

DOI: 10.7652/xjtuxb201501017

一种线性迭代非刚体射影重建方法

彭亚丽^{1,2}, 刘侍刚^{1,2}, 裘国永^{2,3}

(1. 陕西师范大学现代教学技术教育部重点实验室, 710062, 西安; 2. 西安交通大学电子与信息工程学院, 710049, 西安; 3. 陕西师范大学计算机科学学院, 710119, 西安)

摘要: 为了从非标定图像序列中重建出三维非刚体的射影结构,提出了一种线性迭代非刚体射影重建方法。该方法将所有图像点放入一个图像矩阵中,利用图像矩阵具有低秩的特性,通过因式分解构造一个线性迭代算法来求解图像点的深度因子,最终通过奇异值分解实现非刚体的射影重建。该方法的优点是:射影重建过程都是线性求解,且将所有图像及图像点都平等地对待,并没有倚重某些图像及图像点。模拟实验结果表明,所提重建方法比经典的 Alessio 方法重投影误差小 18%,比 Brand 方法小 27%;最后的真实实验结果表明,所提的重建方法重投影误差只有 1.28 个像素,从而验证了所提方法的有效性。

关键词: 射影重建;非刚体;因式分解

中图分类号: TP391.41;P232 **文献标志码:** A **文章编号:** 0253-987X(2015)01-0102-05

A Linearly Iterative Method for Non-Rigid Projective Reconstruction

PENG Yali^{1,2}, LIU Shigang^{1,2}, QIU Guoyong^{2,3}

(1. Key Laboratory of Modern Teaching Technology, Ministry of Education, Shaanxi Normal University, Xi'an 710062, China;
2. School of Electronics and Information Engineering, Xi'an Jiaotong University, Xi'an 710049, China;
3. School of Computer Science, Shaanxi Normal University, Xi'an 710119, China)

Abstract: A linearly iterative method for reconstruction of 3D non-rigid projection is presented to reconstruct the 3D non-rigid projective structure from an un-calibrated image sequence. All the image points are placed into an image matrix, and the linearly iterative method that employs the singular value decomposition is used to get the depth factors of the image points by using the low rank property of the image matrix. Then the projective reconstruction is realized through the singular value decomposition. The innovations of the proposed method are that the projective reconstruction is linear, all the image points are treated fairly, and it does not rely on certain image or image points. Experiment results with simulation data show that the proposed method is 18% and 27% smaller in re-projective error than Alessio's method and Brand's method, respectively. A real experiment shows that the method has only 1.28 pixels error, and hence efficient.

Keywords: projective reconstruction; non-rigid; factorization

基于图像序列的三维重建一直是计算机视觉研究的热点问题之一,而射影重建又是三维重建过程中一个必不可少的阶段,其精度影响最终的重建结果^[1-2]。如果没有任何先验知识,从图像序列中仅能

收稿日期: 2014-06-02。 作者简介: 彭亚丽(1979—),女,讲师;刘侍刚(通信作者),男,副教授。 基金项目: 国家自然科学基金资助项目(61402274);陕西师范大学中央高校基本科研业务费资助项目(GK201402040,GK201302029);陕西师范大学学习科学交叉学科培育计划资助项目。

实现射影重建^[3]。

由于刚体运动的简单性,三维重建的早期研究工作大都基于物体是刚体的假设^[4-5]。然而,现实世界中大多数物体都具有柔性,它们的运动都属于非刚体运动。为了重建非刚体物体,Bregler等人首先提出了非刚体可以由若干个刚体基组成的假设^[6],并重建了非刚体的运动,但是该计算过程非常繁琐,导致其应用受到很大限制。后来的许多非刚体重建方法都是基于Bregler的假设^[7-8],其中Valmadre采用动态规划的方法进行非刚体重建^[9],但是该方法鲁棒性较差。为了提高重建算法的鲁棒性,Tao等学者混合先验知识对非刚体进行了重建^[10-11],但在许多情况下很难获得非刚体的先验知识。有部分学者采用了光流法对非刚体进行重建^[12-14],并获得了较好的重建效果,但这些非刚体重建方法都是假设相机为正投影模型,但正投影模型要求物体的景深远小于相机到物体的距离,当条件不满足时,误差比较大。针孔模型相对于正投影模型更加符合真实情况,因此重建精度更高。

本文假定相机为针孔模型,提出了一种线性迭代非刚体射影重建方法,该方法利用图像矩阵具有低秩的特性,通过因式分解构造一个线性迭代算法来求解图像点的深度因子,最后实现非刚体的射影重建。

1 相机模型

假定相机为针孔模型,成像过程可表示为

$$\lambda \mathbf{m} = \mathbf{K}[\mathbf{R} \quad \mathbf{t}] \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix} \quad (1)$$

式中: $\mathbf{X} = [x \ y \ z]^T$ 表示三维空间点的非齐次坐标; $\mathbf{m} = [u \ v \ 1]^T$ 为对应的图像平面点的齐次坐标; $\mathbf{K}$ 为相机的内参矩阵; $\mathbf{R}$ 、 $\mathbf{t}$ 分别为相机在拍摄位置对应的旋转矩阵和平移向量,即相机的外参矩阵; λ 为深度因子。

若令

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

式中: $\mathbf{P}$ 称为投影矩阵。则式(1)可以简写为

$$\lambda \mathbf{m} = \mathbf{P} \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix} \quad (3)$$

假设有 m 幅图像,每幅图像上有 n 个图像点,对于第 i 幅图像上的第 j 个图像点可表示为

$$\lambda_{i,j} \mathbf{m}_{i,j} = \mathbf{P}_i \begin{bmatrix} \mathbf{X}_{i,j} \\ 1 \end{bmatrix} \quad (4)$$

2 非刚体射影重建

对于第 i 幅图像,由式(4)得

$$[\lambda_{i,1} \mathbf{m}_{i,1} \quad \cdots \quad \lambda_{i,n} \mathbf{m}_{i,n}] = \mathbf{P}_i \begin{bmatrix} \mathbf{X}_{i,1} & \cdots & \mathbf{X}_{i,n} \\ 1 & \cdots & 1 \end{bmatrix} \quad (5)$$

当物体做非刚体运动时, $[\mathbf{X}_{i,1} \ \mathbf{X}_{i,2} \ \cdots \ \mathbf{X}_{i,n}]$ 在运动过程中是变化的,可以认为由 l 个刚性基组成^[6],即

$$[\mathbf{X}_{i,1} \ \mathbf{X}_{i,2} \ \cdots \ \mathbf{X}_{i,n}] = \sum_{k=1}^l \beta_{i,k} \mathbf{B}_k \quad (6)$$

式中: $\beta_{i,k}$ 为权值; $\mathbf{B}_k$ 为刚性基。

将式(6)代入式(5),可得

$$[\lambda_{i,1} \mathbf{m}_{i,1} \quad \cdots \quad \lambda_{i,n} \mathbf{m}_{i,n}] = \mathbf{P}_i \begin{bmatrix} \sum_{k=1}^l \beta_{i,k} \mathbf{B}_k \\ \mathbf{1} \end{bmatrix} \quad (7)$$

令

$$\mathbf{P}_i = [\mathbf{Q}_i \quad \mathbf{p}_i] \quad (8)$$

式中: $\mathbf{Q}_i$ 是 $\mathbf{P}_i$ 的前3列; $\mathbf{p}_i$ 是 $\mathbf{P}_i$ 的第4列。

将式(8)代入式(7),可得

$$\underbrace{[\lambda_{i,1} \mathbf{m}_{i,1} \quad \cdots \quad \lambda_{i,n} \mathbf{m}_{i,n}]}_{\mathbf{w}_i} = [\mathbf{Q}_i \quad \mathbf{p}_i] \begin{bmatrix} \sum_{k=1}^l \beta_{i,k} \mathbf{B}_k \\ \mathbf{1} \end{bmatrix} \quad (9)$$

整理可得

$$\mathbf{W}_i = \underbrace{[\beta_{i,1} \mathbf{Q}_i \quad \cdots \quad \beta_{i,l} \mathbf{Q}_i \quad \mathbf{p}_i]}_{\mathbf{M}_i} \begin{bmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \\ \vdots \\ \mathbf{B}_l \\ \mathbf{1} \end{bmatrix} \quad (10)$$

若将所有的图像点放在一起,则有

$$\mathbf{W} = \mathbf{M} \mathbf{B} \quad (11)$$

$$\text{式中: } \mathbf{W} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_m \end{bmatrix}, \mathbf{M} = \begin{bmatrix} \mathbf{M}_1 \\ \mathbf{M}_2 \\ \vdots \\ \mathbf{M}_m \end{bmatrix}。$$

从式(11)可以看出, $\mathbf{W}$ 矩阵为低秩矩阵,其秩为 $3l+1$ 。同时,对于任何非奇异矩阵 $\mathbf{A}$ 都有

$$\mathbf{W} = \mathbf{M} \mathbf{B} = \mathbf{M}' \mathbf{B}' \quad (12)$$

式中: $\mathbf{M}' = \mathbf{M} \mathbf{A}^{-1}$; $\mathbf{B}' = \mathbf{A} \mathbf{B}$ 。

从式(12)可以看出,如果没有任何先验知识,所有的重建可以相差一个变换 $\mathbf{A}$,因此该重建不是落

在欧氏空间,而是落在射影空间。

3 迭代射影重建算法

如果已知深度因子 $\lambda_{i,j}$,可以对 $\mathbf{W}$ 进行奇异值分解,即

$$\mathbf{W} = \mathbf{S}\mathbf{V}\mathbf{D}^T \quad (13)$$

式中: $\mathbf{S}$ 、 $\mathbf{D}$ 为正交矩阵; $\mathbf{V}$ 为对角矩阵。

由于图像中含有噪声,因此不可能保证 $\mathbf{V}$ 中对角线上刚好只有前 $3l+1$ 个元素为非 0,但可以令对角线元素以外的其他元素为 0。

式(13)可化简为

$$\mathbf{W} = \mathbf{S}'\mathbf{V}'\mathbf{D}'^T \quad (14)$$

式中: $\mathbf{S}'$ 和 $\mathbf{D}'$ 分别为 $\mathbf{S}$ 和 $\mathbf{D}$ 的前 $3l+1$ 列; $\mathbf{V}' = \text{diag}(v_1 \ v_2 \ \cdots \ v_{3l+1})$ 。

因此,令射影重建为

$$\left. \begin{aligned} \mathbf{M} &= \mathbf{S}'\mathbf{V}'^{1/2} \\ \mathbf{B} &= \mathbf{V}'^{1/2}\mathbf{D}'^T \end{aligned} \right\} \quad (15)$$

式(14)表明, $\mathbf{W}$ 的列向量在矩阵 $\mathbf{S}'$ 所生成的子空间的正交补空间的投影为 $\mathbf{0}^{[15]}$,即

$$\mathbf{T} \begin{bmatrix} \lambda_{1,j}\mathbf{m}_{1,j} \\ \lambda_{2,j}\mathbf{m}_{2,j} \\ \vdots \\ \lambda_{m,j}\mathbf{m}_{m,j} \end{bmatrix} = \mathbf{0} \quad (16)$$

式中: $\mathbf{T} = \mathbf{I} - \mathbf{S}'(\mathbf{S}'^T\mathbf{S}')^{-1}\mathbf{S}'^T$ 。

令 $\mathbf{T} = [\mathbf{t}_1 \ \mathbf{t}_2 \ \cdots \ \mathbf{t}_{3m}]$,将式(16)展开得

$$\lambda_{1,j}u_{1,j}\mathbf{t}_1 + \lambda_{1,j}v_{1,j}\mathbf{t}_2 + \cdots + \lambda_{m,j}u_{m,j}\mathbf{t}_{3m-1} + \lambda_{m,j}\mathbf{t}_{3m} = \mathbf{0} \quad (17)$$

整理得

$$\begin{bmatrix} (u_{1,j}\mathbf{t}_1 + v_{1,j}\mathbf{t}_2 + \mathbf{t}_3) & \cdots & (u_{m,j}\mathbf{t}_{3m-2} + v_{m,j}\mathbf{t}_{3m-1} + \mathbf{t}_{3m}) \end{bmatrix} \begin{bmatrix} \lambda_{1,j} \\ \lambda_{2,j} \\ \vdots \\ \lambda_{m,j} \end{bmatrix} = \mathbf{0} \quad (18)$$

式(18)为线性方程,因此 $\lambda_{i,j}$ 可以线性求解。

最初并不知道 $\lambda_{i,j}$ 的值,因此无法通过式(13)完成因式分解,但通过分析知,可以构造一个线性迭代算法来求解 $\lambda_{i,j}$ 。即初始假设所有的 $\lambda_{i,j} = 1$,通过式(14)可以求得 $\mathbf{S}'$,通过式(18)可以求得 $\lambda_{i,j}$,此时的深度因子比上一步的深度因子更加接近真实值,经多次迭代后可以求得 $\lambda_{i,j}$ 。

4 算法总结及运算量分析

本文算法的步骤如下。

步骤 1:令 ϵ_1 为无穷小的正数, $k=1$;

步骤 2:令 $\lambda_{i,j}^{(k)} = 1, i=1,2,\cdots,m, j=1,2,\cdots,n$;

步骤 3:利用式(13),对 $\mathbf{W}$ 奇异值分解,并令 $\mathbf{S}'$ 为 $\mathbf{S}$ 的前 $3l+1$ 列;

步骤 4:利用式(16)构造投影矩阵 $\mathbf{T}$;

步骤 5:利用式(18)线性求取深度因子 $\lambda_{i,j}$;

步骤 6:若 $|v_{3l+2}^{(k)} - v_{3l+2}^{(k-1)}| > \epsilon_1$,则 $k=k+1$,转步骤 3,否则转步骤 7;

步骤 7:利用式(15),实现射影重建。

本文运算量主要来自式(13)对矩阵 $\mathbf{W}$ 奇异值分解及式(18)求取深度因子。对矩阵 $\mathbf{W}$ 奇异值分解的计算量^[15]为 $3mn^2$,式(18)求深度因子的运算量为 $3m^3$,总共有 n 组深度因子,因此求深度因子的总运算量为 $3m^3n$ 。从上面分析可以看出,本文运算量主要来自深度因子的求解,为 $3m^3n$ 。

5 实验

5.1 仿真实验

为了验证本文方法的收敛性能,本文在一个单位球随机产生 200 个空间点,并将这些空间点分别分成 3 个和 4 个非刚体的刚性基。当分成 3 个刚性基时,第 1 个刚性基由前 80 个空间点组成,第 2 个和第 3 个刚性基各由 60 个空间点组成;当分成 4 个刚性基时,第 1 个刚性基由前 80 个空间点组成,后面 3 个刚性基各由 40 个空间点组成。模拟产生相机的内参数,变化相机外参以产生 150 幅图像,并在图像中分别加入 0、0.5、1、1.5 像素的高斯噪声。用本文的方法进行非刚体射影重建,并用 v_{3l+2} 的值来衡量算法的收敛性能,在理想情况下, v_{3l+2} 的值应该为 0,算法收敛性能如图 1、图 2 所示。

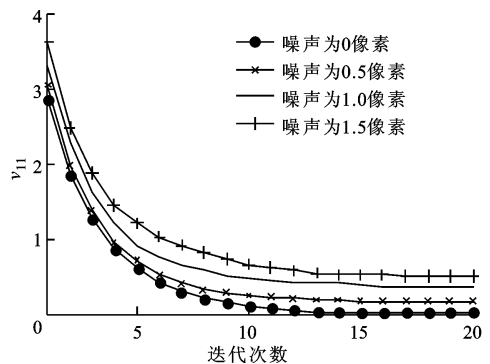


图 1 v_{11} 随迭代次数变化图

从图 1、图 2 可以看出,本文迭代方法具有较好的收敛性能,同时,噪声越大收敛性能越差。其原因

是由于算法收敛速度取决于比值 $r = \frac{v_{3l+1}}{v_{3l+2}}$, r 越大收

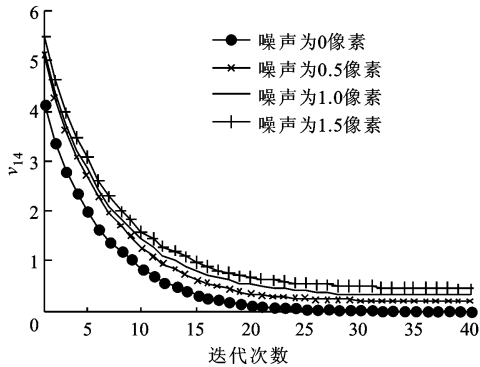


图 2 v_{l+1} 随迭代次数变化图

敛越快。在有噪声的情况下, v_{3l+2} 并不为 0, 噪声越小, v_{3l+2} 越小, 因此 r 越大收敛速度越快。在理想情况下, $v_{3l+2} = 0$, 此时 r 最大, 收敛速度最快。从图 1、图 2 可以看出, 图 1 的收敛速度比图 2 的收敛速度要快。其原因是在理想情况下图 1 对应的秩是 10, 而图 2 对应的秩是 13, 因为秩越大收敛速度越慢。

最后, 为了比较本文方法和 Alessio 方法^[14]、Brand 方法^[16]的重建精度, 用上述方法产生 150 幅图像, 并在图像中加入 0~2 像素的高斯噪声, 分别用本文方法、Alessio 方法和 Brand 方法进行重建, 在每个图像噪声下各运行 200 次, 计算重投影误差, 然后求平均值, 模拟实验结果如图 3 所示。

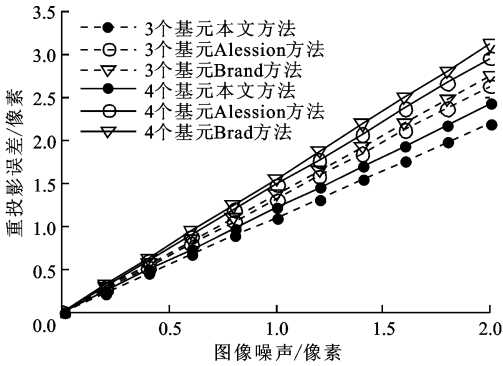


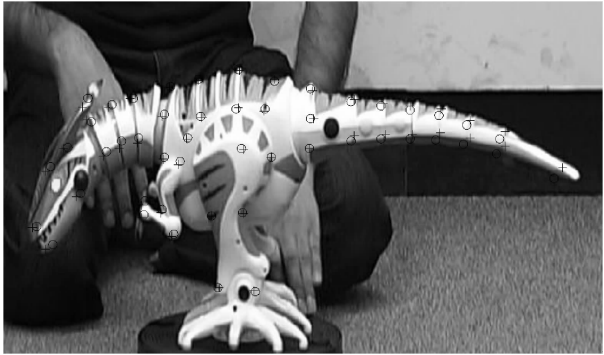
图 3 本文方法、Alessio 方法和 Brand 方法比较

从图 3 可以看出, 在 3 个、4 个刚性基的情况下, 本文方法的重投影误差比 Alessio 方法和 Brand 方法要小, 约比 Alessio 方法小 18%, 比 Brand 方法小 27%, 说明本文方法比 Alessio 方法和 Brand 方法重建精度都要高。原因是由于 Alessio 方法和 Brand 方法假设相机为正投影模型, 当物体的景深远小于相机到物体的距离时, 正投影模型能够较好地表示物体的成像过程, 当该条件不满足时, 重建误差比较大, 而本文假设相机为针孔模型, 该模型相对正投影模型更加符合实际情况, 因此本文方法的重

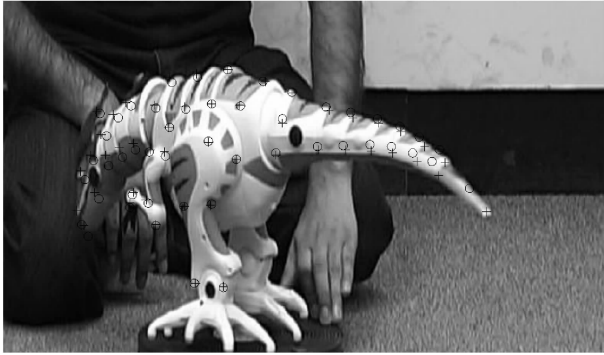
建精度要比 Alessio 方法的精度高, 但非刚性基越多, 两种方法的重投影误差都会增加。

5.2 真实实验

为了验证本文方法的正确性, 本文获得一个 200 帧的恐龙图像序列, 其中的 2 帧如图 4 所示。在该图像序列中, 提取并跟踪了 49 个特征点(如图中+所示)。从图像序列中可以看出, 由于恐龙的运动比较复杂, 故它的运动不能当作刚体运动, 只能当作非刚体运动。利用这些特征点, 用本文方法进行非刚体射影重建, 而且为了衡量本文方法的重建精度, 本文对这些重建点进行重投影(如图中○所示)。



(a) 第 75 帧



(b) 第 125 帧

图 4 恐龙图像序列

从图 4 可以看出, 提取的特征点和重投影点基本重合, 说明本文方法具有较高的重建精度。同时, 计算出平均重投影误差为 1.28 像素, 具有较小的重投影误差。

6 结 论

本文提出了一种线性迭代非刚体射影重建方法, 该方法利用图像矩阵具有低秩的特性, 通过构造一个线性迭代算法求解图像点的深度因子, 最终实现非刚体的射影重建。该方法的优点是: 射影重建过程都是线性求解, 且将所有图像及图像点都平等地对待。实验结果表明, 该方法具有良好的收敛性,

且具有较高的重建精度。本文方法非常适合于图像点位置为高斯噪声的情况,而且现实中提取的图像点位置也非常符合高斯噪声分布,但对于图像点位置为其他噪声的情况(如出格数据等),本文方法性能急骤下降,这是目前三维重建的一个研究热点问题,也是本文下一步的研究重点。

参考文献:

- [1] SHEN S. Accurate multiple view 3D reconstruction using patch-based stereo for large-scale scenes [J]. IEEE Transactions on Image Processing, 2013, 22(5): 1901-1914.
- [2] LEE M, CHOI C. Real-time facial shape recovery from a single image under general, unknown lighting by rank relaxation [J]. Computer Vision and Image Understanding, 2014, 120(3): 59-69.
- [3] WU F, ZHANG M, HU Z. Self-calibration under the Cayley framework [J]. International Journal of Computer Vision, 2013, 103(3): 372-398.
- [4] BASTIAN G, MATHIEU A, KALIN K, et al. A super-resolution framework for high-accuracy multiview reconstruction [J]. International Journal of Computer Vision, 2014, 106(2): 172-191.
- [5] PENG Y, LIU S, LIU F. Projective reconstruction with occlusions [J]. Opto-Electronics Review, 2010, 18(2): 150-154.
- [6] BREGLER C, HERTZMANN A, BIERMANN H. Recovering non-rigid 3D shape from image streams [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2000: 690-696.
- [7] NOGUER F, FUA P. Stochastic exploration of ambiguities for nonrigid shape recovery [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2013, 35(2): 463-475.
- [8] LIU S, PENG Y, ZENG Z, et al. An iterative method based on 1D subspace for projective reconstruction [J]. Opto-Electronics Review, 2011, 19(1): 89-94.
- [9] VALMADRE J, ZHU Y, SRIDHARAN S, et al. Efficient articulated trajectory reconstruction using dynamic programming and filters [C] // 12th European Conference on Computer Vision. Berlin, Germany: Springer, 2012: 72-85.
- [10] TAO L, MEIN S, QUAN W, et al. Recursive non-rigid structure from motion with online learned shape prior [J]. Computer Vision and Image Understanding, 2013, 117(10): 1287-1298.
- [11] VALMADRE J, LUCEY S. General trajectory prior for non-rigid reconstruction [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2012: 1394-1401.
- [12] AKHTER I, SHEIKH Y, KHAN S, et al. Trajectory space: a dual representation for nonrigid structure from motion [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(7): 1442-1456.
- [13] GARG R, ROUSSOS A, AGAPITO L. A variational approach to video registration with subspace constraints [J]. International Journal of Computer Vision, 2013, 104(3): 286-314.
- [14] ALESSIO D. Adaptive non-rigid registration and structure from motion from image trajectories [J]. International Journal of Computer Vision, 2013, 103(2): 226-239.
- [15] 王永茂, 刘德友. 矩阵分析基础 [M]. 北京: 清华大学出版社, 2012.
- [16] BRAND M. A direct method for 3D factorization of non-rigid motion observed in 2D [C] // IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2005: 122-128.

[本刊相关文献链接]

- 陈德祥, 徐自力. 多重网格在黏性流动最小二乘等几何模拟中的应用. 2014, 48(11): 122-127. [doi:10.7652/xjtuxb201411021]
- 杜辉, 王宇平, 董晓盼. 采用万有引力定律自动确定类数的 K 均值算法. 2014, 48(10): 115-119. [doi:10.7652/xjtuxb201410018]
- 蒲伟, 王家序, 周广武, 等. 卷吸速度方向与椭圆短轴成一夹角的弹流润滑渐近网格加密算法. 2014, 48(9): 95-100. [doi:10.7652/xjtuxb201409016]
- 郭涛, 李国君. 内嵌晃荡液体减振的流固耦合分析. 2014, 48(9): 117-122. [doi:10.7652/xjtuxb201409020]
- 任茂栋, 梁晋, 唐正宗, 等. 数字图像相关法中的优化插值滤波器. 2014, 48(7): 65-70. [doi:10.7652/xjtuxb201407012]
- 刘骁, 王小鑫, 胡红利, 等. 改进型离线迭代在线重构算法的电容层析成像技术研究. 2014, 48(4): 35-40. [doi:10.7652/xjtuxb201404007]
- 张思文, 吴九汇, 刘彰宜. 黏弹阻尼对一维杆状声子晶体能带结构频移的影响. 2014, 48(3): 22-27. [doi:10.7652/xjtuxb201403005]
- 吴仁斌, 姚敏立, 贾维敏, 等. 采用幅度响应约束的鲁棒自适应波束形成算法. 2014, 48(4): 109-114. [doi:10.7652/xjtuxb201404019]

(编辑 赵炜)