

基于标记点丢失的多幅自标定图像的 3 维重建和相机姿态恢复

张维中 张丽艳 王小平 周玲 周儒荣

(南京航空航天大学 CAD/CAM 工程研究中心, 南京 210016)

摘要 为了对标记点丢失的多幅自标定图像进行精确重建,提出了一种基于标记点丢失的多幅自标定图像的 3 维重建和相机姿态恢复的方法。该方法与原来方法的不同之处在于,该方法是利用标记点(编码点和非编码点)的方式,即用编码点进行单 CCD 相机的自标定和姿态恢复,而非编码点进行 3 维点的 3 维重建。该方法有以下 3 个主要特点:(1)由于该方法采用了标记点的自动识别匹配,所以避免了手工交互选择图像点对(point correspondences)费工费时的问题;(2)由于标记点匹配精确,提高了 3 维点的重建精度,故符合工程要求;(3)由于噪音对标记点的像点影响较小,因此该方法比以前的方法具有更好的鲁棒性。实验结果表明,利用该方法产生的 3 维重建点精确可靠,能够满足逆向工程等应用的要求。

关键词 基于运动的 3 维重建 标记点 相机自标定 数据丢失

中图法分类号: TP391.41 TP274.2 文献标识码: A 文章编号: 1006-8961(2006)09-1282-06

Reconstruction and Camera Poses Recovery from Multi Self-calibration Images with Marked Point Occlusions

ZHANG Wei-zhong, ZHANG Li-yan, WANG Xiao-ping, ZHOU Ling, ZHOU Ru-rong
(Research Center of CAD/CAM Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)

Abstract In order to reconstruct accurately for multi self-calibration images with marked point occlusions, this paper proposes a method for reconstruction and camera poses recovery from multi self-calibration images with marked point occlusions. Compared to previous methods, this method uses a way of marked point: coded point and non-coded point. The coded points are utilized for a single CCD camera self-calibration and camera poses recovery and the non-coded points are utilized for 3D points reconstruction. The novelties of the proposed method are three-fold. Firstly, the problem of manually selected image point correspondences is avoided by using auto match of marked points. Secondly, the proposed method is based on marked points, the precision of 3D reconstruction is improved since match of marked point is accurate. Thirdly, the robustness is markedly increased since the effect of noise on the marked points of images is small. The practice shows that the results are accurate and reliable so as to meet the requirements of reverse engineering applications.

Keywords structure from motion, marked point, camera self-calibration, missing data

1 引言

如今人们面临着如何用一个单 CCD(charge coupled device) 摄像机,以自由拍摄方式通过获得多

角度场景的多幅图像来恢复场景中物体形状的问题。这就是众所周知的基于运动的 3 维重建(structure from motion problem, SFM)问题。但在这种情况下,既未进行相机标定,也未知场景图形。目前,这方面问题的研究热点是①虚拟现实、视频会议

基金项目:国家自然科学基金项目(50475041)

收稿日期:2005-04-21 改回日期:2005-10-23

第一作者简介:张维中(1963~),男,青岛大学信息工程学院教授,现为南京航空航天大学 CAD/CAM 工程研究中心博士研究生。主要研究方向为计算机视觉、模式识别和图像处理。发表论文 20 余篇,其中 EI 收录 3 篇。E-mail: zhangwz_01@163.com

及医学方面的3维造型^[1];②工业中的逆向工程,即根据实物测量来重建实物的数字化模型;③产品质量检测及机器人技术在航空航天和军事领域的应用。

1992年,Tomasi和Kanade提出了一种算法^[2],即通过计算多幅图像来实现场景的3维重建,并获得相机的内、外部参数,但该方法必须利用所有图像上的所有点。该算法是利用测量矩阵分解来实现3维重建,而且是在正交投影相机模型基础上实现的。后来又被Poelmaman等扩展到弱透视和准透视投影。Sturm和Triggs又将这种方法扩展到仿射投影和透视投影^[3]。Heyden用不同于前面的方法^[4],即利用子空间法来代替多线性约束方法。然而,该方法不可避免地,当相机视点发生改变时,由于有些点被隐藏(occlusions)起来,即某些点在某些图像上是不可见的,因此,某个点仅在某几幅图像上是可见的。

Rother和Carlsson提出了一种算法^[5]较好地解决了上述问题。该算法是通过增加一个简单的假定,即一个平面上的4个点在所有的图像上都可见。这样3维重建和相机恢复问题就可以被表达为矩阵零空间问题而得以解决。由于参考平面的存在,致使图像点、场景中的3维点和相机中心三者间的投影关系由原来的3维点与相机位置的双线性关系变为线性关系。当在所有图像中不是所有点都可见的情况下,这种线性关系就能同时恢复3维点和相机的位置。

Hartley提出了一种基于纯旋转相机光轴的办法^[6]用来解决相机的自标定问题;文献[5]解决了不是所有点在所有图像上都可见的情况下,恢复空间3维点和相机位置问题。

本文提出了一种基于文献[6]和文献[5]的新方法,它是采用编码点和非编码点等标记点方式,即用编码点进行单CCD照相机的自标定及确定单CCD照相机的外部姿态参数,而非编码点确定空间3维点的3维坐标,从而实现了在不是所有的标记点在所有图像上都可见的情况下,恢复3维空间点和相机位置。由于采用了标记点方式,故本方法有以下3个主要特点:(1)克服了以往手工交互选择图像点对(point correspondences)的缺陷;(2)提高3维点重建的精度,符合工程要求;(3)降低了噪声的影响和提高了算法的鲁棒性。

2 预备知识

2.1 相机模型

本文假定相机模型为经典的针孔模型。在该模型下,相机按透视射影变换将空间点 $X=(x, y, z, 1)^T$ 投影到像点 $u=(u, v, 1)^T$,使得写成如下形式:

$$u \sim PX \quad (1)$$

" \sim "表示相差一个常数因子下的相等。投影矩阵 P 可进一步分解成

$$P = K[R | -Rt]K = \begin{pmatrix} f & \beta & u_0 \\ 0 & \alpha f & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

这里 R, t 分别是世界坐标系到相机坐标系的旋转矩阵和平移向量,分别代表相机的方位和位置,被称为相机外参数; K 是一个上三角矩阵,由于将相机坐标系变换到图像坐标系所需的只与相机自身有关的参数矩阵,因此可称之为相机内参数矩阵,亦称为相机标定矩阵; $[R | -Rt]$ 代表一个刚体运动(旋转和平移); (u_0, v_0) 为主点坐标(principal point); f 为有效焦距(effective focal length); α 为纵横比(aspect ratio); β 为倾斜因子(skew factor)。相机自标定的目的就是要求解矩阵 K 。

2.2 标记点

在场景中额外布置一些标记点,这些标记点分为编码点(coded point)和非编码点(non-coded point)。如图2(a)和图3(a)所示,带有图案的大方块是编码点,带有白点的黑圆圈是非编码点。编码点经相机成像后为椭圆。根据图像特征提取算法,首先提取椭圆中心的图像坐标,然后根据编码点的特征识别其编号。

最后利用编码点的编号就可以方便地建立多幅图像中编码点间的对应匹配关系;另外,根据图像特征提取算法,也可以方便地建立多幅图像中非编码点间的对应匹配关系。

3 单CCD相机的自标定

在精确地恢复出物体的位置、形状等3维信息之前,通常需要对相机进行标定,而对于多幅序列图像常采用自标定方式,即确定相机的内参数 K 。本文采用编码点的方式及绕相机光轴旋转拍摄的4幅图像(每幅图像是前一幅相对后一幅绕单CCD相机

自身光轴旋转 90°得到的)来进行相机的自标定。

3.1 确定变换矩阵

根据 Hartley 的标定方法^[6] 标定矩阵满足如下方程 $(KK^T \chi H_i^{-1})^T = H_i(KK^T)$ 其中 H_i 为射影变换矩阵。由此可见 若要求标定矩阵 K 则必须先求出 H_i 。

根据文献 [7], 一对图像点对 $u \hat{u}$ 满足如下关系 $u \sim H\hat{u}$ H 为 3×3 非奇异射影变换矩阵, 也称为单应矩阵。一般确定一个射影变换矩阵, 只需 4 个图像点对。在本文的方法中, 在每幅图像中, 至少有 5 个编码点可见, 即两幅图像中至少有 5 对编码点匹配。这样就可以利用其冗余信息来提高算法的鲁棒性。由于多于 4 个点对, 因此可通过解过约束方程组来求出最小二乘意义下的最优解。通过如下方程即可求解各个像点对的射影变换矩阵 H_i ^[6]。

$$\begin{aligned} h_3^T u_i \hat{u}_i &= h_1^T u_i \\ h_3^T u_i \hat{v}_i &= h_2^T u_i \end{aligned} \quad (2)$$

其中 $u_i = (u_i, v_i, 1)^T$ $\hat{u}_i = (\hat{u}_i, \hat{v}_i, 1)^T \cdot h_1^T h_2^T h_3^T$ 表示射影变换矩阵 H 的各行向量。

3.2 标定矩阵的确定

假定已经拍摄了 $\chi(i=1 \dots m)$ 幅图像, 每幅图像上有 $\chi(j=1 \dots n)$ 个点, 则标定矩阵满足如下方程:

$$(KK^T \chi H_i^{-1})^T = H_i(KK^T) \quad (3)$$

该方程可以用于求矩阵 KK^T 的未知量。 KK^T 可以表示为

$$C = KK^T = \begin{pmatrix} a & b & c \\ b & d & e \\ c & e & f \end{pmatrix} \quad (4)$$

由于 C 中含有 6 个独立未知量, 因此式 (3) 可以产生 9 个线性方程。显然在已知单个变换矩阵 H_i 的情况下, 用等式 (3) 的 9 个方程来解 C 是不够的, 然而, 若两个或更多的 H_i 已知的话, 则可以利用过约束方程组来求最小二乘意义下 C 的解。这样为了提供两个独立的变换矩阵, 就需要 3 幅图像。一旦 $C = KK^T$ 求出, 就可以用 Cholesky 分解法求出上三角矩阵 K 。

根据以上讨论, 结合本文的方法, 本文提出的算法如下:

(1) 从单 CCD 中心位置记录 4 幅标定图像, 而且每幅图像都是前一幅相对后一幅绕单 CCD 自身光轴旋转 90°得到的;

(2) 根据编码点的自动识别、匹配算法(这里不进行讨论), 图像间至少建立 5 对对应点(point correspondences);

(3) 对于每幅图像 $I_i(i=2 \ 3 \ 4)$, 计算第 1 幅与第 i 幅的匹配的射影变换矩阵 H_i ;

(4) 根据 H_i (令 $C = KK^T$) 求解一个上三角阵。

若 C 求出, 则利用 Cholesky 分解法即可确定 K 。

由于编码点的自动识别和匹配非常准确, 故本文算法受噪音的影响较小, 一般会解出 C , 进而可求解出 K 。

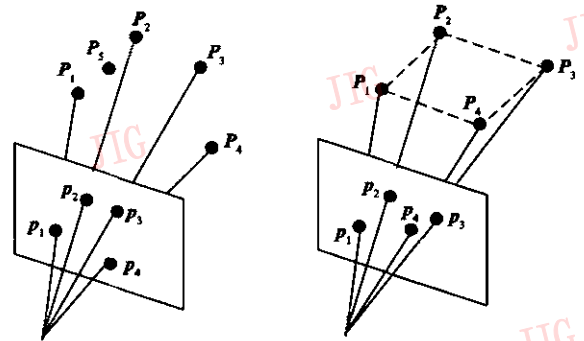
4 3 维重建和相机姿态恢复

4.1 射影重建和相机位置恢复

通常, 对一个带有齐次坐标的图像点 u 的透视投影可以表达为

$$u \sim H(I| - \hat{T})X \sim H(\hat{X} - \hat{T}) \quad (5)$$

X, \hat{X} 分别是 3 维点的齐次和非齐次扩展的笛卡儿坐标, \hat{T} 是相机中心扩展的笛卡儿坐标, H 为单应矩阵(如图 1 所示)。



(a) 一般配置 (b) 参考面配置

图 1 射影 3 维基点 $P_1 \dots P_5$ 和图像基点 $p_1 \dots p_4$

Fig. 1 Projective 3D basis points $P_1 \dots P_5$ and image basis points $p_1 \dots p_4$

若用 P_1, P_2, P_3, P_4, P_5 5 点作为基, 则任何 3 维点 X 和相机中心 T 可用射影坐标(下角 c 代表 center)表达为

$$X \sim X\bar{P}_1 + Y\bar{P}_2 + Z\bar{P}_3 + W\bar{P}_4 \quad (6)$$

$$T \sim X_c\bar{P}_1 + Y_c\bar{P}_2 + Z_c\bar{P}_3 + W_c\bar{P}_4$$

同样地, 任何图像点坐标也可用作为基的 4 个图像点来表示:

$$u \sim u\bar{p}_1 + v\bar{p}_2 + w\bar{p}_3 \quad (7)$$

通过选择正则化 \bar{P} 和 \bar{p} , 以便于基点 P_5 和像基点 p_4 分别有射影坐标 $(1, 1, 1, 1)^T$ 和 $(1, 1, 1)^T$ 。对于一般情况和 4 个共面点在参考面上的情况, 其映射关系是

$$M: (X, Y, Z, W)^T \rightarrow (x, y, w)^T$$

W, w 分别 3 维空间点齐次坐标和 2 维平面点齐次坐标的权因子(下同)。

4.1.1 一般点配置

这种情况在文献 [8] 中应用过。如图 1(a) 所示 3 维基点 P_1, P_2, P_3, P_4 的投影是 p_1, p_2, p_3, p_4 。由此可将射影坐标系下的 3 维点、相机中心和像点坐标写成如下两个约束方程：

$$\begin{aligned} w \frac{X}{X_c} - u \frac{Z}{Z_c} + (u - w) \frac{W}{W_c} &= 0 \\ w \frac{Y}{Y_c} - v \frac{Z}{Z_c} + (v - w) \frac{W}{W_c} &= 0 \end{aligned} \quad (8)$$

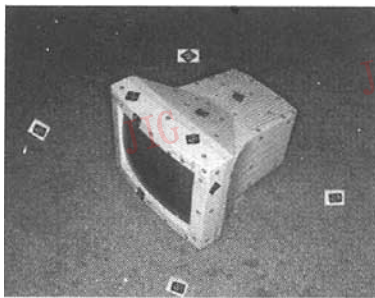
上述方程反映了 3 维点齐次射影坐标 $(X, Y, Z, W)^T$ 与相机中心坐标的倒数 $(X_c^{-1}, Y_c^{-1}, Z_c^{-1}, W_c^{-1})$

是双线性关系。

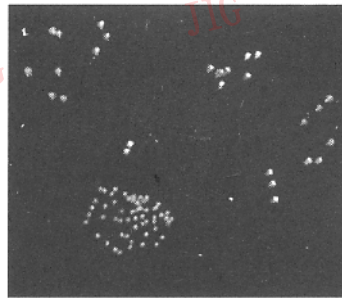
4.1.2 4 个点在一参考面上^[9]

4 个点在一参考面上情况, Heyden 和 Aström 讨论过。若 4 个点 P_1, P_2, P_3, P_4 (本文算法, 是用同一个平面上的 4 个编码点表示, 如图 2(a) 和图 3(a) 所示) 在一个参考面上, 并在所有图像上都可见, 那么它们的图像可用作图像的坐标基(如图 1(b) 所示) 此时式 (8) 变为

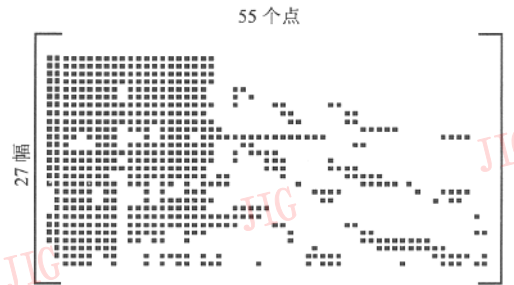
$$\begin{aligned} u \left(\frac{Z}{W} - \frac{Z_c}{W_c} \right) - v \left(\frac{X}{W} - \frac{X_c}{W_c} \right) &= 0 \\ v \left(\frac{Z}{W} - \frac{Z_c}{W_c} \right) - w \left(\frac{Y}{W} - \frac{Y_c}{W_c} \right) &= 0 \\ u \left(\frac{Y}{W} - \frac{Y_c}{W_c} \right) - v \left(\frac{X}{W} - \frac{X_c}{W_c} \right) &= 0 \end{aligned} \quad (9)$$



(a) 一幅显示器图像



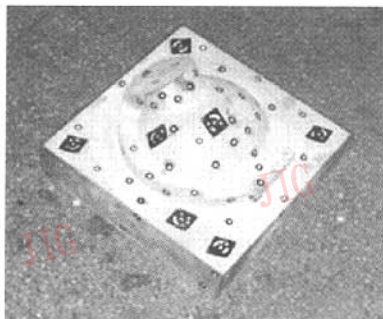
(b) 显示器标记点(非编码点)的 3 维重建和相机姿态恢复



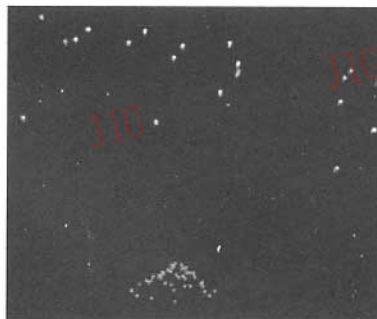
(c) 27 幅显示器的图像和 55 个非编码点的可见度矩阵

图 2 显示器的实验结果

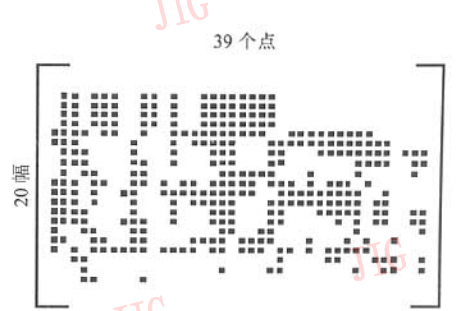
Fig. 2 The experimental results of the monitor



(a) 一幅零件模具图像, 其 4 个共面编码点在零件的 4 个角上



(b) 标记点(非编码点)的 3 维重建和相机姿态恢复



(c) 20 幅零件模具图像和 39 个非编码点的可见度矩阵

图 3 零件模具的实验结果

Fig. 3 The experimental results of the mould

该方程反映了在射影坐标系里 3 维点的非齐次坐标 $(X/W$ 等) 和相机中心的非齐次坐标 $(X_c/W_c$ 等) 关系变为线性的。若 $W = 0$, 则相当于参考面处

于无穷远处, 从数学意义上讲, 这相当于纯平移一个已标定的相机^[5]。

因此, 若通过简单地增加 4 个共面点为一个参

考面,则通常情况下的从多幅图像中的多个点恢复 3 维点和相机中心的双线性问题就转化成线性问题,也就是说,变为求解某一矩阵的零空间问题,该矩阵含有所有可见图像中的像点的齐次坐标。

当参考面上的 4 个点在所有图像中都可见的情况下,这个矩阵可由任意幅图像中的任意数目的点来构造。对于 m 幅图像中的 n 个点的线性方程组可采用如下形式^[5]：

$$\begin{pmatrix} S_{1,1} & 0 & 0 & \dots & 0 & 0 & -S_{1,1} & 0 & \dots & 0 \\ S_{1,2} & 0 & 0 & \dots & 0 & 0 & 0 & -S_{1,2} & \dots & 0 \\ \vdots & & & & & & \vdots & & & \\ S_{1,m} & 0 & 0 & \dots & 0 & 0 & 0 & 0 & \dots & -S_{1,m} \\ 0 & S_{2,1} & 0 & \dots & 0 & 0 & -S_{2,1} & 0 & \dots & 0 \\ 0 & S_{2,2} & 0 & \dots & 0 & 0 & 0 & -S_{2,2} & \dots & 0 \\ \vdots & & & & & & \vdots & & & \\ 0 & S_{2,m} & 0 & \dots & 0 & 0 & 0 & 0 & \dots & -S_{2,m} \\ & & & & \vdots & & & & & \\ 0 & 0 & 0 & \dots & 0 & S_{n,1} & -S_{n,1} & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 & S_{n,2} & 0 & -S_{n,2} & \dots & 0 \\ & & & & & \vdots & & & & \\ 0 & 0 & 0 & \dots & 0 & S_{m,n} & 0 & 0 & \dots & -S_{m,n} \end{pmatrix} \begin{pmatrix} \hat{X}_1 \\ \hat{Y}_1 \\ \hat{Z}_1 \\ \vdots \\ \hat{X}_n \\ \hat{Y}_n \\ \hat{Z}_n \\ \hat{X}_1^c \\ \hat{Y}_1^c \\ \hat{Z}_1^c \\ \vdots \\ \hat{X}_m^c \\ \hat{Y}_m^c \\ \hat{Z}_m^c \end{pmatrix} = 0 \quad (10)$$

非齐次射影点的坐标可用 $\hat{X} = \frac{X}{W}, \hat{Y} = \frac{Y}{W} \dots$ 等表示,相机中心的坐标用 $\hat{X}_c = \frac{X_c}{W_c}, \hat{Y}_c = \frac{Y_c}{W_c}$ 等表示。

$$S_{i,j} = \begin{pmatrix} 0 & w_{i,j} & -v_{i,j} \\ -w_{i,j} & 0 & u_{i,j} \\ v_{i,j} & -u_{i,j} & 0 \end{pmatrix} \quad (11)$$

是一个 3×3 矩阵,其由第 i 幅图像中的第 j 个可见点的图像坐标所构成。当某些点在图像中不可见时,如第 j 个点在第 i 幅图像中不可见,则 $S_{i,j}$ 在式(10)中将被忽略。

对上述矩阵进行奇异值分解(singular value decomposition SVD),即可确定该矩阵的零空间,亦即确定了 3 维点和相机中心的射影坐标。

4.2 欧氏重建

当相机自标定完成后,即求出 K 后,本文就可先采用文献[10]的方法来确定旋转矩阵 R , 然后进行 3 维点的欧氏重建。

5 算法

本文的算法如下：

(1) 根据第 3 节介绍的方法进行单 CCD 相机

标定,即确定单 CCD 相机的内部参数 K ;
 (2) 根据第 4 节介绍的方法进行 3 维点和单 CCD 摄像机中心的射影重建 ;
 (3) 根据文献[10]的方法确定旋转矩阵 R , 然后进行 3 维点的欧氏重建。
 由以上步骤即可实现基于标记点丢失的多幅透视图图像的 3 维重建和相机姿态恢复。

6 算法实现与实验结果

本文算法是用 VC++ 语言并连接 IPL(image processing library)图像处理库,在 PC 机上实现的。本实验用的单 CCD 相机(Nikon 数码相机),其分辨率为 4256×2848 , 实验图像是以自由拍摄方式和任意多角度获得的场景中的 27 幅显示器图像及 20 幅某一零件模具图像。实验时,首先利用自动识别匹配软件完成了一系列图像对应标记点(编码点和非编码点)的匹配 ;然后采用本文方法得到了两个测试实例的实验结果(如图 2、图 3 所示)。图 2(a)和图 3(a)分别为它们其中的一幅,图 2(c)的可见度矩阵表示 55 个标记点(非编码点)在 27 幅图像中部分可见情况。若第 j 点在第 i 幅图像中可见,则用小黑方块表示。前 4 幅图像用于相机的自标定,其可见

度矩阵形式一样,说明相机位置、拍摄场景的范围、光照等均未变化。矩阵的左半部可见点多,说明显示器的上表面的非编码点在大部分图像中可见,而矩阵的右部可见点稀疏,说明显示器的前、后及侧面上的非编码点在有些图像中出现隐藏(occlusions)现象。

由此可见,采用本文方法,在不是所有标记点(非编码点)在所有图像上均可见的情况下,也能够恢复标记点(非编码点)的3维重建和相机姿态。由于标记点的3维重建具有较高的精度,因此完全能够满足逆向工程等应用要求。

本文的方法只能间接地恢复相机姿态位置,而不能解决摄影盲区的问题。因此,该问题有待于进一步研究。此外,由于本文的算法为线性的,为此下一步工作还将研究如何进一步优化的问题。

参考文献(References)

- 1 Han M, Kanade T. Creating 3D models with uncalibrated cameras [A]. In: IEEE Computer Society Workshop on the Application of Computer Vision [C], Palm Springs California, USA, 2000: 178~185.
- 2 Tomasi C, Kanade T. Shape and motion from image streams under orthography: a factorization method [J]. International Journal of Computer Vision, 1992, 9(2): 137~154.
- 3 Sturm P, Triggs B. A factorization based algorithm for multi-image projective structure and motion [A]. In: European Conference on Computer Vision [C], Cambridge, U. K., 1996: 709~720.
- 4 Heyden A, Berthilsson R, Sparr G. An iterative factorization method for projective structure and motion from image sequences [J]. Image and Vision Computing, 1999, 17(13): 981~991.
- 5 Rother C, Carlsson S. Linear multi view reconstruction and camera recovery using a reference plane [J]. International Journal of Computer Vision, 2002, 49(2/3): 117~141.
- 6 Hartley R. Self-calibration of stationary cameras [J]. International Journal of Computer Vision, 1997, 22(1): 5~23.
- 7 Faugeras O, Luong Q. Geometry of multiple images [M]. Cambridge, MA, USA: Massachusetts Institute of Technology Press, 2001.
- 8 Carlsson S. Duality of reconstruction and positioning from projective views [A]. In: Proceedings of IEEE Workshop on Representation of Visual Scenes [C], Boston, MA, USA, 1995: 85~92.
- 9 Heyden A, Åström K. A canonical framework for sequences of images [A]. In: Proceedings of IEEE Workshop on Representation of Visual Scenes [C], Boston, MA, USA, 1995: 45~52.
- 10 Oliensis J, Genc Y. Fast algorithms for projective multi frame structure from motion [A]. In: International Conference on Computer Vision [C], Kerkyra, Greece, 1999: 536~542.