

〈图像处理与仿真〉

## 基于几何约束下区域搜索的红外可见光双目配准算法

王贺松, 张 灿, 蔡 朝, 黄 琨, 樊 凡

(武汉大学 电子信息学院, 湖北 武汉 430072)

**摘要:** 针对相对位置固定的红外和可见光双目相机配准任务, 现有算法没有考虑到两者相对位置固定的先验知识, 存在配准精度低、几何定位差异大等问题, 适用性差。提出一种基于几何约束下区域搜索的红外可见光双目图像配准方法。首先借助红外和可见光双目相机的标定信息对红外和可见光图像进行立体校正使二者处于同一高度之上。接着借助于相位一致性计算红外与可见光的边缘特征图, 然后在红外边缘图上提取特征点, 最后提出两阶段的同名特征点搜索方法, 以红外特征点为基准在可见光边缘图局部区域内搜索同名特征点。在第一阶段以归一化互相关(Normalized cross-correlation, NCC)为相似性度量计算两边缘图的整体水平偏移, 预测同名特征点初始位置, 在第二阶段提出多尺度加权NCC作为相似性度量, 在初始同名特征点位置周围精确搜索同名特征点。在构造的真实环境数据集上进行实验, 实验结果表明相对于其他对比算法, 在特征点匹配数量和准确率以及主观视觉上的配准效果都优于其他对比算法。

**关键词:** 图像配准; 红外图像; 可见光图像; 相位一致性; 双目相机

**中图分类号:** TP391

**文献标志码:** A

**文章编号:** 1001-8891(2024)03-0269-11

## Infrared and Visible Binocular Registration Algorithm Based on Region Search Under Geometric Constraints

WANG Hesong, ZHANG Can, CAI Zhao, HUANG Jun, FAN Fan

(Electronic Information School, Wuhan University, Wuhan 430072, China)

**Abstract:** For the registration task of infrared and visible binocular cameras with fixed relative positions, existing algorithms do not consider the prior fixed relative positions of the two cameras, resulting in problems, such as low registration accuracy, large differences in geometric positioning, and poor applicability. An infrared and visible binocular image registration method based on region search under geometric constraints. First, stereo correction was performed on the infrared and visible images using the calibration information of the infrared and visible binocular cameras, such that they were at the same height. Second, infrared and visible edge maps were obtained using phase congruency and feature points were extracted from the infrared edge map. Finally, a two-stage feature point search method is proposed to search for feature points with the same name in the local area of the visible edge map based on the infrared feature points. In the first stage, normalized cross-correlation (NCC) was used as a similarity metric to calculate the overall horizontal offset of the two edge maps, and the initial positions of feature points with the same name were predicted. In the second stage, a multiscale-weighted NCC was proposed as a similarity metric to accurately search for feature points with the same name around the initial location of feature points of the same name. Then, experiments were performed on the constructed real-environment dataset. The experimental results show that compared with other comparison methods, the number and accuracy of matching points and registration results in subjective vision are better.

收稿日期: 2022-12-06; 修订日期: 2023-03-17.

作者简介: 王贺松(1997-), 男, 安徽亳州人, 硕士研究生, 研究方向为红外与可见光配准。E-mail: wanghesong@whu.edu.cn.

通信作者: 樊凡(1989-), 男, 江西南昌人, 博士, 副教授, 研究方向为红外图像处理。E-mail: fanfan@whu.edu.cn.

基金项目: 国家自然科学基金项目(62003247, 62075169, 62061160370); 湖北省重点研发计划(2021BBA235); 珠海市基础与应用基础研究基金(ZH22017003200010PWC)。

**Key words:** image registration, infrared image, visible image, phase congruency, binocular camera

## 0 引言

在电力巡检、视频监控、无人驾驶、机器人视觉等领域中<sup>[1-4]</sup>,常利用红外和可见光双目成像,将红外和可见光图像进行信息互补,这样可以提高后续应用的性能,如提高检测、识别等任务的精度<sup>[5]</sup>。而通过准确高效地配准将两幅图像进行对齐是有效利用红外和可见光互补信息的一个重要的前提条件,因此,研究红外和可见光双目成像中红外和可见光图像配准具有重要意义。

红外图像与可见光图像由于模态不同,图像存在显著差异,配准难度大。红外图像根据目标的热辐射成像,能够排除光线和天气条件的影响,但通常分辨率较低,缺少纹理细节。可见光图像利用光反射成像,成像分辨率比较高,纹理细节丰富,明暗对比度高,但同时也更容易受到环境因素的影响。目前,已有的红外与可见光图像配准算法中,主要可以被分为基于区域、基于特征和基于学习3类<sup>[6-7]</sup>。

基于区域的配准方法通常给定预变换模型。利用相似性度量和优化方法来估计模型的变换参数,通过优化总体损失函数来对齐两幅图像的公共区域。例如,Yu等人<sup>[8]</sup>通过灰度加权窗口策略从红外和可见光图像中检测边缘结构,通过归一化互信息提高配准性能。Yedukondala等人<sup>[9]</sup>同样将红外和可见光图像转化为边缘图,使用仿射和自由形态变换(Free-Form Deformation, FFD)来解释粗配准和精配准的几何关系,通过最大化互信息(Mutual Information, MI)度量将红外和可见光图像进行配准。上述基于区域算法的算法很大程度受图像噪声和几何畸变的影响,红外和可见光之间严重的非线性强度差异,导致基于区域的方法配准精度不高,目前在实际中应用较少。

基于特征的方法是在提取图像特征点的基础上,通过特征描述算子描述特征,利用相似度量准则获得特征点集合间的匹配关系达到配准的目的。目前应用最为广泛,也最为精确。例如,Wang等人<sup>[10]</sup>提出了SURF-PIIFD-RPM的鲁棒点匹配框架,该方法结合SURF<sup>[11]</sup>(Speeded Up Robust Features, SURF)特征提取与部分强度不变特征描述符(Partial Intensity Invariant Feature Descriptor, PIIFD)<sup>[12]</sup>,利用单高斯鲁棒点匹配模型获得了更好的匹配性能。Jiang等人<sup>[13]</sup>提出的CAO-C2F算法利用曲率尺度空间(Curvature Scale Space, CSS)角点检测器轮廓中的特征点,借助SIFT描述符进行特征点的匹配,实现了红外与可见光

电力图像的配准。Li等人<sup>[14]</sup>提出了一种基于辐射变化不敏感特征的多模配准方法(Radiation-Variation Insensitive Transform, RIFT)。在基于相位一致性得到的角点和边缘特征点上,建构造多个方向的最大索引图实现特征旋转不变性。这些算法在匹配红外和清晰的可见光图像时表现出了良好的性能。但当图像质量下降如暗光环境时,图像细节不清晰,特征点之间相似性减弱,配准性能则会产生明显下降。

近年来,深度学习网络也被应用到红外和可见光图像配准任务之中。例如,Wang等人<sup>[15]</sup>提出了一种两阶段的对抗网络,包括模态转换和几何变换阶段,将可见光图像转换为红外图像并获得精确的配准图像;Arar等人<sup>[16]</sup>提出了一种基于GAN(Generative Adversarial Network)网络的无监督的配准框架,在红外和可见光上同时训练图像模态转换网络和配准网络,借助图像模态转换进而通过使用成熟的单模态度量来训练网络。Deng等人<sup>[17]</sup>利用深度学习进行特征点检测和描述,提出了ReDFeat网络,将多模态特征的检测和描述与互加权策略结合起来提高了多模态特征训练的稳定性和性能。由于红外和可见光数据集的缺少,基于深度学习的配准算法存在着泛化能力不足的问题。

在红外和可见光双目相机中,由于两个相机相对位置固定,即红外和可见光相机成像的相对角度、相对位置等固定。而现有的红外和可见光配准算法没有考虑到两者相对位置固定的先验知识,在应用于红外可见光双目相机时,存在配准精度低、几何定位差异大等多模态配准的问题。

针对上述问题,本文提出一种基于几何约束下区域搜索的红外可见光双目图像配准框架。首先对红外相机和可见光双目相机进行双目标定,通过立体校正将两幅图像约束到同一水平高度,接着利用相位一致性特征提取红外和可见光的边缘图像,在边缘图上提取稳定的红外特征点,然后提出了两阶段的同名特征点搜索方法进行特征点搜索,第一阶段以归一化互相关作为相似度量估计两幅图像的整体水平偏移,从而预测同名特征点位置,第二阶段提出多尺度加权NCC作为相似性度量方式,以红外图像特征点为基准在可见光边缘图中进行区域同名特征点搜索得到匹配点,最后通过迭代式细化的方式剔除匹配误差较大的点,得到更加精确的配准模型。

## 1 配准算法

考虑到红外和可见光相对位置固定, 两个相机的内参和外参可以通过双目相机标定获取, 从而约束红外和可见光同名特征点的相对位置, 即可以预测待匹配图像同名特征点的相对位置, 进而可以在预测的同名特征点周围局部区域精确搜索得到最终的匹配特

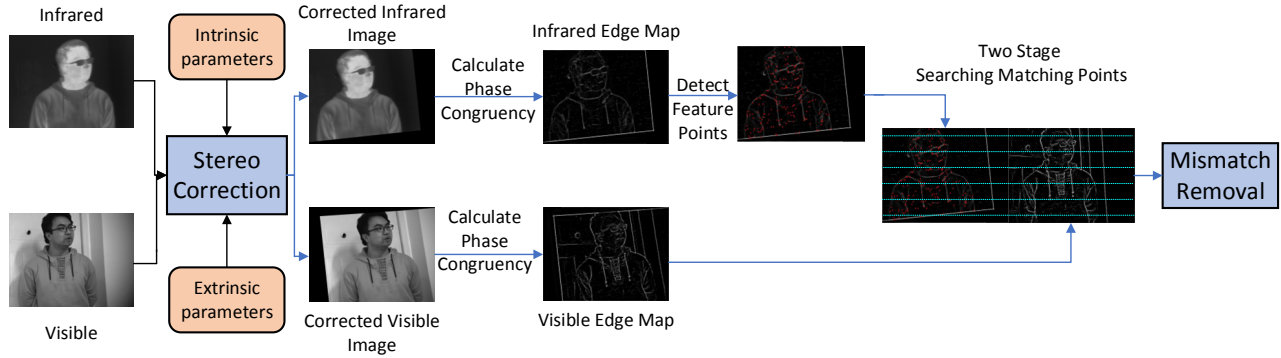


图 1 基于几何约束下区域搜索的红外可见光图像配准方法框架

Fig.1 Pipeline of infrared and visible image registration based region search under geometric constraints

2) 基于相位一致性获取图像的边缘特征图。在 1) 校正后图像的基础上, 计算红外和可见光图像的相位一致性响应图, 并计算其最大矩获得鲁棒的边缘特征图像。

3) 红外边缘图特征点提取。在红外边缘图像进行 FAST 特征点的提取, 并剔除掉原红外图像中背景中的无用特征点。同时使用非极大值抑制策略, 使得到的特征点分布更加均匀。

4) 两阶段特征点搜索匹配。第一阶段以 NCC 为衡量相似性指标计算两边缘图整体水平偏移参数, 从而预测同名特征点初始位置, 进一步减少后续同名特征点的搜索范围, 降低时间复杂度。接着以 3) 得到的红外特征点为基准, 在可见光边缘图的搜索区域内, 提出多尺度加权 NCC 作为相似性衡量指标, 搜索最佳匹配点。

5) 误匹配的剔除。首先通过最小二乘法建立所有匹配特征点的变换模型, 然后计算所有匹配点的误差和均方根误差 (Root mean square error, RMSE), 剔除误差较大的匹配点, 迭代上述过程提高配准精度。

### 1.1 红外相机与可见光双目相机标定

在红外与可见光双目成像中, 两个相机相对位置固定, 双模图像间的旋转平移等几何变换关系相对固定, 可以通过双目相机标定预先估计。换言之, 我们可以通过双目相机标定方法获取两个相机之间的内部光学参数 (内参), 以及相机之间的旋转和平移参数 (外参)。为了计算红外可见光双目相机的内参和外参, 我们可以参考典型的相机标定方法——张正友

征点, 提高匹配精度与效率。因此, 本文提出几何约束下基于区域搜索的红外和可见光图像配准方法, 算法流程如图 1 所示。

1) 红外和可见光图像立体校正。利用相机标定得到的红外和可见光双目相机的内参和外参, 通过立体校正可将图像约束到同一高度之上。

标定法<sup>[18]</sup>, 该方法借助不同角度的多张黑白棋盘格图片, 利用最小二乘法计算双目相机的内参和外参。双目相机的内参和外参只和相机本身的光学特性以及相对位置有关, 所以针对一套红外和可见光双目相机只需要标定一次, 因此相机内参和外参的计算并不会增加红外与可见光双目配准的处理时间。

虽然张正友标定法操作简单、精度高, 但主要用于可见光相机或者深度相机中。而在红外和可见光双目成像中, 两者的模态不同, 直接在纸上打印的黑白棋盘格无法同时在两个相机中都顺利成像。因此我们采用特制的氧化铝黑白棋盘格标定板, 前面板是氧化铝做成的黑色棋盘格, 后面板是玻璃基板。当对标定板背面均匀加热时, 两种材料导热性与反射率不同, 使得标定板在红外图像和可见光图像上都会形成黑白分明的棋盘格图像, 如图 2 所示。然后就可通过张正友标定法获取红外与可见光双目相机的内参与外参。

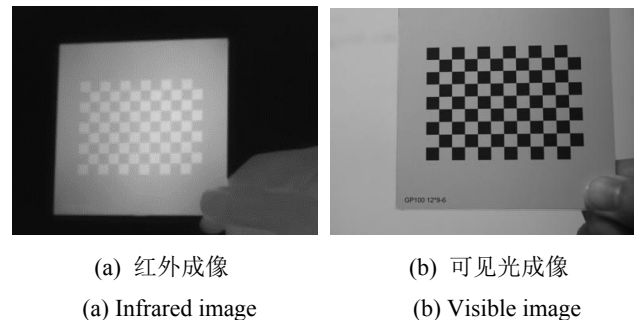


图 2 棋盘格标定板红外和可见光成像

Fig.2 Infrared and visible images of chess-board

在得到了红外和可见光相机的内参和外参后，我们可以进一步对图像进行立体校正，将两者的图像校正到相近的水平高度。具体来说，红外和可见光图像的立体校正的计算根据文献[19]可以式(1)和式(2)表示：

$$I_{ir}' = \text{rectify}(I_{ir}, K_{ir}, R, T) \quad (1)$$

$$I_{vis}' = \text{rectify}(I_{vis}, K_{vis}, R, T) \quad (2)$$

式中： $I_{ir}$ 和 $I_{vis}$ 表示未进行立体校正的红外和可见光图像； $I_{ir}'$ 和 $I_{vis}'$ 分别表示立体校正后的红外和可见光图像； $K_{ir}$ 和 $K_{vis}$ 分别表示红外相机和可见光相机的内参； $[R, T]$ 表示双目相机外参； $R$ 表示两个相机的相对旋转参数； $T$ 表示相对平移参数。 $\text{rectify}$ 表示校正函数，具体细节可参考文献[18]中同名函数，立体校正前和校正后的图像如图3和图4所示。通过图中的水平线可以看出，立体校正后红外和可见光图像基本处于同一水平高度。



(a) 可见光图像 (b) 红外图像  
(a) Visible image (b) Infrared image

图3 立体校正前的可见光和红外图像

Fig.3 Visible and infrared images before stereo rectification



(a) 可见光图像 (b) 红外图像  
(a) Visible image (b) Infrared image

图4 立体校正后的可见光和红外图像

Fig.4 Visible and infrared images after stereo rectification

## 1.2 基于相位一致性的边缘特征图

红外图像和可见光图像相似性小，灰度差异大，直接在灰度图上进行特征点匹配工作难以寻找一致性的特征。考虑到边缘是红外图像和可见光图像都较为稳定的特征，不会因为图像之间的旋转、平移和缩放而引起形状的实质性变化，所以本文先提取红外和可见光图像的边缘图。

考虑到可见光图像对比度、光照会受到环境变化的影响而剧烈变化，而红外与可见光图像模态差异

大，要求边缘具备对图像的灰度、对比度、光照等的不变性。为此本文采用Kovesi<sup>[20]</sup>提出的基于相位一致性(Phase congruency, PC)边缘检测算法分别提取红外和可见光的边缘图。该方法选取 $o$ 个方向和 $s$ 尺度的Log-Gabor小波变换，计算每个方向上的相位一致性图即PC图。本文选取 $o=6$ ， $s=4$ ，计算方式简要表示如式(3)和式(4)，具体细节可参考文献[20]中的式(1)~(3)：

$$PC_{ir}(\theta) = \text{calPC}(I_{ir}, \theta, s_1, \dots, s_4) \quad (3)$$

$$PC_{vis}(\theta) = \text{calPC}(I_{vis}, \theta, s_1, \dots, s_4) \quad (4)$$

式中： $\theta$ 表示选取的Log-Gabor滤波器方向， $\theta_1 \sim \theta_6$ 分别为 $(0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ)$ ， $s_1 \sim s_4$ 表示4个频率尺度。

利用6个方向的相位一致性信息，通过矩分析方法得到PC图的最大矩 $M$ ，计算过程如式(5)所示：

$$M = \frac{1}{2} \left( c + a + \sqrt{b^2 + (a - c)^2} \right) \quad (5)$$

式中：

$$a = \sum_{i=1}^6 (PC(\theta_i) \cos(\theta_i))^2$$

$$b = 2 \sum_{i=1}^6 (PC(\theta_i) \cos(\theta_i))(PC(\theta_i) \sin(\theta_i))$$

$$c = \sum_{i=1}^6 (PC(\theta_i) \sin(\theta_i))^2$$

PC图的最大矩 $M$ 代表了图像的边缘特征，具有亮度、对比度不变性，若像素点 $(x, y)$ 处的最大矩 $M(x, y) > T$ ，则该位置为边缘点， $T$ 表示边缘点检测阈值。对于红外和可见光异源图像，可采用相同的阈值 $T$ 得到细节程度相当的红外边缘图像 $I_{iredge}$ 和可见光边缘图像 $I_{visedge}$ 。获取图像边缘图的过程如图5所示。

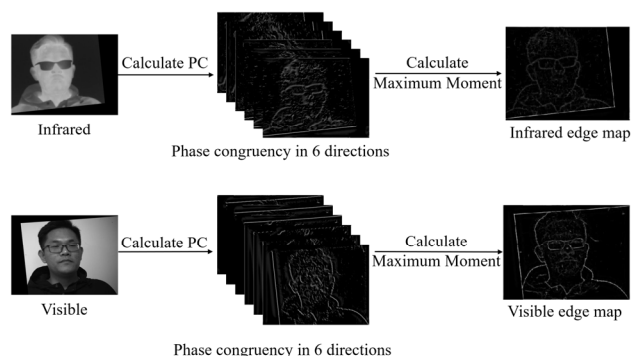


图5 红外与可见光图像边缘图计算过程

Fig.5 Infrared and visible edge maps calculation process

## 1.3 红外特征点提取

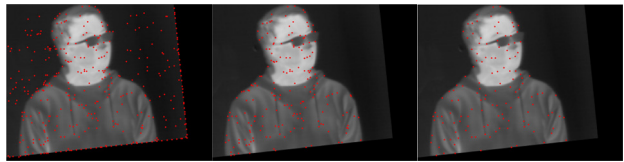
本文针对红外图像和可见光图像之间的配准，仅在红外边缘图上进行特征点检测。原因有两个方面，一方面本文采用局部区域搜索的方式寻找匹配特征



点, 所以仅需要在一幅图像上检测特征点, 在另一幅图像上的局部区域进行搜索即可。另一方面, 具有热源的目标是红外可见光配准任务中的主要目标, 所包含的信息最为丰富也最有价值, 这类目标在红外图像中更加突出。

由于在众多特征点检测算法中, FAST 特征检测算法<sup>[21]</sup>相对于 SIFT, SURF 等特征检测检测算法相比, 能够仅依靠周围像素快速地检测特征点, 且不进行特征点方向等冗余信息的计算。因此, 在红外边缘图中的特征点检测中, 本文在红外边缘图  $I_{\text{ir-edge}}$  的基础上, 进行基于 FAST 的特征点检测, 基于此得到的候选特征点集  $V_{\text{irCandidate}}$  结合了角点和边缘信息, 更加稳定。

通过 FAST 检测出特征点之后, 由于红外图像背景所含信息量极少, 处于背景中的特征点特征信息同样很少, 而且若将所有特征点用于配准会影响主体目标配准的准确性, 因此, 本文在处理红外特征点时, 主要保留目标主体上的特征点, 而舍弃背景中无用的特征点。红外图像中热源目标表现的很“亮”, 背景则很“暗”, 因此保留的候选特征点集  $V'_{\text{irCandidate}}$  满足其灰度大于  $I_{\text{ir}}$  的灰度均值。处理后的特征点显示到原红外图像中如图 6(b)所示。



(a) 原始特征点 (b) 剔除背景后 (c) 非极大值抑制后  
(a) Original feature points (b) After removing background (c) After NMS

图 6 特征点提取结果

Fig.6 Results of detecting feature points

针对候选特征点集  $V'_{\text{irCandidate}}$ , 为了防止特征点过于聚集, 导致配准结果偏向于局部区域, 本文以特征点  $V'_{\text{irCandidate}}$  为中心, 在  $5 \times 5$  的邻域内采用非极大值抑制 (Non-Maximum Suppression, NMS)<sup>[22]</sup>策略, 只保留邻域内 FAST 特征响应最大的特征点, 基于此得到分布较为均匀的特征点集  $V_{\text{ir}}$ 。由图 6(c)可以看出经过非极大值抑制, 减少了一些冗余的特征点, 特征点的分布更加均匀。

#### 1.4 两阶段特征点搜索匹配

上文利用相机标定信息进行立体校正, 可以将红外图像和可见光图像约束到同一高度之上。因此可在同一水平线上搜索同名特征点。考虑到在同一水平线上直接搜索同名特征点, 需要在整个图像宽度范围内进行搜索, 计算量偏大, 本文提出两阶段的同名特征

点搜索方式以红外特征点为基准在可见光边缘图中搜索同名特征点。

##### 1) 第一阶段

通过以 NCC 作为相似度的度量方式, 粗略估计红外图像边缘图  $I_{\text{ir-edge}}$  和可见光图像边缘图  $I_{\text{vis-edge}}$  的水平偏移量  $d_x$ , 使两幅图像在水平方向上大致对齐, 进而预测同名特征点的位置, 减小同名特征点的搜索范围, 提高配准效率,  $d_x$  的计算方式如下式:

$$d_x = \arg \max_d \text{NCC}(I_{\text{ir-edge}}, I_{\text{vis-edge}}^d) \quad (6)$$

式中:  $I_{\text{vis-edge}}^d$  表示水平偏移量为  $d$  时的可见光边缘图像。NCC 表示互相关一致性算子, 表示两幅图像的相似程。计算图像  $X$  和  $Y$  之间的 NCC 的计算方式如式 (7)所示:

$$\text{NCC}(X, Y) = \frac{\sum_{i=1}^m \sum_{j=1}^n [X(i, j) - \bar{X}][Y(i, j) - \bar{Y}]}{\sqrt{\sum_{i=1}^m \sum_{j=1}^n [X(i, j) - \bar{X}]^2} \sqrt{\sum_{i=1}^m \sum_{j=1}^n [Y(i, j) - \bar{Y}]^2}} \quad (7)$$

式中:  $m$  和  $n$  分别表示两幅图像的宽和高;  $\bar{X}$  和  $\bar{Y}$  分别表示图像  $X$  和  $Y$  的均值。

##### 2) 第二阶段

针对红外图像特征点集  $V_{\text{ir}}$  中的每个特征点  $p(x, y)$ , 其位于可见光边缘图中的初始同名点位置为  $p'(x + d_x, y)$ 。由于红外图像根据热辐射成像, 相对于可见光图像, 边缘不清晰是其固有缺陷, 在相机标定以及立体校正过程中, 无法严格将红外图像和可见光图像束到同一高度之上, 存在一定误差, 因此对应的同名点搜索区域为以  $p'$  为中心  $w \times h$  ( $w > h$ ) 大小的长方形区域  $S$ , 即图 7 右侧长方形区域, 而不是仅在水平线上进行搜索, 提高了准确性。 $S$  过大会影响效率, 过小会影响准确性, 经过实验本文搜索范围设为  $11 \times 7$ 。

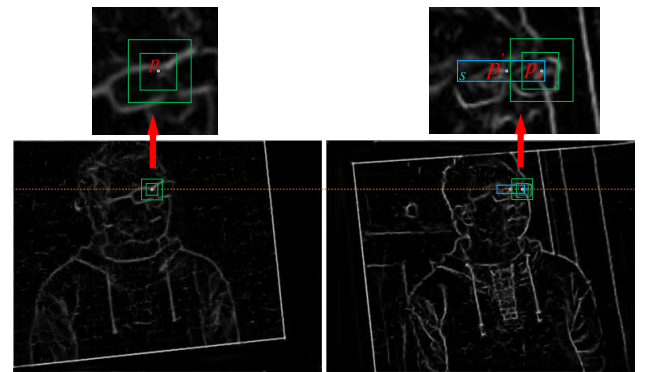


图 7 局部区域同名特征点搜索

Fig.7 Search for feature points with the same name in local area

NCC 作为衡量相似性的指标,一般来说模板越大准确性也就越高,同时时间复杂度也会上升,而边缘图像相对于灰度图像所含信息较少,使用较大尺寸模板对匹配精度来说更有益处,因此为了提高匹配准确性同时保证匹配效率,本文提出多尺度加权 NCC (wncc) 作为搜索同名特征点过程中的相似性衡量指标。即以两个尺寸的模板计算 NCC, 将其加权结合。具体来说首先构建大尺寸模板, 然后将其进行下采样, 这样保留了大模板内的主体结构信息, 同时也不至于增加计算复杂度。然后构建小模板作为信息补充, 弥补了大模板下采样时丢失的特征点周围的信息。最后将用两个模板计算得到的 NCC 进行加权结合, 大模板包含更多的结构信息, 用其计算的 NCC 衡量相似性更高, 故其权值也更大。将双尺寸模板加权结合, 提高了准确性, 更加适用于边缘图像。具体计算方式如式(8)所示:

$$\begin{aligned} \text{wncc}(p, p_s) = \\ \alpha \text{NCC}[D_2(N_p^{F_1}), D_2(N_{p_s}^{F_1})] + \beta \text{NCC}(N_p^{F_2}, N_{p_s}^{F_2}) \quad (8) \\ \text{s.t. } (\alpha + \beta) = 1 \end{aligned}$$

式中:  $p$  表示红外特征点;  $p_s$  表示可见光边缘图搜索区域内的点,  $N_p^{F_1}$  和  $N_{p_s}^{F_1}$  表示  $p$  周围  $F_1 \times F_1$  大小邻域和  $F_2 \times F_2$  大小邻域, 即图 7 左侧同心的两个正方形区域。本文选取  $F_1=24$ ,  $F_2=12$ 。  $N_p^{F_2}$  和  $N_{p_s}^{F_2}$  表示以同样方式构建的  $p_s$  的邻域, 如图 7 右侧同心的两个正方形区域,  $D_2$  表示 2 倍降采样算子。  $\alpha$  和  $\beta$  表示权重因子, 二者的取值范围为  $0 < \beta < 0.5 < \alpha$ , 本文取  $\alpha=0.6$ ,  $\beta=0.4$ 。

对于最终搜索得到的同名特征点  $p_{\text{correspond}}$  有:

$$p_{\text{correspond}} = \arg \max_{p_s} \text{wncc}(p, p_s) \quad (9)$$

式中:  $p_s \in S$ 。得到匹配的可见光图像特征点集  $V_{\text{vis}}$ 。

### 1.5 误匹配剔除

针对匹配后的特征点, 为了更进一步提高配准的精度, 需要将错误匹配的一些点进行剔除。本文采用迭代细化的方式剔除错误的匹配点。具体流程如下:

1) 根据已经匹配的点集  $V_{\text{ir}}$  和  $V_{\text{vis}}$ , 利用最小二乘法计算出初始仿射变换模型  $H$ 。

2) 根据变换模型  $H$  计算出红外图像中的特征点经过变换后在可见光图像中对应的点  $V_{\text{ir}}'$ :

$$V_{\text{ir}}' = V_{\text{ir}} \times H \quad (10)$$

3) 计算所有匹配点的误差和总的 RMSE, RMSE 代表了变换模型对整个点集的拟合程度, RMSE 越小说明模型拟合的越好, 具体计算方式如式(14)所示。

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum (V_{\text{ir}}' - V_{\text{vis}})^2} \quad (11)$$

式中:  $N$  表示匹配特征点对的数量。

4) 剔除点集中误差较大的点, 重复 1) ~ 3) 步骤, 直到 RMSE 小于阈值  $T_{\text{RMSE}}$ 。当  $T_{\text{RMSE}}$  小于 3 时可认为配准模型较好地拟合了特征点集之间的变换关系, 故本文选取  $T_{\text{RMSE}}=3$ 。

## 2 实验结果与分析

### 2.1 实验设置

本文实验采用自行组装的红外和可见光双目相机。红外相机采用高德 plug 417 测温型红外相机, 分辨率为  $384 \times 288$ 。可见光相机使用 FLIR 公司生产的型号为 BFS-U3-17S7M-C 的工业灰点相机。双目相机实物如图 8 所示, 左侧为可见光相机, 右侧为红外相机。实验环境为: Intel Core i5 10200H CPU, 16GB RAM; 软件平台为: Matlab R2016b。



图 8 红外可见光双目相机

Fig.8 Infrared and visible binocular cameras

为了评价本文方法配准的准确性, 采用主观评价和客观评价相结合的方式。主观评价通过观察配准后的红外可见光融合图像中目标的对齐程度直观地感受算法的配准效果。客观评价采用总匹配点对数 NUM、匹配准确率 (Correct Matching Rate, CMR) [23]。  $\text{CMR} = \text{MC}/N \times 100\%$ , 其中 MC 表示正确匹配点对数。当匹配点对数  $\text{NUM} < 4$  时视为配准失败。

### 2.2 实验分析

为了验证本文算法针对红外和可见光双目图像配准任务有效性, 选取了近年较为先进的 3 个异源图像配准算法 SURF-PIIFD-RPM[10]、CAO-C2F[13]、RIFT[14] 以及一个基于深度学习的配准算法 ReDFeat[17]与本文算法进行对比实验。

使用上述红外与可见光双目相机采集不同场景的 3 组图像共 30 对进行综合对比实验, 数据分为 3 组, 每组包含十对图像。第一组图像包含不同距离的以“人”为主体的图像。图像环境, 光线条件较好, 图像质量比较高, 纹理相对丰富。第二组图像主要包含不同距离的以物为主体的图像, 透明水壶装有热水, 红外与可见光图像差异性更大, 红外图像中水壶后面

的图像信息完全看不见，而可见光图像中则可以看到透明水壶后面的场景内容，且图像纹理较差。第三组图像包含在弱光环境下的图像。在弱光环境下，红外图像由于是捕获热辐射成像，成像质量基本不受影响，但依靠光反射成像的可见光图像的质量大大下降，主体变得不再清晰，噪点也相应增多，配准难度增大。每组数据各算法客观评价指标平均值如表 1 所示。

表 1 实验结果对比

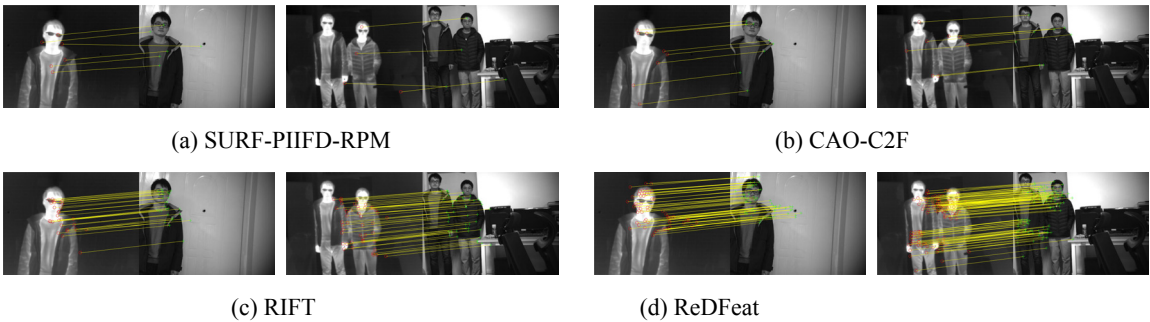
Table 1 Comparison of experimental results			
Image Group	Algorithm	NUM	CMR
First Group	SURF-PIIFD-RPM	5.7	66.0%
	CAO-C2F	6.3	76.8%
	RIFT	42.2	83.7%
	ReDFeat	107.3	80.0%
	OURS	67.6	91.1%
Second Group	SURF-PIIFD-RPM	2.8	59.2%
	CAO-C2F	7.3	69.5%
	RIFT	26.4	83.2%
	ReDFeat	127.4	76.8%
	OURS	34.0	91.6%
Third Group	SURF-PIIFD-RPM	1.8	28.3%
	CAO-C2F	4.6	27.5%
	RIFT	10.8	82.5%
	ReDFeat	22.3	62.2%
	OURS	37.3	88.1%

第一组图像的各个算法的特征点匹配结果和对应的配准融合结果如图 9 和图 10 所示，第二组图像的特征点匹配结果以及配准融合结果如图 11 和图 12 所示。由表 1 可以看出 SURF-PIIFD-RPM 和 CAO-C2F 匹配的特征点较少，准确率偏低，前者在第一组图像中平均特征点匹配数量仅 5.7 对，且准确率较低，这说明配准失败概率很大，如图 10(a)左图，配准已经失败。在第二组纹理较弱的图像特征点匹配数量仅 2.8 对，说明其在图像纹理较弱的情况下无法完成配准任务。CAO-C2F 算法在第一组图像中能匹配少量的特征点完成配准任务，但由于其在轮廓图中构建 SIFT 描述符，在第二组纹理较弱的图像中匹配准确率则明显下降，如图 12(b)的右图，配准结果都已经出现较大的偏移。

RIFT 算法由表 1 可以看出得到较为丰富的匹配特征点，且匹配准确率高，由图 10 和图 12 的(c)图中可以看出能较好的配准图像，配准后的红外与可见光图像都能大体对齐。这说明 RIFT 算法面对常规情况下红外图像和可见光图像能够比较好地提取其共同特征并成功匹配特征点。但 RIFT 存在的一个问题是特征点分布过于集中如图 9(c)的左图集中于人物右上侧区域，图 11(c)的右图集中于左侧水壶，这会导致配准有一定的偏移风险。如图 10(c)中的头部区域和图 12(c)两张图的右侧区域，都明显出现了偏移，这是因为 RIFT 通过独立提取红外和可见光的特征点，在点集中寻找匹配点，要求双方都提取了对应特征点才可能匹配成功，这会筛选掉大量特征点，造成特征点集中于纹理最丰富的区域。

ReDFeat 通过深度学习进行特征点的检测和描述，在特征点匹配数量上明显高于其他对比算法和本文算法，但特征点存在大量重复和冗余，并且比 RIFT 更加聚集于小部分区纹理丰富的区域，如图 11(d)左图，绝大多数特征点都聚集在了杯盖附近。同时由表 1 可以看出特征点匹配的准确率不高，第一组仅有 80%，第二组下降到 76.8%。这在第一组纹理丰富的图像中依靠特征点数量优势尚能得到较好的配准结果如图 10(d)。但在第二组纹理较弱的图像中的配准结果则出现了扭曲错位，如图 12(d)。

本文算法由于是在红外边缘图上进行特征检测，使得特征点围绕主体边缘分布，并且结合非极大值抑制策略，使得特征点分布更加均匀，如图 9 和图 11 的(e)图所示。在特征点匹配准确率上本文算法明显高于其他对比算法，这说明本文两阶段特征点搜索匹配方法的优越性。体现到配准结果中，由图 10 和图 12 的(e)图所示，在第一组和第二组图像中，都得到准确的配准结果。





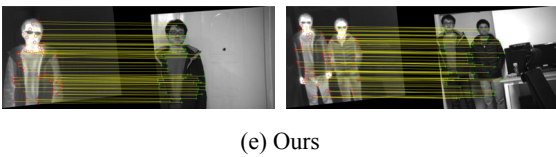


图 9 第一组图像各算法特征点匹配结果

Fig.9 The matching results of the feature points of each algorithm in the first group of images

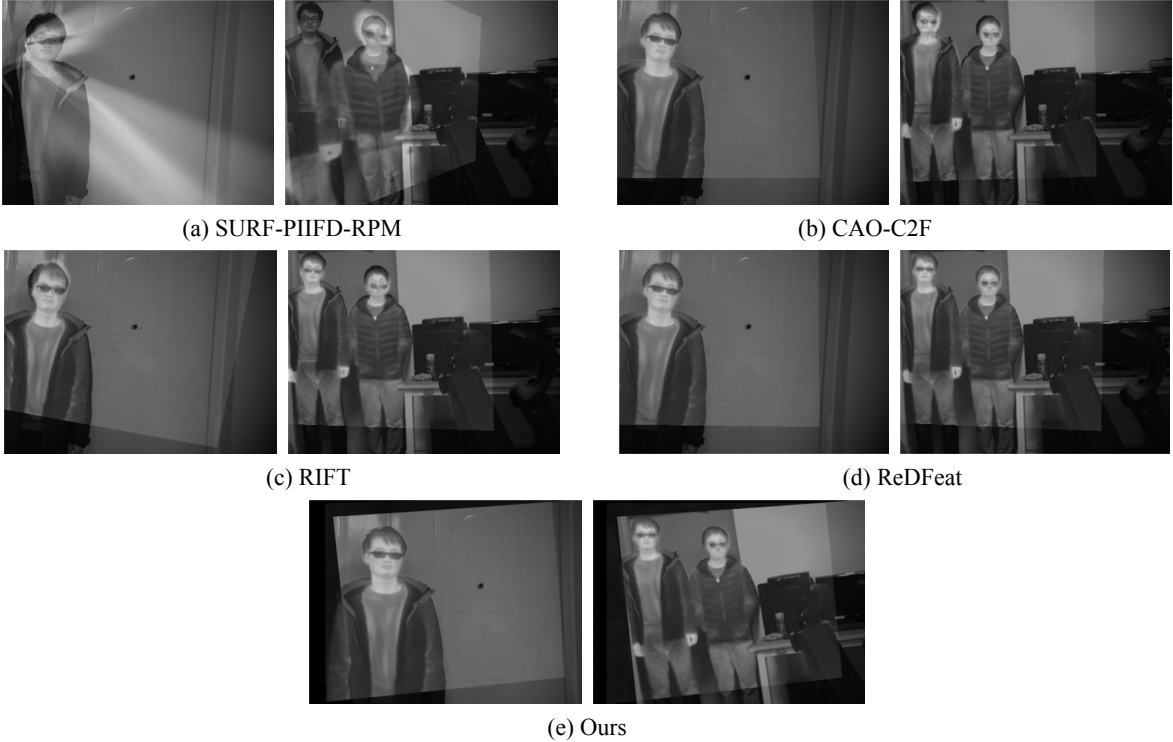


图 10 第一组图像配准结果

Fig.10 Registration results of the first group of images

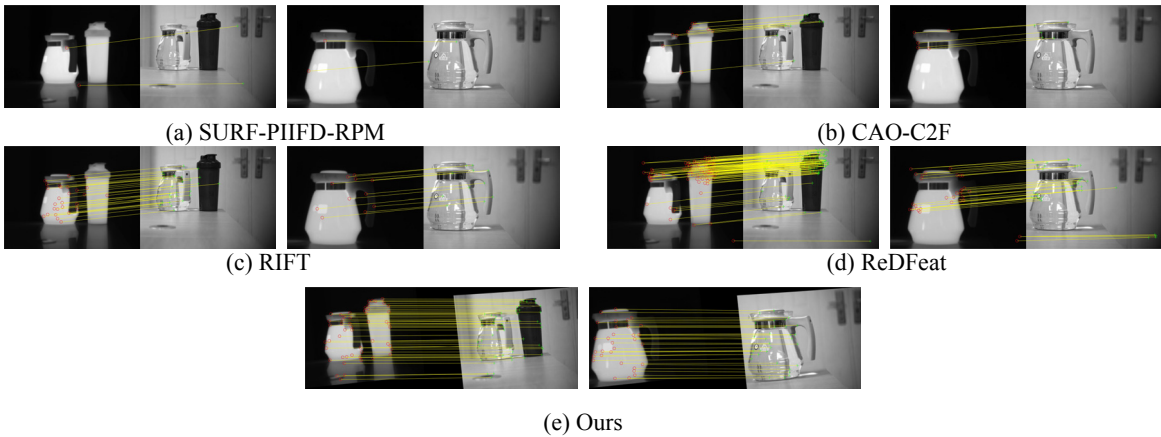


图 11 第二组图像各算法特征点匹配结果

Fig.11 The matching results of the feature points of each algorithm in the second group of images





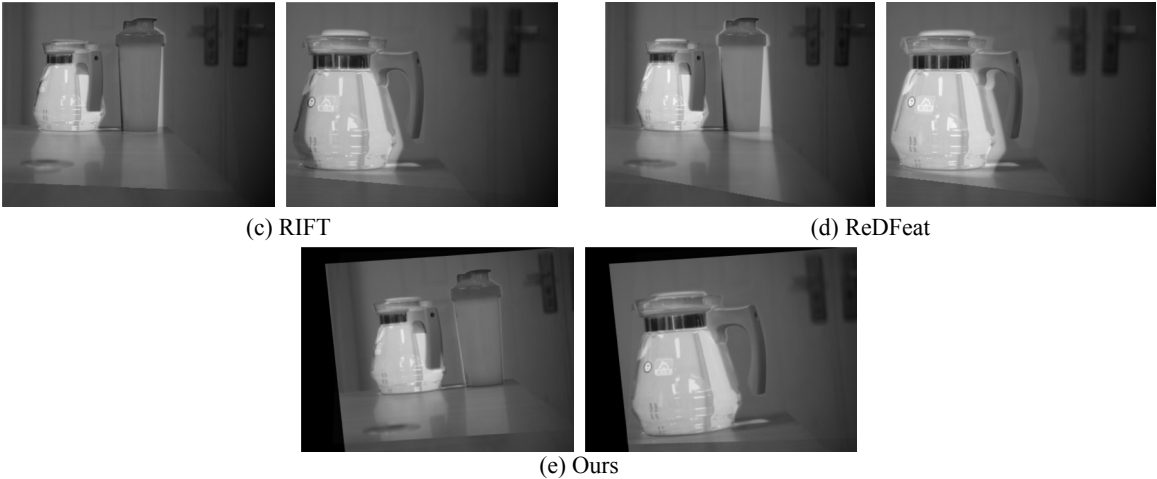


图 12 第二组图像配准结果

Fig.12 Registration results of the second group of images

第三组图像主要针对的弱光环境下的图像。从表 1 和图 13 以及图 14 可以看出本文算法特征点匹配数量和准确率上相较于对比算法表现出显著的优越性。从表 1 可以看出，SURF-PIIFD-RPM 算法特征点匹配数量仅有 1.8 对，无法完成配准任务。CAO-C2F 算法特征点匹配数量和准确率都处于较低水平，说明其无法适用于弱光环境，多数图像的匹配准结果都如图 14(b)右图一样，出现无法辨识的扭曲。RIFT 由于需要通过可见光图像构建对应的描述子，在可见光质量大大下降的情况下，从表 1 可以看出得到匹配特征点数量明显下降，配准效果也随之下降，如图 14(c)中的两张图像，配准结果都产生了明显错

位。而 ReDFeat 在弱光环境下，特征点数量多的优势也无法再继续保持，由表 1 可以看出在第三组图像中特征点匹配数量和准确率相较于第一组和第二组图像都出现了明显的下降。配准结果如图 14(d)所示，也出现了明显错位扭曲。本文算法则体现出了较大优势，在弱光环境可见光图像质量下降的情况下，本文算法通过两阶段特征点搜索匹配能够仅依靠局部有限的信息获取同名特征点，保证了匹配的准确性。同时由于在红外边缘图上进行特征点检测，由表 1 可以看到，依然能够得到丰富的匹配特征点。由图 14(e)的配准结果可以看出，在弱光环境下，本文算法依然能够准确得进行配准。

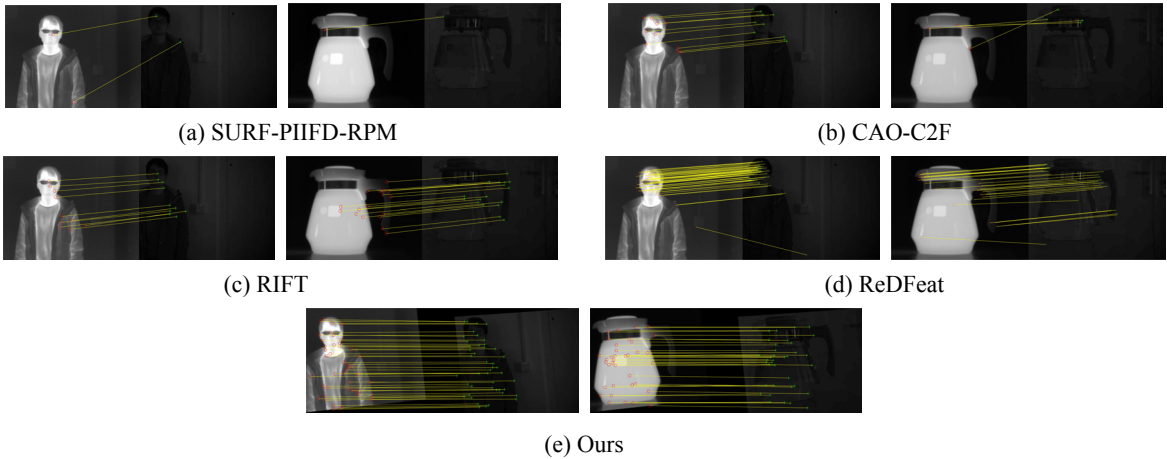


图 13 第三组图像各算法特征点匹配结果

Fig.13 The matching results of the feature points of each algorithm in the third group of images



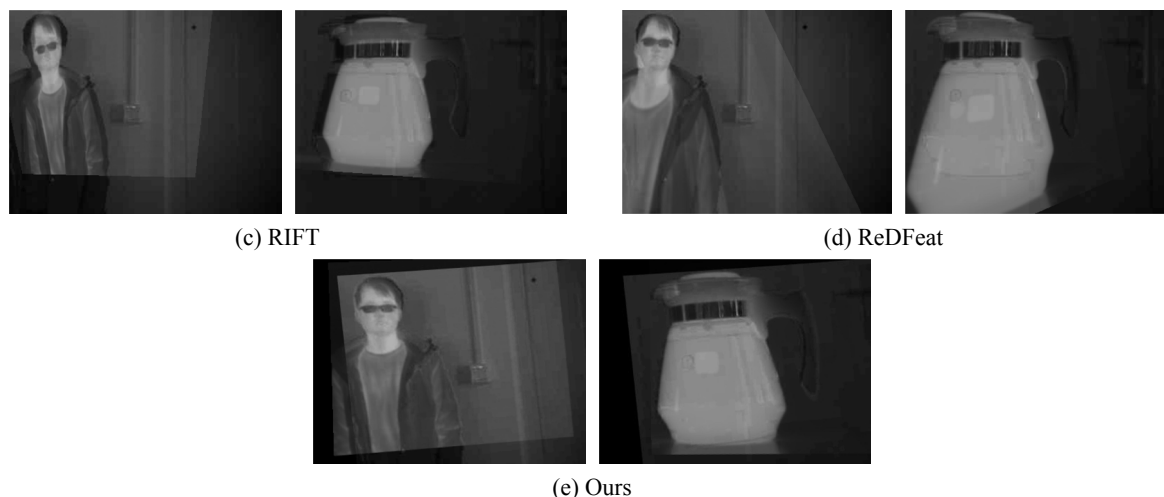


图14 第三组图像配准结果

Fig.14 Registration results of the third group of images

### 3 结语

本文针对相对位置固定的红外可见光双目成像应用场景,提出了一种几何约束下基于区域搜索图像配准方法。算法利用相机标定信息对红外和可见光图像进行立体校正,使二者处于同一高度之上,接着利用相位一致性计算鲁棒的边缘图,在边缘图的基础上,进行红外特征点检测,最后提出两阶段同名特征点搜索方法完成红外可见光图像配准。通过与其他算法在不同场景上的数据进行实验,结果表明相较于其他异源图像配准算法,本文算法能够适用于红外可见光双目成像的配准,在不同场景的图像都保证了特征点匹配数量和配准效果,并且能很好适应弱光环境下的图像配准工作,相对于其他算法取得了明显的优势。

### 参考文献:

- [1] LI Y, YU F Y, CAI Q, et al. Image fusion of fault detection in power system based on deep learning[J]. *Cluster Computing the Journal of Networks Software Tools and Applications*, 2019, **22**(4): 9435-9443.
- [2] 付添, 邓长征, 韩欣月, 等. 基于深度学习的电力设备红外与可见光图像配准[J]. *红外技术*, 2022, **44**(9): 936-943.  
FU T, DENG C, HAN X, et al. Infrared and visible image registration for power equipments based on deep learning[J]. *Infrared Technology*, 2022, **44**(9): 936-943.
- [3] 黄颖杰, 梅领亮, 王勇, 等. 基于红外与可见光图像融合的无人机探测研究[J]. *电脑知识与技术*, 2022, **18**(7): 1-8.  
HUANG Y, MEI L, WANG Y, et al. Research on unmanned aerial vehicle detection based on fusion of infrared and visible light images[J]. *Computer Knowledge and Technology*, 2022, **18**(7): 1-8.

- [4] 王戈, 钟如意, 黄浩, 等. 基于轻量级人脸识别的智慧地铁云支付系统搭建[J]. *湖北大学学报: 自然科学版*, 2021, **43**(4): 437-442.  
WANG G, ZHONG R, HUANG H, et al. Construction of intelligent metro cloud payment system based on lightweight face recognition[J]. *Journal of Hubei University: Natural Science*, 2021, **43**(4): 437-442.
- [5] Banuls A, Mandow A, Vazquez-Martin R, et al. Object detection from thermal infrared and visible light cameras in search and rescue scenes[C]// *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2020: 380-386.
- [6] JIANG X, MA J, XIAO G, et al. A review of multimodal image matching: Methods and applications[J]. *Information Fusion*, 2021, **73**: 22-71.
- [7] 李云红, 刘宇栋, 苏雪平, 等. 红外与可见光图像配准技术研究综述[J]. *红外技术*, 2022, **44**(7): 641-651.  
LIU Y, LIU Y, SU X, et al. Review of Infrared and visible image registration [J]. *Infrared Technology*, 2022, **44**(7): 641-651.
- [8] YU K, MA J, HU F, et al. A grayscale weight with window algorithm for infrared and visible image registration[J]. *Infrared Physics & Technology*, 2019, **99**: 178-186.
- [9] Yedukondala D C, Pejman G, Pfefer T, et al. Free-Form deformation approach for registration of visible and infrared facial images in fever screening[J]. *Sensors*, 2018, **18**(2): 125.
- [10] WANG G, WANG Z, CHEN Y, et al. Robust point matching method for multimodal retinal image registration[J]. *Biomedical Signal Processing and Control*, 2015, **19**: 68-76.
- [11] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF)[J]. *Computer Vision and Image Understanding*, 2008, **110**(3): 346-359.
- [12] CHEN J, TIAN J, Lee N, et al. A partial intensity invariant feature descriptor for multimodal retinal image registration[J]. *IEEE Transactions on Biomedical Engineering*, 2010, **57**(7): 1707-1718.

- [13] JIANG Q, LIU Y, YAN Y, et al. A contour angle orientation for power equipment infrared and visible image registration[J]. *IEEE Transactions on Power Delivery*, 2020, **36**(4): 2559-2569.
- [14] LI J, HU Q, AI M. RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform[J]. *IEEE Transactions on Image Processing*, 2019, **29**: 3296-3310.
- [15] WANG L, GAO C, ZHAO Y, et al. Infrared and visible image registration using transformer adversarial network[C]// 25th *IEEE International Conference on Image Processing (ICIP)*. IEEE, 2018: 1248-1252.
- [16] Arar M, Ginger Y, Danon D, et al. Unsupervised multi-modal image registration via geometry preserving image-to-image translation[C]// *Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition*, 2020: 13410-13419.
- [17] DENG Y, MA J. ReDFeat: Recoupling detection and description for multimodal feature learning[J/OL]. *IEEE Transactions on Image Processing*, 2022: 591-602. <https://arxiv.org/abs/2205.07439>.
- [18] ZHANG Z. A flexible new technique for camera calibration[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2000, **22**(11): 1330-1334.
- [19] Fusiello A, Trucco E, Verri A. A compact algorithm for rectification of stereo pairs[J]. *Machine Vision & Applications*, 2000, **12**(1):16-22.
- [20] Peter Kovesi. Phase congruency detects corners and edges[C]//*Digital Image Computing: Techniques and Applications*, 2003: 309-318.
- [21] Edward R, Tom D. Machine learning for high-speed corner detection[C]//*Computer vision - ECCV*, 2006: 430-443.
- [22] Neubeck A, Van Gool L. Efficient non-maximum suppression[C]// *Pattern Recognition, ICPR*, 2006: 850-855.
- [23] YE Y, SHAN J, Bruzzone L, et al. Robust registration of multimodal remote sensing images based on structural similarity[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, **55**(6): 2941-2958.