

视觉位姿测量中单目视觉与双目立体视觉 测量精度比较*

朱枫^{1,2}, 何雷^{1,2}, 郝颖明¹

(1 中国科学院沈阳自动化研究所 沈阳 110016; 2 中国科学院研究生院 北京 100080)

摘要: 本文结合实际模型,以仿真实验的方式对两种典型的视觉位姿测量方法的测量精度进行了比较分析。误差统计结果表明,在广泛的应用背景下,基于模型的单目视觉方法由于在特征点坐标计算过程中引入了目标模型的约束,计算精度和对误差的鲁棒性都明显优于双目立体视觉方法。仿真结果对实际工程中如何选取测量算法具有一定的指导意义。

关键词: 单目视觉; 双目视觉; 位姿测量; 测量精度

Contrast of calculated accuracy between monocular vision and binocular vision without modeling

Zhu Feng^{1,2}, He Lei^{1,2}, Hao Yingming¹

(1 Shenyang Institute of Automation, Chinese Academy of Science, Shenyang 110016, China;

2 Graduate School of the Chinese academy of Sciences, Beijing 100080, China)

Abstract: In this paper, combining the factual model, we have simulated two representative methods of visual calculation in the computer vision domain. The statistical result of error denotes that the model-based monocular vision behaves more outstanding than the binocular vision on the computational accuracy and robust because the monocular vision introduces the known object model in the process of computing the coordinates of feature points. The simulated result has certain guiding significance for choosing the measurement algorithm in the actual project.

Key words: monocular vision; binocular vision; visual calculation; model-based

1 引言

在实际工程应用中,有许多场合要求测量一个物体相对于另一个物体的位置与姿态,即所谓的位姿测量。从数学意义上说即测量两个坐标系间的平移与旋转变换关系,包括3个位置和3个旋转角共6个位置量。理论上,只要已知空间不共线的3点在两个坐标系下的坐标,就能唯一确定两坐标系间的位姿关系,因此,位姿测量的关键就是如何得到特征点在这两个坐标系下的坐标。

视觉位姿测量作为一种典型的非接触式测量方法

已经在工业、医疗、航天等领域得到了广泛的应用。最简单的视觉测量方法即在一个物体(称为目标物体)上布置一组点目标,并精确测量出这些点目标在目标物体坐标系(简称目标坐标系)下的三维坐标,在另一个物体上布置一定数目的摄像机,摄像机对目标点成像后,先通过图像处理得到这些目标点在图像中的成像坐标,再利用计算机视觉的方法计算出各目标点在测量物体坐标系(简称测量坐标系)下的三维坐标,从而实现目标坐标系与测量坐标系间相对位置与姿态的测量^[1]。

为求出目标点在测量坐标系下的坐标,典型的视觉测量方法有2种:一种是只用一台摄像机对目标成

* 基金项目:国家自然科学基金(60575042)资助项目

像,同时利用目标点间固有的几何约束关系进行求解,这种方法在计算机视觉领域称为 PNP 问题,是该领域的研究热点之一^[2]。另一种更为典型方法的是双目立体视觉方法,即采用两台摄像机,通过同一目标点在两台摄像机中的成像进行交会三角计算,得出目标点在测量坐标系下的三维坐标^[3]。基于模型的单目视觉由于只使用一个摄像机,因此具有系统简单的特点,但需要已知目标点间的几何约束。双目立体视觉虽然系统复杂,但不需要已知目标点间的几何约束,可直接求解目标点的坐标。在位姿测量领域,为实现位姿计算必须已知物体上至少 3 点的坐标,即目标模型必须已知,因此双目立体视觉的这种优势已不复存在。因此,选择何种方法主要取决于这两种方法的测量精度及对输入参数误差的鲁棒性,在测量精度和鲁棒性相当的前提下,基于模型的单目视觉方法比双目立体视觉方法更具有优势。

视觉位姿测量系统的测量精度取决于 2 个方面,输入参数的误差和测量系统对误差的传递。对于这两种方法来说,输入参数都包括以下 3 类:特征点在模型坐标系下的三维坐标,特征点在图像坐标系下的图像坐标,摄像机模型参数。视觉位姿测量系统的精度分析即分析相对位置和相对姿态的误差与这些输入参数误差的关系。系统的鲁棒性则体现了测量结果对输入参数误差的敏感性,它与测量系统的设计参数以及测量方法有直接关系。

目前,对双目立体视觉测量精度的研究很多,主要集中于特征点三维坐标的测量误差与测量系统参数之间的关系研究方面。Jeffrey J. Rodriguez 等人^[4]和 Chienchung Chang 等人^[5]分别针对图像量化误差研究了双目立体视觉的测距精度与测量系统参数之间的关系。吴彰良等人^[6]和刘佳音等人^[7]分析了多种误差源的影响,重点研究了三维坐标测量误差与测量系统设计参数间的关系特性。关于基于模型的单目视觉的研究目前还主要集中于计算方法方面,由于其求解过程复杂,对误差的分析还处于仿真阶段。Larry Davis 等人^[8]从目标跟踪的角度给出了目标在世界坐标系下的坐标的测量误差与目标模型三维坐标误差的关系并给出了仿真实验结果。而对于两种方法的测量精度比较,则只有 Hassan Mostafavi^[9]从刚体位姿估计出发,得到了除目标到相机的距离外,单目视觉的测量精度可以同双目立体视觉的测量精度相比的结论。

本文以视觉位姿测量为背景,以仿真实验为手段,在一定的限制条件下进行基于模型的单目视觉与双目

立体视觉方法的测量精度和鲁棒性的比较研究。所得到的结论不仅具有一定的理论意义,同时对实际应用系统的设计也将具有指导意义。

正文部分将首先简述位姿计算方法,然后介绍分析方法与实验条件,最后给出实验结果与分析。

2 位姿计算方法

两个坐标系间相对位置和姿态的计算要求已知三个以上的非共线点分别在两个坐标系下的坐标,同时还要求目标点在两个坐标系间的变换满足刚体变换,即特征点在两个坐标系下的三维坐标要满足目标模型的约束。只要满足了这两个条件,就可以用文献[1]中的旋动理论计算出相对位置和姿态。

基于模型的单目视觉和双目立体视觉分别用不同的方法计算特征点在测量坐标系下的三维坐标。

2.1 基于模型的单目视觉

基于模型的单目视觉由于只有一个摄像机,三维坐标计算时只能利用目标模型的约束,其计算结果必然满足目标模型的约束。

本文采用我们在文献[10]中给出的最优两分迭代法,可根据特征点的图像坐标和模型约束很快计算出 3 个目标点在测量坐标系下的坐标。

2.2 双目立体视觉

双目立体视觉方法可以通过 2 台摄像机间的约束直接得到特征点的三维坐标,若同时考虑模型的约束,就会有冗余信息,表现为方程的个数多于未知量的个数,当图像坐标有误差时,则这一方程组无解,通常采用在某种意义下的最优解作为计算结果。信息的冗余宏观上提高了系统对误差的鲁棒性。以本研究内容为例,简单地分别对 2 个摄像机进行 P3P 计算,再求平均值,就会在统计意义下提高系统的精度。关于基于模型的双摄像机优化算法及其与单摄像机精度的比较,我们进行了研究,相关内容参见文献[11]。

本文要讨论的是立体视觉算法的精度,首先按 3 个目标点在 2 台摄像机中的成像坐标,采用立体视觉方法,计算出 3 个目标点在测量坐标系下的坐标 $P_w = (x_w, y_w, z_w)$ 。由于图像坐标量化误差的存在,用立体视觉方法计算出的 3 个点的坐标是存在误差的,已经无法满足目标模型的约束,也就无法直接利用旋动理论求解相对位置和姿态。必须对求得特征点空间坐标进行修正,使其满足模型的约束。本文以目标模型为约束条件,以修正点坐标 $p'_w = (x'_w, y'_w, z'_w)$ 与 p_w

的距离平方和最小作为目标函数,通过约束条件下的最小二乘法求解 p'_w ,使之满足模型约束。

3 分析方法与实验条件

3.1 分析方法

不论是基于模型的单目视觉还是立体视觉方法,在计算过程中都存在着迭代运算或非线性优化过程,其测量结果误差都很难或不能表示为以输入参数误差为变量的解析式,从理论上进行分析有一定的困难。所以本文采用实验的方法进行比较,得到结果后,再给出直观定性的解释。

由于测量误差不仅与输入参数误差有关,还与目标与摄像机间的相对位姿关系有关,不同位姿下的测量误差是不同的,因此为全面考察两种方法的误差只能在统计意义下进行分析比较。而在统计意义下进行的比较,必须有各种条件下的足够充分的实验数据,这是实际实验所无法实现的。因此,本文采用仿真实验的方法进行误差分析。

3.2 仿真实验条件

事实上,基于模型的单目视觉计算精度除与图像坐标点的误差有关外,还与目标间的几何关系、目标当前的位置与姿态、摄像机的内外参等多种条件相关。而双目立体视觉的计算精度除与上述条件有关外,还与立体视觉系统的另一重要参数基线长度有关,基线越长,则精度越高。因此很难在完全平等的条件下进行两种方法计算精度的比较。本文仅在一组给定的基本条件下进行精度的比较。

图1为双目立体视觉的仿真模型,假设摄像机视场角已知,通过计算,保证三点模型在两台摄像机的公

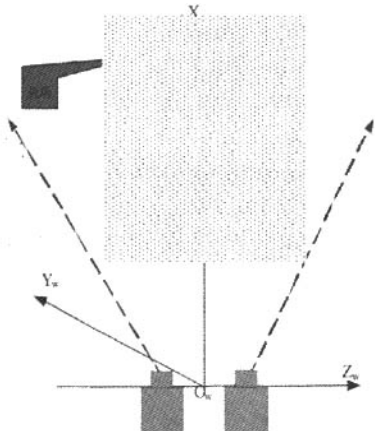


图1 立体视觉仿真模型

共视场内移动。仿真实验中设测量坐标系为 $O_w X_w Y_w Z_w$,两摄像机的基线与 Z_w 轴共线,光心关于 O_w 对称位于 Z_w 上。两摄像机完全相同,基线长度为 0.4 m ,分辨率为 512×512 像素,摄像机内参数为 $u_0 = v_0 = 256$,焦比 $f_u = f_v = 1024$ 。单目视觉仿真模型与图1近似,只是少了一个摄像机,视场和坐标系的选取和立体视觉相同,这样能够尽可能保证两种方法计算精度比较的合理性。

结合实际视觉测量应用,仿真实验中采用的目标模型如图2所示,3个目标点构成的三角形 ABC 的边长分别是 $(AB=0.8\text{ m}, AC=0.5\text{ m}, BC=0.5\text{ m})$ 。目标坐标系的原点 O 位于 AB 边的中点, AB 边所在方向为 Z 轴正方向, OC 方向为 X 轴正方向。这样三个特征点在目标坐标系下的坐标为: $A(0, 0, -0.4), B(0, 0, 4), C(0, 3, 0, 0)$ 。

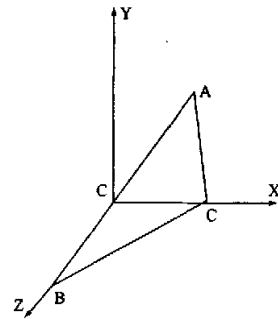


图2 三点目标模型

为简化计算,对于输入参数的误差仅考虑图像坐标的量化误差,不考虑其它误差的影响。由于在同等条件下进行2种方法的比较,这一假设不影响基本结论。仿真实验中取 0.25 像素量化误差对图像点进行提取。

仿真实验时,首先根据特征点的空间坐标求出其对应的图像坐标,然后对图像坐标作模拟量化误差处理,最后通过处理后的图像坐标和已知的目标模型运用前述两种测量方法进行位姿求解。实验中 x 轴(光轴)方向目标的移动范围根据实际工程中的一般需求设定为 $4\text{ m} \sim 20\text{ m}$,借助于单个摄像机的视场角,可以从 x 轴方向的距离计算出 y, z 轴方向的目标移动范围,同时结合实际测量应用,设姿态角的变化范围为 $-6^\circ \sim 6^\circ$ 。在上述目标移动区域中不断改变目标坐标系的位置和姿态角达到模拟不同位姿情况下的位姿求解,从而实现测量系统的误差统计。

3.3 实验结果与分析

图3给出了单目视觉和双目立体视觉位姿计算误

差的比较结果,其中菱形标志表示单目视觉,正方形标志表示立体视觉。横坐标为目标与摄像机间的距离,单位为米,纵坐标表示对应距离的6个位姿参数的统计误差,位置误差单位为毫米,姿态误差单位为度。图中给出的误差为均方根误差,所有误差均以绝对误差表示。

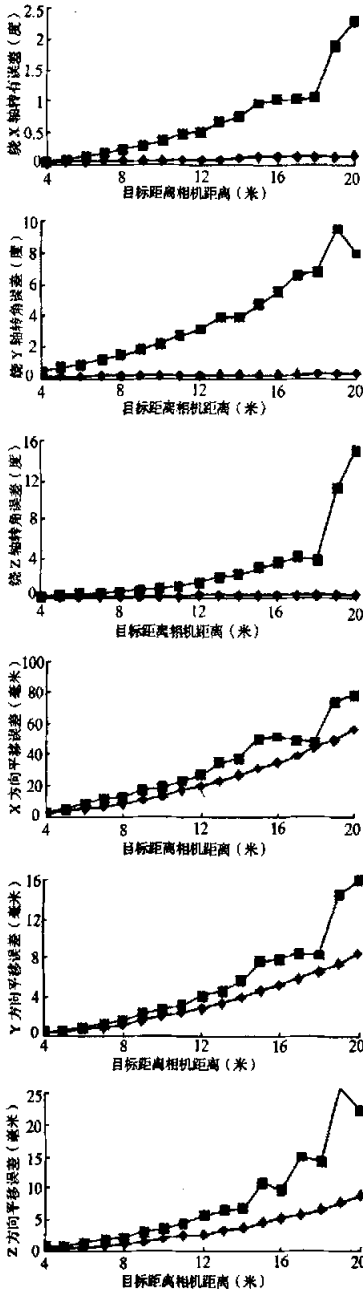


图3 两种方法位姿计算误差

从图中可以看出,在本文的仿真条件下,单目视觉

计算精度在姿态角误差上明显优于双目立体视觉,同时,位置误差也比双目立体视觉要小。从转角误差来看,双目立体视觉计算的误差太大,已经不能达到姿态测量的目的,这反映了在特征点坐标计算过程中不引入模型的弊病。从位置误差来看,在比较合理的基线长度情况下,双目立体视觉测量误差要大于单目视觉。当然理论上,可以通过增大基线长度来提高双目立体视觉的计算精度,并且能做到比单目视觉误差更小,但在实际应用中,两摄像机的基线太长,就意味着更大的空间开销和更小的视场。

另外,从误差图中可以发现,无论转角还是位置误差,在鲁棒性方面,单目视觉要优于双目立体视觉,特别是当目标离摄像机较远的时候。

4 结 论

本文对视觉位姿测量的两种典型方法在相对较公平的情况下,结合实际模型进行了大量仿真实验,仿真结果表明,在广泛的应用背景下,单目视觉算法由于在特征点坐标计算过程中引入了目标模型(必须引入),计算精度和对误差的鲁棒性要明显优于通常我们更为熟悉的双目立体视觉算法(通常我们不需要引入模型计算特征点坐标)。这个结论对实际应用中如何选取测量算法具有一定的指导意义。

参考文献

- [1] 刘伟军,董再励,郝颖明,等.基于旋动理论的移动机器人自定位方法研究[J].中南工业大学学报. 2000,31(专辑):468-470.
- [2] WOLFE W J, MATHIS D, WEBER C, et al. The perspective view of three points [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1991, 13(1): 66-73.
- [3] KELSEY J M, BYRNE J, COSGROVE M, et al. Vision-Based Relative Pose Estimation for Autonomous Rendezvous And Docking, Aerospace Conference, 2006 IEEE: 1-20.
- [4] RODRIGUEZ J J, AGGARWAL J K. Quantization error in stereo imaging [C]. 1988: 153-158.
- [5] CHANG C C, CHATTERJEE S. Quantization error analysis in stereo vision [J]. Signals, Systems and Computers, 1992: 1037-1041.
- [6] 吴彰良,卢荣胜,胡鹏浩,等.双目立体视觉传感器精度分析与参数设计[J].郑州轻工业学院学报(自然科学版),2006,21(3):32-35.

- [7] 刘佳音,王忠立,贾云得. 一种双目立体视觉系统的误差分析方法[J]. 光学技术, 2003, 29(3): 354-357, 360.
- [8] DAVIS L, CLARKSON E. Predicting accuracy in pose estimation for marker-based tracking [C]. Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality, 2003, 1-8.
- [9] MOSTAFAVI H. Rigid body pose estimation accuracy: a comparison of single-camera and stereo techniques [J]. SPIE, 1994, 2350: 206-217.
- [10] 周鑫,朱枫. 关于3P问题解的唯一性条件的几点讨论[J]. 计算机学报, 2003, 26(12): 1696-1701.
- [11] 郝颖明,朱枫,欧锦军. 目标位姿测量中的三维视觉方法[J]. 中国图像图形学报, 2002, 7(12): 1247-1251.