

深度学习和APAP模型结合的大视差图像拼接算法

熊禹^{1,2}, 尹希庆^{1*}, 吴志运¹

¹广西科技大学自动化学院, 广西 柳州

²东风柳州汽车有限公司, 广西 柳州

收稿日期: 2023年3月18日; 录用日期: 2023年4月8日; 发布日期: 2023年4月20日

摘要

图像间位置偏差较大时, 图像拼接容易出现错位、重影问题。提出一种深度学习和APAP模型结合的大视差图像拼接算法, 该算法使用基于学习的SuperPoint网络同时提取图像特征点和描述符, 采用SuperGlue网络对特征点进行筛选和最优匹配, 最后通过APAP模型求取局部投影变换完成拼接。实际结果显示, 在Mikolajczyk数据集上, 基于学习的SuperPoint和SuperGlue网络在特征点提取和匹配方面相比传统算法重复率提升20%左右, 准确率达到99%, 鲁棒性更强, 准确率更高。最终拼接图相比传统算法图像质量评价指标NIQE降低6.5%左右, 基本消除错位、重影问题, 更符合视觉效果。

关键词

图像拼接, 深度学习, APAP模型, SuperPoint网络, SuperGlue网络

Large Parallax Image Mosaic Algorithm Based on Deep Learning and APAP Model

Yu Xiong^{1,2}, Xiqing Yin^{1*}, Zhiyun Wu¹

¹College of Automation, Guangxi University of Science and Technology, Liuzhou Guangxi

²Dongfeng Liuzhou Automobile Co., Ltd., Liuzhou Guangxi

Received: Mar. 18th, 2023; accepted: Apr. 8th, 2023; published: Apr. 20th, 2023

Abstract

When the position deviation between images is large, image Mosaic is easy to appear dislocation

*通讯作者。

and double image problems. A large parallax image stitching algorithm combining deep learning and APAP model is proposed. The algorithm uses learn-based SuperPoint network to simultaneously extract image feature points and descriptors, and uses SuperGlue network to screen and optimize the feature points. Finally, APAP model is used to obtain local projection transform to complete the stitching. The actual results show that on the Mikolajczyk dataset, the repetition rate of the SuperPoint and SuperGlue networks based on learning is improved by about 20% compared with the traditional algorithm in feature point extraction and matching, and the accuracy rate is up to 99%, with stronger robustness and higher accuracy. Compared with the image quality evaluation index NIQE of the traditional algorithm, the final Mosaic image is about 6.5% lower, basically eliminating the dislocation and double image problems, and is more in line with the visual effect.

Keywords

Image Mosaic, Deep Learning, APAP Model, SuperPoint Network, SuperGlue Network

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

图像拼接在视觉 SLAM、三维重建[1] [2]、全景图像拼接以及医学等诸多领域被广泛应用[3]，是数字图像处理技术和计算机视觉领域的重要技术。在实际应用中，以 Auto-Stitching 等为代表的图像拼接商用软件得到广泛应用，这些软件大多基于全局的对齐方法[4] [5] [6]，当相机拍摄角度位置发生较大偏差、不满足单点透视要求时，很难实现准确对齐，导致拼接图像发生模糊和错位。

解决该问题的关键在于拼接过程中的特征点检测和对齐能力两方面，国内外学者对这两方面进行了研究。Lowe [7]等提出了尺度不变特征变化(scale-invariant feature transform, SIFT)算法，通过构造差分金字塔的形式和利用高斯微分函数检测和描述特征点，该算法能够在图像旋转、尺度变化、光照变化和噪声等情况下提取出稳定的特征点，但存在误匹配。Bay [8]等提出了加速稳健特征(speeded up robust features, SURF)算法，通过积分图的形式检测特征点，在提取效率上对 SIFT 算法进行了优化，但在检测特征点的尺度不变性方面有所降低。Rublee [9]等提出 ORB (oriented fast and rotated brief)算法，该算法在特征点检测上使用快速检测(features from accelerated segment test, FAST)算法，并通过灰度质心法为特征点添加方向，在描述符生成上提出快速定向二进制描述符，有效加速了算法的效率，但该算法不能充分解决尺度不变性问题，且在光照变化较大时，提取的特征点效果不佳。Gao [10]等使用双单应性变换(dual-homography warping, DHW)的方法，

将图像分为远平面和地平面，对应区域分别投影，该方法相比全局投影方法在对齐能力上有所提升，但对大视差图像拼接效果依然较差。许越[11]等通过全局单应性先对图像进行对齐和网格划分，根据特征点分布调整变换矩阵，提高图像的对齐能力。Julio Zaragoza [12]等提出的 APAP (As-Projective-As-Possible Image Stitching)算法，利用传统的 SIFT 算法提取兴趣点，使用随机采样一致算法(Random Sample Consensus, RANSAC)筛选错误点，进行局部平面变换将图像划分为网格形式，通过移动线性变换(moving direct linear transformation, MDLT)对每个网格形变矩阵进行高斯加权，这种方法在很大程度上解决了由视差引起的全景图错位和重影问题，但对特征点和正确匹配率要求较高。Chang [13]等提出在重叠区域采用

投影变化,在非重叠区域采用相似变换的版投影变换方法,能够解决在重叠区域扭曲现象,但面对大视差图像时仍然需要结合 APAP 模型进行拼接。Lin [14]等在 APAP 模型的基础上提出了尽可能投影(Adaptive As-Natural-As-possible, AANAP)算法,将投影变换、相似变换和仿射变换相结合,使拼接的图像在视觉上效果更佳。文献[15]使用改进的 ORB 算法和 APAP 模型进行图像拼接,提升了拼接的效率,但 ORB 算法提取的特征点不具备尺度不变性,并且在光照变化较大时不能够提取出稳定的特征点。

以上对大视差图像拼接的研究,大多使用传统的特征点提取方法进行特征点检测或是对现有模型 APAP 模型进行改进,但 APAP 模型非常依赖提取点的准确性和重复性。近年来随着深度学习技术不断发展,基于深度学习的特征提取在鲁棒性、稳定性方面都优于传统特征提取,因此,本文提出一种深度学习和 APAP 模型结合的大视差图像拼接算法。在特征点提取上采用基于学习的 SuperPoint 网络[16]进行特征提取,使用基于学习的 SuperGlue 网络[17]寻找最优匹配点和剔除垃圾点,最后通过 APAP 模型进行局部投影变换进行拼接,利用加权融合方法[18]进行融合。该方法能够融合深度学习在特征点提取上的鲁棒性、正确匹配率优势,提高最后拼接图的质量。

2. 总体技术框架

基于学习和 APAP 模型结合的图像拼接算法流程图如图 1,分为图像采集、图像预处理、图像匹配和图像融合几部分。

图像预处理环节解决图像包含噪声、尺寸不一等问题,将输入的两张图像,先进行缩放处理,将大小统一为 640×480 ,其次对图像进行灰度化和高斯滤波处理,最后得到相应的灰度化图像。

图像配准包含特征点检测、特征点匹配、图像拼接三部分,是研究的重点。特征点检测使用基于学习的 SuperPoint 模型进行提取,对图像进行卷积池化,得到特征点和描述子。特征点匹配使用 SuperGlue 网络,通过在得分矩阵最后增加一层垃圾通道得到增产矩阵,将表现能力低的特征点进行剔除,将得分能力强、匹配率高的特征点进行保留,最小化特征点间的代价矩阵,得出最优匹配关系。最后在 APAP 模型中将原始图像进行网格划分,估计出图像间对应网格中局部投影变换模型,对局部变换模型进行高斯加权,最后通过局部变换模型实现图像的拼接。

为解决拼缝还使用加权融合算法对图像进行融合。

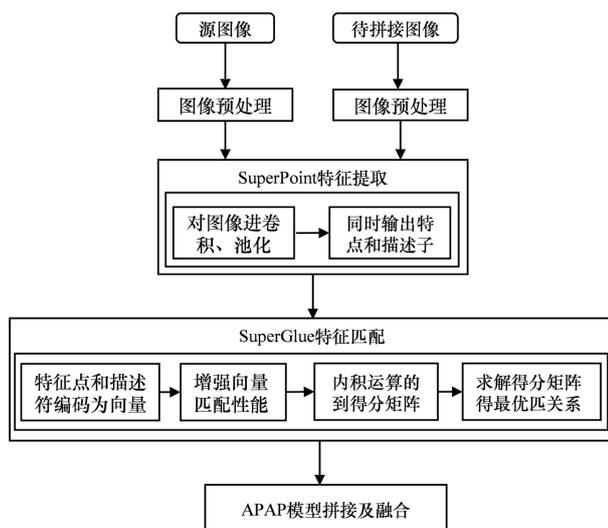


Figure 1. Overall technical flow chart of image mosaic

图 1. 图像拼接总体技术流程图

3. 基于深度学习的大视差图像拼接算法

3.1. 特征点检测

文使用 DeTone D [16]等设计的 SuperPoint 网络结构进行特征点检测，SuperPoint 是一种轻量级、端到端的特征点及描述子提取网络，网络框架主要由共享编码器和解码器两部分组成，如图 2 所示，其中解码器由兴趣点解码器和描述符解码器组成。

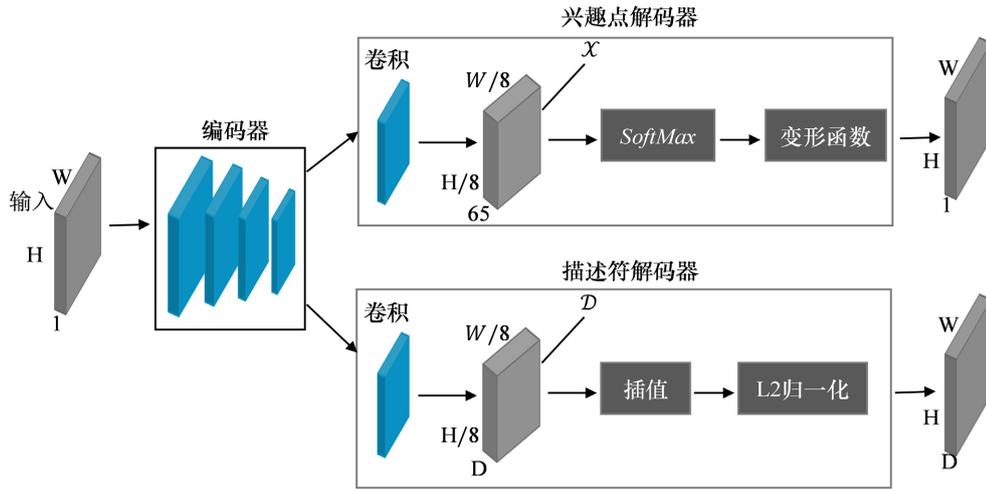


Figure 2. SuperPoint Network Structure

图 2. SuperPoint 网络结构

编码层采用 VGGNet [19]架构中的卷积层，当输入图像的尺寸为 $H \times W$ ，经过编码层的卷积、池化后输出的张量尺寸为 $H_c \times W_c$ ，其中 $H_c = H/8$ ， $W_c = W/8$ ，将低维输出的张量 $H_c \times W_c$ 组成的像素称为单元。编码层将输入的图像 $I \in R^{H \times W}$ 映射到较小空间尺寸和具较大深度通道 F 的中间张量 $B \in R^{H_c \times W_c \times F}$ ，其中 $H_c < H$ ， $W_c < W$ ， $F > 1$ 。

兴趣解码器将输入的 $\chi \in R^{H_c \times W_c \times 65}$ (65 代表 8×8 的局部区域外加一个额外的无兴趣点的通道)张量进行计算处理，输出尺寸为 $R^{H \times W}$ 的张量，并输出尺寸为 $H \times W \times D$ 的张量， D 表示描述符解码器的通道深度，然后解码器执行描述符的双三次插值，最后执行 L_2 标准化输出描述符。为了恢复为 $R^{H \times W}$ ，在特征点检测和描述符的解码器出均使用上采样。

3.2. 特征点匹配

SuperGlue 是通过卷积神经网络进行图像特征点匹配，对两张图像中提取到的特征点和描述子输入到网络中，输出两张图像之间的匹配关系，特征点主要是通过求解可微分最优传输问题来进行匹配。如图 3 为 SuperPoint 网络结构。

SuperGlue 网络主要由注意力神经网络和最优匹配层组成，如图 3 所示。其中注意力神经网络层由关键点编码和注意力聚合网络构成，关键点编码将关键点和描述子编码为一个向量，通过多层感知机 (Multilayer Perceptron, MLP) 将低维向量映射为高维向量。为了得到表现更强的描述符 f_i ，使用自注意力层和交叉注意力层交替更新。关键点位置和描述子的融合后的每个特征点 i 的初始表示为 x_{i0} 。

$$x_{i0} = d_i + MLP_{enc}(P_i) \quad (1)$$

其中 P_i 表示图像上第 i 个特征点， d_i 表示第 i 个描述子， MLP 表示多层感知机。

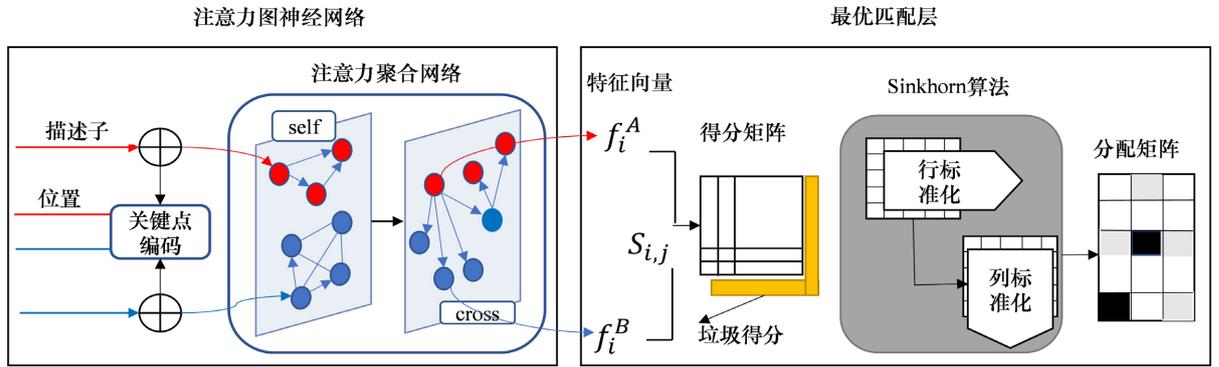


Figure 3. SuperGlue network structure
图 3. SuperGlue 网络结构

注意力聚合网络中模拟人来回浏览匹配过程，使用自注意力和交叉注意力交替聚合。在图像 I 上的更新公式可以表示为如下：

$$(l+1)x_i^l = lx_i^l + MLP\left\{\left[lx_i^l \parallel M_{\varepsilon \rightarrow l}\right]\right\} \quad (2)$$

其中 $\varepsilon \in \{\varepsilon_{self}, \varepsilon_{cross}\}$ 、 lx_i^l 表示图 I 上第 i 个特征点在 L 层的中间表达式， $M_{\varepsilon \rightarrow l}$ 表示自注意力和交叉注意力特征聚合，在图像 I' 上采用类似的方式进行更新。

在最优匹配层最大化总得分 $\sum_{i,j} S_{i,j} P_{i,j}$ ，得到分配矩阵 P ，其中 $S_{i,j}$ 为匹配分数值，对注意力聚合网络中得到的描述符 f_i 和 f_j' 进行内积运算得到匹配得分如式 3 所示。利用 Sinkhorn 算法[20]求解分配矩阵，得到最佳匹配。

$$S_{i,j} = \langle f_i^l, f_j^{l'} \rangle, \forall (i,j) \in I \times I' \quad (3)$$

3.3. APAP 模型拼接

完成特征点的提取和匹配后，使用 APAP 模型求取图像间的局部投影变换模型，然后完成拼接和融合。假设图像 A 和 B 为一组需拼接的图像，则它们检测出的特征点对为 $p = [x, y, 1]^T$ 和 $q = [x', y', 1]$ ，特征点对之间的对应关系为：

$$\begin{aligned} x' &= \frac{h_1 x + h_2 y + h_3}{h_7 x + h_8 y + h_9} \\ y' &= \frac{h_4 x + h_5 y + h_6}{h_7 x + h_8 y + h_9} \end{aligned} \quad (4)$$

齐次坐标 p 和 q 对应的变换关系表示为：

$$q \sim Hp \quad (5)$$

在上式中矩阵 $H \in R^{3 \times 3}$ ， $H = [h_1, h_2, h_3; h_4, h_5, h_6; h_7, h_8, h_9]$ ，由于 p 和 q 在同一个方向，满足 $o_{3 \times 1} = q \times Hq$ ，因此，将特征点 p 和 q 带入可求得矩阵 H ：

$$q \times H \cdot p = \begin{bmatrix} y'h_3 p - h_2 p \\ h_1 p - x'h_3 p \\ x'h_2 p - y'h_1 p \end{bmatrix} = o_{3 \times 1} \quad (6)$$

将上式可以转化为 $A \cdot h = o$ ，即

$$\begin{bmatrix} o_{3 \times 1} & p^T & y'p^T \\ p^T & o_{1 \times 3} & -x'p^T \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = o_{3 \times 1} \quad (7)$$

求出的 h 可以表示为:

$$h = \arg \min \sum_{j=1}^N \|Ah\|^2 \quad (8)$$

求出单应性矩阵 H 后, 通过 Moving DLT 算法对每个小方格中的局部单应性矩阵进行高斯加权, 则网格内的局部单应性矩阵可以表示为:

$$h_j = \arg \min \sum_{j=1}^N \|W_j Ah\|^2 \quad (9)$$

其中, j 为将图像划分为 $n \times n$ 的图像块的数量, W_j 为每个网格中的高斯权重, 得到每个网格中的局部单应性矩阵后, 通过变换参数完成图像拼接。

3.4. 图像融合

由于视差等因素, 图像配准后存在拼接痕迹, 因此, 采用加权平均融合算法[18]对图像进行融合。通过取权重的方式对重叠区域中待融合点赋予贡献值, 使拼接后的画面自然平滑过渡, 提升拼接图质量。加权融合算法公式如式 10 所示。

$$I(x, y) = \begin{cases} I_1(x, y) & (x, y) \in S_1 \\ w_1(I_1(x, y)) + w_2(I_2(x, y)) & (x, y) \in S \\ I_2(x, y) & (x, y) \in S_2 \end{cases} \quad (10)$$

其中 $I(x, y)$ 为融合后图像的灰度值, $I_1(x, y)$ 为左侧图像的灰度值, S_1 为 $I_1(x, y)$ 在图像上除去重叠区域后的区域, $I_2(x, y)$ 为右侧图像的灰度值, S_2 为在 $I_2(x, y)$ 图像上除去重叠区域后的区域, 拼接后重叠区域为 S , w_1 和 w_2 为重叠区域的一个权值, 并且 $w_1 + w_2 = 1$, $w_1 \in (0, 1)$ 。本文采用渐入渐出法计算权值, 权值计算公式如式 11 所示:

$$w_1 = -\frac{1}{D}d + 1 \quad (11)$$

其中, D 为两幅图像中重叠区域的长度, d 为重叠区域中待融合点与左边界点的距离。

3.5. 算法性能评价

对特征匹配中的匹配性能, 采用重复率和正确匹配率对算法性能进行评价, 重复率在整个匹配过程中占据重要位置, 重复率的高低以及稳定程度都会影响最终的匹配结果和特征点的鲁棒性。对重复性和正确率的性能测试都是基于 Mikolajczyk 标准数据集测试, 对于传统的匹配算法, 使用 SIFR、ORB、SUFR 算法提取特征点, 然后用快速最近邻算法进行匹配, 最后用随机抽样一致(random sample consensus, RANSAC)算法[21]进行提纯。

重复率的计算[22]定义如下。

$$\text{Repeatability\%} = \frac{\#\text{correspondences}}{\min(n_1, n_2)} \quad (12)$$

其中, $\#\text{correspondences}$ 为待匹配图像中对应的特征点个数, n_1 和 n_2 分别为源图像和待匹配图像中检出的特征点个数。

正确率的[22]定义如下:

$$\begin{aligned} \varepsilon_s &= \text{dish}(a_c \cdot H, b_c) = \sqrt{\sum_{i=1}^n [(a_c \cdot H)_i - (b_c)_i]^2} < \varepsilon \\ \text{MatchScore} &= \frac{\varepsilon_s}{\#\text{correspondences}} \end{aligned} \quad (13)$$

其中, MatchScore 为正确匹配率, ε_s 为正确匹配点的个数, #correspondences 为对应特征点的个数。 a_c 和 b_c 为源图像 A 和待拼接图像 B 中对应特征点的坐标, H 为图像 A 映射到图像 B 的投影变换模型, $a_c \cdot H$ 表示图像 A 映射到图像 B 中对应特征点的坐标, 对映射后的 $a_c \cdot H$ 和 b_c 求取欧式距离, 距离小于 ε ($\varepsilon = 0.2$) 的点为正确匹配的点。

为了评价拼接图的质量, 本文采用通用型无参考图像质量评价算法, NIQE (natural image quality evaluator, NIQE)算法[23]对拼接图的质量进行评价, NIQE 指标越小, 图像质量越高。NIQE 算法评价图像质量主要是对原图像进行图像特征提取, 通过提取的特征拟合多元高斯(MGV)模型, 计算模型参数 v 和 Σ , 计算公式如下:

$$f_x(a_1, a_2, \dots, a_k) = \frac{1}{(2\pi)^{k/2} |\Sigma|^2} \exp\left(-\frac{1}{2}(x-v)^T \Sigma^{-1}(x-v)\right) \quad (14)$$

其中 $f_x(a_1, a_2, \dots, a_k)$ 为提取出的图像特征, v 和 Σ 为 MG V 模型的均值和方差。再进行待评估图像特征的提取, 拟合多元高斯(MGV)模型, 计算模型参数 v' 和 Σ' , 通过计算评估图像与多元模型(MGV)之间的距离来对图像质量进行评价[23]:

$$D(v, v', \Sigma, \Sigma') = \sqrt{(v-v')^T \left(\frac{\Sigma + \Sigma'}{2}\right)^{-1} (v-v')} \quad (15)$$

其中 v, v', Σ, Σ' 为原图像和待评估图像的 MG V 模型均值和方差, NIQE 指标越小, 图像质量越高。

4. 实验结果及分析

4.1. 测试数据及环境

为了验证本文算法的有效性, 分为两部分进行实际验证, 分别为特征匹配性能和拼接性能, 并用传统的算法进行比较。在特征匹配中采用 Mikolajczyk 标准数据集进行匹配性能测试, 该数据集包含图像模糊(bikes), 光照变化(leuven), 图像压缩(ubc), 视点变化(wall), 在图像拼接性能中采用公用大视差图像拼接实验数据集进行测试, 在进行实验前, 将两组数据集图像的分辨率调整为 640×480 。

实验硬件测试环境为 Intel(R)Core(TM)i7-11800H 2.30GHz, 内存为 16.0GB, 编程平台为 PyCharm2020, 编程语言是 python 语言。并调用 OpenCV 等所需库。SIFR、SUF R 和 ORB 算法是调用 OpenCV 中相应的算法函数。

4.2. 特征匹配性能分析

对特征匹配中的匹配性能, 采用重复率和正确匹配率对算法性能进行评价, 并与传统 SIFT、SUF R 和 ORB 算法进行对比分析, 其中使用 Repeatability% 表示图像中提取体征点的重复率, 该值越稳定, 表示提取特征点鲁棒性越强。MatchScore 评价图像间的正确匹配率, MatchScore 值越大, 正确匹配率越高。

图 4 和图 5 为本文算法和传统算法提取出的特征点可视化图以及特征点匹配可视化图, 图像为 Mikolajczyk 数据集中模糊变化图像, 上面为原图像上特提的特征点, 下面为模糊图像上提取的特征点, 可以看出本文所提算法提取出的特征点是一些纹理剧烈变化、结构变化明显的点, 分布比较均匀。在匹配方面对不同数据集都能达到稳定的匹配效果。

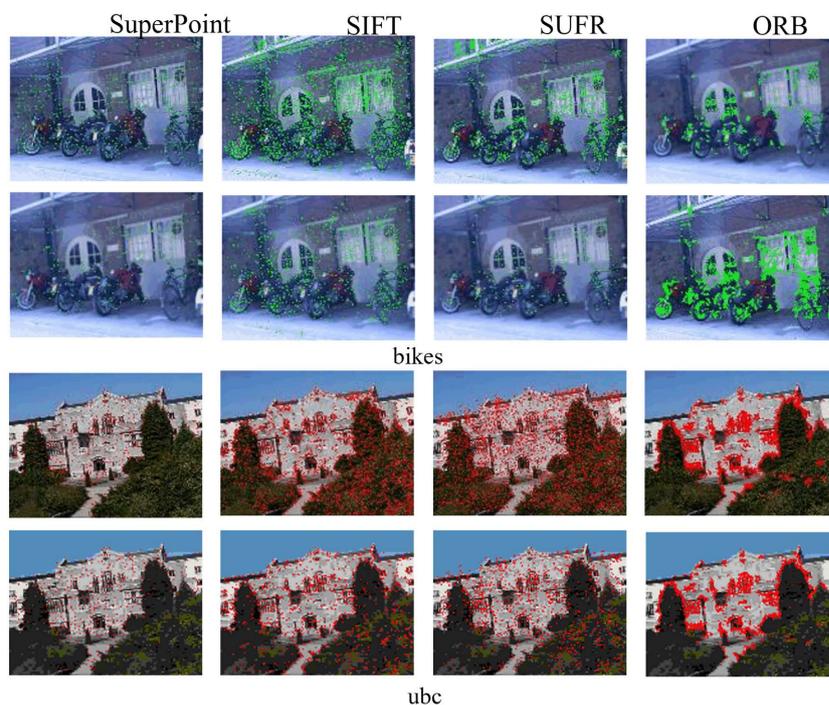


Figure 4. Visualization of feature points
图 4. 特征点可视化图

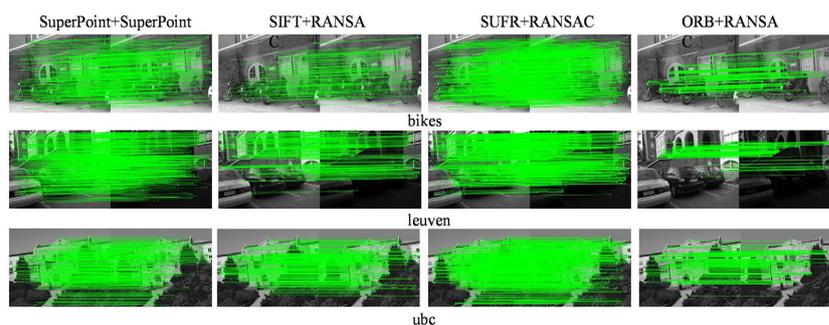


Figure 5. Visualization of feature point matching
图 5. 特征点匹配可视化图

如图 6 和图 7 为本文算法和传统匹配算法在重复率和正确匹配率的比较图，其中(a)、(b)、(c)及(d)分别为 Mikolajczyk 标准数据集中图像模糊变化(bikes)、图像光照变化(leuven)、图像压缩(ubc)和图像视点变化(wall)。1_2 至 1_6 分别表示第一张图和后面每张图像匹配。

从图 6(a)、图 6(b)以及图 6(c)可以看出，SuperPoint + SuperGlue 算法面对 bikes、leuven 以及 ubc 数据集时，提取出特征点的重复率都比较稳定，并且重复率也比较高，而传统的 ORB 算法重复率下降比较明显，SIFT、SUFR 算法在 bikes 和 leuven 数据集上重复率虽然能够维持稳定，但重复率相比本文算法要低 20%左右。在 ubc 数据集上，传统算法的重复率下降非常明显，本文算法依然维持稳定。在 wall 数据集上，本文算法重复率虽然有所变化，但相比传统算法本文算法依然趋于稳定，重复率高出传统算法 30%左右。说明本文算法相比传统特征匹配算法对不同场景下获取的图像都能提取出稳定的特征点，特征点的鲁棒性较强，在进行图像匹配时能够在一定程度上消除匹配误差，提高匹配的正确匹配率。

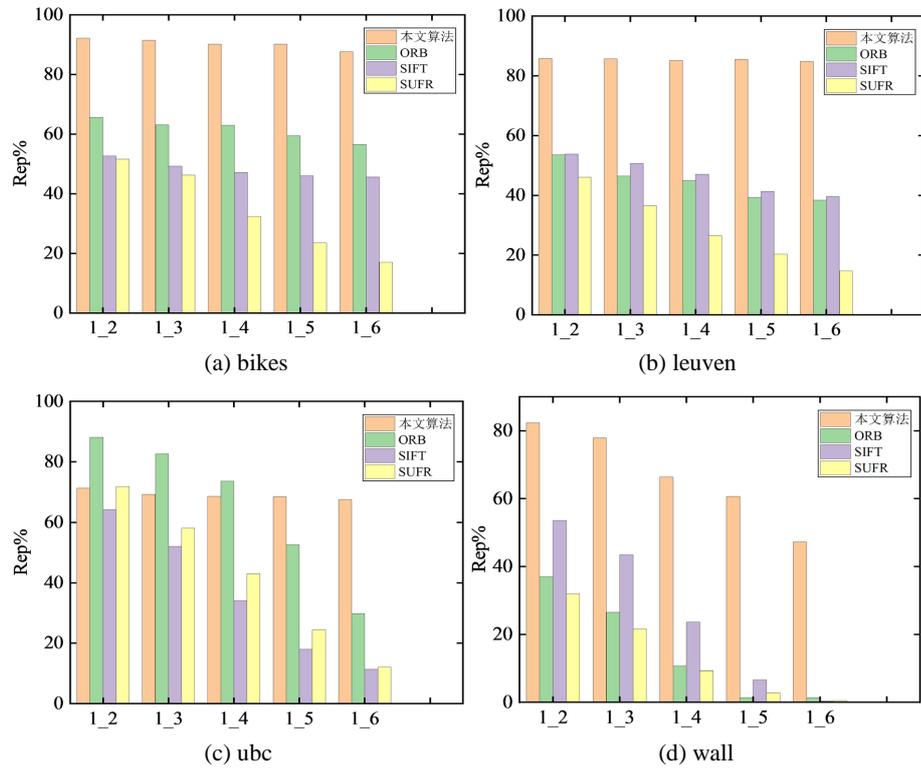


Figure 6. Comparison of repetition rates
图 6. 重复率比较

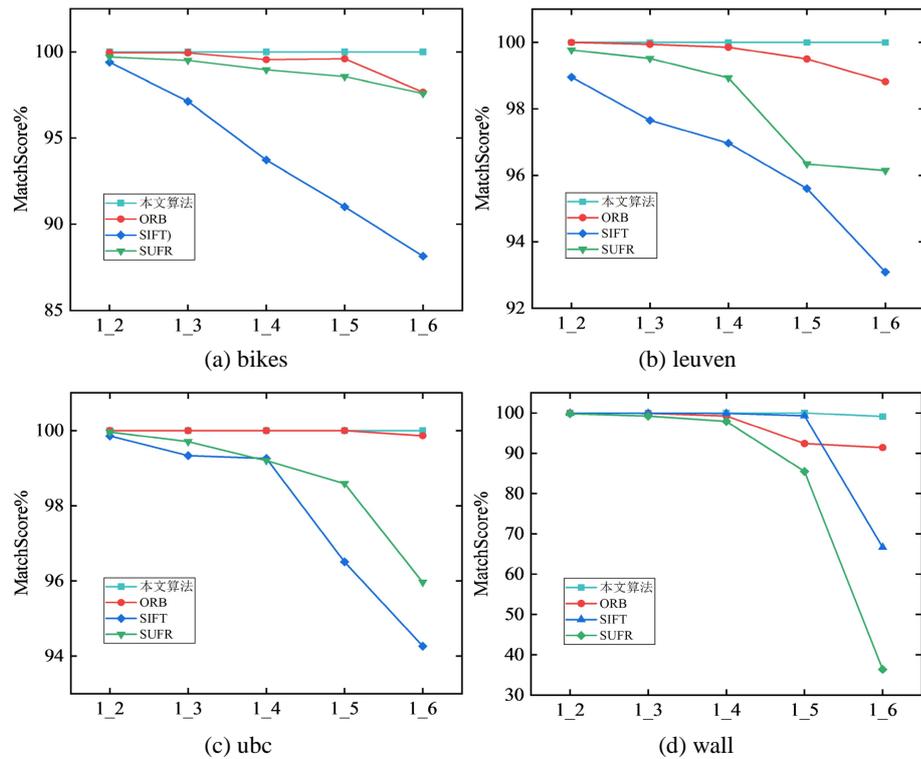


Figure 7. Comparison of correct matching rates
图 7. 正确匹配率比较

在正确匹配率方面，图 7 显示本文所使用算法在四种数据集上几乎能够达到 99% 的正确匹配率，匹配率维持稳定，变化不大，能够在不同场景下进行正确匹配，虽然在有些数据集上和传统算法正确匹配率相差不大，但整体上优于传统算法。可见本文所提算法的描述能力更强，匹配率更高。

4.3. 图像拼接性能

本文采用公共图像拼接数据集，通过 ORB、SIFT、SUFR 算法和本文的深度学习算法对待拼接的两幅图像进行特征提取和匹配，然后用 APAP 模型计算出图像间的局部投影变换模型，完成拼接。

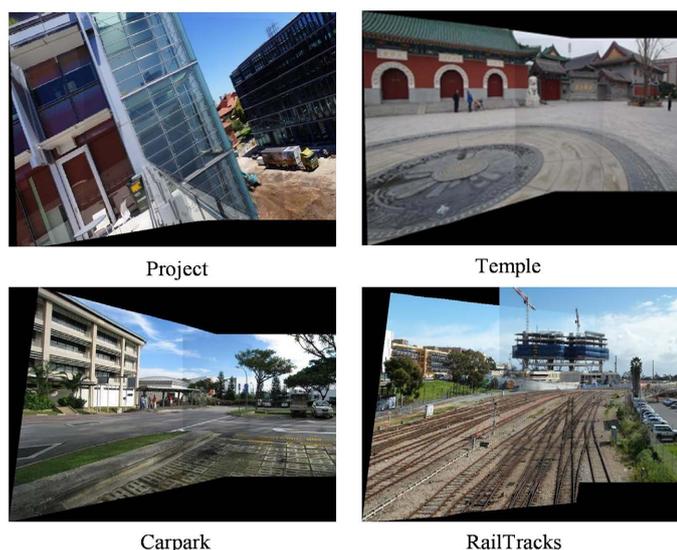


Figure 8. Splicing effect
图 8. 拼接效果图

如图 8 为本文算法用上面四种数据集图像拼接出的效果图，显示拼接出的效果图也符合人眼视觉效果，基本上消除了由于视差带来的重影问题。

Table 1. NIQE values of different algorithms

表 1. 不同算法的 NIQE 值

不同算法	不同类型数据集			
	Project	Temple	Carpark	RailTracks
ORB + RANSAC + APAP	6.900	7.215	8.077	7.357
SIFT + RANSAC + APAP	6.997	7.123	7.544	7.392
SUFR + RANSAC + APAP	7.946	7.230	7.821	7.142
本文算法	6.625	6.854	7.252	6.822

表 1 为本文算法和传统算法进行拼接后图像质量 NIQE 结果，可以看出，在进行大视差图像拼接时本文算法(SuperPoint + SuperGlue)在匹配的准确率和重复性方面都略优于传统的算法，在进行大视差图像拼接时，本文算法(SuperPoint + SuperGlue + APAP)最后拼接出的图像质量评价指标 NIQE 相比传统(ORB, SIFT, SUFR + RANSAC + APAP)算法拼接出的图像质量评价指标 NIQE 降低 6.5% 左右。

5. 结论

本文提出的深度学习和 APAP 模型相结合的大视差图像拼接方法, 采用 SuperPoint 网络进行特征提取和描述符, 用 SuperGlue 进行特征点筛选和最优匹配, 利用筛选出的特征点求取原图像和待匹配图像之间的全局单应性矩阵, 使用 APAP 模型将图像进行网格划分, 将全局单应性通过 Moving DLT 算法对每个小方格进行高斯加权, 求出图像间的局部投影变换模型, 最后完成拼接。

通过实验表明, 本文所提算法和传统的 ORB、SIFT 和 SURF 算法相比, 重复率高出传统算法 20% 左右, 在各种数据集上提取的特征点趋于稳定, 有很强的鲁棒性。正确匹配率方面也优于传统算法, 本文算法达到 99% 的正确匹配率。面对具有大视差的图像时, 拼接出的图像质量效果也较好, 基于深度学习的算法(SuperPoint + SuperGlue)在匹配的准确率和重复性方面都优于传统的算法。

基金项目

广西重点研发计划项目(2021AB01021)。

参考文献

- [1] Tafti, A.P., Baghaie, A., Kirkpatrick, A.B., *et al.* (2018) A Comparative Study on the Application of SIFT, SURF, BRIEF and ORB for 3D Surface Reconstruction of Electron Microscopy Images. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, **6**, 17-30. <https://doi.org/10.1080/21681163.2016.1152201>
- [2] Anderson, R., Gallup, D., Barron, J.T., *et al.* (2016) Jump: Virtual Reality Video. *ACM Transactions on Graphics*, **35**, Article No. 198. <https://doi.org/10.1145/2980179.2980257>
- [3] Lin, H.S., Chang, C.C., Chang, H.Y., *et al.* (2018) A Low-Cost Portable Polycamera for Stereoscopic 360 Imaging. *IEEE Transactions on Circuits and Systems for Video Technology*, **29**, 915-929. <https://doi.org/10.1109/TCSVT.2018.2818186>
- [4] 李加亮, 蒋品群, 夏海英. 基于网格变形和余弦函数权重的图像拼接方法[J]. 广西师范大学学报(自然科学版), 2020, 38(4): 42-53. <https://doi.org/10.16088/j.issn.1001-6600.2020.04.005>
- [5] 薛佳乐, 赵萌, 张哲, 程徐, 陈胜勇. 针对大视差图像拼接的显性子平面配准[J]. 中国图象图形学报, 2018, 23(3): 323-332.
- [6] 何川, 周军. 具有直线结构保护的网格化图像拼接[J]. 中国图象图形学报, 2018, 23(7): 973-983.
- [7] Lowe, D.G. (2004) Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, **60**, 91-110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [8] Bay, H., Ess, A., Tuytelaars, T., *et al.* (2008) Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*, **110**, 346-359. <https://doi.org/10.1016/j.cviu.2007.09.014>
- [9] Rublee, E., Rabaud, V., Konolige, K., *et al.* (2011) ORB: An Efficient Alternative to SIFT or SURF. 2011 *International Conference on Computer Vision*, Barcelona, 6-13 November 2011, 2564-2571. <https://doi.org/10.1109/ICCV.2011.6126544>
- [10] Gao, J., Kim, S.J. and Brown, M.S. (2011) Constructing Image Panoramas Using Dual-Homography Warping. *CVPR 2011*, Colorado Springs, CO, 20-25 June 2011, 49-56. <https://doi.org/10.1109/CVPR.2011.5995433>
- [11] 许越, 徐之海, 冯华君, 李奇, 陈跃庭, 徐毅, 赵洪波. 双场景类型遥感图像的配准拼接优化[J]. 浙江大学学报(工学版), 2019, 53(1): 107-114.
- [12] Zaragoza, J., Chin, T.J., Brown, M.S., *et al.* (2013) As-Projective-As-Possible Image Stitching with Moving DLT. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, 23-28 June 2013, 2339-2346. <https://doi.org/10.1109/CVPR.2013.303>
- [13] Chang, C.H., Sato, Y. and Chuang, Y.Y. (2014) Shape-Preserving Half-Projective Warps for Image Stitching. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Columbus, OH, 23-28 June 2014, 3254-3261. <https://doi.org/10.1109/CVPR.2014.422>
- [14] Lin, C.C., Pankanti, S.U., Natesan, R.K., *et al.* (2015) Adaptive As-Natural-As-Possible Image Stitching. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Boston, MA, 7-12 June 2015, 1155-1163. <https://doi.org/10.1109/CVPR.2015.7298719>

-
- [15] 屈玳辉, 谢益武. 基于 APAP 模型的大视差图像拼接算法[J]. 测控技术, 2021, 40(1): 34-39. <https://doi.org/10.19708/j.ckjs.2021.01.005>
- [16] DeTone, D., Malisiewicz, T. and Rabinovich, A. (2018) Superpoint: Self-Supervised Interest Point Detection and Description. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Salt Lake City, UT, 18-22 June 2018, 337-33712. <https://doi.org/10.1109/CVPRW.2018.00060>
- [17] Sarlin, P.E, DeTone, D., Malisiewicz, T., *et al.* (2020) Superglue: Learning Feature Matching with Graph Neural Networks. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Seattle, WA, 13-19 June 2020, 4938-4947. <https://doi.org/10.1109/CVPR42600.2020.00499>
- [18] 马东岭, 吴鼎辉, 毛力波, 石壮. 基于最优接缝线的加权融合影像拼接算法[J]. 测绘工程, 2020, 29(6): 1-7. <https://doi.org/10.19349/j.cnki.issn1006-7949.2020.06.001>
- [19] Simonyan, K. and Zisserman, A. (2014) Very Deep Convolutional Networks for Large-Scale Image Recognition. <https://arxiv.org/abs/1409.1556>
- [20] Schonberger, J.L. and Frahm, J.M. (2016) Structure-from-Motion Revisited. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, 27-30 June 2016, 4104-4113. <https://doi.org/10.1109/CVPR.2016.445>
- [21] Fischler, M.A. and Bolles, R.C. (1981) Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, **24**, 381-395. <https://doi.org/10.1145/358669.358692>
- [22] Mikolajczyk, K. and Schmid, C. (2005) A Performance Evaluation of Local Descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 1615-1630. <https://doi.org/10.1109/TPAMI.2005.188>
- [23] Mittal, A., Soundararajan, R. and Bovik, A.C. (2013) Making a “Completely Blind” Image Quality Analyzer. *IEEE Signal Processing Letters*, **20**, 209-212. <https://doi.org/10.1109/LSP.2012.2227726>