

## 多视角红外图像目标识别方法

赵璐<sup>1</sup>, 熊森<sup>2</sup>

(1. 西安思源学院 工学院, 陕西 西安 710038;  
2. 西安斯坦仪器股份有限公司, 陕西 西安 710065)

**摘要:** 随着红外传感器的性能提升和应用普及, 获取同一场景下同一目标的多视角图像成为可能。为此, 提出联合多视角红外图像的目标识别方法。首先对多视角红外图像进行聚类分析, 获取多个视角子集。在每个视角子集中, 红外图像具有较强的相关性。对于不同的视角子集, 它们相对独立。为充分利用这种相关性和独立性, 采用联合稀疏表示 (JSR) 对单个视角子集进行决策。特别地, 对于只包含一个视角的子集, 则直接采用经典的稀疏表示分类 (SRC) 进行处理。对于不同视角子集获取的决策结果, 基于线性加权的思想进行融合处理, 并根据融合后的决策变量判决多视角红外图像所属的目标类别。因此, 所提方法在分析多视角红外图像内在关联性的基础上, 分别对局部相关性和整体的独立性进行考察, 并通过决策层的融合将两者融为一体, 提高了最终决策的可靠性。实验中, 在采集的多类交通车辆红外图像上进行识别, 分别在原始图像、加噪声图像以及部分遮挡图像上对方法进行测试和验证, 经过对比分析验证了提出方法的有效性。

**关键词:** 红外图像; 目标识别; 多视角聚类; 稀疏表示

**中图分类号:** TP391      **文献标志码:** A      **DOI:** 10.3788/IRLA20210206

## Target recognition based on multi-view infrared images

Zhao Lu<sup>1</sup>, Xiong Sen<sup>2</sup>

(1. Department of Engineering, Xi'an Siyuan University, Xi'an 710038, China;  
2. Xi'an Sitan Instrument Co., Ltd., Xi'an 710065, China)

**Abstract:** With the improvement of the performance of infrared sensors and the popularization of applications, it becomes possible to obtain multi-view images of the same target in the same scene. Therefore, a target recognition method combining multi-view infrared images was proposed. First, the clustering analysis on multi-view infrared images was performed to obtain multiple view-view subsets. In each view subset, the infrared images shared high correlations. For different view subsets, they were relatively independent. In order to make full use of the correlation and independence, the joint sparse representation (JSR) was used to make decisions on single view subsets. In particular, for the subset with only one view, the classical sparse representation-based classification (SRC) was directly used for decision. For the decision results obtained by different view subsets, the fusion processing was carried out based on the idea of linear weighting. And the target category was determined according to the fused results. Therefore, on the basis of analyzing the inner correlation of the multi-view infrared images, the proposed method separately examined the local correlations and overall independence, and integrated them through the fusion on the decision-making layer, which improved the reliability of the final decision. Experiments were performed on the collected infrared images of multiple types of traffic vehicles. The proposed

收稿日期: 2021-03-30; 修订日期: 2021-04-20

基金项目: 陕西省教育厅 2019 年度专项科学研究计划 (19JK0691)

作者简介: 赵璐, 女, 讲师, 博士, 主要研究方向为图像处理、语音处理。

method was tested and verified on the original, noisy, and occluded samples. The effectiveness of the proposed method is verified by comparison with other methods.

**Key words:** infrared images; target recognition; multi-view clustering; sparse representation

## 0 引言

相比可见光观测, 红外成像可在夜间等复杂场景下工作, 为全天时监测提供有力工具, 因此在军事和民用领域有着十分广泛的运用。在军事领域, 利用红外成像及处理方法, 可辅助开展夜间战场环境监测, 实现目标识别和精确打击<sup>[1-4]</sup>。在民用领域, 红外成像可用于夜间交通管控, 根据不同类别车辆的热效应不同进行准确分析和识别, 为驾驶员夜间行车提供辅助决策。因此, 典型车辆的分类识别问题在军事、民用领域均有着重要的研究价值。当前, 关于红外图像车辆识别的研究主要基于经典的模式识别问题思路, 一般区分两个阶段: 特征提取和分类器学习<sup>[5]</sup>。在特征提取阶段, 研究人员提出了各种方法, 例如尺度不变特征变换 (SIFT)<sup>[6]</sup>、定向梯度直方图 (HOG)<sup>[7]</sup>、加速鲁棒特征 (SURF)<sup>[8]</sup> 和其他特征提取算法<sup>[9-10]</sup> 等。通常, 所有这些方法都涉及手工提取的特征, 需要对图像特征进行更深入的了解, 以获取合适的特征进行可靠的识别。此类特征一般需要专业的邻域知识, 具有一定的针对性。但其设计过程具有一定的经验性, 因此鉴别力往往有限。在分类器学习方面, 红外图像识别邻域与其他模式识别问题一样主要借鉴经典、稳健的分类器, 如支持向量机 (SVM)<sup>[11]</sup>、神经网络、稀疏表示分类 (SRC)<sup>[12]</sup>、联合稀疏表示 (JSR) 等。随着深度学习理论的发展, 基于网络学习的深度特征学习和分类器设计也在红外图像目标识别领域得到应用, 并验证了有效性<sup>[13-15]</sup>。

红外传感技术的发展和成熟为目标观测提供了丰富的样本。针对识别问题, 获取同一目标不同视角的红外图像成为可能。为此, 通过联合多个不同视角的红外图像进行综合分析提高识别精度和稳健性成为一种有效的技术途径。文中基于多视角红外图像开展目标识别方法研究。首先, 对于同一目标的多视角红外图像进行聚类分析。以图像相关性为基本相似度度量准则, 将包含多个视角的红外图像集合区分为多个子集。在每个子集中, 各个图像具有较高的相关性。不同子集之间则保持一定的独立性。为此, 针

对不同子集, 分别采用 SRC(当前子集仅包含一幅图像) 或 JSR<sup>[16-19]</sup>(当前子集仅包含多幅图像) 进行分类。针对不同子集输出的结果 (即重构误差矢量), 采用线性加权融合的思路进行处理, 最终根据融合后的重构误差判定多视角红外图像所述的目标类别。所提方法充分考虑了多视角红外图像的内在分布特性, 通过聚类分析、联合稀疏表示和决策融合予以体现, 因此可以充分利用多视角的互补性提高识别性能。实验中, 在几类交通车辆目标的红外图像集上对方法进行测试和对比分析, 结果反映了方法的有效性。

## 1 基于相关性准则的多视角聚类

进行多视角红外图像的聚类主要为了正确分析它们之间的关联性, 从而有针对性地制定分类准则。文中以图像相关性为聚类算法的基本准则, 定义如下:

$$c = \max \left( \frac{\sum_k \sum_l [I_1(k, l) - m_1][I_2(k - \Delta k, l - \Delta l) - m_2]}{\left[ \sum_k \sum_l [I_1(k, l) - m_1]^2 [I_2(k - \Delta k, l - \Delta l) - m_2]^2 \right]^{1/2}} \right) \quad (1)$$

式中:  $I_1$  和  $I_2$  为多视角红外图像中的任意两幅;  $m_1$  和  $m_2$  代表它们的均值;  $\Delta k$  和  $\Delta l$  为沿图像两个方向的平移量;  $c$  为两者的相关系数。

假设有  $N$  个视角的红外图像, 记为  $V = \{I_1, I_2, \dots, I_N\}$ , 首先按照公式 (1) 中的度量准则计算它们两两之间的相似度, 获取如表 1 所示的矩阵, 其中  $c_{ij}$  表示第  $i$  和  $j$  个视角之间的图像相关。在此基础上, 多视角聚类算法的实施过程描述如表 1 所示。

表 1 多视角红外图像相似性矩阵

Tab.1 Similarity matrix of multi-view infrared images

	$I_1$	$I_2$	...	$I_N$
$I_1$	$c_{11}$	$c_{12}$	...	$c_{1N}$
$I_2$	$c_{21}$	$c_{22}$	...	$c_{2N}$
$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$
$I_N$	$c_{N1}$	$c_{N2}$	...	$c_{NN}$

Step 1: 初始化聚类中心  $I_C = I_1$ , 迭代次数  $t = 1$ , 当前视角子集  $S_t = \{1\}$ ;

Step 2: 执行以下操作

for  $j = \{1, 2, \dots, N\} \setminus S_t$

if  $c_{1S_t} \geq T_c$

$S_t = S_t \cup j$

end

end

直至获取第一个视角子集;

Step 3: 更新待聚类的多视角图像为  $V = V \setminus S_t$ , 迭代次数  $t = t + 1$ , 重复上述步骤, 直到所有视角均被归于某一视角子集。

上述过程中, “ $\setminus$ ”代表余集操作。对于同一视角子集中若干幅红外图像, 它们两两之间的相似度均高于预设门限, 从而体现了内在关联。对于不同视角集之间, 它们的关联性相对较多, 更多体现为独立性。因此, 通过在适当门限下进行多视角红外图像的聚类分析, 可进一步考察它们之间的内在特性, 从而为后续分类策略的制定提供更为有力的支撑。

## 2 基于稀疏表示的目标识别

### 2.1 稀疏表示原理

稀疏表示以线性表示理论为基础, 通过引入稀疏约束提高表示精度。具体在目标识别领域, 稀疏表示分类以训练样本构建全局字典  $A = [A_1, A_2, \dots, A_C] \in R^{d \times N}$ , 其中  $A_i$  表示第  $i$  类中  $N_i$  个训练样本对应的原子, 以测试样本  $y$  为被表示对象, 获取如下所示的基本表达形式:

$$\begin{aligned} \hat{x} &= \arg \min_x \|x\|_0 \\ \text{s.t. } &\|y - Ax\|_2^2 \leq \varepsilon \end{aligned} \quad (2)$$

式中:  $x$  为待求解的稀疏系数矢量;  $\varepsilon$  为设定的误差的门限。

对公式 (2) 进行求解, 获取稀疏表示系数矢量。在此基础上, 分别针对各个类别进行重构误差计算, 如下式所示:

$$r(i) = \|y - A_i x_i\|_2^2 \quad (i = 1, 2, \dots, C) \quad (3)$$

式中:  $x_i$  为对应第  $i$  类上的系数矢量。在此基础上, 通过比较不同类别的误差大小即可进行目标类别的决策。

### 2.2 联合稀疏表示

针对多个存在关联的稀疏表示问题, 若按照传统稀疏表示进行独立求解, 则会损失这种相关性信息, 导致重构结果不是最优的。为此, 研究人员提出联合

稀疏表示模型, 在统一框架下同时求解多个稀疏表示问题<sup>[16-18]</sup>。以三个存在关联的输入为例, 记为  $y^{(1)}, y^{(2)}, y^{(3)}$ , 对它们进行统一稀疏表示的问题初步表示如下:

$$\min_{\beta} \left\{ g(\beta) = \sum_{k=1}^3 \|y^{(k)} - A^{(k)} \alpha^{(k)}\| \right\} \quad (4)$$

式中:  $A^{(k)}$  为对应第  $k(k=1,2,3)$  个输入的全局字典;  $\alpha^{(k)}$  为相应的系数矢量;  $\beta = [\alpha^{(1)} \alpha^{(2)} \alpha^{(3)}]$  为系数矩阵。

从公式 (4) 的形式可以看出, 尽管统一考察了不同问题的表示过程, 但其结果与独立求解并无差别, 缺少它们之间相关性的考虑。为此, 联合稀疏表示模型对稀疏矩阵的结构和分布进行约束, 更新目标函数如下式所示:

$$\min_{\beta} g(\beta) + \lambda \|\beta\|_{2,1} \quad (5)$$

式中:  $\lambda$  为大于零的参数; 通过  $\ell_1/\ell_2$  范数对  $\beta$  的约束, 可反映不同输入之间的关联性。

采用正交匹配追踪或多任务压缩感知等算法可对公式 (5) 中的优化问题进行求解。在此基础上, 分别计算各类别对三个输入的重构误差, 并根据误差最小的原则进行决策:

$$\text{identity}(y) = \min_i \sum_{k=1}^3 \|y^{(k)} - A_i^{(k)} \alpha_i^{(k)}\| \quad (6)$$

式中:  $A_i^{(k)}$  为第  $i$  类中对应第  $k$  个输入的字典;  $\alpha_i^{(k)}$  为相应的系数矢量。

### 2.3 决策融合

文中以稀疏表示为基本手段, 结合多视角红外图像进行目标识别。对于只包含一幅图像的视角子集, 直接采用传统稀疏表示进行处理, 获取对应的重构误差矢量; 对于包含多幅图像的视角子集, 基于联合稀疏表示进行处理, 获取对应的重构误差矢量。假设经聚类处理, 获取  $P$  个视角集, 各视角集对应的重构误差为  $f(i) (i = 1, 2, \dots, P; i = 1, 2, \dots, C)$ , 按照线性加权的思想对它们进行融合处理, 如下式所示:

$$e(i) = \omega_1 f_1(i) + \omega_2 f_2(i) + \dots + \omega_P f_P(i) \quad (7)$$

式中:  $\omega_j (j = 1, 2, \dots, P)$  为不同视角子集的权值;  $e(i)$  为融合后的误差矢量。根据不同视角子集包含的图像数量, 文中设定  $\omega_j = p_j/P$ , 其中  $p_j$  为第  $j$  个视角子集中的图像数量。根据加权融合后各类别的重构误差大小, 按照最小原则判定当前多视角红外图像所属的目标类别。

## 2.4 识别流程

根据提出方法实施的具体步骤,设计如图 1 所示的识别流程。首先,对于同一目标的多视角红外图像进行聚类分析,按照相关性强弱划分为若干个视角子

集。然后,对于每一个视角子集,基于稀疏表示或联合稀疏表示进行处理,获得对应的重构误差矢量。最终,基于线性加权融合的思想,对不同视角子集的结果进行融合处理,根据融合后的结果判定目标所属类别。

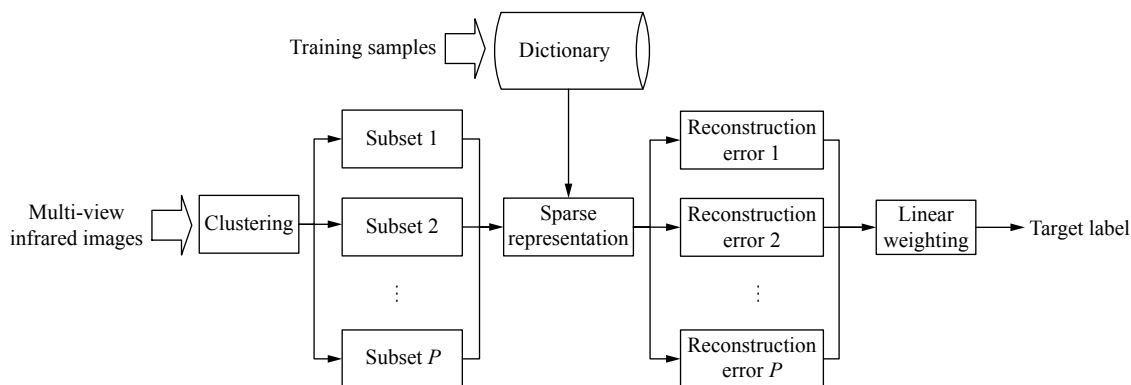


图 1 多视角红外图像目标识别流程图

Fig.1 Flowchart of infrared image target recognition based on multi-view

实际应用过程中,在进行单个视角集决策时均可以获得对应的识别结果,若所有视角子集的决策类别是一致的,则无需进行下一阶段的融合处理,便可获得目标的类别。

## 3 实验及分析

### 3.1 实验准备

基于交通车辆红外数据集对提出方法进行试验。经预处理,获取公交车图像 3200 幅,轿车图像 5400 幅,卡车图像 1200 幅,小货车图像 1800 幅,部分示意图如图 2 所示。四类目标的图像均采集自现实生活环境,来自多个角度的不同传感器。具体实验中,随机选取各类目标的一半样本用于训练,剩余一半为测试样本。

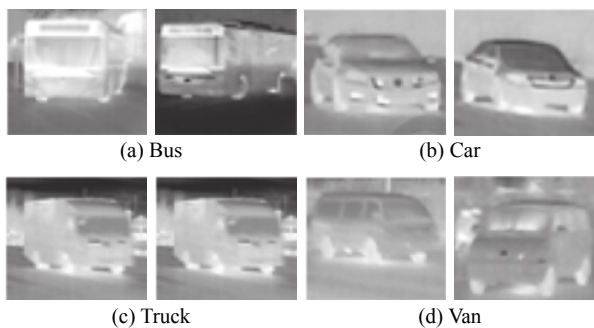


图 2 四类交通车辆红外图像示意图

Fig.2 Illustration of infrared images of the four vehicles

实验过程中,对于提出方法,选用六个视角为典型多视角条件对方法性能进行测试。根据多次重复实验的结果,设定多视角聚类算法中的门限值  $T_c=0.4$ 。同时,从现有文献中选取四类对比方法,分别为基于 SRC 的方法、基于 SVM 的方法,基于 HOG 特征的方法以及基于深度特征 (Deep features) 的方法。采用平均识别率作为识别精度的度量准则,定义为正确识别的样本数占全部测试样本的比例。

### 3.2 结果与分析

#### 3.2.1 原始样本

首先对四类目标的原始图像进行识别。如图 2 所示,现实采集的目标红外图像经预处理后具有良好的可视性,各类目标之间区分性较强,因此识别难度相对较低。表 2 给出了提出方法对四类目标的具体识别结果,公交车、轿车、卡车以及货车的识别率分别为 98.14%、98.14%、98.14% 和 98.14%。经计算,获得提出方法在当前条件下的平均识别率为 98.69%。表 3 列出了各类方法的平均识别率。对比可知,所提方法具有最高的平均识别率。一方面,文中利用了多个不同视角的互补信息,相比传统单一视角具有更充分的鉴别力。另一方面,文中还对多视角图像的内在关联做了更进一步的挖掘,因此可进一步提高识别性能。在四类对比方法中,联合深度特征的方法具有一定优势,表明了深度网络以及深度特征对

表 2 所提方法对四类目标原始样本的识别结果

Tab.2 Recognition results of the proposed method on the original samples of four targets

Class	Bus	Car	Truck	Van	Recognition rate
Bus	1 576	5	6	13	98.50%
Car	9	2 674	7	10	99.04%
Truck	4	1	592	3	98.67%
Van	2	3	6	889	98.78%
Average recognition rate					98.81%

表 3 不同方法在原始样本上的性能对比

Tab.3 Performance comparison of different methods on original samples

Methods	Average recognition rate
Proposed	98.81%
SRC	97.68%
SVM	97.04%
HOG	97.23%
Deep features	98.02%

红外目标识别的有效性。

3.2.2 噪声干扰

与其他各类图像一样, 红外图像获取过程中同样易受到噪声影响, 导致整体信噪比 (SNR) 降低, 给正确识别带来障碍。实验中, 笔者通过噪声添加的方式获取不同 SNR 条件下的测试集。具体地, 计算待处理图像的整体能量, 按照预设的信噪比生成高斯白噪声矩阵, 然后将其与原始图像叠加, 获取相应的噪声样本。对于不同噪声水平的测试集, 提出方法和四类对比方法分别进行识别实验, 统计结果如图 3 所示。

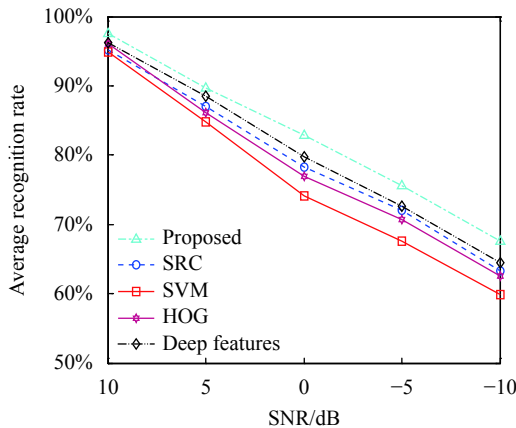


图 3 不同方法在噪声样本上的性能对比

Fig.3 Performance comparison of different methods on noisy samples

从图可见, 噪声干扰对各类方法的性能均有着明显影响。对比来看, 提出方法在各个噪声条件下均保持了最高的平均识别率, 显示其稳健性。从四类对比方法的结果来看, 采用稀疏表示分类机制的性能更优, 体现其噪声稳健性。文中方法正是结合了多视角的互补性以及稀疏表示的噪声稳健性, 进一步提升了整体识别性能。

3.2.3 目标遮挡

除了噪声影响外, 现实中由于遮挡影响, 目标并不一定能够完整体现在获取的红外图像中。为此, 识别方法在遮挡条件下的有效性十分重要。文中在模拟过程中以测试集中完整目标为参考, 分别对其部分区域进行遮挡处理, 按照不同比例区分遮挡程度, 具体构造 10%、30%、50% 和 70% 四类情形。在此基础上, 统计获得不同方法的平均识别率如图 4 所示。与噪声干扰的结果类似, 提出方法在当前测试条件的性能最佳。从对比方法的结果也可以看出稀疏表示对于部分遮挡的情形更具稳健性。文中多个视角的互

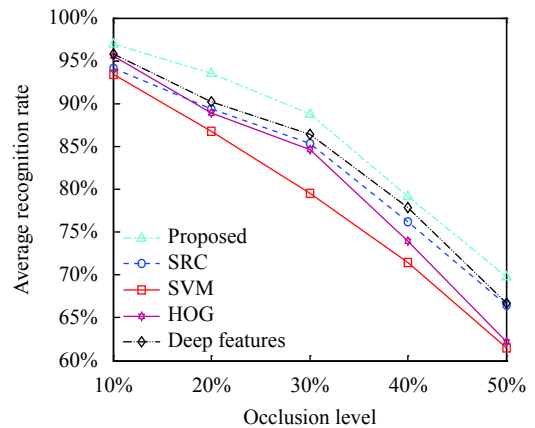


图 4 不同方法在遮挡样本上的性能对比

Fig.4 Performance comparison of different methods on occluded samples

补分析,能够有效提升对于遮挡条件的稳健性。进一步,稀疏表示增强了提出方法的整体稳健性。

## 4 结 论

文中提出一种多视角红外图像目标识别方法。同一目标的不同视角红外图像不同可以反映目标不同侧面的特性,因而具有良好的互补性。提出方法首先以图像相关为准则对多个视角的红外图像进行聚类分析,获取多个视角子集。采用联合稀疏表示对不同视角子集中的图像进行分析,通过考察内在相关性提高整体表示精度。最终,基于线性加权融合对不同视角子集的结果进行联合分析,获取可靠的识别结果。实验以四类交通车辆的多视角红外图像为训练和测试集开展,分别对原始样本和经处理获取的噪声、遮挡样本进行识别。根据实验结果,提出方法相比现有几类方法具有更强的有效性。

### 参考文献:

- [1] Zhou J M, Chen C, Tu W B, et al. Composite layer defect detection method based on infrared heat wave technology, finite element and SVM [J]. *Chinese Journal of Scientific Instrument*, 2020, 41(3): 29-38. (in Chinese)
- [2] Wang H B, Wan L J. Small infrared target detection based on Contourlet transformation and target characteristic analysis [J]. *Journal of Electronic Measurement and Instrumentation*, 2019, 33(3): 120-125. (in Chinese)
- [3] Jiang W, Wang X C, Yang Y. Infrared image enhancement algorithm based on structural feature priori and multi-scale non-sharpening masking mechanism [J]. *Journal of Electronic Measurement and Instrumentation*, 2019, 33(4): 1-9. (in Chinese)
- [4] Wang Y Y, Zhao Y H, Luo H B, et al. Dynamic range compression and detail enhancement of sea-surface infrared image [J]. *Infrared and Laser Engineering*, 2019, 48(1): 0126003. (in Chinese)
- [5] Wang H Q, Xu T F, Sun X L, et al. Infrared-visible video registration with matching motion trajectories of targets [J]. *Optics and Precision Engineering*, 2018, 26(6): 1533-1541. (in Chinese)
- [6] Qiao T Y, Cai L H, Li N, et al. Opposite target measurement based on infrared radiation characteristic system [J]. *Chinese Optics*, 2018, 11(5): 804-811. (in Chinese)
- [7] Wu C L, Hao Y P, Zhang L, et al. Infrared target recognition algorithm based on multi-feature fusion [J]. *Journal of Projectiles, Rockets, Missiles and Guidance*, 2019, 39(3): 39-44. (in Chinese)
- [8] Su J, Yang L, Zhang Y Y. Infrared target recognition algorithm based on fragment matching and graph searching [J]. *Acta Armamentarii*, 2015, 36(5): 854-860. (in Chinese)
- [9] Xie Z H, Liu G D. Infrared face recognition based on co-occurrence histogram of multi-scale local binary patterns [J]. *Infrared and Laser Engineering*, 2015, 44(1): 391-397. (in Chinese)
- [10] Zhao A G, Wang H L, Yang X G, et al. An affine invariant method of forward looking infra-red target recognition [J]. *Laser & Optoelectronics Progress*, 2015, 52: 071501. (in Chinese)
- [11] Zhang D F, Zhang J S, Yao K M, et al. Infrared ship-target recognition based on SVM classification [J]. *Infrared and Laser Engineering*, 2016, 45(1): 0104004. (in Chinese)
- [12] Yang C W, Wang S C, Liao S Y, et al. An infrared target recognition method based on kernel sparse coding [J]. *Infrared Technology*, 2016, 38(3): 230-235. (in Chinese)
- [13] Zhu X X, Tuia D, Mu L, et al. Deep learning in remote sensing: A comprehensive review and list of resources [J]. *IEEE Geoscience and Remote Sensing Magazine*, 2017, 5(4): 8-36.
- [14] Huang L H, Cao L H, Li N, et al. A state perception method for infrared dim and small targets with deep learning [J]. *Chinese Optics*, 2020, 13(3): 527-536. (in Chinese)
- [15] D'Acromont A, Fablet R, Baussard A, et al. CNN-based target recognition and identification for infrared imaging in defense systems [J]. *Sensors*, 2019, 19: 2040.
- [16] Shi G J. Target recognition method of infrared imagery via joint representation of deep features [J]. *Infrared and Laser Engineering*, 2021, 50(3): 20200399. (in Chinese)
- [17] Zhang H C, Nasrabadi M N, Zhang Y N, et al. Multi-view automatic target recognition using joint sparse representation [J]. *IEEE Transactions on Aerospace and Electronic System*, 2012, 48(3): 2481-2497.
- [18] Ding B Y, Wen G J. Exploiting multi-view SAR images for robust target recognition [J]. *Remote Sensing*, 2017, 9(11): 1150.
- [19] Ji S, Dunson D, Carin L. Multitask compressive sensing [J]. *IEEE Transactions on Signal Processing*, 2009, 57(1): 92-106.