

文章编号: 2095-2163(2020)08-0093-04

中图分类号: TP391

文献标志码: A

无人机图像拼接技术的应用与探究

陈小青¹, 王力^{1,2}, 周娇¹

(1 贵州大学 大数据与信息工程学院, 贵阳 550025; 2 贵州工程应用技术学院 信息工程学院, 贵州 毕节 5551700)

摘要: 无人机因其灵活性高、体积小、时效性强、成本低、风险小等优点, 无人机遥感技术受到很多领域的青睐, 如农业、侦查情报、城市规划、航拍、灾难救援、测绘等。因单个图片无法满足实际需求, 为了扩大视野获取拍摄区域的全景图像, 常常需要将相邻的两幅图或者多幅图进行拼接。文章简单阐述了无人机图像拼接技术的研究现状及面临的问题, 主要对无人机图像的图像配准、图像融合技术方法进行了比较详细的阐述。最后就图像拼接技术未来发展与展望提及一些改进方向, 为图像拼接技术的更深层次探究提供一些参考。

关键词: 图像拼接; 图像配准; 图像融合; SIFT 算法

Application and research of UAV image mosaic technology

CHEN Xiaoqing¹, WANG Li^{1,2}, ZHOU Jiao¹

(1 College of Big Data and Information Engineering, Guizhou University, Guiyang 550025, China;

2 College of Information Engineering, Guizhou College of Engineering and Applied Technology, Guizhou 5551700, Bijie, China)

[Abstract] UAV remote sensing technology is favored in many fields because of its high flexibility, small size, high timeliness, low cost and low risk, such as reconnaissance information, urban planning, aerial photography, disaster relief, mapping and so on. However, a single image can not meet the actual needs. In order to expand the field of view to obtain the panoramic image of the region, it is often necessary to join two or more adjacent images. In this paper, the research status and problems of UAV image mosaic technology are briefly described, and the methods of UAV image preprocessing, image registration and image fusion are discussed in detail. Finally, the future development of image mosaics and some improvement direction is proposed to provide some reference for the image mosaics technology in a deeper level.

[Key words] image mosaic; image registration; image fusion; sift algorithm

0 引言

随着科技的不断进步和人们的不断需求, 无人机因其体积小, 便于携带, 灵活性强的优点, 近年来越来越受国内外研究者的青睐。无人机技术被用于许许多多的领域研究, 其中一个国内外的热点研究便是无人机图像拼接技术。图像拼接技术是指将两张或者两张以上相互有重叠部分的图像(可能是不同传感器、不同视角或者不同时间)拼接成一幅宽视野高分辨率的全景图像。图像拼接技术的两个重要环节是图像配准和图像融合。

1 图像拼接技术的研究现状

随着图像拼接技术方法的日趋完善和成熟, 许许多多相关的研究正在如火如荼的相继展开。其中基于无人机图像拼接课题一直受到国内外学者的青睐。无人机图像拼接技术主要包括图像预处理、图像配准和图像融合。由于图像配准和图像融合是图像拼接过程中的重要环节, 因此本文仅对图像配准

和图像融合过程做详细介绍。

David G. Lowe^[1] 提出了 SIFT (Scale Invariant Feature Transform, 尺度不变特征变换) 特征提取算法。对 SIFT 匹配算法进行改进, 使得该算法在国内外特征匹配算法领域占有很重要的位置^[2]。因此, 在基于 SIFT 的算法基础上衍生出来许多图像配准算法。KeY. Sukthankar^[3] 提出的 PCA-SIFT 算法; Herber Bay 等人^[4] 提出了 SURF (Speed Up Robust Features, 加速稳健特征) 特征提取算法; Morel J M 等人^[5] 提出的 ASIFT 算法; Fischler MA^[6] 提出了 RANSAC (Random Sample Consensus, 随机样本一致性) 快速图像匹配算法; Yingen 和 Kari Pulli 等人^[7] 提出了基于 SIFT 算法应用于手机上的快速图像拼接技术; 于建国等人^[8] 为了提高图像拼接速度提出的 ORB 和改进的 RANSAC 相结合的图像拼接算法。

近年来, 无人机图像拼接技术更是快速发展, 新

作者简介: 陈小青(1991-), 女, 硕士研究生, 主要研究方向: 图像处理; 王力(1971-), 男, 学士, 教授, 主要研究方向: 信息系统分析、设计与开发、数据挖掘等; 周娇(1994-), 女, 硕士研究生, 主要研究方向: 图像处理。

通讯作者: 陈小青 Email: 710742342@qq.com

收稿日期: 2020-06-19

方法新技术不断涌现,使得无人机图像拼接技术在很多领域都取得了广泛的应用。图像拼接技术主要流程如图1所示。

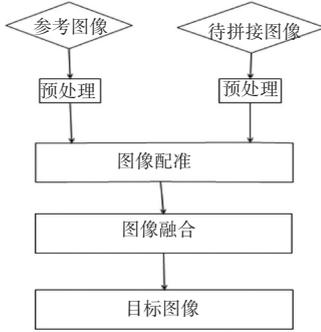


图1 图像拼接流程图

Fig. 1 Image mosaic flowchart

2 图像配准

图像配准的目的是通过几何运动模型找到图像之间的联系,将图像转换到统一的坐标系。配准算法技术依照算法原理可分为3类:基于特征信息的配准、基于变换域的配准、基于灰度值的配准。图像配准的算法有很多,如ORB算法、SURF算法和SIFT算法等。本文以SIFT(Scale Invariant Feature Transform,尺度不变特征变换)算法为例。SIFT算法常用于提取图像的局部特征和描述。SIFT算法可分为以下四步:

(1) 高斯尺度空间的极值检测。尺度空间是通过高斯函数与原图像做卷积计算得到的,定义为:

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

其中, $G(x, y, \sigma)$ 为尺度可变的高斯函数:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}. \quad (2)$$

σ 为尺度空间因子,是高斯正太分布的方差,代表原图像被高斯模糊的次数,描述的是图像细节特征,其值越小则表示原始图像被模糊的次数越少,也就是说图像的清晰程度越高;反之,其值越大则表示原始图像被模糊的次数越多,此时表示图像的概要特征,也就是说图像的清晰程度越低。

为了能更加高效地在尺度空间检测到稳定的特征点,提出了高斯差分尺度空间(Difference-of-Gaussian)DOG来检测极值点,表达式如下:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma) * I(x, y)) = L(x, y, k\sigma) - L(x, y, \sigma), \quad (3)$$

其中, k 指相邻两层的尺度因子比例,是一个常数,一般取 $\sqrt{2}$ 。尺度空间的建立也就是DOG图像

金字塔的生成过程,每一阶相邻尺度空间的高斯图像相减就得到了高斯差分图像。

为了找到高斯差分图像中的极值点,样本像素点需要与它同层的8个像素点和上下相邻两层中各9个像素点,总共26个像素点进行比较。如图2所示,当被检测的像素点在邻域内的26个像素点的DOG值是最大值或者最小值时,则样本像素点被视为极值点。可把这个极值点作为候选特征点提取出来,如若不是极值点则继续按此过程比较其它像素点。

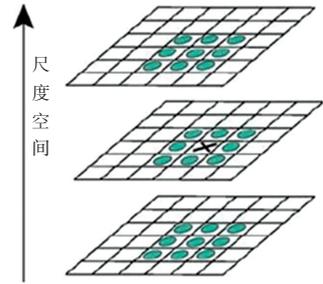


图2 尺度空间极值点检测示意图

Fig. 2 Extreme point detection in scale space schematic

(2) 特征点位置确定。由于在上述DOG尺度空间中提取的候选特征点并不稳定,它对噪声和边缘比较敏感。因此应将候选特征点中对比相对较低和位于边缘的点过滤掉。剔除低于对比度的特征点是将尺度空间函数在关键点按泰勒级数展开:

$$D(X) = D + \frac{\partial D^T}{\partial X} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X^2} X. \quad (4)$$

其中:

$$X = (x, y, \sigma)^T, \quad \frac{\partial D^T}{\partial X} = \begin{bmatrix} \frac{\partial D^T}{\partial x} & \frac{\partial D^T}{\partial y} & \frac{\partial D^T}{\partial \sigma} \end{bmatrix}, \quad (5)$$

$$\frac{\partial^2 D}{\partial X^2} = \begin{bmatrix} \frac{\partial^2 D}{\partial x^2} & \frac{\partial^2 D}{\partial xy} & \frac{\partial^2 D}{\partial \sigma} \\ \frac{\partial^2 D}{\partial yx} & \frac{\partial^2 D}{\partial y^2} & \frac{\partial^2 D}{\partial y\sigma} \\ \frac{\partial^2 D}{\partial \sigma x} & \frac{\partial^2 D}{\partial \sigma y} & \frac{\partial^2 D}{\partial \sigma^2} \end{bmatrix}. \quad (6)$$

对式(4)两边求导,令导数为0即可得到极值点:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2} \frac{\partial D}{\partial X}. \quad (7)$$

将式(7)代入式(4)得:

$$D(\hat{X}) = D + \frac{1}{2} \frac{\partial D^T}{\partial X}. \quad (8)$$

如果 $D(\hat{X}) \geq 0.03$ 的极值点为图像的极值点并保留, 反之剔除该特征点。

过滤边缘特征点可以借助 Hessian 矩阵。边缘点的特征是: 图像边缘处得到的主曲率结果较大, 边缘垂直方向上得到的主曲率结果相对较小。Hessian 矩阵定义如下:

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}, \quad (9)$$

其中, D_{xx}, D_{xy}, D_{yy} 分别代表不同方向上的二阶导数, 设 α, β 分别是 Hessian 矩阵 H 的最大和最小特征值, 使 $\gamma = \alpha/\beta$, 则有:

$$Tr(H) = D_{xx} + D_{yy} = \alpha + \beta, \quad (10)$$

$$Det(H) = D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta. \quad (11)$$

其中, $Tr(H)$ 是指矩阵 H 的迹, 表示的是矩阵 H 的主对角元素和; $Det(H)$ 是指矩阵 H 行列式的

$$\begin{cases} m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2}, \\ \theta(x, y) = \tan^{-1} \frac{(L(x, y+1) - L(x, y-1))}{(L(x+1, y) - L(x-1, y))}. \end{cases} \quad (14)$$

其中, (x, y) 是特征点; $m(x, y)$ 表示特征点 (x, y) 梯度模值; $\theta(x, y)$ 表示特征点 (x, y) 梯度方向; $L(x, y)$ 表示特征点 (x, y) 的尺度, 所用的尺度为每个关键点各自所在的尺度。以特征点为中心的领域窗口内采样, 并用直方图来统计领域像素的梯度方向。梯度直方图的范围是 $0-360^\circ$, 宽度为 10° , 总共 36 个。一个特征点可能会被指具有多个方向 (一个主方向, 许多的辅方向), 其中主方向是指直方图中梯度值的峰值对应的横坐标值。

(4) 生成 SIFT 特征描述符。为了使图像保持旋转不变性, 首先要将坐标轴的方向旋转至特征点的方向, 通过高斯窗口加权。如图 3 所示, 左图中在特征点为中心的领域内选取一个 8×8 大小的窗口, 并将 8×8 的一个大窗口细化分成 4×4 的 4 个子窗口, 特征点所在的行和列不选, 每个子窗口里面有 16 个小格子, 计算出选中领域区域内每个像素点的梯度值和方向。图中的每一个小方框表示一个像素点, 方框里面的箭头和长短表示该像素的梯度和方向大小。统计每一个子窗口里的所有像素的梯度值和方向, 生成右图中的 4 个描述符, 每个描述符有 8 个方向。因为特征点周围有 4 个这样的窗口, 所以这种方法就生成 $4 \times 4 \times 8 = 128$ 维特征向量, 也就是说一个关键点用 128 个数据来描述。左图中圆圈代

值。

$$\frac{Tr(H)}{Det(H)} = \frac{(\alpha + \beta)^2}{\alpha\beta} = \frac{(\alpha\beta + \beta)^2}{\alpha\beta^2} = \frac{(\gamma + 1)^2}{\gamma}. \quad (12)$$

当 $\alpha = \beta$ 时, 式 (12) 取得的值最小, 即公式 $(\gamma + 1)^2/\gamma$ 的值最小, 随着 γ 的增加, 上述式子的比值也随之增大。为了剔除边缘响应点, 常采用另外一种形式:

$$\frac{Tr(H)}{Det(H)} < \frac{(\gamma + 1)^2}{\gamma}. \quad (13)$$

将相应的参数代入式 (13) 中, 如若满足式 (13), 则表示关键点符合要求应该保留, 否则滤去。一般取 $\gamma = 10$ 。

(3) 特征点方向的确定。上述两个过程确定了特征点的尺度和位置, 满足了尺度不变性。为了让 SIFT 算子具备旋转不变性, 则需要对特征点分配方向。

表高斯加权的范围。

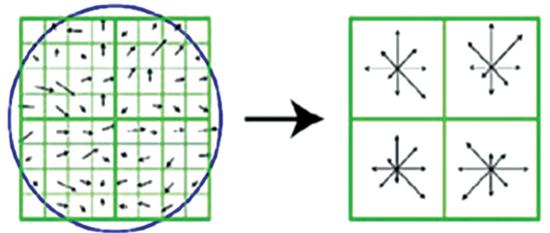


图 3 生成 SIFT 特征描述符

Fig. 3 Generation of SIFT descriptor

3 图像融合

图像融合的目的是通过处理图像相互重叠的部分达到拼接边界的平滑过渡, 从而让拼接效果尽量实现无缝拼接。图像融合之前还需要将图像配准之后的图像用 RANSAC (Random Sample Consensus, 随机样本一致性) 算法去除错误匹配点, 进而达到提纯的目的。图像融合是图像拼接的最后一个步骤, 其实就是将配准之后的图像进行全景图像呈现的过程。

3.1 图像融合的分类

按照图像融合的层次可以将图像融合分为三个层次: 像素级融合、特征级融合和决策级融合。

3.1.1 像素级融合

像素级融合是最低层次的融合, 在精准配准的

前提之下,对不同的图像直接进行信息融合的过程,也称像素融合为数据集融合。这种操作可以让原图像的信息尽可能保留,得到的图像可以有更多的细节信息。其优点是融合更加准确、可靠,但是数据量大,实时性差。

3.1.2 特征级融合

特征融合则先对图像的特征进行提取(包括边缘、方向、纹理、轮廓等),然后对提取的特征信息进行全面的处理和分析。这种融合是在中间层进行融合,优点是实时性能高,但是匹配精度有所降低。

3.1.3 决策级融合

决策级融合是图像融合里面的最高层,它的原理是基于认知。这种融合优点是决策能力强,但是得到的图像清晰度不高。

3.2 图像融合的算法

图像融合的方法有很多,如直接平均法、加权平均法和中值滤波法等,本文以加权平均融合法为例。加权平均融合法的过程是把两张图像重叠区域的像素灰度值,先进行加权计算,再对其进行叠加计算出平均值,最后得到新的图像。令 $f_1(x,y)$ 和 $f_2(x,y)$ 为待拼接的两幅图像, $f(x,y)$ 为融合以后的图像,则:

$$f(x,y) = \begin{cases} f_1(x,y), & (x,y) \in R_1; \\ \omega_1(x,y)f_1(x,y) + \omega_2(x,y)f_2(x,y), & (x,y) \in R_{12}; \\ f_2(x,y), & (x,y) \in R_2. \end{cases} \quad (15)$$

其中, R_1 表示参考图像中没有和待配准图像重叠的区域; R_2 表示待匹配图像中没有和参考图像重叠的区域; R_{12} 表示两幅图像重叠的区域; ω_1 和 ω_2 则分别代表两幅图像在它们重叠部分所对应像素的权值,并且 $\omega_1 + \omega_2 = 1, 0 < \omega_1 < 1, 0 < \omega_2 < 1$ 。

4 图像拼接技术面临的问题

随着国内外诸多研究和探索图像拼接方法的日益更新,这些拼接技术也越来越完善。但是,目前的图像拼接技术还存在以下问题。

(1) 大多算法仅仅针对某一个或者是某一类应用场景,还没有能够满足所有应用的服务。

(2) 绝大部分拼接方法研究重点是匹配精度,而忽略了匹配效率。提高匹配精度虽然能提高全景

图像的质量,但是降低了匹配的效率和。

(3) 对于数据量大、高分辨率的图像拼接处理效率会大大地降低,拼接技术还有待完善。

(4) 当待拼接的图像没有明显特征时进行拼接,会出现匹配错误等,如蓝天和海水。

5 结束语

图像拼接技术还有很多方面值得国内外学者继续研究探索:

(1) 目前拼接技术在配准算法上有很多丰富的经验,但是在分辨率比较高的图像处理时效率还是比较低。高效率的图像拼接技术有很大的实用价值和前景,如雷达目标跟踪、识别等。希望未来的图像拼接技术效率能够大幅度的提高,便于应用到生产技术中去。

(2) 在图像融合过程中优化算法,降低噪声的影响提高融合之后的图像质量。

(3) 创建一个实时高效的图像融合系统以便更好地应用到实际工作之中。

参考文献

- [1] Lowed G Lowe. Object recognition from local scale-invariant features [C]//IEEE Conference on Computer Vision, 1999: 1150-1157.
- [2] David G Lowe. Distinctive Image Features from Scale-Invariant Key points[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [3] Ke Y. Sukthankar. PCA-SIFT: A more distinctive representation for local image descriptors [J]. Proc. cvpr. volume, 2004: 506-513.
- [4] Bay H, Tuvtellars T, Gool L Van. SURF: Speeded Up Robust Feature[C]//Conference on Computer Vision, 2006.
- [5] ABDEL-HAKIM A E, FARAG A. CSIFT: A SIFT Descriptor with Color Invariant Characteristics[C]// IEEE Computer Society Conference on Computer Vision & Pattern Recognition. IEEE, 2006.
- [6] FISCHLER M A, BOLLES R C. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography [J]. Readings in Computer Vision, 1987, 24:726-740.
- [7] Yingen Xiong, Kar i Pulli. Fast Image Stitching and Editing for Panorama Painting on Mobile Phones [C]// IEEE Workshop on Mobileis ion, Conjunction with CVPR' 2010 (IWMV2010), San Francisco, 2010, IEEE Computer Society.
- [8] 余建国, 徐仁桐, 陈宁. 基于 ORB 和改进 RANSAC 算法的图像拼接技术 [J]. 江苏科技大学学报: 自然科学版, 2015, 29(2): 164-169.