旋转尺度不变的实时高精度场景匹配算法

姚 $\overline{\mathbf{h}}^1$, 贾克斌¹, 萧允治²

(1. 北京工业大学电子信息与控制工程学院,北京 100124; 2. 香港理工大学电子与信息工程系,香港 999077)

摘 要:针对场景匹配技术中以二进制稳健基元独立特征(binary robust independent elementary features ,BRIEF)为 代表的实时算法匹配错误率高的问题,提出一种基于局部梯度二值化的特征描述算法.该算法利用重心向量方向 归一化特征描述区域,保证了特征描述符的方向不变性.同时,融合基于局部梯度二值化的区域纹理信息以降低特 征匹配错误率.使用国际通用数据库对算法进行了验证,实验结果表明:提出的场景匹配算法其平均匹配准确率比 BRIEF 算法高 44.59% 具有较高的鲁棒性.

关键词:场景识别;特征描述;特征匹配;梯度二值化
 中图分类号:TP 391.4
 文献标志码:A
 doi: 10.11936/bjutxb2016020007

文章编号: 0254-0037(2016)11-1634-09

Orientation and Scale Invariant Scene Matching With High Speed and Performance

YAO Meng1 , JIA Kebin1 , SIU Wanchi2

(1. College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing 100124, China;2. Department of Electronic and Information Engineering, Hong Kong Polytechnic University, Hong Kong 999077, China)

Abstract: For real-time scene feature extraction and matching , conventional binary descriptors improve the speed of the descriptor generation and matching procedure while the false matching rate is high , such as Binary Robust Independent Elementary Features (BRIEF) which is only based on pixel intensity comparisons. To solve this problem , an improved binary descriptor was proposed in this paper , which preserved not only the pixel intensity information , but also the local texture information based on the gradient value. Additionally , the orientation of the centroid vector was also used in the descriptor calculation process , so that the binary descriptors were orientation-invariant. Image Sequences dataset was used to evaluate the performance of the proposed method , and the average matching accuracy rate of the proposed method was 44. 59% , higher than that of the BRIEF algorithm. Experimental results show that the proposed descriptors have high accuracy and robustness when dealing with image rotation and scale transformation.

Key words: scene recognition; feature description; feature matching; gradient binarization

收稿日期: 2016-02-18

基金项目:北京市自然科学基金暨教委重点科技项目(KZ201310005004)

作者简介:姚 萌(1988—),男,博士研究生,主要从事图像特征描述、基于视觉的移动机器人定位技术方面的研究, Email: yaomeng@ emails. bjut. edu. cn

通信作者: 贾克斌(1962—),男,教授,主要从事图像/视频内容检索、视频编码及处理技术方面的研究, E-mail: kebinj@ bjut. edu. cn

场景图像的匹配与识别作为视觉定位技术的基础,被广泛应用于室外车辆辅助定位系统以及室内机器人导航系统,在交通、救灾抢险和勘探等领域也有着广泛的应用^[1-2].该类实时系统往往对算法性能要求较为苛刻,在保证场景匹配精度以及识别准确率的同时,算法需要具有较高的运算速度.图像特征点匹配作为场景匹配方法之一,通过匹配当前图像与候选图像的特征点从而识别当前场景,其计算复杂度低且可靠性较高,满足了系统对实时性与 准确性的需求^[3].

图像特征点匹配通常包括特征点检测、特征点 描述、描述符匹配以及匹配过滤4个步骤 近年来有 诸多算法被提出. Lowe 等^[4] 提出的尺度不变特征 变换(scale invariant feature transform ,SIFT) 算法具 有良好的尺度以及旋转不变性. Bay 等^[5] 为降低 SIFT 计算复杂度而提出加速鲁棒性特征(speeded up robust features SURF) 算法,虽然 SURF 算法提高 了特征点匹配运算速度并具有良好的鲁棒性,但仍 然不能满足实时系统的需求.为进一步降低计算复 杂度 ,Calonder 等^[6]提出二进制稳健基元独立特征 (binary robust independent elementary features , BRIEF) 大幅提升了图像特征描述和特征匹配计算 速度 满足了实时系统的速度需求. 由于 BRIEF 特 征不具有旋转不变性 ,Leutenegger 等^[7] 提出二进制 鲁棒不变尺度关键点(binary robust invariant scalable keypoints ,BRISK) 特征描述算法,使用长距离采样 点对计算特征点方向并归一化描述区域方向,使得 二值化特征描述符具有方向不变性. 2011年, Rublee 等^[8]提出的定向二进制简单描述符(oriented FAST and rotated BRIEF ,ORB) 算法通过机器学习获 得采样点对 进一步提高了特征点匹配运算速度以 及鲁棒性. Alahi 等^[9] 根据视网膜成像原理提出的 快速视网膜特征点(fast retina keypoint ,FREAK) 描 述符,在一定程度上实现了尺度旋转不变.此外, Gao 等^[10] 提出对 BRIEF 的改进算法局部三元描述 符(local ternary descriptor LTD) ,用3个值表示点对 的灰度值关系.这些二值化特征描述符被广泛用于 目标跟踪等实时系统中[11-3]. 然而,基于像素灰度 值关系获得的二值化描述符只记录了采样区域的灰 度信息而缺少纹理信息,匹配错误率较高.因此, Yang 等^[14] 提出局差分部二进制(local difference binary LDB) 描述符,该算法通过特征点周围的平均 梯度信息得到二值化描述符,但由于该算法不具有 尺度与方向不变性,对场景图像旋转、缩放等几何变

换鲁棒性较低.

本文针对实时场景识别系统中场景匹配问题, 提出一种尺度及旋转不变的实时局部梯度二值化特 征描述符(rotation and scale invariant local difference binary ,RSI-LDB).针对传统二值化描述符只计算灰 度信息而丢失采样区域纹理信息这一问题,该算法 首先在不同空间尺度图像序列中提取采样区域,并 使用重心向量方向归一化该区域,之后计算采样区 域二值化的梯度信息,最终得到具有尺度及旋转不 变性的图像特征描述符.实验结果表明该算法具有 较高的匹配准确率和鲁棒性.

1 二值化特征描述

由于传统的非二值化特征描述符采用欧式距离 进行匹配计算 因此 过高的算法复杂度成为该类描 述符提升匹配速度的瓶颈.二值化特征描述算法生 成的描述符使用汉明距离进行匹配计算 ,而汉明距 离可以通过向量之间进行逐位异或运算获得 ,因此 , 二值化特征描述符大大降低了距离计算复杂度 ,大 幅提高了特征匹配速度^[6].二值化特征描述符的典 型应用是 BRIEF 特征描述符.

BRIEF 特征描述算法首先使用高斯模糊核去除 特征点周围方形描述区域内的噪声干扰,之后在该 区域内按照事先定义的随机采样规则抽取若干点 对,用0与1表示每个点对中2个点之间的灰度大 小关系,最终生成二值化向量.定义 7 为特征点周 围 *S*×*S* 区域内点对 *p* 的灰度比较函数,则有

$$\tau(p; x_{\theta}) = \begin{cases} 1 , p(x) < p(y) \\ 0 , \notin \theta \end{cases}$$
(1)

式中: $x \cdot y$ 为点对 p 中的 2 个点; $p(x) \cdot p(y)$ 是点对 p 中 2 个点的灰度值. 按照采样规则反复随机采样 n_d 个点对 p 得到的计算结果组成 n_d 维二值化描述 符 $f_{a_d}(p)$ 月

$$f_{n_d}(p) = \sum_{1 \le i \le n_d} 2^{i-1} \tau(p; x \ y)$$
 (2)

随机采样规则包含 n_d 个点对坐标,这些点对坐标事先由随机方法生成,通过串联这些点对的灰度 值比较结果最终获得 BRIEF 描述符,并使用汉明距离进行特征点匹配,因此基于 BRIEF 的图像特征点匹配算法复杂度低且运算速度快.但是 BRIEF 描述符的采样区域没有归一化特征点方向因而不具有方向不变性,不适用于旋转情况下的图像匹配.另一方面,BRIEF 描述符仅仅在像素间进行灰度值比较, 无法保留采样区域的纹理信息,因而降低了描述符 的描述能力.

2 尺度与旋转不变二值化梯度描述符

2.1 方向归一化

为保证特征描述符具有方向不变性,需要计算 特征点主方向,并对描述符进行方向归一化.本文 使用重心向量方向作为特征点主方向,重心向量由 采样区域的几何中心指向灰度重心,记灰度重心位 置为(*G_x*,*G_y*)则水平分量*G_x*和垂直分量*G_y*计算分 别为

$$G_{x} = \frac{\sum_{x,y} xI(x,y)}{\sum_{x,y} I(x,y)}$$
(3)

$$G_{y} = \frac{\sum_{x,y} yI(x,y)}{\sum I(x,y)}$$
(4)

式中: *I*(*x*,*y*)为位于(*x*,*y*)位置像素的灰度值;(*x*, *y*)为当前像素与采样中心的相对坐标.重心向量如图1所示.



图 1 重心向量示意 Fig. 1 Centroid vector

记重心向量为 $\nu_e(G_x,G_y)$,则特征方向为该向量所指方向,记特征方向与横坐标方向夹角为

$$\theta = \arctan\left(\frac{G_{y}}{G_{x}}\right) \tag{5}$$

由方向夹角 θ 计算得到描述区域旋转矩阵

$$\boldsymbol{R} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$
(6)

旋转矩阵 R 与特征描述区域 S × S 内的像素位 置相乘得到旋转后的特征描述区域 ,有

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$$
(7)

式中:(x y)^T为原始采样区域像素位置;(x´y´)^T 为旋转后对应像素的位置. 经过方向归一化,所有 特征描述区域都旋转到了重心向量方向,使得描述 符对图像旋转具有一定的鲁棒性. 2.2 尺度归一化

为保证特征描述符具有尺度不变性,RSI-LDB 特征描述符在多尺度空间图像序列中计算. 多尺度 空间图像序列通常从图像金字塔(image pyramid)中 获得^[4].如图 2 所示,图像金字塔由若干幅尺寸不 一的图像组成,第0 层为原始图像,之后每层图像都 由前一层图像降采样获得.如原始图像乘以尺度因 子 1/σ 得到第1 层图像,由此可得,第 k 层图像的 尺度为 1/σ^k,最终得到的图像金字塔由 k + 1 幅不 同尺度的图像序列组成.根据特征点携带的尺度信 息,RSI-LDB 在图像金字塔对应的图层中计算特征 描述符.



在实际应用中,若特征点检测算法在多尺度图 像金字塔中进行,如 SIFT 或 ORB 等特征点检测算 法,则 RSI-LDB 的图像金字塔降采样尺度因子 1/σ 以及金字塔层数 k 应与特征点检测算法中对应参数 相同.

2.3 梯度二值化

经过旋转归一化与尺度归一化后的采样区域具 有方向及尺度不变性,该区域的纹理信息由梯度计 算获得. 与 LDB 描述符类似,本文的梯度采样将特 征点周围 *S*×*S* 区域划分成*n*×*n*个单元,如图 3 所 示,分别计算每个单元内部的平均灰度、水平梯度以 及垂直梯度,进而比较 2 个单元之间这 3 个分量的 大小关系并生成二值描述符.

参考 BRIEF 的二值化公式 修改得到

$$\tau(F(x),F(y)) = \begin{cases} 1, & F(x) > F(y) \\ 0, & \ddagger \psi \end{cases}$$
(8)

式中: x、y 分别为 2 个单元区域; F(·) 为单元内部 3 个分量.

平均灰度分量为采样单元区域内部所有像素点 的平均灰度,该分量反映了该单元的灰度信息,但是 平均灰度信息忽略了单元内部的亮度变化情况,因 此需要计算单元内部的梯度以保留灰度变化信息, 即纹理信息. 梯度包括水平梯度和垂直梯度两部 分,水平梯度由单元块内部右侧所有像素的平均灰 度与左侧所有像素的平均灰度之差计算得到; 垂直 梯度由单元内部下方所有像素的平均灰度与上方所 有像素的平均灰度之差计算得到.



Fig. 3 Gradient binarization

3 个分量的计算公式为

$$F_{\text{Intensity}} = \frac{1}{s} \sum_{k \in S} I(k)$$
 (9)

$$F_{\text{GradientX}} = \frac{2}{s} \sum_{k \in S_{\text{right}}} I(k) - \frac{2}{s} \sum_{k \in S_{\text{left}}} I(k)$$

$$(10)$$

$$F_{\text{GradientX}} = \frac{2}{s} \sum_{k \in S_{\text{right}}} I(k) - \frac{2}{s} \sum_{k \in S_{\text{left}}} I(k)$$

$$\mathcal{T}_{\text{Gradient}Y} = \frac{\mathcal{L}}{s} \sum_{k \in S_{\text{down}}} I(k) - \frac{\mathcal{L}}{s} \sum_{k \in S_{\text{up}}} I(k)$$
(11)

S 分别代表单元内部所有像素集合 *S*_(.) 代表指 定区域内的像素集合 ,如 *S*_{right}代表单元内部右侧所 有像素.将式(9)~(11)代入式(8),通过串联所有 单元对中 3 个分量的比较结果进而得到二值化的特 征描述向量.

3 实验结果与分析

3.1 实验环境与评价方法

为评估所提出的算法性能,本节在实现了其他 经典图像特征描述算法的基础上,使用通用数据库 测试了本文算法与其他图像特征描述算法,并对比 分析了匹配结果.实验中的测试数据库采用国际权 威数据库 Image Sequences^[15],包括 SURF 在内的多 种著名算法都使用了该数据库进行算法评估.如图 4 所示,数据库中包含 6 个场景(Boat、Bike、Trees、 Leuven、Ubc、Wall),涵盖了真实环境中常见的图像 变化以测试算法的鲁棒性.测试类型包括:图像旋 转与图像尺度变化,对应 Boat 场景;图像模糊,对应 Trees 与 Bikes 场景;光照变化,对应 Leuven 场景;图



图 4 Image Sequences 数据库 Fig. 4 Image Sequences dataset

像压缩,对应 Ubc 场景;视角变化,对应 Wall 场景.

数据库中的每个场景序列有6幅图像,对于每 个场景序列,本文使用第1幅图像分别与其余5幅 图像进行匹配.这5幅图像按照匹配的难易程度排 列:随着序号升高,Boat场景旋转与尺度变化越来越 剧烈;Wall场景的视角差异越来越明显;Trees与 Bike场景的图像模糊程度逐渐加强;Leuven与Ubc 场景的图像亮度变化以及质量变化越来越显著.每 个场景生成5个匹配对,例如1-2代表第1幅图像 与第2幅图像的匹配结果.

本文实现了 SIFT^[4]、SURF^[5]、BRIEF^[6]、 BRISK^[7]、FREAK^[9]、ORB^[8]、LDB^[14]以及 RSI-LDB 图像特征匹配算法 其中 SIFT 和 SURF 描述向量的 匹配基于欧氏距离,其余二值化特征描述向量基于 汉明距离. 首先使用不同算法生成每个场景中所有 图像的特征描述向量集合,并用暴力匹配方法得到 各图像对的匹配对集合,即为每个图像对中第1幅 图像的 N 个描述向量在第 2 幅图像的描述集合中 找到最小距离的邻近向量,将这个组合作为匹配对 保存,生成包含 N 个匹配对的集合. 由于数据库提 供了2张图像之间的标准变换矩阵 因此可以利用 该变换矩阵验证匹配结果. 记匹配对 M 中第1 幅图 像的特征点为 P_A ,第2 幅图像的特征点为 P_B 将 P_A 的坐标与标准变换矩阵相乘得到 P_A 在第 2 幅图像中对应的 P_A / 点 P_A / 点是 P_A 在第2 幅图 像中的实际位置 若 P A'与 P B 点之间的误差距离 在可接受范围内(本文为10像素),则可以认为这 个匹配对 M 为优匹配. 统计所有优匹配数量记为 n_c 将匹配正确率 n_c/N 作为衡量描述算法准确度 与鲁棒性的标准.

3.2 匹配正确率分析

图 5 展示了在测试数据库中,不同图像特征匹 配算法的匹配正确率.不同直方图显示了不同算法 在每个场景的匹配正确率,每组中的 5 个图像对分 别来自第 1 张图像与之后的 5 张图像的匹配统计结 果 统计结果包括了前文所述 7 种算法以及本文的 RSI-LDB 算法,其中 LDB 采用了 4 × 4 的分块方式, 即 LDB 16.直方图的横坐标为不同的图像对,纵坐 标为匹配正确率 n_c/N. 由图 5 可以看出,除 Wall 场景的 1-6 图像对之外,本文提出的 RSI-LDB 算法 相对于以 ORB 为代表的旋转不变二值化匹配算法 有着更高的匹配正确率.

3.2.1 尺度与旋转变化

Boat 场景中的图像发生了较明显的旋转变换和

尺度变换,以测试匹配算法对方向和尺度变化的鲁 棒性. 在本组测试中,相对于 ORB、BRISK、FREAK 等改进的二值化特征描述算法,本文提出的 RSI-LDB 算法具有更高的鲁棒性,其中 RSI-LDB 64 在 5 次匹配中表现出了优异的准确率.以1-6 图像对为 例,在图像发生较大旋转以及缩放的情况下,其他二 值描述算法的匹配正确率已经不足 7%,而 RSI-LDB 64 算法为 75.00%,其正确率比 BRIEF 算法高 74.94%,比 ORB 算法高 68.33%.

3.2.2 场景模糊变化

在 Bike 场景中,当 BRIEF 算法匹配正确率为0时,同组测试图像对中 RSI-LDB 16 匹配正确率保持在 81%以上,而 RSI-LDB 64 匹配正确率为 100%. 与 Bike 场景类似,在 Trees 场景中 RSI-LDB 64 的匹配正确率也远高于 BRIEF. 与 BRIEF 以及更成熟的 ORB 相比, RSI-LDB 64 的平均正确率分别高 16.10%和18.97%.

3.2.3 光照强度与图像压缩质量变化

在 Leuven 场景测试中,图像光照发生变化. 在 变化最大的 1-6 图像对中 RSI-LDB 64 匹配正确率 为 96.46%,比 BRIEF 算法高 8.78%,比 ORB 算法 高 12.95%. 在 Ubc 组测试中,当图像质量严重下降 时,如 1-6 图像对,RSI-LDB 64 匹配正确率为 94.71%,比 BRIEF 算法高 0.71%,比 ORB 算法高 4.26%.

3.2.4 视角变化

Wall 场景主要测试算法对视角变化的鲁棒性. 本文提出的 RSI-LDB 64 算法在 1-2~1-5 测试中表 现出明显优势,在这 4 组测试中其匹配正确率一直 保持在 98.43% 以上,只有在 1-6 图像对中,丢失了 匹配对象. 从平均匹配正确率分析,RSI-LDB 64 较鲁 棒性较强的 ORB 算法高 6.15%.

3.2.5 分块影响

由 RSI-LDB 16 与 RSI-LDB 64 的结果对比可以 看出 随着特征描述区域单元块划分数量上升 特征 描述符匹配正确率随之升高.当场景发生较大尺度 变化与旋转变化时,如图 5 中 Boat 场景所示,高密 度块划分方式能够显著提高特征匹配正确率.这种 高密度块划分方式提高了特征描述符的描述能力, 即特征描述符的分辨力,降低了误匹配率.另一方 面,Wall 场景的测试结果表明高分辨力的特征描述 符对投影变化较为敏感.在该测试场景的前4 组实 验中,RSI-LDB 64 仍旧保持了较高匹配正确率,而 在第5 组实验中,场景视角发生较大变化,该描述符



图 5 各场景中算法匹配正确率



表现弱于其他方法.因此,块划分参数应根据具体 应用调整设置.

3.3 鲁棒性分析

图 6 统计了 6 个场景中各算法的平均正确率和 正确率标准差.平均正确率反映了算法在某场景中 匹配正确率的总体表现,平均正确率高.代表该算法 在相应类型图像变化中具有较高的鲁棒性.正确率 标准差一定程度上反映了算法对某种类型图像变换 的敏感程度,在平均正确率相同的情况下,其标准差 越低,代表该算法在图像发生不同程度变换时稳定 性更高.

由图 6(a) 可以看出,本文提出的 RSI-LDB 相对 于 LDB 具有较高的旋转不变性和尺度不变性,与其 他算法相比,RSI-LDB 64 在增加了单元分块后具有 更高的匹配正确率. 由图 6(b) 可知,除 Wall 场景 外,RSI-LDB 64 的正确率标准差较小,说明本文提 出的方法对于图像的大尺度旋转、缩放、模糊、光照 变化以及图像降质等具有较高的鲁棒性,在各种图 像变化情况下匹配结果更稳定.

3.4 时间分析

通过图 5 以及表 1 可知 ,SIFT 以及 SURF 作为 经典的非二值描述算法 ,在图像发生旋转、模糊、尺 度以及视角变换的情况下均表现出了较好的匹配性 能 ,但是二者远远无法满足实时系统对速度的需求. 本文提出的 RSI-LDB 算法通过计算单元块平均灰 度与梯度生成二值化描述符,并通过汉明距离进行 特征描述向量匹配,大大提高了图像特征匹配计算 速度.实验结果表明:计算单个 RSI-LDB 描述符的 时间比 SIFT 算法节约 90.84%,比 SURF 算法节约 69.31%; RSI-LDB 特征匹配耗时比 SIFT 算法节约 82.82%,比 SURF 算法节约 65.82%.

表 1 SIFT、SURF、RSI – LDB 描述符时间性能对比 Table 1 Time performance comparison of SIFT, SURF and RSI-LDB

项目	SIFT/ms	SURF/ms	RSI-LDB/ms	Δt (RSI-LDB 与 SIFT)	Δt (RSI-LDB \models SURF)
单个描述符计算	0. 502 1	0. 149 9	0.0460	-90.84%	-69.31%
单个描述符匹配	1. 181 4	0. 594 0	0.203 0	- 82. 82%	-65.82%





Fig. 6 Average and standard deviation of percentage of correct matches

3.5 距离分布

为深入分析本文提出的 RSI-LDB 特征点匹配 的可靠性 本节统计了描述符集合在不同图像对之 间的汉明距离分布情况 ,数据来自每个场景第1幅 图像与第5幅图像(即1-5)的匹配结果.图7呈现 了归一化之后的匹配点对以及未匹配点对的出现频 率 蓝线代表已匹配点对的距离分布 红线代表未匹 配点对的距离分布 横坐标是归一化之后的汉明距 离 纵坐标是对应距离的匹配点对的出现频率. 理 想情况下 2 条曲线的峰值应该完全独立 意味着在 进行图像匹配过程中,通过选定一个适当的阈值可 以区分出2个特征点是否是匹配点对. 由图7可 知 在6个测试场景中,蓝线与红线的峰值在阈值 35 处被分割开来,说明 RSI-LDB 算法能够通过距离 很好地区分匹配点对以及未匹配点对. 在其他场景 匹配计算中,可以将阈值35作为是否是正确匹配的 重要依据之一.

3.6 算法应用实例

图像特征匹配算法在机器人视觉定位系统中有 着广泛应用.本文采用 RSI-LDB 对室外机器人采集 到的场景图像进行匹配,并通过匹配的点对计算场 景间的变换关系,由单应性矩阵(homography matrix)表示,该矩阵为后续算法估算室外机器人的 运动情况提供重要参数.

对于已有的场景图像以及当前场景图像,分别 生成 RSI-LDB 二值化特征描述向量集合,其中特征 点检测使用 ORB 算法中改进的 FAST^[16]的计算结 果,通过暴力匹配算法匹配 2 幅图像的特征描述向





量. 在匹配过程中为保证匹配质量降低误匹配率, 当且仅当最邻近距离与次邻近距离的比率小于 0.8 时,才能将此匹配保存在匹配集合*M* 中.

第11期

为进一步筛选出优质的匹配对,提高场景变换 矩阵的精确度,使用随机抽样一致性(random sample consensus,RANSAC)^[17]算法对所获得的匹 配集合进行过滤最终得到符合统一约束模型的匹配 点对,其中约束模型为单应性矩阵.经过RANSAC 筛选出的匹配对都满足同种单应性变换关系,对该 集合使用最小二乘法即可获得场景变换矩阵,从而 估算出摄像机运动向量.图8展示了RSI-LDB 64 描述符在室外机器人视觉定位系统中的应用,红线 两端为使用 RANSAC 算法筛选出符合统一变换约 束模型的匹配点对. 实验结果表明:本文算法在应 用实例中表现出良好的性能,在保证匹配正确率的 同时满足了系统对实时性的需求. 图9为通过匹配 结果计算出的单应性矩阵,该矩阵包含了2个场景 间缩放、旋转以及斜切等变换参数,后续算法利用该 矩阵中的参数可以估算出机器人移动以及转向等运 动参数,结合运动传感器获取的运动向量最终可以 实现机器人场景识别定位与地图构建.



图 8 RSI-LDB 在室外机器人视觉定位系统中作为场景匹配算法的应用 Fig. 8 Application of RSI-LDB in visual-based localization system of robot for outdoor environment



图 9 2 幅图像之间的变换矩阵

Fig. 9 Transformation matrix between two images

4 结论

 1) 该算法作为二值化特征描述符具有较高的 匹配速度.

2)该算法通过计算特征采样区域的梯度保存 纹理信息 结合重心向量方向归一化特征描述区域, 并在图像金字塔中计算特征描述符,因而具有较高 的旋转与尺度不变性,相比传统二值化特征描述符 鲁棒性更高.

3) 实验结果表明: 当图像发生旋转、模糊以及 降质变换时 本算法具有较高的匹配性能 ,取得了比 BRIEF 以及 ORB 算法更准确的匹配结果. 另一方 面 ,实验结果也表明本文提出的 RSI-LDB 描述符需 要提高对大尺度视角变换的鲁棒性 ,以提高实际使 用中的稳定性.

参考文献:

- [1] RAJA R , ROOMI S M M , DHARMALAKSHMI D. Robust indoor/outdoor scene classification [C] // 8th International Conference on Advances in Pattern Recognition (ICAPR). Kolkata: IEEE , 2015: 1-5.
- [2] YUAN Y, MOU L, LU X. Scene recognition by manifold regularized deep learning architecture [J]. IEEE Transactions on Neural Networks and Learning Systems, 2015, 26(10): 2222-2233.
- [3] GALVEZ-LOÓPEZ D , TARDOS J D. Bags of binary words for fast place recognition in image sequences [J]. IEEE Transactions on Robotics , 2012 , 28(5): 1188-1197.
- [4] LOWE, DAVID G. Distinctive image features from scaleinvariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [5] BAY H , TUYTELAARS T , VAN GOOL L. Surf: speeded up robust features [J]. Computer Vision and Image Understanding , 2008 , 110(3): 346-359.
- [6] CALONDER M , LEPETIT V , STRECHA C , et al. BRIEF: binary robust independent elementary features [C] //

Computer Vision-ECCV 2010. Berlin: Springer, 2010: 778-792.

- [7] LEUTENEGGER S , CHLI M , SIEGWART R Y. BRISK: binary robust invariant scalable keypoints [C] // IEEE International Conference on Computer Vision. Barcelona: IEEE , 2011: 2548-2555.
- [8] RUBLEE E, RABAUD V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF [C] // IEEE International Conference on Computer Vision. Barcelona: IEEE, 2011: 2564-2571.
- [9] ALAHI A, ORTIZ R, VANDERGHEYNST P. FREAK: fast retina keypoint [C] // IEEE International Conference on Computer Vision and Pattern Recognition. Providence: IEEE , 2012: 510-517.
- [10] GAO Y Q , QIAO Y , LI Z F , et al. LTD: local ternary descriptor for image matching [C] // IEEE International Conference on Information and Automation. Yinchuan: IEEE , 2013: 1375–1380.
- [11] WANG Y , LU K , ZHAI R. Challenge of multi-camera tracking [C] // International Congress on Image and Signal Processing (CISP). Dalian: IEEE , 2014: 32–37.
- [12] BELLO E D, SALVADEO P. An image descriptors extraction hardware-architecture inspired on Human Retina [C] // Southern Conference on Programmable Logic (SPL). Buenos Aires: IEEE, 2014: 1-6.
- [13] LI H, WANG Y, MEI T, et al. Interactive multimodal visual search on mobile device [J]. IEEE Transactions on Multimedia, 2013(15): 594-607.
- [14] YANG X , CHENG K T. LDB: an ultra-fast feature for scalable augmented reality on mobile devices [C] // IEEE International Symposium on Mixed and Augmented Reality. Atlanta: IEEE , 2012: 49-57.
- [15] University of Oxford Visual Geometry Group. Image Sequences [DB/OL]. [2007-07-15]. http://www. robots. ox. ac. uk/~vgg/research/affine.
- [16] ROSTEN E , PORTER R , DRUMMOND T. Faster and better: a machine learning approach to corner detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence , 2010 , 32(1): 105–119.
- [17] FISCHLER M A, BOLLES R C. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6): 381-395.

(责任编辑 吕小红)