

基于 ISAR 图像序列的目标三维重构技术

李敏敏, 杨利红, 吴超

(西安工业大学 光电工程学院, 西安 710021)

摘要: 从目标逆合成孔径雷达 (ISAR) 图像序列中获得目标的三维位置及结构信息, 对目标识别与解译、空间目标监视等技术十分重要; 通过对散射点特征提取与匹配及三维重构算法的研究, 实现基于 ISAR 图像序列对目标的三维重构; 首先利用 LoFTR 特征匹配算法提取目标强散射点关联位置信息, 通过粗匹配与精匹配获得特征点二维坐标并生成观测矩阵; 然后利用正交因式分解法计算目标强散射点三维位置信息, 并通过重构结果融合实现较好重构效果; 通过对序列 ALOS 卫星图像的处理, 得到了目标强散射点的三维位置及形状的重构结果; 结果表明, 该方法能够有效地从 ISAR 二维图像信息中获得目标的三维空间信息。

关键词: ISAR; LoFTR; 特征匹配; 因式分解法; 三维重构

Target 3D Reconstruction Technique Based on ISAR Image Sequence

LI Minmin, YANG Lihong, WU Chao

(School of Optoelectronic Engineering, Xi'an Technology University, Xi'an 710021, China)

Abstract: It is crucial for technologies such as target recognition and interpretation, as well as space target monitoring to obtain three-dimensional information of targets from inverse synthetic aperture radar (ISAR) image sequences. This paper researches the feature extraction, matching and 3D reconstruction technique for the scattering points, achieves the 3D reconstruction for targets based on the ISAR image sequences. Firstly, the feature matching algorithm for local feature based transformer (LoFTR) is used to extract the associated position information of the target strong scattering points. The horizontal and vertical coordinate information of the target is obtained through the coarse matching and fine matching, which generates an observation matrix. Then, the three-dimensional coordinate information of the target's strong scattering points is obtained through the orthogonal decomposition, and the shape and structure of the target are reconstructed, achieving a good reconstruction result through the fusion of reconstruction results. Finally, the three-dimensional position and shape reconstruction results of the target's strong scattering points are obtained by processing the continuous ALOS satellite images. The reconstruction results show that this method can effectively reconstruct the three-dimensional position and shape information of the target from two-dimensional multi-angle ISAR images.

Keywords: ISAR; LoFTR; feature matching; orthogonal decomposition; 3D reconstruction

0 引言

近年来, ISAR (Inverse Synthetic Aperture Radar) 高分辨成像技术获得了快速发展^[1], 因此应用序列 ISAR 图像对目标进行三维重构变得更有可能是。相较于 ISAR 二维图像, ISAR 目标的三维结构更能表现出目标的形状、尺寸、运动姿态等信息, 因此对目标进行重构能够更加充分地提取出目标的丰富信息。关于 ISAR 目标三维重构方面, 目前主要有两个方向的研究方法。第一种是通过 ISAR 成像技术和干涉处理技术结合的干涉逆合成孔径雷达^[2] (In-ISAR) 直接实现三维成像, 这种直接实现三维重构的方法对设备及系统的要求较高, 因此实现难度较大^[3]; 第二种是基于现有的单站雷达实现, 它通过雷达对目标的长时间的观测得到目标二维图像序列^[4], 然后对序列图像进行处

理并通过算法实现对目标的三维重构。该方法直接利用现有的单站雷达系统, 在复杂度和成本方面都低于第一种方法。因此充分利用单站雷达 ISAR 成像结果, 结合图像处理及三维重构技术, 实现目标的三维位置及形状重构, 对 ISAR 三维成像技术具有重要的研究意义也是 ISAR 三维成像技术的一个重要发展方向。通过第二种方法实现 ISAR 目标的三维重构的基础是对序列 ISAR 图像的特征提取与匹配, 需要提取足够多的特征点以及准确的匹配点, 在后续重构时利用这些匹配点的信息才能正确重构出目标的三维位置及形状。

要实现通过图像对目标进行三维重构首先要对序列图像进行特征提取与匹配, 也就是先对每幅图像进行目标的特征点提取, 然后通过特征描述符计算相邻匹配图像之间的关系, 设置阈值对特征点进行匹配。传统的并被广泛应

收稿日期: 2024-01-21; 修回日期: 2024-02-18。

基金项目: 国家自然科学基金(62071359, 62001364)。

作者简介: 李敏敏(1999-), 女, 硕士研究生。

通讯作者: 杨利红(1974-), 女, 博士, 教授。

引用格式: 李敏敏, 杨利红, 吴超. 基于 ISAR 图像序列的目标三维重构技术[J]. 计算机测量与控制, 2024, 32(8): 287-294.

用的特征提取算法有尺度不变特征转换算法 (SIFT, scale invariant feature transform)、快速鲁棒算法 (SURF, speeded up robust feature) 等等^[5], 他们具有不随时间、光照、角度、旋转变换等因素影响的优势, 可以准确提取目标特征点。SIFT 是通过建立不同尺度的差分高斯金字塔图像, 然后对相邻层进行极值检测得到特征点位置, 再通过特征点附近区域进行梯度及方向计算, 最终生成描述向量。然后通过最小欧氏距离计算得到图像间匹配位置信息, 得到不同图像的匹配坐标。SURF 是在 SIFT 算法的基础上做出改进, 在积分图上使用 Hessian 矩阵并且减少了特征描述向量的大小, 因此提高了算法的效率, 减少耗时。近几年随着深度学习的发展, 基于深度学习的特征提取匹配也得到了很多应用。主要是基于卷积神经网络对图像做特征提取, 得到目标比较准确且多量的特征点, 再利用随机采样一致算法进行匹配^[6], 得到不同图像匹配点。采用深度学习的特征提取与匹配相对于传统算法具有计算效率高, 匹配精确等优点。在利用深度学习做图像特征与匹配的方向中, 左潇丽等人^[7]利用深度卷积对不同时间遥感图像进行了特征提取与匹配, 通过 RepVGG 网络提取图像深层次特征, 在特征图中筛选关键点计算描述符, 然后利用反向匹配约束和 RANSAC 相结合的约束策略, 提高匹配准确率和匹配精度。徐梦莹等人^[8]针对同一场景或物体的不同图像利用 SuperPoint 框架提取图像特征点, 在基于欧氏距离匹配的基础上用网络运动统计 (GMS) 算法以及随机采样一致算法去除错误点, 最后在发生变化的图像上的匹配正确率达到 95%。在三维重构方面, 因为 ISAR 图像中的三维目标和二维图像间符合正交分解关系, 故因式分解法也可以进行逆合成的孔径雷达图像序列三维重建。Kanade 等人^[9]首先提出了因式分解法, 他们通过对多幅光学图像处理, 重构出目标的三维形状。汤立波等人^[10]将因式分解法用于 SAR 图像, 重构出运动舰船目标的三维形状。针对 ISAR 图像, 王峰等人^[11]在二维图像特征点提取上利用 KLT (Kanade-Lucas-Tomasi) 特征跟踪算法, 然后基于正交因式分解法进行三维重构, 重构结果和特征点提取数量有关, 实验结果由于特征点数目不足, 重构结果不够理想。尚士泽等人^[12]未采用传统的特征提取匹配方法, 对图像匹配方法做出改进, 主要是首先对图像中边缘及角点特征点进行手动标记, 再利用不同角度图像间投影变换的关系对图像进行匹配, 最后用正交因式分解的方法进行三维重构, 此方法需人为手动标记角点。Di^[13]等人采用基于 SURF 算法的散射中心关联匹配方法, 然后对匹配点采用因式分解法进行重构。

对于采用序列图像进行三维重构来说, 提取的强散射点的数目和散射点匹配精度十分重要。因此如何充分利用目标在不同姿态所成的二维图像序列中, 高效提取图像中散射点信息并充分利用, 对提高三维重构的质量与精度十分重要。本文利用 LoFTR 提取并匹配 ISAR 图像散射点特征, 然后利用因式分解法重构出目标。首先将连续变化的

图像序列利用卷积神经网络提取散射点特征, 通过粗匹配与精匹配得到目标匹配点的二维坐标, 将多幅图像连续匹配特征点坐标, 并调整图像中心生成观测矩阵, 最后通过因式分解法对目标进行重构及融合重构。最后通过对 ALOS 卫星的实验验证了该方法的有效性。

1 方案总体设计

本文的方案流程如图 1 所示, 主要分为 3 个部分: 1) 是对多角度的 ISAR 图像做特征提取与匹配, 得到匹配特征点的横纵坐标; 2) 是通过匹配点坐标生成观测矩阵, 然后调整图像中心、引入约束条件根据正交因式分解原理重构出目标三维结构; 3) 是通过对连续图像两两重构然后通过计算转换矩阵将重构结果融合。第一部分特征提取模块, 选择采用 CNN 卷积神经网络对图像进行不同程度的特征提取, LoFTR 用于特征匹配, 得到匹配点坐标信息, 第二部分是基于正交因式分解原理, 通过矩阵分解计算得到目标三维位置信息。第三部分是对重构结果做融合处理, 采用两两匹配重构, 计算转换矩阵再融合的方法, 得到融合重构结果。

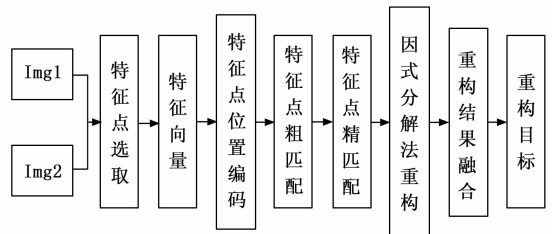


图 1 方案整体流程图

2 图像特征提取与匹配

2.1 基于 LoFTR 的特征提取与匹配

LoFTR (Local Feature based Transformer) 是一种近几年提出的用于局部图像特征匹配的方法, 相对于传统的图像特征点检测、构建描述符、特征点匹配步骤, 它是一种通过两次匹配实现更精确匹配的方法。首先在粗粒度特征图上进行密集匹配然后在精粒度特征图上进一步完善匹配结果。并且在特征向量进行转换时使用了 Transformer 中的自我注意力层和交叉注意力层 (Self and Cross Attention Layers) 来获得两幅图像的特征描述符, 这样可以使得图像在弱纹理区域产生密集匹配, 通过室内室外数据集的实验发现 LoFTR 在很大程度上优于现有技术^[14]。通过 LoFTR 对输入图像的特征提取与匹配流程如图 2 所示, 主要包括 4 部分: 1) 利用卷积对图像做特征提取并计算特征向量; 2) 对提取特征点做位置编码; 3) 对编码后的特征向量通过注意力机制重新计算得到最终用于匹配的特征向量; 4) 对特征向量首先进行粗匹配, 再在粗匹配的基础上做精匹配。

利用卷积神经网络对图像做特征提取, 相比于传统特征提取算法具有更出色的表现, 并且已经应用到各个领域。

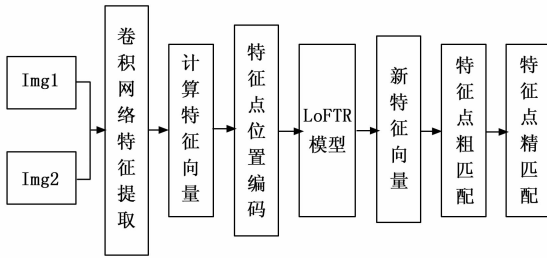


图 2 LoFTR 匹配流程图

卷积神经网络的一个局限是随着网络深度的增加, 到达一定程度后, 再增加网络深度其准确性会降低, 这是卷积神经网络的梯度消失现象^[15]。为了解决这个问题, He 等人^[16]提出了残差神经网络为基础的 ResNet 模型, 主要通过加入快捷连接实现。就是某一层的输出可以跳过一个或多个层的计算, 连接到后续的输入。因为在计算过程中有些层对输出结果不做出贡献, 或者通过某些层后计算结果反而下降, 设计跳过这些层不仅解决了梯度消失现象, 还简化了计算过程。ResNet 的模型根据网络层数不同可以分为多种, 主要有 ResNet18、ResNet34、ResNet50 等。如果需要实现更复杂的内容, 还有更深层的 ResNet 网络, 比如 ResNet101 和 ResNet152。本文使用 ResNet18 对图像做特征提取处理, 其参数设置如表 1 所示。

表 1 ResNet18 参数设置

Layer name	Output size	18-layers
Conv1	112×112	7×7, 64, stride2 3×3 max pool, stride2
Conv2 _x	56×56	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$
Conv3 _x	28×28	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$
Conv4 _x	14×14	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$
Conv5 _x	7×7	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$
	1×1	Average pool, softmax

使用 resnet18 标准卷积神经网络进行特征提取^[17]。从两幅图像中提取多层次特征, 包括表示原始图像 1/8 的粗略特征和表示原始图像 1/2 的精细特征。卷积神经网络 (CNN, convolutional neural network) 具有平移等方差和局部性的归纳偏差等优势, 非常适合用于提取图像局部特征。CNN 引入的下采样同时减少了 LoFTR 模块的输入长度, 减少了计算成本。

通过 2.1 节卷积网络提取图像特征后, 将图像粗略特征通过 LoFTR 模块来提取位置和上下文相关的局部特征。首先利用正弦或者余弦函数的相位信息对特征向量进行编码处理, 这样使得每个特征向量的每个元素提供唯一位置

信息, 通过这一方法处理后转换后的特征向量将和特征点所在图片位置信息有关, 对 LoFTR 在无区分区域产生匹配非常重要。正是这种将特征向量编码转换的方法, 使得最终匹配点增多, 匹配结果更加精确。最后通过 self and cross attention layers 得到用于匹配的特征向量。

LoFTR 中用了两种类型的可微匹配: 一种是最优传输 optimal transport (OT) 层, 另一种是采用 dual-softmax 算子。首先计算特征转换后的得分矩阵 S , 如式 (1) 所示:

$$S(i, j) = \frac{1}{\tau} \tilde{F}_r^A(i) \cdot \tilde{F}_r^B(i) \quad (1)$$

其中: $\tilde{F}_r^A(i)$ 、 $\tilde{F}_r^B(i)$ 为输入两幅图像的特征变换后的粗略特征向量。

当第一种匹配时, $-S$ 可作为部分分配问题的代价矩阵。当使用 dual-softmax 时, 匹配概率矩阵 ρ_c 如式 (2) 所示:

$$\rho_c(i, j) = \text{softmax}[S(i, \cdot)]_j \cdot \text{softmax}[S(\cdot, j)]_i \quad (2)$$

通过计算所有特征向量的得分矩阵后, 统计得到置信度概率矩阵 ρ_c , 在设置阈值信息 θ_c , 如果置信度高于阈值, 则进行特征点匹配, 最后利用互相最近邻 (MNN 标准) 去除离群匹配点。最终可以得到粗匹配的匹配结果为:

$$M_c = \{(\tilde{i}, \tilde{j}) \mid \forall (\tilde{i}, \tilde{j}) \in \text{MNN}(\rho_c), \rho_c(\tilde{i}, \tilde{j}) \geq \theta_c\} \quad (3)$$

产生粗匹配后, 对于粗匹配结果中的每个点, 第一步先找到其在精细特征图上对应的位置, 然后裁剪两组大小为 $w \times w$ 的局部窗口, 对局部区域窗口中每个点的特征经过自注意力和交叉注意力进行重构。最后, 在区域中心计算每个点于中心点的向量关系, 将中心向量与其他向量关联产生一个热力图, 通过计算概率分布上的期望, 得到匹配的最终位置。在实验中选择 5×5 的窗口实现精匹配。

2.2 基于 SIFT 的特征提取与匹配

基于传统的特征提取算法采用 SIFT 算法, 利用高斯卷积核进行尺度变换, 即高斯函数与原图像的卷积可得到图像相应尺度空间, 如式 (4) 所示:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (4)$$

其中: 高斯函数 $G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} \exp\left[-\frac{(x^2 + y^2)}{2\sigma^2}\right]$, $I(x, y)$ 表示输入图像每个位置的灰度信息, σ 为表示图像平滑程度的尺度因子, σ 越大图像平滑程度越高即越模糊, 反之亦然。通过对 σ 的不同取值, 可以得到多张模糊程度不同的图像 $L(x, y, \sigma)$, 高斯尺度空间则是由模糊程度不同的一组图像构成。高斯差分尺度空间, 可由式 (5) 得到:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (5)$$

其中: k 为比例系数, 由上式可得, 将高斯金字塔中每一组里的上下两层模糊程度不同的图像进行相减运算得到高斯差分图像。在对特征点进行提取时主要包括两个步骤, 即关键点位置确定和描述符的生成两部分。

首先对图片阈值化, 然后对上面所构造的高斯差分尺度空间中的每一个像素点进行检测, 其中包括特征点所在

层及相邻上下两层图像中 26 个相邻点，找到初始极值点。在对初始极值点进行精确定位和边缘响应消除。接下来计算关键点局部信息构造特征向量，选取最靠近关键点尺度值的图像中 3σ 的附近窗口，分别由梯度的模值计算公式算出局部梯度和方向的分布信息，梯度的模值和方向计算如式 (6) (7) 所示：

$$m(x,y) =$$

$$\sqrt{[L(x+1,y) - L(x-1,y)]^2 + [L(x,y+1) - L(x,y-1)]^2} \quad (6)$$

$$\theta(x,y) = \text{atan2}\{[L(x,y+1) - L(x,y-1)] / [L(x+1,y) - L(x-1,y)]\} \quad (7)$$

通过上式计算得到关键点的梯度值，通过统计将梯度直方图的最大值定义为主方向。

为了特征点的正确及高效匹配，需建立特征点描述符，就是用一组向量将关键点描述且让其不随其他因素的改变而改变。确定特征点的描述符首先在 4×4 子区域中计算 8 个方向梯度信息，每个子区域为一个种子点，为了保证旋转不变性调整坐标轴为关键点主方向。再对所有采样点通过划分子区域和计算不同方向梯度信息得到关键点特征向量描述符，其大小为 $4 \times 4 \times 8 = 128$ 维。最后根据计算不同角度图像之间特征向量的欧氏距离来对特征点进行匹配。

选取 ALSO 卫星不同角度两幅图像进行实验，将相邻两幅图像分别用 SIFT 算法和 LoFTR 算法做特征提取与匹配，结果如图 3 所示。

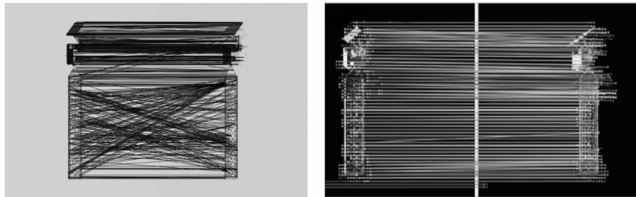


图 3 特征点匹配图

其中：左边为采用 SIFT 算法进行特征提取匹配结果，右边为 LoFTR 匹配结果，SIFT 共匹配 505 对特征点，LoFTR 匹配 487 个特征点，虽然 SIFT 匹配总点数多了 18 对，但从图中可以明显看出 SIFT 算法匹配错误率较高，在卫星下方太阳能版模块存在大部分错误匹配点，这将影响后续三维重构的结果，而基于 LoFTR 匹配结果中大部分的点得到了正确匹配，故采用 LoFTR 对 ISAR 图像序列做特征提取与匹配，更能提高重构结果精度。

2.3 图像连续匹配

将两幅图像输入 LoFTR 算法中可得到两幅图像匹配结果，在对目标的三维重构过程中通常需要对多幅图像进行连续匹配，得到特征点变化的更多信息，这样重构结果才会更加精确，对多幅图像连续匹配流程如图 4 所示。

首先随机选取 3 幅连续变换的图像，第一幅和第二幅匹配点结果记为 PP_{12} 和 PP_{21} ，第二幅和第三幅匹配结果记为 PP_{23} 和 PP_{32} ，其中第二幅图像作为共同匹配图像，其匹

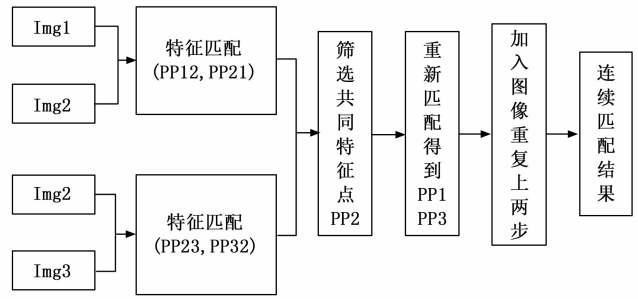


图 4 多幅图像连续匹配流程图

配坐标 PP_{21} 和 PP_{23} 中存在共同坐标，这些坐标是 3 幅图像的共同匹配坐标，也就是我们想要得到的数据，通过对 PP_{21} 和 PP_{23} 筛选相同数据，并记录数据位置，在到 PP_{12} 和 PP_{32} 中找到对应数据位置并选取就得到了 PP_1, PP_2, PP_3 共 3 幅图像匹配结果。同样的方法，再加入新的一幅图像后，与上一幅图像寻找共同特征点做筛选和匹配，通过输入多幅图像最终得到用于三维重构的匹配坐标，这将是初始的观测矩阵。通过对观测矩阵再做处理，利用因式分解法重构出目标三维信息。

3 ISAR 目标三维重构

ISAR 雷达通过发射宽带信号和目标与雷达视线方向转角的累积实现距离维、方向维的高分辨率。通过对序列回波数据的距离维脉冲压缩及方位维压缩可实现 ISAR 二维成像。如图 5 所示，系统观测得到特定角度的回波数据后，再分割为相互重叠的若干个子孔径，由此可以得到若干个连续的 ISAR 图形，并构成了三维重构时所需要的多角度图像。

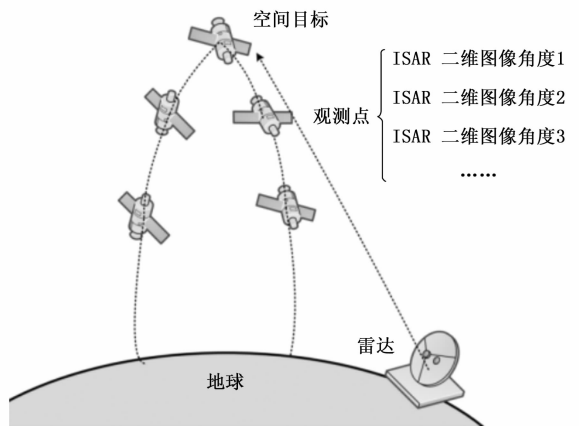


图 5 ISAR 雷达对目标成序列像

通过 ISAR 成像原理可知，ISAR 图像中散射点的位置由雷达的转动向量 ω_f 和雷达视线向量 ρ_f 决定^[18]。将 ISAR 雷达的径向距离和横向多普勒作用记作 $h_f = -\rho_f^T; v_f = -(\omega_f \times \rho_f)^T$ ，ISAR 投影矩阵可表示为：

$$\begin{pmatrix} h_f \\ v_f \end{pmatrix} \begin{pmatrix} x_k \\ y_k \\ z_k \end{pmatrix} = \begin{pmatrix} r_{fk} \\ d_{fk} \end{pmatrix} \quad (8)$$

通过上式可得到目标第 k 个散射点三维坐标和二维投影平面的关系, 那么 Q 个散射点 F 幅图像的运动矩阵和目标三维坐标矩阵分别为:

$$\mathbf{M} = \begin{bmatrix} h_1 \\ \vdots \\ h_f \\ v_1 \\ \vdots \\ v_f \end{bmatrix} \quad \mathbf{S} = \begin{bmatrix} x_1, \dots, x_Q \\ y_1, \dots, y_Q \\ z_1, \dots, z_Q \end{bmatrix} \quad (9)$$

通过推理可将所有序列图像和所有关联散射点写进投影方程 $\mathbf{W} = \mathbf{MS}$, 其中 $\mathbf{M} \in R^{2F \times 3}$, $\mathbf{S} \in R^{3 \times Q}$, $\mathbf{W} \in R^{2F \times Q}$ 。再利用正交因式分解原理, 可从提取的观测矩阵 \mathbf{W} 中重构分解出目标三维位置坐标矩阵。

3.1 因式分解法重构原理

ISAR 成像过程及投影模型从数学角度分析基本满足正交投影关系, 因此通过因式分解法来恢复出目标的三维形状和相对运动信息可以实现^[19-20]。假定利用 ISAR 成像得到了某一目标二维图像序列。一组图像序列中共有 F 帧图像, 每幅图像跟踪到 Q 个共同特征点, 将各个图像特征点位置坐标放入一个矩阵 \mathbf{W} 中可得到如式 (10) 所示的矩阵:

$$\mathbf{W} = \begin{bmatrix} x_{11} & \cdots & x_{1Q} \\ \vdots & \ddots & \vdots \\ x_{F1} & \cdots & x_{FQ} \\ y_{11} & \cdots & y_{1Q} \\ \vdots & \ddots & \vdots \\ y_{F1} & \cdots & y_{FQ} \end{bmatrix} \quad (10)$$

\mathbf{W} 矩阵称为观测矩阵, 其上半部分为每幅图像每个散射点的横坐标值, 下半部分为每幅图像每个散射点的纵坐标值。假设对于提取到的每个散射点, 其空间三维坐标如式 (11) 所示:

$$\mathbf{S}_p = [x_p, y_p, z_p]^T, p = 1, 2, \dots, Q \quad (11)$$

且目标三维坐标系的原点为所有散射点的质心, 则有 $\mathbf{S} \in R^{3 \times P}$:

$$\mathbf{S} = [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \cdots \quad \mathbf{S}_Q] \sum_{Q=1}^Q \mathbf{S}_Q = 0 \quad (12)$$

从目标二维图像来看, 获取的二维图像序列是由其 x 轴、 y 轴的一组单位正交向量 $\mathbf{i}_f, \mathbf{j}_f$ 决定的。每幅图像对应的一组单位向量可组成映射矩阵:

$$\mathbf{M} = [\mathbf{i}_1 \quad \cdots \quad \mathbf{i}_f \quad \mathbf{j}_1 \quad \cdots \quad \mathbf{j}_f]^T \quad (13)$$

因此可以得到如式 (14) 所示的正交投影公式:

$$\mathbf{W} = \mathbf{MS} \quad (14)$$

因为 $2F \times Q$ 维的观测矩阵 \mathbf{W} 是由 $2F \times 3$ 维的映射矩阵 \mathbf{M} 和 $3 \times P$ 维点三维坐标矩阵 \mathbf{S} 相乘, 所以 \mathbf{W} 矩阵的秩最多为 3。基于秩理论可利用子空间计算方法从观测矩阵 \mathbf{W} 中恢复出坐标矩阵 \mathbf{S} ^[21]。

将上节匹配得到的特征点信息作为观测矩阵输入, 并将每幅图像的中心调整一致后生成新的观测矩阵作为因式分解法中的输入矩阵。在三维重构中首先对输入的观测矩

阵做奇异值分解 (SVD, singular value decomposition), 如式 (15) 所示:

$$\mathbf{W} = \mathbf{U}\mathbf{\Delta}\mathbf{V}^T \quad (15)$$

其中: $\mathbf{\Delta} \in R^{2F \times P}$ 为只有对角位置有元素的矩阵, 它的对角元素为观测矩阵所对应的奇异值, $\mathbf{U} \in R^{2F \times 2F}$ 、 $\mathbf{V} \in R^{P \times P}$ 都是列矩阵且满足正交关系。由于在实际雷达观测过程中会存在噪声, 这会导致观测矩阵秩不为 3, 所以要对分解后的观测矩阵进行降秩处理, 将其秩近似为 3。

在降秩的过程中取矩阵 \mathbf{U} 的前三列数据, 记为矩阵 $\mathbf{U}' \in 2F \times 3$, 取矩阵 \mathbf{V} 的前三行数据, 记为矩阵 $\mathbf{V}' \in P \times 3$, 将对角矩阵中的值按照从大到小排列取最大的 3 个值记为 $\mathbf{\Delta}' = \text{diag}(\sigma_1, \sigma_2, \sigma_3)$; 且 $\sigma_1, \sigma_2, \sigma_3$ 为 \mathbf{W} 的 3 个主特征值, 满足 $\sigma_1 > \sigma_2 > \sigma_3$ 。

此时, 令:

$$\begin{aligned} \mathbf{M} &= \mathbf{U}' \sqrt{\mathbf{\Delta}'} \\ \hat{\mathbf{S}} &= \sqrt{\mathbf{\Delta}'} \mathbf{V}'^T \end{aligned} \quad (16)$$

根据式 (16) 可重构出散射点的空间三维位置信息。

实验中首先提取并记录 2.2 节中匹配到的特征点坐标, 将横坐标和纵坐标分开处理, 分别计算所有横坐标的均值和纵坐标的均值, 如式 (17) 所示, 然后用每一个坐标值减去其均值, 得到特征点坐标值, 使所有匹配的图像的中心一致。

$$\begin{cases} \tilde{x}_{fq} = x_{fq} - x_f \\ \tilde{y}_{fq} = y_{fq} - y_f \end{cases} \quad (17)$$

其中: x_{fq} 表示第 F 幅图像中第 q 个匹配特征点横坐标, y_{fq} 为第 F 幅图像中第 q 个匹配特征点纵坐标, x_f, y_f 为所有匹配坐标的均值, 计算如式 (18) 所示:

$$\begin{cases} x_f = \frac{1}{Q} \sum_{q=1}^Q x_{fq} \\ y_f = \frac{1}{Q} \sum_{q=1}^Q y_{fq} \end{cases} \quad (18)$$

通过如上处理将得到的数据输入为因式分解法中的观测矩阵 \mathbf{W} , 通过矩阵分解计算出目标的运动矩阵 \mathbf{M} 和三维坐标矩阵 \mathbf{S} 。

3.2 目标融合重构

通过对输入 ISAR 序列图像特征点提取后利用 2.3 中连续匹配方法进行匹配, 在实验过程中会发现, 由于不停的特征匹配及特征筛选会导致匹配到的特征点越来越少, 随着输入图片越多, 在这些图片中共同存在并匹配的特征点越来越少, 将不利于对目标的三维重构, 会导致不能完整地目标结构重现出来。

为了充分利用提取及匹配到的特征点, 采用一种两两匹配重构再融合的方法, 这种方法将大量保留匹配到的特征点信息, 能够重构出目标的结构。两两匹配重构融合的方法流程如图 6 所示。

首先通过对两幅图像的特征匹配后得到匹配点坐标, 将匹配坐标处理后成为因式分解重构的观测矩阵, 再由矩阵计算重构出三维结果。对第一幅图像和第二幅图像重构

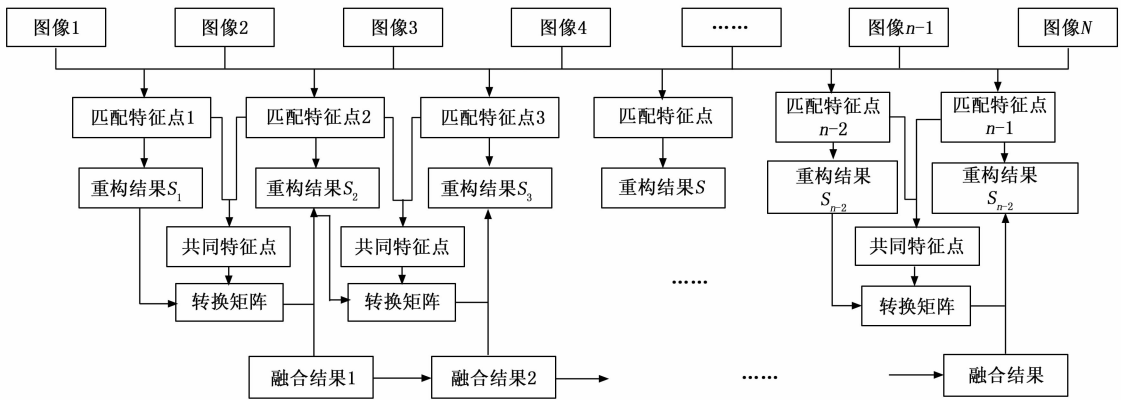


图 6 重构融合流程图

出来结果记为 S_1 ，第二幅图像和第三幅图像重构出来结果记为 S_2 ，以此类推对输入的图像序列分别做两两匹配重构。要对重构结果融合首先要计算重构出的三维结果之间的转换关系，再将所有结果融合到一个坐标系下。

第一步输入 3 幅图像并提取其特征点，得到前两幅图像特征匹配点分别为 PP_{12} ， PP_{21} ，后两幅图像特征点匹配结果为 PP_{23} ， PP_{32} ，将这两对数据作为输入分别利用因式分解法重构出结果，前两幅结果记为 S_1 ，后两幅图像重构结果记为 S_2 。

第二步选择 3 幅图像中共同匹配的特征点，首先从 PP_{21} 和 PP_{23} 中筛选共同点，得到结果记为 PP_2 ，不存在 3 幅图像中的共同匹配点记为 NP_2 ，再利用共同点的位置对 PP_{12} 和 PP_{32} 中的点筛选得到最终匹配点坐标 PP_1 和 PP_3 。再分别将 PP_1 ， PP_2 和 PP_2 ， PP_3 作为因式分解法的输入，重构出匹配点三维结构分别记为 S_{11} 和 S_{22} 。

在第二步的两个重构结果 S_{11} 和 S_{22} 中，这是二维图像中相同特征点在不同坐标系下的重构结果，因此计算出 S_{11} 和 S_{22} 之间的坐标转换关系即可将重构出的坐标点进行融合。因此第三步计算出两个坐标系的转换矩阵 T 。

最后，重构结果 S_{11} 是 S_1 的子集，所以他们俩重构结果是在一个坐标系下的，同理可得 S_{22} 和 S_2 也是同一个坐标系，所以 S_{11} 和 S_{22} 之间的转换矩阵也可用于 S_1 和 S_2 间的转换。最后一步将 S_1 与 S_2 利用转换矩阵 T 转换到同一坐标系

下得到重构融合结果。

再加入一幅图像后依然计算其转换矩阵进行重构结果融合，以此类推达到对输入序列图像的重构融合。如果利用 2.3 中序列图像连续特征提取与匹配的方法，会在匹配过程中失去大量不是共同特征点的坐标，不利于重构结果。利用重构让融合的方法让所有的匹配点都发挥了作用，将筛选掉的不是共同特征点的点集 NP_2 的重构结果利用转换矩阵也融合到重构结果。

4 实验结果与分析

为了验证 LoFTR 算法及因式分解法对 ISAR 目标重构的有效性，本文利用 ALOS 卫星的二维 ISAR 图像序列实现目标三维重构。图 7 为连续变化的一组二维图像序列，其中每幅图像俯仰角转变 5° 航偏角和翻滚角固定不变，图 4 为航偏角 90° 、翻滚角 0° 、俯仰角 $15 \sim 40^\circ$ 时的 ALOS 卫星图像。

可以看到卫星沿一个方向做连续变化，且变化角度不大，因此利用这些图像做特征提取能达到一个较好的匹配结果。实验中分别用基于连续匹配重构的方法和基于先匹配重构再做融合的方法重构出结果。

将卫星 ISAR 图像序列首先通过 LoFTR 进行特征提取与匹配，得到其匹配点坐标矩阵。

通过连续匹配后得到多幅图像匹配坐标生成目标观测矩阵，然后利用 3.1 节中正交因式分解原理对目标进行三

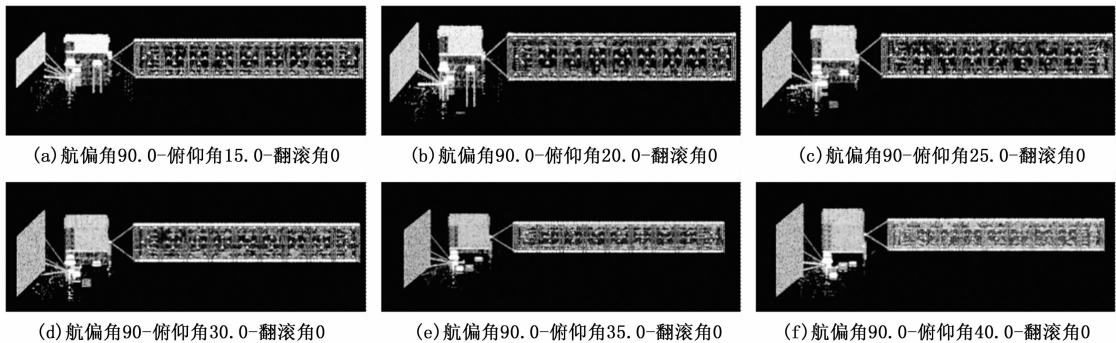
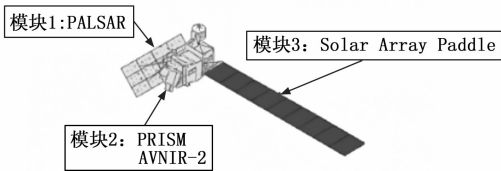
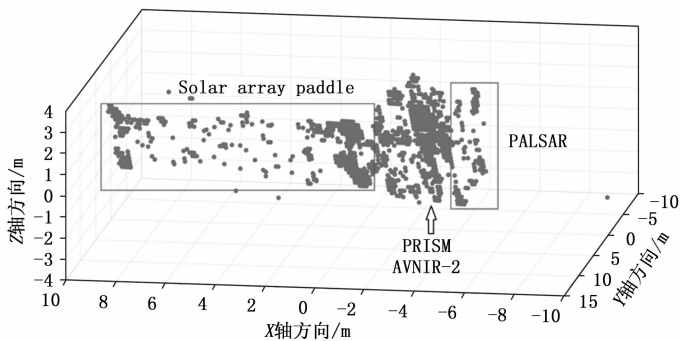


图 7 ALOS 卫星图像序列



(a) ALOS卫星结构图



(b) ALOS卫星重构结果

图 8 卫星结构及重构结果

维重构, 重构结果如图 8 所示。其中图 8 (a) 为 ALSO 卫星结构图, 可以看到卫星主要结构由 3 个模块构成包括 PALSAR、PRISM、AVNIR-2 共 3 个传感器以及太阳能板部分; 图 (b) 为通过 LoFTR 算法连续特征提取与匹配采用正交因式分解法重构目标结果。

从图 8 可以看出, 通过重构卫星 3 部分主要结构已经出现, 因此说明通过因式分解法重构可以得到目标的主要形状结构及其空间三维位置信息。但是不足的是重构结果中卫星太阳能板部分大部分点缺失, 只有轮廓信息, 还存在少量离群点。目标部分结果缺失是由于特征提取中散射点未提取到以及在匹配过程中删掉了未匹配上特征点, 结果导致部分模块没有提取与匹配到的特征点, 故其重构结果结构会有缺失。重构图中存在离目标较远的点, 这是错误重构点, 由于在匹配过程中存在误匹配点以及重构过程中的误差, 导致重构出错误离群点。

在对连续多幅图像采用传统匹配方法, 即寻找多幅图像中共同特征点进行重构的结果, 其效果并不理想是因为在连续的匹配中会导致匹配不上特征点的缺失, 造成最终特征点数目过少、细节信息丢失、重构结果不能准确看出目标结构。所以为了利用尽可能多的匹配点, 减少在连续匹配中特征点的缺失, 采用 3.2 节中先重构再融合的方法, 即首先利用连续两幅图像做特征匹配与重构, 然后将两两重构结果进行融合, 融合重构结果如图 9 所示。

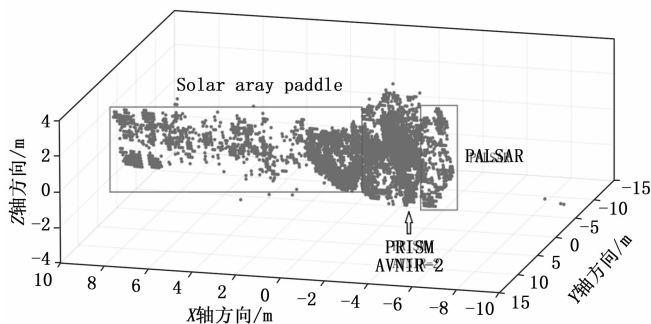


图 9 重构融合结果

结构更加具体, 重构特征点增多。在图中可以看到卫星模块三太阳能板部分结构已经明显, 而图 8 (b) 中太阳能板部分只有少数特征点, 中间部分稀疏。利用先匹配再融合重构的方法, 也使得目标的边缘信息更清晰, 连续匹配重构中缺失的特征点可能会导致检测不到目标边缘。通过重构结果与模型对比可得, 这些散射点的三维信息能够较为准确地表达目标的形状与结构。

5 结束语

本文针对 ISAR 图像序列, 利用卷积神经网络对图像做特征提取与匹配, 然后利用正交投影下矩阵因式分解的方法重构出目标三维空间位置信息, 并利用连续重构融合的方法提升重构效果。选取 ALOS 卫星 ISAR 图像序列, 通过基于 LoFTR 的特征匹配算法得到多幅图像强散射中心匹配坐标, 构建出观测矩阵, 再由三维目标与二维图像之间为正交投影关系利用正交因式分解原理恢复出目标的三维形状。通过对卫星图像的三维重构, 从结果中可见该方法对目标三维结构恢复具有较好的结果, 可以重构出目标强散射点的三维坐标及卫星形状。

参考文献:

- [1] 吴敏, 张磊, 刘松杨, 等. OFDM-ISAR 的稀疏优化成像与运动补偿 [J]. 雷达学报, 2016, 5 (1): 72-81.
- [2] BINGREN J I, YONG W, BIN Z, et al. Multi-static InSAR imaging for ships under sparse aperture [J]. 系统工程与电子技术: 英文版, 2022, 33 (3): 575-584.
- [3] RONG J, WANG Y, LU X, et al. InSAR imaging for maneuvering target based on the quadratic frequency modulated signal model with time-varying amplitude [J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 1-17.
- [4] 杨子聪, 李小勇, 白雪茹. 基于即插即用 2D-FISTA 的高分辨 ISAR 成像方法 [J]. 空军工程大学学报, 2023, 24 (1): 38-44.
- [5] DUSMANU M, ROCCO I, PAJDLA T, et al. D2-Net: A trainable CNN for joint detection and description of local features [J]. Arxiv, 2019, 1905.03561.
- [6] 韦春桃, 李渊哲. 多时相遥感影像的深度卷积匹配算法研究

与连续匹配重构结果对比, 融合重构结果更好, 目标

- [J]. 计算机应用研究, 2023, 40 (3): 932-937.
- [7] 左潇丽, 朱岱寅, 李俊强. 基于特征匹配的空间目标 ISAR 图像横向定标 [J]. 电子设计工程, 2017, 25 (18): 74-78.
- [8] 徐梦莹, 刘文波, 郑祥爱, 等. 基于深度学习的图像特征匹配方法 [J]. 传感器与微系统, 2022, 41 (7): 61-64.
- [9] KANADE T C. Shape and motion from image streams: a factorization method [J]. Proceedings of the National Academy of Sciences, 1993, 90 (21): 9795-9802.
- [10] 汤立波, 李道京, 吴一戎. 单天线 SAR 运动舰船目标三维形状重构 [J]. 遥感学报, 2007 (4): 468-472.
- [11] 王 峰, 徐 丰, 金亚秋. 利用序列 ISAR 图像获取空间目标 3-D 信息的方法 [J]. 遥感技术与应用, 2016, 31 (5): 900-906.
- [12] 尚士泽, 李 明, 侯颖妮, 等. 基于 ISAR 像序列的目标三维重构方法研究 [C] //第四届高分辨率对地观测学术年会论文集, 2017: 997-1009.
- [13] GUOHUI D, FULIN S, HONGXIN Y, et al. ISAR image scattering center association based on speeded-up robust features [J]. Multimedia Tools & Applications, 2020, 79: 5065-5082.
- [14] SUN J, SHEN Z, WANG Y, et al. LoFTR: Detector-free local feature matching with transformers [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 8922-8931.
- [15] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. Arxiv, 2014: 1409-1556.
- [16] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [17] LIN T Y, DOLLÁR P, GIRSHICK R, et al. Feature pyramid networks for object detection [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 2117-2125.
- [18] 毕严先, 王 俊. 基于 ISAR 像序列的目标三维重构 [J]. 太赫兹科学与电子信息学报, 2016, 14 (4): 531-534.
- [19] 许志伟, 张 磊, 邢孟道. 基于特征配准的 ISAR 图像方位定标方法 [J]. 电子与信息学报, 2014, 36 (9): 2173-2179.
- [20] WANG F, EIBERT T F, JIN Y. Simulation of ISAR imaging for a space target and reconstruction under sparse sampling via compressed sensing [J]. IEEE Transactions on Geoscience & Remote Sensing, 2015, 53 (6): 3432-3441.
- [21] 王 昕, 郭宝锋, 尚朝轩. 基于二维 ISAR 图像序列的雷达目标三维重建方法 [J]. 电子与信息学报, 2013, 35 (10): 2475-2480.
- [22] YAMAGUCHI T, MIZUTANI T, NAGAYAMA T. Mapping subsurface utility pipes by 3-D convolutional neural network and Kirchhoff migration using GPR images [J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59 (8): 6525-6536.
- [23] LIU L, YU H, XU H, et al. Underground object classification using deep 3-D convolutional networks and multiple mirror encoding for GPR data [J]. IEEE Geoscience and Remote Sensing Letters, 2022, 19: 1-5.
- [24] KIM N, KIM S, AN Y K, et al. A novel 3D GPR image arrangement for deep learning-based underground object classification [J]. International Journal of Pavement Engineering, 2021, 22 (6): 740-751.
- [25] KANG M S, KIM N, LEE J J, et al. Deep learning-based automated underground cavity detection using three-dimensional ground penetrating radar [J]. Structural Health Monitoring, 2020, 19 (1): 173-185.
- [26] CHRIST P F, ETTlinger F, GRÜN F, et al. Automatic liver and tumor segmentation of CT and MRI volumes using cascaded fully convolutional neural networks [J]. Arxiv Preprint Arxiv: 1702.05970, 2017.
- [27] HU K, GAN Q, ZHANG Y, et al. Brain tumor segmentation using multi-cascaded convolutional neural networks and conditional random field [J]. IEEE Access, 2019, 7: 92615-92629.
- [28] HUA X, WANG X, RUI T, et al. A fast self-attention cascaded network for object detection in large scene remote sensing images [J]. Applied Soft Computing, 2020, 94: 106495.
- [29] ZHANG J, XIE Z, SUN J, et al. A cascaded R-CNN with multiscale attention and imbalanced samples for traffic sign detection [J]. IEEE Access, 2020, 8: 29742-29754.
- [30] ZENG D, ZHAO F, GE S, et al. Fast cascade face detection with pyramid network [J]. Pattern Recognition Letters, 2019, 119: 180-186.
- [31] SZEGEDY C, LIU W, JIA Y, et al. Going deeper with convolutions [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 1-9.
- [32] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C] //Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770-778.
- [33] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722.
- [34] DENG J, DONG W, SOCHER R, et al. Imagenet: A large-scale hierarchical image database [C] //2009 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2009: 248-255.
- [35] YOSINSKI J, CLUNE J, BENGIO Y, et al. How transferable are features in deep neural networks? [J]. Advances in Neural Information Processing Systems, 2014, 27: 3320-3328.

(上接第 256 页)