Journal of Geoscience and Environment Protection, 2015, 3, 17-23
Published Online December 2015 in SciRes. http://www.scirp.org/journal/gep
http://dx.doi.org/10.4236/gep.2015.310003

Automatic Generation of Water Masks from RapidEye Images

Gideon Okpoti Tetteh, Maurice Schönert

BlackBridge AG, Kurfürstendamm 22, Berlin, Germany
Email: gideon.tetteh@blackbridge.com

Received July 2015

Abstract

Water is a very important natural resource and it supports all life forms on earth. It is used by humans in various ways including drinking, agriculture and for scientific research. The aim of this research was to develop a routine to automatically extract water masks from RapidEye images, which could be used for further investigation such as water quality monitoring and change detection. A Python-based algorithm was therefore developed for this particular purpose. The developed routine combines three spectral indices namely Simple Ratios (SRs), Normalized Green Index (NGI) and Normalized Difference Water Index (NDWI). The two SRs are calculated between the NIR and green band, and between the NIR and red band. The NGI is calculated by rationing the green band to the sum of all bands in each image. The NDWI is calculated by differencing the green to the NIR and dividing by the sum of the green and NIR bands. The routine generates five intermediate water masks, which are spatially intersected to create a single intermediate water mask. In order to remove very small waterbodies and any remaining gaps in the intermediate water mask, morphological opening and closing were performed to generate the final water mask. This proposed algorithm was used to extract water masks from some RapidEye images. It yielded an Overall Accuracy of 95% and a mean Kappa Statistic of 0.889 using the confusion matrix approach.

Keywords

Water Mask, Image Threshold, Simple Ratio, Normalized Green Index, Normalized Difference Water Index, Logical and Morphological Operations, RapidEye

1. Introduction

Water, which covers over 70 percent of the earth's surface, sustains all known life forms on earth, making it imperative to monitor the spatial extent and quality of waterbodies over time. Doing this monitoring has been done in the past through direct field measurements, which suffers a draw-back in that it is expensive and lags behind with respect to the frequency at which timely information can be collected. The advent of earth-orbiting satellites over the years has enabled us to remotely monitor waterbodies in a timely manner. This particular research focused essentially on the extraction of the spatial extents of waterbodies from satellite images in an automated fashion.

The automatic extraction of waterbodies from satellite images is well documented. The Normalized Differ-

How to cite this paper: Tetteh, G.O. and Schönert, M. (2015) Automatic Generation of Water Masks from RapidEye Images. *Journal of Geoscience and Environment Protection*, **3**, 17-23. http://dx.doi.org/10.4236/gep.2015.310003

ence Water Index (NDWI), calculated as a ratio of the difference between the green and the NIR band to the sum of those two bands, was first proposed by [1] and was used to extract waterbodies from satellite images. The NDWI was later modified by [2] by simply replacing the NIR band with the MIR band and came up with the name Modified NDWI (MNDWI). Due to the limitations of the NDWI and MNDWI in terms of unstable threshold and the mixture of waterbodies and other dark objects, [3] developed what they call the Automated Water Extraction Index (AWEI) to improve the extraction of waterbodies especially in areas that include shadow and dark surfaces. An automated method for extracting rivers and lakes from Landsat imagery by exploiting the NDWI, MNDWI and AWEI was developed by [4]. The authors then extract waterbodies by applying a threshold value to these indices, extracting both wide and narrow waterbodies and finally combining all the extracted waterbodies.

An automatic object-based waterbody extraction method was developed by [5] based on NDWI using the Locally Excitatory Globally Inhibitory Oscillator Networks (LEGION) as promulgated by [6]. Using a Perceptron Model, [7] developed an automated waterbody extraction routine for Landsat-7 ETM+ by feeding the low reflectance of water in the SWIR band, the relatively lower reflectance in the NIR band to the green band and in the red to the NIR band, and a binary of the summation of NDWI and MNDWI as feature vectors to the model.

The Whole-R Method was developed by [8] to automatically extract waterbodies from Landsat images. This method involves the binarization of the red band based on a threshold value and the spectral angle between the spectral value of the image and that of its spectral library. Two spectral enhancement techniques called Weighted Water Index (WWI) and Weighted Water-Soil Index (WWSI) were introduced by [9] as a means of enhancing waterbodies in satellite images making it easier to extract waterbodies by applying a simple threshold to the enhanced images.

Owing to the limitation of the NDWI and MNDWI in sometimes confusing dark objects like shadow and coal tarred roads with water, [10] came up with an automatic routine to extract a threshold value in the NIR band of some airborne hyperspectral images to mask potential water pixels, and then performs spectral slope analysis to remove possible shadows from the potential water mask to generate the final water mask.

Instead of using a single global threshold value on the NDWI to extract waterbodies, [11] dynamically extracted different threshold values based on the sub-pixel components of each image. Based on empirical analysis of a Landsat TM image of the Huanghe river delta, [12] came up with a decision tree algorithm to automatically extract waterbodies from Landsat image. Through spectral pattern analysis of the green, red, near infrared (NIR) and short wave infrared (SWIR) bands, [13] extracted waterbodies from Spot 5 and Landsat 5 TM images.

In order to come up with an approach which is insensitive to image sensors, [14] introduced an entropy-based method for extracting waterbodies from satellite images based on the justifiable assumption that water normally has a smooth surface in images. Some spectral indices namely Normalized Difference Vegetation Index (NDVI), Normalized Difference Moisture Index (NDMI), Water Ratio Index (WRI), AWEI, NDWI and MNDWI were tested by [15], coming to the conclusion that NDWI gave the best results in extracting waterbodies.

The research presented here adopted the automatic extraction of waterbodies from the NIR band as described by [10], the NDWI as proposed by [1], two proposed simple ratios (SRs), and one newly proposed spectral index called Normalized Green Index (NGI).

2. Materials

2.1. RapidEye Sensor Specifications

The RapidEye satellite system is a constellation of five identical earth observation satellites with the capacity to provide large area, multi-spectral images with frequent revisits in high resolution (6.5 m at nadir). In addition to the blue (440 - 510 nm), green (520 - 590 nm), red (630 - 685 nm) and near-infrared (760 - 850 nm) bands, the sensor has a red-edge (690 - 730 nm) band, which is very suitable for vegetation analysis. The RapidEye Level 3A standard image product covers an area of 25 by 25 km, is radiometrically calibrated to radiance values, and is ortho-rectified and resampled to 5 m spatial resolution.

2.2. Data Used

In order to test the robustness of the developed routine, three geographically varying RapidEye Level 3A images were used. **Table 1** gives a summary of the images used for developing the routines and for the accuracy assessment of the classified waterbodies.

Table 1. Overview of the test images.

Test Site	Location		Water Conditions	Intended test
	Longitude	Latitude		
USA	−89˚50'57.17"	29˚34'59.71"	Black, shallow, turquoise, yellow, green and brownish water	Different inland and coastal water conditions
Brazil	−36˚43'37.19"	−5˚5'54.69"	Variety of greenish and turquoise, greyish violet and turquoise waters	Different inland and coastal water conditions
Mali	−4˚0'21.4632"	15˚18'7.8372"	Greyish-blue, turquoise, black waters and shallow flooded river banks	Flood, different water conditions, flooded AG and vegetation

3. Methodology

The entire methodology adopted in this research is captured in **Figure 1**. The whole process begins by first extracting the solar zenith and acquisition date of each image. These extracted values are used to generate the Top-Of-Atmosphere (TOA) reflectance image from the RapidEye L3A radiance images, which are already corrected for offset. The TOA images serve as base images for all other components of the automation processes.

An initial water mask is created using the low reflectance of water in the NIR band by applying a simple threshold on the NIR band. Water absorbs most of the electromagnetic energy in the NIR region, therefore it has lower reflectance in the NIR region as against other features in an image. The threshold value is extracted automatically from the NIR histogram precisely following the method as proposed by [1]. **Figure 2** shows the typical spectra for six feature types in all five spectral bands of the RapidEye images.

The NIR threshold value is estimated by finding the two local peaks at the left tail of the NIR histogram knowing that the low reflectance surfaces would be at that tail of the NIR histogram. A fifth order polynomial is then fitted on all histogram points between these two peaks, after which the minimum of the polynomial curve is extracted, and finally the corresponding NIR value for this minimum is extracted as the threshold value. This value is then used to threshold the NIR band to obtain the first potential water mask.

Four additional water masks are created using two Simple Ratios (SRs), Normalized Green Index (NGI) and Normalized Difference Water Index (NDWI). Equation (1) to (4) shows the formulas for all the four indices

$$\mathrm{SR1} = NIR / \mathrm{Green} \tag{1}$$

$$\mathrm{SR2} = NIR / \mathrm{Red} \tag{2}$$

$$\mathrm{NGI} = \mathrm{Green} / \left(\mathrm{Blue} + \mathrm{Red} + \mathrm{RedEdge} + NIR\right) \tag{3}$$

$$\mathrm{NDWI} = \left(NIR - \mathrm{Green}\right) / \left(NIR + \mathrm{Green}\right) \tag{4}$$

Sample spectral index values computed for the same aforementioned feature types are shown in **Figure 3**. A close examination of the SRs values shows that the values of both muddy and clear water are lower than all the other feature types and indeed less than one, hence making it easier to extract two additional waterbodies by applying a simple threshold to the two SRs respectively.

The muddy and clear waterbodies had a relatively higher NGI value, 0.25 and above, than all the other feature types, hence that value was used to threshold the NGI image to create the fourth waterbody. An examination of the NDWI values indicated that the waterbodies had the lowest values, 0 and below, making it a straight forward threshold operation to generate the fifth potential waterbodies. There were cases where other dark objects had NGI and NDWI values similar to that of water, hence the next step was to do a *logical intersection* of all the five potential waterbodies to create a single intermediate water mask.

Finally, in order to remove very small waterbodies and any remaining holes within the water mask, morphological opening followed by closing are performed on the intermediate water mask to create the final water mask.

4. Results & Conclusions

The accuracy of the water masking algorithm was assessed using stratified random sampling. 200 points were randomly generated for the water and non-water classes in each classified image. For each point, the ground truth class (non-water and water) was assigned by visual identification with the assistance of Google Earth.

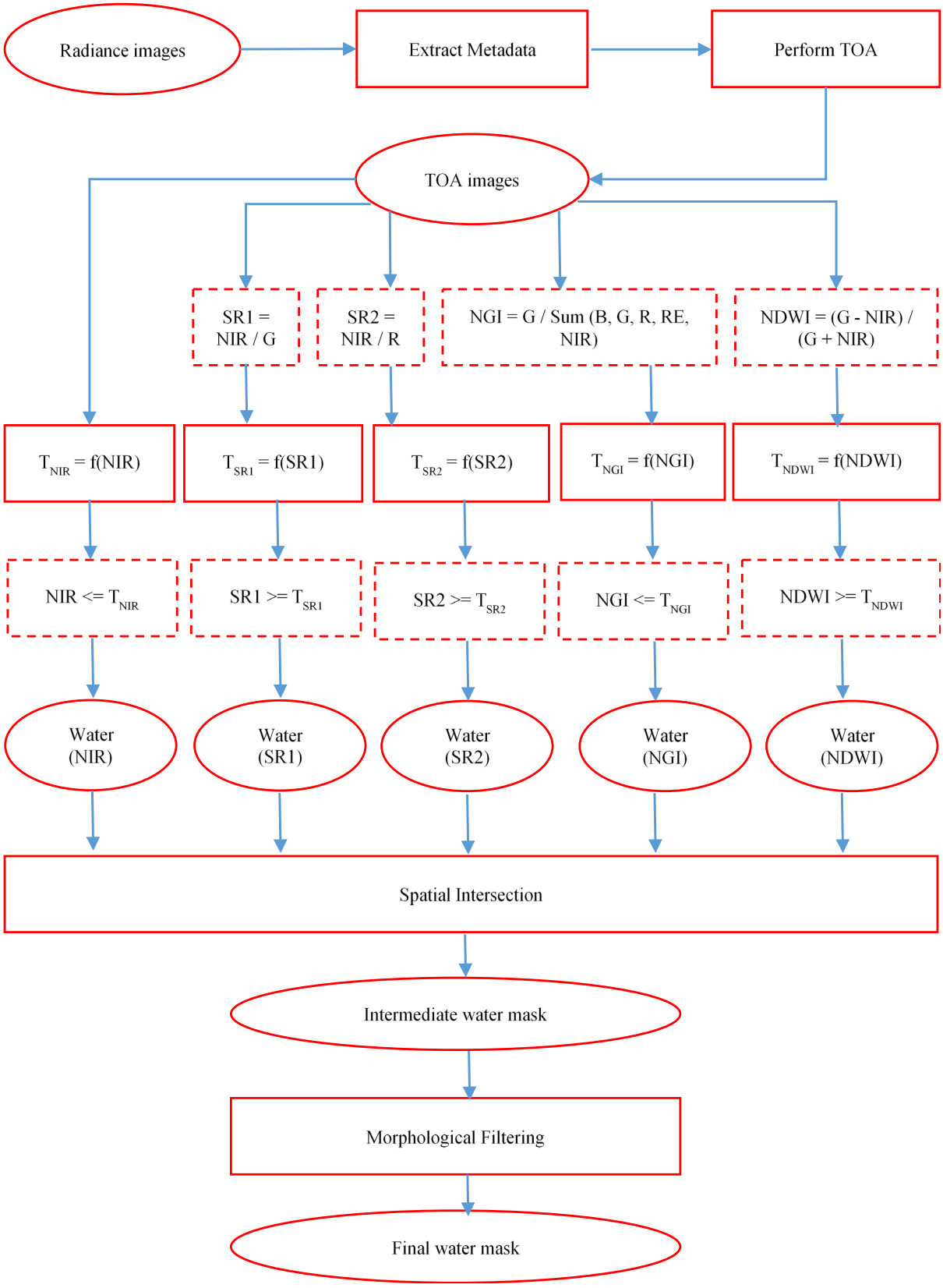


Figure 1. The automated workflow for waterbodies extraction.

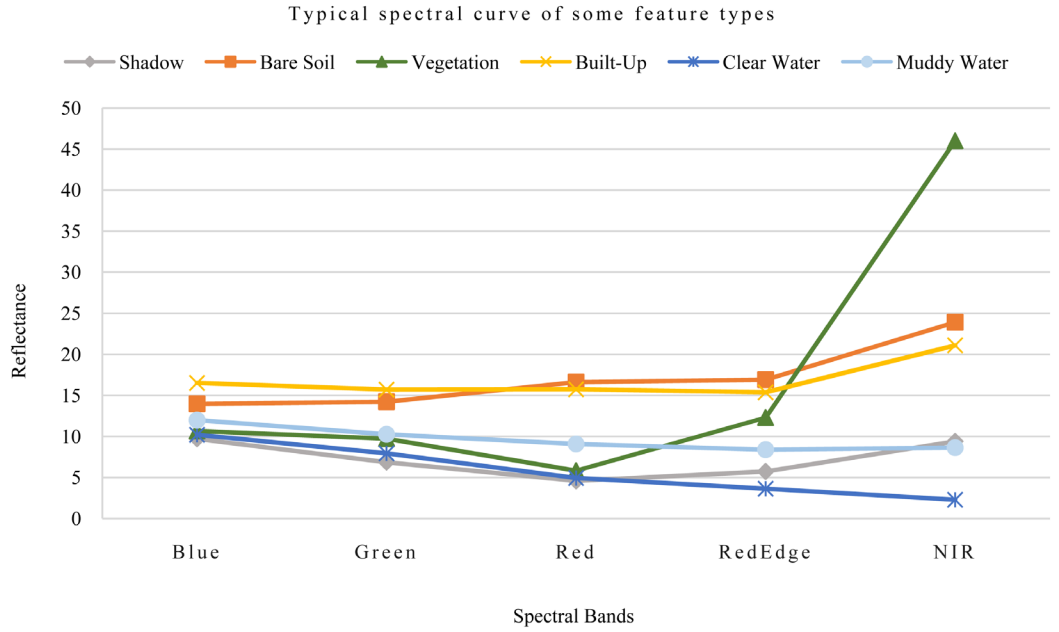


Figure 2. Typical spectral curve of some feature types.

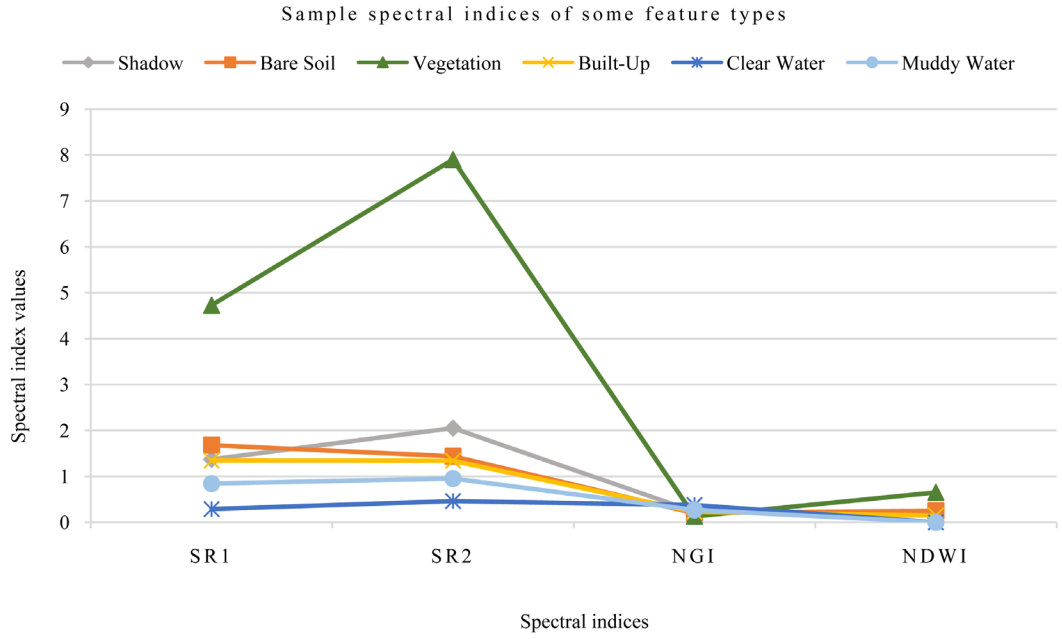
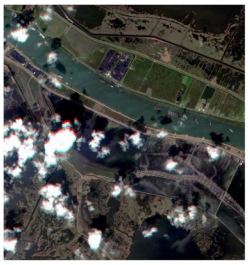


Figure 3. Sample spectral indices of some feature types.

These reference points were compared to those of the generated water masks and the accuracy was assessed using the *Confusion Matrix* approach.

The *Producer's Accuracy*, *User's Accuracy*, *Overall Accuracy* and *Kappa* of the classification result of the implemented algorithm were calculated accordingly. The Producer's Accuracy here refers to the probability that a visually identified and assigned waterbody is classified as such, while the User's Accuracy refers to the probability that a classified waterbody in the image is indeed a waterbody. The Overall Accuracy is the probability of the correctly classified waterbodies. Kappa was used as a measure of agreement between the prediction of the proposed water algorithm and the reality as was done in this case through visual inspection.

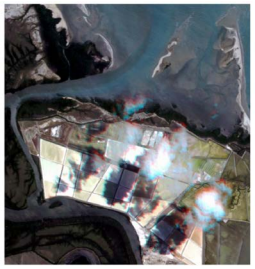
The test images, resultant water masks and Confusion Matrices are shown in **Figure 4-6**.



	Non-water	Water	User's Accuracy
Non-water	109	3	97.3
Water	8	80	90.9
Producer's Accuracy	93.2	96.4	
Overall Accuracy = 94.5%			
Overall Kappa = 0.888			

(a) (b) (c)

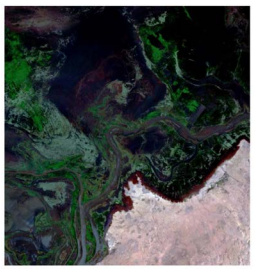
Figure 4. (a) USA image; (b) Extracted water mask; (c) Confusion matrix.

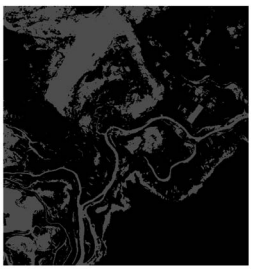


	Non-water	Water	User's Accuracy
Non-water	130	6	95.6
Water	8	56	87.5
Producer's Accuracy	94.2	90.3	
Overall Accuracy = 93.0%			
Overall Kappa = 0.838			

(a) (b) (c)

Figure 5. (a) Brazil image; (b) Extracted water mask; (c) Confusion matrix.





	Non-water	Water	User's Accuracy
Non-water	138	3	97.9
Water	2	57	96.6
Producer's Accuracy	98.6	95.0	
Overall Accuracy = 97.5%			
Overall Kappa = 0.940			

(a) (b) (c)

Figure 6. (a) Mali image; (b) Extracted water mask; (c) Confusion matrix.

In conclusion, the proposed algorithm showed a very good performance on all the water conditions mentioned before. The algorithm detects water with an Overall Accuracy of 95% (mean Kappa of 0.889). For the water class, the mean Producer's Accuracy is 93.9% and the User's Accuracy is 91.7%. The main drawback of this algorithm is that some shadow areas especially the ones over vegetated areas (see **Figure 5**) were also classified as waterbodies. Therefore, this algorithm ought to be further developed to handle those exceptional cases.

References

[1] McFeeters, S.K. (1996) The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Features. *International Journal of Remote Sensing*, **17**, 1425-1432. http://dx.doi.org/10.1080/01431169608948714

[2] Xu, H.Q. (2006) Modification of Normalised Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery. *International Journal of Remote Sensing*, **27**, 3025-3033. http://dx.doi.org/10.1080/01431160600589179

[3] Feyisa, G.L., Meilby, H., Fensholt, R. and Proud, S.R. (2014) Automated Water Extraction Index: A New Technique for Surface Water Mapping Using Landsat Imagery. *Remote Sensing of Environment*, **140**, 23-35. http://dx.doi.org/10.1016/j.rse.2013.08.029

[4] Jiang, H., Feng, M., Zhu, Y., Lu, N., Huang, J. and Xiao, T. (2014) An Automated Method for Extracting Rivers and Lakes from Landsat Imagery. *Remote Sensing*, **6**, 5067-5089. http://dx.doi.org/10.3390/rs6065067

[5] Li, M., Xu, L. and Tang, M. (2011) An Extraction Method for Water Body of Remote Sensing Image Based on Oscil-

latory Network. *Journal of Multimedia*, **6**, 252-260. http://dx.doi.org/10.4304/jmm.6.3.252-260

[6] Terman, D. and Wang, D.L. (1995) Global Competition and Local Cooperation in a Network of Neural Oscillators. *Physica D*: *Nonlinear Phenomena*, **81**, 148-176. http://dx.doi.org/10.1016/0167-2789(94)00205-5

[7] Mishra, K. and Prasad, P.R.C. (2014) Automatic Extraction of Water Bodies from Landsat Imagery Using Perceptron Model. *Journal of Computational Environmental Sciences*, **2015**, 1-9. http://dx.doi.org/10.1155/2015/903465

[8] Nawaz, N., Srinivasulu, S. and Rao, P.K. (2013) Automatic Extraction of Water Bodies Using Whole-R Method. *International Journal of Environmental, Ecological, Geological and Geophysical Engineering*, **7**, 564-567.

[9] Polidorio, A.M., Flores, F.C., Franco, C., Imai, N.N. and Tommaselli, A.M.G. (2010) Enhancement of Terrestrial Surface Features on High Spatial Resolution Multispectral Aerial Images. *Proceedings of the* 23*rd SIBGRAPI Conference on Graphics, Patterns and Images*, Rio Grande do Sul, 30 August-3 September 2010, 295-300. http://dx.doi.org/10.1109/SIBGRAPI.2010.46

[10] Bochow, M., Heim, B., Kster, T., Roga, C., Bartsch, I., Segl, K. and Kaufmann, H. (2012) On the Use of Airborne Imaging Spectroscopy Data for the Automatic Detection and Delineation of Surface Water Bodies. In: Chemin, Y., Ed., *Remote Sensing of Planet Earth*, InTech, Rijeka, 3-22. http://dx.doi.org/10.5772/34073

[11] Ji, L., Zhang, L. and Wylie, B. (2009) Analysis of Dynamic Thresholds for the Normalized Difference Water Index. *Photogrammetric Engineering and Remote Sensing*, **75**, 1307-1317. http://dx.doi.org/10.14358/PERS.75.11.1307

[12] Fu, J., Wang, J. and Li, J. (2008) Study on the Automatic Extraction of Water Body from TM Image Using Decision Tree Algorithm. *Proceedings of SPIE* 6625, *International Symposium on Photoelectronic Detection and Imaging* 2007: *Related Technologies and Applications*, Beijing, 9 September 2007, 1-9.

[13] Nguyen, D.D. (2012) Water Body Extraction from Multi Spectral Image by Spectral Pattern Analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, **XXXIX-B8**, 181-186. http://dx.doi.org/10.5194/isprsarchives-XXXIX-B8-181-2012

[14] Zhaohui, Z., Prinet, V. and Songde, M. (2003) Water Body Extraction from Multi-Source Satellite Images. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, Toulouse, 21-25 July 2003, 3970-3972. http://dx.doi.org/10.1109/igarss.2003.1295331

[15] Rokni, K., Ahmad, A., Selamat, A. and Hazini, S. (2014) Water Feature Extraction and Change Detection Using Multitemporal Landsat Imagery. *Remote Sensing*, **6**, 4173-4189. http://dx.doi.org/10.3390/rs6054173