光电技术及应用

DOI：10.16818/j.issn1001－5868.2023051505

基于多模态的车辆行人检测技术研究

裘一鸣[1]，李范鸣[2*]

（1. 上海科技大学 信息科学与技术学院，上海 201800；2. 中国科学院上海技术物理研究所，上海 200083）

摘　要： 针对单一传感器在复杂路况以及恶劣天气情况下车辆行人检测效果不佳，搭建了一套可见光、可见光偏振、短波红外和长波红外多模态数据采集系统，构建了一个多模态数据集，并提出了一种多模态车辆行人检测算法。首先，提出了一种基于改进型SIFT特征点的多尺度部分强度不变特征的异源图像配准算法；然后，提出基于YOLOv5多模态数据目标检测网络。最终实现了平均精度在日间数据集1.0%的提升，日间夜间混合数据集10.9%的提升。

关键词： 多模态；红外与偏振；图像配准；多模态目标检测

中图分类号： V323.19；TN911.7　**文章编号：** 1001－5868(2023)05－0775－07

Research on Vehicle and Pedestrian Detection Technology Based on Multimodal Image

QIU Yiming[1]，LI Fanming[2]

（1. School of Information Science and Technology，ShanghaiTech University，Shanghai 201800，CHN；
2. Shanghai Institute of Technical Physics，Chinese Academy of Sciences，Shanghai 200083，CHN）

Abstract： Aiming at the problem of poor detection of vehicles and pedestrians by a single sensor in complex road conditions and bad weather conditions，a set of visible light，visible polarization，short-wave infrared and long-wave infrared multimodal data acquisition system was built to construct a multimodal dataset，and a multimodal pedestrian detection algorithm for vehicle pedestrians was proposesd. Firstly，a heterologous image registration algorithm based on the multi-scale partial intensity invariant features of improved SIFT feature points was proposed. Then，for target detection，a multimodal data object detection network based on YOLOv5 was proposed. Finally，the average accuracy is improved by 1.0% in the daytime dataset and 10.9% in the daytime and nighttime mixed dataset.

Key words： multimodal；infrared and polarization；image registration；multimodal target detection

0　引言

视觉感知作为自动驾驶车辆安全行驶环境感知模块之一，在车辆行驶过程中检测周围车辆和行人有着十分重要的作用。早期的目标检测方法通常利用低级视觉特征，检测性能受制于手工特征与真实目标的相似性。目标检测方法主要分为单帧型算法和多帧型算法，其中单帧型算法又可以分为局部信息类单帧型算法[1-5]和非局部信息类单帧型算法[6-10]，多帧型算法可以分为关联交验类多帧型算法[11]和直接求取类多帧型算法[12-16]等。此外，早期的检测算法大多基于单一模态数据，面对复杂情况检测能力有限。比如，基于可见光的检测方法，在日间有着不错的表现，但是在夜间对车辆与行人识别效果不佳。为了解决夜间检测性能不佳的问题，近年来，国内外学者应用红外传感器对车辆、行人进行

收稿日期：2023－05－15.
基金项目：国家十四五预研基金项目(514010405).
* **通信作者：**李范鸣　**E-mail：**lfmjws@163.com

检测，在夜间场景的目标识别率上有了很大的提升。大部分研究成果建立在特定场景下，缺乏通用场景下对车辆、行人检测算法的研究。在日间场景，目标在可见光下的特征明显，红外特征被制约，在夜间或其他低照度场景下，可见光特征被制约。针对单一模态图像进行训练得到的目标检测网络无法在通用场景下识别目标。

近年来，国内外学者在通用场景下，对基于多模态数据基础的目标检测开展了研究，并取得初步成果。Hwang 等[17]建立了可见光红外融合的 KAIST 行人检测数据集，但未对多模图像检测进行研究。Xu 等[18]采用深层卷积神经网络对可见光和红外图像关系进行建模，将学习到的特征迁移到新的神经网络中，能在低能见度下对可见光图像进行检测，但该方法仍强烈依赖于可见光图像。Wagner 等[19]首先应用 CNN 对 KAIST 行人数据集进行检测，并提出前融合和后融合两种模型，但不能区分出不同方法在模型中的作用。Liu 等[20]设计了四种不同的卷积神经网络融合策略，在不同阶段将两个分支的卷积神经网络融合，分别对应前中后以及置信阶段，但其主要针对低照度场景下的行人目标进行研究，局限性大。Konig 等[21]基于 Faster R-CNN 模型进一步提高检测精度，但仍为针对低照度行人的研究。Gupta 等[22]基于可见光图像和深度图检测图像中物体的轮廓，将深度图信息编码进 RGB 图像中，并生成包含像素视差、高度、倾斜角的 2.5D 图像。陈晓智等[23]提出了一种新的检测方法，对于每一个 3D 包围框用 x，y，z，θ，c，t 来表示中心，θ 表示方位角，c 表示类别，t 表示模板集合。采用深度图像的方法在实现时需要将二维图像三维化，其计算量过大，且由于三维图像数据的稀疏性，大量运算为无效运算，因此使用深度图像的方法在当前阶段不具有实用性。

本文通过采集可见光、可见光偏振、短波红外、长波红外的多模态（四种）数据集。多模态相机的成像原理和成像方式各不相同，因此不同模态图像中的特征点存在梯度方向相反的现象，导致现有配准算法，如 SIFT，SURF，ORB，FAST 等方法在对不同图像中的特征点进行特征描述时得到不一致的结果，进而影响特征点匹配的准确性。

部分强度不变特征描述符[24]（Partial Intensity Invariant Feature Descriptor，PIIFD）通过合并对称方向梯度来减小特征点在不同模态图像中特征向量的差异，更适合应用在多模态图像的特征点匹配中。但 PIIFD 特征向量受图像尺度的影响，无法配准不同尺度下的图像。

为了解决 PIIFD 算法不具备尺度不变性的问题，本文建立了多尺度局部强度不变特征描述符（Multi-Scale Partial Intensity Invariant Feature Descriptor，MS-PIIFD），在多尺度空间上对特征点进行特征描述，使用基于 SIFT 特征点的 MS-PIIFD 配准算法对异源图像进行配准。基于 YOLOv5 算法设计了多模态数据目标识别网络，使用 FLIR ADAS 和自建多模态数据集进行验证，实现了对多模态数据的车辆行人检测，提升了目标检测算法在通用环境下的鲁棒性。

1 基于 SIFT 特征点的 MS-PIIFD 配准算法

1.1 算法总体流程

本章提出的基于 SIFT 特征点的 MS-PIIFD[24]配准算法流程图如图 1 所示。以 SIFT 控制点作为特征点。SIFT 特征查找到的关键点是一些不会因为光照、仿射变换、噪声等因素变化的特殊点，如角点、边缘点、暗区亮点和亮区暗点等。通过使用 MS-

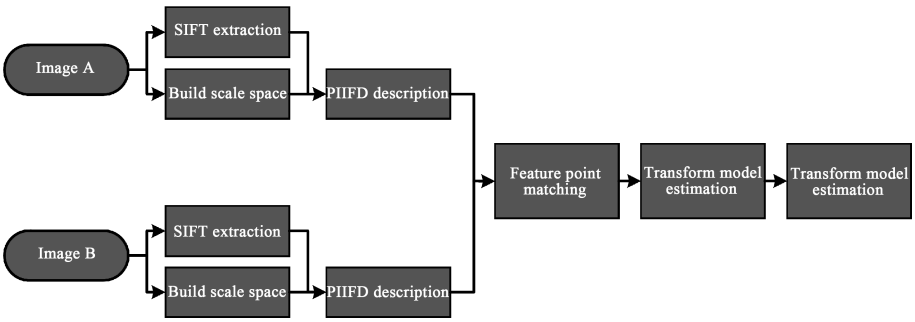


图 1 配准算法流程图

PIIFD 特征描述符代替 SIFT 特征描述符，通过在不同模糊程度的不同尺度空间上建立 MS-PIIFD 特征描述，接着根据特征点的 MS-PIIFD 特征描述对特征点进行匹配。最后根据匹配的特征点，使用仿射变换模型进行估计，计算得到仿射变换矩阵后，对图像进行配准。

1.2　SIFT 特征点

常见的特征点检测算法有 Harris 算法[25]、SIFT 算法[26]、ORB 算法[27]和 FAST 算法[28]等。SIFT 算法具有平移、旋转、缩放不变性，而 Harris 算法、ORB 算法、FAST 算法并不具有尺度不变性。在多模态数据配准场景下，尺度存在明显差异，造成无法配准的问题。为了解决这个问题，本文使用 SIFT 特征点代替 Harris 特征点作为构建 MS-PIIFD 特征向量所用的特征点。

对于待配准图像进行下采样，生成每一组图像的第一层，并通过将图像与一个高斯核进行卷积，改变图像的尺度大小，构建原始图像的多尺度空间。对于下采样后的图像，使用相同方法检测特征点时，感受野更大，更加关注大尺度上的特征，忽略细节上的特征。对于高斯平滑后的图像，使用相同方法检测特征点，更加关注图像的低频特征，而不注重高频特征。通过将每个像素点与其在高斯金字塔空间上相邻空间内的 26 个点比较，初步检测极值点。因为边缘点会导致错误的特征点，所以，本文通过设定阈值来剔除边缘点。

1.3　MS-PIIFD 特征描述

通过在尺度空间上对不同平滑程度的图像使用 PIIFD，使得特征描述符具有多尺度性。在多模态数据配准中，多尺度性对于配准多模态有着极为重要的意义。首先，将输入的单模态图像转化为灰度图像以统一不同模态图像的维度，通过创建高斯金字塔作为多尺度空间，组合不同尺度、不同模糊程度的图像。比如，取一个 SIFT 特征点 p，然后，对特征点 p 大小为 $S\times S$ 的邻域内的像素，计算特征主方向，并将邻域内像素的梯度方向和梯度幅值做归一化操作，最后，统计生成 16 个方向的梯度方向直方图。根据特征点 p 的特征主方向，将邻域分为 4×4 共 16 个面元。考虑到长波红外、短波红外、可见光偏振、可见光图像像素的梯度方向可能相反，因此，将对称方向的梯度相加，使得不同图像中方向相反的梯度能落在直方图的同一方向上。此时，16 个梯度方向合并成了 8 个。梯度方向直方图矩阵可以表示为：

$$\boldsymbol{H}=\begin{bmatrix} O_{11} & O_{12} & O_{13} & O_{14} \\ O_{21} & O_{22} & O_{23} & O_{24} \\ O_{31} & O_{32} & O_{33} & O_{34} \\ O_{41} & O_{42} & O_{43} & O_{44} \end{bmatrix} \tag{1}$$

其中，O_{ij} 为第 ij 面元内的梯度方向直方图。由于 16 个方向通过对称相加的方法消除了跨模态图像梯度相反的问题，O_{ij} 实际为 8 个合并方向的梯度直方图。$\boldsymbol{H}'$为旋转 180°后的梯度方向直方图：

$$\boldsymbol{H}'=\begin{bmatrix} O_{44} & O_{43} & O_{42} & O_{41} \\ O_{34} & O_{33} & O_{32} & O_{31} \\ O_{24} & O_{23} & O_{22} & O_{21} \\ O_{14} & O_{13} & O_{12} & O_{11} \end{bmatrix} \tag{2}$$

为了使特征向量具有旋转不变性，通过对特征点 p 邻域内的梯度方向直方图矩阵求共模和差模，差模增益计算方法如式(3)所示：

$$\boldsymbol{D}=\begin{bmatrix} O_{44}+O_{11} & O_{43}+O_{12} & O_{42}+O_{13} & O_{41}+O_{14} \\ O_{34}+O_{12} & O_{33}+O_{22} & O_{32}+O_{23} & O_{31}+O_{24} \\ s\mid O_{24}-O_{13}\mid & s\mid O_{23}-O_{23}\mid & s\mid O_{22}-O_{33}\mid & s\mid O_{21}-O_{34}\mid \\ s\mid O_{14}-O_{14}\mid & s\mid O_{13}-O_{24}\mid & s\mid O_{12}-O_{43}\mid & s\mid O_{11}-O_{44}\mid \end{bmatrix} \tag{3}$$

其中，$s=\dfrac{\max(O_i+O'_i)}{\max(O_i+O'_i)}$，$i=3,4$。

将特征点的特征描述符变换成一维行向量，得到特征点 p 长度为 128 的特征向量，在高斯金字塔的多尺度空间下，特征向量可以消除不同模态图像之间尺度不同带来的影响，进而匹配不同模态图像的特征点。

1.4　特征点匹配

在构建完各特征点的 MS-PIIFD 特征向量后，本文使用最佳箱优先(Best Box First，BBF)算法来实现不同图像的特征点集的匹配。首先计算特征向量间的余弦相似度，根据余弦相似度进行匹配。

假设待匹配的点分别为 P_x，P_y，对应的特征向量为 $\boldsymbol{D}_x$，$\boldsymbol{D}_y$，$\boldsymbol{D}_x=[x_1,x_2,\cdots,x_{128}]$，$\boldsymbol{D}_y=[y_1,y_2,\cdots,y_{128}]$，两特征向量的距离可以用余弦相似度进行计算，如式(4)所示：

$$\mathrm{dist}(\boldsymbol{D}_x,\boldsymbol{D}_y)=\cos^{-1}(x_1y_1+x_2y_2+\cdots+x_{128}y_{128}) \tag{4}$$

余弦相似度越小表示两个特征向量距离越短。当 P_x 与 P_y 的特征向量距离为 P_x 与所有待匹配特征点的最小距离，且最小距离和次小距离的比值大于阈值时，则将 P_y 其视为 P_x 的单边匹配点。

同理，当 P_x 也是 P_y 的单边匹配点时，则这组特征点间为双边匹配。

只有满足双边匹配的点为正确的匹配关系，其余都为错误匹配。

1.5 配准结果及分析

本文采用了可见光、可见光偏振、短波红外、长波红外四通道相机进行图像采集。可见光相机使用手机镜头进行拍摄，分辨率为 1 920×1 080，帧率为 30 f/s。可见光偏振相机采用 Triton 500 万像素偏振千兆以太网工业相机，分辨率为 2 448×2 048，最大帧速率为 22 f/s。短波红外相机采用 LD-SW6401715-C-G 短波机芯，分辨率为 640×512，帧率为 50 f/s。长波红外相机采用，分辨率为 640×512，帧率为 50 f/s。多通道图像采集平台实物如图2所示。采集得到的图像经辐射校正、单点校正后，图像样本如图 3 所示。其中，图 3(a)为可见光偏振图像，(b)为长波红外图像，(c)为短波红外图像，(d)为可见光图像。

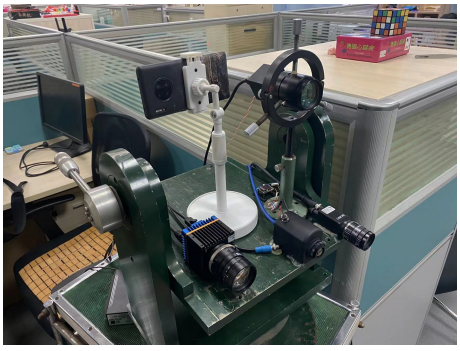
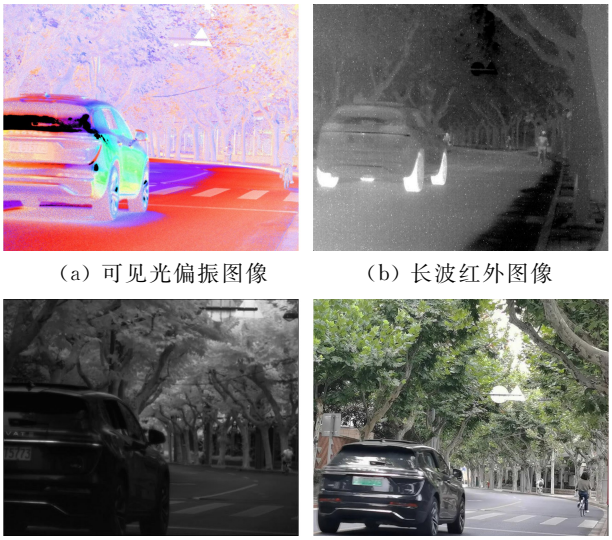


图 2 四模态图像采集系统



(a) 可见光偏振图像 (b) 长波红外图像

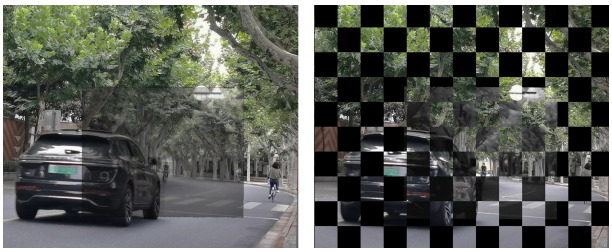
(c) 短波红外图像 (d) 可见光图像

图 3 四模态图像样本

如图 3 所示，目标在不同模态内的位置不同。实验时由于短波红外的视野最小，故将可见光图像、长波红外图像和可见光偏振图像配准至短波红外图像。图 4 为可见光图像与短波红外图像中的特征点匹配关系图。图 5(a)为可见光图像与短波红外图像融合图像，图 5(b)为可见光图像与短波红外图像的棋盘格拼接图像。通过观察可以看到，经配准后，车辆目标基本处于同一位置，远处非机动车在棋盘格图像中拼接正常。



图 4 可见光-短波红外图像特征点匹配结果



(a) 融合图像 (b) 棋盘格图像

图 5 可见光-短波红外图像融合结果

2 多模态图像目标检测算法

YOLOv5 算法在目标检测算法中模型相对较小，训练速度快，识别效果较好。然而，YOLOv5 算法仅针对传统目标检测算法进行设计，并且只支持单一模态图像作为数据输入，未考虑到多模态图像的场景。因此本文基于 YOLOv5 算法构建了多模态图像目标检测算法，对 YOLOv5 网络模块进行改进，以此来适应多模态图像目标检测场景。

2.1 基于 YOLOv5 的多模态目标检测网络

网络的整体结构如图 6 所示。本文重新设计了数据输入部分，以兼容多模态输入。利用多模态数据融合来丰富目标信息，增强数据，使得检测网络性能提升，并且该检测算法具有良好的场景通用性。

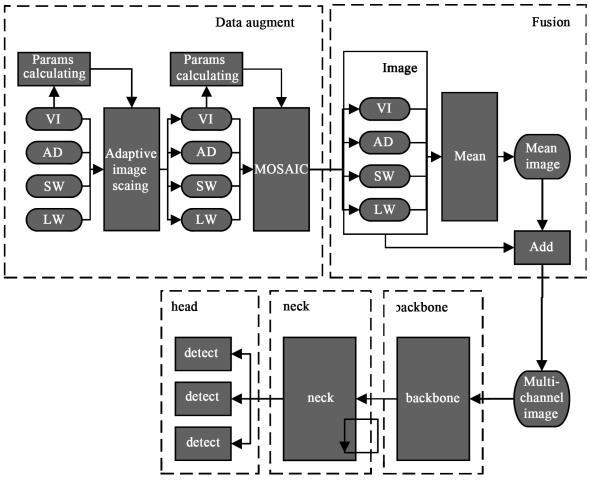


图 6　网络整体结构

2.2　融合模块

本文算法采用前期融合的方法融合多模态图像。融合模块主要分为均值与加权融合两部分，均值模块求得各模态图像强度图的均值，随后将均值图像与各模态图像加权相加后叠加成多模态图像。当输入图像为单模态图像时，本文算法退化为标准 YOLOv5 网络。

2.3　损失函数

损失函数由三部分加权得到，分别为矩形框损失（box loss）、置信度损失（confidence loss）和分类损失（classification loss）。矩形框损失使用 CIOU 作为损失函数，置信度损失和分类损失使用 BCEloss 作为损失函数。得到最终的损失函数为：

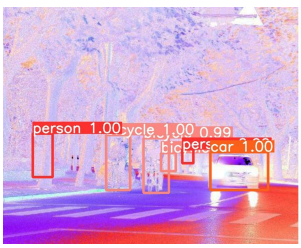
$$L_{\text{total}}=\lambda_{\text{bbox}}L_{\text{bbox}}+\lambda_{\text{cls}}L_{\text{cls}}+\lambda_{\text{obj}}L_{\text{obj}} \tag{5}$$

2.4　评价指标

本文使用 $\text{mAP}_{0.5}$，$\text{mAP}_{0.5:0.95}$ 作为评价指标，mAP 是准确率与召回率的综合衡量，是目标检测算法的基本评价指标。

2.5　实验结果

图 7 为验证集中图像在本文算法中的验证结果，图 7(a)～(d)分别为可见光偏振、可见光、短波红外、长波红外图像数据。图 8 为验证集中图像在 YOLOv5s 算法中的验证结果。



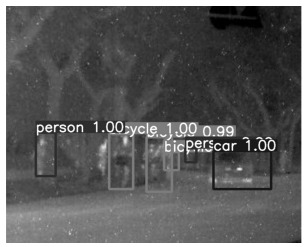
(a) 可见光偏振



(b) 可见光



(c) 短波红外



(d) 长波红外

图 7　在本文算法下的验证结果



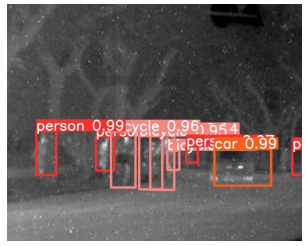
(a) 可见光偏振



(b) 可见光



(c) 短波红外



(d) 长波红外

图 8　在 YOLOv5s 中的验证结果

表 1 和表 2 分别为单模态图像及多模态图像的训练结果。对比单模态图像作为数据输入的 YOLOv5 算法和使用多模图像作为输入的本文算法，取不同深度和宽度下的同种模型的最优指标作为当前模态训练结果，本文算法在 $\text{mAP}_{0.5}$ 和 $\text{mAP}_{0.5:0.95}$ 上分别比单模态最优训练结果（本训练结果中为可见光图像训练结果）有 0.4%和 1.0%的

提升。其中,单模态图像为使用 YOLOv5 网络进行训练得到的结果,多模态图像为使用相应深度宽度的 YOLOv5 网络作为基础网络进行训练得到的结果,max mAP 为不同深度和宽度下的 YOLOv5 模型及其作为基础网络的多模态模型训练结果的最优值。

表 1 不同输入图像在不同宽度深度下 $mAP_{0.5}$ 对比

mAP0.5	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv5x	Max mAP0.5
visible	0.952	0.887	0.81	0.826	0.952
polarized	0.89	0.778	0.813	0.808	0.89
SWIR	0.893	0.669	0.688	0.671	0.893
LWIR	0.799	0.598	0.671	0.67	0.799
Multi-model	0.925	0.956	0.954	0.956	0.956

表 2 不同输入图像 $mAP_{0.5:0.95}$ 对比

mAP	YOLOv5s	YOLOv5m	YOLOv5l	YOLOv5x	Max mAP
visible	0.733	0.613	0.529	0.552	0.733
polarized	0.663	0.494	0.529	0.539	0.663
SWIR	0.637	0.409	0.425	0.405	0.637
LWIR	0.509	0.34	0.399	0.399	0.509
Multi-model	0.657	0.731	0.732	0.743	0.743

为进一步验证本文网络的有效性,本文使用公开数据集 FLIR 日间夜间混合数据集作为数据输入进行性能对比,结果如表 3 所示。由表 3 可知,本文算法训练得到的网络,其准确率、召回率、$mAP_{0.5}$、$mAP_{0.5:0.95}$ 均明显高于由可见光或红外图像训练得到的网络。

表 3 FLIR 数据集的结果对比

模型	数据	准确率	召回率	mAP0.5	mAP0.5:0.95
YOLOv5	visible	0.794	0.609	0.673	0.319
YOLOv5	IR	0.809	0.725	0.786	0.404
本文	visible+IR	0.903	0.832	0.883	0.513

3 结论

针对单一模态图像目标检测难以适应通用场景的不足,提出了一种多模态图像目标检测算法。该算法首先使用基于 SIFT 特征点的 MS-PIIFD 算法配准多模态图像,使其在空间上统一;然后引入了基于多模态数据的 YOLOv5 目标检测算法。该算法通过早期融合对多模态图像进行数据增强。实验结果表明,与现有的结果对比,本文算法在检测精度、鲁棒性上均有提升,但本文算法仍然有一些可以改进的地方,如配准算法耗时过长,不具有实时性等。因此,未来的研究工作可以在提高算法效率等方面进行进一步的研究。

参考文献:

[1] Yang W P, Shen Z K. Small target detection and preprocessing technology in infrared image sequences[J]. Infrared and Laser Engineering, 1998, 27(1): 23-28.

[2] Deshpande S D, Meng H E, Venkateswarlu R, et al. Max-mean and max-median filters for detection of small targets[J]. Proc. of SPIE, 1999, 3809: 74-83.

[3] Cao Y, Liu R M, Yang J. Small target detection using two-dimensional least mean square (TDLMS) filter based on neighborhood analysis[J]. Inter. J. of Infrared and Millimeter Waves, 2008, 29(2): 188-200.

[4] Lu R T, Yang X G, Li W P, et al. Robust infrared small target detection via multidirectional derivative-based weighted contrast measure[J]. IEEE Geoscience and Remote Sensing Lett., 2020, 19: 1-5.

[5] Bi Y G, Chen J Z, Sun H, et al. Fast detection of distant, infrared targets in a single image using multiorder directional derivatives[J]. IEEE Trans. on Aerospace and Electronic Systems, 2020, 56(3): 2422-2436.

[6] Gregoris D J, Yu S K, Tritchew S. Wavelet transform-based filtering for the enhancement of dim targets in FLIR images [J]. Proc. of SPIE, 1994, 2242: 573-583.

[7] Kong X Y, Liu L, Qian Y S, et al. Automatic detection of sea-sky horizon line and small targets in maritime infrared imagery[J]. Infrared Phys. & Technol., 2016, 76: 185-199.

[8] 郑 胜, 柳 健, 田金文. 基于支持向量机的红外小目标分割和聚类方法研究[J]. 信号处理, 2005, 21(5): 515-519.
Zheng Sheng, Liu Jian, Tian Jinwen. Research of SVM-based infrared small object segmentation and clustering method[J]. Signal Processing, 2005, 21(5): 515-519.

[9] Wu Y Q, Zhou Y, Long Y L. Small target detection in hyperspectral remote sensing image based on adaptive parameter SVM[J]. Acta Optica Sinica, 2015, 35(9): 0928001.

[10] Hou Q Y, Wang Z P, Tan F J, et al. RISTDnet: Robust infrared small target detection network[J]. IEEE Geoscience and Remote Sensing Lett., 2022, 19: 7000805.

[11] Wang B, Xu W H, Zhao M, et al. Antivibration pipeline-filtering algorithm for maritime small target detection[J]. Optical Engineering, 2014, 53(11): 113109.

[12] Reed I S, Gagliardi R M, Shao H M. Application of three dimensional filtering to moving target detection[J]. IEEE Trans. on Aerospace and Electronic Systems, 1983, 19(6): 898-905.

[13] Choi J H, Rajala S A. Three-dimensional location estimation of trajectories of point targets using a projection-based transformation method[J]. Optical Engineering, 1995, 34 (3): 933-939.

[14] Wu B, Ji H B. Improved power-law-detector-based moving small dim target detection in infrared images[J]. Optical Engineering, 2008, 47(1): 010503.

[15] Bae T W. Small target detection using bilateral filter and temporal cross product in infrared images[J]. Infrared Phys.

& Technol.，2011，54(5)：403-411.

[16] Liu J，Ji H B. An improved robust estimation algorithm for small IR target detection [C]// IEEE Symposium on Industrial Electronics & Applications，2009：394-398.

[17] Hwang S，Park J，Kim N，et al. Multispectral pedestrian detection：Benchmark dataset and baseline[C]// 2015 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)，2015：1037-1045.

[18] Xu D，Ouyang W L，Ricci E，et al. Learning cross-modal deep representations for robust pedestrian detection[C]// 2017 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)，2017：4236-4244.

[19] Wagner J，Fischer V，Behnke S，et al. Multispectral pedestrian detection using deep fusion convolutional neural networks[C]// 24th European Symposium on Artificial Neural Networks，Computational Intelligence and Machine Learning (ESANN)，2016：1-6.

[20] Liu J J，Zhang S T，Wang S，et al. Multispectral deep neural networks for pedestrian detection [EB/OL]. 2016-11-8. https://arxiv.org/pdf/1611.02644.pdf.

[21] König D，Adam M，Jarvers C，et al. Fully convolutional region proposal networks for multispectral person detection [C]// 2017 IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW)，2017：243-250.

[22] Gupta S，Girshick R，Arbeláez P，et al. Learning rich features from RGB-D images for object detection and segmentation[C]// Computer Vision-ECCV 2014，2014：345-360.

[23] Chen X Z，Kundu K，Zhu Y K，et al. 3D object proposals for accurate object class detection [C]// NIPS'15：Proceedings of the 28th International Conference on Neural Information Processing Systems，2015：424-432.

[24] Gao C Z，Li W. Multi-scale PIIFD for registration of multi-source remote sensing images[J]. J. of Beijing Institute of Technol.，2021，30(2)：113-124.

[25] Harris C G，Stephens M J. A combined corner and edge detector[C]// Alvey Vision Conference，1988：147-151.

[26] Lowe D G. Object recognition from local scale-invariant features[J]. Proc. of the Seventh IEEE Inter. Conf. on Computer Vision，1999，2：1150-1157.

[27] Rublee E，Rabaud V，Konolige K，et al. ORB：An efficient alternative to SIFT or SURF[C]// 2011 Inter. Conf. on Computer Vision，2011：2564-2571.

[28] Rosten E，Drummond T. Machine learning for high-speed corner detection[C]// Computer Vision-ECCV 2006，2006：430-443.

作者简介：

裘一鸣(1995—)，男，上海市人，硕士，主要从事多模态目标识别方面的研究；

李范鸣(1969—)，男，研究员，博士生导师，主要从事红外系统工程、红外成像技术、多维信息获取及微弱目标探测技术方面的研究。

(上接第 746 页)

[9] Chen P，Li X N，Zhang G. Rapid detection to long ship wake in synthetic aperture radar satellite imagery [J]. J. of Oceanology and Limnology，2019，37(5)：1523-1532.

[10] 李　海，贾宏光，陈在斌. 共轴双旋翼系统的气动特性分析及试验[J]. 光学精密工程，2021，29(9)：2140-2148.
Li Hai，Jia Hongguang，Chen Zaibin. Analysis and experiment on aerodynamic characteristics of coaxial rotor system [J]. Optics and Precision Engineering，2021，29(9)：2140-2148.

[11] 张　群，王英民. 尾流中多气泡模型及有限元分析[J]. 鱼雷技术，2014，22(4)：316-320.
Zhang Qun，Wang Yingmin. Multi bubble model and finite element analysis in wake[J]. Torpedo Technology，2014，22(4)：316-320.

[12] 张银波，李思宁，姜　鹏，等. PCA 特征提取和弹性 BP 神经网络的水下气泡识别[J]. 红外与激光工程，2021，50(6)：209-215.
Zhang Yinbo，Li Sining，Jiang Peng，et al. PCA feature extraction and elastic BP neural network for underwater bubble recognition[J]. Infrared and Laser Engineering，2021，50(6)：209-215.

[13] 高　江，张静远，杨　力. 舰船气泡尾流特性研究现状[J]. 舰船科学技术，2008(4)：27-32.
Gao Jiang，Zhang Jingyuan，Yang Li. Research status of ship bubble wake characteristics[J]. Ship Science and Technology，2008(4)：27-32.

[14] 田恒斗，金良安，王　涌，等. 考虑单气泡运动特性的舰船尾流气泡分布研究[J]. 兵工学报，2011，32(9)：1126-1130.
Tian Hengdou，Jin Liangan，Wang Yong，et al. Study on bubble distribution in ship wake considering single bubble motion characteristics[J]. Acta Armamentarii，2011，32(9)：1126-1130.

[15] Zhang X X，Wu Z S，Su X. Influence of breaking waves and wake bubbles on surface-ship wake scattering at low grazing angles[J]. Chin. Phys. Lett.，2018，35(7)：32-35.

[16] Guo C Y，Wu T C，Zhang Q，et al. Numerical simulation and experimental research on wake field of ships under off-design conditions[J]. China Ocean Engineering，2016，30(5)：821-834.

[17] Wang B L，Guo X Y，He C. Numerical simulations of wake signatures around high-speed ships [J]. J. of Hydrodynamics，2014，26(6)：986-989.

作者简介：

宗思光(1979—)，男，博士，副教授，主要从事水下激光探测、激光通信、激光致声等方面的研究。