

引用格式: Zhang Yuzhuo, Li Zhiwei, Shen Huanfeng, *et al.* Threshold Adaptive Cloud Detection for FY-3D Images Using CALIPSO Data as Reference[J]. Remote Sensing Technology and Application, 2022, 37(5): 1109-1118. [张宇卓, 李志伟, 沈焕锋, 等. 激光雷达数据辅助的FY-3D影像阈值自适应云检测方法[J]. 遥感技术与应用, 2022, 37(5): 1109-1118.]

DOI: 10.11873/j.issn.1004-0323.2022.5.1109

激光雷达数据辅助的FY-3D影像阈值自适应云检测方法

张宇卓, 李志伟, 沈焕锋, 彭小元

(武汉大学 资源与环境科学学院, 湖北 武汉 430079)

摘要: 国产风云系列卫星可为全球范围内大气、陆地和海洋的遥感监测提供重要数据支撑, 由于光学卫星影像不可避免受到云覆盖的影响, 通过云检测获取准确的云掩膜是风云系列卫星影像精细处理与应用的关键。现有的云检测方法大多采用简单高效的阈值法, 然而由于传感器光谱响应以及不同场景云覆盖下垫面的辐射差异, 在缺少大量真实云覆盖标记情况下, 现有方法往往难以确定最优的检测阈值。鉴于此, 提出了一种阈值自适应的云检测方法(TACD), 顾及传感器波段特性以及云覆盖下垫面差异, 设置不同场景下的多通道阈值测试, 包括反射率及反射率组合测试、亮度温度测试、亮度温度差异值测试、卷云测试等, 联合具有高精度云层信息的激光雷达数据构建全球范围的云检测样本集, 实现基于样本集真实云标记的迭代阈值优化, 最终基于最优的阈值进行云检测。以风云三号(FY-3D)MERSI-II影像为例, 联合CALIOP云层数据构建全球范围的云检测数据集, 并将所提出的TACD方法云检测结果与官方云掩膜产品进行对比, 结果表明该方法较官方云检测算法精度有明显提高, 其中平均交并比从80.35%提升至84.09%, 召回率可达92.67%, 具有业务化应用的潜力。

关键词: 云检测; 阈值自适应; 风云三号; CALIPSO; TACD

中图分类号: TP75; TN958.98 **文献标志码:** A **文章编号:** 1004-0323(2022)05-1109-10

1 引言

风云三号(FY-3)气象卫星是我国的第二代极轨气象卫星, 被广泛应用于大气、陆地、海洋特性参数的遥感监测^[1-3]。但由于光学卫星影像不可避免受到云覆盖的影响, 据统计云覆盖着全球67%以上的地表^[4], 云覆盖阻碍了光学卫星获取有效地表信息, 限制了风云气象卫星影像的利用率和后续应用的精度。云检测是光学卫星影像应用中必不可少的预处理步骤。通过云检测获取光学卫星影像的高精度云掩膜, 区分云与非云的晴空地表区域, 将

有助于提升光学卫星影像的可用性, 并减少其在对地观测应用中由于云覆盖导致的偏差。因而, 实现高精度云检测对于光学卫星影像精细处理与应用至关重要。

现有的云检测方法可大致分为阈值法、机器学习法和多时相法三类。阈值法基于云的物理与光谱特性而构建, 如云具有高反射率和低温的辐射特性, 云与晴空地表在特定波长处具有显著的光谱特性差异以及云与地表的纹理特征差异^[5-8]。该类方法根据云与晴空地表的光谱差异设计阈值测试, 以达到区分云与非云地表的目的。经典云检测算法

收稿日期: 2021-12-28; 修订日期: 2022-06-28

基金项目: 国家自然科学基金项目(41971303、42101357), 中国博士后科学基金项目(2021M692462、2020TQ0229)。

作者简介: 张宇卓(1995-), 女, 河北石家庄人, 硕士研究生, 主要从事遥感图像处理方面的研究。E-mail: yuzhuozhang816@whu.edu.cn

通讯作者: 沈焕锋(1980-), 男, 河北河间人, 教授, 主要从事资源环境遥感方面的研究。E-mail: shenhf@whu.edu.cn

ISCCP (International Satellite Cloud Climatology Project) 算法^[9-12] 和 APOLLO (The AVHRR Processing scheme Over cloud Land and Ocean)^[13-15] 算法基于 AVHRR 数据反射率和发射率波段信息实现阈值法云检测; MODIS 团队^[16-19] 充分利用 MODIS 数据丰富的光谱信息, 在多种先验数据的辅助下, 设置阈值测试, 生成较高精度的云掩膜产品; Zhu 等^[20-21] 在自动云覆盖评估 (ACCA) 方法的基础上, 提出了 Fmask (Function of mask) 方法用于 Landsat 系类影像的云与云阴影检测。Li 等^[22] 针对仅有可见光与近红外波段的高分一号影像, 提出了多特征联合的云与云阴影检测方法, 通过充分挖掘影像中的多种光谱和空间特征, 实现了波段信息有限情况下的高精度云检测。

机器学习方法通常基于构建的云检测数据集进行模型训练, 然后将预训练的模型应用至其他数据进行云检测。应用于云检测的传统机器学习方法包括支持向量机 (SVM)^[23,24]、随机森林^[25]、决策树^[26] 和神经网络^[27,28]。作为机器学习方法的延伸, 深度学习方法是近年来被广泛应用于云检测并取得具有前景性的结果, 如卷积神经网络^[29-31]、对抗生成网络^[32,33] 等。在有足量具有代表性训练样本支撑情况下, 基于机器学习云检测方法通常能取得较高的精度, 但由于该类方法的云检测结果较大程度上依赖于训练样本, 因而现阶段在实际业务化生产中应用较少。

多时相云检测方法通过检测影像中像元的反射率在时域上的陡增来实现云像元的判别。Zhu 等^[34] 针对多时相 Landsat 影像提出了 Tmask 算法, 通过比对影像中像元原始反射率与拟合的反射率结果, 实现了 Landsat 影像云与云阴影的自动提取; Liu 等^[35] 基于云会导致影像中像元反射率在时序上的拐点的假设, 发展了 MOD09 数据的多时相云检测算法; Chen 等^[36] 提出一种不依赖热红外波段的增强型多时相云检测算法, 实现 Landsat-8 影像高精度云检测; Zhang 等^[37] 提出一种青藏高原地区多时相多通道阈值组合的云检测方法, 用于 FY-4A 影像的云检测。尽管多时相方法通常能比基于单时相的方法获得更高的云检测精度, 但其对输入影像在时序上具有更高的要求, 并在地表覆盖发生显著变化区域可能存在较大误差。

综上所述, 尽管基于机器学习和多时相的云检测方法可以获得相对较高的精度, 但其各自存在局

限性, 而阈值法目前仍然是业务应用最多的云检测方法。然而, 由于不同传感器的波段设置和光谱响应存在差异, 且固定阈值与全局阈值难以适应不同场景之间的差异, 传统阈值法受传感器和场景差异影响较大, 导致最优的阈值参数难以确定。考虑到激光雷达主动探测器可精确探测云层与气溶胶的垂直结构, 与影像匹配后可作为其高精度真实云覆盖标记, 应用于云检测算法和结果定量评估。因此, 以 FY3D MERSI-II 卫星影像为例, 提出了一种阈值自适应的云检测方法, 充分利用风云三号影像的多波段光谱信息, 结合激光雷达数据, 通过分场景确定多通道测试的最优阈值, 生成高精度的云掩膜。

2 多通道阈值自适应云检测算法

实验提出的多通道阈值自适应云检测算法 (Threshold Adaptive Cloud Detection, TACD) 主要包括 3 个步骤, 首先构建光学-LiDAR 云检测数据集; 其次基于云的物理和光谱特性, 通过多类型特征提取和多通道阈值测试, 其中, 本文选取 FY-3D MERSI-II 红、绿、蓝 3 个可见光波段, 1.03 μm 近红外波段, 1.38 μm 、1.64 μm 短波红外波段, 3.8 μm 、7.2 μm 、8.55 μm 和 10.8 μm 热红外波段单波段或组合波段作为输入数据提取不同类型特征; 最后, 以激光雷达数据为真实云标记进行阈值的自适应优化, 最终用于生成影像的云掩膜, 流程图如图 1 所示。

2.1 光学-LiDAR 配对数据集构建

配对数据集由原始影像经预处理后的全部波段数据和对应像元点参考云标记组成。考虑到人工标记云像元不仅存在主观性, 同时费时费力难以满足大量标记样本需求。实验将在时间和空间上与原始影像匹配的 CALIPSO 激光雷达云层数据中的云层数量数据集作为真实云标记, 用于阈值优化和精度对比验证。

配对数据集的构建主要包括时空匹配和异常值剔除。激光雷达数据与目标影像的匹配是构建数据集的关键^[38-40]。数据的匹配考虑时间、空间两个维度, 必要时需对目标影像采样以匹配激光雷达数据的空间分辨率。时间上, 考虑时间误差影响和像元点数量, 筛选一定时间差范围内的像元点。空间上, 采用最邻近法遍历搜索经纬度最接近的对应像元位置, 同时获取经纬度匹配的像元坐标, 保留经度差与纬度差较小的像元点, 减小空间位置差距

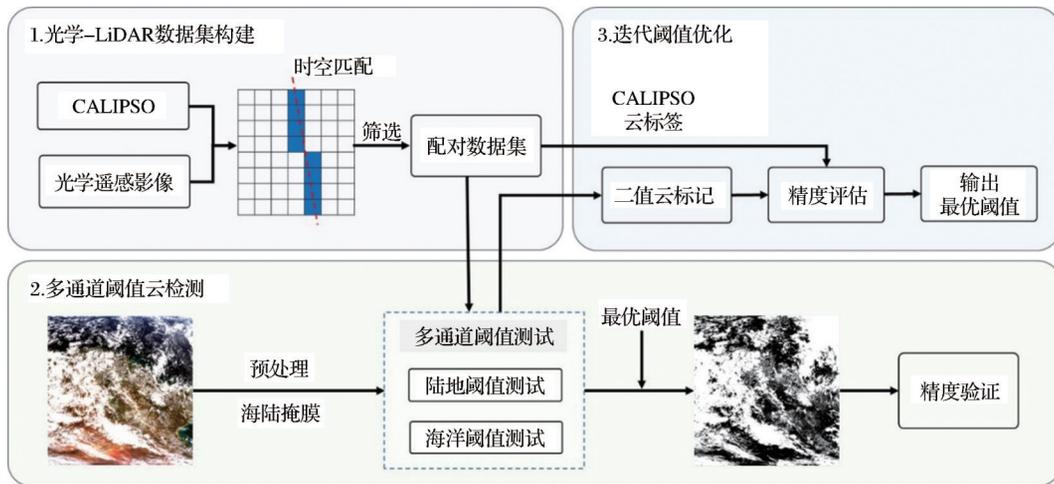


图1 多通道阈值自适应云检测算法流程图

Fig.1 Flowchart of the Threshold Adaptive Cloud Detection (TACD) algorithm

带来的误差,空间匹配过程如图2所示。

由于原始数据质量问题,数据集构建后需对数据集像元点进行筛选,剔除异常点与突变点。光学卫星影像存在数据缺失及数据异常等情况,体现为辐射定标后反射率为负值或亮度温度过低。CALIPSO数据为沿廓线分布的探测足迹点,数据点在空间上连续分布,观测足迹连续,连续数据中存在连续标记为晴空点突变为云像元的点与连续标记为云像元点突变为晴空点。上述情况均影响数据集质量从而干扰后续算法步骤,因此需剔除异常点与突变点,以确保数据集质量。

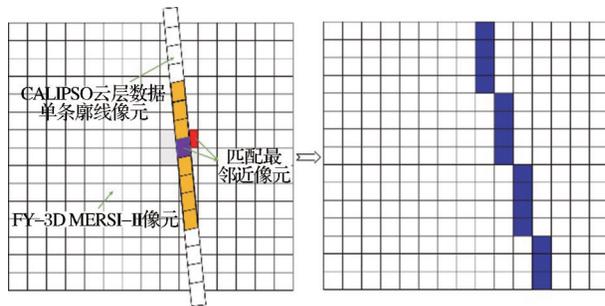


图2 CALIPSO数据与FY-3D MERSI-II影像空间匹配

Fig.2 Collocation of CALIPSO data and FY-3D MERSI-II image

2.2 多通道阈值测试

云相对于所覆盖下垫面,具有高反射率以及低温的特性,阈值测试以此为依据设置。如图3所示,所提出方法的多通道阈值测试主要包括反射率测试、波段组合测试、亮度温度测试、薄卷云测试和亮度温度差异值测试。此外,考虑到海洋表面均匀稳定,云和海洋的光谱差异易分离云和下垫面,而陆地表面地表覆盖类型丰富,地表反射率受下垫面类

型的影响而变化多样,因而,TACD通过FY-3D GEO地理信息数据中的海陆掩膜将影像分为海陆两个场景,根据场景特性分别设置不同测试,通过特定波段和波段的组合实现云与非云地表的区分。

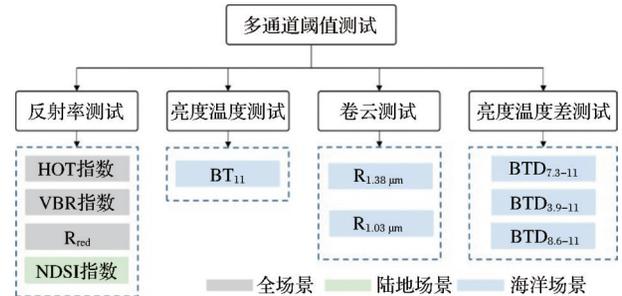


图3 多通道阈值测试

Fig.3 Multi-channel threshold tests

(1)HOT指数。HOT(Haze Optimized Transformation)指数依据云在红波段与蓝波段之间光谱响应差异区分云与非云地物^[41]。由于在晴空条件下,地表的可见光波段高度相关,但对云和雾的光谱响应在蓝和红波段之间存在不同,因此可以通过HOT指数提取云雾,简化后的HOT指数计算公式为:

$$HOT = Blue - 0.5 \times Red \quad (1)$$

(2)VBR指数。VBR(Visible Band Ratio)指数为红、绿、蓝波段的最小和最大反射率的比值,可剔除具有显著色彩特性的非云地物像元,当VBR指数接近于1时,灰度越高,像元点为云的概率越大^[22]。VBR指数计算公式如下:

$$VBR = \frac{\min(Blue, Green, Red)}{\max(Blue, Green, Red)} \quad (2)$$

(3)NDSI指数。反射率及反射率波段组合对

大多陆地场景都有较好的检测结果,但存在一些较为特殊的情况,如冰雪覆盖区域,冰雪所覆盖地表具有高反射率的特点,与云像元难以区分,因此需要结合NDSI指数纠正冰雪像元错检,NDSI指数可由绿波段和1.64 μm短波红外波段计算得到^[42,43],公式如下:

$$NDSI = \frac{(\text{Green} - \text{SWIR}_{1.64\mu\text{m}})}{(\text{Green} + \text{SWIR}_{1.64\mu\text{m}})} \quad (3)$$

反射率阈值测试为以上测试结果的综合结果,并加入对白色像元敏感的红波段测试,红波段反射率越大,是云的可能性越高,增加红波段反射率测试,提高云的检出率,反射率阈值测试的结果 CM_R 计算公式如下:

$$CM_R = (\text{HOT} > T_1) \cap (\text{VBR} > T_2) \cap (\text{Red} > T_3) \cap (T_4 < \text{NDSI} < T_5) \quad (4)$$

(4)亮度温度测试。海水表面较为稳定,具有均一性。云在11 μm波段和所覆盖海洋下垫面温度差异明显。当11 μm波段亮度温度 BT_{11} 低于270 K,该像元点可判定为潜在云像元^[19]。11 μm波段对低温厚云有较好的检测效果,但无法准确检测海面高空薄卷云,因此需要其他阈值测试检测薄卷云,海洋测试的结果 CM_{or} 计算公式如下:

$$CM_{or} = BT_{11} < T_6 \quad (5)$$

(5)薄卷云测试。1.38 μm波段靠近强水汽吸收带,对日间卷云十分敏感,常作为薄云检测的条件之一。除此之外,根据传感器的不同,检测卷云的方法可根据传感器波段设置而作调整,如FY-3D MERSI-II传感器的1.03 μm波段为MERSI-II特有的卷云波段,卷云检测结果 CM_{cir} 计算公式如下:

$$CM_{cir} = R_{cirrus} > T_7 \quad (6)$$

(6)亮度温度差异值BTD测试。云和晴空的亮度温度差异值是由普朗克函数的非线性和云的单一散射特性的光谱变化引起的,因此可以通过亮度温度差异值区分云和晴空像元,计算公式如下:

$$BTD_{i-j} = BT_i - BT_j \quad (7)$$

其中: BTD_{i-j} 为*i,j*波段亮度温度差异值, BT_i 为*i*波段的亮度温度, BT_j 为*j*波段的亮度温度。

中云由微小水滴、过冷水滴或者冰晶、雪晶混合而组成。晴空条件下,7.3 μm对大气层中层的温度和湿度很敏感,而11 μm对温暖表面较为敏感,云层会减少这一差异的绝对值,因此可以利用7.3 μm和11 μm亮度温度差异值 $BTD_{7.3-11}$ 检测中高层云^[19]。

低云水滴、过冷水滴、冰晶混合组成。日间3.9 μm波段反射太阳能量,在低空水云处3.9 μm和11 μm波段的亮度温度差异值 $BTD_{3.9-11}$ 很大且为负值,有较好的检测表现,因此也常用来检测对于空间高度较低的低云像元^[44]。

存在水汽吸收带的波段对云像元也有较好的检测效果。8.6 μm和11 μm波段亮度温度差异值 $BTD_{8.6-11}$,也作为检测云像元的一个判断条件,陆地表面8.6 μm波段辐射变化较大,具有不确定性,因此8.6 μm和11 μm波段亮度温度差异值只应用于海洋表面,亮度温度差异值阈值测试结果 CM_{BTD} 计算公式如下:

$$CM_{BTD1} = BTD_{7.3-11} > T_8$$

$$CM_{BTD2} = (BTD_{3.9-11} > T_9) \cup (BTD_{8.6-11} > T_{10}) \quad (8)$$

由于FY-3D MERSI-II数据无3.9 μm、7.3 μm、8.6 μm和11 μm波段,由近似波段,即3.8 μm、7.2 μm、8.55 μm和10.8 μm波段代替。

海洋场景下云像元判定计算公式如下:

$$CM_{Ocean} = (CM_R \cap CM_{OT} \cap CM_{cirrus1.38} \cup CM_{cirrus1.03}) \cap CM_{BTD1} \cup CM_{BTD2} \quad (9)$$

在分别获得陆地和海洋场景下的云检测结果后,可将其合并为最终的云掩膜CM,公式如下:

$$CM = CM_{Ocean} \cup CM_{Land} \quad (10)$$

2.3 迭代阈值优化

在TACD算法阈值优化过程中,参考不同测试的经验阈值来设定一定阈值范围,针对不同测试的阈值设置适当的迭代变化间隔,针对海洋和陆地场景分别进行多通道阈值测试,基于所构建的云检测数据集,评估每次阈值测试的精度,经多次循环迭代,确定总精度最高时的阈值参数为各场景的最优阈值。

3 实验结果与分析

3.1 实验数据

实验以CALIPSO激光雷达数据作为辅助,以空间分辨率1 km的FY-3D MERSI-II遥感影像为例,构建全球范围的云检测数据集。所使用的真实云标记为CALIPSO荷载的云-气溶胶正交极化激光雷达(CALIOP)的云层数据(Cloud Layer)。其中,云层数量数据集(Number_Layers_Found)可作为真实的云和晴空标记,数据集仅记录探测廓线云层个数,本文参考相关研究^[18,38,40,45],将探测结果为无云层的像元点标记为晴空,一层及多层云标记为有云。此外,本文所用FY-3D MERSI-II与FY-

3系列A、B、C 3颗卫星的第一代MERSI传感器相比,D星MERSI-II传感器更为稳定,波段信息更为丰富且数据质量明显提高^[46]。

在CALIPSO云层数据与FY-3D MERSI-II影像的匹配过程中,云层探测廓线足迹点与影像像元间需满足时间差在25 min内,经度和纬度差异小于0.02°。最后筛选掉云层探测廓线足迹中的异常点和突变点,以确保配对数据集的数据质量和可信度。

为保证实验结果和验证的可靠性,考虑时间与季节特征,选取了2019至2021年期间分布在全球范围并覆盖包括海洋、植被、裸土、极地和沙漠多类型地表场景的215景FY-3D MERSI-II影像作为实验数据,包括经过辐射定标的FY-3D MERSI-II一级全波段数据、FY-3D MERSI-II地理数据中的时空信息数据、角度数据和海陆掩膜。此外,还包括FY-3D官方云掩膜产品数据。将FY-3D MERSI-II影像与CALIPSO云层数据进行匹配,构成成对的云检测样本点。选取其中181景影像共127 264个样本点,分海洋和陆地两种场景分别进行阈值优化;另外34景影像中的35 255个样本点作为验证数据。

3.2 实验结果

基于所构建验证数据的云检测样本点,对FY-3D云掩膜产品和TACD算法云检测结果进行精度评价。采用总体精度(Overall Accuracy)、召回率(Recall)、准确率(Precision)、平均交并比(mIoU)、F1分数(F1 score)5个指标进行精度统计。5个指标的值越大,表明精度越高,其中召回率和准确率可分别反映漏检和错检情况。通过对不同阈值设置情况下云检测结果的精度测试,综合考虑以上精度指标,将取得最高精度时对应的阈值作为TACD算法的最优阈值。最终,在全场景取得最高总体精度84.94%的情况下,确定的海洋和陆地场景的最优阈值如表1所示。

将最优阈值作为TACD算法中最终的阈值参数,生成验证数据的云掩膜,并与FY-3D官方云掩膜产品对比。结果示例如图4所示。其中,FY-3D官方云掩膜产品结果分为云、可能有云、可能无云以及晴空4种类别,为便于进行定量的精度评价,将官方云掩膜划分为两种结果:FY-3D CLM¹将可能有云的情况划分至云类别,可能无云划分至晴空类别;FY-3D CLM²将可能有云和可能无云均划分至

表1 阈值优化所确定的最优阈值

Table 1 Optimal thresholds acquired by threshold optimization

	海洋场景	陆地场景
HOT	>0.11	>0.11
VBR	>0.45	>0.40
NDSI	N/A	>-0.3 & <0.59
R _{red}	>0.06	>0.06
R _{1.38μm}	>0.001	N/A
R _{1.03μm}	>0.05	N/A
BT ₁₁	<290 K	N/A
BT _{7.2} -BT ₁₁	>-29 K	N/A
BT _{3.8} -BT ₁₁	>6.5 K	N/A
BT _{8.6} -BT ₁₁	>-1 K	N/A

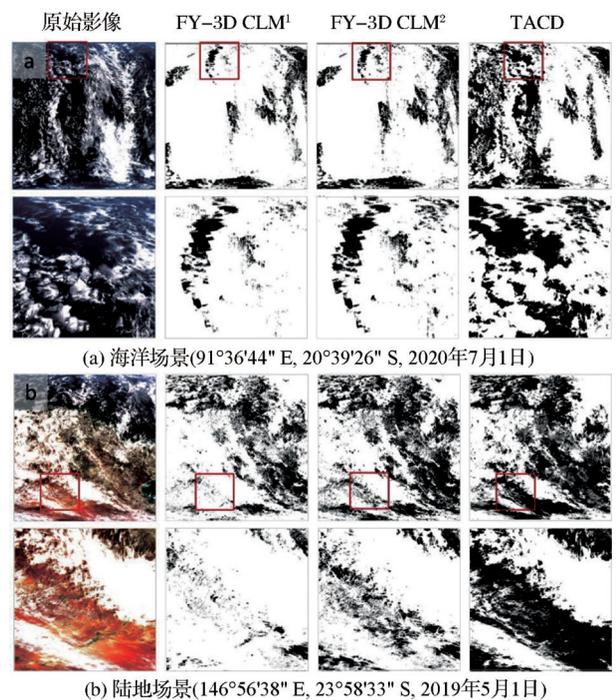


图4 云检测结果示例

Fig.4 Examples of cloud detection results

晴空类别。

目视对比结果表明,与TACD云检测结果相比,官方云掩膜产品受影像质量(条带噪声等)影响较大,存在较多晴空像元被错误分类为云像元。TACD算法生成的云掩膜云边界清晰,受原始影像质量影响较小,在海洋和陆地场景下的云检测结果均优于官方产品,对于容易漏检的海面薄云、碎云也有较好的检测效果。

此外,基于34景影像中35 255个样本点,对TACD算法和FY-3D官方云掩膜产品进行定量的精度评价,结果如表2所示。

表2 结果精度验证(单位:%)
Table 2 Validation accuracy(Unit:%)

	总体精度	召回率	准确率	平均交并比	F1分数
FY-3D CLM ¹	84.91	91.31	89.28	82.29	90.28
FY-3D CLM ²	83.62	87.24	91.05	80.35	89.10
TACD算法	86.54	92.67	90.08	84.09	91.35
TACD(海洋)	86.53	92.64	90.11	84.09	91.36
TACD(陆地)	87.28	95.56	86.69	83.33	90.91

定量评价结果表明,与官方方法相比,TACD算法在总体指标上均优于官方方法,其中总体精度可达86.54%,高于官方结果的83.62%和84.91%。TACD算法云检测召回率可达92.67%,优于官方云掩膜产品的91.31%和87.24%。根据准确率指标结果,TACD算法的云掩膜中错检高于FY-3D CLM²,与FY-3D CLM¹接近。此外,在海洋和陆地的分场景精度评估结果表明,TACD算法在陆地场景的云检测总体精度与海洋场景接近,但在陆地场景的查全率为95.56%,显著高于海洋区域的92.64%。

3.3 误差分析

云的漏检通常出现在薄云和碎云区域,特别是在地表反射率变化较大的非均质地区。云的错检通常受下垫面类型影响,对于具有高地表反射率的沙漠、极地等区域,云像元易与下垫面混淆而出现误检。TACD方法在一些区域也存在误差,图5展示了更多场景下的云检测对比结果。在图5(a)和5(b)所示的海岸线和高亮的陆地地区,官方云掩膜存在较大面积的错检,TACD算法云掩膜具有较好的目视效果,但存在少量的薄云漏检;在图5(c)所示的沙漠区域,TACD云掩膜与官方云掩膜都存在错检和漏检问题,这可能由于受限于匹配到数据集的空间局限性,在沙漠区域样本点少于其他场景,导致所得到的阈值参数对于该区域并不是最优;在极地区域的云检测具有较大的挑战性,图5(d)中影像为假彩色合成图,其中白色为云,蓝色为冰雪,目视上TACD算法结果优于官方云掩膜,但仍有提升空间。

FY-3D影像与CALIPSO数据时空匹配的误差不可避免,导致参考的云标签存在偏差而对总精度产生影响。此外,由于FY-3D与CALIPSO轨迹和过境时间差异,导致两者仅能匹配到较少数据,海岸线区域因面积相对较小难以匹配到样本点,沙漠和高原区域受匹配限制样本点少导致训练不够充分,极地区域云检测难度较大,因而仍有改进与

提升的空间。本文为匹配到更多样本点并涵盖多类型场景,适当增大了两者匹配允许的空间误差与时间误差。然而,这也可能导致在风速较大云速较快的情况下,由于云位置的偏移而产生匹配误差,给TACD算法的阈值优化和精度评价带来不利影响。

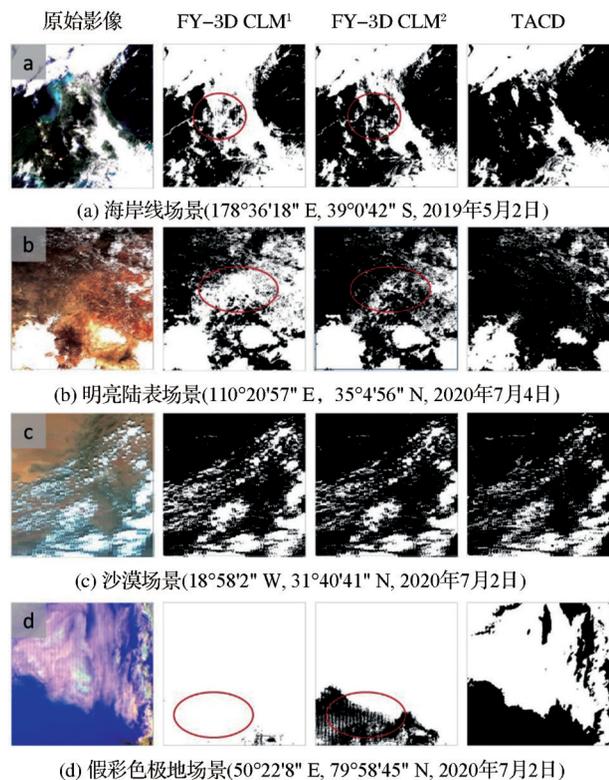


图5 多场景云检测结果示例

Fig.5 Comparison of results in multiple scenes

3.4 云掩膜后处理

对于云检测实验区域结果中存在的漏检问题,可利用引导滤波进一步精细化云掩膜^[47]。以原始RGB影像作为引导图,以二值云掩膜作为输入,通过引导滤波基于局部窗口处理可生成引导羽化后的云灰度图,经阈值分割后可获得精细化的云掩膜。本文对基于引导滤波的云掩膜后处理进行了测试,其中窗口大小设置为10,分割阈值为0.12,图6展示了其中的一组测试结果。目视结果表明引导滤波后云的漏检有明显改善,尤其是在云边缘处,此外还削弱了原始影像质量问题给云掩膜带来的影响。

云掩膜经引导滤波后处理,可有效减少厚云边缘薄云的漏检,然而可能会增加对厚云周围非云高亮地表的错检。根据基于34景影像中35 255个样本点的定量精度评价结果,经过引导滤波精细化处理的云掩膜总精度略微下降,但召回率提升明显,

从92.67%提高至97.07%,表明漏检情况得到较大改善。然而,准确率由90.08%下降至85.95%,这部分归因于极地区域冰雪的错检。因此,在实际应用中,基于引导滤波的云掩膜后处理适合于对云掩膜中漏检具有较低容忍度的应用,尽管其在一定程度上会造成总体精度略微下降。

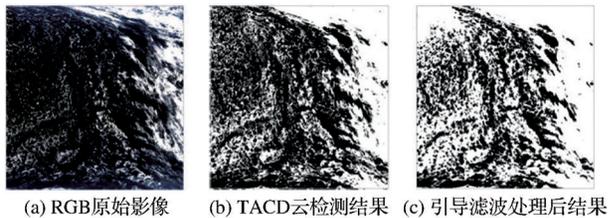


图6 基于引导滤波的云掩膜后处理示例

Fig 6 Example results of cloud mask postprocessing based on guided filtering

4 讨论

尽管实验所提出的TACD算法具有业务化应用的潜力,但在实际应用中可能仍面临以下问题。一方面,由于FY-3D影像自身的质量问题和低分辨率影像的局限性,在一定程度上限制了TACD算法的性能。FY-3D影像存在的数据质量问题包括部分数据缺失、反射率或亮度温度差异值异常、条带噪声等,使得能成功匹配的样本点数量减少,导致云检测结果产生误差,并影响精度评估。此外,由于低分辨率影像中混合像元问题,单一像元包含多于一种地类的情况较为普遍,尤其是在云边缘处和薄云覆盖区域,云的光谱特征不够显著,容易出现漏检与错检问题;另一方面,雾、霾、沙尘与云具有较为类似的特征,对于类似空间分辨率较低的光学卫星影像,难以实现其与云的准确区分,一定程度上影响最终检测结果的准确性^[48]。因此在进行云检测前可通过去噪等预处理提高影像质量,针对混合像元和雾、霾、沙尘等可借助先验数据等辅助信息排除干扰,从而改善最终的云检测结果。

5 结论

本文提出了一种激光雷达数据辅助的阈值自适应云检测方法,以FY-3D影像为例,基于在全球范围内不同区域选取的215景影像,构建了光学-LiDAR云检测数据集,进行算法的阈值自适应优化和精度验证。所提出的TACD算法充分利用低分辨率影像的多波段光谱信息,针对不同场景设计了多种阈值测试,通过迭代测试获取了全局最优阈

值,最终实现了FY-3D影像的高精度云检测。定量评价结果表明TACD算法生成的云掩膜精度明显优于官方云掩膜产品,其总体精度从官方云掩膜的84.91%和83.62%提升至86.54%。在海洋、植被和裸土区域,TACD算法经大量样本训练具有较好的表现,表现出业务化应用的潜力。在将来,TACD算法可从更精细的影像场景划分等方面改进,充分挖掘其应用潜力。除FY-3D影像外,TACD云检测算法因简单高效,未来还将拓展应用到其他低分辨率卫星影像的云检测,如MODIS、Himawari-8等。

参考文献(References):

- [1] Lin Xiaohong, Zhang Wenjuan, Fan Nengzhu, *et al.* Lightning activity in the Pre-TC squall line of Typhoon Lekima (2019) observed by FY-4A LMI and its relationship with convective evolution [J]. *Remote Sensing Technology and Application*, 2021, 36(4): 873-886. [林小红, 张文娟, 范能柱, 等. FY-4A LMI观测的利奇马(2019)台前飑线闪电活动及其与对流演变的关系[J]. *遥感技术与应用*, 2021, 36(4): 873-886.]
- [2] Wang Yazheng, Yang Yuanjian, Liu Chao, *et al.* Analysis on the applicability of Fengyun-3 satellite microwave remote sensing soil moisture products in Shandong [J]. *Chinese Journal of Agrometeorology*, 2021, 42(4): 318-329. [王雅正, 杨元建, 刘超, 等. 风云三号卫星微波遥感土壤水分产品在山东地区的适用性分析[J]. *中国农业气象*, 2021, 42(4): 318-329.]
- [3] Dou Fangli, Shang Jian, Wu Qiong, *et al.* Retrieval of cloud liquid water content over global oceans using FY-3C/3D microwave imager [J]. *Journal of Remote Sensing*, 2018, 24(6): 766-775. [窦芳丽, 商建, 吴琼, 等. 风云三号卫星被动微波反演海洋上空云液态水含量[J]. *遥感学报*, 2018, 24(6): 766-775.]
- [4] King M D, Platnick S, Menzel W P, *et al.* Spatial and temporal distribution of clouds observed by MODIS onboard the Terra and Aqua satellites [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2013, 51(7): 3826-3852. DOI: 10.1109/TGRS.2012.2227333.
- [5] Xia Deshen, Jin Sheng, Wang Jian. Fractal dimension and GGCM based meteorology cloud pictures recognition-description of texture complexity and roughness by fractal dimension (I) [J]. *Journal of Nanjing University of Science and Technology*, 1999, 23(3): 90-93. [夏德深, 金盛, 王健. 基于分数维与灰度梯度共生矩阵的气象云图识别(I)-分数维对纹理复杂度和粗糙度的描述[J]. *南京理工大学学报*, 1999, 23(3): 90-93.]
- [6] Xia Deshen, Jin Sheng, Wang Jian. Fractal dimension and GGCM based meteorology cloud pictures recognition-description of texture statistical characteristics by gray level-gradient co-occurrence Matrix (II) [J]. *Journal of Nanjing University of Science and Technology*, 1999, 23(4): 289-292. [夏德深, 金

- 盛,王健.基于分数维与灰度梯度共生矩阵的气象云图识别(II)—灰度梯度共生矩阵对纹理统计特征的描述[J].南京理工大学学报(自然科学版),1999,23(4):289-292.]
- [7] Feng Jianhui, Yang Yujing. Study of texture images extraction based on gray level co-occurrence matrix[J]. Beijing Surveying and Mapping, 2007 (3): 19-22.[冯建辉,杨玉静.基于灰度共生矩阵提取纹理特征图像的研究[J].北京测绘,2007(3):19-22.]
- [8] Sun L, Mi X, Wei J, *et al.* A cloud detection algorithm-generating method for remote sensing data at visible to short-wave infrared wavelengths[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2017, 124: 70-88. DOI: 10.1016/j.isprsjprs.2016.12.005.
- [9] Rossow W B. Measuring cloud properties from space: A review [J]. Journal of Climate, 1989, 2(3): 201-213. DOI: 10.1175/1520-0442(1989)0022.0.CO;2.
- [10] Rossow W B, Garder L C, Lacin A A. Global, seasonal cloud variations from satellite radiance measurements. Part I: Sensitivity of analysis[J]. Journal of Climate, 1989, 2(5):419-458. DOI: 10.1175/1520-0442(1989)002<0419:GSCVFS>2.0.CO;2.
- [11] Seze G, Rossow W B. Time-cumulated visible and infrared radiance histograms used as descriptors of surface and cloud variations[J]. International Journal of Remote Sensing, 1991, 12(5): 877-920. DOI: 10.1080/01431169108929702.
- [12] Rossow W B, Garder L C. Cloud detection using satellite measurements of infrared and visible radiances for ISCCP[J]. Journal of Climate, 1993, 6(12): 2341-2369. DOI: 10.1175/1520-0442(1993)006<2341:cdusmo>2.0.co;2.
- [13] Saunders R W, Kriebel K T. An improved method for detecting clear sky and cloudy radiances from AVHRR data[J]. International Journal of Remote Sensing, 1988, 9(1): 123-150. DOI: 10.1080/01431168808954841.
- [14] Kriebel K T, Saunders R W, Gesell G. Optical properties of clouds derived from fully cloudy AVHRR pixels[J]. Beiträge zur Physik der Atmosphäre, 1989, 62: 165-171.
- [15] Gesell G. An algorithm for snow and ice detection using AVHRR data an extension to the APOLLO software package [J]. International Journal of Remote Sensing, 1989, 10(4-5): 897-905. DOI: 10.1080/01431168908903929.
- [16] Ackerman S A, Strabala K I, Menzel W P, *et al.* Discriminating clear sky from clouds with MODIS[J]. Journal of Geophysical Research: Atmospheres, 1998, 103 (D24) : 3214132157. DOI: 10.1080/01431168908903929.
- [17] Frey R A, Ackerman S A, Liu Y, *et al.* Cloud detection with MODIS. Part I: Improvements in the MODIS cloud mask for collection 5[J]. Journal of Atmospheric and Oceanic Technology, 2008, 25(7): 1057-1072. DOI: 10.1175/2008jtecha1052.1.
- [18] Ackerman S A, Holz R E, Frey R, *et al.* Cloud detection with MODIS. Part II: validation[J]. Journal of Atmospheric and Oceanic Technology, 2008, 25(7): 1073-1086. DOI: 10.1175/2007JTECHA1053.1.
- [19] Ackerman S, Strabala K, Menzel P, *et al.* Discriminating clear-sky from cloud with MODIS algorithm theoretical basis document (MOD35[C])// MODIS Cloud Mask Team, Cooperative Institute for Meteorological Satellite Studies, University of Wisconsin. 2010.
- [20] Zhu Z, Woodcock C E. Object-based cloud and cloud shadow detection in Landsat imagery[J]. Remote Sensing of Environment, 2012, 118: 83-94. DOI: 10.1016/j.rse.2011.10.028.
- [21] Zhu Z, Wang S, Woodcock C E. Improvement and expansion of the Fmask algorithm: Cloud, cloud shadow, and snow detection for Landsats 4-7, 8, and Sentinel 2 images[J]. Remote Sensing of Environment, 2015, 159: 269-277. DOI: 10.1016/j.rse.2014.12.014.
- [22] Li Z, Shen H, Li H, *et al.* Multi-feature combined cloud and cloud shadow detection in GaoFen-1 wide field of view imagery[J]. Remote Sensing of Environment, 2017, 191: 342-358. DOI: 10.1016/j.rse.2017.01.026
- [23] Ishida H, Oishi Y, Moritak, *et al.* Development of a support vector machine based cloud detection method for MODIS with the adjustability to various conditions[J]. Remote Sensing of Environment, 2018, 205: 390-407. DOI: 10.1016/j.rse.2017.11.003.
- [24] Bai T, Li D, Sun K, *et al.* Cloud detection for high-resolution satellite imagery using machine learning and multi-feature fusion[J]. Remote Sensing, 2016, 8(9): 715. DOI: 10.3390/rs8090715.
- [25] Wei J, Huang W, Li Z, *et al.* Cloud detection for Landsat imagery by combining the random forest and superpixels extracted via energy-driven sampling segmentation approaches [J]. Remote Sensing of Environment, 2020, 248: 112005. DOI: 10.1016/j.rse.2020.112005.
- [26] Xu L, Niu R, Fang S, *et al.* Cloud detection based on decision tree over tibetan plateau with MODIS data[J]. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2017, 8929. DOI: 10.1117/12.2030399.
- [27] Hughes M J, Hayes D J. Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing [J]. Remote Sensing, 2014, 6(6): 4907-4926. DOI: 10.3390/rs6064907.
- [28] Sun L, Yang X, Jia S, *et al.* Satellite data cloud detection using deep learning supported by hyperspectral data[J]. International Journal of Remote Sensing, 2020, 41(4): 1349-1371. DOI: 10.1080/01431161.2019.1667548.
- [29] Xie F, Shi M, Shi Z, *et al.* Multilevel cloud detection in remote sensing images based on deep learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017, 10(8): 3631-3640. DOI: 10.1109/JSTARS.2017.2686488.
- [30] Li Z, Shen H, Cheng Q, *et al.* Deep learning based cloud detection for medium and high resolution remote sensing images of different sensors [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2019, 150: 197-212. DOI: 10.1016/j.

- isprsiprs.2019.02.017.
- [31] Mohajerani S, Saeedi P. Pragmatic augmentation algorithms for deep learning-based cloud and cloud shadow detection in remote sensing imagery [J]. *IEEE Geoscience and Remote Sensing Letters*, 2021.DOI:10.1109/LGRS.2021.3084932.
- [32] Zou Z, Li W, Shi T, *et al.* Generative adversarial training for weakly supervised cloud matting [C] // *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2019: 201-210.DOI: 10.1109/ICCV.2019.00029.
- [33] Li J, Wu Z, Hu Z, *et al.* Automatic cloud detection method based on generative adversarial networks in remote sensing images[J]. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2020, 5(2): 885-892. DOI: 10.5194/ISPRS-ANNALS-V-2-2020-885-2020.
- [34] Zhu Z, Woodcock C E. Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: An algorithm designed specifically for monitoring land cover change[J]. *Remote Sensing of Environment*, 2014, 152: 217-234. DOI: 10.1016/j.rse.2014.06.012.
- [35] Liu R, Liu Y. Generation of new cloud masks from MODIS land surface reflectance products[J]. *Remote Sensing of Environment*, 2013, 133: 21-37.DOI: 10.1016/j.rse.2013.01.019.
- [36] Chen Xidong, Zhang Xiao, Liu Liangyun, *et al.* Enhanced multi-temporal cloud detection algorithm for optical remote sensing images[J]. *Journal of Remote Sensing*, 2019, 23(2): 280-290.[陈曦东, 张肖, 刘良云, 等. 增强型多时相云检测[J]. *遥感学报*, 2019, 23(2): 280-290.]
- [37] Zhang Yonghong, Yang Chenyang, Tao Runzhe, *et al.* Multi-temporal cloud detection method for Qinghai-Tibet Plateau based with FY-4A Data[J]. *Remote Sensing Technology and Application*, 2020, 35(2): 389-398.[张永宏, 杨晨阳, 陶润喆, 等. 基于FY-4A数据的青藏高原多时相云检测方法[J]. *遥感技术与应用*, 2020, 35(2): 389-398.]
- [38] Kittaka C, Winker D M, Vaughan M A, *et al.* Intercomparison of column aerosol optical depths from CALIPSO and MODIS-Aqua[J]. *Atmospheric Measurement Techniques*, 2011, 4(2): 131-141.DOI: 10.5194/amt-4-131-2011.
- [39] Tan Z, Liu C, Ma S, *et al.* Detecting multilayer clouds from the geostationary advanced himawari imager using machine learning techniques[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, (99): 1-12. DOI: 10.1109/TGRS.2021.3087714.
- [40] Heidinger A, Foster M, Botambekov D, *et al.* Using the NASA EOS A-Train to probe the performance of the NOAA PATMOS-x cloud fraction CDR[J]. *Remote Sensing*, 2016, 8(6): 511.DOI: 10.3390/rs8060511.
- [41] Zhang Y, Guindon B, Cihlar J. An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images[J]. *Remote Sensing of Environment*, 2002, 82(2-3): 173-187.DOI: 10.1016/S0034-4257(02)00034-2.
- [42] Hall D K, Riggs G A, Salomonson V V. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data[J]. *Remote sensing of Environment*, 1995, 54(2): 127-140.DOI: 10.1016/0034-4257(95)00137-P.
- [43] Luo Y, Trishchenk A P, Khlopenkov K V. Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial resolution for the seven MODIS land bands over Canada and North America[J]. *Remote Sensing of Environment*, 2008, 112(12): 4167-4185.DOI: 10.1016/j.rse.2008.06.010.
- [44] Shang H, Chen L, Letu H, *et al.* Development of a daytime cloud and haze detection algorithm for Himawari-8 satellite measurements over central and Eastern China[J]. *Journal of Geophysical Research: Atmospheres*, 2017, 122(6): 3528-3543.DOI: 10.1002/2016JD025659.
- [45] Vaughan M A, Young S A, Winker D M, *et al.* Fully automated analysis of space-based LiDAR data: An overview of the CALIPSO retrieval algorithms and data products[C] // *Laser Radar Techniques for Atmospheric Sensing*. International Society for Optics and Photonics, 2004, 5575: 16-30.DOI: 10.1117/12.572024.
- [46] Ding Y, Hu X, He Y, *et al.* Cloud detection algorithm using advanced fully convolutional neural networks in FY3D-MERIS imagery[C] // *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*. Springer, Cham, 2020: 615-625.DOI: 10.1007/978-3-030-60633-6_51.
- [47] He K, Sun J, Tang X. Guided image filtering[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012, 35(6): 1397-1409.DOI:10.1109/TPAMI.2012.213.
- [48] Li Z, Shen H, Weng Q, Zhang Y, Dou P, Zhang L. Cloud and cloud shadow detection for optical satellite imagery: Features, algorithms, validation, and prospects[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2022, 188: 89-108.DOI: 10.1016/j.isprsiprs.2022.03.020.

Threshold Adaptive Cloud Detection for FY-3D Images Using CALIPSO Data as Reference

Zhang Yuzhuo, Li Zhiwei, Shen Huanfeng, Peng Xiaoyuan

(School of Resource and Environmental Sciences, Wuhan University, Wuhan 430079, China)

Abstract: FY series satellites can provide important data support for remote sensing monitoring of the atmosphere, land, and ocean on a global scale. As optical satellite images are inevitably affected by cloud coverage, obtaining accurate cloud masks through cloud detection is the key to the processing and application of FY series satellite images. Most of the existing cloud detection methods use simple and efficient threshold methods, however, the optimal threshold in the traditional threshold method is difficult to determine in the absence of a large number of cloud and clear sky labels due to differences in sensor spectral response and radiance differences between different underlying surfaces. Therefore, a Threshold Adapted Cloud Detection (TACD) method is proposed in this paper, which has taken the band characteristics and underlying surfaces differences into consideration comprehensively, then sets up multi-channel threshold tests consisting of reflectance and reflectance combination test, brightness temperature test, brightness temperature difference test and cirrus cloud test under different scenarios, and establish global Optical-LiDAR cloud detection dataset to achieve iteratively optimize thresholds in TACD algorithm, and finally perform cloud detection based on the optimal thresholds. We take FY-3D MERSI-II images as an example to establish a high-precision global cloud detection sample dataset collocated with CALIOP cloud layer data, compare the cloud detection results of the proposed TACD method with the official cloud mask products. The evaluation results show that the accuracy of the cloud masks produced by TACD is significantly improved compared with the official masks, in which the mIoU is increased from 80.35% to 84.09% and the recall can reach 92.67%. In conclusion, TACD has great potential for application.

Key words: Cloud detection; Threshold adaptive method; FY-3D; CALIPSO; TACD