Research Article

A Study on Coastline Extraction and Its Trend Based on Remote Sensing Image Data Mining

Yun Zhang,1,2,3 Xueming Li,1 Jianli Zhang,2,3 and Derui Song2,3

1 Liaoning Normal University, Dalian 116029, China
2 National Marine Environmental Monitoring Center, Dalian 116023, China
3 Key Laboratory of Sea Areas Management Technology, SOA, Dalian 116023, China

Correspondence should be addressed to Yun Zhang; cloud208@163.com

Received 31 January 2013; Accepted 21 March 2013

Academic Editor: Jianhong (Cecilia) Xia

Copyright © 2013 Yun Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, data mining theory is applied to carry out the field of the pretreatment of remote sensing images. These results show that it is an effective method for carrying out the pretreatment of low-precision remote sensing images by multisource image matching algorithm with SIFT operator, geometric correction on satellite images at scarce control points, and other techniques; the result of the coastline extracted by the edge detection method based on a chromatic aberration Canny operator has a height coincident with the actual measured result; we found that the coastline length of China is predicted to increase in the future by using the grey prediction method, with the total length reaching up to 19,471,983 m by 2015.

1. Introduction

The coastline is an adjacent transition zone between land and ocean that is the result of dynamic change processes involving natural changes and human activities. At present, 60% of the world’s population is gathered near a coastal zone rich in metallic mineral resources, oil and gas resources, tidal and wave energy sources, and other renewable energy sources. The coastline is an important place for human economic and social activities. A quick and accurate determination of the coastline length variation trend is not only a necessary technical activity for the study of land-ocean interactions, coastal reclamation, port development, and urban expansion, but also an important subject for marine economics and marine multidisciplinary research.

There are two coastline extraction methods. One is the field detection method, which is the more common method, that is, the photogrammetric technology and GPS technology. The other is the remote sensing image coastline automatic extraction technology and image interpretation. Such software and technologies for coastline extraction have gradually become more mature, as discussed by Li and Jigang, Lee et al., Manavalan et al., Holman et al., Rudin et al., Chan et al., Donoho, Xiaofeng et al., and Chaoyang [1–9]. The automatic remote sensing image classification have been well researched, such as the proposal by Jiang et al. for an automatic scheme for the classification of land use based on change detection and a semisupervised classifier [10]. Stavrakoudis et al. [11] built a boosted genetic fuzzy classifier for land cover classification of remote sensing imagery. Although the coastline can be extracted by these two methods, they fail to include a systematic and comprehensive analysis of coastline length and type variation trend, which means that the automatic sensing image classifications do not meet the shoreline classification. The spatial data mining theory presented in this study is aimed at the image processing of remote sensing image data and is focused on the discovery of potential, hidden, and useful models and rules between the image targets from remote sensing images as discussed by Xiaocheng and Xiaoqin [12]. It combines the spatial data mining method as discussed by Dengke, Deren et al., Zhanquan, Kaichang et al., and Longshu et al. [13–17] with classic remote sensing image processing technology as discussed by Aihua et al., Deren and Juliang, Guobao et al., and Tao and Bin [18–21]. A new coastline extraction and analysis model was put forward to analyze...
the future variation trend of coastline length by means of grey prediction algorithm to achieve the resulting conclusions and to propose feasible recommendations for the exploitation of the future coastal zone.

2. **Coastline Extraction Process**

Data mining refers to the process as discussed by Kantardzic et al. [22], in which implicit, unknown, potential, and useful information and knowledge can be extracted from a large number of incomplete, noisy, fuzzy, and random databases. According to the knowledge structure of data mining, the coastline extraction processes can be divided into four stages as follows in Figure 1: image selection; image registration, correction, toning, mosaic, and other pretreatment processes; coastline extraction and mining; and the coastline result analysis and trend prediction.

3. **Study on the Remote Sensing Image Coastline Extraction Technology**

3.1. **Data Preparation.** The remote sensing image data presented and selected in this paper were collected at five stages from 2007 to 2011. The space range covered the coastal zones of the Chinese Mainland (including Hong Kong and Macao) and Hainan Island encompassing 290 scenic spots. Details of the images are as in Table 1.

3.2. **Data Preprocessing**

3.2.1. **Automatic Registration of the Coastal-Zone Image.** The multisource image matching algorithm of a SIFT operator was presented in this paper, which provided different spectral band setting, imaging model, reflectivity, dimension, and other aspects of the multisource data. RANSAC and the identical point triangulation network construction method were used for eliminating the mismatching points in order to achieve the automatic matching of multisource remote sensing images and control point image databases, after which the automatic image registration was completed.

3.2.2. **Geometric Correction of the Coastal-Zone Image.** It was difficult to find the control points due to the large water area on the remote sensing image at the coastal zone. Meanwhile, a small number of control points were used to achieve the geometric correction of the low-precision remote sensing images in order to reduce the workload of the field measurements. Combined with satellite orbit parameters and the geometric correction model with correction precision in different satellite data sources and terrain conditions, the geometric correction of multiscene images was jointly carried out to achieve the geometric correction of satellite images at several control points in the use of the correlation between subsence overlapping regions.

3.2.3. **Automatic Toning and Mosaic of Coastal-Zone Image.** Image data toning processing was carried out by using the color-transferring algorithm to maintain the consistency of typical surface features. In the fuzzy classification method, the overlapping region calculation method and image multiscale segmentation method were used for the study on the color transferring algorithm to maintain the consistency of typical surface features. The coastal-zone remote sensing image was spliced by using the automatic matching-line generation image mosaic technology.

3.3. **Data Mining**

3.3.1. **Coastline Extraction Based on Edge Detection.** Edge detection methods were widely used for the automatic extraction of the water sideline so as to significantly reduce the quantity of data, remove the irrelevant information, and retain the important structural properties of the images. The methods consisted of Roberts Cross operator, Prewitt operator, Sobel operator, Kirsch operator, and Compass operator, in which the Canny operator can be regarded as the most common edge detection method. In order to compensate for the insufficient remote sensing image detection of complex surface features, edge features, or broken linear characteristics by the conventional Canny
operator, the chromatic aberration Canny operator was used instead to extract the coastline.

(1) Calculation of the Chromatic Aberration Amplitude by Means of Chromatic Aberration Canny Operator. In order to improve on deficiencies of the traditional Canny operator, an improved Canny operator was used based on the LAB color model as the main technical route in this paper, that is, the chromatic aberration Canny operator color difference canny (CDC) algorithm. This algorithm was able to detect the edge through the color gradient (chromatic aberration) between pixels and determine the pixel amplitude as per the calculation of the first-order partial derivative finite difference at pixels and determine the pixel amplitude as per the calculation of the first-order partial derivative finite difference at the x-direction, y-direction, 135° direction, and 45° direction within the 8 adjacent zones of pixels. Its main advantages include high edge-positioning accuracy and effective noise suppression. It was able to accurately determine the color differences in accordance with the vision of human eye, and good results were achieved during practical application. The mathematical expression of the chromatic aberration Canny operator (CDC) is as follows.

Chromatic aberration in the x-direction:

\[ D_x[i, j] = CD(I[i + 1, j], i - 1, j), \]  

(1)

chromatic aberration in the y-direction:

\[ D_y[i, j] = CD(I[i, j + 1], i, j - 1), \]  

(2)

chromatic aberration in the 135° direction:

\[ D_{135°}[i, j] = CD(I[i + 1, j + 1], I[i - 1, j - 1]), \]  

(3)

chromatic aberration in the 45° direction:

\[ D_{45°}[i, j] = CD(I[i - 1, j + 1], I[i + 1, j - 1]), \]  

\[ CD(A, B) = \left[ (L_A - L_B)^2 + (A_A - A_B)^2 + (B_A - B_B)^2 \right]^{1/2}, \]  

(4)

where \( CD(A, B) \) represents the chromatic aberration between the pixel point \( A \) and the pixel point \( B \).

The pixel chromatic aberration amplitude and direction were calculated as per the coordinate conversion equation from the rectangular coordinates to the polar coordinates. The chromatic aberration amplitude was calculated as per the second-order norm equation as follows:

\[ CDC[i, j] = \sqrt{D_x[i, j]^2 + D_y[i, j]^2 + D_{135°}[i, j]^2 + D_{45°}[i, j]^2}. \]  

(5)

Chromatic aberration direction:

\[ \theta[i, j] = \arctan\left( \frac{D_y[i, j]}{D_x[i, j]} \right). \]  

(6)

(2) Adaptive Calculated Dynamic Threshold. A whole image can be divided into several subimages. There might be a certain overlapping region between subimages so as to achieve the continuous profile and calculate the parameters for the proportion between the overlapping region and the subimages. The high and low threshold values of subimages were adaptively set according to the nonmaximum suppression results. The calculation equation should be as follows:

\[ \tau_{high} = (1 - \beta) \tau_H + \beta \tau_L, \]  

\[ \tau_{low} = (1 - \beta) \tau_L + \beta \tau_T, \]  

(7)

where \( \tau_H \) and \( \tau_L \) represent the global high and low threshold values of the whole image, respectively; \( \tau_T \) and \( \tau_L \) represent the local high and low threshold values of subimage zone, respectively. \( 0 < \beta < 1 \) represents the threshold adjustment rate. If \( \beta = 0 \), the whole image should not be adjusted. If \( \beta = 1 \), the whole image should be divided fully in accordance with the local features of the subimages.

(3) Boundary Tracking Generated Coastline. When the point with the chromatic aberration amplitude of a certain pixel in the whole image is greater than the high threshold value which was used as the starting point for tracking, the neighborhood pixel (other chromatic aberration amplitudes within the 8 pixel neighborhoods were greater than the high threshold value) should be set as the edge and used as the starting point for tracking. If there were no pixels (chromatic aberration amplitude was greater than the high threshold value) around the pixel point, the pixel (chromatic aberration amplitude was greater than the low threshold value) should be discovered within the 8 pixel neighborhoods and used as the starting point for tracking, up until the edge and the starting point cannot be found, and thus determined as the contour endpoint. The template was checked to determine whether the edge could be connected. In the event of any noise at the isolated point, it should be removed. Finally, the boundary was extracted for fine processing.

(4) Coastline Classification and Analysis Based on the Spectrum Feature Database. Based on the sample collection and field spectrum measurement of the remote sensing image of the sea area, the systematic research on the types and usage methods for development and exploitation of coastal zones, as well as the spectral features and image features in the common star-source data, the remote sensing classification criteria were presented for development and exploitation of coastal zones in this paper. Coastlines can be divided into natural coastlines, artificial coastlines, and estuary coastlines. The natural coastlines can be divided into bedrock coastline, silt-muddy coastline, sandy coastline, and biological coastline (including mangroves and coral reefs). The artificial coastlines can be divided into cofferdam coastline, marine reclamation land coastline, transportation engineering coastline, and protected coastline.

The object-oriented image feature analysis for the sea area was used in this study. Based on the main types and components of coastline development and exploitation, all kinds of target surface features were detected and extracted, including the spectrum, shape, texture, shadow, space, location, and related layout. The standard curve spectrum was obtained corresponding to the component. The spectral characteristics
of each component were analyzed, the coastline category information extraction was achieved by using a fuzzy classification algorithm, and the length of Chinese coastline was obtained (as shown in Table 2) after the analysis of the spectrum feature database.

3.3.2. Validation of Coastline Extraction Results. After the completion of the coastline extraction algorithm model, it was necessary to test and validate the model extraction results. The validation methods of the model extraction result should include the point-by-point comparison method and the overall comparison method. Point-by-point inspection refers to the superposition of extracted results and actual conditions with point-by-point comparison as to its accuracy. Overall comparison should refer to the evaluation of extracted results by means of the indexes for the overall spatial pattern as discussed by Wu [23]. The Kappa coefficient was selected to quantitatively reflect the model operation for extracting the accuracy of coastline. This was mainly based on the different-stage coastline extraction results and the corresponding remote sensing image data to carry out the Kappa coefficient calculation of the adjacent grid map, to determine the average coefficient value of results at each stage, and to obtain the Kappa coefficient for the accuracy of coastline extraction results over the timeframe. The calculation equation should be as follows:

\[
K = \frac{(P_o - P_C)}{(1 - P_C)},
\]

where \(P_o\) represents the percentage of consistent types of parts on the comparative grid map, that is, the observation consistency ratio; \(P_C\) represents the expectation consistency ratio, in which \(P_o = s/n, P_C = (a1 * b1 + a0 * b0) / (n * n)\) (the total pixel number of grid: \(n\); the pixel number of grid (1): \(a1\); the pixel number of grid (0): \(a0\); the pixel number of extracted grid (1): \(b1\); the pixel number of extracted grid (0): \(b0\); the pixel number of two grids with the equal corresponding pixel value: \(s\)). The different Kappa coefficients showed consistency to varying extents.

Shoreline information from the National Marine Data & Information Service was used to test the accuracy of the extracted shoreline. The calculation equation was as follows:

\[
I = \frac{(L - L')}{L},
\]

where \(I\) is the accuracy coefficient, \(L\) is the actual length of shoreline, and \(L'\) is the extracted length of shoreline.

<table>
<thead>
<tr>
<th>Year</th>
<th>Coastline length (m)</th>
<th>Natural coastline length (m)</th>
<th>Artificial coastline length (m)</th>
<th>Estuary coastline length (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>18,501,296</td>
<td>9,383,095</td>
<td>8,980,905</td>
<td>137,297</td>
</tr>
<tr>
<td>2008</td>
<td>18,515,830</td>
<td>9,730,865</td>
<td>8,651,800</td>
<td>133,165</td>
</tr>
<tr>
<td>2009</td>
<td>18,599,902</td>
<td>10,167,232</td>
<td>8,301,830</td>
<td>130,839</td>
</tr>
<tr>
<td>2010</td>
<td>18,644,768</td>
<td>10,378,360</td>
<td>8,136,654</td>
<td>129,754</td>
</tr>
<tr>
<td>2011</td>
<td>18,946,699</td>
<td>10,733,161</td>
<td>8,085,561</td>
<td>127,977</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
<th>2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kappa coefficient</td>
<td>0.69</td>
<td>0.72</td>
<td>0.78</td>
<td>0.81</td>
<td>0.86</td>
</tr>
<tr>
<td>Accuracy coefficient</td>
<td>0.018</td>
<td>0.013</td>
<td>0.023</td>
<td>0.015</td>
<td>0.012</td>
</tr>
</tbody>
</table>

It can be seen from Table 3 that all the Kappa coefficients obtained in the previous five years were greater than 0.6 and that the accuracy coefficients obtained in the previous five years were less than 0.023. Previous studies about Kappa by Blackman and Koval [24] and Landis and Koch [25] have shown that the coastline extraction results achieved using the chromatic aberration Canny operator edge detection method should be consistent with the actual coastline heights, and the coastline extraction results in 2010 and 2011 should be almost completely consistent with the actual coastline heights.

3.3.3. Validation of Coastline Classification Results. In order to ensure the accuracy of the coastline classification results, the supervised classification method was used to complete the automatic classification of the coastline remote sensing images and to carry out sampling secondary interpretation and confirmation by means of the artificial visual interpretation method, which ensured the accuracy of coastline properties and the distribution boundary of the classification results. From the coastline extraction results over the timeframe, 2,000 samples were extracted for analysis. As per the probability calculation, the ratio of conforming samples was up to 95% or above, which met the required standards.

4. Analysis of Coastline Length Changes and Trends

4.1. Variation of Coastline Length over the Years. Variations in coastline length reflect the overall impact of coastline resource utilization, natural erosion and siltation, oceanic dynamics, and other factors. Changes and trends in the coastline length depend on the development speed and trends between natural coastline, artificial coastline, and estuary coastline.

It can be seen from Table 2 that over the past five years, except for when the coastline length decreased slightly in 2009, the total length of China’s coastline increased. From 2007-2008 and 2009-2011, the rate of increase of China’s
artificial coastline was greater than the rate of decrease of natural coastline and estuary coastline, so the overall coastline length increased. In 2009, the artificial coastline increased by 436,368 m, less than the rate of increase in 2008. The natural coastline and the estuary coastline decreased by 349,970 m, more than the rate of decrease in 2008. Analysis showed that the main reason for this was the formal implementation of China’s Sea Reclamation Scheme and Management System in 2009. The area of sea reclamation (mostly including the cutoff type of reclamation) was increased by 6887.38 hectares over 2008 levels, while the coastline length slightly decreased in 2009. After 2010, when the coastal marine administrative departments at all levels began making considerable efforts on sea reclamation management and paid more attention to sea reclamation methods, which resulted in larger proportions of offshore type (island type) and convex barrier type reclamations, and combined with the enhancement of environmental protection and ecological awareness, the rate of decrease of the natural coastline and estuary coastline slowed, while the coastline length increased.

4.2. Coastline Changes and Trends in the Future

4.2.1. Prediction Methods. There was a big difference between the method of endpoint rate and the average rate as calculating the coastline change rates. It was known from the coastline change rate calculation method and the impact factor analysis made by Jingfu et al. [26] that a change in coastline length should be regarded as a periodic and oscillating change and that the rate calculation method should not be used for prediction of future variation trends. Having only five stages of data, the total change in length of the coastline should be predicted using the grey sequence prediction method.

4.2.2. Future Variation Trend. Using the grey prediction method, the coastline length values extracted from 2007-2011 were included into the grey sequence prediction model. The development factor \((a = -0.01)\) and the grey action \((b = 18357.74)\) should be calculated, and the grey prediction model GM \((1, 1)\) of coastline should be obtained:

\[
X^{(1)}_{(t+1)} = [X^{(1)}_{(0)} - 2637356.24] e^{(0.01)} - 2637356.24, \tag{10}
\]

where \(X^{(1)}_{(t+1)}\) represents the predicted accumulated value of \(t + 1\) time; \(X^{(1)}_{(0)}\) represents the original value of starting time. The accumulative length of coastline (in 2008–2015) should be calculated as per (10). The former term should be subtracted by the later term, and the coastline prediction results were shown in Table 4. The variance ratio was 0.391 and the small error probability was 1, which was bigger than 0.95. It was shown that the accuracy of predicted results was consistent with the requirements. Therefore, the predicted value of the mainland coastline length obtained by the grey prediction method was in accordance with the actual conditions. The predicted results showed that the future mainland coastline length would increase and China’s mainland coastline length would be up to about 19,471,983 m by 2015.

5. Conclusions and Discussions

Through the data mining algorithm, the preprocessing of China’s low-precision remote sensing images taken in the previous five years, the extraction of coastline lengths and types from 2007–2011, and the analysis of coastline changes and trends, the following conclusions can be made.

(1) The multisource imaging matching algorithm of a SIFT operator, the geometric correction of a few control point satellite images, the color-transferring algorithm for maintaining the consistency of typical surface features, toning processing technology, and the automatic matching-line generation image mosaic technology were used to preprocess the low-precision environmental disaster reduction satellite and CBERS-01 satellite images. Combined, these are all effective methods.

(2) The coastline can be extracted using the chromatic aberration Canny operator edge detection method. The Kappa coefficient should be greater than 0.6 throughout the calculation, which shows that the extraction results are consistent with the actual measured results. By analysing the coastline type using the spectrum feature database, trends identify a decreased natural coastline and estuary coastline, but an increased artificial coastline were found. These trends are consistent with the current utilization of the coastline area in China.

(3) The total length of China’s coastline increased overall. This is despite the coastline decreasing slightly in 2009 due to the formal implementation of China’s Sea Reclamation Scheme. By using the grey prediction method, China’s coastline length is predicted to increase in the future, with the total length reaching 19,471,983 m by 2015.

With the continuous increase of artificial coastline length and related economic benefits, the balance of the marine

<table>
<thead>
<tr>
<th>Year</th>
<th>Coastline length (m)</th>
<th>Prediction length (m)</th>
<th>Residual (m)</th>
<th>Relative error</th>
</tr>
</thead>
<tbody>
<tr>
<td>2007</td>
<td>18,501,296</td>
<td>18,501,296</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>2008</td>
<td>18,515,830</td>
<td>18,549,380</td>
<td>-33550.33</td>
<td>0.00</td>
</tr>
<tr>
<td>2009</td>
<td>18,599,902</td>
<td>18,678,455</td>
<td>-78552.97</td>
<td>0.00</td>
</tr>
<tr>
<td>2010</td>
<td>18,644,768</td>
<td>18,808,427</td>
<td>-163659.73</td>
<td>0.01</td>
</tr>
<tr>
<td>2011</td>
<td>18,946,699</td>
<td>18,939,304</td>
<td>7394.30</td>
<td>0.00</td>
</tr>
<tr>
<td>2012</td>
<td>18,501,296</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2013</td>
<td>113,250,456</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2014</td>
<td>19,337,425</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2015</td>
<td>19,471,983</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
ecosystem will be disturbed, affecting the harmony between humans and nature. In order to ensure the sustainable exploitation and utilization of coastal resources and the reasonable adjustment of coastal development methods, the impact of coastline change trends on the marine ecological environment and the changes of spatial position should be studied further.

Acknowledgments

This work was financially supported by the Public Science and Technology Research Funds Projects of Ocean (no. 201005011) and Foundation of Key Laboratory of Sea Areas Management Technology SOA (201107).

References


Submit your manuscripts at
http://www.hindawi.com