

文章编号: 0253-2239(2009)10-2832-05

一种单摄像机虚拟鼠标的标定方法

李伟仙 周富强 张鲁闽 杨晓科

(北京航空航天大学仪器科学与光电工程学院, 北京 100191)

摘要 建立单摄像机虚拟鼠标的数学模型,提出一种虚拟鼠标的现场标定方法。以二维实体靶标为中介,利用计算机屏幕生成的虚拟靶标,确定摄像机坐标系到屏幕坐标系的变换,建立空间三维控制点图像坐标到鼠标指针的映射,实现了单摄像机虚拟鼠标的现场标定。提出的方法无需高成本的辅助设备,现场操作简单。在 30 cm 作用距离内采用 640 pixel×480 pixel 图像中 50 个特征点,当图像噪声方差达到 1 pixel 时,试验中控制点映射为屏幕坐标的均方根(RMS)误差小于 3 pixel。该方法应用于研发的虚拟鼠标演示程序,表明切实可行,适用于虚拟鼻尖鼠标的现场标定。

关键词 机器视觉; 虚拟鼠标; 单目视觉; 标定; 鼻尖鼠标

中图分类号 TP391.41 **文献标识码** A **doi**: 10.3788/AOS20092910.2832

Calibration Approach to Single-Camera Virtual Computer Mouse

Li Weixian Zhou Fuqiang Zhang Luming Yang Xiaoke

(School of Instrumentation Science and Opto-Electronics Engineering, Beihang University, Beijing 100191, China)

Abstract The mathematical model of a single-camera virtual computer mouse sensor is established and a field calibration approach to the single-camera virtual mouse is proposed. The mapping from the camera coordinates in working status to the screen coordinates can be calibrated through one media planar target in the camera's field of view and one virtual target on the screen. Calibrating the single-camera virtual mouse is completed by efficiently establishing an available correspondence between the three-dimensional control point in space and the point of the mouse in the operating system. The proposed approach has an easy operation, and demands no extra costly tools. In the work range of 30 cm and with 50 features from image in 640 pixel×480 pixel, the RMS errors of the control point coordinates in the screen under test are within 3 pixels when the image noise variance reaches 1 pixel. The demo program shows the proposed approach is practicable and is applicable to field calibration for virtual mouse based on the tip of the user's nose.

Key words mechanical vision; virtual mouse; monocular vision; calibration; nose's mouse

1 引 言

随着人机交互技术的不断发展,有线及无线的有形鼠标暴露越来越多的缺点,极大限制了人机交互的速度、距离及舒适度。近年来,采用摄像机拍摄空间控制点,通过跟踪、分析控制点的图像坐标,并与计算机屏幕坐标相对应,作为输入提供给计算机,起到真实鼠标的功能,这类虚拟鼠标引起了国内外专业人士的广泛重视^[1~5]。该类基于机器视觉原理

的虚拟鼠标无辐射、非接触,大大地降低了人机交互的认知负荷,有效地实现了交互的快速性和自然性,可广泛应用于为残疾人士服务、多媒体教学、娱乐、辐射控制区应用及公共设施的传染病预防与控制等领域。现有基于视觉的虚拟鼠标主要分两大类,一类是激光笔虚拟鼠标^[6~7]。刘芳等^[1]提出的基于激光笔的人机交互技术,通过提取光点图像坐标建立与计算机屏幕坐标的对应,此类虚拟鼠标采用了求

收稿日期: 2008-11-28; 收到修改稿日期: 2009-01-20

基金项目: 教育部新世纪优秀人才支持计划(NCET-05-0194)资助课题。

作者简介: 李伟仙(1985—),女,博士研究生,主要从事计算机视觉方面的研究。E-mail: liweixian@aspe.buaa.edu.cn

导师简介: 周富强(1972—),男,教授,博士生导师,主要从事计算机视觉方面的研究。

E-mail: zfq@buaa.edu.cn(通信联系人)

解二维平面间单应性映射矩阵的方法进行标定。另一类是基于手势、眼睛视线或人脸特征等身体特征^[8~13]的虚拟鼠标。D. O. Gorodnichy 等^[2]研发的 Nouse 基于双目立体视觉技术和极线几何关系,完成鼻尖三维空间定位,实现对计算机的简单控制。该类虚拟鼠标系统的成本较高、视场范围小、标定过程复杂、立体匹配难、计算量大、实时性较差。

提出一种单摄像机虚拟鼠标的标定方法,仅需一个普通 USB 摄像机,一个二维实体靶标和二维虚拟靶标及一台计算机,设备简单,现场操作方便。该方法以实体靶标为中介,利用计算机屏幕生成的虚拟靶标,确定摄像机坐标系到非视场范围内的屏幕坐标系的变换,建立空间控制点到计算机鼠标指针的映射关系,提高了系统实时性,增大了有效视场范围。

2 单摄像机虚拟鼠标的数学模型

如图 1 所示, $o_c - x_c y_c z_c$ 和 $o_{c2} - x_{c2} y_{c2} z_{c2}$ 是观测和中介位置的摄像机坐标系, $o_s - x_s y_s z_s$ 是以屏幕左上角顶点 o_s 为原点的屏幕坐标系, $o_t - x_t y_t z_t$ 是以实体靶标平面左上角顶点 o_t 为原点的实体靶标坐标系, $o_u - x_u y_u$ 和 $o_{u2} - x_{u2} y_{u2}$ 是摄像机观测和中介位置处二维图像像素坐标系。 π_u 和 π_{u2} 分别为摄像机观测和中介位置处的理想成像平面, π_t 和 π_s 分别为实体靶标和虚拟靶标(计算机屏幕)所在平面, o_p 和 o_{p2} 为两处摄像机的主点。 P 为虚拟鼠标在三维空间的实体控制点, 在观测和中介位置处摄像机的理想像点分别为 p_u 和 p_{u2} 。 设 P 的观测处摄像机坐标为 $\mathbf{X}_c = [x_c \ y_c \ z_c \ 1]^T$, 中介位置处摄像机坐标为 $\mathbf{X}_{c2} = [x_{c2} \ y_{c2} \ z_{c2} \ 1]^T$, 实体靶标坐标为 $\mathbf{X}_t = [x_t \ y_t \ z_t \ 1]^T$, 屏幕坐标为 $\mathbf{X}_s =$

$[x_s \ y_s \ z_s \ 1]^T$, 其像点 p_u 和 p_{u2} 的理想图像坐标为 $\mathbf{X}_u = [x_u \ y_u \ 1]^T$ 和 $\mathbf{X}_{u2} = [x_{u2} \ y_{u2} \ 1]^T$ 。 图 1 中像点 p_u 和摄像机中心 o_c 确定的直线 $o_c p_u$ 称为投影射线。

观测位置摄像机透视投影模型^[14]可表示为

$$\rho \mathbf{X}_u = \mathbf{A} \mathbf{X}_c$$

其中

$$\mathbf{A} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad (1)$$

其中 ρ 是非零比例系数, \mathbf{A} 是摄像机 3×4 内参数矩阵, f_x 和 f_y 是摄像机在 x 和 y 方向上的有效焦距, (u_0, v_0) 是摄像机主点坐标。 因为两处采用同一摄像机, 所以中介位置摄像机投影模型可表示为

$$\lambda \mathbf{X}_{u2} = \mathbf{A} \mathbf{X}_{c2}, \quad (2)$$

λ 是非零比例系数。 假设像点 p_u 和 p_{u2} 的实际图像坐标为 $\mathbf{X}_d = [x_d \ y_d \ 1]^T$ 和 $\mathbf{X}_{d2} = [x_{d2} \ y_{d2} \ 1]^T$, 考虑摄像机镜头的一次径向畸变系数 k_1 , 则理想图像坐标到实际图像坐标的变换为

$$\begin{cases} x_d = x_u + k_1 [(x_u - u_0)^2 + (y_u - v_0)^2] (x_u - u_0) \\ y_d = y_u + k_1 [(x_u - u_0)^2 + (y_u - v_0)^2] (y_u - v_0) \end{cases}, \quad (3)$$

$$\begin{cases} x_{d2} = x_{u2} + k_1 [(x_{u2} - u_0)^2 + (y_{u2} - v_0)^2] (x_{u2} - u_0) \\ y_{d2} = y_{u2} + k_1 [(x_{u2} - u_0)^2 + (y_{u2} - v_0)^2] (y_{u2} - v_0) \end{cases}, \quad (4)$$

根据摄像机坐标系和屏幕坐标系及靶标坐标系的关系, 有

$$\mathbf{X}_c = \mathbf{M}_s^c \mathbf{X}_s, \quad (5)$$

$$\mathbf{X}_c = \mathbf{M}_t^c \mathbf{X}_t, \quad (6)$$

$$\mathbf{X}_{c2} = \mathbf{M}_s^{c2} \mathbf{X}_s, \quad (7)$$

$$\mathbf{X}_{c2} = \mathbf{M}_t^{c2} \mathbf{X}_t, \quad (8)$$

其中 \mathbf{M} 为 4×4 摄像机外参数矩阵, 其下标指示被表示的坐标系, 而上标指示用来表示对象的坐标系。 设 \mathbf{R} 为 3×3 单位正交旋转矩阵, \mathbf{T} 为 3×1 平移矢量, 则 \mathbf{M} 形如

$$\mathbf{M} = \begin{bmatrix} \mathbf{R} & \mathbf{T} \\ 0 & 1 \end{bmatrix}.$$

3 单摄像机虚拟鼠标的标定

3.1 摄像机内参数的标定

调整好摄像机的焦距及光圈后, 在视场范围内自由、非平行移动实体靶标至少有 3 个位置, 每移动一个位置拍摄一幅摄像机标定图像。 参照文献^[15]

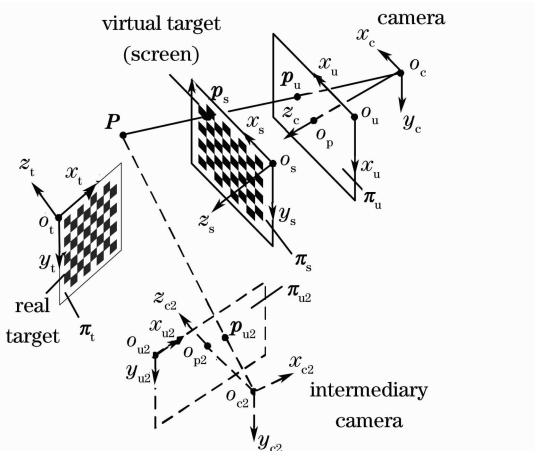


图 1 单摄像机虚拟鼠标的数学模型

Fig. 1 Mathematical mode of single-camera virtual mouse

的特征点提取方法提取所有摄像机标定图像的特征点图像坐标,根据其对应的世界坐标标定摄像机内参数^[14],包括摄像机有效焦距、主点以及畸变系数。

3.2 摄像机坐标系和屏幕坐标系的变换

标定好摄像机内参数后,标定摄像机坐标系和屏幕坐标系的变换关系。首先,实体靶标置于屏幕前方,计算机屏幕生成与实体靶标规格大小完全一致的图像,称为虚拟靶标。靶标平面与屏幕成 $90^\circ\sim 150^\circ$ 的角度,保证摄像机同时看到实体靶标和屏幕虚拟靶标上的所有特征点,此时摄像机位于中介位置。在此处拍摄一幅包含实体靶标和虚拟靶标的图像,称为双靶标标定图像。然后,在保证实体靶标和虚拟靶标位置不动的情况下,将摄像机置于观测位置(要能看到实体靶标的所有特征点),拍摄一幅包含实体靶标的图像,称为单靶标标定图像。

根据摄像机模型和已标定的内部参数,提取双靶标标定图像中的实体靶标特征点的图像坐标,利用对应的实体靶标坐标,计算实体靶标坐标系到中介摄像机坐标系的变换 M_t^c ;提取双靶标标定图像中的虚拟靶标特征点的图像坐标,利用对应的屏幕坐标,计算屏幕坐标系到中介摄像机坐标系的变换 M_s^c ;提取单靶标标定图像中的实体靶标特征点的图像坐标,利用对应的靶标坐标,计算实体靶标坐标系到摄像机坐标系的变换 M_t^c 。

因为虚拟屏幕不在观测位置的摄像机视场中,所以不能直接标定屏幕坐标系到摄像机坐标系的变换 M_s^c 。但是联立(5)式~(8)式,得变换

$$M_s^c = M_t^c (M_t^{c2})^{-1} M_s^{c2}, \quad (9)$$

根据(9)式,容易得到屏幕坐标系到摄像机坐标系的旋转矩阵 R_s^c 和平移矢量 T_s^c ,坐标系之间的变换关系如图2所示。

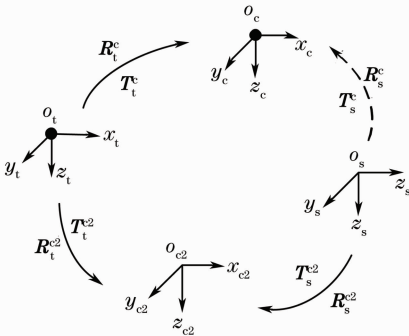


图2 坐标系之间的变换关系

Fig. 2 Transformation between different coordinate frames

3.3 单摄像机控制点到鼠标位置的对应

完全标定好单摄像机虚拟鼠标的数学模型仍不

能确定控制点 P 的三维坐标,但是虚拟鼠标旨在建立控制点 P 到屏幕平面 π_s 上鼠标指针的映射关系。取 P 的映射就是 p_s ,即控制点 P 的映射就是投影射线 $o_c p_u$ 与平面 π_s 的交点 p_s 。设点 p_s 的三维屏幕坐标为 $X_{ps} = [x_{ps} \ y_{ps} \ z_{ps}]^T$, $z_{ps} = 0$, $[x_{ps} \ y_{ps}]$ 直接作为鼠标坐标提供给计算机。在屏幕坐标系下计算映射点的过程如下:

从拍摄的图像提取像点 p_u 的实际图像坐标经校正得理想图像坐标 X_u ,则投影射线 $o_c p_u$ 在屏幕坐标系下直线方程为

$$\rho X_u = AM_s^c X_{ps}, \quad \rho \neq 0, \quad (10)$$

(10)式同时也是直线 $o_c p_s$ 的方程,即

$$\rho X_u = AM_t^c X_{ps}, \quad \rho \neq 0, \quad (11)$$

又有屏幕平面方程(忽略实际屏幕的弯曲度)

$$z_{ps} = 0, \quad (12)$$

(11)~(12)式联立,解得 $[x_{ps} \ y_{ps}]$,建立了单摄像机系统下实体控制点的图像坐标与鼠标坐标的映射。

值得注意的是,有时交点 p_s 不一定刚好在屏幕物理尺寸范围内,需要全局统一作合适比例缩放保证屏幕上任何一个位置都有空间控制点与之对应。

4 实 验

采用罗技 QuickCam Pro 5000 USB2.0 彩色摄像机和 15 inch(38.1 cm) LCD 液晶显示器构成单摄像机虚拟鼠标系统,摄像机镜头为广角镜头,工作距离约为 30cm,摄像机图像传感器为 CMOS 芯片,30 万像素分辨率,动态分辨率 640×480 ,最大帧速 15 frame /s。计算机显示分辨率 1024×768 ,可视物理尺寸 12×9 inch(30.48 cm \times 22.05 cm),屏幕显示 8×8 个黑白相间棋盘格虚拟靶标(每格为 60 pixel \times 60 pixel);打印同规格实体靶标平面,即每格实际大小为

$$25.4 \times \left(\frac{60 \times 12}{1024} \times \frac{60 \times 9}{768} \right) \text{mm}.$$

根据 3.1~3.3 节的标定方法,对单摄像机虚拟鼠标系统进行标定。其中转换矩阵 M_s^c 的旋转矩阵 R_s^c 和平移矢量 T_s^c 标定结果如下:

$$R_s^c = \begin{bmatrix} -0.848 & 0.426 & 0.316 \\ 0.529 & 0.721 & 0.448 \\ -0.037 & 0.547 & -0.837 \end{bmatrix}$$

$$T_s^c = [129.317 \ 122.604 \ 408.971]^T$$

采用高斯噪声污染双靶标和单靶标标定图像的特征点图像坐标,取噪声均值为 0,方差 $\sigma=0\sim 1.5$ pixel,进行 500 个独立排列顺序实验,图 3 是 R_s^c 和 T_s^c 的 3 个

自由度的 RMS 偏差随图像噪声方差的变化曲线。旋转矩阵 R_s^c 的 3 个自由度是欧拉角的 3 个分量 (θ, φ, ϕ) ，平移矢量 T_s^c 的 3 个自由度是 (t_x, t_y, t_z) 。下面分析像点图像噪声给屏幕映射点带来的影响。有代表性选择 5 个图像坐标作为控制像点，分别靠近图像边缘的 4 个顶点和图像中心。对每个控制像点用均值为 0，方差

$\sigma=0\sim 1.5$ pixel 的高斯噪声污染，进行 500 个独立排列顺序实验，图 4 分别是 5 个控制像点的屏幕映射 x 和 y 坐标的 RMS 偏差随图像噪声方差的变化曲线。实验表明当图像噪声方差达 1 pixel 水平时屏幕坐标 RMS 偏差小于 3 pixel。

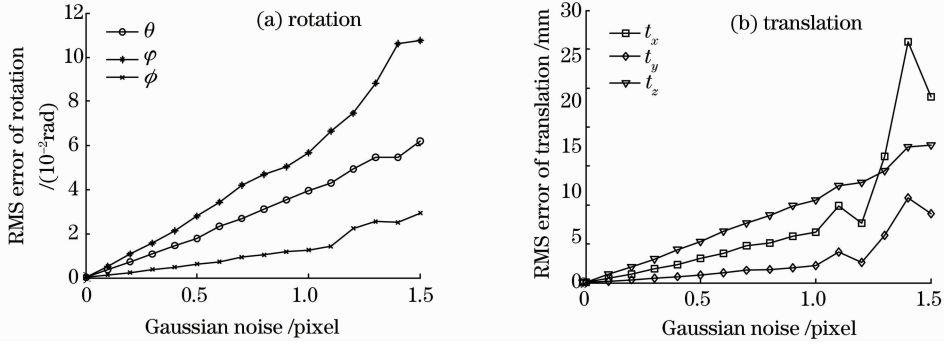


图 3 旋转矩阵和平移矢量 RMS 偏差变化曲线

Fig. 3 RMS error of rotation and translation w.r.t noise

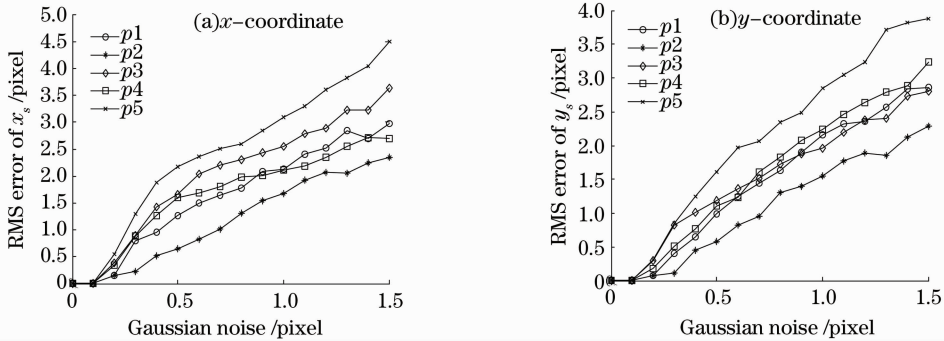


图 4 5 个控制像点的映射屏幕 x 和 y 坐标 RMS 偏差变化曲线

Fig. 4 RMS error of screen-coordinate along x and y for five control points w.r.t noise

虚拟鼠标旨在简化人机交互以实现大众化应用，现场标定方法没有苛刻的专业精度要求，因此主要有以下 3 种不可避免的误差来源：采用摄像机为普通 USB 摄像机，镜头成像质量不高，一阶径向畸变系数不足以校正像差；打印的平面靶标精度低，而且屏幕虚拟靶标尺寸和实际平面靶标尺寸不一定严格相等；平面靶标在试验时不一定严格在一个平面。尽管存在上述误差，但是基于该标定方法研发的单摄像机虚拟鼠标系统能实现控制计算机光标稳健地移动和操作，可很好的进行人机交互，比如浏览网页。

空间控制点坐标到屏幕坐标的映射。在 30 cm 作用距离内采用 640 pixel \times 480 pixel 的采集图像，当图像噪声方差达到 1 pixel 时，屏幕坐标的 RMS 误差小于 3 pixel。

参 考 文 献

- 1 Liu Fang, Lin Xueyan, Shi Yuanchun. Remote human-computer interaction using laser pointer[J]. *J. Image and Graphics*, 2003, 8A(11): 1356~1360
- 刘 芳, 林学闫, 史元春. 基于激光笔的远程人机交互技术[J]. *中国图像图形学报*, 2003, 8A(11): 1356~1360
- 2 Dmitry O. Gorodnichy, Gerhard Roth. Nose use your nose as a mouse perceptual vision technology for hands-free games and interfaces[J]. *Image and Vision Computing*, 2004, 22(12): 931~942
- 3 C. Gr tzel, T. Fong, S. Grange *et al.*. A non-contact mouse for surgeon-computer interaction [J]. *Technology and Health Care*, 2004, 12(3): 245~257
- 4 Jun Rekimoto. SmartSkin: an infrastructure for freehand manipulation on interactive surfaces [C]. *Proceedings of the SIGCHI conference on Human factors in computing systems: Changing our world, changing ourselves*, 2002, 4(1): 113~120

5 结 论

提出的单摄像机虚拟鼠标的标定方法已经成功应用于研发的鼻子虚拟鼠标，通过非视场范围内的屏幕坐标系和摄像机坐标系的变换，有效地实现了

- 5 C. Fagiani, M. Betke, J. Gips. Evaluation of tracking methods for human-computer interaction [C]. *Proceedings of the IEEE Workshop on Applications in Computer Vision*, 2002, 121~126
- 6 D. Pasquariello, M. Vissenberg, C. J. M *et al.*. Remote-touch: a laser input user-display interaction technology [J]. *J. Display Technology*, 2008, **4**(1): 39~46
- 7 Yan Haitao, Wang Ming. Orientation of mouse using digital speckle correlation method [J]. *Acta Optica Sinica*, 2008, **28**(3): 467~471
闫海涛, 王 鸣. 数字散斑相关方法实现鼠标定位原理 [J]. *光学学报*, 2008, **28**(3): 467~471
- 8 N. Berci, P. Szolgay. Vision based human-machine interface via hand gestures [C]. *18th European Conference on Circuit Theory and Design*, 2007, 496~499
- 9 Gihan Shin, Junchul Chun. Vision-based multimodal human computer interface based on parallel tracking of eye and hand motion [C]. *International Conference on Convergence Information Technology*, 2007, 2443~2448
- 10 M. Betke, J. Gips, P. Fleming. The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities [J]. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 2002, **10**(1): 1~10
- 11 Huang Ying, Wang Zhiliang, Qi Ying. Development and application of human-computer interaction simulation system based on gaze tracking [J]. *J. System Simulation*, 2008, **20**(24): 6678~6682
黄 莹, 王志良, 戚 颖. 基于视线追踪的人机交互仿真系统的研制与应用 [J]. *系统仿真学报*, 2008, **20**(24): 6678~6682
- 12 Sun Xuemei, Su Fei, Cai Anni. Illumination normalization for face recognition based on inherent texture features within micro-neighborhood [J]. *Acta Optica Sinica*, 2008, **28**(11): 2083~2089
孙雪梅, 苏 菲, 蔡安妮. 基于微观本义复原的人脸图像光照正规范化 [J]. *光学学报*, 2008, **28**(11): 2083~2089
- 13 Wang Xiangjun, Wang Yan, Li Zhi. Fast target recognition and tracking method based on characteristic corner [J]. *Acta Optica Sinica*, 2007, **27**(2): 360~364
王向军, 王 研, 李 智. 基于特征角点的目标跟踪和快速识别算法研究 [J]. *光学学报*, 2007, **27**(2): 360~364
- 14 Z. Zhang. A flexible new technique for camera calibration [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330~1334
- 15 Zhou Fuqiang. Study on Key Techniques for Binocular Stereo Vision Inspection [D]. Beijing: Beihang University, 2002
周富强. 双目立体视觉检测的关键技术研究 [D]. 北京: 北京航空航天大学, 2002