

基于跨尺度相似先验的遥感图像时空融合算法

方帅^{1,2,3}, 万旗¹, 曹洋⁴

(1. 合肥工业大学计算机与信息学院, 安徽合肥 230000; 2. 工业安全与应急技术安徽省重点实验室, 安徽合肥 230000;
3. 大数据知识工程教育部重点实验室(合肥工业大学), 安徽合肥 230000; 4. 中国科学技术大学, 安徽合肥 230000)

摘要: 遥感卫星图像在空间分辨率和时间分辨率之间权衡导致图像序列的时空矛盾. 时空图像融合提供了一个生成高空间分辨率和高时间分辨率图像的解决方案, 以满足各种地球观测应用. 基于稀疏表示的时空融合算法通过联合训练字典和稀疏编码表示建立高低空间分辨率图像之间的关系, 为物候变化、类型变化等各种情况提供了统一的融合框架. 然而, 多源遥感图像来自于不同的传感器, 高低空间分辨率图像之间关系模型暗含有传感器映射关系, 导致模型设备依赖. 针对该问题, 本文提出将多源遥感图像时空融合过程分解为传感器偏差校正和时空融合两个子问题, 即设备依赖部分和设备无关部分. 传感器偏差校正部分可以作为时空融合预处理模块, 提高融合精度, 并且使得后续的融合模型更加具有普适性. 当高低空间分辨率图像空间分辨率差异较大时, “高低空间分辨率图像稀疏系数一致”的假设带来的融合误差非常突出. 针对该问题, 本文提出基于跨尺度相似先验的遥感图像时空融合算法, 利用跨尺度相似块构建稀疏结构先验的正则项, 优化稀疏表示的目标函数, 并构建中间尺度图像, 降低跨尺度相似块的二义性. 本文分别使用3组典型场景的实验数据集与其他算法进行对比, 实验结果表明, 在BOREAS数据集上, 与次优的指标相比, 本文算法的结构相似度 (Structural SIMilarity, SSIM) 提高了4.2%, 光谱角 (Spectral Angle Mapper, SAM) 提高了4.6%; 在CIA数据集上, 与次优的指标相比, 本文算法的SSIM提高了2.7%, SAM提高了12.8%; 在LGC数据集上, 与次优的指标相比, 本文算法的SSIM提高了7.1%, SAM提高了16.3%; 证明本文算法在空间和光谱特性上表现出优秀的特性.

关键词: 遥感; 时空融合; 稀疏表示; 跨尺度相似

基金项目: 国家自然科学基金 (No.61872327, No.61175033)

中图分类号: TP751.1

文献标识码: A

文章编号: 0372-2112(2024)06-2037-16

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20221147

A Spatiotemporal Fusion Algorithm of Remote Sensing Images Based on Cross-Scale Similarity Prior

FANG Shuai^{1,2,3}, WAN Qi¹, CAO Yang⁴

(1. School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230000, China;

2. Anhui Province Key Laboratory of Industry Safety and Emergency Technology, Hefei, Anhui 230000, China;

3. Key Laboratory of Knowledge Engineering with Big Data (Hefei University of Technology), Ministry of Education, Hefei, Anhui 230000, China;

4. University of Science and Technology of China, Hefei, Anhui 230000, China)

Abstract: The trade-off between spatial and temporal resolution of satellite images leads to spatial and temporal contradictions in image sequences. Spatiotemporal image fusion provides a solution to generate high spatial resolution and high temporal resolution images to satisfy various earth observation applications. The spatiotemporal fusion algorithm based on sparse representation establishes the relationship between high and low spatial resolution images by jointly training the dictionary and sparse coding representation, which provides a unified fusion framework for phenological change and type change. However, the multi-source remote sensing images come from different sensors, and the relationship model between high and low spatial resolution images implies the sensor mapping. This inevitably leads to that the model is device dependent. To solve the problem, we decompose the multi-source remote sensing spatiotemporal fusion process into two sub-problems, device dependent sensor bias correction and device independent spatiotemporal fusion. The sensor bias correction

can be used as a preprocessing module to improve the universality and accuracy of subsequent fusion models. When there are large space scale gaps between high and low spatial resolution image, the assumption that “the sparse coefficients of high and low spatial resolution images are the same” will bring about very significant fusion errors. To solve the problem, we optimize the objective function of sparse representation using cross-scale similarity prior. Intermediate-scale images are constructed to reduce ambiguity of cross-scale similar patches and improve the accuracy of cross-scale similar patches. Experimental results in three typical scenarios demonstrate the generalization ability of our algorithm. The contrastive experiments show that on the BOREAS dataset, compared to suboptimal indicators, SSIM (Structural SIMilarity) is improved by 4.2%, SAM (Spectral Angle Mapper) is increased by 4.6%; On the CIA dataset, compared to suboptimal indicators, SSIM is increased by 2.7%, and SAM is increased by 12.8%; On the LGC dataset, compared to suboptimal indicators, SSIM is increased by 7.1%, and SAM is increased by 16.3%. Our algorithm is superior to other compared methods in spatial and spectral performance.

Key words: remote sensing; spatiotemporal fusion; sparse representation; cross-scale similarity

Foundation Item(s): National Natural Science Foundation of China (No.61872327, No.61175033)

1 引言

全球地表覆盖分布及变化反映人与自然的相互作用,是生态环境监测、气候变化研究、可持续发展规划的重要基础信息^[1-3]. 在过去的几十年中,全球卫星发射数量稳步增长,捕获地球表面的多时相图像. 但是,由于硬件技术和预算的限制,多时相图像面临着空间分辨率和回访频率之间的权衡. 例如,以 Landsat 系列卫星和 Terra/AQUA 卫星为例, Landsat 卫星搭载的增强型专题制图仪 (Enhanced Thematic Mapper plus, ETM+) 可以捕捉到 15~60 m 的高空间分辨率遥感图像,但回访周期较长(16天). 考虑到云层污染,可用于对地观测的无云图像只占总数据量的 65%. 而 Terra/AQUA 卫星搭载的中分辨率成像光谱仪 (MODerate-resolution Imaging Spectroradiometer, MODIS) 的回访周期是 12 h,意味着

可以获得高时间分辨率的遥感图像序列,但是 MODIS 的空间分辨率较低,只有 250~1 000 m. 在硬件无法直接获取高空间分辨率、高时间分辨率图像的现状下,时空融合技术将来自多源遥感卫星传感器的低空间分辨率高时间分辨率图像和高空间分辨率低时间分辨率图像进行融合,生成具有高回访频率的高空间分辨率遥感图像序列. 如图 1 所示,时空融合问题是已知先验日期高低空间分辨率图像对和预测日期低空间分辨率图像,在时间维度上插值生成高空间分辨率图像,从而提高高空间分辨率图像的采样频率. 其中,图 1(a)使用 1 对先验日期高低空间分辨率图像对,图 1(b)使用 2 对先验日期高低空间分辨率图像,图 1(b)对数据的需求更加严格,融合精度会高一些,但由于传感器、云层等原因,2 对先验在实际中难以获取,所以也有很多研究仅使用 1 对先验. 本文算法采用 1 对先验.

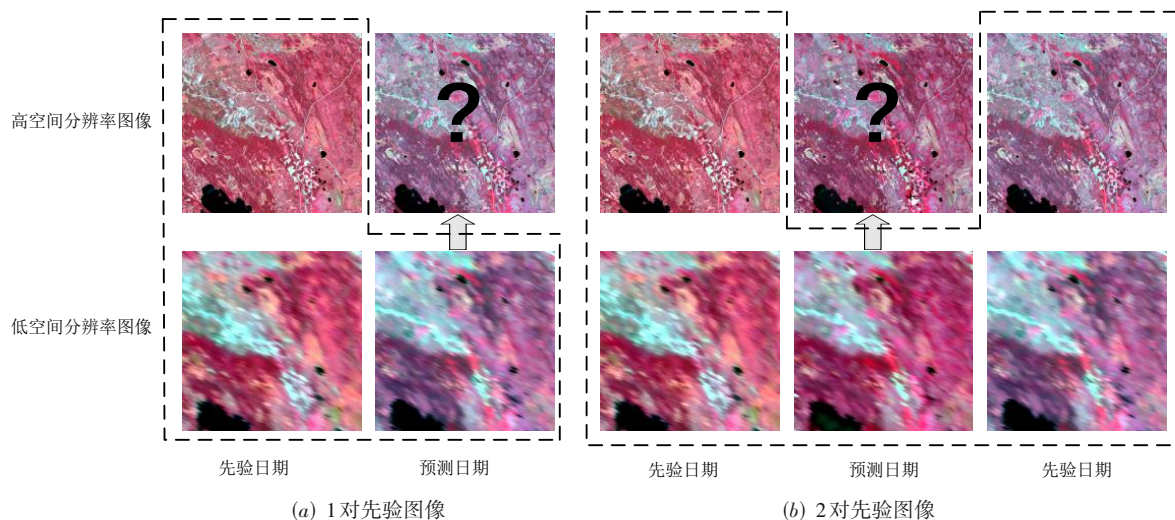


图 1 时空融合原理示意图

近年来,时空融合方法的研究经历了快速的发展,涌现出大量研究成果. 现有的时空融合方法通常可分

为两类:基于传统的方法和基于学习的方法.

基于传统的方法主要包括基于加权方法和基于解

混方法.前者通常使用加权函数和滑动窗口从已知图像中获得融合图像.典型的算法有时空自适应反射融合模型(Spatial and Temporal Adaptive Reflectance Fusion Model, STARFM)^[4]、反应反射率变化的时空自适应算法(Spatial Temporal Adaptive Algorithm for mapping Reflectance CHange, STAARCH)^[5]和增强型时空自适应反射融合模型(Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model, ESTARFM)^[6].这类方法具有较好的稳定性,但是融合精度较低.在“低空间分辨率图像是高空分辨率图像的线性混合”的假设基础上,基于解混方法通过对低空间分辨率图像进行光谱分解来预测融合图像.这类算法^[7-11]通常首先使用诸如k均值聚类算法(k-means)、迭代自组织的数据分析算法(Iterative Self Organizing Data Analysis Techniques Algorithm, ISODATA)等无监督方法对高空分辨率图像进行分类,确定低空间分辨率图像的端元.这些无监督分类算法对初始参数设置敏感,初始参数的设置会影响分类精度,进而影响光谱分解.同时,基于解混的时空融合算法要求类别端元在观测时间段内不能增加,故而限制了其应用范围.此外,上述两类算法均将图像融合模型定义为线性模型,难以适应复杂的环境预测.

基于学习的方法可分为基于稀疏表示的方法和基于深度学习的方法.与传统方法相比,基于学习的方法在准确性和鲁棒性方面有着突出的表现.基于深度学习的方法^[12-18]通过设计网络结构、学习算法和成本函数、提取高质量特征和构造有效监督,从而获得高融合性能.另外,基于深度学习的方法还需要大量的数据集进行模型训练,为了获得较高的融合精度,很多方法在训练阶段采用3对高低空间分辨率图像,在测试阶段采用2对高低空间分辨率图像^[12-14],而这样的数据需求在实际中很难被满足.为此,部分模型采用2对高低空间分辨率图像训练,1对高低空间分辨率图像测试^[15-18],精度则会有明显降低.

基于稀疏表示的方法^[19-26]在图像域或梯度域训练高低空间分辨率图像字典对,利用高低空间分辨率图像共享稀疏表示来重建融合图像.基于稀疏表示的方法基于两个假设:(1)先验日期图像训练的字典能适用于预测日期;(2)同一场景高低空间分辨率图像的稀疏系数一致.依据这两个假设,提出了基于耦合字典^[19-21]和半耦合字典^[22-24]的时空融合算法.由于对字典和稀疏系数的假设过于严格,导致融合结果并不理想.Wu等人^[22]提出了误差有界的半耦合字典学习(Error-Bound-regularized Semi-Coupled Dictionary Learning, EBSDL)模型,构建了关于字典的误差有界正则化,稀疏系数采用半耦合方式,允许字典进行微调,同时系数

可以存在线性差异.通过放松两个假设,提高了融合精度.稀疏系数中隐含结构信息,研究者们利用这一特性提出一些新的结构稀疏先验,比如Wei等人^[23]通过增加通道内的聚类结构稀疏和通道间的联合结构稀疏约束,使融合的稳定性和精度都有提高.但是当高低空间分辨率图像的分辨率差距过大时,低空间分辨率图像遭受严重退化,仅依赖低空间分辨率图像估计稀疏系数,高空分辨率图像上的很多细节无法重构.Liu等人^[21]用先验日期高空分辨率图像稀疏系数约束预测日期的稀疏系数,提高融合图像细节,但导致类型变化区域引入错误细节.

对于时空融合问题而言,从先验日期到预测日期这段时间内场景会发生一些变化,从空间上可划分为物候变化、类型变化和无变化区域,其中,类型变化区域融合难度是最大的,很多文献^[4,9-11]致力于解决类型变化区域场景的融合问题.基于稀疏表示的方法把时空融合作为超分辨率问题来解决,将这3种区域统一在一个框架下,即利用先验日期高低空间分辨率图像学习高低空间分辨率图像字典对,在预测日期利用低空间分辨率图像进行稀疏系数估计,重建融合图像.从理论上讲,该类方法本身不会受类型变化区域的“困扰”,这是它最大的优势.因此,本文采用基于稀疏表示的融合方法.然而,时空融合与常规的超分辨率有着不可忽视的特殊性.首先,时空融合的高低空间分辨率图像来自不同传感器,因此,时空融合的超分辨率模型还要包含不同传感器之间的映射关系,这不仅使得模型复杂,且模型与具体传感器密切相关.其次,当高低空间分辨率图像空间分辨率差距过大时,相对高空分辨率图像,低空间分辨率图像退化严重,高空分辨率图像上部分空间信息在低空间分辨率图像上几乎消失,因此,当前基于稀疏表示的时空融合方法无法仅利用预测日期在低空间分辨率图像精确重建高空分辨率图像.以上两点正是本文要解决的问题.

基于以上两个要解决的问题,本文主要贡献有以下两点.(1)提出将多源遥感图像时空融合问题分解为传感器偏差校正和时空融合两个子问题.一方面,将具有强设备依赖性的传感器映射关系,从多源遥感图像时空融合模型中独立出来,使得时空融合具有更好的普适性;另一方面,传感器偏差校正模型能够为现有的多源遥感图像时空融合提供一个预处理,从而提高时空融合的精度.(2)提出基于跨尺度相似先验的遥感图像时空融合算法.解决了高低空间分辨率图像的空间分辨率差异大时,依据“同一场景高低空间分辨率图像稀疏系数一致”的假设,不能实现高空分辨率图像的精确求解的问题.一方面构建了基于跨尺度相似块的稀疏结构先验的正则项,优化了基于稀疏表示时空融

合的目标函数. 另一方面, 构建中间尺度图像距离度量, 降低了跨尺度相似块的二义性, 提高了跨尺度相似块的精度.

2 问题分析

基于稀疏表示的时空融合面临着两个关键问题: 一是传感器偏差会导致模型复杂且设备依赖; 二是“高低空间分辨率图像稀疏系数一致”的假设, 随着高低空间分辨率图像尺度因子的增大, 这一假设的负面效果越明显.

2.1 传感器偏差

本节以 MODIS 图像与 Landsat 图像为例, 描述传感器偏差. 它们来自不同卫星, 有着不同的轨道参数、波段范围、光谱响应、拍摄时间、大气参数、校正模式等, 这导致不同卫星图像存在明显差异. 如图 2 所示, 图 2(a) 和图 2(b) 为相同分辨率的 MODIS 和 Landsat 图像, 图 2(c) 和图 2(d) 是两者在近红外波段的字典, 图 2(e) 和图 2(f) 是两者在近红外波段的像素值的直方图. 从视觉感知, 直方图分布和字典表示等方面均呈现出明显的不同.

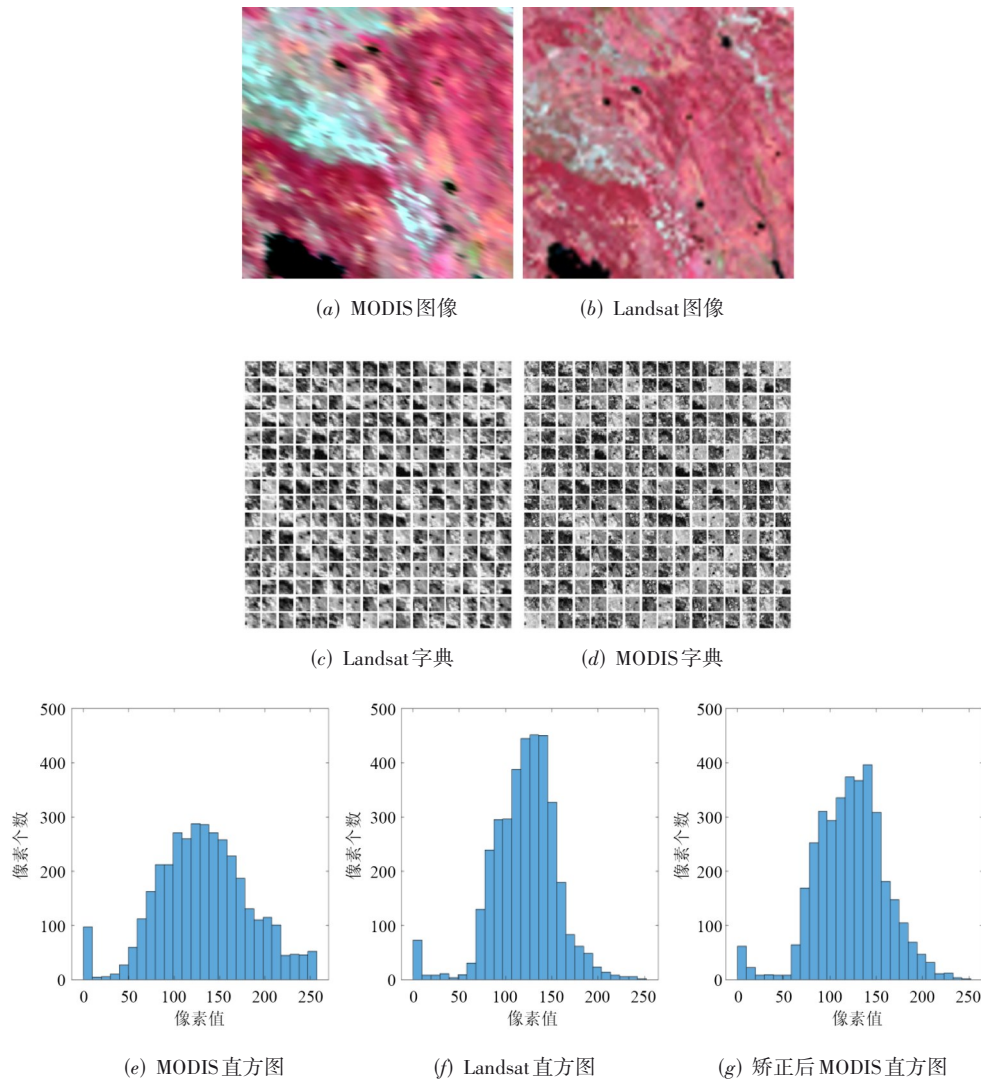


图 2 同一场景相同分辨率 MODIS 和 Landsat 图像对分析

传感器偏差对时空融合会有两方面的负面影响. 第一, 遥感时空融合的原图像是 MODIS 图像序列, 目标图像是 Landsat 图像, 在像素分布和特征分布方面, 源图像和目标图像相差越大越不利于目标图像的生成. 第二, 基于稀疏表示的时空融合方法将时空融合问题当作图像超分辨率问题来解决, 传感器偏差的存在, 会导

致基于稀疏表示的时空融合模型不仅要表示高低空间分辨率图像之间的关系, 还要暗含 MODIS 到 Landsat 的传感器映射关系, 这使得模型复杂, 而且设备依赖.

因此, 也有一部分研究将 Landsat 图像进行下采样, 代替 MODIS 图像数据源进行时空融合研究^[11], 从而排除传感器偏差所带来的负面影响, 提高融合精度. 但是

传感器偏差是实际存在的,在实际应用中无法回避.传统方法^[10,27]和深度学习方法^[14]都有相关文献解决传感器偏差的问题.但在目前基于稀疏表示的时空融合算法中却忽略了这一点.另外,传感器偏差的存在还导致“高低空间分辨率图像稀疏系数一致”这个假设更加不满足.

2.2 稀疏系数一致性假设

基于稀疏表示的时空融合方法中“同一场景高低空间分辨率图像的稀疏系数一致”假设过于严格,尤其当高低空间分辨率图像空间尺度差异大时问题更为突出.在目前常用的耦合字典学习方法^[19-21]中,仅从低空间分辨率图像中提取的稀疏系数无法有效重构高空间分辨率图像上的细节信息.为了方便,以下先验日期用 t_1 表示,预测日期用 t_2 表示.

低空间分辨率图像遭受严重退化,很多高空间分辨率图像上的细节在低空间分辨率图像上完全消失了.若采用系数一致性假设,仅从低空间分辨率图像中提取系数,无法重构那些在高空间分辨率图像上消失的细节信息.如图3所示,图3(a)和图3(b)是 t_1 时刻的高低空间分辨率图像对,高低空间分辨率图像之间有着 16 倍分辨率差异,依此学习得到高低空间分辨率图像字典对.从 t_2 时刻低空间分辨率图 3(c) 中提取的稀疏系数,与高空间分辨率图像字典相乘得到预测结果图 3(e),与真实图像图 3(d) 相比,可以看出图 3(e) 严重缺少空间细节.利用 t_2 时刻高空间分辨率图 3(d) 中提取的稀疏系数,与高空间分辨率图像字典相乘得到预测结果图 3(f),与真实图像图 3(d) 非常接近.图 3(e) 和图 3(f) 的生成使用相同的字典,它们之间的差异主要来自稀疏系数的不同,可以看出低空间分辨率图像中提取的稀疏系数并不能精确地表示高空间分辨率图像,即“同一场景高低空间分辨率图像的稀疏系数一致”只是便于问题的求解,不能实现问题的精确求解.

半耦合字典学习将高低空间分辨率图像的稀疏系数之间的关系进行线性建模,放宽了“高低空间分辨率图像稀疏系数一致”这一假设.半耦合字典学习方法的引入使得融合精度有一定的提升,但仍然无法有效重构高空间分辨率图像上的细节信息.针对半耦合字典学习方法, Yang 等人^[25]进行了显著性水平为 5% 的 F 检验,检验结果表明高低空间分辨率图像的稀疏系数之间的线性关系不显著.

Liu 等人^[21]用 t_1 时刻高空间分辨率图像稀疏系数约束 t_2 时刻的稀疏系数,这一举措给重构高空间分辨率图像上的细节信息带来了可能,但是若观测时间段内发生类型变化,变化区域则会缺失有效的先验指导,最终导致将错误细节引入预测结果.本文方法搜索与 t_2 时刻低空间分辨率图像块相似的 t_1 时刻低空间分辨率图像块,将这些相似块按坐标投影到 t_1 时刻高空间分辨

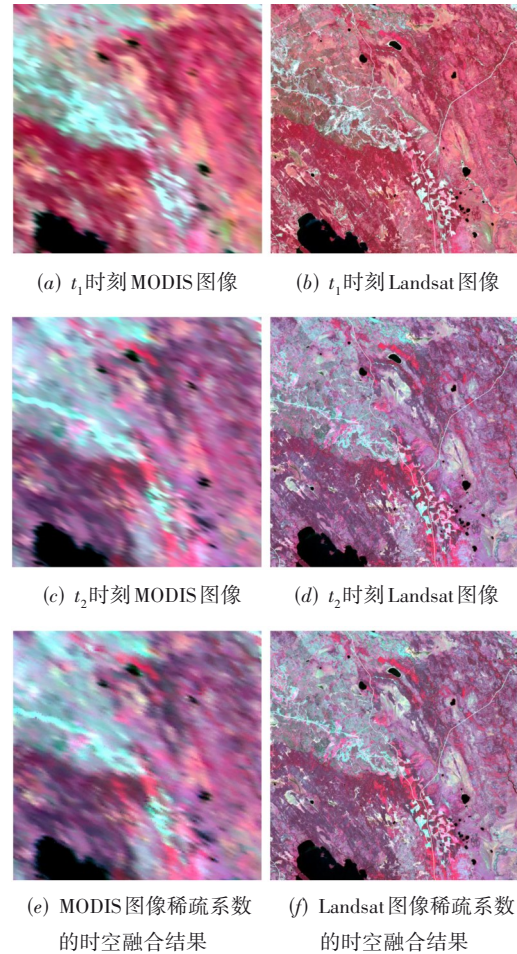


图3 不同稀疏系数的时空融合结果分析

率图像上,从而获得跨尺度的相似块.然而,由于高低空间分辨率图像的空间分辨率差距巨大,低空间分辨率图像分辨率下的相似块,其对应的高空间分辨率图像块可能并不相似,即跨尺度相似块具有二义性,具体操作时面临着跨尺度的高空间分辨率图像相似块难以确定的问题.

2.3 解决思路

针对以上两个问题,本文工作提出相应的解决方案.

(1)针对传感器偏差问题,本文提出建立偏差映射模型,解决思路是利用同一场景采集的 MODIS 和 Landsat 图像对,学习耦合字典,通过稀疏表示将 MODIS 图像转换为相同分辨率下 Landsat 风格图像.

(2)针对低空间分辨率图像退化过于严重,仅依据低空间分辨率图像中提取的稀疏系数,无法有效重构高空间分辨率图像的问题,本文提出基于跨尺度相似先验的稀疏系数求取算法.解决思路是对于 t_2 时刻低空间分辨率图像块,在 t_1 时刻的高空间分辨率图像上找多个跨尺度相似块;为了降低跨尺度相似块的二义性,

利用中间尺度图像间的距离度量,构建目标函数,对多个相似块以加权的方式进行挑选,并合成一个高质量的跨尺度相似块;使用合成的跨尺度相似块稀疏系数的结构(非零位置)来约束 t_2 时刻低空间分辨率图像块的稀疏系数,以获得更高质量的高空间分辨率图像预测结果.

3 时空融合算法

本文的时空融合算法主要解决传感器偏差问题,以及当高低空间分辨率图像空间分辨率差异过大时,“高低空间分辨率图像稀疏系数一致”的假设导致融合精度低的问题.因此,本文选择 Landsat 图像和 MODIS 图像,它们的空间尺度相差 16 倍.

3.1 传感器偏差映射

从图 2 可以看出,相同分辨率的 MODIS 图像和 Landsat 图像从视觉感知、直方图分布和字典表示等方面均呈现出明显的不同.因此,在进行 t_2 时刻图像重建之前有必要消除图像间的传感器偏差.

如图 4 所示,本文采用耦合字典学习的方法建立 Landsat 下采样图像和 MODIS 图像的关系,从 MODIS 图像中提取系数,然后与 Landsat 下采样图像的字典相乘,最后获得经过偏差校正的 MODIS 图像.

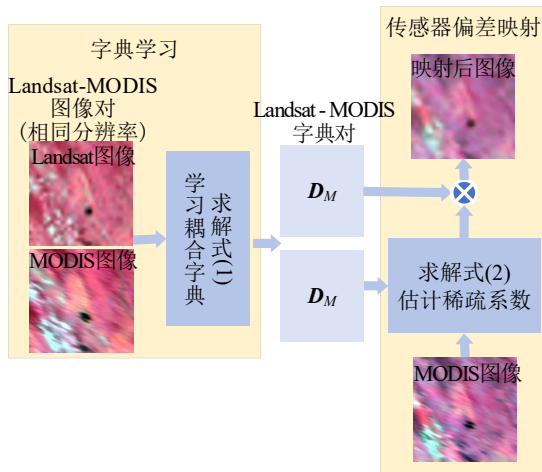


图 4 传感器偏差映射模块示意图

首先,根据式(1)进行耦合字典学习:

$$\arg \min_{\{D_L, D_M\}} \left\{ \left\| \begin{bmatrix} M_L \\ M \end{bmatrix} - \begin{bmatrix} D_L \\ D_M \end{bmatrix} \alpha \right\|_2^2 + \lambda \|\alpha\|_1 \right\} \quad (1)$$

其中, M_L 是下采样的 Landsat 图像; M 是 MODIS 图像; D_L 和 D_M 分别表示 M_L 和 M 的过完备字典; α 为 M_L 和 M 的共享稀疏系数. 求解式(1)获得 M_L - M 图像的字典对 D_L 和 D_M .

其次,对任意输入的 MODIS 图像 M , 通过式(2)和式(3)实现传感器偏差校正:

$$\arg \min_{\{\alpha_M\}} \left\{ \|M - D_M \alpha_M\|_2^2 + \lambda \|\alpha_M\|_1 \right\} \quad (2)$$

$$M_R = D_L \alpha_M \quad (3)$$

其中, α_M 为 M 的稀疏系数; M_R 表示经过偏差映射后的 MODIS 图像.

图 2(e)~(g) 的 3 个直方图分别为 M 、 M_L 和 M_R 的像素直方图,从 3 幅直方图可以看出, M_R 的直方图整体趋势比 M 的更接近 M_L 直方图. 同时,以 M_L 图像作为先验图像,对 M_R 进行了客观评价,如表 1 所示,与 M 相比, M_R 从结构、光谱、相关性等方面都更接近 M_L 图像. 偏差映射使得 MODIS 图像转换为更加逼近下采样的 Landsat 图像,降低了传感器差异导致融合模型的复杂性,后续融合模型更加专注于空间细节的表示.

表 1 偏差映射前后 MODIS 图像的客观评价 单位: %

图像	CC	SAM	SSIM	RMSE
M	79.840	637.520	58.240	3.320
M_R	93.880	387.150	89.940	1.490

3.2 基于加权跨尺度相似块先验的稀疏表示

为了方便,本节 MODIS 图像是指经过第 3.1 节算法校正后的 MODIS 图像,也称为低空间分辨率图像,标记为 M^R .

当高低空间分辨率图像之间空间分辨率差距过大时,低空间分辨率图像退化严重,很多高空间分辨率图像上的细节在低空间分辨率图像上完全消失了,这导致仅依赖低空间分辨率图像估计稀疏系数,在目前常用的耦合、半耦合等先验下,无法有效重构高空间分辨率图像上的细节信息.

本文提出利用跨尺度相似块,提供更加有效的稀疏系数的估计. 第一个关键问题是如何利用跨尺度相似块提高稀疏系数的求解精度. 由于跨尺度相似块有二义性,因此,第二个关键问题是如何获取高质量的跨尺度相似块.

首先,解决如何利用跨尺度相似块的问题,这里假设已经得到跨尺度相似块,将其用于约束稀疏系数的求解.

在已知高低空间分辨率图像字典对 (D_{High} 和 D_{Low}) 的情况下,构建基于跨尺度相似块结构约束稀疏系数目标函数:

$$\arg \min_{\{\alpha_{Low}^i\}} \left\{ \left\| (M_2^R)^i - D_{Low} \alpha_{Low}^i \right\|_2^2 + \lambda_1 \|\alpha_{Low}^i\|_1 + \gamma \left\| \left(1 - \text{sign}(|\alpha_{Low}^i|) \right) \odot |\beta_i| \right\|_2^2 \right\} \quad (4)$$

$$\text{sign}(|x|) = \begin{cases} 1, & |x| \neq 0 \\ 0, & |x| = 0 \end{cases} \quad (5)$$

其中, $(M_2^R)^i$ 表示 t_2 时刻的低空间分辨率图像的第 i 个

图像块; α_{Low}^i 为 $(\mathbf{M}_2^{\text{R}})^i$ 对应的稀疏系数; x 表示向量 α_{Low}^i 中的某项值; $|\cdot|$ 表示绝对值; \odot 表示点乘. 式(4)中前两项分别是常规的保真项和系数稀疏正则项, 第三项是利用 t_1 时刻的跨尺度相似块构建的对稀疏系数的约束项.

$(\mathbf{M}_2^{\text{R}})^i$ 跨尺度相似块的稀疏系数为 β_i , 用 β_i 的结构约束要求解的稀疏系数 α_{Low}^i . 这里期望求解的 α_{Low}^i 在满足前两项的同时, 也要与跨尺度相似块的稀疏系数 β_i 的结构相似.

式(4)中的第三项表示, 当 α_{Low}^i 与 β_i 的非零位置相同时, 此项取得最小值 0; 当 α_{Low}^i 与 β_i 的非零位置完全相反时, 此项取得最大值, 最大值为 $|\beta_i|$ 的各项平方和. 因此, 该项能够约束 α_{Low}^i 与 β_i 具有一致的稀疏结构.

其次, 解决如何确定跨尺度相似块的问题, 由于高低空间分辨率图像空间分辨率差距巨大, 导致低空间分辨率图像块可能只有细微差异, 而对应的高空间分辨率图像块可能完全不同. 因此, 从多个跨尺度相似块, 通过权重方式对图像块进行筛选, 最终合成一个高质量的跨尺度相似块.

对于 t_2 时刻低空间分辨率图像中任意图像块 $(\mathbf{M}_2^{\text{R}})^i$, 在 t_1 时刻低空间分辨率图像中搜索 N 个相似块 $(\mathbf{M}_1^{\text{R}})^{i,q}$ ($q=1, 2, \dots, N$), 然后将其按坐标投影到 t_1 时刻高空间分辨率图像上, 从而获得一组跨尺度相似块 $L_1^{i,q}$ ($q=1, 2, \dots, N$), 利用 N 个 $L_1^{i,q}$ 加权构建 $(\mathbf{M}_2^{\text{R}})^i$ 的跨尺度相似块 $(\mathbf{M}_2^{\text{L}})^i$. 为了消除跨尺度相似块的二义性, 得到更加精确的权重, 构建如下目标函数:

$$\arg \min_{\{\omega_i\}} \left\| (\mathbf{M}_2^{\text{R}})^i - \sum_{q \in \Omega_i} \omega_{i,q} (\mathbf{M}_1^{\text{R}})^{i,q} \right\|_2^2 + \tau \left\| \text{dist} \odot \omega_i \right\|_2^2 + \lambda_2 \|\omega_i\|_1 \quad (6)$$

其中, \odot 表示向量的点乘; Ω_i 表示相似块集合 $(\mathbf{M}_1^{\text{R}})^{i,q}$ 的上标索引 q 的集合; $\omega_{i,q}$ 即为 $(\mathbf{M}_1^{\text{R}})^{i,q}$ 所对应的权重. 求得权重后, 构建 $(\mathbf{M}_2^{\text{R}})^i$ 的跨尺度相似块 $(\mathbf{M}_2^{\text{L}})^i$:

$$(\mathbf{M}_2^{\text{L}})^i = (\mathbf{M}_2^{\text{R}})^i + \sum_{q \in \Omega} \left(L_1^{i,q} - (\mathbf{M}_1^{\text{R}})^{i,q} \right) \quad (7)$$

式(6)的第一项是保真项, 保证相似块加权能表示 $(\mathbf{M}_2^{\text{R}})^i$, 第三项为稀疏正则项, ω_i 的稀疏性能剔除伪相似块. 由于跨尺度相似块具有二义性, 第一项和第三项只能实现在低分辨率 ($(\mathbf{M}_1^{\text{R}})^{i,q}$) 下挑选相似块, 不能保证挑选准确的高分辨率相似块 ($L_1^{i,q}$). 式(6)第二项的作用是定义一种距离, 以提高跨尺度相似块的可靠程度.

本文的实验数据高低空间分辨率图像空间分辨率是 16 倍的关系. 如果在低分辨率上计算距离, 则与高分辨率下的距离不具有 consistency; 如果在高分辨率下定义距离, 则没有可靠的数据用来计算距离. 因此, 本文构建了在中等分辨率, 即 4 倍分辨率下计算距离, 兼顾了分辨率和数据可靠性.

计算距离的做法如下. (1) 利用基于耦合字典的时空融合方法^[19]求解出提升 4 倍分辨率的 \mathbf{M}_2^{R} , 标记为 \mathbf{M}_2^{M} ; 将 L_1 下采样 4 倍的图像标记为 \mathbf{M}_1^{M} . (2) 图像块 $(\mathbf{M}_2^{\text{R}})^i$, 以及其 N 个相似块 $(\mathbf{M}_1^{\text{R}})^{i,q}$ ($q=1, 2, \dots, N$), 按坐标分别投影到 \mathbf{M}_2^{M} 和 \mathbf{M}_1^{M} , 标记为 $(\mathbf{M}_2^{\text{M}})^i$ 和 $(\mathbf{M}_1^{\text{M}})^{i,q}$ ($q=1, 2, \dots, N$). (3) 计算距离公式如下:

$$\text{dist}_q = \left\| (\mathbf{M}_2^{\text{M}})^i - (\mathbf{M}_1^{\text{M}})^{i,q} \right\|_2^2 \quad (8)$$

其中, dist_q 表示中等分辨率 (4 倍) 下图像块的相似程度, 相对于低分辨率下的比较, 减少了跨尺度相似块的二义性; 相对于高分辨率 (16 倍) 下的比较, 中等分辨率下的 \mathbf{M}_2^{M} 更可靠.

利用 $(\mathbf{M}_2^{\text{M}})^i$ 和 $(\mathbf{M}_1^{\text{M}})^{i,q}$ 的相似性, 对权重的求解进行约束, 当 $(\mathbf{M}_2^{\text{M}})^i$ 和 $(\mathbf{M}_1^{\text{M}})^{i,q}$ 不相似时, 随着 dist_q 值增大, $\omega_{i,q}$ 会减小, 在稀疏性的约束下, 会将较大 dist_q 对应的 $\omega_{i,q}$ 置 0, 将 $(\mathbf{M}_1^{\text{M}})^{i,q}$ 剔除.

求得 $\omega_{i,q}$ 后, 带入式(7), 可求得跨尺度相似块 $(\mathbf{M}_2^{\text{L}})^i$, 构建跨尺度相似块的步骤如算法 1 所示.

在已知跨尺度相似块 $(\mathbf{M}_2^{\text{L}})^i$ 和高空间分辨率图像字典 \mathbf{D}_{High} 的情况下, 求解稀疏表示的基本目标函数 (参考式(2)) 即可获得 $(\mathbf{M}_2^{\text{L}})^i$ 的稀疏系数 β_i .

算法 1 基于加权的跨尺度相似块构建

输入: t_2 时刻低空间分辨率图像块 $(\mathbf{M}_2^{\text{R}})^i$, t_1 时刻低空间分辨率图像上的相似块 $\{(\mathbf{M}_1^{\text{R}})^{i,q}\}_{q \in \Omega_i}$, t_1 时刻下采样 4 倍的高空间分辨率图像块 $\{(\mathbf{M}_1^{\text{M}})^{i,q}\}_{q \in \Omega_i}$, $(\mathbf{M}_2^{\text{R}})^i$ 4 倍分辨率差异下的预测结果 $(\mathbf{M}_2^{\text{M}})^i$, t_1 时刻高空间分辨率图像块 $\{L_1^{i,q}\}_{q \in \Omega_i}$.

根据式(8), 计算 $(\mathbf{M}_2^{\text{M}})^i$ 与 $\{(\mathbf{M}_1^{\text{M}})^{i,q}\}_{q \in \Omega_i}$ 的距离 dist_q .

根据式(6), 求解最优权重 ω_i .

输出: 根据式(7), 计算 $(\mathbf{M}_2^{\text{R}})^i$ 的跨尺度相似块 $(\mathbf{M}_2^{\text{L}})^i$.

3.3 算法总体框架

本文提出的基于加权的跨尺度相似块先验的稀疏表示算法, 提供更加有效的稀疏系数的估计. 算法流程如图 5 所示, 从功能上可分为左边的传感器偏差校正和右边的图像融合, 从算法上可以分为上面的字典学习

和下面的稀疏表示两个部分。

(1)字典学习. 第3.1节给出了相同分辨率的MODIS-Landsat字典对 D_M 和 D_L 的学习,用于传感器偏差校正;使用 t_1 时刻低空间分辨率图像 M_1^R 和高空间分辨率图像 L_1 ,通过求解式(9)进行耦合字典学习,获得时空融合部分所需的高低空间分辨率图像字典 D_{High} 和 D_{Low} .

(2)时空融合. 首先,对于 t_2 时刻低空间分辨率图像 M_2^R 上的任意图像块,在 t_1 时刻低空间分辨率图像 M_1^R 中,寻找多个相似块;其次,生成中间尺度图像 M_1^M

和 M_2^M ,并将 M_1^R 和 M_2^R 上的图像块按坐标分别投影到 M_1^M 和 M_2^M ;然后,对式(6)所表示的目标函数进行求解,得到权重 ω_i ,并根据式(7)合成一个高质量的跨尺度相似块;接着,求取合成跨尺度相似块的稀疏系数,并作为式(4)的重要约束项,提高低空间分辨率图像块稀疏系数的求解精度;最后,将低空间分辨率图像块稀疏系数与高空间分辨率图像字典 D_{High} 相乘得到最终的融合结果.

$$\arg \min_{\{D_{High}, D_{Low}\}} \left\{ \left\| \begin{bmatrix} L_1 \\ M_1^R \end{bmatrix} - \begin{bmatrix} D_{High} \\ D_{Low} \end{bmatrix} \alpha \right\|_2^2 + \lambda \|\alpha\|_1 \right\} \quad (9)$$

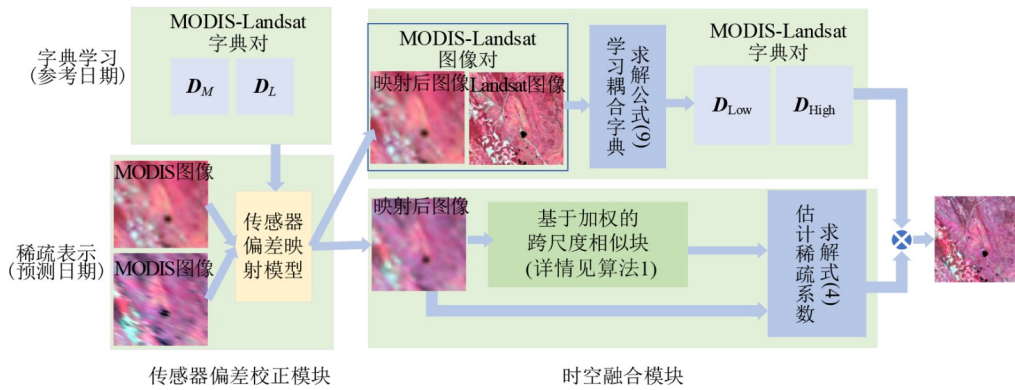


图5 融合算法整体流程图

4 实验结果与分析

4.1 实验数据集

为了验证本文算法对各种场景的通用性,在实验中选择3组Landsat-MODIS数据集和1组Sentinel-Landsat数据集场景图像,4组数据集分别代表了3种典型场景:(1)物候变化;(2)异质景观;(3)类型变化. 以下是实验数据集说明.

数据集1如图3所示,该数据集选取于2001年5月24日和2001年8月12日的美国北方生态系统大气研究(BOREAS)南部研究区(54°N,104°W),该研究区植被生长季节短,物候变化明显.

数据集2(CIA)如图6所示,该数据集是澳大利亚新南威尔士州图像,拍摄于2001年10月至2002年5月之间,本文选取2002年1月11日和2月12日的图像作为研究数据. 该区域农田规模小且不规则,涵盖了许多不同类型作物,属于典型的异质景观区域,兼具物候变化特征.

数据集3(LGC)如图7所示,该数据集覆盖的区域同样位于澳大利亚新南威尔士州,选取2004年11月26日和12月12日的图像作为研究数据. 该时间段内LGC发生了洪水,发生了剧烈的类型变化.

数据集4如图8所示,该数据集覆盖的区域位于中国江西省鄱阳湖,选取2022年7月10日和8月11日的

图像作为研究数据. 该时间段内鄱阳湖水水位退去,发生了类型变化.

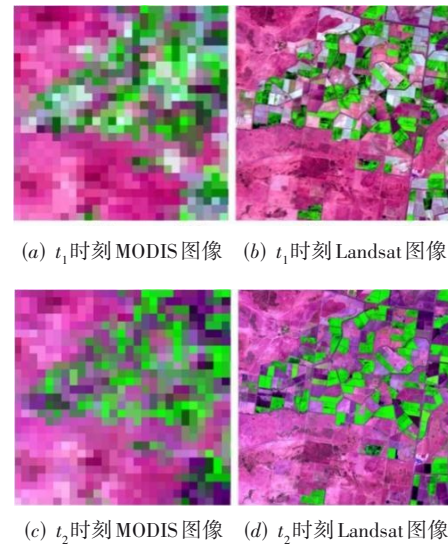


图6 CIA地区Landsat-MODIS图像对

4.2 对比算法

本文选取了5种基于一对图像的时空融合算法进行对比实验,分别是基于加权的方法时空自适应反射

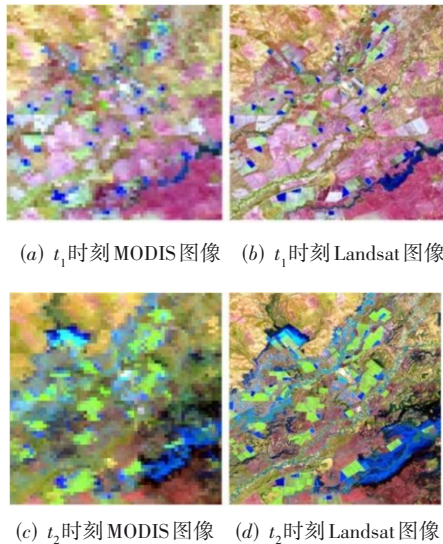


图7 LGC地区 Landsat-MODIS 图像对

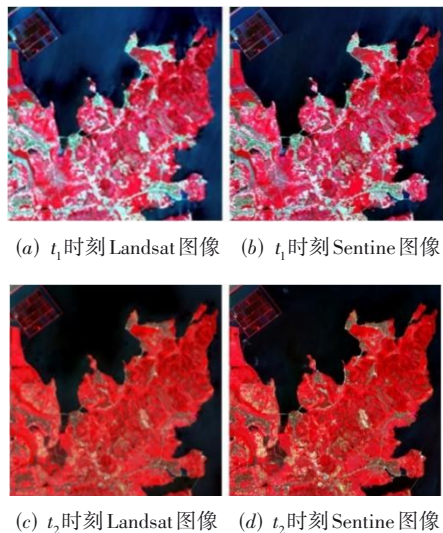


图8 鄱阳湖地区 Sentinel-Landsat 图像对

融合模型(STARFM)^[4](Gao等,2006),基于解混的方法灵活的时空融合算法(Flexible Spatiotemporal Data Fusion, FSDAF)^[11](Zhu等,2016),基于稀疏表示的方法通过一对图像学习的时空卫星图像融合算法(spatiotemporal satellite image fusion through One-pair image learning, Onepair)^[20](Huang等,2013),快速迭代的软阈值算法(Fast Iterative Shrinkage-Thresholding Algorithm, FISTA)^[21](Liu等,2016),以及基于深度学习的方法基于GAN的时空融合模型(GAN-based SpatioTemporal Fusion Model, GANSTFM)^[28](Tan等,2021).其中,STARFM算法是优秀的基于加权的方法,构建了高低空间分辨率图像中心像素的反射率的线性模型,从时间、位置和光谱差3个角度计算权重来估计 t_2 时刻高空间分辨率图像的中心像素,能够较好地预测物候变化;

FSDAF基于解混和加权的思想,在解混精度得到保证的前提下,能够较好地预测类型变化区域;Onepair和FISTA是优秀的基于稀疏表示的时空融合算法,Onepair提出了两层时空融合框架,减少了因高低空间分辨率图像空间分辨率的巨大差异带来的重建压力,提高了融合质量,能够较好地处理物候变化和类型变化.FISTA在耦合字典学习的基础上使用 t_1 时刻高空间分辨率图像的稀疏系数约束 t_2 时刻的稀疏系数.GANSTFM是一个基于条件生成对抗网络(CGAN)的时空融合算法,同时将可切换归一化技术引入时空融合问题中.GANSTFM可以减少模型输入的数量,即GANSTFM只需要 t_2 时刻的低空间分辨率图像和相同场景下任意时间的另一个高空间分辨率图像用于模型输入,且打破 t_1 时刻参考图像选择的时间限制,可以在少量样本的情况下,通过这些样本随机组合,达到训练所需的数据规模,适用于样本较少的情况.

4.3 实验参数设置

本文实验中,高低空间分辨率图像字典原子数为512,稀疏表示重建的效果随着字典原子数的增加会有一些的提升,但是也会增加计算的负担.在搜索跨尺度相似块时,以欧式距离和SSIM为依据,选前30个进行权重的求解.同时为了确定本文算法中的关键参数的最优数值范围,如式(4)中的正则化参数 λ_1 、 γ ,式(6)中的 λ_2 、 τ ,以及图像块大小(进行字典学习和在搜索相似块时采取分块策略,两者在实验中采用同样大小的窗口尺寸),本文用控制变量法进行实验.图9展示了在CIA和LGC数据集上相关系数(CC)随各参数变化曲线.从图9(a)中可以看出,窗口大小为5时,本文算法在两个数据集上有更好的融合效果.从图9(b)可以看出,当 τ 为0.01时,本文算法在CIA数据集上有更好的融合效果; τ 为0.001时,本文算法在LGC数据集上有更好的融合效果.从图9(c)和图9(d)可以看出,当 λ_2 为0.1~10时,当 γ 为0.1~10时,融合图像的质量最好.从图9(e)可以看出,当 λ_1 为0.01~1时,本文算法在两个数据集上的融合效果均较好.且图9中橙色和蓝色曲线的走势具有一定的相似性,说明本文算法在不同的数据集上具有稳定的性能.

4.4 实验结果分析

实验1 为验证本文传感器偏差校正模块在不同传感器数据集上的有效性,针对Landsat卫星、TERRA/AQUA卫星和Sentinel卫星传感器进行了实验;利用Landsat-MODIS数据集和Sentinel-Landsat数据集,验证本文算法对校正Landsat卫星传感器和TERRA/AQUA卫星传感器(MODIS)差异的有效性,以及校正Sentinel卫星传感器和Landsat卫星传感器差异的有效性.

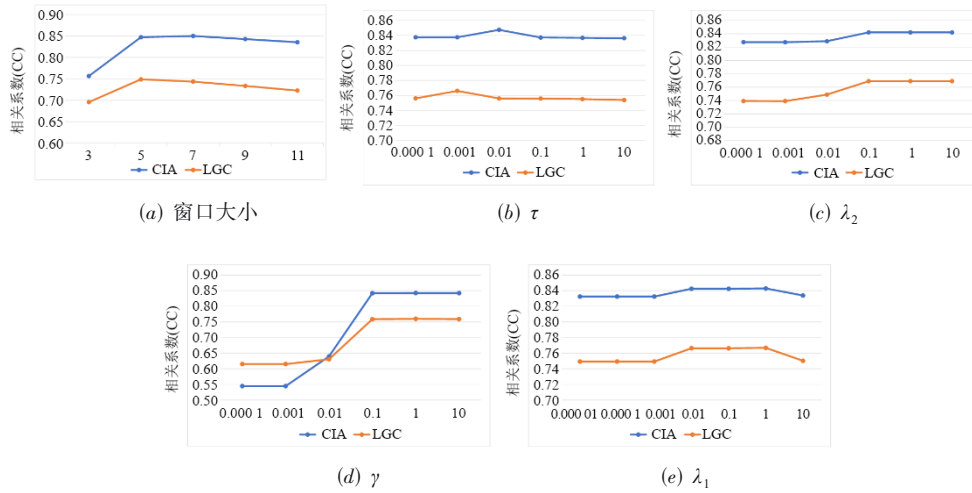


图9 随参数变化的CC曲线

图 10 为将原始 MODIS 图像和校正后的 MODIS 图像作为低空间分辨率图像进行时空融合的实验结果对比,对比较正前后的融合结果图 10(b)和图 10(d),以及差值图 10(c)和图 10(e),图像整体上视觉效果差别不大,但从图 10(b)和图 10(d)的放大区域可以看出,图 10(d)的图像轮廓更加清晰明了,从图 10(c)和图 10(e)的放大区域可以看出,差值图 10(c)轮廓边缘锐利清晰,表明未校正传感器偏差融合结果该部分未准确预测,与理想图像差异较大。

图 11 为将原始 Landsat 图像和校正后的 Landsat 图像作为低空间分辨率图像进行时空融合的实验结果进行对比. 对比较正前后的融合结果图 11(b)和图 11(d),图像整体上视觉效果差别不大. 但从差值图 11(c)和图 11(e)的放大区域可以看出,图 11(e)的残差图像剩

余更少,即图 11(d)与图 11(a)更为接近. 通过对比可以看出,图 11(d)的融合结果更优. 定量评价如表 2 所示,可以看出 Landsat-MODIS 数据集融合结果图 10(d)在指标上高于融合结果图 10(b), Sentinel-Landsat 数据集融合结果图 11(d)在指标上也高于融合结果图 11(b).

本文采用基于跨尺度相似先验的遥感图像时空融合算法,由于高尺度相似块先验引导大大提高了融合精度,导致传感器偏差校正的效果没那么突出. 如果融合算法不采用先验结构引导,传感器校正的优势会更加明显. 另外,后面如第 3.2 节所述需要计算中间尺度图像块的距离,若未校正传感器偏差,则会导致距离度量产生额外误差,稀疏结构引导不够精确,从而融合结果出现误差.

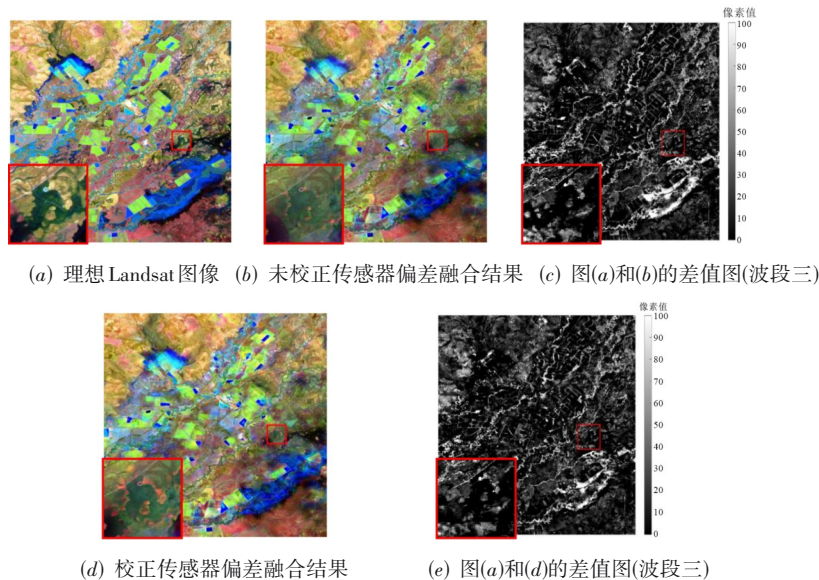
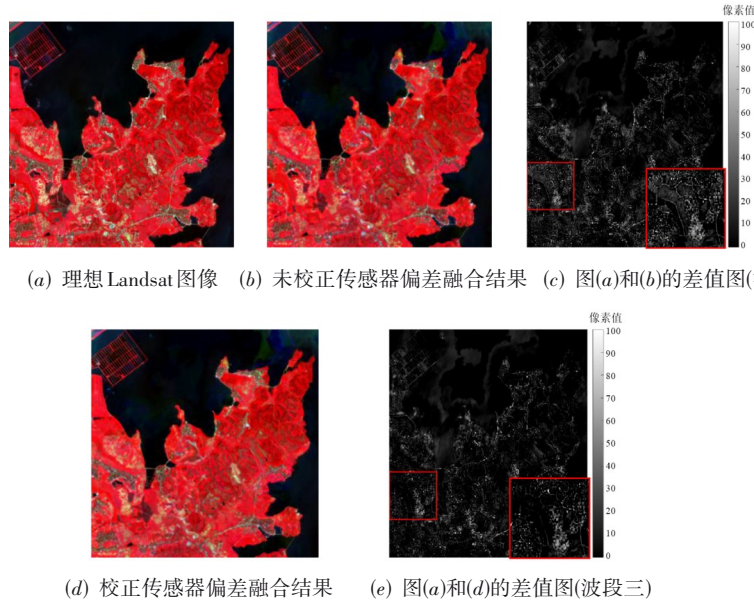


图 10 在 Landsat-MODIS 数据集上的传感器偏差校正消融实验



(a) 理想 Landsat 图像 (b) 未校正传感器偏差融合结果 (c) 图(a)和(b)的差值图(波段三)

(d) 校正传感器偏差融合结果 (e) 图(a)和(d)的差值图(波段三)

图 11 在 Sentinel-Landsat 数据集上的传感器偏差校正消融实验

表 2 不同传感器之间的偏差校正消融实验

单位:%

数据集	不同传感器	PSNR	CC	SAM	ERGAS	SSIM	RMSE
Landsat-MODIS 数据集	未校正传感器偏差融合结果	1 697.690	75.954	900.670	31.170	55.348	3.630
	校正传感器偏差融合结果	1 713.590	76.612	875.210	29.973	56.844	3.557
Sentinel-Landsat 数据集	未校正传感器偏差融合结果	2 337.480	91.309	1 288.410	48.060	74.946	1.731
	校正传感器偏差融合结果	2 410.740	92.215	1 224.160	44.346	76.372	1.589

实验 2 本文的传感器偏差校正是一个独立模块,可作为预处理模块,提高现有时空融合算法的精度.本实验对比将原始 MODIS 图像和校正 MODIS 图像分别作为输入时,对其他融合算法的影响.这里选用了基于稀疏表示的时空融合算法 Onepair 和 FISTA.

比较图 12(b)和图 12(c),以及图 12(d)和图 12(e),可以看出,校正后的融合结果在细节上有着明显的提高.由于原始 MODIS 图像有着明显的块效应,导致未预处理的融合结果也具有一定的块效应;而经过传感器偏差校正的预测结果,在图像的不同区域之间过渡得更加自然平滑,基本上消除了块效应.从表 3 可以看出,Onepair 和 FISTA 的 SSIM 比之前分别提高了 7.8% 和 7.2%,SAM 比之前分别提高了 4.7% 和 4.4%.

实验 3 为了验证本文算法的优越性,分别在 3 个具有代表性的数据集上,与其他算法进行对比实验.为

为了更好地对比各算法的效果差异,左下角添加了局部放大图.

场景 1 在物候变化场景的算法性能.

从图 13 可以看出,STARFM 采用加权的方式,图 13(c)体现出明显的模糊;Onepair 方法采用两层融合模型和高通调制策略,其细节有所增强;FISTA 采用了 t_1 时刻对应高空间分辨率图像块的稀疏系数做约束,增强了空间细节,其融合结果视觉效果很好,但是在变化区域部分会受 t_1 时刻图像影响而引入错误,与前两者相比,其 SAM 和 SSIM 有所改善;FSDAF 和本文算法在视觉上都达到了相对较好的效果,FSDAF 通过解混和 TPS 插值两种方式估计预测两个时刻的场景变化,与 t_1 时刻高空间分辨率图像相加得到融合结果.由于该场景以物候变化为主,模型很好地估计出光谱变化,而场景细节区域变化较小,因此 FSDAF 融合结果的空

表 3 传感器偏差校正模块在 Onepair 和 FISTA 方法中的消融实验

单位:%

方法	PSNR		CC		SAM		SSIM		RMSE	
	原算法	预处理 + 原算法	原算法	预处理 + 原算法	原算法	预处理 + 原算法	原算法	预处理 + 原算法	原算法	预处理 + 原算法
Onepair	1 532.170	1 576.030	72.731	74.399	1 045.830	996.420	47.367	51.041	4.404	4.185
FISTA	1 468.390	1 508.770	71.734	73.663	1 252.080	1 196.410	47.431	50.851	4.756	4.542

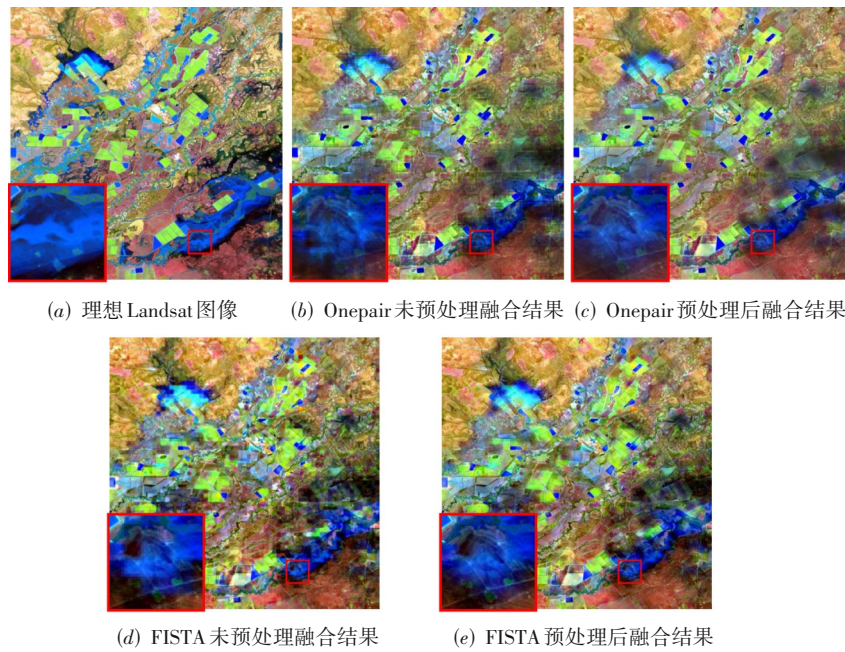


图 12 传感器偏差校正模块在 Onepair 和 FISTA 方法中的消融实验

间和光谱性能都比较好. 本文算法由于跨尺度相似块
的先验引导, 在细节和光谱上表现最优. 从表 4 可以

看出, 本文算法在各项指标上都是最优, FSDAF 表现
次之.

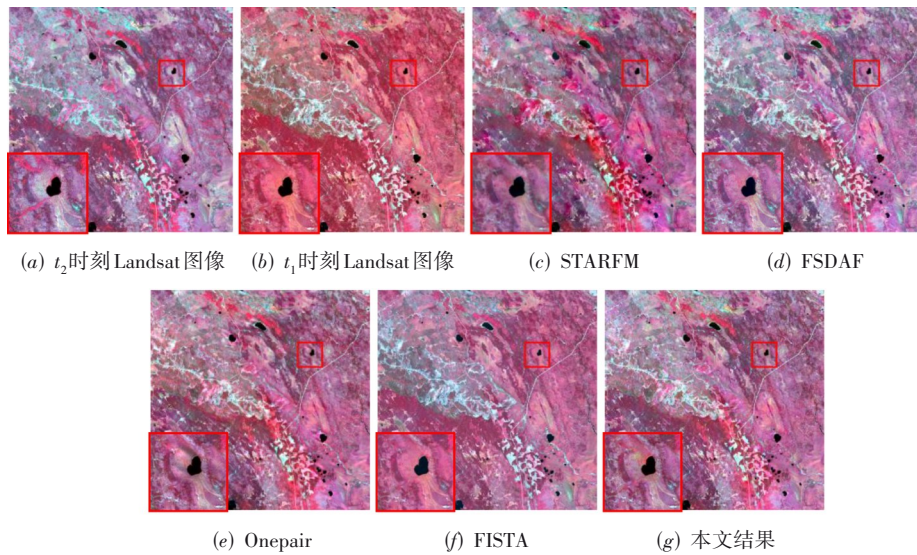


图 13 BOREAS 数据集融合结果比较

表 4 BOREAS 数据集融合结果评价指标比较

单位: %

指标	STARFM	FSDAF	Onepair	FISTA	本文算法
PSNR	1 773.430	<u>1 967.920</u>	1 852.960	1 885.590	2 028.86
CC	81.540	<u>84.910</u>	83.636	82.940	87.382
SAM	923.510	<u>752.560</u>	905.410	847.330	718.210
ERGAS	30.530	<u>20.460</u>	25.379	22.860	19.960
SSIM	72.730	<u>76.360</u>	73.397	75.160	79.546
RMSE	3.450	<u>2.790</u>	3.124	2.920	2.547

场景2 在异质景观场景的算法性能.

图 14 是典型的异质区域和物候变化区域,由局部放大区域可以看出,STARFM 在大的轮廓和区域与理想图像接近,但有一定程度的模糊,在非常细小的区域受退化的低空间分辨率图像影响,权重估计误差变大,导致融合结果不佳. 由于高低空间分辨率图像尺度差异非常大,而本场景又属于异质区域,导致解混和插值的效果都比较差,FSDAF 算法整体效果较差. 本实验场景的两个时刻的细节基本没变,Onepair 采用了高通调制,FISTA 采用了 t_1 时刻细约束,对融合结果体现以积极

的作用为主,因此两者细节上表现较好,但在放大区域还是可以看出先验带来的负面影响,融合结果在细节上明显更接近 t_1 时刻图像,偏离理想结果. GANSTFM 通过设计网络结构和损失函数,提取高质量特征,构造有效监督,从而获得较好的预测效果,由于只有两幅图像作为输入,在样本不足的情况下,成本函数的拟合效果稍有欠缺,导致整体结果更接近 t_1 时刻图像. 本文由于精细地构建跨尺度相似块,融合结果更加逼近理想状态. 由表 5 可以看出,本文方法在各项指标都表现出最优,GANSTFM 整体表现次之.

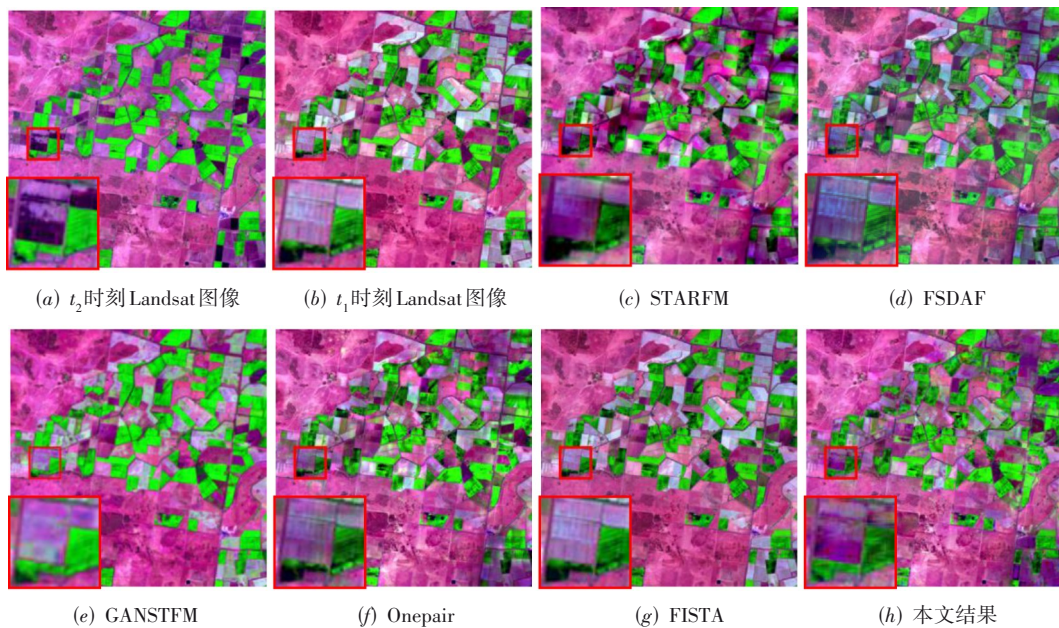


图 14 CIA 数据集融合结果比较

表 5 CIA 数据集融合结果评价指标比较

单位: %

指标	STARFM	FSDAF	GANSTFM	Onepair	FISTA	本文算法
PSNR	1 620.040	1 628.680	<u>1 769.860</u>	1 697.920	1 569.770	1 841.410
CC	79.210	78.280	<u>83.104</u>	81.250	75.950	84.648
SAM	946.420	1 096.090	850.980	<u>819.570</u>	854.820	715.000
ERGAS	39.420	33.240	<u>33.186</u>	34.005	38.100	27.373
SSIM	61.890	63.060	<u>68.794</u>	64.616	66.900	70.680
RMSE	4.180	4.040	<u>3.625</u>	3.841	4.250	3.141

场景3 在类型变化场景的算法性能.

从图 15 可以看出,FISTA 和 Onepair 都利用了 t_1 时刻的细节做优化,由于两个时刻场景发生了巨大变化,先验优化对融合结果以负面作用增大,他们在 SSIM 上表现不好. 由于无法从低空间分辨率图像中准确地捕捉变化信息,尤其是细小的变化,因此 STARFM 的融合效果表现不佳. 由于场景类型变化预测难度大,导致 FSDAF 融合效果相对于物候变化效果大幅度降低. 因

此,在 LGC 数据集上表现不理想. GANSTFM 从原理上讲,可以选择 t_2 时刻的低空间分辨率图像和相同场景下任意时间的另一个高空间分辨率图像用于模型输入,但是 LGC 数据集不同时刻场景变化较大,损失函数的拟合难度增大,GANSTFM 的预测效果表现不佳. 本文算法理论上不受场景变化的限制,且采用跨尺度相似块的稀疏系数约束. 由表 6 可以看出,本文算法优于其他算法的融合结果.

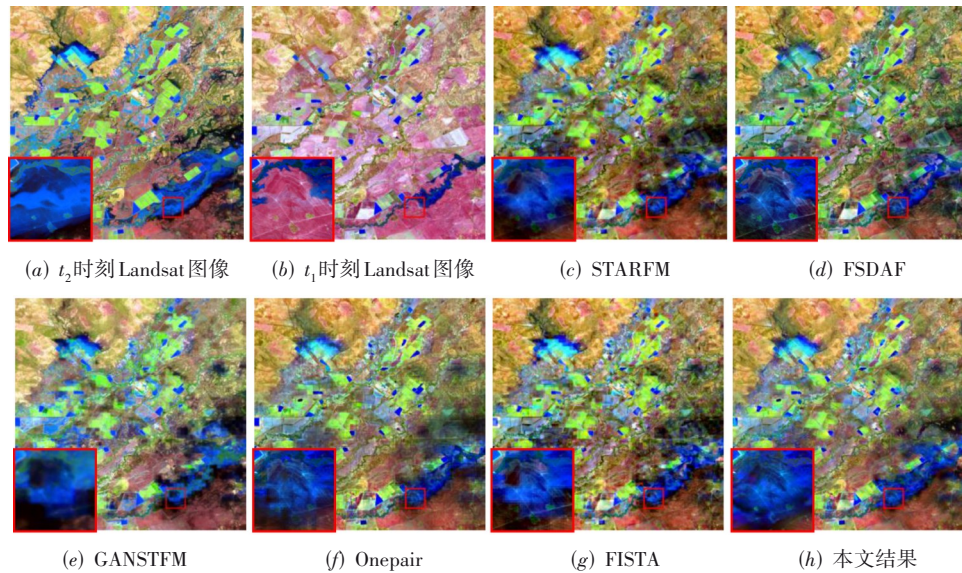


图 15 LGC 数据集融合结果比较

表 6 LGC 数据集融合结果评价指标比较

单位: %

指标	STARFM	FSDAF	GANSTFM	Onepair	FISTA	本文算法
PSNR	1 445.000	1 470.010	<u>1 547.440</u>	1 532.170	1 468.390	1 713.590
CC	72.246	70.369	<u>73.874</u>	72.731	71.734	76.612
SAM	1 290.220	1 190.730	1 123.400	<u>1 045.830</u>	1 252.080	875.210
ERGAS	48.953	42.323	<u>35.489</u>	37.823	42.889	29.973
SSIM	47.534	50.556	<u>53.079</u>	47.367	47.431	56.844
RMSE	4.905	4.777	<u>4.367</u>	4.404	4.756	3.557

5 结论

本文将多源遥感图像时空融合问题分解为设备相关的传感器偏差校正和设备无关的时空融合两个子问题,消除了时空融合模型的设备依赖,提高模型通用性,且传感器偏差校正可以作为预处理模型. 模型两部分可以一起使用,也可以独立使用,非常灵活. 针对时空融合部分,本文提出基于跨尺度相似先验的遥感图像时空融合算法,利用相似块构建稀疏结构先验的正则项,优化稀疏表示的目标函数,并构建中间尺度图像,降低跨尺度相似块的二义性,解决了高低空间分辨率尺度差异过大时,细节难以精确重建问题. 本文分别使用 3 组典型场景的实验数据集与其他算法进行对比,实验结果表明本文算法能够很好地适应这 3 种环境,在空间和光谱特性上均表现出优秀的特性.

本算法的一个核心问题是跨尺度相似块的确定,我们需要构建中间尺度图像,利用中间尺度图像的距离度量,进行跨尺度相似块的挑选和加权,合成精度较高的跨尺度相似块,提供比较准确的稀疏系数的估计. 未来,进一步统计分析跨尺度相似块的统计特性,使得跨尺度相似块的选取有理论支撑,提高跨尺度相似块的准确性,减少运算成本.

参考文献

- [1] VOGELMANN J E, HOWARD S M, YANG L, et al. Completion of the 1990s National Land Cover Data set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources[J]. Photogrammetric Engineering and Remote Sensing, 2001, 67(6): 650-662.
- [2] MASEK J G, HUANG C Q, WOLFE R, et al. North American forest disturbance mapped from a decadal Landsat record[J]. Remote Sensing of Environment, 2008, 112(6): 2914-2926.
- [3] WOODCOCK C E, OZDOGAN M. Trends in land cover mapping and monitoring[M]//Land Change Science. Dordrecht: Springer, 2012: 367-377.
- [4] GAO F, MASEK J, SCHWALLER M, et al. On the blending of the Landsat and MODIS surface reflectance: Predicting daily Landsat surface reflectance[J]. IEEE Transactions on Geoscience and Remote Sensing, 2006, 44(8): 2207-2218.
- [5] HILKER T, WULDER M A, COOPS N C, et al. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MO-

- DIS[J]. *Remote Sensing of Environment*, 2009, 113(8): 1613-1627.
- [6] ZHU X L, CHEN J, GAO F, et al. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions[J]. *Remote Sensing of Environment*, 2010, 114(11): 2610-2623.
- [7] ZHUKOV B, OERTEL D, LANZL F, et al. Unmixing-based multisensor multiresolution image fusion[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, 37(3): 1212-1226.
- [8] ZURITA-MILLA R, CLEVERS J G P W, SCHAEPMAN M E. Unmixing-based Landsat TM and MERIS FR data fusion[J]. *IEEE Geoscience and Remote Sensing Letters*, 2008, 5(3): 453-457.
- [9] WU M Q, NIU Z, WANG C Y, et al. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model[J]. *Journal of Applied Remote Sensing*, 2012, 6(1): 063507.
- [10] WU M Q, HUANG W J, NIU Z, et al. Generating daily synthetic Landsat imagery by combining landsat and MODIS data[J]. *Sensors*, 2015, 15(9): 24002-24025.
- [11] ZHU X L, HELMER E H, GAO F, et al. A flexible spatiotemporal method for fusing satellite images with different resolutions[J]. *Remote Sensing of Environment*, 2016, 172: 165-177.
- [12] SONG H H, LIU Q S, WANG G J, et al. Spatiotemporal satellite image fusion using deep convolutional neural networks[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(3): 821-829.
- [13] ZHANG H Y, SONG Y Y, HAN C, et al. Remote sensing image spatiotemporal fusion using a generative adversarial network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(5): 4273-4286.
- [14] LI Y F, LI J, HE L, et al. A new sensor bias-driven spatiotemporal fusion model based on convolutional neural networks[J]. *Science China Information Sciences*, 2020, 63(4): 140302.
- [15] TAN Z Y, YUE P, DI L P, et al. Deriving high spatiotemporal remote sensing images using deep convolutional network[J]. *Remote Sensing*, 2018, 10(7): 1066.
- [16] TAN Z Y, DI L P, ZHANG M D, et al. An enhanced deep convolutional model for spatiotemporal image fusion[J]. *Remote Sensing*, 2019, 11(24): 2898.
- [17] LI W S, ZHANG X Y, PENG Y D, et al. DMNet: A network architecture using dilated convolution and multi-scale mechanisms for spatiotemporal fusion of remote sensing images[J]. *IEEE Sensors Journal*, 2020, 20(20): 12190-12202.
- [18] JIA D, CHENG C X, SONG C Q, et al. A hybrid deep learning-based spatiotemporal fusion method for combining satellite images with different resolutions[J]. *Remote Sensing*, 2021, 13(4): 645.
- [19] HUANG B, SONG H H. Spatiotemporal reflectance fusion via sparse representation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(10): 3707-3716.
- [20] SONG H H, HUANG B. Spatiotemporal satellite image fusion through one-pair image learning[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(4): 1883-1896.
- [21] LIU X, DENG C W, ZHAO B J. Spatiotemporal reflectance fusion based on location regularized sparse representation[C]//2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). Piscataway: IEEE, 2016: 2562-2565.
- [22] WU B, HUANG B, ZHANG L P. An error-bound-regularized sparse coding for spatiotemporal reflectance fusion [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(12): 6791-6803.
- [23] WEI J B, WANG L Z, LIU P, et al. Spatiotemporal fusion of remote sensing images with structural sparsity and semi-coupled dictionary learning[J]. *Remote Sensing*, 2017, 9(1): 21.
- [24] WEI J B, WANG L Z, LIU P, et al. Spatiotemporal fusion of MODIS and Landsat-7 reflectance images via compressed sensing[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2017, 55(12): 7126-7139.
- [25] YANG W M, YUAN T R, WANG W, et al. Single-image super-resolution by subdictionary coding and kernel regression[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017, 47(9): 2478-2488.
- [26] ASEFPOUR VAKILIAN A, SARADJIAN M R. An object-based sparse representation model for spatiotemporal image fusion[J]. *Scientific Reports*, 2022, 12: 5021.
- [27] SHEN H F, WU P H, LIU Y L, et al. A spatial and temporal reflectance fusion model considering sensor observation differences[J]. *International Journal of Remote Sensing*, 2013, 34(12): 4367-4383.

- [28] TAN Z Y, GAO M L, LI X H, et al. A flexible reference-insensitive spatiotemporal fusion model for remote sensing images using conditional generative adversarial network[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5601413.

作者简介



方 帅 女,合肥工业大学计算机与信息学院教授,博士生导师.研究方向为计算机视觉、图像复原等.主持国家自然科学基金青年基金1项,面上基金1项,博士专项基金1项,学院杰出青年科学基金人才培育计划1项,企业合作项目多项,作为主要成员参加了国防973专项1项,国家自然科学基金3项,博士学科点专项科研基金1项,安徽省自然科学基金1项,近年来

共发表论文30余篇.

E-mail: fangshuai@hfut.edu.cn



曹 洋 男,中国科学技术大学信息科学技术学院自动化系副教授,博士生导师.研究方向为底层视觉处理及表征学习等.近3年来,在CVPR、ICCV、NeuRIPS、ICLR等人工智能领域顶尖学术会议以及IJCV、TNNLS、TIP等国际权威期刊上发表文章30余篇,承担国家重点研发计划、国家自然科学基金、安徽省重大专项等科研项目10余项,作为主要完成人获安徽省技术发

明奖一等奖、安徽省自然科学奖二等奖和中国自动化学会科学技术一等奖.

E-mail: forrest@ustc.edu.cn

万 旗 男,合肥工业大学计算机与信息学院硕士.研究方向为计算机视觉.

E-mail: 2020171223@mail.hfut.edu.cn