

# 面向宽基线立体影像匹配的高质量仿射不变特征提取方法

程亮<sup>1</sup>, 龚健雅<sup>1</sup>, 宋小刚<sup>2</sup>, 杨晓霞<sup>1</sup>

1. 武汉大学 测绘遥感信息工程国家重点实验室, 湖北 武汉 430079; 2. 武汉大学 遥感信息工程学院, 湖北 武汉 430079

## A Method for Affine Invariant Feature Extraction with High Quality for Wide Baseline Stereo Image Matching

CHENG Liang<sup>1</sup>, GONG Jian-ya<sup>1</sup>, SONG Xiao-gang<sup>2</sup>, YANG Xiao-xia<sup>1</sup>

1. State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China; 2. Remote Sensing and Information Engineering School, Wuhan University, Wuhan 430079, China

**Abstract:** Wide baseline stereo image matching is a challenging problem in image processing and computer vision. For the purpose of wide baseline stereo image matching, a filter strategy of affine invariant feature based on information content and spatial distribution constraints is put forward. Moreover, a new feature extraction algorithm integrating MSER, SIFT and this filter method is discussed, in order to realize high quality (high information content, good spatial distribution, high repeatability) affine invariant feature extraction. The experiments by using of actual wide baseline images sequence with different texture features prove that this method guarantees much better repeatability, which benefits the subsequent stereo image matching.

**Key words:** wide baseline stereo matching; affine invariant; feature extraction; information content; spatial distribution dispersion

**摘要:** 针对宽基线影像匹配问题, 提出基于信息量和空间分布均衡性双重约束的多层次特征筛选方法, 并在此基础上探讨集成该特征筛选方法、MSER (Maximally Stable Extremal Region detector)、SIFT (Scale Invariant Scale Transformation) 的特征提取算法, 以达到提取高质量 (高信息量、空间分布均衡、高重复率) 仿射不变特征的目的。实验表明, 相对现有最优算法, 本文方法具有更高的重复率和匹配成功率, 更有利于后续的 (宽基线) 立体匹配。  
**关键词:** 宽基线立体匹配; 仿射不变; 特征提取; 信息量; 空间分布离散度

## 1 引言

影像匹配一直是计算机视觉、数字摄影测量等领域所共同关心的学术前沿和难点问题。目前, 多数算法是针对短基线条件的。宽基线条件下, 由于视点的改变、光照条件的变化等因素, 影像自动匹配远比短基线困难。影像不变特征提取技术的成熟使得宽基线影像立体匹配成为可能<sup>[1]</sup>。影像不变特征指对图像变化 (旋转、缩放、仿射变换、视角变化等) 保持稳定的局部区域 (local region) 及其特征描述。仿射不变 (affine invariant) 特征提取技术目前是该领域的最新技术, 相

关的研究异常活跃。文献[2]介绍了当前常用的尺度不变特征提取算子和仿射不变特征提取算子。文献[3]对常用仿射不变特征提取算子性能进行了对比分析, 指出在大多数情况下, MSER (Maximally Stable Extremal Region detector) 算子<sup>[4]</sup>效果最佳。文献[5]指出 SIFT (Scale Invariant Scale Transformation)<sup>[6]</sup>是目前最佳的特征描述符。将 MSER 与 SIFT 结合起来, 无疑是一项值得研究的工作, 然而不论是 MSER、SIFT, 还是将两者简单结合, 都存在着一些问题。

基于此, 本文提出基于信息量和空间分布均衡性双重约束的多层次特征筛选方法, 并在此基

收稿日期: 2007-02-06; 修回日期: 2007-06-01

基金项目: 国家重点基础研究发展计划项目 (2006CB701300)

作者简介: 程亮 (1978-), 男, 江苏南京人, 博士生, 主要从事数字摄影测量、3 维重建等方面的研究。E-mail: geocl@163.com

基础上探讨将该筛选方法与 MSER 算子、SIFT 特征描述符相结合的特征提取算法,以达到提取高质量(高信息量、空间分布均衡、高重复率)仿射不变特征的目的。整个解决方案的关键是特征筛选方法,其核心思想是在特征区域(region)被检测出后,进行特征描述前,以信息量和空间分布均衡性为衡量标准,对提取出的特征进行评估,过滤掉信息量不佳、空间分布均衡性差的特征,只保留高质量特征,以提高特征的重复率以及后续的立体影像匹配的成功率,进而有利于后续的(宽基线)立体匹配。

## 2 现有仿射不变特征提取算法分析

文献[3]介绍了当前常见的仿射不变特征提取算法:① Harris-Affine 算子;② Hessian-Affine 算子;③ MSER 算子(Maximally Stable Extremal Region detector);④ EBR算子(Edge-Based Region detector);⑤ IBR 算子(Intensity extremal-Based Region detector);⑥ salient regions 算子。文献[3]通过对比试验,进一步分析了这些仿射不变特征提取算法,得出如下结论:随着视角变化的增大,所有这些算子的性能都以大致类似的速率衰减;没有一个算子可以很好地处理所有类型的影像、所有类型的图像变化;在大多数情况下,MSER 算子性能最佳,Hessian-Affine 其次;MSER 算子对于含有相似区域并有明显边界的影像效果较好,IBR 算子也具有类似的特性;Hessian-Affine 算子和 Harris-Affine 算子相对其他算子,可以检测出更多的特征区域,这一特性有利于有遮挡或紊乱情况的影像匹配;EBR 算子适用于交叉边界较多的影像匹配;Salient Region 算子性能相对较低。

文献[5]对比了 SIFT, steerable filters, differential invariants, complex filters, moment invariants 等特征描述符。试验结论是 SIFT 对图像缩放、旋转甚至仿射变换保持不变性,是目前最佳的特征描述符。

将现有最佳的特征提取算子 MSER 和最佳的特征描述符 SIFT 结合起来,无疑是一项值得研究的工作。然而试验表明:即使特征提取算子 MSER,在视点角度变化达到或大于  $30^{\circ} \sim 40^{\circ}$  时,特征的重复率和匹配成功率将迅速衰减,导致影像匹配效果不佳。特征描述符 SIFT 采用 128 维矢量表征检测出的区域,描述特征的方法复杂,构建特征向量及后续应用时计算量大。两者的简单结合,无法发挥取长补短的作用,所取得的

影像匹配效果也仍不能令人满意,因而研究进一步提高仿射不变特征的提取质量有着十分重要的意义和应用价值。

## 3 高质量仿射不变特征提取方法

### 3.1 特征区域(local region)的信息量

文献[7]指出影像信息量与影像可匹配性之间存在很强的相关性,尽管由于地表景观的不同,表现出的具体规律略有差异,但都表现出匹配正确率随影像信息量的增大而增大的趋势。因此,本文采用信息熵作为衡量特征区域信息量的定量指标,根据特征区域信息熵的大小,对被提取特征进行筛选。仿射不变特征提取算法中,不同的算法提取出的特征形状各异,如 Harris-Affine 算子,Hessian-Affine 算子提取的是椭圆区域,MSER 算子提取的是不规则区域,因此采用规则网格评估信息熵的方式<sup>[7]</sup>在此已不再适用。

本文采用基于任意形状特征区域计算影像信息熵的方法。具体步骤如下:

1. 以特征区域的形状为感兴趣区域,创建与之对应的掩模(mask);
2. 利用掩模作用原图像,屏蔽非感兴趣区域;
3. 采用信息熵表征感兴趣区域的信息量,利用式(1)计算。

$$H = - \sum_{i=1}^M \sum_{j=1}^N P_{ij} \log_2 P_{ij} \quad (1)$$

式中, $P_{ij}$ 是不同的灰度值在整个图像中出现的概率。

该方法对于规则区域(矩形、圆、椭圆等)、不规则区域等均可以快速地计算出信息熵,将在后续的特征筛选方法得到应用。

### 3.2 信息量和空间分布约束下的多层次特征筛选方法

特征点的空间分布对于影像匹配的有效性、可靠性、精度有着直接影响<sup>[8-10]</sup>。文献[10]利用金字塔影像进行多层次的影像匹配,指出种子点的选取方式直接决定了金字塔影像匹配的有效性和可靠性,种子点必须合理地分布在整幅影像上。文献[9]指出立体影像匹配生成的 DEM 精度并没有随种子点数量的增多而显著提高,但与这些种子点在影像上的几何分布有关。

本文将特征的空间分布问题抽象为特征区域中心点的空间分布问题。点的空间分布可以通过式(2)来计算,本文称之为空间分布离散度。式(2)计算出的  $Disp$  值越小,说明空间离散度小,即

点在空间分布上越集中; $Disp$  值越大,说明空间离散度大,即点在空间分布上越分散。

$$Disp = \sqrt{\frac{\sum_{i=1}^n (x_i - x_{wmc})^2 + \sum_{i=1}^n (y_i - y_{wmc})^2}{n}}$$

$$(\bar{x}_{wmc}, \bar{y}_{wmc}) = \left( \frac{\sum_{i=1}^n w_i x_i}{\sum_{i=1}^n w_i}, \frac{\sum_{i=1}^n w_i y_i}{\sum_{i=1}^n w_i} \right) \quad (2)$$

式中, $w_i$  是点  $i$  的权重,本文中采用特征区域的信息熵作为权。

鉴于信息量的大小与特征的空间分布状况对于影像匹配均具有重要意义,本文提出基于信息量和空间分布双重约束的特征筛选方法。采用“整体控制、局部优化”的策略,将特征筛选扩展为多层次的筛选过程。本筛选方法利用信息熵表征特征区域的影像信息量,以特征信息量为参照,采用双阈值的策略进行两层筛选,同时从规则格网划分的整体控制到单元格中空间分布离散度阈值的约束,保证了筛选结果的空间分布均衡性。具体算法如下:

1. 确定特征选取的数量,设为  $M$ ;
2. 计算全局熵下限(阈值  $T_1$ ),对所有特征进行过滤,只保留熵大于  $T_1$  的特征;
3. 划分  $3 \times 3$  的规则格网,共有 9 个单元格;
4. 确定每个单元格中包含的特征,计算单元格的平均熵  $E_{avg}$ ,按平均熵的比例确定各单元格的特征提取数量,称为配额,即将特征提取工作分配到每个单元格中;
5. 对于每个单元格  $i$ ,做如下计算:
  - (1) 根据特征区域的中心点,确定单元格  $i$  包含的特征;
  - (2) 统计单元格  $i$  中特征的熵,设定单元格  $i$  的局部熵下限(阈值  $T_{local_i}$ );
  - (3) 在单元格  $i$  中,过滤熵小于  $T_{local_i}$  的特征;
  - (4) 根据配额,在剩余特征中选择,计算选择集的空间分布离散度  $Disp_i$ ;
  - (5) 如果  $Disp_i$  大于空间分布离散度阈值,则单元格  $i$  中的特征提取工作完成,如果小于阈值,则更新选择(用剩余集中熵最大的特征替换距离中心最近的特征),重新计算,直至满足阈值;
6. 循环所有单元格,选出每个单元格的局部占优特征;
7. 累加所有单元格特征,计数为  $F$ ,如果  $F$

$< M$ ,则从全局剩余特征中检索出熵值最大的前  $M-F$  个特征。特征选取结束。

本文提出的特征筛选方法特点在于“双重约束”和“多层次”。双重约束是指基于高信息量和空间分布均衡性的双重约束。多层次则是双重约束下的多层次,即一方面是“基于信息量的多层次筛选”,另一方面是“从整体到局部的多层次空间分布约束”。“基于信息量的多层次筛选”是指首先通过全局熵下限过滤信息量明显差的特征,从整体上保证了高信息量,再通过单元格中的局部熵下限,精选出局部占优特征。“从整体到局部的多层次空间分布约束”是指先利用规则格网进行整体划分,从整体上杜绝了极端空间分布情况的出现,再针对每个单元格采用空间分布离散度的阈值进行约束,实现局部的空间分布优化。

### 3.3 特征提取算法

在上述特征筛选方法的基础上,笔者将其与 MSER 算子、SIFT 特征描述符集成,提出一种高质量仿射不变特征的组合提取算法。本算法可以归纳成三大步骤:① 利用标准 MSER 算子检测仿射不变特征;② 特征筛选,评估和选择标准 MSER 算子检测出的特征,计算每个特征的信息量和空间分布离散度,保留信息量和空间分布离散度符合要求的特征,去除不符合要求的,以有利于后续的立体匹配;③ 利用 SIFT(Lowe, 1999) 作为筛选出的特征的描述符。其中的第二步特征筛选是本算法关键之所在,具体又可分为四个步骤:计算特征信息熵、基于全局熵阈值进行特征全局过滤、格网化并计算个单元格的特征配额数、每个单元格的处理等。其中,每个单元格的处理是指针对每个单元格进行局部过滤(基于局部熵阈值),再结合单元格配额根据信息熵选择特征,计算选择集的空间离散度,如果空间离散度小于阈值则更新选择集重新计算,如果大于阈值则该单元格中的特征提取工作结束。本特征提取算法的详细实现流程可参见图 1。

值得一提的是特征提取数量的确定问题。针对宽基线影像配准的具体应用,用于仿射不变特征的几何形状往往是椭圆或不规则区域,几何定位精度不高,直接用来做精确匹配尚有很大技术难度。以现有的技术水平,仿射不变特征更适合用于确定立体像对之间的大致对应关系,即“粗匹配”,因此本文认为特征选取的数量没有必要设定过多。同时,通过多次试验,笔者发现特征选取的数量也不

可过少,因为宽基线条件下特征对应的技术难度很大,特征误匹配的情况不可避免,过少的特征在匹配时会因为误匹配造成匹配结果的不稳定,降低匹配的可靠性和鲁棒性。一般情况下,建议特征选取的数目在20~100左右为宜。

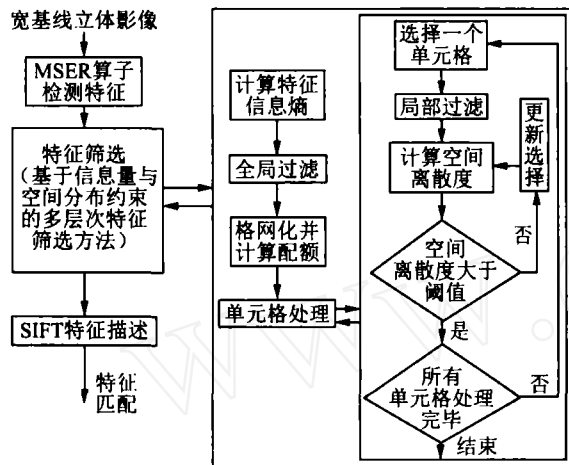


图1 本文提出的仿射不变特征提取算法基本流程图  
Fig. 1 Flowchart of affine invariant feature extraction algorithm

本算法的特点在于,充分利用当前最优的算子和特征描述符,并将两者组合使用,在此基础上融入特征筛选方法,提高了特征提取的质量。同时由于大量冗余特征已被过滤,算法的总体效率上也有较大提高。本算法的实现结果图(图2)反映出效果明显:标准MSER算子特征提取(未经筛选)结果图(图2(a))中存在大量的、重叠的特征,本算法特征提取(经过筛选)的结果图(图2

(b)中特征纹理细节丰富、空间分布均衡。



图2 本文算法与标准MSER算子特征提取效果比较图  
Fig. 2 A comparison of feature extraction between standard MSER and our algorithm on a same image

## 4 试验与分析

### 4.1 试验数据

为了验证本文所探讨方法的有效性,将本文方法与现有最优算子MSER和次优算子Hessian-Affine进行对比试验。试验使用了与文献[3]相同的影像组。为了检测本文算法面向宽基线序列影像匹配的效果,选择了其中的视点变化影像组。为了验证本方法对于不同纹理特征影像的适用性,选择两种类型纹理特征的影像组,分别是壁画影像和砖墙影像,其中壁画带有明显的边界特征,纹理丰富;砖墙影像有大量重复的纹理,纹理具有高度的相似度。对于这两组影像没有进行任何的预处理,如去噪声、归一化等。为了验证特征的重复率和匹配的成功率,原始影像与其他5个视点变化影像之间的对应关系已经测定,由单应性矩阵(homography)表示。每张影像大小约为 $800 \times 640$ ,见图3。

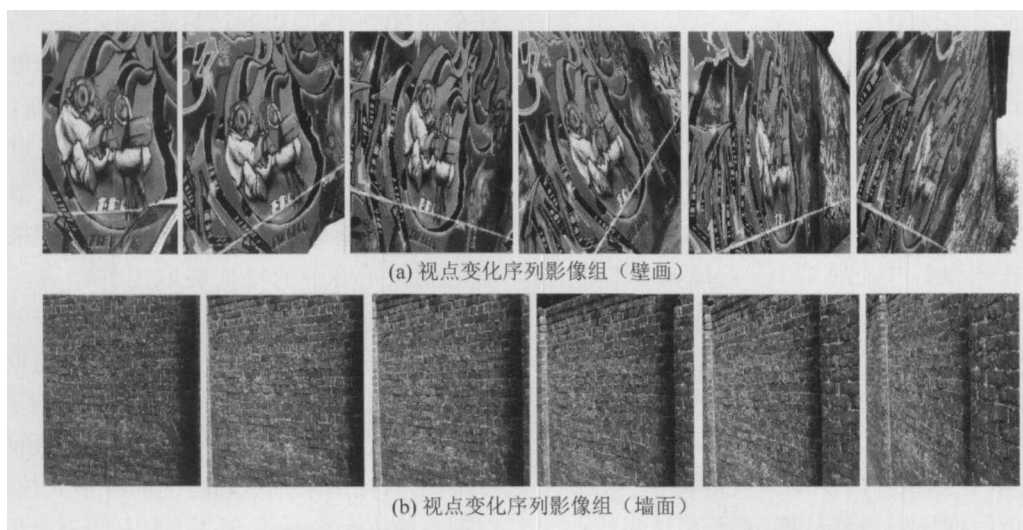


图3 视点变化序列影像组,最左侧是参考影像,其余从左到右是 $20^\circ$ , $30^\circ$ , $40^\circ$ , $50^\circ$ , $60^\circ$ 影像  
Fig. 3 Viewpoint change image sequences, the most left image is used as the reference; the others are respectively  $20^\circ$ ,  $30^\circ$ ,  $40^\circ$ ,  $50^\circ$ ,  $60^\circ$  image from left to right

4.2 评价指标

文献[11]指出,对于影像匹配问题,应该从重复率(repeatability)和信息量(information content)两个方面来对特征提取算子进行评价。由于本文的方法中已经采用了信息量技术,为了避免循环论证的嫌疑,本文只采用重复率作为算法性能对比的评价指标。重复率的计算基于单应性矩阵这一前提,利用单应性矩阵将待匹配影像投影到参考影像,对于同名特征而言,如果两者的区域重叠误差(overlap error)小于阈值,则认为两者一一对应,即重复。可重复的特征占有所有提取出的特征数目的百分比即重复率,重复率清楚地表达了对同一个场景在不同成像条件下所获得影像上所提取特征的几何稳定性。重复率可利用式(3)计算:

$$1 - \frac{R_{\mu a} \cap R_{(H^T \mu b H)}}{R_{\mu a} \cup R_{(H^T \mu b H)}} < \varepsilon \quad (3)$$

式中, $a$ 与 $b$ 表示一个像对, $H$ 是该像对的单应矩阵, $R_{\mu a}$ 表示影像 $a$ 上某局部区域面积, $R_{(H^T \mu b H)}$ 表示影像 $b$ 上与 $a$ 对应的局部区域面积。 $R_{\mu a} \cap R_{(H^T \mu b H)}$ 表示对应局部区域的交集, $R_{\mu a} \cup R_{(H^T \mu b H)}$ 表示对应局部区域的并集,局部区域交集或并集的面积由数值计算可得。

4.3 试验结果与分析

本文分别利用壁画序列影像和墙面序列影像进行试验,设定特征选取数量为50,分别在覆盖误差阈值为30,40的条件下,进行对比试验。试验共对比了四种仿射不变特征提取算子,其中本文方法分为两个方案,分别是顾及空间分布的方案(称“方案一”);未顾及空间分布的方案(称“方案二”),所以共有方案一、方案二、MSER算子、Hessian-Affine算子四种算法。图4显示了四种算法对于壁画影像、覆盖误差阈值为30时的重复率;图5显示了四种算法对于壁画影像、覆盖误差阈值为40时的重复率。图6显示了四种算法对于墙面影像、覆盖误差阈值为30时的重复率;图7显示了四种算法对于墙面影像、覆盖误差阈值为40时的重复率。分析认为:

1. 对于壁画、墙面这两组纹理特征明显不同的影像,不论覆盖误差阈值是30还是40,本文方法的两种方案在重复率方面均明显优于MSER算法、Hessian-Affine算法,尽管随着视点变化的不同情况表现出的优势有所不同,但整体的优越性十分明显。

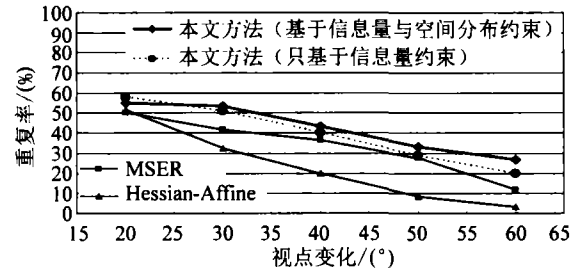


图4 重复率随视点变化状况图(壁画组序列影像,图3(a),覆盖误差阈值为30%)

Fig. 4 Repeatability score for viewpoint change (Graffiti image sequences, Fig. 3(a), overlap error: 30%)

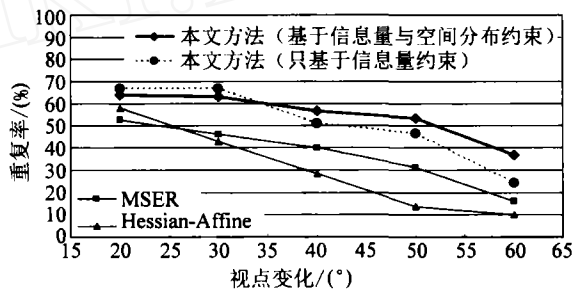


图5 重复率随视点变化状况图(壁画组序列影像,图3(a),覆盖误差阈值为40%)

Fig. 5 Repeatability score for viewpoint change (Graffiti image sequences, Fig. 3(a), overlap error: 40%)

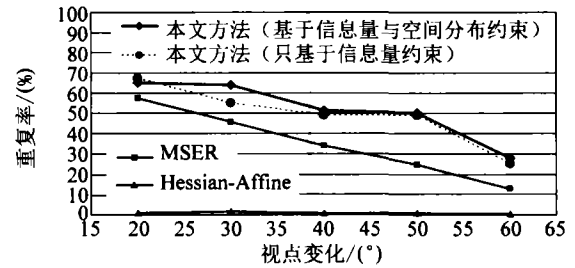


图6 重复率随视点变化状况图(墙面组序列影像,图3(b),覆盖误差阈值为30%)

Fig. 6 Repeatability score for viewpoint change (Wall image sequences, Fig. 3(b), overlap error: 30%)

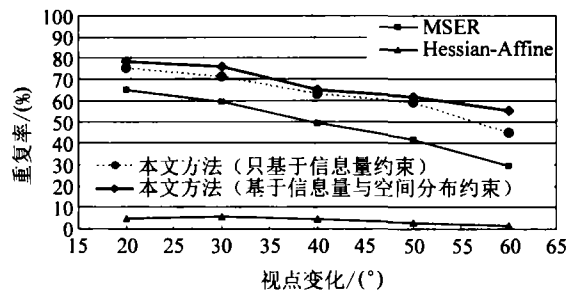


图7 重复率随视点变化状况图(墙面组序列影像,图3(b),覆盖误差阈值为40%)

Fig. 7 Repeatability score for viewpoint change (Wall image sequences, Fig. 3(b), overlap error: 40%)

2. 大多数情况下,方案一要优于方案二,进一步分析如下。从图4,图5可以看出,当视点变化不大(小于 $30^\circ$ )时,两个方案性能相当;当视点变化较大(大于 $30^\circ$ )时,随着视点变化的增大,方案一优势逐渐明显。可见,对于视点变化较大的影像,顾及空间分布的方案具有较大的优越性。从图6,图7可以看出,相对方案二,方案一在效果上提升有限,分析认为这是由于此墙面序列影像为高度相似的重复纹理影像(其空间不同区域的纹理都很相似),因而从理论上顾及空间分布的方案无法大幅提高性能,试验结果与理论是一致的。

## 5 结论

本研究聚焦于影像仿射不变特征提取的问题,提出一种高质量仿射不变特征的提取方法。该方法的核心在于基于信息量和空间分布双约束的多层次特征筛选技术,“双重约束”和“多层次”的策略,剔除了大量信息量不佳、空间分布均衡性差、空间重叠严重的冗余特征,仅保留适量高质量、高稳定性的特征,增强了对图像仿射变换等的抗干扰能力,提高了后续影像匹配的可靠性和鲁棒性。文章在此筛选方法的基础上,进一步探讨了将其与MSER算子、SIFT特征描述符相结合的组合算法,实现了提取高质量特征的目的。以重复率为评价标准,分别采用不同纹理特征的影像进行试验,对比分析本文方法与现有最佳算法MSER的表现。试验证明结果,相对MSER及其他算法,本文方法明显具有更高的重复率和匹配成功率,更适合于后续的影像匹配工作。

## 参考文献:

- [1] STRECHA C, FRANSENS R, VAN GOOL L. Wide-baseline Stereo from Multiple Views: A Probabilistic Account[A]. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition[C]. Washington: [s. n.], 2004. 552-559.
- [2] MIKOLAJCZYK K, SCHMID C. Scale & Affine Invariant Interest Point Detectors[J]. International Journal of Computer Vision, 2004, 60(1): 63-86.
- [3] MIKOLAJCZYK K, TUYTELAARS T, SCHMID C. A Comparison of Affine Region Detectors[J]. International Journal of Computer Vision, 2005, 65(1): 43-72.
- [4] MATAS J, CHUM O, URBAN M. Robust Wide-baseline Stereo from Maximally Stable Extremal Regions[A]. Proceedings of the British Machine Vision Conference[C]. Cardiff: [s. n.], 2002. 384-393.
- [5] MIKOLAJCZYK K, SCHMID C. A Performance Evaluation of Local Descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1 615-1 630.
- [6] LOWE D G. Distinctive Image Features from Scale-invariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [7] ZHU Qing, WU Bo, WAN Neng, XU Zhi-xiang, TIAN Yi-xiang. An Interest Point Detect Method to Stereo Images with Good Repeatability and Information Content[J]. Acta Electronica Sinica, 2006, 34(2): 205-209. (朱庆, 吴波, 万能, 徐志祥, 田一翔. 具有良好重复率与信息量的立体影像点特征提取方法[J]. 电子学报, 2006, 34(2): 205-209.)
- [8] ZHU Q, WU B, XU Z X. Seed Point Selection Method for Triangle Constrained Image Matching Propagation[J]. IEEE on Geoscience and Remote Sensing Letters, 2006, 3(2): 207-211.
- [9] JONATHAN H. SPOT Digital Elevation Model (DEM) Creation Using the Otto and Chau Method[DB/OL]. [http://www.jon-claudi.co.uk/ee/index.php/geofiction/extended\\_text/spot\\_digital\\_elevation\\_model\\_dem\\_creation\\_using\\_the\\_otto\\_and\\_chau\\_method/](http://www.jon-claudi.co.uk/ee/index.php/geofiction/extended_text/spot_digital_elevation_model_dem_creation_using_the_otto_and_chau_method/), 2004. 9.
- [10] CHEN Ai-jun, XU Guang-you, SHI Yuan-chun. Automated 3D Building Modeling Based on Urban Aerial Stereo[J]. Acta Geodaetica et Cartographica Sinica, 2002, 31(1): 54-59. (陈爱军, 徐光祐, 史元春. 基于城市航空立体像对的全自动三维建筑物建模[J]. 测绘学报, 2002, 31(1): 54-59.)
- [11] SCHMID C, MOHR R, BAUCKHAGE C. Evaluation of Interest Point Detectors[J]. International Journal of Computer Vision, 2000, 37(2): 151-172.

(责任编辑:丛树平)