

基于改进 LCCD 算法的高分六号 WFV 数据云检测研究

王永吉¹, 明艳芳^{1*}, 梁天辰¹, 周雪莹², 贾臣¹, 王权¹

¹ 山东科技大学测绘科学与工程学院, 山东 青岛 266590;

² 武汉大学遥感信息工程学院, 湖北 武汉 430072

摘要 高分六号 WFV 是搭载在我国高分六号卫星上的高空间分辨率多光谱传感器, 该传感器实现了高分辨率和宽覆盖的结合。精确识别高分六号 WFV 数据的云像元对于农业资源监测、林业资源调查以及防灾减灾等行业具有重要意义。基于全球土地覆盖产品 FROM-GLC10 数据, 改进地表类型支持的云检测算法 (LCCD 算法), 开展了高分六号 WFV 数据的云检测工作。以 FROM-GLC10 作为先验数据, 充分考虑不同地表类型反射率的变化, 在每种地表类型上分别采用不同的方法设置阈值。通过目视解译的方法对云检测结果进行精度评价, 云正确率整体达到了 92.46%, 其中植被类、水体类、高亮地表类的云正确率分别为 93.09%、95.60% 和 88.70%。结果表明, 改进的基于地表类型的云检测算法有效提高了高分六号 WFV 数据云检测的精度。

关键词 遥感; 云检测; FROM-GLC10; 高分六号 WFV 数据

中图分类号 P237

文献标志码 A

doi: 10.3788/AOS202040.2128001

GF-6 WFV Data Cloud Detection Based on Improved LCCD Algorithm

Wang Yongji¹, Ming Yanfang^{1*}, Liang Tianchen¹, Zhou Xueying², Jia Chen¹, Wang Quan¹

¹ College of Geomatics, Shandong University of Science and Technology, Qingdao, Shandong 266590, China;

² School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei 430072, China

Abstract GF-6 WFV is a high-spatial resolution multi-spectral sensor loaded on Chinese GF-6 satellite, which realizes the combination of high spatial resolution and wide coverage. Accurately identifying the cloud pixels of GF-6 WFV data is of great significance for supporting agricultural resources monitoring, forestry resources investigation, disaster prevention and mitigation and other industry applications. Based on the global land cover product—FROM-GLC10 (Finer Resolution Observation and Monitoring-Global Land Cover 10) data, the LCCD (Land Cover-based Cloud Detection) algorithm is improved to carry out cloud detection of GF-6 WFV data in the paper. Taking FROM-GLC10 data as a priori data, fully considering the change of reflectivity of different surface types, different methods are used to set thresholds for each surface type. The accuracy of cloud detection results was evaluated by visual interpretation, and the cloud accuracy rate as a whole reached 92.46%, among which the cloud accuracy rates of vegetation type, water type and highlighted surface type were 93.09%, 95.60% and 88.70%, respectively. The results show that the improved cloud detection algorithm based on surface type effectively improves the accuracy of cloud detection of GF-6 WFV data.

Key words remote sensing; cloud detection; FROM-GLC10; GF-6 WFV data

OCIS codes 280.4788; 100.3008; 280.1100

1 引 言

云会严重阻碍电磁波的传播, 从而在很大程度上影响了遥感信息获取的质量, 降低了数据利用率。对于定性遥感来说, 云会阻挡地表信息的表达, 使地表信息模糊甚至缺失, 导致判别失败或判别困难, 降低判别精度; 对于定量遥感来说, 云检测是反演大气

和地表参数必需的预处理工作, 云检测结果的好坏直接影响到其他参数的反演结果^[1-2]。云的种类十分丰富, 厚云在可见光波段具有高反射率, 易于检测; 而薄云由于结合了下垫面地表的信息, 与下垫面结合形成了混合像元, 往往难以实现高精度检测^[3-4]。因此, 进行快速高效的云检测是十分必要的。

收稿日期: 2020-05-14; 修回日期: 2020-06-14; 录用日期: 2020-07-15

基金项目: 国家自然科学基金(41771408)、山东省自然科学基金(ZR201702210379)

* E-mail: myf414@163.com

目前,国内外已经进行了大量的云检测工作,采用的方法主要包括统计学方法和阈值法。阈值法是应用较为广泛的云检测方法^[5-6]。针对不同的传感器,国外已经发展了多种阈值云检测算法,如 ISCCP (The International Satellite Cloud Climatology Project) 算法^[7-8]、APOLLO (AVHRR Processing Over Cloud Land and Ocean) 算法^[9-10]和 CLAVR (The NOAA Cloud Advanced Very High Resolution Radiometer) 算法^[11-12]。

前期的云检测主要是基于云与典型地表在不同波段的表观反射率或亮度温度的差异设定固定的阈值来识别有云像元。由于地表类型的复杂性以及传感器波段的局限性,通常难以用有限波段的固定阈值将云与大部分典型地表区分开,特别是高亮度背景上空的云检测,以及薄云和碎云的检测。为了提高云检测的精度,Sun 等^[13-15]提出了一种先验地表反射率数据库支持的动态阈值云检测算法 (UDTCDA),该算法以现有的 MOD09 地表反射率数据库为支撑,设定与反射率大小相关的动态阈值,改变了过去对所有像元设定同样阈值的状况。该算法在波段较少的高分辨率影像云检测中表现出了较高的精度,但该算法对数据库的要求较高,需要以短周期范围内的大量清洁图像作为支撑,前期工作量比较大,而且容易受反射率数据库中没有完全剔除的云像元的影响^[16]。为了解决以上问题,Sun 等^[17]使用地表类型支持的云检测算法 (LCCD) 来替代地表反射率支持的卫星数据云检测,他们选择 GlobeLand30 地表类型数据集作为支撑,用于支持 Landsat 8 OLI 数据的云检测;LCCD 算法充分考虑了不同地表类型反射率的变化,将地表类型分为反射率慢变型和快变型两种类型,分别采用不同的方法设定阈值;结果表明,该算法达到了较好的云识别效果。

云检测作为定量遥感关键的一步,快速高效是最基本要求,因此本文将 LCCD 算法应用于高分六号 WFV 数据云检测。高分六号 WFV 数据的空间分辨率为 16 m,空间分辨率为 30 m 的 GlobeLand30 数据与高分六号 WFV 数据的空间分辨率差异较大,加之 GlobeLand30 数据产品的时间为 2010 年,代表 2010 年的土地覆盖状况,与高分六号 WFV 数据时间差距过大,在支持高分六号 WFV 数据云检测时会产生一定误差。为降低尺度差异以及时间差异的影响,本文采用了具有 10 m 空间分辨率且产品时间为 2017 年的 FROM-GLC10 (Finer

Resolution Observation and Monitoring-Global Land Cover 10) 数据作为先验数据来支持高分六号 WFV 数据的云检测,并在每种地表类型上分别采用不同的方法设置阈值,逐像元进行云检测。

2 数据及预处理

2.1 FROM-GLC10 数据及其处理

Gong 等^[18-19]基于 2017 年获取的 10 m 分辨率的 Sentinel-2 全球影像,开发出了世界首套 10 m 分辨率的全球土地覆盖产品——FROM-GLC10。该产品提供了耕地、森林、草地、灌木、湿地、水体、苔原、不透水面、裸地和雪/冰等 10 种类型。验证实验显示,FROM-GLC10 的总体精度为 72.76%。

FROM-GLC10 数据的空间分辨率为 10 m,高分六号 WFV 数据的空间分辨率为 16 m,采用双线性内插法将 FROM-GLC10 数据重采样到 16 m 分辨率,并赋予其与高分六号 WFV 数据相同的投影坐标。

2.2 高分六号 WFV 数据及其处理

2018 年 6 月 2 日 12 时 13 分,中国在酒泉卫星发射中心用长征二号丁运载火箭成功发射了高分专项高分六号卫星。高分六号是一颗低轨光学遥感卫星,也是我国首颗精准农业观测高分卫星,具有高分辨率、宽覆盖、高质量成像、高效能成像、国产化率高等特点,设计寿命 8 年。高分六号卫星配备有 2 m 全色/8 m 多光谱高分辨率相机、16 m 多光谱中分辨率宽幅相机,2 m 全色/8 m 多光谱相机的观测幅宽为 90 km,16 m 多光谱相机的观测幅宽为 800 km。

表 1 高分六号 WFV 数据介绍

Table 1 Introduction of GF-6 WFV data

Band number	Spectral range / μm	Spatial resolution /m
B01	0.45-0.52	16
B02	0.52-0.59	16
B03	0.63-0.69	16
B04	0.77-0.89	16
B05	0.69-0.73	16
B06	0.73-0.77	16
B07	0.40-0.45	16
B08	0.59-0.63	16

首先将高分六号 WFV 数据进行正射校正处理,然后将影像的 DN (digital number) 值转化为具

有实际物理意义的辐射亮度值和表观反射率,以进行遥感的定量提取。公式为

$$L_{\lambda} = G \cdot D_N + B, \quad (1)$$

$$\rho^* = \frac{\pi \cdot L_{\lambda} \cdot D^2}{N_{\lambda} \cos \theta_s}, \quad (2)$$

式中: L_{λ} 为传感器入瞳处的光谱辐亮度; G 为增益值; B 为偏移值; D_N 为传感器接收记录的电压信号或数字值; ρ^* 为大气顶层表观反射率; D 为日地平均距离; N_{λ} 为大气顶层太阳光谱辐照度; θ_s 为太阳天顶角。

3 改进的 LCCD 算法

3.1 云检测原理

云粒子的直径(5~100 μm)远大于太阳的入射波长,可发生无选择性散射。云对各个波长的可见光及近红外散射强度是相同的,因而云在可见光以及近红外波段呈现白色,且具有较高的反射率^[20]。阈值法云检测是在可见光以及近红外波段根据云与植被、水体、不透水面等典型地物的反射率差异进行检测的。对于厚云,尤其是暗地表上方的云,传统阈值法可以达到良好的识别效果。然而薄云的云层比较薄,反射率信息通常结合了云与下垫面组成的混合像元的信息,传统的阈值法在进行薄云检测时往往得不到较好的结果^[21-22]。因此,以高精度的先验地表类型数据为支撑,在每种已知地表类型上分别设置阈值进行云检测是十分必要的。需要指出的是,本文在使用 FROM-GLC10 先验数据进行云检测时,并没有进行苔原以及雪、冰区域的云检测,这是由于受高分六号 WFV 数据波段设置的限制,波段范围主要集中在可见光、近红外波段,没有对云与雪、冰进行识别的有效波段。另外,本文研究区域为中国大陆,因此并没有进行苔原地区的云检测。

3.2 关键技术

根据 FROM-GLC10 地表覆盖数据,本文将地表类型分为时空序列地表和非时空序列地表。时空序列地表包括耕地、森林、草地和灌木,非时空序列地表则为湿地、水体、不透水面和裸地。

设置时空序列地表阈值时,根据不同的植被类型、不同的时间以及空间,依据混合像元的原理,将植被类地表看作是植被与裸土的混合像元^[23-24],按照季节与纬度分别赋予不同的植被裸土比例;非时空序列地表则无需考虑时空变化的影响,根据云与各地物单波段光谱的差异以及波段比值等特征将云

识别出来。

云通常具有较低的 NDVI (normalized difference vegetation index) 值,但在薄云与浓密植被区域形成的混合像元区域,NDVI 值可能偏高^[25-26]。NDVI 的计算公式为

$$\alpha_{\text{NDVI}} = \frac{\rho_{\text{NIR}} - \rho_{\text{R}}}{\rho_{\text{NIR}} + \rho_{\text{R}}}, \quad (3)$$

式中: α_{NDVI} 代表归一化植被指数; ρ_{NIR} 、 ρ_{R} 分别为近红外与红光波段的反射率值,它们在高分六号 WFV 数据上分别对应波段 4 和波段 3 的反射率值。

云与亮地表的区分一直是云检测的难点,云和亮地表在可见光波段都具有较高的反射率,且差异不大。亮地表在近红外波段仍具有较高的反射率,但云由于水汽的吸收作用,在近红外波段的反射率有下降的趋势。根据此特性,可将绿光、近红外比值指数(α_{GNI})作为云与亮地表区分的有效因子^[27-30]。 α_{GNI} 的计算公式为

$$\alpha_{\text{GNI}} = \frac{\rho_{\text{G}}}{\rho_{\text{NIR}}}, \quad (4)$$

式中: ρ_{G} 、 ρ_{NIR} 分别代表绿光与近红外波段的反射率值,它们在高分六号 WFV 数据上分别对应波段 2 和波段 4 的反射率值。

3.3 非时空序列地表

非时空序列地表包括水体、湿地、裸地和不透水面,其属性相对比较稳定,随时空变化较小或基本不变,因此在设置阈值时不必考虑时空变化对其反射率的影响。本文将此类地表阈值设置为固定阈值,每种地表类型分别设置不同的阈值进行云检测。

图 1(a)为水体的反射光谱曲线,水体在可见光与近红外波段的反射率都较低,随着水体中浮游植物和泥沙浓度增加,其反射率在可见光波段有明显的增大趋势,但仍显著低于云的反射率,因此选取蓝、绿、红波段进行水体上方的云识别。若下垫面为水体时的表观反射率满足(5)式,则被识别为云像元,否则为非云像元。

$$\rho_{\text{B}}^* > 0.16 \cup \rho_{\text{G}}^* > 0.13 \cup \rho_{\text{R}}^* > 0.11, \quad (5)$$

式中: ρ_{B}^* 、 ρ_{G}^* 、 ρ_{R}^* 分别代表蓝波段、绿波段和红波段的表观反射率。

图 1(b)为湿地的反射光谱曲线。湿地上分布的多为湿地水生植物,其背景为浅水,其反射率因受水的影响而在可见光与近红外波段普遍偏低,但高于水体的反射率,因此,其阈值设置方法与水体的相似。当下垫面为湿地时的表观反射率满足(6)式时,被识别为云像元,否则为非云像元。

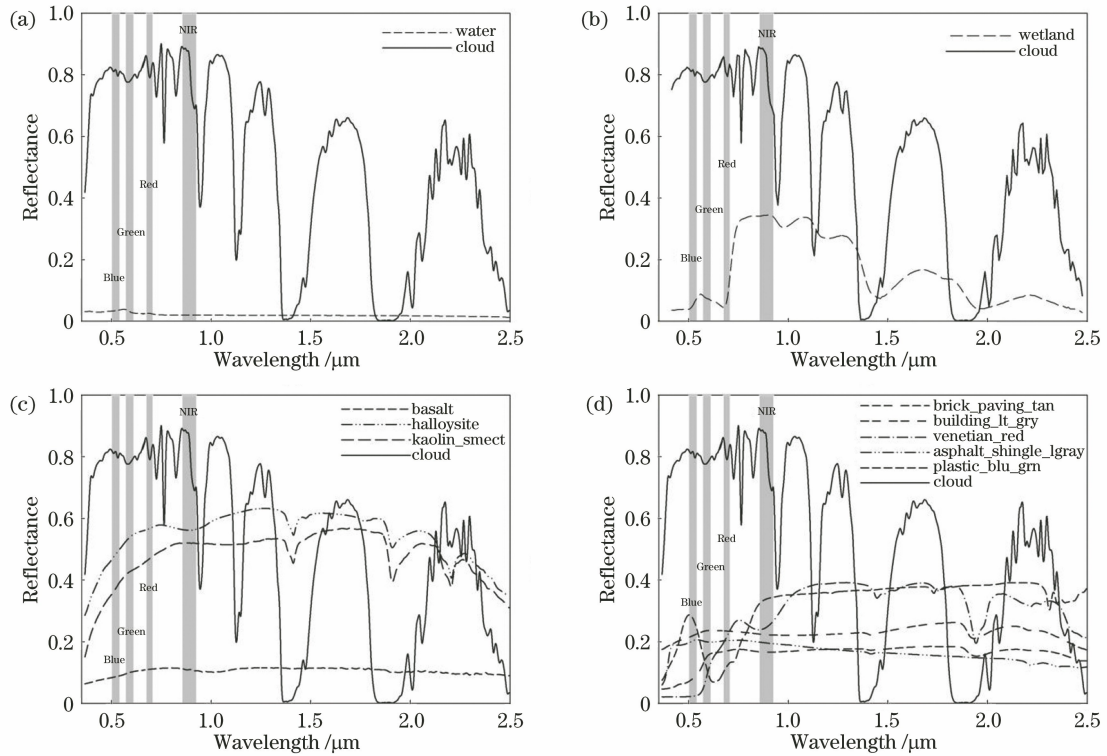


图 1 非时空序列地表的反射率特征。(a)水体;(b)湿地;(c)裸地;(d)不透水面

Fig. 1 Surface reflectance characteristics of non-spatiotemporal series. (a) Water; (b) wetland; (c) bareland; (d) impervious surface

$$\rho_B^* > 0.18 \cup \rho_G^* > 0.20 \cup \rho_R^* > 0.18. \quad (6)$$

图 1(c)为各种类型裸地的反射光谱曲线,可以看出,随着波长的增大,不同类型裸地的反射率增大,且与云的反射率差异逐渐变小。因此选用蓝、绿、红波段进行裸土上方的云识别,并选用绿光、近红外波段比值指数 α_{GNI} 以及归一化植被指数 α_{NDVI} 对薄云识别进行进一步限制。具体方法是适度降低裸地上方云检测的阈值,然后用两种指数进行高亮裸地的去除,以减少误提现象。当下垫面为裸地时的表观反射率满足(7)式时,被识别为云像元,否则为非云像元。

$$\rho_B^* > 0.17 \cup \rho_G^* > 0.17 \cup \rho_R^* > 0.22 \cap \alpha_{GNI} > 0.85 \cap \alpha_{NDVI} < 0.3. \quad (7)$$

图 1(d)为城市各种类型不透水面的反射光谱曲线,可以看出不同材质的不透水面的反射率差异比较大,但整体有一定的规律性:在可见光波段随着波长的增加而逐渐增大。在高分六号 WFV 数据上,云与不透水面的光谱差异在可见光波段最明显,因此选取蓝、绿、红波段进行云与不透水面的区分,同时选取绿光、近红外波段比值指数 α_{GNI} 和归一化植被指数 α_{NDVI} 对云检测结果进行进一步限制。当下垫面为不透水面时的表观反射率满足(8)式时,被

识别为云像元,否则为非云像元。

$$\rho_B^* > 0.22 \cup \rho_G^* > 0.27 \cup \rho_R^* > 0.32 \cap \alpha_{GNI} > 0.85 \cap \alpha_{NDVI} < 0.3. \quad (8)$$

3.4 时空序列地表

时空序列地表包括耕地、森林、草地和灌木,其属性随着时间和空间而改变,因此在阈值设置时要考虑时空的变化。除耕地外,将其他时空序列地表看作是植被和裸土的混合像元,根据不同的季节和经纬度分别赋予不同的混合像元比值,再根据各典型地物的反射光谱得到云检测阈值。

耕地的情况比较复杂,整体上耕地一般处于两种状态:耕作期和休耕期。在耕作期,耕地多被植被覆盖,反射率较低,因此红、绿、蓝波段可以区分云与耕作期的耕地;在休耕期,耕地几乎没被植被覆盖,大多表现为裸地的形式,因此采用绿光、近红外波段比值指数 α_{GNI} 和归一化植被指数 α_{NDVI} 进行裸地的去除。经过以上两步之后得到最终的云像元。当下垫面为耕地时的表观反射率满足(9)式时,被识别为云像元,否则为非云像元。

$$\rho_B^* > 0.20 \cup \rho_G^* > 0.25 \cup \rho_R^* > 0.20 \cap \alpha_{GNI} > 0.85 \cap \alpha_{NDVI} < 0.3. \quad (9)$$

图 2(a)、(b)、(c)分别为不同类型的森林、草

地、灌木的光谱曲线,可以看出:植被类时空序列地表与云的反射率差异在可见光波段最为明显,因此选取蓝、绿、红波段进行森林、草地上方的云识别;而灌木属性的不确定性较高,因此选取蓝、绿波段进行灌木上方的云检测。植被类时空序列地表在一定空间和时间范围内的长势比较稳定,因此森林、草地、灌木类的阈值设置要综合考虑时间和空间的影响:在时间上,以春(3—5月)、夏(6—8月)、秋(9—11月)、冬(12—2月)四季进行划分;在空间上,以热带、温带、寒带进行划分。

森林的分布受时空的影响比较明显:热带地区的森林主要为常绿林,植被覆盖浓密且四季变化不分明,因此假设热带地区森林与裸地的比值为0.9/0.1;寒带地区的森林主要为针叶林,叶片面积较小,也无明显的四季特征,因此假设寒带地区森林与裸地的比值为0.7/0.3;温带地区的森林主要为落叶林,有明显的四季变化特征,因此要综合考虑四季的

影响,以春季为例,假设温带地区森林与裸地的比值为0.4/0.6。

草地类植被在不同的温度带类型下差异不明显,因此不考虑草地的种类。热带地区的草地全年生长茂盛,覆盖浓密,因此假设草地与裸地的比值为0.9/0.1;在寒带地区,假设草地与裸地的比值为0.3/0.7;温带以春季为例,假设草地与裸地的比值为0.4/0.6。

灌木没有明显的主干,呈丛生状态,是比较矮小的树木,一般可分为观花、观果、观枝干等类型。如图2(c)所示,某些品种的灌木在红光波段的反射率有明显上升的趋势,因此采用蓝、绿波段进行灌木上方的云检测。热带地区的灌木生长旺盛,因此假设灌木与裸地的比值为0.9/0.1;寒带地区灌木的混合像元比应介于森林和草地之间,假设灌木与裸地的比值为0.6/0.4;温带地区以春季为例,假设灌木与裸地的比值为0.4/0.6。

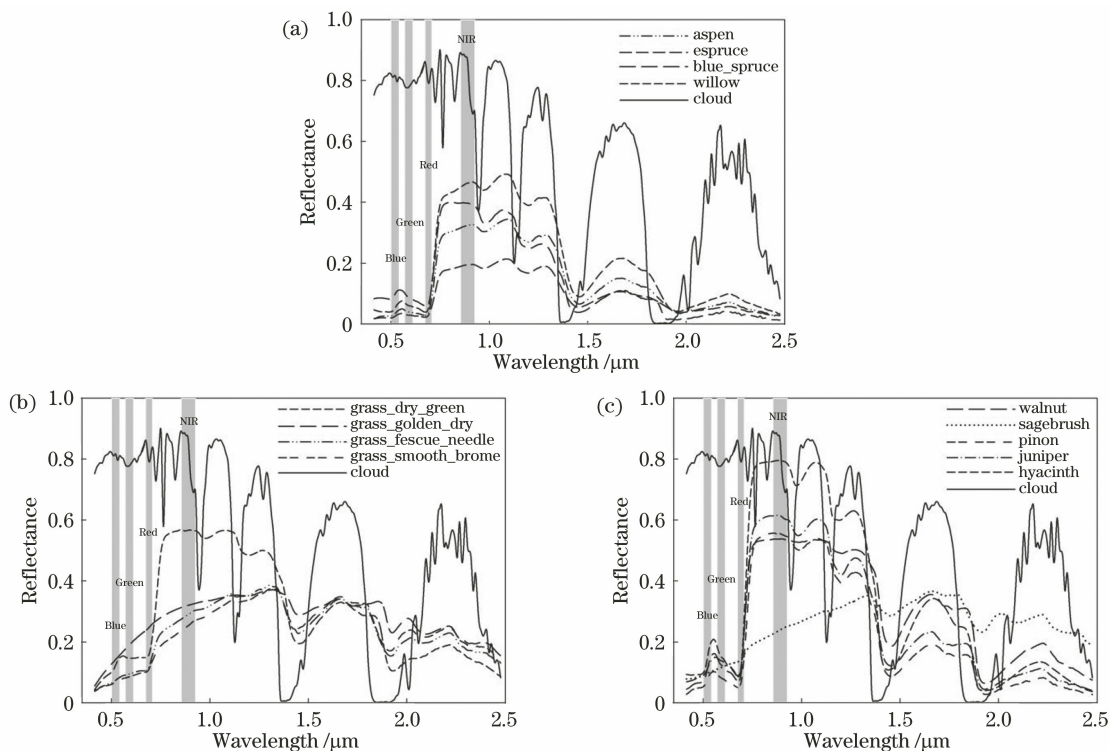


图 2 时空序列地表的反射率特征。(a)森林;(b)草地;(c)灌木

Fig. 2 Surface reflectance characteristics of spatiotemporal series. (a) Forest; (b) grassland; (c) shrubland

表 2 列出了时空序列地表在不同季节以及不同温度带设置的混合像元的植被裸地比值以及各波段阈值。

表 2 中森林、草地、灌木列为各典型植被在各个波段的反射率值,裸土列为裸土在各个波段的反射率值, T_B 、 T_G 、 T_R 分别为时空序列地表在蓝、绿、红

波段的云检测阈值, N_s/N_f 、 N_s/N_g 、 N_s/N_b 分别为假定的裸土与森林、草地、灌木混合像元的比值。当下垫面为森林和草地时的表观反射率满足(10)式时,被识别为云像元,否则为非云像元;当下垫面为灌木时的表观反射率满足(11)式时,被识别为云像元,否则为非云像元。

表 2 森林、草地、灌木的阈值
Table 2 Thresholds of forests, grasslands and shrublands

Parameter	Forest	Bareland	Tropical	Frigid zone	Temperate zone			
					Spring	Summer	Autumn	Winter
N_s/N_f			0.1/0.9	0.3/0.7	0.6/0.4	0.1/0.9	0.4/0.6	0.9/0.1
T_B	0.12	0.18	0.126	0.138	0.156	0.126	0.144	0.174
T_G	0.18	0.20	0.182	0.186	0.192	0.182	0.188	0.198
T_R	0.13	0.25	0.142	0.166	0.202	0.142	0.178	0.238

Parameter	Grassland	Bareland	Tropical	Frigid zone	Temperate zone			
					Spring	Summer	Autumn	Winter
N_s/N_g			0.1/0.9	0.7/0.3	0.6/0.4	0.1/0.9	0.4/0.6	0.9/0.1
T_B	0.20	0.18	0.198	0.186	0.188	0.198	0.192	0.182
T_G	0.23	0.20	0.227	0.209	0.212	0.227	0.218	0.203
T_R	0.30	0.25	0.295	0.265	0.270	0.295	0.280	0.255

Parameter	Shrubland	Bareland	Tropical	Frigid zone	Temperate zone			
					Spring	Summer	Autumn	Winter
N_s/N_b			0.1/0.9	0.4/0.6	0.6/0.4	0.1/0.9	0.4/0.6	0.9/0.1
T_B	0.16	0.18	0.162	0.168	0.172	0.162	0.168	0.178
T_G	0.18	0.20	0.182	0.188	0.192	0.182	0.188	0.198

$$\rho_B^* > T_B \cup \rho_G^* > T_G \cup \rho_R^* > T_R, \quad (10)$$

$$\rho_B^* > T_B \cup \rho_G^* > T_G. \quad (11)$$

3.5 云检测结果的修正

某些特殊材质的亮地表在太阳照射下会发生镜面反射,反射的太阳辐射能量几乎全部到达传感器,导致这类地物在遥感影像上具有较高的反射率,与云在可见光、近红外波段光谱的差异较小,容易造成误提现象。由于影像分辨率的关系,这些地物在遥感影像上所占的像元数量很少,在云检测结果中呈细小的零星分布,需要对其进行去除。

开运算是一个基于几何运算的滤波器,能够对图像上的零星像元进行去除,同时能够保证原始图像的总位置 and 形状不变。开运算是一个先腐蚀后膨胀的过程。本文利用开运算去除云检测结果中的零碎像元,将卷积核设置为 3×3 的矩阵,从而实现影像中零碎像元的去除。

本文所提先验地表类型库的云检测方法大大提高了云的识别精度,但当厚云边缘存在大量薄云时,薄云与下垫面尤其是亮地表区域的反射率特征极为相似,二者之间的反射率差异较小,导致这部分薄云被漏判。当薄云出现在厚云中间或高云阴影落于低云上时,容易导致云检测结果产生空洞,增大了云覆

盖范围判定的不确定性,使后续遥感影像的应用产生误差。因此,本文在上述云检测完成之后进行薄云检测结果的优化。具体方法为:遍历初步云检测结果中的所有无云像元,搜索该像元 3×3 邻域内的其他 8 个像元,并进行判断,若邻域内云像元的个数大于 5,则重新定义该像元为云像元。图 3 所示为云检测流程图。

4 实验结果与分析

4.1 目视解译结果

实验结果如图 4~6 所示。为了突出显示待检测影像中的云,对原始影像进行假彩色合成,左图为假彩色合成影像,波段组合为 432,右图为本文云检测结果的二值图,白色代表云,黑色代表非云。

图 4 为水体上方的云检测结果,三幅子区域云类型分别为厚云、碎云、薄云。图 4(a)中所示的厚云导致太阳辐射无法到达水面,云的反射率远远高于水体,因此厚云易于识别,且厚云边界细节描述较好;图 4(b)中所示的碎云分布比较分散,增加了云检测的难度,本文基于单个像元的云检测方法可以对水体上方的碎云进行较好的检测,且云边缘较为完整;图 4(c)中所示的薄云由于结合了下垫面水体

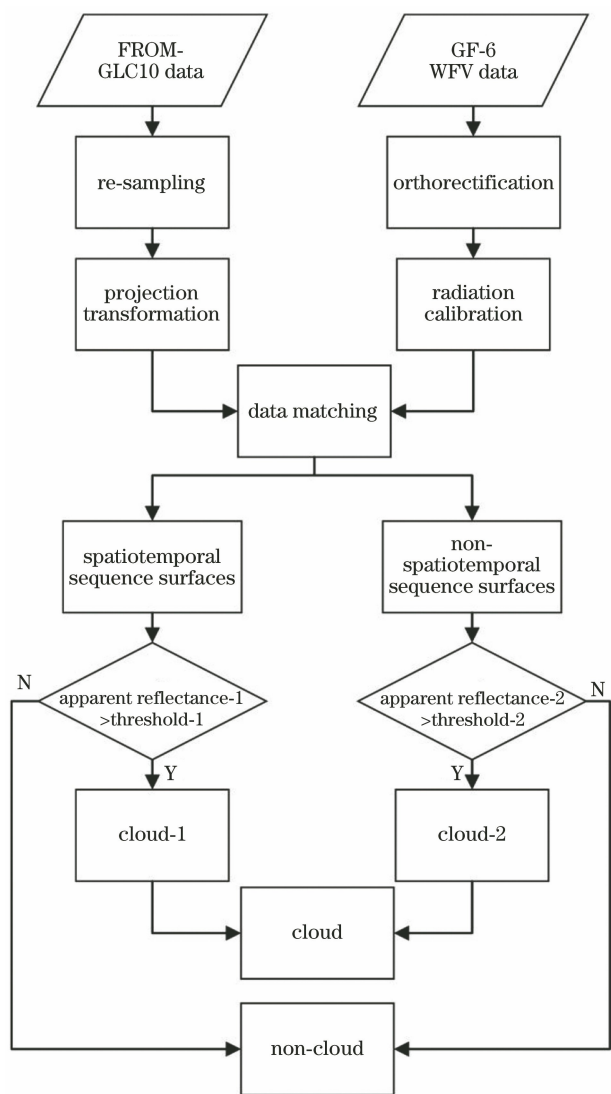


图 3 云检测流程图

Fig. 3 Flow chart of cloud detection

的信息,反射率高于水体且低于厚云,是云种类中最难检测的类型。可以看出,本文方法在准确已知地表覆盖为水体的情况下,对水体上方的薄云具有良好的检测效果,且具有较好的云边缘细节描述。

图 5 为植被上方的云检测结果,图 5(a)、(b)、(c)分别选取了不同经纬度、不同时相的影像,有厚云、薄云、碎云三种云类型。可以看出,通过混合像元原理以及不同经纬度和季节设置阈值的方法可以有效检测植被类地表上方的云,厚云识别准确并且边界完整;基于像元的云检测方法同样可以有效识别植被上方的碎云;与云相比,植被类地表的反射率整体上比较低,因此易于检测,尤其是图 5(b)左上方特别薄的云,此算法也能有效地响应。

图 6 为高亮地表上方的云检测结果,图 6(a)、(b)、(c)分别为薄云、厚云、碎云的检测结果。整

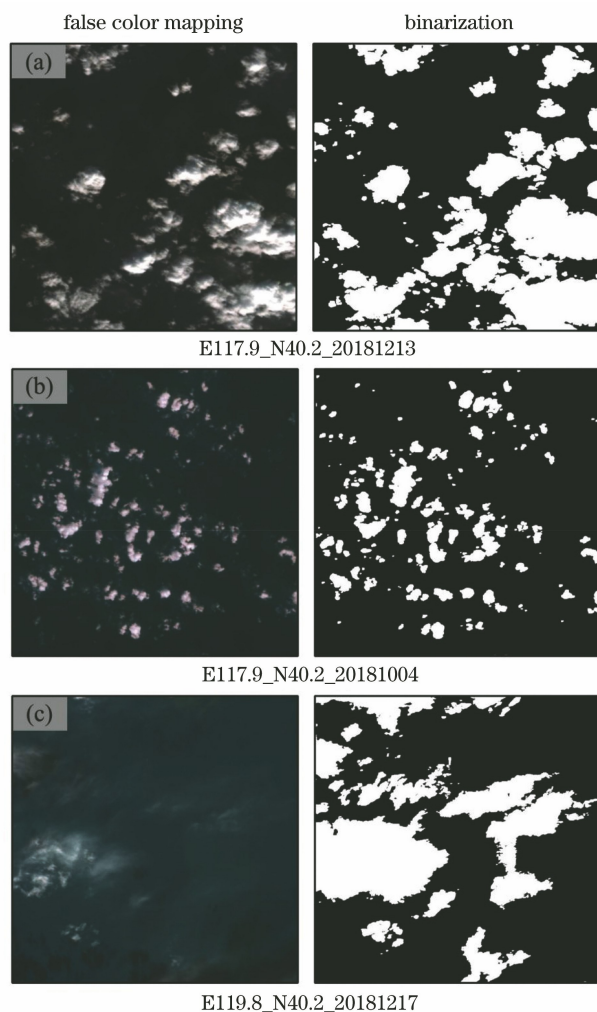


图 4 水体上方的云检测结果。(a)厚云;(b)碎云;(c)薄云
Fig. 4 Cloud detection results above water.

(a) Thick cloud; (b) broken cloud; (c) thin cloud

体上来看,高亮地表上方得到了较为理想的云检测结果,云边界也较为完整,但存在一定的误提现象,有少量的亮地表被误识别为云。由图 1(c)、(d)中的光谱曲线可以看到,亮地表与云在可见光、近红外波段的光谱特征很相似,且不同材质的亮地表的光谱特征差异较大,增大了亮地表上方云检测的难度。

LCCD算法在 Landsat8 OLI 影像亮地表区域的云检测公式为

$$\rho_B^* > 0.25 \cup \rho_G^* > 0.20 \cap T_B < 298 \cap I_{CC} < 0.95 \cup \rho_{Cirrus}^* > 0.0025, \quad (12)$$

$$I_{CC} = \frac{\rho_{Coastal}^* - \rho_{Cirrus}^*}{\rho_{Coastal}^* + \rho_{Cirrus}^*}, \quad (13)$$

式中: T_B 代表了由 Landsat8 OLI 影像热红外波段计算得到的亮温; $\rho_{Coastal}^*$ 和 ρ_{Cirrus}^* 分别代表了 Landsat8 OLI 影像 Coastal 波段以及 Cirrus 波段的

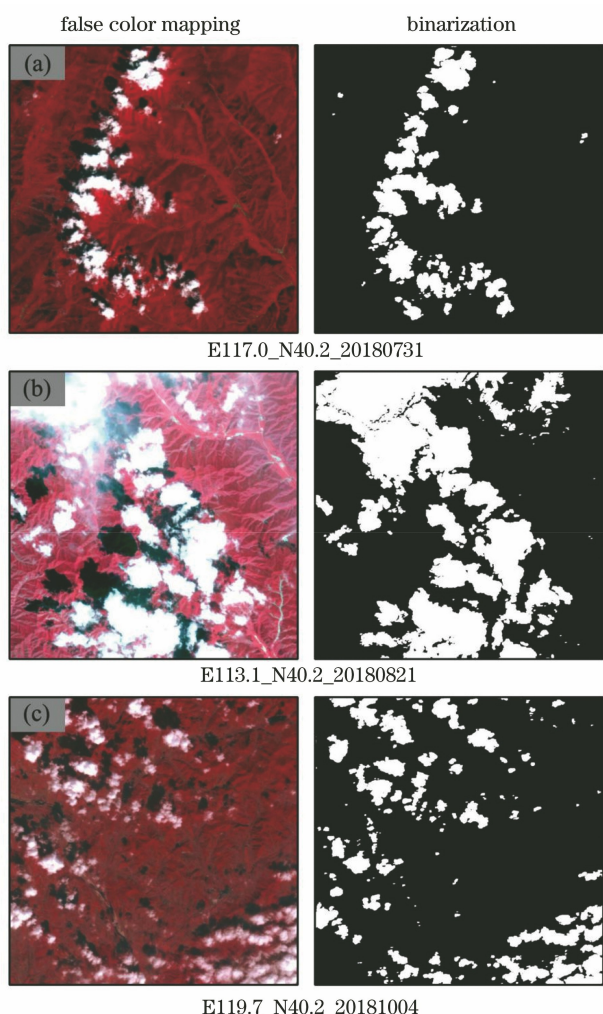


图 5 植被上方的云检测结果。(a)厚云;(b)薄云;(c)碎云
Fig. 5 Cloud detection results above vegetation.

(a) Thick cloud; (b) thin cloud; (c) broken cloud

表观反射率; I_{CC} 为深蓝-卷云波段指数, 主要用来区分卷云及薄云。

图 7 为 LCCD 算法在 Landsat8 OLI 影像上的云检测结果, 左图为假彩色合成影像, 波段组合为 543, 中间图为使用 Landsat8 OLI 热红外波段、Coastal 波段以及 Cirrus 波段的云检测结果, 右图为仅用可见光波段进行云检测的结果。图 7(a)、(b)、(c)结果显示出较为一致的规律, 在不使用热红外波段、Coastal 波段以及 Cirrus 波段的情况下, 云检测结果不理想, 误提现象普遍存在, 过多亮的地表被误识别为云, 云边界检测效果不佳。可以看出, Landsat8 OLI 热红外波段、Coastal 波段以及 Cirrus 波段在基于地表类型的云检测算法中发挥着重要作用, 这也间接解释了本文基于地表类型支持的高分六号 WFV 数据云检测方法在亮地表区域存在的误提现象。

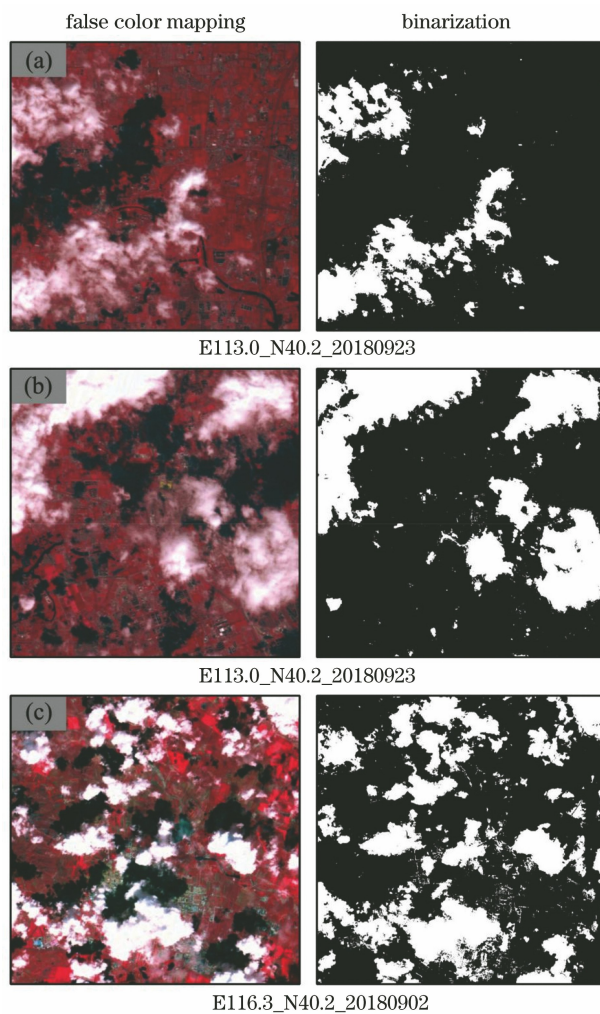


图 6 高亮地表上方的云检测结果。(a)薄云;(b)厚云;(c)碎云
Fig. 6 Cloud detection results above bright surfaces.

(a) Thin cloud; (b) thick cloud; (c) broken cloud

4.2 定量验证

为了更好地评价本文算法的适用性, 对高分六号 WFV 数据云检测结果进行了精度评价。由于高分六号 WFV 数据图幅过大, 因此本文选取有代表性的 800×800 的子区域进行精度验证, 要求子区域覆盖不同的地表类型且地表上空覆盖不同类型的云, 通过目视解译的方法对云进行矢量勾画, 并将其作为参考真值, 对云检测结果的精度进行评价。选取总正确率 (β_{CRT})、云正确率 (β_{CRC})、晴空正确率 (β_{CRS})、错判率 (β_{CER})、漏判率 (β_{OER}) 作为精度验证的指标, 这些指标的计算公式分别为

$$\beta_{CRT} = \frac{N_s + N_c}{N_{sv} + N_{cv}} \times 100\%, \quad (14)$$

$$\beta_{CRC} = \frac{N_c}{N_{cv}} \times 100\%, \quad (15)$$

$$\beta_{CRS} = \frac{N_s}{N_{sv}} \times 100\%, \quad (16)$$

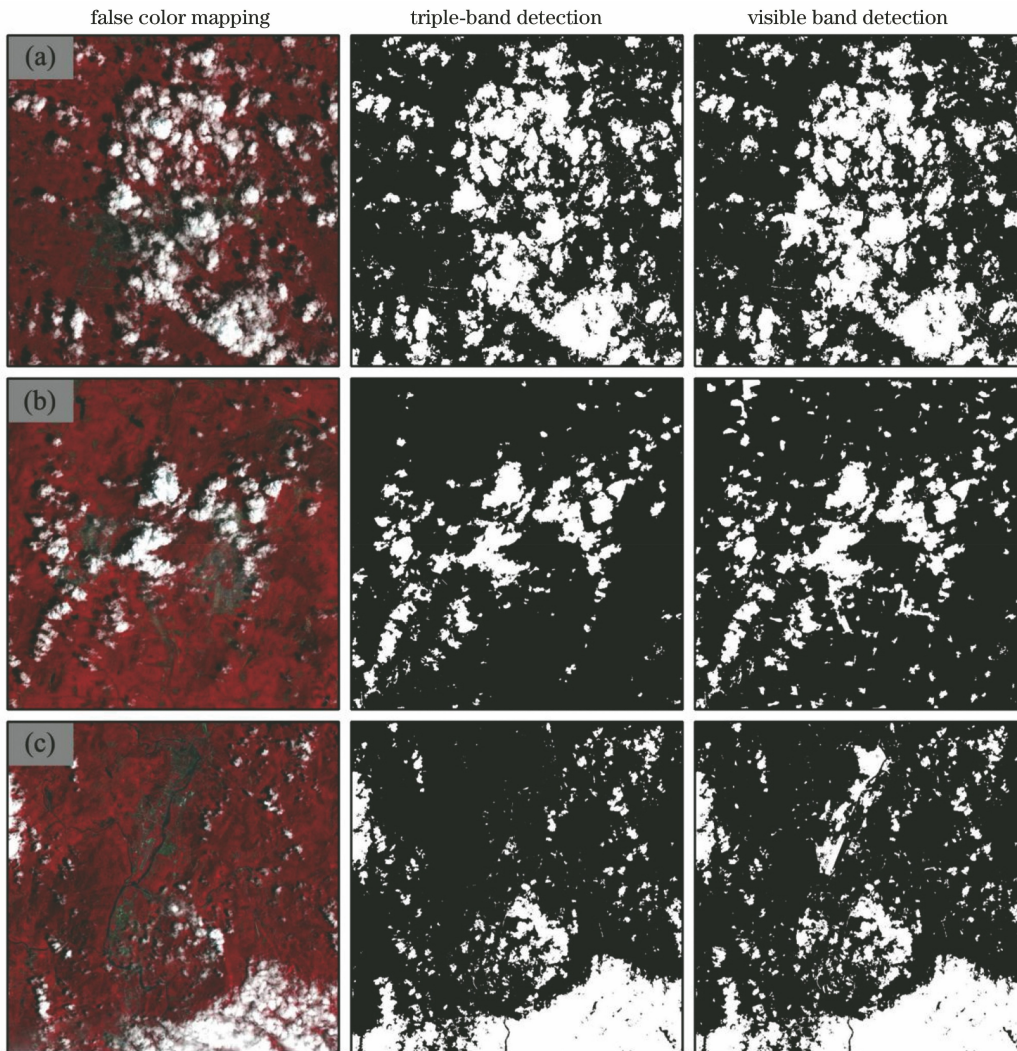


图 7 Landsat8 高亮地表上方的云检测结果。(a)碎云;(b)厚云;(c)薄云

Fig. 7 Landsat8 cloud detection results above bright surfaces. (a) Broken cloud; (b) thick cloud; (c) thin cloud

$$\beta_{\text{CER}} = \frac{N_{\text{CE}}}{N_{\text{SV}}} \times 100\%, \quad (17)$$

$$\beta_{\text{OER}} = \frac{N_{\text{OE}}}{N_{\text{CV}}} \times 100\%, \quad (18)$$

式中： N_s 代表正确检测为晴空的像元个数； N_c 代表正确检测为云的像元个数； N_{sv} 代表参考晴空像元的个数； N_{cv} 代表参考云像元的个数； N_{ce} 代表参考晴空像元检测为云像元的个数； N_{oe} 代表参考云像元检测为晴空像元的个数。

图 8 分别为各地表类型 10 个样本区域的正确率指标分布情况,可以看出,水体的三个指标分布得较为集中,植被和高亮地表的三个指标则分布得较为分散,说明水体上方的云检测效果最为稳定。植被云检测的总体正确率均在 92% 以上,水体云检测的总体正确率均在 94% 以上,高亮地表云检测的总体正确率则均在 89% 以上;植被、水体、高亮地

表的云正确率分别在 89%、92%、85% 以上;植被的晴空正确率均高于 92%,水体的晴空正确率均高于 94%,高亮地表的晴空正确率均高于 93%。水体的反射率与云的反射率具有较大差异,前者在遥感影像上为暗色调,后者为亮色调,水中泥沙等悬浮物以及叶绿素浓度的变化会对水的反射率产生影响,但整体上水与云的反射率还是存在很大差异,因此水体上方的云易于检测且检测结果相对最稳定;植被类地表由于植被种类的多样性且受不同生长期植被叶片以及花蕾等的影响,加之不同区域的土壤不同,与水体相比云检测的难度更大,云检测结果的正确率以及算法稳定程度较水体低;高亮地表主要包括裸地和不透水面,其本身与云的反射率差异不大,加之裸地土壤类型不同以及各种不透水面材质的差异导致此类地表上云检测的不确定性最高。

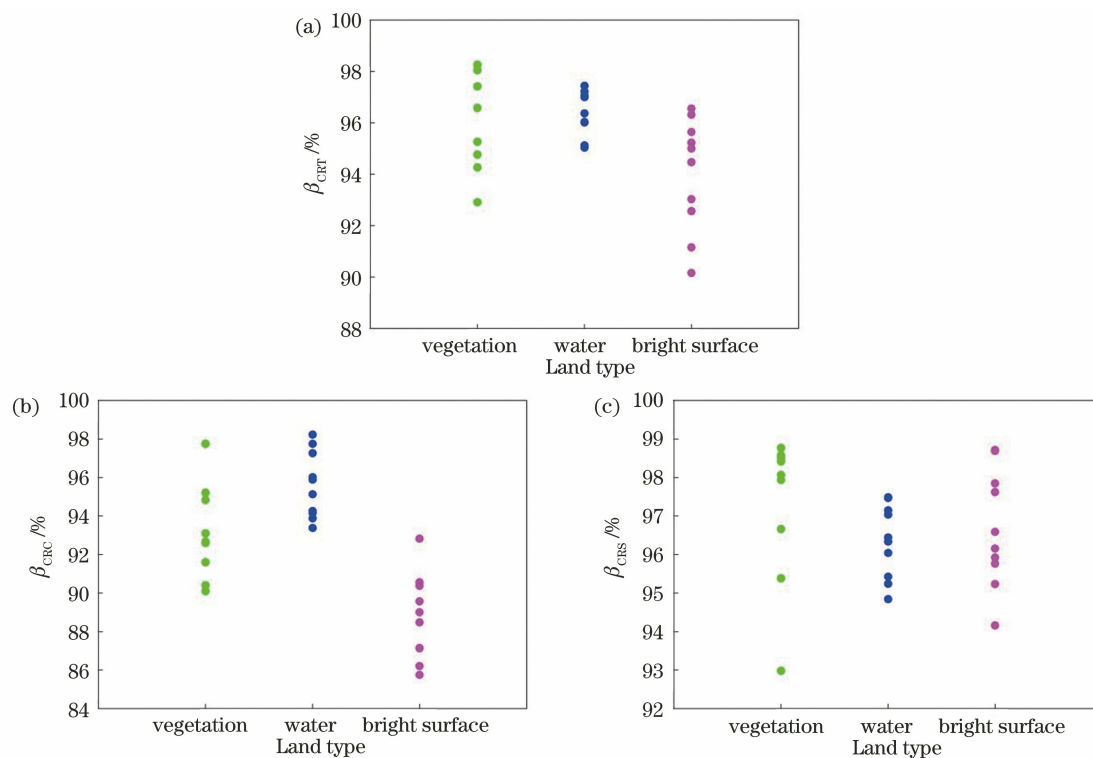


图 8 各样本的正确率指标。(a) β_{CRT} ; (b) β_{CRC} ; (c) β_{CRS}

Fig. 8 Accuracy index for each sample. (a) β_{CRT} ; (b) β_{CRC} ; (c) β_{CRS}

图 9 分别为各地表类型 10 个样本区域的错误率指标分布情况,可以看出植被的错误率分布最为分散,这是因为不同区域不同植被的反射率差异较大,如灌木类植被不同种类的花期,以及花、果实颜色会有较大差距,从而增大了云检测的不稳定性。水体的错误率相对较低,均在 5% 以下。根据漏判率指标的分布情况可以看出:水体的漏判率最低,均在 8% 以下,植被次之,均在 11% 以下,高亮地表的漏判率相对来说较高。高亮地表上方的薄云结合了下垫面的信息,与高亮地表在可见光、近红外波段的光谱差异较小,加之高分六号 WFV 数据波段设置

的限制,导致薄云难以准确检测。

云检测整体精度评价如表 3 所示。总正确率 (β_{CRT})、云正确率 (β_{CRC})、晴空正确率 (β_{CRS}) 分别为 95.60%、92.46% 和 96.80%, 错判率 (β_{CER}) 以及漏判率 (β_{OER}) 分别为 2.92% 和 7.16%。总体来讲,改进的 LCCD 云检测算法应用于高分六号 WFV 数据云检测时可以达到较高的检测精度,为后期的定量遥感研究打下了坚实的基础。

5 结 论

高分六号是我国首颗精准农业观测的高分卫

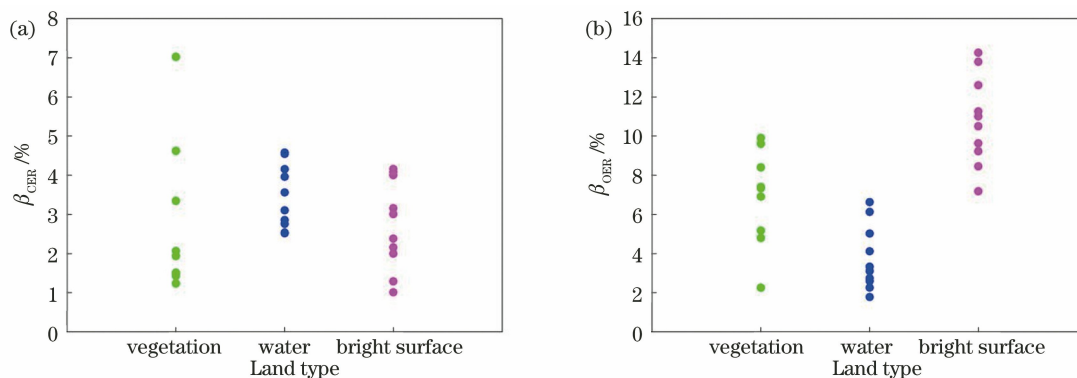


图 9 各样本的错误率指标。(a) β_{CER} ; (b) β_{OER}

Fig. 9 Error rate index of each sample. (a) β_{CER} ; (b) β_{OER}

表 3 云检测整体精度评价

Table 3 Overall accuracy evaluation of cloud detection

Accuracy	Vegetation	Water	Bright surface	Average
$\beta_{\text{CRT}}/\%$	96.36	96.43	94.01	95.60
$\beta_{\text{CRC}}/\%$	93.09	95.60	88.70	92.46
$\beta_{\text{CRS}}/\%$	97.38	96.35	96.67	96.80
$\beta_{\text{CER}}/\%$	2.59	3.46	2.72	2.92
$\beta_{\text{OER}}/\%$	6.91	3.77	10.79	7.16

星,高分六号 WFV 数据具有高空间分辨率和宽覆盖相结合的特点。精确识别高分六号 WFV 数据的云像元对于支撑农业资源监测、林业资源调查以及防灾减灾等具有重要意义。本文将高精度的地表覆盖数据——FROM-GLC10 作为先验数据,采用改进 LCCD 算法快速高效地开展高分六号 WFV 数据云检测工作。考虑到不同地表类型反射率的变化,本文将地表类型分为时空序列地表和非时空序列地表,在每种地表类型上分别采用不同的方法设置阈值。

本文结果表明,此算法在水体类和植被类地表上可以得到较为理想的云检测结果,厚云可以准确识别,薄云、碎云边界的识别也较为完整,云正确率分别达到 95.60%和 93.09%。对于高亮类地表,由于高分六号 WFV 数据波段设置的限制,薄云存在一定的漏提现象,且少量亮地表被误提为云,平均云正确率达到 88.70%。总体来讲,本文算法在高分六号 WFV 数据上得到了较为理想的云检测结果,此算法具有原理简单、操作方便快捷等特点,适合批量生产云产品,为后续的定量遥感研究提供了稳定可靠的基础。但仍存在以下不足:1)此算法依赖地表类型数据的精度,尽管 FROM-GLC10 数据整体上已达到很高的精度,但仍会在一定程度上影响云检测的精度。而且,地表类型数据定性地地将地表定义为某种类型,无过渡的概念,这与实际情况存在一定偏差,会导致不同地表类型交界处云检测效果不佳的情况发生。2)由于高分六号 WFV 数据波段设置的限制,高亮地表区域存在一定的漏提和误提现象。

参 考 文 献

- [1] Harshvardhan U, Randall D A, Corsetti T G, et al. Earth radiation budget and cloudiness simulations with a general circulation model [J]. *Journal of the Atmospheric Sciences*, 1989, 46(13): 1922-1942.
- [2] 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.
- [3] Zhang X, Liu L Y, Chen X D, et al. A novel multitemporal cloud and cloud shadow detection method using the integrated cloud Z-scores model [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12(1): 123-134.
- [4] Liu H, Du H, Zeng D, et al. Cloud detection using super pixel classification and semantic segmentation [J]. *Journal of Computer Science and Technology*, 2019, 34(3): 622-633.
- [5] Yu C H, Yuan Y, Miao M J, et al. Cloud detection method based on feature extraction in remote sensing images [J]. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2013, XL-2/W1: 173-177.
- [6] Jedlovec G J, Haines S L, LaFontaine F J. Spatial and temporal varying thresholds for cloud detection in GOES imagery [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2008, 46(6): 1705-1717.
- [7] Rossow W B, Schiffer R A. ISCCP cloud data products [J]. *Bulletin of the American Meteorological Society*, 1991, 72(1): 2-20.
- [8] Rossow W B, Schiffer R A. Advances in understanding clouds from ISCCP [J]. *Bulletin of the American Meteorological Society*, 1999, 80(11): 2261-2287.
- [9] 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.
- [10] Kriebel K T, Gesell G, Kästner M, et al. The cloud analysis tool APOLLO: improvements and validations [J]. *International Journal of Remote Sensing*, 2003, 24(12): 2389-2408.
- [11] Lovell J L, Graetz R D. Filtering pathfinder AVHRR land NDVI data for Australia [J]. *International Journal of Remote Sensing*, 2001, 22(13): 2649-2654.
- [12] Stowe L L, Davis P A, McClain E P. Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer [J]. *Journal of Atmospheric and Oceanic Technology*, 1999, 16(6): 656-681.
- [13] Sun L, Wei J, Wang J, et al. A Universal Dynamic Threshold Cloud Detection Algorithm (UDTCDA)

- supported by a prior surface reflectance database[J]. *Journal of Geophysical Research: Atmospheres*, 2016, 121(12): 7172-7196.
- [14] Sun L, Zhou X Y, Wang R L, et al. A comparison of the cloud detection results between the UDTCD mask and MOD35 cloud products [C]//2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS). July 23-28, 2017, Fort Worth, TX, USA. New York: IEEE Press, 2017: 25-28.
- [15] Chi Y L, Sun L, Wei J. Improved dynamic threshold cloud detection algorithm for Suomi-NPP visible infrared imaging radiometer[J]. *Acta Optica Sinica*, 2019, 39(5): 0528005.
迟雨蕾, 孙林, 韦晶. Suomi-NPP 卫星可见光红外成像辐射仪的改进动态阈值云检测算法[J]. *光学学报*, 2019, 39(5): 0528005.
- [16] Wang Q, Sun L, Wei J, et al. Improvement of universal dynamic threshold cloud detection algorithm and its application in high resolution satellite [J]. *Acta Optica Sinica*, 2018, 38(10): 1028002.
王权, 孙林, 韦晶, 等. 动态阈值云检测算法改进及在高分辨率卫星上的应用[J]. *光学学报*, 2018, 38(10): 1028002.
- [17] Sun L, Zhou X Y, Wei J, et al. A new cloud detection method supported by GlobeLand30 data set [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(10): 3628-3645.
- [18] Gong P, Liu H, Zhang M N, et al. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017[J]. *Science Bulletin*, 2019, 64(6): 370-373.
- [19] Gong P, Wang J, Yu L, et al. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data [J]. *International Journal of Remote Sensing*, 2013, 34(7): 2607-2654.
- [20] Sun L, Mi X T, 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.
- [21] Clark R N, Swayze G A, Wise R A, et al. USGS digital spectral library splib06a [R]. [S. l.]: US Geological Survey, 2007.
- [22] Boardman J W, Kruse F A, Green R O. Mapping target signatures via partial unmixing of AVIRIS data [C]//Summaries of JPL Airborne Earth Science Workshop. [S.l.:s.n.], 1995: 23-26.
- [23] Plaza A, Martinez P, Perez R, et al. A new approach to mixed pixel classification of hyperspectral imagery based on extended morphological profiles[J]. *Pattern Recognition*, 2004, 37(6): 1097-1116.
- [24] Miao L D, Qi H R, Szu H. A maximum entropy approach to unsupervised mixed-pixel decomposition [J]. *IEEE Transactions on Image Processing*, 2007, 16(4): 1008-1021.
- [25] Ishida H, Nakajima T Y. Development of an unbiased cloud detection algorithm for a spaceborne multispectral imager [J]. *Journal of Geophysical Research: Atmospheres*, 2009, 114(D7): D07206.
- [26] Chen P Y, Srinivasan R, Fedosejevs G, et al. An automated cloud detection method for daily NOAA-14 AVHRR data for Texas, USA [J]. *International Journal of Remote Sensing*, 2002, 23(15): 2939-2950.
- [27] Zhu Z, Woodcock C E. Object-based cloud and cloud shadow detection in Landsat imagery [J]. *Remote Sensing of Environment*, 2012, 118: 83-94.
- [28] Zhu Z, Wang S X, 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.
- [29] Qiu S, Zhu Z, Qiu S, He B, et al. Fmask 4.0: cloud and cloud shadow detection in Landsats 4-8 and Sentinel-2 [J]. *Remote Sensing of Environment*, 2019, 231(15): 111205.
- [30] Liu X Y, Sun L, Yang Y K, et al. Cloud and cloud shadow detection algorithm for Gaofen-4 satellite data [J]. *Acta Optica Sinica*, 2019, 39(1): 0128001.
刘心燕, 孙林, 杨以坤, 等. 高分四号卫星数据云和云阴影检测算法[J]. *光学学报*, 2019, 39(1): 0128001.