

单目视觉定位方法研究综述

李荣明, 芦利斌, 金国栋

(第二炮兵工程学院 602 教研室, 西安 710025)

摘要: 根据单目视觉定位所用图像帧数不同把定位方法分为基于单帧图像的定位和基于双帧或多帧图像的定位两类。单帧图像定位常利用已知的点特征、直线特征或曲线特征与其在图像上的投影关系进行定位, 其中用点特征和直线特征的定位方法简单有效而应用较多; 基于双帧或多帧图像的定位方法因其操作复杂、精度不高而研究的还较少。通过对各方法的介绍和评述, 为单目视觉定位问题的研究提供参考。

关键词: 单目视觉; 视觉定位; 位姿估计

0 引言

近年来, 视觉传感器因能采集丰富的环境信息且价格低廉、使用方便而受到了普遍的关注, 基于视觉传感器的定位方法也成为了研究的热点。根据使用视觉传感器数目的不同, 视觉定位方法可分为单目视觉定位、双目视觉(立体视觉)定位和多目视觉(全方位视觉)定位。

单目视觉定位就是仅利用一台摄像机完成定位工作。它具有简单易用和适用广泛等特点, 无需解决立体视觉中的两摄像机间的最优距离和特征点的匹配问题, 也不会像全方位视觉传感器那样产生很大的畸变^[1]。在机器视觉研究领域, 如何在单目视觉条件下, 完成位置与姿态的求解已成为一个重要的研究方向。单目视觉定位技术可应用在多个方面, 例如摄像机标定、机器人定位、视觉伺服、目标跟踪和监测等。单目视觉定位的方法有很多, 但还没有一个明确的分类标准。

1 基于单帧图像的定位方法

基于单帧图像的定位就是根据一帧图像的信息完成目标定位工作。因为仅采用一帧图像, 信息量少, 所以必须在特定环境内设置一个人工图标, 图标的尺寸以及在世界坐标系中的方向、位置等参数一般都是已

知的, 从预先标定好的摄像机实时拍摄的一帧图像中提取图标中某些特征元素的像面参数, 利用其投影前后的几何关系, 求解出摄像机与人工图标的相对位置和姿态关系。如何快速准确地实现模板与投影图像之间的特征匹配问题是其研究的重点。该方法具有形式简单、算法实现容易、硬件要求低等优点, 但是也存在着鲁棒性、实时性较差和对人工路标依赖性强等缺点。常用的特征元素有点、直线、二次曲线等。

1.1 基于点特征的定位

基于点特征的定位又称为 PnP(Perspective-n-Point)问题^[2], 是机器视觉领域的一个经典问题。它是根据物体上 n 个特征点来确定照相机的相对位置和姿态, 具体描述为: 假定摄像机为小孔模型且已标定好, 摄取一幅在物体坐标系下坐标已知的 n 个空间点的图像, 且这 n 个图像点的坐标已知, 来确定这 n 个空间点在摄像机坐标系下的坐标。

对 PnP 问题的研究基本围绕解的确定性和求解算法的线性两方面展开, 多年来研究者们主要针对 P3P、P4P 和 P5P 问题作了大量有益的探索, 得到以下结论^[3-5]: 当 $n \leq 2$ 时有无限组解, 即仅有两个点不能确定点在摄像机坐标系下的位置; 当 $n=3$ 且三个控制点决定的平面不通过光心时, 最多有 4 组解且解的上限可以达到;

收稿日期: 2011-04-21 修稿日期: 2011-05-21

作者简介: 李荣明(1984-), 男, 江苏徐州人, 硕士研究生, 研究方向为视觉定位

当 $n=4$ 时, 4 个空间点在同一平面时解是唯一的, 4 个空间点不共面时, 则可能出现多个解; 当 $n=5$ 时, 若 5 个控制点中任意 3 点不共线, 则 P5P 问题最多可能有两个解, 且解的上限可以达到; 当 $n \geq 6$ 时, PnP 问题就成为经典的 DLT(Direct Linear Transformation)问题, 可以线性求解。在目标上设置点特征定位具有测量系统精度高、测量速度快的特点, 在陆上、空间、水下定位计算中得到了广泛的应用^[6-8]。由于通过同一平面不共线的 4 个空间点可以得到摄像机的唯一确定位置, 所以用点特征进行定位多应用 P4P 方法; 为了提高特征点提取的鲁棒性, 一般设计采用多于 4 个特征点的人工图标。

1.2 基于直线特征的定位

基于直线特征的定位研究也不少, 因为直线特征在自然环境中存在的比较多, 且其抗遮挡能力强, 易于提取。在文献[9~12]中应用图像中直线和摄像机光心构成的投影平面的法向量和物体直线垂直来构建定位数学模型。这种方法要求确定物体位姿的三条直线不同时平行且不与光心共面, 进而建立由三条直线构成的三个非线性方程。它有效地解决了利用直线特征如何进行视觉定位的问题, 但是非线性方程组较为复杂, 定位误差偏大。也有一些学者根据空间不平行于像面的平行线投影到像面交于灭影点的原理, 利用灭影点在像面上的位置, 可以计算出代表该组直线 3 维方向(相对于摄像机坐标系)的矢量, 从而可获得摄像机与人工图标的相对位姿参数。文献[13~15]就采用灭影点方法实现了无人机的位姿估计。该算法的计算建立在分析性结论的基础上, 无需迭代, 计算量小, 但必须准确提取像面直线和灭影点的位置参数, 这使图像的处理变得比较复杂。文献[16~19]针对四条直线组成的平面四边形(例如矩形、平行四边形等)特征提出了一些视觉定位算法, 求解过程简单, 具有较高的求解精度和较大的应用价值。

1.3 基于曲线特征的定位

基于曲线特征的定位一般需要对复杂的非线性系统进行求解。比较经典的如文献[20~21], 分别利用共面曲线和非共面曲线进行定位, 都需要对几个高次多项式进行求解, 算法比较复杂。但是当两个空间曲线共面时, 可以得到物体姿态的闭式解。

圆是很常见的图形, 作为二次曲线的一种, 也引起

研究人员的关注。一般情况下, 圆经透视投影后将在像面上形成椭圆, 该椭圆的像面参数与圆的位置、姿态、半径等存在着对应的函数关系, 采用一定的方法对相应的关系求解即可得到圆与摄像机的相对位置和姿态参数。文献[22~24]分别运用不同的方法对圆特征进行了定位。利用圆特征进行定位可以摆脱匹配问题, 提高定位速度, 但其抗干扰能力欠佳。

2 基于双帧或多帧图像的定位

虽然使用特殊的人工图标实现定位可以容易找到匹配特征, 但是使用人工图标本身制约了视觉定位的应用场合, 所以利用摄像机获取的自然图像信息来进行定位是该领域的研究趋势之一。基于双帧或多帧图像的定位方法, 就是利用摄像机在运动中捕捉同一场景不同时刻的多帧图像, 根据拍摄图像像素间的位置偏差实现目标的定位。实现多帧投影图像之间的对应特征元素匹配是该定位算法的关键。这类算法一般相对比较复杂, 精确性和实时性不高, 但不依赖人工标志, 通过拍摄自然图像就可实现定位。

在文献[25]中介绍了一种双帧图像定位估计方法。采用事先已标定的摄像机在运动中拍摄目标, 利用目标在前后相邻两帧图像上的投影点形成多个匹配点对估计出基本矩阵, 由基本矩阵和本质矩阵的关系进一步求出本质矩阵, 并经过分解获得单目摄像机的外部运动参数(旋转和平移), 利用坐标系转换从而获得目标的三维信息。文中针对定位中基本矩阵对定位精度的影响问题, 提出一种新的基本矩阵迭代估计算法, 结合 RANSAC 算法实现了基本矩阵的鲁棒性估计。该定位方法类似立体视觉定位原理, 可以获取较多的周围环境信息, 但是需要获得摄像机运动的平移距离和投影图像之间至少八对匹配点, 局限性大。文献[26~27]利用尺度不变特征变换(SIFT)具有尺度、旋转不变性的特点, 采用 SIFT 算法进行图像特征的提取和匹配, 计算出目标的三维信息, 实现目标的定位。该算法很好地解决了拍摄图像对应点的自动匹配问题, 但由于其图像采样频率偏低, 不适合摄像机在快速运动状态下的应用。

文献[28]利用了图像之间的拼接技术实现摄像机的定位。摄像机通过平移或是旋转, 可以获取两幅相邻的有着重叠区域的图像, 且重叠区域中的相同像点的

位置发生了改变。通过图像拼接过程中的图像配准技术,利用仿射变换求得相邻图像之间的特征变化关系,最后进一步推得摄像机的运动情况,从而初步实现摄像机的定位。该方法采用基于灰度信息的拼接方法对两幅图像进行图像配准。实验表明该方法有效,可初步实现摄像机定位的要求。但因为缺乏场景中景物到摄像机光心的实际距离,无法推得摄像机的位移量,而只能计算得出摄像机的运动方向以及摄像机绕光轴的旋转角度。

文献[29]提出了一种将单目视觉测量中的离焦法和聚焦法相结合的摄像机定位方法。通过移动摄像机,目测找出图像近似最清晰的位置,在其前后各取等间距的两个位置,并在以上三个位置拍摄图像;然后应用离焦定位算法进行计算,得到近似的峰值点位置。在近似峰值点位置附近取若干个测点并拍摄图像,然后应用聚焦定位算法进一步求得精确的峰值点位置,从而确定摄像机相对于被测点的聚焦位置。实验证明了其方法的正确性和准确性。该方法将离焦法和聚焦法结合起来,使两者的优缺点互相弥补,避免了建立复杂的数学模型,同时减少了由于简化的假设与模型而造成的误差,提高了测量精度,适合于针对实际被测物体复杂图像的测量。

3 结 语

本文根据单目视觉定位所用的图像帧数不同把定位方法进行了分类,并全面而简要地介绍了各个方法的研究现状。目前,基于单帧图像的定位方法研究的比较多,特别是利用点特征和直线特征的方法,已有许多成功的应用;而基于双帧或多帧图像的定位方法研究的还较少,其在实时性和准确性方面还需进一步提高。

参考文献

- [1]周娜. 基于单目视觉的摄像机定位技术研究[D]. 南京:南京航空航天大学,2007
- [2]吴朝福,胡占义. PNP 问题的线性求解算法[J]. 软件学报, 2003,14(3):682~688
- [3]Fishler 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
- [4]Abdel-Aziz Y, Karara H. Direct Linear Transformation from Comparator to Object Space Coordinates in Close-Range Photogrammetry [C]. In:ASP Symp.Close-Range Photogrammetry. Urbana,Illinois.1971:1~18
- [5]胡占义,雷成,吴福朝. 关于 P4P 问题的一点讨论[J]. 自动化学报,2001,27(6):770~776
- [6]邹伟,喻俊志,徐德等. 基于 ARM 处理器的单目视觉测距定位系统[J]. 控制工程,2010,17(4): 509~512
- [7]徐筱龙,徐国华,陈俊. 水下机器人的单目视觉定位系统[J]. 传感器与微系统,2010,29(7): 109~111
- [8]任沁源. 基于视觉信息的微小型无人直升机地标识别与位姿估计研究[D]. 浙江:浙江大学,2008
- [9]R.Horaud,B.Conio,O.Leboulleux,et al. An Analytic Solution for the Perspective 4-Point Problem [J]. Computer Vision Graphics and Image Processing,1989,47(1):33~44
- [10]Dhome,M.Ricchi,J.T.Laprest,et al. Determination of the Attitude of 3D Objects from a Single Perspective View [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1989,11(12):1266~1278
- [11]H.Chen. Pose Determination from Line to Plane Correspondences:Existence Solution and Closed form Solutions [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1993, 13(6): 530~541
- [12]Y.Liu,T.S.Huang,O.D. Faugeras. Determination of Camera Location from 2D to 3D Line and Point Correspondences [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,1990,12(1):28~37
- [13]Mukundan R, Raghu Narayanan R V, Philip N K.A Vision Based Attitude and Position Estimation Algorithm for Rendezvous and Docking[J]. Journal of Spacecraft Technology,1994,4(2):60~66
- [14]刘士清,胡春华,朱纪洪. 一种基于灭影线的无人直升机位姿估计方法[J]. 计算机工程与应用,2004, 9: 50~53
- [15]刘晓杰. 基于视觉的微小型四旋翼飞行器位姿估计研究与实现[D]. 吉林:吉林大学,2009
- [16]王晓剑,潘顺良,邱力为等. 基于双平行线特征的位姿估计解析算法[J]. 仪器仪表学报,2008, 29(3):600~604
- [17]吴福朝,王光辉,胡占义. 由矩形确定摄像机内参数与位置

- 的线性方法[J]. 软件学报,2003,14(3):703~712
- [18]Sun Fengmei,Wang Weining. Pose Determination from a Single Image of a Single Parallelogram[J]. Acta Automatica Sinica,2006,32(5):746~752
- [19]秦丽娟,胡玉兰,魏英姿等. 一种基于平面四边形的视觉定位算法[J]. 沈阳理工大学学报,2009,28(2):66~69
- [20]D.A.Forsyth, J.L.Munday, A.Zisserman, et al. Projective Invariant Representation Using Implicit Algebraic Curves[J]. Image and Vision Compute.1991,9(2):130~136.
- [21]S.D.Ma,S.H.Si,Z.Y.Chen. Quadric Curve Based Stereo [C]. Proceedings of IAPR Conference. Hague:The Netherlands, 1992:234~235
- [22]Safaei-Rad,I.Tchoukanov,K.C.Smith,et al. Three-Dimension of Circular Features for Machine Vision[J]. IEEE Transaction on Robotics and Automation, 1992,8:624~640
- [23]Zen Chen, JenBin Huang. A Vision-Based Method for the Circle Pose Determination with a Direct Geometric Interpretation [J]. IEEE Transaction on Robotics and Automation. 1999,15(6):1135~1140
- [24]张广军,周富强. 基于双圆特征的无人机着陆位置姿态视觉测量方法[J]. 航空学报,2005,26(3):344~347
- [25]张治国. 基于单目视觉的定位系统研究[D]. 武汉:华中科技大学,2009
- [26]刘立. 基于多尺度特征的图像匹配与目标定位研究[D]. 武汉:华中科技大学,2008
- [27]任沁源,李平,韩波. 基于视觉信息的微型无人直升机位姿估计[J]. 浙江大学学报(工学版),2009,43(1):18~21
- [28]沈慧杰. 基于单目视觉的摄像机定位方法的研究[D]. 吉林:吉林大学,2009
- [29]祝世平,强锡富. 用于摄像机定位的单目视觉方法研究[J]. 光学学报,2001,21(3):339~342

Research Overview of Location Method for Monocular Vision

LI Rong-ming , LU Li-bin , JIN Guo-dong

(Faculty 602, The Second Artillery Engineering University, Xi'an 710025)

Abstract: The location methods based on monocular vision fall into two classes, which are the location method from single image and two or more images, based on the different number of the images needed. The location methods from single image use the projection of the point features, straight line features, curved line features to estimate the pose of the camera. The location algorithm based on point features and straight line features is simple, effective and greatly used. There are few studies on the location method from two or more images, because the location algorithm is complex in computing and less exact. Introduces and reviews all kinds of location methods, which provides reference for the research of location method based on monocular vision.

Keywords: Monocular Vision; Visual Location; Pose Estimation