

# 基于陆地成像仪影像和主成分分析的水体信息提取——以鄱阳湖区为例

周小莉<sup>1,2</sup> 郭加伟<sup>2,3\*</sup> 刘锬铭<sup>2,3</sup>

<sup>1</sup>四川水利职业技术学院测绘工程系, 四川 成都 611231

<sup>2</sup>国土资源部地学空间信息技术重点实验室, 四川 成都 610059

<sup>3</sup>西南交通大学地球科学与环境工程学院, 四川 成都 611756

**摘要** 利用卫星影像快速准确地提取地球表面水体分布一直是一个重要的研究课题,其对水体灾害监测、水资源利用等具有重要意义。利用陆地成像仪(OLI)遥感影像,通过对图像进行辐射定标、大气校正获得影像的地表反射率值;然后通过分析典型水体指数的构建方法及地物的主分量空间特征,构建了一种主成分水体指数(PCWI)。以鄱阳湖区为研究区,选取枯水期两个不同时间的 OLI 影像,利用 PCWI 进行了水体提取,提取总体精度分别为 95.92%、95.52%。与其他已有的 5 种典型水体指数的水体信息提取效果对比显示,PCWI 对两期 OLI 影像水体提取的总体效果最好。综上表明,利用 PCWI 进行水体提取是可行且有效的。

**关键词** 海洋光学; 陆地成像仪影像; 主成分水体指数; 水体提取; 鄱阳湖

**中图分类号** P23 **文献标识码** A

**doi:** 10.3788/LOP53.080104

## Water Body Information Extraction Based on Operational Land Imager Imagery and Principal Component Analysis: A Case Study of Poyang Lake Region

Zhou Xiaoli<sup>1,2</sup> Guo Jiawei<sup>2,3</sup> Liu Kunming<sup>2,3</sup>

<sup>1</sup>Department of Surveying and Mapping Engineering, Sichuan Water Conservancy Vocational College, Chengdu, Sichuan 611231, China

<sup>2</sup>Key Laboratory of Geo-Spatial Information Technology, Ministry of Land and Resources, Chengdu, Sichuan 610059, China

<sup>3</sup>Faculty of Geosciences and Environmental Engineering, Southwest Jiaotong University, Chengdu, Sichuan 611756, China

**Abstract** The use of satellite images to quickly and accurately extract the distribution of the earth's surface water body has been an important research topic all the time, which has important significance for water disaster monitoring, water resource utilization, etc. The operational land imager (OLI) remote sensing imagery is used to obtain the surface reflectance values of the images by radiometric calibration and atmospheric correction. Then principal components water index (PCWI) is constructed by analyzing the construction method of the typical water index and the principal component space features of ground objects. Taking Poyang Lake as the research area, two different time OLI images are selected during the dry season to extract water body information by using PCWI. The overall accuracy is 95.92% and 95.52% respectively. Compared with the water body extraction effects of other existing five typical water index, the overall effect of water extraction based on PCWI for the two different time OLI images is the best. In conclusion, PCWI of water extraction is feasible and effective.

**Key words** oceanic optics; operational land imager imagery; principal components water index; water extraction; Poyang Lake

**OCIS codes** 010.0280; 280.4788; 300.6550

**收稿日期:** 2016-03-29; **收到修改稿日期:** 2016-05-16; **网络出版日期:** 2016-07-20

**基金项目:** 国土资源部地学空间信息技术重点实验室开放基金(KLGST2015-03)

**作者简介:** 周小莉(1982—),女,硕士,讲师,主要从事摄影测量与遥感、工程测量方面的研究。E-mail: zxlscsl@126.com

\* **通信联系人.** E-mail: guojiaweio@126.com

# 1 引言

水体是地球表面最重要的一种资源,水体信息对水资源、气候和农业研究都非常重要<sup>[1]</sup>。水体变化可能会引起洪涝灾害、水资源枯竭、水源性疾病爆发等,因此,精确迅速地实时监测水体是必要的。由于具有观测范围大、更新时间快、信息丰富等特点,遥感数据已在水资源变化监测、水质评估和监测、洪涝灾害监测和损失评估等方面得到了广泛的应用。

Landsat 数据具有较高的时空分辨率,且可以免费获取,已成为全球变化监测及地球资源调查的重要光学遥感数据源。Landsat 8 是美国地质调查局(USGS)及美国国家航空航天局(NASA)于 2013 年 2 月 11 日发射的卫星,该卫星携带了陆地成像仪(OLI)和热红外传感器(TIRS),与 Landsat 5/7 相比,OLI/TIRS 在波段设置、辐射分布率性能和扫描方式上都有很大的改进。

目前,针对 Landsat 数据进行水体提取的方法已经有大量的研究。Moller-Jensen L<sup>[2]</sup>根据专题制图仪(TM)图像的中红外和近红外波段水体光谱反射特性设定经验阈值来识别水体。文献[3-7]通过扩大水体的最强反射波段与最弱反射波段之间的差距,增强水体与背景地物之间的反差,进而实现水体的自动提取,先后构建了归一化差异水体指数(NDWI)、改进型归一化差异水体指数(MNDWI)、增强型水体指数(EWI)、新型水体指数(NWI)、自动水体提取指数(AWEI)。Kloiber 等<sup>[8]</sup>基于 TM 影像和航空影像数据利用非监督分类提取了水体范围。庞晓燕<sup>[9]</sup>根据水体在增强型专题制图仪(ETM)数据上的光谱特征,通过将 TM5 影像与(TM2+TM3)-(TM4+TM4)结果图像叠加后进行聚类分割,实现了水体的半自动提取。由于水体指数法简单易行,效果佳,因此已成为最常用的水体识别方法。由于 Landsat 8 OLI 数据具有 9 个波谱段,地物波谱空间复杂,且数据存在一定的冗余,因此会对水体提取精度造成影响。主成分分析(PCA)能够有效压缩数据量,起到降维、去噪的效果,便于地物特征提取。

本文以 Landsat 8 为数据源,针对其波段设置特点,分析典型地类在波段空间的波谱特征及在主分量空间的分量特征,结合现有典型水体指数模型的构建方法,构建了一种新型水体指数—主分量水体指数(PCWI),选取鄱阳湖两个不同时期的 OLI 数据进行水体信息提取,并与利用其他指数提取的水体信息进行了精度对比。

## 2 研究区域及数据预处理

### 2.1 研究区概况

鄱阳湖是我国第一大淡水湖,位于江西省北部,长江中下游南岸,即北纬 28°22'~29°45',东经 115°47'~116°45'。由于受赣江、抚河、信江、饶河、修水(简称“五河”)等江河的影响,鄱阳湖的水位、边界范围常年变化显著,尤其在“五河”与鄱阳湖交汇处,形成了大量湖滩、沼泽和沙洲。一般 4~6 月为“五河”主汛期,10 月至翌年 3 月是鄱阳湖的枯水期,汛期时湖水浩渺,枯水时草滩毕露。以枯水期的鄱阳湖区为研究对象,因为在枯水期,水体不连通,地物分布复杂,对探究水体信息提取方法具有一定的代表性。

### 2.2 数据来源及预处理

选用 Landsat 8 OLI 多波段影像,OLI 传感器波段设置如表 1 所示。与 ETM 传感器数据相比较,OLI 传感器数据的近红外波段 5 的光谱范围为 0.85~0.88  $\mu\text{m}$ ,排除了 0.825  $\mu\text{m}$  处水汽吸收的影响;Band8 波段范围变窄,从而可以更好地区分植被和非植被。新增加的 Band1 深蓝色波段主要应用于海岸带观测,Band9 短波红外应用于云检测。

表 1 Landsat 8 OLI 传感器波段设置  
Table 1 Landsat 8 OLI sensor band settings

Band	Wavelength/ $\mu\text{m}$	Spatial resolution/m
1-aerosol	0.43~0.45	30
2-blue	0.45~0.51	30
3-green	0.53~0.59	30
4-red	0.64~0.67	30
5-near infrared	0.85~0.88	30

续表 1

Band	Wavelength/ $\mu\text{m}$	Spatial resolution/m
6-mid infrared	1.57~1.65	30
7-mid infrared	2.11~2.29	30
8-panchromatic	0.50~0.68	15
9-cirrus	1.36~1.38	30

使用不同时间的 OLI 影像数据是在 USGS 网站上免费获取的,条带号为 121-40,时相 1 成像时间为 2013 年 10 月 18 日,时相 2 成像时间为 2016 年 2 月 16 日。包括整个鄱阳湖区的实验假彩色影像如图 1 所示。影像包含水体、建筑用地、植被、阴影、裸地等地物。

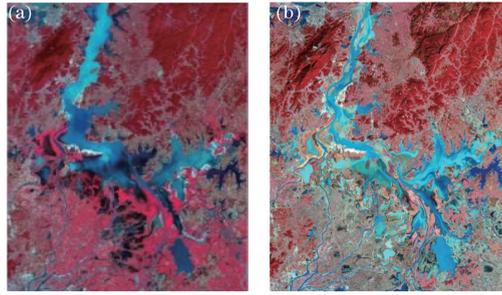


图 1 研究区不同时相假彩色合成影像。(a)时相 1; (b)时相 2

Fig. 1 False color composite images of study area for different phases. (a) Phase 1; (b) phase 2

由于获取的 L1T 数据产品已经经过系统辐射校正和几何校正,因此本研究对数据的预处理主要包括辐射定标和大气校正两个部分。辐射定标是根据数据头文件中给出的波段辐射重定义因子将影像的像元亮度 (DN) 值转换为辐射亮度。大气校正正是为了消除大气的吸收和散射的影响,将辐射亮度转换为地表的反射率,本研究选用 6S 模型对影像进行大气校正。6S 校正前后的光谱曲线对比如图 2 所示。

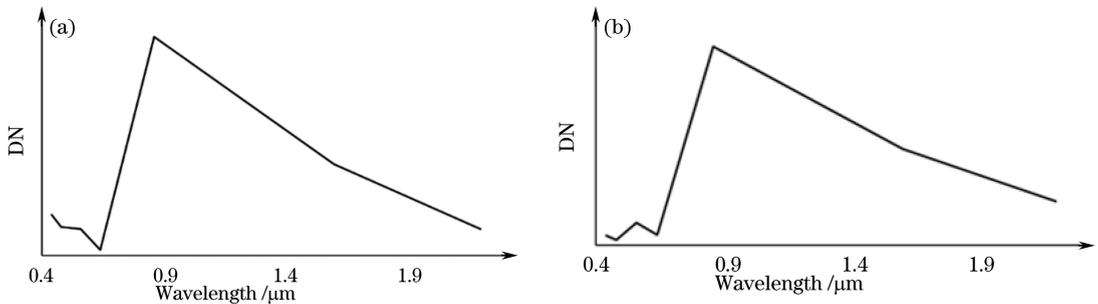


图 2 大气校正前后光谱曲线对比。(a)校正前; (b)校正后

Fig. 2 Comparison of spectral curves before and after atmospheric correction. (a) Before correction; (b) after correction

### 3 水体信息提取

基于遥感的水体信息提取方法主要是以水体在电磁波谱的各波段上反映的特征为基础。由于不同水体的水面性质、水体中悬浮物的性质和含量、水深和水底特性等不同,遥感传感器所接收到的反射光谱特征存在一定的差异。对于一般清澈的水体,在可见光范围内,水体的反射率总体上比较低,并随着波长的增大逐渐降低,在近红外和短波红外波段,水体的反射率显著衰减,接近零。而植被、土壤、城市在近红外波段反射率相对较高,尤其是植被,在近红外波段形成一个反射峰。因此,水体的光谱递减规律以及水体与其他地物在波段上光谱特征的差异就成为水体自动提取方法的重要理论基础。

#### 3.1 光谱特征分析

为了分析水体与不同背景地物在 Landsat 8 OLI 图像中的光谱特征差异,在研究区内选择 6 种典型地物,包括水体、建筑用地、植被、阴影、裸土和湖滩,然后均匀选择各类别地物样本点,统计出不同地物在 OLI 7 个波段(除去全色和卷云波段)的反射率,绘制的波谱特征曲线如图 3 所示。

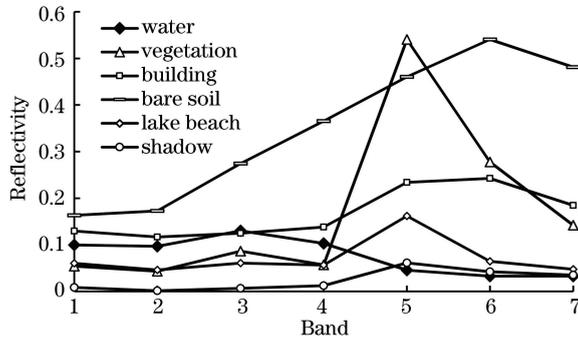


图3 典型地物波谱特征曲线

Fig. 3 Spectral characteristic curves of typical ground objects

由图3可以看出,水体在可见光波段的反射率明显大于近红外波段及短波红外波段,且在近红外波段及短波红外波段的反射率明显低于其他地物。这一特性显著且是其他地物所不具备的。

### 3.2 典型水体指数

#### 1) 归一化差分植被指数(NDVI)

$$I_{\text{NDVI}} = \frac{\rho_5 - \rho_4}{\rho_3 + \rho_4}, \quad (1)$$

式中  $I_{\text{NDVI}}$  为归一化差分植被指数,  $\rho_5$ 、 $\rho_4$  分别为近红外波段5和红光波段4的反射率值。由于在红光波段水体的反射率高于植被,而近红外波段植被的反射率明显高于水体,因此采用NDVI来处理可增强水陆反差。

#### 2) NDWI

$$I_{\text{NDWI}} = \frac{\rho_3 - \rho_5}{\rho_3 + \rho_5}, \quad (2)$$

式中  $I_{\text{NDWI}}$  为归一化差分水体指数,  $\rho_3$ 、 $\rho_5$  分别代表绿光波段3和近红外波段5的反射率。由于水体反射从可见光到红外波段逐渐减弱,在近红外和短波红外波长范围内吸收性最强,而植被在近红外波段反射率最强,因此,用绿光波段和近红外波段的反差构成NDWI,可以最大程度地抑制植被信息,突出水体特征,从而提取水体。

#### 3) MNDWI

$$I_{\text{MNDWI}} = \frac{\rho_3 - \rho_6}{\rho_3 + \rho_6}, \quad (3)$$

式中  $I_{\text{MNDWI}}$  为改进型归一化差分水体指数,  $\rho_3$ 、 $\rho_6$  分别代表绿光波段3和短波红外波段6的反射率。由于建筑物的反射率从近红外波段到短波红外波段骤然转强,因此采用短波红外波段替换近红外波段计算出来的建筑物的指数值将明显减小;另外,由于水体在短波红外波段的反射率持续降低,因此,替换后得出的指数值将会增大,使得水体与建筑物的反差明显增强,有利于城镇中水体信息的准确提取。

#### 4) AWEI

$$I_{\text{AWEI1}} = 4 \times (\rho_3 - \rho_6) - (0.25 \times \rho_5 + 2.75 \times \rho_7),$$

$$I_{\text{AWEI2}} = \rho_2 + 2.5 \times \rho_3 - 1.5 \times (\rho_5 + \rho_6) - 0.25 \times \rho_7, \quad (4)$$

式中  $I_{\text{AWEI1}}$  为无阴影覆盖区域的自动水体提取指数( $I_{\text{AWEI1}_{\text{nsh}}}$ ),  $I_{\text{AWEI2}}$  为有阴影覆盖区域的自动水体提取指数( $I_{\text{AWEI2}_{\text{sh}}}$ ),  $\rho_2$ 、 $\rho_3$ 、 $\rho_5$ 、 $\rho_6$ 、 $\rho_7$  分别为蓝光波段2、绿光波段3、近红外波段5、短波红外波段6和短波红外波段7的反射率。利用水体在蓝波段和绿波段的强反射性,以及近红外波段和短波红外波段的强吸收性,通过精确测定典型地表覆盖类型的光谱反射值,确定不同波段的权重系数,以差值区分水体与非水体信息。

上述4种典型水体指数的一个共同点为:利用水体和典型背景地物在多个波谱段的反射差异,通过扩大这种差异,增强目标地物信息,抑制背景地物的信息,从而达到突出目标地物的目的。

### 3.3 新型水体指数的构建

PCA是一种重要的图像增强方法,通过主成分变换,可以减少冗余信息,保留影像主要信息。对时相1

的 OLI 数据 1~8 波段进行主成分变换后得到 8 个主分量的特征值及贡献率如表 2 所示。

表 2 各主分量的特征值和贡献率

Table 2 Eigenvalues and contribution rate of each principal component

Principal component	Characteristic value	Contribution rate /%	Cumulative contribution /%
P1	1204.28	70.48	70.48
P2	398.65	23.33	93.81
P3	85.87	5.03	98.84
P4	11.76	0.69	99.53
P5	3.35	0.19	99.72
P6	2.68	0.16	99.88
P7	1.82	0.11	99.99
P8	0.22	0.01	100.00

由表 2 可以看出,影像绝大部分的地物信息主要集中在前三个主分量,其累计贡献率可达 98.84%。后面 5 个主分量的累积贡献率仅为 1.16%,且大多数是由原始波谱的噪声引起的。选用前三个主分量进行典型地物波谱特征分析,得到的波谱曲线如图 4 所示。从图 4 可以看到,地物在第一个主分量上具有最好的可分离性,而在第三个主分量上可分离性最差。第三个主分量仅为 5.03%,地物可分离性又极差,可以认为第三主分量对地物分类没有帮助。仔细观察地物在前两个主分量的波谱特征曲线,可以看到,水体的波谱曲线斜率为负,而其他非水体的波谱曲线斜率均为正,且水体在第一个主分量上的波谱值最低。根据这一特点,利用典型水体指数的构建思路,构建新型水体指数,进行水体分离,称为 PCWI,其计算公式为

$$I_{\text{PCWI}} = \frac{P_2 - P_1}{P_2 + P_1}, \quad (5)$$

式中  $I_{\text{PCWI}}$  为主成分水体指数,  $P_1$  和  $P_2$  分别是第一个主分量的光谱值和第二个主分量的光谱值。

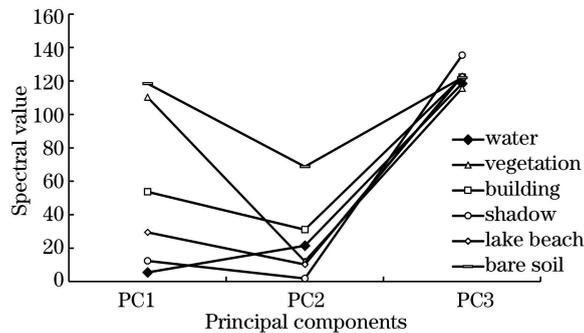


图 4 典型地物在前三个主分量上的光谱特征

Fig. 4 Spectral characteristics of typical ground objects in the first 3 principal components

## 4 结果分析与精度评价

### 4.1 水体提取

对选取的两个时相的 OLI 影像进行辐射定标、大气校正及主成分变换后,利用 ENVI 5.0 软件构建新型水体指数 PCWI 模型进行水体提取。得到的水体指数灰度图像具有典型的双峰直方图,通过最大类间方差法(Otsu)进行阈值最优化自动选择及阈值分割,即可得到水体信息。

两个时相 OLI 影像的 PCWI 水体提取结果如图 5 所示,为验证 PCWI 的有效性,利用多种典型水体指数对研究区进行了水体提取实验,得到的水体提取效果比较如图 6 所示。

### 4.2 精度评价

精度评价是遥感监测中重要的一步,也是检验图像信息提取方法是否可靠的一种有效手段。可以通过精度评价来确定所使用的模型方法是否有效,并根据精度评价的结果进一步改进所构建的模型方法,提高信息提取精度。本研究的水体信息提取,可以认为是将所有地物划分成水体和非水体两大类,利用图像分类评

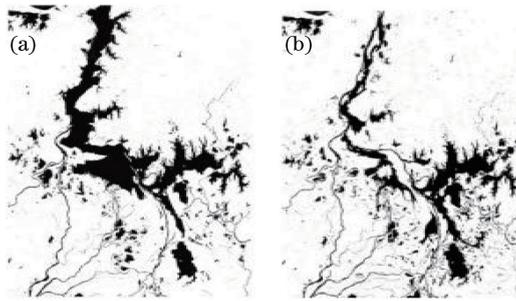


图 5 不同时相的 PCWI 水体提取结果。(a)时相 1; (b)时相 2

Fig. 5 Water extraction results based on PCWI for different phases. (a) Phase 1; (b) phase 2

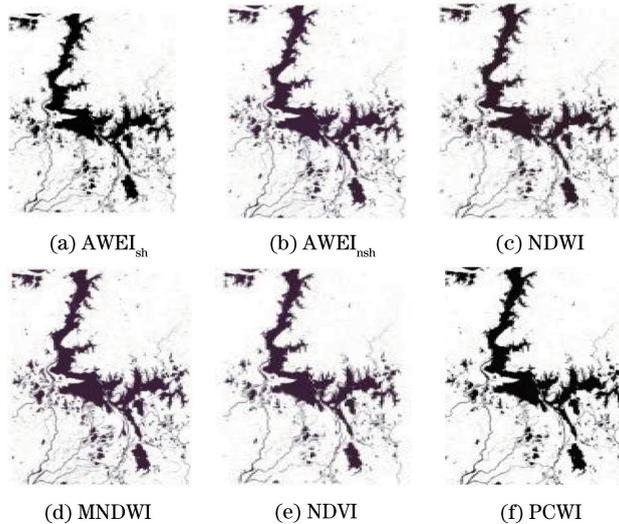


图 6 时相 1 各种水体指数水体提取结果

Fig. 6 Water extraction results of each water index for phase 1

价的方法进行评价。图像分类精度评价是把分类结果图像与标准数据或地面实测值进行比较,以正确分类的百分比来确定分类的准确程度。目前,通常是建立混淆矩阵(或者称为误差矩阵)来计算各种分类精度指标,如生产者精度、用户精度、Kappa 系数和总精度。

在 Landsat 8 OLI 假彩色(5、4、3 波段)合成图像上,水体呈现蓝黑色,植被为鲜红或浅红色,火烧地呈现深黑色,可以通过目视解译获取地面水体与非水体的真实数据。本研究在实验区内随机均匀选取多个水体与非水体参考数据,通过与相应分类结果建立误差矩阵,利用总体分类精度、Kappa 系数、水体漏分误差、水体错分误差定量评价各种水体指数模型提取水体的精度。鄱阳湖地区两个不同时相各水体指数提取水体的精度评价结果如表 3 所示。

表 3 各水体指数提取水体的精度评价

Table 3 Water extraction accuracy evaluation of each water index

Index	Overall accuracy / %		Kappa coefficient		Water misclassification error / %		Water leakage error / %	
	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2	Phase 1	Phase 2
NDVI	94.10	93.16	0.82	0.80	2.37	2.64	25.57	24.73
NDWI	94.86	94.88	0.84	0.84	18.29	19.34	6.35	6.01
MNDWI	75.54	80.57	0.54	0.60	38.34	38.11	0.76	0.62
AWEI <sub>sh</sub>	95.92	94.63	0.92	0.90	6.98	7.06	1.54	1.62
AWEI <sub>nsh</sub>	87.39	89.04	0.75	0.75	23.87	23.16	1.25	1.33
PCWI	95.68	95.25	0.91	0.90	6.56	7.05	3.94	3.28

#### 4.3 结果分析

从不同时相的 PCWI 水体提取结果(图 5)可以清晰地看到,构建的新型水体指数 PCWI 可以成功地将

研究区的水体信息提取出来,提取结果与假彩色影像目视解译的结果基本一致,时相 1 影像水体提取总体精度可达 95.68%,时相 2 影像水体提取总体精度可达 95.25%。图 6 列出了几种水体指数模型对时相 1 影像的提取结果,可以直观地看到这些经典指数都可以成功地提取大片水体,但是从局部对比(图 7)可以看到,提取效果还是存在一定的差异。结合表 3,不难看出 6 种指数提取水体的总体精度由大到小的顺序依次为:PCWI、 $AWEI_{sh}$ 、NDWI、NDVI、 $AWEI_{nsh}$ 、MNDWI。

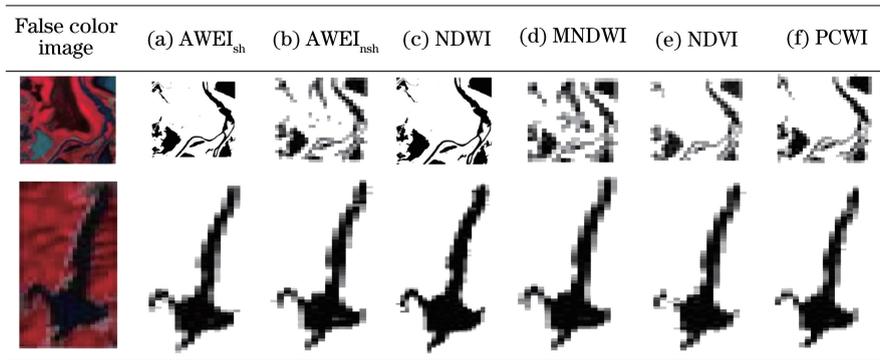


图 7 各种水体指数提取水体结果局部对比

Fig. 7 Local comparison of water extraction results of each water index

实验采用的两个时相的 OLI 影像均为研究区枯水季节的影像。湖泊水位下降,出现了含水量高的湖滩,有些地方还长出了植被,利用短波红外进行水体识别会增大错分概率,实验中两个时相 MNDWI 指数提取水体的错分误差分别为 38.34%、38.11%。虽然 NDVI 能够识别水体,也是较早提出的识别水体的指数,但是存在较多的信息丢失问题,两个时相的水体漏分误差分别达到 25.57%、24.73%。NDWI、 $AWEI_{nsh}$  水体指数受山体阴影和建筑物阴影的影响较大。 $AWEI_{sh}$  的总体精度较高,但是城市区域的一些强反射屋顶被识别为水体。而所提出的新型水体指数 PCWI,只使用了影像的两个主分量,在一定程度上降低了实验运算量和复杂度,且均保持较高的水体提取精度,两期数据的水体提取 Kappa 系数分别达到 0.91 和 0.90,能够更加有效地识别水体。

## 5 结 论

以枯水期的鄱阳湖区为研究区,通过分析典型地物在 Landsat 8 OLI 数据的波谱特征及水体信息识别的典型指数的构建思路,提出一种利用 PCA 主成分分量进行水体识别的指数 PCWI。通过对两个不同时期的鄱阳湖区的 OLI 影像进行 PCWI 水体提取,验证了该指数的可行性。然后与多种典型水体指数在水体信息识别时的效果进行对比,总体上所提出的新型水体指数 PCWI 精度最高,与  $AWEI_{sh}$  精度相当,明显比 NDWI、MNDWI、NDVI、 $AWEI_{nsh}$  各指数的水体提取精度高。新型水体指数 PCWI 利用水体在主分量空间与其他地物的差异来提取水体,影像数据经主成分变换后,大量有用信息都集中在前三个主分量,便于特征分析,减少了实验运算的数据量和复杂度,便于水体信息的快速、准确提取。

## 参 考 文 献

- 1 Roberts N, Taieb M, Barker P, *et al.* Timing of the younger dryas event in East Africa from lake-level changes [J]. Nature International Weekly Journal of Science, 1993, 366(6451): 146-149.
  - 2 Moller-Jensen L. Knowledge-based classification of an urban area using texture and context information in Landsat-TM imagery [J]. Photogrammetric Engineering and Remote Sensing, 1990, 56(6): 475-479.
  - 3 McFeeters S K. The use of normalized difference water index (NDWI) in the delineation of open water features [J]. International Journal of Remote Sensing, 1996, 17(7): 1425-1432.
  - 4 Xu Hanqiu. A study on information extraction of water body with the modified normalized difference water index (MNDWI) [J]. Journal of Remote Sensing, 2005, 9(5): 589-595.
- 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究 [J]. 遥感学报, 2005, 9(5): 589-595.

- 5 Yan Pei, Zhang Youjing, Zhang Yuan. A study of information extraction of water system in semi-arid regions with the enhanced water (EWI) and GIS based noise remove techniques[J]. Remote Sensing Information, 2007, (6): 62-67.  
闫 霏, 张友静, 张 元. 利用增强型水体指数(EWI)和 GIS 去噪音技术提取半干旱地区水系信息的研究[J]. 遥感信息, 2007, (6): 62-67.
- 6 Ding Feng. A new method for fast information extraction of water bodies using remotely sensed data[J]. Remote Sensing Technology and Application, 2009, 24(2): 167-171.  
丁 凤. 一种基于遥感数据快速提取水体信息的新方法[J]. 遥感技术与应用, 2009, 24(2): 167-171.
- 7 Feyisa G L, Meilby H, Fensholt R, *et al.* Automated water extraction index: a new technique for surface water mapping using Landsat imagery[J]. Remote Sensing of Environment, 2014, 140(1): 23-35.
- 8 Kloiber S M, Brezonik P L, Bauer M E. Application of Landsat imagery to regional-scale assessments of lake clarity[J]. Water Research, 2002, 36(17): 4330-4340.
- 9 Pang Xiaoyan. Application of satellite remote sense technique to investigation of forest resources of Daxinanling mountains in Inner Mongolia[J]. Inner Mongolia Forestry Investigation and Design, 2009, 32(3): 59-60.  
庞晓燕. 卫星遥感技术在内蒙古森林资源调查中的应用[J]. 内蒙古林业调查设计, 2009, 32(3): 59-60.