

未标定多眼装置的欧氏重构

Euclidean Reconstruction of Uncalibrated Omni-Rig

章权兵¹ 韦 穗¹ 杨尚骏²

(安徽大学计算智能与信号处理教育部重点实验室 合肥230039)¹

(安徽大学数学系 合肥230039)²

Abstract The paper investigates a multiple cameras rig, where the cameras can rotate and change its internal parameters freely. This rig can be applied to inspect system, robot guides and military affairs and so on. On the assumption that the cameras' centers relative to each other have been determined, we describe a new linear method of Euclidean reconstruction from projective reconstruction, and also show that the maximal generality is reached when the rig consists of 5 cameras whose centers of projection are given. The advantages of this method are that the cameras can rotate and change its internal parameters freely, and the algorithm is linear.

Keywords Projective reconstruction, Euclidean reconstruction, Internal parameters, Camera center

1. 引言

在计算机视觉领域中,由多幅图像恢复摄像机运动参数和空间物体三维几何形状的问题称为多视图三维重构,它一直是计算机视觉研究的重点和热点之一。该问题的研究成果可以直接应用于机器人导航、精密工业测量、物体识别、虚拟现实以及军事等方面。近年来,该问题又引起多媒体和计算机图形学界的极大关注,为了生成已知景物和虚拟物体的合成环境,必须进行三维重构,否则用虚拟技术合成的环境只能用于浏览,而不能进入其中漫游。因此对三维重构理论和方法的研究无论是在计算机视觉中还是在工程应用中都具有十分重要的理论意义和实用价值。

早期的三维重构方法都必须对摄像机进行预标定,这就大大限制了其应用范围:只能应付静止和已知环境下的重构工作。1992年 Faugeras^[1]和 Hartley^[2]提出的利用未标定图像序列计算射影结构的理论,使这方面的工作有了显著的提高。在这种理论框架下,可以从未标定图像序列进行三维射影重构。但要得到欧氏结果,必须给出进一步的信息(如摄像机内参数等)。这种先进行射影重构,再利用额外信息恢复到欧氏重构的方法称为分层重构^[3,4]。

在传统的计算机视觉技术中,至少需要两台摄像机来观察物体。目前用于机器人对目标进行识别、跟踪、抓取等多是称为双眼装置的两个摄像机。最近,Amnon Shashua^[5]和 Assaf Zomet et al.^[6]分别对摄像机中心固定和旋转固定的多摄像机装置进行了研究,给出了一种从射影到欧氏重构的方法。但 Amnon Shashua 用到了很不稳定的基本矩阵和极点。本文则避开了基本矩阵和极点,直接对摄像机中

心固定的多摄像机装置(其中每个摄像机都可以自由旋转和改变内参数)进行了研究,给出了一种从射影重构到欧氏重构的线性算法,并证明了最小配置数是5。

2. 射影重构

针孔摄像机将三维空间中的点 P 映射到二维空间中的点 p,当 P 和 p 都用齐次矢量表示时,这个射影变换可以用一个秩为3的3×4齐次(摄像机)矩阵 M 表示为:

$$p \cong MP$$

其中“ $\cong$ ”表示在相差一个尺度因子下成立,其中 M 在欧氏坐标系下可以分解为

$$M = KR[I - C] \quad (1)$$

其中 R 表示摄像机旋转, C 是摄像机中心, K 是摄像机内参数。

假设 n 个三维空间点 $P_j (j=1, 2, \dots, n)$ 通过 m 个3×4射影矩阵 $M^i (i=1, 2, \dots, m)$ 投射到 m 幅图像上,得到的对应图像点为 $p_{ij} (i=1, 2, \dots, m, j=1, 2, \dots, n)$ 。如果这些都是齐次表示,那么我们可以得到下面的射影方程

$$W = \begin{bmatrix} \lambda_{11}p_1^1 & \lambda_{12}p_1^2 & \dots & \lambda_{1n}p_1^n \\ \lambda_{21}p_1^1 & \lambda_{22}p_1^2 & \dots & \lambda_{2n}p_1^n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_{m1}p_m^1 & \lambda_{m2}p_m^2 & \dots & \lambda_{mn}p_m^n \end{bmatrix} = \begin{bmatrix} M^1 \\ M^2 \\ \vdots \\ M^m \end{bmatrix} (P_1 \ P_2 \dots P_n)^T = MP \quad (2)$$

其中 $3m \times n$ 矩阵 W 称为测量矩阵, λ_j 称为射影深度。

当以某种方法求得射影深度 λ_j 后,就可对 W 进行奇异值分解(SVD)得到:

$$W = U \text{diag}(\sigma_1, \dots, \sigma_n) V^T$$

其中 $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_N \geq 0$ 为 W 的奇异值, $N = \min(3m, n)$, U, V 分别为 $3m \times N, m \times N$ 的次正交矩阵。如果图像坐标无噪声 (即 m 幅图像精确匹配), 则 W 的秩为 4, 故 $\sigma_5 = \sigma_6 = \dots = \sigma_N = 0$, 从而 $U_1 \text{diag}(\sigma_1, \dots, \sigma_4) V_1^T$ 为 W 的一个满秩分解, 其中 U_1 和 V_1 分别为 U 和 V 的前 4 列。结合 (2) 式可得:

$$M=U_1, P=\text{diag}(\sigma_1, \dots, \sigma_4) V_1^T$$

但上面的分解并不唯一, 因为(2)式在相差任意非零常数因子的情况下可以写为:

$$W = MHH^{-1}P \quad (3)$$

其中 H 为任意 4×4 非奇异矩阵, 因此均为 W 的分解。这就表明可以在相差一个射影变换下恢复摄像机运动和空间物体形状, 并可以任意选择一个 3D 射影坐标系来描述这种运动和形状, 这种恢复称为射影重构。

目前,由图像进行射影重构已有很多方法(这里不予重述),但要上升到欧氏重构必须添加额外的信息。下面我们就摄像机的中心的相对位置已知的情形,讨论由射影到欧氏重构的方法。我们将得到一种非常简单的线性算法。

3. 欧氏重构

由(3)式可知,对于任意一个 4×4 非奇异矩阵 H , MH 和 $H^{-1}P$ 均可以用来描述射影空间下的摄像机运动和空间物体形状。为了从射影重构恢复欧氏重构,我们选取 H 满足欧氏变换的条件,即(1)式,因此有:

$$\begin{cases} \mathbf{M}^1\mathbf{H} = \alpha^1\mathbf{K}^1\mathbf{R}^1[\mathbf{I} & -\mathbf{C}^1] \\ \vdots \\ \mathbf{M}^n\mathbf{H} = \alpha^n\mathbf{K}^n\mathbf{R}^n[\mathbf{I} & -\mathbf{C}^n] \end{cases} \quad (4)$$

在一定的约束条件下,由此方程组可以解得一个 H , 这样 MH 和 $H^{-1}P$ 即为欧氏空间下的摄像机运动和物体形状,这种重构称为欧氏重构。下面我们讨论摄像机的中心的相对位置已知,即 $C^i (i=1, 2, \dots, m)$ 已知时,如何求解 H 。

记 $H = [\hat{H} | H_4] = [H_1, H_2, H_3, H_4]$, 其中 H_i 为 H 的第 i 列, 则对任意 $i = 1, 2, \dots, m$ 有

$$M^i H_i = -\alpha^i K^i R^i C^i$$

不妨假设第一个摄像机的中心为世界坐标原点,即 $C^1 = (0, 0, 0)^T$, 则

$$M^1 H_i = 0$$

这是一个关于 H_4 的元素的齐次线性方程组, 因为 $\text{rank}(M^1)=3$, 所以上面方程有唯一的齐次解 $H_4=\delta(d_1, d_2, d_3, d_4)^T=\delta D$, 其中 $\delta \neq 0$ 是任意的尺度因子, $d_j=(-1)^j \det M^{1(j)}$, 其中 $M^{1(j)}(j=1, \cdots, 4)$ 是 M^1 去掉第 j 列所得到的 3×3 子矩阵。

再由(4)式得,对任意 $i=2, \dots, m$, 有

$$\mathbf{M}'\mathbf{H}_i = -\alpha^i \mathbf{K}'\mathbf{R}'\mathbf{C}'$$

$$M^i \hat{H} = \alpha^i K^i R^i$$

因此

$$M^i H_i = -M^i \hat{H} C^i, i=2, \cdots, m \quad (5)$$

设

$$\begin{aligned} \mathbf{A}' &= [\mathbf{M}_1^i, \mathbf{M}_2^i, \mathbf{M}_3^i]^T \mathbf{C}^T, \mathbf{h} = [\mathbf{H}_1^T, \mathbf{H}_2^T, \mathbf{H}_3^T]^T, \\ \mathbf{b}^i &= -[\mathbf{M}_1^i, \mathbf{M}_2^i, \mathbf{M}_3^i]^T \mathbf{H}_4, \end{aligned} \quad (6)$$

其中 M_i^T 是 M^i 的第 i 行, 则(5)式可以改写为:

$$A^i h = b^i, i = 2, \dots, m \quad (7)$$

进一步,记

$$\begin{pmatrix} A^2 \\ \vdots \\ A^m \end{pmatrix} = A, \begin{pmatrix} b^2 \\ \vdots \\ b^m \end{pmatrix} = (M_1^2, M_2^2, M_3^2, \dots, M_1^m, M_2^m,$$

$$M_i^m)^T H = -\delta B D$$

则(7)式可以联立得矩阵方程:

$$A_h = -\delta BD \quad (8)$$

因为 A、B、D 都已知, 所以上式是关于 h 的 12 个元素的线性方程组。当 $m=5$ 时, A 是一个 12×12 方阵。若 A 可逆, 则该方程组有唯一解, $h = -\delta A^{-1} BD$ 。如果 $m>5$, 那么 (8) 是一个超定方程组, 可用最小二乘法求得唯一解 h 。求得的 h 和 $H_4 = \delta D$ 一起就确定了射影矩阵 H。因此当 $m \geq 5$ 时, 我们就可以得到欧氏重构。

对于我们所讨论的多眼装置,欧氏重构的一般步骤总结如下:

1. 根据图像信息进行射影重构, 求出射影摄像机矩阵 M^i 和点的射影坐标 P_i 。

2. 求出齐次矢量 $D = (d_1, d_2, d_3, d_4)^T$, $d_j = (-1)^j \det M^{(j)}$, 其中 $M^{(j)}$ ($j=1, \dots, 4$) 是 M^{-1} 去掉第 j 列所得到的 3×3 子矩阵, 从而得到 H 的第四列 $H_4 = -\delta D$.

3. 按照(6)式定义未知矢量 h , 求出系数矩阵 A' 和常数项 b' , 并构作方程组(7), 进一步联立第2个到第 m 个摄像机的方程组得到方程组(8)。

4. 求解方程组(8), 得到 H 的后三列元素。

5. MH 和 $H^{-1}P$ 就是欧氏空间下的摄像机运动和物体形状, 也即得到了欧氏重构。

4. 最小配置数——5摄像机

由上面的讨论我们得到下面的结论:

结论 当仅知道摄像机中心的相对位置,而没有其他的额外信息时,5个一般位置的摄像机就足以从射影到欧氏重构,但摄像机少于5个则不行。

证明:由上面的算法知道,当给定一般位置的5个摄像机时,我们可以求出射影矩阵 H ,从而得到欧氏重构。

但如果摄像机少于5个,那么就无法确定射影矩阵 H ,理由如下: 4×4 矩阵 H 有16个元素,除去一个

2779, and the recognition rate is 70.8333%.

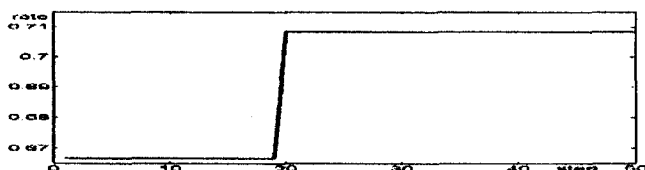


Figure 3 the training process of the unbalanced attention parameter training algorithm based on Genetic Algorithm

③The improved algorithm

Its training process is demonstrated in figure 4, where $\lambda_1 = \lambda_2 = 1$, $\delta = 0.01$. It takes about 20 min-

utes, the result is $\lambda_1 = 1.04$, $\lambda_2 = 1$, and the recognition rate is 70.8333%.

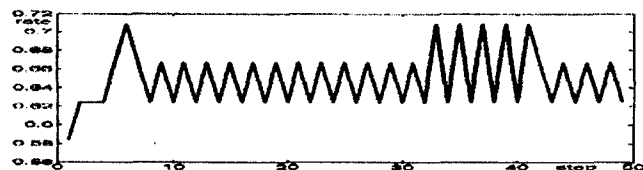


Figure 4 the training process of the improved algorithm

The above experimental results show that the novel algorithm is better than the original one under the same circumstance and less time-consuming than the unbalanced attention parameter training algorithm based on Genetic Algorithm.

References

- 1 H. Haken. Synergetic Computers and Cognition, Vol. 50 of Springer Series Synergetics. Springer, Berlin (1991)
- 2 F.-Y. Wang, P. J. A. Fever, B. Fu. A Robotic Vision System for

Object Identification and Manipulation Using Synergetic Pattern Recognition. Robot. Comput. Integrated Manufacturing, 1993, 10 (6): 445~459

- 3 Wang Hai Long, Qi Feihu, Zhang jinfeng. A Genetic-Synergetic Learning Algorithm Under Unbalanced Attention Parameters. Journal of Electronics, 2000, 28(11): 15~17
- 4 Wagner T, Bobel F G. Testing Synergetic Algorithms with Industrial Classification Problems. Neural Networks, 1994, 17(8): 1313~1321
- 5 Trevor Hogg, Habib Talhami, David Rees. An Improved Synergetic Algorithm for Image Classification. Pattern Recognition, 1998, 31(12): 1893~1903

(上接第187页)

全局尺度因子,剩下15个自由度,而每个摄像机中心有3个元素,至多提供3个自由度,所以至少需要5个摄像机。

总结及展望 我们给出了对于 m 摄像机装置 ($m \geq 5$),从射影重构恢复到欧氏重构的一种简单的线性算法,这里需要添加的额外信息仅仅是摄像机中心的相对位置已知。并且证明了这种最小配置是5个一般位置的摄像机。这种方法的优点是摄像机可以自由旋转和改变内参数,并且算法是线性的。

今后的工作中,我们将进一步研究这种算法的鲁棒性和稳定性以及退化情形。

参考文献

- 1 Faugeras O. What can be seen in three dimensions with an-

calibrated stereo rig? G. Sandini. Proc. ECCV'92. Santa Margherita Ligure, Italy: Springer-Verlag, 1992. 563~578

- 2 Hartley R I. Stereo from uncalibrated cameras, E. Grimson Proc. CVPR'92. Urbana-Champaign, Illinois: IEEE Computer Society, 1992. 761~764
- 3 Francois G. Hierarchical visual perception with calibration: [Report of research]. INRIA, France. No: 3002, 1996
- 4 Hartley R I, Zisserman A. Multiple View Geometry in Computer Vision. Cambridge University press, 2000
- 5 Shashua A. Omni-rig Sensors: What Can be Done With a Non-rigid Vision Platform? Proc. of the Workshop on Applications of Computer Vision (WACV), Princeton, 1998
- 6 Zomet A, Wolf L, Shashua A. Omni-rig: Linear Self-recalibration of a rig with Varying Internal and External Parameters, ICCV'2001 Vancouver, Canada, July, 2001