

基于计算机视觉的三维重建技术

王 剑 王 影 卫飞飞

(东北大学信息科学与工程学院 沈阳 110004)

摘 要 利用照相机成像模型从获取的二维图像序列恢复物体的三维结构是计算机视觉研究领域的经典问题之一,三维重建也是决定视觉三维测量精度的关键的因素之一。首先介绍立体视觉三维重建的基本原理,然后介绍视觉系统中的一个重要约束条件——极线约束。最后提出了射影几何意义下的重建。这种方法既不需要精密机械定位和光学跟踪定位,精度也不完全取决于标志点视觉测量。

关键词 三维重建,立体视觉,计算机视觉,极线约束

1 引言

从获取的二维图像序列恢复物体的三维结构是计算机视觉研究领域的经典问题之一,多年来学者们提出了恢复三维信息的各种算法,这些算法归纳起来又可分为三类:(1) 假设建立了图像间的对应关系,同时已知照相机内参数矩阵和照相机之间的相对位置关系,确定观察点的三维位置,如立体视觉系统;(2) 假设建立了图像间的对应关系,已知照相机内参数矩阵,先计算各视角图像坐标系之间的位置关系,进而获取三维信息,如由运动恢复结构(Structure from Motion,即 SFM)^[1];(3) 已知图像间特征点的匹配关系,要求提取照相机内参数,然后计算各视角图像坐标系之间的相对位置,最后获取三维信息,如近景摄影测量^[2]。

假设特征点的提取和匹配具体算法详见文[3,4],照相机标定具体实现详见文[5,6]和图像的畸变矫正都是在欧式几何意义下进行处理并且已经完成,所用的坐标系都是直角坐标系,坐标系之间的变换都是平移或正交变换。欧式几何描述了点集在单位正交变换与平移变换下的几何不变量(以下称为欧式变换),在此变换下,任意两点间的距离不变,因此,欧式几何最适合描述刚体的形状和运动。但是空间一个点集经照相机的中心射影变换为图像平面的一个点集并不能保证上述不变性,称这样的变换为射影变换。在射影几何意义之下,最基本、最重要的是交比(cross ratio)不变性,也就是由相交于空间某个三维点的直线所定义的线束交比保持不变。这个交比与照相机的内外参数(也就是射影变换参数、坐标系的选择)均无关,只与真实的三维点的分布有关,所以可以利用这个不变量来重建空间的三维点,并通过特征点来识别特定的物体。

2 立体视觉与三维重建

立体视觉三维测量一般是在已知多照相机(通常是两个)相对位置的情况下,由多幅图像恢复被测

物体三维形状。本节以两照相机组成的立体视觉测量系统为例说明其二维测量基本算法。图1所示为立体视觉的几何结构图。 C_1 、 C_2 为两个照相机, O_1 、 O_2 为相机的中心, I_1 、 I_2 为相机的图像平面。 P 为空间点, m_1 、 m_2 为点 P 在 I_1 、 I_2 上的投影,它们互称为对应点。 l_1 、 l_2 为 $\triangle O_1PO_2$ 所在平面与 I_1 、 I_2 的交线,称为点 m_1 、 m_2 分别在 I_1 、 I_2 上的对应极线, l_1 、 l_2 互称为共轭极线。从图中可以看出,对应点必然位于其对应极线上。由于空间点 P 为直线 O_1m_1 和 O_2m_2 的交点,在 m_1 、 m_2 的对应关系已知后,就可以确定点 P 的位置,这就是由立体视觉恢复三维结构的基本原理。

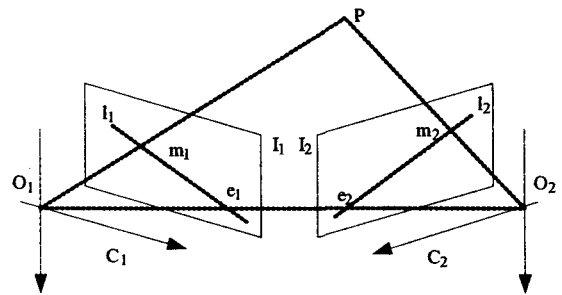


图1 立体视觉的几何结构图

设两照相机的投影矩阵分别为 M_1 和 M_2 ,因为立体视觉系统两照相机位置关系一般是固定的,通过对照相机和立体视觉系统的精确标定可以确定 M_1 和 M_2 ,于是有

$$z_{c1} \begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = M_1 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^1 & m_{12}^1 & m_{13}^1 & m_{14}^1 \\ m_{21}^1 & m_{22}^1 & m_{23}^1 & m_{24}^1 \\ m_{31}^1 & m_{32}^1 & m_{33}^1 & m_{34}^1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (1)$$

$$z_{c2} \begin{bmatrix} u_2 \\ v_2 \\ 1 \end{bmatrix} = M_2 \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11}^2 & m_{12}^2 & m_{13}^2 & m_{14}^2 \\ m_{21}^2 & m_{22}^2 & m_{23}^2 & m_{24}^2 \\ m_{31}^2 & m_{32}^2 & m_{33}^2 & m_{34}^2 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \quad (2)$$

其中, $(u_1, v_1, 1)$ 与 $(u_2, v_2, 1)$ 分别为 P_1 与 P_2 在各自图像坐标系中像点的齐次坐标; $(X_w, Y_w, Z_w, 1)$ 为 P 点在世界坐标系中的齐次坐标; m_{kj}^i 为 M_k 的第 i 行、第 j 列元素。在式(1)和(2)中消去 Z_{c1} 和 Z_{c2} , 得到关于 X_w, Y_w, Z_w 的四个线性方程

$$(u_1 m_{31}^1 - m_{11}^1) X_w + (u_1 m_{32}^1 - m_{12}^1) Y_w + (u_1 m_{33}^1 - m_{13}^1) Z_w = m_{14}^1 - u_1 m_{34}^1 \quad (3)$$

$$(v_1 m_{31}^1 - m_{21}^1) X_w + (v_1 m_{32}^1 - m_{22}^1) Y_w + (v_1 m_{33}^1 - m_{23}^1) Z_w = m_{24}^1 - v_1 m_{34}^1$$

$$(u_2 m_{31}^2 - m_{11}^2) X_w + (u_2 m_{32}^2 - m_{12}^2) Y_w + (u_2 m_{33}^2 - m_{13}^2) Z_w = m_{14}^2 - u_2 m_{34}^2$$

$$(v_2 m_{31}^2 - m_{21}^2) X_w + (v_2 m_{32}^2 - m_{22}^2) Y_w + (v_2 m_{33}^2 - m_{23}^2) Z_w = m_{24}^2 - v_2 m_{34}^2 \quad (4)$$

3 外极线几何

如图 1 所示, m_1 与 m_2 为对应点, 由于 m_1, m_2, P, O_1, O_2 在同一平面上, m_1 与 m_2 分别位于该平面与两个图像的交线极线 l_1 和极线 l_2 上。由双目成像的透视关系可知, 左图像 I_1 平面内的任意点在右图像 I_2 平面内的外极线都要经过外极点 e_2 , 反之亦然, 这种关系具有对称性。

根据照相机线性模型, 令 $X = [x, y, z]^T$ 是空间中的任意一点, X 的齐次坐标为 $\tilde{X} = [x, y, z, 1]^T$, 在像平面上该点的成像为 $m = [u, v]^T$, 对应齐次坐标为 $\tilde{m} = [u, v, 1]^T$, 则有

$$\tilde{s}m = M\tilde{X} \quad (5)$$

式中: s 是比例因子; M 为透视矩阵, $M = A \cdot [R \ T]$; A 为内参数矩阵。该矩阵变换将照相机系下的点转换到图像坐标系下。

4 基本矩阵 F

设 $m_1(u_1, v_1), m_2(u_2, v_2)$ 为两图像平面上的对应匹配点, 由外极线约束方程可得包含基本矩阵 F 九个未知分量的线性齐次方程

$$u^T f = 0 \quad (6)$$

其中:

$$u = (u_1 u_2, v_1 u_2, u_2, u_1 v_2, v_1 v_2, v_2, u_1, v_1, 1)^T$$

$$f = (F_{11}, F_{12}, F_{13}, F_{21}, F_{22}, F_{23}, F_{13}, F_{23}, F_{33})^T$$

如果已知 n 对匹配点, 将它们的坐标代入式(6), 则得到线性方程

$$U_n f = 0 \quad (7)$$

其中: $U_n = (u_1^T, u_2^T, \dots, u_n^T)^T$ 为一 $n \times 9$ 的矩阵。在满足基本矩阵 F 秩 2 的情况下, 通过求解线性方程组(7)就可得到 F 。基本矩阵 F 的求解算法分线性算法和非线性算法, 在线性算法中又有 7 点法、8 点法、线性最小二乘法; 在非线性算法中还有鲁棒法、再投影法等。

由于 7 点法、8 点法、线性最小二乘法均是假设图像对应点匹配已完成且匹配是正确的, 但实际上图像对应点匹配不可避免存在像点定位误差和匹配错误, 导致上述算法不是鲁棒的, 严重影响参数估计的精度。最常用的鲁棒算法是 M-估计(M-Estimators)和最小平方中值法(Least-Median-Squares, Lmeds)。

M-估计 令 r_i 为第 i 个数据的残差, 即第 i 个像点与计算值之差, 标准的最小二乘法是对 $\sum r_i^2$ 最小化, 当有错误匹配数据时, 该算法不稳定。M-估计将残差的平方 r_i^2 替换成另一个残差的函数 $\min \sum \rho(r_i)$, 这里 $\rho(r)$ 是个对称、正定函数, 其最小值在零处, $\rho(r_i) = \omega_i r_i^2$, 常用的权重函数是

$$\omega_i = \begin{cases} 1 & |r_i| \leq \sigma \\ \sigma/|r_i| & \sigma < |r_i| \leq 3\sigma \\ 0 & 3\sigma < |r_i| \end{cases} \quad (8)$$

其中 σ 计算结果的误差标准方差, 最初的误差标准方差可以通过标准的最小二乘法估计得到。该算法对图像数据的非高斯噪声是鲁棒的, 但对误匹配数据无效^[7]。

Lmeds 算法 最小平方中值法是通过解非线性最小化问题来优化参数

$$\min \text{med} r_i^2 \quad (9)$$

基本思想就是对全部数据所有可能的残差平方中值求解, 这个方法给出的平方中值最小。文^[7]表明该算法对图像数据的非高斯噪声和错误匹配数据的鲁棒性都很好。

关于如何选取样本数 m , 可参考文^[7]。

5 运动视觉三维重建

如果仅用一个照相机通过多次拍摄来重建空间物体的三维信息, 那么假设建立了图像间的特征点对的对应关系, 通过照相机标定确定了照相机内参数矩阵, 就可以计算各视角图像坐标系之间的位置关系, 进而获取匹配特征点的三维信息, 这就是由运动恢复结构(SFM)的算法。本文在 Hartley 和 Zhang^[1]提出的方法基础上进行了改进, 实现运动视觉特征点三维测量。

(1) 求解本质矩阵 E

求解本质矩阵是 SFM 算法的第一步, 从两幅图像的匹配结果出发, 鲁棒性算法估计出基础矩阵 F 。结合已经获得的照相机内参数矩阵 A , 计算出本质矩阵 E , 并归一化得到

$$E = A^T F A \quad (10)$$

(2) 计算照相机外参数

从本质矩阵 E 出发, 对 E 进行奇异值分解得到 $E = U D V^T$ (11)

其中: $D = \text{diag}(a, b, c)$, 理论上如果匹配点准确, 所

得到 a, b, c 的关系应为 $a=b, c=0$, 但是由于实际问题中总是存在误差, 因此得到的情况是 $a>b>c$, 于是令 $\tilde{D}=\text{diag}((a+b)/2, (a+b)/2, 0)$, 重新得到新的本质矩阵 $\tilde{E}=\tilde{U}\tilde{D}\tilde{V}^T$, 对 \tilde{E} 再进行奇异值分解, $\tilde{E}=\tilde{U}\tilde{K}\tilde{V}^T$, 这样就可以得到下面照相机运动参数 $(R|T)$ 的候选值: $(\tilde{U}\tilde{K}\tilde{V}^T|\tilde{U}(0,0,1)^T)$ 。

(3) 计算空间点三维坐标

在获得运动参数 $(R|T)$ 后, 给定一对匹配对, 就可以直接利用线性最小二乘法求出相应的空间三维点坐标。

于是根据最小二乘法得

$$X=(Z^T Z)^{-1} Z^T z \quad (12)$$

对每一个匹配对采用如上方法可以获得所有的空间点的三维坐标。

6 三维点阵配准拼接

假定来自一个物体上的两个点集(或两个网面)分别表示为 A 和 B , 其中 A 和 B 在物体表面上有足够的重叠以便实现可靠对准。点集 B 中的一点 q 和点集 A 的距离是

$$d(q, A)=\min_{p \in A} \|q-p\| \quad (13)$$

点集 B 中的一个子集 $Q=\{q_i\}$ 与点集 A 中的一个子集 $P=\{p_i\}$ 构成最近距离点对, 形成共轭对集 $\{(p_1, q_1), (p_2, q_2), \dots, (p_n, q_n)\}$, 这个共轭对集可以用来求解绝对定位问题。如果这些点集不是真正的共轭对, 也可以获得两个网面之间变换的逼近解法。两个网面之间变换以后, 重新求解最近共轭点集, 使得该共轭对集更接近真正的共轭对集, 这样的过程可以重复进行, 直到最近点之间的距离均方差低于某一阈值。

结论 本文首先介绍立体视觉三维重建的基本原理, 然后介绍了视觉系统中的一个重要约束条

件——极线约束。由于欧式几何意义下的重建要求的条件过多(如需要严格的照相机定标等), 而有些信息在某些场合下是不必要的, 因此就提出了射影几何意义下的重建, 可以做到只在解决基本矩阵 F 的照相机弱标定条件下, 得到识别物体所需要的射影不变量, 从而既简化了计算和对系统的要求, 又满足相当一部分视觉任务的需要。本文使用的重建算法收敛的条件是为对应图像中点对的相对误差范数小于一个临界值。但是需要指出的是, 由于点云数据中的噪声影响, 若干误差比较大的点对于配准拼接算法精度有很大的影响, 鲁棒性方面有待提高。配准拼接算法和目前普遍采用的方法及 Reich 提出的集成测量方法相比, 既不需要精密机械定位和光学跟踪定位, 其精度也不完全取决于标志点视觉测量。

参 考 文 献

- 1 Hartley R, Gupat R, Chang T. Stereo from uncalibrated cameras [J]. In: Proc. IEEE Conf. on Computer vision and pattern Recognition. 1992. 761~764
- 2 Ganci G, Brown J. Development in non-contact measurement using video grammetry [J]. In: Boeing Large Scale Metrology Conf. Long Beach, 2000
- 3 Moravec, Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover by H. Moravec [J]. Stanford AI memo AIM-340, 1980
- 4 Harris C, Stephens M. A combined corner and edge detector [J]. In: the Fourth Alvey Vision Conference, 1988
- 5 Zhang Z Y. Flexible camera calibration by viewing a plane from unknown orientations [J]. In: International Conference on Computer Vision (ICCV99), Greece, 1999. 666~673
- 6 孟晓桥. 摄像机自标定和三维重建中的若干问题研究[D]. [中国科学院研究生院硕士论文]. 2000
- 7 马颂德, 张正友. 计算机视觉[M]. 科学出版社, 1998

(上接第 118 页)

存储空间扩展、高性能传输需求, 并且单台外置存储系统的容量, 已经发展到了 2TB 以上, 随着大容量硬盘的推出, 单台外置存储系统容量还会上升。此外, DAS 还可以构成基于磁盘阵列的双机高可用系统, 满足数据存储对高可用的要求。从趋势上看, DAS 仍然会作为一种存储模式, 继续得到应用。

SAN 对于大容量块状级数据传输具有明显的优势, 而 NAS 则更加适合文件级别上的数据处理。尽管二者存在根本特性上的差异, 但 SAN 和 NAS 实际上也是能够相互补充的。SAN 擅长块数据传输、极易扩展且管理设备有效。用户可以使用 SAN 运行关键应用, 比如数据库、备份等, 以进行数据的集中存取与管理; 而 NAS 支持若干客户端之间文

件共享, 所以我们可以使用 NAS 作为日常办公中需要经常交换小文件的地方, 比如网页、公文等。SAN 和 NAS 在实际情况中是可以并存在一个系统中的, 从而构建一个功能强大, 结构灵活的存储系统。

参 考 文 献

- 1 袁文莉. 数字图书馆存储体系架构探析. 图书馆论坛, 2004(8)
- 2 SAN 与 NAS 和谐统一. 网络世界, 2005. 10. 13
- 3 存储在线. 深入了解网络存储, SAN、NAS 技术的比较. PCOnline, 2003. 10. 30
- 4 杨子伍. 探讨数字化图书馆的存储策略. 现代图书情报技术, 2003(3)
- 5 高先锋, 张洪滔. 图书馆海量存储系统架构与接口的选择. 现代图书情报技术, 2003(4)