

# 基于扩展 Kalman 预测模型的双目立体视觉 图像特征点跟踪方法

任少盟 魏振忠 张广军

(北京航空航天大学精密光机电一体化技术教育部重点实验室 北京 100191)

**摘要:** 针对双目立体视觉动态测量中图像特征点快速跟踪问题,提出了一种利用 Kalman 预测模型进行图像点跟踪的方法。该方法以图像特征点坐标值为观测向量,将坐标值和图像特征点运动速度作为状态向量,首先建立了线性 Kalman 预测递推模型,在此基础上引入了极线约束条件,进一步建立了图像特征点非线性扩展 Kalman 预测递推模型。实验表明该方法得到的特征点的预测轨迹和真实轨迹符合度好,预测精度高、速度快、通用性强。

**关键词:** Kalman 预测模型;极线约束;圆斑中心提取;图像点跟踪;双目立体视觉

**中图分类号:** TP391.41 **文献标识码:** A

## Tracking method for stereo vision image feature points based on extend Kalman prediction model

Ren Shaomeng Wei Zhenzhong Zhang Guangjun

(Key Laboratory of Precision Opto-mechatronics Technology, Ministry of Education, Beihang University, Beijing 100191)

**Abstract:** Aiming at the problem of rapid tracking for stereo-camera image feature points, a forecast method base on extend Kalman prediction model is proposed. It defines that the observed vector is the coordinates of the image feature points, and the state vector is the coordinates and its speed. Therefore a linear Kalman prediction iterative model is derived. Based on that, a nonlinear extend Kalman prediction model is proposed when taking into account the epipolar constraints in stereo-camera vision system. At last experimental results prove that it's well accordant between the prediction trace and the real one and this method has the advantages of high forecast precision, rapid tracking speed and wide availability.

**Keywords:** Kalman prediction; epipolar constraints; circular facula extraction; image points tracking; stereo vision

### 0 引言

视觉动态测量系统在众多领域中具有广泛应用,如飞机运动姿态<sup>[1]</sup>以及舵面转角动态测量<sup>[2]</sup>,航空视觉导航<sup>[3]</sup>,飞机构件风洞实验动态形变测量<sup>[4]</sup>等。而要实现动态测量需在保证测量精度的条件下保证测量速度,进行特征跟踪是提高动态测量性能的关键技术之一。

目前,图像特征跟踪的方法主要有全帧搜索方法和设定感兴趣矩形区域(即 ROI)方法。全帧搜索数据量和运算量大,影响测量速度。ROI 图像特征跟踪方法主要思想为在前一帧图像的位置设定一个感兴趣矩形区域(ROI),在 ROI 区域内利用图像处理技术进行特征提取实现跟踪,该类方法主要有:基于模板的跟踪、基于区域的跟踪、基于轮廓的跟踪、基于特征的跟踪和基于运动特性的跟踪<sup>[5-6]</sup>。具

体的算法有:均值平移法<sup>[7]</sup>(mean shift),目标跟踪算法(CAMSHIFT)<sup>[8]</sup>,基于特征点的跟踪方法<sup>[9]</sup>等。该类方法要求目标特征在 2 帧图像之间坐标位置变化不能太大,因而限制了其实际应用范围。

对于点特征来说,所占图像像素较少,如果为了保证目标始终在 ROI 区域内,需要设定较大的区域范围,就会降低测量速度。如果区域范围设定过小,图像点又会超出区域,导致跟踪失效。因此,针对图像特征点跟踪问题,提出利用 Kalman 预测模型根据前一帧的位置首先进行图像点坐标预测,在预测位置设定适当 ROI 区域,然后进行特征点精确提取。

当前的 Kalman 滤波原理在立体视觉中主要应用于结合当前其他已知信息得到被测物体位置和姿态<sup>[10-11]</sup>等,主要是基于 Kalman 滤波方程。而所提方法是基于预测方程

加入了预测环节,可以在跟踪过程中缩小 ROI 区域的大小,从而提高了测量速度,并且在 ROI 区域内采取的基于图像处理的亚像素级特征提取又不会降低测量精度,因而能够在保证测量精度的前提下提高测量速度。

## 1 双目立体视觉图像特征点跟踪原理

为实现双目立体视觉系统对特征点的跟踪,采用基于 Kalman 预测原理进行跟踪预测。考虑到估计值的精度并不能满足测量的要求,因此在估计出的图像点位置的 ROI 区域内采用光斑中心的亚像素级提取方法<sup>[12-13]</sup>,得到图像点更加精确的坐标  $m_i^*$  和  $m_r^*$ ,以此坐标值更新观测向量来对下一时刻进行预测。因此特征点跟踪的步骤为:

- 1) 在  $k$  时刻利用亚像素提取方法精确定位图像特征点中心  $m_i^*(k), m_r^*(k)$ ;
- 2) 结合  $m_i^*(k), m_r^*(k)$ , 根据 Kalman 预测模型得到  $k+1$  时刻最优估计值  $\hat{m}_i(k+1), \hat{m}_r(k+1)$ ;
- 3) 根据估计值,设置 ROI 区域,进入第 1) 步,以此循环下去。

## 2 图像特征点 Kalman 预测模型

### 2.1 图像特征点线性 Kalman 预测模型

设图像特征点  $m = (u, v)$ ,  $u, v$  分别为像素横纵坐标,令观测向量  $Z = (u, v)$ , 状态向量  $X = (u, v, \dot{u}, \dot{v})$ , 则离散状态方程与观测方程分别为<sup>[14]</sup>:

$$X(k+1) = \Phi(k+1, k)X(k) + \Gamma(k+1, k)W(k) \quad (1)$$

$$Z(k) = H(k)X(k) + V(k) \quad (2)$$

$$\text{式中: } \Phi(k+1, k) = \begin{bmatrix} 1 & 0 & T & 0 \\ 0 & 1 & 0 & T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, W(k) = (\ddot{u}, \ddot{v})\Gamma(k+1, k)$$

$$= \begin{bmatrix} T^2/2 & 0 \\ 0 & T^2/2 \\ T & 0 \\ 0 & T \end{bmatrix}, V(k) = \begin{pmatrix} \Delta u \\ \Delta v \end{pmatrix} H(k) =$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, T \text{ 为等时间采样间隔, 并且设 } W(k), V(k)$$

为相互独立且均值为零的白噪声。

由于 Kalman 滤波的状态方程和观测方程都是线性的, 因此给定初值  $X(0|0_-) = X(0)$ , 和误差方差阵  $P_0, R_0, Q_0$ , 使用如下预测递推方程<sup>[15]</sup>:

$$K(k) = \Phi(k+1, k)P(k|k-1)H^T(k) * [H(k)P(k|k-1)H^T(k) + R_k]^{-1} \quad (3)$$

$$\hat{X}(k+1|k) = \Phi(k+1, k)X(k|k-1) + K(k) * [Z(k) - H(k)X(k|k-1)] \quad (4)$$

$$P(k+1|k) = \Phi(k+1, k)P(k|k-1)\Phi^T(k+1, k) - K(k)H(k)P(k|k-1)\Phi^T(k+1, k) + \Gamma(k+1, k)Q_k\Gamma^T(k+1, k) \quad (5)$$

式中:  $P(k)$  为估计误差方差阵,  $R(k)$  为  $V(k)$  协方差阵,  $Q(k)$  为  $W(k)$  协方差阵。

在双目立体视觉测量模型中, 左右图像坐标分别为  $m_1 = (u_1, v_1), m_2 = (u_2, v_2)$ , 则只需要将以上的观测向量和预测向量分别相应扩展为:  $Z = (u_1, v_1, u_2, v_2), X = (u_1, v_1, u_2, v_2, \dot{u}_1, \dot{v}_1, \dot{u}_2, \dot{v}_2)$ , 式(1)和(2)进行相应扩展即可。

### 2.2 图像特征点扩展 Kalman 非线性预测模型

对于双目立体视觉测量模型来说, 线性 Kalman 预测的方法并没有充分利用双目系统的结构与内部参数的约束。设空间点相对应的左右图像点齐次坐标分别为  $M_1 = (u_1, v_1, 1)^T, M_2 = (u_2, v_2, 1)^T$ , 它们满足极线约束关系<sup>[16]</sup>:

$$M_2^T F M_1 = 0 \quad (6)$$

式中:  $F$  为基本矩阵,  $F = A_2^{-T} S R A_1^{-1}, A_1, A_2$  为摄像机内参数,  $R$  为双目立体视觉系统旋转矩阵,  $S$  为由系统平移矢量  $T$  定义的反对称矩阵。

将极线约束方程式(6)作为观测方程的一部分, Kalman 预测的状态方程和状态向量  $X$  不改变, 观测向量扩展为  $Z = (u_1, v_1, u_2, v_2, 0)$ , 则此时的观测方程组为<sup>[17-18]</sup>:

$$\begin{cases} u_1(k) = u_1(k) \\ v_1(k) = v_1(k) \\ u_2(k) = u_2(k) \\ v_2(k) = v_2(k) \\ 0 = M_2^T F M_1 = h[X(k)] = \\ h(u_1(k), v_1(k), u_2(k), v_2(k)) \end{cases} \quad (7)$$

由于观测方程是非线性的, 需要对其线性化。对于扩展 Kalman 最优预测线性化, 只有在误差要选取的泰勒展开点  $X^*(k)$  必须和真实点  $X(k)$  的误差  $\Delta X(k) = X(k) - X^*(k)$  较小时才能使用, 因此, 这里选取  $X^*(k) = \hat{X}(k|k-1)$ , 即为  $k$  时刻的最优估计值, 既能保证整个预测过程为无偏最优估计, 又能使得预测误差最小<sup>[19]</sup>。因此有:

$$h[X(k)] = h[\hat{X}(k|k-1)] + H(k)[X(k) - \hat{X}(k|k-1)] \quad (8)$$

其中:

$$H(k) = \frac{\partial h}{\partial X} \Big|_{X=\hat{X}(k|k-1)} \quad (9)$$

在此基础上, 根据式(3)中的递推关系来进行最优估计。

非线性预测模型是在线性模型基础上引入了非线性约束, 优点在于极线约束的补偿会让预测值更加准确。然而由于线性化使得式(9)中  $H(k)$  也包含了一定的误差。仿真实验表明引入非线性约束后预测精度有了提高。

### 2.3 图像特征点跟踪误差补偿方法

在刚体周期性旋转测角任务过程中, 在转动边界处由于转动方向的变化会导致离散采样产生转动方向突变的情况, 即在旋转边界处产生的误差最大。但是在旋转边界处由于要改变方向就需要先减速再加速, 则此阶段的平均运动速度最小, 等间隔采样点最密集。因此在旋转边界处取预测值为前一时刻的观测值, 即:  $\hat{X}(k+1) = Z(k)$ 。设旋转

的左右边界角度分别为  $\theta_l, \theta_r$ , 则补偿方法为:

$$\begin{cases} \hat{\mathbf{X}}(k+1) = \mathbf{Z}(k), & \theta < \theta_l + \zeta \text{ or } \theta > \theta_r - \zeta \\ \hat{\mathbf{X}}(k+1) = \hat{\mathbf{X}}(k+1 | k), & \text{else} \end{cases} \quad (10)$$

式中:  $\zeta$  为正的角度阈值, 在动态测量过程中可以根据实际情况确定补偿参数。虽然进行了以上补偿处理, 但为了保证 Kalman 预测模型的无偏性, 其递推过程还需要进行, 只是在取值时根据补偿条件判断取舍。

### 3 实验结果

#### 3.1 数字仿真实验

##### 1) 实验数据设定

设空间中任意一点  $P$ , 在左相机坐标系中的坐标为:  $\mathbf{X}_p = [-13.708 \ 0.346 \ 728.153]^T$ , 它绕着固定轴  $\mathbf{U}_r = [-0.0785 \ -0.9969 \ 0.0058]^T$  转动, 转动角度  $\theta(t) = 5 + 20\sin(\pi t/6)$ ,  $t \in [0, 30]$ , 等间隔  $T = 0.2$  取值。  $P$  点的空间轨迹如图 1。

摄像机的内外参数设置为:

① 左右摄像机内参数分别为  $\mathbf{A}_1$  和  $\mathbf{A}_2$ :

$$\mathbf{A}_1 = \begin{bmatrix} 2210.44 & 0 & 646.47 \\ 0 & 2210.12 & 501.94 \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{A}_2 = \begin{bmatrix} 2196.99 & 0 & 596.13 \\ 0 & 2196.46 & 448.84 \\ 0 & 0 & 1 \end{bmatrix}$$

② 左右摄像机旋转向量  $u_1, u_2$ , 平移向量  $T_1, T_2$ :

$$u_1 = [2.063 \ 2.028 \ 0.283]^T, u_2 = [-2.071 \ -2.002 \ 0.376]^T, T_1 = [-82.518 \ -45.517 \ 805.860]^T, T_2 = [-30.658 \ -31.576 \ 870.445]^T.$$

图像点加上均值为零, 标准差为 0 到 2.0 pixel 的高斯白噪声。

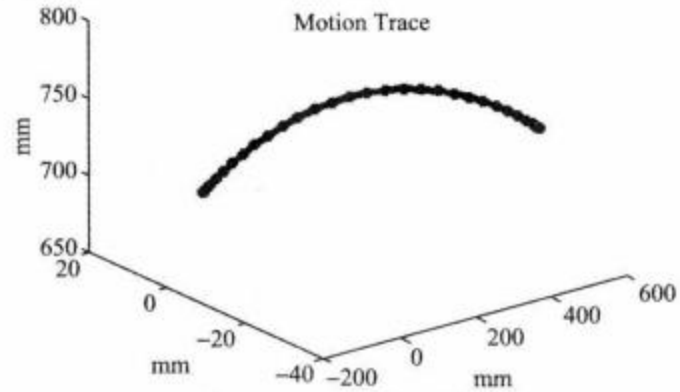


图 1 空间点运动轨迹

##### 2) 初值的选取

考虑到状态方程误差大于观测方程, 因此不妨取  $\mathbf{P}_0 = \mathbf{R}_0 = \mathbf{I}, \mathbf{Q}_0 = 2\mathbf{I}$ , 其中  $\mathbf{I}$  为单位矩阵。取状态向量初值为:

$$\mathbf{X}_0 = [u_1(0), v_1(0), u_2(0), v_2(0), u_1(T) - u_1(0), v_1(T) - v_1(0), u_2(T) - u_2(0), v_2(T) - v_2(0)]^T \quad (11)$$

$$v_1(T) - v_1(0), u_2(T) - u_2(0), v_2(T) - v_2(0)]^T \quad (12)$$

对于线性 Kalman 预测估计模型, 取观测向量初值  $\mathbf{Z}_0 = [u_1(0), v_1(0), u_2(0), v_2(0)]^T$ , 非线性预测模型观测向量初值  $\mathbf{Z}_0 = [u_1(0), v_1(0), u_2(0), v_2(0), 0]^T$ 。

##### 3) 实验结果

图 2 和图 3 分别给出图像点标准差为 2.0 pixel 时, 使用线性和非线性预测模型的误差对比图。

表 1 2 种估计模型得到的预测结果

时刻	20			30			40		
结果	线性/pixel	非线性/pixel	真实/pixel	线性/pixel	非线性/pixel	真实值/pixel	线性/pixel	非线性/pixel	真实/pixel
ul	569.12	522.04	501.47	1077.64	1075.86	1079.34	1761.70	1761.12	1751.11
vl	505.83	507.77	509.60	463.98	465.41	465.38	410.09	410.67	412.43
ur	482.07	442.75	424.02	993.63	997.14	1004.96	1837.75	1835.40	1825.53
vr	497.13	499.52	501.90	443.34	446.47	445.42	350.58	356.63	359.05
时刻	50			60			70		
结果	线性/pixel	非线性/pixel	真实/pixel	线性/pixel	非线性/pixel	真实值/pixel	线性/pixel	非线性/pixel	真实/pixel
ul	1848.16	1850.67	1834.61	1226.81	1229.22	1227.29	564.47	569.05	578.50
vl	411.98	404.55	405.75	454.72	454.08	453.84	508.08	504.20	503.78
ur	1962.56	1964.23	1943.22	1172.86	1168.11	1169.84	483.06	485.97	496.24
vr	351.43	343.31	346.13	426.28	427.84	428.65	494.19	495.21	495.10

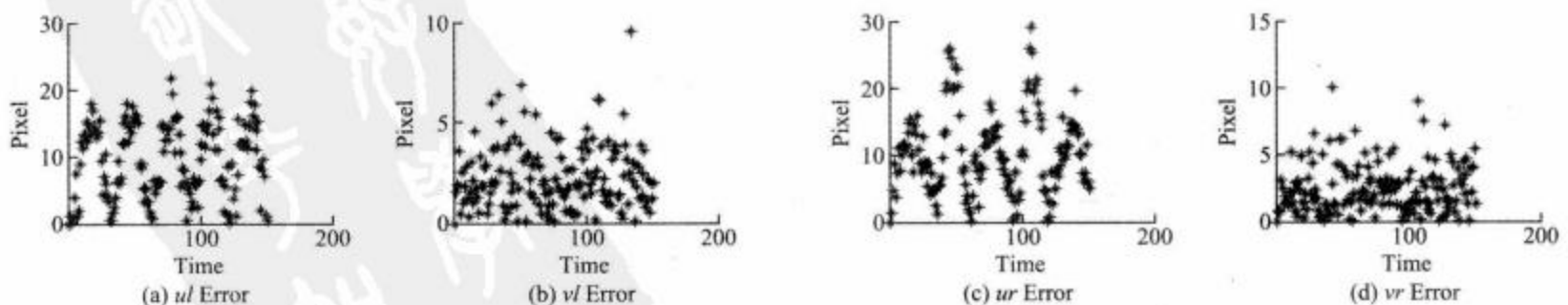


图 2 线性预测误差



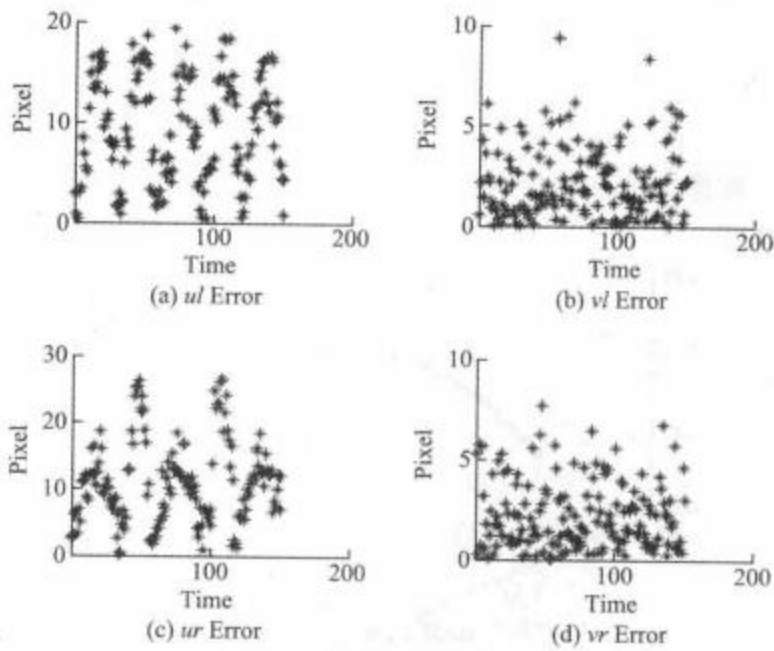


图 3 非线性预测误差

如表 1 给出分别使用线性和非线性估计模型得出的在运动过程中的 6 个时刻(20,30,40,50,60,70)的预测结果(其中  $ul, vl, ur, vr$  分别表示左图像特征点的  $u, v$  坐标和右图像特征点的  $u, v$  坐标, 以下所有图表中该符号意义与此相同)。图 4 为非线性预测模型得到的图像坐标点 RMS 误差图。

由图 2 和图 3 误差对比结果及表 2 中不同预测方法得到的误差值可以看出, 非线性模型预测精度高于线性预测精度, 说明极线约束的引入提高了预测精度。

图 4 中图像坐标 RMS 误差表明当图像点误差在 0~2.0 变化时, RMS 误差的变化较小。这是由于在整个跟踪过程中预测误差主要来源于运动的反向, 图像点较小的误差对于反向位置误差影响较小。

表 2 图像特征点预测误差对比

误差(pixel)	$u1$	$v1$	$u2$	$v2$
无预测	50.46	4.15	60.12	6.16
线性	16.36	1.92	19.08	2.45
非线性	10.47	1.03	12.17	1.35

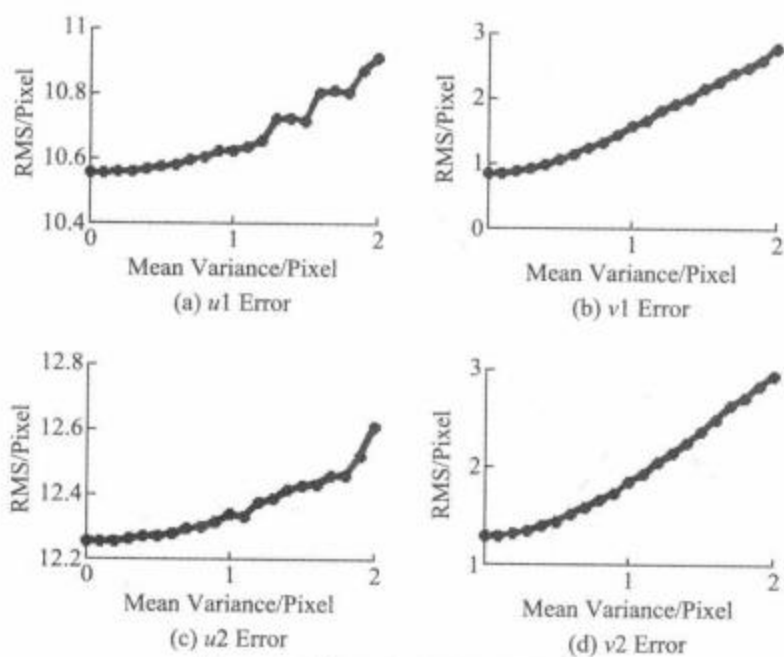


图 4 图像坐标点 RMS 误差

### 3.2 实验

#### 1) 实验条件

在实验室环境下搭建双目立体视觉测量系统。如图 5 所示, 该系统由 4 个部分组成: 1 为高精度转台, 2 为模拟舵面, 3 为计算机软件界面, 4 为双目立体视觉传感器。其中摄像机的分辨率为  $1280 \times 960$  pixel, 帧频为 7.5 Hz。转台转动速度为  $20^\circ/\text{s}$ , 转角范围为  $-30^\circ \sim 30^\circ$ 。2 个摄像机利用平面方格靶标已经标定<sup>[14]</sup>, 标定结果如表 3, 图 6 为测量过程中左右摄像机拍到的 2 幅真实图像:

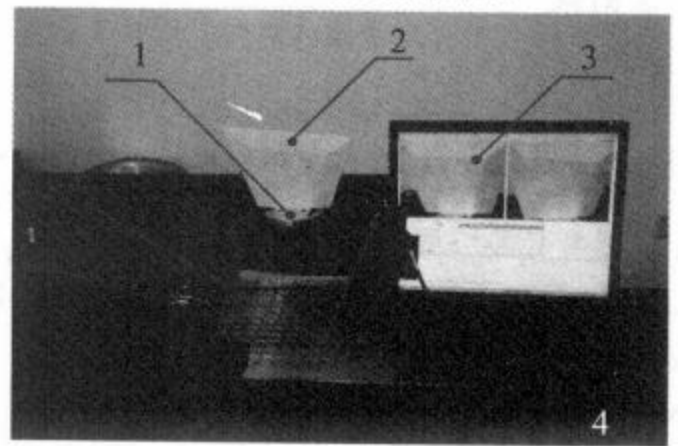


图 5 实际实验系统图

表 3 摄像机标定结果

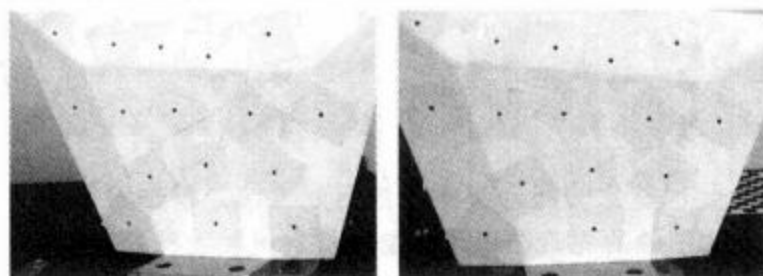
名称	结果
左摄像机内参数	$A_1 = \begin{bmatrix} 3562.27 & 0 & 811.91 \\ 0 & 3562.16 & 633.40 \\ 0 & 0 & 1 \end{bmatrix}$
右摄像机内参数	$A_2 = \begin{bmatrix} 3551.78 & 0 & 772.68 \\ 0 & 3552.03 & 619.58 \\ 0 & 0 & 1 \end{bmatrix}$
结构参数	$R = \begin{bmatrix} 0.7820 & -0.0004 & 0.6232 \\ 0.0008 & 0.9999 & -0.0004 \\ -0.6232 & 0.0008 & 0.7820 \end{bmatrix}$ $T = \begin{bmatrix} -381.12 \\ -0.75 \\ 144.87 \end{bmatrix}$

#### 2) 实验结果

这里取式(9)中  $\xi$  为  $0.5^\circ$ ,  $P_0 = R_0 = I$ ,  $Q_0 = 2I$ , 对图像中的 16 个特征点进行预测, 图 7 为测量过程中其中一帧左右图像内标记特征点在图像上的坐标预测结果(图中左下角为标记点 2 和 14 的局部放大图)。图 8 为标记点 16 在整个转动过程中扩展 Kalman 预测模型的预测误差图。

从图 8 误差结果可以看出, 预测误差偏大的位置在横坐标轴上对应的时刻是相同的, 这是因为在这些时刻舵面处于转动边界状态。在转动的边界位置, 舵面转动方向会产生突变, 因此预测误差偏大。并且明显可以看出非线性模型预测误差小于线性预测模型。如果设置 ROI 区域的

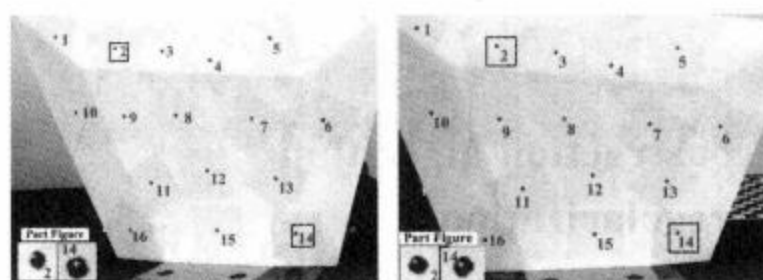
大小为  $35 \times 35$ , 在此区域内采用图像处理技术进行特征点精确提取, 就能准确地实现图像特征点的跟踪。表 4 为特征点提取时间对比结果, 表明加入的预测环节使得 ROI 区域缩小, 从而提高了图像特征点的跟踪速度。



(a) 左相机图像

(b) 右相机图像

图 6 真实图像



(a) 左摄像机

(b) 右摄像机

图 7 左右摄像机图像预测结果

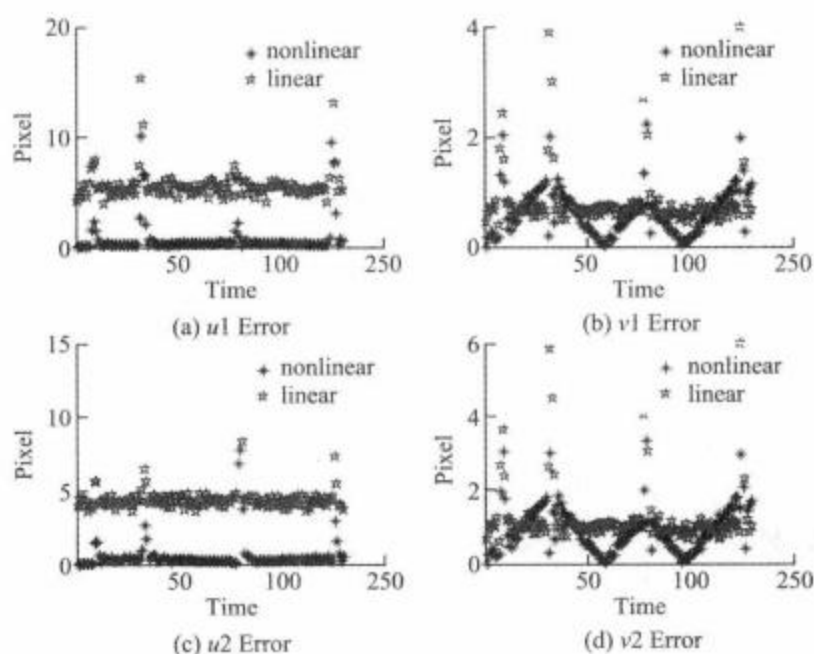


图 8 预测误差对比

表 4 图像特征点跟踪时间对比

方法	ROI/(Pixel×Pixel)	时间(ms)
有预测环节	$35 \times 35$	62
线性预测	$41 \times 41$	94
非线性预测	$51 \times 51$	156

## 4 结 论

针对双目立体视觉图像特征点动态跟踪问题, 提出了

一种基于 Kalman 预测模型的图像特征点跟踪方法, 并利用极线约束条件进行了非线性扩展。该方法中图像特征点亚像素提取方法保证了测量精度, 基于 Kalman 预测模型的预测方法缩小了 ROI 区域, 提高了测量速度。实验结果表明该方法具有速度快、精度高的特点。

## 参 考 文 献

- [1] 李沛, 罗武胜, 李冠章. 一种基于双目立体视觉原理的飞机尾旋运动姿态测量方法[J]. 国防科技大学学报, 2008, 30(2): 107-111.
- [2] 张广军, 刘震. 基于线结构光的四通道舵偏角同步动态测量系统[J]. 仪器仪表学报, 2010, 31(8): 1910-1915.
- [3] 张广军, 周富强. 基于双圆特征的无人机着陆位置姿态视觉测量方法[J]. 航空学报, 2005, 26(3): 344-348.
- [4] 张孝梯, 蒋甲利. 视频测量方法在风洞模型姿态角测量中的应用[J]. 实验流体力学, 2005, 19(3): 21-25.
- [5] 刘宏鼎, 秦世引. 双目轮式移动机器人的运动目标识别与跟踪[J]. 智能系统学报, 2007, 2(3): 19-24.
- [6] 王进花, 曹洁. 一种基于特征融合的点特征目标跟踪算法[J]. 电子测量与仪器学报, 2010, 24(6): 536-541.
- [7] CHENG J, YANG J. Real-Time infrared object tracking based on meanShift[J]. Institute of Image Processing & Pattern Recognition, 2004: 45-52.
- [8] 刘冬冬. 基于双目立体视觉和 Camshift 算法的目标检测与跟踪[D]. 山东: 山东大学硕士论文, 2006.
- [9] 郑林, 刘泉. 一种基于特征点的跟踪算法[J]. 武汉理工大学学报, 2006, 30(3): 398-400.
- [10] QIN Y Y, ZHANG H Y, et al. Kalman filter and Combined Navigation[M]. Northwestern Polytechnical University Press, 1998.
- [11] YANG M, ZHONG X Q, et al. 3D motion estimation from a stereo image sequence using dual-sequential-kalman-filter[C]. International Workshop on Imaging Systems and Techniques, 2009(11): 114-118.
- [12] 邱伟, 刘向东, 靳永强, 等. 基于双目视觉的相对状态自主确定[J]. 上海航天, 2010, 27(2): 9-13, 28.
- [13] 张广军, 杨珍. 一种圆形光斑亚像素中心提取方法及装置[P]. 中国专利: 2008.
- [14] STEGER C. An unbiased detector of curvilinear structures[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(2): 113-125.
- [15] 王洪斌, 郑瑾. 双目立体视觉基于修正卡尔曼滤波器的运动估计[J]. 控制工程, 2007, 14(2): 220-223.
- [16] 周凤岐, 卢晓东. 最优估计理论[M]. 北京: 高等教育出版社, 2009.

(下转第 84 页)