大仿射场景的混合特征提取与匹配

佟国峰¹,李勇²,刘楠^{3*},纪光旭⁴ 东北大学信息科学与工程学院,辽宁沈阳 110819

摘要 为了提高大规模场景三维重建中的精度,在保证算法效率的前提下,提取两种局部稳定不变特征,并采用多 特征融合方法进行匹配。针对基于航拍影像和城市街景图像联合建模存在的问题,提出了一种两种局部稳定特征 匹配的方法。其步骤为:先提取 ASIFT(Affine Scale Invariant Feature Transform)特征点和 MSER(Maximally Stable Extremal Regions)特征区域,并对 MSER 算法进行改进,得到这两种稳定的图像特征;再用 SIFT(Scale Invariant Feature Transform)特征描述符对其进行描述;最后利用基于单应矩阵的特征匹配算法进行匹配。利用 图形处理单元(GPU)对特征匹配环节进行并行优化处理。大量实验及对比结果表明,本文算法可以得到两种单一 算法两倍以上的正确匹配对。

关键词 机器视觉; 三维重建; 特征融合; 特征匹配; 仿射场景 中图分类号 TP751 文献标识码 A doi: 10.3788/AOS201737.1115003

Mixed Feature Extraction and Matching for Large Affine Scene

Tong Guofeng¹, Li Yong², Liu Nan³, Ji Guangxu⁴

College of Information Science and Engineering, Northeastern University, Shenyang, Liaoning 110819, China

Abstract In order to improve accuracy of large-scale scene model in three-dimensional (3D) reconstruction, we extract two kinds of partial stable invariant features, under the premise of ensuring the efficiency of the algorithm, and use a multi-feature fusion method to match image features. Considering both problems of the joint modeling based on aerial and urban street images, we propose a matching method based on the two kinds of partial stable features. The method comprises the following steps. Firstly, we extract ASIFT (Affine Scale Invariant Feature Transform) feature points and MSER feature areas, and improve the MSER (Maximally Stable Extremal Regions) algorithm to get the two stable features described by SIFT (Scale Invariant Feature Transform) feature descriptor; secondly, we use the homography matrix to match features by the feature matching algorithm; finally, we parallelly optimize feature matching by using graphics processing unit(GPU). A large number of experiments and comparison results show that more than twice correct matching pairs can be obtained by the proposed algorithm than other two algorithms.

Key words machine vision; three-dimensional reconstruction; feature fusion; feature matching; affine scene OCIS codes 150.0155; 150.1135; 150.4232

1 引 言

近年来,基于真实场景的三维重建模型逐渐成为研究热点,无论是传统地理信息行业、军事模拟、城市规划,还是新兴的虚拟现实(VR)和增强现实(AR),都可利用点云模型作为真实场景的补充。在利用二维图像 重建的精细化三维点云模型中,二维图像中的特征提取与匹配是关键。现阶段大规模场景三维模型重建方 法主要分为两种:一种是基于无人机的航拍影像;另一种是基于街景车的城市影像。而这两种方法采集的信 息均会有一定的缺失:基于航拍的重建结果缺乏底部区域的重建信息,对底部信息重建精度较低;基于街景

收稿日期: 2017-05-16; 收到修改稿日期: 2017-06-14

基金项目:国家 863 计划 (2012AA041402)、中国科技部国家重点科技研究发展计划 (2015BAF13B00-5)

作者简介:佟国峰(1973-),男,博士,教授,主要从事基于移动视场的全景重建和虚拟现实技术等方面的研究。

E-mail:tongguofeng@ise.neu.edu.cn

^{*} 通信联系人。E-mail: 1597196180@qq.com

车的重建结果缺乏较高位置的重建信息。因此利用这两种方法均不能得到精细的三维模型。基于二者进行 联合建模(即空地一体化)可以获得丰富的信息,但存在几何变换较大、图像亮度不同、图像畸变等问题。因此,在大仿射场景的图像中提取出大量的稳定特征,从而得到强匹配特征的图像序列,对将二维图像重建为 精细化的三维模型具有重要的意义。

对于特征匹配,不仅需要保证效率,还需要保证匹配的效果。常见的特征匹配算法主要有穷举法或基于 KD 树^[1]的最近邻搜索(BBF)算法,利用这些方法对全局特征、局部特征进行匹配,如直方图特征^[2]、Harris 角点^[3]、SIFT(Scale Invariant Feature Transform)^[4]、ASIFT(Affine Scale Invariant Feature Transform)^[5] 和 MSER(Maximally Stable Extremal Regions)^[6]等特征。然而,不同的场景应选择适合自己的特征^[7]。考 虑到 ASIFT 特征点较 SIFT 特征还具有完全的仿射不变性^[8],MSER 特征区域对于灰度图像具有仿射不变 性和稳定性,且这两个特征之间具有一定的相互独立性和互补性,因而将 ASIFT 特征点和 MSER 特征区域 融合可以丰富大仿射场景的特征并提高特征的仿射不变性。但如果 MSER 特征数量过多,会出现多个特征 区域交叠而使得特征区域被重复描述的现象,而且多特征的提取与融合存在计算量大、处理时间长的问题。 因此,对最大稳定极值区域算法进行改进,通过对图像进行网格划分,利用极值区域的信息熵来剔除低丰富 度的极值区域,使得区域特征在图像中分布均匀。因此,本文针对空地一体化三维重建中特征匹配正确率低 的问题,提出了基于改进的 MSER 特征区域和 ASIFT 特征点进行匹配(大仿射场景的混合特征匹配)的方 法。并且在特征提取阶段^[9]利用图形处理单元(GPU)进行并行处理^[10],以减少算法的处理时间。

2 基本原理

2.1 方法概述

基于多特征融合的特征提取与匹配算法可分为3个阶段,即特征提取、特征匹配和 GPU 加速。1)特征 提取:首先提取图像的 ASIFT 特征点,然后提取 MSER 特征区域,并利用 SIFT 特征描述符对其进行描述^[11]。2)特征匹配:利用基于 KD 树的 BBF 算法对提取到的稳定特征进行匹配。3)GPU 加速:仿射尺度不 变特征变换算法利于并行化,使用 GPU 处理 ASIFT 特征点的提取。该流程如图1所示。





Fig. 1 Flow chart of mixed feature extraction and matching algorithm

2.2 特征提取

关于图像的局部不变特征点的检测算法有很多,ASIFT 特征提取算法对于图像尺度、亮度等均具有较高的稳定性;MSER 特征区域提取算法对于仿射变换具有较强的稳健性,但如果提取得到的区域较多,会存在特征区域重叠的现象。针对此问题,提出了基于网格划分的 MSER 改进算法,使得到的特征区域不但稳定而且分布均匀。本文在常用的 ASIFT 特征点和 MSER 特征区域的基础上,提出了一种混合特征提取与匹配方法。图 2(a)为航拍图像序列中的一幅,对其提取 ASIFT 特征,可视化结果如图 2(b)所示。图 3(a)为利用无人机航拍得到的图像序列中的一幅,对其提取 MSER 特征,可视化结果如图 3(b)所示;对其提取改进的 MSER 特征,可视化结果如图 3(c)所示。





图 2 航拍图像 ASIFT 特征提取。(a) 原图;(b) ASIFT 特征 Fig. 2 ASIFT feature extraction of aerial image. (a) Original image; (b) ASIFT features



图 3 航拍图像 MSER 特征提取。(a) 原图; (b) MSER 特征; (c) 改进后的 MSER 特征

Fig. 3 MSER feature extraction of aerial image. (a) Original image; (b) MSER features; (c) improved MSER features

2.3 ASIFT 特征提取

仿射尺度不变特征变换算法是在尺度不变特征变换算法的基础上引入了由仿射变换产生的模拟图像。 模拟不同视角的情况,该算法具有良好的抗仿射变换性能^[12]。该算法主要通过模拟所有可能的经纬度变化 来实现图像的仿射变换,在采样足够多、间隔足够小的情况下,分别对原始图像应用不同的旋转和切变,可以 得到多幅仿射变换后的图像,然后对这些模拟的图像进行 SIFT 特征点提取,就可以得到具有仿射不变性的 局部特征,即 ASIFT 特征。

假设图 2(a)图像上某一点的坐标为(x,y),通过对图像进行仿射变换处理,构建出模拟图像。变换公式为

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{S}(\lambda) \mathbf{R}(\psi) \mathbf{T}(\theta) \mathbf{R}(\varphi) \begin{pmatrix} x \\ y \end{pmatrix} + \mathbf{\Gamma}(e, f) ,$$
 (1)

式中 $R(\varphi)$ 表示光轴的经度变化引起的图像旋转; $T(\theta)$ 表示光轴的纬度变换引起的图像切变; $R(\varphi)$ 表示相 机绕光轴自转引起的图像旋转; $S(\lambda)$ 表示相机距物体的距离改变引起的图像缩放; $\Gamma(e, f)$ 表示物体中心距 光轴距离的改变引起的图像平移。本文中, 纬度 θ 的采样按照切变 $t = 1/\cos\theta$ 的等比数列来采样, t = 1, a, a^2 , ..., a^n , 公比 $a = \sqrt{2}$ 满足足够的经度, 采用 $n \ge 5$, $\theta_{max} \approx 79.8^\circ$, 能够满足实际的使用范围。经度 φ 的采样 范围为 $[0^\circ, 180^\circ)$, 采样取等差数列 $\varphi = 0, b, 2b, ..., kb$, 公差 $b \approx 72^\circ/t$ 且公差能够被 360° 整除, 保证有整数 个均匀分布的采样点, k 为满足 $kb < 180^\circ$ 的最大整数。

然后,利用 SIFT 算法提取模拟图像中的特征点,将特征点的坐标依据仿射变换关系换算到原始图像坐标系,将经纬度信息记录到特征点信息中,最后剔除边缘特征点。将其可视化,得到如图 2(b)所示的结果。

2.4 MSER 特征提取

MSER 特征是当使用不同的灰度阈值对图像进行二值化时得到的最稳定的区域。不同于 SIFT 和 SURF(Speed up Robust Features)特征^[13],利用 MSER 特征对区域的旋转和尺寸归一化,可以实现仿射不 变性,因此,MSER 特征对于获得较多大仿射场景下稳定的特征匹配对具有重要意义。

图 3(a)所示的图像 MSER 特征的提取步骤如下:

1) 灰度值排序。首先对图像进行灰度化处理,并采用快速排序算法对灰度值进行排序,得到排序完毕的 灰度值。 2)查找联通区域。利用合并查找算法来寻找图像中的联通区域,通常从灰度值的最小值、图像的左上角 开始搜寻。

3)确定极值区域。用 Q_i 表示任意连通区域的面积, i 表示从最小值开始变换的灰度值, 当灰度值在 [$i-\Delta$, $i+\Delta$]内变化时,连通区域的面积相应的变化为 $Q_{i-\Delta}$ 和 $Q_{i+\Delta}$,则在该变化区域内若满足 $q(i) = (Q_{i+\Delta}-Q_{i-\Delta})/Q_{i+\Delta} < T, T$ 一般取值为 0.25,则可初步认定该区域是稳定极值区域。之后需要将该原始图 像进行反转,假设图像上的像素点的灰度值可以用 p_i 表示,则将图像反转,即计算 $p'_i = p_{max} - p_i$,得到新的 图像灰度值,再重复 1)、2)、3)步的方法得到反转之后的极值区域。假设原始图像的面积可用 Q_{Image} 表示,若 $Q_i/Q_{Image} < a_{min}$ 或者 $Q_i/Q_{Image} > a_{max}$,说明该极值区域过小或者过大,将会被剔除。

2.5 改进的 MSER 特征提取

利用最大稳定极值区域算法进行特征区域提取,如果得到的区域较多,会造成多个特征区域互相交叠, 使得许多特征区域被重复描述,影响特征匹配效率。针对这个问题,本文提出了一种基于网格划分的改进算 法,通过对图像划分网格,筛选丰富度低的特征,保留信息熵高的特征^[14],并且根据特征分布离散度对特征 再次筛选,使得区域特征分布均匀,具体算法如下。

1)利用最大稳定极值区域算法检测图像中的特征区域,得到的特征区域数量为 M_{max}。确定图像中筛选 后的最大稳定极值区域数量为 0.8M_{max}。

2)依照图像大小划分网格,选择划分 3×3 的矩形规则网格,共9个单元格。依照极值区域中心点所在 位置确定每个单元格中所属极值区域的索引值,确定每个单元格包含的极值区域及数量 m_{ci}。

3) 对单元格内所有稳定极值区域计算区域内的像素灰度出现的概率值, 然后计算该区域对应的信息熵 H_i。并且利用(2)式计算该单元格的平均信息熵

$$\bar{H}_{ci} = \sum_{i=1}^{m_{ci}} H_i / m_{ci} \,.$$
⁽²⁾

4)按照新计算得到的每个单元格的平均信息熵,利用(3)式计算出每个单元格需保留的极值区域的数目

$$m'_{ci} = \bar{H}_{ci} M / \sum_{i=1} \bar{H}_{ci}$$
 (3)

5)根据 m_{ci} 与 m'_{ci} 的关系对特征区域进行剔除处理。如果 $m_{ci} \leqslant m'_{ci}$,说明该单元格内的极值区域数量较少,无需剔除。若数量大于 m'_{ci} ,则将该单元格内的极值区域依照信息熵排序,剔除信息熵小于 1/5 均值的极值区域。而后计算剩余极值区域的离散度,若该离散度大于设定的阈值,说明该单元格内的极值区域分布较均匀,此时需要判断筛选后的特征区域数量,若小于等于 m'_{ci} ,说明达到要求,停止计算。否则需要继续剔除信息熵小于 1/5 均值的极值区域,然后继续计算特征分布离散度,此处有一个迭代次数,若满足该迭代次数则停止计算,不再判断极值区域数量及特征分布离散度是否满足要求,直接进行下一步运算。

6)该单元格内计算完毕后,需要统计单元格内极值区域数量 mⁿ_{ci}与筛选前的特征区域预保留的数量 m^{'_{ci}} 的比值∂。若该值小于 90%,说明该单元格内被筛选掉的特征区域较多,将已筛选掉的特征区域按照信息熵 的大小排序后,取值较大的区域加入到该单元格内,使其数量满足最低要求。否则该单元格计算完毕,继续 完成其他单元格内的计算。待所有的单元格遍历完毕后,可以得到近似均匀分布的稳定极值区域,将其统一 进行特征描述,以便对区域进行匹配。分别利用 MSER 算法和改进的 MSER 算法对街景图像和航拍图像 进行特征提取,可视化结果如图 4、图 5 所示。

从以上两组图像中可以看出,筛选后的极值区域分布更加均匀,剔除了大量面积较大的椭圆区域和信息 熵较小的椭圆区域,说明基于网格划分的最大稳定极值区域算法能够获得较好的效果。

SIFT 特征向量不受尺度变化、旋转等几何变形因素的影响,因此将以上两种特征均采用 SIFT 特征描述符进行描述,这样不但结合了两种算法的优点,而且简化了计算,降低了算法的复杂度。

3 特征匹配

在获得局部稳定的极值区域和特征点后,需要进行特征匹配,具体匹配过程分为以下 4 个步骤: 1)利用基于 KD 树的 BBF 算法对参考图像和待匹配图像的极值区域进行粗匹配,遍历所有极值区域并



图 4 模拟街景图像筛选前后极值区域对比。(a) 原图; (b) MSER 特征; (c) 改进的 MSER 特征 Fig. 4 Contrast of extreme value regions before and after simulated street view image screening. (a) Original image; (b) MSER features; (c) improved MSER features



图 5 航拍图像筛选前后极值区域对比。(a) 原图;(b) MSER 特征;(c) 改进的 MSER 特征 Fig. 5 Contrast of extreme value regions before and after aerial image screening. (a) Original image; (b) MSER features; (c) improved MSER features

计算出匹配区域。

2)利用 RANSAC(Random Sample Consensus)算法去除误匹配^[15]。首先随机采样特征匹配对,估计 得到约束模型,本文采用单应矩阵作为几何约束模型,然后利用约束模型评价匹配对,得到新模型和正确匹 配对,进而去除错误匹配。

3)点特征去除误匹配。直接利用已经求得的单应矩阵来划分正确匹配点,因为在上一步中已经根据正确的匹配对求得了单应矩阵,所以就无须再估计两幅图像间的几何变换关系。

4)去除相似匹配对。大量的点特征是通过模拟图像中的特征点坐标转换到原始图像的坐标系获得的, 由于转换后大量特征点的坐标相近,可以看作同一特征点,因此需要去除相似匹配对。加上极值区域获得的 正确匹配对,最后得到了基于多特征融合的特征匹配结果,如图 6 所示。



图 6 多特征融合的特征提取与匹配结果 Fig. 6 Feature extraction and matching result of multi-feature fusion

4 并行优化

并行优化技术能够显著提高计算效率,降低算法计算成本。在传统的处理方法中,特征的提取均利用中央处理器(CPU)进行处理,而本文在特征提取阶段,根据两种特征的特点分别利用 GPU 和 CPU 进行处理。

光 学 授 报

基于网格划分的最大稳定极值区域算法由于逻辑判断及逻辑计算较多,利用 CPU 进行处理。而仿射尺度 不变特征变换算法,利用 GPU 进行并行处理。示意图如图 1 所示,从图 1 中可以看出,本文的算法使用了 CPU 和 GPU 协同处理,分线程同时执行,充分利用了 CPU 和 GPU 数据处理的差异,有效地提高了效率。

5 实验与分析

本文的实验环境为:计算机配置为内存 8 GB 的 64 位 Windows10 操作系统,VS2012 下运行。为了验证 本文提出的混合特征的有效性,对利用 ASIFT、MSER 和 ASIFT+改进 MSER 特征进行匹配的方法进行了 对比。实验用到的数据集是在中午和傍晚采集到的 25 张东北大学图书馆图像,如图 7 所示。其中 11 张是 利用无人机模拟街景车高度采集的图像,14 张是无人机采集的航拍图像,两种图像的视角差异在 50°~60°, 图 8 为利用不同算法对大仿射图像处理得到的正确匹配数,图 9 为利用不同算法得到的正确匹配率。由图 8 和图 9 可知,基于混合特征的方法比基于单一特征的方法得到的正确匹配对数更多,匹配率也更高。通过对比 本文方法与其他算法部分统计结果可知,本文算法基本上可以得到两种单一算法 2 倍以上的正确匹配对。



图 7 东北大学图书馆航拍图像与模拟街景图像





图 8 大仿射图像不同算法正确匹配数对比

Fig. 8 Contrast of correct matching numbers of large affine image with different algorithms

依据以上获得的大仿射场景的图像匹配成果可以进行三维重建。图 10 为东北大学图书馆航拍图像和 地面模拟采集图像进行空地一体联合建模重建的三维点云,从中可以看出,利用本文的算法能够在三维重建 中获得足够的正确匹配对,正确匹配对越多,重建后的三维模型精度越高,但是由于仿射变换较大,使用的图 像较少,重建的点云模型仍有不少空洞。在实验过程中,图像的处理耗费了大量时间,如果重建使用的图像



图 9 大仿射图像不同算法正确匹配率对比

Fig. 9 Contrast of correct matching rates of large affine image with different algorithms



图 10 东北大学图书馆空地一体联合建模稠密点云模型

Fig. 10 Dense point cloud model of Northeastern University library joint space-ground modeling

增加,处理时间势必也会大量延长,因此本文采用 GPU 加速的方式对多特征融合的算法进行优化。优化前 后运行时间对比如表1所示。对比结果表明,GPU 并行优化能够大幅度提升算法的效率,并且不影响精度。

Image number	Serial running time $/\mathrm{s}$	Parallel running time $/\mathrm{s}$	Acceleration ratio $/ \frac{0}{\sqrt{0}}$
A_01&S_01	224.55	14.78	15.19
A_02&S_02	229.53	15.21	15.09
A_03&S_05	226.08	14.93	15.14
A_04&S_01	226.35	15.60	14.51
A_05&S_04	231.70	15.91	14.56
A_06&S_04	227.37	14.00	16.24
A_08&S_08	226.21	15.19	14.89
A_08&S_09	226.87	14.81	15.32
A_10&S_10	231.15	14.05	16.45

表 1 串并行特征提取运行时间统计对比

Table 1 Contrast of running time statistics of serial-parallel feature extraction

6 结 论

本文介绍的基于多特征融合的特征提取与匹配算法虽然实现了在仿射、尺度等几何变换下的图像匹配, 并且比单一的特征提取与匹配算法得到更多的正确匹配对,但是仍然存在着一些需要改进之处。本文针对 灰度图像采用的最大稳定极值区域算法,利用的算法由于处理特征多、计算量大,只能对图像进行离线处理。 由于目前对于实景三维模型进行实时重建的需求与日俱增,且要求实现移动端场景的实时三维建模,因此如 何在减少特征数量的同时又保证正确的匹配对数量不至于大量减少,便成为日后研究的主要问题。

参考文献

- Liu Y H, Jia Q X, Sun H X, et al. Wide-baseline stereo matching based on multiple views [J]. Journal of Harbin Institute of Technology, 2010, 17(2): 225-228.
- [2] Wang Qiwei. Study on image histogram feature and application[D]. Hefei: University of Science and Technology of China, 2014.

汪启伟.图像直方图特征及其应用研究[D].合肥:中国科学技术大学,2014.

- [3] Moravec H P. Rover visual obstacle avoidance [C]. International Joint Conference on Artificial Intelligence, 1981: 785-790.
- [4] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [5] Morel J M, Yu G. ASIFT: A new framework for fully affine invariant image comparison[J]. SIAM Journal on Imaging Sciences, 2009, 2(2): 438-469.
- [6] Sun Jing. Research on the local invariant feature extraction of images and its application[D]. Dalian: Dalian University of Technology, 2009.

孙晶.图像局部不变特征提取技术研究及其应用[D].大连:大连理工大学,2009.

[7] Wang Gang, Sun Xiaoliang, Shang Yang, et al. A robust template matching algorithm based on best-buddies similarity
 [J]. Acta Optica Sinica, 2017, 37(3): 0315003.

王刚,孙晓亮,尚洋,等.一种基于最佳相似点对的稳健模板匹配算法[J].光学学报,2017,37(3):0315003.] Yu Xiaoyu, Guo Yubo, Chen Gang, *et al*. Real-Time point feature extraction based on connected components labeling

- [8] Yu Xiaoyu, Guo Yubo, Chen Gang, et al. Real-Time point feature extraction based on connected components labeling and distributed computing[J]. Acta Optica Sinica, 2015, 35(2): 0210001.
 于潇宇,郭玉波,陈刚,等. 基于点目标连通域标记的实时特征提取及其分布式运算[J]. 光学学报, 2015, 35(2): 0210001.
- [9] Bao Yongqiang, Ma Dongyang, Wu Zhongnan, et al. The comparison of two image registration algorithm(SIFT and ASIFT)[J]. Image Technology, 2013, 25(2): 50-52.

保永强,马东洋,伍中楠,等.两种图像配准算法(SIFT 与 ASIFT 算法)比较[J].影像技术,2013,25(2):50-52.

[10] He Tingting, Rui Jianwu, Wen La. Accelerating ASIFT based on CPU/GPU synergetic parallel computing [J]. Computer Science, 2014, 41(5): 14-19.

何婷婷, 芮建武, 温腊. CPU-GPU 协同计算加速 ASIFT 算法 [J]. 计算机科学, 2014, 41(5): 14-19.

- [11] Li Lingling, Li Cuihua, Zeng Xiaoming, et al. An automatic image registration method based on SIFT and Harris-Affine features matching [J]. Journal of Huazhong University of Science and Technology(Nature Science Edition), 2008, 36(8): 13-16.
 李玲玲,李翠华,曾晓明,等. 基于 Harris-Affine 和 SIFT 特征匹配的图像自动配准[J]. 华中科技大学学报(自然科学版), 2008, 36(8): 13-16.
- [12] Xu S, Peng Q C, He S. An invariant descriptor design method based on MSER[J]. Information Technology Journal, 2010, 9(7): 1345-1352.
- [13] Pang Y W, Li W, Yuan Y, et al. Fully affine invariant SURF for image matching [J]. Neurocomputing, 2012, 85: 6-10.
- [14] Choi H. Localization and regularization of normalized transfer entropy[J]. Neurocomputing, 2014, 139: 408-414.
- [15] Rothganger F, Lazebnik S, Schmid C, et al. 3D object modeling and recognition using affine-invariant patches and multi-view spatial constraints[C]. Proceedings of the 2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003: 272-277.