Edge-Aware Segmentation in Satellite Imagery: A Case Study of Shoreline Detection

Ümit Ruşen Aktaş, Gülcan Can, Fatoş T. Yarman Vural
Department of Computer Engineering
Middle East Technical University
06800, Ankara, Turkey
{rusen,gulcan,vural}@ceng.metu.edu.tr

Abstract

Shoreline extraction algorithms from multispectral imagery depend on threshold selection over spectral values and segmentation in general. Although this method gives high performance values for water delineation, error is accumulated on pixels near shoreline and complicates detection of nearby ships, docks etc. Water-shadow spectral mixing and spectral difference in water regions are two of the reasons for such untrustworthy shoreline results. With only four bands available, improvement in water detection depending only on pixel values is not very promising. Therefore, segmentation gains importance. By an edge-aware segmentation method, we aim to improve overall water and shoreline detection performances.

In this study, a robust three-stage shoreline extraction algorithm is proposed. In the first stage, segmentation is applied over spectral values and then, some segments are combined according to edge information. In the second stage of the algorithm, pixel-based water information is combined with segmentation. The last step consists of enhancement of water regions based on local optimization by merging regions near shore boundary. Additionally, two new boundary-sensitive performance metrics are introduced for measuring the accuracy of the detected boundaries.

1. Introduction

Accurate shoreline detection has great importance as a prior step of land use/land cover (LULC) monitoring and planning. In order to extract shoreline accurately, a robust water detection algorithm is crucial. In multispectral images, water areas give reflectance near to zero in near-infrared (NIR) band and by taking difference of NIR value and green band value, they can be differentiated from green land or soil areas [2]. Nowadays, normalized difference water index (NDWI) analysis or NIR band analysis are used for this task [1, 2]. In this study, dynamic threshold selection is applied on NIR values for IKONOS images, and on NDWI values for GEOEYE images for detecting water.

Recently, region-based approaches have taken place of pixel-based ones in literature [3]. However, unlike Di et al. in [3], we do not need human interaction to extract shoreline correctly. Another comprehensive study that uses segmentation for coastline extraction is presented in [4]. We embrace and use the idea of local boundary optimization rather than global thresholds, yet in a much simpler concept.

In this paper, we propose an edge-aware segmentation method based on popular mean shift algorithm and steerable filter responses [5, 6]. First, images are segmented by mean shift algorithm according to their spectral values. Then, segments are merged if there is no strong steerable filter response separating them and their spectral characters are similar. Steerable filter response, which is obtained from water index, gives maximum response at borders of light to dark area transitions. Thus, it gives a continuous high response over shoreline. Information about dataset used in this study can be found in next section. The shoreline detection algorithm is explained in detail in third part, while proposed performance metrics can be found in fourth section. Then, results are presented, and paper is concluded with an evaluation of the suggested method.

2. Dataset

Dataset contains three IKONOS and seven GEOEYE images, most of which have ships along shoreline, making the process more challenging. All images in the dataset have a 2-meter resolution. The total length of shoreline tested is approximately 160 km. Figure 1 shows a sample GEOEYE image.

Figure 1. GEOEYE sample (RGB)

Ground truth data for all images have also been prepared to be used during performance evaluation.

3. Proposed Algorithm

3.1 Pixel-based Water Detection

In the literature, a constant threshold value which is tuned according to dataset is used in many studies. However, to adapt to image-specific reflectance characteristics, dynamic threshold selection is preferred in this study. Two different strategies are applied while determining dynamic thresholds. NIR histogram is analyzed as described in a previous study for IKONOS images [7], while Otsu's method is applied on NDWI histogram for GEOEYE images [8] computed from the following equation:

$$NDWI = \frac{Green - NIR}{Green + NIR} \quad (1)$$

The result of applying Otsu's threshold over NDWI output of sample image can be seen in Figure 2.

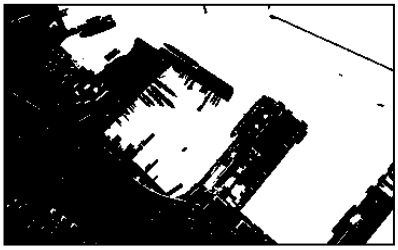


Figure 2. Water mask of sample after applying dynamic threshold

3.2 Edge-aware Segmentation

First, an over-segmented map is obtained by mean shift algorithm on red-green-blue (RGB) spectral values. Mean shift segmentation parameters are selected as range = 3, spatial = 3, min. size = 50. Edge information is integrated in the following manner:

1) *NDWI output is processed using steerable filters, giving high responses especially over the shores. The response image of the sample can be seen in Figure 3. This image holds information about the boundaries which should be preserved.*

2) *Each segment pair is given a score based on their color information, average steerable filter response value inside them and on their common boundary, as follows:*

$$Sc_{xy} = w_1 |I_x - I_y| + w_2 |S_x - S_y| + w_3 S_{xy}, \quad x, y \in S \quad (2)$$

Notation of the formula (2) *is:*

$Sc_{xy} = Score\ of\ segment\ pair\ x\ and\ y,$
$S = Set\ containing\ all\ segments,$
$I_x = Average\ intensity\ for\ segment\ x,$
$S_x = Average\ steerable\ filter\ response\ for\ segment\ x,$
$S_{xy} = Average\ steerable\ filter\ response\ of\ the\ common\ boundary\ between\ segment\ x\ and\ segment\ y,$

w_1, w_2 and w_3 are given equal weights.

3) *A 100-bin histogram of the given scores is calculated, and their distribution is examined. A score threshold T is learned from the dataset.*

4) *All segment pairs having a score below T are merged to obtain larger, homogenous regions.*

Segments obtained after this process are homogenous in terms of spectral values and do not contain high steerable filter edges inside them. Raw and processed segmentation maps of sample image can be found in Figure 4.

The reader should be aware that not all water regions can be merged together. The fact is that finding an optimal threshold T that works for all images is impossible, and we choose to set it a low value (9/100) to stay on the safe side. The optimality of the result is obtained using a second region merging method that works completely locally in Step 3.4.

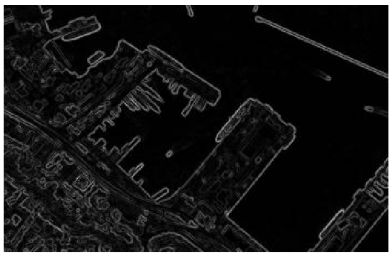


Figure 3. Steerable filter response

3.3 Combining Pixel-based Water Output and Segmentation

After obtaining edge-aware segmentation, water percentage in each segment is analyzed. If the percentage of water pixels among all the pixels in a segment is greater than T_{Water}, then the segment is marked as water region. The threshold T_{Water} is learnt from dataset by cross-validation, and assigned as 0.9.

3.4 Region Merging through Boundaries

While the mask obtained previously contains the main water bodies, it still fails to manage segments on the shore correctly. Due to the differences between deep and shallow water areas, these segments may not be merged with others. A second round of region merging that works over the boundary obtained from Step 3.3 optimizes the output. This optimization is accomplished as follows:

1) *Coarse water regions coming from Step 3.3 form the basis water mask. Their neighboring land segments are extracted using the initial mean-shift segmentation, forming our search space.*

2) *Each land segment from step 1 is labeled as water, if it passes a test. A sample is added to its accompanying water region if and only if it improves the region's shoreline quality. The decision formula for "improvement" is given in formula* (3).

$$L_S = Water, \quad iff \quad Score_{SL\,(With\ S)} > Score_{SL\,(Without\ S)}\ , \ where \quad Score_{SL} = SF_{SL} + ID_{SL} \tag{3}$$

L_S is the label of the segment S, SF_{SL} is the average steerable filter response under the boundary, and ID_{SL} is the average intensity difference of the two sides of shoreline, calculated within a 5-pixel band.

If $Score_{SL}$ increases after the addition of the sample to the water area, it is labeled as water.

3) *If the shoreline has not changed in the last run, terminate. If it did, form the basis water mask including changes in step 2 and go back to step 1.*

The purpose of this procedure is to remove the need to find an optimal score threshold T in Step 3.2. The previously applied region merging procedure gives a global point of view to the algorithm. The local application performed in this step further refines the result. In effect, the boundary is grown into land until it is met with a high resistance. Output of this local procedure is close to the global optimum, since main body of the water region was already detected.

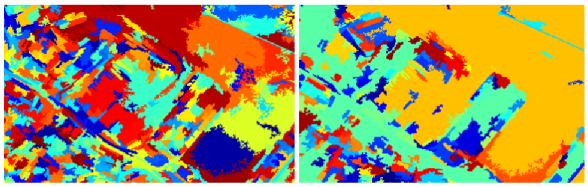


Figure 4. Raw mean shift segmentation (left), edge-aware segmentation (right)

4. Performance Metrics

To estimate the system's performance, a 200 meter-wide band near the actual shoreline is selected for application of precision, recall and accuracy methods. In addition to these, two new performance metrics are proposed. These complementary metrics focus on how well the shore boundary is extracted, rather than the overall performance of the water detection algorithm.

The proposed metrics, Shore Line Precision (SLP) and Shore Line Recall (SLR) only work on boundaries rather than pixel-wise water assignments. Their mathematical representations are given below:

$$\text{SLP} = \frac{\sum_{i \in T} f(i)}{N_T}\ , \qquad \text{SLR} = \frac{\sum_{i \in G} f(i)}{N_G} \tag{4}$$

f(i) function above is defined as:

$$f(i) = ||x(i) - x(j)||$$

While j is the nearest pixel to pixel i in G for SLP, and in T for SLR,

$x(i)$ is pixel position in 2D space,

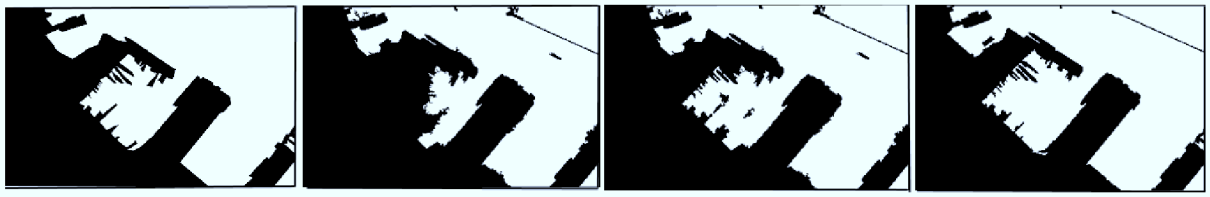
G is the set of shoreline pixels in ground truth data,

N_G is the number of pixels in G.

T is the set of shoreline pixels in test data,

N_T is the number of pixels in T.

SLP and SLR are designed to see the amount of deviation between detected shoreline and actual boundaries. The analogy with precision and recall metrics is straightforward. SLP estimates the "correctness" of a detected pixel by searching for nearest actual boundary pixel in ground truth, and SLR



a. Ground Truth *b. SL-1 Result* *c. SL-2 Result* *d. SL-3 Result*

Figure 5. Algorithm Results for Sample Image

similarly detects how far a ground truth pixel is detected in test mask. Both of these metrics are helpful in the assessment of the shoreline algorithm's quality.

The proposed metrics prevent the water region's size from creating a bias on the overall performance. Smaller values indicate better performance.

5. Conclusion

In addition to the classical performance measures of accuracy, recall, precision, we define and employ the proposed SLP and SLR metrics in this study. First three measures are computed in a pixel-wise manner according to matching of water mask with ground truth. Performance values per unit length of shoreline throughout the dataset can be seen in Table 1. Performance values are computed for algorithms using raw segmentation (SL-1), edge-aware segmentation (SL-2) and proposed two-stage segmentation (SL-3).

Proposed algorithm gives improved results for most of the images according to all metrics but precision. This is due to small segments near shoreline which have fewer detected water pixels. These segments do not contain enough water evidence and are labeled as land in SL-1, thus increasing precision. However, as observed in Figure 5.b, SL-1 misses many water segments. Figure 5 presents the results of proposed algorithms, which are much more accurate than SL-1.

Table 1: Performance Values

	Precision	Recall	Accuracy	SLP	SLR
SL-1	**98.34**	89.18	97.76	19.07	6.31
SL-2	98.41	89.09	97.76	17.11	6.28
SL-3	96.53	**95.17**	**98.50**	**14.26**	**4.05**

Water regions near shoreline or with wavy surfaces have different spectral values than deeper water areas. These regions make water mask quite noisy and are found as separate segments by raw mean-shift segmentation. Thus, it becomes hard to label them as water due to missing evidence. They are combined with nearby water segments in the proposed algorithm, and effects of noisy water results can be eliminated.

Proposed SLP and SLR metrics are closer to human visual perception as well. When compared to other metrics, they capture the deviations near the shoreline better. Algorithms tend to give very similar results in precision, recall and accuracy metrics, as they only change by 1-5 percent. SLP and SLR, on the other hand, differ by as much as 30-35 percent. Their power of discrimination is much more than usual methods.

As future work, errors related to shadows of ships or objects near shore are planned to be eliminated by the help of polygon fitting to the detected shoreline.

References

[1] Bowker, D.E., Davis, R.E., Myrick, D.L., Stacy, K., and Jones, W.T., "Spectral Reflectances of Natural Targets for use in Remote Sensing Studies," *NASA Ref. Pub.*, 1139, 1985.

[2] McFeeters, S.K., "The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features," *International Journal of Remote Sensing*, 17(7):1425-1432, 1996.

[3] Kaichang Di, Ruijin Ma, Jue Wang, Ron Li. "Automatic shoreline extraction from high-resolution IKONOS satellite imagery," *In Proceedings of the 2003 annual national conference on Digital government research, Boston, MA*, 130:1-4,2003.

[4] Liu H., Jezek K.C. , (2004). "Automated extraction of coastline from satellite imagery by integrating Canny edge detection and locally adaptive thresholding methods". *International Journal of Remote Sensing, 25(5)*:937-958.

[5] Comaniciu, D.; Meer, P.; , "Mean shift: a robust approach toward feature space analysis," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* , 24(5):603-619, May 2002.

[6] M. Jacob, M.Unser, "Design of Steerable Filters for Feature Detection Using Canny-Like Criteria", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 8, pp. 1007-1019.

[7] Bayram, U.; Can, G.; Yuksel, B.; Duzgun, S.; Yalabik, N.; , "Unsupervised land use - land cover classification for multispectral images," *Signal Processing and Communications Applications (SIU), 2011 IEEE 19th Conference on* , 574-577, 2011.

[8] Otsu, N.; "A Threshold Selection Method from Gray-Level Histograms," *Systems, Man and Cybernetics, IEEE Transactions on* , 9(1):62-66, Jan.