

文章编号:1001-9081(2014)02-0576-04

doi:10.11772/j.issn.1001-9081.2014.02.0576

## 基于视觉同时定位与地图构建数据关联优化算法

赵亮\*, 陈敏, 李洪臣

(电子科技大学 航空航天学院, 成都 611731)

(\*通信作者电子邮箱 dfasfzhao@126.com)

**摘要:**数据关联的复杂程度随着地图规模的不断扩大而增加是导致机器人同时定位与地图创建(SLAM)实时性差的一个主要原因。在SLAM系统中,主要应用尺度不变特征变换(SIFT)算法提取自然路标。提出两种方法来改进数据关联的实时性:1)提取感兴趣区域;2)引入当前路标的物理位置信息作预判断。实验结果表明,所提的改进方法是可靠的,改善算法复杂度的效果是显而易见的。

**关键词:**同时定位与地图构建; 数据关联; 边缘提取; 区域裁剪; 特征提取

**中图分类号:** TP391.4    **文献标志码:**A

### Optimized data association algorithm based on visual simultaneous localization and mapping

ZHAO Liang\*, CHEN Min, LI Hongchen

(School of Aeronautics and Astronautics, University of Electronic Science and Technology of China, Chengdu Sichuan 611731, China)

**Abstract:** The scale of data association increases as the map grows, which is one of the major reasons for the poor real-time performance of robot in the process of Simultaneous Localization And Mapping (SLAM). In visual SLAM system, SIFT (Scale Invariant Feature Transform) algorithm was used to extract the natural landmarks. Two improvements were introduced to improve the real-time of data association: firstly, extracted the "interest region"; secondly, took into account the physical location of current landmarks. The experimental results indicate that this kind of improvement method is reliable, and the capability of reducing computational complexity is obvious.

**Key words:** Simultaneous Localization And Mapping (SLAM); data association; edge extraction; regional cutting; feature extraction

### 0 引言

同时定位与地图构建(Simultaneous Localization and Mapping, SLAM)问题是近年来移动机器人领域的研究热点。该问题是将移动机器人放置于未知环境中,机器人通过自身携带的传感器增量式地创建该未知环境的地图,同时利用自身传感器和外部信息确定它在地图中的位置。随着机器人应用环境的不断复杂化和非结构化,人们对机器人的自主化要求不断提高,因此SLAM问题正逐渐成为移动机器人自主化的关键技术。

大部分研究者认为SLAM问题分为两个部分<sup>[1]</sup>:状态估计和数据关联。状态估计是指机器人通过自身携带传感器测量未知环境特征(路标或障碍物)位置并通过环境特征信息估计自身位置;而数据关联问题作为状态估计的输入是实现状态正确估计的前提条件,所以数据关联对SLAM的正确运行是至关重要的。

数据关联是指在不同时间、地点建立一套机器人传感器量测之间、地图库特征点之间以及传感器量测与地图库特征点之间的对应关系,并确定这些量测和特征点是否为环境中的同一物理特征点<sup>[1-3]</sup>的过程,同时在这一过程中将未匹配的点作为新发现的点加入地图库中。数据关联是SLAM问题

的重点难点问题之一,关联失败将直接导致SLAM整体算法的发散,从而导致算法失败。为了保证SLAM算法的有效性,必须保证数据关联的正确性和实时性,这就要求关联算法满足计算复杂度小、判断准则正确率高这两条要求。而在实际SLAM问题中,建立的地图库中的特征点(路标)数量随着时间的积累将变得异常的巨大,每个特征点又具有一个庞大的特征描述矩阵和位置信息矩阵,进行匹配时,新检测到的特征点需要和库中的每一个特征进行比较,可想而知这是一个巨大的计算过程,将严重影响关联的实时性。

本项目使用的机器人外部传感器为摄像头,通过摄像头采集图像,从图像中提取特征点。为了确保关联速度,本文提出从两个方面改善数据关联的实时性:1)从传感器获取的原始图像中提取特征明显(感兴趣区域)的特征点;2)将感兴趣区域匹配成功的路标点位置信息与欧氏距离相结合进行关联。通过以上两点一方面可以减慢地图库数据量的增加速度;另一方面可以使当前匹配的路标点与地图库中少量的符合条件的路标点进行关联,而不必与所有路标一一关联,从而改善整个关联过程的实时性和准确性。

### 1 SLAM 特征提取及数据关联

特征提取和数据关联是SLAM问题的两个重要难点。特

收稿日期:2013-08-09;修回日期:2013-11-14。    基金项目:总装预研基金资助项目(9140A01010111DZ0205)。

作者简介:赵亮(1987-),男,河南项城人,硕士研究生,主要研究方向:移动机器人同时定位与地图创建、模式识别; 陈敏(1974-),男,四川成都人,副研究员,博士,主要研究方向:音频信号处理、智能系统; 李洪臣(1987-),男,山东菏泽人,硕士研究生,主要研究方向:数字图像处理、智能机器人控制与应用、嵌入式系统。

征提取为 SLAM 系统提供稳定的特征路标,而特征稳定路标是移动机器人能够准确定位和创建地图的前提条件;而不正确的数据关联往往会导致地图特征估计的发散和特征定位的不准确,同时随着时间的积累,地图库的不断扩大,数据关联的运算复杂度将不断提高,关联的实时性将受到很大的挑战。

### 1.1 视觉 SLAM 的特征提取

近年来,利用视觉传感器的 SLAM 系统逐渐成为研究热点。视觉 SLAM(VisualSLAM)兴起的原因主要是其与传统的声呐、激光和超声传感器相比具有成本低廉、获取特征丰富等优势。VisualSLAM 常用的特征提取算法有基于形状识别、纹理、角点提取和尺度不变特征转换(Scale-Invariant Feature Transform,SIFT)等算法。其中 SIFT 算法以其对尺度、旋转具有不变性,同时对光线变化、视角变换和噪声也有很好的容忍度<sup>[4-6]</sup>等优势被广泛采用。

SIFT 算法的特征提取过程如图 1 所示。该算法由检测极值点、精确定位极值点、消除低对比度关键点、确定主梯度方向和生成特征描述符等过程组成。

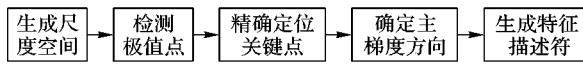


图 1 SIFT 算法工作流程

#### 1.1.1 生成 DoG 尺度空间

为了有效地提取稳定的特征点,Lowe 提出高斯差分尺度空间(DoG scale space)<sup>[7]</sup>,该空间利用不同尺度的高斯差分核与图像卷积生成。

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

其中: $\sigma$  为正态分布标准差,\* 为卷积符号, $I(x, y)$  为原图像, $G(x, y, \sigma)$  为原图像高斯函数。

#### 1.1.2 极值点检测

为了得到尺度空间的极值点,采样点要与和该点相邻的所有点进行比较,即与同尺度的 8 个相邻点和上下邻域的 2×9 个点共 26 个点进行比较,看该点是否为最大或最小的极值点。

#### 1.1.3 关键点精确定位

为提高匹配准确性和抗干扰能力。需对特征点作如下处理:1)通过拟合三维二次函数以精确定位关键点的位置和尺度(达到亚像素精度);2)通过设置合适的阈值(一般取 0.03)滤除低对比度的关键点;3)利用阈值(一般取 10)滤除不稳定边缘响应的关键点<sup>[7]</sup>。

#### 1.1.4 确定主梯度方向

根据已经检测到的特征点和该特征点的尺度值 $\sigma$ ,可以得到接近这一尺度值的高斯图像:

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

则可得到该特征点的梯度大小和方向如下:

$$m(x, y) = [(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2]^{1/2} \quad (3)$$

$$\theta(x, y) = \arctan \frac{L(x+1, y) - L(x-1, y)}{L(x, y+1) - L(x, y-1)} \quad (4)$$

其中: $m(x, y)$  和  $\theta(x, y)$  分别高斯金字塔  $(x, y)$  处的梯度大小和方向, $L$  所用的方向为所有特征点所在的尺度。

得到  $m(x, y)$  和  $\theta(x, y)$ ,以关键点为中心的邻域窗口内进行采样,然后用直方图的形式统计邻域像素的梯度方向。

梯度直方图一般取 10°为一个直方柱,范围为 0~360°,共 36 个柱。图 2 中,峰值为主方向,其余达到峰值 80% 以上的作为辅助方向。

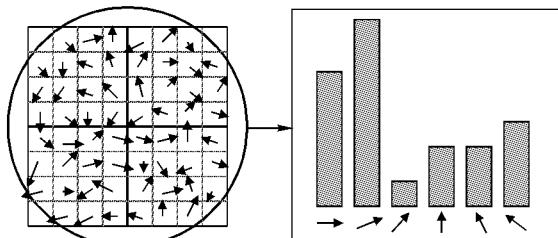


图 2 关键点邻域梯度统计

#### 1.1.5 生成特征描述符

关键点的描述通过统计其邻域内各点的梯度方向来构造,即以关键点为中心取  $16 \times 16$  的采样窗口,然后将该窗口分为  $4 \times 4$  个种子点,每个种子点对应一个  $4 \times 4$  的像素区域,将该采样窗口内的采样点与特征点的相对方向通过高斯加权后归入包含 8 个 bin 的方向直方图(每个方向为 45°),最后获得  $4 \times 4 \times 8$  的 128 维特征描述子。

### 1.2 SLAM 数据关联问题

SLAM 系统的数据关联问题是环境中的两个特征进行匹配,判断二者是否为实际环境中的同一个物体或点<sup>[8]</sup>。因此在 SLAM 系统中,数据关联的主要工作过程是:在地图库中对传感器当前检测到的特征进行匹配,若能够找到与之对应的关系,则将该特征融入 SLAM 系统以改善机器人位置和对特定路标的估计;若没有找到与之有对应关系的特征点,则判断该点是新特征点还是噪声点,如是新特征点则将其加入到地图库中,如是噪声点则将其屏蔽掉。

在 SLAM 问题的初期研究中,研究者们一般是将目标跟踪的数据关联技术应用于 SLAM 系统中,也取得了不错的效果。然而,随着移动机器人应用领域的不断扩展、应用环境的日益非结构化,简单的数据关联算法已经无法满足需求了。近年来,研究者们提出了一些适用于 SLAM 的数据关联算法<sup>[9-12]</sup>。常用的数据关联算法包括最近邻数据关联、概率数据关联、联合相容性数据关联等算法。这些算法大多是对当前观测值进行预处理,然后根据一定的规则(一般为马氏距离)选出一些候选点进行匹配。然而随着地图库的不断扩维,数据量不断增大,对算法的实时性有较大的影响。

## 2 数据关联算法的改进

本文使用视觉传感器实现 SLAM 系统,一般采用 SIFT 算法提取图像特征点信息,生成 128 维特征向量,匹配算法使用欧氏距离判定。这样在数据关联过程中,当前观测到的特征点需要与库中所有特征点进行矩阵相乘算法,设库中共有  $n$  个特征点,每个特征点有一个 128 维的描述向量,所以当前特征点需要与进行  $n \times 128$  次乘法运算,然后进行判断。随着  $n$  的不断增大,运算量将迅速增大,影响算法的实时性;本文从减少特征点数量和缩小当前路标与地图库中路标的关联范围两方面入手来改善算法实时性。本方法的基本思想如下:

1)从传感器获取的原始图像中提取特征明显(感兴趣区域)的特征点。利用边缘提取算法和灰度统计的方法将原始图像中边缘明显的物体或灰度与周边变化明显的区域提取出来,对这些区域进行特征提取。由于这些区域多为特征稳定、

随时间变化影响较小的特征点,因此既减小了无用点的数量又提高了匹配的准确性。

2) 将由步骤 1) 匹配成功的路标点位置信息与欧氏距离相结合进行关联。首先利用 SIFT 算法提取感兴趣区域的特征点,生成一个 128 维的特征点描述向量,利用摄像机成像原理得到特征点世界坐标;然后将这些信息加入到地图库矩阵中;最后通过位置信息以新采集到的特征点为中心设定一个圆形区域,区域内的点被认为是可能与新特征点匹配的点,区域外的点直接被屏蔽,这样就不需要将新特征点与库中的所有点进行欧氏距离匹配,减小算法的复杂度。

### 3 仿真实验分析

本文采用 Matlab 建立仿真环境,对本文算法进行仿真实验。实验中,人为设定好机器人的行进路径,路标(特征点)由算法在获取的环境图像中提取。

#### 3.1 特征点提取数量对比

本实验主要是提取图像特征点进行匹配,将匹配成功的特征点作为路标加入到地图库中。通过提取感兴趣区域的预处理,如图 3(a)、(b) 为处理前后提取特征点数量对比图,表 1 为特征点提取时间及数量对比。

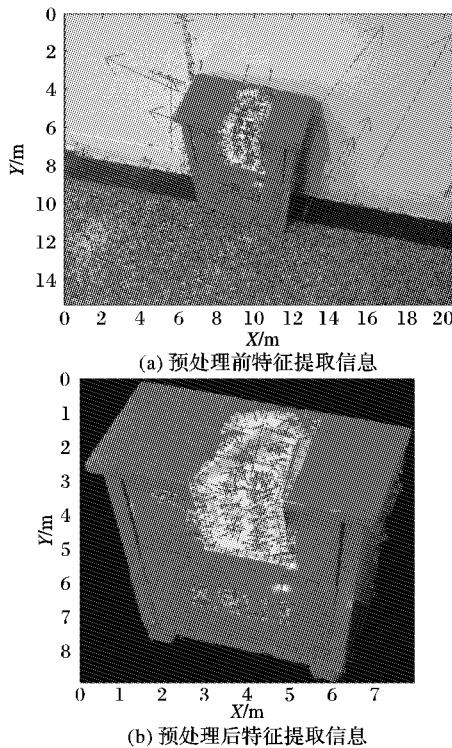


图 3 预处理前后提取对比

表 1 图 3(a)、(b) 特征点提取个数及时间

图例	特征点数量	提取时间/s
图 3(a)	9 650	26.582 977
图 3(b)	1 415	4.683 670

图 3 中箭头表示提取到的特征点及其特征。由图 3 和表 1 可知,预处理前图 3(a) 特征点数量(9 650)远远多于图 3(b)(1 415)且多出的部分多为大理石地面纹路上的点,这些点特征相近已发生误匹配;预处理后提取时间缩短很多。

#### 3.2 路标位置信息判定

判断由 3.1 节得到当前路标与地图库路标的位置关系,

图 4 为在  $100 \text{ m} \times 140 \text{ m}$  的区域内地图库中已有的路标及其位置信息。

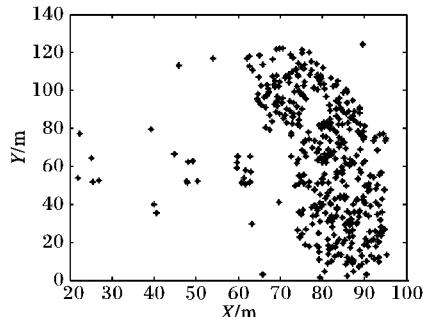


图 4 当前地图库路标位置信息

由图 5 可知,当前时刻机器人采集到的范围在  $0 \sim 40 \text{ m}$ ,其中圆形为该路标点的辐射范围,其半径为所选阈值,主要用来与地图库中的路标位置信息进行对比,看是否有库中的路标落在圆形区域内,若有,则认为可能是与之关联的点,将图 5 中的路标置于地图库中如图 6 所示。

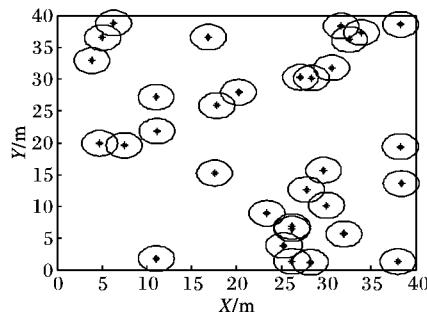


图 5 当前匹配成功的路标位置信息

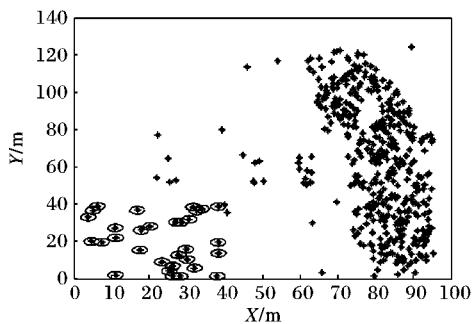


图 6 当前路标在地图库中示意图

由图 6 可知,当前时刻机器人所处位置为一个新的区域,与之前的位置不同,因此采集到的新路标从位置信息上即可大致判断出库中可能与之关联的路标,这样就省去了新路标与库中路标一一匹配的复杂过程。

#### 3.3 改进前后算法对比

本节主要通过关联时间和关联错误率两方面对改进前后的算法进行对比。

由图 7 可知,随着路标数的增加,两种算法的匹配时间都逐渐变大,改进后的算法在匹配时间上较改进前的算法有一定的减小;同时,改进后的算法的关联误差率也有一定的改进,如图 8 所示。

由图 8 可知,改进后的算法在匹配误差上有较大的改进。其原因主要在于:利用位置信息对路标进行预关联度判断的同时,算法将那些物理位置相差较远的路标直接屏蔽掉了,这就减小了那些不可能相关联的路标的匹配误差。

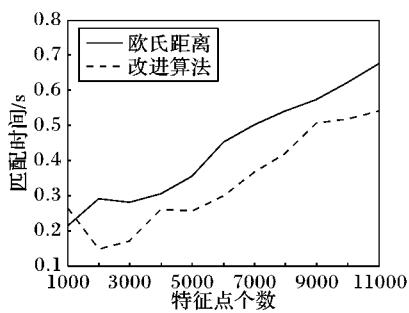


图7 两种算法关联时间对比

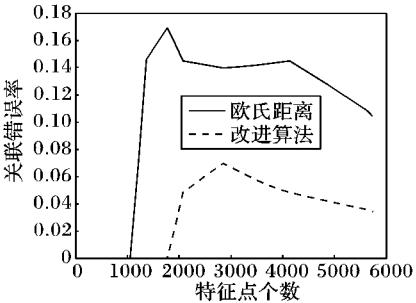


图8 两种算法的关联误差对比

#### 4 结语

本文通过仿真实验分析了视觉 SLAM 系统中数据关联的一种改进方法,该方法将路标位置信息与欧氏距离相结合的方式对地图库路标与当前采集到的图像匹配特征点进行数据关联,该方法在关联时间和关联准确性上都有一定的改善。但是随着机器人工作范围的扩大,机器人的自身位姿和路标位置的计算都将产生一些误差,这种误差可能会对路标位置的判断产生一些影响,造成一些点关联点不在圆形范围之内,错失一些匹配点。因此,机器人自身位姿与路标位置的精确计算是值得进一步研究的问题。

#### 参考文献:

- [1] LEONARD J J, DURRANT-WHYTE H F. Mobile robot localization by tracking geometric beacons[ J]. IEEE Transactions on Robotics and Automation, 1991, 7(3): 376 – 382.
- [2] TRIEBEL R, BURGARD W. Improving simultaneous mapping and localization in 3D using global constraints[ EB/OL]. [2013-06-20].
- [3] PENG L, SHI S, HONG X. A SLAM algorithm based on monocular vision and odometer[ J]. Computer Simulation, 2008, 25(10): 172 – 175. (王彭林,石守东,洪小伟. 基于单目视觉和里程计的SLAM 算法研究[J]. 计算机仿真, 2008, 25(10): 172 – 175.)
- [4] MIKOLAJCZYK K, SCHMID C. Performance evaluation of local descriptors[ J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615 – 1630.
- [5] KE Y, SUKTHANKAR R. PCA-SIFT a more distinctive representation for local image descriptors[ C]// Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC: IEEE Computer Society, 2004, 2: 506 – 513.
- [6] BICEGO M, LAGORIO A, GROSSO E, et al. On the use of SIFT features for face authentication[ C]// Proceedings of the 2006 Conference on Computer Vision and Pattern Recognition Workshop. Piscataway: IEEE, 2006: 35 – 35.
- [7] DAVISON A J, MURRAY D W. Simultaneous localization and map-building using active vision[ J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 865 – 880.
- [8] LOWE D G. Object recognition from local scale-invariant features [ C]// Proceedings of the 7th IEEE International Conference on Computer Vision. Piscataway: IEEE, 1999, 2: 1150 – 1157.
- [9] DURRANT-WHYTE H, RYE D, NEBOT E. Localization of autonomous guided vehicles[ M]. London: Springer, 1996: 613 – 625.
- [10] LINDEBERG T. Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention[ J]. International Journal of Computer Vision, 1993, 11(3): 283 – 318.
- [11] LINDEBERG T. Feature detection with automatic scale selection [ J]. International Journal of Computer Vision, 1998, 30(2): 79 – 116.
- [12] F Y, SONG B, WU J. An improved scale invariant feature transform algorithm [ J]. Journal of Harbin Engineering University, 2010, 31(5): 632 – 636. (付永庆,宋宝森,吴建芳. 边缘分类SIFT 算法[J]. 哈尔滨工程大学学报, 2010, 31(5): 632 – 636.)

(上接第 561 页)

- [6] HAN Y, WANG G, YANG Y. Speech emotion recognition based on MFCC [ J]. Journal of Chongqing University of Posts and Telecommunications: Natural Science, 2008, 20 (5): 597 – 602. (韩一,王国胤,杨勇. 基于MFCC 的语音情感识别[J]. 重庆邮电大学学报:自然科学版,2008,20(5):597 – 602.)
- [7] YUAN Z, XIAO W. Improved speech recognition algorithm based on MFCC feature[ J]. Computer Engineering and Applications, 2009, 45 (33): 108 – 110. (袁正午,肖王辉. 改进的混合MFCC 语音识别算法研究[J]. 计算机工程与应用, 2009, 45 (33): 108 – 110.)
- [8] KANEDERA N, ARAI T, HERMANSKY H. On the importance of various modulation frequencies for speech recognition [ C]// Proceedings of Eurospeech. Rodos, Greece:[ s. n. ], 1997, 9: 1079 – 1082.
- [9] ZHEN B, WU X, LIU Z, et al. On the importance of components of the MFCC in speech and speaker recognition [ J]. Journal of Peking University:Natural Sciences, 2001, 37 (5): 371 – 378. (甄斌,吴玺宏,刘志敏,等. 语音识别和说话人识别中各倒谱分量的相对重要性[J]. 北京大学学报:自然科学版, 2001, 37(5): 371 – 378.)

- [10] ZHANG Y, LI X, ZHENG Y, et al. Speaker recognition based on Fisher discriminant criterion [ J]. Journal of Lanzhou University: Natural Sciences, 2007, 43 (2): 92 – 97. (张芸,李昕,郑宇,等. 一种基于Fisher 准则的说话人识别方法研究[J]. 兰州大学学报:自然科学版, 2007, 43(2): 92 – 97.)
- [11] HU X, ZHANG X. Speaker recognition method based on Mel frequency cepstrum coefficient and inverted Mel frequency cepstrum coefficient[ J]. Journal of Computer Applications, 2012, 32(9): 2542 – 2544. (胡雪峰,张璇. 基于梅尔频率倒谱系数与翻转梅尔频率倒谱系数的说话人识别方法[J]. 计算机应用, 2012, 32(9): 2542 – 2544.)
- [12] HU Y, LOIZOU P C. Subjective evaluation and comparison of speech enhancement algorithms [ J]. Speech Communication, 2007, 49(7/8): 288 – 601.

[http://www.researchgate.net/publication/221604028\\_Improving\\_Simultaneous\\_Mapping\\_and\\_Localization\\_in\\_3D\\_Using\\_Global\\_Constraints/file/79e415062d878ccb61.pdf](http://www.researchgate.net/publication/221604028_Improving_Simultaneous_Mapping_and_Localization_in_3D_Using_Global_Constraints/file/79e415062d878ccb61.pdf).