An automated technique to determine spatio-temporal changes in satellite island images with vectorization and spatial queries

SÜLEYMAN EKEN* and AHMET SAYAR

Department of Computer Engineering, Kocaeli University, 41380 Kocaeli, Turkey
e-mail: suleyman.eken@kocaeli.edu.tr

MS received 9 November 2013; revised 16 April 2014; accepted 21 September 2014

Abstract. For spatio-temporal and topologic analyses, vectorial information (carrying coordinate values defined as point sets) gives better information than its raster (grid of pixels) counterpart. The study presented in this paper is based on (i) recognition and extraction of an island object in a set of digital images captured by LandSat-7 satellite and (ii) modelling it as a polygon (vectorial) and making it easy to process and easy to understand for computers and information science applications. Polygon representations of island images then can be stored and manipulated through object-relational spatial databases. Spatial databases have built-in functions and services for spatial objects defined with geometry types such as points, lines, and polygons. By this way we will be utilizing the rapidly changing and developing object-relational database communities' studies and discoveries in spatio-temporal and topological analysis for the investigation of digital satellite images. This approach also enables service qualities as well as a better performance. The efficiency and feasibility of the proposed system will be examined by various scenarios such as earthquake, erosion and accretion. Scenarios are based on measuring the effects of the natural phenomena on a selected island on satellite images.

Keywords. Vectorization; raster data; vector data; spatial databases; spatial queries; data modelling.

1. Introduction

The proposed study is based on creating polygon-vector model of real island satellite imageries through vectorization based on edge-detection, and storing island objects into Object-Relational spatial databases and making them available for spatio-temporal and topological analysis. Spatio-temporal analyses aim at determining the changes in shape and location over the time. Topological analyses study location-based relations of geographic objects. Examples of

*For correspondence

topological relations and properties are overlapping, touching, intersection and neighbourhood. The reason behind selecting islands is that they are visually distinguishable objects not like the others such as countries, cities and roads whose boundaries are not easy to define visually. However, this study is going to form a basis for investigation of those objects through satellite images for the future studies.

There are two kinds of data models — raster and vector. Raster graphics are also called bitmap images, and consist of a grid of individual pixels where each pixel can be of a different colour or a shade. However, to describe an image, vector graphics use mathematical relationships between points and connection paths. To produce a raster image file (for example, TIFF image) accurately, graphics software keeps track of a large amount of information, including the exact location and color of each pixel in the collection of pixels. This results in huge file sizes for raster graphics. Vector counterparts of the same raster files are thousands, even tens of thousands, times smaller in size. It is because of the fact that vector data carries only point coordinates and mathematical relations between them. Raster-to-vector conversion, i.e., vectorization, has the key functionalities in Geographic Information Systems (GIS) and image processing for Remote Sensing. Using vector representations of raster images in spatial queries gives more accurate results, lower computation costs and better response times. Examples of spatial queries are calculating area and perimeter, and finding the relations of objects in terms of overlapping, proximity and neighbourhood.

The research presented in this paper is a two-phase process. The first phase develops a vector-based representation of a raster island image through edge detection-based algorithms. Island images are extracted from the whole image and vectorized as polygons. The second phase stores the island image into an object relational database as a polygon object and analyses its attributes with spatial queries. Creating geometric modelling of an island image enables more efficient analyses in terms of spatio-temporal and topological aspects.

The proposed vectorization method consists of four major processing steps: (i) Median filtering, (ii) colour space segmentation, (iii) region grouping and labelling, and (iv) line tracing. Median filtering is used to reduce noise. After median filtering, raw satellite image is segmented into two regions, island and water, by means of colour space segmentations. Third step is the application of region grouping-labelling algorithm, which groups and labels both island and water regions as raster objects. By this way, image will only be consisted of two large continuous raster objects, island and water regions. In the line tracing stage, the pre-defined border pixels are traced with the 8-connectivity algorithm and edge pixels island images are stored into spatial databases by aid of PostgreSQL and its extension PostGIS. The technique is applied on an island satellite image. This enables storing satellite images as vector objects into spatial database systems.

Spatial databases have built-in functions and services for spatial objects defined with geometry types such as points, lines, and polygons. By this way we will be utilizing the rapidly changing and developing object-relational database communities' studies and discoveries in spatio-temporal and topological analysis for the investigation of digital satellite images. This approach also enables service qualities as well as better performance. The efficiency and feasibility of the proposed system will be examined by the various scenarios such as earthquake, erosion and accretion. Scenarios are based on measuring the effects of the natural phenomena on a selected island through the satellite images.

The remainder of this paper is organized as follows. Section 2 presents relevant research. Section 3 introduces the concept of colour space segmentation and region grouping-labelling based model broadly. Section 4 gives some experimental results regarding the stability and the

performance of the proposed approach. Section 5 provides a conclusion and suggests some future works.

2. Related research

Vectorization is a process to convert a graphic object in a raster image into a vector form. Vectorization processes focus on grouping pixels of a raster image into raw line segments, and forming lines, polylines or polygons. There are many vectorization approaches developed in various domains for different purposes (Wenyin & Dori 1999). These methods are roughly classified in six groups: (1) Hough Transform based methods (Duda & Hart 1971; Gates *et al* 2003; Jiaxin *et al* 2009; Xu *et al* 2011), (2) skeletonization based methods (Tamura 1978), (3) contour based methods (Martínez-Pérez *et al* 1987; Kirsanov *et al* 2010), (4) graph-structure based methods (Zenko & Morelli 1989), (5) mesh pattern based methods (Lin *et al* 1985), and (6) pixel tracking based methods (Dori & Liu 1999). All these vectorization approaches have some common characteristics. They all have axis representation acquisition, chain coding, and line segment approximation or polygonization.

Most of the vectorization algorithms are applied on straight line images (Mena 2006; Xumin & Xue 2011). As we deal with island images which are not composed of only straight lines like engineering drawings, we need a novel method or a composition of methods. The details on the proposed architecture are given in section 3. The aforementioned vectorization techniques have been used in many earth related researches working on object determination and extraction through satellite map images.

Although there is no such study conducted on extraction and vectorization of satellite island image, there are alternative ways suggested by researchers to extract objects from satellite images. These researches basically consist of three different extractions: roads, rivers, and coastlines etc. Dial *et al* (2001) suggested a method to extract roads from IKONOS satellite imageries. The technique is based on discovering non-road elements. To achieve this, they define a parameter from the image texture for unmasked pixels. Then, they group the cells as roads and others. Zhang (2003) has presented a knowledge-based image analysis system for extraction of 3-D road networks from stereo aerial colour images. His system consists of following components: feature and cue extraction, road primitive generation and grouping, road junction and road network construction. The experiments and the results show that more than 93% of the roads in rural areas are correctly reconstructed by the system. Zhang & Couloigner (2004) proposed a technique containing a multi-resolution analysis approach based on wavelets, road junction detection, and grouping. Mangala & Bhurud (2010) have presented ANN-Based system to extract rural roads from satellite imagery. The images are segmented by using region scalable fitting model in the segmentation step. The next step, dominant objects are extracted by means of two-stage morphological operation. Satisfactory results have been obtained by using well-trained neural network at the last step. Chaudhuri *et al* (2012a) proposed semi-automatic approach for road detection and evaluated the proposed algorithm on a variety of high resolution panchromatic remotely sensed imageries of IKONOS, QuickBird, CARTOSAT-2A satellites. Their approach includes some basic image processing algorithms such as image enhancement, segmentation, hole filling, and area based filtering. Ahmed & Rahman (2011) presented a cost effective and modified method for road intersection from high resolution satellite images on urban, semi-urban and rural area taken from internet (Google Map). Their modified method has two sub-steps which one of them is automatic road extraction and the other one is road intersection detection. Proposed method able to detect single, multiple, intersected, and branched roads efficiently.

Luo *et al* (2007) proposed an automatic knowledge-based extraction tool to extract bridges over water from high-resolution panchromatic IKONOS satellite images. Chaudhuri & Samal (2008) suggested an approach for detecting bridges over water bodies from multispectral imagery. Firstly, multispectral image is first classified by means of multi-seed supervised classification technique and then classified image is categorized into a three-level image like water, concrete, and background. Finally, bridges are recognized by using a knowledge-based approach. Shah *et al* (2011) have suggested a method utilizing algorithms like edge detection, image erosion and feature extraction to extract water content and algorithms such as K-means clustering, Hill Climbing algorithm, colour histogram and image thresholding to detect the condition of river like normal, drought or flood. Chaudhuri *et al* (2012b) proposed a statistical approach to detect dark curvilinear features due to ocean disturbances caused by wind, movements of surface or underwater objects, and oil spill from SAR images. The efficiency and feasibility of the proposed approach is examined on both synthetic images with various levels of added Gaussian noise and on actual SAR images from ERS-2, SEASAT, ENVISAT, and RADARSAT. Liu and Jezek (2004) suggested a comprehensive approach to effectively and accurately extract coastlines (land/water boundaries) from satellite imagery. Their method consists of a sequence of image processing algorithms (Levenberg-Marquardt method, Canny edge detector, local thresholding, and a vector-based line coverage of the coastline). Despite the successful applications of their method on both radar and optical satellite images, the success of their method still depends on whether considerable contrast exists between water and land masses.

In our previous work (Eken & Sayar 2012), vectorization is realized through the use of edge detector. The proposed method is based on applying Canny operator (Canny 1986) to greyscale image to find out the edges, and then, converting the edges into binary-level with Otsu threshold technique (Otsu 1979). In the next step, it applies 4-way scan to find out the border pixels of the image, and create a convex hull covering those pixels by aid of Gift Wrapping algorithm (Jarvis 1973). Convex representation of the island image includes some parts which are not existed in the actual island image. So, results of spatial queries such as area and length are not close to the real results.

3. Methodology

The proposed architecture to determine spatio-temporal changes in island satellite imagery with vectorization and spatial queries are summarized in the following steps:

- Removing noise from the image to improve the results of later processing
- Eliminating objects that are not islands (pixels of water regions)
- Region grouping and labelling
- Tracing boundaries for polygon creation
- Modelling a polygon representation of island image and storing into the spatial database
- Testing efficiency of the proposed system for sample scenarios

3.1 Data Pre-processing and smoothing

Most smoothing techniques remove noise, but they blur edges. It is important to maintain edges, because edges are critically important to the visual appearance of images. We utilize median filtering to reduce noise over the sea surface. The median filtering preserves certain edge shapes. Applying multiple median filters lessens noise and brightness. To avoid the computational

expense, median filtering was used with small size (3×3) window (Schalkoff 1992). The filter is considered as following 3×3 image window:

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{vmatrix},$$

where $a_1 - a_9$ are grey levels of each pixel in the filter window.

3.2 Filtering objects with colour model conversions

The satellite images used in the framework are LandSat-7 satellite images in full colour forms. In this study, three different colour models are used. These are RGB (Red, Green and Blue), YCbCr and HSV. RGB colour model is used to obtain the characteristic features of the images. YCbCr model is mostly used in video and image compression schemes such as MPEG and JPEG. Y stands for luminance component that represents the brightness in an image. Cb and Cr stand for the difference between a reference value and the blue and red components, respectively. In HSV colour model, H (Hue) is an attribute of a colour which distinguishes it from other colours, S (Saturation) is the ratio of colourfulness relative to its own brightness and V (Value) is average of Red, Green and Blue values of a specific pixel. These parameters can be derived from RGB colour model as follows.

$$Y = 65.481R + 128.553G + 24.966B + 16 \quad (1)$$

$$C_b = -37.797R - 74.203G + 112.000B + 128 \quad (2)$$

$$C_r = 112.000R - 93.786G - 18.214B + 128 \quad (3)$$

$$H = \begin{cases} \theta, & B \leq G \\ 360 - \theta, & \text{otherwise} \end{cases} \quad (4)$$

$$\theta = \cos^{-1} \left\{ \frac{0.5 [(R - G) + (R - B)]}{\sqrt{(R - G)^2 (R - B) (G - B)}} \right\} \quad (5)$$

$$S = 1 - \frac{3}{(R + G + B)} a \quad (6)$$

where 'a' is the minimum of R, G and B.

$$V = \frac{(R + G + B)}{3}. \quad (7)$$

In any satellite image, there are various kinds of coloured regions having different characteristic features. These regions can be grouped and named as roads, buildings, water regions, trees and sandy regions. This classification and characteristic features of the regions in satellite maps are investigated by (Bong *et al* 2009). Also, Erick (2009) suggested a method of high-resolution image segmentation using graphs and morphological operations. After this segmentation, he could extract significant cartographic elements such as urban areas, streets or green areas. Illustration of the classification is given in figure 1.

Figure 1. Five groups of regions in a satellite image.

Those regions shown in figure 1 can be easily differentiated with naked eyes by their luminance and colour values. Computers can also be programmed and given ability to analyse the colour space components (hue value, saturation and luminance) of the satellite images. HSV and YCbCr colour space components are obtained from RGB values of corresponding full colour satellite image by applying the formulae, which are given in equations numbered 1 through 7 above. The classified regions and their component values in HSV and YCbCr colour spaces are given in table 1 (Bong *et al* 2009). Cb, Cr and Value (V) components are not needed in the proposed architecture. Therefore, they are ignored and not listed in the table.

Table 1. Five groups of regions and ranges of their component values.

	Trees	Roads	Buildings	Sandy regions	Water regions
Luminance	25–100	110–160	50–150	100–200	50–80
Hue	0.05–0.3	0.05–0.2	0–0.1	0.05–0.1	0.1–0.45
Saturation	0.05–0.4	0.1–0.25	0.1–0.5	0.3–0.5	0.05–0.2
Intensity	0.05–0.5	0.15–0.8	0.2–0.6	0.4–1	0.2–0.3

In this study, our aim is to extract out the island object from the satellite images by eliminating water regions. By using table 1, we determine that we need to

- (i.) Set a threshold value for luminance $50 \leq$ and ≤ 80 .
- (ii.) Set a threshold value for hue $0.1 \leq$ and ≤ 0.45 .
- (iii.) Set a threshold value for saturation $0.05 \leq$ and ≤ 0.2 .
- (iv.) Set a threshold value for intensity $0.2 \leq$ and ≤ 0.3 .

After converting satellite image (in RGB) into YCbCr and HSV colour models, the pixel elimination process is performed. Each pixel belongs to only one of the classified components listed as columns in table 1. Elimination process is iteration and at each iteration an unwanted pixel (a pixel of water regions) is filtered out. An unwanted pixel is determined by applying the below boolean statement.

$$(50 \leq Y \text{ and } Y \leq 80) \text{ and } (0.1 \leq H \text{ and } H \leq 0.45) \text{ and } (0.05 \leq S \text{ and } S \leq 0.2) \\ \text{and } (0.2 \leq V \text{ and } V \leq 0.3) \quad (8)$$

At the end of iteration, a binary image $b(i, j)$ is created.

$$b(i, j) = \begin{cases} 255, & \text{if passes iteration} \\ 0, & \text{otherwise} \end{cases} \quad (9)$$

$b(i, j)$ for a sample image is illustrated at figure 2a.

Figure 2. Region grouping and labelling: (a) segmented image with island pixels (black) and water pixels (white); (b) four groups of water raster objects (1, 2, 3 and 4); (c) noisy raster objects (1, 2 and 3) in island are removed; (d) four groups of island raster objects (5, 6, 7 and 8); (e) noisy island raster objects (5, 7 and 8) are removed, and only 2 raster objects (water numbered as 4) and (island numbered as 6) are left; and (f) border pixels are determined.

Water regions are successfully separated from roads, buildings and sandy regions, but it is not possible to differentiate the pixels representing trees from the pixels representing water regions completely. This is due to the fact that there is an overlap in the range of possible luminance, hue, saturation and intensity values. See table 1 for the columns labelled as trees and water regions. Because of that, there is still a possibility of misclassification, and tree regions and water regions still can be overlapping. To minimize this misclassification, another feature (R, G, B) can be added for filtering. In accordance with RGB model, green threshold value is set to the range of [50–255], and red and blue values are set to zero to obtain tree regions accurately. Also, border regions of both buildings and water area overlap due to their hue features, but it is in very small range and can be ignored.

3.3 Region grouping and labelling

$b(i, j)$ is created by the image segmentation process given in section 3.2 and consists of multiple image regions. At the end of the process, there are two types of raster objects, water and island. Any contiguous regions of island or water in $b(i, j)$ can be grouped into a higher level of representation—raster objects (figure 2). Raster objects are characterized by their shape, size and positions and labelled with unique id-numbers. Defining them as raster objects enables them to be differentiated from other objects easily.

The implementation of region grouping-labelling technique is based on flood-fill concept. The flood-fill method enables defining connected regions, and filling with a colour. Algorithm runs recursively until all the inner pixels are coloured or labelled. It is applied on the bitmap image $b(i, j)$.

In the island area, there might be small water areas such as streams, lakes, wetlands, or any other water-like image noises. Due to the fact that they are never going to be included in contour lines, they should be dissolved in to the island area. Dissolution is illustrated in figure 2. When figure 2b and c are compared, dissolved regions can be seen as numbers 1, 2 and 3. Similarly, there might be small noisy raster objects in water area. Examples of these small areas can be given as icebergs, ships, very small islands, rocky islets and any other image noises. Again these are not true raster objects and need to be dissolved into the water area. However, it might be reasonable for someone to wonder if such a small object is actually a very small island (for example, a rocky islet on the sea) and if its removal from the image will eventually produce errors. The study in this paper covers vectorization of large islands, in which only the outer boundary pixels are calculated. Since only the outer boundaries are calculated, small noisy objects in watery regions are expected not to affect the result negatively. This case is also illustrated in figure 2. When you compare figure 2d and e dissolved regions can be seen as numbers 5, 7 and 8. At the end of the region grouping-labelling process, two large continuous raster objects (island and water) are obtained (figure 2e).

After having done with region grouping and labelling, the morphological closing operation (Parker 2010) is applied on the image to generalize the jagged boundaries of image object. It makes the boundary morphologically smoother.

3.4 Tracing border pixels and polygon creation

The contour is a boundary of object, a population of points (pixels), and separating object from a background. Contours are formed as polygons and created by tracing the border pixels.

Background pixels are assumed to be water pixels, and when one of the neighbours of any island pixel is water pixel, then that island pixel is marked as boundary pixel. Boundary pixels are traced by a 3×3 neighbourhood window. Boundary pixel values are set to 255. If it is not set to 255, it takes the default value 0 (background pixel). At the end of this process, boundary pixels are extracted for the island image.

$$c(i, j) = \begin{cases} 255, & \text{if } b(i, j) = 255 \text{ \& } (b(i-1, j) \parallel b(i+1, j) \parallel b(i, j-1) \parallel b(i, j+1) = 0) \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

After setting each pixel either 0 (background pixels) or 255 (border pixels), a raster image of border pixel is created. $c(i, j)$ for a sample image is illustrated in figure 2f. The next step does tracing of the border pixels (also called edge-pixels) and recording their coordinate values to create a polygon object. The aim is to create a polygon from the border pixels calculated in the earlier steps. To create a correct polygon for an image, points need to be listed in a correct order and starting and ending points in the list should be the same.

Algorithm starts with a randomly selected border pixel and selects a direction (figure 3). Determining the next pixel along the tracing path is based on the principle of minimizing angular change. The angle of the line connecting the previous and current pixel is used as a reference value for selecting the next pixel. The next pixel is going to be one causing the minimum angular change compared to the reference line. If there is no such pixel, the algorithm tries to find out another one according to the principles of minimum angular change. The tracing process goes on recursively until there is no candidate pixel available (figure 3b).

Odd numbers (1, 3, 5, 7) represent horizontal and vertical directions, even numbers (0, 2, 4, 6) represent diagonal directions as shown figure 3a. The angle of the line connecting the previous and current pixel is used as a reference value for selecting the next pixel. If there is no change in angle, reference value will be 0; otherwise this value will be some factors (1–8) of 45 degree. So, zero value represents the smallest value of angle change. In other words; if the considering pixels are in same line, the value of minimum angular change will be zero. So, this considering pixel must be chosen as a next pixel. For example; in figure 3b top left-most pixel and its 4-number directed pixel are in same line or all vertical pixels are in same line.

3.5 Computational complexity and efficacy of the proposed algorithm

Computational complexity of each steps of the proposed algorithm is as follows:

- In median filtering step, the median of pixel values inside a filter window has to be calculated. In its straight-forward implementation the pixel values are sorted and the value located at $(k^2/2)^{\text{th}}$ position is taken, where k is the filter size, i.e., the size of one side of the square filter window. If quicksort is used, the computational complexity of filter is $O(k \lg k)$ operations per pixel of result image. Total cost for entire image is $O(n^2 k \lg k)$, but we can eliminate $k \lg k$ because $k \ll n$. Finally, median filtering step costs $O(n^2)$.
- Filtering objects with colour model conversions requires test of Equation (8) on entire image. So its cost is $O(n^2)$.
- In the region grouping and labelling step, connected components between pixels in binary images are assigned a unique label to separate it from others. Algorithm performs two

passes. In the first pass, the image is processed from left to right and top to bottom to generate labels for each pixel and all of the equivalent labels are stored in a pair of arrays. In the second pass, each label is replaced by the label assigned to its equivalence class. This algorithm can be performed in $O(n^3)$.

- To trace the contour (also called border and boundary) of a single object, we can use a procedure considering minimum angular change. Because this procedure performs on entire image, its cost is $O(n^2)$.

Total computational complexity of proposed algorithm is $O(n^2 + n^2 + n^3 + n^2) = O(n^3)$.

3.6 Storing vector island images into spatial databases

Spatial databases store and query data for spatial objects in accordance with object relational database model. While typical databases can only store and manipulate various numeric and character types of data, spatial databases can analyse spatial data types such as point, line, and polygon. Spatial data can be geo-spatial maps, rivers, roads, and islands that can be represented with points and lines. How many towns Kocaeli city has and which district has the highest number of population in Istanbul city can be given as examples of typical SQL queries. On the other

Figure 3. Tracing edge pixels (black cells): (a) 8-connectivity neighbours; and (b) edge lines - island shoreline.

hand, a spatial query is a query that returns features based on their spatial relationship with query geometry. This includes searching for features given an extent (e.g., finding all of the streets within an envelope) or searching for features based on their relation to other features (e.g., finding all farmland intersected by a particular stream). There are some other features that can be queried through spatial queries such as proximity, containment, adjacency, intersection and overlap.

For the proof of correctness and analysis of the proposed technique, PostgreSQL and its PostGIS extension is used. PostgreSQL is an open source relational database management system that is initially developed at University of California, Berkeley. It is under the PostgreSQL License (TPL). PostGIS is an open source software program that adds support for geographic objects to the PostgreSQL database.

A sample scenario for creating a relational table and putting a spatial polygon object into the table is given below. These are regular SQL queries enhanced with spatial capabilities.

A table with a spatial dataset is created below:

```
CREATE TABLE polygon_region (name varchar, coordinate geometry);
```

Here, `polygon_region` represents name of the table, and 'name' and 'coordinate' are the names of two separate columns. First one consists of characters and it represents names of island objects and the second one represents spatial data with geometry type. As shown in this example, a polygon region table is created with a polygon geometry column and a name. Once the table with spatial data attributes have been created, GIS data (geometry) can be inserted to the table by using an SQL insert statement as below:

```
INSERT INTO polygon_region VALUES ('island', 'POLYGON((0 10,15 10,15 -5,10 -5,10 0,0 10))');
```

To view GIS data in the table, an SQL select statement can be used. In the same way, any other database query can be constructed as combination of return values, functions, and Boolean tests.

After storing the GIS data (e.g., island image in a vector form) into a spatial database, spatial operators listed in table 2 can be applied on the image. In other words, we utilize spatial functions such as Distance, ST_Intersects, ST_Contains and ST_Within to narrow down the results of our search. Most spatial queries include spatial functions to test and analyse spatial data.

Table 2. Some spatial operators.

Operators	Operator functions
Envelope	Returns the minimum bounding rectangle.
Closed	Returns true if the given point set forms a polygon. In other words, start and end points are the same.
Boundary	Returns the boundary of the geometry.
Intersection	Returns true if geometries are not disjoint.
Contains	Returns true if the given geometry object enclose another geometry object.
Overlap	Tests whether the interiors of 2 geometric objects have non-empty intersections.
Area	If the object has surface, such as polygon, then the function returns the area of the object. Returns the area of the surface if it is a polygon or multi-polygon.
Length	If the object is line or line-string, then the function returns the length of the object.

4. Experimentation and analysis

The LandSat-7 island images, which are used in this study, were captured by LandSat-7 satellites. LandSat-7 data products are available from the USGS (U.S. Geological Survey) Centre for Earth Resources Observation and Science (EROS). Images were obtained from online web site (<http://glovis.usgs.gov/>). The web site offers satellite images of different locations and at different time intervals for registered users.

This chapter presents an application of the proposed methods given in section 3 on to the Cyprus Island. Cyprus is the island with the third largest area in the Mediterranean Sea, measuring 240 km in length and 100 km in width at its widest point. The actual area of the island is 9251 km² and it has 648 km of coastline. Cyprus is located within the latitudes of 34° and 36° N, and longitudes 32° and 35° E. The selected region of the island is shown in (figure 4).

Figure 5 illustrates the steps of the proposed technique (see section 3) applied on the satellite imagery given in figure 4. First, median filter is applied to reduce noise over the sea surface. Then, colour space conversions (from RGB model to HSV and YCbCr colour model) are applied to filtered raster satellite image to obtain the characteristic features of the satellite image. According to HSV and YCbCr component values, each pixel in the image is set to either 255 or 0. The number 255 represents an island pixel, and 0 represents a water pixel. Figure 5b shows the segmented image.

The segmented image goes through the region grouping and labelling operations. This process groups and labels water pixels as individual raster objects and resulting tiny noisy images are removed (see figure 5c). Similarly, the same process is applied on the island pixels as shown in figure 5d. This step in the algorithm is done for getting over the small noisy raster objects in the island and water regions. At the end of the process we get consistent water and island raster objects, in which tiny noisy images are removed as shown in figure 5e.

In the next step, boundary pixels are determined. Boundary pixels are coded as 255 and other pixels are called background pixels and coded as 0. By this way, boundary pixels are extracted

Figure 4. Location of study area.

Figure 5. Automated of vectorization LandSat-7 satellite imagery: (a) raw LandSat-7 image in raster form; (b) segmented image after colour space conversion; (c) image after region grouping and labelling – showing island image object; (d) image after region grouping and labelling – showing water image object; (e) image obtained after c and d are applied successively; and (f) island/water boundary lines as a polygon.

(figure 5f). Finally, the coordinates of island boundary pixels are converted to a vector format, which is a polygon, and stored into a spatial database.

Some scenarios such as earthquake (Scenario 1), erosion (Scenario 2) and accretion (Scenario 3) are applied on a satellite image for an island and their effects are calculated via proposed techniques. The effects of these phenomena named as scenarios are measured by comparing the values of circumferences and area of the island before and after the simulations. The area and circumference values are obtained by spatial queries on vector representation of the image. Image is stored as a polygon in an object relational PostgreSQL database.

The real area and length (or circumferences) are calculated from the polygon coordinates according to the formula given below:

$$\text{Real_length} = \text{Map scale value} * \text{Length calculated from polygon in database}$$

$$\text{Real_area} = (\text{Map scale value})^2 * \text{Area calculated from polygon in database}$$

Sample spatial queries performing area and length of vector island image are listed below:

```
SELECT Round(ST_Area(formerpolygon)), Round(ST_Area(latterpolygon)), area FROM dbpolygon WHERE ad='Cyprus'
```

```
SELECT Round(ST_length(ST_Boundary(formerpolygon))), Round(ST_length(ST_Boundary(latterpolygon))), contour FROM dbpolygon WHERE ad='Cyprus'
```

In Scenario 1, we obtained artificial earthquake with translocation operation on the island of Cyprus image. The magnitude of slip was assessed by comparing centroids of polygons obtained

Figure 6. Some scenarios applied on satellite image: (a) original vector satellite image; (b) slipping; (c) erosion; (d) accretion.

from images before and after translocation (see the arrow in figure 6b). The centroid is defined as a location (c_x, c_y) and coordinate values c_x and c_y are calculated as below.

$$c_x = \frac{1}{6A} \sum_{i=0}^{N-1} (x_i + x_{i+1}) (x_i y_{i+1} - x_{i+1} y_i) \quad (11)$$

$$c_y = \frac{1}{6A} \sum_{i=0}^{N-1} (y_i + y_{i+1}) (x_i y_{i+1} - x_{i+1} y_i), \quad (12)$$

where A is the area of the polygon and it is defined as follows.

$$A = \frac{1}{2} \sum_{i=0}^{N-1} (x_i y_{i+1} - x_{i+1} y_i). \quad (13)$$

Here, N is the number of vertices represented as (x_i, y_i) . Since the vertices are assumed to form a polygon, the last vertex is ignored. In a polygon, last and first vertices are the same.

In scenario 2, artificial erosion has been applied on the island by deleting a portion of the island of Cyprus image. Then, we have investigated how much reduction has occurred by querying the area and the length of the polygon. (see the circle in figure 6c)

In Scenario 3, artificial accretion was obtained by adding a portion to the island of Cyprus. Then, we investigated how much expansion has occurred by querying the area and the length on polygons (see the circle in figure 6d).

The real area covered by Cyprus is indeed 9251 km^2 and its length is 648 km . Calculated area of the island that was recognised on the raster image by the proposed technique is 9745 km^2 and calculated length is 712 km . The accuracy of the proposed method is about 95% in terms of area and 90% in terms of length/circumference. When we look at the values in table 3 we see that there is no change in the area and circumference values of the island image before and after the application of scenario 1 (slipping). In this scenario, degree of slip was determined as 25 km by comparing the centroids of the polygon representing the image before and after the translocation. In scenario 2 (erosion), it is shown that the area and circumference of the image has decreased

Table 3. Results of simulations.

SCENARIOS	Area (km^2)		Length (km)	
	Before	After	Before	After
Slipping	9745	9745	712	712
Erosion	9745	9311	712	689
Accretion	9745	9876	712	726

around 434 km² and 23 km, respectively. When it comes to scenario 3 (accretion), it is seen in the table that area of the image has expanded 131 km² and a length of the image has increased 14 km. In addition, there is no slipping in scenario 2 and scenario 3 due to the lack of translocation.

Compared to our initial work on the same problem (Eken & Sayar 2012), the work presented in this paper gives much better results. In the previous work, the area and circumference of Cyprus were calculated as 15715 km² and 543 km respectively. Here, the calculated values are 9745 km² and 712 km. Therefore, this paper represents more accurate results compared to the actual values, which are 9251 km² and 648 km.

5. Summary and conclusion

Edge detection and vectorization processes on island satellite images are more difficult than expected. Due to the inconsistent intensity contrast between water and island regions, vectorization of island raster objects is a difficult job when the general-purpose image segmentation or edge detectors are used. This paper proposed a technique to convert a raster image into a vector data as an object by using colour space conversion, region grouping, and line tracing. This enabled storing satellite island images into spatial database systems as polygons and utilizing built-in spatial tools and functions developed by spatial database community. The proposed technique was applied on a real satellite image and effectiveness of the proposed technique was examined with the simulations. We have applied three different kinds of simulations; earthquake, erosion and accretion. We have tested queries on the image before and after the simulation and determined the effects of the natural phenomena on the island of Cyprus through the spatial queries. The island of Cyprus is stored in relational spatial databases as a polygon, which is created with the proposed technique. Spatial queries are summarized as finding out how much slip has occurred after the earthquake, how much reduction has occurred after the erosion, and how much expansion has occurred after the accretion. After the applications of the proposed techniques on some scenarios, we saw that accuracy of the system is about 95% in terms of area and 90% in terms of length/circumference.

Although the proposed technique proved to be effective on extracting island images objects from the LandSat-7 satellite imageries, it should be noted that the quality of image sources still remains as an important factor. In such an application, level of contrast between the water and island regions should be considerably different. Moreover, the homogeneity degrees of the regions also affect the success of the results.

We plan to make further enhancements on the proposed technique in the future versions. Currently, we assume objects to be extracted are continuous and identified by grouping and labelling process to differentiate the true island border from other object boundaries. In the future, other attributes of the raster objects such as shape, texture and relative position can also be used. The resolution of the source data affects the success rate of the proposed methodology. We will experiment with higher resolution data sets smaller than 15 m.

Acknowledgements

This work was supported by Kocaeli University Scientific Research and Development Support Program (BAP) in Turkey (project number is 2013/012). Also, we thank the referees for their valuable comments to improve our paper.

References

- Ahmed B and Rahman Md F 2011 Automatic Road Extractions from High Resolution Satellite Imagery Using Road Intersection Model in Urban Areas. *Computer Eng. and Intelligent Syst.* 2: 82–93
- Bong B L, La K C and Joseph A 2009 Automatic Road Network Recognition and Extraction for Urban Planning. *World Academy of Science, Eng. and Technol.* 29: 209–215
- Canny J 1986 A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 8: 679–698
- Chaudhuri D and Samal A 2008 An Automatic Bridge Detection Technique for Multispectral Images. *IEEE Transactions on Geoscience and Remote Sensing* 46(9): 2720–2727
- Chaudhuri D, Kushwaha N K and Samal A 2012a Semi-Automated Road Detection From High Resolution Satellite Images by Directional Morphological Enhancement and Segmentation Techniques. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5: 1538–1544
- Chaudhuri D, Samal A, Agrawal A, Sanjay Mishra A, Gohri V and Agarwal R C 2012b A Statistical Approach for Automatic Detection of Ocean Disturbance Features From SAR Images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 5(4): 1231–1242
- Dial G, Gibson L and Poulsen R 2001 Ikonos satellite imagery and its use in automated road extraction. In: A. Gruen, O. Kuebler and P. Agouris (Eds.), *Automated Extraction of Man-Made Objects from Aerial and Space Images (III)*. Publishing, Lisse, The Netherlands, pp. 349–358
- Dori D and Liu W 1999 Sparse Pixel Vectorization: An Algorithm and Its Performance Evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 21: 202–215
- Duda R O and Hart P E 1971 Use of the Hough Transformation to Detect Lines and Curves in Pictures. *Communications of the ACM* 15: 11–15
- Eken S and Sayar A 2012 Vectorization and Spatial Query Architecture on Island Satellite Images. *Procedia Comput. Sci. J.* 2: 37–43
- Erick L O 2009 High Resolution Images: Segmenting, Extracting Information and GIS Integration. *World Academy of Science, Engineering and Technology* 54: 172–177
- Gates J W, Haseyama M and Kitajima H 2003 An Algorithm for extracting lines and circles with High-speed and accuracy. *16th International Conference on Vision Interface*, Canada
- Jarvis R A 1973 On the identification of the convex hull of a finite set of points in the plane. *Information Processing Letters* 2: 18–21
- Jiaxin S, Dong D, Xinjie Z and Li W 2009 Weld Slim Line Defects Extraction Based on Adaptive Local Threshold and Modified Hough Transform, *2nd International Congress on Image and Signal Processing*, China, pp 1–5
- Kirsanov A, Vavilin A and Jo K H 2010 Contour-based algorithm for vectorization of satellite images, *International Forum on Strategic Technology*, Moscow, Russia, 241–245
- Lin X, Shimotsuji S, Minoh M and Sakai T 1985 Efficient Diagram Understanding with Characteristic Pattern Detection. *Computer Vision, Graphics and Image Processing* 30: 84–106
- Liu H and 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
- Luo J, Ming D, Liu W, Shen Z, Wang M and Sheng H 2007 Extraction of bridges over water from IKONOS panchromatic data. *International Journal of Remote Sensing* 28(16): 3633–3648
- Mangala T R and Bhirud S G 2010 An Effective ANN-Based Classification Syatem for Rural Road Extraction in Satellite Imagery. *European J. Scientific Res.* 47: 574–585
- Martínez-Pérez M P, Jiménez J and Navalón J L 1987 A thinning algorithm based on contours. *Computer Vision, Graphics, and Image Processing* 39: 186–201
- Mena J B 2006 Automatic vectorization of segmented road networks by geometrical and topological analysis of high resolution binary image. *Knowledge-Based Systems* 19: 704–718
- Otsu N 1979 A threshold selection method from gray-level histograms. *IEEE Transactions on Systems and Cybernetics* 9: 62–66
- Parker J R 2010 *Algorithms for Image Processing and Computer Vision*. 2 Edition ed. Wiley.

- Schalkoff R J 1992 *Digital Image Processing and Computer Vision*. Wiley
- Shah V, Choudhary A and Tewari K 2011 River extraction from satellite image. *Int. J. Comput. Sci. Issues* 8: 386–391
- Tamura H 1978 A Comparison of line thinning algorithms from digital geometry viewpoint, *4th International Conference on Pattern Recognition* Kyoto, Japan, 715–719
- Wenyin L and Dori D 1999 From Raster to Vectors: Extracting Visual Information from Line Drawings. *Pattern Pattern Analysis and Applications* 2: 10–21
- Xu S, Liu J, Wang Y, Han L and Zhang Y 2011 Straight line extraction via multi-scale Hough transform based on pre-storage weight matrix. *International Journal of Remote Sensing* 32: 8315–8330
- Xumin L and Xue Y 2011 Research on Vectorization of Weather Radar Image. *Procedia Eng.* 15: 1298–1302
- Zenzo S D and Morelli A 1989 A useful image representation. V. Cantoni (Ed.), *5th International Conference on Image Analysis and Processing*, 170–178
- Zhang C 2003 *Updating of Cartographic Road Databases by Image Analysis*. Unpublished Ph.D. Dissertation, Swiss Federal Institute of Technology Zurich
- Zhang Q and Couloigner I 2004 A Wavelet Approach to Road Extraction from High Spatial Resolution Remotely-Sensed Imagery. *GEOMATICA* 58: 33–39