

文章编号: 2095-2163(2021)07-0060-06

中图分类号: TP391.41

文献标志码: A

基于多特征优选的图像拼接算法及系统设计

姜月武, 路东生, 党良慧, 杨永兆, 施建新

(上海工程技术大学 电子电气工程学院, 上海 201620)

摘要:以无人机为平台的桥梁巡检方案能够减少人工、避免安全隐患、提高效率,为了解决该系统中图像拼接实时性和拼接质量之间的矛盾,本文提出一种基于多特征优选的图像拼接算法,并设计开发出了一套基于无人机桥梁巡检的图像采集拼接子系统,先利用简单特征对图像流进行快筛,然后利用复杂特征来保持准确率,并提高系统的鲁棒性,最后通过性能指标对图像拼接质量进行调节。在实验中,针对处理速度、准确率和鲁棒性等各方面因素,本文算法平均达到了83.34%的特征提取准确率。在调节系统性能指标后,图像拼接取得了满意的效果,具有一定实用价值。

关键词:无人机;桥梁巡检;图像拼接;特征提取;图像拼接系统

Image stitching algorithm and system design based on multi-feature optimization

JIANG Yuewu, LU Dongsheng, DANG Lianghui, YANG Yongzhao, SHI Jianxin

(School of Electronic and Electrical Engineering, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] The bridge inspection program based on drones can reduce labor, avoid safety hazards, and improve efficiency. In order to solve the contradiction between the real-time image splicing and the splicing quality in the system, this paper proposes an image splicing algorithm based on multi-feature optimization. The algorithm has also designed and developed a set of image collection and splicing subsystems based on UAV bridge inspection. First, simple features are used to quickly screen the image stream, and then complex features are used to maintain accuracy and improve the robustness of the system. Finally, the image mosaic quality is adjusted through performance indicators. In the experiment, for the processing speed, accuracy and robustness and other factors, the algorithm in this paper reaches an average feature extraction accuracy of 83.34%. After adjusting the system performance indicators, image stitching has achieved satisfactory results and has certain practical value.

[Key words] UAV; bridge inspection; image mosaic; feature extraction; image stitching system

0 引言

图像拼接技术通过图像空间上的匹配、映射以及重叠区域融合将2幅或多幅图像进行拼接,获得包含原始图像所有信息的高分辨率、宽视角的合成图。该项技术在医学成像、计算机视觉、卫星数据、军事目标自动识别等领域具有重要的研究意义与使用价值。

目前,图像拼接技术按照是否使用卷积神经网络分为2类。一类是传统的方法:以灰度、尺度变化,判别相关性的基于区域相关的拼接算法,以Harris^[1]、FAST^[2-3]、SIFT^[4]、BRISK^[5]、SURF^[6]、FREAK^[7]、AKAZE^[8]、ORB^[9]等提取特征的基于特征的拼接算法,其中以AutoStitch^[10]为经典,适合没有视差或极小视差的场景;另一类是卷积神经网络方法:DHW^[11]、APAP^[12]、AANAP^[13]等以空间域变化绘制适合小视差场景的算法、Parallax-tolerant^[14]、seam-driven^[15]等以缝合线主导适合大

视差场景的算法。

与卷积神经网络算法相比,图像拼接的传统算法在准确率方面略显不足,但是在计算量和计算速度上则占据着优势。根据具体的无人机桥梁巡检项目实际运用在满足精度的同时还要兼顾系统实时性要求,本文提出了一种基于多特征优选的图像拼接算法及系统设计,经过实验仿真取得了满意的效果。

1 相关简介

1.1 Harris角点检测

使用角点检测算子,对图像的每个像素计算角点响应函数(Corner Response Function),阈值化角点响应函数,根据实际情况选择阈值,对阈值化的角点响应函数进行非极大值抑制,并获取非零点作为角点。通过一个小的滑动窗口在邻域检测角点的任意方向上移动窗口,若窗口内的灰度值都有剧烈的变化,则窗口的中心就是角点。

作者简介:姜月武(1995-),男,硕士研究生,主要研究方向:人脸检测、人脸表情识别、计算机视觉等。

通讯作者:姜月武 Email:jiangyuewu@sues.edu.cn

收稿日期:2020-12-20

假设图像 $I(x, y)$, 从点 (x, y) 处平移 $(\Delta x, \Delta y)$ 后的自相似性, 可由自相关函数表示:

$$c(x, y; \Delta x, \Delta y) = \sum_{(u, v) \in W(x, y)} w(u, v) (I(u, v) - I(u + \Delta x, v + \Delta y))^2 \quad (1)$$

$W(x, y)$ 是以点 (x, y) 为中心的窗口, 加权函数计为 $w(u, v)$ 。对图像 $I(x, y)$ 在平移 $(\Delta x, \Delta y)$ 后进行泰勒展开并进行一阶近似得到:

$$c(x, y; \Delta x, \Delta y) \approx \sum_w (I_x(u, v)\Delta x + I_y(u, v)\Delta y)^2 = [\Delta x, \Delta y] M(x, y) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \quad (2)$$

其中,

$$M(x, y) = \sum_w \begin{pmatrix} I_x(x, y)^2 & I_x(x, y)I_y(x, y) \\ I_x(x, y)I_y(x, y) & I_y(x, y)^2 \end{pmatrix} \quad (3)$$

最后可以将自相关函数简化近似为二项函数:

$$c(x, y; \Delta x, \Delta y) \approx A\Delta x^2 + 2C\Delta x\Delta y + B\Delta y^2 \quad (4)$$

其中, $A = \sum_w I_x^2, B = \sum_w I_y^2, C = \sum_w I_x I_y$

研究发现式 (4) 本质上就是一个椭圆函数。

$M(x, y)$ 的特征值 λ_1, λ_2 决定了椭圆的扁率大小和尺寸大小, $M(x, y)$ 的特征矢量决定了椭圆的方向, 如式 (5) 所示, 椭圆方程为:

$$[\Delta x, \Delta y] M(x, y) \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} = 1 \quad (5)$$

Harris 方法是通过计算一个角点的响应值 R 来判断角点, 并不需要计算具体的特征值。 R 的计算公式为:

$$R = \det M - \alpha (\text{trace} M)^2 \quad (6)$$

其中, $\det M$ 为矩阵 $M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 的行列式;

$\text{trace} M$ 为矩阵 M 的直迹; α 为常数, 其取值范围为 0.04~0.06。

1.2 FAST 特征点

图 1 即是 FAST 特征点的图像。假设以某点像素 p 为圆心, 半径大小为 3 做一个离散化的圆, 使得这个圆的边界上正好有 16 个像素。

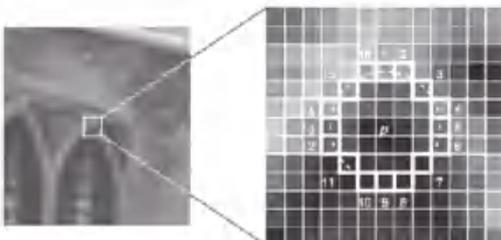


图 1 FAST 特征点的示意图

Fig. 1 Schematic diagram of FAST feature points

如果在上述 16 个像素点的圆上存在着 N (N 为 9 或 12) 个连续的像素点, 并且该像素值要么都大于 $I(p) + \varepsilon_d$, 要么都小于 $I(p) - \varepsilon_d$, 那么定义 p 为一个特征点, 如式 (7) 所示:

$$N = \sum_{z \in \mathcal{V}(\text{circle}(p))} |I(x) - I(p)| > \varepsilon_d \quad (7)$$

其中, $I(x)$ 表示圆环模板上的任意一点像素灰度值的大小; $I(p)$ 表示中心的像素点的灰度值大小; ε_d 为设定的阈值大小。

1.3 FREAK 特征检测

FREAK 采样模式如图 2 所示。图 2 中, 采用以特征点为中心的采样模式, 对每层圆环均匀采样 6 个点, 最终构成 7 层同心圆。为了获取足够多的图像信息让圆和圆之间发生交叠。使用采样点和特征点之间的距离来刻画采样点的密度, 采样点距离特征点越近, 采样点圆的半径就越小, 密度也就越大, 距离特征点若越远, 采样点圆的半径就越大, 密度也就越小。对每个采样点进行高斯平滑处理来降低噪声影响, 每个圆圈的半径刻画了高斯模糊的标准差。

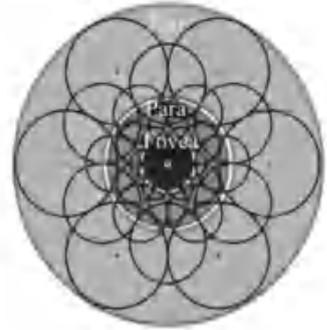


图 2 FREAK 采样模式

Fig. 2 FREAK sampling mode

FREAK 描述子由采样点对的强度比较结果级联组成, 形成二进制位串描述子。这里用 F 表示, 则有:

$$F = \sum_{0 \leq a < N} 2^a T(P_a)$$

$$T(P_a) = \begin{cases} 1, & (I(P_a^1) - I(P_a^2) > 0) \\ 0, & \text{其他} \end{cases} \quad (8)$$

其中, P_a 表示一个采样点, P_a^1 表示在经过高斯模糊后得到的采样点的灰度值大小。

1.4 SIFT 特征检测

图像的尺度空间用 $L(x, y, \sigma)$ 函数表示, 具体是由一个变尺度的高斯函数 $G(x, y, \sigma)$ 与图像 $I(x, y)$ 通过卷积产生, 即:

$$L(x, y, \sigma) = G(x, y, \sigma) \otimes I(x, y) \quad (9)$$

其中, \otimes 表示在 x 和 y 两个方向上进行卷积操

作,而 $G(x, y, \sigma)$ 为:

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \quad (10)$$

其中,决定着图像模糊平滑处理的程度的尺度因子为 σ 。 σ 值较大时代表的是大尺度下图像的概貌信息, σ 值较小时代表的是小尺度下图像的细节信息。因此大尺度对应着低分辨率,小尺度对应着高分辨率。 (x, y) 则表示在 σ 尺度下的图像坐标。

为了在连续的尺度下检测图像的特征点,需要建立 DoG 金字塔,而 DoG 金字塔的建立又离不开高斯金字塔的建立,如图 3 所示,左侧为高斯金字塔,右侧为 DoG 金字塔。

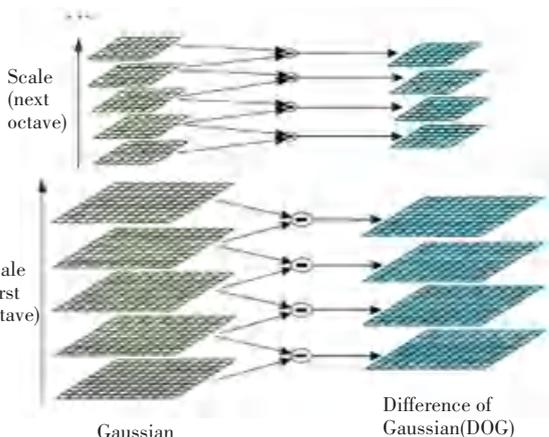


图 3 高斯金字塔和 DoG 金字塔

Fig. 3 Gaussian pyramid and DoG pyramid

1.5 SURF 特征检测

给定图像 I 中的某点 $\mathbf{x} = (x, y)$, 在该点 \mathbf{x} 处, 尺度为 σ 的 Hessian 矩阵 $\mathbf{H}(\mathbf{x}, \sigma)$ 定义为:

$$\mathbf{H}(\mathbf{x}, \sigma) = \begin{pmatrix} \hat{L}_{xx}(\mathbf{x}, \sigma) & \hat{L}_{xy}(\mathbf{x}, \sigma) \\ \hat{L}_{xy}(\mathbf{x}, \sigma) & \hat{L}_{yy}(\mathbf{x}, \sigma) \end{pmatrix} \quad (11)$$

其中, $L_{xx}(\mathbf{x}, \sigma)$ 是高斯二阶微分 $\frac{\sigma^2 g(\sigma)}{\sigma x^2}$ 在点

$\mathbf{x} = (x, y)$ 处与图像 I 的卷积, $L_{xy}(\mathbf{x}, \sigma)$ 和 $L_{yy}(\mathbf{x}, \sigma)$ 具有相似的含义,这些微分就是 LoG。

1.6 ORB 特征检测

ORB 算法结合了 Fast^[2-3] 和 Brief^[16] 算法,提出了构造金字塔,为 Fast 特征点添加了方向,从而使得关键点具有了尺度不变性和旋转不变性。具体流程描述如下:

(1) 构造尺度金字塔,金字塔共有 n 层,与 SIFT 不同的是,每一层仅有一幅图像。第 s 层的尺度为 $\sigma_s = \sigma_0^s$, σ_0 是初始尺度,默认为 1。

(2) 原图在第 0 层。第 s 层图像的大小如下:

$$SIZE = (H * \frac{1}{\sigma_s}) \times (W * \frac{1}{\sigma_s}) \quad (12)$$

在不同的尺度上利用 Fast 算法检测特征点,采用 Harris 角点响应函数,根据角点的响应值排序,选取前 N 个特征点,作为本尺度的特征点。

计算特征点的主方向,计算以特征点为圆心,半径为 r 的圆形邻域内的灰度质心位置,将从特征点位置到质心位置的方向做特征点的主方向。为了解决旋转不变性,将特征点的邻域旋转到主方向上利用 Brief^[16] 算法构建特征描述符,至此就得到了 ORB 的特征描述向量。

2 算法框架

图像拼接流程见图 4。由图 4 可知,图像拼接分为 4 个步骤:图像的匹配(Registration)、重新投影(Reprojection)、拼接缝合(Stitching)和图像融合(Blending)。

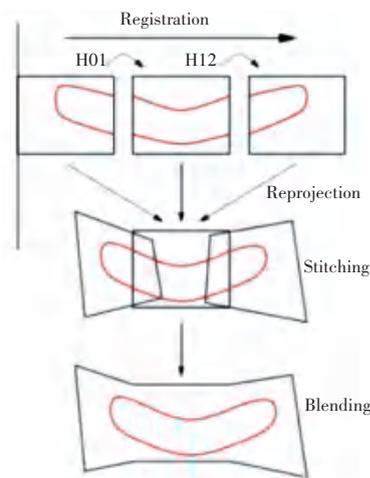


图 4 图像拼接流程

Fig. 4 Image stitching process

其中,图像匹配是指通过一定的匹配算法在 2 幅或多幅图像之间识别同名点。2 幅或者多幅图像可以由不同的时间、不同的位置拍摄,或者由多个传感器同时拍摄。重新投影是指通过图像的几何变换,把所有的图片都转换到一个共同的坐标系下。拼接缝合是指通过合并重叠部分的像素值,并保持没有重叠的部分使之生成更大分辨率的图像。图像融合是指通过几何和光度偏移错误通常导致对象的不连续,并在 2 个图像之间的边界附近产生可见的接缝。图像的配准程度决定了图像的拼接质量,因此拼接算法的核心和关键就是图像的配准。

本文提出的基于多特征优选的图像拼接算法的流程如图 5 所示。

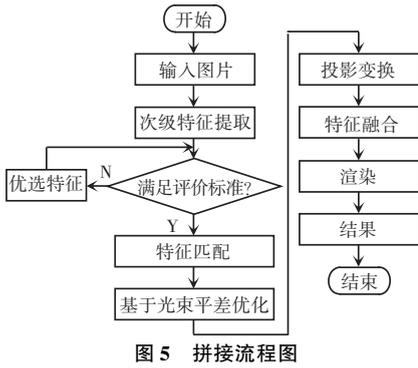


图 5 Stitching flowchart

其中,对于基于无人机拍摄的视频流数据,可以使用 Harris^[1]、AKAZE^[8]、FAST^[2-3]、SIFT^[4]、SURF^[6]、BRISK^[5]、FREAK^[7]、ORB^[9]等方法对相邻两帧之间进行特征提取,但是为了在拼接速度和拼接效果中找到平衡点,本文提出了评价标准,通过将传统特征的提取方法分为 2 类,按照特征点数量和拼接速度 2 个维度,将使用 Harris^[1]、AKAZE^[8]、ORB^[9]、FREAK^[7]算法进行特征提取的方法记为次级特征提取方法,将使用 BRISK^[5]、SIFT^[4]、SURF^[6]算法进行特征提取的方法记为优选特征提取方法,这样做的目的是可以使用调节评价策略(特征点、角点、斑点、帧间差分值),可以先通过次级特征点提取剔除一部分质量较差的图像或者对质量满足设定标准的图像直接进行拼接,这样就可以在实时性和最终的拼接效果中找到平衡点。针对具体的基于无人机桥梁巡检凭借项目,通过实验部分可以看到策略的鲁棒性。

3 实验结果

3.1 实验结果分析

无人机在预先设定飞行的区域范围内,随机抽取 5 个采集点的视频流数据,每个点取 20 帧相邻图片,对其中每相邻的 2 帧图片进行实验,分别将其分为 A 组和 B 组各 50 张图片,首先对每张图片使用不同的算法进行特征提取,然后记录 A 组和 B 组的每一对平均特征点数量,如图 6 所示。

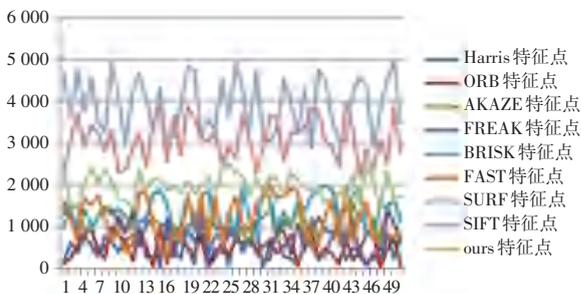


图 6 特征点提取数量对比

Fig. 6 Comparison of the extracted number of feature points

对于图 6 提取特征点,记录其所消耗时间如图 7 所示。

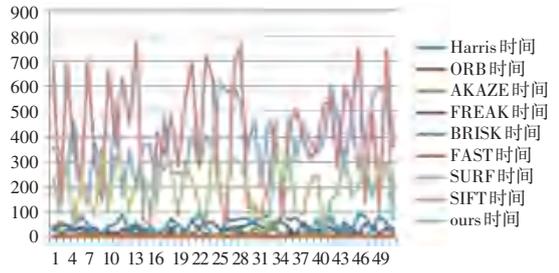


图 7 提取特征点消耗时间对比

Fig. 7 Comparison of time consumption to extract feature points

将每组图片中提取的最多的特征点作为基数,每一组中匹配成功的数量作为衡量标准,定义匹配成功率=最多匹配成功数量对/最大提取特征数,可以得到图 8。

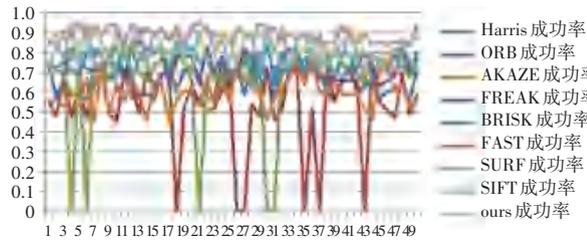


图 8 成功率对比

Fig. 8 Comparison of success rate

对于每种算法,统计了 50 对实验的总计消耗时间、提取的总特征数量和最多匹配点数量见表 1。

表 1 对比实验

Tab. 1 Comparative experiment

算法种类	总共时间/ ms	图片数量/ 对	特征点/个	最多匹配 点/对
Harris ^[1]	509	50	25 426	24 957
FAST ^[2-3]	228	50	20 562	20 458
SIFT ^[4]	20 352	50	152 637	149 774
BRISK ^[5]	2 534	50	68 418	67 542
SURF ^[6]	18 342	50	146 736	154 874
FREAK ^[7]	1 829	50	54 880	55 145
AKAZE ^[8]	1 176	50	42 351	42 754
ORB ^[9]	292	50	24 800	25 741
ours	9 984	50	99 841	97 945

通过实验,由表 1 得出结论:可以将实验的算法分为 3 个层次。第一层是速度最快,准确率和鲁棒性最差的,如 Harris、ORB、FAST;第二层是速度、准确率和鲁棒性适中,如 BRISK、FREAK、AKAZE;第三层是速度最慢,准确率和鲁棒性最好的,如 SIFT、SURF、ours。进而可推得:

(1) 速度比较: FAST>ORB>Harris>AKAZE>FREAK>BRISK>ours>SURF>SIFT。

(2) 准确率和鲁棒性: SIFT>SURF>ours>BRISK

>FREAK>AKAZE>Harris>ORB >FAST。

其中,针对快速提取特征算法,总体准确度和鲁棒性较差,但是对于个别实验组会出现优于慢速提取特征算法。

本文算法通过设计优选特征环节,一方面通过快速特征提取,筛除一部分数据,另一方面结合了准

确率和鲁棒性较高的特征来提高图像拼接质量,对于整个无人机桥梁巡检项目系统,算法实现了各项性能指标的综合平衡。

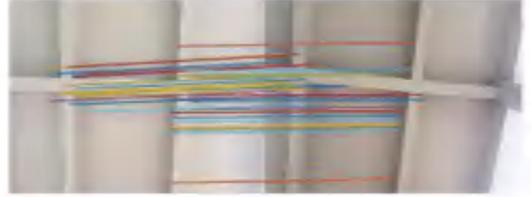
对于具体的实验效果,如图 9 所示,通过调节评价指标中的匹配特征点,可以控制拼接时间和图片的拼接质量。

36(18.95%) inliner matches out of 190



(a) 18.95%特征点匹配

40(16.19%) inliner matches out of 247



(b) 16.19%特征点匹配



(c) 18.95%特征点耗时 318 ms 融合结果



(d) 16.19%特征点耗时 303 ms 融合结果

图 9 实验结果对比

Fig. 9 Experimental results comparison

3.2 系统实现截图

本项目开发的系统界面及功能展示如下。

(1)登录界面。登录界面包含登录和注册功能,通过点击“注册”可以新建用户,输入已有账户的用户名和密码,点击“登录”就可以成功登录系统,登录系统界面如图 10 所示。



图 10 登录界面

Fig. 10 Login interface

(2)输入拼接图片。在菜单栏中点击“输入图像”按钮,可以通过文件路径选择需要拼接的图像,左边可以显示 6 张选择的图像,右边窗口显示选择图像的详细信息,系统界面截图如图 11 所示。



图 11 输入拼接图片

Fig. 11 Input stitched picture

(3)拼接进行中。在菜单栏中点击“拼接图像”按钮,系统后台会调用拼接算法,界面显示如图

12 所示,进度条进度显示了算法的运行进度。

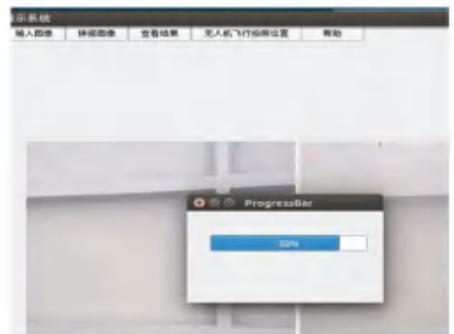


图 12 拼接进行中

Fig. 12 Stitching in progress

(4)拼接完成。当拼接算法运行结束时,系统界面会弹出如图13所示的“拼接完成”界面,提醒用户图像已经拼接完成。



图13 拼接完成

Fig. 13 Stitching completed

(5)拼接结果展示。在收到图像拼接完成的提示后,点击菜单栏的“查看结果”,系统会自动打开图像拼接结果查看器,界面截图如图14所示,在查看器菜单栏中,设置了对图像放大和缩小以及调整到和显示器合适的尺寸等功能,还可以对图片结果进行打开、保存等操作,只需要点击对应的功能按钮即可。

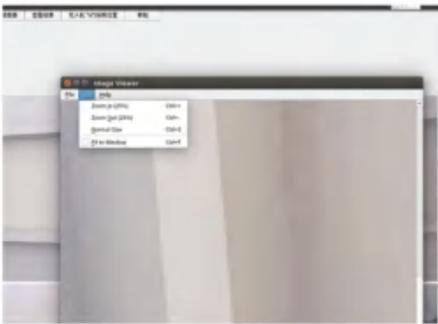


图14 拼接结果展示

Fig. 14 Stitched result display

4 结束语

在无人机桥梁巡检的实际项目中,对于采集的图像存在特征点不均、光影交替、图片抖动模糊和拼接时效性等问题。本文提出的基于多特征优选的图像拼接方案,在实际测试中体现出较好的鲁棒性,能够适应桥梁底面复杂的环境,证明了该算法可以在嵌入式设备良好运用。但是,该算法在精度上还需要进一步优化,包括特征点提取和匹配的优化。

参考文献

[1] HARRIS C G, STEPHENS M. A combined corner and edge

- detector [C] // Proceedings of 4th Alvey Vision Conference. Manchester England; University of Manchester, 1988; 189-192.
- [2] ROSTEN E, DRUMMOND T. Machine learning for high speed corner detection [C] // 9th European Conference on Computer Vision. Graz, Austria; Springer, 2006, 1; 430-443.
- [3] ROSTEN E, PORTER R, DRUMMOND T. Faster and better: A machine learning approach to corner detection [J]. IEEE transactions on pattern analysis and machine intelligence, 2008, 32(1): 105-119.
- [4] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [5] LEUTENEGGER S, CHLI M, SIEGWART R Y. BRISK: Binary robust invariant scalable keypoints [C] // 2011 International Conference on Computer Vision. Barcelona, Spain; IEEE, 2011: 2548-2555.
- [6] BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF) [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.
- [7] ALAHI A, ORTIZ R, VANDEREGHEYNST P. Freak: Fast retina keypoint [C] // 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA; IEEE, 2012: 510-517.
- [8] ALCANTARILLA P F, NUEVO J, BARTOLI A. Fast explicit diffusion for accelerated features in nonlinear scale spaces [C] // British Machine Vision Conference. Bristol; QIalcomm, 2013: 13.1-13.11.
- [9] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: An efficient alternative to SIFT or SURF [C] // 2011 International Conference on Computer Vision. Barcelona, Spain; IEEE, 2011: 2564-2571.
- [10] BROWN M, LOWE D G. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007, 74(1): 59-73.
- [11] GAO J, KIM S J, BROWN M S. Constructing image panoramas using dual-homography warping [C] // 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Colorado; IEEE, 2011: 49-56.
- [12] ZARAGOZA J, CHIN T J, TRAN Q H, et al. As-projective-as-possible image stitching with moving DLT [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36: 1285-1298.
- [13] LIN C C, PANKANTI S U, RAMAMURTHY K N, et al. Adaptive as-natural-as-possible image stitching [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA; IEEE, 2015: 1155-1163.
- [14] ZHANG F, LIU F. Parallax-tolerant image stitching [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA; IEEE, 2014: 3262-3269.
- [15] GAO Junhong, LI Yu, CHIN T J, et al. Seam-driven image stitching [C] // Eurographics (Short Papers). Spain; Eurographics, 2013: 45-48.
- [16] CALONDER M, LEPETIT V, STRECHA C, et al. Brief: Binary robust independent elementary features [M] // DANIILIDIS K, MARAGOS P, PARAGIOS N. Computer Vision - ECCV 2010. ECCV 2010. Lecture Notes in Computer Science. Berlin/Heidelberg; Springer, 2010, 6314: 778-792.