

机器人视觉定位中的路口场景识别方法研究

高庆吉¹⁾ 李娟¹⁾ 马乐²⁾ 梁言贺²⁾

¹⁾(中国民航大学机器人研究所,天津 300300) ²⁾(东北电力大学自动化系,吉林 132022)

摘要 针对室外巡逻机器人的视觉定位问题,提出了一种基于尺度不变特性变换(scale invariant feature transform, SIFT)和颜色特征的路口场景识别方法。该方法首先提取路口场景图像的SIFT和颜色特征,并计算其在HSI颜色空间中的颜色直方图;然后采用K-D树和Bhattacharyya距离进行特征匹配;最终用决策公式对路口场景进行识别。为了提高SIFT算法进行场景匹配时的速度,还对场景地图库采用基于阈值分割的聚类方法进行了预处理。实验结果表明,该方法对环境光照变化、动态干扰和自身旋转有较强的鲁棒性,并能很好地识别出路口,以实现定位。

关键词 尺度不变特征变换 颜色直方图 聚类 Bhattacharyya距离

中图法分类号: TP242.62 文献标识码: A 文章编号: 1006-8961(2009)12-2510-07

Road Crossing Scene Recognition for Robot Vision_based Location

GAO Qing-ji¹⁾, LI Juan¹⁾, MA Le²⁾, LIANG Yan-he²⁾

¹⁾(Robot Institute, Civil Aviation University of China, Tianjin 300300)

²⁾(Department of Automation, Northeast Dianli University, Jilin 132022)

Abstract According to the problem of vision based localization for the outdoor patrol robot, an approach of road crossing scene recognition based on scale invariant feature transform (SIFT) and color features is proposed in this paper. Firstly, the SIFT features are extracted and the color histogram in HSI space is calculated. Secondly, the K-D trees algorithm is used to match SIFT features of images in road crossing images database, and the Bhattacharyya distance match result is calculated using color histogram. Finally, the SIFT features match result and the Bhattacharyya distance match result are combined together to confirm the suitable image in database. The image pre-classified idea is also adopted to accelerate the SIFT features matching. The experiment results demonstrate that the algorithm is robust to the various illumination, dynamic disturbance and self-circumrotating, and can be used for robot location.

Keywords scale invariant feature transform (SIFT), color histogram, clustering, Bhattacharyya distance

1 引言

目前在机器人视觉定位方法中,较为典型的方法有模型匹配法^[1]、路标识别法^[2]和场景识别法^[3]。

模型匹配法即根据先验的环境模型,通过将由传感器信息得到的模型与之比较来进行定位。由于

该方法需要精确的环境模型,其适用于静态环境,且由于数据量大、处理速度慢,因此,不适合应用到动态的室外环境中。

路标识别法是目前最为常用的定位方法,大致有以下两类思路:一类是先人为环境中加入具有明显特征的路标,然后根据路标的已知位姿信息,利用三角形法和卡尔曼滤波等进行定位,但定位的前提是路标能被很好地从场景图像中分割出来;另一

基金项目:国家自然科学基金项目(60776811)

收稿日期:2008-05-26;改回日期:2008-11-18

第一作者简介:高庆吉(1966~),男,教授。2006年获哈尔滨工业大学计算机应用技术专业博士学位。主要研究方向为机器视觉理论与技术、机器人导航控制。发表论文50余篇,其中20篇论文被EI收录。E-mail:gaoqingji@vip.sohu.com

类是利用已知的环境特征,对室内而言是指墙面直线、墙角等,对室外而言(主要是针对道路的检测)是指道路的中心线、边沿直线等特征进行定位。

场景识别方法与路标识别大体上是相同的,一般不需要通过对场景图像进行分割来提取其中的路标,而是对整个场景图像(可能含有多种自然路标)先进行特征提取、匹配,进而确定机器人的位置,以实现逻辑上的定位。

无论是采用路标识别方法,还是场景识别的方法,都需对图像的特征进行提取,常见的特征有颜色、纹理、形状和特征点等。其中前三者都是全局特征。

另外,基于局部不变量描述子的方法在目标识别和匹配方面取得了显著的效果。Mikolajczyk 和 Schmid 针对不同的场景,对光照变化、图像几何变形、分辨率差异、旋转、模糊和图像压缩等 6 种情况,采用多种最具代表性的描述子,如尺度不变特性变换(scale invariant feature transform, SIFT)、方向滤波器、矩、互相关等 10 种描述子进行了实验和性能比较,结果表明,在以上各种情况下, SIFT 描述子的性能最好^[4]。之后, Ke 等人和 Abdel-Hakim 等人分别提出了 SIFT 的改进算法 PCA-SIFT^[5] 和 CSIFT^[6],前者将其描述子部分,用 PCA 代替直方图的方式进行改进;后者在 RGB 颜色空间上,对其进行了改进。实验证明,这两种方法虽然减少了 SIFT 的计算量,但实际效果并不如 SIFT。基于 SIFT 描述子的匹配方法已被成功地应用到了很多领域,如目标识别^[7]和全景图拼接^[8]等。

室外巡逻机器人在对小区非结构化的已知环境进行巡逻时,路口是其必经之处。由于路口包含丰富的场景(楼房、路灯和树等自然路标)信息,因此可以通过对路口进行识别来实现机器人定位。本文首次提出了一种基于 SIFT 和颜色特征的路口场景识别方法,该方法先提取路口场景图像的尺度不变特征和颜色特征,并计算其所在的 HSI 颜色空间中 H 分量所对应的颜色直方图;然后采用 K-D 树和巴氏(Bhatta charyya)距离进行特征匹配;接着用阈值分割方法对场景地图库进行聚类,以提高场景匹配的速度;最后实现机器人的实时定位。针对光照变化、动态干扰和自身旋转这 3 方面的影响,在保安巡逻机器人上进行实验,验证了该方法在机器人定位中具有较强的鲁棒性。

2 场景特征提取与匹配

2.1 SIFT 特征提取与匹配

SIFT 是 Lowe 等人在 1999 年提出的一种尺度不变特征变换^[9],采用这种变换提取的特征被称为尺度不变特征,并在 2004 年完善^[7]。SIFT 特征对图像的比例缩放、旋转、3 维视角、噪声和光强的变化具有较好的不变性。

SIFT 特征是指对图像进行不同程度的模糊与缩放,首先产生具有不同比例的图像,然后从这些图像中分别提取特征。它是连续 3 幅高斯差分图像中不处于图像中边缘附近的极点。由于高斯差分图像中的极点是灰度值比它周围像素点都大或者都小的像素点,是在图像进行高斯模糊后变化最剧烈的像素点,因此具有很好的稳定性。也就是说, SIFT 特征点是从极点中选取出来的,满足一定条件的点。

SIFT 主要分为以下 4 步:(1)检测尺度空间极值点;(2)精确定位极值点;(3)为每个关键点指定方向参数;(4)关键点描述子的生成。描述子的构造过程为:先对任意一个关键点,在其所在的尺度空间,取以关键点为中心的 16×16 pixels 大小的邻域,再将此邻域均匀地分为 4×4 个子区域(子区域大小为 4×4 pixels);最后对每个子区域计算梯度方向(8 方向)直方图。这样就构成了一个 $4 \times 4 \times 8 = 128$ 维的向量,该向量就是 SIFT 描述子。

原始图像及 SIFT 特征点图如图 1 所示,其高斯金字塔分解的部分图片如图 2 所示,金字塔为 6 组,每组有 6 层($O = 0, 1, \dots, 5, S = 0, 1, \dots, 5$)。当两幅图像的 SIFT 特征向量生成后,即可以采用关键点特征向量(128 维)的欧氏距离来作为两幅图像中的关键点相似性的判定度量。特征提取匹配时,首先取当前观测到的图像 I_1 中的某个关键点;然后找出其与场景地图库中图像 I_2 的欧氏距离最近的两个关

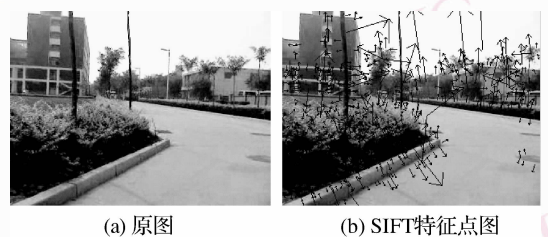
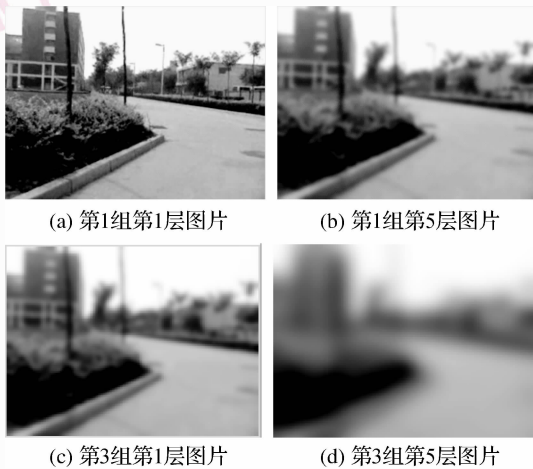


图 1 原图及 SIFT 特征点图

Fig. 1 Original image and SIFT keypoints image

键点,在这两个关键点中,如果最近距离与次近距离的比值小于某个阈值,则接受这一对匹配点。上述过程中,算法的时间复杂度为 $o(N)$,其中 N 为图像 I_2 中特征点的个数。



(a) 第1组第1层图片 (b) 第1组第5层图片

(c) 第3组第1层图片 (d) 第3组第5层图片

图 2 高斯金字塔分解的部分图片

Fig. 2 Some Gaussian images of scale space

本文采用文献[9]提到的 K-D 树进行匹配,其时间复杂度变为 $o(\ln N)$ 。Andrew 在文献[10]中提出基于 K-D 树的最近邻算法。K-D 树是一种存储 k 维空间点集的数据结构,它是一二叉树。本文是采用最邻近点搜索算法和深度优先的启发式搜索策略来进行匹配,并用一种超矩形数据结构来描述 k 维空间采样点集合所占据的空间范围。

首先对场景地图库中的图像 I_2 的所有 SIFT 特征点建立一棵 K-D 树,并对当前观测到的图像 I_1 中的每个 SIFT 特征点 k_{SIFT} 与 K-D 树进行最近邻算法操作;然后找到特征点 k_{SIFT} 的最邻近点 a 和次邻近点 b ,以及最近距离 dis_a 与次近距离 dis_b 。如果 $\text{dis}_a < \text{dis}_b \times \lambda, 0 \leq \lambda \leq 1$,则认为点 a 为匹配点。如果降低这个比例阈值,则 SIFT 匹配点数目会减少,但匹配结果会更加稳定。

SIFT 的不足之处是,它处理的是灰度图像,没有用到颜色信息^[11]。然而在场景匹配过程中,由于匹配数目相近,难以做出决策,所以针对这个问题,本文通过引入颜色特征来增强算法的识别能力。

2.2 颜色特征提取与匹配

路口所在的场景(楼房、路灯和树等这些自然路标)含有丰富的颜色信息,而这些颜色信息可以用颜色直方图的形式表达。由于颜色直方图对图像旋转、缩放和变形等具有不变性,因此被广泛地运

用。它的横轴表示颜色等级,纵轴表示在某一个颜色等级上所具有的该颜色的像素数目,两个图像的相似性可以通过各自的颜色直方图,采用欧氏距离、直方图求交等传统方法来得到。但后来被 Bhattacharyya 系数的方法所取代^[12]。

在众多的颜色空间中,HSI (hue, saturation, intensity) 颜色空间最能满足人的视觉感受。而 RGB 颜色空间则更易被机器接受。在 RGB 颜色空间中的某个像素的 R,G,B 分量的灰度值,经下面的公式^[13]变换后,可得到在 HSI 颜色空间中该像素的 H,S,I 分量值。

$$H = \cos^{-1} \left\{ \frac{(R-G) + (R-B)}{2 \sqrt{(R-G)^2 + (R-B)(G-B)}} \right\} \quad (1)$$

$R \neq G$ 或 $R \neq B$

若 $B > G$, 则 $H = 2\pi - H$ (2)

$$S = \max(R, G, B) - \min(R, G, B) \quad (3)$$

$$I = (R + G + B) / 3 \quad (4)$$

在 HSI 颜色空间中, H 分量表示颜色信息。颜色特征提取时,首先提取当前场景和场景地图库中图片的 H 分量的 ID 直方图,用于进行特征匹配;然后通过计算它们之间的 Bhattacharyya 距离来判断两幅图像的相似性。已知两个颜色的直方图分别为 $P = \{p_i\}_{i=1,2,\dots,k}$ 和 $Q = \{q_i\}_{i=1,2,\dots,k}$, 它们的 Bhattacharyya 距离^[14]为

$$d(P, Q) = \sqrt{1 - \sum_i \frac{p_i \times q_i}{\sqrt{p_i \times q_i}}} \quad (5)$$

其中, $\sum_i p_i = 1, \sum_i q_i = 1$ 。若 P 和 Q 为同一颜色的直方图,则 $d(P, Q) = 0$ 。

3 场景判断

3.1 场景地图库的创建及聚类

在创建场景地图库时,首先将机器人停放到路口的路中线的交点处,待图像稳定后抓取图片,使其按顺时针方向每隔 30° 抓一次图片,图片分辨率为 320×240 ;然后依此方法创建相应路口的场景地图库,它包括场景图片和与场景图片对应的特征文件(SIFT 特征向量的 txt 文件)。由于 SIFT 的计算量很大、运行速度慢,因此为节省 SIFT 场景的匹配时间,可先对场景地图库采用阈值分割的方法进行聚类。因为场景地图库的聚类过程属于离线操作,算法时间复杂度要求不高,聚类时间不受限制,且事先

并不知道要分几类聚类中心合适,故可采用阈值分割的方法进行聚类。在实际的操作过程中,要对场景地图库中与场景图片对应的 n 个特征文件 $Z = \{Z_1, Z_2, Z_3, \dots, Z_n\}$ 进行聚类,聚类中心集合记为 $C = \{C_1, C_2, \dots, C_k\}$, k 是聚类中心的个数,聚类方案如下:

- (1) 初始化 $k = 1, i = 1$ 。对于 $Z = \{Z_1, Z_2, Z_3, \dots, Z_n\}$, 可取样本 Z_i 作为第 1 类的聚类中心, 令 $C_k = Z_i$ 。
- (2) 令 $i = i + 1$, 计算 Z_i 与 $C = \{C_1, C_2, \dots, C_k\}$ 的 SIFT 的匹配点数目集 $M = \{M_{1,i}, M_{2,i}, \dots, M_{k,i}\}$ 。令 $j = \arg \max_{m=1,2,\dots,k} (M_{m,i})$, 若 $M_{m,i} \geq 5$, 则 Z_i 属于以 C_j 为中心的那一类, 否则 Z_i 为新的类中心 C_{k+1} , 令 $k = k + 1$ 。
- (3) 转到步骤(2), 直到对全体样本做完处理。

3.2 路口场景识别框架

图 3 中的虚线框处代表场景地图库的创建和聚类过程。当摄像机捕捉到当前场景图像后, 即可先提取其 SIFT 特征和颜色特征, 然后将其与聚类后的场景地图库进行特征匹配, 再经决策公式判断出当前场景所在的路口, 即可实现机器人的定位。

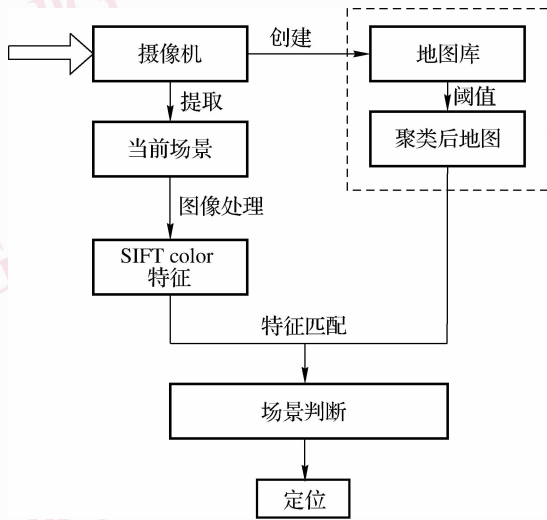


图 3 路口场景识别框架图

Fig. 3 The Frame of crossing scene recognition

具体步骤如下:

- (1) 初始化 $n = 0$, 清空链表 List。
- (2) 提取摄像机的第 n 帧图像, 并对第 n 帧图像进行 SIFT 特征提取。
- (3) 依次读取聚类中心 $C = \{C_1, C_2, C_3, \dots, C_n\}$, 计算 n 与第 i 个聚类中心 C_i 的 SIFT 的匹配点

数目 $M_{I,i}$ 。

(4) 若 $M_{I,i} \geq N_0$, 则提取 C_i , 并将其存放到链表 List 中。若所有 $M_{I,i} < N_0$, 则令 $n = n + 1$, 转到步骤(2)。

(5) 依次读取 C_i 中的样本 $X = \{x_1, x_2, \dots, x_k\}$ 。计算 n 与 X_j 的 SIFT 的匹配点数目 $M_{I,j}$ 。若 $M_{I,j} \geq N_1$, 则提取 x_j 存放到链表 List 中。

(6) 将链表 List 中所有的 C_i 和 x_j 按照 SIFT 匹配点数目由大到小排序, 若链表 List 中只存在 1 个满足条件的样本, 则直接判断其所在的路口, 并进行定位。反之, 取链表 List 中的前 N_2 个样本集 X 和聚类中心集合 C 的匹配点数的最大值 $P_{\max}^{cs} = \{P_1^{cs}, P_2^{cs}, \dots, P_{N_2}^{cs}\}$ 。

(7) 令 $F = \{f_1, f_2, \dots, f_{N_2}\}$ 是与 P_{\max}^{cs} 对应的图片集合。先将 n 与 F 从 RGB 颜色空间转换成 HSI 颜色空间, 并提取 H 分量; 然后分别计算出 H 分量的颜色直方图 R 与 $Q = \{Q_1, Q_2, \dots, Q_{N_2}\}$ 。

(8) 根据式(5), 计算出 R 与 Q_i 的相似度 S , 其计算式如下:

$$S = 1 - d(R, Q_i) = 1 - \sqrt{1 - \sum_j \frac{k}{j} \sqrt{p_j q_j}} \quad (6)$$

(9) 经式(7)决策后, 即可判断出机器人当前位于哪个路口。

由上可知, 颜色特征的加入并未影响算法的时间复杂度。在步骤(6)中, 先将满足条件的 SIFT 特征样本存到链表 List 中, 然后取前 N_2 个, 把 P_{\max}^{cs} 映射到图片集 F 中, 并提取图片集的相关属性(如表 1 所示)。

表 1 图片集的相关属性

Tab. 1 Some attribute of picture-set

图片编号	f_1	f_2	...	f_{N_2}
SIFT 特征匹配点的数目	$M_{I,1}$	$M_{I,2}$...	M_{I,N_2}
Bhattacharyya 距离	d_1^{Bh}	d_2^{Bh}	...	$d_{N_2}^{Bh}$
所在路口	1	3	...	4

由属性表得到的决策公式为

$$Des = \max \left\{ \alpha \frac{M_{I,1}}{M_{\max}} + \beta \frac{d_1^{Bh}}{d_{\max}^{Bh}}, \alpha \frac{M_{I,2}}{M_{\max}} + \beta \frac{d_2^{Bh}}{d_{\max}^{Bh}}, \dots, \alpha \frac{M_{I,N_2}}{M_{\max}} + \beta \frac{d_{N_2}^{Bh}}{d_{\max}^{Bh}} \right\} \quad (7)$$

其中, α 和 β 是决策权重因子, 并且 $\alpha + \beta = 1$, $M_{\max} = \max \{M_{I,1}, M_{I,2}, \dots, M_{I,N_2}\}$, Des 的值所对应的路口序号即为机器人的当前位置。

4 实验与结果分析

该算法在课题组研发的保安巡逻机器人(如图 4所示)系统中进行了实验,实验环境为作者所在的校园。在 Windows 平台下,用 VC 6.0 开发环境完成了算法实现。



图 4 保安巡逻机器人

Fig. 4 Patrol robot

在实验过程中,选取了有代表性的两个十字型路口和一个 T 字型路口。路口场景地图库中的场景图片为 36 张,聚类算法的阈值选为 5,聚类中心为 16 个。 α 选为 0.5, β 选为 0.5。实验中,高斯金字塔的组数 O 取为 6,层数 L 取为 6。 λ 选取为 0.32。 N_0 选为 5, N_1 选为 10, N_2 选为 3。文中 1-5 代表路口 1 中编号为 5 的场景图片。本文从环境光照变化、动态干扰(汽车、行人)和角度旋转影响这 3 个方面进行了实验,结果表明,本文算法具有鲁棒性和稳定性。

4.1 光照变化影响

光照变化对路口识别的影响如图 5 所示,图 5(a)和图 5(b)两幅图像的信息对比如表 2 所示,图 5(c)和图 5(d)两幅图像的信息对比如表 3 所示,表 4 是场景地图库中满足当前图 5(c)和图 5(d)两场景的前 3 个图片的信息表。

表 2 图 5(a)和图 5(b)信息对比表

Tab. 2 Information contrast for Pic. (a) and Pic. (b)

图像编号	图 5(a)	图 5(b)
当前帧号	2	2
SIFT 特征点个数	440	542
匹配点个数	213	37
相似度	0.91	0.67
决策判断	路口 1	路口 1
满足场景的图片个数	1	1

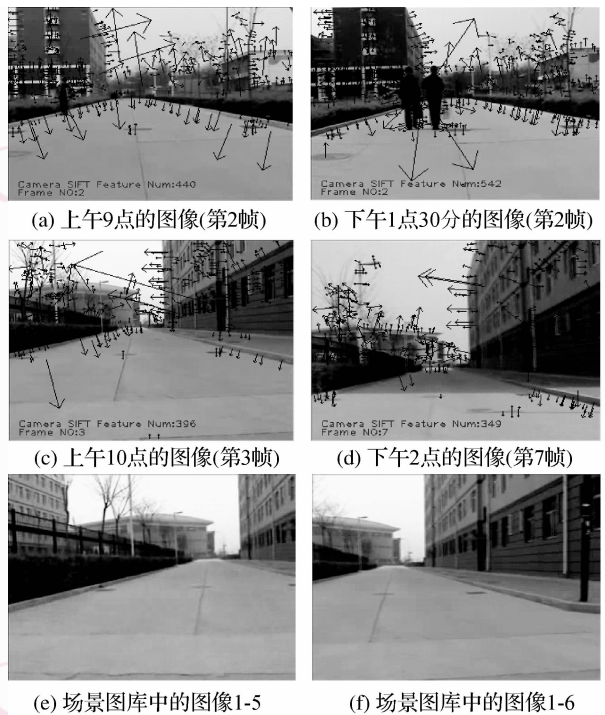


图 5 光照影响

Fig. 5 Influence of illumination

表 3 图 5(c)和图 5(d)信息对比表

Tab. 3 Information contrast for Pic. (c) and Pic. (d)

图像编号	图 5(c)	图 5(d)
当前帧号	3	7
SIFT 特征点个数	396	349
匹配点个数	79	16
相似度	0.84	0.50
决策判断	路口 1	路口 1
满足场景的图片个数	4	4

表 4 场景图片信息表

Tab. 4 Information of scene image

图像编号	图库中的 图片编号	匹配点个数	相似度	决策值
图 5(c)	1-5	81	0.76	0.95
	1-6	79	0.84	0.98
	1-1	60	0.72	0.79
图 5(d)	1-6	16	0.50	0.98
	1-5	15	0.52	0.96
	1-1	9	0.51	0.77

由表 2 知,场景受光照变化影响时,SIFT 匹配点的个数和相似度变化差异虽然很大,但仍可判断出机器人在路口 1。由表 3 知,在地图库中满足当前场景的图像个数是 4 个,由于 N_2 的取值为 3,故在表 4 的信息详情表中,只对比了 3 个图片。在图 5(c)

和图 5(d) 场景中,地图场景库中的图片 1-5, 图片 1-6 的匹配点的个数很相近, 仅用单一特征很难决策, 但结合相似度后, 即可判断出场景图 5(c) 和图 5(d) 与图 5(f) 最相近。

4.2 动态干扰影响

4.2.1 汽车影响

行驶车辆对路口识别的影响如图 6 所示:

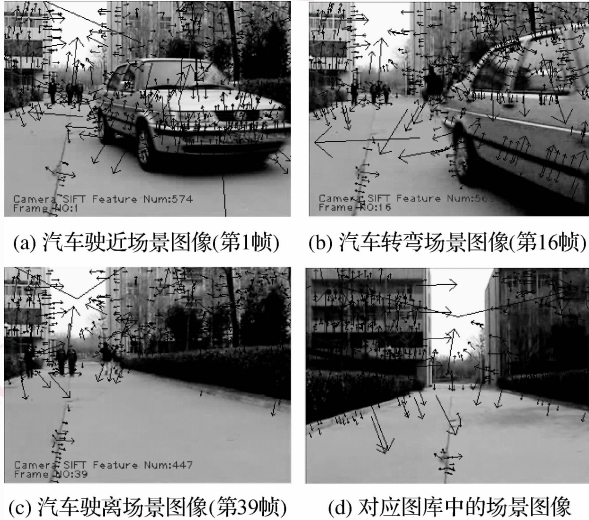


图 6 汽车影响
Fig. 6 Influence of car

图 6(a) 是机器人到达路口 2 后, 当前场景的第 1 帧图像, 此时汽车离机器人已经很近; 图 6(b) 是当前场景的第 16 帧图像, 汽车正在转弯; 图 6(c) 是汽车转弯后的第 39 帧图像; 图 6(d) 是创建场景地图库时的当前场景图片, 它的特征点数目是 445 个。图 6(a) ~ 图 6(c) 3 幅图像的信息对比如表 5 所示。由表 5 可见, 机器人对当前路口的识别不受行驶汽车的影响, 经判断都在路口 2。

表 5 汽车影响信息对比表

Tab. 5 Information contrast for influence of car

图片编号	图 6(a)	图 6(b)	图 6(c)
当前帧号	1	16	39
SIFT 特征点个数	574	569	447
匹配点个数	54	47	125
相似度	0.74	0.70	0.87
决策判断	路口 2	路口 2	路口 2
满足场景的图片个数	3	3	3

4.2.2 行人影响

对路口识别的行人影响如图 7 所示。图 7(a) 和图 7(b) 两幅图像的信息对比如表 6 所示。两次判断都在路口 2, 可见机器人对当前路口的识别不

受行人干扰影响。

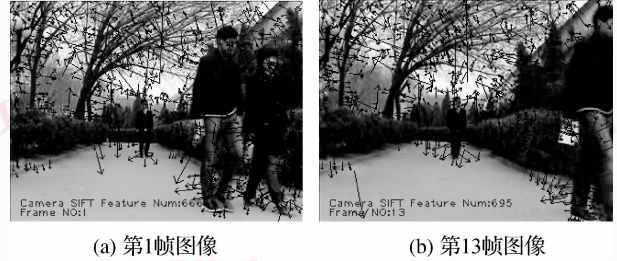


图 7 行人影响
Fig. 7 Influence of passer

表 6 行人影响信息对比表

Tab. 6 Information contrast for influence of passer

图片编号	图 7(a)	图 7(b)
当前帧号	1	13
SIFT 特征点个数	666	695
匹配点个数	166	153
相似度	0.73	0.75
决策判断	路口 2	路口 2
满足场景的图片个数	1	1

4.3 角度旋转影响

角度旋转时对路口识别的影响如图 8 所示:

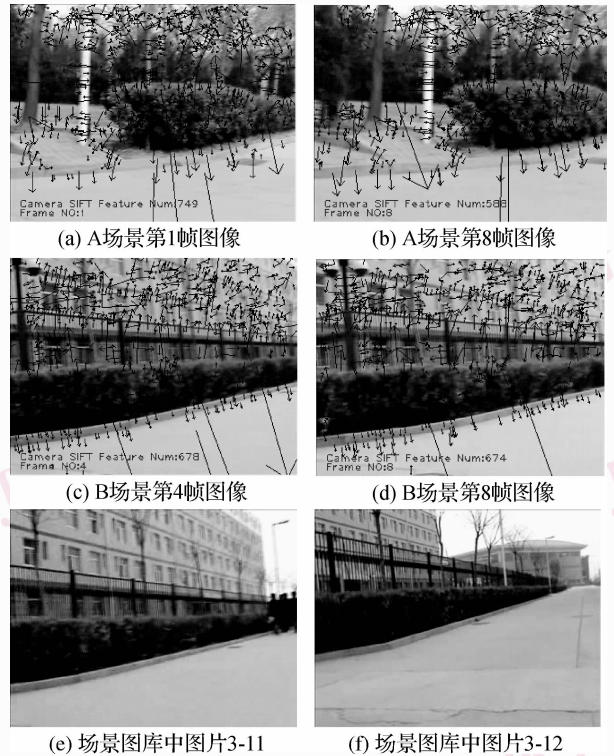


图 8 角度旋转影响

Fig. 8 Influence of circumrotating

图 8(b)是当前场景第 8 帧,机器人顺时针旋转了 26° ,图 8(a)和图 8(b)的信息对比如表 7 所示。图 8(d)机器人顺时针旋转了 12° ,图 8(c)和图 8(d)的信息对比如表 8 所示。由表 7 知,机器人对路口的识别不受角度旋转影响。由表 8 知,在匹配点数目相似的情况下,颜色直方图的相似性判断起到了很大作用。

表 7 图 8(a)和图 8(b)信息对比表

Tab. 7 Information contrast for Pic(a) and Pic(b)

图片编号	图 8(a)	图 8(b)
当前帧号	1	8
SIFT 特征点个数	749	588
匹配点个数	312	388
相似度	0.93	0.95
决策判断	路口 3	路口 3
满足场景的图片个数	2	2

表 8 图 8(c)和图 8(d)信息对比表

Tab. 8 Information contrast for Pic. (c) and Pic. (d)

图片编号	图片库中 图片编号	匹配点个数	相似度	决策值
图 8(c)	3-11	6	0.72	1.0
	1-1	5	0.57	0.70
	3-12	5	0.43	0.61
图 8(d)	1-4	5	0.63	0.92
	2-8	5	0.61	0.91
	3-11	5	0.74	1.0

5 结 论

本文提出了一种进行机器人视觉定位的路口场景识别方法,该方法结合了 SIFT 特征和颜色特征(颜色特征的加入,不会提高算法的时间复杂度),并对场景库进行了聚类。当满足场景个数的 SIFT 特征匹配点数目相近时,则颜色直方图的相似性判断起到了一定作用。聚类后的算法可较好地提高场景的匹配速度。实验结果表明,本文算法对光照变化、动态干扰和角度旋转变化都有很好的鲁棒性。如何进一步减少其算法识别路口场景的时间复杂度和提高颜色特征所起的作用,将是本文继续研究的问题。

参考文献 (References)

- Grimson W E L, Lozano-perez T. Model-based recognition and localization from sparse range or tactile data[J]. *International Journal of Robotics Res*, 1984, **3**(3):3-35.
- Betke M, Gurdits L. Mobile robot localization using landmarks[J]. *IEEE Transactions on Robotics and Automation*, 1997, **13**(2):251-263.
- Li Gui-zhi, An Cheng-wan. Scene recognition for mobile robot localization [J]. *ROBOT*, 2005, **27**(2):28-32. [李桂芝,安成万. 基于场景识别的移动机器人定位方法研究[J]. *机器人*, 2005, **27**(2):28-32.]
- Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10):1615-1630.
- Ke Y, Sukthankar R. PCA-SIFT: A more distinctive representation for local image descriptors [A]. In: *Proceedings of Conference on Computer Vision and Pattern Recognition [C]*, Washington, DC, USA, 2004:506-513.
- Abdel-hakim A E, Farag A A. CSIFT: A SIFT descriptor with color invariant characteristics [A]. In: *Proceedings of Conference on Computer Vision and Pattern Recognition [C]*, New York, DC, USA, 2004: 506-513.
- Lowe D G. Distinctive image features from scale-invariant keypoints [J]. *International Journal of Computer Vision*. 2004, **60**(2):91-110.
- Bromm M, Lowe D G. Recognising panoramas[A]. In: *Proceedings of the 9th International Conference on Computer Vision [C]*, Nice, France, 2003:1218-1225.
- Lowe D G. Object recognition from local scale-invariant features[A]. In: *Proceedings of International Conference on Computer Vision [C]*, Corfu, Greece, 1999:1150-1157.
- Moore A W. An introductory tutorial on KD-Trees[R]. Technical Report No. 209, Computer Laboratory, University of Cambridge, London, 1991:1-20.
- Ancuti C, Bekaert P. SIFT-CCH: Increasing the SIFT distinctness by color co-occurrence histograms [A]. In: *Proceedings of the 5th International Symposium on Image and Signal Processing and Analysis [C]*, Istanbul, Turkey, 2007:130-135.
- Perez P, Hue C, Vermaak J, et al. Color-based probabilistic tracking [A]. In: *European Conference on Computer Vision [C]*, London, U K, 2002:661-675.
- Gonzalez R C, Woods R E. *Digital Image Processing [M]*. 3rd Edition. Addison-wesley, 1992.
- Thacker N A, Aherne F J, Rockett R E. The Bhattacharyya metric as an absolute similarity measure for frequency coded data [J]. *Kybernetika*, 1998, **34**(4):363-368.