

文章编号 1004-924X(2011)06-1353-07

工业增强现实中的相机跟踪

潘绍松*, 左洪福

(南京航空航天大学 民航学院, 江苏 南京 210016)

摘要:提出了利用工业对象表面共面圆进行相机位姿跟踪的解决方案。首先,利用共面圆构建对象坐标系;用扩展卡尔曼滤波器预测相机位姿限定椭圆轮廓点搜索的位置及范围,计算梯度归一化互相关系数来形成轮廓点候选数据集;然后,剔除候选数据集中的野值,用最大似然估计法进行高精度椭圆拟合;最后,根据二次曲线的射影理论计算相机位姿并更新滤波器的状态。实验表明,本文方法跟踪相机位姿的最大误差角度为 1.4° ,平移为 3.5 mm ,跟踪速度为 $10\sim 12\text{ frame/s}$,满足工业增强现实中相机位姿跟踪的应用需求。

关键词:计算机视觉;相机跟踪;共面圆特征;扩展卡尔曼滤波;归一化互相关系数;最大似然估计

中图分类号:TP391.1 **文献标识码:**A **doi:**10.3788/OPE.20111906.1353

Camera tracking in industrial augmented reality

PAN Shao-song*, ZUO Hong-fu

(College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China)

* Corresponding author, E-mail: panshaosong@yahoo.cn

Abstract: A solution for camera tracking based on coplanar circles from industrial objects was presented. Firstly, a coordinate system for an object frame was constructed based on coplanar circles. Then, the Extended Kalman Filter(EKF) was exploited to predict camera pose to restrict the positions and regions for searching ellipse contours, and the candidate contour points of ellipse were formed by calculating normalized cross coefficient of gradient. With maximum likelihood estimation, a high accurate ellipse was obtained after eliminating outliers from candidate dataset. Finally, the camera pose was estimated using projective theory of conics, and the filter state was updated. Experiments show that the maximum pose deviation of camera tracking is 1.4° for rotation and 3.5 mm for translation, and the tracking rate is $10\sim 12\text{ frame/s}$. The results can meet the requirement of pose tracking in industrial Augmented Reality(AR).

Key words: computer vision; camera tracking; coplanar circle feature; Extended Kalman Filter (EKF); normalized cross coefficient; maximum likelihood estimation

1 引言

增强现实(Augmented Reality, AR)以数据

可视化的方式开拓了全新的应用领域,如设备维修、CFD显示、文化遗产再现等。AR应用的前提条件之一是确定并跟踪真实对象与相机间的相对位姿(AR中简称为相机位姿)。用人工标志确定

相机位姿虽然简单可靠但制约了 AR 的实际应用,因此人们研究利用自然特征估算并跟踪相机位姿,用自然特征确定相机位姿也是机器视觉、机器人导航等领域中的常见问题。

基于对象 3D 模型和自然特征确定相机位姿是常见方法,它通过构建对象模型与图像点或边特征的对应关系来估算相机位姿。点特征法将图像中稳定的特征点,用随机抽样、随机树法^[1]或粒子滤波^[2]形成 2D-3D 点对应,如 Bleser^[3]实现了基于 SIFT 点的相机位姿跟踪系统。边特征法利用 2D-3D 边对应估算对象位姿,例如 Drome^[4]最早提出了利用 3 条边对应估算对象位姿的方法,文献^[5]讨论了点、边组合求解位姿的方法。SoftPosit^[6]法通过搜索目标函数的极小值同时完成点对应及位姿估算,但该算法在遮挡情形或者对象特征模式有竞争性时失效。

在工业环境中,用点或边特征估计并跟踪相机位姿通常会遇到困难:油污和灰尘会导致对象表面点特征不稳定,弧状外形也使得边特征不明显。本文针对缺少点、边特征,但具有共面圆特征的工业对象提出了相机位姿跟踪的解决方案。本文所要求的共面圆特征模式在工业对象中普遍存在,如圆柱齿轮减速机的壳体,冲压模具的导向孔等。

如文献^[7-9]所述,共面圆特征一般只用于相机内参数标定,很少用于工业环境中估算相机位姿,原因在于快速、稳定地检测椭圆本身就是机器视觉领域中一个独立的课题,常规的椭圆检测与拟合方法会导致较大的位姿累积误差。本文将共面圆特征引入工业增强现实中用于相机位姿跟踪,综合应用扩展卡尔曼滤波(EKF)、轮廓搜索、野值剔除、最大似然估计等方法实施椭圆的检测与拟合,最终实现了估算相机位姿。这些方法的综合应用有效地解决了在工业环境中用共面圆进行位姿跟踪所遇到的困难。另外,本文假设相机内参数均为已知量。

2 相机位姿

如图 1(a)所示, M_0 和 M_1 是对象表面共面圆 Ω_0 和 Ω_1 的圆心且圆心间距离已知, M_0 为对象坐标系原点, M_0M_1 为 X 轴。图 1(b)示意出了共面圆的像 ω_0 和 ω_1 、圆心的像 m_0 和 m_1 。 v_x, v_y 是 X 和 Y 轴的消隐点。

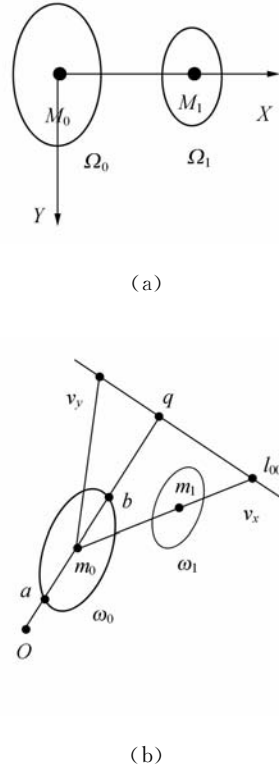


图 1 对象坐标系及其射影图像示意图
Fig. 1 Illustration for object frame and its projection

在该坐标系下相机位姿可以通过求解方程组(1)并辅以圆心间距离加以估计:

$$\begin{cases} \omega_0 : a_0x^2 + b_0xy + c_0y^2 + d_0x + e_0y + f_0 = 0 \\ \omega_1 : a_1x^2 + b_1xy + c_1y^2 + d_1x + e_1y + f_1 = 0 \\ \omega_a : x^2 + y^2 + 1 = 0 \end{cases} \quad (1)$$

其中 ω_a 为绝对二次曲线的射影图像,位姿估计的原理和具体计算过程参见文献^[10]。

3 椭圆轮廓检测

随机边缘片断选择^[11]、边分组^[12]等新颖的椭圆检测方法,虽然性能优越但并不适用于本文的应用。这是因为,当图像中有多个椭圆时,检测所有的椭圆不但耗时颇多,而且增加了区分被跟踪轮廓的难度;圆柱孔的内壁所产生的椭圆弧段通常会与被跟踪的椭圆轮廓连成一体而很难区分,如图 2 中的白色圆弧。因此,本文用计算梯度归一化互相关系数的方式隐式地搜索椭圆轮廓,同时用 EKF 提高搜索效率。



图 2 不在同一平面的椭圆轮廓
Fig.2 Non-coplanar ellipse contour

3.1 轮廓位置预测

EKF 广泛应用于视觉预测与目标跟踪^[13]。在 AR 应用中,相机位姿变化大都是局部线性、小增量的,因此,本文利用 EKF 预测相机位姿并限定椭圆轮廓变化范围。设前一时刻相机位姿为 $p = [\omega, t]^T$, 当前时刻相对于前一时刻相机位姿增量为 $\Delta p = [\Delta\omega, \Delta t]^T \in \mathcal{R}^{6 \times 1}$ (旋转分量用 Rodrigues 表达), 位姿增量的协方差矩阵为 Σ , 由 EKF 的时间更新方程预测当前相机位姿及 3D 点 x_i 投影的协方差为:

$$p^+ = p + \Delta p, \quad (2)$$

$$\Sigma^+ = J(x_i) \cdot \Sigma \cdot J^T(x_i), \quad (3)$$

式中 $J(x_i)$ 是投影 $prj(p^+, x_i)$ 的雅克比矩阵。若某个轮廓点在当前时刻实际及预测位置分别为 u_i 和 v_i , 则二者满足:

$$(v_i - u_i)^T \Sigma^+ (v_i - u_i) \leq k^2(\alpha), \quad (4)$$

$k^2(\alpha)$ 表示自由度 $n=2$ 、置信度期望为 α 的 χ_n^2 分布的逆, 如 $k^2(0.95)=5.99$ 。EKF 动态地调整轮廓搜索的起始位置和搜索范围, 避免了过大或过小地设置搜索范围带来的效率或性能的损失。根据搜索到的轮廓点完成当前相机位姿估算之后, EKF 的状态更新方程如下:

$$\Delta p^+ = \Delta p + G \cdot (m_0 - prj(p^+, M_0)), \quad (5)$$

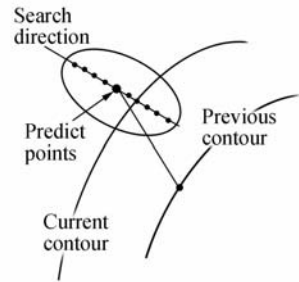
$$\Sigma^+ = (I_6 - G \cdot J(m_0)) \Sigma. \quad (6)$$

滤波器增益的更新方程较为简单, 这里没有列出。由于方程组(1)的根轨迹在某些点具有较大的跃变, 因此预测的相机位姿也作为方程组(1)求解的迭代初值, 以加快求解速度。

3.2 轮廓搜索

本文通过计算梯度归一化互相关系数的方式

搜索并构建椭圆轮廓集 D 。具体步骤如下:(1)用等间距分布的 3D 点 x_i 模拟共面圆轮廓;(2) $prj(p, x_i)$ 在前一帧图像上的位置称为采样点, 计算每一个采样点的法线方向 β ;(3)以 $[1 \ 0 \ -1]$ 为掩码沿 β 方向计算前一帧和当前帧图像的梯度, 以减轻光照变化对结果的影响;(4)在采样点沿 β 方向取矩形区域的梯度作为模板;(5)在当前图像帧中, 从预测位置 $prj(p^+, x_i)$ 开始, 在限定的区域内沿 β 方向以 0.5 像素为步长, 取与模板同方向、同尺寸的矩形区域作为目标区域;(6)依次计算每个目标区域与模板区域的归一化互相关系数, 值最大的位置作为椭圆轮廓加入集合 D ;(7)对共面圆的所有的采样点重复(4)-(6)。图 3 为轮廓点预测、搜索方向、搜索区域限制的示意图。



(a) 搜索方向及搜索范围

(a) Search direction and range

1	1	1
0	0	0
-1	-1	-1

90°

0	1	1
-1	0	1
-1	-1	0

45°

(b) 2 个不同方向的梯度算子

(b) Gradient operators of different directions

图 3 椭圆轮廓搜索示意图

Fig.3 Illustrations for contour searching

为获得亚像素精度的目标轮廓点, 矩形区域的位置采用双线性插值。本文利用 EKF 和轮廓搜索的方式替代显式的轮廓提取, 不但减轻了计算负担, 而且有效地避免了区分图像平面上多个椭圆轮廓所带来的模式竞争。

4 椭圆拟合

搜索到的数据集 D 中不可避免地含有野值,它则降低椭圆拟合精度,严重时甚至会错误地拟合出双曲线或抛物线^[14]。本文采用 RANSAC 法剔除 D 中的野值以获得最大一致集,再利用数据回归的方式提高椭圆拟合精度。

4.1 野值剔除

每次从 D 中随机抽取 6 个点形成椭圆拟合矩阵,取其 SVD 分解后的最小奇异向量作为椭圆方程的系数,这实际上是右零空间法^[15]的数值实现,拟合结果应满足 $b^2 - 4ac < 0$,否则重新抽样拟合。投票过程就是计算 D 中每个点到椭圆的 Sampson 距离,若距离大于某个阈值则认为野值。重复上述随机抽样与投票过程,直至集合 s 的势最大。

本文采用 Sampson 距离^[15]作为点到椭圆的距离度量,Sampson 距离是几何距离的高度近似,但它的计算却要简单得多。

4.2 数据回归

数字图像的光栅作用以及搜索过程中采用 0.5 步长使得 s 中的数据点含有不同程度的噪声,而非准确地位于椭圆理论曲线上。不同位置的噪化点对椭圆拟合精度的影响是各点异性的:理论轮廓内部点对结果的影响要大于外部点;大曲率处点对结果的影响要大于小曲率处点。大多数的椭圆拟合算法未考虑各点异性,因而拟合误差较大,误差累积会很快导致位姿跟踪出现较大的偏差。以高精度的椭圆拟合为目标,文献^[16-18]分别给出了不同的数据回归与椭圆拟合方法。综合比较之后,本文采用^[16]中的最大似然估计法对噪化数据进行异性回归并实施椭圆拟合。重写式(1)中一般椭圆的方程如下:

$$f + \mathbf{z}^T(\mathbf{u}_{i_0})\boldsymbol{\theta} = 0, \quad (7)$$

式中 $\mathbf{u}_{i_0} = (x_{i_0}, y_{i_0})$ 是椭圆曲线上不含噪声的未知真值点, $\mathbf{z}(\mathbf{u}_{i_0}) = (x_{i_0}^2, x_{i_0}y_{i_0}, y_{i_0}^2, x_{i_0}, y_{i_0})$ 为无噪声点构成的列向量, $\boldsymbol{\theta} = (a, b, c, d, e)^T$ 。噪化点列向量为 $\mathbf{z}_i = \mathbf{z}_{i_0} - \boldsymbol{\delta}z_i$, 其中 $\boldsymbol{\delta}z_i \sim GI(0, \sigma^2 \mathbf{C}_z)$ 是均值为零、协方差矩阵对称且半正定的噪声分布。当 \mathbf{u}_i 所含噪声为正态分布时, $f, \boldsymbol{\theta}$ 以及各 \mathbf{u}_{i_0} 可通过最小化最大似然估计式(8)获得:

$$[\hat{f}, \hat{\boldsymbol{\theta}}, \hat{\mathbf{u}}_i] = \arg \min_{f, \boldsymbol{\theta}, \mathbf{u}_{i_0}} \sum_{i=1}^n \boldsymbol{\delta}u_i^T \mathbf{C}_u^+ \boldsymbol{\delta}u_i. \quad (8)$$

实验表明,用式(8)进行数据回归,只需 3~5 次迭代即可获得满意效果,因此而增加的计算量也非常小。

4.3 算法流程

综上所述,本文提出的基于共面圆相机位姿跟踪算法可总结为图 4 所示的流程,图中的循环体即对应着连续的位姿跟踪计算。在位姿计算过程中需要的圆心距离,既可以直接测量获得,也可以从对象的 3D 模型获得,当这两种方法都不可行时,则可以通过离线 3D 重构获得。重构方法较为简单:用相机从不同位置拍摄 2~3 幅图像,然后利用平面单应检测共面椭圆特征^[19],再利用参考文献^[10]中确定圆心的方法,3D 重构后即可得到圆心间距离。此时的相机位姿可以通过其它简单方式获得,如人工标志等。



图 4 跟踪算法流程图

Fig. 4 Flow chart for tracking algorithm

5 实验与结果

本文的算法用 C 语言实现,图像处理调用了 OpenCV 库函数,向量、矩阵运算调用了 CLAPACK 库函数以提高效率。程序运行硬件环境为 Xeon(TM)CPU 为 2.80 GHz, RAM 为 2 GM。

选择的工业对象是飞机翼肋上的传动操纵杆,如图 6 所示,白色轮廓为被跟踪的共面圆。该零件表面几乎没有稳定的点特征,且边缘特征也不明显,但具有共面圆特征。实验中由 Panasonic WVCZ302 摄像头、嘉恒 MC10A 采集卡完成图像采集,六自由度位置跟踪器 Fastrak (摄像头与 Fastrak 固连)的读数经滤波处理后作为位姿真

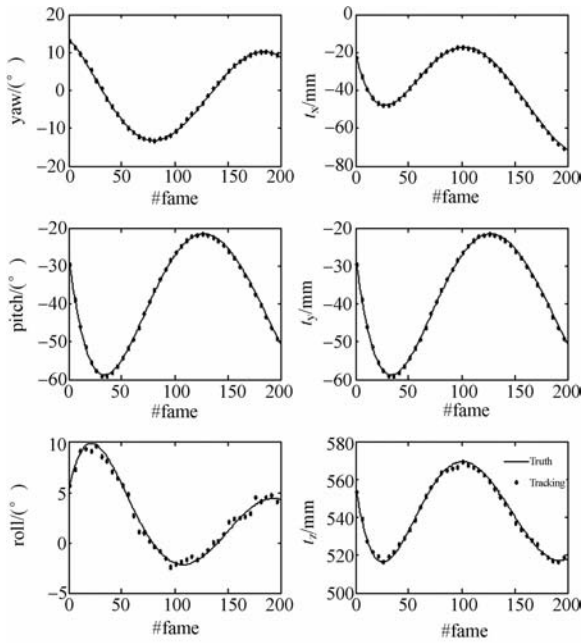


图 5 跟踪性能曲线

Fig. 5 Tracking performance

值,图像尺寸为 480×640 。

共面圆轮廓用 50 个等间距点模拟,为剔除野值设定的 Sampson 距离阈值为 3.5。图 5 给出连续 200 帧视频图像,本文方法跟踪相机位姿结果,从上到下,左侧依次为 yaw, pitch, roll 角(单位:°),右侧为 X, Y, Z 轴的平移(单位:mm),为清晰起见,位姿真值全部画出,而跟踪结果按 5 : 1 做了部分显示。

实验还对比了基于 SIFT 特征点的位姿跟踪性能,由于所选择的对象表面缺少显著特征点,作者在其表面用随机画“X”的方式生成了若干 SIFT 点。这些点的 3D 坐标由 3D 重构获得, SIFT 匹配采用了具有线性匹配时间性能的改进型方法^[20],位姿计算则采用文献[4]的方法。

表 1 中 a, b 分别列出了本文算法与 SIFT 特征法跟踪相机位姿误差的最大值、均值及方差,由表 1 知本文位姿跟踪角度最大误差为 1.4° , 平移为 3.5 mm。数据对比发现两种方法在单项跟踪误差上各有长短,但本文方法提供了缺少 SIFT 特征、但具有共面圆特征的工业对象的位姿跟踪解决方案。此外本文方法在跟踪速度上具有明显优势:实验表明按表 4 计算一次相机位姿耗时 80~100 ms,即处理视频图像的速度为 10~12 frame/s,而一幅图像仅提取 SIFT 特征耗时约为 200 ms。

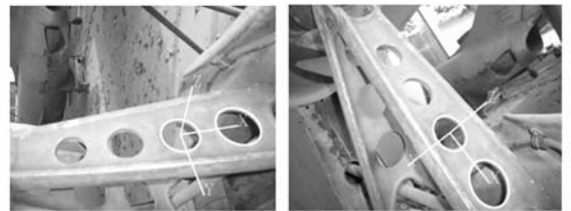
表 1 位姿跟踪误差

Tab. 1 Deviation of camera tracking

类别	最大误差		平均误差		方差	
	a	b	a	b	a	b
Yaw/(°)	0.17	0.77	-0.17	-0.02	0.23	0.35
Pitch/(°)	1.20	1.28	-0.01	-0.21	0.50	1.22
Roll/(°)	1.40	1.14	0.30	0.12	1.16	1.11
T_x /mm	0.82	0.64	0.02	0.32	0.38	0.34
T_y /mm	0.96	1.03	0.16	0.03	0.42	0.43
T_z /mm	3.53	2.11	0.23	-0.25	1.30	1.04

实验中还发现用 EKF 进行轮廓位置预测较固定设置搜索范围要节约 1 倍以上的时间,这说明本文使用 EKF 预测轮廓位置是有效的。

图 5 给出了位姿跟踪过程中对部分视频帧注册出对象坐标系坐标轴的情形,而且从 200 帧开始,将被跟踪的共面圆由相邻位置切换至非相邻位置,由注册结果可以看出:位姿跟踪结果与共面圆的相互位置无关;本文的方法在位姿跟踪上具有较高的精度,否则切换跟踪对象将直接导致跟踪失败。



(a) Frame 37

(b) Frame 134



(c) Frame 201

(d) Frame 265

图 6 AR 跟踪注册实例

Fig. 6 Tracking and registry instance for AR

6 结 论

本文针对缺少点、线特征,但具有共面圆特征的一类工业对象,提出了利用共面圆建立对象坐

标系,进行连续视频相机位姿跟踪的方法。该算法用相机位姿预测和轮廓搜的方式检测椭圆轮廓,避免了轮廓对应的难题;剔除野值数据后用最大似然估计得到高精度的椭圆拟合;基于射影二次曲线的位姿计算代价非常小。视频图像实验表

明,本文方法位姿跟踪的最大角度误差为 1.4° ,最大平移误差为 3.5 mm,跟踪速度为 $10\sim 12$ frame/s。这表明针对工业增强现实中具有共面圆特征的对象,本文的相机位姿跟踪方法可以满足 AR 的应用需求。

参考文献:

- [1] 王涌天,林精敦. 随机树特征匹配算子性能研究[J]. 北京理工大学学报,2009,29(11): 988-993.
WANG Y T, LIN J D. A study on the feature matching performance of randomized trees [J]. *Transactions of Beijing Institute of Technology*, 2009,29(11):988-993. (in Chinese)
- [2] 李利军,刘伟,等. 基于粒子滤波的虚实配准方法研究[J]. 工程图学学报,2009(1):119-124.
LI L J, LIU W. A new registration method based on particle filter [J]. *Journal of Engineering Graphics*, 2009(1):119-124. (in Chinese)
- [3] BLESER G, PASTARMOV Y. Real-time 3D camera tracking for industrial augmented reality applications [J]. *Journal of WSCG*, 2005:47-54.
- [4] DHOME M, RICHTIN M, LAPRESTE J T. Determination of the attitude of 3D objects from a single perspective view [J]. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1989,11(12):1265-1278.
- [5] ANSAR A, DANIILIDIS K. Linear pose estimation from points or lines [J]. *IEEE Trans. Pattern Anal. Mach. Intell.*, 2003,25(5):578-589.
- [6] PHILIP DAVID, DANIEL DEMENTHON. Soft-POSIT: simultaneous pose and correspondence determination [J]. *Intl. Journal of CV*, 2004, 59(3):259-284.
- [7] CHEN Q, WU H Y. Camera calibration with two arbitrary coplanar circles [C]. *Proc. European Conf. CV*, 2004:521-532.
- [8] WANG G, WU Q. Pose estimation from circle or parallel lines in a single image [J]. *Lecture Notes in Computer Science*, 2007,4844:363-372.
- [9] GURDJOS P. Euclidean Structure from $N\geq 2$ parallel circles: theory and algorithms [C]. *Proceedings of the 9th European Conference on Computer Vision, Graz, Austria*, 2006(1):238-252.
- [10] 潘绍松,左洪福. 利用共面圆特征实现增强现实图像注册[J]. 光电工程,2010(8):81-85.
PAN SH S, ZUO H F. Registration in augmented reality from coplanar circles [J]. *Opto-Electronic Engineering*, 2010(8):81-85. (in Chinese)
- [11] KAEWAPICHAJ W, KAEWTRAKULPONG P. Robust ellipse detection by fitting randomly selected edge patches [C]. *World Academy of Science, Engineering and Technology*, 2008,48:30-34.
- [12] THANH M N, SIDDHANT A. A real-time ellipse detection based on edge grouping [C]. *Proceedings of the 2009 IEEE international conference on Systems, Man and Cybernetics*, 2009: 3280-3286.
- [13] 赵其杰,屠大维. 基于 Kalman 滤波的视觉预测目标跟踪及其应用[J]. 光学 精密工程, 2008,16(5):937-942.
ZHAO Q J, TU D W. Kalman filter based vision predicting and object tracking method and its application [J]. *Opt. Precision Eng.*, 2008,16(5): 937-942. (in Chinese)
- [14] 孙宏伟,房建成. 椭圆拟合方法在磁罗盘罗差校准中的应用[J]. 光学 精密工程, 2008,17(12): 3034-3039.
SUN H W, FANG J CH. Application of ellipse fitting method to calibration of magnetic compass deviation [J]. *Opt. Precision Eng.*, 2008,17(12): 3034-3039. (in Chinese)
- [15] HARTLEY R, ZISSERMAN A. *Multiple View geometry in Computer Vision* [M]. Second Edition. London: Cambridge University Press, 2003.
- [16] BOGDAN M P M. Reduction of bias in maximum likelihood ellipse fitting [C]. *Proc. Int. Conf. Patt. Recog.*, 2000(13):3802-3086.
- [17] KANATANI. Compact algorithm for strictly ML ellipse fitting [C]. *Proc. Int. Conf. Patt. Recog.*, 2008, Tampa, FL, 2008:1-4.
- [18] KENICHI, PRASANNA. Hyperaccurate ellipse fitting without iterations [J]. *Memoirs of the Faculty of Engineering*, 2010(44):42-49.
- [19] VINCENT E, LAGANIERE R. Detecting planar homographies in an image pair [C]. *Proceedings*

of the 2nd International Symposium on Image and Signal Processing and Analysis, Pula, Croatia, 2001:182-187.

[20] OFIR P, WERMAN M. A linear time histogram

metric for improved SIFT matching [C]. Proceedings of the 10th European Conference on Computer Vision, Berlin, Heidelberg, 2008, 5304:495-508.

作者简介:



潘绍松(1973—),男,安徽滁州人,博士研究生,1996年于北京航空航天大学获学士学位,2006年于东南大学获硕士学位,主要从事增强现实维修诱导、人机交互等方面的研究。E-mail: pan-shaosong@yahoo.cn

导师简介:



左洪福(1959—),男,湖南茶陵人,教授,国务院学位委员会学科评议组成员,民航局特聘专家,1989年于徐州矿业大学获博士学位,主要从事民用航空器健康管理、数字化维修等方面的研究。E-mail: rms@nuaa.edu.cn

●下期预告

水平式经纬仪指向误差统一补偿技术

薛向尧^{1,2},高云国¹,韩光宇¹,邵帅¹,乔健¹

(1.中国科学院长春光学精密机械与物理研究所,吉林长春130033;

2.中国科学院研究生院,北京100039)

为修正水平式经纬仪的指向误差,提出了一种视轴指向的统一补偿模型。根据水平式经纬仪的光机结构,建立照准坐标系和地平坐标系,并根据两坐标系的几何关系得到目标在地平坐标系下的坐标方程,并对该方程进行全微分,得到地平坐标误差与指向误差的关系式。针对设备主要误差源之一的3轴误差,进行5次线性变换,推导出目标在地平坐标系中关于3轴误差的统一地平坐标误差方程,结合全微分所得关系式,经整理最终求出指向误差关于3轴误差的统一补偿模型。将该模型与编码器误差模型线性叠加后获得水平式经纬仪指向误差的统一补偿模型。最后对全天分布较为均匀的46颗恒星进行观测,得到观测误差的实验数据,利用最小二乘法对该模型进行拟合,得到模型中各待定系数。实验结果表明:采用该模型进行修正后,设备总指向精度由修正前的提高到,满足系统总体提出的精度要求。