

基于联合相机路径的红外视频稳像算法

尹丽华¹, 杭娟¹, 康亮¹, 刘士建²

(1. 上海第二工业大学 工程训练中心, 上海 201209;

2. 中国科学院上海技术物理研究所 中国科学院红外探测与成像技术重点实验室, 上海 200083)

摘要: 针对由于视差变化而导致的红外视频稳像技术难题, 文中提出了一种基于联合相机路径的红外视频稳像算法, 该算法主要包括: 预处理、联合相机路径求解、多路径优化、运动补偿四个步骤。首先, 要对红外图像进行直方图均衡化、特征点提取和匹配处理、预映射。接着, 要将每帧图像分为 $m \times n$ 个网格单元, 利用基于网格的映射运动表征, 将每一帧中对应网格得到的局部单应性矩阵进行相乘得到联合相机路径。然后, 对联合相机路径采用“多路径优化”策略进行平滑处理。最后, 利用平滑后路径进行运动补偿稳定视频。实验结果表明, 该方法能有效地处理由于视差而引起的非线性运动, 相比传统的稳像算法效果更好, 当特征点存在部分遮挡时, 也能取得不错的稳像效果。

关键词: 视频稳像; 视差; 网格单元; 多路径; 非线性

中图分类号: TP391.41 **文献标志码:** A **DOI:** 10.3788/IRLA20200405

Infrared video image stabilization algorithm based on joint camera path

Yin Lihua¹, Hang Juan¹, Kang Liang¹, Liu Shijian²

(1. Engineering Training Center, Shanghai Polytechnic University, Shanghai 201209, China;

2. Key Laboratory of Infrared System Detection and Imaging Technology, Chinese Academy of Sciences, Shanghai Institute of Technical Physics, Chinese Academy of Sciences, Shanghai 200083, China)

Abstract: To solve the problem of infrared video image stabilization caused by parallax variation, the infrared video image stabilization algorithm based on joint camera path was proposed, which included four steps: preprocessing, joint camera path solving, multi-path optimization and motion compensation. Firstly, histogram equalization, feature point extraction, matching processing and pre-mapping were needed for infrared image. Then, each frame of image was divided into $m \times n$ grid cells, and local homography matrixs obtained from the corresponding grid in each frame were multiplied by the mapping motion representation based on the grid to obtain the joint camera path. Then, the path of the joint camera was smoothed by the strategy of ‘multi-path optimization’. Finally, the smooth path was used to stabilize the video. The experimental results show that this method can effectively deal with the nonlinear motion caused by parallax, which is better than the traditional image stabilization algorithm, and can achieve good image stabilization even when the feature points have partial occlusion.

Key words: video image stabilization; parallax; grid unit; multi-path; nonlinear

收稿日期: 2021-01-02; 修订日期: 2021-04-22

基金项目: 上海第二工业大学学科基金 (XXKZD1605); 上海第二工业大学校基金 (EGD19XQD11); 校企合作汽车电子联合实验中心建设项目 (A11NH190704)

作者简介: 尹丽华, 女, 讲师, 博士, 主要研究方向为红外光电系统工程及信息与图像处理技术。

0 引言

基于运动平台的各种类型成像系统在红外探测方面发挥了重要作用^[1-2],受到平台振动的影响,会影响后续的目标识别或跟踪,所以,迫切需要通过稳像技术^[3]来去除视频抖动。数字视频稳像的步骤一般包括:运动估计->路径平滑->利用平滑路径合成稳定的视频^[4]。根据运动模型的不同,传统视频稳像算法分为两类,即:

(1) 基于 2D 的视频稳像方法

通过估计连续视频帧之间的 2D 线性转换,并随时间对其进行平滑以生成稳定的视频,目前主要着重于对路径平滑算法的研究。最常用的技术手段是采用低通滤波器或卡尔曼滤波器法用于相机运动路径的平滑;Chen 采用多项式拟合法对运动路径进行平滑^[5];后来, Matsushita 等人提出利用高斯加权函数对全局运动参数进行运动平滑^[6],减少了累积误差。最近, Grundmann 应用 L1 范数^[7]优化生成一个包含常量、线性和抛物线运动的摄像机路径,这些运动遵循了电影的摄影规则。

国内对于 2D 视频稳像技术也有一定研究^[8-10],如朱娟娟提出了抗前景干扰的电子稳像算法^[8],通过自上向下的金字塔进行由粗到细的运动估计;邱家涛提出基于 ORB^[9-12]特征的稳像算法^[9],相对于传统的 SIFT、SURF 特征点提取算法,实时性更好;2018 年谢亚晋等人提出基于最小生成树和运动矢量对的视频稳像算法^[10],能剔除运动前景的干扰。

(2) 基于 3D 的视频稳像方法

通过估计 3D 空间中的相机运动以保持稳定。Beuhler 使用未经校准的摄像机对场景进行 3D 投影

重建来实现稳定化^[13]。Liu 开发了 3D 视频稳定化系统^[14],引入“内容保留”变形来稳定视频。由于 3D 重建很难,后来采用二维特征轨迹上应用低秩子空间约束^[15],或加入极间转移技术^[16],这些方法将三维重建的要求放宽到二维长特征跟踪。

当相机拍摄视频时不可避免地存在拍摄角度的变化,比如大幅度的旋转、缩放或前进,均会导致图像的深度发生变化,引起视差,对应的图像运动模型也更加复杂。3D 视频稳像方法虽然能处理视差,但对各种退化(如特征跟踪失败)鲁棒性较低,2D 视频稳像方法更加稳健^[17-18],但单一 2D 线性运动模型无法从根本上处理视差^[19-20]。因此,文中提出了一种基于联合相机路径的红外视频稳像算法,创新点主要体现在两个方面:一、运动估计:对红外图像进行预处理,将图像进行网格划分,采用基于网格映射的运动表征图像运动信息,这是一种更强大的 2D 运动模型,并加入“形状保持约束”来增加泛化能力;二、路径平滑:通过多路径优化策略来优化相机路径,是一种时空路径的优化方法,综合解决了由于视差变化导致的稳像技术难题。

1 基本原理

1.1 算法流程

图 1 为基于联合相机路径的红外视频稳像算法流程。该算法主要包括:预处理、联合相机路径求解、多路径优化、运动补偿,基本步骤如下:

(1) 输入红外视频,对红外图像进行直方图均衡化、ORB 特征点提取和匹配,利用 RANSAC 算法^[12]从两种尺度上剔除异常的匹配点对,并进行预映射。

(2) 将原始视频帧分成 $m \times n$ 的网格单元,利用基

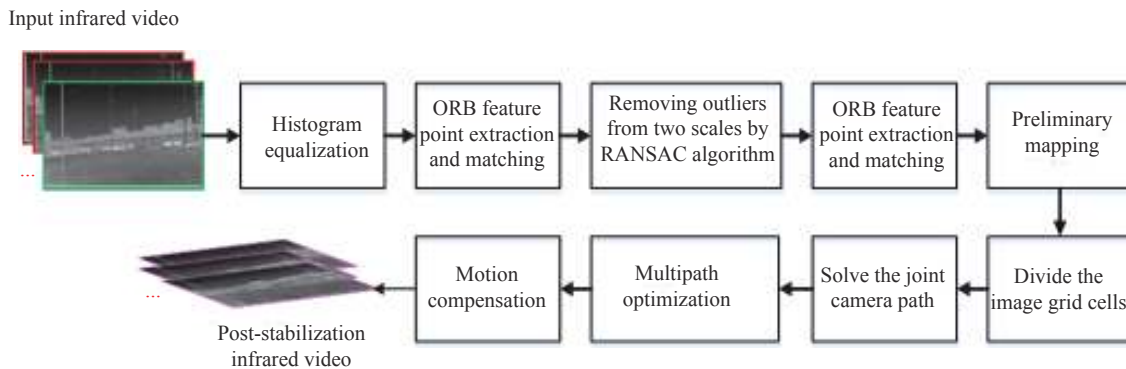


图 1 基于联合相机路径的红外视频稳像算法流程

Fig.1 Process of infrared video image stabilization algorithm based on joint camera path

于网格映射的运动模型来表征相机的运动,在目标函数中增加“形状保持”约束,估算网格的顶点向量;然后,利用映射前、后的顶点向量,求解出单应性矩阵 $F_i(t)$,通过累乘得到联合相机路径,假设 $C_i(t)$ 为第 t 帧图像第 i 个网格的联合相机路径,其公式为:

$$C_i(t) = C_i(t-1)F_i(t-1), \Rightarrow C_i(t) = F_i(0)F_i(1) \cdots F_i(t-1) \quad (1)$$

式中: $\{F_i(0), \cdots, F_i(t-1)\}$ 表示第 i 个网格在不同时刻的局部单应性矩阵,这些能描述空间变化的路径称之为“联合相机路径”,而路径经过平滑后才能用于后图像稳定。

(3) 对联合相机路径采用“多路径优化”策略进行平滑处理。假设初始的相机路径为 $C = \{C(t)\}$, 经过平滑后得的相机路径为 $P = \{P(t)\}$, 其中 $C(t)$ 、 $P(t)$ 分别表示 t 时刻的初始相机路径,平滑后相机路径。

(4) 对输入视频进行运动补偿,利用平滑后相机路径和原始路径之间的差异,计算出每一帧中每个网格的运动补偿量 $B_i(t)$, $B_i(t) = C_i^{-1}(t)P_i(t)$, 然后,对图像中每个网格进行反向补偿得到稳定后的红外图像。

1.2 预处理

由于红外图像反映的是目标及背景向外辐射能量的差异,属于被动成像,对比度比较低,信噪比较差,所以,文中首先采用直方图均衡技术来提高图像对比度;然后,对相邻两帧进行 ORB 特征点的提取和匹配;接着,利用 RANSAC 算法剔除异常匹配点对。

由于后续运动模型估计是基于特征点匹配对的,错误匹配点的存在会直接影响运动估计的精度,因此文中采用了 RANSAC 算法从两种尺度上剔除异常匹配点对:首先,在整幅图像上使用 RANSAC 算法拟以相对较大的拟合误差阈值(建议阈值:8% 图像宽度)丢弃特征;然后,在 4×4 的局部子图像上,使用 RANSAC 算法以相对较小的阈值(建议阈值:2% 图像宽度)丢弃特征。

1.3 联合相机路径求解

1.3.1 基于网格映射的运动模型描述

将视频帧分为 $m \times n$ 的网格后,采用基于网格映射的运动模型^[17]来表征图像运动,每帧中定义一个统一的网格,基于网格映射的运动模型如图 2 所示。

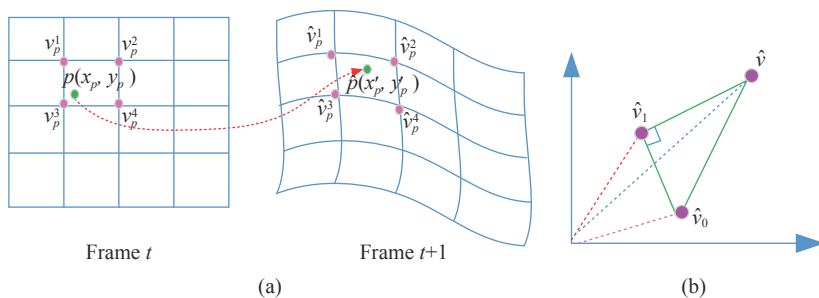


图 2 基于网格映射的运动模型。(a) 两帧图像网格之间的对应关系;(b) 形状保持约束

Fig.2 Motion model based on grid mapping. (a) Correspondence between two frames of image grids; (b) Shape retention constraints

在图 2(a) 中, $\{p, \hat{p}\}$ 表示一对匹配的特征点对,特征点 p 用封闭网格顶点 $V_p = [v_p^1, v_p^2, v_p^3, v_p^4]^T$ 的二维双线性插值表示, $p = V_p w_p$, $w_p = [w_p^1, w_p^2, w_p^3, w_p^4]^T$ 中各参数和为 1。求解对应特征点 \hat{p} 的映射网格顶点向量 $\hat{V}_p = [\hat{v}_p^1, \hat{v}_p^2, \hat{v}_p^3, \hat{v}_p^4]^T$ 的目标函数如下:

$$E_d(\hat{V}) = \sum_p \| \hat{V}_p w_p - \hat{p} \|^2 \quad (2)$$

式中: \hat{V}_p 为待求参量; w_p 和 \hat{p} 为已知参量; \hat{V} 中包括所有的映射网格顶点。

在图 2(b) 中,由顶点 \hat{v} 、 \hat{v}_1 、 \hat{v}_0 构成的三角形要遵

循相似变换。当每个单元格中可能没有足够的特征时,可以在目标函数中加入“形状保持约束”^[14],引入正则项 $E_s(\hat{V})$,使每个输出网格单元与其对应输入网格单元的相似度偏差最小:

$$E_s(\hat{V}) = \sum_{\hat{v}} \| \hat{v} - \hat{v}_1 - s R_{90} (\hat{v}_0 - \hat{v}_1) \|^2, R_{90} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \quad (3)$$

式中: 参数 s 根据初始网格进行估算 $s = \|v - v_1\| / \|v_0 - v_1\|$, 加入正则化项有助于保持网格的平滑变化。因此,最终的目标函数为 $E(\hat{V})$:

$$E(\hat{V}) = E_d(\hat{V}) + \alpha E_s(\hat{V}) \quad (4)$$

式中： α 为正则化参数，由于 $E(\hat{V})$ 是一个二次线性规划问题，可以利用一个稀疏线性系统求解器来求解 \hat{V} 。

文中方法采用基于网络映射的运动模型进行运动估计，该模型本质上是一系列基于 2D 网格的、空间变化的单应性模型，比传统单一的 2D 线性变换有更强的建模能力，能处理视差问题。另外，该模型通过在能量函数中增加“形状保持”约束，使文中算法对于图像特征点较少（由于无纹理区域或存在遮挡）情况时仍具有运动估计的泛化能力，所以该算法能较好地处理图像区域特征不足的问题。

那么，应该如何选择网格单元的数目呢？网格单元数目的多少对稳像精度和实时性均有影响。网格单元的数目直接决定了整帧图像中待求解的单应性矩阵个数，进而会影响到运动估计、路径平滑以及运动补偿环节，最终影响视频稳像的结果。一般情况下，对于同一帧图像，设置的网格单元数目越多，运动估计的精度越高，路径平滑的复杂度越高，运动补偿的效果更好，对应的稳像精度越高，但同时也会牺牲一定的实时性。因此，具体要综合考虑工程应用的实时性和稳像精度需求进行选择。为此，文中选用 640×320 分辨率的红外视频图像进行实验，并分别比较了六组不同网格尺寸下的视频稳像结果，结果如表 1 所示。

由表 1 可知，当网格单元尺寸较少时，运动模型估计时间较少，但 Stability 值也较低；而当网格单元数目增多时，Stability 值会提高，但运动模型估计时间增加。当网格单元尺寸为 16×16 时，运动模型估计时间为 2.165 s，同时稳定度可以达到 0.86，完全能满足视频稳像的需要；而当网格单元尺寸变为 32×32

和 64×64 时，Stability 值提高不明显，但是算法的运动时间却会明显增多。因此，文中建议将图像帧分为 16×16 的网格单元，这样能同时保证稳像算法的视频稳定性和运行效率。

表 1 不同网格单元尺寸下的视频稳像结果

Tab.1 Video image stabilization results under different grid cell sizes

Grid cell size	Run model estimate time/s	Stability
3×3	0.988	0.65
7×7	1.155	0.72
9×9	1.443	0.78
16×16	2.165	0.86
32×32	13.261	0.87
64×64	370.538	0.83

1.3.2 模型运动估计

当映射网格的顶点向量估算完成后，需要求解出第 t 帧图像第 i 个网格的初始局部单应性矩阵为 $F_i'(t)$ ，等式如下：

$$\hat{V}_i = F_i'(t) V_i \quad (5)$$

式中： V_i 和 \hat{V}_i 分别为映射前后的顶点向量，均为已知量； $F_i(t)$ 为未知量。

1.3.3 联合相机路径迭代

对于第 t 帧图像的第 i 个网格，将初始局部单应性矩阵 $F_i'(t)$ 和全局单应性变换矩阵 $\bar{F}(t)$ 相乘可以得到最终的局部单应性矩阵 $F_i(t)$ ，即 $F_i(t) = F_i'(t) \times \bar{F}(t)$ 。然后，通过公式 (1) 估算出描述空间变化的联合相机路径 $C_i(t)$ ， $C_i(t)$ 为第 t 帧图像中第 i 个网格的初始相机路径，联合相机路径求解的过程如图 3 所示。

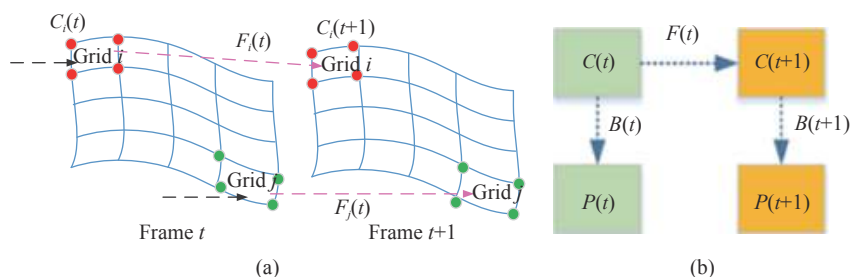


图 3 联合相机路径求解过程的示意图。(a) 网格的路径迭代；(b) 初始相机路径 $C(t)$ 、平滑路径 $P(t)$ 、补偿路径 $B(t)$ 之间的关系

Fig.3 Schematic diagram of the joint camera path solution process. (a) Path iteration of the grid; (b) Relationship among the initial camera path $C(t)$, smooth path $P(t)$, and compensation path $B(t)$

图 3(a) 为网格的路径迭代, 其中 $C_i(t+1) = C_i(t)F_i(t)$, $C_j(t+1) = C_i(t)F_j(t)$ 分别为第 t 帧图像中第 i 个、第 j 个网格单元的路径迭代公式。

图 3(b) 为初始相机路径 $C(t)$ 、平滑路径 $P(t)$ 、补偿路径 $B(t)$ 之间的关系, 初始相机路径 $C(t)$ 是由路径迭代获得的, 而平滑路径结合后续第 1.4 节的多路径优化后得到, 则最终图像需要补偿的路径为 $B(t)$, 并且 $B(t) = C^{-1}(t)P(t)$, 利用利用矩阵 $B(t)$ 对图像进行反向补偿, 得到稳定后的图像。

文中算法的关键就是采用网格映射模型进行运动估计, 在能量函数中加入正则项增加运动模型估计

的泛化能力。为了验证文中算法图像对特征点较少情况时的运动估计泛化能力, 文中对图像特征点集丰富和较少情况下的稳像结果进行了对比, 结果如图 4 所示。

图 4(a) 为图像特征点集丰富的情况, 左侧为特征点集, 绿色点表示特征点, 右侧为稳像结果, 网格表示利用所有特征点估计的映射网格; 图 4(b) 为图像特征点集较少的情况, 左侧为特征点集, 红色矩形框代表其中的特征点被剔除掉, 右侧为稳像结果。图 4(a) 和图 4(b) 达到的稳像结果相似, 这说明该算法能较好地处理图像区域特征不足的问题。

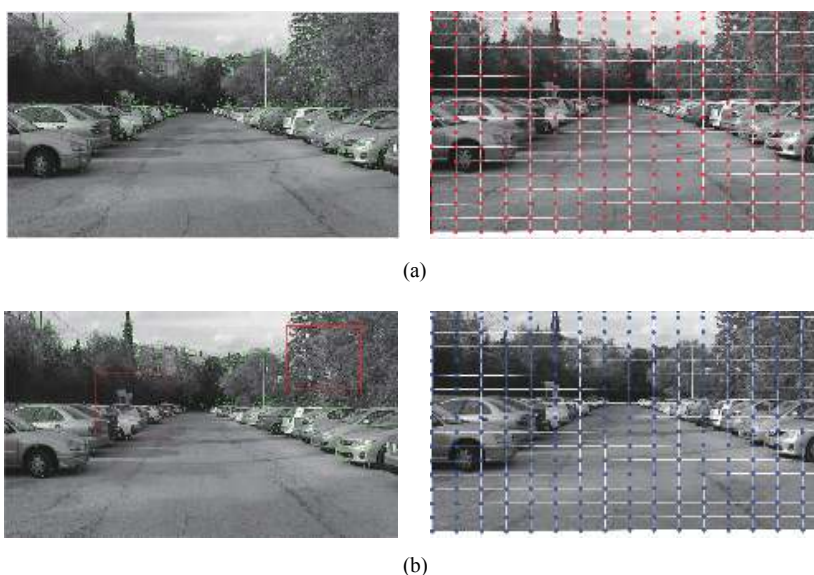


图 4 图像特征点集丰富和较少情况下的稳像结果对比。(a) 特征点集丰富; (b) 特征点集较少

Fig.4 Comparison of image stability results in the case of abundant feature points and few feature points. (a) Rich feature point set; (b) Less feature point set

1.4 多路径优化原理

对联合相机路径采用“多路径优化”策略进行平滑处理。相机路径平滑问题应考虑多个因素: 要消除抖动, 避免过度裁剪, 最小化各种几何畸变(剪切/倾斜, 抖动)。

1.4.1 单路径优化

每个网格对应生成一条相机路径, 首先, 对单个相机路径进行平滑。对于给定的初始相机路径 $C = \{C(t)\}$, 希望得到平滑后得的相机路径为 $P = \{P(t)\}$, 通过极小化能量函数得到:

$$f(P) = \sum_t \left(\|P(t) - C(t)\|^2 + \lambda \sum_{r \in \Omega_t} \|P(t) - P(r)\|^2 \right) \quad (6)$$

式中: λ 表示自定义参数; Ω_t 表示第 t 帧的相邻帧, 第一项 $\|P(t) - C(t)\|^2$ 是为了使平滑后的相机路径能尽可能地接近初始相机路径, 避免稳定后的视频出现过度的裁剪和扭曲, 第二项 $\|P(t) - P(r)\|^2$ 是平滑项, 能去除原始视频中的抖动。

1.4.2 多路径优化

如果只对单个网格的路径进行优化, 则相邻路径的一致性可能会降低, 稳定视频中会产生失真。因此, 对单个路径进行平滑的同时, 需要考虑多个网格路径之间的空间关系, 通过最小化以下目标函数对多个路径进行时空优化:

$$\sum_{i=1}^{m \times n} f(P_i) + \beta \sum_t \sum_{j \in N(t)} \|P_i(t) - P_j(t)\|^2 \quad (7)$$

式中: β 表示自定义的参数; $N(i)$ 包括第 i 个网格单元的八个领域网格, 该式中第一项是单个路径的目标函数, 引入的第二项 $\sum_{j \in N(i)} \|P_i(t) - P_j(t)\|^2$ 是为了加强相邻路径之间的平滑性, 即保证每一个网格的路径其他领域网格的路径尽可能接近, 减少扭曲现象。

通过公式 (7) 求解得到最终的平滑后相机路径 $\{P_i\}_{i=1, \dots, m \times n}$, 由于优化问题是二次型的, 最优结果能通过求解一个稀疏线性系统求解器得到, 同样也可以通过基于雅可比矩阵^[18]的迭代更新求解。

第 t 帧图像的每个网格的路径经过平滑后, 计算第 i 个网格单元的补偿矩阵 $B_i(t)$, $B_i(t) = C_i^{-1}(t)P_i(t)$, 然后应用 $B_i(t)$ 在第 t 帧图像处扭曲第 i 个网格单元,

以生成最终的输出视频, 通常直接应用 $B_i(t)$ 会产生不错的稳像效果。图 5 为红外图像序列对应的联合相机路径和路径优化后结果。

公式 (7) 中第二项是为了加强相邻路径之间的平滑性, 如果去掉此项约束, 稳像后图像容易出现扭曲现象。接下来将对第二项约束加入前、后的平滑效果进行具体分析, 结果如图 6 所示。通过对比图 6(b) 和 6(c) 可以看出, 当未加入第二项约束时, 图像相邻路径之间的差异性较大, 图像容易出现扭曲, 影响稳像效果, 图中用箭头标识出; 而文中算法避免了图像的扭曲, 加强了相邻路径之间的平滑性, 从而保证了整体的稳像效果。

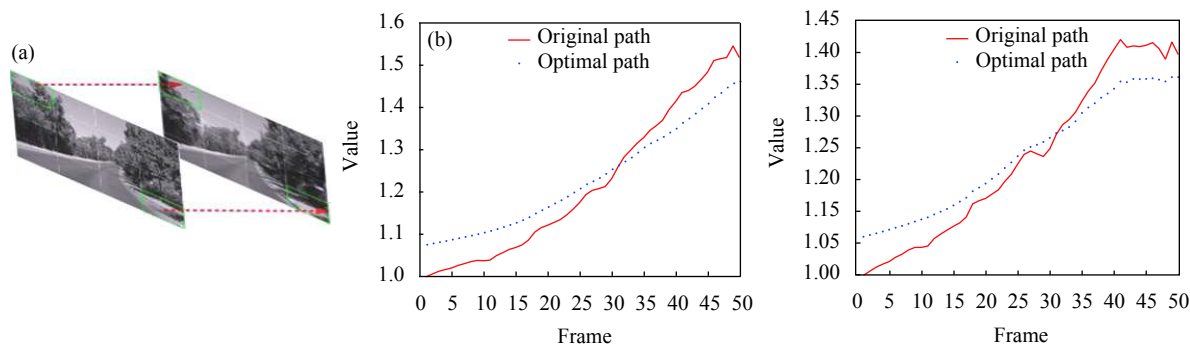


图 5 图像对应的联合相机路径和优化后的路径。(a) 图像帧间的对应网格; (b) 路径优化前后的结果: 左边第 (1, 1) 个网格, 右边第 (1, 1) 个网格
Fig.5 Image corresponds to the joint camera path and the optimized path. (a) Corresponding grids between image frames; (b) Results before and after path optimization: the left side is the (1, 1) grid, and the right side is the (1, 1) grid

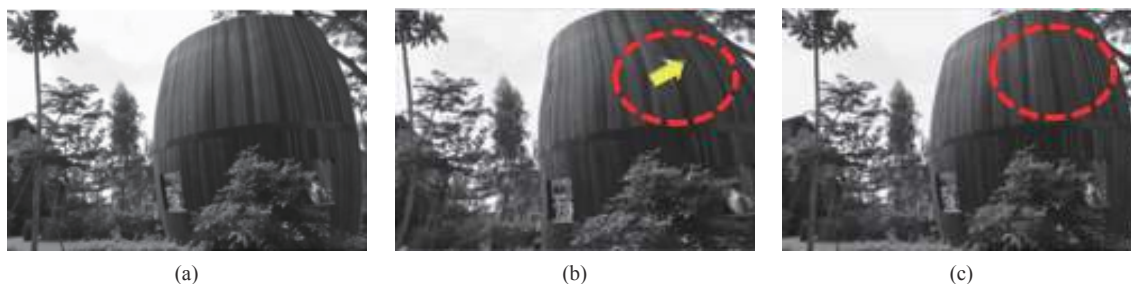


图 6 第二项约束加入前、后的平滑效果。(a) 原图像; (b) 未加第二项约束; (c) 文中的结果

Fig.6 Smoothing results before and after adding the second restraint. (a) Original image; (b) Without the second restraint; (c) Result in this paper

2 分析与讨论

2.1 实验条件与数据集介绍

文中的验证实验采用了如下的两大数据集:

数据集 1: 采用了参考文献 [19] 中提供的公共视频数据集, 这是常采用的基准数据集, 图像的分辨率为 640×320 , 总共包含了七种不同的拍摄情况: (I) Sim-

ple 类, 摄像机的运动比较简单; (II) Rotation 类, 摄像机存在大幅的旋转运动; (III) Zooming 类: 摄像机存在大幅的旋转运动; (IV) Parallax 类: 摄像机进行了扫描拍摄; (V) Driving 类: 利用基于车载的摄像机拍摄; (VI) Crowd 类: 拍摄的场景存在运动前景; (VII) Running 类: 利用快速前进的摄像机拍摄。

数据集 2: 利用红外线列探测器采集的视频数据

Algorithm :The infrared video image stabilization algorithm based on joint camera path

Input: inputDir, outputDir, nFrames, MeshSize

Output: stabilized image I_t' , ($1 \leq t \leq nFrames$)

Read image from path inputDir

Step1: Compute original camera path C

Initialize regularization parameter α ;

for t = 2 to nFrames **do**

Image pairs $\{I_t, I_{t-1}\}$ through **image preprocessing**, obtain global homography $\bar{F}(t)$;

Using the formula (4)(5), calculate t local homography matrix $F_i(t)$,

then homograph matrix result: $F_i(t) \leftarrow F_i(t) \times \bar{F}(t)$

For the **Mesh grid** (i, j) , $1 \leq i, j \leq MeshSize$, using formula(1), to calculate the

bundled path **C**: $C_i(t) \leftarrow F_i(0)F_i(1) \cdots F_i(t-1)$

end for

Step2: Optimize the camera path

Set $P = C$; Initialize the parameters λ and β ;

for t = 2 to nFrames **do**

For the **Mesh grid** (i, j) , using formula(6)(7) to calculate the optimize path **P**;

end for

Step3:Render the stabilzied frames

for t = 1 to nFrames **do**

$I_t' \leftarrow \text{render1}(I_t, t, P, C)$;

Save stabilized image I_t' to **outputDir**;

end for

Form a stable video file

图 7 文中算法实现的伪代码

Fig.7 Pseudo code of the algorithm in this paper

集,图像分辨率为 282×600,图像深度为 16 bit,总共包含七种不同的拍摄情况:(I) Simple 类,(II) Rotation 类,(III) Zooming 类,(IV) Parallax 类,(V) Driving 类,(VI) Crowd 类,(VII) Running 类。

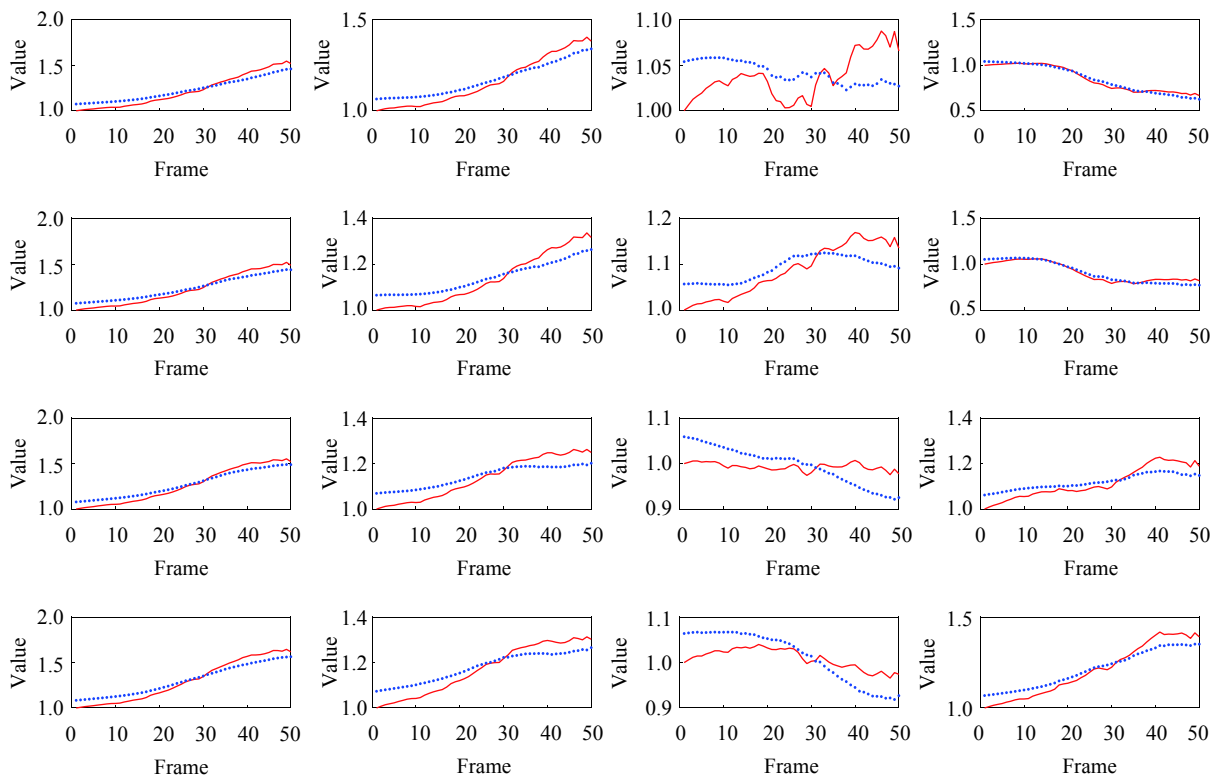
文中的验证实验均在 Intel(R) Core(TM)i5-5200U CPU@2.20 GHZ, 8 GB 内存, 64 位 Win10 系统上进行,利用 Matlab2016a 软件实现算法程序,并且结合“ffmpeg”媒体处理工具对视频/图像数据进行转化,文中算法框架方案可详见图 7 的算法伪代码。验证实验中,利用上述的数据集,将每一帧图像被分成了 16×16 的网格单元,通过预处理、联合相机路径求解、多路径优化、运动补偿步骤实现了稳像处理,并结合

客观指标进行评价。

2.2 直观评价结果

2.2.1 单相机路径与联合相机路径稳像结果的对比

传统的红外稳像算法中,整帧图像通常采用单一的相机路径,如仿射或单应性矩阵,但是图像各局部运动差异导致误差较大。而文中算法将视频帧分成 $m \times n$ 的网格单元,联合相机路径就是同一个网格单元随着时间推移的同构连接,将联合路径作为一个整体进行平滑,能保持空间和时间的一致性,所以稳像的效果更好。为了说明文中算法的优势,文中对单相机路径和联合相机路径的稳像结果进行了对比,结果如图 8 所示。



(a)



(b)

图 8 单路径和联合路径的稳像结果对比。(a) 路径优化结果; (b) 稳像后图像; 左边为单路径, 右边为联合路径

Fig.8 Comparison of image stability results between single path and joint path. (a) Path optimization results; (b) Image after image stabilization: the left side is the single path, and the right side is the joint path

图 8(a) 为路径优化结果, 原始和优化后路径, 各路径之间保持了空间和时间的一致性, 所以优化效果更好。图 8(b) 为稳像后的图像, 分别为单一和联合相机路径的稳像结果, 通过对比可以看出, 联合相机路径的稳像后图像, 图像过渡更加平滑, 视觉效果更好, 而单相机路径的稳像后图像存在扭曲变形, 边缘区域留白较大, 影响了视觉的效果。

2.2.2 对于含有视差情况的视频, 文中算法与传统算法的稳像结果对比

传统的稳像算法当场景存在视差的变化时, 适用性较差, 为了说明文中算法的稳像优势, 文中选取了

两种不同视差变化情况下的视频进行实验, 将文中算法分别和传统的稳像算法: L1 算法^[7]、AE 系统^[16](著名商业系统 Adobe After Effects CS6 中的“Warp Stabilize”功能, 简称 AE)、朱娟娟电子稳像算法^[8]进行对比, 实验结果如图 9~10 所示。

通过图 9~10 可以看出, AE 系统的视频稳定器容易出现视频的过渡裁剪, 导致大量像素缺失; 而朱娟娟电子稳像算法适用于 Parallax 类的视频, 对含有视差变化的视频, 稳像效果较差; L1 算法稳像后的视频图像相对于原图像差异较大, 容易出现视觉的突变, 稳像效果较差; 而文中算法由于采用网格化映射和多

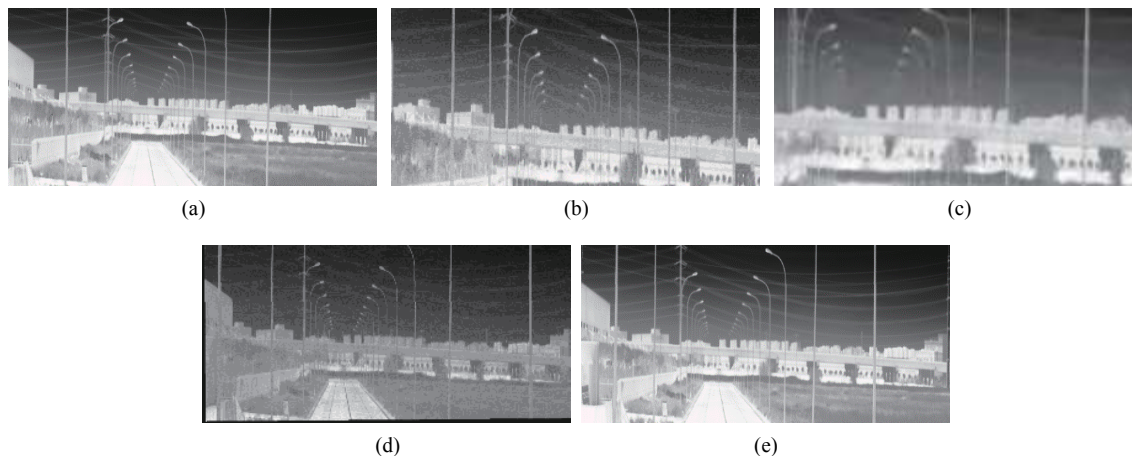


图 9 Driving 类视频下多种稳像算法的结果对比。(a) 原图像; (b) L1 算法; (c) AE 系统; (d) 朱娟娟电子稳像算法; (e) 文中算法

Fig.9 Comparison of results of various image stabilization algorithms under the Driving type video. (a) Original image; (b) L1 algorithm; (c) AE system; (d) Zhu Juanjuan's electronic image stabilization algorithm; (e) Proposed algorithm

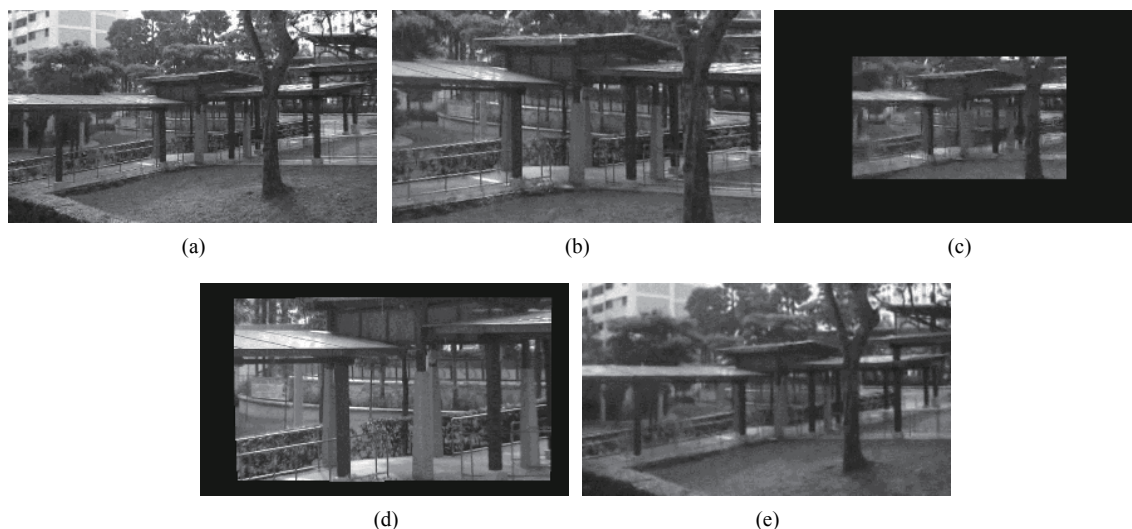


图 10 Zooming 类视频下多种稳像算法的结果对比。(a) 原图像; (b) L1 算法; (c) AE 系统; (d) 朱娟娟电子稳像算法; (e) 文中算法

Fig.10 Comparison of results of various image stabilization algorithms under the Zooming type video. (a) Original image; (b) L1 algorithm (c) AE system; (d) Zhu Juanjuan's electronic image stabilization algorithm; (e) Proposed algorithm

路径优化,能尽量保留稳定后图像的像素信息,而且图像过渡比较平滑,充分说明了文中算法的稳像结果更好。

2.3 客观评价结果

为了定量地评价和度量视频稳像结果,文中采用三个客观度量标准^[19-20]:即裁剪率(Cropping Ratio)、失真率(Distortion)、稳定度(Stability),定义如下:

裁剪率:衡量稳定后图像帧剩余的有效区域占原图像帧的比例,其值越大,说明视频稳像的效果越好,整个视频的裁剪率=average{所有视频帧裁剪率}。

失真率:衡量稳定后视频的失真程度,可以用映

射矩阵 $B(t)$ 的各向异性比例来衡量失真率,每一帧图像的失真率=映射矩阵 $B(t)$ 的两个最大特征值的比值,整个视频的失真率= $\min\{\text{所有帧图像的失真率}\}$ 。

稳定度:衡量稳定后视频的稳定度,采用频率分析法对视频中的 2D 运动进行估算,基于如下假设:运动分量中低频分量所占的越多,说明视频越稳定。

为了能客观说明文中稳像算法的优势,文中利用了七类视频数据进行实验,分别采用 L1 算法^[7]、AE 算法^[16]、文中算法,计算出了每种算法对应的稳像指标。

图 11 为几种稳像算法的定量评价结果。对于类别 (I),因为该类视频具有相对平稳的摄像机运动和较

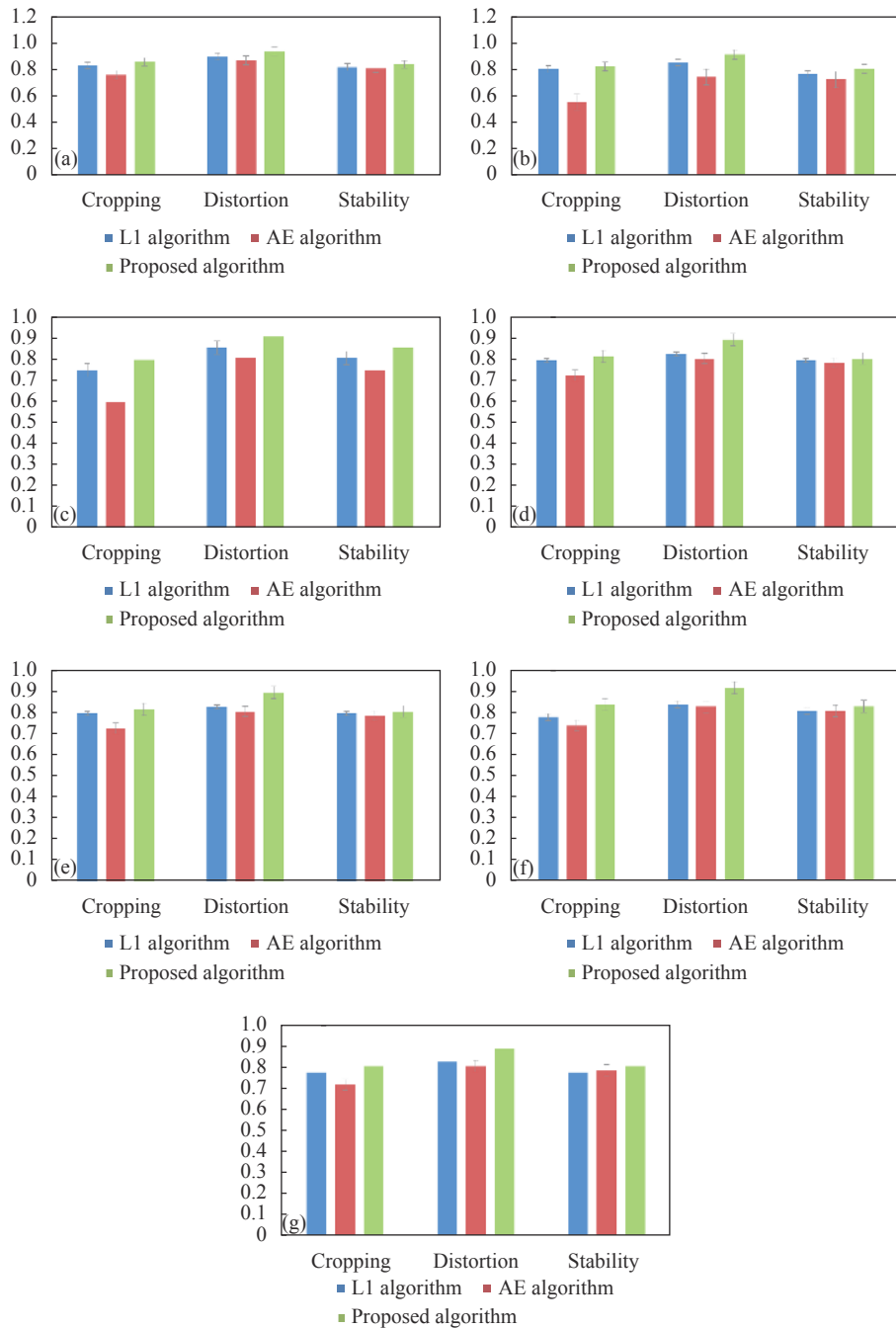


图 11 几种稳像算法在七组视频下的定量评价结果。(a) Simple; (b) Rotation; (c) Zooming; (d) Parallax; (e) Driving; (f) Crowd; (g) Running

Fig.11 Quantitative evaluation results of several image stabilization algorithms under seven sets of videos. (a) Simple; (b) Rotation; (c) Zooming; (d) Parallax; (e) Driving; (f) Crowd; (g) Running

小的深度变化,虽然这三种算法的稳像效果均不错,但是文中算法的裁剪率、失真率和稳定度指标更好。对于类别 (II~IV), AE 算法的裁剪率最低,主要是因为 AE 算法出现过度裁剪,特别是视差变化时。对于类别 (V~VII),这三种算法可以产生类似的稳定度指标,但文中算法在裁剪率或失真控制方面始终较好。

2.4 局限性讨论

相对于传统的基于 3D 的视频稳像方法,文中的算法更加稳定,因为不需要长时间特征跟踪,却可以产生类似的或者只是稍微差一点的结果。文中算法对于含有视差变化的视频,仍能取得较好的稳像效果,同时保留了传统 2D 视频稳像算法的优势。

但在某些情况下,文中算法仍存在局限性,比如当视频存在严重遮挡时,稳像的效果较差,主要是因为文中的运动模型估计仍是基于特征点提取的,当特征点严重不足时,会严重影响运动模型的估计精度。

3 结 论

文中提出了一种基于联合相机路径的红外视频稳像算法,该算法能在一定程度上解决针对红外视频稳像技术中存在的视差难题。相对于传统稳像算法,该算法的效果更好,对特征点存在部分遮挡的情况,也取得了不错的稳像效果。但是,当特征点严重不足或遮挡严重时,稳像的效果较差,主要是因为文中的运动模型估计仍是基于特征点提取;另外,也可以结合现在热门的神经网络技术融入视频稳像,旨在解决多场景下的稳像难题,这些都是今后重要的研究方向。

参考文献:

- [1] Li Bing. Research on key technology of infrared panoramic searching system[D]. Beijing: University of Chinese Academy of Sciences, 2017: 32-56. (in Chinese)
- [2] Zhang Haojun. Research on key technologies for infrared image restoration, video stabilization and imaging system based on embedded system[D]. Beijing: University of Chinese Academy of Sciences, 2012: 63-74. (in Chinese)
- [3] Yin Lihua. Study on key technology of panoramic image stabilization based on vehicle-mounted platform[D]. Beijing: University of Chinese Academy of Sciences, 2018: 10-50. (in Chinese)
- [4] Chen Xiaolun. Study on digital image stabilization technology for aerial opto-electric imaging system[D]. Beijing: University of Chinese Academy of Sciences, 2014: 57-89. (in Chinese)
- [5] Chen Y B, Lee K Y, Huang W T, et al. Capturing intention based full-frame video stabilization [J]. *Computer Graphics Forum*, 2008, 27(7): 1805-1814.
- [6] Matsushita Y, Ofek E, Tang X, et al. Full-frame video stabilization with motion inpainting. [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006, 28(7): 1150-1163.
- [7] Grundmann M, Kwatra V, Essa I, et al. Auto-directed video stabilization with robust L1 optimal camera paths[C]//IEEE Conference on Computer Vision and Pattern Recognition 2011. IEEE Computer Society, 2011: 225-232.
- [8] Zhu Juanjuan, Fan Jing, Guo Baolong. Adaptive electronic image stabilization algorithm resistant foreground moving object [J]. *Acta Photonica Sinica*, 2015, 44(6): 0610002. (in Chinese)
- [9] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]//IEEE International Conference on Computer Vision, 2011: 2564-2571.
- [10] Yu Jianguo, Xu Rentong, Chen Yu. Image stitching technology based on ORB improved RANSAC algorithm [J]. *Journal of Jiangsu University of Science and Technology (Natural Science Edition)*, 2015, 29(2): 164-169.
- [11] Qiu Jiatao. A study of electronic image stabilization algorithms and visual tracking algorithm[D]. Xi'an: Xidian University, 2013: 60-82. (in Chinese)
- [12] Xie Yajin, Xun Zhihai, Feng Huajun, et al. Real-time video stabilization based on minimal spanning tree and modified kalman filter [J]. *Acta Photonica Sinica*, 2018, 47(1): 65-74. (in Chinese)
- [13] Buehler C, Bosse M, McMillan L. Non-metric image-based rendering for video stabilization[C]//Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, 2001: 991019.
- [14] Liu F, Gleicher M, Jin H, et al. Content-preserving warps for 3D video stabilization[J]. *ACM Transactions on Graphics*, 2009, 28(3): 80032.
- [15] Liu F, Gleicher M, Wang J et al. Subspace video stabilization [J]. *ACM Transactions on Graphics*, 2011, 30(1): 4-10.
- [16] Huynh L, Choi J, Medioni G. Aerial implicit 3D video stabilization using epipolar geometry constraint[C]//2014 22nd International Conference on Pattern Recognition: IEEE, 2014: 3487-3492.
- [17] Liu S C, Yuan L, Tan P, et al. Bundled camera paths for video stabilization [J]. *ACM Transactions on Graphics*, 2013, 32(4): 1-10.
- [18] Guo H, Liu S, Zhu S, et al. Joint bundled camera paths for stereoscopic video stabilization[C]//2016 IEEE International Conference on Image Processing (ICIP), IEEE, 2016: 1071-1075.
- [19] Liu S, Yuan L, Tan P, et al. Steady flow: spatially smooth optical flow for video stabilization[C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014: 4209-4216.
- [20] Wang M, Guo Y, Jin K, et al. Deep online video stabilization with multi-grid warping transformation learning [J]. *IEEE Transactions on Image Processing*, 2019, 28(5): 2283-2292.