

文章编号: 1000-6893(2004)01-0005-04

# 图像序列中机动目标三维运动和结构的计算

陈震<sup>1</sup>, 高满屯<sup>2</sup>, 杨声云<sup>1</sup>, 沈允文<sup>2</sup>

(1. 南昌航空工业学院 人工智能与图像处理研究中心, 江西 南昌 330034)  
(2. 西北工业大学 机电工程学院, 陕西 西安 710072)

Calculation of 3D Motion and Structure from Optical Flow in Image Sequence

CHEN Zhen<sup>1</sup>, GAO Man2tun<sup>2</sup>, Yang Sheng2yun<sup>1</sup>, SHEN Yun2wen<sup>2</sup>

(1. The Center of Artificial Intelligence and Image Processing, Nanchang Institute of Aeronautical Technology, Nanchang 330034, China)

(2. College of Mechanical and Electrical Engineering, Northwestern Polytechnical University, Xi'an 710072, China)

**摘要:** 综合视觉运动分析中的2类处理方法, 选取图像中的角点作为特征点, 在理论上证明了图像序列的光流场可以近似地用角点的位移场代替。利用已有文献中的建模思想, 详细推导出递归计算机目标三维运动和结构的非线性计算模型, 采用广义卡尔曼滤波(EKF)递归地计算图像序列中机动目标的三维运动和结构。合成图像序列和真实图像序列实验结果表明该算法能取得较好的效果。

**关键词:** 光流场; 角点位移场; 广义卡尔曼滤波; 三维运动和结构; 图像序列; 计算机视觉

**中图分类号:** V243 **文献标识码:** A

**Abstract:** The paper presents a new method for optical flow estimating, which combines feature based with flow based method. By using the corner points as feature points and estimating the optical flow from image sequence, the optical flow is estimated by measuring the displacement of sparse located corner points between consecutive frames. It is proved theoretically that optical flow can be replaced approximately by using the displacement field. The paper also estimates 3D motion and structure from optical flow in image sequence. In this paper, the implementation of a nonlinear algorithm is described, whose uniform observability, minimal realization and stability are proved analytically. Experimental results show that the new method provide a good estimation of the optical flow and 3D motion and structure.

**Key words:** computer vision; optical flow; corner point displacement field; extended Kalman filter; 3D motion and structure; image sequence

目前视觉运动分析的两类方法中, 基于特征的方法已取得了一些很好的结果<sup>[1, 2]</sup>, 但图像序列中特征匹配技术的鲁棒性问题是薄弱环节。基于光流场的方法在近年取得很大的进展<sup>[3]</sup>, 但由于图像噪声的干扰, 光流场计算的误差较大, 而且光流场方法用于图像特征的跟踪和匹配还未能取得满意的结果。

从图像序列中估计场景的三维运动和结构通常看成是一个非线性问题。利用广义卡尔曼滤波器可以递归地求解场景的三维运动和结构<sup>[4]</sup>。Chiuso<sup>[5]</sup>等人提出的非线性动态系统被证明是目前最有效的、最稳定的方法。

## 1 角点检测及跟踪

记时刻  $t$  时, 图像上一点  $m$  处的灰度值为

$E_t(m)$ ,  $m = [x \ y]^T$ 。在时刻  $t + S$  时, 这一点运动到新位置, 其在图像上的灰度值变为  $E_{t+S}(m + D)$ , 假定它与  $E_t(m)$  相等, 即

$$E_{(t+S)}(m + D) = E_t(m) \quad (1)$$

其中:

$$D = Dm + d \quad (2)$$

$D$  表示变形矩阵;  $d$  是特征窗口中心的平移量。

把式(2)代入式(1), 则有

$$E_{t+S}(m + Dm + d) = E_t(m) \quad (3)$$

在以  $m$  为中心的一个  $W$  特征窗口(区域)内求解式(3), 即可得到跟踪的角点。为了求解式(3), 利用微分平方和技术使其最小化

$$E = \iint_Q [E_{t+S}(m + Dm + d) - E_t(m)]^2 R(m) dm \quad (4)$$

其中:  $R(m)$  为高斯加权函数。出于简化考虑, 通常情况下设  $R(m) = 1$ 。

求最小值  $E$  对未知量的偏导数并用泰勒公式

展开,则生成 6 @6 阶的线性等式,表示如下

$$Lh = a \tag{5}$$

其中:  $h^T = [D_{11} \ D_{12} \ D_{21} \ D_{22} \ d_1 \ d_2]$ , 向量  $a$  可表示为

$$a = \iint_Q [E_x(s(m)) - E_t(m)] \begin{bmatrix} xE_x \\ xE_y \\ yE_x \\ yE_y \\ E_x \\ E_y \end{bmatrix} R(m) dm \tag{6}$$

其中:  $E_x$  和  $E_y$  分别为图像灰度  $E_t(m)$  对  $x$  轴方向和  $y$  轴方向的偏导数。

矩阵  $L$  表示为

$$L = \iint_Q \begin{bmatrix} A & B \\ B^T & G \end{bmatrix} R(m) dm \tag{7}$$

式中:  $A, B$  和  $G$  均是  $E_x, E_y$  和  $m$  的函数。

矩阵  $G$  近似用图像灰度的海赛形式表示

$$G = \begin{bmatrix} E_x^2 & E_x E_y \\ E_y E_x & E_y^2 \end{bmatrix} \tag{8}$$

当相邻图像帧之间的时间间隔很短,角点的运动距离很小时,式(2)作为角点运动的超越方程,求解时可能导致较大的误差。通过求解下列方程,可以得到角点跟踪的精确解

$$Gd = [a_5 \ a_6]^T \tag{9}$$

这实际上相当于假定角点运动  $D$  是一个简单的平移运动,即  $D = d$ 。

通过分析矩阵  $G$  的特征值可以遴选出正确的跟踪角点。

### 2 光流场计算

摄像机几何模型为透视投影模型,不失一般性,设摄像机焦距为  $f = 1$ ,则

$$\begin{cases} x = X/Z \\ y = Y/Z \end{cases} \tag{10}$$

$M = [X \ Y \ Z]^T$  为机动目标上点的坐标;  $m = [x \ y]^T$  为其投影的坐标。

在三维空间中的机动目标旋转运动用一个 3 @3 阶的正交矩阵描述,记作  $R$ ; 平移运动用一个 3 @1 阶矩阵描述,记作  $T = [t_1 \ t_2 \ t_3]^T$ 。则

$$M(t + \Delta t) = R(t + \Delta t)M(t) + T(t + \Delta t) \tag{11}$$

由于相邻图像帧之间的时间间隔  $\Delta t$  很小,可对式(11)作下列近似<sup>[6]</sup>

$$\left. \begin{aligned} T(t + \Delta t) &\approx V\Delta t \\ R(t + \Delta t) &= e^{\Delta t X_{@}} \approx I + \Delta t X_{@} \end{aligned} \right\} \tag{12}$$

其中:  $I$  为单位矩阵;  $V = (M \ M \ M)^T$  分别表示机动目标沿  $X, Y, Z$  方向的平移速度。用  $X = [X_1 \ X_2 \ X_3]^T$  分别表示机动目标绕  $X, Y, Z$  方向的旋转速度,  $X_{@}$  记为  $X$  的反对称矩阵。

在作上述近似时,实际上是默认下式成立

$$\left. \begin{aligned} X_{@}^2 \Delta t \ll X_{@} \Delta t \\ V\Delta t \ll V\Delta t \end{aligned} \right\} \tag{13}$$

即二阶项远小于一阶项。上式意味着运动是平滑的,也就是说,不允许有突变的运动速度。

把式(12)代入式(11),可得

$$M(t + \Delta t) = (I + \Delta t X_{@})M(t) + V\Delta t \tag{14}$$

定义位移场的方程

$$\frac{m(t + \Delta t) - m(t)}{\Delta t} = U(x, y, t) \tag{15}$$

其中:  $U(x, y, t)$  表示时刻为  $t$  时的光流场。

实际上,不能由式(14)得出位移场的方程。展开式(14)的左右各项可知,只有当下面的式子

$$\begin{aligned} Z(t + \Delta t) &= Z(t) + X_3 \Delta t Y(t) - \\ &X_2 \Delta t X(t) + M_3 \Delta t = Z(t) \end{aligned} \tag{16}$$

成立时,才能由式(14)得出位移场的方程。以统一的  $Z$  表示前后时刻机动目标上点的  $Z$  坐标,式(16)两边同除以  $Z$ ,可得

$$1 + (X_3 \Delta t Y - X_2 \Delta t X + M_3 \Delta t) / Z = 1 \tag{17}$$

从式(17)中可以看出,光流场并不等于位移场。为了获得满意的近似,必须使等式(17)中左边项接近于 1。为此,作下列假设

$$\left. \begin{aligned} X_1 Y \Delta t \ll Z \\ X_2 X \Delta t \ll Z \\ M_3 \Delta t \ll Z \end{aligned} \right\} \tag{18}$$

式(18)表示沿着光轴  $Z$  的运动相比于机动目标到摄像机之间的距离非常小。在满足式(13)和式(18)两个前提的情况下,光流场就可以近似用角点的位移场来代替。

### 3 机动目标三维运动和结构计算

设定一个离散时间的非线性动态系统

$$\left. \begin{aligned} N(t + 1) &= f(N(t)) + Q(t) \\ Q(t) &\sim N(0, 2\sigma(t)) \\ y(t) &= h(N(t)) + n(t) \\ n(t) &\sim N(0, 2n_0(t)) \end{aligned} \right\} \tag{19}$$

其中:  $y(t)$  为观测向量,初始条件是  $t_0$  时刻的  $N_0$ , 写成  $y(t; t_0, N_0)$ ; 向量函数  $h_i(\#)$  和  $f_i(\#)$  分

别为该系统的观测函数和状态函数;  $\mathbf{N}$  为状态向量;  $\mathbf{N}$  表示状态向量的观测值;  $\mathbf{G}$  和  $\mathbf{n}$  分别为系统动态噪声和观测噪声。

设滤波误差在  $t$  时刻的协方差为  $\mathbf{P}(t)$ 。

初始条件为

$$\left. \begin{aligned} \mathbf{N}(0) &= \mathbf{N}_0 \\ \mathbf{P}(0) &= \mathbf{P}_0 > \mathbf{0} \end{aligned} \right\} \quad (20)$$

设与非线性动态方程对应的线性动态方程为

$$\left. \begin{aligned} \mathbf{N}(t+1) &= \mathbf{F}(t) \mathbf{N}(t) \quad \mathbf{N}(t_0) = \mathbf{N}_0 \\ y(t) &= \mathbf{H}(t) \mathbf{N}(t) \end{aligned} \right\} \quad (21)$$

线性化过程如下

$$\left. \begin{aligned} \mathbf{F}(t) &= \frac{\partial \mathbf{f}}{\partial \mathbf{N}}(\mathbf{N}(t|t)) \\ \mathbf{H}(t+1) &= \frac{\partial \mathbf{h}}{\partial \mathbf{N}}(\mathbf{N}(t+1|t)) \end{aligned} \right\} \quad (22)$$

则递归过程由下列系统控制

$$\left. \begin{aligned} \mathbf{N}(t+1) &= \mathbf{f}(\mathbf{N}(t)) + \\ &\quad \mathbf{K}(t)[y(t+1) - \mathbf{h}(\mathbf{N}(t))] \\ \mathbf{P}(t+1) &= \mathbf{R}(\mathbf{P}(t), \mathbf{F}(t), \\ &\quad \mathbf{H}(t), 2\mathbf{n}, 2\mathbf{G}) \end{aligned} \right\} \quad (23)$$

其中:  $\mathbf{K}$  为卡尔曼增益;  $\mathbf{R}$  为线性化系统的 Riccati 等式。

对于从运动到结构的递归过程来讲, 常常把时刻  $t$  的位置作为时刻  $t+1$  递归的初始值, 即

$$\mathbf{M}_i(t+1) = \mathbf{M}_i(t) \quad \mathbf{M}(0) = \mathbf{M}_0 \quad (24)$$

在透视投影模型下, 点的投影用函数  $\mathbf{P}$  表示

$$\mathbf{m}_i(t) = \mathbf{P}(\mathbf{M}_i(t)) = (X_i/Z_i \quad Y_i/Z_i)^T \quad (25)$$

上式可以写成如下 2 个式子

$$\mathbf{m}_i(t+1) = \mathbf{m}_i(t) \quad i = 1, 2, \dots, N \quad \mathbf{m}(0) = \mathbf{m}_0 \quad (26)$$

$$Z_i(t+1) = Z_i(t) \quad i = 1, 2, \dots, N \quad Z_i(0) = Z_{i0} \quad (27)$$

其中:  $Z_i$  表示点  $\mathbf{M}_i$  处的结构, 即深度。

又因为机动目标上的点满足式 (11) 运动方程, 因此点的投影关系可以表示成

$$\mathbf{m}_i(t) = \mathbf{P}(\mathbf{R}(t) \mathbf{M}_i(t) + \mathbf{T}(t)) \quad (28)$$

刚性机动目标在三维空间运动, 空间相邻点之间的运动关系用罗德里格斯公式表示为

$$\left. \begin{aligned} \mathbf{T}(t+1) &= \mathbf{e}^{\mathbf{X}_@^{(t)}} \mathbf{T}(t) + \mathbf{V}(t) \\ \mathbf{R}(t+1) &= \mathbf{e}^{\mathbf{X}_@^{(t)}} \mathbf{R}(t) \end{aligned} \right\} \quad (29)$$

另外如果用  $\mathbf{8}$  表示三维真实向量, 则旋转矩阵  $\mathbf{R}$  可以表示成

$$\mathbf{R} = \mathbf{e}^{\mathbf{8}_@} \quad (30)$$

这样关于旋转矩阵  $\mathbf{R}$  的方程就写成

$$\mathbf{e}^{\mathbf{8}_@^{(t+1)}} = \mathbf{e}^{\mathbf{X}_@^{(t)}} \mathbf{e}^{\mathbf{8}_@^{(t)}} \quad (31)$$

综上所述, 机动目标三维运动和结构计算的模型可以完整表述为

$$\left. \begin{aligned} \mathbf{m}_i(t+1) &= \mathbf{m}_i(t) \quad i = 1, \dots, N \quad \mathbf{m}(0) = \mathbf{m}_0 \\ Z_i(t+1) &= Z_i(t) \quad i = 1, \dots, N \quad Z(0) = Z_0 \\ \mathbf{T}(t+1) &= \mathbf{e}^{\mathbf{X}_@^{(t)}} \mathbf{T}(t) + \mathbf{V}(t) \quad \mathbf{T}(0) = \mathbf{T}_0 \\ \mathbf{e}^{\mathbf{8}_@^{(t+1)}} &= \mathbf{e}^{\mathbf{X}_@^{(t)}} \mathbf{e}^{\mathbf{8}_@^{(t)}} \quad \mathbf{8}(0) = \mathbf{8}_0 \\ \mathbf{V}(t+1) &= \mathbf{V}(t) + \mathbf{A}(t) \quad \mathbf{V}(0) = \mathbf{V}_0 \\ \mathbf{X}(t+1) &= \mathbf{X}(t) + \mathbf{A}_x(t) \quad \mathbf{X}(0) = \mathbf{X}_0 \\ \mathbf{m}_i(t) &= \mathbf{P}(\mathbf{e}^{\mathbf{8}_@^{(t)}} \mathbf{m}_i(t) Z_i(t) + \\ &\quad \mathbf{T}(t)) + \mathbf{n}_i(t) \quad i = 1, \dots, N \end{aligned} \right\} \quad (32)$$

文献 [5] 从理论上证明了上述非线性动态方程是最小的, 且具有非常高的稳定性和精确性。

#### 4 算例与分析

**算例 1** 材质球合成图像序列。材质球作平移运动, 平移运动速度为 (0.1, 0.1) 像素/帧。取图像序列中的连续 95 帧验证本文提出的算法, 实验结果如图 1 所示。其中图 1(a) 为图像序列中的第 20 帧, 图 1(b) 表示角点跟踪的结果。最后匹配出的有效角点 118 个, 图中矩形区域外的小短线表示角点跟踪的过程。图 1(c) 为图像序列的真实光流场。图 1(d) 为利用本文基于角点的光流场算法得到的结果。

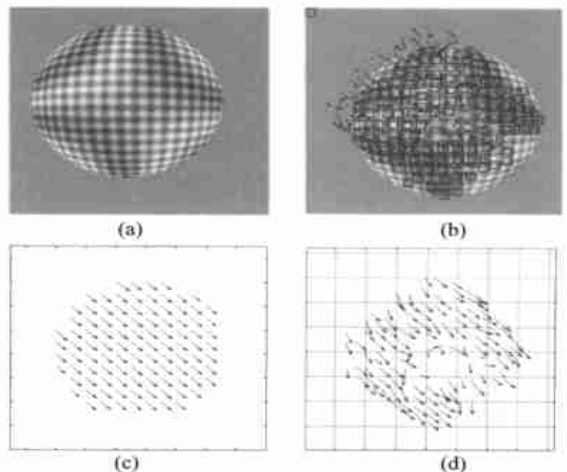


图 1 材质球合成图像序列光流场计算结果

Fig 1 The calculated results of the optical flow of synthetic sphere image sequence

图 2 为材质球合成图像序列运动和结构的计算结果。图 2(a) 为特征角点相对结构的计算结果。相对结构很快收敛到稳定值, 自 20 帧后, 相对结构基本没有波动; 图 2(b) 表示相对平移速度

和旋转角速度的计算结果。平移速度快速收敛,最后稳定于 01093 像素/帧。旋转角速度收敛于 0。实验结果与图像序列真实运动完全吻合。

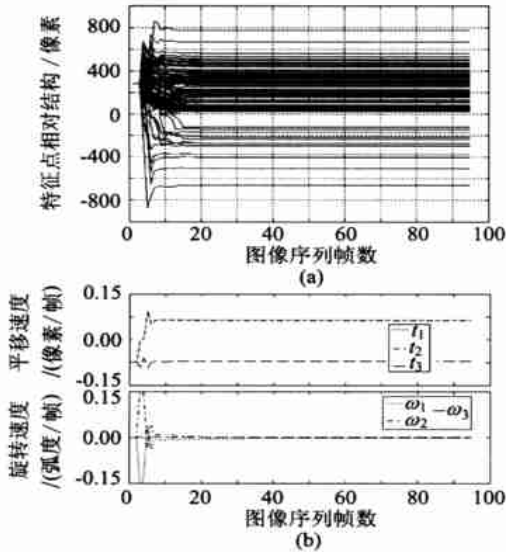


图2 材质球合成图像序列二维运动和结构计算结果  
Fig 2 The calculated results of the 3D motion and structure of synthetic sphere image sequence

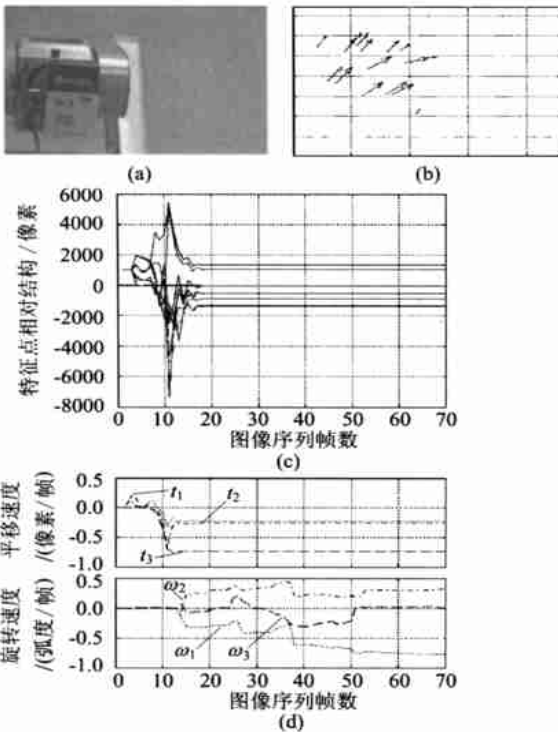


图3 红河烟盒图像序列实验结果

Fig 3 Experiment results image sequence of Honghe box

**算例 2 红河烟盒图像序列。**取图像序列中的连续 70 帧验证本文的算法,实验结果如图 3 所示。图 3(a)为图像序列中的第 1 帧;图 3(b)为计算所得第 9 帧的光流场。本实验中最后匹配出的角点 16 个。设摄像机的焦距为 1000 像素。经历

70 帧后,共得到有效的匹配角点 8 个;图 3(c)表示计算出的机动目标各角点的结构图。各角点的结构在经历 15 帧后基本稳定,而且由于有 2 个角点位于机器人手臂上,烟盒是固定在机器人手臂上的,烟盒与机器人手臂的运动相同,所以这 2 个角点的深度与其它 6 个位于烟盒上角点的深度明显不同;图 3(d)为三维运动计算结果。该算例共有 5 个运动,而且平移运动收敛较快。

5 结束语

本文提出了一种基于特征和基于光流场相结合的光流场算法,并采用广义卡尔曼滤波递归地计算图像序列中机动目标的三维运动和结构。

用本文的算法对不同复杂场景的合成图像和真实图像序列进行实验,均取得了较好的效果。

参 考 文 献

[1] Tomasi C, Kanade T. Detection and tracking of point features [R]. CMU/CS-91132, Pittsburgh: Carnegie Mellon University School of Computer Science, 1991.  
 [2] Shi J, Tomasi C. Good features to track[A]. In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR. 94) [C]. Seattle, Washington, 1994. 593- 600.  
 [3] 陈震,高满屯,沈允文.图像光流场计算技术研究进展[J].中国图像图形学报,2002,7A(5): 434- 439.  
 (Chen Z, Gao M T, Shen Y W. The current situation and trends of optical flow estimation[J]. Journal of Image and Graphics, 2002, 7A(5): 434- 439.)  
 [4] Azarbayejani A, Pentland A. Recursive estimation of motion, structure and focal length[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1995, 17(6): 562- 575.  
 [5] Chiuso A, Soatto S. 3D motion and structure from 2D motion causally integrated over time. Part I: theory[A]. IEEE Conference on Robotics and Automation[C]. San Francisco, Canada, 2000. 734- 750.  
 [6] Kanatani K. Geometric computation for machine vision[M]. Oxford: Oxford Science Publications, 1993.

作者简介:



陈震 (1969- ) 男,博士,副教授。主要研究方向为计算机视觉、图形图像处理。

高满屯 (1962- ) 男,教授。主要研究方向为计算机视觉、图形图像处理、人工智能和模式识别。

杨声云 (1971- ) 女,硕士,讲师。主要研究方向为现代控制理论和智能控制。

沈允文 (1936- ) 教授,博导。主要研究方向为非线性动力学、振动、冲击和噪声。

(责任编辑:俞敏)