

基于字典学习和拓展联合动态稀疏表示的 SAR 目标识别

曹娜¹ 王永利¹ 孙建红² 赵宁¹ 宫小泽³

摘要 提出了一种基于字典学习和拓展联合动态稀疏表示的合成孔径雷达 (Synthetic aperture radar, SAR) 图像的目标自动识别 (Automatic target recognition, ATR) 方法. 首先, 在图像预处理时, 分割出目标区域和目标遮挡地面形成的阴影区域, 将这两个区域的信息结合起来能更好地表示图像. 其次, 将字典学习方法 LC-KSVD (Label consistent k-singular value decomposition) 引入到训练阶段中, 分别学习目标区域和阴影区域的特征字典, 而不是直接将所有训练样本作为固定字典. 最后, 在测试阶段提出了拓展联合动态稀疏表示算法, 使图像数据中的两个特征共享相似但不完全相同的稀疏模式, 还可处理图像噪声遮挡损坏问题. 标准数据集上的实验结果表明, 该方法使不同类别更具区分性, 有效地提高了 SAR 图像的目标识别准确度.

关键词 字典学习, 拓展联合动态稀疏表示, 目标识别, 合成孔径雷达图像

引用格式 曹娜, 王永利, 孙建红, 赵宁, 宫小泽. 基于字典学习和拓展联合动态稀疏表示的 SAR 目标识别. 自动化学报, 2020, 46(12): 2638–2646

DOI 10.16383/j.aas.c180228

SAR Target Recognition Based on Dictionary Learning and Extended Joint Dynamic Sparse Representation

CAO Na¹ WANG Yong-Li¹ SUN Jian-Hong² ZHAO Ning¹ GONG Xiao-Ze³

Abstract This paper proposes a target automatic recognition (ATR) method for synthetic aperture radar (SAR) images based on dictionary learning and extended joint dynamic sparse representation. First of all, in the step of image pre-processing, the target area and the shadow area formed by the target obstructing the ground are segmented. Combining the information of these two areas can represent the image better. Secondly, instead of directly using all the training samples as fixed dictionaries, a dictionary learning method, LC-KSVD (Label consistent k-singular value decomposition), is introduced into the training phase to learn the feature dictionaries of target area and shadow area. Finally, the extended joint dynamic sparse representation algorithm, EJDSR (Extended joint dynamic sparse representation), is proposed in the testing phase. It allows two features in the image data to share similar but not identical sparse patterns. It can also handle the noise, occlusion, damage problems of images. The experimental result on the standard data set shows that this method makes different categories more distinguishable and effectively improves the target recognition accuracy of SAR images.

Key words Dictionary learning, extended joint dynamic sparse representation, target recognition, synthetic aperture radar (SAR) image

Citation Cao Na, Wang Yong-Li, Sun Jian-Hong, Zhao Ning, Gong Xiao-Ze. SAR target recognition based on dictionary learning and extended joint dynamic sparse representation. *Acta Automatica Sinica*, 2020, 46(12): 2638–2646

收稿日期 2018-04-19 录用日期 2018-11-05
Manuscript received April 19, 2018; accepted November 5, 2018
国家自然科学基金 (61170035, 61272420, 81674099), 中央高校基本科研业务费专项资金项目 (30916011328, 30918015103), 南京市科技计划项目 (201805036) 资助

Supported by National Natural Science Foundation of China (61170035, 61272420, 81674099), Fundamental Research Fund for the Central Universities (30916011328, 30918015103), and Nanjing Science and Technology Development Plan Project (201805036)

本文责任编辑 桑农
Recommended by Associate Editor SANG Nong
1. 南京理工大学计算机科学与工程学院 南京 210094 2. 南京理工大学电子工程与光电技术学院 南京 210094 3. 63850 部队 白城 137 001

1. School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing 210094 2. School of Electronic and Optical Engineering, Nanjing University of

合成孔径雷达 (Synthetic aperture radar, SAR) 图像的目标自动识别 (Automatic target recognition, ATR) 对军事和民用应用具有重要价值. 随着合成孔径雷达成像技术的发展, SAR ATR 受到了越来越多的关注. 但是, 由于目标具有散射特性、所处场景复杂以及对成像参数敏感等问题, SAR ATR 依然是个亟待解决的难题.

许多方法已经广泛应用于 SAR 图像自动识别, 如模板匹配^[1]、支持向量机 (Support vector machine, SVM)^[2]、卷积神经网络 (Convolutional neural networks, CNN)^[3] 等. 在模板匹配中, 将测试样

Science and Technology, Nanjing 210094 3. 63850 Troops, Baicheng 137001

本与训练样本进行匹配, 以最小平方误差作为识别准则. 文献 [2] 先使用主成分分析法将样本投影到另一个子空间以获得独特的特征, 并且用训练好的内核支持向量机识别测试样本. 文献 [3] 使用改进的卷积神经网络提取 SAR 图像特征, 再使用支持向量机对这些特征进行分类. 它们对 SAR 图像取得了较好的识别效果, 但这些方法在目标变体的识别方面识别性能仍不够理想. 目标的变体和目标属于同一类别, 但两者的结构或配置存在区别.

近年来, 稀疏表示逐渐深入应用于图像分析和分类识别等领域. 稀疏表示最早被用于人脸识别^[4], 即使在遮挡和噪声的情况下也能获得鲁棒的结果. 之后, 稀疏表示也被引入到了 SAR ATR 中. Zhang 等^[5] 提出了一种新的 SAR ATR 方法, 即联合稀疏表示 (Joint sparse representation, JSR) 方法, 该方法利用了同一个目标的多种信息或特征之间的相关性, 使稀疏表示可以将多种信息结合起来, 这是对稀疏表示的一种改进, 获得了令人印象深刻的识别结果. 与传统识别方法相比, 稀疏表示相关算法优化选择特征与测试样本最相似的一组原子表示测试样本, 原子是训练样本组成的稀疏字典的元素即稀疏字典的列向量. 换句话说, 测试样本可以由与它同类别的训练样本近似表示, 这一优势使得稀疏表示在识别 SAR 目标变体时的适应性更好.

虽然已经取得了良好的效果, 但是应用这些方法时还存在一些问题. 1) 由于传统方法中只使用目标区域的信息, ATR 性能受到限制; 2) 用于稀疏表示的字典通常由所有训练样本组成, 导致存储和计算的成本很高; 3) 文献 [5] 中的 JSR 方法约束条件非常严格, 以致多个视图的非零稀疏系数处于完全相同的位置. 此外, 获取同一目标的多个视图的难度限制了该方法在现实中的应用.

稀疏表示识别主要利用的是 SAR 图像中强度值较大的区域的信息, 这种区域是由目标散射回波形成的, 除了这种区域, SAR 图像中还包括目标遮挡地面形成的阴影区域, 但阴影区域的信息未被稀疏表示识别方法加以利用. 而文献 [6] 提出阴影信息对 SAR 目标识别也是可用的.

文献 [7] 提出了 K 奇异值分解 (K-singular value decomposition, K-SVD) 算法来学习紧凑字典. 但是, 作为无监督的字典学习方法, K-SVD 不能利用标签信息. Jiang 等^[8] 提出了标签一致 K-SVD (Label consistent k-singular value decomposition, LC-KSVD) 算法来学习辨识字典以进行识别. 更具体地说, 该算法将标签信息与每个字典原子相关联以在字典学习过程中加强识别效果.

文献 [9] 提出了一种基于联合动态稀疏表示 (Joint dynamic sparse representation, JDSR) 的

视觉识别方法. 该方法强调了多个稀疏表示向量在类级别上共享联合稀疏模式, 同时允许每个类别内的原子级别上的不同稀疏模式, 以便于灵活表示.

受文献 [6–10] 的启发, 本文提出了一种基于字典学习和拓展联合动态稀疏表示的 SAR 目标识别方法, 将其命名为 DL-EJDSR (Dictionary learning-extended joint dynamic sparse representation). 主要贡献如下: 1) 识别过程同时使用了目标区域和阴影区域的信息, 充分利用了图像信息; 2) 利用 LC-KSVD 为联合动态稀疏表示学习紧凑而独特的词典, 学习的字典尺寸较小, 不同类别之间更具特色, 可以加快识别速度, 同时提高准确度; 3) 针对图像的噪声遮挡损坏问题, 提出改进的联合稀疏表示方法, 拓展联合稀疏表示法 EJDSR, 采用灵活的原子选择过程, 使目标特征和阴影特征共享相似但不完全相同的稀疏模式. 在移动和静止目标获取和识别 (Moving and stationary target acquisition and recognition, MSTAR) 公开数据集上的实验结果显示了所提出的方法在 SAR 图像自动识别上的效率较高.

本文其他部分组织如下, 第 1 节介绍了 SAR 图像目标识别整体过程; 在第 2 节介绍了 SAR 图像预处理; 第 3 节和第 4 节分别介绍了字典学习和改进的联合动态稀疏表示算法; 第 5 节是实验部分, 通过实验来测试方法的性能, 并与其他方法进行了比较; 最后, 在第 6 节对本文进行了总结.

1 识别流程

由图 1 可见, 本文提出的基于字典学习和拓展联合动态稀疏表示的 SAR 目标识别方法包括训练和测试两个阶段.

训练阶段由预处理和字典学习两部分组成, 预处理包括对 SAR 图像进行降噪、分割、目标区域处理; 预处理完成后, 采用 LC-KSVD 方法分别对目标区域和阴影区域特征进行字典学习.

测试阶段包括预处理、拓展联合动态稀疏表示及目标识别, 首先对测试样本进行预处理; 然后利用拓展联合动态稀疏表示模型, 结合训练阶段学习到的字典, 求解测试样本目标区域信息和阴影区域信息在字典下的稀疏表示系数; 最后, 目标识别部分利用求得的稀疏系数对测试样本的目标区域信息和阴影区域信息进行重构, 根据重构误差判断测试样本的类别.

2 SAR 图像预处理

MSTAR 提供的原始 SAR 图像包含相干斑噪声和背景杂波, 为了克服噪声和背景杂波的干扰, 减

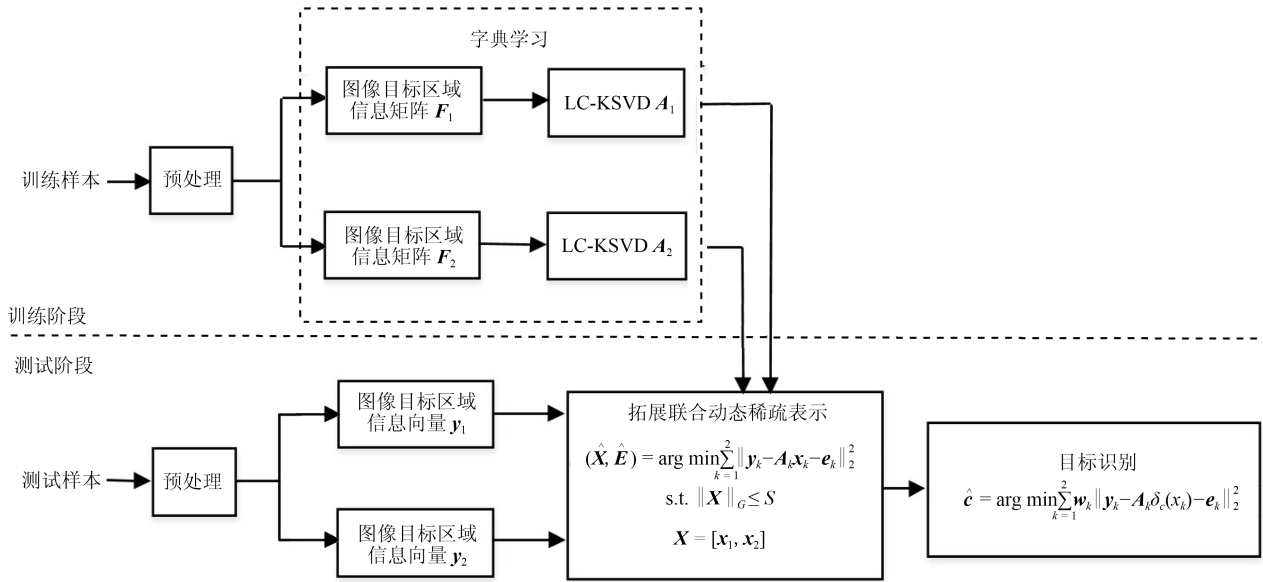


图 1 基于字典学习和拓展联合动态稀疏表示的 SAR 目标识别方法流程图
Fig.1 Flowchart of SAR target recognition based on dictionary learning and extended joint dynamic sparse representation

少分布不均匀的目标散射中心给识别性能带来的影响, 本文将对 SAR 图像进行必要的预处理, 以获得利于识别的样本, 预处理流程如图 2 所示.

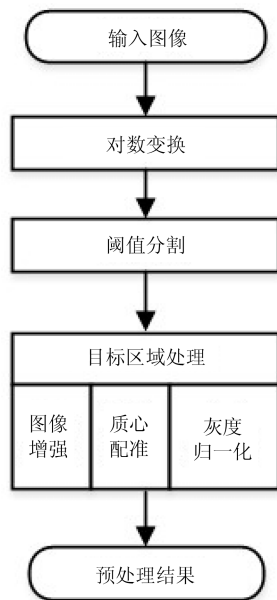


图 2 SAR 图像预处理流程
Fig.2 The preprocessing process of SAR images

首先对原始图像进行对数变换, 这可以抑制斑点噪声, 并使图像直方图更近似于高斯分布.

稀疏表示分类方法主要利用了目标散射回波形成的图像强度值较大的区域的信息, 而忽略了阴影区域的信息. 文献 [6] 表明阴影部分信息也可用于

SAR 目标识别. 分割 SAR 图像目标区域与阴影区域的方法有很多^[11], 而选择最佳的分割方法不是本文的关注点, 所以本文选择了比较简单的二维最大熵阈值分割法, 得到了 SAR 图像的目标图像和阴影图像.

图像增强可以通过某种技术提高图像质量, 本文应用幂次变换来对图像进行增强. MSTAR 数据中的 SAR 目标在图像中的位置是不定的, 这会对特征提取结果造成影响. 另外, 整个图像中目标区域面积较小, 背景部分占的区域较大, 过多的背景可能削弱目标中的分类信息, 这对目标识别是不利的. 对 MSTAR 数据, 正方形边长为 50 个像素即足以包含完整的重要区域, 因此, 分别以目标图像和阴影图像的质心为中心, 向四周扩展出 50 × 50 像素的正方形^[12], 将这一步称为质心配准. 此外, 还需要对 SAR 图像进行灰度归一化处理, 因为雷达作用距离不同时, 回波强度不同, SAR 图像整体明暗程度也会有差异.

预处理去除了 SAR 图像中的背景区域, 削弱了 SAR 图像中无关信息对分类的影响, 增强了重要的分类信息, 配准了 SAR 目标位置, 还统一了图像灰度. 这些操作将使 SAR 目标识别的分类信息更加精炼.

3 字典学习

基于稀疏表示的 SAR 目标识别要求每类目标有足够多的训练样本数以便能够张成类别子空间, 训练样本集按列排列组合而成的矩阵可以直接用作

联合动态稀疏表示的字典。但是, 它们包含很多冗余, 且需要大量的内存空间, 因此需要获得更紧凑的字典。选择一个超完备字典有两种方法。第一种方法是使用预先指定的变换矩阵, 如不完全小波、曲波、轮廓线和短时傅里叶变换, 这就为稀疏表示的评估提供了简单而快速的算法。这些字典是否具有好的性能取决于它们能否恰当并稀疏地描述信号。第二种方法是设计基于训练的词典, 如 PCA、MOD 和 K-SVD, 这种词典被认为有潜力超越固定字典。

由于 K-SVD 算法得到了广泛的认可, 因此本文使用 K-SVD 算法学习字典。经典的 K-SVD 算法^[7]可以写成

$$\begin{aligned} \langle \mathbf{A}, \mathbf{X} \rangle = \operatorname{argmin} & \|\mathbf{F} - \mathbf{A}\mathbf{X}\|_2^2 \\ \text{s. t. } \forall i, & \|\mathbf{x}_i\|_0 \leq T \end{aligned} \quad (1)$$

其中, $\|\cdot\|_2$ 表示 l_2 范数, $\|\cdot\|_0$ 表示 l_0 范数, $\mathbf{F} = [\mathbf{f}_1, \mathbf{f}_2, \dots, \mathbf{f}_N] \in \mathbf{R}^{d \times N}$ 是输入训练样本矩阵, \mathbf{f}_i 是第 i 个样本的向量, $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_M] \in \mathbf{R}^{d \times M}$ 代表学习到的字典, M 为字典尺寸, \mathbf{a}_m 是第 m 个字典原子, $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N] \in \mathbf{R}^{M \times N}$ 代表稀疏系数矩阵, \mathbf{x}_i 代表第 i 个样本的稀疏系数, $\|\mathbf{F} - \mathbf{A}\mathbf{x}\|_2^2$ 代表重构误差。作为一种无监督的字典学习方法, K-SVD 算法不能利用训练样本的标签信息。在本文中, 为了利用字典学习过程中的标签信息, 使用了 LC-KSVD 算法^[8]。该算法引入了新的标签一致性约束, 称为“判别性稀疏码误差”, 另外还加入了“分类误差”, 并将它们与重构误差相结合, 形成一个统一的目标函数, 即

$$\begin{aligned} \langle \mathbf{A}, \mathbf{X}, \mathbf{D}, \mathbf{W} \rangle = \operatorname{argmin} & \|\mathbf{F} - \mathbf{A}\mathbf{X}\|_2^2 + \\ & \alpha \|\mathbf{Q} - \mathbf{D}\mathbf{X}\|_2^2 + \beta \|\mathbf{H} - \mathbf{W}\mathbf{X}\|_2^2 \\ \text{s. t. } \forall i, & \|\mathbf{x}_i\|_0 \leq T \end{aligned} \quad (2)$$

其中, α 和 β 是控制相应项相对贡献的标量。 $\mathbf{D} \in \mathbf{R}^{M \times M}$ 是一个线性变换矩阵, 将原始的稀疏系数矩阵 \mathbf{X} 变为最具区分性的。新引入的项 $\|\mathbf{Q} - \mathbf{D}\mathbf{X}\|_2^2$ 即代表判别式稀疏编码误差, 而 $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N] \in \mathbf{R}^{M \times N}$ 用于控制稀疏模式, $\mathbf{q}_i = [q_i^1, q_i^2, \dots, q_i^M]^T \in \mathbf{R}^M$ 与训练样本 \mathbf{f}_i 有关, 当 \mathbf{f}_i 和第 m 个字典原子 \mathbf{a}_m 是同一类时, $q_i^m = 1$; 否则 $q_i^m = 0$ 。

$\|\mathbf{H} - \mathbf{W}\mathbf{X}\|_2^2$ 代表分类误差, \mathbf{W} 代表分类器参数, $\mathbf{H} = [\mathbf{h}_1, \mathbf{h}_2, \dots, \mathbf{h}_N] \in \mathbf{R}^{K \times N}$ 是输入信号 \mathbf{F} 的类别标签, K 为类别总数, $\mathbf{h}_i = [0, \dots, 0, 1, 0, \dots, 0]^T \in \mathbf{R}^K$ 是训练样本 \mathbf{f}_i 的标签向量, 非零项的位置为 \mathbf{f}_i 的类别。

4 改进的联合动态稀疏表示

通过字典学习得到了紧凑的目标图像字典和阴影图像字典后, 接着本文利用这两个字典和改进的 JDSR 模型联合动态表示测试图像数据的目标区域特征和阴影区域特征, 再根据重构残差识别测试图像, 详细介绍如下。

4.1 联合动态稀疏表示模型

为了稀疏表示从单个图像中提取的两种特征, 可以利用联合稀疏表示模型 JSR 将这两个稀疏表示问题放在一起, 即

$$\begin{aligned} \hat{\mathbf{X}} = \operatorname{argmin} & \sum_{k=1}^2 \|\mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k\|_2^2 \\ \text{s. t. } & \|\mathbf{x}_k\|_0 \leq S, \quad \forall 1 \leq k \leq 2 \end{aligned} \quad (3)$$

其中, $\hat{\mathbf{X}} = [\mathbf{x}_1, \mathbf{x}_2]$ 是测试样本 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2]$ 的稀疏系数矩阵, \mathbf{A}_k 是通过式 (2) 学习到的第 k 个特征的字典, S 为稀疏度, 这就是联合稀疏表示模型^[13]。

该目标函数对待测试样本的重构策略并不是最优的, 该模型独立地对测试样本图像中的目标图像和阴影图像进行稀疏重构, 即在重构的过程中并没有考虑到目标和阴影图像间的潜在相关性。

图 3 为不同稀疏模型的系数矩阵 \mathbf{X} 的图示, 每个列表示一个稀疏表示向量, 每个方块表示一个系数值, 白色块表示零值, 有颜色的块表示非零值。图 3 (a) 分离稀疏: 单独处理每个稀疏表示向量; 图 3 (b) 联合稀疏: 稀疏表示向量共享相同的稀疏模式; 图 3 (c) 联合动态稀疏: 多个稀疏表示向量共享相同的类级稀疏模式, 但在原子级别的稀疏模式可能是不同的。

给定从单个图像中提取的两种特征, 两个观测值的稀疏表示向量应该共享同一个类级的稀疏度模式, 而原子级稀疏度模式可能是不同的, 即具有共同的动态稀疏度。但在式 (3) 中, 两个特征的稀疏系数是独立计算的, 如图 3 (b) 所示。为了结合不同的特征, 引入 JDSR 模型^[9] 同时计算不同特征的稀疏系数, 如图 3 (c) 所示。使用动态活跃集来表示不同特征之间的关系, $\mathbf{g}_s \in \mathbf{R}^2, s = 1, 2, \dots, S$ 。在稀疏系数矩阵 \mathbf{X} 中, 每个动态活跃集指的是属于同一类的一组系数的行索引, 如 $\mathbf{g}_1 = [1, 2]$ 。在联合稀疏表示 JSR 的过程中, 多个动态活跃集被联合激活。对每个动态集合 \mathbf{g}_s 应用 l_2 范数来合并动态集合内所有原子的强度, 并且在 l_2 范数之间应用 l_0 范数来确保 JSR 中只涉及少量的动态集合。因此, JDSR 模型的规范化项可以表示为

$$\|\mathbf{X}\|_G = \left\| \left[\|\mathbf{X}_{\mathbf{g}_1}\|_2, \|\mathbf{X}_{\mathbf{g}_2}\|_2, \dots, \|\mathbf{X}_{\mathbf{g}_s}\|_2 \right] \right\|_0 \quad (4)$$

其中, $\mathbf{X}_{g_s} = \mathbf{X}(g_s) = [\mathbf{X}(g_s(1), 1), \mathbf{X}(g_s(2), 2)]^T \in \mathbf{R}^2$ 指的是与第 s 个动态活跃集 g_s 有关的系数的集合形成的向量。

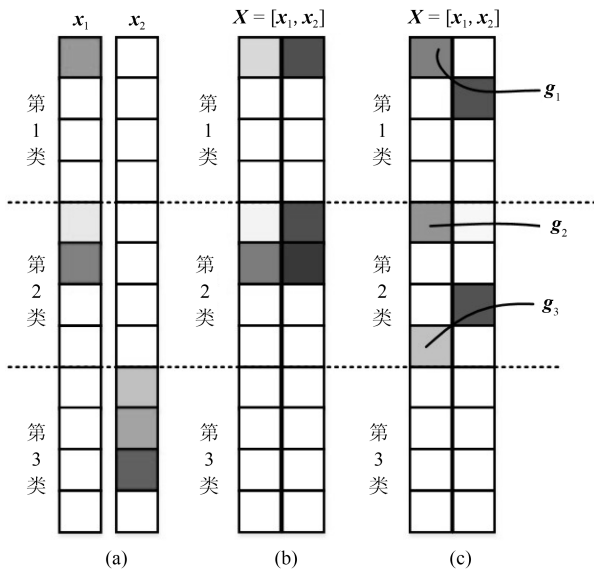


图3 不同稀疏模型的系数矩阵

Fig. 3 Illustration of coefficient matrix \mathbf{X} of different sparse models

那么, JDSR 模型可以重新表示为

$$\hat{\mathbf{X}} = \arg \min \sum_{k=1}^2 \|\mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k\|_2^2$$

$$\text{s.t. } \|\mathbf{X}\|_G \leq S \quad (5)$$

其中, S 是稀疏度, 即表示在每一个 \mathbf{x}_k 中非零元素的个数. 混合范数的规范化项不仅结合了测试图像中的目标特征和阴影特征, 同时也促进了联合动态稀疏模式在类别上的共享. 使用动态稀疏归一化项 $\|\mathbf{X}\|_G$ 有如下好处^[9]:

1) l_2 范数被用于每一个动态活跃集, 那么就可以在联合稀疏表示时将不同特征的信息结合在一起; 此外动态活跃集在从字典中选择同一类的原子时十分灵活, 它们能更好地表示多信息图像.

2) l_0 范数被用于所有动态活跃集之间, 那么就会选择最简约和有代表性的动态活跃集, 这促进了类级的联合稀疏模式, 同时允许类别间的稀疏模式是不同的, 以便于按类别选择每个特征中最具代表性的原子.

4.2 改进的联合动态稀疏模型

上述联合动态稀疏表示模型将测试样本向量分解成少数原子的线性组合, 即

$$\mathbf{y} = \mathbf{A} \mathbf{x} \quad (6)$$

但在实际应用中, 上述优化问题中低维线性模型的前提会被图像遮挡、损坏、噪声等问题打破, 因此, 本文将式 (6) 拓展为

$$\mathbf{y} = [\mathbf{A} \quad \mathbf{I}] \begin{bmatrix} \mathbf{x} \\ \mathbf{e} \end{bmatrix} = \mathbf{A} \mathbf{x} + \mathbf{e} \quad (7)$$

其中, \mathbf{I} 为单位矩阵, \mathbf{e} 是低维线性模型的系数修正向量, 表示遮挡、损坏的像素或噪声. 那么稀疏表示的优化目标函数就变为了

$$\langle \hat{\mathbf{x}}, \hat{\mathbf{e}} \rangle = \arg \min \|\mathbf{y} - \mathbf{A} \mathbf{x} - \mathbf{e}\|_2^2$$

$$\text{s.t. } \|\mathbf{x}\|_0 \leq S \quad (8)$$

理想情况下, 训练样本足够多且图像质量较高, 同一类的训练样本可以很好地表示测试样本, 稀疏系数足够稀疏, 但通常情况下图像存在噪声遮挡损坏等问题, 造成特征局部稀疏, 非零项的数目大大超过理想情况下的, 因此拓展稀疏表示模型比原始的稀疏表示模型更好. 将式 (8) 与联合动态稀疏表示结合成一个统一的框架, 记为拓展联合稀疏表示模型 (Expanded joint dynamic sparse representation, EJDSR), 该模型可表示为

$$\langle \hat{\mathbf{X}}, \hat{\mathbf{E}} \rangle = \arg \min \sum_{k=1}^2 \|\mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k - \mathbf{e}_k\|_2^2$$

$$\text{s.t. } \|\mathbf{X}\|_G \leq S \quad (9)$$

4.3 模型优化求解

本文提出的求解式 (9) 的方法类似于贪婪的匹配追踪算法, 首先初始化重构误差 $\mathbf{R} = \mathbf{Y}$, 初始动态活跃集 $\mathbf{I} = \emptyset$, 然后经过多次迭代优化求解, 直至满足停止条件.

算法 1 给出了具体的算法流程, 包括以下步骤:

步骤 1. 根据当前误差, 选择新的候选集;

步骤 2. 将新选择的动态活跃集合结合到之前已选择的动态活跃集当中;

步骤 3. 利用步骤 2 得到的动态活跃集求解相应的稀疏系数;

步骤 4. 基于最新估计的稀疏表示系数, 将混合集的稀疏度约减到一个特定的稀疏级别;

步骤 5. 更新误差;

步骤 6. 判断是否满足停止条件, 若满足则停止, 若不满足, 则转至步骤 1.

本文规定的停止条件为相邻两次迭代的重构误差的差值小于给定的阈值或达到预定义的最大迭代次数.

算法 1. 拓展联合动态稀疏表示算法 (EJDSR)

输入. 测试样本矩阵 \mathbf{Y} , 字典 \mathbf{A} , 稀疏度 S , 特征个数 K , 图像维度 d

输出. 稀疏系数矩阵 \mathbf{X}

```

01: Initialize  $\mathbf{R} \leftarrow \mathbf{Y}$ ,  $\mathbf{I} \leftarrow \emptyset$ ,  $\mathbf{E} \leftarrow \text{zeros}(d, K)$ 
02: While stopping criteria false Do
03:    $\mathbf{B} = \mathbf{A}^T \mathbf{R}$ 
04:   //通过 JDS mapping 进行原子选择
05:    $\mathbf{I}_{\text{new}} \leftarrow \mathbb{P}_{\text{JDS}}(\mathbf{B}, 2S)$ 
06:   //更新索引矩阵
07:    $\mathbf{I} \leftarrow [\mathbf{I}^T, \mathbf{I}_{\text{new}}^T]^T$ 
08:   //更新稀疏表示系数
09:   For  $k = 1, 2, \dots, K$ 
10:      $\mathbf{i} \leftarrow \mathbf{I}(:, k)$ 
11:      $\mathbf{C}(\mathbf{i}, k) \leftarrow (\mathbf{A}^T(:, \mathbf{i})\mathbf{A}(:, \mathbf{i}))^{-1} \mathbf{A}^T(:, \mathbf{i}) \times$ 
       $(\mathbf{Y}(:, k) - \mathbf{E}(:, k))$ 
12:      $\mathbf{E}(:, k) \leftarrow \mathbf{Y}(:, k) - \mathbf{A}(:, \mathbf{i})\mathbf{C}(:, k)$ 
13:   End For
14:   //通过 JDS mapping 进行原子剪枝
15:    $\mathbf{I} \leftarrow \mathbb{P}_{\text{JDS}}(\mathbf{C}, 2S)$ 
16:    $\mathbf{X} \leftarrow \mathbf{0}$ 
17:   For  $k = 1, 2, \dots, K$ 
18:      $\mathbf{i} \leftarrow \mathbf{I}(:, k)$ ,  $\mathbf{X}(\mathbf{i}, k) \leftarrow \mathbf{C}(\mathbf{i}, k)$ 
19:   End For
20:    $\mathbf{R} = \mathbf{A}\mathbf{X} - \mathbf{Y}$  //更新误差
21: End While
22:  $\mathbf{X} \leftarrow \mathbf{0}$ 
23: For  $k = 1, 2, \dots, K$ 
24:    $\mathbf{i} \leftarrow \mathbf{I}(:, k)$ 
25:    $\mathbf{X}(\mathbf{i}, k) \leftarrow (\mathbf{A}^T(:, \mathbf{i})\mathbf{A}(:, \mathbf{i}))^{-1} \mathbf{A}^T(:, \mathbf{i}) \times$ 
       $(\mathbf{Y}(:, k) - \mathbf{E}(:, k))$ 
26:    $\mathbf{E}(:, k) \leftarrow \mathbf{Y}(:, k) - \mathbf{A}(:, \mathbf{i})\mathbf{C}(:, k)$ 
27: End For

```

EJDSR 每次迭代 (步骤 1 和步骤 4) 时, 对于给定的系数矩阵 $\mathbf{Z} \in \mathbf{R}^{N \times K}$, 需要从中选择 L 条候选动态活跃集, 用它来构建与 \mathbf{Z} 最相似的矩阵 $\hat{\mathbf{Z}}_L$. 通过以下目标函数求解这个问题:

$$\begin{aligned} \hat{\mathbf{Z}}_L &= \arg \min \|\mathbf{Z} - \mathbf{Z}_L\|_F \\ \text{s. t. } \|\mathbf{Z}_L\|_G &\leq L \end{aligned} \quad (10)$$

式 (10) 可以通过名为联合动态稀疏匹配 (Joint dynamic sparsity mapping, JDS mapping) 的程序进行求解

$$\mathbf{I}_L = \mathbb{P}_{\text{JDS}}(\mathbf{Z}, L) \quad (11)$$

其中, 索引矩阵 $\mathbf{I}_L \in \mathbf{R}^{L \times 2}$ 包含 2 个特征的 top- L 动态活跃集. 由于 JDS mapping 算法比较常用, 其实现过程在此不做赘述, 可参考文献 [14]. 每一次迭代更新 \mathbf{X} 和 \mathbf{E} (步骤 3) 的过程为

$$\mathbf{x}'_k = (\mathbf{A}_k^T \mathbf{A}_k)^{-1} \mathbf{A}_k^T (\mathbf{y}_k - \mathbf{e}_k) \quad (12)$$

$$\mathbf{e}'_k = \mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k \quad (13)$$

式 (12) 和式 (13) 的具体推导过程如下:

$$\mathbf{x}'_k = \arg \min \|\mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k - \mathbf{e}_k\|_2^2 =$$

$$\begin{aligned} &\arg \min \{ \text{tr}((\mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k - \mathbf{e}_k)^T \times \\ &(\mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k - \mathbf{e}_k)) \} = \\ &\arg \min \{ \text{tr}(\mathbf{y}_k^T \mathbf{y}_k + \mathbf{x}_k^T \mathbf{A}_k^T \mathbf{A}_k \mathbf{x}_k + \\ &\mathbf{e}_k^T \mathbf{e}_k - 2\mathbf{x}_k^T \mathbf{A}_k^T \mathbf{y}_k - \\ &2\mathbf{e}_k^T \mathbf{y}_k + 2\mathbf{x}_k^T \mathbf{A}_k^T \mathbf{e}_k) \} \end{aligned} \quad (14)$$

对 \mathbf{x}_k 求偏导, 将偏导数置为 0, 得

$$-2\mathbf{A}_k^T \mathbf{y}_k + 2\mathbf{A}_k^T \mathbf{A}_k \mathbf{x}_k + 2\mathbf{A}_k^T \mathbf{e}_k = 0 \quad (15)$$

由式 (15) 可求得

$$\mathbf{x}_k = (\mathbf{A}_k^T \mathbf{A}_k)^{-1} \mathbf{A}_k^T (\mathbf{y}_k - \mathbf{e}_k) \quad (16)$$

同样地, 对 \mathbf{e}_k 求偏导, 将偏导数置为 0, 得

$$2\mathbf{e}_k - 2\mathbf{y}_k + 2\mathbf{A}_k \mathbf{x}_k = 0 \quad (17)$$

由式 (17) 可求得

$$\mathbf{e}_k = \mathbf{y}_k - \mathbf{A}_k \mathbf{x}_k \quad (18)$$

上述方法与同步正交匹配追踪方法 (Simultaneous orthogonal matching pursuit, SOMP)^[15] 和压缩同步正交匹配追踪方法 (Compressive simultaneous orthogonal matching pursuit, CoSOMP)^[16] 的算法结构相似. 它与 CoSOMP 的主要区别在于步骤 1 的原子选择和步骤 4 的原子剪枝. CoSOMP 方法中原子选择时, $\mathbf{I}_{\text{new}} \leftarrow \text{supp}(\mathbb{A}(\mathbf{B}, 2S))$, 原子剪枝时 $\mathbf{I} \leftarrow \text{supp}(\mathbb{A}(\mathbf{C}, S))$, 其中 $\mathbb{A}(\mathbf{x}, S)$ 指向量 \mathbf{x} 的 S 个最近邻, 而 EJDSR 使用的是 JDS mapping^[14], 原子选择 $\mathbf{I}_{\text{new}} \leftarrow \mathbb{P}_{\text{JDS}}(\mathbf{B}, 2S)$, 原子剪枝 $\mathbf{I} \leftarrow \mathbb{P}_{\text{JDS}}(\mathbf{C}, 2S)$.

4.4 识别标准

对于一个测试样本 $\mathbf{Y} = [\mathbf{y}_1, \mathbf{y}_2]$, 用 EJDSR 模型联合表示测试图像数据的目标区域特征和阴影区域特征, 在求得稀疏系数矩阵 $\hat{\mathbf{X}} = [\mathbf{x}_1, \mathbf{x}_2]$ 和系数修正向量 $\hat{\mathbf{E}} = [\mathbf{e}_1, \mathbf{e}_2]$ 后, 本文用最小重构残差作为评估其类别标签的标准, 可以表示为

$$\hat{c} = \arg \min \sum_{k=1}^2 \mathbf{w}_k \|\mathbf{y}_k - \mathbf{A}_k \delta_c(\mathbf{x}_k) - \mathbf{e}_k\|_2^2 \quad (19)$$

其中, \hat{c} 是测试样本的类别标签, \mathbf{w}_k 是第 k 个特征的重构参数. 算子 δ_c 代表保存与 c 类相对应的行并将所有其他行设置为零的操作.

5 实验

5.1 实验数据和机器配置

实验采用美国国防高等研究计划署公开发布的实测 SAR 地面静止军用目标数据, 即 MSTAR 数

据集. 该数据集由不同方位角和俯仰角下的各种 SAR 目标组成, 包括 3 大类: 装甲车 BMP2、装甲车 BTR70、主战坦克 T72, 每幅图像的分辨率为 $0.3\text{ m} \times 0.3\text{ m}$, 图像大小为 128×128 像素. 目标的方位角从 0° 到 360° 不等, 以俯仰角 17° 的成像数据作为训练数据, 而俯仰角 15° 的成像数据作为测试数据. 需要注意的是, 在测试集中目标 BMP2 和 T72 存在不同的变体 (BMP2-SN9563, BMP2-SN9566, BMP2-SNC21, T72-SN132, T72-SN812 和 T72-SNS7). 然而, 训练集仅具有变体 BMP2-SNC21 和 T72-SN132, 因为分类器应具备识别训练集中不存在的变体的能力. 表 1 给出了训练样本和测试样本的类别及其对应的数量.

表 1 样本的类别和数量
Table 1 Categories and quantities of samples

训练样本	样本数量	测试样本	样本数量
BMP2-SN9563	233	BMP2-SN9563	195
-	-	BMP2-SN9566	196
-	-	BMP2-SNC21	196
BTR70-SNC71	233	BTR70-SNC71	196
T72-SN132	233	T72-SN132	196
-	-	T72-SN812	195
-	-	T72-SNS7	191
总数	698	总数	1 365

本文实验的机器配置如下: Intel core i5-4570 CPU 3.2 GHz 处理器, 4 GB 内存, Windows 10 64 位操作系统.

5.2 结果分析

利用图 2 所示的预处理流程对 SAR 图像进行预处理, 实验效果如图 4, 图 4(a) 为原始图像, 图 4(b) 为对数变换后图像, 图 4(c) 为目标图像, 图 4(d) 为阴影图像, 图 4(e) 为目标区域处理后的图像, 图 4(f) 为阴影区域处理后的图像.

好的预处理结果才能保证之后的 SAR 目标识别能获得良好的识别率. 从图像分割结果可以看出, SAR 图像阈值分割可以得到较为完整的目标区域, 且分割结果中的干扰小区域数量较少; 从预处理的最终结果可以看出, 图像中孤立的伪目标区域被消除, 获得的目标轮廓和阴影轮廓比较完整, 目标的结构和边缘等细节也被保留了下来.

将本文提出的方法记为 DL + EJDSR, 按照图 1 所示的方法识别 SAR 图像目标, 表 2 为稀疏度 $S = 20$, 字典尺寸 $M = 429$ 时的识别结果, 正确识别出的测试样本总数为 1 284, 总体识别率为

94.07%, 非变体的识别率达到 99.15%, 识别效果较好.

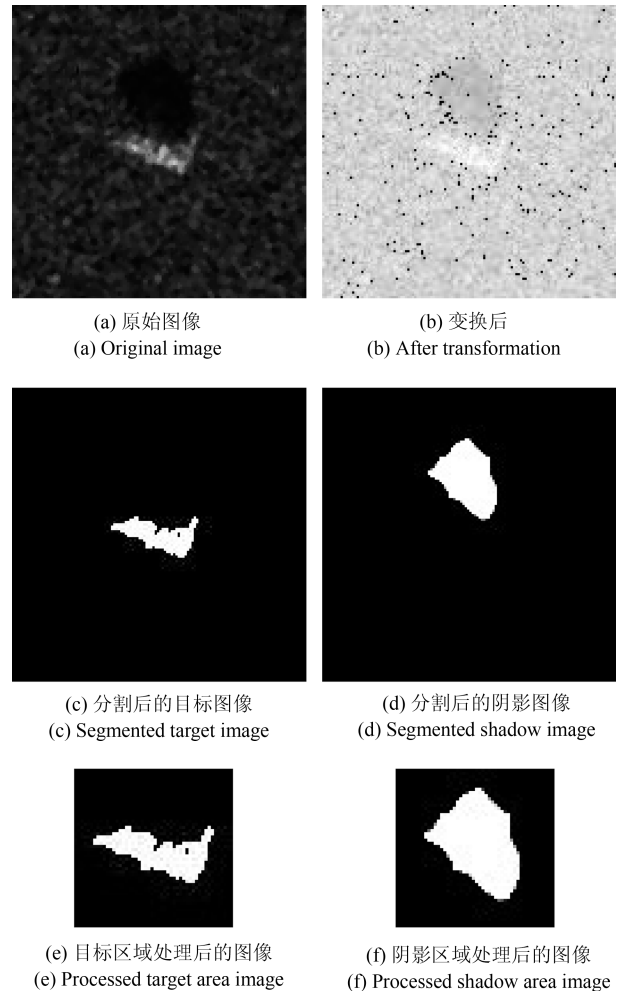


图 4 SAR 图像 (以 BMP2 为例) 预处理结果

Fig. 4 Preprocessing result of SAR image (take BMP2 as an example)

表 2 DL + EJDSR 方法的识别结果

Table 2 The identification result of DL + EJDSR

型号	BMP2	BTR70	T72	识别正确率 (%)
BMP2-SN9563	190	5	0	97.44
BMP2-SN9566	181	6	9	92.35
BMP2-SNC21	182	4	10	92.85
BTR70-SNC71	0	196	196	100
T72-SN132	0	0	196	100
T72-SN812	9	12	174	89.23
T72-SNS7	20	6	165	86.39

5.3 对比实验

将本文提出的方法与其他几种方法在字典尺寸

较小时比较识别正确率, 将字典尺寸设为 75, 实验对比结果见图 5. 这几种方法分别为:

方法 1. 文献 [4] 中提出的用稀疏表示实现 SAR 目标识别方法 (Sparse representation-based classification, SRC).

方法 2. 文献 [8] 中提出的用字典学习实现 SAR 目标识别的方法 (LC-KSVD).

方法 3. 文献 [9] 中基于联合动态稀疏表示实现 SAR 目标识别的方法 (Joint dynamic sparse representation classification, JDSRC).

方法 4. 同时利用目标区域和阴影区域信息, 结合字典学习和联合动态稀疏表示的方法 (DL + JDSR).

方法 5. 本文提出的同时利用目标区域和阴影区域信息, 结合字典学习和拓展联合动态稀疏表示的方法 (DL + EJDSR).

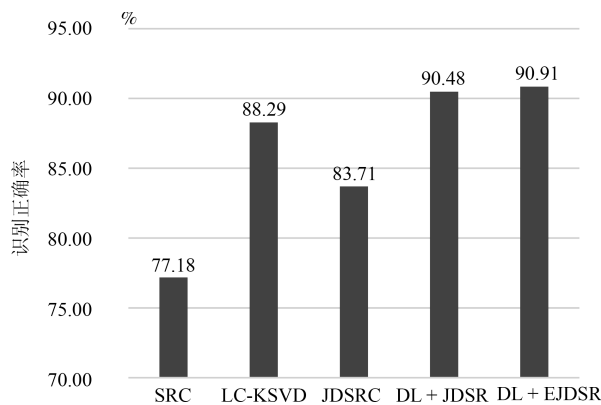


图 5 不同方法的识别正确率

Fig. 5 Recognition accuracy of different methods

此处需要说明的是, 因为图像预处理效果也会影响识别率, 为了单纯地比较对比实验中的几种 SAR 目标识别方法, 实验时这 5 种方法采用了相同的预处理过程, 而不是各方法相应参考文献的预处理过程, 因此识别率会与相关参考文献中有所出入.

从图 5 中可以看出, LC-KSVD 由于经过了鉴别字典学习, 在字典尺寸较小的情况下, 字典仍然能较好地包含图像信息, 识别率比 SRC 高很多. 而 JDSRC 的识别率也比 SRC 高, 这说明相比使用单一信息, 联合利用目标和阴影的信息能取得更好的识别效果. 与其他三种方法相比, DL + JDSR 先在训练阶段对 SAR 图像进行了字典学习, 在测试阶段又利用学习到的字典对目标特征和阴影特征进行了联合动态稀疏表示, 这两个优势使得 DL + JDSR 方法对 SAR 图像的识别率比前 3 种高. 这些方法中识别率最高的是本文提出的 DL + EJDSR 方法, 它不仅具有 DL + JDSR 的优势, 还可以处理图像噪声遮挡损坏问题, 因此识别率也相对稍微有所提

高.

5.4 参数影响

为了研究稀疏度和字典尺寸对识别性能的影响, 将字典尺寸设置为 [69, 75, 123, 153, 207, 261, 369, 429], 稀疏度设为 [10, 15, 20], 进行了多次实验. 图 6 中 3 条折线展示了稀疏度分别为 10、15、20 时, 本文方法的识别率随字典尺寸大小变化的情况. 从 3 条折线相互间的距离来看, 本文方法识别率受稀疏度变化的影响不大. 从每条折线的整体变化趋势来看, 随着字典尺寸的增加, 总识别率缓慢增长, 因此当字典尺寸较大时, 本文方法具更好的识别性能, 但由于后面增长幅度较缓, 也不需设置过大, 这样反而增加计算量.

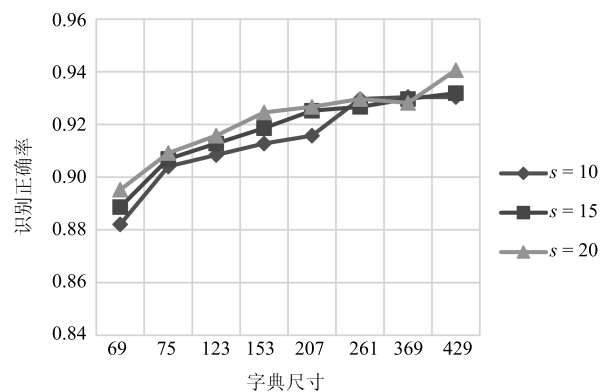


图 6 不同稀疏度下识别正确率随字典尺寸变化的情况
Fig. 6 The variety trend of recognition accuracy with the dictionary size under different sparsity

6 结束语

本文提出了一种基于字典学习和改进的联合动态稀疏表示的 SAR 图像识别方法, DL + EJDSR. 该方法利用 LC-KSVD 算法来学习紧凑而独特的词典. 此外, 它不仅利用了 SAR 图像的两个特征 (即目标区域和阴影区域) 之间的相关性, 而且还提供了更灵活的、自适应的原子选择过程, 并且它还可以处理图像噪声遮挡损坏问题. 在 MSTAR 数据集上的实验结果表明, 该方法比其他方法具有更好的性能.

References

- 1 Yang H Y, Cao Z J. Target detection performance analysis of SAR image with different resolutions based on template matching. In: Proceedings of the 3rd International Conference on Communications, Signal Processing, and Systems. Switzerland: Springer, 2015. 322: 177-185
- 2 Zhou P, Guo G, Xiong F. Research on modified SVM for classification of SAR images. In: Proceedings of the 2017 International Conference on Frontiers of Manufacturing Science and Measuring Technology. France: Atlantis Press, 2017

- 3 Zhan R H, Hu J M, Zhang J. SAR ATR based on convolutional neural network. *Journal of Radars*, 2016, **5**(3): 229–243
- 4 Wright J, Yang A Y, Ganesh A, Sastry S S, Ma Y. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2009, **31**(2): 210–227
- 5 Zhang H C, Nasrabadi N M, Zhang Y, Huang T S. Multi-view automatic target recognition using joint sparse representation. *IEEE Transactions on Aerospace and Electronic Systems*, 2012, **48**(3): 2481–2497
- 6 Ding Jun, Liu Hong-Wei, Wang Ying-Hua, Wang Zheng-Jue, Qi Hui-Jiao, Shi Li-Hui. SAR Target recognition by combining images of the shadow region and target region. *Journal of Electronic Information*, 2015, **37**(3): 594–600 (丁军, 刘宏伟, 王英华, 王正珏, 齐会娇, 时荔蕙. 一种联合阴影和目标区域图像的 SAR 目标识别方法. *电子与信息学报*, 2015, **37**(3): 594–600)
- 7 Aharon M, Elad M, Bruckstein A. K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Processing*, 2006, **54**(11): 4311–4322
- 8 Jiang Z L, Lin Z, Davis L S. Label consistent K-SVD: learning a discriminative dictionary for recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2013, **35**(11): 2651–2664
- 9 Zhang H C, Nasrabadi N M, Zhang Y, Huang T S. Joint dynamic sparse representation for multi-view face recognition. *Pattern Recognition*, 2012, **45**(4): 1290–1298
- 10 Qi Hui-Jiao, Wang Ying-Hua, Ding Jun, Liu Hong-Wei. SAR target recognition based multi-information dictionary learning and sparse representation. *System Engineering and Electronic Technology*, 2015, **37**(6): 1280–1287 (齐会娇, 王英华, 丁军, 刘宏伟. 基于多信息字典学习及稀疏表示的 SAR 目标识别. *系统工程与电子技术*, 2015, **37**(6): 1280–1287)
- 11 Lian X J, Wu Y, Zhao W, Wang F, Zhang Q, Li M. Unsupervised SAR image segmentation based on conditional triplet markov fields. *IEEE Geoscience and Remote Sensing Letters*, 2014, **11**(7): 1185–1189
- 12 Fang Qing. Research on SAR image target recognition technology based on compressive Sensing [Master thesis], University of Electronic Science and Technology, China, 2012. (方庆. 基于压缩感知的 SAR 图像目标识别技术研究 [硕士学位论文]. 电子科技大学, 中国, 2012.)
- 13 Zhang H C, Nasrabadi N M, Huang T S, Zhang Y N. Joint sparse representation based automatic target recognition in SAR images. In: Proceedings of the 2011 SPIE — The International Society for Optical Engineering. Washington, USA: SPIE, 2011. **8051**(5): 361–372
- 14 Zhou Wei. Research on face recognition based on locally constrained joint dynamic sparse representation [Master thesis], Northeast Normal University, China, 2015. (周唯. 基于局部约束联合动态稀疏表示的人脸识别研究 [硕士学位论文]. 东北师范大学, 中国, 2015.)
- 15 Tropp J A, Gilbert A C, Strauss M J. Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit. *Signal Processing*, 2006, **86**(3): 572–588
- 16 Duarte M F, Cevher V, Baraniuk R G. Model-based compressive sensing for signal ensembles. In: Proceedings of the 2009 Allerton Conference on Communication, Control, and Computing. New York, USA: IEEE, 2009. 244–250



曹娜 南京理工大学计算机科学与工程学院硕士研究生. 主要研究方向为模式识别和稀疏表示.

E-mail: 116106000721@njjust.edu.cn

(CAO Na Master student at the School of Computer Science and Engineering, Nanjing University of Science and Technology. Her research interest covers pattern recognition and sparse representation.)



王永利 博士, 南京理工大学计算机科学与工程学院教授. 主要研究方向为海量数据分析, 机器学习, 自然语言处理, 模式识别. 本文通信作者.

E-mail: yongliwang@njjust.edu.cn

(WANG Yong-Li Ph.D, professor at the School of Computer Science and Engineering, Nanjing University of Science and Technology. His research interest covers massive data analysis, machine learning, natural language processing, and pattern recognition. Corresponding author of this paper.)

His research interest covers massive data analysis, machine learning, natural language processing, and pattern recognition. Corresponding author of this paper.)



孙建红 南京理工大学电子工程与光电技术学院副教授. 主要研究方向为电路信号检测与处理, 电路故障研究, 目标识别. E-mail: sunjh1210@sina.com

(SUN Jian-Hong Associate professor at the School of Electronic and Optical Engineering, Nanjing University of Science and Technology. Her research

interest covers circuit signal detection and processing, circuit fault research, and target recognition.)



赵宁 南京理工大学计算机科学与工程学院硕士研究生. 主要研究方向为网络异常检测.

E-mail: johnny_mail@foxmail.com

(ZHAO Ning Master student at the School of Computer Science and Engineering, Nanjing University of Science and Technology. His main research interest is network anomaly detection.)

His main research interest is network anomaly detection.)



宫小泽 高级工程师. 陆军安全委员会专家库成员. 主要研究方向为数据治理, 弹药毁伤评估, 网络空间安全.

E-mail: gongxiaoze2003@sina.com

(GONG Xiao-Ze Senior engineer, member of the Army Security Council Expert Pool. His research interest covers data management, ammunition damage assessment, and cyberspace security.)

His research interest covers data management, ammunition damage assessment, and cyberspace security.)