

一种面向图像拼接的改进 PCA-SIFT 算法

杨炳坤,程树英,郑茜颖

(福州大学 物理与信息工程学院,福建 福州 350108)

摘 要: 针对图像拼接中尺度不变特征变换(SIFT)算法没有充分考虑特征点的分布情况且计算复杂、耗时较长等问题,提出了一种基于改进的 PCA-SIFT 算法.该算法首先在空间极值点检测阶段引入改进的非极大值抑制法对初始特征点进行优选,得到分布更加均匀的特征点集;然后在构建描述符阶段基于圆形领域提取 64 维 SIFT 描述符,并使用主成分分析(PCA)法对描述符进一步降维,减少描述符的数据复杂度;最后在特征匹配阶段引入基于 K-D 树的 BBF 搜索策略,采用随机抽样一致性(RANSAC)算法剔除误匹配点,从而提高了匹配速度与匹配精度.在 10 组图像拼接实验中,本文算法的拼接速度是传统 SIFT 算法的 1.6~2.2 倍.实验结果表明,本文算法具有较高的精度、较好的鲁棒性,较强的实时性.

关键词: 图像拼接;尺度不变特征变换;非极大值抑制;主成分分析;随机抽样一致性

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2018)12-0070-06

An Improved PCA-SIFT Algorithm for Image Mosaics

YANG Bing-kun, CHENG Shu-ying, ZHENG Qian-ying

(College of Physics and Information Engineering, Fuzhou University, Fuzhou 350108, China)

Abstract: Aiming at the problem that the SIFT algorithm does not fully consider the distribution of feature points in the image splicing and the calculation is complex and takes a long time, an improved PCA-SIFT algorithm is proposed. The algorithm firstly introduces an improved non-maximum suppression method to optimize the initial feature points so as to obtain a more even distribution of feature point sets. Then the 64-dimension SIFT descriptor is extracted based on the circular neighborhood, and the descriptor is further reduced using PCA to reduce the data complexity of the descriptor. Finally, the BBF search strategy based on K-D tree was introduced. The RANSAC was used to eliminate the false matching points, which improved the matching speed and matching accuracy. In the 10 sets of image stitching experiments, the stitching speed of this algorithm is 1.6~2.2 times that of the traditional SIFT algorithm. Experimental results show that the proposed algorithm has higher accuracy, better robustness, and stronger real-time performance.

Key words: image mosaic; seale invariant feature transform (SIFT); non-maximal inhibition; principal component analysis (PCA); random sample consensus (RANSAC)

1 引言

图像拼接是指将几幅有重叠的图像,拼接成一幅宽视野、无缝的高分辨率图像^[1].目前图像拼接技术广泛应用在医疗图像诊断、军事、遥感影像处理等领域^[2-4].图像拼接中最重要的一个步骤就是图像配准,图像配准算法的好坏直接决定图像拼接的质量

以及效率.Mikolajczyk^[5]对几种有代表性的图像配准算法就尺度缩放、图像压缩、视角变换等方面进行了性能评估,结果表明 SIFT^[6]是目前图像处理领域中最为有效的图片配准算法,具有较强的鲁棒性.但是该算法没有考虑特征点的分布情况,在细节复杂的区域会提取过多的特征点,这样容易产生误匹配,SIFT 算法计算复杂、时间开销大,难以满足实时性.

收稿日期: 2018-03-01; **修回日期:** 2018-03-25

基金项目: 国家自然科学基金(61471124);福建省科技厅工业引导性重点项目(2016H0016,2015H0021)

本文提出一种面向图像拼接的改进 PCA-SIFT 算法,以提高 SIFT 图像拼接的速度与精度. 该算法对特征点进行优选,得到数量更少,分布更加均匀的特征点集;同时对描述符进行改进,加快构建描述符的速度和匹配速度,降低描述符的维度,最后采用随机抽样一致性算法剔除误匹配点.

2 SIFT 算法

SIFT 特征是一种基于尺度不变空间理论的局部特征描述子,对图像的尺度缩放、旋转、亮度等变换具有较好的鲁棒性^[7],它利用高斯差分算子构建尺度空间,并在尺度空间中寻找极值点,然后基于图像的局部梯度方向,为每个关键点分配一个或多个方向角度,形成 SIFT 特征点,最后通过统计特征点领域的梯度信息与方向,生成 128 维 SIFT 描述符. 其主要步骤如下.

2.1 尺度空间极值检测

把待处理的图像转换为灰度图像,与不同尺度下的高斯核 $G(x, y, \sigma)$ 进行卷积操作如式(1),构建高斯金字塔. 对于获得的尺度空间函数 $L(x, y, \sigma)$ 利用 DoG(Difference of Gaussians)对两个相邻高斯尺度空间的图像进行相减,构建高斯金字塔. 为了寻找高斯函数的极值点,每个像素点要和它同尺度的 8 个邻域点以及上下相邻尺度各 9 个邻域点相比较取极值点.

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (1)$$

2.2 特征点的精确定位

检测到的极值点并不是所求的特征点,还需要对其进一步筛选,我们需要精确定位特征点的位置和尺度,使其达到亚像素精度. 使用泰勒级数展开式作为拟合函数,对极值点进行拟合处理. 设定阈值 γ ,利用 Hessian 矩阵求特征点的主曲率. 记 $\text{Tr}(H) = D_{xx} + D_{yy}$, $\text{Det}(H) = D_{xx}D_{yy} - (D_{xy})^2$. 如果满足 $\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(\gamma+1)^2}{\gamma}$,则保留,反之剔除(根据 Lowe 的建议, γ 的值取 10).

$$H(x, y) = \begin{bmatrix} D_{xx}(x, y) & D_{xy}(x, y) \\ D_{yx}(x, y) & D_{yy}(x, y) \end{bmatrix} \quad (2)$$

2.3 确定特征点的主方向

为使生成的描述子具有旋转不变性,根据特征点的幅角统计幅值的方式来对特征点的方向进行确定. 对特征点所在的邻域窗口内,计算每个像素点的幅值 $m(x, y)$ 与幅角 $\theta(x, y)$.

$$m(x, y) = ((L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2)^{\frac{1}{2}} \quad (3)$$

$$\theta(x, y) = \arctan \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \quad (4)$$

根据幅角以 10 度为一柱,划分为 36 柱(柱所代表的方向为像素点梯度方向,柱的长短代表了梯度幅值),直方图的峰值即为该特征点的主方向. 为增强算法的鲁棒性,大于峰值 80% 的值选为特征点辅方向.

2.4 生成 SIFT 特征点描述符

以特征点为中心取一个 16×16 的正方形窗口(为了保证特征点具有旋转不变性,将坐标轴旋转到特征点所在的主方向),把窗口划分为 4×4 个子窗口. 然后在每一个子窗口上根据幅值统计 8 个方向的梯度累加值,共形成 $4 \times 4 \times 8$ (即 128)维的特征向量. 最后对特征向量进行亮度归一化处理,完成描述符设计.

3 改进的 PCA-SIFT 算法

3.1 改进的非极大值抑制法

SIFT 算法在尺度空间极值检测阶段提取了大量的特征点,但却没有充分考虑到特征点的分布情况,导致在特征点密集区域会产生一定数量的不稳定特征点,影响匹配的正确率. 同时,对于图像拼接来说,只需要少量有效的特征点就足够完成变换矩阵的计算,如果特征点过多,不仅会增加匹配耗时,有时甚至会影响拼接效果. 为了获得更加稳定且分布均匀的特征点,本文采用改进的非极大值抑制的方法. 该方法把 SIFT 算法在每个高斯金字塔层检测到的局部极大值点和局部极小值点分别存放进两个有序队列中. 然后对于两个有序队列的每个极值点进行遍历,以极大值点序列为例:对于每个极大值点以其本身为中心建立圆形窗口,在窗口内搜索其他极大值点,如果窗口内不存在其他极大值点,则此点为极值点,予以保留;如果窗口内存在其他极大值点,则与其进行比较,若此点的值小于其他极值点则认为其不是极值点,予以剔除,若此点的值大于其他极值点则认为其是圆形窗口内的极值点,予以保留,同时把圆形窗口内的其他极值点剔除. 对于极小值点序列同样执行此操作,通过这种方式,每个极值点窗口内仅存在自身一个极值点. 最后对于两个序列进行合并,得到空间离散分布的极值点.

3.2 特征描述符设计

SIFT 算法采用方形特征邻域来生成描述符,其缺点是:特征点方向的测量以一定的角度步进,计

算过程中不可避免会产生误差,如图 1(a)所示,假设理想图像应旋转到 (x', y') 的位置,而由于主方向角度误差的原因,实际图像只旋转到 (x'', y'') 的位置.这样,原本需要统计的灰色区域就被忽略掉,而原本不需要统计的黑色区域反而被包含进来,从而降低了旋转后的区域相似性.相比之下,采用圆形区域可以很好的解决这类问题.从图 1(b)可以看出,圆形区域旋转前后覆盖的像素是一致的.本文把圆的直径选为正方形的边长,这样做的优点是:圆形区域比正方形区域含有更少的像素点,剔除了边缘易受干扰的像素点,使得计算梯度描述子的时间有所缩短,相似性加强.

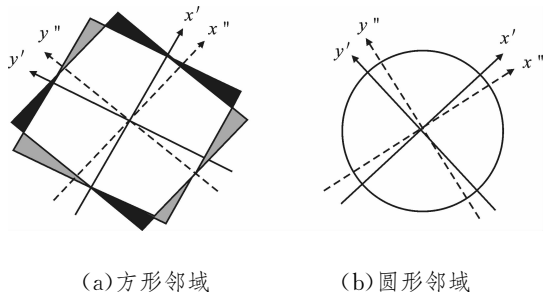


图 1 方形与圆形邻域误差比较

曾峦等^[8]基于圆形邻域采用扇形区域分割构建描述符,把圆形邻域分割成 8 个扇形区域提取 64 维梯度描述符.杨恒等^[9]基于圆形邻域采用圆环区域分割构建描述符,把圆形邻域分割成 6 个圆环区域 48 维提取梯度描述符.两种方法在抗干扰能力上都强于 SIFT 算法,但是没有充分考虑特征邻域的空间分布情况.本文采用以上两种方式相结合的方法分割构建梯度描述符,分割方式如图 2 所示,主要步骤如下:

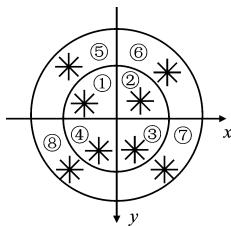


图 2 基于圆形区域的梯度描述子

Step1 以特征点为中心,提取半径为 $8s$ 的圆形区域,并对该区域以 s 为采样间隔进行等间隔采样,其中 s 为特征点所在图层的尺度.为了实现旋转不变性,将圆形区域根据特征点的方向进行旋转.

Step2 把圆形区域划分为两环,半径分别为 5 和 8,并根据 X 轴与 Y 轴围成的 4 个象限,再依次分为四个部分,得到八个小区域(如图 2 序号①到序号⑧所示).

Step3 用式(3)、式(4)计算每个像素点的幅值与幅角.

Step4 对于每个小区域的像素点,把 360° 分为 8 个柱,即 45° 为一个柱,每个像素点根据其幅角值确定其在哪一个柱上,然后对其幅值进行高斯加权运算并把结果累加在相应柱上.加权系数为:

$$W_{ij} = \frac{1}{2\pi\sigma_0^2} \exp\left(-\frac{(i-i_0)^2 - (j-j_0)^2}{2\sigma_0^2}\right) \quad (5)$$

式中, (i, j) 和 (i_0, j_0) 为所求像素点和中心点在待描述图像区域中的坐标; σ_0 为选定的常数.

Step 5 以①号区域内的八个柱的柱值编码为 1~8 号向量,②号区域内的八个柱的柱值编码为 9~16 号向量,依此类推,⑧号区域内的八个柱的柱值编码为 57~64 号向量.对 64 维向量进行亮度归一化处理形成 64 维梯度描述子.

3.3 主成分分析法降维

提取的 64 维梯度描述符虽然相对于 SIFT 算法的 128 维描述符在维数上有一定的降低,但是仍然偏大,我们需要对其进一步降维,以降低数据的复杂度.本文采用主成分分析(PCA)法对前一步骤提取特征描述符进行降维^[10],具体过程如下:

Step1 将图像中 n 个特征点的 64 维梯度描述符 $x_1, x_2, \dots, x_n$ 作为样本构成一个 $n \times 64$ 的矩阵 $\mathbf{X}$,并对这 n 个向量计算 64×64 的协方差矩阵 $\mathbf{R}$.

Step2 计算协方差矩阵 $\mathbf{R}$ 的 64 个特征值 λ 与特征向量 α ,并把特征值与特征向量根据特征值从大到小进行排序,即 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_{64}$,对应的特征向量 $(\alpha_1, \alpha_2, \dots, \alpha_{64})$.

Step3 选择前 32 个特征值对应的特征向量组成一个 64×32 的投影矩阵 $\mathbf{W}$.

Step4 将 $n \times 64$ 的梯度描述符矩阵 $\mathbf{X}$ 与 64×32 的投影矩阵 $\mathbf{W}$ 相乘,得到 $n \times 32$ 的矩阵 $\mathbf{Z}$ 如式(6),即降维描述符向量组成的矩阵.此时 n 个特征点的描述符向量均为 k 维.

$$\mathbf{Z} = \mathbf{X} \cdot \mathbf{W} \quad (6)$$

这样就把上一步提取的 64 维梯度描述符将成了 32 维的 PCA-SIFT 描述符,极大的减少了描述符的数据复杂度.

4 图像匹配

4.1 匹配点初筛

对于浮点型特征描述向量的匹配一般采用欧氏距离作为描述符之间的相似性度量.匹配方式主要有穷举法和建立数据索引两类, SIFT 算法采用穷

举法进行特征点匹配,即将数据集中的点与查询点逐一计算距离,这种方法显然效率低下,耗时较多.因此本文采用建立数据索引的方法,利用数据集中的点建立 K-D 树,使用 BBF(Best Bin First)搜索策略来寻找最近邻和次近邻点. BBF 通过建立优先队列确保查询机制能优先包含最近邻点的空间,极大的缩短了匹配时间,并且总能获得较好的结果.

计算 BBF 搜索策略寻找到的最近邻点 D_{first} 与次近邻点 D_{second} 的比值,如果其比值小于或等于某个阈值 $T^{[11]}$,则认为是判定为候选匹配点.

$$\frac{D_{\text{first}}}{D_{\text{second}}} \leq T$$

(7)

4.2 匹配点提纯

随机抽样一致性(RANSAC)算法是一种鲁棒性较高的变换估计算法^[12].它可以在一组包含“外点”的数据集中,寻找最优参数模型,剔除不符合最优模型的“外点”,极大的提升匹配正确率.

RANSAC 方法的主要思想是:对于一组候选匹配点集,随机抽取 4 个不共线样本,计算其变换矩阵(记模型 M).把匹配点集中所有点与模型 M 计算投影误差,若误差小于阈值,记为内点;若误差大于阈值,记为外点.通过不断迭代的方式,寻找最优参数模型,即内点最多.根据寻找到的最优参数模型,把所有外点剔除,从而优化匹配效果,提升配准率.

5 实验与结果分析

为验证本文算法的有效性和可行性,实验采用的性能评价指标包括:

- (1)离散度,评价特征点的分布情况;
- (2)正确匹配率,评价特征描述符的匹配性能;
- (3)时间代价,评价图像拼接的速率.

所有的实验都运行在 2.20 GHz CPU 和 4GB 内存的 PC 机环境上,在 VS2010 上采用 VC++ 编程语言实现.

5.1 特征点分布情况分析

为了描述特征点的分布情况,引入离散度的概念来定义特征点分布的离散程度.令图像中所有特征点构成的集合为 F ,则每一个特征点可以用 $f_i = (x_i, y_i)$ 表示.假设图像中的特征点的数目为 n ,那么图像中的所有特征点离坐标原点的平均距离为:

$$m = (\frac{1}{n} \sum_{f_i \in F} x_i, \frac{1}{n} \sum_{f_i \in F} y_i)$$

(8)

为了描述离散程度,定义离散度 S 为:

$$S = \frac{1}{n} \sum_{f_i \in F} ((f_i - m)(f_i - m)^T)^{1/2}$$

(9)

对于一副图像中特征点分布.离散度 S 越大,表明特征点的分布越离散越均匀; S 越小,表明特征点的分布越密集越不均匀.

本文采用 490×490 的 lena 图片作为离散度分析的实验图像. SIFT 算法可以提取在图像中提取到 1 127 个特征点,特征点的分布情况如图 3 (a)所示,特征点的离散度为 126.628,本文算法可以提取到 701 个特征点,离散度为 133.706.通过图 3 可以很发现,用本文算法处理过的图片特征点分布比用 SIFT 算法处理过的更加均匀.在特征点分布稀疏的区域, SIFT 算法和本文算法提取的特征点位置基本一样,但在特征点分布密集的区域,本文算法提取的特征点比 SIFT 算法更加稀疏均匀,这种特征点分布方式更加有利于图像之间的配准.



(a)SIFT 算法 (b)本文算法

图 3 lena 图片特征点分布情况

5.2 匹配性能对比分析

为了验证所提算法的匹配性能,选取 10 组具有一定重叠度的待拼接图像分别使用 SIFT, SURF 与本文算法进行特征匹配.使用正确匹配率对匹配结果进行比较.正确匹配率定义为:

$$\text{正确匹配率} = \frac{\text{正确匹配对数}}{\text{总匹配对数}}$$

图 4~图 6 是 10 组实验图像中的前三组匹配结果, SIFT, SURF 和本文算法的阈值 T (式(7))都设为 0.6.对三组实验图像的实验结果就特征点数与正确匹配率进行统计,结果如表 1 所示.

针对图 4(a)、图 5(a)和图 6(a)所示的测试图像,从表 1 可以看出虽然本文算法匹配点对数较少,但是正确匹配率大大优于 SIFT 和 SURF 算法.这是由于本文采用了改进的非极大值抑制法剔除了部分不稳定的特征点,使得特征点之间分布较为均匀,同时本文采用圆形区域构建描述符,具有更好的鲁



图 4 实验组 1 测试图像及三种算法的匹配结果

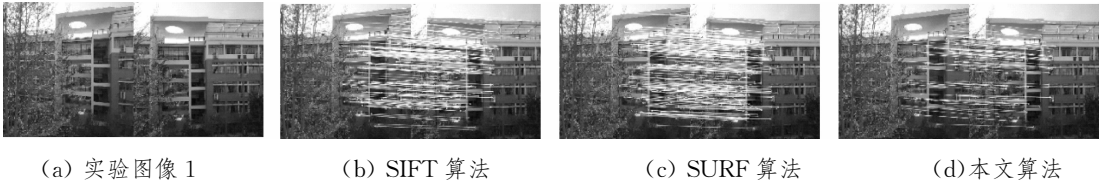


图 5 实验组 2 测试图像以及三种算法的匹配结果

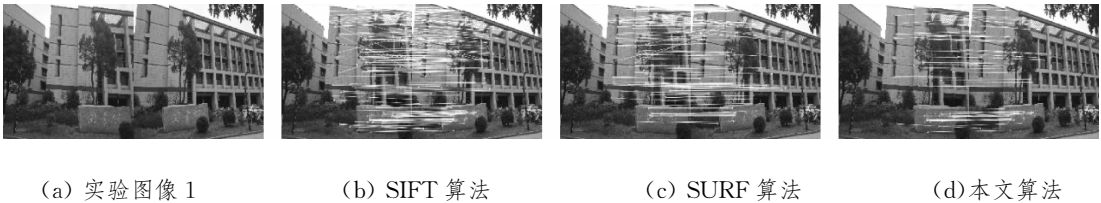


图 6 实验组 3 测试图像以及三种算法的匹配结果

棒性,从而降低了误匹配点对数.

表 1 正确匹配率比较

测试图像	算法	匹配点对数	误匹配点对数	正确匹配率/%
图 4(a)	SIFT	305	18	94.10
	SURF	292	28	90.41
	本文	95	1	98.95
图 5(a)	SIFT	174	17	90.23
	SURF	195	25	87.18
	本文	72	2	97.22
图 6(a)	SIFT	168	25	85.12
	SURF	117	19	83.76
	本文	61	2	96.72

5.3 时间复杂度分析

实验中,在各项参数相同情况下,使用 SIFT 算法和本文算法分别对图 4(a)、图 5(a)和图 6(a)进行拼接,本文算法的拼接结果如图 7 所示.同时对两种算法就特征点数量和拼接时间上进行比较,结果如

表 2 所示.

表 2 时间复杂度比较

测试图像	算法	特征点数	拼接耗时/s
图 4(a)	SIFT	2 094/2 030	11.92
	本文	1 069/1 164	6.08
图 5(a)	SIFT	3 578/2 254	15.47
	本文	1 926/1 347	7.58
图 6(a)	SIFT	1 690/2 541	12.18
	本文	1 088/1 505	6.73

从表 2 可以看出,本文算法检测到的特征点数量相对于 SIFT 算法大为减少,这加速了本文算法构建描述符和匹配用时,所以在总体拼接速度上,本文算法提高较为明显.而且从图 7 可以看出本文算法的拼接效果较为出色.实验中,统计 10 组实验图像的拼接耗时,如图 8 所示.



图 7 三组实验图像拼接结果

从图 8 可以看出,本文算法的拼接速度是传统的 SIFT 算法的 1.6~2.2 倍.使用本文算法可以实

现高质量、高效率的图像拼接.

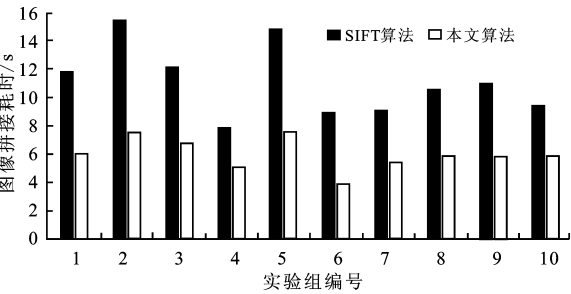


图 8 实验图像拼接耗时

6 结束语

本文通过分析传统 SIFT 算法构建描述子和匹配时间过长导致拼接效率低的不足,提出了一种面向拼接的改进 PCA-SIFT 算法. 该算法比 SIFT 算法和 SURF 算法具有更好的鲁棒性,同时进一步提升了拼接的速度. 实验结果表明本文算法的拼接速度是 SIFT 算法的 1.6~2.2 倍,是一种有效、可靠的快速图像拼接方法.

参考文献:

[1] 张勇,王志锋,马文. 基于改进 SIFT 特征点匹配的图像拼接算法研究[J]. 微电子学与计算机, 2016, 33(3):60-64.

[2] 蒋晓玲,章慧. 基于多特征融合的激光医学图像拼接[J]. 激光杂志, 2015, 36(12):55-58.

[3] Yu Qing H E, Xue W, Wang S Y, et al. Modified SIFT descriptor and key-point matching for fast and robust image mosaic[J]. 北京理工大学学报(英文版), 2016, 25(4):562-570.

[4] Dong Q, Liu J. Seamline determination based on PKGC segmentation for remote sensing image mosa-

icking[J]. Sensors, 2017, 17(8):1721.

[5] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630.

[6] Lowe D G. Distinctive image features from scale invariant key points [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.

[7] 李玉峰,李广泽,谷绍湖,等. 基于区域分块与尺度不变特征变换的图像拼接算法[J]. 光学精密工程,2016,24(5): 1197-1205.

[8] 翟优,曾峦,熊伟. 不同局部邻域划分加速鲁棒特征描述符的性能分析[J]. 光学精密工程, 2013, 21(9): 2395-2404.

[9] 杨恒,王庆. 一种新的局部不变特征检测和描述算法[J]. 计算机学报, 2010, 33(5):935-944.

[10] 蒋波,翟旭平. 基于 PCA-SIFT 特征匹配的图像拼接算法[J]. 计算机应用, 2016, 36(2):143-145.

[11] Luan H, Law N F. A novel dual-threshold SIFT-based copy-move forgery detection[C]// Signal and Information Processing Association Summit and Conference. Kuala Lumpur, Malaysia, IEEE, 2017:1-5.

[12] Zhang Z, Kuang L, Han X, et al. An improved RANSAC algorithm for image mosaic[J]. Computer Measurement & Control, 2014, 22(6): 1856-1858.

作者简介:

杨炳坤 男,(1993-),硕士研究生. 研究方向为计算机视觉.

程树英 女,(1966-),博士,教授,博士生导师. 研究方向为薄膜太阳能电池和太阳能光伏及其应用.

郑茜颖(通讯作者) 女,(1970-),博士,副教授. 研究方向为人工智能、计算机视觉. E-mail: zhengqy@vip. sina. com.