



## 基于背景建模的VideoSAR动目标阴影检测方法

王鑫, 田甜, 田金文

引用本文:

王鑫, 田甜, 田金文. 基于背景建模的VideoSAR动目标阴影检测方法[J]. 智能系统学报, 2022, 17(1): 59–68.

WANG Xin, TIAN Tian, TIAN Jinwen. Moving target shadow detection in VideoSAR based on background modeling[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(1): 59–68.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202107019>

## 您可能感兴趣的其他文章

### 深度自编码与自更新稀疏组合的异常事件检测算法

Abnormal event detection method based on deep auto-encoder and self-updating sparse combination

智能系统学报. 2020, 15(6): 1197–1203 <https://dx.doi.org/10.11992/tis.202007003>

### 区域损失函数的孪生网络目标跟踪

Regional loss function based siamese network for object tracking

智能系统学报. 2020, 15(4): 722–731 <https://dx.doi.org/10.11992/tis.201910005>

### 一种改进的深度学习道路交通标识识别算法

An improved deep learning algorithm for road traffic identification

智能系统学报. 2020, 15(6): 1121–1130 <https://dx.doi.org/10.11992/tis.201811009>

### 多层卷积特征的真实场景下行人检测研究

Research on pedestrian detection based on multi-layer convolution feature in real scene

智能系统学报. 2019, 14(2): 306–315 <https://dx.doi.org/10.11992/tis.201710019>

### 一种鲁棒的Multi-Egocentric视频中的多目标检测及匹配算法

A robust multi-object detection and matching algorithm for multi-egocentric videos

智能系统学报. 2016, 11(5): 619–626 <https://dx.doi.org/10.11992/tis.201603050>

### 高斯核函数卷积神经网络跟踪算法

Convolutional neural network tracking algorithm accelerated by Gaussian kernel function

智能系统学报. 2018, 13(3): 388–394 <https://dx.doi.org/10.11992/tis.201612040>

微信公众平台



关注微信公众号, 获取更多资讯信息

DOI: 10.11992/tis.202107019

网络出版地址: <http://kns.cnki.net/kcms/detail/23.1538.TP.20211217.1715.006.html>.

# 基于背景建模的 VideoSAR 动目标阴影检测方法

王鑫<sup>1</sup>, 田甜<sup>1,2</sup>, 田金文<sup>1,2</sup>

(1. 华中科技大学人工智能与自动化学院, 湖北 武汉 430074; 2. 华中科技大学多谱信息处理技术国家级重点实验室, 湖北 武汉 430074)

**摘要:** 针对视频合成孔径雷达 (video synthetic aperture radar, VideoSAR) 数据进行地面运动目标检测的问题, 本文提出了一种基于单高斯背景模型的 VideoSAR 动目标阴影检测方法。该方法使用一个时间维度的滑窗对视频序列进行处理: 首先使用 RED20 深度神经网络模型抑制 VideoSAR 图像的斑点噪声, 随后使用帧间配准算法快速配准窗口内的图像序列, 然后对序列进行建模和差分得到窗口末帧的二值化前景, 最后通过连通区域筛选和区域生长剔除虚假目标。采用美国 Sandia 国家实验室公布的 VideoSAR 视频对本文算法进行了验证, 实验表明, 该算法能实现对动目标阴影的准确检测。

**关键词:** 视频合成孔径雷达; 动目标检测; 卷积神经网络; 乘性噪声; 去噪; 图像配准; 单高斯模型; 区域生长  
**中图分类号:** TP391 **文献标志码:** A **文章编号:** 1673-4785(2022)01-0059-10

中文引用格式: 王鑫, 田甜, 田金文. 基于背景建模的 VideoSAR 动目标阴影检测方法 [J]. 智能系统学报, 2022, 17(1): 59-68.

英文引用格式: WANG Xin, TIAN Tian, TIAN Jinwen. Moving target shadow detection in VideoSAR based on background modeling[J]. CAAI transactions on intelligent systems, 2022, 17(1): 59-68.

## Moving target shadow detection in VideoSAR based on background modeling

WANG Xin<sup>1</sup>, TIAN Tian<sup>1,2</sup>, TIAN Jinwen<sup>1,2</sup>

(1. School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China; 2. National Key Laboratory of Science and Technology on Multi-spectral Information Processing, Huangzhong University of Science and Technology, Wuhan 430074, China)

**Abstract:** Aiming at the problem of ground moving target detection based on Video Synthetic Aperture Radar (VideoSAR) data, a VideoSAR moving target shadow detection method based on single Gaussian background model is proposed in this paper, which uses a time-dimensional sliding window to process the video sequence: The RED20 deep neural network model is first used to suppress the speckle noise of VideoSAR image, and then the interframe registration algorithm is applied to quickly register the image sequence of the window. After that, the binary foreground of the last frame of the window is obtained by sequence modeling and background subtraction. Finally, false targets are eliminated by connected region screening and region growing. The proposed approach is validated on the VideoSAR video published by Sandia National Laboratory, and experimental results show that the algorithm can accurately detect the shadow of moving targets.

**Keywords:** VideoSAR; moving target detection; convolutional neural network; multiplicative noise; denoising; image registration; single Gaussian model; region growing

针对传统合成孔径雷达 (synthetic aperture

radar, SAR) 成像帧率较低且最小可检测速度较大的问题, 美国 Sandia 国家实验室 2003 年提出了 VideoSAR<sup>[1]</sup> 成像模式, 该模式的成像结果类似于视频, 能够实现对地面场景的高帧率、高分辨率

收稿日期: 2021-07-12. 网络出版日期: 2021-12-21.

基金项目: 国家自然科学基金项目 (42071339).

通信作者: 田甜. E-mail: [ttian@hust.edu.cn](mailto:ttian@hust.edu.cn).

成像,显著提升了对观测场景的动态感知能力,随后学者们对 VideoSAR 成像算法进行了较多的研究<sup>[2-12]</sup>。然而,在 VideoSAR 成像结果中,运动目标方位向速度分量会使它的像产生方位向的散焦,距离向速度分量会使它的像产生方位向的频移。但是,在合成孔径成像时间内,部分时段或者全程被目标遮挡的区域由于回波信号弱,在图像中呈现为黑色区域,也就是目标遮挡地物而产生的阴影。图像中的阴影反映了真实目标的存在性,因此可以通过检测运动目标阴影实现对观测场景中运动目标的检测。

当前的 VideoSAR 动目标阴影检测算法主要分为单帧检测算法和多帧检测算法。单帧检测算法基于动目标阴影的灰度特性来进行检测,又可分为传统方法和深度学习的方法。文献[13]提出了一种基于动目标阴影局部特征的检测方法。该方法先用改进的 OTSU 算法对图像进行阈值分割,然后使用形态学操作和连通域分析剔除明显的虚警,最后根据检测结果的局部信杂比大小剔除和目标近似、但与背景特征不一样的虚警。文献[14]提出一种基于改进 Faster-R-CNN 的动目标检测方法。该方法先用 K-means 确定 anchor box 的长宽和长宽比,然后训练以 FPN 和 Resnet-101 作为特征提取器的 Faster-R-CNN 对目标进行检测。多帧检测算法使用相邻的多帧图像对背景进行建模,然后使用背景差分的方式提取出前景,即运动目标阴影。文献[15]提出使用单高斯模型提取动目标阴影,该方法包括 SIFT+RANSAC 配准、单高斯背景建模获取前景、连通区域筛选、形态学操作几个步骤。文献[16]提出使用中值背景建模和三帧差分提取动目标阴影,该方法包括 SIFT+RANSAC 配准、CattePM 降噪、Tsallis 灰度熵最大化阈值分割、中值背景建模和三帧差分提取前景、形态学操作、连通区域标记等步骤。文献[17]提出使用检测前跟踪的方式检测 VideoSAR 视频序列中的运动目标。文献[18]提出了一种 VideoSAR 视频序列的配准方法,该方法先使用一个网络估计刚性变换参数,并使用这些参数进行双线性插值进行初步配准,然后将初步配准结果输入到另一个网络,估计出每个像素的位移矩阵,再次应用双线性插值得到配准结果。

VideoSAR 图像中动目标阴影是典型的弱目标,其特征较为简单,与其他的弱反射区域没有明显的区别,因此基于传统方法的单帧检测算法较难取得较好的效果。基于深度学习的单帧检测算法具有很强的学习能力,但由于 VideoSAR 视

频数据较少,深度神经网络的训练及其泛化能力测试也是困难的问题。此外,单帧算法没有利用 VideoSAR 的高帧率成像优势和动目标阴影的运动特性;而当前的多帧检测算法后处理普遍过于简单,难以有效地消除虚警。为此,本文提出了一种基于单高斯模型的 VideoSAR 动目标阴影检测方法,使用一个时间维度的滑窗对视频序列进行处理,它主要包括深度神经网络去噪、窗口序列快速配准、建模差分提取前景、后处理消除虚警四个部分,该算法结合多帧图像的信息实现对动目标阴影的有效检测。

## 1 动目标阴影检测算法

本文算法单个时间窗口的处理流程如图 1 所示,对于窗口内的所有图像,首先使用 RED20 深度神经网络模型<sup>[19]</sup>抑制图像中的斑点噪声,然后使用帧间配准<sup>[20]</sup>的方式,快速配准当前窗口中的所有图像;接着使用单高斯模型对窗口中的图像序列进行背景建模和差分,得到窗口中最后一帧的前景;最后是后处理,先剔除前景中明显不是阴影的较亮部分,再使用连通区域筛选和区域生长消除虚警,得到当前窗口中最后一帧的动目标阴影检测结果。窗口每往后移动一帧便使用上述方法检测窗口末帧的运动目标阴影,以此实现对 VideoSAR 图像每一帧的处理。

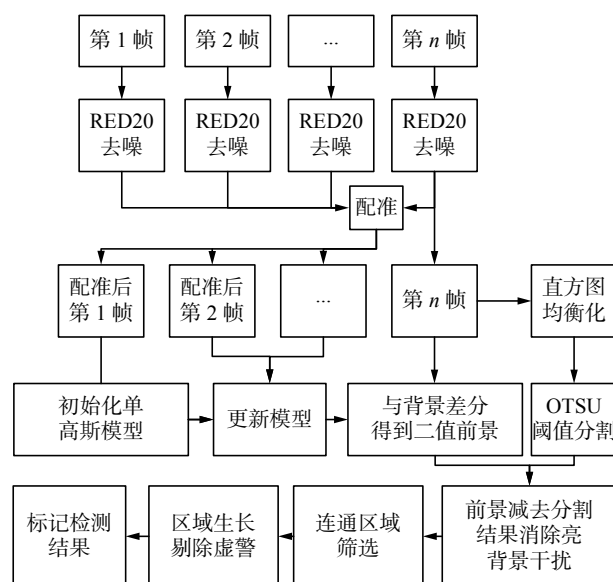


图 1 算法流程

Fig. 1 Flowchart of the proposed algorithm

### 1.1 RED20 网络去噪

斑点噪声是 VideoSAR 成像系统中固有的一种噪声,它表现为相同区域的像素值在某一平均值附近随机波动,造成相同区域不同像素点亮度



不一样,使得原本平滑的区域产生许多的亮点或者暗点。如图 2 所示,上面的红框中动目标阴影周围的道路区域和下面的红框中的区域都很不光滑,斑点噪声的存在严重影响了图像质量,继而影响了检测、识别等任务。

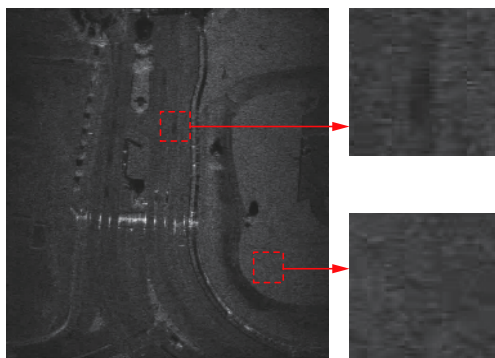


图 2 VideoSAR 图像中的斑点噪声  
Fig. 2 Speckle noises in VideoSAR image

目前 VideoSAR 图像去噪大多是基于一定的假设或者先验知识设计的算法:文献 [16] 使用了 cattePM<sup>[21]</sup> 模型抑制斑点噪声,文献 [22] 采用的是小波分析去噪<sup>[23-25]</sup>。上述传统方法在进行自适应滤波时,对噪声的感知能力有限,无法精确地获得噪声的分布情况,且对于噪声的处理方式比较单一,难以应对复杂的噪声。

为了更好地抑制斑点噪声、提高图像质量,本文率先将深度学习应用到 VideoSAR 图像去噪当中。RED20 网络<sup>[19]</sup> 原本是设计用来对可见光图像进行去噪处理的,该网络的结构如图 3 所示,网络前 10 层为卷积层,后 10 层为反卷积层,卷积层和反卷积层的核的尺寸都是 3×3,后面连接 ReLU 作为激活函数。因为池化层会丢失掉图像中一些有用的细节信息,因此网络中没有使用池化层。网络中卷积层用于提取特征消除噪声,反卷积用于恢复图像细节。在卷积层中,每隔一层就将该层连接到对称的反卷积层,因此可以直接进行正向和反向传播。卷积层与反卷积层的连接,一方面允许信号直接反向传播到底层,能解决梯度消失问题,使得深度网络更加容易实现,另一方面将图像细节从卷积层传到反卷积层,有利于恢复原始图像。而隔层连接的方式使得网络收敛更快。

RED20 使用 MSE 作为损失,其计算方式如式 (1) 所示:

$$L(\theta) = \frac{1}{N} \sum_{i=1}^N \|F(X_i; \theta) - Y_i\|_2^2 \quad (1)$$

式中:  $\theta$  表示模型参数;  $N$  表示样本对数量;  $X_i$  和  $Y_i$

分别表示含有噪声的图像及其对应的无噪声真值图。

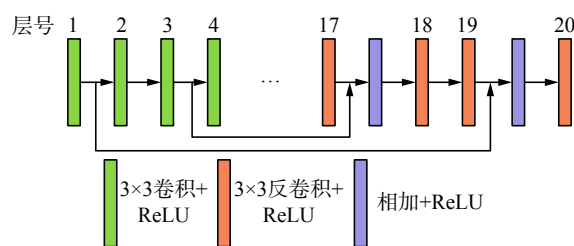


图 3 RED20 网络结构  
Fig. 3 Structure of RED20

VideoSAR 视频资源少,且没有不含噪声的真值图,因此无法构建数据集。考虑到 VideoSAR 图像与可见光灰度图具有一定的相似性,理论上来说,可以通过向可见光灰度图中添加乘性噪声来构建训练集,只要添加的乘性噪声分布涵盖了 VideoSAR 中的噪声分布,就可以将使用该数据集训练的模型迁移到 VideoSAR 去噪任务当中,所以需要估计一下 VideoSAR 视频序列中的乘性噪声的分布。对于每一个滑动窗口,具体步骤为:使用 1.2 节中的帧间配准思想,配准该窗口中的图像序列;计算配准后序列的均值,作为不含噪声的真值图;对所有配准后的图片,计算有效区域每个像素点的噪声,即“像素点的值/该像素点的均值”,将该噪声值保存下来。最后利用保存的噪声数据计算噪声的均值和方差。在 Sandia 实验室公布的视频上使用该方法计算出来的乘性噪声均值为 1.052,标准差为 0.096。考虑到噪声估计存在一定的误差,仿真数据集中添加的是均值为 1,标注差为 0.2 的服从高斯分布的乘性噪声,即使实际噪声分布和估计的噪声相似,那在强噪声数据集上训练好的模型也能较好地处理更弱的实际噪声;反之如果仿真数据集中添加的是估计出的噪声分布,一旦噪声被低估,模型的泛化能力就存疑了。

VOC2012 数据集作为基准数据之一,常被用于对象检测、图像分割网络对比实验与模型效果评估中,该数据集包含了丰富的场景,因此从 VOC2012 数据集(共 17 125 张)的每张图片中各随机裁剪出 50×50 的图像块,转成灰度图,并逐像素加入均值为 1,标准差为 0.2 的乘性噪声构成数据集,然后以 6:2:2 的比例划分为训练集、验证集和测试集,来训练 RED20 模型。随后应用该模型抑制 VideoSAR 图像中斑点噪声,第二部分的对比实验数据表明该方法是可行的。

## 1.2 基于帧间配准的序列配准

VideoSAR 系统的成像方式分为圆迹式、聚束

式和条带式,不管是哪种模式,都会使得雷达对场景的观测角度和观测距离发生变化,体现在图像上就是序列之间会发生平移、旋转、缩放、透视等几何畸变。因此在进行像素级背景建模之前,需要通过配准将同一窗口中的所有图像的背景进行对齐。

配准一组图像的传统方法是直接计算所有图像相对于参考帧的配准矩阵,假设滑动窗口的总数是  $n$ , 长度是  $s$ , 则每个窗口需要计算  $s-1$  次配准矩阵, 一共需要计算  $n \times (s-1)$  次配准矩阵, 计算量随着窗口长度的线性增加。

因此这里采用文献 [20] 提出的一种 VideoSAR 图像序列帧间配准方式: 对于所有帧, 只计算其与下一帧配准的透视变换矩阵, 任意不相邻的两帧的配准矩阵通过累乘它们之间的所有的配准矩阵得到。如图 4 所示,  $H_0 \sim H_3$  为图像序列的帧间配准矩阵, 当窗口滑动到红色框所示位置时,  $\text{img}_0$ 、 $\text{img}_2$  与  $\text{img}_3$  的配准矩阵分别为  $H_0 \times H_1 \times H_2$ 、 $H_1 \times H_2$ 、 $H_2$ , 同理当窗口滑动到蓝色框位置时配准矩阵分别为  $H_1 \times H_2 \times H_3$ 、 $H_2 \times H_3$ 、 $H_3$ 。该方法只需要计算  $n+s-1$  次配准矩阵, 远少于直接配准的  $n \times (s-1)$  次。

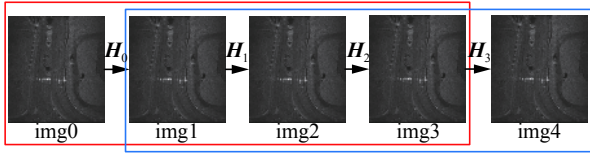


图 4 帧间快速配准示意图

Fig. 4 Schematic diagram of fast inter frame registration

如图 5 红框标出区域所示, VideoSAR 是对三维场景进行二维成像, 对于具有一定高度的物体, 不同的观测角度会得到不同的像, 且观测角度差别越大, 像的差别越大。理论上两张图像帧号间隔越小, 计算配准矩阵受到干扰越小, 因此帧间配准会比直接配准具有更高的精度, 第 2 部分的实验数据证明确实如此。



图 5 不同观测角度的成像结果差异

Fig. 5 Difference of imaging results from different observation angles

此外, 配准结果的部分区域没有像素值, 如图 6 红框标出的区域所示, 这些区域无法用于模型初始化和更新。因此, 窗口中每一帧与末帧配准时都需要计算出一个有效区域的掩码, 将窗口中所有掩码求与, 得到窗口掩码, 模型初始化、更新、背景差分均只对窗口掩码区域进行。



图 6 配准产生的无效区域

Fig. 6 Invalid region generated by registration

### 1.3 背景建模

窗口中图像序列进行背景对齐后, 使用单高斯模型进行背景建模, 背景像素  $i$  的均值和方差分别用  $\mu_i$  和  $\sigma_i^2$  表示, 初始方差用  $\sigma_{init}^2$  表示, 学习率用  $\alpha$  表示, 更新阈值和前景阈值分别用  $t_1$  和  $t_2$  表示, 窗口中第  $t$  帧中像素点  $i$  的值用  $I(x_i, y_i, t)$  表示。本文中  $\sigma_{init}^2 = 100$ ,  $t_1 = 1.35$ ,  $t_2 = 3$ ,  $\alpha = 0.1$ , 滑动窗口长度  $n=20$ 。

首先用窗口中第一帧图像初始化背景, 即  $\mu_i = I(x_i, y_i, 1)$ , 并初始化方差  $\sigma_i^2 = \sigma_{init}^2$ 。

然后用第 2 ~  $n-1$  帧更新背景和方差, 对于  $I(x_i, y_i, t)$ ,  $t \in [2, n-1]$ , 如果满足:

$$(I(x_i, y_i, t) - \mu_i)^2 < (t_1 \sigma_i)^2$$

则对  $\mu_i$  和  $\sigma_i^2$  进行如下更新, 否则不更新:

$$\begin{cases} \sigma_i^2 = \alpha(I(x_i, y_i, t) - \mu_i)^2 + (1 - \alpha)\sigma_i^2 \\ \mu_i = \alpha(I(x_i, y_i, t) - \mu_i) + (1 - \alpha)\mu_i \end{cases}$$

最后用第  $n$  帧与模型均值进行差分, 如果像素点  $I(x_i, y_i, n)$  满足:

$$\begin{cases} (I(x_i, y_i, n) - \mu_i)^2 > (t_2 \sigma_i)^2 \\ I(x_i, y_i, n) < \mu_i \end{cases}$$

则判断为前景, 否则判断为背景。至此, 初步得到了窗口最后一帧中的动目标阴影。

### 1.4 后处理

#### 1) 去除前景中过亮的部分

因为观测角度的变化, 静止场景在不同帧之间的回波强度也会有所不同。图 7 展示了一个窗口中的两帧图像, 右侧是末帧图像, 它既是配准的参考图, 也是需要提取前景的帧; 左侧是首帧图像, 它已经和右图配准并用于初始化模型的均值。红框标出了两个同名点, 它们在两帧中的灰

度差别较大,如果左图中两个同名点在背景更新后灰度值变化不大,那么背景差分时,右图比左图暗的部分(虚线围起的部分)很有可能会被判定为动目标阴影。

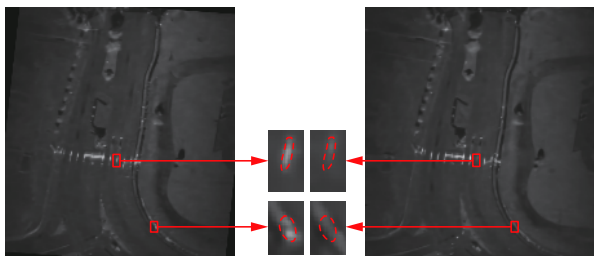


图 7 容易被误分为前景的背景

Fig. 7 Background areas that are likely to be mistakenly divided as foreground

为了消除这种干扰,先将窗口末帧进行直方图均衡化,然后再进行 OTSU 阈值分割,以得到场景中亮背景的掩码,最后用前景减去该掩码,即可消除亮背景区域的干扰。图 8 给出了图 7 右侧图像的亮背景掩码获取过程,虚线围起的区域即使在差分的时候被视为动目标阴影提取了出来,减去该掩码后即可消除此类干扰。

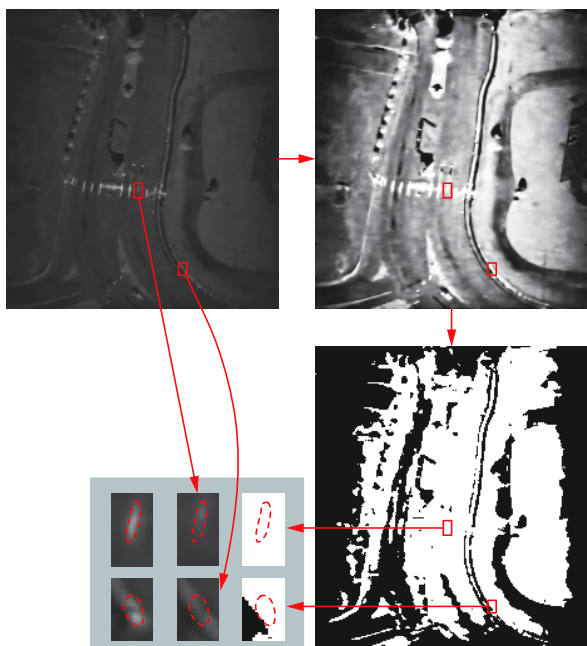


图 8 获取亮背景区域掩码示意

Fig. 8 Schematic diagram of getting the mask of the bright background area

## 2) 连通区域筛选

单高斯模型是像素级的背景建模,尽管前面已经进行了去噪、配准、滤除亮背景干扰的处理,得到的前景中仍然会存在一些离散噪点和非动目标阴影的小面积连通区域。为此对上一步得到的前景进行连通区域标记,然后根据待检测目标的尺寸、速度范围、雷达分辨率等参数,计算出动目

标阴影的面积范围,并据此对连通区域进行筛选,以消除部分虚警。

## 3) 区域生长剔除假前景

前面解释了亮背景带来的干扰,并提出了消除此类干扰的方法,同理暗背景也会产生干扰,且这种干扰无法用消除亮背景干扰的方式去除。考虑到动目标阴影与周围区域的对比度较大,而暗背景干扰与周围区域的对比度通常较小,甚至多数情况下它们与周围区域是一体的,因此我们采用区域生长的方式来消除暗背景干扰。

对于前一步得到的每一个连通区域,选择它里面灰度接近该区域灰度均值的点作为种子点。接下来是确定生长的上下限,首先计算连通区域的最小包围矩形,并适当扩大该矩形,计算矩形范围内的标准差,该标准差乘以一个常系数,得到区域生长的上下生长范围。对于暗背景,该方式计算到的标准差比较接近其真实标准差,因此区域生长会将其与周围背景连成一片,而动目标阴影因为与背景对比度较大就不会。剔除区域生长前后面积差值过大和生长结果不符合动目标阴影尺寸的连通区域,可以一定程度上消除暗背景的干扰,不过此方法对于那种孤立的暗背景干扰则效果不佳。本文采样矩形区域尺寸是连通区域的最小包围矩形尺寸的两倍,标准差系数为 5。

# 2 实验结果和讨论

## 2.1 RED20 去噪效果

网络训练的学习率为 0.0001, batch\_size 设为 16,采用的优化算法为自适应矩估计 (adaptive moment estimation, ADAM),共训练了 110 代,第 100 代的时候在验证集中的损失最小,且第 100 代的模型在测试集上测试时去噪效果良好,为此将第 100 代得到的模型应用到 VideoSAR 图像去噪中,并与文献 [16] 使用的 cattePM<sup>[21]</sup> 模型、文献 [22] 采用小波分析去噪<sup>[23-25]</sup> 进行客观指标和视觉质量的对比。cattePM 模型采用和文献 [16] 一样的参数,小波分析去噪的分解层数为 3,小波函数为 rbio1.1。

本文分别使用 3 种方法对 Sandia 实验室公布的 VideoSAR 视频的 899 帧图像进行了去噪处理,因为没有不含噪声的真值图,所以不能采用峰值信噪比 (peak signal-to-noise ratio, PSNR) 和结构相似度 (structural similarity, SSIM) 作为评价指标,因此采用等效视数 (equivalent number of looks, ENL) 作为评价标准,ENL 值越大,代表图像越平滑。表 1 是 3 种方法的 ENL 值对比,从数据可以



看出, RED20 去噪结果最平滑, 其次是 cattePM, 最后是小波分析。

表 1 不同方法的去噪结果  
Table 1 Denoising results of different methods

| 不同去噪方法  | ENL    |
|---------|--------|
| 原图      | 46.95  |
| 小波分析    | 134.04 |
| cattePM | 146.82 |
| RED20   | 173.16 |

接下来从视觉质量方面对去噪图像进行对比。首先是总体上进行对比, 因为图像比较暗, 不太容易看出它们的差别, 因此对去噪结果进行了直方图均衡化, 如图 9 所示。可以看到原图颗粒感非常严重; 小波分析能一定程度上抑制斑点噪声, 但是保留了一些噪声的纹理, 像是盖了一层毛玻璃; cattePM 让图像平滑了一些, 但是还是存在颗粒感, 只不过颗粒比原图更少且更大了; 而 RED20 则因为能够获取到不同尺度的信息和具备复杂的映射关系, 具有远强于前两者的去噪能力, 去噪结果中同质区域灰度差距小, 图像比前两者更加平滑。

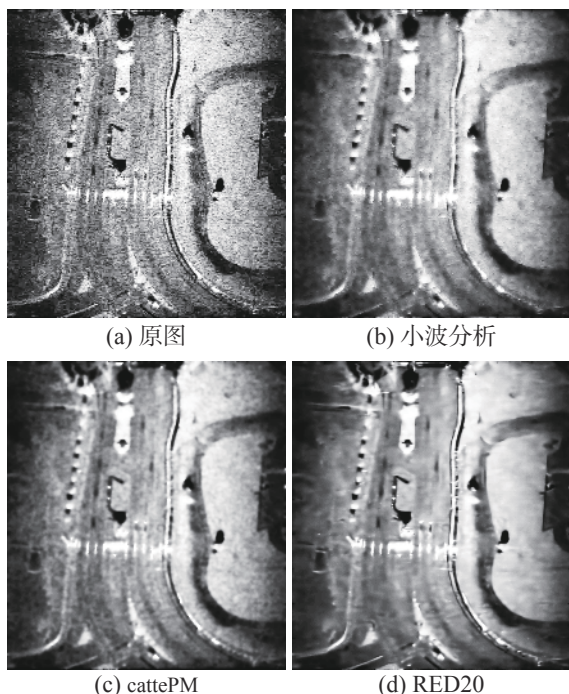


图 9 去噪效果总体对比  
Fig. 9 Overall comparison of denoising effect

接下来再从细节上进行对比, 截取图 9 中两个感兴趣区域并放大, 如图 10 所示。第一行图片中的黑色区域为动目标阴影, 其周围是道路区域, 原图不管是动目标阴影还是周围的道路区域, 内部的像素点都存在较大的差异; 小波分析

虽然抑制了部分斑点噪声, 但是像前面所说仍保留了一些噪声的纹理; cattePM 相比之下使得让阴影和道路都平滑了很多, 但道路上仍然存在较大的斑块; 而 RED20 则使得道路区域更加的平滑, 几乎看不到斑点的存在。第 2 行图像选取的是场景中一块同质背景区域, 可以看出 RED20 的去噪结果也是最佳的。综上, RED20 能在保持目标边缘的情况下, 比传统去噪算法更有效地抑制斑点噪声, 这表明, 用可见光灰度图构建数据集来训练模型, 并将其迁移到 VideoSAR 中是可行的。

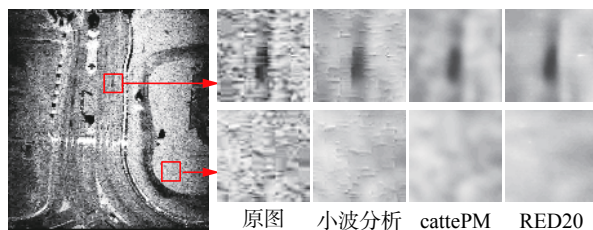


图 10 去噪效果细节对比  
Fig. 10 Detail comparison of denoising effect

## 2.2 相邻帧配准的效果

前面算法部分的分析说明, 计算帧间配准矩阵可以显著减少配准的计算量, 且理论上精度更高。这里以长度为 20 的窗口对整个视频序列滑窗进行试验, 计算有效区域掩码中参考图与配准结果的 PSNR, 将所有配准的平均 PSNR 作为评价指标。PSNR 原本是用于衡量去噪结果和真值图的差别的, 这里之所以能用它衡量配准性能, 是因为前面分析已经说明 RED20 能够有效地抑制斑点噪声, 可以认为一个窗口中的图像序列的同名点像素值是相似的, 因此配准精度越高, PSNR 会越高。表 2 是实验结果, 其中 PSNR 是 880 个窗口共计 16 720 次配准 PSNR 的平均值, 实验环境为 windows10 x64 位, CPU 为 i7-8700, 内存为 16 GB, 程序基于 opencv 3.4.6 编写。从表中可以看到帧间配准比直接配准速度快了很多, 且精度也略高于直接配准, 这与前面的分析是一致的。

表 2 不同配准方式的对比  
Table 2 Comparison of different registration methods

| 配准方法 | PSNR  | 平均配准时间/ms |
|------|-------|-----------|
| 直接配准 | 34.18 | 4071.01   |
| 帧间配准 | 34.33 | 127.85    |

## 2.3 动目标阴影检测性能对比

为了量化评估本文算法的检测性能, 将本文算法与文献 [13] 提出的传统的单帧检测算法、文献 [14] 提出的基于深度学习的单帧检测算法、文

献 [16] 提出的传统的多帧检测算法进行量化对比。Sandia 实验室公布的 VideoSAR 视频共有 899 帧, 标注了前 660, 其中第 1~400 帧用于训练文献 [14] 使用的改进 Faster-RCNN 模型, 第 401~660 帧用于测试各个算法的性能, 测试的统计结果如表 3 所示, 图 11 是各算法第 401、450、500、

550、600、650 帧的检测结果。因为原图对比度较小, 检测结果标在了直方图均衡化后的图上, 红色的框表示误检, 绿色的框表示正确检测, 蓝色的框表示漏检。文献 [16] 算法是在配准好的图像上进行的, 为了方便对比, 将其检测结果进行了坐标逆变换再标记到了配准前的图像上。

表 3 不同算法的检测结果

Table 3 Detection results of different algorithms

| 算法     | 正确检测/个 | 误检/个 | 漏检/个 | 准确率/% | 召回率/% |
|--------|--------|------|------|-------|-------|
| 文献[13] | 714    | 4322 | 709  | 14.18 | 50.18 |
| 文献[16] | 1023   | 4586 | 400  | 14.51 | 72.94 |
| 文献[14] | 1021   | 113  | 402  | 90.04 | 71.75 |
| 本文     | 1232   | 56   | 191  | 95.65 | 86.58 |

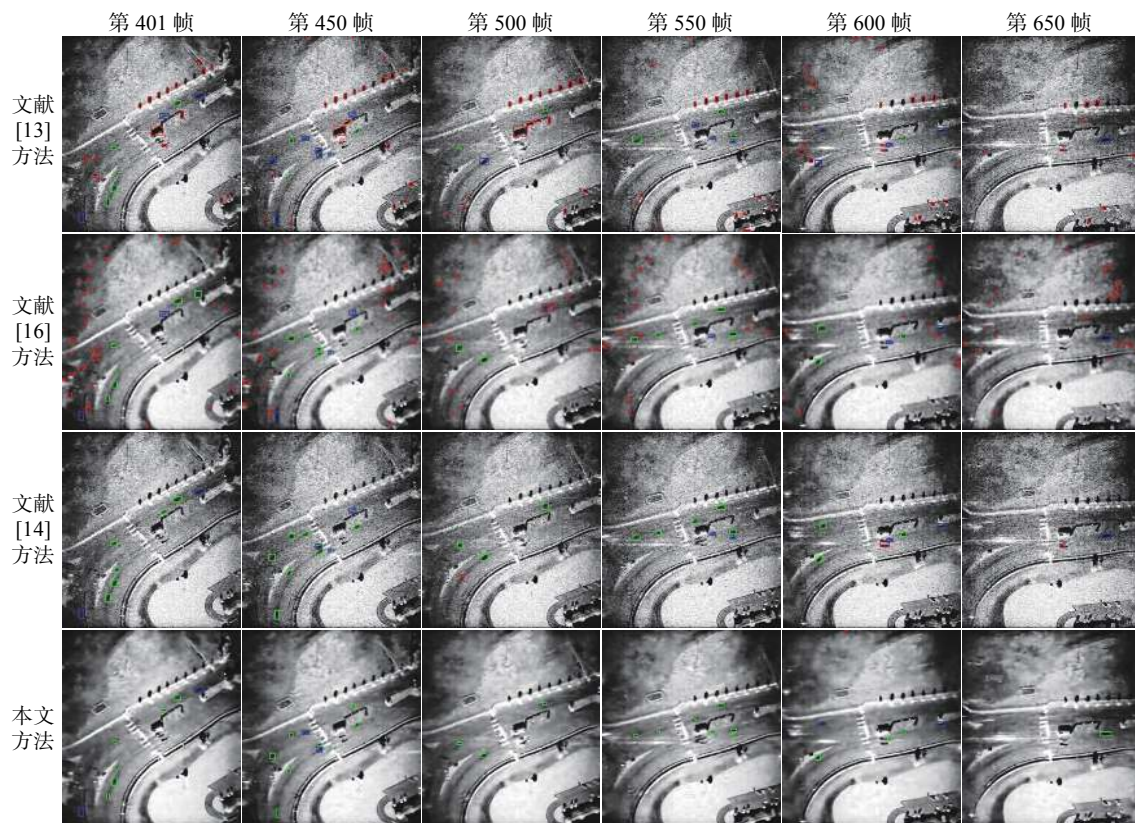


图 11 各算法第 401、450、500、550、600、650 帧的检测结果

Fig. 11 Detection results of the 401st, 450th, 500th, 550th, 600th, 650th frames of each algorithm

从表 3 和图 11 可以看出, 文献 [13] 算法的准确率和召回率都偏低, 误检和漏检严重, 它是一个传统的单帧检测算法, 流程分为 OTSU 阈值分割、形态学操作、连通区域筛选、局部信杂比筛选几个步骤。图 12 是其对第 401 帧进行检测时各个步骤的结果, 从图中可以看到, 经过连通区域筛选后仍然存在很多虚警, 而局部信杂比的筛选只能去掉其中一部分, 最终结果仍存在较多的误

检, 使得该算法的准确率较低。另外, OTSU 分割结果对召回率起着决定性作用, 图 12 最终结果中左下角和右上角的漏检是因为分割后目标与其他区域连成一片造成的, 中间的漏检是因为分割时分错类别造成的, 这就使得该算法召回率较低。

文献 [16] 算法的召回率不是很低, 但是准确率不高, 它是一个传统的多帧检测算法, 算法流程分为 SIFT+RANSAC 配准、cattePM 去噪、Tsal-



lis 阈值分割、中值建模和三帧差提取前景、形态学操作几个步骤。图 13 是其检测第 401 帧时各个步骤的结果, 可以看到该算法通过中值建模和三帧差, 能够将大部分动目标阴影提取出来, 但是用二值化结果进行中值建模还是丢失了大量的信息, 导致前景中存在许多的干扰, 而后续的形态学操作和连通区域筛选并不能很好地消除虚警, 这就导致该算法虚警率高, 影响了检测精度。

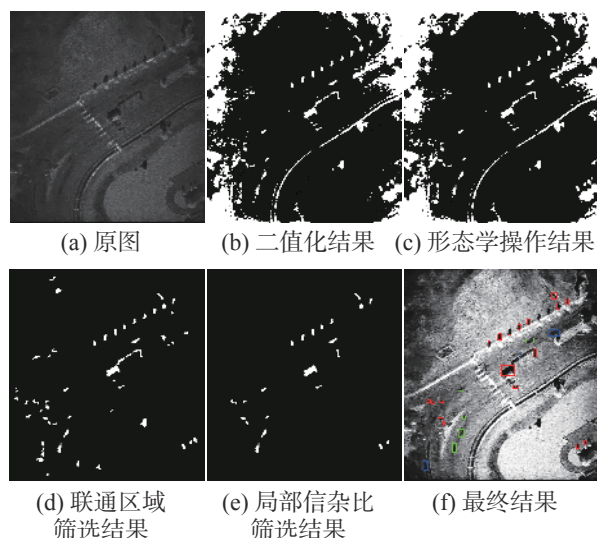


图 12 文献 [13] 算法第 401 帧检测结果

Fig. 12 Detection result of the 401th frame of the algorithm in reference [13]

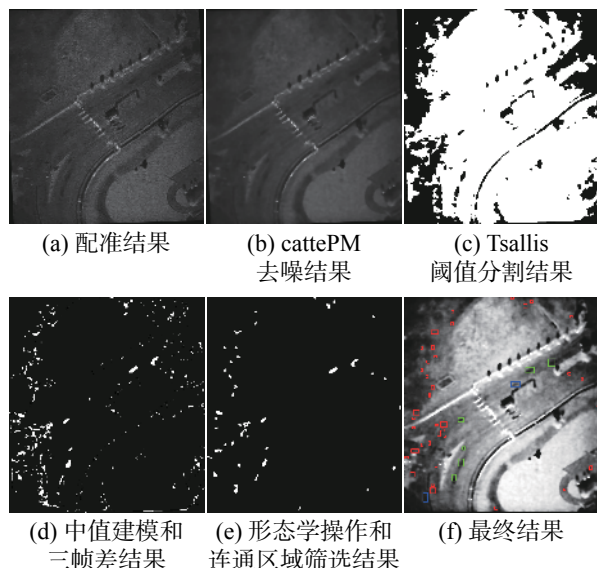


图 13 文献 [16] 第 401 帧检测结果

Fig. 13 Detection result of the 401st frame of the algorithm in reference [16]

而文献 [14] 使用的改进 Faster-RCNN 算法因为能够自动学习目标特征, 具有很强的目标检测识别能力, 表 3 和图 11 的结果证实了这一点, 但是可能由于目标特征单一且训练数据有限, 该算

法召回率不是特别高, 另外该算法是单帧检测算法, 未利用目标的运动信息, 导致仍存在一定的虚警, 影响了算法的准确率。

从上面的分析我们知道传统算法中, 初步提取的前景决定了召回率的高低, 而后处理决定了虚警的多少。图 14 是本文算法各个步骤的处理结果。从图中可以看到, 单高斯模型提取的前景中存在的干扰较少, 且大部分动目标阴影被提取出来了; 消除亮背景干扰这一步能够在保留住动目标阴影的同时, 有效地消除前景中的亮背景; 接下来的连通区域筛选去掉了小的连通区域; 而最后的区域生长消除了两个虚警 (图中彩色标记的两块区域), 不过左下角和右上角由于背景区域较暗, 动目标阴影与周围对比度太小, 各有一个目标没有被检测到, 左下角的目标四种算法都没检测到, 右上角的目标只有一个算法检测到了。

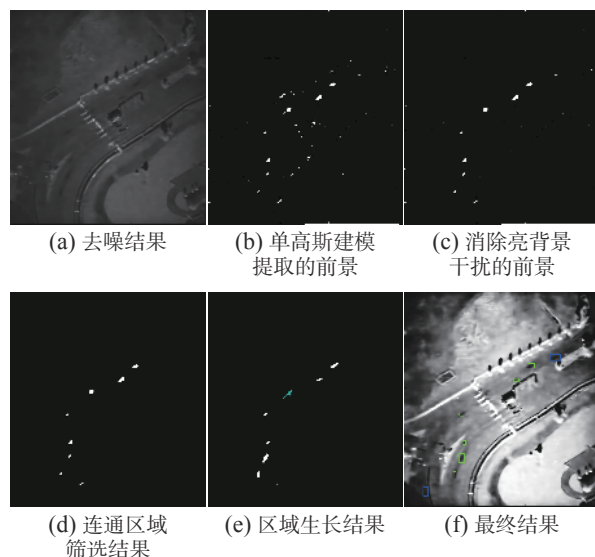


图 14 本文算法第 401 帧检测结果

Fig. 14 Detection result of the 401st frame of the proposed algorithm

从表 3 和图 12 中可以看到, 本文算法无论是准确率还是召回率, 均优于其他 3 种算法, 该算法能够在准确率极高的情况下, 拥有较高的召回率。

### 3 结束语

本文提出了一种 VideoSAR 动目标阴影检测方法, 该方法使用一个时间维度的滑动窗口结合多帧信息来检测 VideoSAR 序列中的动目标阴影。本文率先将深度学习应用到 VideoSAR 图像去噪, 表明使用可见光灰度图像训练的去噪模型能够成功迁移到 VideoSAR 图像去噪任务当中,

且该模型能够比传统方法更加有效地抑制 VideoSAR 中的斑点噪声;此外,帧间配准方式的应用使得配准速度大大提升且能略微提高配准精度;最后,区域生长的应用,弥补了许多传统算法仅仅依靠连通区域筛选来剔除虚警的不足,更好地利用了图像的空间和灰度信息。实验结果证明,该方法与传统单帧、多帧检测算法和基于深度学习的单帧检测算法相比,能够以更高的精度和召回率检测出 VideoSAR 序列中的动目标阴影。

不过本文算法在动目标阴影与背景对比度较低的帧可能会存在漏检,因此后续还需要研究如何关联帧间检测结果,以便对于目标未能被检测到的帧,可以通过历史轨迹预测出其最可能存在的位置,以此提高召回率。

## 参考文献:

- [1] WELLS L, SORENSEN K, DOERRY A, et al. Developments in sar and ifsar systems and technologies at Sandia national laboratories[C]//2003 IEEE Aerospace Conference Proceedings. New York, USA: IEEE, 2003: 1085–1095.
- [2] ZHAO Songtao, CHEN Jie, YANG Wei, et al. Image formation method for spaceborne video SAR[C]//2015 IEEE 5th Asia-Pacific Conference on Synthetic Aperture Radar. New York, USA: IEEE, 2015: 148–151.
- [3] KIM C K, AZIM M T, SINGH A K, et al. Doppler shifting technique for generating multi-frames of video SAR via sub-aperture signal processing[J]. *IEEE transactions on signal processing*, 2020, 68: 3990–4001.
- [4] MOSES R L, ASH J N. Recursive SAR imaging[J]. *Proceedings of SPIE - The International Society for Optical Engineering*, 2010, 6970: 69700P-1–69700P-12.
- [5] MOSES R L, ASH J N. An autoregressive formulation for SAR backprojection imaging[J]. *IEEE transactions on aerospace and electronic systems*, 2011, 47(4): 2860–2873.
- [6] HAWLEY R W, GARBER W L. Aperture weighting technique for video synthetic aperture radar[C]//Algorithms for synthetic aperture radar imagery XVIII. Orlando, USA: SPIE, 2011, 8051: 67–73.
- [7] LINNEHAN R, MILLER J, BISHOP E, et al. An autofocus technique for video-SAR[C]//Algorithms for synthetic aperture radar imagery XX. Baltimore, USA: SPIE, 2013, 8746: 56–65.
- [8] MILLER J, BISHOP E, DOERRY A. An application of backprojection for video SAR image formation exploiting a subaperture circular shift register[C]//Algorithms for synthetic aperture radar imagery XX. Baltimore, USA: SPIE, 2013, 8746: 66–79.
- [9] BISHOP E, LINNEHAN R, DOERRY A. Video-SAR using higher order Taylor terms for differential range[C]//2016 IEEE Radar Conference. New York, USA: IEEE, 2016.
- [10] SONG Xiaoshen, YU Weidong. Processing video-SAR data with the fast backprojection method[J]. *IEEE transactions on aerospace and electronic systems*, 2016, 52(6): 2838–2848.
- [11] HU Ruizhi, MIN Rui, PI Yiming. Interpolation-free algorithm for persistent multi-frame imaging of video-SAR[J]. *IET radar, sonar & navigation*, 2017, 11(6): 978–986.
- [12] GARREN D A. SAR focus theory of complicated range migration signatures due to moving targets[J]. *IEEE geoscience and remote sensing letters*, 2018, 15(4): 557–561.
- [13] LIU Zhongkang, AN Daoxiang, HUANG Xiaotao. Moving target shadow detection and global background reconstruction for VideoSAR based on single-frame imagery[J]. *IEEE access*, 2019, 7: 42418–42425.
- [14] 闫贺, 黄佳, 李睿安, 等. 基于改进快速区域卷积神经网络的视频 SAR 运动目标检测算法研究 [J]. *电子与信息学报*, 2021, 43(3): 615–622.  
YAN He, HUANG Jia, LI Ruian, et al. Research on video SAR moving target detection algorithm based on improved faster region-based CNN[J]. *Journal of electronics & information technology*, 2021, 43(3): 615–622.
- [15] 聊蕾, 左潇丽, 云涛, 等. 基于图像序列的 VideoSAR 动目标检测方法 [J]. *雷达科学与技术*, 2016, 14(6): 563–567, 573.  
LIAO Lei, ZUO Xiaoli, YUN Tao, et al. An approach to detect moving target in VideoSAR imagery sequence[J]. *Radar science and technology*, 2016, 14(6): 563–567, 573.
- [16] 张营, 朱岱寅, 俞翔, 等. 一种 VideoSAR 动目标阴影检测方法 [J]. *电子与信息学报*, 2017, 39(9): 2197–2202.  
ZHANG Ying, ZHU Daiyin, YU Xiang, et al. Approach to moving targets shadow detection for VideoSAR[J]. *Journal of electronics & information technology*, 2017,

- 39(9): 2197–2202.
- [17] TIAN Xiaoqing, LIU Jing, MALLICK M, et al. Simultaneous detection and tracking of moving-target shadows in ViSAR imagery[J]. *IEEE transactions on geoscience and remote sensing*, 2021, 59(2): 1182–1199.
- [18] HUANG Xuejun, DING Jinshan, GUO Qinghua. Unsupervised image registration for video SAR[J]. *IEEE journal of selected topics in applied earth observations and remote sensing*, 2020, 14: 1075–1083.
- [19] MAO Xiaojiao, SHEN Chunhua, YANG Yubin. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections[J]. *Neural information processing systems*, 2016, 29: 2810–2818.
- [20] LI Zihan, DONG Zhen, YU Anxi, et al. A robust image sequence registration algorithm for VideoSAR combining surf with inter-frame processing[C]//2019 IEEE International Geoscience and Remote Sensing Symposium. New York, USA: IEEE, 2019: 2794–2797.
- [21] ALVAREZ L, LIONS P L, MOREL J M. Image selective smoothing and edge detection by nonlinear diffusion. II[J]. *SIAM journal on numerical analysis*, 1992, 29(3): 845–866.
- [22] WANG Hongshuo, ZHAO Baojun, YANG Xingsha. Moving target detection based on Circular Video SAR[C]//2019 IEEE International Conference on Signal, Information and Data Processing. New York, USA: IEEE, 2019.
- [23] NASON G P, SILVERMAN B W. The stationary wavelet transform and some statistical applications[M]//*Wavelets and Statistics*. New York, USA: Springer New York, 1995: 103: 281–299.
- [24] COIFMAN R R, DONOHO D L. Translation-invariant de-noising[M]//*Wavelets and Statistics*. New York, USA: Springer New York, 1995: 125–150.
- [25] PESQUET J C, KRIM H, CARFANTAN H. Time-invariant orthonormal wavelet representations[J]. *IEEE transactions on signal processing*, 1996, 44(8): 1964–1970.

#### 作者简介:



王鑫, 硕士研究生, 主要研究方向为 VideoSAR 运动目标检测。



田甜, 副教授, 博士, 主要研究方向为遥感图像处理、计算机视觉、深度学习。



田金文, 教授, 博士, 主要研究方向为人工智能、计算机视觉。