

doi:10.19306/j.cnki.2095-8110.2020.03.009

基于局部单应性矩阵的图像拼接与定位算法研究

迟龙云¹, 张 海¹, 赵晨旭²

- (1. 北京航空航天大学自动化科学与电气工程学院, 北京 100083;
2. 北京航空航天大学中法工程师学院, 北京 100083)

摘要:针对单个摄像机的视野范围有限导致在大场景下监控效果不够理想的问题,提出了一种改进的图像拼接与目标定位算法。该算法以多个摄像机获取的具有共视区域的监控图像为基础,通过对图像进行网格划分后分别计算多个局部单应性矩阵完成初步对准,然后对网格顶点进行微调优化完成最后配准。最后对图像进行融合形成无缝、自然的大视角图像,并利用场景信息在获取的全景图像上对目标进行快速定位,以满足监控人员对场景中目标的全景捕捉分析功能。实验结果表明,该算法能显著提高大场景下图像拼接结果的质量并实现目标的快速定位。

关键词:大场景; 图像拼接; 网格划分; 全景图像; 目标定位

中图分类号: TP751.1

文献标志码: A

开放科学(资源服务)标识码(OSID):

文章编号: 2095-8110(2020)03-0062-08



Local Homography Matrix Based Image Stitching and Location Algorithm

CHI Long-yun¹, ZHANG Hai¹, ZHAO Chen-xu²

- (1. School of Automation Science and Electrical Engineering, Beihang University, Beijing 100083, China;
2. Sino-French Engineer School, Beihang University, Beijing 100083, China)

Abstract: Due to the limited field of view of a single camera, the actual monitoring effect in large scenes is not ideal. An improved image stitching and target location algorithm is put forward for the problem. The algorithm is based on the surveillance images with overlapping regions acquired by multiple cameras. The initial alignment is completed by calculating local homography matrix after dividing the input images into uniform grids, and then the final alignment is achieved by fine-tuning the vertices of the grids. Finally, these images are fused to form a seamless and natural large-view image, and the scene information is used to rapidly locate the target in the panoramic image to satisfy the needs of monitoring and analyzing the target in the scene. The experimental results show that the proposed method can improve the quality of image stitching results effectively in large scenes and achieve fast target localization.

Key words: Large scenes; Image stitching; Grid partition; Panoramic image; Target localization

收稿日期: 2019-05-29; 修订日期: 2019-07-16

基金项目: 国家重点研发计划(2017YFC0821102, 2016YFB0502004); 北京市科技计划项目(Z171100000517006)

作者简介: 迟龙云(1993-), 男, 硕士, 主要从事图像处理及组合导航方面的研究。E-mail: lychi@buaa.edu.cn

0 引言

为了保证大型场景的安全稳定,视频监控作为一种安防的有效手段,已经被广泛应用在各种复杂的公共场合。由计算机代替人眼对视频进行分析处理,不仅可以减轻监控人员的劳动强度、提高监控识别的准确率,而且还可以提供许多人工监控不能实现的功能。目前的监控技术大多都采用一个固定安装的摄像机对某个场景进行监控。对于场景比较大的区域如机场、火车站、海关、停车场等,由于传统的摄像机的视野范围有限,不能同时获取周围环境中所有目标的位置,监控人员需要对着多个监控画面分别进行观察,对目标的监控定位只能限定在有限的摄像机监控范围内^[1],这样不仅丧失了对目标监控的连续性,更不利于从整体上把握分析监控目标的状态。因此,需要一种经济的广角图像生成技术实现对大场景的广角监控,以多摄像机拍摄图像为基础的图像拼接技术应运而生。

早期的图像拼接技术以 Brown 等^[2]的 Auto-Stitch 算法为代表,但是该算法有 2 个应用前提:拍摄到场景大致在一个平面上和拍摄图片时的光心位置必须相同。当拍摄场景或拍摄条件不符合以上情况时,拼接结果会产生明显的重影。Gao 等^[3]提出了基于双单应性矩阵假设的拼接算法,即假设待拼接图像可以分为背景和地面 2 个基本平面,通过计算 2 个平面对应的单应性矩阵,实现图像的精确配准。Lin 等^[4]使用仿射变换对齐图像,能够处理一定视差下的图像拼接问题。网格优化模型是一套灵活的图像优化框架,网格优化方法以 Zaragoza 等^[5]提出的尽可能投影(As Projective As Possible Image Stitching, APAP)算法为代表,将图像划分为均匀的网格,并分别为每个网格计算一个单应矩阵,使用局部单应性对齐图像。Lin 等^[6]提出了使用缝合线引导的局部单应对齐(Seam-guided Local Alignment, SEAGULL)方法,通过聚类产生多条局部单应性矩阵假设,然后计算最佳缝合线生成一系列的候选拼接结果,并提出了一种评价标准来评估缝合效果,最后选择得分最高的结果作为最终的全景图像。Chang 等^[7]提出了形状保护半投影(Shape-Preserving Half-Projective, SPHP)算法,通过引入结构保护方法,使得图像的非重叠区域到重叠区域,由相似变换逐渐过渡到透视变换,不仅可以使重叠区域准确对齐,并且能最大限度地减少非

重叠区域的投影失真。Chen 等^[8]提出的具有全局相似性先验的自然图像拼接(Natural Image Stitching with the Global Similarity Prior, NISwG-SP)算法使用 APAP 初始化网格,并同时利用局部相似项和全局相似项对网格优化结果进行约束,减小了投影失真,大大提升了多张图像的拼接效果。

针对存在视差的图像拼接问题,现有算法虽然能得到一定的拼接结果,但是仍会出现重影、鬼影和形状扭曲变形等问题。因此,需要研究全局配准方法的局部优化问题,尽可能地消除投影畸变,实现重叠区域的平滑过渡。

论文基于 NISwGSP 算法进行改进,针对该算法特征匹配数量较少并且分布不均匀、特征稀疏区域配准方法不合理及直线结构容易遭到破坏等问题,对特征点筛选算法加以改进,利用超像素区域辅助网格匹配点的筛选,并对网格优化过程添加直线保持项,以实现部分重叠区域的合理配准。最后采用最佳缝合线的方法避免了未配准区域出现重影,提升了大视差复杂场景下的拼接性能并保持了拼接结果的合理性。在由上述方法得到大场景下的全景图像后,调用最优摄像机配置资源实现对图像中目标的快速定位,方便监控人员从整体上直观把握前景目标的位置。

1 基于网格变形的图像拼接算法

1.1 特征点匹配与筛选

该算法首先对所有待拼接的图像进行快速特征点提取和描述(Oriented FAST and Rotated BRIEF, ORB)^[9]并进行暴力匹配得到候选匹配信息。传统的随机抽样一致性(Random Sample Consensus, RANSAC)算法^[10]针对纯旋转条件下的错误匹配点剔除十分有效,但是当拍摄 2 张待拼接图像的相机光心不再重合时,图像之间对应的转换关系不能单纯地利用一个单应性矩阵来表达,该算法便无法获得足够多的正确匹配点对。针对此问题,本文利用鲁棒的多平面 RANSAC 算法^[11]实现了大视差图像间的匹配点对的有效筛选。通过计算多组可能的假设模型,每组假设模型对应一个全局单应性矩阵或全局基础矩阵,找到使得匹配点投影误差小于阈值数量最多的假设模型即为最终合理的模型。模型对应的投影残差小于阈值的候选匹配点即为筛选后的匹配点。经过筛选后的正确匹配点为后续的网格匹配点生成及图像配准过程提供支持。

1.2 网格匹配点生成与筛选

为了使图像匹配点位置分布更加均匀,在得到正确匹配特征点后,将输入图像分别划分为均匀的网格,然后用 APAP 算法将重叠区域的网格顶点利用局部单应性矩阵投影得到相应的网格匹配点。

假设具有待拼接的图像 I_i 和 I_j , 分别将源图像 I_i 和目标图像 I_j 划分为 $M \times N$ 个均匀的网格。假设 $\{\mathbf{V}_{ik}\}_{k=1}^{(M+1) \times (N+1)}$ 和 $\{\mathbf{V}_{jk}\}_{k=1}^{(M+1) \times (N+1)}$ 分别为图像 I_i 和 I_j 的网格顶点集。 M^{ij} 为顶点 \mathbf{V}_{ik} 与该点经过局部单应性矩阵投影后的点组成的匹配点集合,用来描述图像 I_i 中的网格点投影到图像 I_j 后得到的匹配关系,形式如下

$$M^{ij} = \{(\mathbf{V}_{i1}, P(\mathbf{V}_{i1})), (\mathbf{V}_{i2}, P(\mathbf{V}_{i2})), \dots, (\mathbf{V}_{ik}, P(\mathbf{V}_{ik}))\} \quad (1)$$

其中, $P(\cdot)$ 为网格顶点经局部单应性矩阵投影后的图像坐标,并且 $P(\mathbf{V}_{jk})$ 位于图像 I_j 内。同理定义集合 M^{ji} 为对于顶点 \mathbf{V}_{jk} 满足 $P(\mathbf{V}_{jk})$ 位于图像 I_i 内的匹配点对集合,即 $(P(\mathbf{V}_{jk}), \mathbf{V}_{jk}) \in M^{ji}$ 。以 $M^{ij} \cup M^{ji}$ 表示图像 I_i 和图像 I_j 中所有区域初步得到的网格点匹配对集合。

考虑到在特征点较少的区域,局部单应性矩阵会退化为全局单应性矩阵,导致某些生成的网格匹配点不符合真实特征的分布特性,因此利用 SEEDS 超像素分割算法^[12]对区域进行划分完成对网格匹配点的筛选。假设经过筛选后的特征匹配点集合 $C = \{(\mathbf{p}, \mathbf{q}) \mid \mathbf{p} \in I_i, \mathbf{q} \in I_j\}$, 图像 I_i 经过超像素分割获得的区域集合为 $\{S_{ik}\}_{k=1}^M$, $\{B_{ik}\}_{k=1}^M$ 为该区域对应的置信度集合,其定义为

$$B_{ik} = \begin{cases} 1 & \exists (\mathbf{p}, \mathbf{q}) \in C \text{ and } \mathbf{p} \text{ in } S_{ik} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

当区域 S_{ik} 内没有匹配特征点时,该区域的置信度 B_{ik} 为 0, 否则该区域的置信度为 1。同理可以得到图像 I_j 的区域集合 $\{S_{jk}\}_{k=1}^M$ 和置信度集合 $\{B_{jk}\}_{k=1}^M$ 。

在得到所有图像的区域集合和置信度集合后,利用这些信息对上文得到的网格点匹配集合 M^{ij} 和 M^{ji} 进行筛选。假设集合 M'^{ij} 为经过筛选的网格点匹配对集合,其元素 $(\mathbf{V}'_{ik}, P(\mathbf{V}'_{ik}))$ 满足

$$(\mathbf{V}'_{ik}, P(\mathbf{V}'_{ik})) = \begin{cases} (\mathbf{V}_{ik}, P(\mathbf{V}_{ik})) & \text{if } \mathbf{V}_{ik} \in S_{ij} \text{ and } B_{ij} = 1 \\ \text{null} & \text{if } \mathbf{V}_{ik} \in S_{ij} \text{ and } B_{ij} = 0 \end{cases} \quad (3)$$

同理可得筛选后的集合 M'_{ji} , 最终利用相对均

匀的网格匹配点 $M'_{ij} \cup M'_{ji}$ 代替特征点匹配集合 C 进行后续的网格优化。

1.3 全局尺度项与全局旋转项计算

全局尺度项:通过网格匹配点生成阶段 APAP 算法的计算,每个网格均可以求取一个局部单应性矩阵,进而得到相机的焦距估计值^[13-14]。最后选择焦距值的平均值 \bar{f}_i 作为图像 I_i 的焦距估计值。令 \bar{f}_0 为选定的基准图像的焦距估计值平均值,则图像 I_i 的全局尺度项 s_i 为

$$s_i = \frac{\bar{f}_0}{\bar{f}_i} \quad (4)$$

全局旋转项:对于 2 幅具有重叠视野的图像 I_i 和 I_j , 利用直线检测 (Line Segment Detector, LSS) 算法^[15]分别对 2 幅图像的重叠区域进行直线检测得到直线集合 $\{L_i\}$ 和 $\{L_j\}$ 。对于直线集合 $\{L_i\}$ 中的直线 l_{ik} , 若该直线的长度和宽度大于一定的阈值,则该直线相对于水平线的角度 θ_{ik} 参与图像 I_i 相对水平线的旋转角度 θ_i 的计算^[16];若参考图像的旋转角度为 θ_0 , 则图像 I_i 的全局旋转角度 $\hat{\theta}_i$ 计算方式如下

$$\hat{\theta}_i = \theta_0 - \theta_i \quad (5)$$

1.4 基于网格优化的图像配准方法

为了提高图像拼接结果的自然性及准确性,算法根据不同的规则通过对网格添加约束项的方法对每个网格的顶点坐标进行微调,以提升图像配准的精度。令 $\{\mathbf{V}_i\}$ 和 $\{E_i\}$ 表示图像 I_i 经过网格点生成得到的网格顶点集合和网格边的集合, $\{\mathbf{V}\}$ 表示所有待拼接图像的所有顶点的集合。网格优化算法的目标是寻找变形后的最优网格顶点集 $\{\hat{\mathbf{V}}\}$ 使得能量函数 $\Psi(\mathbf{V})$ 达到最小。因此,网格优化问题可以表示成如下形式

$$\underset{\mathbf{V}}{\operatorname{argmin}} \Psi(\mathbf{V}) = \underset{\mathbf{V}}{\operatorname{argmin}} (\Psi_a(\mathbf{V}) + \lambda_1 \Psi_1(\mathbf{V}) + \lambda_g \Psi_g(\mathbf{V}) + \lambda_s \Psi_s(\mathbf{V})) \quad (6)$$

在式(6)中, $\Psi_a(\mathbf{V})$ 为区域配准项, $\Psi_1(\mathbf{V})$ 为局部相似项, $\Psi_g(\mathbf{V})$ 为全局相似项, $\Psi_s(\mathbf{V})$ 为直线保持项, $\lambda_1, \lambda_g, \lambda_s$ 分别为各项对应的系数。

区域配准项:在得到网格匹配点后,区域配准项通过最小化匹配点之间的像素坐标误差,求取最优的网格顶点位置集合。具体可以表示为

$$\Psi_a(\mathbf{V}) = \sum_{i=1}^n \sum_{(i,j) \in J} \sum_{p_k^{ij} \in Q^{ij}} \|\Phi(p_k^{ij}) - \Phi(P(p_k^{ij}))\|^2 \quad (7)$$

式中, J 为具有视野重叠区域的图像集合; Q^{ij}

为图像 I_i 和 I_j 的匹配点对集合; $P(p)$ 表示与点 p 对应的匹配点; $\Phi(p)$ 为点 p 在其所处网格顶点下的双线性插值结果, 可表示为: $\Phi(p) = \sum_{i=1}^4 \alpha_i \tilde{V}_i$, 其中 \tilde{V}_i 为 4 个顶点的坐标, α_i 表示插值权重。

局部相似项: 为了保证匹配点附近的网格在配准时不会发生局部扭曲的现象, 该项通过局部相似变换的方式最小化网格的形变量, 可以使变换合理地由重叠区域过渡到非重叠区域。具体定义如下

$$\Psi_1(\mathbf{V}) = \sum_{i=1}^n \sum_{(j,k) \in E_i} \left\| (v_k^i - v_j^i) - T_{jk}^i (v_k^i - v_j^i) \right\|^2 \quad (8)$$

式中, v_k^i 和 v_j^i 为图像 I_i 中原始网格的网格点坐标, v_k^i 和 v_j^i 表示对应的变形后的网格顶点坐标。 T_{jk}^i 表示边 e_{jk} 对应的相似变换矩阵, 可以通过文献[17]中的方法求得, 具体形式为

$$T_{jk}^i = \begin{bmatrix} c(e_{jk}) & s(e_{jk}) \\ -s(e_{jk}) & c(e_{jk}) \end{bmatrix} \quad (9)$$

全局相似项: 该项通过上文求取的全局尺度项和全局旋转项保证待拼接的图像能够自然地投影到最终的全景图像平面上。能量函数定义如下所示

$$\Psi_g(\mathbf{V}) = \sum_{i=1}^n \sum_{e_k^i \in K_i} \omega(e_k^i)^2 \left[(c(e_k^i) - s_i \cos \hat{\theta}_i)^2 + (s(e_k^i) - s_i \sin \hat{\theta}_i)^2 \right] \quad (10)$$

式中, $c(e_k^i)$ 和 $s(e_k^i)$ 为图像 I_i 的边 e_k^i 对应的相似变换矩阵 T_k^i 的参数, K_i 为图像 I_i 中网格边的集合, s_i 为图像 I_i 的全局尺度项, $\hat{\theta}_i$ 为全局旋转项。 $\omega(e_k^i)$ 为针对边 e_k^i 的权重函数, 具体定义如下

$$\omega(e_k^i) = \beta + \frac{\gamma}{|Q(e_k^i)|} \sum_{q_i \in Q(e_k^i)} \frac{d(q_i, M^i)}{\sqrt{R_i^2 + C_i^2}} \quad (11)$$

式中, β 和 γ 为常数, 其数值大小决定全局相似项对于网格优化的总体重要性。 $Q(e_k^i)$ 为边 e_k^i 两端网格四边形的集合, M^i 为图像 I^i 的重叠区域内的网格四边形 q_k 的集合, $d(q_k, M^i)$ 表示四边形 q_k 与重叠区域的距离, R_i 和 C_i 则分别表示图像 I_i 网格划分的行数和列数。

直线保持项: 在 NISwGSP 算法中, 当图像中的直线结构穿越多个网格时, 由于每个网格进行的变换均不相同, 可能会导致直线在经过变形之后失去原本的特性。针对此问题, 算法设计了直线保持项。在计算该项之前, 首先利用 LSD 算法分别检测待拼接图像中的直线, 并根据直线的长度对其进行筛选, 当直线长度大于网格对角线长度时, 以该直

线起点作为起始采样点对直线上的点进行均匀采样, 最后将采样点作为优化项的输入参与网格顶点的优化过程。该项的具体形式为

$$\Psi_s(\mathbf{V}) = \sum_{i=1}^n \sum_{j \in L_i} \sum_{k=1}^{m-2} \left\| (\Phi(d_{k+2}^j) - \Phi(d_{k+1}^j)) - (\Phi(d_{k+1}^j) - \Phi(d_k^j)) \right\| \quad (12)$$

式中, L_i 为图片 I_i 上所有直线的集合, L_i^j 为集合 L_i 中的一条直线, $d_1^j, d_2^j, \dots, d_m^j$ 为该直线上的采样点, 经过变形后的采样点为 $d_1^j, d_2^j, \dots, d_m^j$, $\Phi(d_k^j)$ 为点 d_k^j 在其所处网格的 4 个顶点下的双线性插值结果。

由式(6)可以得到变形后网格顶点的求解方程, 其中待优化顶点的系数由稀疏矩阵存储, 待求解项为优化后的网格顶点。由于直接使用最小二乘法计算量太大, 因此采用共轭梯度最小二乘法 (Least-Square Conjugate Gradient, LSCG) 最小化能量误差, 计算得到最优顶点集坐标 $\tilde{\mathbf{V}}$ 。得到全局顶点坐标 $\tilde{\mathbf{V}}$ 后, 将每个网格划分为 2 个三角形, 对于每个三角形分别求得唯一的仿射变换矩阵。利用得到的仿射变换矩阵对三角形内的所有像素进行投影, 即可得到经过变形对齐之后的图像。

1.5 最佳缝合线融合

待拼接图像经过对准后必须通过融合才能形成完整的全景图像。传统的图像融合算法有多分辨率融合算法、加权平均法、泊松融合算法和最佳缝合线算法等^[18]。多分辨率融合算法具有算法复杂度低、速度快等特点, 但是当图像重叠区域不能严格对齐时, 融合效果并不理想; 加权平均法速度较快, 但是容易出现伪影和曝光等问题; 泊松融合算法融合效果较好, 但是计算量较大并且用时太长; 最佳缝合线算法虽然能够较好地处理大视差场景下的图像融合问题, 但是当图像存在明显的曝光差异时, 融合区域不能自然地进行过渡, 造成全景图像出现明显的接缝。

为了实现重叠区域的自然过渡, 得到更加准确、合理的融合图像, 本文将最优缝合线搜索算法与多分辨率融合算法相结合。首先根据待拼接图像相邻像素的相似性构造重叠区域的加权图, 利用最大流-最小割算法得到待拼接图像的掩码; 然后构造图像的拉普拉斯金字塔, 将原图像分解到不同的空间频带上, 并在每个频带上采用不同的融合算子进行融合; 最后根据最佳缝合线信息对图像进行重构得到最终的全景图像。其中加权图的具体构造方法如下

$$E(p, q) = \frac{\|I_i(p) - I_j(p)\| + \|I_i(q) - I_j(q)\|}{2} \quad (13)$$

式中, p 和 q 代表图像 I_i 和图像 I_j 重叠区域的 2 个相邻像素, $E(p, q)$ 表示像素 p, q 之间的边的权重。若 p, q 在 2 幅图之间的内容均具有较高的相似性, 则认为该部分得到了准确的对齐, 相应边的权重较小, 反之则权重较大。

2 基于全景图像的目标定位算法

2.1 基于标定参数的定位方法

当目标在某一时刻出现在多个摄像机的视野重叠区域时, 可以通过摄像机配置信息联合三角测量的方式实现目标空间位置的快速定位, 具体介绍如下:

令观测到目标 S_j 的某一个相机为 c_i , 其内参矩阵为 K_i , 相机坐标系相对于世界系的旋转矩阵为 $R_w^{c_i}$, 平移矢量为 $t_w^{c_i}$, 运动目标在相机 c_i 图像平面的坐标为 C_j^i , 在世界系的坐标为 C_j^w , 则由相机透视投影模型可得

$$C_j^i = K_i (R_w^{c_i} C_j^w + t_w^{c_i}) \quad (14)$$

进一步可得

$$K_i R_w^{c_i} C_j^w = C_j^i - K_i t_w^{c_i} \quad (15)$$

联立多个观测到该目标的相机给出的方程可得

$$\begin{bmatrix} K_i R_w^{c_i} \\ K_{i+1} R_w^{c_{i+1}} \\ \vdots \end{bmatrix} C_j^w = \begin{bmatrix} C_j^i - K_i t_w^{c_i} \\ C_j^{i+1} - K_{i+1} t_w^{c_{i+1}} \\ \vdots \end{bmatrix} \quad (16)$$

为了简便, 式(16)可以表示为

$$H_j C_j^w = b_j \quad (17)$$

由此可知, 求解 C_j^w 等价于求解线性最小二乘问题

$$\operatorname{argmin}_{C_j^w} \frac{1}{2} \|H_j C_j^w - b_j\|_2^2 \quad (18)$$

运动目标在世界系的坐标可以直接通过求解伪逆的方式获得

$$C_j^w = (H_j^T H_j)^{-1} H_j^T b_j \quad (19)$$

2.2 基于单应性矩阵的快速定位方法

当场景中存在如地面和操场之类的平面或者监控摄像机的内参或外参矩阵不再可信时, 可以通过平面在世界系与图像坐标系下的对应关系实现前景目标的快速定位。具体方法如下:

由摄像机透视投影关系可知, 平面在不同坐标系间的转换关系可以由一个单应性矩阵 H 表达

$$H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \quad (20)$$

已知平面上的点 P 在世界系下的坐标 P^w 和图像平面的齐次坐标 p , 则两者的关系如下

$$P^w = \begin{bmatrix} x^w \\ y^w \\ z^w \end{bmatrix} = H \times p = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (21)$$

一般情况下, 运动目标都会与地面等场景中的主平面有接触, 因此可以选取目标在平面上的点作为待定位点。由于单目视觉在无外界测量辅助或先验知识的情况下不具备距离无尺度模糊恢复能力^[19]。因此, 需要选取至少 4 组已知世界系坐标的图像点作为控制点, 计算该平面在射影变换中的投影矩阵 H , 并且控制点与待定位的目标点都处于最佳缝合线的同一侧。

在获取目标在图像平面的像素坐标后, 可以通过单应性矩阵 H 求取前景目标在世界坐标系的大致位置, 实现目标的快速定位。

3 实验验证

由于图像拼接没有公用的数据集, 为了验证算法对大场景下图像拼接结果的合理性以及对目标定位结果的精确性, 本文在相关论文中的图像以及真实采集的图像的基础上进行了一系列实验, 并与 AutoStitch、APAP、SPHP、NISwGSP 算法进行了比较。具体结果如图 1 所示。

如图 1 所示, A 列和 B 列输入图像来自于文献[6], 明显可以看出, AutoStitch、APAP、SPHP 和基于平均融合的 NISwGSP 算法均出现了明显的鬼影问题, 除此之外 AutoStitch、APAP 和 SPHP 算法也出现了不同程度的尺度失真问题。基于最佳缝合线的 NISwGSP 算法虽然解决了鬼影和尺度失真的问题, 但是由于重叠区域对齐效果不理想, 拼接结果在接缝处出现了结构不连续的现象。本文的方法则得出了相对较好的结果。

图 1 的 C 列是由实际采集的图像得到的拼接结果。由于条件的限制, 实验以操场看台上手机拍摄的图像模拟监控摄像机拍摄的真实操场场景。输入图像存在重复的纹理结构(草坪、跑道等), 拥有多个复杂平面, 存在较大的视差并且共视区域包含特征稀疏区域和特征丰富区域等繁杂的场景信息。由 C 列可以看出, 基于平均融合的 NISwGSP 算法的拼接结果

存在明显的失真情况,如严重的鬼影、损失了操场边缘的直线特性及左侧路灯及建筑发生了明显的变形等状况。虽然基于最佳缝合线的 NISwGSP 算法消

除了鬼影问题,但是图中的结构信息仍然遭到了严重的破坏。本文中的算法则相对较好地解决了这些问题,提升了拼接结果的自然性与准确性。

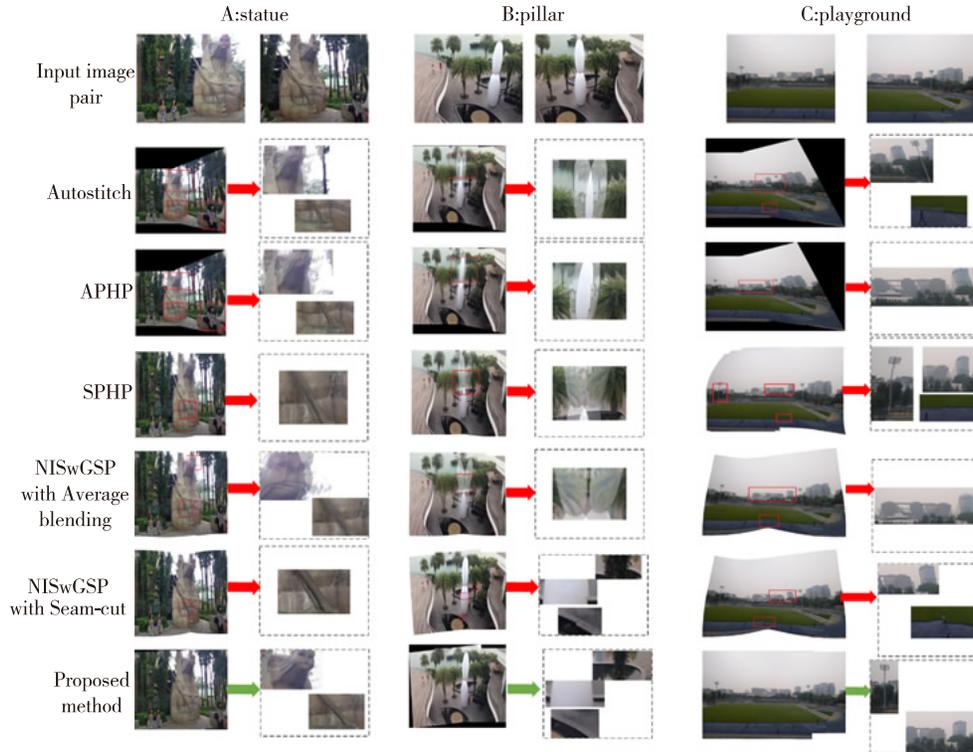


图 1 输入图像及结果对比

Fig. 1 Input image pair and results comparison

实验在 Windows10 环境 (2.80GHz CPU, 8GB RAM) 下进行,算法运行时间对比如表 1 所示。将融合算法由平均融合改为最佳缝合线算法后,算法耗时大幅增加。相对于基于最佳缝合线的 NISwGSP 算法来说,本文提出的改进算法在时间上稍有增加,但是能取得良好的拼接效果。

表 1 运行时间

Tab. 1 Running time

图片	基于平均融合的 NISwGSP/ms	基于最佳缝合线的 NISwGSP/ms	论文提出的算法/ms
Statue	5724	9568	10284
pillar	4570	7006	8883
playground	7503	11396	12056

对于前景目标定位问题,如图 2 所示,首先选取已知世界坐标的 4 个点(任意 3 点不共线)作为控制

点。其中以点 1 处作为世界系坐标原点建立直角坐标系,然后基于此求取操场平面与图像平面的单应性矩阵 H 。最后利用求得的单应性矩阵 H 对场景中的一些平面点进行定位的结果及误差如表 2 所示,其中估计误差由目标点的世界坐标估计值与实际值的欧氏距离表示。



图 2 控制点选取

Fig. 2 Control points selection

表2 定位结果及误差

Tab. 2 Location results and errors

序号	图像坐标/ pixel	估计的世界 坐标/m	真实的世界 坐标/m	估计误 差/m
1	(717,1857)	(8.66,8.72)	(9.35,7.85)	1.11
2	(2043,1827)	(53.67,-0.50)	(51.20,-2)	2.89
3	(257,1889)	(0.47,16.16)	(0,15.70)	0.66
4	(1986,2464)	(50.19,71.47)	(53.25,69.10)	3.87
5	(827,1820)	(6.07,-0.33)	(6.35,-2)	1.77

由实验结果可以看出,利用该方法可以实现监控角度下目标的快速定位,最终估计误差在 2.06m 左右。

4 结论

本文针对 NISwGSP 图像拼接算法中网格匹配点生成结果不符合图像真实信息以及直线结构容易遭到破坏等情况导致的全景图像失真问题,提出了相关的改进方案。实验总结及展望如下:

1)特征匹配及网格点匹配的结果严重影响图像的对齐质量。本文利用多平面 RANSAC 算法获得了较多的合理匹配点,并且基于超像素区域对网格匹配点进行了筛选,在充分利用图像特征的同时,减少了由错误的匹配信息带来的影响,对局部单应性矩阵的生成和网格优化提供了一定的支持。

2)几何结构的保持对全景图像的自然性至关重要。本文提出的基于直线约束的网格变形算法对网格变形起到了一定的抑制作用,使网格变形后的结果更加合理。

3)基于局部单应性矩阵的目标定位算法能够根据图像信息对监控场景下的目标进行快速定位,增强了监控端对目标信息的整体把控能力。

4)本文提出的算法能够有效提升图像配准的精确性,后续可以利用直线和轮廓等其他结构信息进一步提升重叠区域的配准效果。对于基于全景图像的定位算法,如何在统一的坐标系下实现目标的高精度定位则是后续的改进方向。

参考文献

[1] 邓中亮,尹露,唐诗浩,等.室内定位关键技术综述[J].导航定位与授时,2018,5(3):14-23.
Deng Zhongliang, Yin Lu, Tang Shihao, et al. A survey of key technology for indoor positioning[J]. Navigation Positioning and Timing, 2018, 5(3): 14-

23 (in Chinese).

- [2] Brown M, Lowe D G. Automatic panoramic image stitching using invariant features[J]. International Journal of Computer Vision, 2007, 74(1): 59-73.
- [3] Gao J, Kim S J, Brown M S. Constructing image panoramas using dual-homography warping [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2011: 49-56.
- [4] Lin W Y, Liu S, Matsushita Y, et al. Smoothly varying affine stitching[C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2011: 345-352.
- [5] Zaragoza J, Chin T J, Brown M S, et al. As-projective-as-possible image stitching with moving DLT [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2013: 2339-2346.
- [6] Lin K, Jiang N, Cheong L F, et al. SEAGULL: seam-guided local alignment for parallax-tolerant image stitching[C]// Proceedings of European Conference on Computer Vision. Springer, Cham, 2016: 370-385.
- [7] Chang C H, Sato Y, Chuang Y Y. Shape-preserving half-projective warps for image stitching[C]// Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition(CVPR). IEEE, 2014.
- [8] Chen Y S, Chuang Y Y. Natural image stitching with the global similarity prior[C]// Proceedings of European Conference on Computer Vision. Springer, Cham, 2016: 186-201.
- [9] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF[C]// Proceedings of 2011 International Conference on Computer Vision. IEEE, 2012.
- [10] Fischler M A. Random Sample Consensus: a paradigm for model fitting with applications to image analysis and automated cartography[J]. Readings in Computer Vision, 1981: 726-740.
- [11] Chin T J, Yu J, Suter D. Accelerated hypothesis generation for multi-structure robust fitting [C]// Proceedings of European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2010: 533-546.
- [12] Van Den Bergh M, Boix X, Roig G, et al. SEEDS: superpixels extracted via energy-driven sampling[C]// Proceedings of European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2012: 13-26.
- [13] Szeliski R, Shum H Y. Creating full view panoramic

- image mosaics and environment maps[C]// Proceedings of 24th Annual Conference on Computer Graphics and Interactive Technology. 1997: 251-258.
- [14] Szeliski R. Image alignment and stitching: a tutorial [J]. Foundations and Trends[®] in Computer Graphics and Vision, 2007, 2(1): 1-104.
- [15] Von Gioi R G, Jakubowicz J, Morel J M, et al. LSD: a fast line segment detector with a false detection control[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(4): 722-732.
- [16] Zhao C, Zhang H, Chen J, et al. Region-based parallax-tolerant image stitching[C]// Proceedings of 2018 10th International Conference on Graphics and Image Processing(ICGIP). International Society for Optics and Photonics, 2019, 11069: 1106909.
- [17] Igarashi T, Igarashi Y. Implementing as-rigid-as-possible shape manipulation and surface flattening[J]. Journal of Graphics GPU and Game Tools, 2009, 14(1): 17-30.
- [18] 谷雨, 周阳, 任刚, 等. 结合最佳缝合线和多分辨率融合的图像拼接[J]. 中国图象图形学报, 2017, 22(6): 842-851.
Gu Yu, Zhou Yang, Ren Gang, et al. Image stitching by combining optimal seam and multi-resolution fusion [J]. Journal of Image and Graphics, 2017, 22(6): 842-851(in Chinese).
- [19] 张海, 吴克强, 张晓鸥. 视觉导航技术的发展[J]. 导航定位与授时, 2017, 4(2): 1-8.
Zhang Hai, Wu Keqiang, Zhang Xiaou. The development of visual navigation[J]. Navigation Positioning and Timing, 2017, 4(2): 1-8(in Chinese).