

本质流形上的自适应运动估计*

朱良家,董国华,周宗潭,王正志

(国防科技大学机电工程与自动化学院,湖南长沙 410073)

摘要 分析了计算机视觉中一种新的运动估计模型,将运动估计作为一个参数定义在“本质流形”上的非线性系统参数辨识问题。针对基于离散处理和连续处理方法得到的“本质流形”,提出了一种新的、统一的自适应推广卡尔曼滤波算法,在“本质流形”的局部坐标系上进行运动估计。仿真结果验证了该算法的正确性。

关键词 运动估计;本质流形;自适应卡尔曼滤波

中图分类号:TP391 文献标识码:A

Adaptive Motion Estimation on the Essential Manifold

ZHU Liang-jia, DONG Guo-hua, ZHOU Zong-tan, WANG Zheng-zhi

(College of Mechatronics Engineering and Automation, National Univ. of Defense Technology, Changsha 410073, China)

Abstract A novel formulation of the visual motion estimation problem is analyzed in terms of identification of the nonlinear implicit system with parameters on the so called “essential manifold”; and a new and unified implicit adaptive extended kalman filter (IAEKF) is proposed to perform the estimation in the local coordinates, for both discrete and continuous cases. The correctness of the algorithm is illustrated by numerical simulations.

Key words motion estimation; essential manifold; IAEKF

运动视觉估计问题是计算机视觉中的主要研究问题之一。它主要研究如何从变化着的一系列不同时刻的图像中,提取出物体的形状、位置和运动信息。研究运动视觉的方法大体可以分为两类:一类是基于特征的方法,另一类是基于光流场的方法(也称作离散处理的方法和连续处理的方法)。但是,大多数都可以归结为基于射影几何的方法。随着计算机视觉在无人车自主导航、机器人、视觉伺服等领域的广泛应用,必须使视觉和控制更好地结合起来。近年来,一些学者提出了一种基于微分几何的方法来解决运动视觉的问题,并在无人机着陆和移动机器人自主导航等领域中得到了成功的应用^[1,2]。

1 本质流形

1.1 基于离散方法得到的本质流形

考虑三维空间一个由特征点集合来描述的刚体。选取如下坐标系:投影中心为坐标系的原点, Z 轴沿光轴方向, X 、 Y 轴平行于成像平面,且与 Z 轴组成右手坐标系。则一个特征点就可以用坐标 $X = [X, Y, Z]^T \in R^3$ 来表示。考虑透射投影模型,则有特征点 X 在成像平面的坐标为 $x = [\frac{X}{Z}, \frac{Y}{Z}, 1]^T$ 。如果某特征点在 t 时刻的坐标为 X ,在 $t+1$ 时刻的坐标为 X' ,刚体在 t 和 $t+1$ 之间的平移向量为 T ,那么该特征点在给定坐标系中的坐标和在成像平面中的坐标的约束关系^[3]

$$X'^T T A R X = x'^T T A R x @ x'^T Q x = 0 \quad (1)$$

称为极线约束或者本质约束。

考虑点 $g = (T, R) \in SE(3)$,其中 $T \in R^3$, R 为转动矩阵,也称为 $SO(3) \times R^3$ 中的特殊正交变换

* 收稿日期:2003-10-19

基金项目:国家自然科学基金重点资助项目(60234030)

作者简介:朱良家(1979—),男,硕士生。

群); $SE(3)$ 为 R^3 中的特殊欧氏群; $T\Lambda \in SO(3)$ 是一个反对称矩阵($T\Lambda$ 矩阵是向量 T 的外积)^[4]。不失一般性,假设 $\|Q\|_2 = \|T\| = 1$,定义“本质矩阵”空间为:

$$E^d @ \{SR | R \in SO(3), S = T\Lambda \in SO(3), \|T\| = 1\} \subset R^{3 \times 3} \tag{2}$$

其中, E^d 中用上标 d 表示基于离散方法得到的本质流形,以和后面基于光流场(或连续方法)得到的本质流形相区分。

这样,就可以把 E^d 看做一个五维的微分流形^[2],即所谓的本质流形。考虑如下在 E^d 和 R^5 之间定义的一个局部映射:

$$\begin{cases} \Phi^d: E^d \rightarrow R^2 \times SO(3) \rightarrow R^2 \times R^3 \\ Q^\alpha \begin{bmatrix} \pm U \cdot 3 \\ UR_z(\pm \frac{\pi}{2})V^T \end{bmatrix} = \begin{bmatrix} T \\ e^{\Omega\Lambda} \end{bmatrix} \alpha \begin{bmatrix} T \\ \Omega \end{bmatrix} \end{cases} \tag{3}$$

其中, Φ^d 中上标 d 的意义同 E^d , U, V 是由 $Q = U\Sigma V^T$ 的奇异值分解(SVD)得到的矩阵; $U \cdot 3$ 表示矩阵 U 的第三列, $R_z(\frac{\pi}{2})$ 表示绕 001^T 轴转动 $\frac{\pi}{2}$ 得到的矩阵, T 为单位球面上通常的球坐标。如果不考虑映射中的正负号,则映射 Φ^d 定义了本质流形的一个局部坐标系。通过加入“正深度”的约束,也就是每一个可见点都在观察者的前方,就可以消除这种不确定性^[6,7]。因此,就能够得到 R^5 中一个惟一的局部同胚,给本质流形确定了一个局部坐标卡。 Φ^d 的逆映射为:

$$(\Phi^d)^{-1}: R^3 \times R^3 \rightarrow E, \begin{bmatrix} T \\ \Omega \end{bmatrix} \alpha (T\Lambda) e^{(\Omega\Lambda)} \tag{4}$$

更多的细节可见文献[5]。

1.2 基于连续方法得到的本质流形

由离散处理得到的本质流形来估计运动参数的方法只适用于相邻时刻特征点的位移(特别是平移向量)相对比较大的情形。在实时应用中,往往会由于照相机的位移较小而使得算法变得具有奇异性,估计结果不够准确。而用连续的方法进行运动估计就是从光流场中恢复照相机(或者是物体)的运动信息。

设 $u = \dot{x}$ 为光流场在坐标 x 处的值(v, ω)为照相机的运动速度,那么类似地可有连续情况下的极线约束方程为^[1]:

$$[u^T, x^T] \begin{bmatrix} \omega\Lambda \\ s \end{bmatrix} x @ [u^T, x^T] Q'x = 0 \tag{5}$$

其中 $v\Lambda$ 的意义同 $T\Lambda$, s 是由 $s = \frac{1}{2}((\omega\Lambda \times v\Lambda) + (v\Lambda \times \omega\Lambda))$ 定义的对称矩阵。由于方程(5)是关于 v 的齐次方程,因此,可以假设 v 为单位向量。则连续情况下的本质流形定义为:

$$E^c = \left\{ \begin{bmatrix} v\Lambda \\ s \end{bmatrix} \mid \omega \in R^3, v \in S^2 \right\} \subset R^{6 \times 3} \tag{6}$$

可以证明,对于 $Q' \in E^c$ 存在惟一的 (v, ω) 为其解^[1]。类似也可以把 E^c 看做一个五维的微分流形。由于 $Q', -Q' \in E^c$ 都满足方程(5),因此 $(\pm v, \omega)$ 都为方程的解。类似离散的情形,通过加入“正深度”的约束,就可以消除这种不确定性^[8]。这样就可以得到 R^5 中一个惟一的局部同胚,给本质流形确定一个局部坐标卡。

2 本质流形上的运动估计

2.1 离散本质流形上的运动估计

在刚体上选取 N 个特征点,那么当刚体在空间运动时,在成像平面就可以得到一系列对应的坐标 x^i 和相应光流的值 u^i 。这样,对于每一个点,极线约束方程(1)可以写成如下形式:

$$x^i(t+1)^T Qx^i(t) = 0, \forall i = 1KN \tag{7}$$

通常,只能得到关于 x^i 的一个带有误差的观测值, $y^i = x^i + n^i, n^i \in (0, R_{n^i})$ 其中 R_{n^i} 为观测误差的方差。

因此,对于离散的情形,运动估计问题就转化为如下模型辨识问题:

$$\begin{cases} x^i(t+1)^T Q x^i(t) = 0, Q \in E^d \\ y^i = x^i + n^i, \forall i = 1, \dots, KN, n^i \in N(0, R_n^i) \end{cases} \quad (8)$$

在每一时刻,可以得到如下形式的 N 组约束方程^[2,7]:

$$F_{x^i(t), n^i(t)}(Q(t)) = 0 \quad (9)$$

随着时间的变化,点 $Q(t)$ 在本质流形 E^d 上的轨线满足方程:

$$Q(t+1) = Q(t) + n_Q(t) \quad (10)$$

同时,在 E^d 的局部坐标系中也相应地确定了一条轨线。其中, $n_Q(t)$ 为未知量,要对 Q 进行估计就必须对 $n_Q(t)$ 进行一些假设。将方程(10)和(8)联立,可有:

$$\begin{cases} Q(t+1) = Q(t) + n_Q(t); Q \in E^d \\ 0 = F_{x^i(t), n^i(t)}(Q(t)) \\ y^i = x^i + n^i; \forall i = 1, \dots, KN \end{cases} \quad (11)$$

由此可知,视觉运动估计问题可以看做上述定义在本质流形上包含隐式观测和未知输入的动力学系统的状态估计问题。可以证明上述模型在一定条件下是全局可观的。这个条件是观察者的轨迹和刚体上的特征点不能恰好同时含于嵌入 R^3 中的一个二次曲面上,是一个通有的条件^[9]。

2.2 连续本质流形的运动估计

类似地,对于连续情形,运动估计问题也可以转化为如下动力学系统的状态估计问题:

$$\begin{cases} [x^i(t)^T, n^i(t)^T] Q^i x^i(t) = 0, Q^i \in E^c \\ 0 = F'_{x^i(t), n^i(t)}(Q^i(t)) \\ y^i = x^i + n^i, \forall i = 1, \dots, KN, n^i \in N(0, R_n^i) \end{cases} \quad (12)$$

3 运动估计问题的求解

采用滤波的方法,用统一的形式来求解方程(11)和(12)所确定的运动参数。下面将不区分 E^d 和 E^c , 统一用 E 来表示,而且 $F_{x^i(t), n^i(t)}(Q(t))$ 和 $F'_{x^i(t), n^i(t)}(Q^i(t))$ 统一记为 $F(Q(t))$, 观测值不加区分地记为 $y(t)$ 和 $y'(t)$, 在 E 和 R^5 之间定义如下映射:

$$\begin{cases} \Phi: E \rightarrow S^2 \times R^3 \rightarrow R^5 \\ Q a \xi @ \begin{bmatrix} T \\ \Omega \end{bmatrix} \end{cases} \quad (13)$$

其中, T 为单位球面上通常的球坐标。那么系统(11)和(12)在局部坐标系上可以表示为:

$$\begin{cases} \xi(t+1) = \xi(t) + n_\xi(t); \xi(t_0) = \xi_0 \\ 0 = F(\Phi^{-1}(\xi(t))) + m(t) @ \bar{F}(\xi(t), t) + m(t) \end{cases} \quad (14)$$

这样,就得到了一个 R^5 中运动的非线性动力学模型。由于这是一个未知的运动,可以假设运动的动力学模型为 R^5 中的一阶随机行走模型: $n_\xi(t) \in N(0, R_n)$, R_n 为模型误差的方差。 $y(t)$ 和 $y'(t)$ 是方差为 R_x 的零均值高斯白噪声。文献[2]中给出了上述模型在考虑误差 $m(t)$ 的相关性情形时的估计方法,即所谓的 Implicit Extended Kalman Filter (IEKF), 并且提出一种简化的方法,假设 $m(t)$ 是一个白色过程。该方法假设模型误差和观测误差的方差都已知。而在实际应用中,观测误差的方差 R_x 相对比较容易确定^[10]。模型误差的方差 R_n 通常是通过参数调整的方法获得。但是一般 R_n 不是常数,而且有时候还可能发生剧烈的变化。可以通过采用自适应滤波的方法,在估计运动参数的同时对 R_n 进行辨识来解决这个问题。我们采用了文献[11,12]中的方法估计 R_n 。记 $C @ (\frac{\partial \bar{F}}{\partial \xi})$, $D @ (\frac{\partial \bar{F}}{\partial x})$, 则具体的估计方程如下。

预测方程:

$$\hat{\xi}(t+1|t) = \hat{\xi}(t|t); \hat{\xi}(0|0) = \xi_0 \quad (15)$$

$$P(t+1|t) = P(t|t) + \hat{R}_n(t|t); P(0|0) = P_0 \quad (16)$$

滤波方程：

$$\hat{\xi}(t+1|t+1) = \hat{\xi}(t+1|t) - L(t+1)\bar{F}(\hat{\xi}(t+1|t), t) \quad (17)$$

$$P(t+1|t+1) = \Gamma(t+1)P(t+1|t)\Gamma^T(t+1) + L(t+1)R_m(t+1)L^T(t+1) \quad (18)$$

增益矩阵：

$$L(t+1) = P(t+1|t)C^T(t+1)\Lambda^{-1}(t+1) \quad (19)$$

$$\Gamma(t+1) = I - L(t+1)C(t+1) \quad (20)$$

$$\Lambda(t+1) = C(t+1)P(t+1|t)C^T(t+1) + R_m(t+1) \quad (21)$$

残差协方差矩阵：

$$R_m(t+1) = D(t+1)R_x D^T(t+1) \quad (22)$$

R_n 的估计方程：

$$R_n^* = \Delta\xi(t)\Delta\xi(t)^T + P(t+1|t) - P(t+1|t+1) - \hat{R}_n(t+1|t) \quad (23)$$

$$\Delta\xi(t) = \hat{\xi}(t+1|t+1) - \hat{\xi}(t+1|t) \quad (24)$$

$$\hat{R}_n(t+1|t+1) = \hat{R}_n(t+1|t) + (R_n^* - \hat{R}_n(t+1|t))/L_{R_n}; \hat{R}_n(0|0) = R_{n0} \quad (25)$$

4 实验结果及分析

在离观察者距离 2m 的单位立方体内 随机产生 N 个特征点 (对于离散处理的方法, $N=20$; 连续处理的方法, $N=80$)。场景以透射投影的方式投影到分辨率为 500×500 个像素的像平面上。其中, 照相机的焦距是 1m, 视角为 50° 。观测噪声为高斯噪声, 其标准差为 1 个像素。刚体的运动为匀速运动, 序列的长度为 50 帧。其中, 离散处理的方法处理的是一个绕 X 轴转动和沿 Y 轴平动的合成运动, 连续处理的方法处理的是一个绕 X 轴转动和沿 Z 轴平动的合成运动。在图 1(a)(b) 中分别给出了由离散处理的方法得到的平移和转动速度的 6 个分量的估计结果。当初值在真值的 40% 以内变化时, 算法在 30 帧以内收敛 (误差 $< 5\%$)。初值利用 Longuet-Higgins 算法求取。图 2(a)(b) 中给出了由连续处理的方法得到的平移和转动速度的 6 个分量的估计结果。当初值在真实值的 40% 以内变化时, 平移速度在 20 帧以内收敛 (误差 $< 5\%$), 而转动速度收敛较慢。初值利用文献 [1] 中的算法求取。由实验结果可以看出, 离散处理的方法对转动速度估计的效果要好于对平移速度的估计, 而连续处理的方法则恰好相反。而且连续处理的方法对观测噪声相对比较敏感。

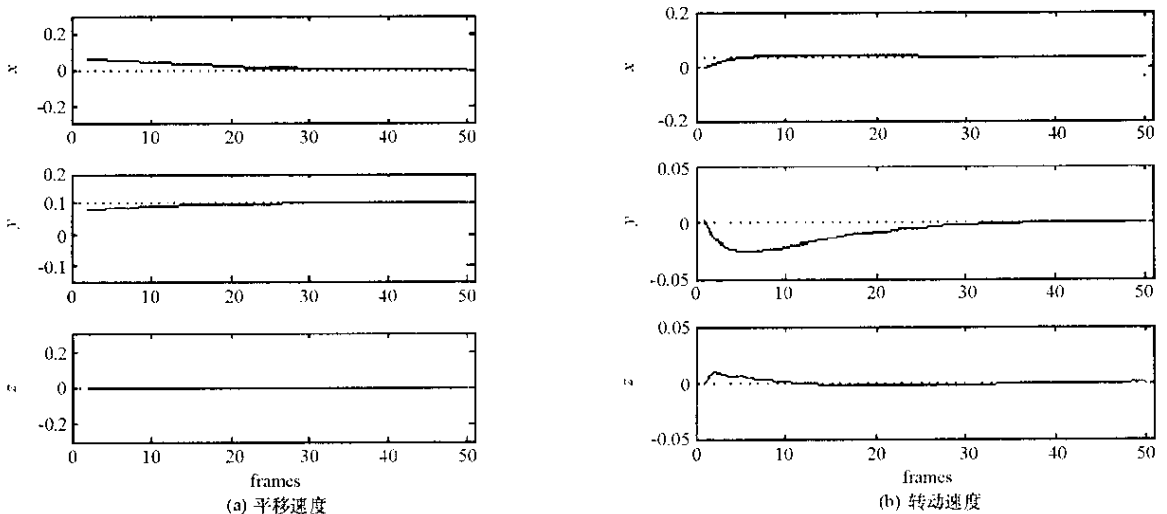


图 1 离散处理的方法得到的运动估计结果

Fig. 1 Motion estimation results from feature points based method

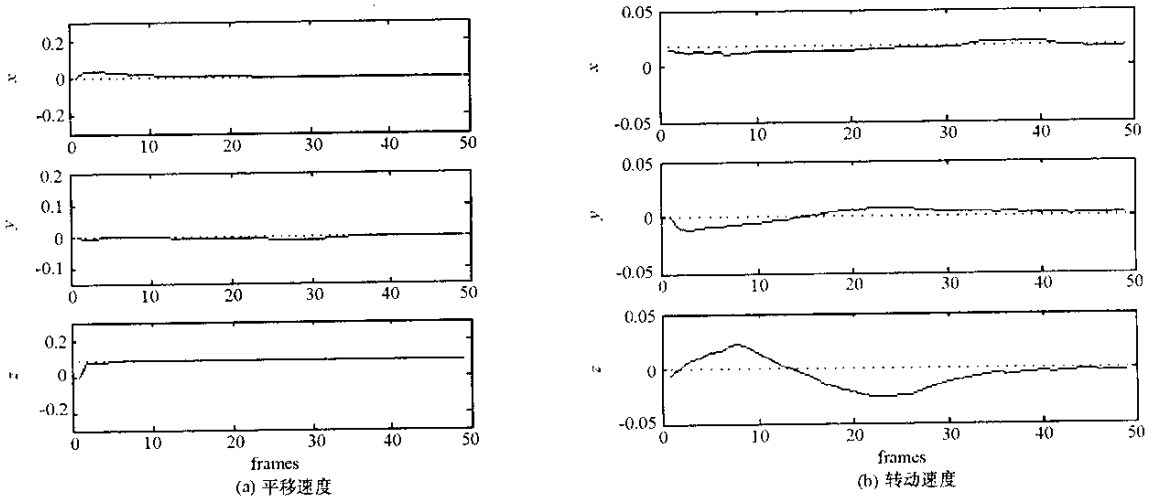


图 2 连续处理的方法得到的运动估计结果
Fig.2 Motion estimation results from optical flow based method

5 结论

分析了计算机视觉中一种新的运动估计模型,该模型将运动估计问题转化成参数定义在本质流形上的非线性动力学模型的状态估计问题。针对基于离散处理和连续处理方法得到的本质流形,提出了一种新的、统一的自适应推广卡尔曼滤波算法,在本质流形的局部坐标系上进行运动估计。仿真结果显示出了预期的优点。算法的收敛性、鲁棒性、不同实验条件下各种算法性能的进一步比较和全局的运动估计方法及在实际工程问题中的应用将是下一阶段的工作。

参考文献:

- [1] Ma Y. Differential Geometric Approach to Computer Vision and Its Applications in Control [D]. University of California at Berkeley, Fall 2000.
- [2] Soatto S, Freza R, Perona P. Motion Estimation via Dynamic Vision [J]. IEEE Transactions on Automatic Control, 1996 41(3):393-41.
- [3] Longuet-higgins. A Computer algorithm for Reconstructing a Scene from Two Projections [J]. Nature, 293:133-135, 1981.
- [4] Murray R M, Li Z, Sastry S S. A Mathematical Introduction to Robotic Manipulation [M]. Boca Raton: CRC Press, 1994.
- [5] Huang T, Faugeras. Some Properties of the E matrix in Two-view Motion Estimation [J]. IEEE PAMI, 1989 11(12):1310-12.
- [6] Hartley R. Estimation of Relative Camera Positions for Uncalibrated Cameras [A]. In Proc. 2nd ECCV [C], G. Sandini (Ed.), Vol. 588, Springer Verlag, 1992.
- [7] 吴立德. 计算机视觉 [M]. 上海: 复旦大学出版社, 1993.
- [8] Zucchelli M. Optical Flow Based Structure From Motion [D]. Royal Institute of Technology, 2002.
- [9] Soatto S. Observability/Identifiability of Rigid Motion under Perspective Projection [A]. In Proc. of the 33rd IEEE Conf. On Decision and Control, 1994: 3235-3240.
- [10] Perona P, Soatto S. Motion and Structure from 2 Perspective Views of p Localized Features: Noise Sensitivity Analysis [A]. Technical Report CIT/CNS n.23 - California Institute of Technology, April 1993.
- [11] Maybeck P. Stochastic Models, Estimation and Control [M]. Volume 2, New York: Academic Press, 1982.
- [12] Busse F D. Demonstration of Adaptive Extended Kalman Filter for Low Earth Orbit Formation Estimation Using CDGPS [A]. Institute of Navigation GPS Meeting, Portland OR, September 2002.

