

基于 1 维子空间线性迭代射影重建

刘侍刚¹, 吴成柯¹, 李良福², 宁纪锋¹

(1. 西安电子科技大学综合业务网国家重点实验室, 陕西西安 710071; 2. 陕西师范大学计算机科学学院, 陕西西安 710065)

摘 要: 提出了一种基于 1 维子空间线性迭代的射影重建方法. 该方法利用所有图像序列的行向量生成的子空间之和, 与射影重建空间点的行向量生成的子空间是同一线性子空间, 同时, 由第 1 幅图像的 3 个行向量及另外一个行向量可以构成该线性子空间的一个基底的特性, 线性迭代求取这个行向量及图像深度因子, 最后完成射影重建. 模拟实验和真实实验数据结果表明, 该射影重建方法具有鲁棒性好、收敛性好以及重投影误差小等优点.

关键词: 射影重建; 分解算法; 线性迭代

中图分类号: TP391.41; P232 **文献标识码:** A **文章编号:** 0372-2112 (2007) 04-0692-05

A Linear Iterative Method for Projective Reconstruction Based on 1D Subspace

LIU Shi-gang¹, WU Cheng-ke¹, LI Liang-fu², NING Ji-feng¹

(1. National Key Laboratory of ISN, Xidian University, Xi'an, Shaanxi 710071, China;

2. School of Computer Science, Shaanxi Normal University, Xi'an, Shaanxi 710065, China)

Abstract: A linear iterative method for projective reconstruction based on 1D subspace is presented in the paper. It relies on the facts that the rows in the matrix including all the image points span the same linear subspace as the rows in the matrix including space points and that the basis of the subspace can consist of the three rows in the matrix comprising the first image points and a row vector which is orthogonal the former. The row vector and the depth factors are obtained based on linear iteration, and lastly the projective reconstruction is accomplished. The experiments with both simulate and real data show that the method is efficient, robust and has good property of convergence and small re-projection errors.

Key words: projective reconstruction; factorization method; linear iteration

1 引言

从图像序列中重建出三维场景结构是计算机视觉的主要目标之一. 同时, 射影重建是欧氏重建的必要阶段, 其精确性将影响最终的重建结果. 目前, 它仍然是计算机视觉领域中的研究热点之一. Faugeras 指出^[1], 如果没有任何先验知识, 从图像序列中只能得到射影重建.

现有的射影重建方法大部分都是基于多线性约束关系^[2]. 当只有 2 幅、3 幅和 4 幅图像时, 分别存在一个双线性约束 (也称为对极几何约束)、三线性约束和四线性约束关系; 当图像数大于 4 时, 则不存在这样的几何约束关系^[3]. Vidal 等人利用对极几何约束关系对多个物体进行射影重建^[4]; Martinec 和 Pajdla 利用三线性约束关系对存在遮挡的物体进行射影重建^[5]. 利用多线性约束关系进行射影重建的缺点是它仅能够利用 2 幅至 4 幅图像, 而不能利用更多的图像.

在射影重建过程中, 有许多利用景物中的先验知识 (例如: 平行线、直角、平面等) 来实现射影重建的方法^[6~8], 这些方法的最大缺点是要利用景物中的先验知识, 但是在许多情况下, 这些景物中的先验知识很难知道, 因此这些方法的应用受到一定的限制. 当景物的深度远小于到相机的距离时, 文献^[9, 10]将所有空间点的深度认为相同, 利用仿射模型进行重建. 在实际情况下, 空间各点的深度很难保证一致, 因此基于该模型的重建精度较低. 为了克服上述缺点, 基于 Tomasi-Kanade^[11] 正交投影分解算法的思想, Peter 提出了利用分解算法来进行射影重建^[12]. 该方法的缺点是, 在进行分解之前, 各个图像点的深度因子必须要事先求到, 而深度因子的求解又要利用对极几何约束关系. 这样, 又遇到了前面所述的问题.

本文针对上述缺点, 利用所有的图像序列生成的子空间之和与射影空间点生成的子空间是同一子空间,

同时,由第 1 幅图像点构成的 3 个行向量及另外一个行向量可以构成该子空间的一个基底的特性,线性迭代求到这个行向量及图像深度因子,最后完成射影重建.模拟实验和真实实验数据结果表明,该射影重建方法具有鲁棒性好、收敛性好等优点.

2 相机模型及投影矩阵

本文假定摄像机模型为经典的针孔模型,即成像过程可以用下列方程表示:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} [R \quad t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \Leftrightarrow m = K [R \quad t] X \quad (1)$$

式中 $X = [x \ y \ z \ 1]^T$ 表示三维空间点的齐次坐标, $m = [u \ v \ 1]^T$ 为对应的图像平面点的齐次坐标; K 为摄像机的内参矩阵,其中, (u_0, v_0) 为主点坐标, f_u 和 f_v 分别为图像 u 轴和 v 轴的尺度因子, s 为畸变因子; R, t 分别为摄像机在拍摄位置对应的旋转矩阵和平移矩阵,即摄像机的外参矩阵.

若令

$$P = K [R \quad t] \quad (2)$$

式中 P 称为摄像机的投影矩阵.则式(1)可以简写成:

$$m = PX \quad (3)$$

3 子空间描述

假设有 m 幅图像,每幅图像上有 n 个对应的图像点,则式(3)为:

$$m_{i,j} = P_i X_j \quad i=1, 2, \dots, m \quad j=1, 2, \dots, n \quad (4)$$

式中 i 表示第 i 幅图像, j 表示第 j 个图像点.

将所有空间点的齐次坐标表示成一个矩阵

$$X = \begin{bmatrix} x_1 & x_2 & x_3 & \dots & x_n \\ y_1 & y_2 & y_3 & \dots & y_n \\ z_1 & z_2 & z_3 & \dots & z_n \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \\ \mathbf{1} \end{bmatrix} \quad (5)$$

则对于第 i 幅图像上所有点及深度因子也有下式

$$m_i = \begin{bmatrix} u_{i,1} & u_{i,2} & u_{i,3} & \dots & u_{i,n} \\ v_{i,1} & v_{i,2} & v_{i,3} & \dots & v_{i,n} \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} = \begin{bmatrix} u_i \\ v_i \\ \mathbf{1} \end{bmatrix} \quad (6)$$

$$i = \text{diag}(i_1 \quad i_2 \quad \dots \quad i_n) \quad (7)$$

根据式(4),则有:

$$m_i = P_i X \quad (8)$$

用 i 表示由 m_i 的行向量生成的 3 维线性子空间,即:

$$i = \text{span}\{u_i, v_i, \mathbf{1}\} \quad (9)$$

式中 u_i 和 v_i 分别表示由第 i 幅图像上的所有图像点和 v 分量组成的行向量, $\mathbf{1}$ 表示由 n 个 1 组成的行向量,即:

$$u_i = (u_{i,1} \quad u_{i,2} \quad \dots \quad u_{i,n}) \quad (10)$$

$$v_i = (v_{i,1} \quad v_{i,2} \quad \dots \quad v_{i,n}) \quad (11)$$

$$\mathbf{1} = (1 \quad 1 \quad \dots \quad 1) \quad (12)$$

同样,用 i 表示由 X 的行向量生成的 4 维线性子空间,即:

$$i = \text{span}\{x, y, z, \mathbf{1}\} \quad (13)$$

式(13)中 x, y 和 z 分别表示由空间点的 x 轴、 y 轴和 z 轴的分量组成的行向量,即:

$$x = (x_1 \quad x_2 \quad \dots \quad x_n) \quad (14)$$

$$y = (y_1 \quad y_2 \quad \dots \quad y_n) \quad (15)$$

$$z = (z_1 \quad z_2 \quad \dots \quad z_n) \quad (16)$$

对于第 i 幅图像的 m_i , 同样可以生成一个 3 维的线性子空间,即:

$$i = \text{span}\{u_i, v_i, \mathbf{1}\} \quad (17)$$

由式(8)可得

$$i \subset i \quad (18)$$

$$i_1 + i_2 + \dots + i_m \subseteq i \quad (19)$$

若摄像机作特殊运动,式(19)中取 \subset 号,对于一般情况下取等号.在本文中,并不考虑摄像机作特殊运动,因此本文取等号.从式(19)中可以看出,在一般情况下,至少需要 2 幅图像才能够求到线性子空间 i ,即完成射影重建.

同时,从式(8)中可以看出,对于任何一个非奇异矩阵 T ,都有

$$m_i = P_i X = P_i T^{-1} T X = P_i X \quad (20)$$

式中

$$P_i = P_i T^{-1}, \quad X = T X \quad (21)$$

即,对于一组图像点序列 m ,如果存在一组重建 P 和 X ,则同样存在另一组重建 P 和 X .对于无穷多的非奇异矩阵 T ,存在无穷多组重建满足要求.如果没有任何先验知识,不能够知道所求得的重建对应的是哪一个变换 T .从而这个重建落在射影空间而不是欧氏空间.但从式(21)可以看出,由于 T 为非奇异矩阵,因此,重建 X 和 X 的行向量生成的线性子空间为同一子空间,只是它们的基底不一样,而且 X 和 X 的行向量都可以作为这个子空间的一个基底.

4 基于 1 维子空间线性迭代射影重建

定理 在由空间点 X 生成的线性子空间 i 中,任何一个基底都可以作为射影重建 X 的结果.

证明 设 $\{i_1, i_2, i_3, i_4\}$ 是由 X 生成的子空间中的任意一个基底,则存在一个非奇异矩阵 T ,使

$$(\quad \quad \quad \quad)^T = TX \quad (22) \quad \quad \quad \subseteq \quad (33)$$

由式(21)可得,若令 $P_i = P_i T^{-1}$, $X = TX = (\quad \quad \quad \quad)^T$, 则 P 和 X 可以认为是一组射影重建. 证毕

由上述定理可知,线性子空间 \mathcal{L}_i 中的任何一组基底都当可以作射影重建结构 X 的结果. 而由式(18)可知, $\mathcal{L}_1 \subset \mathcal{L}$, 因此, \mathcal{L}_1 空间中的任意一组基底加上一个行向量 w , 它既在 \mathcal{L}_1 正交补空间, 又在 \mathcal{L} 空间, 就可以构成 \mathcal{L} 空间中的一组基, 即完成了射影结构重建 X . 而由式(17)可知, $\{u_{1,1}, v_{1,1}, \mathbf{1}_{1,1}\}$ 是线性子空间 \mathcal{L}_1 的一组基底, 因此, $\{u_{1,1}, v_{1,1}, \mathbf{1}_{1,1}, w\}$ 可以构成线性子空间 \mathcal{L} 的一组基底, 即可令射影结构重建为

$$X = [u_{1,1} \quad v_{1,1} \quad \mathbf{1}_{1,1} \quad w]^T \quad (23)$$

下面来求线性子空间 \mathcal{L} 中的 w 向量. 为了表示方便, 令

$$\hat{m}_i = m_i \quad (24)$$

将空间一个向量投影到由第 1 幅图像点及深度因子所构成的矩阵 \hat{m}_1 行所生成的线性子空间的正投影矩阵 T 为^[13]

$$T = \hat{m}_1^T (\hat{m}_1 \hat{m}_1^T)^{-1} \hat{m}_1 \quad (25)$$

投影到其正交补空间中的投影矩阵 T 为

$$T = I - \hat{m}_1^T (\hat{m}_1 \hat{m}_1^T)^{-1} \hat{m}_1 \quad (26)$$

令

$$m_i = \hat{m}_i T \quad (27)$$

$$X = XT \quad (28)$$

式(27)和式(28)表示线性子空间 \mathcal{L}_i ($i = 2, 3, \dots, m$) 和 \mathcal{L} , 经投影矩阵 T 变换后, 得到新的线性子空间 \mathcal{L}_i 和 \mathcal{L} , 子空间 \mathcal{L}_i 和 \mathcal{L} 分别由 m_i 和 X 的行向量所生成.

定理 经投影矩阵 T 变换后的线性子空间 \mathcal{L}_i 为 1 维子空间, 即 $\dim(\mathcal{L}_i) = 1$ 且有: $\mathcal{L}_i \subseteq \mathcal{L}$, $\mathcal{L}_i \perp \mathcal{L}_j$.

证明: 由于 $\dim(\mathcal{L}) = 4$ 和 $\dim(\hat{m}_1) = 3$, 且 $\mathcal{L}_1 \subset \mathcal{L}$, 则:

$$\dim(\mathcal{L}) = \dim(\mathcal{L}) - \dim(\hat{m}_1) = 1 \quad (29)$$

由式(19)可得:

$$\mathcal{L}_i \subseteq \mathcal{L} \quad (30)$$

由 \mathcal{L}_1 的定义可知, \mathcal{L}_1 为零空间, 因此有:

$$\mathcal{L}_i \subseteq \mathcal{L} \quad (31)$$

若摄像机作特殊运动, 式中取 $\mathcal{C} = \mathcal{L}$.

由于 $\mathcal{L}_1 \subset \mathcal{L}$, 则有:

$$\mathcal{L}_1 \perp \mathcal{L}_i \quad (32)$$

即有:

由上面式(29)和式(31)就可以构造一个算法来求射影重建, 由于 m_i 行向量生成的子空间 \mathcal{L}_i 之和与由 X 行向量生成的子空间 \mathcal{L} 是同一子空间, 且维数为 1, 令

$$\bar{m} = \begin{bmatrix} m_2 \\ m_3 \\ \dots \\ m_m \end{bmatrix} \quad (34)$$

对 \bar{m} 进行奇异值分解(SVD), 即

$$\bar{m} = UVD^T \quad (35)$$

式中 V 是一个对角阵, 在没有噪声的情况下, 对角阵 V 只有第一个元素 $v_{1,1}$ 为非零, 其它元素都为零. 可令 d_1 为 D 第 1 列元素.

定理 $d_1^T \in \mathcal{L}_1$, $d_1^T \in \mathcal{L}$.

证明: 由式(35)可得:

$$\bar{m} = UVD^T = U \begin{bmatrix} v_{1,1} & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ \dots & \dots & \ddots & \dots \\ 0 & 0 & \dots & 0 \end{bmatrix} [d_1 \quad d_2 \quad \dots \quad d_m]^T$$

$$= U \begin{bmatrix} v_{1,1} d_1^T \\ \mathbf{0} \\ \dots \\ \mathbf{0} \end{bmatrix} \quad (36)$$

由于 U 为非奇异矩阵, 所以由矩阵 \bar{m} 的行向量生成的子空间和 d_1^T 生成的子空间是同一子空间. 因此, $d_1^T \in \mathcal{L}$. 同时, 由式(33)可得 $d_1^T \in \mathcal{L}_1$.

同时, 由 \bar{m} 的定义可得 $d_1^T \in \mathcal{L}_1$. 证毕

因此, 式(23)中的 w 可令为 d_1^T , 即可求到射影重建结构 X .

由于图像中含有噪声, 同时, 由于在求解过程中, 深度因子也不知道, 只能够近似的假设, 因此矩阵 V 中的对角元素不可能只有 $v_{1,1}$ 为非零, 可以令其它元素为零.

若已经射影结构 X , 可利用式(8)线性地求到深度因子及投影矩阵, 但本文还是利用投影的方法来求解深度因子 i .

投影到 \mathcal{L}_i 正交补空间中的投影矩阵 T_d 为

$$T_d = I - d_i (d_i^T d_i)^{-1} d_i^T \quad (37)$$

由于 $\mathcal{L}_i \subseteq \mathcal{L}$, 可得

$$m_i T_d = \mathbf{0} \quad (38)$$

利用式(38)可以线性求解深度因子 i .

由前面分析可以构造一个迭代算法, 通过线性交替地求取 w 及深度因子 i , 最后达到收敛, 完成射影重

建.

在上面的求解过程中,还存在一个问题,那就是深度因子不能唯一的确定^[14],为了解决这个问题,可以假设 $\lambda = \text{diag}(1 \ 1 \ \dots \ 1)$ 来求解 $\lambda_i (i=2,3, \dots, m)$.

5 算法总结

1. 假设所有的深度因子都为 1,令 ϵ 为任意小的一个正数;
2. 利用式(35)对 \bar{m} 进行 SVD 分解,求解到 w ,若 $v_{2,2} > \epsilon$,则转至第 4 步,否则至第 3 步;
3. 利用式(38)求解到深度因子 $\lambda_i (i=2,3, \dots, m)$,返回到第 2 步;
4. 利用式(23)求到射影重建结构 X ,利用式(8)求到投影矩阵 P_i ,最后完成射影重建.

6 模拟与真实图像实验

6.1 模拟实验

为了检验本文所提出的射影重建方法的收敛性,用计算机模拟在单位球内随机产生 100 个空间点,模拟摄像机的内参数 $f_u = 1000, f_v = 1000, s = 0, u_0 = 320, v_0 = 240$,变换摄像机的外参数以产生 8 幅大小为 $640 \times$

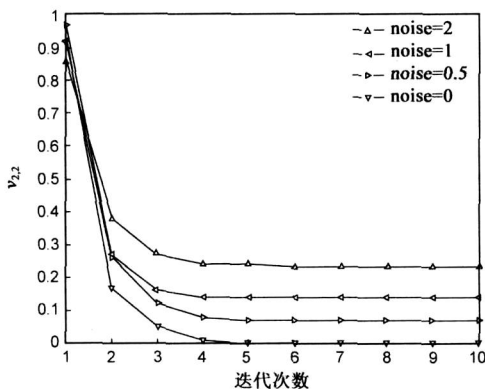


图 1 $v_{2,2}$ 的值随迭代次数变化图

480 的图像,并利用这些模拟图像点用本文提出的方法进行射影重建,模拟结果如图 1 和图 2 所示,其中横坐

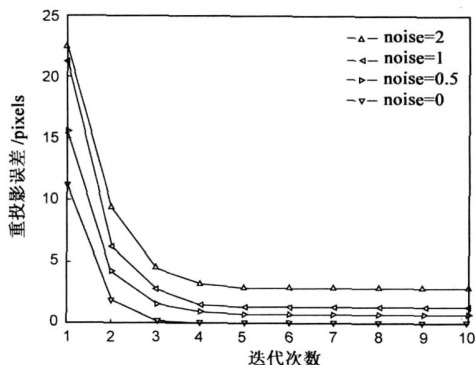


图 2 重投影误差随迭代次数变化图

标轴为迭代次数,图 2 的纵坐标轴为重投影误差,其定义如下:

$$= \frac{1}{mn} \sum_{j=1}^m \sum_{i=1}^n P_i x_j - u_{ij} \quad (39)$$

从图 1 及图 2 可以看出,本文所提出的射影重建方法迭代 5 次以内就能够收敛,因此该方法具有良好的收敛性.

同时,为了研究图像数对本文方法的影响,本文用上述方法模拟产生 2 至 16 幅图像,并在每幅图像中加入 1 个像素的高斯噪声,在不同图像数下各进行 100 次实验,利用式(39)求得重投影误差,然后取均值,实验结果如图 3 所示.

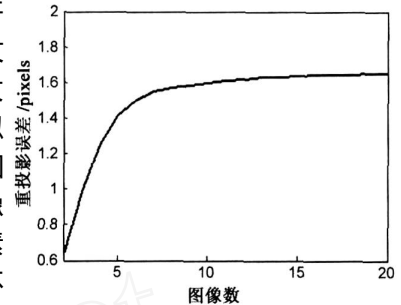


图 3 重投影误差随图像噪声变化图

从图 3 可以看出,随着图像数的增加,重投影误差会增大,而且增加的速度越来越慢,最后达到稳定.

为了比较本文方法和 Peter 的分解方法^[12]及 Quan 的三线性约束方法^[9]的重建性能,模拟产生 8 幅图像,在图像像素中加入高斯噪声,并在每种噪声

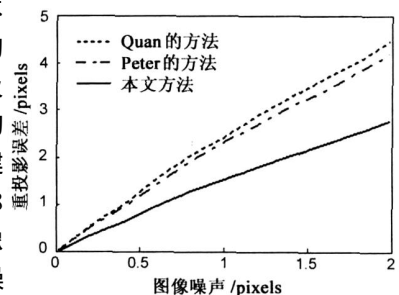


图 4 重投影误差随图像噪声变化图

水平下用上述三种射影重建方法各进行 100 次实验,然后取均值,实验结果如图 4 所示.

从图 4 中可以看出,本文方法的重投影误差比 Peter 的分解方法和 Quan 的三线性约束方法重投影误差都要小,原因是 Peter 的方法和 Quan 的方法都要计算多线性约束关系,而多线性约束关系的计算对图像噪声非常敏感^[3],本文方法却不需要计算多线性约束关系,因此本文方法具有更小的重投影误差.从图中还可以看出,在 2 个像素的噪声以内,本文方法求到重投影误差只比原来的图像的噪声约大 40%,具有较小的重投影误差,因此,具有较好的重建结果.

6.2 真实图像实验

为了验证本文方法的正确性及重建效果,本文获得一个由 10 幅图像组成的房子图像序列,图像的大小为 512×480 ,其中两幅图像如图 5 所示.我们用 Harris 算子提取了 50 个特征点(如图 5 中 * 所示),并利用这

些特征点进行射影重建.为了显示重建效果,本文将重建后的三维空间点重投影到原来图像中,重投影点如图5中所示,从图中可以看出,重投影点(○)和特征点(*)基本重合.同时,利用式(39)求到重投影误差为0.36个像素,因此本文提出的射影重建方法具有较小的重投影误差,也说明该方法具有较好的重建效果.

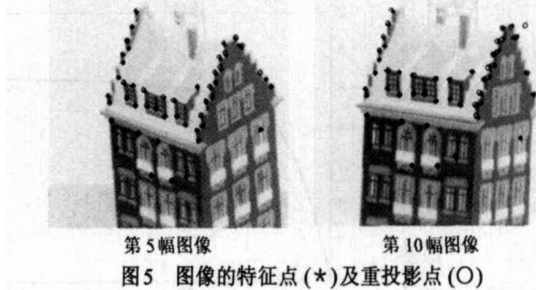


图5 图像的特征点(*)及重投影点(○)

7 结束语

本文提出了一种利用1维子空间线性迭代来进行射影重建,模拟实验结果表明,该方法只需要5次以内迭代就可以收敛,同时,重投影误差约大于图像的噪声的40%,因此,该方法具有良好的收敛性、鲁棒性及重投影误差小等优点.真实图像实验结果表明,本文提出的射影重建方法是非常有效的.

参考文献:

- [1] O Faugeras. What can be seen in three dimensions with an uncalibrated stereo rig? [A]. ECCV, Lecture Notes in Computer Science [C]. London: Springer-Verlag, 1992, 588: 563 - 578.
- [2] R Hartley, A Zisserman. Multiple View Geometry in Computer Vision [M]. Cambridge Univ. Press, Cambridge, UK, 2000.
- [3] O Faugeras, Q Luong. The Geometry of Multiple Images [M]. MIT Press, Boston, 2004.
- [4] R Vidal, Y Ma, S Soatto, S Sastry. Two-view multibody structure from motion [J]. International Journal of Computer Vision, 2006, 68(1): 7 - 25.
- [5] D Martinec, T Pajdla. Structure from many perspective images with occlusions [A]. ECCV, Lecture Notes in Computer Science [C]. Berlin: Springer-Verlag, 2002, 2351: 355 - 369.
- [6] P Baker, Y Aloimonos. Structure from motion of parallel lines [A]. ECCV, Lecture Notes in Computer Science [C]. Berlin: Springer-Verlag, 2004, 3024: 229 - 240.
- [7] Y Liu, C Wu, H Tsui. A practical approach for 3D building modeling from uncalibrated video sequences [J]. International Journal of Image and Graphics, 2002, 2(2): 287 - 307.
- [8] M Pollefeys, F Verbiest, L Gool. Surviving dominant planes in uncalibrated structure and motion recovery [A]. ECCV, Lecture Notes in Computer Science [C]. Berlin: Springer-Verlag, 2002, 2351: 837 - 851.
- [9] L Quan, M Lhuillier. Structure from motion from three affine views [A]. ICPR [C]. Quebec City: IEEE Press, 2002, 1 - 6.
- [10] Q Ke, T Kanade. Robust L1 norm factorization in the presence of outliers and missing data by alternative convex programming [A]. CVPR [C]. San Diego, CA: IEEE Press, 2005, 739 - 746.
- [11] C Tomasi, T Kanade. Shape and motion from image streams under orthography: a factorization method [J]. International Journal of Computer Vision, 1992, 9(2): 137 - 154.
- [12] P Sturm, B Triggs. A factorization based algorithm for multi-image projective structure and motion [A]. ECCV, Lecture Notes in Computer Science [C]. Cambridge: Springer-Verlag, 1996, 1064: 709 - 720.
- [13] 杨明, 刘先忠. 矩阵论 [M]. 武汉: 华中科技大学出版社, 2003.
- [14] G Sparr. Simultaneous reconstruction of scene structure and camera locations from uncalibrated image sequences [A]. ICPR [C]. Vienna: IEEE Press, 1996, 328 - 333.

作者简介:



刘侍刚 男, 1973 年生于江西省峡江县, 于 1997 年和 2001 年在哈尔滨工程大学分别获得学士学位和硕士学位, 现为西安电子科技大学综合业务网国家重点实验室博士研究生. 主要研究方向为计算机视觉、图像处理、三维重建、虚拟现实等.

E-mail: xdlsg@hotmail.com



吴成柯 男, 1938 年生于安徽黄山, 教授, 博士生导师. 主要研究方向为计算机视觉、三维重建、图形图像处理、视频编码和图像通信等. 出版科技论著 4 本, 在国内外杂志上发表论文 100 余篇.

李良福 男, 1977 年生于江西省丰城市, 于 2005 年获得西安交通大学博士学位, 现为陕西师范大学计算机科学学院教师, 主要研究方向为计算机视觉、模式识别、机器人控制等.

宁纪锋 男, 1975 年生于陕西省渭南市, 现为西安电子科技大学综合业务网国家重点实验室博士研究生. 主要研究方向为计算机视觉、三维重建等.