

文章编号: 1671-7449(2026)04-0435-14

J. Test Measur. Technol., 2026, 40(4): 435-448.

基于多相机阵列型仿生复眼的图像拼接方法

田科¹, 沈寅松², 刘晓晨^{1*}, 王晨光³, 赵慧俊⁴, 申冲^{1*}

(1. 中北大学 仪器与电子学院, 山西 太原 030051; 2. 中北大学 机电工程学院, 山西 太原 030051;
3. 中北大学 信息与通信工程学院, 山西 太原 030051; 4. 中北大学 电气与控制工程学院, 山西 太原 030051)

摘要: 随着自动驾驶、安防监控和虚拟现实技术的快速发展, 为了获得更广的视觉感知范围, 大视场成像系统的需求日益增长。受自然界昆虫复眼启发, 设计了一种基于多相机阵列型仿生复眼的成像系统, 该系统由8个相机组成, 采用球形圆环排列以模拟自然复眼结构, 可实现 $105^{\circ} \times 100^{\circ}$ 的大视场。针对大视场成像的图像拼接过程中易出现的拼接质量较差、拼接重影与色差等问题, 提出了一种基于凝聚型层次聚类(Agglomerative Hierarchical Clustering, AHC)引导的由粗到精图像拼接方法AHC-LightGlue。首先, 利用SuperPoint提取图像的关键点与描述符, 在特征匹配之前, 引入AHC算法构建特征点的拓扑结构, 在簇级别进行粗匹配以锁定对应区域; 随后, 在对应区域内调用LightGlue进行精细匹配, 解决误匹配率高、拼接错位的问题; 之后, 通过边缘化采样一致性算法计算单应性矩阵; 最后, 通过提出的自适应加权融合方法实现参考图像与目标图像间的自然过渡, 在亮度差异较大的情况下也能达到更好的融合效果, 减少大范围色差的问题, 进一步提升图像拼接效果。实验表明, 所提算法在图像质量和拼接精度方面均得到了提升, 并且能够有效地减少图像拼接中的颜色差异, 从而产生更平滑自然的拼接图像, 相较传统方法具有更强的鲁棒性与自适应性。

关键词: 仿生复眼; 大视场; 图像匹配; 图像融合

中图分类号: TP391.41 文献标识码: A DOI: 10.62756/csjs.1671-7449.2026058

引用格式: 田科, 沈寅松, 刘晓晨, 等. 基于多相机阵列型仿生复眼的图像拼接方法[J]. 测试技术学报, 2026, 40(4): 435-448.

Tian Ke, Shen Yinsong, Liu Xiaochen, et al. Image stitching method based on multi-camera array bionic compound eye[J]. Journal of Test and Measurement Technology, 2026, 40(4): 435-448.

Image Stitching Method Based on Multi-Camera Array Bionic Compound Eye

Tian Ke¹, Shen Yinsong², Liu Xiaochen^{1*}, Wang Chenguang³, Zhao Huijun⁴, Shen Chong^{1*}

(1. School of Instrument and Electronics, North University of China, Taiyuan 030051, China;
2. School of Mechanical and Electrical Engineering, North University of China, Taiyuan 030051, China;
3. School of Information and Communication Engineering, North University of China, Taiyuan 030051, China;
4. School of Electrical and Control Engineering, North University of China, Taiyuan 030051, China)

Abstract: With the rapid advancement of autonomous driving, security surveillance, and virtual reality

【收稿日期】2026-01-07 【修回日期】2026-03-02 【录用日期】2026-03-11

【基金项目】国家自然科学基金资助项目(62503437, 62503438); 山西省重点研发计划资助项目(202202020101002)

【作者简介】田科(2001—), 男, 硕士生, 主要从事图像处理研究。E-mail: sz202306155@st.nuc.edu.cn。

*【通信作者】刘晓晨(1995—), 男, 讲师, 博士, 主要从事深度学习视觉导航研究。E-mail: liuxiaochen@nuc.edu.cn。

申冲(1986—), 男, 教授, 博士, 主要从事仿生传感与智能导航研究。E-mail: shenchong@nuc.edu.cn。

technologies, the demand for wide-field-of-view imaging systems continues to grow to achieve broader visual perception ranges. Inspired by the compound eyes of insects in nature, a bionic compound eye imaging system is designed based on a multi-camera array. Comprising eight cameras arranged in a spherical ring to mimic the natural compound eye structure, a wide field of view of $105^{\circ} \times 100^{\circ}$ is achieved in this system. To address common issues in wide-field image stitching, such as poor quality, ghosting, and chromatic aberration, a coarse-to-fine image stitching method guided by agglomerative hierarchical clustering, named AHC-LightGlue, is proposed. Firstly, key points and descriptors are extracted from the images using SuperPoint. Before feature matching, the topological structure of feature points is constructed by the agglomerative hierarchical clustering algorithm. Coarse matching is performed at the cluster level to lock corresponding regions, followed by fine matching within these regions using LightGlue to address high mismatch rates and misalignment issues. Subsequently, the homography matrix is calculated by marginalizing sample consensus. Finally, a natural transition between the reference and target images is obtained by the proposed adaptive weighting fusion method. In this approach, superior fusion results are delivered even under significant brightness differences, large-scale color discrepancies are reduced, and image stitching quality is further enhanced. Experiments demonstrate that both image quality and stitching accuracy are improved by the proposed algorithm and color discrepancies are effectively reduced in image stitching to form smoother and more natural stitched images. Compared to traditional methods, the image stitching method based on a multi-camera array bionic compound eye has great robustness and adaptability.

Key words: bionic compound eye; wide field of view; image matching; image fusion

0 引言

大视场成像系统在自动驾驶、全景摄影和安防监控等应用中变得越来越重要^[1]。实现大视场的传统方法通常采用鱼镜头或超广角光学元件,这不可避免地引入严重的几何畸变,并损害边缘的图像质量。蜻蜓、蜜蜂等昆虫的复眼,不仅具备优异的视觉性能,包括大视场、对运动目标的高灵敏度,还拥有结构紧凑、体积小巧的形态优势^[2]。Tanida等^[3]的早期工作设计了名为TOMBO的仿生复眼系统,该系统由微透镜阵列、分离层和光电探测器阵列组成,形成一个小视场成像系统,是一款并置型仿生复眼。Floreano等^[4]提出了一种新型的CurvACE复眼系统,将透镜阵列、图像传感器和电路系统集成在一起,使得该成像系统的体积更小。除此之外,该复眼成像系统还具有镜头畸变小和成像视场较大的特点。Afshari等^[5]提出了PANOPTIC仿生复眼相机阵列,采用半球面CMOS阵列构建多视点光场传感系统,并基于FPGA实现了实时全向图像重建。Liu等^[6]搭建了由19个独立低成本相机组成的多相机成像系统,相机按特定角度分布,以中心相机为参考,通过旋转角度调整实现“无死角覆盖”,最终合成近极限视场。Qiao等^[7]采用了球面多排排列的

方式来模拟昆虫复眼结构的成像系统,系统由9个相机组成,排列在一个球面上,最终通过这种排列实现了大范围视场成像。Zheng等^[8]提出了一种基于双端协同优化的快速变焦、高分辨率稀疏复眼相机,通过4个相同的子相机找同一角度曲面排布构成复眼阵列,实现大视场快速变焦成像。现有复眼成像系统存在着大视场覆盖和高分辨率成像难以同时兼顾的问题,同时基于复眼相机阵列的高鲁棒图像拼接算法也相对缺乏。

如何准确对齐并融合多视角拍摄的重叠图像是为复眼相机阵列开发有效的图像拼接算法所面临关键技术挑战之一^[9-11]。传统图像拼接通常包含4阶段流程:特征检测、特征匹配、图像变换和图像融合。Lowe^[12]提出的尺度不变特征转换(Scale Invariant Feature Transform, SIFT)是最具代表性的算法,其对图像尺度变化与旋转变换均具备良好的不变性,但存在计算复杂度高的缺陷。加速稳健特征算法(Speeded Up Robust Features, SURF)是利用积分图等技术对SIFT算法实现了加速优化,在保持特征鲁棒性的同时提升了运算效率。定向FAST与旋转BRIEF特征算法结合了FAST角点检测和BRIEF描述子,运算速度进一步提升,但其不具备尺度不变性。此外,还有KAZE、BRISK等多种特征算法。在特征匹配环

节,常用的匹配策略包括暴力匹配与基于 FLANN 的快速近似最近邻搜索策略^[13]。为剔除错误匹配对,随机抽样一致算法(Random Sample Consensus, RANSAC)被广泛采用,可有效从含大量噪声的数据中精准估计图像变换模型。为了剔除错误的匹配对, RANSAC 被广泛用于从包含大量噪声的数据中估计变换模型。为了使拼接缝尽可能不明显,常用的融合技术包括线性加权、多频段融合和泊松融合^[14]。深度学习的最新进展显著改进了特征检测和匹配。Detone 等^[15]提出了一种自监督学习的 CNN 网络 SuperPoint,可以同时检测兴趣点并生成描述子,与传统方法相比,该方法表现出优异的可重复性和匹配精度。Sun 等^[16]提出了局部特征变换算法(Local Feature Transformer, LoFTR),是一种无检测器的匹配方法,该算法使用 Transformer 直接在粗略的特征图上建立像素级的密集匹配,对于无纹理区域的匹配效果尤为出色。Sarlin 等^[17]提出了 SuperGlue 网络,该网络对 SuperPoint 提取出的特征进行匹配,能结合上下文信息,提升在弱纹理、视角变化大等困难场景下的匹配准确率。Lindenberger 等^[18]通过 LightGlue 进一步改进了 SuperGlue,引入了自适应匹配深度,在保持精度的同时实现了更快的推理。Zheng 等^[19]提出了局部重构网络(Localin Reshuffle Net, LRNet),可以保持混合图像的平滑梯度域,并在拼接过程中传递局部纹理和光照。Liu 等^[6]基于 19 个相机组成的多相机成像系统,提出了一种实时、高精度且全视场的图像合成算法。但现有图像拼接算法在面对大视差条件及剧烈光照变化时,仍然易出现拼接质量较差、拼接重影与色差等问题。

受上述研究现状启发,本文针对图像拼接过程中易出现的问题,提出了一种为多相机阵列型仿生复眼设计的基于凝聚型层次聚类(Agglomerative Hierarchical Clustering, AHC)引导的由粗到精图像拼接方法 AHC-LightGlue。通过本文所提图像拼接方法得到的大视场成像可以适用在自动驾驶汽车的智能感知和公共场所的大范围安全监控等真实应用场景。

本文的主要创新点包括:

1) 借鉴昆虫复眼的成像原理,分析了仿生复眼大视场成像的特点,结合多相机成像特点,设计了一种基于多相机阵列型仿生复眼的成像系统。该系统由 8 个分辨率为 $1\,920 \times 1\,080$ 的相机

组成,球形圆环排列,模拟复眼结构。

2) 在 LightGlue 特征匹配之前,引入 AHC 算法,提出了一种基于 AHC 引导的由粗到精图像拼接方法 AHC-LightGlue,解决了误匹配率高、拼接错位的问题。

3) 提出了一种自适应加权融合方法(Adaptive Weighted Fusion, AWF),在图像亮度差异较大的情况下也能达到更好的融合效果,解决了大范围色差的问题,进一步提升了图像拼接效果。

1 复眼相机阵列系统

为了实现大视场成像,本文借鉴昆虫复眼的成像原理,构建了复眼视觉系统。多个单孔径成像系统以特定顺序排列以捕捉具有重叠视场角(Field of View, FOV)的子图像。考虑到相机的视场是矩形的,为了获得更大的整体视场,本文采用了球形圆环排列的多相机系统,其中每个子相机作为独立的光学系统运行^[2]。所有摄像机的光轴在球体中心相交,每个子眼捕获的图像被拼接在一起,最大限度地减少冗余信息,获得大视场图像,子眼的光轴和它们的 FOV 之间的角度决定了整个 FOV。

图 1 为球形圆环排列子眼的模拟示意图。

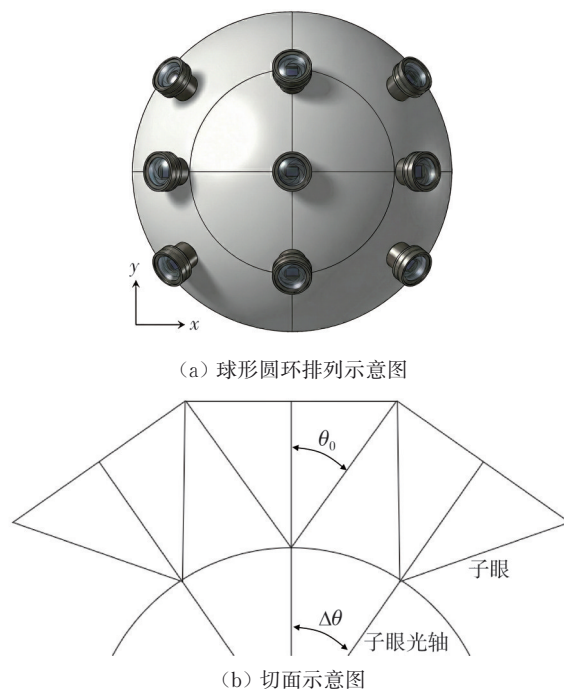


图 1 子眼排列设计

Fig. 1 Sub-eye arrangement design

如果相邻子眼之间的夹角为 $\Delta\theta$, 视场为 θ_0 , 则相邻子图像的重叠条件应满足 $\Delta\theta < 2\theta_0$ 。大视

场拼接的质量受拍摄对象到相机的距离的影响。如果对象太靠近系统,相互遮挡可能会降低图像缝合质量。为了实现有效的图像拼接和获得较宽的视场,物距应满足

$$Z \geq R \frac{\sin \theta_0 \cos \frac{\Delta \theta}{2}}{\sin \left(\theta_0 - \frac{\Delta \theta}{2} \right)}, \quad (1)$$

式中: R 为球壳的半径。子图像的重叠率 η_θ 由式(2)给出。

$$\eta_\theta = 1 - \frac{\Delta \theta}{2\theta_0 - 2\arcsin \frac{R \sin \theta_0}{Z}}. \quad (2)$$

为了获得较大的视场,本文设计并模拟了一种带有8个相机的球形排列,并使用SolidWorks软件对8个相机的视场进行了测量。图2为8个相机的FOV模型。图2(a)为具有32°垂直视场和60°水平视场的单个相机的视场。图2(b)和2(c)分别展示了水平和垂直两个方向上的重叠FOV模型。根据式(1),即当 Z 满足该条件时,整个成像系统获得了105°×100°的视场。

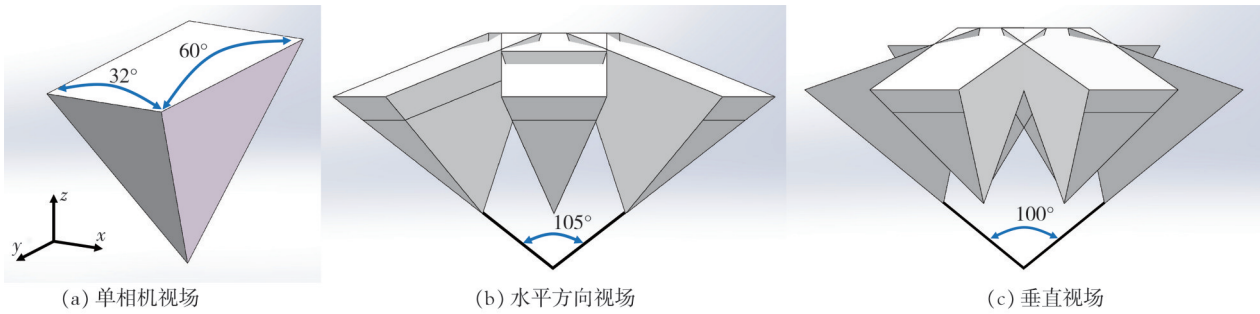


图2 相机视场模拟图

Fig. 2 Camera field of view simulation diagram

基于球面多排排列的原理,本文设计了一种仿生多相机复眼光学系统,图3为该系统的实物图,系统的球面半径为46 cm,外形尺寸为30 cm×30 cm。系统实物外壳通过光固化树脂3D打印成型,各个相机通过螺丝固定在相应的安装位置上。

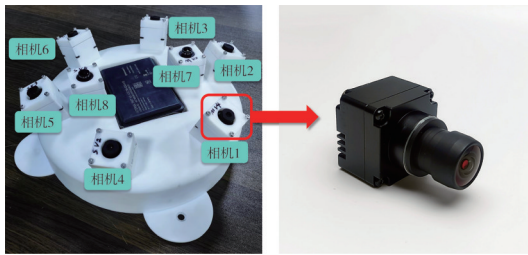


图3 复眼相机阵列系统

Fig. 3 Compound eye camera array system

2 图像匹配与拼接算法

面向上述设计的仿生复眼阵列系统,并针对仿生复眼图像中各子眼间视差大、色差严重且纹理重复度高的问题,以及将全图海量特征点输入匹配网络计算量大,且容易在重复纹理区域产生全局误匹配,本文提出一种由粗到精图像拼接方法AHC-LightGlue。在不改变LightGlue网络结构的前提下,通过引入AHC算法构建特征点的拓扑结构,先在簇级别进行粗匹配以锁定对应区域,随后在对应区域内调用LightGlue进行精细匹配。该方法有效地将全局搜索问题转化为若干个局部匹配子问题。

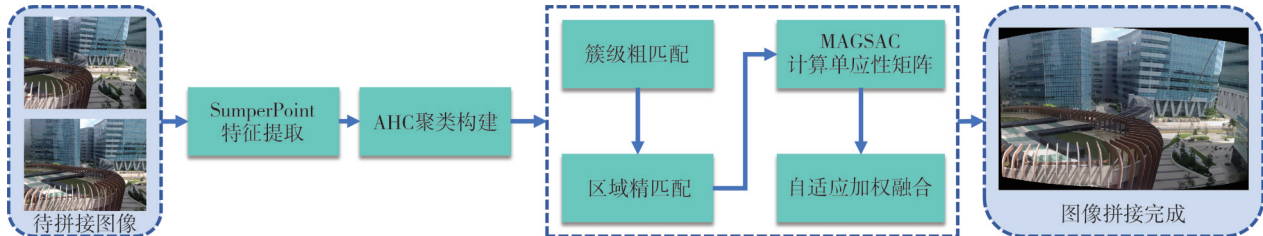


图4 整体算法流程图

Fig. 4 Overall algorithm flowchart

整体算法如图4所示,流程如下:

1) 特征提取: 利用SuperPoint提取图像的关键点与描述符;

2) AHC聚类构建: 对两幅图像的特征点分别执行结合几何与语义信息的AHC,生成特征簇;

3) 簇级粗匹配: 计算簇中心描述符的相似

度,建立簇与簇之间的对应关系,过滤无重叠背景;

4) 区域精匹配:提取匹配簇对内的原始特征点,输入标准LightGlue模型获取高精度匹配对;

5) 全局一致性优化:合并所有子块的匹配结果,通过边缘化采样一致性算法(Marginalizing Sample Consensus, MAGSAC)计算单应性矩阵。最后,通过AWF实现参考图像与目标图像间的自然过渡,完成拼接。

2.1 SuperPoint 特征提取模型

SuperPoint 网络采用全卷积架构实现端到端的特征点检测与描述符生成,结构图如图 5 所示。该网络采用 VGG 风格的编码器-解码器架构,主要由一个共享编码器和两个任务特定的解码器组成。SuperPoint 网络使用一个共享的编码器对输入图像 $I \in \mathbb{R}^{H \times W}$ 进行降维处理。其中,编码器包含卷积层、最大池化层和非线性激活函数,经过 3 次最大池化操作后,图像的空间分辨率被降低为

原始尺寸的 $1/8$ 。在特征点检测分支中,特征点检测解码器负责计算每个像素作为特征点的概率。为避免传统上采样带来的棋盘格伪影并减少计算量,SuperPoint 采用了显式解码设计。首先,输出一个维度为 $\mathbb{R}^{H_c \times W_c \times 65}$ 的张量,其中 64 个通道对应输入图像 8×8 像素的局部网格,1 个通道为“非特征点”丢弃通道,用于过滤背景区域。随后,通过通道维度的 Softmax 和子像素卷积操作,将特征图重塑为 $\mathbb{R}^{H \times W}$ 的全分辨率热力图。描述符生成分支中,描述子解码器生成半稠密的描述子图 $D \in \mathbb{R}^{H_c \times W_c \times D}$ 。为了在保持运行效率的同时获得全分辨率输出,网络使用双三次插值将描述子恢复至 $H \times W$ 尺寸,并进行 L2 归一化以生成单位长度的描述子向量。这种双任务协同设计使得模型能够同时学习定位鲁棒性和描述区分度。该网络采用轻量化的卷积模块构建,在保持较高特征质量的同时显著降低计算复杂度。全卷积的特性使其能够适应不同分辨率的输入,并通过单次前向传播完成所有计算。

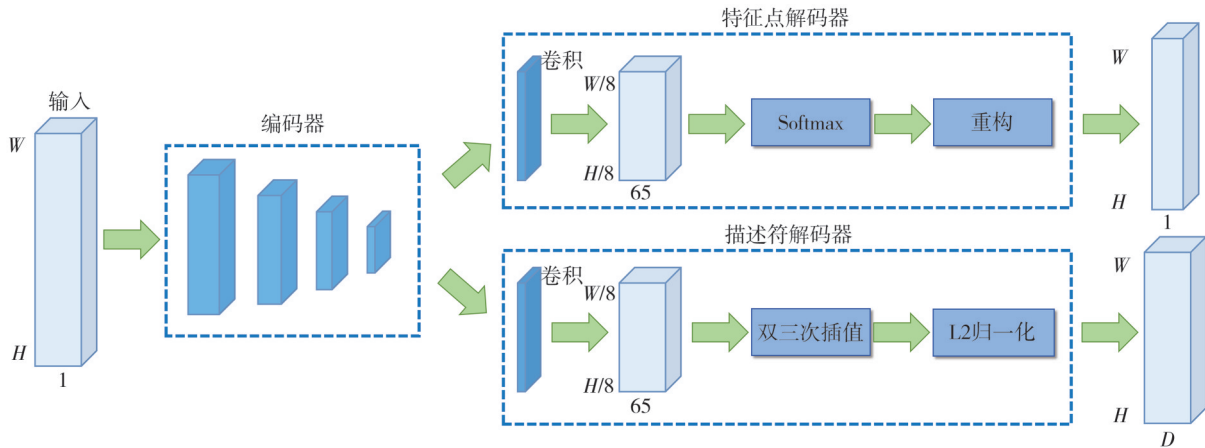


图 5 SuperPoint 网络结构图

Fig. 5 SuperPoint network structure diagram

2.2 基于 AHC 引导的由粗到精图像匹配方法

2.2.1 AHC 预处理

为了使聚类结果既在空间上紧凑,又在纹理上一致,定义特征点 i 和 j 之间的联合差异度 $D(i, j)$ 为

$$D(i, j) = \alpha \cdot \|d_i - d_j\|_2 + (1 - \alpha) \cdot \frac{\|p_i - p_j\|_2}{S_{\text{img}}}, \quad (3)$$

式中: d_i 和 d_j 为 SuperPoint 提取的 256 维描述符; p_i 和 p_j 为归一化像素坐标; S_{img} 为图像尺度因子; α

为平衡权重的超参数。该度量确保了同一簇内的点不仅物理距离相近,且属于同一语义物体。

采用 AHC 算法构建特征结构,与 K 均值聚类算法不同, AHC 不需要预设簇的数量,而是通过构建树状结构来适应复眼图像中大小不一的视场区域。聚类过程如下:首先初始化,将每个特征点视为一个独立簇;之后进行递归合并,每次寻找簇间距离 D_c 最小的两个簇进行合并,采用 Ward 连接准则以最小化簇内方差;最后进行动态剪枝,为了获取适应当前场景的“特征块”,设置距离阈值 τ 对层次树进行水平切割,得到最终的簇集合 $C = \{C_1, C_2, \dots, C_K\}$ 。

2.2.2 AHC-LightGlue 匹配

LightGlue算法源自Transformer注意力模型,其网络结构如图6所示。它主要分3个部分:第1部分是由自注意力机制和交叉注意力机制组

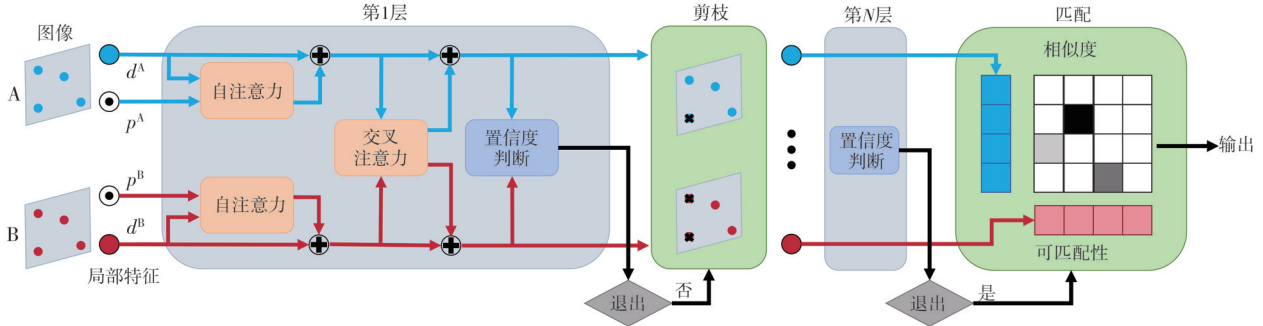


图6 LightGlue网络结构图

Fig. 6 LightGlue network structure diagram

簇级粗匹配:在调用LightGlue之前,先进行簇级别的快速关联。对于参考图的簇集合 $\{C_A\}$ 和目标图的簇集合 $\{C_B\}$,计算簇中心描述符之间的余弦相似度矩阵。若两个簇 C_A^m 和 C_B^n 的中心相似度高与阈值 T_{coarse} ,且满足互为最近邻约束,则判定这对簇为候选匹配区域对。这一步能够迅速剔除大量非重叠区域和明显的错误背景区域,大幅缩小了后续搜索空间。

局部LightGlue精匹配:针对每一对候选匹配簇 (C_A^m, C_B^n) ,提取其内部包含的原始特征点子集 P_A^m 和 P_B^n 。

将这两个子集输入到LightGlue网络中,

$$M_{mn} = \text{LightGlue}(P_A^m, P_B^n). \quad (4)$$

由于LightGlue擅长处理上下文信息,它能在这些局部子集中精确地找到点对点的对应关系。最后,进行匹配对融合,将所有局部匹配结果进行并集操作,得到全局匹配集。最终,通过本文方法可规避全图匹配时因纹理重复造成的远距离误匹配。

2.3 图像配准与自适应融合

经过AHC-LightGlue网络的特征匹配后,获得了一组高置信度的特征点对集合 $M = \{(p_i, p_j)\}_{k=1}^K$ 。然而,由于仿生复眼边缘区域的非线性畸变,直接计算单应性矩阵可能导致配准偏差。因此,本文采用基于USAC_MAGSAC的鲁棒估计策略,并设计AWF算法以生成无缝全景图。

2.3.1 基于USAC_MAGSAC的单应性矩阵估计

为了从匹配点集中精确解算出单应性矩阵

成的特征处理层,共有 N 层;第2部分是置信度判断层,位于特征处理层之后,主要负责决定是否退出特征处理层的迭代;第3部分是特征匹配层,通过轻量级检测头计算匹配矩阵并输出特征点的匹配关系。

$H \in \mathbb{R}^{3 \times 3}$,构建两幅子眼图像之间精准的二维平面射影映射关系。传统的RANSAC算法依赖于固定的内点阈值,这在噪声水平变化的复眼图像中往往难以选取最优值。

本文采用的USAC_MAGSAC算法结合了USAC的通用采样框架和MAGSAC的无阈值评分机制,其核心在于对噪声方差进行边缘化处理,从而消除了对硬阈值参数的依赖。

模型假设与采样:算法首先使用渐进一致采样策略,优先采样具有高匹配得分的特征点对,以加速模型假设的生成。

边缘化似然评分:对于生成的每一个假设模型 H_{hyp} ,MAGSAC不再通过简单的0/1计数来判断内点,而是计算所有数据点在不同噪声水平下的似然积分。对于第 k 个匹配点对,其残差为 $r_k = \|p_j^k - H_{hyp} p_i^k\|_2$ 。模型质量得分定义为

$$L(H_{hyp}) = \sum_{k=1}^K \int_0^{\sigma_{max}} P(r_k | \sigma) f(\sigma) d\sigma, \quad (5)$$

式中: σ 为特征点坐标的高斯噪声标准差; $P(r_k | \sigma)$ 为假设残差服从高斯分布时的概率密度; $f(\sigma)$ 为噪声标准差的先验分布。

模型精炼与输出:通过最大化上述似然函数,算法筛选出最优模型,并基于所有的加权内点进行最终的最小二乘优化,得到高精度的全局单应性矩阵 $\hat{H}$ 。

2.3.2 自适应加权融合方法

在复眼相机阵列系统中,每个子相机捕获的图像是碎片化的,只包含部分所需目标场景,因

此需要对其进行融合拼接来实现大视场图像输出。对于两个相邻子相机采集的图像,融合算法在重叠区域将对应的像素点按比例融合,对于非重叠区域则进行拼接操作。在实际应用中,不同相机位姿难以调节,视场重叠区域不一致,无法用一组权重系数实现所有子相机图像的拼接。为了解决这些问题,本文在传统渐变和非线性融合算法的基础上提出了AWF方法,利用自适应机制,通过动态权重融合两幅图像重叠区域的像素值,并根据亮度差异调整权重。该方法能有效消除拼接缝,在亮度差异较大的情况下也能达到更好的融合效果,进一步提升图像拼接精度。重叠区域的权重可以根据待融合两幅图像的亮度差异进行设置和调整。加权融合方法如图7所示。

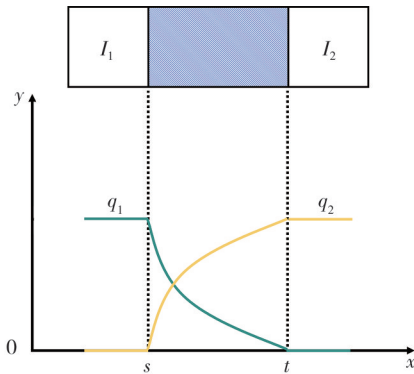


图7 自适应加权融合函数

Fig. 7 Adaptive weighted fusion function

其中, I_1 和 I_2 代表待拼接的两幅图像, 假设两幅待拼接图像的融合区域为 $[s, t]$ 。两幅图像的权重分别为 q_1 和 q_2 , 其融合函数如式(6)所示。

$$I(x, y) = \begin{cases} I_1(x, y), & 0 \leq x < s, \\ q_1 I_1(x, y) + q_2 I_2(x, y), & s \leq x \leq t, \\ I_2(x, y), & x > t. \end{cases} \quad (6)$$

在式(6)中, 权重 q_1 和 q_2 表示为

$$q_1 = \left(1 - \frac{x-s}{w}\right)^\beta, \quad (7)$$

$$q_2 = 1 - \left(1 - \frac{x-s}{w}\right)^\beta, \quad (8)$$

式中: w 为融合宽度; β 为一个随待合并图像亮度差异变化的参数, 可以表示为

$$\beta = 0.5 + 0.001 \times \gamma, \quad (9)$$

式中: γ 为待融合图像的亮度差异。在 RGB 颜色模式下,

$$\gamma = \frac{299 \times \Delta R + 587 \times \Delta G + 114 \times \Delta B}{1000}, \quad (10)$$

式中: $\Delta R, \Delta G, \Delta B$ 分别为两幅图像的 3 个原色差值。当两幅图像的亮度差异 γ 变化时, β 值也同时变化。与固定权重的非线性融合算法相比, 该权重不仅能根据非线性变化, 还能根据待融合图像间的亮度差异进行调整, 可用于亮度差异图像间的融合。可以看出, 改进后的权重系数在图像重叠区域内随着所处位置不同而平滑渐变, 从而实现参考图像与目标图像间的自然过渡。

3 实验和结果

3.1 数据集

在图像配准阶段, 本文使用经典的 HPatches 数据集^[20]进行实验, HPatches 数据集包含光照组和视角组图片, 并且提供真实单应性矩阵, 用于评估特征匹配在视角和光照变化下的鲁棒性。在图像拼接阶段, 使用 SEAGULL 数据集^[21]和本文所设计的仿生复眼系统所采集的图像进行实验。SEAGULL 数据集包含 24 对具有大视差、光照轻微变化特点的待拼接图像, 用于测试不同方法在真实复杂场景中的表现。本文的实验数据处理均在搭载 Intel Xeon (R) Platinum 8357B CPU @ 2.70 GHz 和 NVIDIA GeForce RTX 4090 的台式工作站进行。模型训练与参数设置方面, 本文基于 PyTorch 深度学习框架搭建网络。在训练模型时, 本文采用 Adam 优化器, 初始学习率设置为 1×10^{-4} , 并采用余弦退火策略进行学习率衰减, Batch Size 设为 8, 共训练 100 轮。

3.2 对比实验

3.2.1 图像配准实验

将本文方法与 SIFT、D2Net^[22] 和 SuperPoint+SuperGlue 方法进行比较。分别使用这 4 种方法处理 HPatches 的图像, 通过这 4 种方法获取特征点对进行参数估计, 获得单应性矩阵, 并与 HPatches 数据集中的单应性矩阵真值误差进行比较。通过比较单应性变换将原图像的 4 个角点映射到目标图像上的效果来进行评估。将原图像的 4 个角点定义为 c_1, c_2, c_3, c_4 。然后, 应用真实单应性矩阵 H 计算出在目标图像中的真实角点位置 c'_1, c'_2, c'_3, c'_4 。同时, 利用估计得到的单应性矩阵

$\hat{H}$ 计算出估计的角点位置 $\hat{c}'_1$ 、 $\hat{c}'_2$ 、 $\hat{c}'_3$ 、 $\hat{c}'_4$ 。设定像素误差阈值 ϵ 来判定估计出的单应性是否正确,单应性估计正确率(CorrH)的范围为0~1,数值越高表示效果越好,误差越小。

$$\text{CorrH} = \frac{1}{N} \sum_{i=1}^N \left(\left(\frac{1}{4} \sum_{j=1}^4 \|c'_{ij} - \hat{c}'_{ij}\| \right) \leq \epsilon \right). \quad (11)$$

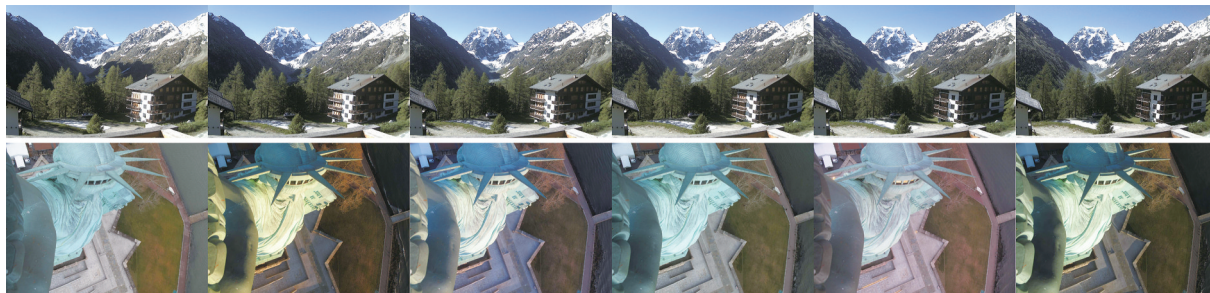


图8 HPatches数据集中的不同光照组示例

Fig. 8 Examples of different illumination groups in HPatches dataset



图9 HPatches数据集中的不同视角组示例

Fig. 9 Examples of different view groups in HPatches dataset

如表1所示,基于深度学习的D2Net、Super-GLue和本文方法优于基于SIFT的经典检测器方法。

表1 HPatches数据集上的图像匹配算法比较

Tab. 1 Comparison of image matching algorithms on HPatches dataset

方法	精度($\epsilon < 1/3/5$ pixel)			整体组精确率/%	整体组召回率/%	时间/s
	整体组	光照组	视角组			
SIFT	0.38/0.77/0.85	0.49/0.88/0.93	0.27/0.67/0.77	68.5	63.2	0.84
D2Net	0.44/0.77/0.86	0.66/0.94/0.97	0.23/0.61/0.75	78.9	74.5	1.64
SuperPoint+SuperGLue	0.51/0.81/0.88	0.62/0.91/0.95	0.41/0.72/0.81	85.4	89.7	0.14
本文方法	0.55/0.83/0.89	0.69/0.92/0.97	0.42/0.75/0.82	86.9	89.5	0.15

依托密集特征提取与匹配机制,即使在图像存在几何变换和光照差异的情况下,仍能实现特征点检测与匹配。在这些方法中,D2Net算法在 H 矩阵估计中精度较低,而本文方法比SuperPoint+SuperGLue方法在 H 矩阵估计中精度高,且在所对比的方法中精度最高。精度越高,意味着特征提取和图像匹配后获得的单应性矩阵越接近两者之间的实际位置关系,复原后的两幅图像也越接近真值图像。在表1中可以看出,本文方法比D2Net方法的精确率高约8%,召回率高约

HPatches数据集可以分为光照组和视角组。光照组是在同一视角和不同光照条件下拍摄的,光照组内同一组之间的 H 矩阵应为单位矩阵(如图8所示)。视角组是在同一光照条件下从不同视角拍摄的,视角组内同一组的 H 矩阵应不同(如图9所示)。

15%;比SuperGLue方法的精确率高约1.5%,召回率大致相当。因此,本文方法具有较高的精确率和召回率,匹配的质量更高。

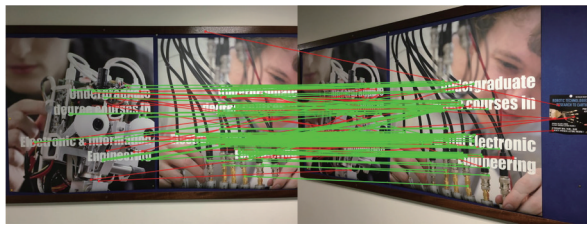
图10为HPatches数据集上的匹配效果对比图,图中绿色线表示正确匹配的特征点对,红色线表示误匹配的特征点对,可以看出,本文算法能保持较高数量的正确匹配点,且误匹配较少,具有匹配精度优势。

SuperPoint能够准确提取图像中的有效特征点,而LightGlue能够准确快速地匹配提取的特征

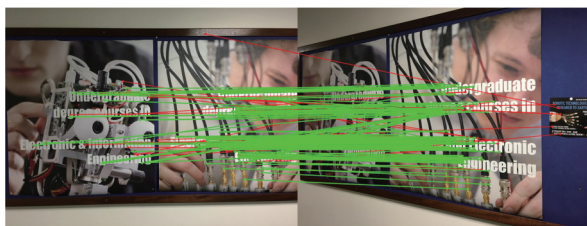
点以实现配准。这两种优秀算法的结合确保了图像在只有少量重叠部分时,融合图像也能被准确匹配。准确的配准为后续的图像融合奠定了良好的基础。



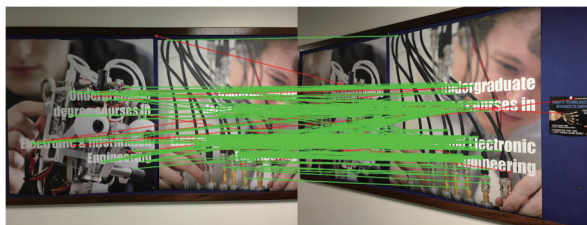
(a) SIFT



(b) D2Net



(c) SuperPoint+SuperGlue



(d) 本文方法

图 10 匹配效果对比

Fig. 10 Comparison of matching results

3.2.2 图像拼接实验

将本文方法与单透视扭曲算法(Single-Perspective Warps, SPW)^[23]、抗视差的无监督深度图像拼接算法(Parallax-Tolerant Unsupervised Deep Image Stitching, UDIS++)^[24]和LightGlue算法进行了定性和定量的比较。使用SEAGULL数据集和本文所设计的仿生复眼系统所采集的图像进行实验。

定量分析:在定量比较中,使用峰值信噪比(Peak Signal to Noise Ratio, PSNR)和结构相似性指数(Structural Similarity, SSIM)来比较算法。

PSNR是信号最大功率与信号噪声功率的比值,是评估图像质量最常用和最广泛的客观测量方法。两幅图像之间的PSNR值越大,表示两幅图像越相似。SSIM是衡量两幅图像相似度的指标,是从亮度、对比度和结构3个方面衡量两幅图像之间的相似性。SSIM的值范围为0~1,值越大表示图像失真越小。

计算了多种方法在SEAGULL数据集上的PSNR和SSIM的平均值,具体结果如表2所示。可以看出,本文方法在SEAGULL图像拼接数据集上的PSNR和SSIM均优于SPW、UDIS++和LightGlue算法。与UDIS++算法相比,PSNR指标平均提高了0.510,SSIM指标平均提高了0.020;与LightGlue算法相比,PSNR指标平均提高了0.166,SSIM指标平均提高了0.023,这表明本文方法所进行的优化有利于图像拼接的恢复。

表 2 SEAGULL数据集上的PSNR和SSIM

Tab. 2 PSNR and SSIM on SEAGULL dataset

	SPW	UDIS++	LightGlue	本文方法
PSNR	9.897	13.285	13.629	13.795
SSIM	0.418 8	0.455 2	0.451 8	0.475 2

定性分析:将本文方法的结果与SPW和UDIS++算法进行了比较,如图11所示。可以看出,由SPW拼接的图像会出现重影的现象,如图中红框标注的区域内房屋存在重影;由UDIS++拼接的图像会出现部分物体错位的现象,如图中红框标注的区域内树枝上下错位,严重影响了观感;由LightGlue拼接的图像会出现部分物体扭曲和错位的现象,如图中红框标注的区域内屋顶错位及树枝的变形扭曲,严重影响了观感;而本文算法拼接的图像在同样的区域不会出现重影或错位等影响图像拼接质量的问题。

图12为本文所设计的仿生复眼系统所采集的部分图像,其中左右各两幅图像分别为相机1和相机3所拍摄的图像。图13为在这些图像上进行SPW、UDIS++和本文方法的对比实验。从图13中可以看出,对于复眼所采集的图像,由SPW算法拼接的结果会出现明显的重影与畸变;由UDIS++和LightGlue算法拼接的结果也会出现明显的物体扭曲;而本文通过引入AHC算法和AWF方法,提高了图像拼接质量,因此,由本文算法拼接的图像在同样的区域不会出现重影或物体扭曲等问题。

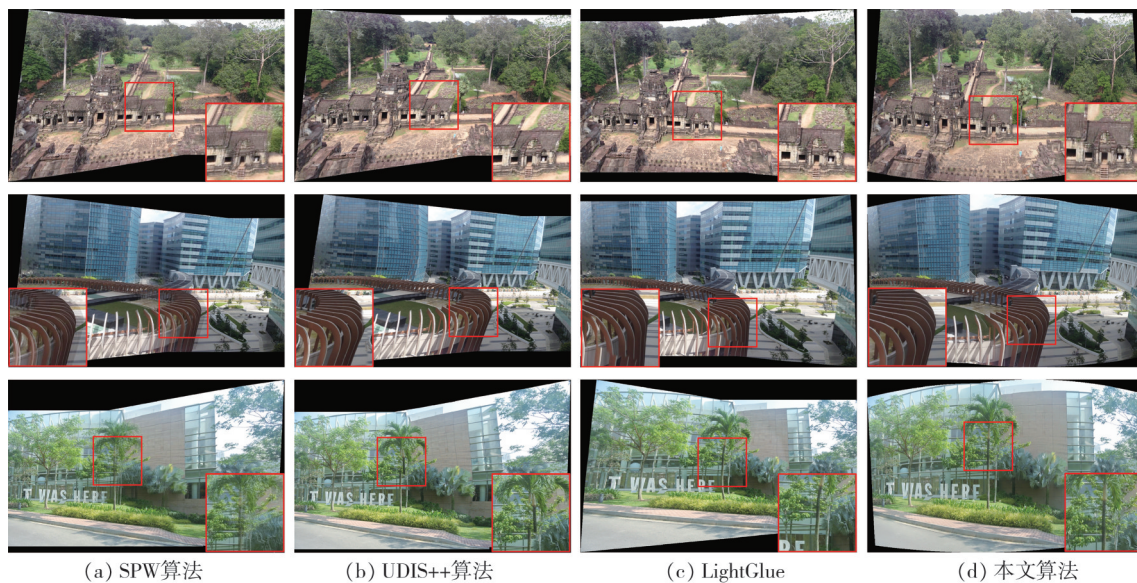


图 11 不同算法在SEAGULL数据集中对比结果

Fig. 11 Comparison results of different algorithms in SEAGULL dataset



图 12 仿生复眼系统所采集的部分图像

Fig. 12 Images captured by the bionic compound eye system

计算多种方法在某校园环境复眼相机数据集上的PSNR和SSIM的平均值,具体结果如表3所示。

可以看出,本文方法在仿生复眼图像上的PSNR和SSIM均优于SPW、UDIS++和LightGlue算法。与LightGlue算法相比,PSNR指标平均提高了0.485,SSIM指标平均提高了0.0236,这表明本文所做的优化有利于通过图像拼接恢复原图像。

表 3 仿生复眼图像中的PSNR和SSIM

Tab. 3 PSNR and SSIM on bionic compound eye images

	SPW	UDIS++	LightGlue	本文方法
PSNR	10.963	13.272	13.841	14.326
SSIM	0.4318	0.4625	0.4885	0.5121

综上所述,无论是定性还是定量分析实验中,都可以看出使用本文方法获得的拼接结果优于经典SPW算法,以及基于深度学习的UDIS++和LightGlue算法。

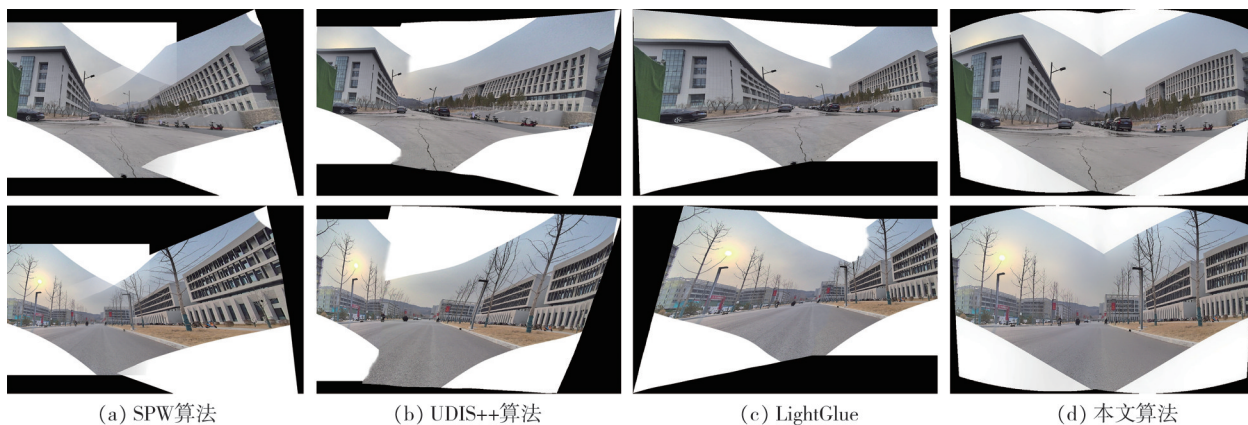


图 13 不同算法在仿生复眼图像中对比结果

Fig. 13 Comparison results of algorithms in bionic compound eye images

3.3 消融实验

为了证明本文所提算法各部分对整体算法优化的贡献,本节进行消融实验,将算法的各部分与原始算法的其余部分相结合,并进行了定量和定性实验。

定量分析:定量实验结果如表4和图14所示。算法分为3部分:SuperPoint+LightGlue、凝AHC和AWF算法,通过不同算法组合设计了4个定量分析实验。

从表4中的结果可以看出:在SuperPoint和LightGlue的基础上加入AHC后,通过先粗后精的匹配方法,剔除误匹配点,使得PSNR和SSIM都有所提高,表明聚类算法对于图像匹配具有积极作用。

在SuperPoint和LightGlue的基础上加入AWF后,通过动态权重融合,使得SSIM也有所提高。在SuperPoint和LightGlue的基础上同时加入AHC和AWF算法后,PSNR和SSIM数据均优于单独加入这两种算法的情况。这表明本文加入的改进算法,能够相互补充,获得更好的拼接结果。

定性分析:定性实验结果如图14所示。

表4 在SEAGULL数据集上进行消融实验的PSNR和SSIM
Tab. 4 PSNR and SSIM for ablation experiments on the SEAGULL dataset

方法	PSNR	SSIM
SuperPoint、LightGlue	13.629	0.451 8
SuperPoint、LightGlue+AHC	13.693	0.463 8
SuperPoint、LightGlue+AWF	13.603	0.471 3
SuperPoint、LightGlue+AHC+AWF	13.795	0.475 2



图14 SEAGULL数据集上的消融实验

Fig. 14 Ablation experiment on SEAGULL dataset

从图14中可以看出,仅使用SuperPoint和LightGlue算法时,拼接结果存在明显的色差(红框)和拼接错位(黄框);在单独加入AHC算法后,通过先粗后精的匹配方法优化,黄框中的错位拼接得到了明显的修复,但是图像仍然存在色差;在单独加入AWF算法后,通过动态权重融合,图像中的色差得到了明显改善,但仍存在拼接错位

的问题;使用本文算法时,即在SuperPoint和LightGlue的基础上同时加入AHC和AWF算法,通过误匹配点剔除和动态权重融合,图像拼接错位和色差都得到了明显改善。

为验证所提方法在实际环境下的实用性和有效性,本文采集了复眼相机数据集,数据集的采集地点为校园环境内,采集数量为25组,每组包

含8个子眼所拍摄的图像。在仿生复眼所采集的

图像上进行的消融实验结果如图 15 所示。



(a) SuperPoint、LightGlue



(b) SuperPoint、LightGlue+AHC



(c) SuperPoint、LightGlue+AWF



(d) SuperPoint、LightGlue+AHC+AWF

图 15 仿生复眼图像上的消融实验

Fig. 15 Ablation experiment on bionic compound eye images

从图 15 中可以看出,仅使用 SuperPoint 和 LightGlue 算法时,拼接结果存在明显的拼接重影错位(红框),并且图像中左右两侧存在色差;在单独加入 AHC 算法后,红框中的房屋建筑错位拼接得到了改善,但是图像中左右两侧仍然存在色差;在单独加入 AWF 算法后,图像中的色差得到了改善,但是红框中的房屋建筑仍存在拼接错位的问题;使用本文算法时,通过 AHC 和 AWF 算法,有效地减少了图像拼接中的颜色差异,生成了更平滑自然的拼接图像,提高了图像拼接质量。

综上所述,本文方法通过逐步引入 AHC 和 AWF 模块,使得图像拼接精度逐步提升,表明所提出的各个模块具有一定的有效性和可行性。

3.4 复眼系统拼接结果图

最后利用本文算法对复眼相机拍摄的图像进行图像拼接,图 16 为复眼中各个子眼所拍摄的图像,其中,图中从上至下,从左至右分别是相机 1、2、3、4、5 和 6 所拍摄的图像,图 17 为经过本文算法拼接完成的图像。

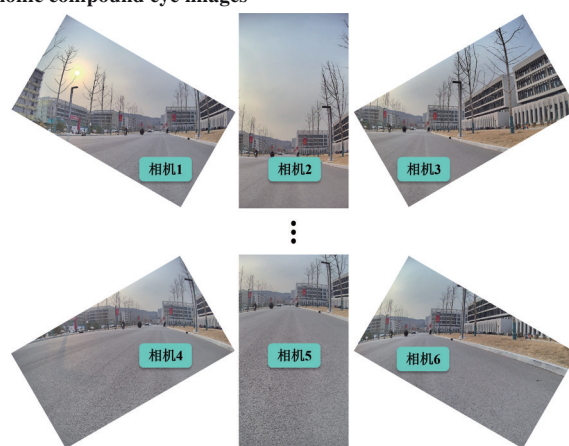


图 16 复眼中各个子眼所拍摄的图像

Fig. 16 Images captured by sub-eye in compound eye

从图 17 中可以看出,通过本文所提 AHC-LightGlue 图像拼接算法,图像中没有明显的拼接错位且图像整体色彩均衡,没有大范围色差,图像融合自然,进一步证明了本文提出的基于多相机阵列型仿生复眼的图像拼接方法的可行性和有效性。



图 17 复眼系统拼接结果图

Fig. 17 Stitching result image of compound eye system

4 结 论

针对现有图像拼接算法在复眼场景下存在的误匹配率高、拼接错位和大范围色差的问题,本文提出了一种基于AHC引导的由粗到精图像拼接方法AHC-LightGlue,采用球形圆环排列的多摄像机系统,构建了复眼视觉系统,实现了 $105^{\circ} \times 100^{\circ}$ 的大视场。在图像拼接部分使用了能高效准确提取特征点的SuperPoint。在LightGlue特征匹配之前,引入AHC算法构建特征点的拓扑结构。首先,在簇级别进行粗匹配以锁定对应区域;随后,在对应区域内调用LightGlue进行精细匹配;之后,通过MAGSAC计算单应性矩阵;最后,通过AWF方法实现参考图像与目标图像间的自然过渡,完成拼接。实验结果表明,本文所提方法在图像质量和拼接精度方面均得到了提升,并且能够有效地减少图像拼接中的颜色差异,从而生成了更平滑自然的拼接图像。

虽然本文方法在图像拼接质量上取得了一定进展,但在实际应用中仍存在以下局限性:深度学习图像拼接框架的计算强度对硬件成本要求较高,以及在剧烈光照变化等恶劣环境下的适用性有限。在未来的工作中,需要进一步探索模型轻量化技术,可采用模型剪枝、量化等方法,或引入光照不变特征提取模块,以降低计算复杂性,使算法可部署在低成本的边缘设备上,并增强算法的环境鲁棒性,以将其应用于更广泛的场景中。

利益冲突声明/Conflict of Interests

所有作者声明不存在利益冲突。

All authors declare no relevant conflict of interests.

参考文献

- [1] Fu M, Liang H, Zhu C, et al. Image stitching techniques applied to plane or 3-D models: a review[J]. IEEE Sensors Journal, 2023, 23(8): 8060-8079.
- [2] Qi Q, Fu R, Shao Z, et al. Multi-aperture optical imaging systems and their mathematical light field acquisition models[J]. Frontiers of Information Technology & Electronic Engineering, 2022, 23(6): 823-844.
- [3] Tudela R, Brückner A, Duparré J, et al. An image restoration approach for artificial compound eyes[J]. Image Processing: Algorithms and Systems VI, 2008, 6812: 68120O.
- [4] Floreano D, Pericet-Camara R, Viollet S, et al. Miniature curved artificial compound eyes[J]. Proceedings of the National Academy of Sciences of the United States of America, 2013, 110(23): 9267-9272.
- [5] Afshari H, Jacques L, Bagnato L, et al. The PANOPTIC camera: a plenoptic sensor with real-time omnidirectional capability[J]. Journal of Signal Processing Systems, 2013, 70(3): 305-328.
- [6] Liu S, Wang J, Yuan R, et al. Real-time and ultrahigh accuracy image synthesis algorithm for full field of view imaging system[J]. Scientific Reports, 2020, 10: 12389.
- [7] Qiao Z, Hao Q, Liu M, et al. Multicamera wide-field-of-view compound eye imaging system [C]// Optoelectronic Imaging and Multimedia Technology XI, 2024, 13239: 132391R.
- [8] Zheng Y, Zhang H, Li X, et al. Fast-zoom and high-resolution sparse compound-eye camera based on dual-end collaborative optimization [J]. Opto-Electronic Advances, 2025, 8(6): 240285.
- [9] Ge R, Liu D, Zhou H, et al. Real-time instance segmentation based on contour learning [J]. Journal of Measurement Science and Instrumentation, 2024, 15(3): 328-337.
- [10] 郭亚楠, 陈燕, 王康谊, 等. 基于内容自适应分块的眼底图像无损压缩算法[J]. 中北大学学报(自然科学版), 2025, 46(4): 430-437.
Guo Yanan, Chen Yan, Wang Kangyi, et al. Lossless fundus image compression algorithm based on content adaptive blocking [J]. Journal of North University of China (Natural Science Edition), 2025, 46(4): 430-437. (in Chinese)
- [11] 姬文芳, 朱子文, 邓德志, 等. 基于改进SIFT算法的城市航拍图像快速拼接方法[J]. 测试技术学报,

- 2024, 38(5): 500-505.
- Ji Wenfang, Zhu Ziwen, Deng Dezhi, et al. Rapid urban aerial image stitching method based on improved SIFT algorithm[J]. Journal of Test and Measurement Technology, 2024, 38(5): 500-505. (in Chinese)
- [12] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [13] Muja M, Lowe D G. Scalable nearest neighbor algorithms for high dimensional data[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 36(11): 2227-2240.
- [14] Xiu C, Fang J, Zhang J. Image stitching method based on adaptive weighted fusion[C]//2021 33rd Chinese Control and Decision Conference (CCDC), IEEE, 2021: 3099-3103.
- [15] Detone D, Malisiewicz T, Rabinovich A. SuperPoint: self-supervised interest point detection and description [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2018: 337-33712.
- [16] Sun J, Shen Z, Wang Y, et al. LoFTR: detector-free local feature matching with transformers [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021: 8918-8927.
- [17] Sarlin P E, Detone D, Malisiewicz T, et al. Superglue: learning feature matching with graph neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020: 4937-4946.
- [18] Lindenberger P, Sarlin P E, Pollefeys M. Lightglue: local feature matching at light speed[C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 17581-17592.
- [19] Zheng C, Xia S, Robinson J, et al. Localin reshuffle net: toward naturally and efficiently facial image blending [C]//Computer Vision-ACCV 2020. Cham: Springer, 2020: 206-222.
- [20] Balntas V, Lenc K, Vedaldi A, et al. HPatches: a benchmark and evaluation of handcrafted and learned local descriptors[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 3852-3861.
- [21] Lin K, Jiang N, Cheong L F, et al. SEAGULL: seam-guided local alignment for parallax-tolerant image stitching [C]//Computer Vision-ECCV 2016. Cham: Springer, 2016: 370-385.
- [22] Dusmanu M, Rocco I, Pajdla T, et al. D2-net: a trainable CNN for joint description and detection of local features [C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 8084-8093.
- [23] Liao T, Li N. Single-perspective warps in natural image stitching[J]. IEEE Transactions on Image Processing, 2020, 29: 724-735.
- [24] Nie L, Lin C, Liao K, et al. Parallax-tolerant unsupervised deep image stitching [C]//2023 IEEE/CVF International Conference on Computer Vision (ICCV), 2023: 7365-7374.