

基于影子的乒乓球机器人单目视觉系统标定

张远辉 韦巍 虞旦 彭博

(浙江大学电气工程学院, 杭州 310027)

摘要 视觉系统的场景标定问题是乒乓球机器人研究中要解决的首要任务。针对传统的双目视觉方案中图像处理运算量大、成本高、高速同步采集实现较为困难的缺点,探索了一种单目摄像机下的基于影子辅助成像的视觉标定方法。从单目摄像机模型出发,简化立体映射为平面映射,利用视觉场景中灯光、乒乓球和影子之间的几何关系,结合最小二乘法和最小误差估计,准确地计算视觉系统中灯与摄像机的位置,为图像识别和3维目标定位奠定基础。该方法既标定了摄像机的参数,同时也标定了辅助灯的参数,而且从平面图像坐标计算3维坐标过程简便,实验结果证明其具有较高的精度($<4\text{ mm}$),可应用于快速图像处理的场合($>60\text{ fps}$),有望成为一种低成本、便捷而有效的方法。

关键词 单目 乒乓球机器人 标定 影子

中图法分类号: TP391.4 **文献标识码:** A **文章编号:** 1006-8961(2009)09-1895-05

Shadow Based Single Camera Vision System Calibration

ZHANG Yuan-hui, WEI Wei, YU Dan, PENG Bo

(College of Electrical Engineering, Zhejiang University, Hangzhou 310027)

Abstract Vision system calibration is the first and foremost task in ping-pong robot. To reduce the deficiency of the traditional binocular vision cameras method in heavy computation and difficulty in image synchronization, a new calibration algorithm is proposed to deal with a monocular vision system. Based on the monocular camera model, the stereo mapping is transferred to the planar mapping. Using the projection geometry relationship between light, ball and shadow, both the least square method and the optimized error estimation method are used to evaluate the parameters of the vision system. Furthermore, it lays a foundation of 3D position calculation and trace prediction. This method has been proved to have high accuracy to locate the object in the scene ($<4\text{ mm}$) and high-speed image processing ($>60\text{ fps}$). It is an effective and reliable way in computer vision compared with binocular stereo vision in structure size, measurement speed, and cost. Also it has a wide application prospect.

Keywords monocular camera, ping-pong robot, calibration, shadow

1 引言

乒乓球机器人,是一个典型的实时运动伺服系统,需要融合机器视觉、运动学、智能控制等多学科知识,具有重要的研究价值。它属于一类以计算机

视觉为传感器的机器人,这类机器人首先需要对其视觉系统进行标定,才能根据摄像机捕获的图像数据,准确地获取空间物体的3维坐标信息,在此基础上实现进一步的功能开发。乒乓球机器人更有赖于视觉系统对乒乓球目标进行角度、速度和旋转方向的辨识判断,因此,视觉标定就显得尤为重要。

基金项目:浙江省人才基金项目(R105341)

收稿日期:2008-03-13;改回日期:2008-06-27

第一作者简介:张远辉(1982~),男。浙江大学控制理论与控制工程专业博士研究生。主要研究方向为机器人智能控制,计算机视觉,图像处理等。E-mail: zhangyh23@gmail.com

关于这方面的研究,最早始于美国 AT&T 实验室,由 Adnersson 领导团队开发的乒乓球机器人^[1],该装置采用 4 个摄像机和多台分布式处理器的设计,以复杂的网络并行完成图像的处理和运动控制。目前,以计算机视觉为传感器的机器人多采用双目摄像机方案^[2-3],1997 年,日本大阪大学 Miyazaki 设计的双摄像机乒乓球机器人^[4],即以双目方案定位乒乓球 3 维空间的坐标和旋转角度等。但是该双目系统的标定,不仅需要同时确定两个摄像机的内参数和外参数^[5],而且需要采用 Quick-MAG 专用图像处理^[4],往往数学运算复杂、耗时,运算量大,成本也大大增加。此外,从图像处理的实时性考虑,捕获高速运动的乒乓球(速度达到 3.5 m/s 以上)必须要求两个摄像机具有较高的采集速度(每秒 60 帧以上)^[4],且两个摄像机同步采集,这在物理实现上较为困难,亦势必增加了图像处理和定位算法的时间。2001 年,Acosta 等人在一个小型的桌面(500 × 500 mm)上,实现了巧妙应用单目摄像机方案的乒乓球机器人^[6]。该系统创立了一种单目摄像机和一个辅助灯获取乒乓球 3 维坐标的视觉系统,简化了 3 维场景到平面图像坐标的变换,简便了乒乓球 3 维坐标的计算,提高了时效性。但是,Acosta 的系统仍存在明显的不足,其视觉范围相对狭窄,对乒乓球运动轨迹识别精度要求低,且假定可直接测量得到摄像机和灯的光心位置而并未实际对其进行标定,同时该机器人在击球前不需要预测乒乓球和桌面碰撞过程。基于上述局限,本文中采用了国际标准的乒乓球桌(长 2 740 mm,宽 1 525 mm,离地面高 760 mm),图像识别区域大大增加,而 Acosta 的系统中简单的测量方法是无法满足本系统的精度要求的。本文提出的标定方法,应用单目视觉系统,借助灯、摄像机、乒乓球实体和乒乓球影子的几何关系,采用最优的误差估计手段,使得标定方法方便快捷,对于场景的改变亦具有较强的适应能力。

2 单目视觉系统的实现

2.1 单目视觉系统的场景设计

乒乓球机器人系统的场景如图 1 所示,左上角为普通白炽灯,右上角为数字摄像机(A602fc 的数字摄像机, Basler 公司)。

在世界坐标系中,灯、球、摄像机和影子之间的位置关系如图 2 所示, F 表示摄像机的光心, B 表

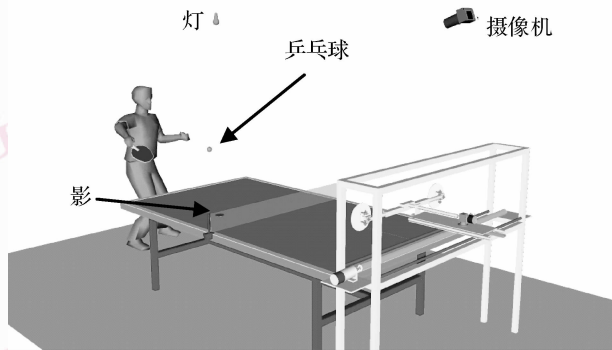


图 1 乒乓球机器人

Fig. 1 Ping-pang robot

示球, L 表示灯, S 表示影, I 表示光心 F 和 B 的延长线和桌面的交点。

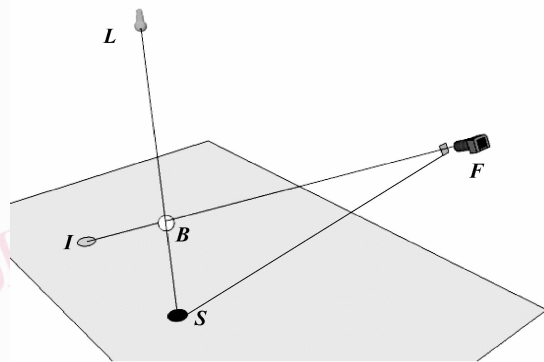


图 2 3 维的光学系统视觉模型

Fig. 2 Robot vision model

单目视觉系统的实现,基于如下假设:

- (1) 摄像机采用小孔模型^[5,7]。
- (2) 机器人场景中只有一个发光物体,即为悬挂于天花板的灯 L ,且假设这是一个点光源,即所有光线的反向延长线交于点 L 。
- (3) 乒乓球桌面为理想的水平平面。
- (4) 摄像机的成像失真可忽略不计。

对于 B 和 I ,在像素坐标系中具有相同的 (x, y) 位置。基于上述假设和几何关系阐述,根据投影原理,点 L, B, S 处于同一直线。点 F, B, I 处于同一直线。

2.2 平面映射参数标定,桌面点 I, S 3 维坐标估计

综合坐标变换和小孔成像模型^[8],从世界坐标系到像素坐标系的变换可用式(1)表示,该式反映的是空间 3 维坐标向图像 2 维坐标的一个映射关系。

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & m_3 & m_4 \\ m_5 & m_6 & m_7 & m_8 \\ m_9 & m_{10} & m_{11} & m_{12} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

其中, x_w, y_w, z_w 为世界坐标系中的3维坐标, u, v 为物体在图片中的像素坐标, z_c 为物体在摄像机坐标系下的 z 坐标。系数 $m_1 - m_{12}$ 构成投影矩阵, 该矩阵包含摄像机的内参数和外参数^[7-9]。

考察式(1), 定义世界坐标系的原点位于乒乓球桌面上, z 轴向上(如图3)。则乒乓球桌面即 xy 平面, 上面的点满足 $z_w = 0$, 可得到式(2)如下:

$$z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} m_1 & m_2 & m_4 \\ m_5 & m_6 & m_8 \\ m_9 & m_{10} & m_{12} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ 1 \end{bmatrix} \quad (2)$$

消去 z_c 可得如下公式:

$$x_w m_1 + y_w m_2 + m_4 - u x_w m_9 - u y_w m_{10} = u m_{12} \quad (3)$$

$$x_w m_5 + y_w m_6 + m_8 - v x_w m_9 - v y_w m_{10} = v m_{12} \quad (4)$$

上式中, m_{12} 是不可能为0的, 否则当平面坐标为(0,0)时, 式(2)的第3个方程不能满足。所以可将式(3)、式(4)两式的两边同时除以 m_{12} , 并重新定义参数, 得到新方程:

$$x_w r_1 + y_w r_2 + r_3 - u x_w r_7 - u y_w r_8 = u \quad (5)$$

$$x_w r_4 + y_w r_5 + r_6 - v x_w r_7 - v y_w r_8 = v \quad (6)$$

通过求解式(5)、式(6)联立的方程组, 得到如下公式:

$$x_w = \frac{(r_3 - u)(vr_8 - r_5) - (r_6 - v)(ur_8 - r_2)}{(ur_7 - r_1)(vr_8 - r_5) - (ur_7 - r_4)(ur_8 - r_2)} \quad (7)$$

$$y_w = \frac{(r_3 - u)(vr_7 - r_4) - (r_6 - v)(ur_7 - r_1)}{(ur_8 - r_2)(vr_7 - r_4) - (ur_8 - r_6)(ur_7 - r_1)} \quad (8)$$

式(7)、式(8)即是平面映射的公式, 表明了 xy 平面上的视角范围内的任意点, 可以和图片中的任意点确定一一对应的关系, 其中需要标定的参数为 r_1 到 r_8 共8个。

将 xy 平面上的一个标定点的平面坐标和像素坐标代入式(5)、式(6), 就可以确定两个线性方程。故只需4个标定点就可以确定8个方程, 从而解出8个需要标定的参数。如果采用冗余的标定点, 则可以应用最小二乘法求出需标定参数, 减少标定数据测量误差对标定的影响, 即增加了鲁棒性。计算参数后, 按式(7)、式(8), 根据图片中物体的像素坐标 u, v , 计算得到对应于球桌平面的点 I 的3维坐标。同理, 根据乒乓球影子在图片中的坐标, 可以计算点 S 的3维坐标。

2.3 基于影子的 $F, L, 3$ 维坐标的估计

将乒乓球固定于圆柱体物体的上端(如图3)构成本实验中的标定物体, 圆柱体的高度是一定的, 摆放位置垂直于桌面。灯的发光中心 L , 根据投影原理必然落在标定物体 B 和影子 S 的延长线上。将标定物体沿指定的网格位置放置, 则可以得到一系列连接点 B, S 的直线, 这些直线均通过灯的中心 L 。同理可知, 连接点 I, B 的直线也必然通过摄像机的光心 F 。

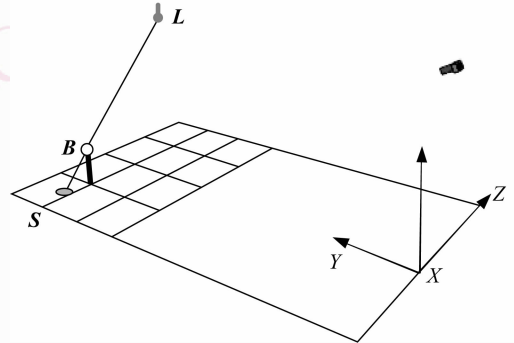


图3 标定物体的示意图

Fig.3 Object for calibration

3维欧氏空间中, 如图4所示, 通过点 X , 模为1的向量 $vector$ 上面的点 (x_1, y_1, z_1) 描述为

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \lambda \begin{bmatrix} u_x \\ u_y \\ u_z \end{bmatrix} \quad (9)$$

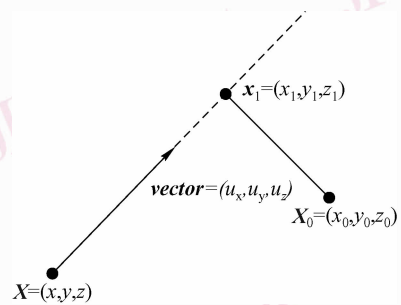


图4 点到空间直线的距离计算

Fig.4 Point to Line distance

则, 3维空间中任意点 X_0 到直线的距离为

$$d^2 = (x + \lambda u_x - x_0)^2 + (y + \lambda u_y - y_0)^2 + (z + \lambda u_z - z_0)^2 \quad (10)$$

当 d^2 取得最小值时, 求解出 λ 的最优解为

$$\lambda = \frac{u_x(x_0 - x) + u_y(y_0 - y) + u_z(z_0 - z)}{u_x^2 + u_y^2 + u_z^2} \quad (11)$$

由于 u_x, u_y, u_z 均定义为归一化向量, 即

$$u_x^2 + u_y^2 + u_z^2 = 1 \quad (12)$$

将 λ 代入式(10)得到:

$$\begin{aligned} \text{mind}^2 = & (x + x_0(\mathbf{u}_x^2 - 1) + y_0 \mathbf{u}_x \mathbf{u}_y + z_0 \mathbf{u}_x \mathbf{u}_z + \\ & \mathbf{u}_x(-x\mathbf{u}_x - y\mathbf{u}_y - z\mathbf{u}_z))^2 + (y + y_0(\mathbf{u}_y^2 - 1) + \\ & x_0 \mathbf{u}_x \mathbf{u}_y + z_0 \mathbf{u}_y \mathbf{u}_z + \mathbf{u}_y(-x\mathbf{u}_x - y\mathbf{u}_y - z\mathbf{u}_z))^2 + \\ & (z + z_0(\mathbf{u}_z^2 - 1) + x_0 \mathbf{u}_x \mathbf{u}_z + y_0 \mathbf{u}_y \mathbf{u}_z + \\ & \mathbf{u}_z(-x\mathbf{u}_x - y\mathbf{u}_y - z\mathbf{u}_z))^2 \end{aligned} \quad (13)$$

为描述方程方便,引入如下替代变量,进一步简化计算,令

$$\begin{aligned} K &= x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z \\ Cxy &= \mathbf{u}_x \mathbf{u}_y \\ Cxx &= \mathbf{u}_x \mathbf{u}_x \end{aligned}$$

代入式(13),基于最速下降求解距离最小值的原理,对 x_0, y_0, z_0 偏微分后,得到式(14)

$$\begin{cases} \frac{\partial \text{mind}^2}{\partial x_0} = -2[x_0(Cxx - 1) + y_0(Cxy) + z_0(Cxz) + (x - \mathbf{u}_x K)] \\ \frac{\partial \text{mind}^2}{\partial y_0} = -2[x_0(Cxy) + y_0(Cyy - 1) + z_0(Cyz) + (y - \mathbf{u}_y K)] \\ \frac{\partial \text{mind}^2}{\partial z_0} = -2[x_0(Cxz) + y_0(Cyz) + z_0(Czz - 1) + (z - \mathbf{u}_z K)] \end{cases} \quad (14)$$

上式描述了对于已知的点 $\mathbf{X} = (x, y, z)$ 和向量 $\mathbf{vector} = (\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z)$,空间中任意点 $\mathbf{X}_0 = (x_0, y_0, z_0)$ 到该直线的最短距离 d 满足如式(14)所示的方程组。

2.4 等效交点最优估计

对于任意一条直线,均可以构造式(14)的由 3 个线性方程组成的方程组,对于空间中任意一点到 n 条直线的距离平方和,将式(14)中各系数累加,则距离的最优解可简化为如下,即该点到所有直线的距离平方和为最短。

$$\begin{cases} x_0 \sum_{i=1}^{i=n} (C_i x x - 1) + y_0 \sum_{i=1}^{i=n} C_i x y + z_0 \sum_{i=1}^{i=n} C_i x z + \sum_{i=1}^{i=n} (x_i - u_{ix} K_i) = 0 \\ x_0 \sum_{i=1}^{i=n} C_i x y + y_0 \sum_{i=1}^{i=n} (C_i x y - 1) + z_0 \sum_{i=1}^{i=n} C_i y z + \sum_{i=1}^{i=n} (y_i - u_{iy} K_i) = 0 \\ x_0 \sum_{i=1}^{i=n} C_i x z + y_0 \sum_{i=1}^{i=n} C_i y z + z_0 \sum_{i=1}^{i=n} (C_i z z - 1) + \sum_{i=1}^{i=n} (z_i - u_{iz} K_i) = 0 \end{cases} \quad (15)$$

式中,符号 $\sum_{i=1}^{i=n}$ 表示从第 1 根直线到第 n 根直线对应系数求和。

求解式(15),可得到最优的 \mathbf{X}_0 值。采用此方法,使得目标函数取得最小值,得到乒乓球系统中,点 \mathbf{F} 和 \mathbf{L} 的最优估计。选取图 4 中点 \mathbf{X}_0 到 \mathbf{X}_1 的距离,作为点 \mathbf{F} 和 \mathbf{L} 的估计误差。测量过程中将标定物摆放于 n 处不同的位置,则可以得到通过点 \mathbf{F} 或者 \mathbf{L} 的 n 条直线,同时也得到 n 个误差的统计特征。

2.5 乒乓球 3 维坐标计算

图 2 中已经描述了点 $\mathbf{L}, \mathbf{B}, \mathbf{S}$ 处于同一直线,点 $\mathbf{F}, \mathbf{B}, \mathbf{I}$ 处于同一直线。但在实际计算中,连接 \mathbf{L} 和 \mathbf{S}, \mathbf{F} 和 \mathbf{I} 的直线,并不是相交于点 \mathbf{B} 。两条直线往往呈空间异面直线的情况,根据 \mathbf{F} 和 \mathbf{L} 的位置,以及平面映射关系求解 \mathbf{S} 和 \mathbf{I} 的位置,就获得了两直线(直线 \mathbf{FI} 和直线 \mathbf{LS})的方程。乒乓球的 3 维坐标,即为两条异面直线最短距离的线段中点^[9]。

3 实验结果和误差分析

3.1 求解等效的 \mathbf{F} 结果的示意

如图 5 所示,通过点 \mathbf{B}, \mathbf{I} 的一系列直线,相交于摄像机光心 \mathbf{F} 。点 \mathbf{B} 的位置由标记物体的实际摆放位置测量决定,点 \mathbf{I} 的位置,由式(7)、式(8)从图像坐标计算所得。

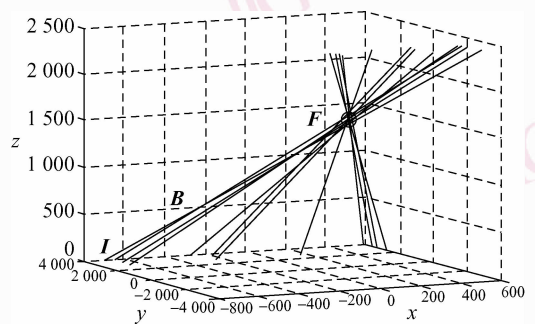


图 5 实验结果示意

Fig. 5 Result of experiment

3.2 验证实验和误差分析

验证标定方法的实验分 3 步进行:(1)采用乒乓球两桌边各 4 个点为平面映射的标定点,根据式(5)、式(6)联立的线性方程组,求解平面映射参数。(2)选取桌面上的 6 个点放置标定物,根据式(15)标定 \mathbf{F} 和 \mathbf{L} 的坐标。(3)另外选取 6 个点放置标定物,验证标定精度,用前两步得到的参数计算标定物的 3

维坐标,将计算结果与实际位置进行比较。

当摄像机工作于高速采集模式时,提供的图片分辨率为 656×490 像素,同时摄像机的观测区域为 $1\,400\text{ mm} \times 1\,400\text{ mm}$ 的平面区域,因此,理论上图片坐标中一个像素的误差引起的 3 维测量误差为 3 mm。

实验后,通过计算,点 F 和点 L 的估算误差小于 2.5 mm,测量 3 维坐标的位置误差均小于 4 mm,该结果略大于理论估计结果,原因来自摄像机畸变误差等因素,但是该结果已经完全满足乒乓球的位置测定和轨迹预测的要求(如表 1 所示)。

表 1 标定后视觉系统的测量误差

Tab. 1 Calibration error list

单位:mm

序号	X_1	Y_1	Z_1	X_2	Y_2	Z_2	误差
1	-162.5	2 428	494	-163.01	2 431.2	496.15	3.925 4
2	437.5	1 528	494	438.41	1 528.2	492.38	1.880 5
3	-462.5	2 128	494	-459.74	2 126	492.45	3.721 7
4	437.5	2 128	494	438.9	2 125.4	493.13	3.051 0
5	-462.5	1 828	494	-463.46	1 829.4	490.67	3.745 1
6	-462.5	1 528	494	-461.47	1 525.3	494.64	2.989 1

注:($X_1 Y_1 Z_1$)是实际标定物体的 3 维坐标;

($X_2 Y_2 Z_2$)是根据图像计算的 3 维坐标。

标定后的参数,可在乒乓球识别轨迹的处理中直接使用,实验证明,用该标定方法得到的参数,可以满足快速图像处理和目标定位的需求(>60 fps)。

4 结 论

针对乒乓球机器人的特定需求,设计了一种较先进的在 PC 机上可以实现的 3 维视觉系统,就此视觉系统的特殊性,提出了一种快捷的标定系统参数的方法。实验结果表明,此标定方法可以达到预期的精度,完全满足实际应用的要求。且与其他方

法相比,该单目视觉系统不需要把投影矩阵再做进一步的分解来计算摄像机的内外参数^[8-9],而只需求解线性方程组,计算量小,更加简便,具有较强的鲁棒性,在成本和控制响应时间上也有很大的优势。

参考文献 (References)

- Andersson R L. A Robot Ping-Pong Player: Experiment in Real-time Intelligent Control[M]. London, England: The MIT Press, 1988. 14-15.
- Bjorkman M, Eklundh J O. Real-time epipolar geometry estimation of binocular stereo heads[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(3): 425-432.
- Negahdaripour S. Epipolar geometry of opti-acoustic stereo imaging [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(10): 1776-1788.
- Rui Qing, Hu Zong-wu, FuMio Miyazaki. Simulation on prediction trace of Ping-pang ball using LWR learning [J]. Robot, 1998, 20(5): 373-377. [芮庆, 胡宗武, 宫崎文夫. 用 LWR 学习进行乒乓球轨道预测的仿真研究 [J]. 机器人, 1998, 20(5): 373-377.]
- Forsyth D A, Ponce J. Computer Vision: A Modern Approach[M]. New Jersey, USA: Prentice Hall, 2002:28-31.
- Acosta L, Rodrigo J J, Mendez J A, et al. Ping-pang player prototype [J]. Robotics & Automation Magazine, IEEE, 2003, 10(4): 44-52.
- Hartley Richard, Zisserman Andrew. Multiple View Geometry in Computer Vision[M]. New York, NY, USA: Cambridge University Press, 2001. 178-193.
- Zhang Z. A flexible new technique for camera calibration [J]. Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(11): 1330-1334.
- Tsai R. A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses [J]. IEEE Journal of Robotics and Automation, 1987, 3(4): 323-344.
- Lourakis Manolis I A, Argyros Antonis A. Efficient, causal camera tracking in unprepared environments[J]. Computer Vision and Image Understanding, 2005, 99(2): 259-290.