

文章编号:1674-2869(2011)04-0085-04

基于特征点相似度的匹配定位算法

甄巍松¹, 李国强¹, 鲁统伟²

(1. 光电控制技术重点实验室, 河南 洛阳 471009;

2. 武汉工程大学智能机器人湖北省重点实验室, 湖北 武汉 430074)

摘 要:前视定位系统中可能存在如视点、方位和距离等误差, 导致在匹配时刻获取的实时目标场景与模板中目标不一致, 从而影响目标定位的准确性. 本文提出了一种基于特征点相似度的匹配定位算法, 首先在图像的尺度空间上提取尺度不变特征点, 然后根据描述子来进行相似度的判别, 得到初始的匹配点集, 然后利用极线约束, 从而消除匹配错误的点. 利用类似 RANSAC 方法估计场景中目标的变换参数, 从而确定场景中目标所在的位置、尺度变化和旋转角度. 实验结果验证了该算法的有效性和鲁棒性.

关键词:尺度不变; 极线约束; 匹配; 定位

中图分类号: TP391

文献标识码: A

doi: 10.3969/j.issn.1674-2869.2011.04.022

0 引 言

对于实际有意义的前视定位系统^[1]而言, 惯导系统总是存在一定的误差的, 导致在匹配定位时刻获取的实时目标场景相对于参考目标场景, 存在视点、方位和距离的不同. 但是, 进一步考虑到惯导误差不可能是任意大的, 因此在匹配定位时刻获取的实时目标场景相对于参考目标场景的视点、方位和距离差异不会任意大. 这意味着只需要考虑一定范围内的视点、方位和距离变化. 当获取目标的三维信息比较困难, 或者不能获取目标的三维信息时, 可以将目标及其周围的图像作为模板, 将实时图作为场景. 采用基于特征的匹配算法, 则匹配运算量仅与特征的个数有关, 从而可以大大提高匹配速度; 同时在基于特征的匹配算法中还易于引入估计匹配图像对之间几何失真的机制, 从而使得算法具有抗图像几何失真的能力, 所以利用特征进行景像匹配定位是一种有效且实用的技术.

1 相关工作

许多研究者利用我种信息进行图像的匹配定位研究, 李相盲^[2]使用面积与周长等信息对国标进行定位. 杨述斌^[3]采用一种基于交替滤波的加权形态边缘检测算法实现边缘特征提取. 刘兴^[4]提出了光对摄像机进行标定, 然后再进行识别的方案, 更多的是采用特征点的办法. Harris^[5]从图像中提取类似角点的特征. 张正友^[6]还添加了约束条件, 即匹配

点应该在两个图像相同的位置附近. Schmid^[7]证明 Harris 角点检测具有良好的可重复性, 但在尺度变化范围大于 2 的情况下其性能很差. Matas^[8]提出了显著区域 (distinguished regions) 算法, 它能随图像的变化而保证稳定性和不变性. 它与角点检测不同, 角点检测关注的是两个方向上的灰度变化或梯度变化, 而它利用连通区域的特性, 能够做到视点不变性. Tuytelaars^[9]提出了两种区域检测算法: 平行四边形和椭圆. 这两种方法都是基于固定点的提取算法, 但是对尺度变化比较敏感. Baumberg^[10]也提出了一个类似的算法, 它先用多尺度的 Harris 算法检测尺度空间上的点, 然后利用二阶矩自适应地去除切变和拉伸带来的影响. Lowe^[11]指出上述方法是以非仿射不变的形式选择位置和尺度, 所以上述方法并不是完全仿射不变的方法, 同时也提出了尺度不变特征变换方法 (Scale Invariant Feature Transformation, SIFT). Yanke 等人提出了 PCA-SIFT 算法, 主要是将 SIFT 中 128 维的描述子降维, 以加快运算速度. Mikolajczyk^[12]也提出了梯度位置朝向直方图 (Gradient Location Orientation Histograms, GLOH) 算法, 采用了圆形的模板, 共 3 个径向和 8 个纬向线条, 分成了 17 个格子, 每个格子用 16 个朝向表示, 则其描述子共有 272 维, 再利用 PCA 降维, 产生 128 维的向量.

2 算法分析与设计

图像匹配定位过程主要包括四个步骤: **a.** 提

取特征点;**b.** 寻找初始匹配点;**c.** 去除错误匹配点;**d.** 目标定位.

2.1 提取特征点

由于距离不同,导致同样的特征点在模板图和场景图上表现出尺度上的差异,因此特征点提取技术必须具有一定的尺度不变性. 这里选择 SIFT 特征点进行匹配,由于 SIFT 使用具有尺度选择性的高斯差分(Difference of gaussian, Dog)滤波器,当特征点的尺度与某 Dog 算子的尺度相同时,该 Dog 算子的滤波输出响应将达到极大或极小(依赖于特征点的极性),这里极大或极小是相对于相邻空间位置及尺度而言的. 因此,通过检测 Dog 滤波器组输出空间位置和尺度空间的极值,就可得到相应的特征点位置及其尺度信息. 进一步地,可以在所检测出的特征点处计算梯度强度和方向,从而得到关于该特征点的更多信息. 同时由于 Dog 算子是各向同性的,因此用于提取特征点时具有旋转不变性.

2.2 寻找初始匹配点

特征点相似度测量的最简单方式就是最近邻方法. 最近邻方法定义为描述子之间的最小欧氏距离. 由于场景图像的背景不同或亮度变化,图像的特征点不容易完全匹配,所以匹配特征点的描述子区别可能很大,如果用一个全局的门限来约束,效果不会很好. 另外一个有效的方法就是比较最近邻和次最近邻的距离之比,如果最近邻的距离与次最近邻的距离的比值小于一定门限则认为是匹配点. 对于错误的匹配点来说,最近邻和次最近邻一般很相似,所以很容易去除,但也会去除掉一些匹配正确的点,如果能牺牲一些正确的匹配而去除大部分的错误点也是值得的.

因此制定特征点相似度匹配的准则,采取以下方法来确定初始匹配对:**a.** 用最近邻方法寻找初始匹配点,将小于一定门限的匹配对作为可能的匹配对;**b.** 用最近邻和次最近邻之比方法再次寻找匹配点,将比值小于一定门限的匹配对也作为可能的匹配对.

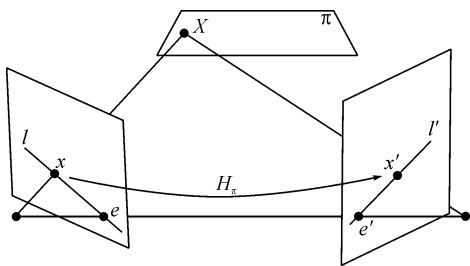


图 1 极线约束

Fig. 1 The epipolar constraint

2.3 去除错误匹配点

经过最近邻与次最近邻之比匹配后,仍会有一些匹配错误的点. 可以利用 F 矩阵^[13]来修正. 根据视觉成像原理, F 矩阵为 3×3 矩阵,且它的秩为 2,共有 7 个自由度. 对于同一物体在不同的视角下成像,可以知道图像中的一点 x 在另外一幅图中的一条直线 l' 上.

在利用 F 矩阵去除错误的匹配对过程中,由于可能有比较多的错误的匹配对,所以求到的 F 矩阵可能不够准确,从而导致特征点对应的极线不够准确. 本文采取迭代方式,设置一个变量 Δ 与特征点到极线的距离进行比较. 首先将 Δ 设为图像宽度和图像高度中最大值的一半,如果对应点到极线的距离小于它,则认为是可能正确的匹配对. 然后根据这些匹配对重新计算 F 矩阵,将 Δ 减小到原来的一半,如果对应点到极线的距离小于 Δ ,则认为是可能正确的匹配对. 如此反复,直到 Δ 小于一定的门限为止. 本文中设定 Δ 的最小值为 2,即对应的特征点到对应极线的距离应该小于 2 个像素.

2.4 目标定位

根据得到的对应匹配点,可以计算出模板图像在场景中的位置. 如果目标的变化,仅仅是旋转,平移和缩放,那么仅仅需要 2 对匹配点就能求出目标在场景中的位置和旋转角度和缩放比例. 但是由于并不能保证所有的匹配点都是正确的,可能会有部分匹配点错误,或者特征点本身提取的精度不够高,所以需要采取拟和的方法计算变换参数.

采用类似 RANSAC 方法来求变换参数. RANSAC 方法是一个在允许有许多野点(outlier)的数据集里寻找模型的方法. 采取以下方法来估计变换参数:

- a. 把匹配对按照特征点描述子之间距离排序,选择 M 个距离最小的匹配对;
- b. 应用 RANSAC 方法寻找 T 个正确的匹配对;
- c. 根据 T 个匹配对,重新计算变换参数 X ;
- d. 对所有的匹配对 m 中满足变换参数 X 的加入到 T 中,重新计算变换参数 X ;

在本文中,考虑到场景图像的变化不仅仅是仿射变化,而是复杂的透视变换,所以变换参数采取 2 次多项式的形式:

$$\begin{cases} u=(x^2 \ y^2 \ xy \ x \ y \ 1)(a_{20} \ a_{02} \ a_{11} \ a_{10} \ a_{01} \ a_{00})^T \\ v=(x^2 \ y^2 \ xy \ x \ y \ 1)(b_{20} \ b_{02} \ b_{11} \ b_{10} \ b_{01} \ b_{00})^T \end{cases} \quad (1)$$

公式(1)中, (x, y) 表示模板中特征点的位置, (u, v) 是场景图中的特征点, 该表达式表示模板图中特征点的位置到场景图中特征点位置的变换, 其参数用 $(a_{20} \ a_{02} \ a_{11} \ a_{10} \ a_{01} \ a_{00})^T$ 和 $(b_{20} \ b_{02} \ b_{11} \ b_{01} \ b_{00})^T$ 表示.

从公式(1)中可以看出, 需要计算的变换参数共包含 12 个, 一个匹配对就有 2 个等式, 那么求出该组参数至少需要 6 个匹配对.

根据求得的变换参数, 可以利用一个线段来计算出模板图像在场景中的位置、旋转角度和尺度变化. 将模板图像中心作为线段的起点 $T_1(x_1, y_1)$, 其水平方向作为模板图像的初始方向, 沿该方向经过一定的距离 T , 选择该点作为线段的终点 $T_2(x_2, y_2)$. 那么, 根据变换参数, 求出该线段的起点和终点在场景图中的坐标 $S_1(u_1, v_1)$ 和 $S_2(u_2, v_2)$, 点 $S_1(u_1, v_1)$ 即为模板图像中心在场景图中的位置. 场景图像相对模板图像的旋转方向为点 $S_1(u_1, v_1)$ 指向 $S_2(u_2, v_2)$ 的朝向. 点 $S_1(u_1, v_1)$ 和 $S_2(u_2, v_2)$ 之间的距离与点 $T_1(x_1, y_1)$ 和 $T_2(x_2, y_2)$ 之间的距离之比即为场景图像相对模板图像的尺度变化.

3 实验结果及分析

首先使用 INRIA 研究中心的 LEAR 研究组使用的数据库, 其中船序列图像的变化包括尺度、旋转变化和部分遮挡. 对其中的船的图像序列计算匹配对的数目, 该序列共有 9 幅图像, 使用第一幅图像作为模板图像, 其它的 8 幅图像作为待定位的图像, 图像间的变化为平移、尺度和旋转变化. 由于后面 8 幅图像的尺度变化越来越大, 所以找到的匹配点数目成下降趋势. 用本文计算匹配对的方法与 Lowe 的方法和 Mikolajczyk 算法进行比较, 其效果如图 2 所示.

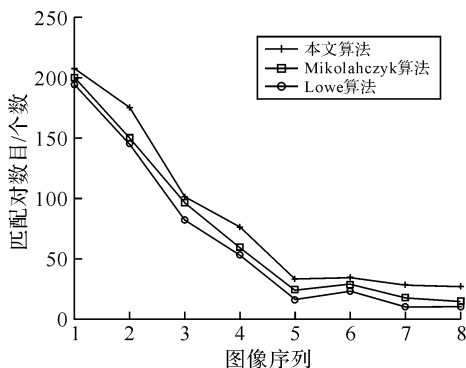


图 2 计算匹配对数目的算法比较

Fig. 2 The comparison of number of matching pairs

该数据库中, 缩放并且旋转的图像共有 9 个场景, 101 幅图像; 旋转图像共 4 个场景, 88 幅图

像; 缩放图像有 6 个场景, 87 幅图像. 对该数据库中所有的图像进行实验, 假设一个场景有 n 幅图像, 任意选取其中的一幅作为模板, 其它的图像作为待匹配定位图像, 则有 $n * (n - 1)$ 次匹配实验. 对同一个场景, 人工选取特征点进行匹配, 将得到的定位结果作为正确的定位结果, 其中包括位置 x 、位置 y 、尺度因子 s 、旋转角度 θ . 然后将计算结果与真实结果进行对比, 得到位置、尺度因子和旋转角度的误差的均值和标准差. 对这 19 个场景进行实验, 其统计结果如表 1 所示.

表 1 实验统计结果
Table 1 The experimental results

位置 x 的误差	均值	2.020 1
	标准差	1.150 3
位置 y 的误差	均值	2.193 9
	标准差	1.153 9
尺度因子 s 的误差	均值	0.119 0
	标准差	0.070 0
旋转角度 θ 的误差	均值	0.952 1
	标准差	0.529 4

另外, 对视点变化的图像进行定位, 其效果如图 3 所示. 图 3 中模板图与场景图的变化为视点变化, 另外图像间深度方向旋转的角度约为 30° . 图 3 中模板图与场景图的变化还存在尺度变化, 场景图的背景还包括杂乱的树叶. 该组图像的定位结果如图 3 所示.



图 3 视点变化加尺度变化下的定位结果

Fig. 3 The location result of changed viewpoint and scale

在图 3 中, 由于存在视点变化, 无法计算出准确的尺度变化和旋转角度, 只能手工指出模板图像的大概定位. 其结果与真实数据的比较如表 2 所示. 实验结果表明该检测方法, 可以克服尺度、旋转、平移、仿射亮度变换带来的影响, 甚至是一定的透视变换.

表 2 图像的的定位结果与真实数据的比较

Table 2 The comparison of location result with true data

定位结果 (坐标 x , 坐标 y , 尺度变化因子, 旋转角度)	真实值 (坐标 x , 坐标 y)
(211.3368, 220.0960, 0.6454, 11.7137)	(225, 212)

4 结 语

本文在介绍多种检测特征点方法的基础上, 选取 SIFT 特征点作为定位算法中的特征点. 综合

利用特征点的梯度朝向信息和极线约束,给出了特征点相似度度量的准则.采用 2 次多项式的形式对模板图像和场景图像的变换参数进行拟和,由类似 RANSAC 方法计算出变换参数.根据变换参数,求出模板图像在场景图像中的位置以及图像间的尺度变化和旋转角度.最后通过实验验证了该算法在尺度变化、平面旋转变化、深度方向旋转变化的场景定位中的有效性.

参考文献:

[1] 鲁统伟.前视目标图像匹配定位技术研究[D].武汉:华中科技大学图像所,2009.

[2] 李相育,刘辉.图像处理技术在染色体自动分析中的应用[J].武汉化工学院学报,2003,25(1):77-79.

[3] 杨述斌,陈艳菲,程莉.交替滤波的加权形态边缘检测算法[J].武汉工程大学学报,2009,31(7):88-90.

[4] 刘兴,刘庆祥.一种实用的车牌图象获取方案[J].武汉化工学院学报,2005,27(5):82-85.

[5] Harris C, Stephens M. A combined Edge and Corner Detector[J]. In 4th Alvey Vision Conference, 1988: 189-192.

[6] Zhang Z, Deriche R, Faugeras O, et al. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry [J]. Artificial Intelligence, 1995, 78:

87-119.

[7] Schmid C, Mohr R. Comparing and evaluating interest points[J]. In Proceedings of the International Conference on Computer Vision,1998:230-235.

[8] Matas J, Koubaroulis D, Kittler J. Colour image retrieval and object recognition using the multimodal neighborhood signature [J]. In Proceedings of the European Conference on Computer Vision, 2000: 48-64.

[9] Tuytelaars T, Van-Gool L. Wide baseline stereo based on local, affinely invariant regions [J]. In Proceedings of the British Machine Vision Conference,2000:121-130.

[10] Baumberg. Reliable feature matching across widely separated views [J]. In Conference on Computer Vision and Pattern Recognition,2000:774-781.

[11] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision,2004,60(2):91-110.

[12] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,2003(2):257-263.

[13] 陈泽志,吴成柯.一种高精度估计的基础矩阵的线性算法[J].软件学报,2002,13(4):840-845.

Match and location algorithm based on similarity of feature point

ZHEN Wei-song¹, LI Guo-qiang¹, LU Tong-wei²

(1. Science and Technology on Electro-optic Control Laboratory, Luoyang 471009, China;
2. Hubei Province Key Laboratory of Intelligent Robot, Wuhan Intitute of Technology, Wuhan 430074, China)

Abstract: The targets in the template and the real-time scene are different duo to the errors in the forward-looking locating system such as the view point, orientation, distant et al, which affect the accuracy of the target location. This paper presents a matching and locating algorithm based on the similarity of the feature points. Firstly, the scale-invariant feature points are extracted in the image scale space. Then the set of the initial matching points is build through the similarity of the descriptor. The wrong matching points are eliminated using the epipolar constraint. Finally, the position, scale factor and rotation angle are determined by estimating the parameters using RANSAC. Experimental results show the validity and robustness of the algorithm.

Key words: scale-invariant; epipolar constraint; matching; location

本文编辑:陈小平