

迭代二次规划遮挡点恢复

彭亚丽^{1,2,3}, 刘侍刚^{1,2}, 孙增国^{2,3}, 洪灵^{1,3}, 曹菡^{1,2,3}

(1. 现代教学技术教育部重点实验室, 陕西西安 710062; 2. 陕西省教学信息技术工程实验室, 陕西西安 710119;
3. 陕西师范大学计算机科学学院, 陕西西安 710119)

摘要: 为了有效地恢复遮挡点, 本文提出一种迭代二次规划遮挡点恢复方法, 该方法首先分别利用图像矩阵的行向量和列向量在图像矩阵生成的正交补空间上的投影为 0 的特性, 构造行和列余差函数, 同时, 对遮挡点分别按行为主序和列为主序进行排列, 利用排列后这两者之间存在一个变换关系, 将行和列余差函数统一表示为一个二次优化目标函数. 该方法同时考虑了遮挡点在行和列两个方向的约束, 而且将遮挡点求解转化为迭代求解一个二次规划问题. 实验结果表明, 本文方法具有收敛速度快, 恢复精度高等优点.

关键词: 遮挡点恢复; 二次规划; 投影

中图分类号: TP391.41; P232

文献标识码: A

文章编号: 0372-2112 (2018)11-2733-05

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2018.11.021

Iterative Quadratic Programming Method for Occlusion Recovery

PENG Ya-li^{1,2,3}, LIU Shi-gang^{1,2}, SUN Zeng-guo^{2,3}, HONG Ling^{1,3}, CAO Han^{1,2,3}

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

Abstract: In order to effectively recover the occlusion, this paper presents an iterative quadratic programming method for occlusion recovery. Based on the characteristic that the projections of the row vector and the column vector of image matrix to the orthogonal complementary subspace spanned by image matrix are zero vectors, the row and the column residual objective functions are respectively defined. At the same time, the occlusion positions are respectively sorted according to the row and the column order, which can be denoted by a transformation matrix. Based on the transformation matrix, a united residual objective function which is quadratic is obtained from the row and the column ones. The method has the advantages that both the row and the column constraints are simultaneously considered and the solution of occlusion is transformed to iterative solution a quadratic programming. The experimental results show that the method has fast convergence speed and high precision.

Key words: occlusion recovery; quadratic programming; projection

1 引言

特征点提取和跟踪是三维重建、目标跟踪与识别、人体行为分析等计算机视觉研究的前提和基础^[1-3]. 由于各种原因, 在特征点的跟踪过程中, 必然会导致部分特征点跟踪丢失, 即遮挡点. 如何恢复这些遮挡点, 就成为当前计算机视觉的研究热点问题之一^[4-6].

为了恢复遮挡点, 部分学者采用子矩阵法对遮挡

点进行恢复^[7,8]. 但子矩阵法每次只能恢复一个遮挡点. 为了克服该方法的缺点, 有些学者利用图像矩阵等于投影矩阵和空间点矩阵之积的特性, 采用 Resection-intersection 技术(即保持空间点不变, 求解投影矩阵, 然后保持投影矩阵不变, 求解空间点, 如此循环迭代求解)进行遮挡点恢复. 在该类方法中, 最典型的就是 Wiberg 方法^[9], 目前它已经成为遮挡点恢复的标准算法之一^[10-12]. 虽然有部分学者在 Wiberg 方法的基础上

收稿日期: 2017-08-30; 修回日期: 2018-07-18; 责任编辑: 马兰英

基金项目: 国家自然科学基金(No. 61672333, No. 61873155, No. 61701290); 陕西省工业科技攻关项目(No. 2016GY-081); 陕西省自然科学基金基础研究计划(No. 2018JM6050); 陕西省重点科技创新团队(No. 2014KTS-18); 现代教学技术教育部重点实验室学习科学交叉学科培育计划资助; 陕西师范大学中央高校基本科研业务费项目(No. GK201803088, No. GK201803059)

进行了改进,但该类方法仅在遮挡率低时比较有效.同时,部分学者利用重投影点和已知点构造目标函数^[13,14],采用牛顿迭代法进行求解,但该目标函数是一个非凸函数.刘等人将图像两两组合,提出了一种线性遮挡点恢复方法^[15],但若两幅变化非常小的图像组合在一起,很容易导致算法失效.

为了解决这个问题,一些学者利用图像矩阵和空间结构点的行向量生成同一子空间的特性^[16,17],对遮挡点进行迭代求解,该方法没有考虑列方向的约束,因此,恢复精度比较低.

为了克服上述缺点,本文提出一种迭代二次规划遮挡点恢复方法,该方法分别利用图像矩阵的行向量和列向量在图像矩阵生成的正交补空间上的投影为0的特性,构造行和列余差函数,同时对遮挡点分别按行为主序和列为主序进行排列,利用排列后这两者之间存在一个变换关系,将行和列余差函数统一为一个二次优化目标函数.该方法由于同时考虑了遮挡点在行和列两个方向的约束,因此具有更高的恢复精度.同时,该方法将遮挡点求解转化为迭代求解一个二次规划问题,而二次规划问题的求解具有良好的数值特性,如可以一次线性求出全局最优解,因此本文具有收敛速度快,而且能够收敛到全局最优.

2 相机模型

在相机为正投影模型下,假设有 N 个三维空间点, Q 幅图像,则有^[17]:

$$\mathbf{W}_{2Q \times N} = \mathbf{S}_{2Q \times 3} \mathbf{X}_{3 \times N} \quad (1)$$

式中 $\mathbf{W}_{2Q \times N}$ 为图像矩阵, $\mathbf{S}_{2Q \times 3}$ 为相机投影矩阵, $\mathbf{X}_{3 \times N}$ 为三维空间点矩阵.

为了叙述方便,令 $M = 2Q$,则上式可以表示为:

$$\mathbf{W}_{M \times N} = \mathbf{S}_{M \times 3} \mathbf{X}_{3 \times N} \quad (2)$$

从式(2)可以看出,图像矩阵 $\mathbf{W}_{M \times N}$ 的秩为3,对 $\mathbf{W}_{M \times N}$ 进行奇异值分解,则有:

$$\mathbf{W}_{M \times N} = \mathbf{U}_{M \times 3} \mathbf{A}_{3 \times 3} \mathbf{V}_{3 \times N} \quad (3)$$

从上式可以看出, $\mathbf{W}_{M \times N}$ 中的列向量和 $\mathbf{U}_{M \times 3}$ 的列向量生成同一线性子空间, $\mathbf{W}_{M \times N}$ 中的行向量和 $\mathbf{V}_{3 \times N}$ 的行向量生成同一线性子空间.

3 迭代二次型规划遮挡点恢复方法

从式(3)可以看出, $\mathbf{W}_{M \times N}$ 和 $\mathbf{U}_{M \times 3}$ 的列向量生成同一线性子空间,根据投影定理可知,任一行向量投影到 $\mathbf{U}_{M \times 3}$ 列向量生成的正交补空间上的投影矩阵为:

$$\mathbf{T}_{M \times M}^{\perp} = \mathbf{E}_{M \times M} - \mathbf{U}_{M \times 3} \mathbf{U}_{M \times 3}^T \quad (4)$$

式中 $\mathbf{E}_{M \times M}$ 为单位阵.

任意行向量投影到 $\mathbf{W}_{M \times N}$ 行向量生成的正交补空间上的投影矩阵为:

$$\mathbf{P}_{N \times N}^{\perp} = \mathbf{E}_{N \times N} - \mathbf{V}_{3 \times N}^T \mathbf{V}_{3 \times N} \quad (5)$$

3.1 列方向约束

$\mathbf{W}_{M \times N}$ 中的任一行向量 $\boldsymbol{\gamma}_n$ 投影至 $\mathbf{W}_{M \times N}$ 的正交补子空间为0,可以转化为求余差 r_n^c 的极小值,即:

$$r_n^c = \boldsymbol{\gamma}_n^T \mathbf{T}_{M \times M}^{\perp} \boldsymbol{\gamma}_n \quad (6)$$

将上式展开,则有:

$$r_n^c = \mathbf{X}_n^T \mathbf{B}_n \mathbf{X}_n + 2\mathbf{b}_n^T \mathbf{X}_n + b_n \quad (7)$$

式中 \mathbf{X}_n 为 $\boldsymbol{\gamma}_n$ 中的未知向量, \mathbf{B}_n 为的 $\mathbf{T}_{M \times M}^{\perp}$ 子矩阵; \mathbf{b}_n 为 $\mathbf{T}_{M \times M}^{\perp}$ 行中的元素和 $\boldsymbol{\gamma}_n$ 中已知元素相乘之和, b_n 为常数项.

在上述推导中,我们仅针对 $\mathbf{W}_{M \times N}$ 矩阵中的第 n 列 $\boldsymbol{\gamma}_n$,若对于所有的列,则总余差为:

$$r^c = \sum_{n=1}^N r_n^c \quad (8)$$

将式(7)代入式(8),则有:

$$r^c = \mathbf{X}^T \mathbf{B} \mathbf{x} + 2\mathbf{b}^T \mathbf{x} + b \quad (9)$$

$$\text{式中 } \mathbf{x} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix}, \mathbf{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_N \end{pmatrix}, \mathbf{B} = \begin{pmatrix} B_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & B_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & B_N \end{pmatrix},$$

$$b = \sum_{n=1}^N b_n.$$

3.2 行方向约束

同上节类似, $\mathbf{W}_{M \times N}$ 中的任一行向量 $\boldsymbol{\lambda}_m$ 投影到 $\mathbf{W}_{M \times N}$ 正交补子空间可以转化为求余差 r_m^r 的极小值,即:

$$r_m^r = \boldsymbol{\lambda}_m \mathbf{P}_{N \times N}^{\perp} \boldsymbol{\lambda}_m^T \quad (10)$$

同理,将上式展开,则有:

$$r_m^r = \mathbf{y}_m^T \mathbf{D}_m \mathbf{y}_m + 2\mathbf{d}_m^T \mathbf{y}_m + d_m \quad (11)$$

式中 \mathbf{y}_m 为未知向量, \mathbf{D}_m 为的 $\mathbf{P}_{N \times N}^{\perp}$ 子矩阵, \mathbf{d}_m 为 $\mathbf{P}_{N \times N}^{\perp}$ 行中的元素和 $\boldsymbol{\lambda}_m$ 中已知元素相乘之和; d_m 为常数项.

对于所有的行,则总余差为:

$$r^r = \sum_{m=1}^M r_m^r \quad (12)$$

将式(11)代入式(12),则有:

$$r^r = \mathbf{y}^T \mathbf{D} \mathbf{y} + 2\mathbf{d}^T \mathbf{y} + d \quad (13)$$

$$\text{式中 } \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_M \end{pmatrix}, \mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ \vdots \\ d_M \end{pmatrix}, \mathbf{D} = \begin{pmatrix} D_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & D_2 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & D_M \end{pmatrix},$$

$$d = \sum_{m=1}^M d_m.$$

3.3 联合行列方向约束

在上面的列方向和行方向约束中,分别考虑了遮挡点处于列或行上的约束,而实际上,遮挡点在行和列

上的约束是同时存在的,因此应该联合考虑.

向量 \mathbf{x} 和 \mathbf{y} 具有以下变换关系:

$$\mathbf{y} = \mathbf{M}\mathbf{x} \quad (14)$$

式中 \mathbf{M} 为变换矩阵.

将式(14)代入式(13),并联合考虑行和列上的余差,则有:

$$\mathbf{r} = \mathbf{r}^c + \mathbf{r}^r = \mathbf{X}^T (\mathbf{B} + \mathbf{M}^T \mathbf{D} \mathbf{M}) \mathbf{x} + 2 (\mathbf{b} + \mathbf{M}^T \mathbf{d})^T \mathbf{x} + (\mathbf{b} + \mathbf{d}) \quad (15)$$

上式中的最小值点位于:

$$\mathbf{x} = (\mathbf{B} + \mathbf{M}^T \mathbf{D} \mathbf{M})^{-1} (\mathbf{b} + \mathbf{M}^T \mathbf{d}) \quad (16)$$

在上述的推导过程中,假设投影矩阵 $\mathbf{T}_{M \times M}^{\perp}$ 和 $\mathbf{P}_{N \times N}^{\perp}$ 都已知,但 $\mathbf{T}_{M \times M}^{\perp}$ 和 $\mathbf{P}_{N \times N}^{\perp}$ 都是来自式(3)对 $\mathbf{W}_{M \times N}$ 矩阵的奇异值分解,而开始时, $\mathbf{W}_{M \times N}$ 矩阵中含有遮挡点,因此,无法实现奇异值分解.但开始时,可以假设遮挡点的初值,然后利用行列向量约束,求到遮挡点,此时求到的遮挡点比开始时给定的初值要更加接近真实值,因此可以对遮挡点进行迭代求解.

4 实验

4.1 仿真实验

为了验证本文方法的收敛性能,本文在 Win7 操作系统下,处理器为 Intel Core2 Duo CPU T6500 @ 2.10GHz,内存为 2GB 的笔记本电脑下,首先在单位球内产生 50 个三维空间点,并变化相机的参数以产生大小为 480×640 的 30 幅图像.同时,随机地把其中的 65% 图像点当作遮挡点,并在剩下的图像点中加入均值为 0,方差分别为 1,2,3,4 个像素的高斯噪声;然后,用同样的方法产生 30 幅图像,分别遮挡 25%、45%、65%、85% 的图像点,并在剩下的图像点中加入均值为 0,方差为 2 个像素的高斯噪声.在每种情况下,分别用本文方法求解遮挡点的值,迭代收敛图如图 1 和图 2 所示.

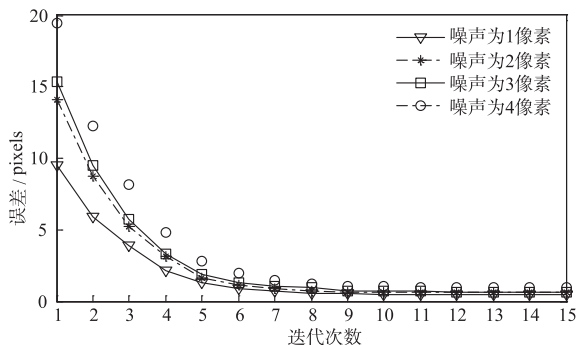


图1 在不同的噪声下,恢复误差随迭代次数变化图

从图 1 和图 2 可以看出,本文方法的收敛速度非常快,只需要 10 次以内迭代就能达到收敛,原因是由于本文将遮挡点的求解转化为迭代二次目标函数的极值求解,而二次目标函数的极值求解可以一步求解完成.同

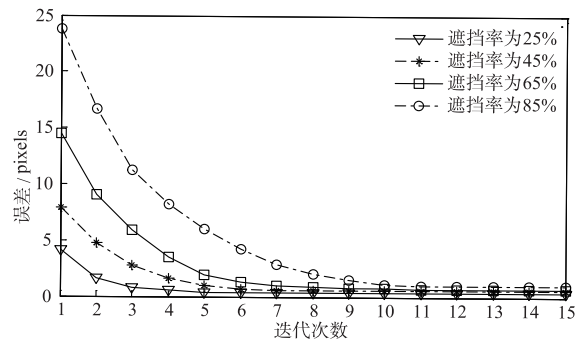


图2 在不同的遮挡率下,恢复误差随迭代次数变化图

时,从图 1 可以看出,图像噪声不会影响收敛速度,只会影响最终收敛值;从图 2 还可以看出,收敛速度和遮挡率成反比,原因是遮挡率越高,约束越少,未知数越多,因此算法的收敛速度就越慢.

为了比较研究遮挡率对 Wiberg 方法和本文方法收敛精度的影响,产生 30 幅图像,并随机遮挡 10% 至 80% 的图像点,并在剩下的图像点中加入均值为 0,方差分别为 1,2,3 的高斯噪声,然后分别用 Wiberg 方法和本文方法进行遮挡点求解.在每种情况下进行 50 次重复实验,求平均误差,实验结果如图 3 所示.

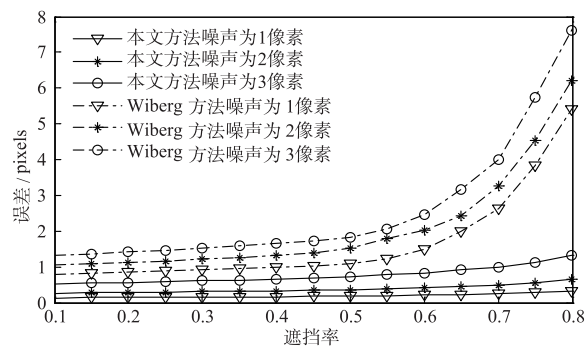


图3 恢复误差随遮挡率变化图

从图 3 可以看出:(1)收敛精度与遮挡率成反比,原因是由于遮挡率越高,可利用的点信息就越少,因此,约束就越少,恢复精度就越低;(2)本文方法的恢复精度要优于 Wiberg 方法,原因主要有两方面,一方面是 Wiberg 方法仅仅考虑了行方向上的约束,没有考虑列方向上的约束,而本文方法既考虑了行方向上的约束,又考虑了列方向上的约束,而且行列方向平等地对待;另一方面是 Wiberg 方法构造的目标函数是非凸的,因此无法保证收敛到全局最优解,而本文方法的目标函数是一个二次型函数,是一个凸函数,能够保证收敛到全局最优解.

最后,为了比较研究 Wiberg 方法和本文方法的抗噪性能,用上述方法产生 30 幅图像,并分别随机遮挡 45%、55%、65% 的图像点,然后在剩下的图像点中加入均值为 0,方差从 0 变化到 3 的噪声,分别用 Wiberg 方

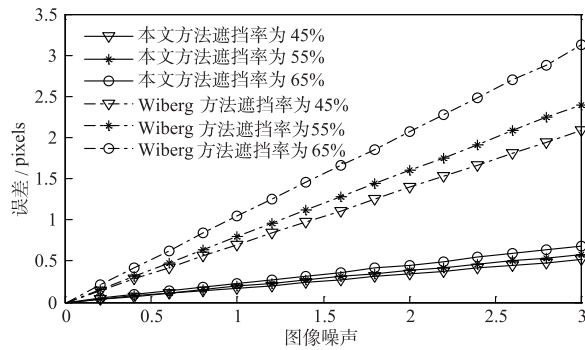


图4 恢复误差随图像噪声变化图

法和本文方法求解遮挡点. 同上述一样, 在每种情况下各重复 50 次, 然后求平均误差, 结果如图 4 所示.

从图 4 可以看出, 本文方法在恢复精度方面要优于 Wiberg 方法, 原因如上所述.

4.2 真实实验

本文用数码相机获得了一个由 22 幅大小为 640×480 组成的西安钟楼图像序列, 其中 2 幅图像如图 5 所示, 该图像序列绕西安钟楼一圈. 本文提取了 396 个特征点 (如图 5 中 * 所示), 由于该图像序列绕西安钟楼一圈, 因此遮挡率高于 50%. 用本文方法和 Wiberg 方法分别恢复遮挡点的位置, 恢复结果如图 5 中 □ 所示, 其中第一行为本文方法的恢复结果, 第二行为 Wiberg 方法的恢复结果. 图 5 中 □ 对应的点大部分为钟楼背面上对应的特征点, 钟楼底座是一个正方形, 从图中可以看出, 本文方法恢复的钟楼底座上的特征点在一条直线上, 而 Wiberg 方法恢复的误差明显大于本文方法的误差, 因此本文方法具有更高的恢复精度.

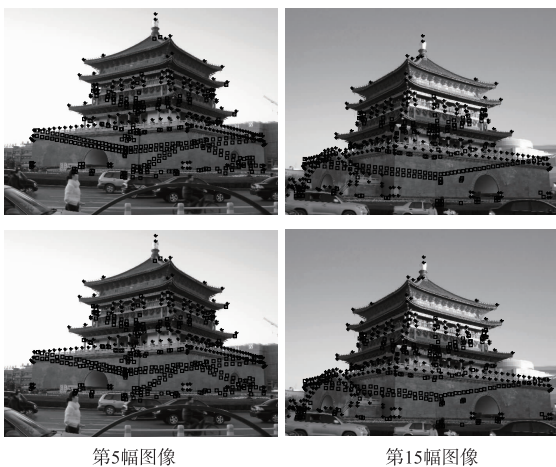


图5 本文方法和Wiberg方法恢复结果对比图 (*为提取的特征点, □为恢复的遮挡点)

5 结束语

本文在相机为正投影模型下, 提出了一种迭代二

次规划遮挡点恢复方法, 该方法首先分别利用图像矩阵的行向量和列向量在图像矩阵生成的正交补空间上的投影为 0 的特性, 构造行和列余差函数, 同时, 对遮挡点分别按行为主序和列为主序进行排列, 利用排列后这两者之间存在一个变换关系, 将行和列余差函数统一表示为一个二次优化目标函数. 该方法由于同时考虑了遮挡点在行和列两个方向的约束, 因此, 具有更高的恢复精度. 同时, 该方法将遮挡点求解转化为迭代求解一个二次规划问题, 而二次规划的求解具有良好的数值特性, 因此本文具有收敛速度快, 而且能够收敛到全局最优. 仿真实验和真实实验结果表明, 本文方法具有收敛速度快, 恢复精度高等优点.

参考文献

- [1] 彭亚丽, 等. 基于非刚体轨迹基的线性自标定方法[J]. 电子学报, 2017, 45(1): 135 - 139.
Peng Yali, et al. Linear camera self-calibration method based on non-rigid trajectory basis [J]. Acta Electronica Sinica, 2017, 45(1): 135 - 139. (in Chinese)
- [2] Peng Y L, et al. Space-frequency domain based joint dictionary learning and collaborative representation for face recognition[J]. Signal Processing, 2018, 147: 101 - 109.
- [3] Liu T, et al. On the performance of Manhattan nonnegative matrix factorization[J]. IEEE Transactions on Neural Networks and Learning Systems, 2016, 27(9): 1851 - 1863.
- [4] Chen G, et al. Multi-modal curriculum learning for semi-supervised image classification[J]. IEEE Transactions on Image Processing, 2016, 25(7): 3249 - 3260.
- [5] Peng Y L, et al. Negative ϵ dragging technique for pattern classification[J]. IEEE Access, 2018, 6(1): 488 - 494.
- [6] Liu T, et al. A large-cone nonnegative matrix factorization [J]. IEEE Transactions on Neural Networks and Learning Systems, 2017, 28(9): 2129 - 2142.
- [7] Jacobs D. Linear fitting with missing data for structure-from-motion[J]. Computer Vision and Image Understanding, 2008, 82(1): 57 - 81.
- [8] Chen P. Optimization algorithms on subspaces: revisiting missing data problem in low-rank matrix [J]. International Journal of Computer Vision, 2008, 80(1): 125 - 142.
- [9] Okatani T, et al. On the Wiberg algorithm for matrix factorization in the presence of missing components [J]. International Journal of Computer Vision, 2007, 72(3): 329 - 337
- [10] Zheng Y, et al. Generalizing Wiberg algorithm for rigid and nonrigid factorizations with missing components and metric constraints [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. Rhode Island, 2012. 2010 - 2017.
- [11] Strelow D, et al. Eriksson. General, nested, and constrained

- Wiberg minimization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(9): 1803 – 1815.
- [12] Jiang X, et al. Robust matrix completion via alternating projection[J]. IEEE Signal Processing Letters, 2017, 24(5): 579 – 583.
- [13] Kennedy R, et al. Online algorithms for factorization-based structure from motion[J]. Computer Vision and Image Understanding, 2016, 150: 139 – 152.
- [14] Agudo A, et al. Sequential non-rigid structure from motion using physical priors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(5): 979 – 994.
- [15] 刘侍刚, 等. 基于正交补空间的遮挡点恢复方法[J]. 电子学报, 2015, 43(5): 911 – 915.
- Liu Shigang, et al. An occlusion recovery method based on orthogonal complement subspace [J]. Acta Electronica Sinica, 2015, 43(5): 911 – 915. (in Chinese)
- [16] Kumar S, et al. Spatio-temporal union of subspaces for multi-body non-rigid structure-from-motion [J]. Pattern Recognition, 2017, 71(2): 428 – 443.
- [17] 刘侍刚, 等. 投影约束的遮挡点恢复方法[J]. 电子学报, 2017, 45(11): 2611 – 2616.
- Liu Shigang, et al. Occlusion recovery method based on projective constraints [J]. Acta Electronica Sinica, 2017, 45(11): 2611 – 2616. (in Chinese)

作者简介



彭亚丽 女, 1979年3月出生, 陕西石泉人. 2005年在哈尔滨工程大学获得硕士学位, 2013年在西安电子科技大学获得博士学位. 现为陕西师范大学副教授. 从事计算机视觉、模式识别等方面的有关研究.



刘侍刚 (通信作者) 男, 1973年11月出生, 江西峡江人, 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, 2005年在西安电子科技大学获得博士学位, 现为陕西师范大学教授. 从事计算机视觉、三维重建等方面的有关研究.

E-mail: shgliu@snnu.edu.cn

孙增国 男, 1980年10月出生, 陕西西安人, 2003年和2010年在西安交通大学分别获得学士和博士学位, 2011年至2015年在华侨大学从事教学科研工作, 现为陕西师范大学副教授. 从事遥感图像处理、计算机视觉、模式识别等方面的有关研究.

洪灵 女, 1986年3月出生, 浙江台州人, 2008年和2015年在西安电子科技大学分别获得学士学位和博士学位, 现为陕西师范大学副研究员. 从事机器学习、模式识别等方面的有关研究.

曹茜 女, 1963年出生, 四川荣县人, 1986年和1989年在西北大学计算机系分别获得学士学位和硕士学位, 2002年武汉大学测绘遥感信息工程国家重点实验室获得博士学位, 现为陕西师范大学教授. 从事图像处理、空间数据挖掘、智慧旅游等方面的有关研究.