

基于时空融合的遥感图像镶嵌方法

何超琦, 李奇泽, 刘华霖, 魏静波*

南昌大学信息工程学院, 江西 南昌 330031

摘要 由于高分辨率的卫星幅宽有限, 涉及大面积的对地观测应用通常需要镶嵌多个拍摄于不同时间或不同传感器的高分辨率图像, 镶嵌图像的过程通常包括 5 个基本步骤, 即排序、配准、色彩平衡、接缝线检测和重叠区融合。现有的色彩平衡方法均未充分利用已有的卫星数据, 且大多数算法只追求视觉上的效果, 不追求数据的保真性, 为此提出一种新的时空融合镶嵌框架。该框架引入基于增强型深度超分辨率网络的时空融合方法, 通过一幅全局的参考图像将所有拍摄时刻的图像调整到统一时刻以保证色彩风格一致。本文方法在 LandSat8 图像的红、绿和蓝波段进行镶嵌测试, 实验结果表明所提方法比现有的镶嵌方法效果更好。

关键词 机器视觉; 遥感图像镶嵌; 色彩平衡; 时空融合; 卷积神经网络

中图分类号 TP751

文献标志码 A

doi: 10.3788/LOP202158.1415002

Remote Sensing Images Mosaicking Method Based on Spatiotemporal Fusion

He Chaoqi, Li Qize, Liu Hualin, Wei Jingbo*

School of Information Engineering, Nanchang University, Nanchang, Jiangxi 330031, China

Abstract Due to the limited width of high-resolution satellites, earth observation applications involving a large area usually require mosaics of multiple high-resolution images taken at different times or different sensors. The process of mosaicking images usually includes 5 basic steps: sorting, registration, color balance, seam line detection, and overlap area fusion. The existing color balance methods do not make full use of the existing satellite data, and most of the algorithms only pursue visual effects, not the fidelity of the data. For this reason, a new spatio temporal fusion mosaic framework is proposed. The framework introduces the spatio temporal fusion method based on the enhanced deep super-resolution network, and the images of all shooting moments are adjusted to the same time through a global reference image to ensure consistent color styles. The method in this paper carries out mosaic testing in the red, green, and blue bands of LandSat8 images. The experimental results show that the proposed method is more effective than the existing mosaic methods.

Key words machine vision; remote sensing image mosaicking; color balance; spatiotemporal fusion; convolutional neural network

OCIS codes 150.0155; 280.4788; 280.4750

1 引言

在对地观测技术的应用过程中, 大面积的区域调查需要大幅面且高分辨率的遥感图像作为数据支

撑。然而, 高分辨率的卫星幅宽有限, 使得高分辨率图像的成像幅面通常较小。为了获得覆盖大区域的图像, 需要拼接多个图像^[1], 而这些图像通常在不同的时间由不同的传感器拍摄而得。拼接过程采用的

收稿日期: 2020-09-16; 修回日期: 2020-10-21; 录用日期: 2020-11-14

基金项目: 江西省水利厅 2016 年水利科技计划(KT201616)

通信作者: * qzlicn@163.com

是图像镶嵌技术,该技术将具有重叠区域的多个不同图像在统一的地理坐标系下拼接成包含所有区域的大幅面图像。

一幅高质量的镶嵌图像应当信息丰富、色彩均匀以及定位准确。基于此,遥感图像的镶嵌过程通常包括 5 个基本步骤,即排序、配准、色彩平衡、接缝线检测和重叠区融合。镶嵌过程如下:1)排列参与镶嵌图像的拼接顺序^[2];2)当几何位置不准确时,应采用图像配准方法对两幅图像中的一幅进行配准以匹配另一幅图像;3)采用色彩平衡算法对镶嵌过程中使用的图像进行色彩调整,使相邻图像之间的色彩差异较小;4)在重叠区域中采用接缝线检测方法来寻找差异最小的路径,使得重叠区域在视觉上更自然;5)沿着最佳的接缝线将重叠区域缝合在一起^[3]。

每个像素位置处的辐射值会根据传感器、土地使用、光线、大气条件和成像时间等差异而发生变化,因此色彩平衡是所有镶嵌过程中最关键的步骤。科研学者已提出了许多方法并用于消除遥感图像在镶嵌过程中的色彩差异,这些方法主要可以分为基于直方图的方法、基于动量匹配的方法和基于辐射归一化的方法。

基于直方图的方法是利用重叠区域的联合直方图来平衡色彩,通常该方法中的映射函数没有特定的参数格式。例如,Jia 等^[4]采用了二维张量投票法来解决投票空间中的强度对齐问题;Oliveira 等^[5]采用了一种概率方法得到局部调色板映射函数,并将其扩展到整个目标图像。基于动量匹配的方法是将颜色从 RGB(Red, Green, Blue)空间转换到 $L\alpha\beta$ (L 为非彩色的亮度信号通道, α 为黄-蓝对抗通道, β 为红-绿对抗通道)空间、YUV(Y 代表灰度值, U 代表蓝色偏量, V 代表红色偏量)空间或 HSV (Hue, Saturation, Value)空间,然后线性匹配颜色分量的均值和标准偏差以使色彩达到一致。例如,Li 等^[6]采用了具有不同增益和偏移量的局部矩匹配法来逐行调整重叠区域的色调;Li 等^[7]采用了结合熵、亮度差和色差来自动选择参考图像的新指标,并构建了灰度校正模型以校正色调;Cresson 等^[8]在 $L\alpha\beta$ 去相关颜色空间中对多幅具有重叠区域的遥感图像进行全局色彩平衡,并提出了二次规划优化的解决方案以计算增益和偏移量。相对辐射归一化的方法通过在其他时间或其他传感器上选择一幅图像作为参考,并对参与镶嵌的图像的灰度值逐波段进行归一化,使其接近所选参考图像的辐射度,从

而达到数据保真性的目的。例如,Zhou 等^[9]使用了归一化水指数来选择原始伪不变特征(Pseudo-Invariant Features, PIF),并使用带有迭代控制的统计规则来确定最终的 PIF。其他提取 PIF 的典型方法包括时间不变聚类^[10]、迭代重加权多变量变化检测变换^[11]和对象图像的直方图建模^[12]等。

基于直方图方法和基于动量匹配方法的目的都是确保视觉质量,因此结果都缺乏数据的真实性,从而导致镶嵌结果不能直接用于定量分析。虽然基于 PIF 的辐射归一化方法使用参考图像来改善数据的真实性,但是由于提取 PIF 的过程中在很大程度上依赖于专业经验,甚至需要人工干预才能获得较高的性能,因此该方法仍然存在很大的局限性。上述方法中没有一种色彩平衡方法能够充分利用已有的卫星数据,如果能够充分利用已有的卫星数据,那么可以将所有参与镶嵌的图像都映射到同一时刻,则镶嵌后的图像色彩自然会很协调,并且达到数据保真性的目的。

基于时空融合的图像重建法可以产生接近目标成像时刻的图像,与辐射归一化算法不同的是,时空融合方法至少需要两个参考因素,即拍摄时间和预测时间。参考图像具有低空间分辨率和高时间分辨率的特征,而且其数据较多且易于获得。时空融合的目的是使数据保真,在 Wang 等^[13]的重建结果中,红、绿和蓝波段的平均辐射误差均小于 30%,视觉感知误差很小。因此,采用时空融合的方法将所有参与镶嵌的图像映射到同一时刻后再进行拼接,将会得到一幅高质量的镶嵌图像。

针对现有的色彩平衡方法均未充分利用已有的卫星数据,以及大多数算法只追求视觉上的效果,不追求数据保真性的问题,本文提出一种利用时空融合的镶嵌方法,该方法将参与镶嵌的图像中不一致的色彩风格转化为统一的时间风格。此外,由于目前缺乏适用于镶嵌的融合方法,本文提出一种基于深度神经网络的时空融合方法。该网络是离线进行训练的,只需要在线训练一个周期就可以适用于参与镶嵌的图像。与其他时空融合方法相比,该方法具有快速和高效的特点,更适用于镶嵌应用。

2 相关工作

2.1 时空融合方法

时空融合方法的目的是融合不同时空分辨率的图像,从而获得高时空分辨率的图像。为了充分利用不同遥感图像中的时空信息,现有文献中已经设

计了许多时空融合方法来合成像素级的高时空分辨率图像。按照不同的优化策略可将方法分为 4 类: 基于加权的方法、基于解混的方法、基于字典对的方法和基于卷积神经网络的方法。

基于加权的方法是假定不同类别或相邻位置的像素共享相似的像素值。已知图像中相似像素的位置, 然后基于相似性规则对预测时间的低分辨率图像中的像素进行加权以预测新像素。Gao 等^[14]提出了时空自适应反射率数据融合模型(Spatial and Temporal Adaptive Reflectance Fusion Model, STARFM), Hilker 等^[15]提出了用于映射反射率变化的时空自适应算法(Spatial Temporal Adaptive Algorithm for mapping Reflectance Change, STAARCH), 二者均是典型的基于加权的方法。Wang 等^[13]提出了一种三阶段的时空融合方法, 该方法首先使用回归模型来拟合以增加在两个时刻获取的观测值之间的相关性, 然后使用空间滤波来去除块状伪像, 最后对滤波后的图像进行残差补偿以充分利用低分辨率图像的光谱信息。

基于解混的方法是利用已知时刻的高分辨率图像来求出端元的分数丰度矩阵, 从而实现时空图像的融合。例如, Wu 等^[16]提出了一种时空数据融合算法, 该算法可提取分数覆盖并根据最小二乘误差规则来预测表面的反射率; Xu 等^[17]提出了一种包含先验光谱的解混方法, 用于平滑每个类别中的 STARFM 预测图像; Zhu 等^[18]提出了基于薄板样条插值的灵活时空数据融合(Flexible Spatiotemporal Data Fusion, FSDAF)算法; Shi 等^[19]采用了增强的 FSDAF 算法来改善其不精细空间细节, 该框架预测的图像具有更多的类内可变性和空间细节。

基于字典对的方法是引入字典学习和非解析优化来预测稀疏域中的缺失图像。对于相似传感器同时捕获的图像, 如果使用精心设计的字典, 则其稀疏编码系数可能非常相似。因此, 当使用设计的字典对用于其他时间的图像对进行编码时, 编码系数通常保持相似性。基于这一理论, 提出了基于稀疏表示的时空反射率融合模型(Sparse-representation-based SpatioTemporal reflectance Fusion Model, SPSTFM)^[20]、误差边界正则化半耦合字典学习(Error-Bound-regularized Sparse Coding Dictionary Learning, EBSCDL)^[21]和快速迭代收缩阈值算法(Fast Iterative Shrinkage Thresholding Algorithm, FISTA)^[22]。

近年来, 卷积神经网络(Convolutional Neural

Networks, CNN)和深度学习已用于时空融合。Dai 等^[23]提出了一种两层融合策略, 在每一层中使用 CNN 来开发图像之间的非线性映射。Song 等^[24]提出了两个 5 层的 CNN, 以解决 MODIS (Moderate-resolution Imaging Spectroradiometer)和 LandSat 图像之间对应关系复杂和空间分辨率差距大的问题。在预测阶段, Dai 等^[23]和 Song 等^[24]均设计了一个融合模型, 其中包含高通调制和加权策略, 可以充分利用先前图像中的信息。

2.2 存在的问题

毫无疑问, 深层神经网络在提升融合性能方面具有潜力, 但面临 CNN 的网络复杂度低和训练数据不足的问题, 提出了一种基于深层神经网络的时空融合方法, 并借助该方法达到遥感图像镶嵌的目的。具体来说, 将增强型的深度超分辨率(Enhanced Deep Super Resolution, EDSR)^[25]网络级联后用于时空融合, 并且使用了迁移学习技术。为了解决训练数据不足的问题, 首先使用了自然图像进行训练以建模空间差异, 并在后续使用遥感图像进行有限训练以建模传感器的差异, 具体内容在下一节详述。

3 方 法

本文提出一个利用时空融合来实现图像块之间色彩平衡的镶嵌框架, 同时介绍一种新的基于深度神经网络的时空融合方法, 用来达到图像镶嵌框架中色彩平衡的目的, 该方法利用已知的低分辨率图像来预测未知的高分辨率图像。

3.1 镶嵌框架

实验的目标是从一组几何校正后的卫星图像中建立一幅单一的镶嵌图像, 这些图像可以具有一个、三个或更多的波段, 而所提方法与传统方法之间存在三个显著差异。首先, 目标是不同的, 所提方法不仅追求视觉效果, 还追求数据的保真性。其次, 光谱范围是不同的, 所提方法不限制于可见光范围, 而传统方法需要红色、绿色和蓝色波段。最后, 在所提方法中不需要利用图像之间的重叠区域。

所提方法的关键是将低分辨率图像作为用于色彩平衡的参考图像, 但低分辨率图像通常有较大幅面, 覆盖了所有参与拼接的区域, 而所提方法是将所有图像映射到拍摄低分辨率图像的时刻再进行拼接。

将用于镶嵌的一组图像定义为 $\{X_{1,t_1}, X_{2,t_2}, \dots, X_{i,t_i}, \dots, X_{N,t_N}\}$, 其中 N 表示图像数量, t_i 表示参与镶嵌的第 i 个图像块 X_{i,t_i} 的拍摄时间。

X_{t_c} 表示镶嵌后的图像,其中 t_c 为给出结果时间, c 为给出结果时间的图像块序号。镶嵌方法可以分为如下步骤。

1) 对于待预测的 X_{t_c} ,提供一幅已知 t_c 时刻的低分辨率图像 Y_{t_c} 并作为低分辨率参考图像。与 X_{t_c} 类似, Y_{t_c} 同样覆盖了完整的镶嵌区域, t_c 可以是 $\{t_1, t_2, \dots, t_N\}$ 中的任何一个。

2) 将低分辨率图像 Y_{t_c} 裁成多个用于参考的图像块 $\{Y_{1,t_c}, Y_{2,t_c}, \dots, Y_{N,t_c}\}$,使得每个低分辨率图像块 Y_{i,t_i} 的地理位置与每个参与镶嵌的 X_{i,t_i} 相匹配。也就是说, Y_{i,t_i} 和 X_{i,t_i} 应该具有相似的拍摄时刻、地理位置和光谱范围。

3) 当每对低分辨率参考图像 (Y_{i,t_c}, Y_{i,t_i}) 与 X_{i,t_i} 的几何位置不准确时,还需要将 (Y_{i,t_c}, Y_{i,t_i}) 配准到图像 X_{i,t_i} 上, $i \in 1, \dots, N$ 。

4) 对所有高分辨率图像在每个波段的累积分布进行统计,用来找出每个波段上统一的上下阈值。根据对应波段的阈值,将所有参与镶嵌的 X_{i,t_i} 非线性且逐波段拉伸到 $0 \sim 255$,即对阈值内的像素值进行线性拉伸,阈值外的像素值则设为 0 或 255。对于所使用的参考图像 (Y_{i,t_c}, Y_{i,t_i}) ,根据 X_{i,t_i} 的阈值进行拉伸。值得注意的是,这一操作是逐个波段进行操作,即每个波段使用自己的阈值。此步骤操作完成之后,所有图像的数值范围都拉伸到了 $0 \sim 255$ 。

5) 使用三幅拉伸后的图像 Y_{i,t_i} 、 X_{i,t_i} 和 Y_{i,t_c} 来预测目标时刻 t_c 的图像 X_{i,t_c} 。直到所有参与镶嵌的图像都映射到 t_c 时刻,最终得到的结果为 $\{X_{1,t_c}, X_{2,t_c}, \dots, X_{N,t_c}\}$ 。

6) 最后将所有预测出的图像 $\{X_{1,t_c}, X_{2,t_c}, \dots, X_{N,t_c}\}$ 拼成一张完整的图像,其中重叠区可以进行简单平均计算。重叠区也可以通过图割算法来找到最优接缝,然后采用金字塔混合方法融合在一起。经过时空融合法得到的图像,色彩一致性非常高,一般无需复杂操作。

步骤 5) 采用时空融合的方法来预测目标时刻图像 X_{i,t_c} ,该方法的关键是从两个低分辨率图像的差异 Y_{ci} 中预测两个高分辨率图像的差异 X_{ci} ,可表示为

$$\begin{cases} Y_{ci} = Y_{t_c} - Y_{t_i} \\ X_{ci} = X_{t_c} - X_{t_i} \end{cases} \quad (1)$$

具体而言,首先将 Y_{ci} 输入时空融合模型中以预测 X_{ci} ;然后将 X_{ci} 与 t_i 时刻的高分辨率图像相

加得到灰度范围为 $0 \sim 255$ 的目标时刻图像;最后将该图像中的每个波段都根据步骤 4) 统计的阈值反向拉伸以映射到原始灰度范围,从而得到最终的目标时刻图像 X_{i,t_c} 。步骤 5) 的流程如图 1 所示。

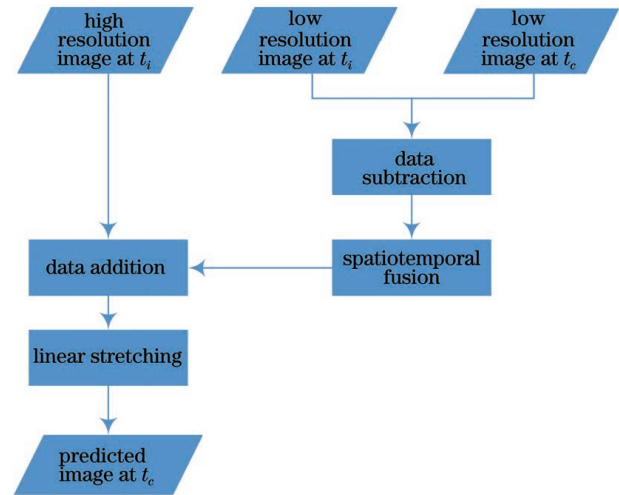


图 1 时空融合方法的流程

Fig. 1 Process of spatiotemporal fusion method

使用所提的镶嵌框架对图像进行大规模镶嵌,要求时空融合方法快速有效,但是目前性能较好的时空融合方法普遍速度较慢,FSDAF 算法和 EBSCDL 算法等处理一幅分辨率为 $1200 \text{ pixel} \times 1200 \text{ pixel}$ 的三波段图像所需要的时间都在 30 min 以上,而 STARFM 算法和 SPSTFM 算法等具有较快的处理速度,但效果又不太理想。为了将时空融合法应用于镶嵌过程,基于深层卷积神经网络和迁移学习提出了一种新的时空融合方法。

在所提的镶嵌框架中,参考图像是高分辨率图像在拍摄时刻和预测时刻的两幅低分辨率图像。这是一个典型的单对图像时空融合问题,预测时刻的高、低分辨率图像间的差异只有空间差异和其他差异(主要是传感器差异,也包括轻微的角度差异和时间差异等)。鉴于空间差异是主要差异,对空间差异进行建模就可以得到较高的融合质量。使用深层神经网络来建模空间差异,使用离线学习技术只需要一次迭代训练就可以使网络适应所需的图像内容。至于其他差异,可以在空间差异训练结束后进行有限学习,以适应由拍摄时刻和卫星高度等轻微不同带来的差异。

3.2 网络架构

图 2 为包含两个 X4 上采样器的级联网络,其中 \oplus 为跳过连接符号。网络的输入和输出是单波段图像,这意味着每个波段都有自己的模型参数。使用 EDSR 网络来建模高分辨率图像和低分辨率图

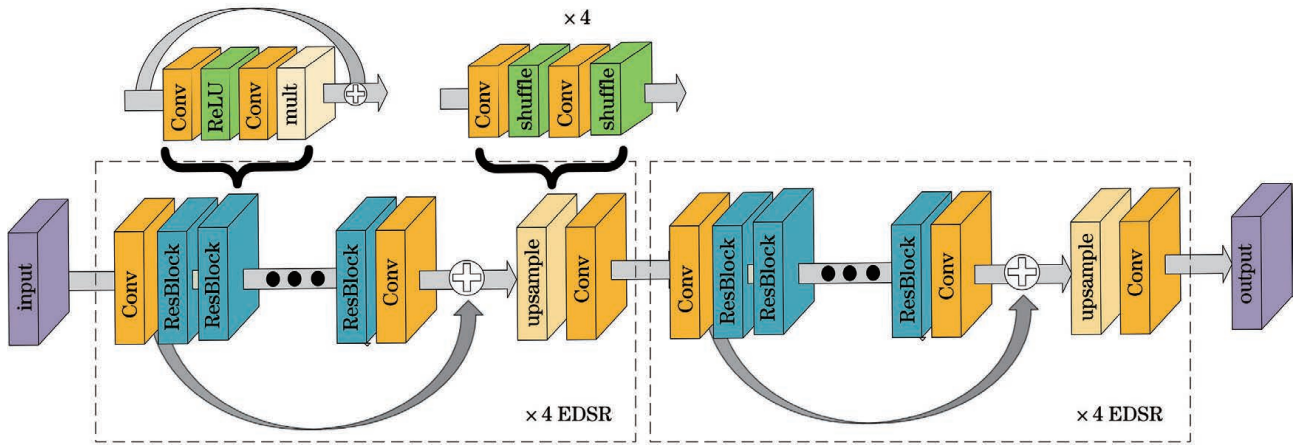


图 2 用于时空融合的级联卷积神经网络

Fig. 2 Cascaded convolutional neural network for spatiotemporal fusion

像间的空间分辨率差异。EDSR 网络的结构类似于 SRResNet^[26], 一个 4 倍上采样的 EDSR 主要包含 32 个残差块和一个上采样模块, 具体模型结构可参考图 2 中的单个 X4 EDSR 模型。每个残差块中包含两个卷积层、一个 ReLU 激活层和一个恒定比例的缩放层。与 SRResNet 不同的是, EDSR 在残差块中移除了批量归一化 (Batch Normalization, BN), 而且残差块之外没有 ReLU 激活层。上采样模块包含两个卷积层和两个像素重组模块, 像素重组模块是一种进行上采样操作的模块, 其将低分辨率的特征图通过卷积和多通道间的重组可以得到高分辨率的特征图。在所有卷积层 (也包括残差块和上采样模块的卷积层) 内, 卷积核的大小统一设为 3×3 , 特征通道数量设为 256。

高空间分辨率图像和低空间分辨率图像之间的分辨率通常差距很大。例如 LandSat8 的空间分辨率为 30 m, MODIS 的空间分辨率为 500 m, 其之间空间差异约为 16 倍, 但是典型的上采样器通常有 4 倍的最大分辨率。如果需要训练一个大于 4 倍的上采样器, 那么必须仔细级联多个低倍上采样器。例如, 为了获得具有高重建性能的 16 倍上采样器, 可以级联两个 4 倍上采样器。

3.3 网络训练

为了完整训练, 将模型分为两步。第一步, 训练一个理想的高倍率上采样器, 其中的训练数据来自高分辨率图像数据集; 第二步, 将一对已知时刻的低分辨率和高分辨率图像分别用作输入和输出, 并对其进行训练, 用来微调整个预训练网络来适应传感器的差异。

第一步的空间差异训练操作缺乏高质量的遥感数据。成像过程中存在大气模糊等因素, 导致遥感

图像的信噪比和细节出现了退化, 远比不上自然图像。为了准确建模空间差异, 这一训练过程使用了自然图像数据集 DIV2K, 其含有 800 张清晰度的图像, 这也可以理解为一种迁移学习的思路。

考虑到高、低分辨率图像之间存在着其他差异, 第二步将使用一对已知时刻的低分辨率和高分辨率图像对级联网络进行训练。该过程可以使整个网络适应传感器、卫星高度、拍摄时刻等与图像内容相关的差异, 最终建模出不同卫星图像间的映射关系。

不同来源的数据具有不同的定量范围。通常, 自然图像使用 8 位量化, 数字范围为 $0 \sim 255$, 为此可以直接将其放入网络中。遥感图像通常量化为 10 位以上, 并且不同传感器使用的位数可能会不同, 即使对同一目标进行同时观察, 其数值范围也会有很大差异。为了统一网络训练和重构的量化范围, 将高、低分辨率图像的数字范围非线性扩展到 $0 \sim 255$ 。为此, 将最亮的 0.1% 像素值设为 255, 最暗的 0.1% 像素值设为 0, 将其之间的像素值线性拉伸为 $0 \sim 255$ 。输出图像后重新拉伸到其原始范围, 利用输入图像在非线形拉伸过程中使用的上限和下限将 $0 \sim 255$ 的输出范围线性地转换为其原始范围。

对网络使用 Adam 优化器进行优化, 训练周期为 300 个。在 $1 \sim 200$ 周期内的学习率为 0.0001, 在 $200 \sim 300$ 周期内的学习率为 0.00005。实验所运行的硬件环境由 6 Mbit 的高速缓存和 2.10 GHz (6 核/12 路多任务处理) 的 Intel(R) Xeon(R) E5-2620 v2 处理器、32 GB 的内存、显存为 6 GB 的 NVIDIA Tesla K20Xm GPU、存储容量为 8 Tbit 东芝硬盘以及 Supermicro X9DRG-QF 主板组成。所

使用的软件环境由 Red Hat Enterprise Linux Server release 6.4×64 操作系统、CUDA 9、cuDNN 7.2、PyThon 3.5 以及使用 PyTorch 框架搭建而成。

4 实 验

在本节中,使用新提出的镶嵌框架对 LandSat8 高分辨率图像进行测试,并与一些最新或常用的算法进行比较。

4.1 实验方案

为了方便起见,将所提方法命名为时空融合镶嵌(MOST),将其与常用或最近提出的镶嵌算法进行比较,这些算法包括 ENVI 软件中的拼接算法、ENBLEND(enblending)^[27]、ENFUSE(enfusing)算法^[28],以及二次规划(Quadratic Programming, QP)算法^[8],所有对比算法都使用各自的默认参数。

镶嵌实验使用三组 LandSat8 高分辨率图像,分别于 2017 年 11 月 1 日、2017 年 12 月 19 日和 2018 年 10 月 3 日拍摄。这些图像在地理上相邻并且具有小范围的重叠区域,整体覆盖范围为 25°36′57″N~28°28′20″N 和 114°58′31″E~116°36′23″E。在比较过程中,所提方法和 ENFUSE 算法均使用相同位置的 MODIS MOD09A1 低分辨率数据作为参考。由于 QP 算法需要自然的颜色空间,所以只测试了红色、绿色和蓝色三个波段。

为了验证数据的保真性,准备一张大幅面的 LandSat8 卫星图像,其由参与镶嵌的第一个图像扩展而来的。也就是说,其覆盖了整个镶嵌区域,可以作为真值图像使用。因此,可以将镶嵌结果与此图像进行比较,用来观察辐射精度和光谱一致性。

4.2 实验结果

4.2.1 客观指标评价

首先评估镶嵌图像的辐射精度、结构相似性和颜色一致性。方均根误差(Root Mean Square Error, RMSE)和相关系数(Correlation Coefficient, CC)用于评价辐射精度,结构相似性(Structural Similarity, SSIM)和互信息(Mutual Information, MI)用于衡量结构的一致性,光谱角映射(Spectral Angle Mapper, SAM)、相对平均光谱误差(Relative Average Spectral Error, RASE)、总体相对误差合成指数(Erreur Relative Globale Adimensionnelle de Synthèse, ERGAS)和 Q4(Quaternion theory-based quality index)用于观察光谱一致性。当计算结构相似性时,彩色图像在计算之前先通过三个波段转换为灰度图像。由于 QP 算法的镶嵌结果为 8 bit,而其他镶嵌结果为 16 bit,因此没有计算 RMSE。

三个图像块和整体镶嵌结果的数字评估结果如表 1 所示。从表 1 可以看到,所提的 MOST 方法具有最高的辐射保真度和光谱一致性,并且明显优于

表 1 镶嵌结果的指标评价

Table 1 Digital evaluation of mosaicking results

Block	Algorithm	RMSE	CC	SSIM	MI	SAM	ERGAS	RASE	Q4
Block 1	ENVI	68	0.986	0.973	4.622	0.141	0.158	0.150	0.983
	ENBLEND	3	1.000	1.000	4.847	0.003	0.007	0.007	1.000
	ENFUSE	113	0.891	0.704	1.099	0.064	0.244	0.238	0.868
	QP	—	0.987	0.896	4.883	0.050	0.311	0.310	0.893
	MOST	2	1.000	1.000	5.217	0.001	0.004	0.003	1.000
Block 2	ENVI	219	0.673	0.861	1.501	0.073	0.516	0.490	0.601
	ENBLEND	210	0.694	0.875	1.646	0.063	0.497	0.473	0.626
	ENFUSE	126	0.817	0.701	0.830	0.062	0.300	0.290	0.781
	QP	—	0.903	0.873	1.698	0.076	0.339	0.334	0.868
	MOST	166	0.778	0.843	1.301	0.059	0.391	0.380	0.738
Block 3	ENVI	132	0.902	0.879	1.858	0.076	0.307	0.294	0.865
	ENBLEND	83	0.939	0.943	2.093	0.047	0.199	0.194	0.930
	ENFUSE	103	0.891	0.869	1.081	0.058	0.256	0.247	0.858
	QP	—	0.934	0.782	2.055	0.140	0.487	0.481	0.775
	MOST	91	0.923	0.909	1.496	0.054	0.229	0.227	0.913

续表

Block	Algorithm	RMSE	CC	SSIM	MI	SAM	ERGAS	RASE	Q4
Total	ENVI	153	0.854	0.900	1.857	0.096	0.357	0.340	0.818
	ENBLEND	130	0.868	0.941	2.266	0.038	0.302	0.289	0.845
	ENFUSE	114	0.870	0.788	1.000	0.061	0.267	0.259	0.842
	QP	—	0.936	0.852	2.242	0.090	0.391	0.388	0.838
	MOST	109	0.898	0.920	1.955	0.039	0.257	0.250	0.885

ENVI 算法、ENFUSE 算法和 QP 算法；ENVI 算法的镶嵌结果呈现出最大的辐射误差和光谱畸变；ENBLEND 算法的镶嵌结果中辐射误差较大，但是结构和光谱方面的一致性较好；ENFUSE 算法的镶嵌结果具有较小的辐射和光谱误差，但是结构相似性较差；QP 算法无法计算辐射保真性，但结构相似性和光谱一致性较好，特别是相关系数显示出计算结果与真实图像具有较高的相关性。仅仅从数值比较上，可以推断出 MOST 方法在定量应用方面具有优势。

4.2.2 视觉评价

对镶嵌结果进行视觉效果评估。考虑到 QP 算法的结果是在 0~255 之间，为了进行视觉比较，在评估前将所有镶嵌图像非线性拉伸至 0~255 范围，拉伸阈值为 2%，在减少异常像素影响的同时保持

数据的真实性。对于 16 bit 量化的 ENVI 算法、ENBLEND 算法、ENFUSE 算法和 MOST 方法，拉伸使用的上限和下限阈值来自真值图像。8 bit 量化的 QP 算法的拉伸使用了自身图像作为基准。

图 3 为不同镶嵌算法的镶嵌结果。从图 3 可以看到，所提的 MOST 方法的三个块具有良好的颜色一致性；ENBLEND 算法的结果显示出图像块之间存在明显的风格差异；在 ENVI 算法和 QP 算法处理的镶嵌图像中，块之间的颜色不一致较弱；ENFUSE 算法的不同块也具有较好的色彩一致性。图 4 为图 3 的局部区域。从图 4 可以看到，ENVI 算法、ENBLEND 算法和 ENFUSE 算法的结构和颜色与真值图 LandSat 差异较为明显，QP 算法和所提的 MOST 方法的结果与 LandSat 较为接近。

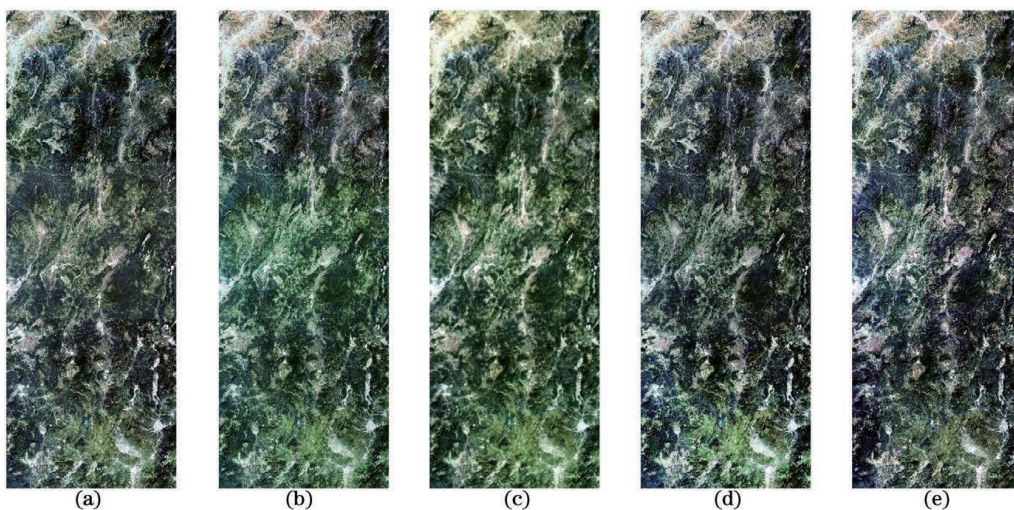


图 3 不同算法的整体镶嵌结果。(a)ENVI 算法；(b)ENBLEND 算法；(c)ENFUSE 算法；(d)QP 算法；(e)MOST 方法
Fig. 3 Integral mosaic results of different algorithms. (a) ENVI algorithm; (b) ENBLEND algorithm; (c) ENFUSE algorithm; (d) QP algorithm; (e) MOST method

不同算法在不同参考时刻的镶嵌结果如图 5 所示。在显示过程中，所有镶嵌图像都使用了 12 月的 LandSat 图像作为基准图像，并对其进行 2% 的非线性拉伸。从图 5 可以看到，不管使用哪个时刻作为参考时刻，所提方法在调整不同块的色调方面都能

与目标图像一致。同时，在调整不同块之间的色调风格一致性方面，所提方法非常稳定。

过渡区的色彩一致性与匀色算法处理的色彩质量高度相关。图 6 为镶嵌结果中的第一张图像和第二张图像过渡区域。从图 6 可以看到，在过渡区域

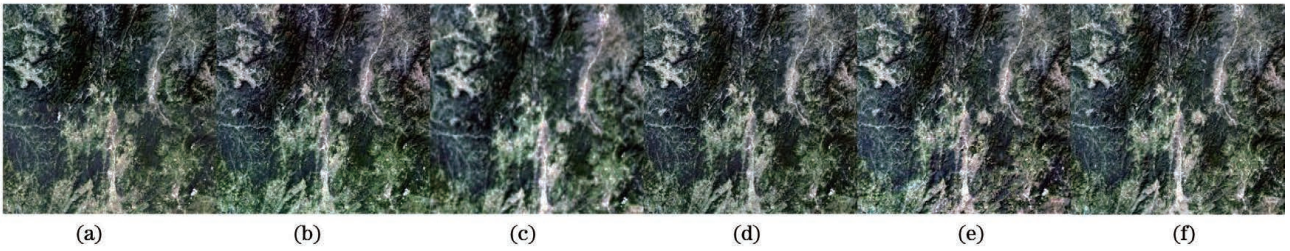


图 4 图 3 的局部区域。(a)ENVI 算法;(b)ENBLEND 算法;(c)ENFUSE 算法;(d)QP 算法;(e)MOST 方法;(f)LandSat
Fig. 4 Local area of Fig. 3. (a) ENVI algorithm; (b) ENBLEND algorithm; (c) ENFUSE algorithm; (d) QP algorithm;
(e) MOST method; (f) LandSat

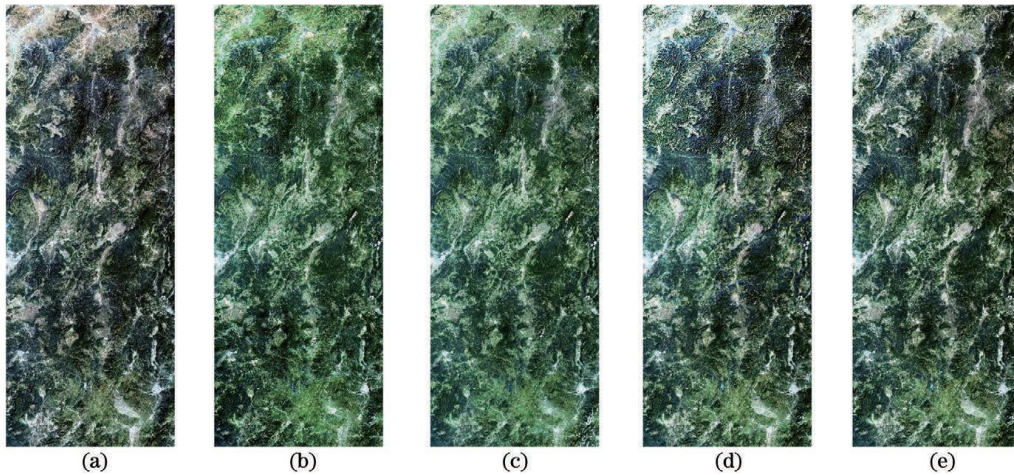


图 5 不同参考时刻的镶嵌结果。(a)LandSat (December);(b)MOST (October);(c)LandSat (October);(d)MOST (November);(e)LandSat (November)

Fig. 5 Mosaic results at different reference moments. (a) LandSat (December); (b) MOST (October); (c) LandSat (October); (d) MOST (November); (e) LandSat (November)

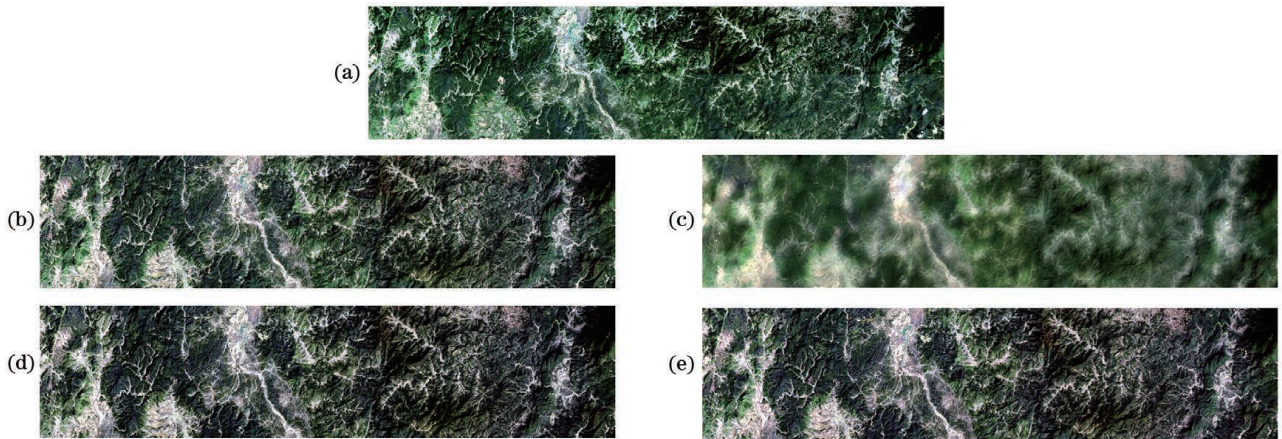


图 6 不同算法在过渡区的镶嵌结果。(a)ENVI 算法;(b)ENBLEND 算法;(c)ENFUSE 算法;(d)QP 算法;(e)MOST 方法
Fig. 6 Mosaic results of different algorithms in transition region. (a) ENVI algorithm; (b) ENBLEND algorithm;
(c) ENFUSE algorithm; (d) QP algorithm; (e) MOST method

的中心,ENVI 算法可以看到清晰的边框,表明其尚未调整块之间的颜色;在重叠区域中,其他算法没有明显的缝合标记,并且相邻块之间存在连续的颜色变化趋势,但是 ENFUSE 算法的结果显示出严重的细节丢失。

综上,QP 算法和所提方法在视觉上显示出令人满意的结果。由于块之间或过渡区的颜色风格不一致,ENVI 算法和 ENBLEND 算法的结果值得商榷,ENFUSE 算法显示出与真实图像严重的颜色偏差与细节损失。

5 结 论

本文提出一种采用时空融合算法消除镶嵌图像之间色彩差异的方法。时空融合是利用增强型深度超分辨率网络来实现的,利用全局的参考图像可以将所有拍摄时刻的图像调整到统一时刻,以保证色彩风格一致。基于神经网络进行时空融合后,不同镶嵌块之间具有整体一致的色彩风格、局部丰富的细节信息和难以觉察的重叠区,可以简化镶嵌过程中的接缝线检测和重叠区融合操作。

所提的 MOST 方法在 LandSat8 图像的红、绿和蓝波段进行镶嵌测试,并与 4 种匀色算法或工具进行比较。实验结果表明,所提方法在辐射、结构和光谱保真度方面可以获得最佳的性能,视觉比较显示出所提方法在整体上具有统一的色彩风格,在过渡区可以无缝过渡。

参 考 文 献

- [1] Li Z Y, Tian Y, Chen F J, et al. Aerial image stitching algorithm for unmanned aerial vehicles based on improved ORB and PROSAC [J]. *Laser & Optoelectronics Progress*, 2019, 56(23): 231003.
李振宇, 田源, 陈方杰, 等. 基于改进 ORB 和 PROSAC 的无人机航拍图像拼接算法[J]. *激光与光电子学进展*, 2019, 56(23): 231003.
- [2] Tian Y M, Sun A F, Wang D. Seam-line determination via minimal connected area searching and minimum spanning tree for UAV image mosaicking [J]. *International Journal of Remote Sensing*, 2018, 39(15/16): 4980-4994.
- [3] Gao X Q, Yang F, Fan H R, et al. Remote sensing image S-type fusion/stitching via low-error matching strategy [J]. *Laser & Optoelectronics Progress*, 2019, 56(21): 211005.
高晓倩, 杨帆, 范海瑞, 等. 低误差匹配策略下的遥感图像 S 型融合拼接[J]. *激光与光电子学进展*, 2019, 56(21): 211005.
- [4] Jia J, Tang C K. Tensor voting for image correction by global and local intensity alignment [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, 27(1): 36-50.
- [5] Oliveira M, Sappa A D, Santos V. A probabilistic approach for color correction in image mosaicking applications [J]. *IEEE Transactions on Image Processing*, 2015, 24(2): 508-523.
- [6] Li X H, Hui N, Shen H F, et al. A robust mosaicking procedure for high spatial resolution remote sensing images[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 109: 108-125.
- [7] Li J Y, Hu Q W, Ai M Y. Optimal illumination and color consistency for optical remote-sensing image mosaicking[J]. *IEEE Geoscience and Remote Sensing Letters*, 2017, 14(11): 1943-1947.
- [8] Cresson R, Saint-Geours N. Naturalcolor satellite image mosaicking using quadratic programming in decorrelated color space[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(8): 4151-4162.
- [9] Zhou H Z, Liu S M, He J J, et al. A new model for the automatic relative radiometric normalization of multiple images with pseudo-invariant features [J]. *International Journal of Remote Sensing*, 2016, 37(19): 4554-4573.
- [10] Chen X X, Vierling L, Deering D. A simple and effective radiometric correction method to improve landscape change detection across sensors and across time[J]. *Remote Sensing of Environment*, 2005, 98(1): 63-79.
- [11] Canty M J, Nielsen A A, Schmidt M. Automatic radiometric normalization of multitemporal satellite imagery[J]. *Remote Sensing of Environment*, 2004, 91(3): 441-451.
- [12] Sadeghi V, Ahmadi F F, Ebadi H. A new automatic regression-based approach for relative radiometric normalization of multitemporal satellite imagery[J]. *Computational and Applied Mathematics*, 2017, 36(2): 825-842.
- [13] Wang Q M, Atkinson P M. Spatio-temporal fusion for daily Sentinel-2 images [J]. *Remote Sensing of Environment*, 2018, 204: 31-42.
- [14] 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.
- [15] 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 MODIS [J]. *Remote Sensing of Environment*, 2009, 113(8): 1613-1627.
- [16] 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.
- [17] Xu Y, Huang B, Xu Y Y, et al. Spatial and temporal image fusion via regularized spatial unmixing [J].

- IEEE Geoscience and Remote Sensing Letters, 2015, 12(6): 1362-1366.
- [18] 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.
- [19] Shi C, Wang X, Zhang W, et al. A comprehensive and automated fusion method: the enhanced flexible spatiotemporal data fusion model for monitoring dynamic changes of land surface [J]. Applied Sciences, 2019, 9(18): 3693.
- [20] Huang B, Song H H. Spatiotemporal reflectance fusion via sparse representation [J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(10): 3707-3716.
- [21] 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.
- [22] 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), July 10-15, 2016, Beijing, China. New York: IEEE Press, 2016: 2562-2565.
- [23] Dai P Y, Zhang H Y, Zhang L P, et al. A remote sensing spatiotemporal fusion model of Landsat and modis data via deep learning[C]//IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, July 22-27, 2018, Valencia, Spain. New York: IEEE Press, 2018: 7030-7033.
- [24] 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.
- [25] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1132-1140.
- [26] Ledig C, Theis L, Huszár F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 105-114.
- [27] Burt P J, Adelson E H. A multiresolution spline with application to image mosaics [J]. ACM Transactions on Graphics, 1983, 2 (4): 217-236.
- [28] Mertens T, Kautz J, van Reeth F. Exposure fusion [C]//15th Pacific Conference on Computer Graphics and Applications (PG'07) 2007, October 29-November 2, 2007, Maui, HI, USA. New York: IEEE Press, 2007: 382-390.