N-1) \delta_{mn} \quad (3-5)$$

其中,  $m, n = 0, 1, \dots, N-1$ 。

$$\rho(n, p, N-1) = (-1)^n \cdot \left(\frac{1-p}{p}\right)^n \cdot \frac{n!}{(-N+1)_n} \quad (3-6)$$

归一化 Krawtchouk 多项式定义如下

$$\hat{\mathbf{K}}_n(x, p, N-1) = \mathbf{K}_n(x, p, N-1) \cdot \sqrt{\frac{\omega(x, p, N-1)}{\rho(n, p, N-1)}} \quad (3-7)$$

正交条件转换为

$$\sum_{x=0}^{N-1} \hat{\mathbf{K}}_m(x, p, N-1) \cdot \hat{\mathbf{K}}_n(x, p, N-1) = \delta_{mn} \quad (3-8)$$

若图像的像素大小为  $N \times M$  个像素, 则图像  $f(x, y)$  的  $(m+n)$  阶 Krawtchouk 矩表示为

$$Q_{mn} = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{K}_m(x, p_1, M-1) \cdot \hat{K}_n(x, p_2, N-1) f(x, y) \quad (3-9)$$

用于图像重构的逆变换定义为

$$f(x, y) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} \hat{K}_m(x, p_1, M-1) \cdot \hat{K}_n(x, p_2, N-1) Q_{mn} \quad (3-10)$$

由公式(3-9)可知，Krawtchouk 矩实际上是图像  $f(x, y)$  和  $\hat{K}_m(x, p, N)$ ， $\hat{K}_n(x, p, N)$  的内积，通过参数  $p_1$  和  $p_2$  的变化，图像低阶 Krawtchouk 矩可用于提取图像的局部特征信息。

对于  $f(x, y)$ ， $(m+n)$  阶几何矩表示如下

$$M_{mn} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^m y^n f(x, y) dx dy, m, n = 0, 1, 2, \dots \quad (3-11)$$

相应的中心矩表示为

$$\mu_{mn} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - x_c)^m (y - y_c)^n f(x, y) dx dy$$

$$x_c = \frac{M_{10}}{M_{00}}, y_c = \frac{M_{01}}{M_{00}} \quad (3-12)$$

构造 Krawtchouk 矩不变量，把图像  $f(x, y)$  平移至质心  $(x_c, y_c)$  的位置，主轴旋

转到水平位置，转化公式是： $\begin{cases} x' = x \cos \theta - y \sin \theta \\ y' = x \sin \theta + y \cos \theta \end{cases}$ 。图像旋转  $\theta$  角：

$$\theta = \frac{1}{2} \arctan \frac{2\mu_{11}}{\mu_{20} - \mu_{02}}。最后，把图像缩放  $\alpha$  倍尺度： $\alpha = \left( \frac{N}{2M_{00}} \right)^{\frac{1}{2}}$ 。得到新图像  $g(x, y)$$$

与原图像  $f(x, y)$  的关系

$$g\left(x + \frac{N}{2}, y + \frac{N}{2}\right) = f\left(x' + \frac{x \cos \theta + y \sin \theta}{\alpha}, y' + \frac{y \cos \theta - x \sin \theta}{\alpha}\right) \quad (3-13)$$

经上述变换，得到归一化图像，此时得到的矩保持平移、旋转、尺度不变特性。

Krawtchouk 矩不变量定义如下

$$\tilde{Q}_{mn} = [\rho(m) \rho(n)]^{-\frac{1}{2}} \sum_{i=0}^m \sum_{j=0}^n a_{i,m,p_1} a_{j,n,p_2} \tilde{v}_{ij} \quad (3-14)$$

$$\tilde{v}_{ij} = \sum_{p=0}^m \sum_{q=0}^n \binom{m}{p} \binom{n}{q} \left( \frac{N'}{2} \right)^{(p+q)/2+1} \left( \frac{N'}{2} \right)^{n+m-p-q} v_{pq} \quad (3-15)$$

式中  $v_{pq}$ ——标准的几何矩不变量。

由公式(3-14)可得出任意阶的 Krawtchouk 矩不变量<sup>[34]</sup>。

### 3.2.2 特征点描述

设  $P_i(x, y, \theta, \sigma)$  为 SURF 算法特征点检测获得的某一特征点,  $(x, y)$  为特征点  $P_i$  在图像中位置坐标信息,  $\theta$  为点  $P_i$  的方向信息,  $\sigma$  为尺度大小。如图 3-1 所示, 通过 Krawtchouk 不变矩分别求取特征点在 X 和 Y 方向上的响应, 则特征点主方向的幅值和角度表示如下

$$m(x, y) = \sqrt{(\tilde{Q}_{p,0}(x, y))^2 + (\tilde{Q}_{0,q}(x, y))^2} \quad (3-16)$$

$$\tilde{\mu}(x, y) = \arctan\left(\frac{\tilde{Q}_{p,0}}{\tilde{Q}_{0,q}}\right) \quad (3-17)$$

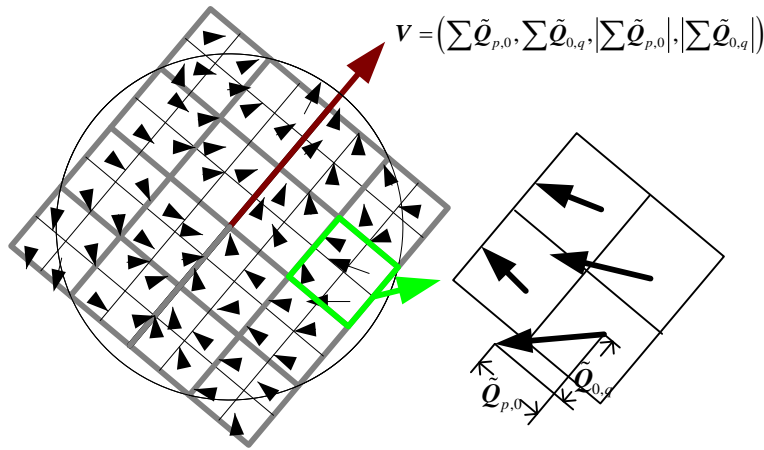


图 3-1 特征描述子表示

以特征点为中心, 沿主方向建立一个  $20\sigma \times 20\sigma$  大小的正方形块区, 划分为  $4 \times 4$  个子块, 每个子块有  $5 \times 5$  个像元, 共进行 25 次采样计算, 采样点在 X 和 Y 方向上的 Krawtchouk 不变矩分别记作  $\tilde{Q}_{p,0}$  和  $\tilde{Q}_{0,q}$ , 并计算每个子区域的  $\sum \tilde{Q}_{p,0}$  和  $\sum \tilde{Q}_{0,q}$ ,  $|\sum \tilde{Q}_{p,0}|$  和  $|\sum \tilde{Q}_{0,q}|$ , 定义关键点特征向量为

$$\mathbf{v} = \left( \sum \tilde{\mathbf{q}}_{p,0}, \sum \tilde{\mathbf{q}}_{0,q}, \left| \sum \tilde{\mathbf{q}}_{p,0} \right|, \left| \sum \tilde{\mathbf{q}}_{0,q} \right| \right) \quad (3-18)$$

### 3.2.3 特征向量匹配

利用改进算法获得模板图像与待匹配图像的特征点后，进行特征匹配。为提高算法匹配正确度，匹配点对的寻找以欧氏距离为准则，以阈值为条件，对于不满足阈值要求的错误匹配点予以去除。以欧氏距离进行判断，计算公式为  $\mathbf{D} = \|\mathbf{f}_m - \mathbf{f}_n\|$ ， $\mathbf{f}_m$ ， $\mathbf{f}_n$  分别为两者的特征向量，以最近邻距离与次近邻距离的比值为准则，将两个邻接点的距离比率与事先假定的阈值进行比较，满足要求则匹配成功，否则失败。即： $r = d_1/d_2$ ，如果  $r > \varepsilon$ ，则匹配成功，反之失败。

## 3.3 多特征分层次的图像匹配算法

### 3.3.1 颜色特征提取

由于 RGB 颜色空间与人类视觉感知差别大，而 HSV 颜色空间直接对应于人类眼睛感知的三个要素，分别用色度(Hue)、饱和度(Saturation)和亮度(Value)来表示。HSV 颜色空间平均，空间各轴在视觉上保持彼此独立的关系，人的肉眼感觉到的色差与颜色值之间的欧氏距离呈现正比例函数关系。因此，用颜色信息表达图像的内容时，通常会采用 HSV 颜色空间<sup>[35]</sup>。彩色图像从 RGB 空间到 HSV 空间的转换关系如下。

设(r, g, b)分别表示一个颜色中的红、绿和蓝坐标，它们的取值范围为(0~1)之间的实数，设 max 定义为这三个值当中的最大值，min 相应表示为三个值当中的最小值。

$$h = \begin{cases} \text{undefined, if } \max = \min \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, \text{ if } \max = r \ \& \ g \geq b \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, \text{ if } \max = r \ \& \ g < b \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, \text{ if } \max = g \\ 60^\circ \times \frac{g-b}{\max-\min} + 0^\circ, \text{ if } \max = b \end{cases} \quad (3-19)$$

$$s = \begin{cases} 0, & \text{if } \max = 0 \\ \frac{\max - \min}{\max} = 1 - \frac{\min}{\max}, & \text{otherwise} \end{cases} \quad (3-20)$$

$$v = \max \quad (3-21)$$

式中  $h$  —— 图像的色度;

$s$  —— 饱和度;

$v$  —— 亮度。

颜色矩是 Stricker 等人提出的一种十分有效的颜色特征表示方法, 把图像中的颜色分布通过矩表达<sup>[36]</sup>。其中, 一阶矩为颜色平均值、二阶矩为颜色方差、三阶矩为颜色的偏斜度。图像的颜色信息分布大部分存在于低阶矩当中, 所以只用颜色的前三阶矩则可以完成图像的颜色信息分布情况的表达, 还可以简化计算。利用颜色矩的表示方法具有无需量化描述图像的颜色特征并且颜色向量维数低的优势, 3 个低阶颜色矩的数学表示为

$$u_i = \frac{1}{N} \sum_{j=1}^N p_{i,j} \quad (3-22)$$

$$\sigma_i = \left( \frac{1}{N} \sum_{j=1}^N (p_{i,j} - u_i)^2 \right)^{\frac{1}{2}} \quad (3-23)$$

$$s_i = \left( \frac{1}{N} \sum_{j=1}^N (p_{i,j} - u_i)^3 \right)^{\frac{1}{3}} \quad (3-24)$$

式中  $u_i$ ,  $\sigma_i$ ,  $s_i$  —— 依次为颜色分量  $i$  的均值, 方差和斜度;

$N$  —— 图像的总像素数;

$p_{i,j}$  ——  $j$  像素值出现在  $i$  通道的概率。

HSV 的 3 阶颜色矩构成一个 9 维向量, 则图像的颜色特征向量为

$$\mathbf{F}_{color} = [u_H, \sigma_H, s_H, u_S, \sigma_S, s_S, u_I, \sigma_I, s_I] \quad (3-25)$$

### 3.3.2 颜色特征匹配

由移动机器人获取图像, 输出相似图像的数目根据需要设定。本文实验中, 设为 5。两幅图像的颜色矩向量生成后, 本文采用欧式距离(Euclidean distance)算法计



算颜色向量之间的相似度。其数学定义为

$$D(I, J) = \left\{ \sum_{k=0}^8 [C_I(k) - C_J(k)]^2 \right\}^{\frac{1}{2}} \quad (3-26)$$

式中  $C_I(k)$  ——图像库中图像 I 的颜色矩向量;

$C_J(k)$  ——当前图像 J 的颜色矩向量。

### 3.3.3 多特征分层次匹配算法

对移动机器人获取的图像采用全局特征与局部特征融合的图像匹配方法进行匹配<sup>[37,38]</sup>。首先提取颜色矩特征对图像库中的图像进行粗匹配,选取最相似的5幅作为待匹配对象存入临时图像库,其余图像则不再进行精确匹配,节省移动机器人 VSLAM 系统的图像处理时间,然后,对临时图像库中的待匹配图像提取改进的 SURF 特征进行精确匹配,得到最佳匹配则匹配结束。多特征匹配流程如图 3-2 所示。

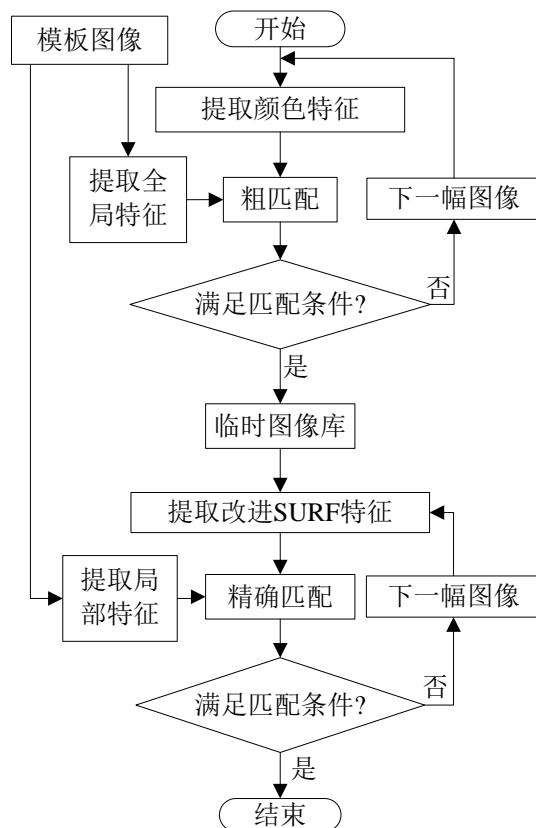


图 3-2 图像匹配流程图

3.4 实验结果与分析

为验证本文算法的有效性，在 Intel(R) Core(TM) i3-2120 CPU 3.30GHz、4GB 内存、matlab R2013a 环境中进行仿真。实验数据为 pioneer 3<sup>TM</sup> 获取的课题室环境的图像序列。

3.4.1 图像特征匹配实验

首先，对移动机器人 pioneer 3<sup>TM</sup> 获取的 10 幅实验室环境图像进行颜色矩的提取及与模板图像进行匹配，按最相似的顺序展示结果，如图 3-3 所示，实验用时如表 3-1 所示。



图 3-3 颜色矩的图像匹配

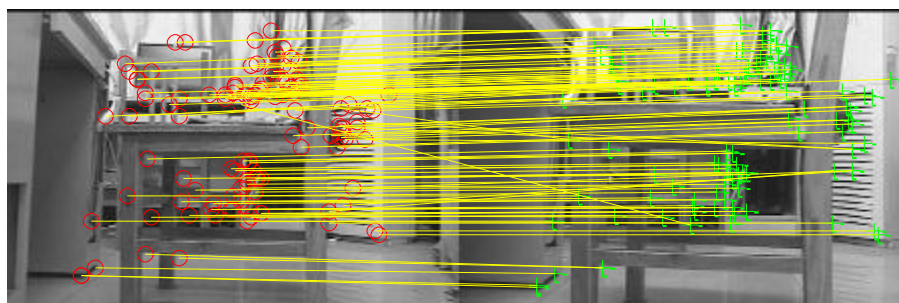
表 3-1 颜色特征提取与匹配耗时

	图1	图2	图3	图4	图5	图6	图7	图8	图9	图10
时间(s)	0.013	0.009	0.009	0.013	0.010	0.008	0.007	0.009	0.008	0.008

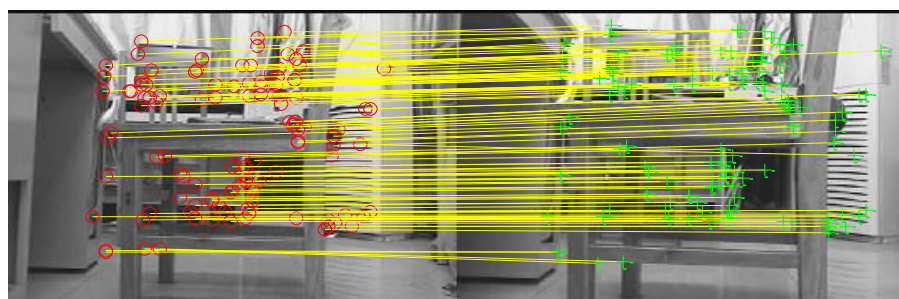
由实验结果可知，模板图像提取时间为 0.247 秒，实验的总耗时为 0.341 秒，处理每幅图像平均用时 0.031 秒。利用颜色矩特征进行粗匹配以少的时耗为后续的精确定匹配缩小了目标图像的搜索范围，排除了与模板图像不相似的图像，减少了后续进行精确匹配的工作量，提高了算法的实时性，同时验证了采用颜色特征进行粗匹配的高效性。

为了验证改进 SURF 算法的性能，从算法运行时间和特征点匹配正确率两个方面着手对改进 SURF 算法与 SURF 算法进行对比。匹配结果如图 3-4 至图 3-8 所示，依次为尺度、视角、模糊及尺度和视角变化的图像匹配实验。图 3-4 中的匹配图像是移动机器人位置走过 0.5m 的拍摄图像，图 3-5 中的匹配图像是移动机器人摄像头

转动  $20^\circ$  的拍摄图像，图 3-6 中的匹配图像是图像发生模糊变化的拍摄图像。图 3-7 中的匹配图像是移动机器人位置前后移动  $0.3\text{m}$  且摄像头转动  $20^\circ$  的拍摄图像。图 3-8 中的匹配图像是移动机器人发生光照变化的拍摄图像。

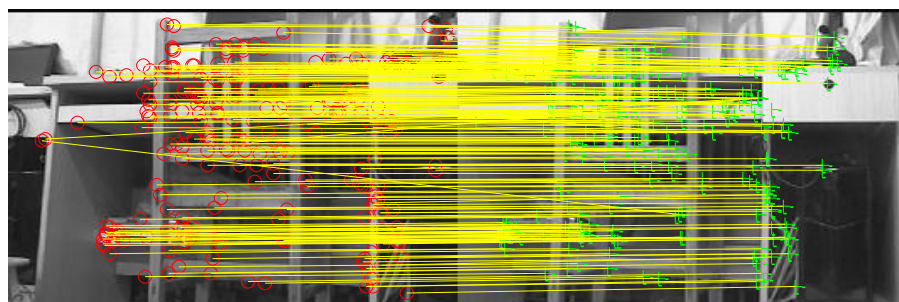


a) SURF 算法

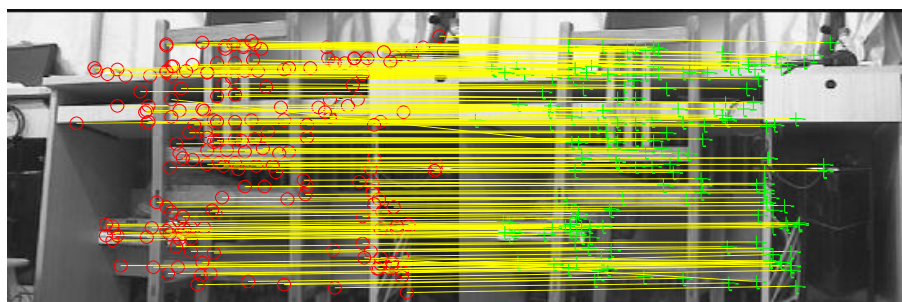


b) ISURF 算法

图 3-4 尺度变化

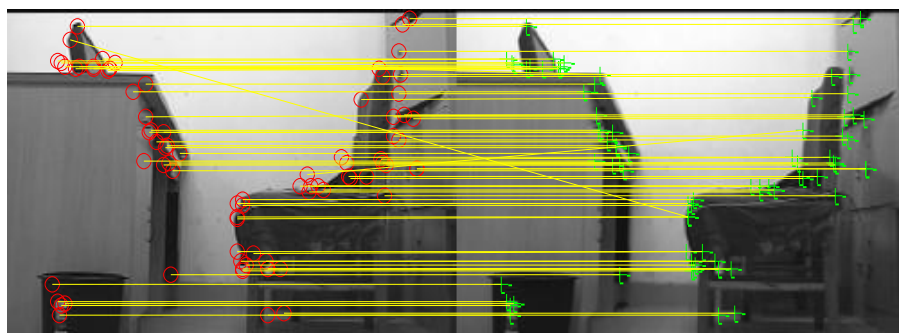


a) SURF 算法

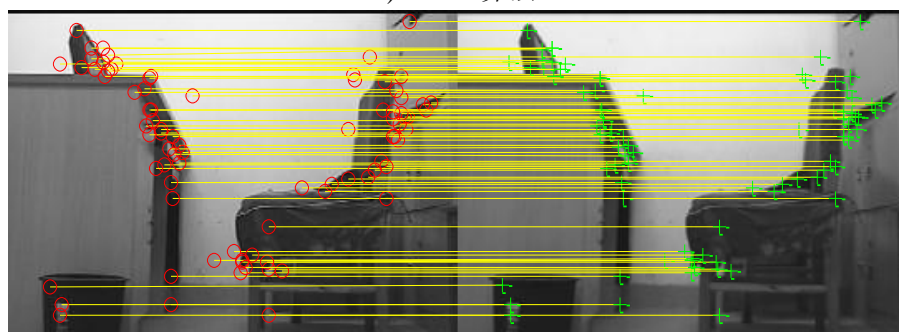


b) ISURF 算法

图 3-5 视角变化

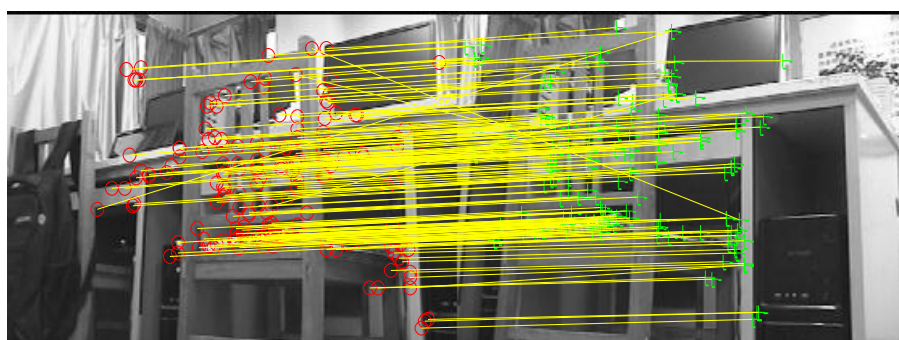


a) SURF 算法

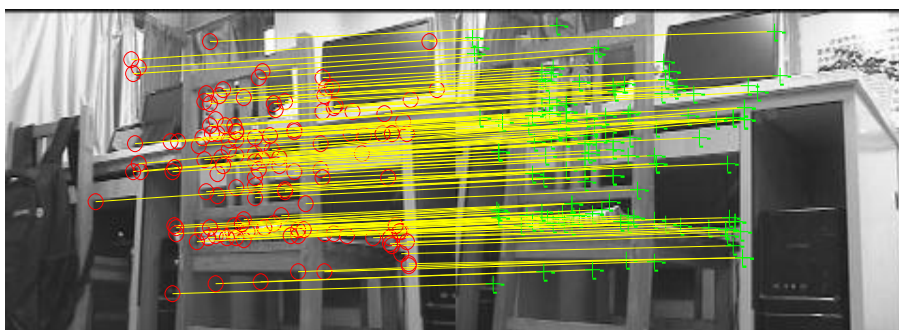


b) ISURF 算法

图 3-6 模糊变化



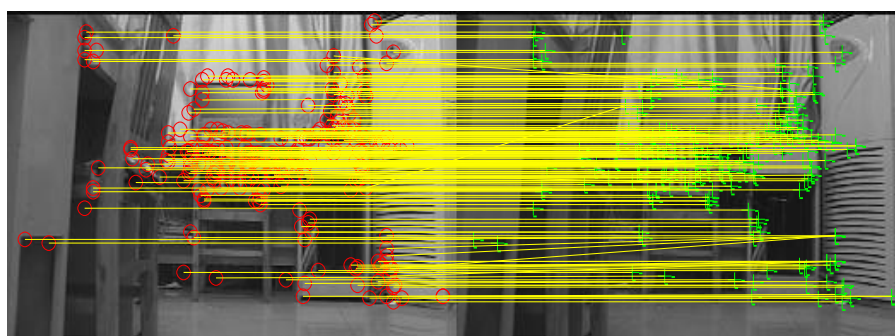
a) SURF 算法



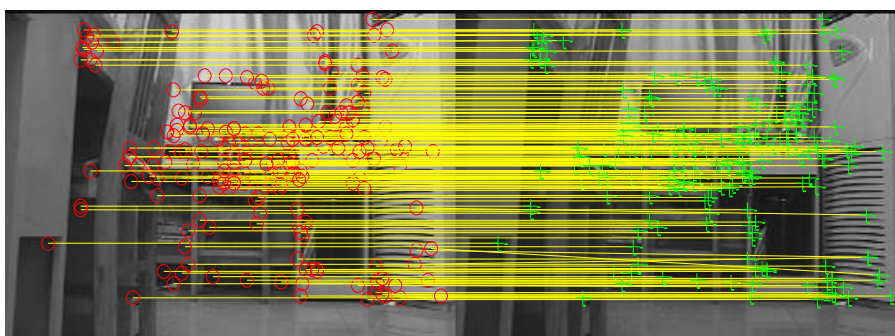
b) ISURF 算法

图 3-7 尺度和视角变化





a) SURF 算法



b) ISURF 算法

图 3-8 光照变化

### 3.4.2 改进 SURF 的性能比较

图像匹配的实验数据如表3-2所示，ISURF 代表改进 SURF 的特征匹配算法。

表 3-2 ISURF 特征匹配算法与改进 SURF、SURF 性能比较

性能		尺度	视角	模糊	尺度 视角	光照
总时间(s)	SURF	1.016	0.981	0.903	0.997	1.149
	改进 SURF	0.872	0.901	0.718	0.863	0.982
	ISURF	0.922	0.887	0.794	0.918	1.057
特征点数	SURF	460	397	110	394	274
	改进 SURF	324	292	112	250	189
	ISURF	460	397	110	394	274
匹配对数	SURF	190	217	92	149	223
	改进 SURF	42	112	88	40	140
	ISURF	163	168	84	106	209
误匹配率(%)	SURF	10.53	11.52	9.78	13.42	12.56
	改进 SURF	4.76	2.05	4.76	7.5	3.57
	ISURF	4.91	3.57	2.38	5.66	4.78

本文算法与 SURF 算法和文献[39]改进 SURF 算法的对比实验，由主观视觉的匹配结果和实验数据可以发现：Krawtchouk 矩具有较强的特征描述能力，在平移、旋转、尺度、光照和模糊图像的变换下均显示了较好的匹配结果。改进算法相对于 SURF 算法耗时减少 10% 左右，匹配点对数基本无差异，但比 SURF 算法的匹配精度提高 8% 左右；在与文献[39]的改进算法程序耗时和误匹配率差异不大的情况下，本文算法提取的特征点数和匹配对数均高于此算法，有利于视觉单词本的构建。

本文算法的优点还体现在处理大规模样本数据时，如表 3-3 所示，分别用 SIFT、SURF 算法和本文多特征分层次算法对 pioneer 3<sup>TM</sup> 获取实验室环境的 15 幅连续图像进行匹配，得到最优匹配结果，统计 5 次不同实验样本的程序累计运行时间。由实验结果可知，本文算法将图像处理的效率提高了近 50%，大大减少了程序运行的时间总和。

表 3-3 15 幅图像程序运行累计时间比较

算法	时间 1(s)	时间 2(s)	时间 3(s)	时间 4(s)	时间 5(s)
SIFT	19.35	20.85	22.14	21.65	20.07
SURF	16.56	18.45	19.23	17.9	17.18
C-ISURF	8.55	9.16	10.08	9.74	8.79

为了验证改进算法的优越性，随机选取 pioneer 3<sup>TM</sup> 获取的 10 组发生几何变化的实验室环境图像，分别利用 SIFT、SURF、和本文算法进行图像匹配。由图 3-9 的统计实验结果可知，改进算法的平均匹配正确率比其它方法提高近 6%。

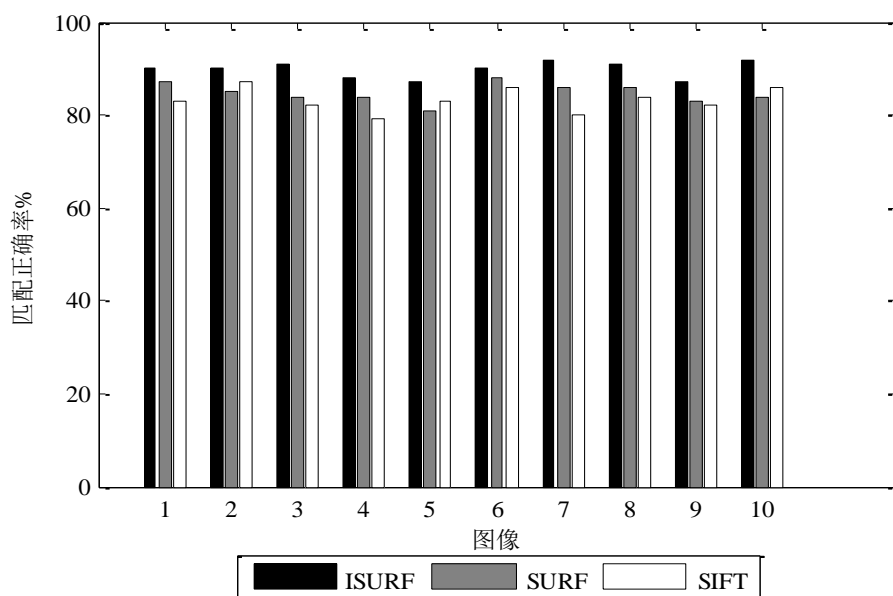


图 3-9 匹配正确率

由实验结果可知,改进算法的优点在于:首先进行粗匹配,过滤掉与模板图像不相似的图像,不再进行精确匹配,提高了图像处理的实时性。匹配准确度的提高在于 Krawtchouk 矩不变量具有更好的描述能力,使得 SURF 特征更具显著性。

### 3.5 本章小结

文章针对移动机器人 VSLAM 中图像匹配实时性和鲁棒性问题,考虑到 Krawtchouk 矩作为一种性能优越的离散矩,能更好的满足图像特征提取的要求,利用基于颜色矩和改进 SURF 算法组合的移动机器人多特征分层次图像匹配算法。结合颜色特征和改进 SURF 特征,使两者得到优势互补,图像描述更加丰富,再通过分层次的图像匹配方法提高了移动机器人 VSLAM 中图像处理的实时性和鲁棒性。实验结果表明该方法有效可行。

## 第 4 章 基于改进 BoVW 算法的研究

### 4.1 引言

机器人在行走时,采集图像和对图像的场景建模至关重要,是实现 SLAM 中闭环检测的基础环节。文献[40,41]等人直接对提取的图像全局或局部特征进行直方图表示,然后进行相似性度量,由于场景图像的复杂和提取描述符的不同,使得相似度的计算不可靠。利用 BoVW(Bag of visual words, BoVW)模型进行场景外观的建模可以有效改善其不足,对提取的描述符进行筛选、整合,从而建立统一和规范的视觉单词本,各个图像都可以用这些具有代表性的“单词们”来表示,如图 4-1 所示。

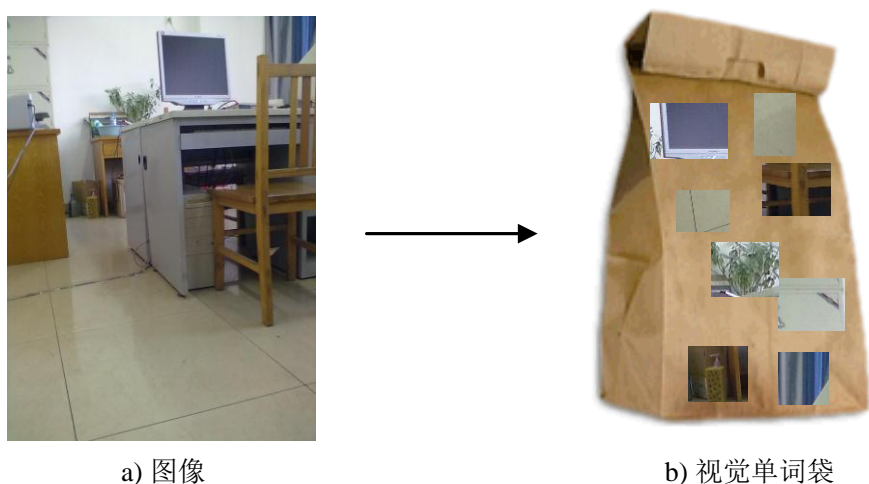


图 4-1 视觉单词示意图

本章首先对 BoVW 模型进行描述,然后基于此模型进行场景建模。根据现有算法的不足与缺陷,提出一种改进的 BoVW 方法。首先,对场景图像进行基于指导滤波(guided image filter)的去雾处理,改善由于机器人摄取的图像质量低或者是外界环境所引起的图像失真,使得图像的边缘等重要信息更突出;利用基于 Krautchouk 矩的 SURF 算子进行场景图像的特征提取,然后,利用基于最大最小距离改进的粒子群(Particle Swarm Optimization, PSO)聚类算法构建视觉单词本,弥补原始 BoVW 模型中聚类算法的聚类结果不稳定和迭代次数过多的缺陷,而后将特征描述子们训练成“单词”,然后通过基于信息熵的空间金字塔的直方图表示场景图像,使得图



像相似度计算准确简单。

## 4.2 BoVW 模型描述

BOW(Bag Of Words, BOW)模型因其简单有效性,一开始用于文本的识别和分类领域,而后引入视觉图像领域,形成了 BoVW 模型。利用 BoVW 模型进行场景图像的建模<sup>[42,43]</sup>。BoVW 模型的基本步骤包括:图像的特征提取,是 BoVW 模型的基本步骤,高质量的图像摄取是基本保证,从场景图像中有效地提取稳定且丰富的特征点,使其具有平移、尺度、旋转不变性;构建视觉单词本,是 BoVW 模型的核心问题,把提取的图像特征点整合成视觉“单词”,选取的视觉“单词”具有良好的表征能力,形成视觉单词本;将图像特征映射到视觉单词本,然后计算视觉单词本之间距离来进行图像相似度度量,以此完成图像分类、图像检索等。如图 4-2 所示,利用 BoVW 模型进行图像分类的主要步骤:图像的特征提取,创建视觉单词本,训练分类器。

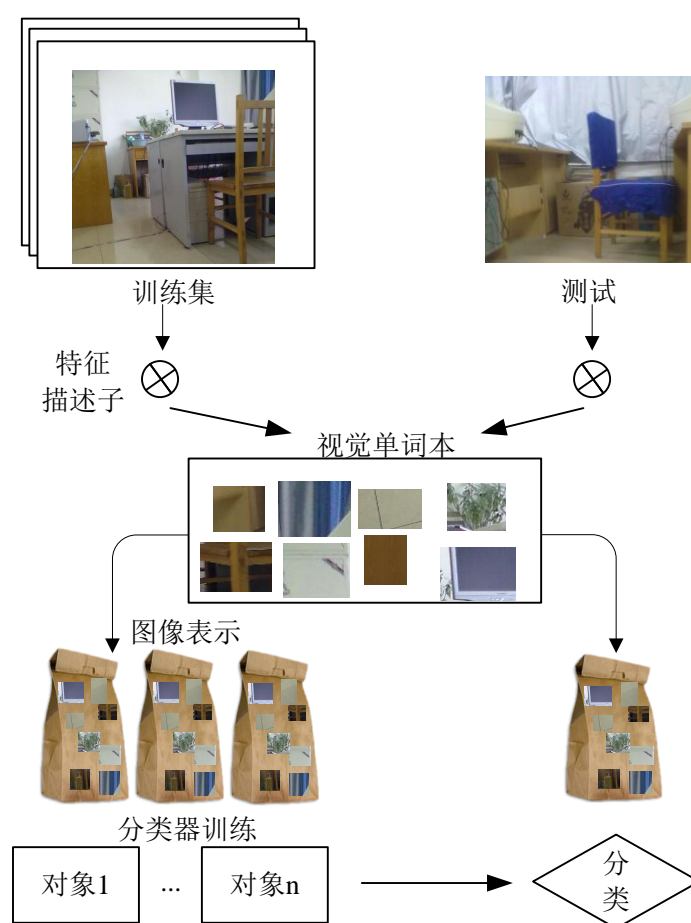


图 4-2 基于 BoVW 模型的图像分类

BoVW 模型是利用特征点来表示一幅图像，所以对图像的质量和特征点提取与描述显得尤为重要，即使在非常复杂、恶劣的条件下，也要求图像清晰无失真，还要求提取的特征点具有高效稳定性，能够抵抗外界因素和噪声、尺度、光照等不利因素带来的干扰，同时，要求特征描述子具有高度可辨识性。如果不能满足以上条件，就会导致错误出现。针对以上不利因素，本章首先对获取的场景图像进行指导滤波去雾，从而得到清晰的图像，然后对清晰图像利用基于 Krautchouk 矩的 SURF 算子进行场景图像的特征提取与描述，对提取的特征向量转换成视觉单词，用最大最小距离改进的 PSO 聚类算法构建视觉词典，最后通过基于信息熵的空间金字塔直方图表示场景图像，计算相似度，从而完成对场景图像进行分类的任务。

### 4.3 改进的 BoVW 算法

#### 4.3.1 基于指导滤波的 BoVW 模型去雾预处理

机器人在行走过程中，由于其所处的环境或自身抖动等原因，可能使得获取的图像存在模糊、失真等问题，由此会导致提取的图像特征不够准确，也会引起此后进行构建词典树和相似性度量一系列的误差等问题。所以，首先对机器人获取的场景图像进行指导滤波去模糊操作，以此得到清晰的图像。

$p$  是原始的输入， $I$  是引导图， $q$  是滤波后的所得图像，则以  $k$  为中心窗口  $\omega_k$  的线性关系为

$$q_i = a_k I_i + b_k, \forall i \in \omega_k \quad (4-1)$$

式中  $\omega_k$  ——半径为  $r$  的正方形窗口；

$a_k$  和  $b_k$  ——窗口  $\omega_k$  的一系列线性因子。

因为线性关系，所以窗口中  $I$  有边缘， $q$  才会相应的产生边缘，引导滤波就是在  $I$  的指引下得出线性因子  $(a_k, b_k)$  的最优解，求得输入图像  $p$  与输出图像  $q$  的最小差值<sup>[44]</sup>。通过线性回归方法得到的代价函数表示为

$$E(a_k, b_k) = \sum_{i \in \omega_k} \left[ (a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2 \right] \quad (4-2)$$

式中  $\varepsilon$  ——避免  $a_k$  取值过大设置的调整系数。

利用公式(4-2)得出 $(a_k, b_k)$ 的窗口代价最小的窗口系数为

$$a_k = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\delta_k^2 + \varepsilon} \quad (4-3)$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (4-4)$$

式中  $|\omega|$ ——窗口 $\omega_k$ 内的像素数;

$\delta_k$ ——引导图像 $I$ 在窗口 $\omega_k$ 的方差;

$\mu_k$ ——引导图像 $I$ 在窗口 $\omega_k$ 的均值;

$\bar{p}_k$ ——输入图像 $P$ 的均值。

因为像素点 $i$ 在不同窗口的 $(a_k, b_k)$ 的值也不相同,所以需要计算出以像素点 $i$ 为中心的窗口内 $(a_k, b_k)$ 求取平均值,然后得到指导滤波的最终表达式为

$$q_i = \frac{1}{|\omega|} \sum_{k \in \omega_k} (a_k I_i + b_k) = \bar{a}_i I_i + \bar{b}_i \quad (4-5)$$

式中  $\bar{a}_i$ ——以像素点 $i$ 为中央的窗口内的平均系数值;

$\bar{b}_i$ ——以像素点 $i$ 为中央的窗口内的平均系数值<sup>[45]</sup>。

图 4-3 中 a)是机器人拍摄的课题室的场景图像, b)是对场景图像进行指导滤波处理后的图像,由实验结果清晰可见,处理后的图像明显细节显示更加突出,噪声减小。



a) 机器人获取图像



b) 处理后的图像

图 4-3 图像去噪前后对比

### 4.3.2 基于改进的聚类算法构建视觉单词树

获取到清晰图像后,利用基于 Krautchouk 矩的 SURF 算子进行特征提取和描述,得到稳定的特征描述向量,然后构建视觉单词本。一般 BoVW 模型都是通过聚类的方法来构建视觉单词本,聚类结果的好坏和聚类中心的选取直接影响着视觉单词本的构建,计算图像相似度的直方图表示也是直接受聚类结果影响的,所以对聚类算法的选择显得十分关键<sup>[46,47]</sup>。

原始 BoVW 模型中采用了简单易行的 k 均值聚类算法,虽然得到了普遍应用,但也存在不足之处。例如:开始选取中心点对聚类结果影响大,如果选中孤立点,就会导致结果特别差;而且,由于 k 均值算法在聚类执行过程中由于初始点的影响,不同的中心点导致的聚类结果不尽相同,使得聚类结果不稳定,随机性大。针对以上问题,在构建视觉单词本时,采用 PSO 算法<sup>[48]</sup>结合距离函数的聚类算法,获得最优的聚类中心点,使得聚类中心更稳定和更具代表性。

改进算法的基本描述如下:设定提取的图像特征点为样本  $X = \{x_1, x_2, \dots, x_n\}$ , 选定初始聚类中心。

步骤一,初始化粒子群,设定参数:权重因子  $\omega$ , 加速系数  $c_1, c_2$ , 随机数  $r_1, r_2$ ;

步骤二,计算粒子的适应度值,适应度函数表示为

$$fitness = \min \sum_{i=1}^n \|x_i - x_j\| \quad (4-6)$$

更新粒子群中粒子的位置和速度的依据如公式(4-7)和公式(4-8),选取适应度值最小的样本点作为第一类聚类中心  $z_1 = x_i$ 。

$$v_i(t+1) = \omega v_i(t) + c_1 r_1 (p_i(t) - x_i(t)) + c_2 r_2 (p_s(t) - x_i(t)) \quad (4-7)$$

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (4-8)$$

式中  $t$ ——当前迭代次数;

$x_i(t)$ ——粒子当前位置;

$v_i(t)$ ——粒子速度向量;

$p_i(t)$ ——单个粒子  $i$  搜索到的最好位置,即个体最优;

$p_s(t)$ ——粒子群中全部粒子的当前最优位置,即全局最优。

步骤三，计算选取距离聚类中心  $z_2$  最远的点作为第二类的中心， $z_2 = x_j$ ， $x_j = \text{MAX} \|x_j - z_1\|$ ， $j = 1, 2, \dots, n$ 。

步骤四，剩余对象分别计算到  $z_1$ ， $z_2$  的距离和，选择和值最大的对象作为第三类的中心， $z_3 = x_j$ ， $x_k = \text{MAX} \{ \text{sum}(\|x_k - z_1\| + \|x_k - z_2\| + \|x_k - z_3\|) \}$ 。

步骤五，重复操作，直至计算得出需求设置的第  $k$  个类的聚类中心， $z_k = x_p$ ， $x_p = \text{MAX} \{ \text{sum}(\|x_p - z_1\| + \|x_p - z_2\| + \|x_p - z_3\| + \dots + \|x_p - z_{k-1}\|) \}$ ， $p = 1, 2, \dots, n$ 。

步骤六，计算提取的所有特征点与  $k$  个聚类中心的距离，以最小准则分类。

步骤七，根据公式(4-9)计算目标函数  $J$ ，使得  $J$  取得最小值。

$$J(k, z) = \sum_{j=1}^k \sum_{x \in S_j} \|x - z_j\|^2 \quad (4-9)$$

式中  $S_j$ ——第  $j$  类中的集合；

$z_j$ ——第  $j$  类的中心。

步骤八，循环执行步骤六和步骤七，直至聚类中心不再发生变化则停止循环操作，即可得到视觉“单词”。

### 4.3.3 基于信息熵的图像相似性度量

BoVW 模型中，对图像提取的特征向量通过聚类后，生成了有代表性的视觉“单词”，形成了视觉单词本；然后，当提取的图像局部特征信息和视觉单词本中的某一个“单词”的相似性超过阈值，则把这个“单词”赋予此图像，统计满足条件的这些“单词”，图像的表达就用这些视觉“单词”的统计直方图来表示<sup>[49]</sup>，示意图如图 4-4 所示。

由于图像的空间位置信息不在原始 BoVW 模型的图像直方图表示的考虑范围，所以引入空间金字塔匹配核方法(spatial pyramid matching kernel method, SPM)对图像进行直方图的表示<sup>[50]</sup>，SPM 是用来计算不同特征数目的图像之间的相似度，进行多分辨率下的图像切割匹配。基本原理为：将一幅图像规范地分成几个子图区域，子图区域细化再分成几个子图块区，计算各个子图块区的“单词”直方图向量，整合形成空间式的直方图向量。

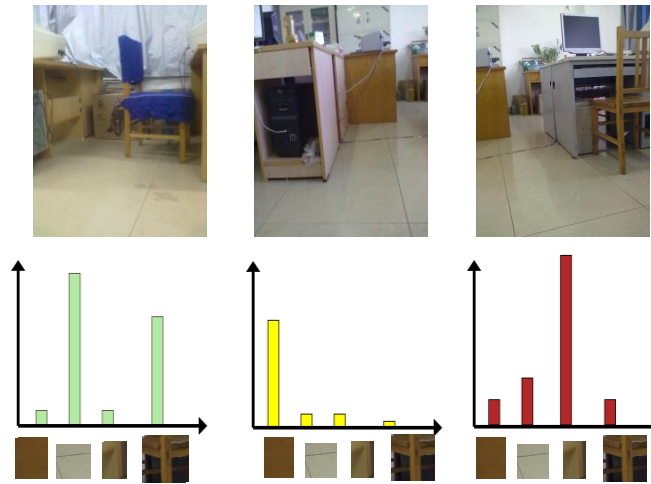


图 4-4 场景图像的直方图表示示意图

不同层次的场景图像，假设有  $\{0, 1, 2, \dots, N\}$  层，将  $l$  层的场景图像按照十字分割的规律划分为相同大小的  $4^l$  个子区域，统计视觉“单词”出现在  $4^l$  个各个子区域的数目，频率数=次数/“单词”的总数，统计子图块区的频率数，以此得到在  $l$  层的场景图像的直方图向量

$$\mathbf{H}_l^l = [\mathbf{H}_l^{l,1}, \mathbf{H}_l^{l,2}, \dots, \mathbf{H}_l^{l,k}, \dots, \mathbf{H}_l^{l,4^l}] \quad (4-10)$$

式中  $\mathbf{H}_l^{l,k}$ ——场景图像  $l$  的第  $l$  层的第  $k$  个子图区域的直方图向量，将所有的子图区域汇总，则可得到第  $l$  层的直方图向量表示；

$\mathbf{H}_l^l$ ——场景图像  $l$  的第  $l$  层的视觉“单词”直方图向量<sup>[51,52]</sup>。

对图像进行分割子图区域后，一些子图区域包含的特征点多，同样的，一些子图区域包含的特征点少或没有。因此采用基于信息熵的空间金字塔匹配核进行场景图像的相似性度量，在统计直方图之前先求取各个子图像的信息熵，以图  $l$  为例，假设共有  $m$  个状态，则图像  $l$  信息熵为

$$H(l) = \sum_{i=1}^m p_i \log \left( \frac{1}{p_i} \right) \quad (4-11)$$

式中  $p_i$ ——状态  $i$  的概率分布。

信息熵越大，包含的特征点数越多。计算出  $4^l$  个子图区域的信息熵，从大到小排序，选取熵值大于设定阈值的子图。然后，统计满足条件的子图的频数，构成新的直方图向量

$$\mathbf{H}_l^l = [\mathbf{H}_l^{l,1}, \mathbf{H}_l^{l,2}, \dots, \mathbf{H}_l^{l,k}, \dots, \mathbf{H}_l^{l,p}] \quad (4-12)$$

由本文构建的直方图向量筛选出了包含图像信息密集的区域，删去了图像信息量少的部分，从而减小了直方图向量的维度，避免了不必要的运算，提高了算法的运算效率。

计算场景图相间的相似度，以图像  $I_1$  和图像  $I_2$  为例，直方图交叉函数在第  $l$  层的定义为

$$\Gamma(\mathbf{H}_{I_1}^l, \mathbf{H}_{I_2}^l) = \sum_{k=1}^{4^l} \sum_{d=1}^D \min(\mathbf{H}_{I_1}^{l,k}(d), \mathbf{H}_{I_2}^{l,k}(d)) \quad (4-13)$$

式中  $\mathbf{H}_{I_1}^{l,k}(d)$ ——场景图像  $I_1$  在第  $l$  层的第  $k$  个子图区域视觉“单词”直方图向量的第  $d$  维的值；

$\mathbf{H}_{I_2}^{l,k}(d)$ ——场景图像  $I_2$  相应的值；

$\Gamma(\mathbf{H}_{I_1}^l, \mathbf{H}_{I_2}^l)$ ——图像  $I_1$  和图像  $I_2$  的第  $l$  层相似度<sup>[53]</sup>。

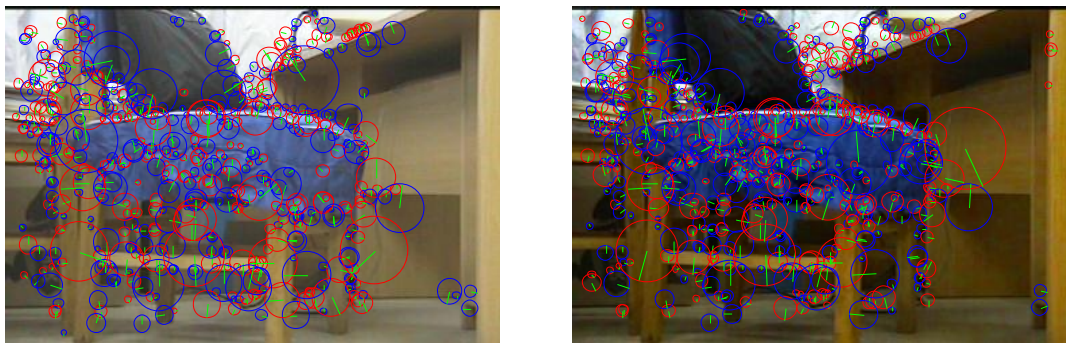
由空间金字塔的匹配原理可知，空间金字塔层数越大，场景图像划分的区域越多，匹配的精确随之越高，所占的权重比值也应该越大<sup>[54]</sup>。所以采用加权的方法表示不同层次场景图像的匹配，计算全部金字塔层次的子图区域，以此得到金字塔匹配核为

$$K_{SPM} = \Gamma^N + \sum_{l=0}^{N-1} \frac{1}{2^{N-l}} (\Gamma^l - \Gamma^{l+1}) = \frac{1}{2^N} \Gamma^0 + \sum_{l=1}^N \frac{1}{2^{N-l+1}} \Gamma^l \quad (4-14)$$

## 4.4 实验研究及分析

### 4.4.1 提取的特征点对比

为了验证对场景图像预处理的必要性，采用 SURF 算法，在相同环境下，对原图和处理后图分别提取的特征点数量作对比实验，实验结果如图 4-5 中 a) 和 b) 所示。



a) 原图特征点提取

b) 处理图特征点提取

图 4-5 图像处理前后特征点提取对比



对原图获得的特征点数量为 298 个，对处理后的图像获得的特征点数量为 341 个。处理后的图像提取的特征点数目的增多，使得图像的描述更丰富，有利于构建的视觉单词本更具区分度。此外，利用指导滤波方法执行速度快，不会增加整个算法的运行负担。

4.4.2 聚类效果对比

为了验证改进聚类算法构造视觉单词本的有效性，通过对随机生成的 120 个仿真实验样本点作为视觉“单词”进行实验，设定分类的类别为 4 类，分类结果用不同的形状显示样本点各个类的分布情况，图 4-6 中 a)和 b)分别为改进算法和 k 均值算法分类的的仿真实验结果，表 4-1 为由仿真实验获得的相应数据。

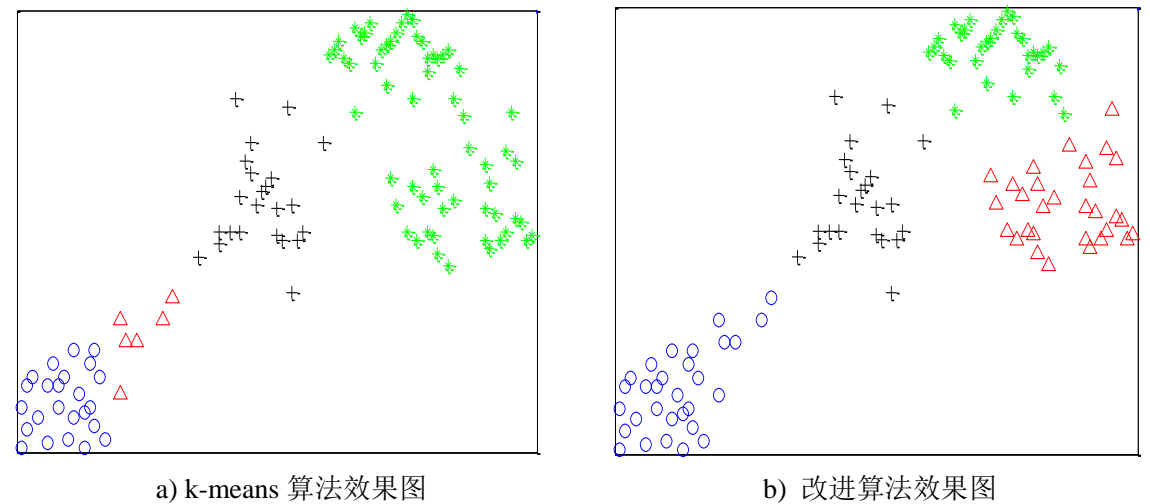


图 4-6 聚类效果对比

表 4-1 聚类效果数据对比

	k-means 算法	改进算法
迭代次数	9	2
聚类准则函数 $J (10^3)$	2.7433	1.8145

聚类准则函数越小，说明聚类结果越好。由实验结果和直观效果可见，改进的算法比 k 均值算法的聚类效果更合理，迭代次数也相应减少。为了验证改进算法的稳定性，用 10 组随机数据样本点对两种算法进行对比测试，在迭代次数指标上的结果如图 4-7 所示。k 均值算法的迭代次数在 4~9 次之间波动，而改进算法的迭代次数始终是 2 次，改进算法使得聚类的迭代次数减少，降低了程序的运行时间，提高了聚类的效率，合理的聚类也使得图像提取的特征更具表达能力。



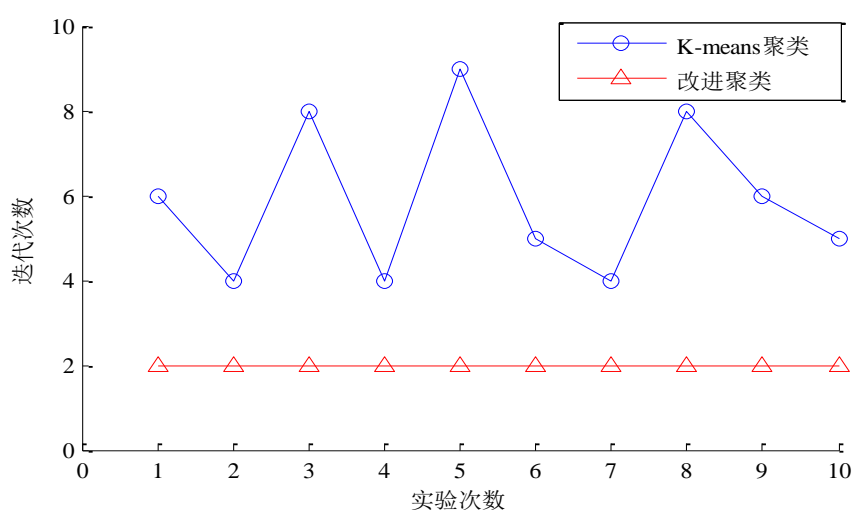
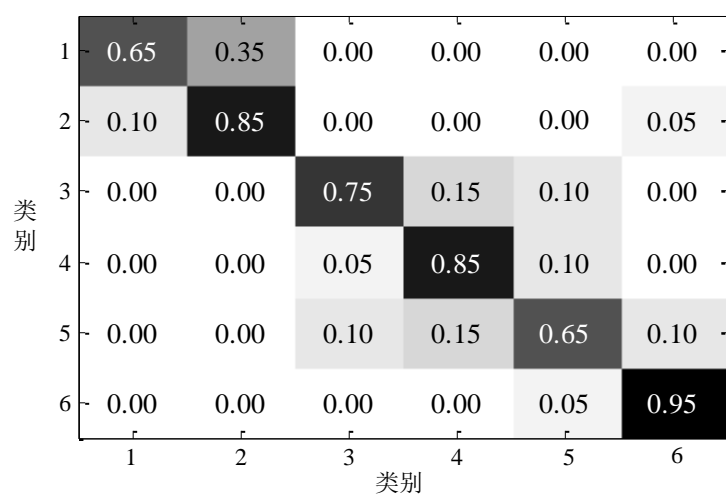


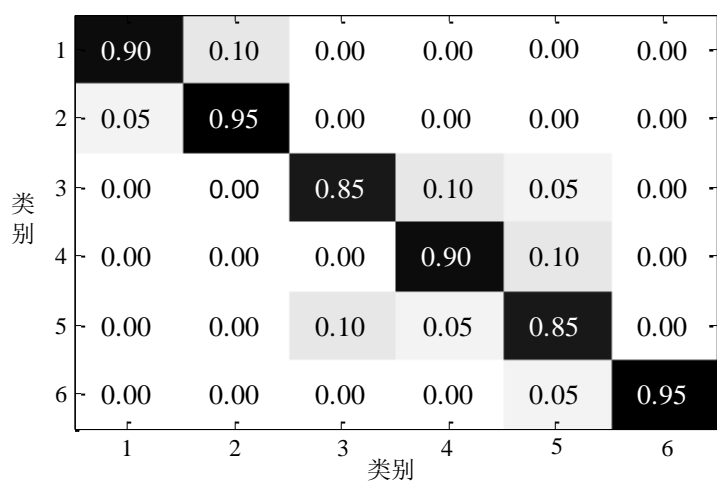
图 4-7 迭代次数对比

### 4.4.3 图像分类的效果对比

为了证明改进 BoVW 模型的有效性，进行场景图像的分类实验。图像库选取 Corel 数据库中<sup>[55]</sup>提供的 6 个相近类别的图像，类别越相近，图像相似性越强，更能体现算法的分类能力。其中，选定 240 幅图像作为训练样本，选取 120 幅图像作为测试样本进行实验。图 4-8 中 a)和 b)分别为经典的 BoVW 模型和改进的 BoVW 模型仿真得到的图像分类混合矩阵呈现分类结果。其中，矩阵中的黑色格子表示图像匹配成功的正确率，灰色格子表示图像匹配的错误率，白色格子表示图像的匹配没有出现错误。经典的 BoVW 算法在场景图像分类的平均正确率为 78.33%，改进 BoVW 算法的平均正确率 90%，改进的 BoVW 算法将准确率整体提高了 10%，准确率有了明显提高。



a) BoVW 算法的检测正确率



b) 改进算法的检测正确率

图 4-8 图像分类正确率对比

为了证明改进算法的有效性，通过对不同类别数目的场景图像集求取准确率，图 4-9 为不同场景类别数目的图像进行分类，横坐标分别包含 3、6、9、12 类图像，纵坐标为分类平均准确度。由统计实验结果可知，改进算法比其他现有算法均具有更高的准确率，使得图像分类的成功率更高。同时，改进算法在实验中正确率的变化率在 5% 以内，基本上趋于平稳，说明改进算法具有较高的鲁棒性，适用于实际应用。

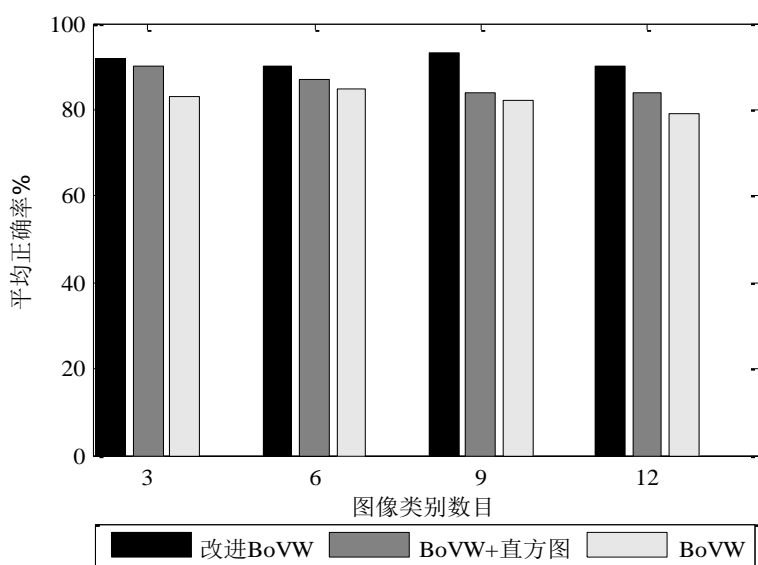


图 4-9 类目不同的场景分类平均正确率

## 4.5 本章小结

图像特征描述是机器人 VSLAM 和闭环检测的关键问题，视觉字典的构建和优

化也是 BoVW 模型的关键环节。本文提出的改进算法首先利用指导滤波变换对图像进行预处理操作，获得清晰的图像，利用基于 Krautchouk 矩的 SURF 算子进行特征提取，然后利用粒子群结合距离函数优化聚类算法建立视觉单词。利用改进 BoVW 方法有效改善了经典算法的聚类结果不稳定的缺陷，使视觉“单词”更具代表性，保证了算法稳定性。通过图像分类的对比实验，证明改进算法提高了正确率。

## 第 5 章 基于后验处理的闭环检测算法研究

### 5.1 引言

目前, 众多学者对闭环检测算法的研究越来越广泛, 梁志伟等人利用 BoW 模型和贝叶斯模型后验概率进行闭环检测, 实现了很好的闭环检测效果, 但由于其用 k 均值算法对特征点聚类, 容易使得聚类结果陷入局部最小, 不能保证算法的稳定性<sup>[56]</sup>。Angeli 等人利用增量式的检测方法, 通过图像的颜色和形状等特征的提取来增量式的构建词典, 增加了算法的复杂度, 加大了计算量, 并没有取得很好的闭环效果<sup>[57]</sup>。

针对现有的闭环检测方法精度不够等问题, 设计了基于后验处理的闭环检测方法。本章首先, 对机器人获取的场景图像, 利用基于 Krautchouk 矩改进的 SURF 算法提取图像的特征, 然后, 构建金字塔式的分层视觉单词树, 通过改进加权词频-逆向文件频率(Term Frequency-Inverse Document Frequency, TF-IDF)熵得分计算方法进行图像之间的相似性度量, 通过与阈值的比较从而确定初始的闭环, 而后利用时间上的约束和极线几何关系的后验处理进行闭环剔除, 最终得到有效正确的闭环。

### 5.2 闭环检测算法

闭环检测是 SLAM 中的基础问题之一, 是提升整体 SLAM 水平的必不可少的组成<sup>[58]</sup>。近几年来, BoVW 算法在闭环检测中有着越来越多的应用。针对 BoVW 模型的闭环检测算法的基本流程如图 5-1 所示。首先是对 SLAM 中的数据图像进行预处理, 获取清晰高质量的场景图像, 然后通过构建视觉字典树简化图像表示, 再次利用距离算法进行图像之间的相似性度量, 最后是通过阈值的设定进行闭环的验证。由于闭环检测对算法的实时性和准确性的高要求, 现有的闭环检测算法仍存在一些缺陷和不足。

对视觉词典的构建并不是分层处理, 而是在同一平面上, 使得词典的数据量大并且冗余, 不能突出各个“单词”的表征能力, 同时也会增加闭环检测系统的计算负担。如果减小词典的数据量的摄入, 就会使得“单词”不具有良好的表达力,

可能会使得获取图像的不同特征投影到同一个“单词”，从而使得相似度的计算出错，最终引发错误的闭环出现；而词典的数据量摄入过大，就会增加整个系统的计算负担，导致实时性差的问题。一些闭环检测的算法虽然在构建视觉词典时，在相似性计算上选取了优化的搜索算法，如：**k-d tree** 算法、**k** 近邻算法和最近邻搜索等算法。但不能从根本上解决闭环检测存在的问题，这些问题进而也会影响 **SLAM** 的实现效果<sup>[59,60]</sup>。

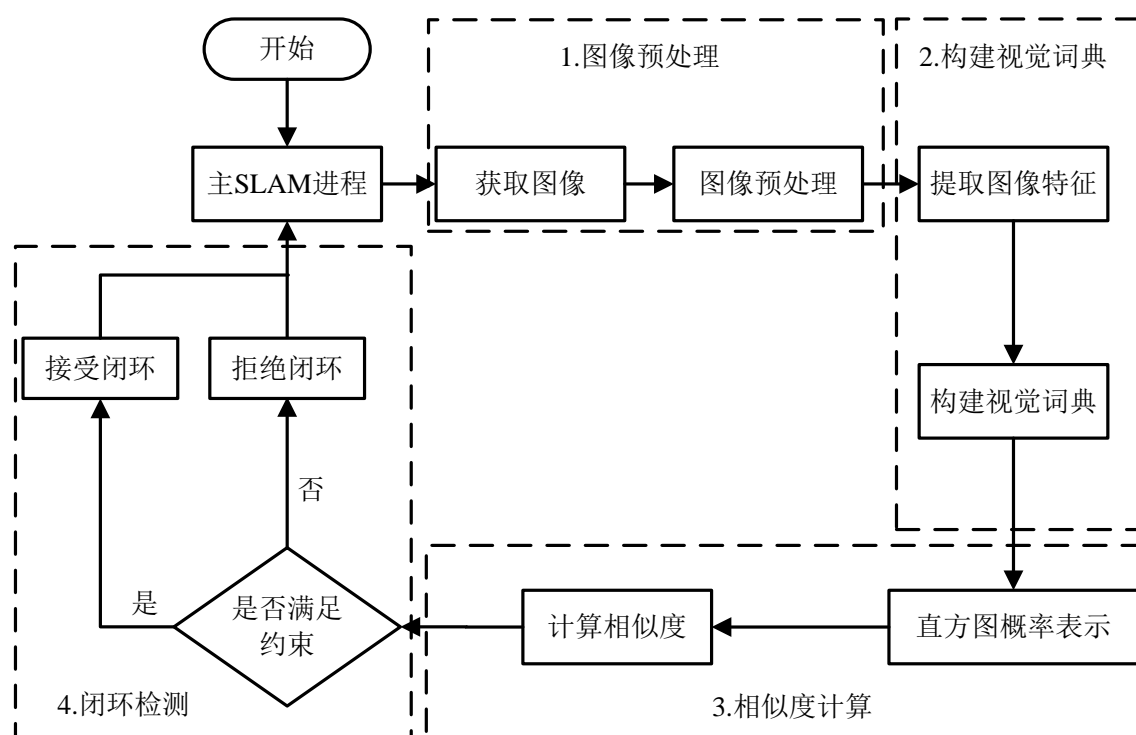


图 5-1 基于 BoVW 模型的闭环检测流程

相似度高的场景图像会产生视觉上的混淆问题，从而会引起感知歧义。闭环检测算法需具备减少感知歧义的能力，才能有效避免错误闭环。机器人是连续地进行图像提取，因此图像之间必然包含很多相同的因子，例如在课题室的室内场景下，桌子、椅子、电脑等物品都会使得机器人获取的图像具有大量相同的信息，尤其在特征提取的优化之后，图像的表达只是通过一些关键的特征信息，使得图像之间的差别变得更小，从而更会增加感知歧义的概率。而现有算法只是通过相似度度量时阈值的设定来避免此问题，如果阈值的设置过大，就可能排除掉正确的闭环，从而导致得不到正确的闭环结果；如果阈值的设置过小，就会出现很多错误的闭环，增加算法的计算量。

在开始出现闭环响应时，可能并不是正确的闭环产生，也可能是出现相似度高

的元素，所以需要后验处理来进行闭环的验证。现有算法大多是通过贝叶斯模型计算最大后验概率的方法进行验证，通过似然估计  $p(Z_t | J_t)$  进行闭环响应的后验概率计算，但由于似然估计自身存在一些估计误差，并且因为是对时间上的连续估计，就会使得误差变得越来越大，因此其不能准确地表达当前时刻图像的内容信息，也不能保证计算得到的后验概率完成正确。

针对以上闭环检测的不足与缺陷，本章提出了基于后验处理的闭环检测算法，首先构建金字塔式的分层视觉单词树改善平面结构下词典的数据量大的问题，提高“单词”的表征能力；通过加权的 TF-IDF 熵和结合反比例函数的得分计算方法进行图像之间的相似性度量，提高计算得分的可靠性，减小闭环检测的感知歧义；通过与阈值的比较从而确定初始的闭环，而后利用时间上的约束和极线几何关系的后验处理进行闭环剔除，最终得到有效正确的闭环。

## 5.3 改进的闭环检测算法

### 5.3.1 视觉单词树的创建和描述

聚类的效果和质量的的好坏程度直接影响视觉单词的生成，视觉单词的生成的快慢与聚类的速度息息相关。传统的视觉字典集由于其平面结构，计算代价高，表征能力差，往往不同的特征会投影到同一个视觉单词，从而导致闭环检测不准确。针对原始 BoVW 算法受单词集大小的制约，本文采用分层次 TF-IDF 熵的得分匹配算法。视觉字典树的构建如图 5-2 所示，一棵树的层数为  $L$ ，各分支的节点数为  $k$ ，对各个分支结构递归的调用  $k$  均值聚类方法，直到最大的  $L$  层停止。将全部分支特征的中心点作为树的节点，就是一个一个的视觉单词，以此构建成单词树，其表征能力远远超过传统的平面的视觉字典。

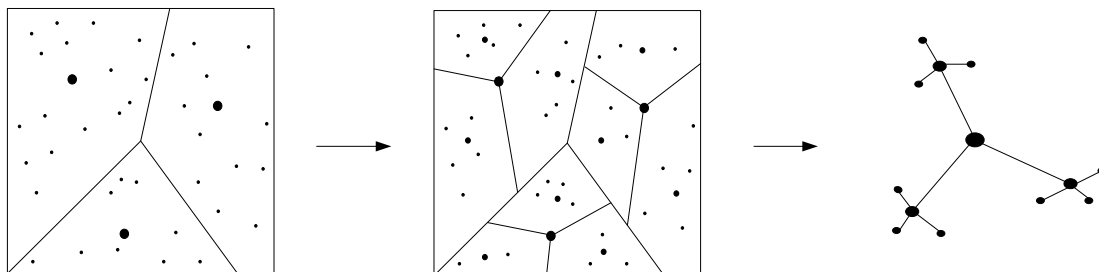


图 5-2 视觉字典树的构建

一颗包括  $L$  层  $k$  个分支结构的树其单词的表征数量为

$$\sum_{i=1}^L k^i = (k^{L+1} - k) / -1 \approx k^L \quad (5-1)$$

根据文献[61]和结合自身经验，在本文中， $L$  和  $k$  的取值分别为 5 和 4。

### 5.3.2 基于改进的 TF-IDF 熵的构造

构造视觉单词树，图像在树的各个节点的 TF-IDF 熵作为得分权重，从而进行场景的描述<sup>[62]</sup>。TF-IDF 熵的定义如下

$$w = \frac{n_i}{n} \log \frac{N}{N_i}, i = 1, 2, 3, \dots, k' \quad (5-2)$$

式中  $n_i$  —— 图像在节点  $i$  的特征个数；

$n$  —— 全部特征总数；

$N$  —— 进行图像处理的总个数；

$N_i$  —— 至少包含一个特征  $i$  的图像数目。

传统的 TF-IDF 熵计算简单，有很多的应用。但也存在不足：例如某一种特征  $k$  在某个类  $C_i$  中出现的频率特别高，也就是说  $k$  特征可以很好地表征  $C_i$  类的属性，应该被赋予大的权重，但又由于特征  $k$  在其他类别中出现的频率低或者为零，就会被赋予一个小的权重。因为特征  $k$  是类  $C_i$  的独有特征，辨识力强，理应被赋给大的权重，但事实却没有。同理，某个特征  $j$  频繁的出现在各个类别中，出现的频率高，就会被赋予很大的权重，但事实上，特征  $j$  的辨识能力差，就会使得权重的赋予不合理。因此，本文中引入基于加权的 TF-IDF 熵，定义如下

$$w_i^l(X) = k \frac{n_i}{n} \log \frac{N}{N_i} \quad (5-3)$$

$$k = \max \left( \frac{tf_i(k)}{\sum_{i=1}^m tf_i(k)} \times \frac{tf_i(k)}{S} \right) \quad (5-4)$$

式中  $w_i^l(X)$  ——  $X$  图像在树上的第  $l, l \in \{0, 1, 2, \dots, L\}$  层中的第  $i$  个节点处的投影得分；

$k$  —— 特征  $i$  在各个类中求取的数值最大的一个；

$tf_i(k)$  —— 特征  $k$  在  $C_i$  类内出现的频率数；

$\sum_{i=1}^m tf_i(k)$ ——特征  $k$  在全部类出现的总次数;

$S$ ——包含特征  $i$  的总图像数量。

假定图像库有三个类别  $C_1$ 、 $C_2$ 、 $C_3$  的图像, 每个类别有 4 幅图像, 只采用其中的三个特征  $k_1$ 、 $k_2$ 、 $k_3$ , 图像中 3 个特征出现的次数如表 5-1 所示。

表 5-1 TF-IDF 熵计算实例

	种类 C1				种类 C2				种类 C3			
	图 1	图 2	图 3	图 4	图 1	图 2	图 3	图 4	图 1	图 2	图 3	图 4
特征 $k_1$	5	5	5	5	0	0	0	0	0	0	0	0
特征 $k_2$	3	0	0	0	3	0	0	0	3	0	0	0
特征 $k_3$	2	0	0	0	0	0	0	0	0	0	0	0

由表 5-1 中的数据可知, 特征  $k_1$  区分能力强, 权重应该大,  $k_2$  的区分能力差, 权重应该小。以图像 1 为例, 计算  $k_1$  的原始 TF-IDF 熵为 0.7,  $k_2$  的原始 TF-IDF 熵为 0.6; 计算  $k_1$  的改进 TF-IDF 熵为 3.5,  $k_2$  的改进 TF-IDF 熵为 0.2。很明显本文的改进 TF-IDF 熵更加合理, 强调了  $k_1$  的重要性, 减弱了  $k_2$  的权重值。相对于原始 TF-IDF 熵, 本文改进的 TF-IDF 熵更加突出金字塔式视觉单词树上的各个“单词”的表征能力, 提高图像的分类效果, 从而提高闭环检测的正确率。

### 5.3.3 基于加权 TF-IDF 熵的得分匹配方法

$X$  图像在视觉树上的得分向量表述为

$$\mathbf{W}(X) = (\mathbf{W}^1(X), \mathbf{W}^2(X), \mathbf{W}^3(X), \dots, \mathbf{W}^L(X)) \quad (5-5)$$

$$\mathbf{W}^l(X) = (w_1^l(x), w_2^l(x), w_3^l(x), \dots, w_{k'}^l(x)) \quad (5-6)$$

由文献[63]可知, 金字塔分层式的 TF-IDF 得分计算方法可以使得闭环检测场景的感知歧义减少很多。计算  $X$  图像和  $Y$  图像在节点  $O_i^l$  处的相似度的有效得分公式为

$$S_i^l(X, Y) = \frac{1}{|w_i^l(X) - w_i^l(Y)| + 1} \quad (5-7)$$

用公式(5-7)代替传统的最小值法的  $H_i^l(X, Y) = \min\{w_i^l(x), w_i^l(y)\}$  公式, 两个图像的在节点  $O_i^l$  处相似性得分与其熵成反比, 其熵值越接近, 两幅图像越相似。有 A 和 B 两幅图像, 存在图像 A 的熵大于图像 B 的熵大于图像 X 的熵且



$H_i^l(X, A) = H_i^l(X, B)$ , 上式的计算结果更合理。通过分母的+1 控制二者的得分在  $(0, 1]$  之间, 排除得分受各个节点的影响太大的可能, 使得得分计算更合理。由此可得, 图像在第  $l$  层的相似性计算公式为

$$S^l(X, Y) = \sum_{i=1}^{k^l} S_i^l(X, Y) = \sum_{i=1}^{k^l} \frac{1}{|w_i^l(X) - w_i^l(Y)| + 1} \quad (5-8)$$

为了使得不重复地计算图像之间的相似性度量, 所以由金字塔底层向顶层计算相似度的增量  $\Delta S^l$ , 表示为

$$\Delta S^l(X, Y) = \begin{cases} S^L(X, Y), l = L \\ S^l(X, Y) - S^{l+1}(X, Y), 1 \leq l < L \end{cases} \quad (5-9)$$

金字塔的匹配核为

$$K(X, Y) = \sum_{l=1}^L \eta_l \Delta S^l(X, Y) \quad (5-10)$$

$\eta_l$  的定义是单词树的第  $l$  层的匹配程度系数。为了避免不同层次间匹配的差异性, 匹配核重新表示为

$$K(X, Y) = K(W(X), W(Y)) = S^L(X, Y) + \sum_{l=1}^{L-1} \frac{1}{k^{L-l}} (S^l(X, Y) - S^{l+1}(X, Y)) \quad (5-11)$$

通过计算各个  $K(X, Y_i)$  的值, 选取得分超过所设阈值的图像  $Y_i$  帧序列的位置作为候选的闭环位置存入数据库进行后验处理。

#### 5.3.4 基于后验处理的闭环检测方法

首先, 机器人在运动过程中, 机器人的 Camera 摄取的场景图像可能会发生场景得翻转、镜头的伸缩等机器外在原因和在传输图像的时候可能存在一些不确定的干扰; 此外, 图像特征提取的精度、图像的尺度和拍摄场景的混淆等原因, 会使得通过单纯计算相似度选取的闭环位置不能保证全部是正确的, 因此本文引入基于后验处理的闭环检测方法。

通过时间上的约束和图像矩阵的极线几何(Epipolar Geometry)关系<sup>[64]</sup>来剔除错误的候选闭环, 从而选定最终正确闭环。由于机器人图像获取的时间都是连续的, 所以在连续的图像帧之间会相应地出现候选闭环, 所以可以通过此约束方式剔除错误的闭环。如图 5-3 所示, 在第 9 帧时刻和第 2 帧时刻、第 10 帧时刻和第 3 帧时刻、

第 11 帧时刻和第 4 帧时刻发生了闭环，然而在第 12 帧时刻和第 3 帧时刻也出现发生的闭环现象，以上这种情况的发生不满足时间约束条件，则不被接受闭环，从而剔除了错误闭环。

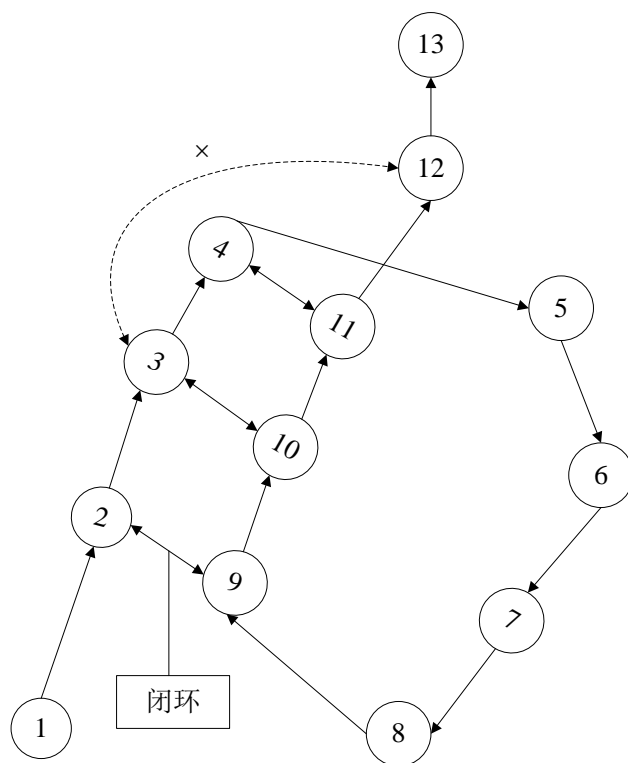


图 5-3 确定闭环的时间约束条件

在产生闭环处的一对图像就是机器人摄像机在位置差别引起的视觉角度不同时对相同场景的成像结果，所以这两幅图像满足极线几何关系。极线几何约束呈现了相应点对重要的约束条件，通过计算这对图像的基础矩阵，根据其内点数的比例与阈值的比较可以确定两幅图像是否满足极线几何关系。如果满足，接受闭环；反之，剔除闭环。

由以上关键步骤可得到本文的基于后验处理的闭环检测流程，如图 5-4 所示。首先进行条件预判断，判断新获取的新图像与历史图像的相似度是否超过设定的阈值；然后对满足预设条件的场景图像进行特征提取，采用基于 **Krautchouk** 矩改进的 **SURF** 算法提取图像的特征，构建金字塔式的分层视觉单词树；通过加权 **TF-IDF** 熵的得分计算方法进行图像间相似性的度量，进行相似性得分计算；通过与阈值的对比来判断是否发生闭环，如果超过了设定阈值，则记录为初始闭环，如果没有超过阈值，则认为没有产生闭环；对产生的初始闭环，利用时间约束和极线几何关系的后验处理进行闭环剔除与确认，最后得到有效的正确闭环。

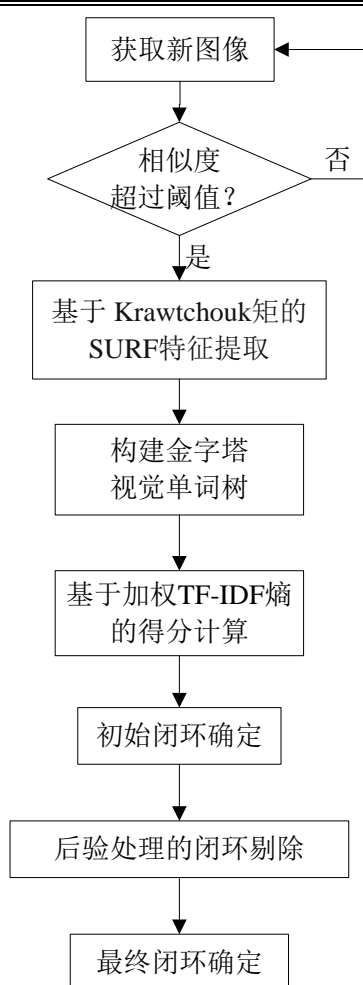


图 5-4 闭环检测流程图

## 5.4 实验研究及分析

为了验证改进算法的有效性,对机器人分别获取的 30 帧室内和室外图像,设定 6 个闭环,利用基于 SIFT 的闭环检测算法<sup>[65]</sup>与本文改进算法作对比实验,结果如表 5-2 所示。

表 5-2 室内、外的闭环检测结果

	本文算法		SIFT 算法	
	室内	室外	室内	室外
实际闭环数目	5	5	6	5
正确闭环数目	5	5	4	4
召回率(%)	100	100	66.67	80
正确率(%)	93.33	91.27	86.67	83.46

其中，召回率 =  $\frac{\text{正确闭环数量}}{\text{实际闭环数量}}$ 。

由于室外环境的复杂性及光照、角度等原因，本文算法的检测成功率没有达到 100%，但本文算法在检测的正确率和召回率的结果都明显优于 SIFT 算法，比 SIFT 算法的准确率平均提高了近 8%，召回率平均提高了 26% 左右。另外，本文算法在室内和室外的召回率达到了 100%，检测正确率达到了 90% 以上，可以满足实时性应用。

为了更好的验证本章算法，以课题室环境为背景，进行闭环检测实验，实验平台如图 5-5 中图(a)和图(b)所示。在 matlab R2013a 环境下，利用机器人采集信息进行实验。图 5-6 是实验室环境的虚拟地图，粗红线代表机器人的真实运动轨迹。



a) 机器人



b) 摄像头

图 5-5 移动机器人平台

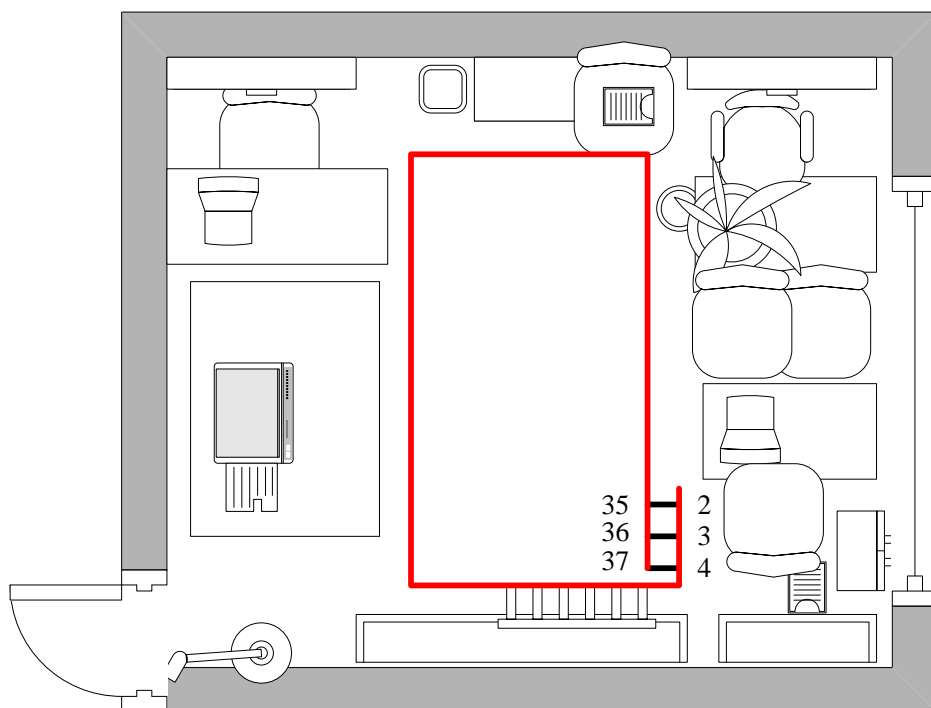


图 5-6 实验室环境虚拟地图

图 5-7 是机器人实际采集图像序列的相应的产生闭环的位置。图 5-8 是进行课题室环境闭环检测后验处理的结果，蓝色圆形是闭环检测的数据，红色三角形表示检测到闭环。

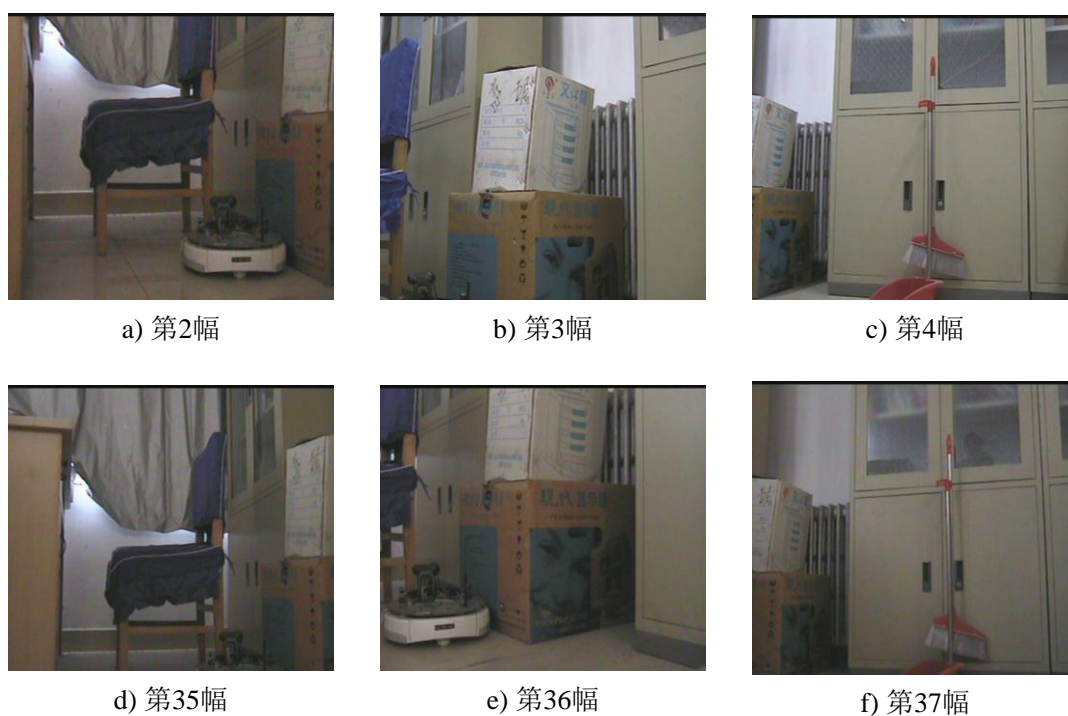


图 5-7 闭环位置的图像序列

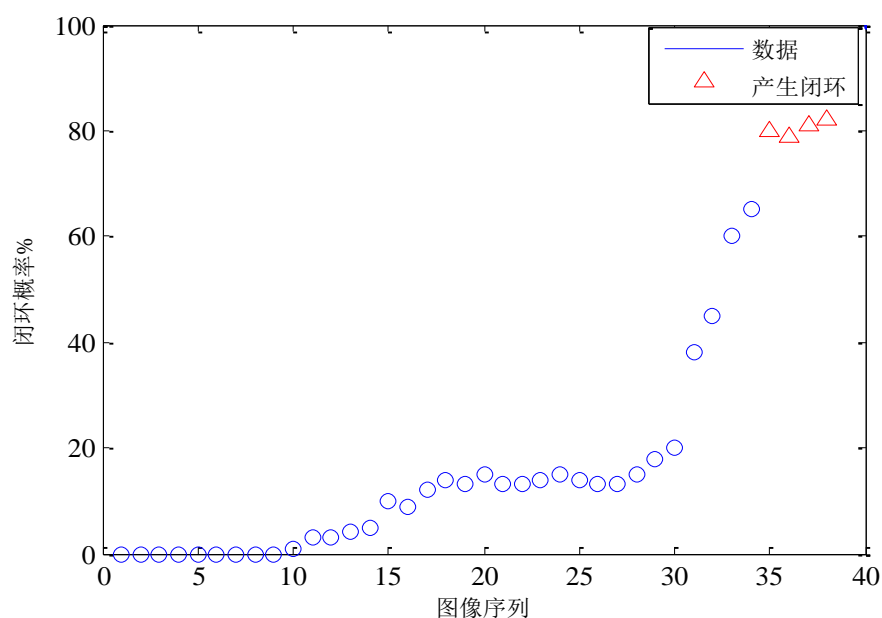


图 5-8 闭环检测结果

本实验中，阈值的设置为 0.8。在图像序列的第 35 幅场景与第 2 幅场景成功检测到了闭环，并且根据时间约束和极线几何关系的后验处理最终成功确定闭环。通

通过对课题室环境的闭环检测实验，本文的闭环检测算法得到了很好的验证，依靠后验处理操作，有效地保证了闭环检测准确度。

## 5.5 本章小结

在本章，对现有闭环检测算法中存在的视觉混淆、错误闭环的问题，提出了基于后验处理的闭环检测算法，构建金字塔式视觉单词树，提高“单词”的表征能力；通过加权 TF-IDF 熵和结合反比例函数的得分计算方法进行图像之间的相似性度量，通过与阈值的比较从而确定初始的闭环，而后利用时间上的约束和极线几何关系的后验处理进行闭环剔除，最终得到有效正确的闭环。

## 结 论

移动机器人闭环检测问题是机器人完成地图构建与即时定位任务的基本和关键。本文针对移动机器人的地图构建与即时定位中闭环检测环节准确率低的问题，对基于视觉的移动机器人 SLAM 中的闭环检测算法进行了研究改进。主要创新性成果如下：

(1)针对场景图像特征点匹配准确度不高、实时性不够的缺陷，采用了颜色特征和局部特征融合的算法，通过粗匹配和精确匹配相结合的方式，利用粗匹配以少的时耗为后续精确匹配排除了与模板图像不相似图像，减少了后续精确匹配的工作量；利用 Krawtchouk 矩不变量改进 SURF 算子，得到描述能力强的特征点进行再匹配，得到了最佳的匹配结果，通过 MATLAB 仿真实验，证明了算法大大提高了图像处理的效率。

(2)针对传统的 BoVW 模型准确率低的问题，提出一种改进 BoVW 方法，采用指导滤波方法获取去雾后清晰的场景图像，利用基于 Krautchouk 矩的 SURF 算子进行特征提取，得到优化的特征向量，最后利用最大最小距离结合粒子群聚类算法构建出高质量、有代表性的视觉单词本。通过聚类对比试验，证明改进算法降低了程序运行时间；进行图像分类的对比实验，证明改进算法提高了正确率。

(3)针对闭环检测算法的正确率不高的问题，设计基于后验处理的闭环检测方法，通过加权 TF-IDF 熵和结合反比例函数的得分计算方法进行相似性度量，使得相似度计算更加准确，构建出金字塔式表征能力强的视觉单词树；通过与阈值比较确定初始闭环后，利用时间上的约束和极线几何关系的后验处理去除错误闭环，最终得到有效正确的闭环。通过实验，验证算法了可行性，保证了闭环检测正确率。

在基于视觉 SLAM 的闭环检测研究中，通过实验结果得到了一些有效可行的进展，还存在需完善的问题：

(1)本文针对图像特征提取的改进算法，提高了图像特征匹配正确率，但在特征提取的时间上没有显著提高，有待进一步改善。

(2)针对闭环检测提出的改进算法，提高了检测的正确率，本文算法在小规模的图像样本实验中体现出了优越性，在大规模的图像情况下，正确率还需验证。

## 参考文献

- [1] Wei H, Chen Y, Zhao Y. Adaptive Neural Network Control of an Uncertain Robot With Full-State Constraints[J]. IEEE Transactions on Cybernetics, 2016, 46(3):620-629.
- [2] 李庆, 郑力新, 潘书万, 等. 使用单目视觉的移动机器人导航方法[J]. 计算机工程与应用, 2017, 53(4):223-227.
- [3] Kuindersma S, Deits R, Fallon M, et al. Optimization-based locomotion planning, estimation, and control design for the atlas humanoid robot[J]. Autonomous Robots, 2016, 40(3):429-455.
- [4] Robin C, Lacroix S. Multi-robot target detection and tracking: taxonomy and survey[J]. Autonomous Robots, 2016, 40(4):729-760.
- [5] Carrasco P L N, Bonin-Font F, Codina G O. Stereo Graph-SLAM for Autonomous Underwater Vehicles[J]. 2016, 302:351-360.
- [6] Ivez-L, Pez D, Salas M, et al. Real-time monocular object SLAM[J]. Robotics & Autonomous Systems, 2015, 75(PB):435-449.
- [7] Aubry C, Desmare R, Jaulin L. Loop detection of mobile robots using interval analysis[J]. Automatica A Journal of Ifac the International Federation of Automatic Control, 2013, 49(2):463-470.
- [8] Liang Z, Gao X, Chen Y, et al. A novel loop closure detection method in monocular SLAM[J]. Intelligent Service Robotics, 2013, 6(2):79-87.
- [9] Zakhor A. Automatic loop closure detection using multiple cameras for 3D indoor localization[C]// IS&T/SPIE Electronic Imaging. International Society for Optics and Photonics, San Francisco, USA, 2012:1-3.
- [10] Liu Y, Zhang H. Visual loop closure detection with a compact image descriptor[C]. Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on IEEE, Chicago, USA, September, 2012:1051-1056.
- [11] Clemente L, Davison A, Reid I, et al. Mapping large loops with a single hand-held camera[C]. In Proceedings of Robotics: Science and Systems, Atlanta, USA, 2007:297-304.
- [12] Williams B, Klein G, Reid I. Automatic Relocalization and Loop Closing for Real-Time Monocular SLAM[J]. IEEE Transactions on Software Engineering, 2011, 33(9):1699-1712.



- [13] Williams B, Cummins M, Neira J, et al. An image-to-map loop closing method for monocular SLAM[C]. In Proceedings of International Conference on Intelligent Robots and Systems, Nice, France, 2008: 2053-2059.
- [14] Angeli A, Filliat D, Doncieux S, et al. Fast and Incremental Method for Loop Closure Detection Using Bags of Visual Words[J]. IEEE Trans on Robotics, 2008, 4(5): 1027-1037.
- [15] Cummins M, Newman P. FAB-MAP: Probabilistic Localization and Mapping in the Space of Appearance[J]. International Journal of Robotics Research, 2008, Vol. 27(6): 647-665.
- [16] Cummins M, Newman P. Appearance-Only SLAM at Large Scale with FAB-MAP 2.0[J]. The International Journal of Robotics Research, 2011, 30(9): 1100-1123.
- [17] 梁志伟, 陈燕燕, 朱松豪, 等. 基于视觉词典的单目视觉闭环检测算法[J]. 模式识别与人工智能, 2013, 26(6): 561-570.
- [18] Zhang H, Liu Y, Tan J. Loop Closing Detection in RGB-D SLAM Combining Appearance and Geometric Constraints[J]. Sensors, 2015, 15(6):14639-14660.
- [19] 陈燕燕. 混合人工自然特征环境的单目视觉 SLAM[D]. 南京: 南京邮电大学控制理论与控制工程硕士学位论文, 2013, 1-4.
- [20] Glover A, Maddern W, Warren M, et al. Open FABMAP: An open source toolbox for appearance-based loop closure detection[C]. Robotics and Automation (ICRA), 2012 IEEE International Conference on IEEE, Karlsruhe, Germany, 2012: 4730-4735.
- [21] 曾新贵, 陶卫, 颜发才, 等. 基于 Harris 算子和方向场的图像配准算法[J]. 计算机应用, 2016, 36(S1):146-148.
- [22] 任伟建, 王子维, 康朝海. 基于改进 SIFT 算法的无人机遥感图像匹配[J]. 计算机科学, 2015(B11):179-182.
- [23] 张志斌, 赵帅领, 罗锡文, 等. 基于 SURF 算法的绿色作物特征提取与图像匹配方法[J]. 农业工程学报, 2015, 31(14):172-178.
- [24] Kadir H A, Arshad M R. Features detection and matching for visual simultaneous localization and mapping (VSLAM)[M].Control System, Computing and Engineering (ICCSCE), Penang, Malaysia. 2013,40-45.
- [25] Bay H, Ess A, Tuytelaars T, et al. Speeded-Up Robust Features (SURF)[J]. Computer Vision & Image Understanding, 2008, 110(3):346-359.
- [26] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transactions

- on Pattern Analysis & Machine Intelligence, 2005, 27(10):1615-1630.
- [27] 刘萍萍, 赵宏伟, 臧雪柏. 移动机器人定位图像匹配的快速局部特征算法[J]. 仪器仪表学报, 2009, 30(8): 1714 -1719.
- [28] 廉蒨, 李国辉, 田昊, 等. 加窗灰度差直方图描述子及其对 SURF 算法的改进[J]. 电子与信息学报, 2011, 33(5): 1042-1048.
- [29] Yap P T, Paramesran R, Ong S H. Image analysis by Krawtchouk moments[J]. Image Processing IEEE Transactions on, 2003, 12(11):1367-1377.
- [30] Valgren C, Lilienthal J A. Appearance-based long-term localization in outdoor environments[J]. Robotics and Autonomous Systems, 2010, 58(2):149-156.
- [31] Wang Kai, Cheng Bo, Ma Lu, Xiaolu Song. Multi-source Remote Sensing Image Registration Based on Normalized SURF Algorithm[J].Computer Science and Electronics Engineering (ICCSEE), 2012:373-377.
- [32] Hye-Jin Kim, Jae Yeon Lee, Jae Hong Kim. Object recognition and pose estimation using KLT Control[J]. Automation and Systems (ICCAS), 2012: 214-217.
- [33] Venkataramana A, Raj P A. Image Watermarking Using Krawtchouk Moments[C]// International Conference on Computing: Theory and Applications. Changsha, China, August, DBLP, 2007:676-680.
- [34] 张力, 肖薇薇, 钱恭斌. 基于 Krawtchouk 不变矩的仿射攻击不变性局部水印算法[J]. 电子学报, 2007, 35(7): 1403-1408.
- [35] Cucchiara R, Grana C, Piccardi M. Improving shadow suppression in moving object detection with HSV color information[C]. Proceedings of the IEEE Intelligent Transportation Systems, Oakland, August, CA, 2001: 334-339.
- [36] 陈卫东, 张兴家, 朱奇光, 等. 基于混合特征的移动机器人图像匹配算法[J]. 中国机械工程, 2015, 26(9):1137-1141.
- [37] 闫自庚, 蒋建国, 郭丹. 基于 SURF 特征和 Delaunay 三角网格的图像匹配[J]. 自动化学报, 2014, 40(6):1216-1222.
- [38] 刘佳, 傅卫平, 王雯, 等. 基于改进 SIFT 算法的图像匹配[J]. 仪器仪表学报, 2013, 34(5):1107-1112.
- [39] 朱奇光, 王佳, 张朋珍, 等. 基于高斯矩改进 SURF 算法的移动机器人定位研究[J]. 仪器仪表学报, 2015, 36(11):2451-2457.

- [40] Kato H, Harada T. Image Reconstruction from Bag-of-Visual-Words[J]. Computer Science, 2015:955-962.
- [41] Purda L, Skillicorn D. Accounting Variables, Deception, and a Bag of Words: Assessing the Tools of Fraud Detection[J]. Contemporary Accounting Research, 2015, 32(3):1193–1223.
- [42] Aldavert D, Rusiñol M, Toledo R, et al. A study of Bag-of-Visual-Words representations for handwritten keyword spotting[J]. International Journal on Document Analysis and Recognition (IJDAR), 2015, 18(3):223-234.
- [43] Bahmanyar, shiyong Cui, Datcu. A Comparative Study of Bag-of-Words and Bag-of-Topics Models of EO Image Patches[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 12(6): 1357-1361.
- [44] 王卫星, 肖翔, 陈良琴. 结合最小滤波和引导滤波的暗原色去雾[J]. 光学精密工程, 2015, 23(7):2100-2108.
- [45] Xiang J C, Liu X B. Research on K-Means clustering algorithm based on density and its application to customer segmentation[J]. Computer Engineering & Applications, 2006, 44(35) :246-248.
- [46] Kanungo T, Mount D M, Netanyahu N S, et al. An efficient k-means clustering algorithm: analysis and implementation[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2002, 24(7):881-892.
- [47] Na S, Liu X, Yong G. Research on k-means Clustering Algorithm: An Improved k-means Clustering Algorithm[C]// International Symposium on Intelligent Information Technology and Security Informatics, Iitsi 2010, Jinggangshan, China, April. DBLP, 2010:63-67.
- [48] Hui Z, Fei Y. A Novel Fuzzy Clustering Recommendation Algorithm Based On Pso[J]. Cybernetics & Information Technologies, 2014, 14(5):108-117.
- [49] 朱道广. 基于视觉词袋模型的图像分类研究[D]. 郑州: 解放军信息工程大学信号与信息处理硕士学位论文, 2013, 15-20.
- [50] 生海迪. 视觉词袋模型的改进及其在图像分类中的应用研究[D]. 济南: 山东师范大学计算机理论与理论硕士学位论文, 2015, 5-12.
- [51] 李丽双, 王意文, 黄德根. 基于信息熵和词频分布变化的术语抽取研究[J]. 中文信息学报, 2015, 29(1):82-87.
- [52] 王小林, 陆骆勇, 邵伟鹏. 基于信息熵的新的词语相似度算法研究[J]. 计算机技术与发展,

- 2015(9):119-122.
- [53] 黄承慧, 印鉴, 侯昉, 等. 一种结合词项语义信息和 TF-IDF 方法的文本相似度量方法[J]. 计算机学报, 2011, 34(5):856-864.
- [54] 王子慕. 一种利用 TF-IDF 方法结合词汇语义信息的文本相似度量方法研究[D]. 吉林: 吉林大学软件工程硕士学位论文, 2015, 21-38.
- [55] Griffin G, Holub A, Perona P. Caltech-256 Object Category Dataset[J]. California Institute of Technology, 2007,1-18.
- [56] 袁方, 周志勇, 宋鑫. 初始聚类中心优化的 k-means 算法[J]. 计算机工程, 2007, 33(3):65-66.
- [57] 崔大成, 曾连荪. 基于视觉字典的移动机器人闭环检测方法研究[J]. 微型机与应用, 2015, 34(9):85-88.
- [58] 李维鹏, 张国良, 姚二亮, 等. 基于空间位置不确定性约束的改进闭环检测算法[J]. 机器人, 2016, 38(3):301-310.
- [59] 李博, 杨丹, 邓林. 移动机器人闭环检测的视觉字典树金字塔 TF-IDF 得分匹配方法[J]. 自动化学报, 2011, 37(6):665-673.
- [60] Botterill T, Mills S, Green R. Bag-of-words-driven, single-camera simultaneous localization and mapping[J]. Journal of Field Robotics, 2011, 28(2):204-226.
- [61] 李永锋, 张国良, 王峰, 等. 一种基于历史模型集的改进闭环检测算法[J]. 机器人, 2015, 37(6):663-673.
- [62] Zhang Yuntao, Gong Ling, Wang Yongcheng. An improved TF-IDF approach for text classification[J]. Journal of Zhejiang University-SCIENCE A, 2005, 6(1):49-55.
- [63] 李博. 基于场景外观建模的移动机器人视觉闭环检测研究[D]. 重庆: 重庆大学计算机科学与技术博士学位论文, 2011, 93-108.
- [64] Jiang L Z. Application of SURF Feature and Preprocessing RANSAC Algorithm in Face Recognition[J]. Computer Science, 2015, 1234:147-158.
- [65] 孙旻喆. 基于 RGB-D 图像的 SLAM 闭环检测方法研究[D]. 秦皇岛: 燕山大学计算机应用技术硕士学位论文, 2014, 30-37.

## 攻读硕士学位期间承担的科研任务与主要成果

### (一)参与的科研项目

[1] 未知复杂非结构环境下的不确定移动机器人导航控制策略研究, 河北省自然科学基金项目. 项目编号: F2016203245.

### (二)发表的学术论文

[1] 朱奇光, 张朋珍, 李昊立, 等. 基于全局和局部特征融合的图像匹配算法研究[J]. 仪器仪表学报, 2016, 37(1):170-176.

[2] 朱奇光, 王佳, 张朋珍, 等. 基于高斯矩改进 SURF 算法的移动机器人定位研究[J]. 仪器仪表学报, 2015, 36(11):2451-2457.

## 致 谢

时光荏苒，白驹过隙。恍惚之中，才发现研究生生活即将画上句号。在此论文付梓之际，我怀着感恩的心向所有给予过我帮助和支持的老师、同学致以最诚挚的感谢。

首先，我要衷心感谢我的导师--朱奇光副教授。他治学严谨，以身作则，在课题研究上始终给予我细心的指导和无私的帮助。每当我的课题遇到难题时，朱老师总会耐心地倾听我的困惑，同我一起讨论、分析，然后找出问题的症结所在。是他，让我从一个对课题一窍不通的新手，转变成有着自己课题见解的准研究者。此外，每当我心态不佳、彷徨迷茫之际，朱老师就会和我谈心，分享作为过来者的经验或是给出建议，总能让我重新燃起斗志。倘若没有朱老师的鼓励和鞭策，我的研究生生活不会有如此多的收获。

其次，我要感谢三年来和我一起奋斗的同伴们。我们来自天南海北，缘分的力量让我们相识。感谢所有实验室的小伙伴们，是你们让我感受到团队的力量，和你们在一起枯燥的生活也变得有趣。还要感谢我的室友，她们一直在生活上给我很大的帮助和精神支持，同时也帮助我克服了很多难关。此外，感谢信息科学与工程学院全体老师的培养。

最后，我要感谢我的家人。漫漫求学路，正是你们的支持才使我一路走来。谢谢你们背后无私的付出，我会尽我所能让你们幸福。