

目标位姿测量中的三维视觉方法

郝颖明 朱 枫 欧锦军

(中国科学院沈阳自动化研究所, 中国科学院机器人学开放研究实验室, 沈阳 110016)

摘 要 要测量出一组特征点分别在两个空间坐标系下的坐标, 就可以求解两个空间目标间的位姿关系, 实现上述目标位姿测量方法的前提条件是要保证该组特征点在不同坐标系下, 其位置关系相同, 但计算误差的存在却破坏了这种固定的位置关系; 为此, 提出了两种基于模型的三维视觉方法——基于模型的单目视觉和基于模型的双目视觉, 前者从视觉计算的物理意义入手, 通过简单的约束迭代求解实现模型约束; 后者则将简单的约束最小二乘法和基于模型的单目视觉方法融合在一起来实现模型约束. 引入模型约束后, 单目视觉方法可以达到很高的测量精度, 而基于模型的双目视觉较传统的无模型立体视觉方法位移精度提高有限, 但姿态精度提高很多.

关键词 立体视觉 单目视觉 目标位姿 三维测量 量化误差 基于模型

中图分类号: TP274.5 TP242.6 **文献标识码**: A **文章编号**: 1006-8961(2002)12-1247-05

3D Visual Methods for Object Pose Measurement

HAO Ying-ming, ZHU Feng, OU Jin-jun

(Robotics Laboratory, Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110016)

Abstract The related pose between two space objects can be determined if the coordinates of a group marks at two reference frames are measured. The precondition of the above method is that the site relationship among the marks must be changeless at the two reference frames, but this relationship will be destroyed by the computational error. So two model-based visual methods include model-based monocular vision method and model-based binocular stereo method are introduced. The model-based monocular vision method starts with the physical meaning of vision, and get the 3D space coordinates of marks by simple restrictive iterative. The model-based binocular stereo method inoculates the model-based monocular vision and advanced least square method to realize the model restriction. The compare analysis between the two methods and the experiment result will be given at last. Using these method, the model-based monocular vision can reach very high measurement precise, and the model based binocular stereo has a little higher measure than the model-based monocular vision. For any method, the size of quantization error will effect the measurement precise of visual method directly. These method has been used in mobile robot visual guided experiment at our laboratory successfully, the speed of image grabbing is 0.3 second per frame.

Keywords Stereo vision, Monocular vision, Object pose, 3D measurement, Quantization error, Model-based

0 引 言

随着机器人技术的发展, 快速准确地确定两个空间目标(如机器人)间的相互关系, 从而实现它们的协调合作显得愈来愈重要. 对于两个空间目标来说, 无论是两个目标都运动, 还是一个静止, 一个运动, 都可以把其中一个作为观察者, 另一个作为目标. 理论上

讲, 只要在目标上取3个以上的点, 并分别测出其在目标坐标系和观察坐标系下的坐标, 就可以确定目标坐标系相对于观察坐标系的位姿^[1], 即两个空间目标间的相互关系. 视觉方法是确定空间目标位姿的常用方法, 该方法首先在观察者上设置一个或两个摄像机, 而在目标上选择一组固定的标志, 通过图象处理技术检出各标志在观察坐标系下的三维坐标, 即可得到目标相对于观察坐标系的位姿.

基金项目: 国家自然科学基金基金项目(69975020)

收稿日期: 2001-12-10; **改回日期**: 2002-05-13

目标位姿测量的首要条件是求出各特征点在观察坐标系下的三维坐标. 在计算机视觉中, 设第 i 个特征点 P_i 的图象坐标为 (u_i, v_i) , 空间坐标为 (x_i, y_i, z_i) , 则有:

$$I = CS \quad (1)$$

其中, $I = \begin{bmatrix} u_i t_i \\ v_i t_i \\ t_i \end{bmatrix}$ 表示图象坐标, t_i 是一个系数;

$C = \begin{bmatrix} c_{00} & c_{01} & c_{02} & c_{03} \\ c_{10} & c_{11} & c_{12} & c_{13} \\ c_{20} & c_{21} & c_{22} & c_{23} \end{bmatrix}$ 为变换矩阵, $S = \begin{bmatrix} x_i \\ y_i \\ z_i \\ 1 \end{bmatrix}$ 表示空

间坐标. 由此可见, 通过某一特征点的空间坐标可以唯一地确定其图象坐标. 在式(1)中, 将 t_i 消去, 可以得到下面的形式

$$\begin{cases} (u_i c_{20} - c_{00})x_i + (u_i c_{21} - c_{01})y_i + (u_i c_{22} - c_{02})z_i = c_{03} - u_i c_{23} \\ (v_i c_{20} - c_{10})x_i + (v_i c_{21} - c_{11})y_i - (v_i c_{22} - c_{12})z_i = c_{13} - v_i c_{23} \end{cases} \quad (2)$$

这是一条直线方程, 它表示图象上的一点对应空间的一条直线, 因此仅仅已知某一特征点的图象坐标是无法计算其空间坐标的. 为求解特征点的三维空间坐标, 一种方式是将特征点间的位置关系作为一种约束条件, 构成单目视觉测量系统^[2]. 另一种方式是引入第 2 个摄像机, 同一特征点分别在两个摄像机中成像, 得到两个空间直线方程, 其交点坐标即为该特征点的空间坐标, 从而构成双目立体视觉测量系统^[3].

特征点三维坐标的求解和两个坐标系间变换关系的计算似乎已经完全解决了目标位姿的测量问题, 其实并不尽然. 因为通过特征点的空间坐标求解两个坐标系间变换关系的前提条件是要保证标志点间的刚性, 即在两个不同的坐标系下, 各标志点间的位置关系要保持一致. 基于模型的单目视觉系统可以满足这一要求, 但一般的双目立体视觉系统却不能满足这一要求. 这是因为图象获取过程中量化误差的存在, 使得通过双目立体视觉计算出的各标志点三维坐标的计算值与实际值间存在着一定的误差^[4]. 这样, 仅仅依靠双目立体视觉方法计算出的特征点间的位置关系就不再满足原来的位置约束, 也就无法确定两个坐标系间的对应关系. 因此无论采用哪种方法求解特征点的三维坐标, 都必须将标志点间的位置关系作为一个严格的约束条件引入计算过程, 在这一前提下使用视觉计算原理, 这实际上是一个基于模型的视觉问题.

1 位姿关系的求解

设有两个坐标系: 观察坐标系 S 和目标坐标系 O . 第 i 个特征点 P_i 在两个坐标系下的坐标分别为: $S_i = (x_i, y_i, z_i)^T$ 和 $O_i = (x_{oi}, y_{oi}, z_{oi})^T$, 则两个坐标系间的变换关系可以通过式(3)来表示, 其中 R 为旋转矩阵, T 为平移矩阵.

$$S_i = RO_i + T \quad (3)$$

这里 T 代表了两个坐标系间的位置变换关系, 而姿态变化则包含在 R 矩阵中, 可以通过下式由 R 矩阵求出绕 x, y, z 轴沿逆时针方向的转角 α, β, γ , 旋转顺序为 $x \rightarrow y \rightarrow z$.

$$\begin{cases} \gamma = -\arctan \frac{r_{01}}{r_{05}} \\ \beta = \arctan \frac{r_{20}}{r_{00} \cos \gamma - r_{10} \sin \gamma} \\ \alpha = \arctan \frac{r_{02} \sin \gamma + r_{12} \cos \gamma}{r_{01} \sin \gamma + r_{11} \cos \gamma} \end{cases} \quad (4)$$

为求解位姿关系, 先求旋转矩阵 R . 式(3)不仅对两个坐标系下的特征点成立, 对向量也同样成立. 当向量为单位向量时, 只有旋转而没有平移, 即为

$$n_{Si} = Rn_{Oi} \quad (5)$$

其中, n_{Si} 和 n_{Oi} 分别为同一单位列向量 n 在两个坐标系下的表示. 当已知 3 个以上特征点时, 可以根据通过任意两个特征点的连线求出两个线性无关(靠特征点的位置来保证)的单位列向量 n_1 和 n_2 . 令 n_3 为 $n_1 \times n_2$ 所构成的单位列向量, 则这 3 个列向量之间都是线性无关的, 且满足

$$n_S = Rn_O \quad (6)$$

其中, $n_S = [n_{S1} \ n_{S2} \ n_{S3}]$, $n_O = [n_{O1} \ n_{O2} \ n_{O3}]$.

根据式(6)和式(7)可以求出旋转阵 R 和平移阵 T .

$$\begin{cases} R = n_S n_O^{-1} \\ T = S_i - RO_i \end{cases} \quad (7)$$

上述求解过程的先决条件是: 在两个坐标系下, 特征点间的位置关系保持不变, 即连接任意两个特征点的线段, 其大小和方向在两个坐标系下都是相同的.

2 基于模型的单目视觉

对于单目视觉来说, 如果有 k 个特征点, 则其空间坐标共有 $3k$ 个未知数. 根据式(2)可以得到 $2k$ 个

线性方程 根据特征点间的距离约束可以得到 C_k^2 个 2 次方程, 当 $k \geq 3$ 时, $C_k^2 \geq k$, 方程数不少于未知数的个数, 最直接的求解方法是将 C_k^2 个 2 次方程组成的方程组作为约束方程组, 对 $2k$ 个线性方程组成的方程组求解, 这实际上是一个约束条件下的线性规划问题, 由于约束条件的非线性, 求解比较困难。

本文提出了一种适合 3 个特征点的求解方法, 该方法避开规划问题, 从视觉计算的物理意义入手, 通过简单的约束迭代, 求解出各特征点的三维空间坐标, 求解过程为

(1) 通过矩阵 C 的分解求出摄像机光心 P_c 的三维坐标 $W_c = (x_c, y_c, z_c)$ 。

(2) 计算光心到各特征点的单位向量 a_i 。

(3) 计算光心到各特征点的距离 R_i 。

(4) 求出各特征点的空间坐标 $S_i = R_i a_i + W_c$ 。

该算法的关键在于距离 R_i 的求解, 这里运用了特征点间的距离约束和迭代求解方法, 计算原理如图 1 所示。其中 P_i 为空间的特征点, P_c 为摄像机光心, R_i 为特征点 P_i 到光心的距离, $d_{i,j}$ 为特征点 P_i 与 P_j 间的距离, R_i 的求解过程为:

首先给定一个 R_1 的初值, 因为 $d_{1,2}$ 和 $d_{1,3}$ 已知, 可以通过几何关系分别求出 R_2 和 R_3 , 从而求出 P_2 与 P_3 间的计算距离 $d'_{2,3}$, 将其与实际距离 $d_{2,3}$ 相比较, 根据比较结果调整 R_1 的初值, 重复上面的迭代过程, 直至 $|d'_{2,3} - d_{2,3}| < \epsilon$ 。

单目视觉的计算结果虽然与实际值有一定的误差, 但特征点间的位置关系一定满足模型的约束, 这就保证了位姿计算过程中的刚性要求。

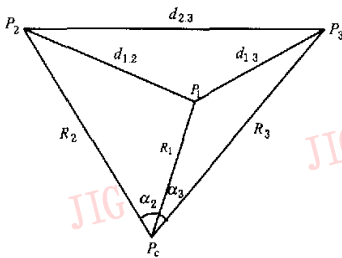


图 1 光心到特征点间距离计算原理图

3 基于模型的双目视觉

对于双目立体视觉来说, 由于增加了一个摄像机, 对于 k 个特征点, 可以得到 $4k$ 个线性方程, 完全

可以求解特征点的空间坐标, 但由于各种测量误差的存在, 使得各特征点空间坐标值的计算值与实际值间存在一定的误差, 这样特征点间的位置关系不再满足模型的约束, 位姿计算的刚性条件得不到保证, 计算结果的误差有时会大得惊人, 因此正确的解法还是在保证模型约束的条件下, 求解由 $4k$ 个线性方程组成的方程组, 文献[5]提出了一种改进的约束最小二乘法^[5]来解决这一问题, 但该方法对空间坐标的初始值要求较高, 如果初始值的误差较大, 会造成结果误差的增大。

本文将传统的立体视觉算法、基于模型的单目视觉算法, 以及约束最小二乘法融合在一起, 利用两个摄像机计算结果来相互约束实现迭代, 具体过程为:

(1) 根据立体视觉, 计算出各特征点的空间坐标 $P_{S_i} = (x_{S_i}, y_{S_i}, z_{S_i})$ 。

(2) 计算该空间坐标在两个摄像机上投影的图象坐标 $I_{L_i} = (u_{L_i}, v_{L_i}), I_{R_i} = (u_{R_i}, v_{R_i})$ 。

(3) 用 I_{L_i} 和 I_{R_i} 根据上节的算法分别计算出两组新的空间坐标

$$P_{L_i} = (x_{L_i}, y_{L_i}, z_{L_i}), P_{R_i} = (x_{R_i}, y_{R_i}, z_{R_i})$$

令

$$E = \sum_{i=1}^3 [(x_{L_i} - x_{R_i})^2 + (y_{L_i} - y_{R_i})^2 + (z_{L_i} - z_{R_i})^2] \quad (8)$$

如果 E 小于给定的误差范围, 则用任意一组空间坐标作为最后结果, 否则重新修正 I_{L_i} 和 I_{R_i} 后, 重复迭代过程。

(4) 如果迭代次数大于一定数目, E 值还仍旧很大, 则以 P_{L_i}, P_{R_i} 的均值作为空间坐标初始值, 用改进的约束最小二乘法作优化处理, 得到满足模型约束的空间坐标。

4 实验结果

以如图 2 所示的系统为例分别进行了量化误差的仿真实验和移动机器人视觉引导的真实实验, 图中被测目标由 3 个特征点 p_1, p_2, p_3 组成, p_1 较 p_2, p_3 距摄像机近, p_1, p_2 间的距离和 p_1, p_3 间的距离都为 0.5m, p_2, p_3 间的距离为 0.6m。

仿真实验时, 假定已知两个坐标系间的变换关系和摄像机的 C 矩阵, 先根据变换关系求出目标特征点在摄像机坐标系下的空间坐标, 再根据 C 矩阵求出特征点的图象坐标; 然后模拟量化误差的影响, 对图象坐标做四舍五入处理(即引进了量化误差);

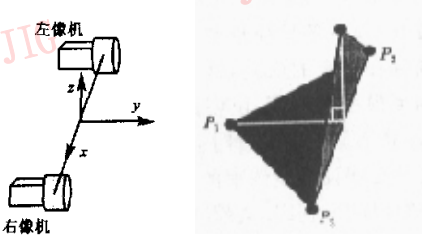


图2 实验系统示意图

最后利用上面的各种方法重新求解出两个坐标系间

的变换关系,与前面的假设值比较即可得到测量误差.实验中, x 方向的变化范围为 $2.4\text{m}\sim 12\text{m}$, y 方向的变化范围为 $6\text{m}\sim 20.4\text{m}$, z 方向始终为 0 .

图3、图4、图5分别给出了3组仿真实验结果.各图中,横坐标表示采样点的顺序,纵坐标表示误差,图中以图形表示 i 个采样点的误差.其中,位移误差为相对误差,表示 x, y, z 三个方向的位移误差与摄像机和目标间距离相对值的均方根值,转角误差为绝对误差,表示 α, β, γ 三个转角误差的均方根值.

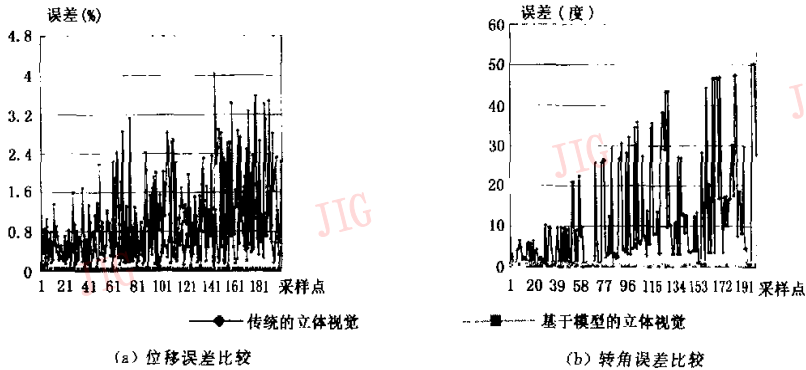


图3 模型约束的作用

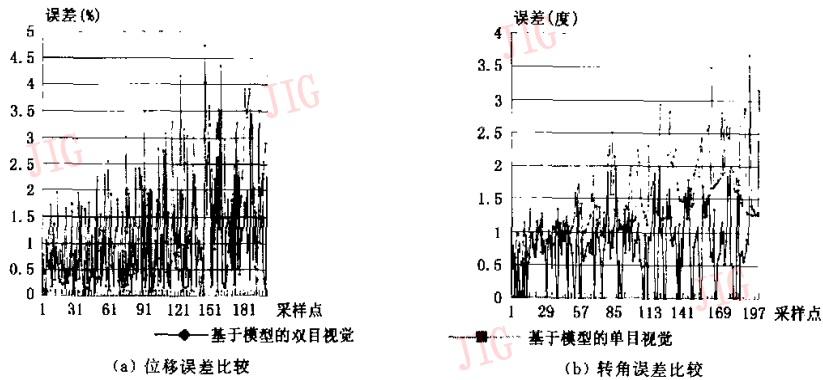


图4 基于模型的双目视觉与单目视觉比较

图3给出了传统的立体视觉与基于模型的立体视觉的比较.由此可见,当以立体视觉方法计算目标位姿时,是否利用模型约束对位移的影响不大,而对转角的影响较大,如果说位移误差尚可接受,则转角误差已经到了无法容忍的地步.

图4给出了基于模型的双目视觉和基于模型的单目视觉的比较,从中可见,由于多了一个摄像机的约束后,位姿测量误差有所减小,但幅度不大.

图5给出了基于模型的立体视觉中量化误差的

幅度对测量结果的影响.图中纵坐标为径向位移误差的相对值(相对于距离).其他几种方法都有类似的结果.表1给出了统计结果.从表中可以看出,测量结果的位移误差与量化误差几乎按同一比例增长.

表1 量化误差与测量误差的关系

量化误差 (像素)	0.25	0.50	0.25	0.50	0.25	0.50
	最大误差(%)		平均误差(%)		均方根误差(%)	
单目视觉	2.559	4.735	0.663	1.362	0.035	0.074
双目视觉	1.930	3.969	0.399	0.901	0.025	0.057

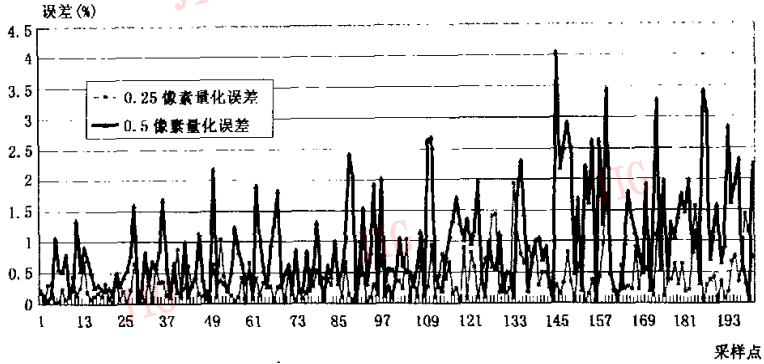


图 5 量化误差的影响

上述方法已经成功地应用于本实验室的移动机器人目标导引实验中,实验机器人为本实验室自行研制的全方位移动机器人,两个摄像机均为敏通 1881 黑白摄像机,采用 6mm 焦距的镜头,摄像机间距离为 500mm,视场角约为 40°。目标由三组发光二极管构成,实验过程中,不断将目标与机器人间的位置关系送给机器人控制器,控制机器人不断地接近目标,直至达到预定位置,目前已经在 30m 的运动范围内,分别利用基于模型的单目视觉和双目视觉方法,在目标静止与运动、不同距离、不同速度条件下进行了多次实验,成功率为 100%。图象采样速度为 0.3s。

5 结 语

本文重点讨论了利用三维视觉方法实现目标位姿测量中的模型约束问题,给出了传统的无模型立体视觉方法,基于模型的单目视觉方法和基于模型的双目视觉方法的比较。传统的无模型立体视觉方法由于无法满足位姿测量中的刚性约束,其测量误差达到了难以忍受的地步,引入模型约束后,基于模型的单目视觉方法可以达到很高的测量精度,而基于模型的双目视觉较单目视觉的测量精度有所提高,但幅度不大。无论哪种方法,图象量化误差的幅度将直接影响到视觉方法的测量精度,如何在给定的量化误差下,提高整个系统的测量精度是三维视觉测量的一个长期而艰巨的任务。

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(专刊):468~470.

2 洪海,姜伟,李金宝. 基于单视觉原理实时提取运动物体三维信息的一种方法[J]. 中国图象图形学报,1997,2(7):464~468.
 3 游素亚,徐光祐. 立体视觉研究的现状与进展[J]. 中国图象图形学报,1997,2(1),17~23.
 4 Behrooz Kamgar-parisi, Behzad Kamgar-parisi. Evaluation of quantization error in computer vision[J]. PAMI, 1989,11(9): 929~940.
 5 Hao Yingming, Zhu Feng. Optimization method of model-based stereo vision [J]. Journal of Image and Graphics, 2000, 5(Supp.):407~410.



郝颖明 1966 年生,副研究员,1987 年于大连理工大学电子工程系获工学学士学位,1990 年于中国科学院沈阳自动化研究所模式识别与智能控制专业获工学硕士学位。主要研究领域为图象处理、计算机视觉、三维检测。



朱 枫 1962 年生,研究员,1985 年于复旦大学计算机科学系获理学学士学位,1988 年于中国科学院沈阳自动化研究所模式识别与智能控制专业获工学硕士学位。主要研究领域为计算机视觉、虚拟现实、多机器人系统。



欧锦军 1975 年生,1997 年于华中理工大学电信系获工学学士学位,2000 年于中国科学院沈阳自动化研究所模式识别与智能控制专业获工学硕士学位,现在中国科学院沈阳自动化研究所工作。主要研究领域为计算机视觉、图象处理。