Comparative Analysis of Automatic Water Identification Method Based on Multispectral Remote Sensing

ZHANG Fang-fang¹,a, ZHANG Bing²,b, LI Jun-sheng, SHEN Qian, Wu Yuan-feng, SONG Yang

¹ Key Laboratory of Geographical Information Science, Ministry of Education, East China Normal University, Shanghai, 200062, China
² Center for Earth Observation and Digital Earth Chinese Academy of Sciences, Beijing, 100094, China
*zhffang6666@163.com,*zb@ceode.ac.cn

Abstract

In this paper, from the water spectral reflectance characteristics, I select five multi-spectral remote sensing water information extraction methods, issue comparative analysis and applicability assessment in five typical experiment areas based on Landsat data. I have made a comprehensive comparison and algorithm applicability analysis on the five kinds of methods. The main conclusions are the following: MNDWI and The spectral relationship method between multi-band have the most widely applicability, MNDWI always use in the building shadow and bare land. The spectral relationship method between multi-band has the special effects on the mountain shadow areas. NDVI has an apparent superiority in the dense forest and grassland, the low-cost of single-band threshold make it have more advantages in the great plains areas, compromise and low noise of NDWI can be used in complicated areas.

1. Introduction

Water bodies identified of Multispectral remote sensing have seven main methods: single-band threshold, vegetation index, water index, improved water index, Multi-band Spectral Relationship, supervision and decision tree classification method. However, most studies have only discussed the superiority of one approach or a variety of methods in a single application of the study area [1] ~ [2], there is no dedicated to a variety of methods in comparative studies of different experimental areas, so this article is for the purpose of a comprehensive comparative analysis of this water information on a variety of remote sensing method for automatic identification verified by experimental analysis of the advantages and
applicability of different methods to explore the most suitable water body identification method of the different natural conditions in different study areas.

2. Recognition Theory and Methods of water

2.1. Spectral reflectance characteristics

1) Spectral reflectance characteristics of water: The main band multi-spectral remote sensing is the visible, near infrared and shortwave infrared bands, in these bands range the water bodies, vegetation, towns, and shadow all has its own reflective properties. Clean ocean water generally lower reflectivity, coastal, lakes and other more turbid water because it contains chlorophyll, non-pigmented particles, reflecting higher rates than clean waters, in the visible spectral characteristics of chlorophyll a certain extent, near-infrared back reflectance dip due to the strong absorption of pure water[3].

2) TM / ETM+ Spectral Characteristics: Landsat TM images have seven bands, water identification common bands are the second to fifth, the spectrum characterized is: TM2 is 0.52 ~ 0.60μm, the green band, this band is near the green reflectance peak, lush plants for health reflex sensitivity. TM3 is 0.63 ~ 0.69μm, the red band, this band is located in the main absorption band of chlorophyll. TM4 is 0.76 ~ 0.90μm, is the near-infrared band, the strong water absorption in the region. TM5 is 1.55 ~ 1.75μm, is the short-wave infrared band, this band is located between two water absorption bands.

2.2. Methods and Algorithms

1) Single-band threshold: Single-band threshold method is used to distinguish different surface features by set a certain threshold to a single band of remote sensing image. The fifth band of TM images is the best band of water identified. This article’s Experiment is use the fifth band to segmentation, less than a given value of the pixel is extracted as water, but mixed with a certain shadow information.

2) Vegetation Index: This paper adopts Normalized Difference Vegetation Index (NDVI):

\[ \text{NDVI} = \frac{(\text{NIR} - \text{RED})}{(\text{NIR} + \text{RED})} \]

NDVI is short for Normalized Difference Vegetation Index, NIR stands for near infrared wave band that is the fourth band of TM, RED stands for red band which is the third band of TM. The principle of the paper is that the vegetation has low reflectance in red band while reflectance is high in near infrared band. However, water has higher reflectance in red band than that of near infrared band. Both of them take on remarkable difference on spectral-reflectance curve, thus while we highlight vegetation, water information also stands out.

3) Water Index: This paper adopts Normalized Difference Water Index (NDWI) to identify water information, the algorithm is:

\[ \text{NDWI} = \frac{(\text{GREEN} - \text{NIR})}{(\text{GREEN} + \text{NIR})} \]

NDWI is short for Normalized Difference Water Index, Green stands for green band, it is the second band of TM, NIR is near infrared band which is the fourth band of TM. The principle is that water has high reflectance in green band while it is low in near infrared band. Through normalizing procession the difference becomes broad, thus water information is highlighted [4].

4) Modified Water Index: Xu has implemented improvement to water index who uses short infrared wave band(the fifth band of TM)to substitute near infrared band to improve Normalized Difference Water Index, that is:

\[ \text{MNDWI} = \frac{(\text{GREEN} - \text{MIR})}{(\text{GREEN} + \text{MIR})} \]
MNDWI is short of Modified Water Index, GREEN is green band (the second band of TM), MIR is short infrared wave band (the fifth band of TM). Water has low reflectance in the fifth band of TM, thus using the fifth band to substitute the fourth band can highlight water information well [5].

5) Multi-band Spectral Relationship: According to the rule that water reflectance decreases from the second band to the fifth band, we design an algorithm:

\[ \text{TM2} + \text{TM3} > \text{TM4} + \text{TM5} \]  

The principle is that water reflectance in green and red band is higher than that of near infrared band and short infrared wave band, after addition, reflectance of the second and the third band is higher than that of the fourth and the fifth band.

3. Water Recognition experiment

3.1. Study area and data

1) Study area: The paper selected five typical experimental area of water body identification to verify the advantages and disadvantages and the adaptability of different geographical conditions. Here are a few basic overview of the study area (Table I).

2) Data: In this paper, the data source is Landsat TM/ETM+. The following table (Table II) is part of the metadata of the data source used:

<table>
<thead>
<tr>
<th>Data item</th>
<th>Dali</th>
<th>Shanghai</th>
<th>Shilianghe Reservoir</th>
<th>Taihang Mountains</th>
<th>Yellow River</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satellite</td>
<td>Landsat-5</td>
<td>Landsat-7</td>
<td>Landsat-5</td>
<td>Landsat-5</td>
<td>Landsat-5</td>
</tr>
<tr>
<td>Sensor</td>
<td>TM</td>
<td>ETM+</td>
<td>TM</td>
<td>TM</td>
<td>TM</td>
</tr>
<tr>
<td>Ranks NO.</td>
<td>131, 042</td>
<td>118, 038</td>
<td>121, 036</td>
<td>124, 035</td>
<td>123, 036</td>
</tr>
</tbody>
</table>

3.2. Research programs

Three parts of water identification processing: first, data preprocessing, the raw remote sensing data for radiometric correction, geometric correction, atmospheric correction, reflectance data by surface features; and then the water body identification, first the band calculation, then the calculated area index products for the water mask, K-Mean Method for unsupervised classification, with the final visual interpretation methods of water screen tracking vector automatic identification of information for water verify the accuracy and precision evaluation algorithm design.

3.3. Water recognition experiment in study area

Based on the above design flow of water identified five research areas of the five methods were used to water body information extraction experiment, the 20 band math data and the fifth band reflectivity data (Figure I), recognition results of water limited space plans are not listed.

4. Accuracy test and analysis

4.1. Accuracy test

In order to accurately and objectively evaluate each method’s advantages of different study area we design a integrated precision evaluation algorithm. The algorithm thoughts are as follows: (1) selected a study area to calculate the overall accuracy, Kappa coefficient, mapping accuracy, user accuracy of each
method in the study area; (2) Four accuracy Evaluation indicators of five methods were sorted by numerical of size as the five levels one to five; (3) Use inverse method to determine the weight of each data value level that the 1, 1/2, 1/3, 1/4, 1/5; (4) Determine the four accuracy evaluation indicators’ weight value, the Kappa coefficient and overall accuracy are the best indicators reflects the classification accuracy, giving Kappa coefficient of the maximum weight of 1, the overall accuracy of the weight of 0.8, mapping accuracy and user accuracy weights are 0.5; (5) According the data-level weights and the accuracy evaluation indicators weight to sum, get each method’s integrated precision $P_i$, as the following formula calculated:

$$P_i = p_t \times 0.8 + p_{Kappa} \times 1 + p_m \times 0.5 + p_u \times 0.5$$  \hspace{1cm} (5)

Where, $P_i$ refers to a water body identification integrated precision, $i$ were single (single-band threshold method), plant (vegetation index), water (water index), change (modified water index), more (multi-band spectral Relations); $P_t$ refers to the overall accuracy of total weight; $P_{Kappa}$ refers to the Kappa coefficient; $P_m$ system refers to the mapping accuracy of weights; $P_u$ refers to the user accuracy weights. (6) To compare the $P_{single}$, $P_{vegetation}$, $P_{water}$, $P_{modify}$, $P_{multi-band}$ than the size of the maximum value that is identified for this study area is the best way to water bodies’ recognition. Accordance with the above method in the study area were calculated on the five comprehensive accuracy were compared for each study area are the best water recognition (Table III).

4.2. The comparative Analysis to the applicability of the water identification methods

1) Mountain areas: Dali and the Taihang mountain are typical of mountain landscape, the ground rugged, cliff lined gully in the mountains, in remote sensing images have many shadows and the shadow of mountains and water’s spectral reflectance characteristics are very similar, so it is difficult to distinguish. By the integrated precision table (Table III) shows, Dali and Taihang Mountains’ multi-band spectral Relationship have reached 2.4 in all methods is the most high accurate, and the heaviest weight on accuracy evaluation index Kappa coefficient reached 0.9844 and 0.7039, it is also the maximum, so the multi-band spectral Relationship is the best way to remove the shadow of mountains for mountain area to extract water.

2) Dense vegetation areas: Vegetation index method’s four accuracy indices are second better method only less to multi-band spectral relationship, and the vegetation index and multi-band spectral Relationship’s various accuracy indicators in Dali is minimal in difference. This is because Dali is in subtropical and tropical edge location, year round warm and humid, mountain vegetation Special lush, dense vegetation to a certain extent reduce the mountain shadows’ intensity in remote sensing images, thereby reduce the confusion between mountain shadows and water bodies. So use of vegetation index in the dense forest can also get very good recognition accuracy.

3) City areas: Shanghai is the typical urban test area, the study area is skyscrapers, roads, occupy the vast majority of artificial ground. As many tall buildings, there is much building shadow in remote sensing images, and building shadows and water bodies are very similar in spectral reflectance characteristics, so it is very difficult for water body identification. Modified water index’s integrated precision reach to 2.8 in this paper five methods selected, and the overall accuracy, Kappa coefficient, mapping accuracy, user accuracy are all the best in five, so the Modified water index is good for identifying water bodies in urban areas.

4) River sand areas: There are many flood lands without plants, most of which are exposed sands. The exposed sandy earth and man-made earth are similar very much in spectral reflectance characteristics. The sands in TM 2, 5 wave bands have high reflection. The water has higher reflection in 2 wave band but lower in 5 wave band, thus modified water index can highlight water itself through inhibiting sand information. The 3 methods in modified water index are the best in the 4 accurate tested indexes. The integrated precision of this method reaches up to 2.55, so modified water index method is the best one in identifying different waters in river sand areas.
5) Plains and lakes areas: Lianyungang reservoir is a typical plain lake area, which is the best representative of most countries in plain. The advanced value of the four accuracy estimated indexes distribute scatter, but modified water index method is the best one, no matter in comprehensive accurate indexes or in monomial accurate indexes. The difference among overall accuracy, Kappa index, mapping accuracy and user accuracy is much little in the area researched, meanwhile, as overall accuracy, the minimum of in this area is the maximum in all the areas researched. So as to plain lake areas, no matter which kind of method we choose, it has competitively high accuracy. However, the advantage of cost and speed of single band threshold rule is the best than any other methods.

5. Conclusion and discussion

We can get a conclusion that there aren’t any universal identified water methods. The highest accuracy is modified water index and multi-band spectral relationship in the 5 researched methods in this experiment. But it doesn’t mean this 2 methods are the best. Because each method has its own advantages and unique scope of applications, users should choose a best one according to their needs and the conditions of areas researched.

The research of this article set an example for water identified methods of multi wave bands, but there are still some problems and disadvantages, which are mainly as follows. Firstly, the numbers of areas researched are small, just 5 areas, which can’t represent all conditions of all land. Secondly, the conclusion of this article needs to be tested and improved in more real applications for perfect.

Acknowledgment

This work has been jointly supported by China major projects on control and rectification of water body pollution (Grant No. 2009ZX07527-006-5), National natural science foundation of China (Grant No. 40901174) and (Grant No. 41001205).

References


[6] TABLE II. OVERVIEW OF THE STUDY AREA

<table>
<thead>
<tr>
<th>Study area</th>
<th>Dali</th>
<th>Shanghai</th>
<th>Shilianghe Reservoir</th>
<th>Yellow River</th>
<th>Taihang Mountains</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position</td>
<td>26N, 100E</td>
<td>31N, 121E</td>
<td>35N, 119E</td>
<td>35N, 115E</td>
<td>36N, 114E</td>
</tr>
<tr>
<td>Max elevation</td>
<td>4295</td>
<td>——</td>
<td>50</td>
<td>——</td>
<td>1675</td>
</tr>
<tr>
<td>Min elevation</td>
<td>730</td>
<td>——</td>
<td>20</td>
<td>——</td>
<td>200</td>
</tr>
<tr>
<td>Fall</td>
<td>3500</td>
<td>——</td>
<td>30</td>
<td>——</td>
<td>1475</td>
</tr>
<tr>
<td>Ground feature</td>
<td>Divide large, steep mountains, dense vegetation</td>
<td>Skyscrapers</td>
<td>Terrain horizon, many rivers and lakes</td>
<td>The majority of the river beach area</td>
<td>Gully horizon, the cliffs, sparse vegetation</td>
</tr>
<tr>
<td>Regional</td>
<td>Mountain(Dense)</td>
<td>City</td>
<td>flatland</td>
<td>River beach</td>
<td>Mountain(Sparse)</td>
</tr>
</tbody>
</table>
### TABLE III. INTEGRATED PRECISION

<table>
<thead>
<tr>
<th>Research method</th>
<th>Single-band threshold</th>
<th>Vegetation index</th>
<th>Water index</th>
<th>Modified water index</th>
<th>Multi-band spectral relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dali</td>
<td>0.625</td>
<td>1.4</td>
<td>0.884</td>
<td>1.075</td>
<td>2.4</td>
</tr>
<tr>
<td>Shanghai</td>
<td>0.625</td>
<td>0.7</td>
<td>0.859</td>
<td>2.8</td>
<td>1.4</td>
</tr>
<tr>
<td>Shilianghe Reservoir</td>
<td>0.71</td>
<td>0.8</td>
<td>0.884</td>
<td>2.425</td>
<td>1.965</td>
</tr>
<tr>
<td>Yellow river</td>
<td>0.71</td>
<td>0.715</td>
<td>0.884</td>
<td>2.55</td>
<td>1.525</td>
</tr>
<tr>
<td>Taihang mountain</td>
<td>0.585</td>
<td>1.4</td>
<td>1.075</td>
<td>0.924</td>
<td>2.4</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>3.255</strong></td>
<td><strong>5.015</strong></td>
<td><strong>4.586</strong></td>
<td><strong>9.774</strong></td>
<td><strong>9.69</strong></td>
</tr>
</tbody>
</table>

Figure 1. Index products

(From left to right is the first 5-band image, vegetation index, water index, modified water index, Multi-band Spectral Relationship; top to bottom as Dali, Shanghai, Shiliang Reservoir, Yellow River Lankao section, the Taihang Mountains)