

Preliminary Paper  
Received: 15-05-2009  
Accepted: 15-06-2009

# Area Preserving Cartographic Line Generalization

**Dražen TUTIĆ, Miljenko LAPAINE**

University of Zagreb, Faculty of Geodesy, Zagreb, Croatia

e-mail: dtutic@geof.hr, mlapaine@geof.hr

84

**Abstract:** Cartographic line generalization is a generalization with the primary purpose of obtaining a line shape suitable for representation on maps at smaller scales than the scale of the given lines. Existing literature also differentiates model generalization and geographic generalization, each with a primary purpose. A new algorithm for cartographic generalization with distinct property of area preservation is described in this paper. The algorithm yields generalized lines with a subset of points from given lines together with new points. For that reason, it can be considered partially as simplification and partially as a smoothing process. Area preservation is not often considered in research on cartographic line generalization. Such research mostly stresses the importance of the line shape. A detailed description of the new algorithm is given in this paper, as well as the results of the application on coastlines. A comparison with manual line generalization, the Douglas-Peucker simplification algorithm, the Boyle smoothing algorithm and the combination of the latter two is also given. The tests show the new algorithm yields satisfactory results for cartographic line generalization. The property of area preservation is especially suitable for application in GIS when attributes are represented or analysed in regard to object areas.

**Key words:** cartographic generalization, lines, area preservation

## 1 Motivation

Results of automatic vectorisation of linear features from raster maps can contain irregularities which cause unnecessary zigzags in vector lines (Fig. 1).

A simple solution of the problem described is to delete segments shorter than a given length and the angle to the one neighbouring segment is larger than  $180^\circ$  and the angle to the other neighbouring segment is smaller than  $180^\circ$ , measured on the same side.

Based on that problem, the idea for applying such a method or a similar one to a more general task of cartographic line generalization emerged.

## 2 Introduction

Cartographic line generalization is one of the major research areas in the more general area of cartographic generalization. The number of existing papers for the area of cartographic line generalization is large, especially since line generalization is also performed with computers. Valuable bibliographic resources for cartographic generalization can be found at (URL1, URL2).

There exist a number of methods for automatic line generalization, each using a different approach and methodology. Cartographic theory differentiates sub-processes in line generalization, such as filtering, simplification, smoothing, etc. An overview of more important algorithms

Prethodno priopćenje  
Primljeno: 15-05-2009.  
Prihvaćeno: 15-06-2009.

# *Kartografska generalizacija linija sa svojstvom čuvanja površina*

**Dražen TUTIĆ, Miljenko LAPAINE**

Geodetski fakultet Sveučilišta u Zagrebu, Zagreb

e-pošta: dtutic@geof.hr, mlapaine@geof.hr

85

**Sažetak:** Pod kartografskom generalizacijom linija podrazumijeva se ona koja ima prvenstvenu namjenu dobivanje takvog oblika linija koji je pogodan za prikaz na kartama sitinijeg mjerila od mjerila zadanih linija. Osim takve generalizacije postojeća literatura razlikuje i modelnu generalizaciju i geografsku generalizaciju koje imaju i odgovarajuću namjenu. U radu je opisan novi algoritam za kartografsku generalizaciju linija koji ima svojstvo čuvanja površina. Osim svojstva čuvanja površina, algoritam daje generalizirane linije koje se općenito sastoje od podskupa polaznih točaka i skupa novih točaka. Zbog toga se algoritam može smatrati kombinacijom pojednostavljenja i izgladivanja. Svojstvo očuvanja površina rjeđe je zastupljeno u istraživanjima kartografske generalizacije linija. Težina takvih istraživanja usmjerena je prema obliku linije. U radu je detaljno opisan novi algoritam te su dani rezultati primjene novog algoritma na primjeru obalnih crta te usporedba s ručnom generalizacijom, Douglas-Peuckerovim algoritmom za pojednostavljenje, Boyleovim algoritmom za izgladivanje i kombinacijom posljednja dva algoritma. Testovi pokazuju da algoritam daje zadovoljavajuće rezultate u smislu kartografske generalizacije, a svojstvo očuvanja površina ističe se kao posebno pogodno za primjenu u GIS-u kada se atributi prikazuju ili analiziraju u odnosu na površine objekata.

**Ključne riječi:** kartografska generalizacija, linija, očuvanje površina

## 1. Motivacija

Prilikom automatske vektorizacije linijskog sadržaja s rasterskih karata rezultat vektorizacije može sadržavati nepravilnosti koje dovode do nepotrebne drhtavosti ili vijugavosti vektorskih linija (slika 1).

Jednostavno rješenje takvog problema je brisanje onih segmenata koji su kraći od neke zadane duljine, a prijelomni kut sa susjednim segmentima je jednom veći od  $180^\circ$ , a drugi put manji od  $180^\circ$ .

Na temelju tog problema proizašla je ideja da se primjeni takva ili slična metoda i na općenitiji zadatak kartografske generalizacije linijskog sadržaja.

## 2. Uvod

Kartografska generalizacija linija jedno je od važnih područja istraživanja kartografske generalizacije. Radovi iz tog područja su brojni, pogotovo od kada se generalizacija linija obavlja i s pomoću računala. Dobri izbori bibliografije na temu kartografske generalizacije nalaze se na (URL1, URL2).

Postoji veći broj metoda za automatsku generalizaciju linija koje upotrebljavaju različite pristupe i metodologije. Kartografska teorija generalizaciju linija dijeli na filtriranje, pojednostavljenje, izgladivanje i sl. Pregled algoritama može se naći npr. u (McMaster 1987), a širi pregled kartografske generalizacije pa tako i generalizacije linija može se naći u (Mackannes et. al. 2007).

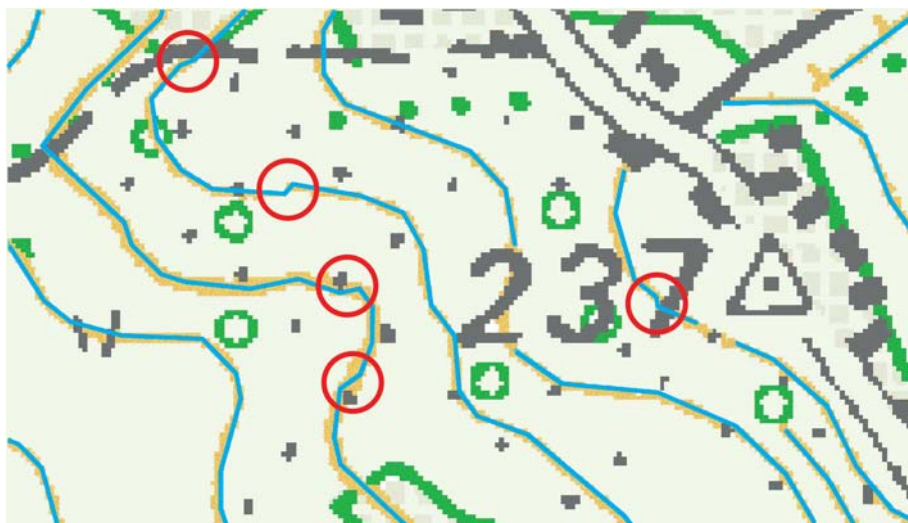


Fig. 1. Automatically vectorised linear features from raster map with marked "irregularities". Red circles indicate the segments which change the curvature of the vectorised lines when it does not correspond to the raster

Slika 1. Prikaz automatski vektoriziranoga linijskog sadržaja karte s naznačenim "nepravilnostima". Crveni kružići označuju segmente koji mijenjaju zakrivljenost vektorizirane linije kada to ne odgovara rasterskoj podlozi

86

can be found in (McMaster 1987), and a wider consideration of cartographic generalization together with line generalization can be found in (Mackannes et al. 2007).

The whole process of cartographic generalization is difficult to define in an exact way. The human role is of great importance for the final estimation of the success of an automated process (Mackannes et al. 2007). McMaster (1986) gave an overview of different statistical measures which can be used to evaluate the results of line generalization.

The shape of a line in derived scale can not be exactly defined or preserved (in that case, there would be no generalization). Most of the methods are based on the best possible preservation of the "line shape" analysing and preserving important or critical points (Mackannes et al. 2007). A discussion of this idea can be also found in (Thapa 1988).

The property which can be preserved is the ratio of areas before and after line generalization. This paper contains a description of a new method of line generalization with the property of preserving the area of regions bounded by polylines. The algorithm is simple, which positively affects the definition, implementation and the processing time. Measures of "line shape" are not going to be considered and a user can do a subjective evaluation of the results on the examples provided.

### 3 Introductory Definitions

Let there be plane  $\Pi$  with a rectangular Cartesian coordinate system. The ordered pair of rectangular

Cartesian coordinates  $T = \{x, y\}$  defines the point in the plane, and it is going to be designated as  $T(x, y)$ .

Let  $V = \{T_i(x_i, y_i) \in \Pi; i = 1, 2, \dots, n\}$  be the ordered set of  $n$  points in the plane  $\Pi$  such that  $T_j \neq T_{j+1}$ ,  $j = 1, 2, \dots, n-1$ ,  $T_j \neq T_{j+2}$ ,  $j = 1, 2, \dots, n-2$  and  $i \geq 4$ . The set  $V$  is going to be called the set of vertices.

Let us define the set

$$S_j = \begin{cases} T(x, y) \in \Pi; y - y_j = \frac{y_{j+1} - y_j}{x_{j+1} - x_j} (x - x_j) \\ \min(x_j, x_{j+1}) \leq x \leq \max(x_j, x_{j+1}) \\ \min(y_j, y_{j+1}) \leq y \leq \max(y_j, y_{j+1}) \end{cases}$$

The set  $S_j$  is going to be called the segment. The segment represents the line in the plane with endpoints  $T_j$  and  $T_{j+1}$ .

Let  $P = \{S_j; j = 1, 2, \dots, n-1\}$  be the ordered set of segments. The set  $P$  represents the polyline.

If  $T_1 \neq T_n$ , the polyline is open, otherwise, if  $T_1 = T_n$ , the polyline is closed (Fig. 2). It should be noted that according to the definition of the set  $V$ , the polyline can not have two identical consecutive vertices nor identical even or odd consecutive vertices and it must have at least four vertices. According to the accepted definition, the polyline can intersect itself (Fig. 3).

A polyline is one of the basic geometric elements used in digital cartography, GIS and spatial databases. It is used for approximation (more or less simplified representation) of different objects in reality (borders, rivers,

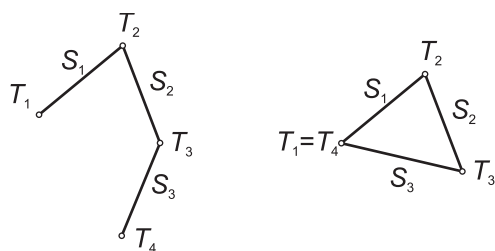


Fig. 2. Open and closed polyline  
Slika 2. Otvorena i zatvorena polilinija

Cjelokupni proces kartografske generalizacije teško je u potpunosti strogo definirati. Prilikom ocjene uspješnosti nekog automatiziranog procesa kartografske generalizacije konačnu prosudbu ipak donosi čovjek (Mackannes et. al. 2007). McMaster (1986) daje pregled različitih statističkih mjera koje se mogu upotrijebiti pri ocjeni generalizacije linije.

Oblik linije u izvedenom mjerilu nije moguće strogo definirati niti sačuvati (inače nema generalizacije). Većina metoda temelji se na što boljem očuvanju "oblika linije" zadržavanjem važnih ili kritičnih točaka (Mackannes et. al. 2007). Osvrt na tu tematiku može se naći i u (Thapa 1988).

Ono što se može sačuvati je odnos površina prije i poslije generalizacije. U ovom radu opisat će se nova metoda generalizacije polilinija koja ima svojstvo čuvanja površina područja omeđenih polilinjama. Algoritam je jednostavan, što uz lakše definiranje i njegovu implementaciju utječe i na vrijeme obrade. Mjere za "oblik linije" neće se razmatrati, a subjektivna se ocjena prepušta čitatelju kroz primjere.

### 3. Uvodne definicije

Neka je u ravnini  $\Pi$  definiran pravokutni Kartezijev koordinatni sustav. Uređeni par pravokutnih Kartezijevih koordinata  $T = \{x, y\}$  zadaje točku u ravnini, a označavat ćemo ju kraće kao  $T(x, y)$ .

Neka je  $V = \{T_i(x_i, y_i) \in \Pi; i = 1, 2, \dots, n\}$  uređeni skup od  $n$  točaka u ravnini  $\Pi$  takav da vrijedi  $T_j \neq T_{j+1}$ ,  $j = 1, 2, \dots, n-1$ ,  $T_j \neq T_{j+2}$ ,  $j = 1, 2, \dots, n-2$  i  $i \geq 4$ . Skup  $V$  nazvat ćemo vrhovima.

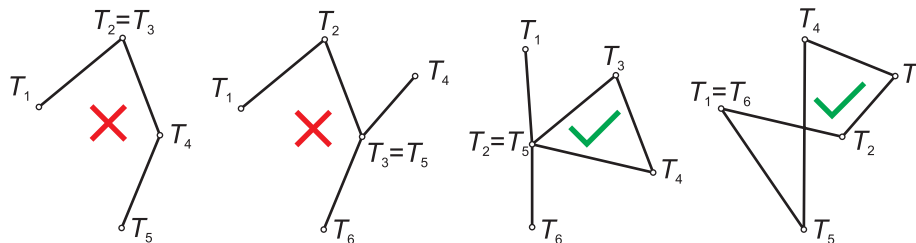


Fig. 3. Unallowed and allowed cases of polylines according to the accepted definition  
Slika 3. Neki nedopušteni i dopušteni oblici polilinije prema usvojenoj definiciji

Definirajmo skup

$$S_j = \begin{cases} T(x, y) \in \Pi; y - y_j = \frac{y_{j+1} - y_j}{x_{j+1} - x_j} (x - x_j) \\ \min(x_j, x_{j+1}) \leq x \leq \max(x_j, x_{j+1}) \\ \min(y_j, y_{j+1}) \leq y \leq \max(y_j, y_{j+1}) \end{cases}$$

Skup  $S_j$  nazvat ćemo segment. Segment predstavlja dužinu u ravnini kojoj su krajnje točke vrhovi  $T_j$  i  $T_{j+1}$ .

Neka je  $P = \{S_j; j = 1, 2, \dots, n-1\}$  uređeni skup segmenata.  $P$  predstavlja izlomljenu crtu u ravnini, tzv. polilinjiju.

Ako je  $T_1 \neq T_n$  polilinjiju ćemo nazvati otvorenom, u protivnom, ako je  $T_1 = T_n$ , polilinjiju ćemo nazvati zatvorenom (slika 2). Primijetimo da je prema definiciji skupa  $V$  polilinjija takva da ne može imati dva uzastopna vrha jednaka, niti jednake parne ili neparne uzastopne vrhove i mora imati barem 4 vrha. Prema toj definiciji polilinjija može sijeći samu sebe (slika 3).

Polilinjija je jedan od osnovnih geometrijskih elemenata koji se upotrebljava u digitalnoj kartografiji, GIS-u i bazama prostornih podataka. Njome se aproksimiraju (više ili manje pojednostavljeno prikazuju) različiti stvarni objekti (razgraničenja, rijeke, prometnice, izohipse itd.) ili prikazuju pojave (razne vrste izolija, planirane ili prevaljene rute putovanja, geografska mreža itd.).

U praksi se javlja potreba da se jedne polilinjije aproksimiraju drugima, u nekom smislu jednostavnijima. Iako se ponekad zahtijeva da se polilinjija zamjenjuje i sa u nekom smislu složenijom polilinjijom, takav zadatak se ovdje neće razmatrati.

### 4. Generalizirana polilinjija

Za zadanu polilinjiju  $P$ , s pripadajućim skupom vrhova  $V$  i segmenata  $S$ , definirajmo drugu polilinjiju  $P_G$ , s pripadajućim skupom točaka  $V_G = \{T_j(x_j, y_j) \in \Pi; j = 1, 2, \dots, m\}$  i segmenata  $S_G$ , takvu da vrijedi:

- broj vrhova u generaliziranoj polilinjiji jednak je ili manji broju vrhova u polaznoj polilinjiji, tj.  $m \leq n$ ,
- duljina generalizirane polilinjije manja je ili jednaka duljini polazne polilinjije, tj.  $d(P_G) \leq d(P)$ , gdje je duljina definirana kao

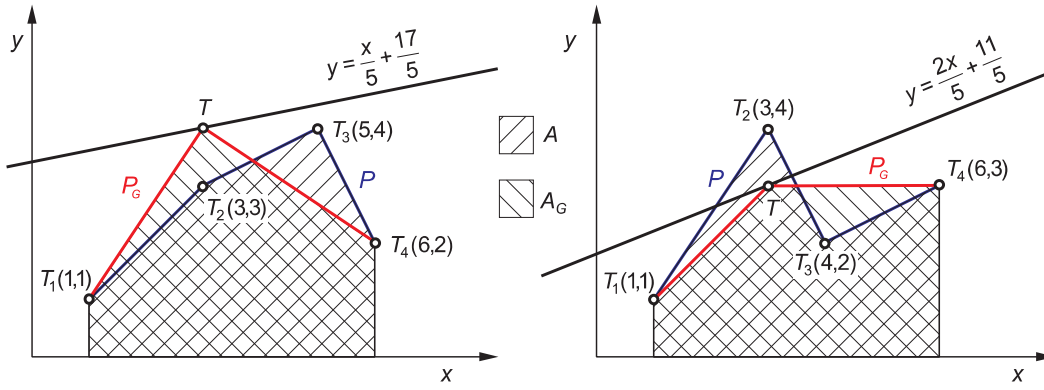


Fig. 4. The application of generalization function *e*  
Slika 4. Primjena funkcije generalizacije *e*

transportation, contours, etc.) or phenomena (isolines, planned routes, graticule, etc.).

88

When working with polylines, there is a greater need to approximate polylines with others. Even though it is sometimes required to approximate given polyline with another, that is in a way more complicated polyline, such a problem is not considered here.

### 4 Generalized Polyline

For a given polyline *P*, with its set of vertices *V* and set of segments *S*, let us define the other polyline *P<sub>G</sub>*, with its set of vertices  $V_G = \{T_j(x_j, y_j) \in \Pi; j = 1, 2, \dots, m\}$  and set of segments *S<sub>G</sub>* for which:

- the number of vertices in the generalized polyline is less than or equal to the number of vertices in the given polyline, i.e.  $m \leq n$ ,
- the length of the generalized polyline is less than or equal to the length of the given polyline, i.e.

$$d(P_G) \leq d(P), \text{ with length defined as } d(P) = \sum_{i=1}^{n-1} d(S_i),$$

$$\text{where } d(S_i) = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2},$$

- if the given polyline *P* is closed, i.e.  $T_1 = T_n \in V$ , then let the polyline *P<sub>G</sub>* be closed, i.e.  $T_1 = T_m \in V_G$ , and generally  $T_1 \in V \neq T_1 \in V_G$ ,
- if the given polyline *P* is open, then  $V_G(1) = V(1)$  and  $V_G(m) = V(n)$ , i.e. the first and last points of both polylines are identical, which means *P<sub>G</sub>* is an open polyline.

Polyline *P<sub>G</sub>* is going to be called the generalized polyline *P*.

If the length of the polyline *P<sub>G</sub>* is equal to the length of the polyline *P*, i.e.  $d(P_G) = d(P)$ , the *P<sub>G</sub>* is going to be called the equidistant generalization of *P*.

$$\begin{aligned} \text{If } A(P) &= \frac{1}{2} \sum_{i=1}^{n-1} (x_{i+1} - x_i)(y_i + y_{i+1}) = \\ &= \frac{1}{2} \sum_{j=1}^{m-1} (x_{j+1} - x_j)(y_j + y_{j+1}) = A_G(P_G) \end{aligned}$$

then *P<sub>G</sub>* is going to be called the *equiareal* generalization of *P*. *A(P)* represents the area of the oriented simple polygon (Macarol 1978, Braden 1986).

Cartographic literature differentiates the terms simplification and smoothing. For simplification  $V_G \subseteq V$ , i.e. the vertices in the simplified polyline are a subset of the vertices of the given polyline, which is not in general the case for smoothing.

If  $P_G \cap P$  is a set with an infinite number of points, then *P* and *P<sub>G</sub>* have segments which overlap at least partly. It is then possible to introduce a measure of overlap as the ratio of the overlapped segments length and the length of the given or generalized polyline. Generally, there are also going to be pairs of segments which only intersect in one point.

Otherwise, if  $P_G \cap P$  is a set of a finite number of points, the pairs of segments in *P* and *P<sub>G</sub>* only intersect with each other.

If  $P_G \cap P = \emptyset$ , then *P* (and *P<sub>G</sub>*) must be a closed polyline (because of the constraint that open polylines *P* and *P<sub>G</sub>* have identical endpoints).

The measure of the overlap and number of intersections of segment pairs can be used to define the measure of closeness of *P* and *P<sub>G</sub>*, but that is beyond the scope of this work.

### 5 Line Generalization Function

What follows is the definition of the generalization function *e* of an open or closed polyline defined in the previous section. Input for this function is an open or closed polyline with a maximum of 4 vertices.

$$d(P) = \sum_{i=1}^{n-1} d(S_i) \text{ i } d(S_i) = \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2},$$

□ ako je zadana polilinija  $P$  zatvorena odnosno  $T_1 = T_n \in V$  tada neka je i polilinija  $P_G$  zatvorena i vrijedi  $T_1 = T_m \in V_G$  i općenito  $T_1 \in V \neq T_1 \in V_G$ ,

□ za zadanu otvorenu poliliniju  $P$  zadajemo uvjet  $V_G(1) = V(1)$  i  $V_G(m) = V(n)$ , tj. prva i zadnja točka obje polilinije su jednake, a samim time i  $P_G$  je otvorena polilinija.

Poliliniju  $P_G$  nazvat ćemo generalizacijom polilinije  $P$ .

Ako je duljina polilinije  $P_G$  jednaka duljini polilinije  $P$ , tj.  $d(P_G) = d(P)$ ,  $P_G$  ćemo nazvati ekvidistantnom generalizacijom od  $P$ .

Ako vrijedi

$$A(P) = \frac{1}{2} \sum_{i=1}^{n-1} (x_{i+1} - x_i)(y_i + y_{i+1}) = \frac{1}{2} \sum_{j=1}^{m-1} (x_{j+1} - x_j)(y_j + y_{j+1}) = A_G(P_G)$$

$P_G$  ćemo nazvati ekvivalentnom generalizacijom od  $P$ . Izraz  $A(P)$  predstavlja izraz za površinu orijentiranoga jednostavnog poligona (Macarol 1978, Braden 1986).

U kartografskoj literaturi mogu se razlikovati dva pojma: pojednostavljenje (*simplification*) i izgladivanje (*smoothing*). Kod pojednostavljenja vrijedi  $V_G \subseteq V$ , dok kod izgladivanja to redovito nije slučaj.

Ako  $P_G \cap P$  sadrži beskonačno mnogo točaka  $P$  i  $P_G$  imaju segmente koji se barem jednim dijelom preklapaju. Moguće je onda uvesti mjeru preklapanja  $P$  i  $P_G$  kao omjer duljine preklopa u odnosu na duljinu polazne ili pojednostavljene polilinije. Općenito će u tom slučaju postojati i parovi segmenata koji se sijeku samo u jednoj točki.

U protivnom, ako  $P_G \cap P$  sadrži konačno mnogo točaka parovi segmenata  $P$  i  $P_G$  se samo sijeku.

Ako je  $P_G \cap P = \emptyset$  onda  $P$  i  $P_G$  mora biti zatvorena polilinija (zbog uvjeta da otvorene polilinije imaju zajedničke krajnje točke).

Mjera preklapanja i broj presjeka parova segmenata u jednoj točki mogu poslužiti za definiciju mjere bliskosti  $P$  i  $P_G$ , međutim to prelazi okvire ovoga rada.

## 5. Funkcija za generalizaciju linije

U nastavku definirat će se funkcija  $e$  za generalizaciju otvorene ili zatvorene polilinije definirane u prethodnom poglavlju. Ulazni argument funkcije je otvorena ili zatvorena polilinija s maksimalno 4 vrha.

Funkcija ima oblik  $P_G = e(P)$ , gdje je  $P$  polazna polilinija, a  $P_G$  generalizirana polilinija.

Neka je  $P$  otvorena ili zatvorena polilinija kojoj pripadajući skup  $V$  ima maksimalno 4 vrha. Ako skup  $V$  ima manje od 4 vrha onda  $P_G = P$ .

Ako je  $P$  zatvorena polilinija čiji skup  $V$  ima 4 vrha (ulazna polilinija je trokut) tada  $P_G = P$ .

Ako je  $P$  otvorena polilinija čiji skup  $V$  ima 4 vrha tada je  $P_G$  otvorena polilinija s 3 vrha za koje vrijedi:

$$V_G(1) = V(1), V_G(3) = V(4) \text{ i } V_G(2) = T \text{ gdje je}$$

$$T \left( x, \frac{2A + x_1 y_1 - x_4 y_4 + x(y_4 - y_1)}{x_4 - x_1} \right);$$

$$A = \frac{1}{2} \sum_{i=1}^3 (x_{i+1} - x_i)(y_i + y_{i+1}).$$

Kako se vidi točka  $T$  za sada nije u potpunosti određena već leži na pravcu. Jednadžba tog pravca određena je iz sljedećeg uvjeta:

$$A = \frac{1}{2} \sum_{i=1}^3 (x_{i+1} - x_i)(y_i + y_{i+1}) = \frac{1}{2} ((x - x_1)(y_1 + y_2) + (x_2 - x)(y_2 + y_3)) = A_G.$$

Gornja formula zapravo govori da će površina područja omeđenog polilinijom  $P$ , osi  $x$  i pravcima  $x = x_1$  i  $x = x_4$  biti jednaka površini područja omeđenog polilinijom  $P_G$ , osi  $x$  i pravcima  $x = x_1$  i  $x = x_4$ . To vrijedi ako izraz za  $A$  prihvatimo kao izraz za orijentiranu površinu.

Neovisnost gornjeg rješenja o izboru koordinatnog sustava može se pokazati na sljedeći način. Zatvorimo polilinije  $P$  i  $P_G$  segmentom koji povezuje točku  $T_4$  i  $T_1$ , tj.  $P' = P \cup S_4 T_4, T_1$  i  $P'_G = P_G \cup S_3(T_4, T_1)$ . Sada izraz

$$A = \frac{1}{2} \sum_{i=1}^n (x_{i+1} - x_i)(y_i + y_{i+1})$$

predstavlja površinu zatvorenog poligona uz odgovarajuću orijentaciju. Taj izraz prihvaća kao izraz za površinu i poligona koji nisu jednostavni, tj. mogu sjeći sami sebe. Uvjet sada glasi

$$\frac{1}{2} (x_1 - x_4)(y_4 + y_1) + \frac{1}{2} \sum_{i=1}^3 (x_{i+1} - x_i)(y_i + y_{i+1}) = \frac{1}{2} (x_1 - x_4)(y_4 + y_1) + \frac{1}{2} \sum_{j=1}^2 (x_{j+1} - x_j)(y_j + y_{j+1})$$

odnosno prvi član s obje strane se poništava i vrijedi  $A = A_G$ .

Konkretan položaj točke  $T$  treba odrediti na temelju dodatnih uvjeta. Slika 3 prikazuje dva jednostavna slučaja primjene ovako definirane funkcije za generalizaciju polilinije. Položaj točke  $T$  na tim slikama odabran je proizvoljno.

Analizirajmo formulu

$$T \left( x, \frac{2A + x_1 y_1 - x_4 y_4 + x(y_4 - y_1)}{x_4 - x_1} \right). \text{ Pravac na kojem leži}$$

točka  $T$  ovom formulom nije određen u slučaju kada je  $x_1 = x_4$ . U tom slučaju treba primijeniti formulu oblika

$$T \left( \frac{2A + x_1 y_1 - x_4 y_4 - y(x_4 - x_1)}{y_1 - y_4}, y \right).$$

The function is defined as  $P_G = e(P)$ , where  $P$  is the given polyline and  $P_G$  is the generalized polyline.

If  $P$  is an open or closed polyline with the set  $V$  which has less than 4 vertices, then  $P_G = P$ .

If  $P$  is a closed polyline with the set  $V$  which has 4 vertices (the input polyline represents the triangle), then  $P_G = P$ .

If  $P$  is an open polyline with the set  $V$  which has 4 vertices, then  $P_G$  is an open polyline with the set  $V$  which has 3 vertices for which:

$V_G(1) = V(1)$ ,  $V_G(3) = V(4)$  and  $V_G(2) = T$ , where

$$T \left( x, \frac{2A + x_1y_1 - x_4y_4 + x(y_4 - y_1)}{x_4 - x_1} \right) \text{ and}$$

$$A = \frac{1}{2} \sum_{i=1}^3 (x_{i+1} - x_i)(y_i + y_{i+1}).$$

The point  $T$  has not been exactly defined yet, it is stated that it lies on the straight line. The equation of that straight line is derived from the constraint:

$$\begin{aligned} A &= \frac{1}{2} \sum_{i=1}^3 (x_{i+1} - x_i)(y_i + y_{i+1}) = \\ &= \frac{1}{2} ((x - x_1)(y_1 + y) + (x_2 - x)(y + y_2)) = A_G. \end{aligned}$$

The above equation defines that the area of the region bounded by polyline  $P$ , axis  $x$  and straight lines  $x = x_1$  and  $x = x_4$  is going to be the same as the area of the region bounded by polyline  $P_G$ , axis  $x$  and straight lines  $x = x_1$  and  $x = x_4$ . This is true if we accept the expression  $A$  as the area of the oriented polygon.

The independence of the solution above from the selection of the coordinate system can be shown in the following way. Let us close polylines  $P$  and  $P_G$  with a segment whose endpoints are  $T_4$  and  $T_1$ , i.e.  $P' = P \cup S_4(T_4, T_1)$  and  $P'_G = P_G \cup S_3(T_4, T_1)$ . Now the expression

$$A = \frac{1}{2} \sum_{i=1}^n (x_{i+1} - x_i)(y_i + y_{i+1})$$

represents the area of the polygon with certain orientation. We are going to accept this expression for the area of non-simple polygons, too. Now the constraint is

$$\begin{aligned} &\frac{1}{2} (x_1 - x_4)(y_4 + y_1) + \frac{1}{2} \sum_{i=1}^3 (x_{i+1} - x_i)(y_i + y_{i+1}) = \\ &= \frac{1}{2} (x_1 - x_4)(y_4 + y_1) + \frac{1}{2} \sum_{j=1}^2 (x_{j+1} - x_j)(y_j + y_{j+1}), \end{aligned}$$

the first parts of the equation on both sides cancel each other out and  $A = A_G$ .

The specific position of the point  $T$  should be defined using additional constraints. Fig. 4 shows two simple cases of the application of a generalization function of

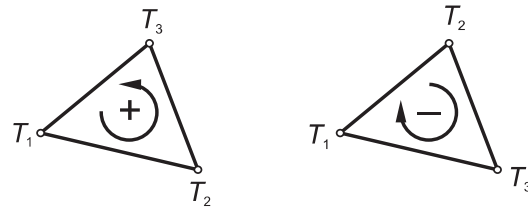


Fig. 5. Positive and negative triangle orientation  
Slika 5. Pozitivno i negativno orijentirani trokut

polyline defined in such a way. The position of the point  $T$  in Fig. 4 was chosen arbitrarily.

Let us analyze the expression

$$T \left( x, \frac{2A + x_1y_1 - x_4y_4 + x(y_4 - y_1)}{x_4 - x_1} \right).$$

The straight line on which the point  $T$  lies is not defined for the case  $x_1 = x_4$ . In that case, the expression

$$T \left( \frac{2A + x_1y_1 - x_4y_4 - y(x_4 - x_1)}{y_1 - y_4}, y \right) \text{ can be used.}$$

If  $x_1 = x_4$  and  $y_1 = y_4$ , then the first and the last point in the given polyline are identical, the polyline is closed, and as above,  $P_G = P$ .

The slope of the straight line containing the point  $T$  is defined by  $\frac{y_4 - y_1}{x_4 - x_1}$ . The straight line on which the point  $T$  lies is parallel to the straight line passing through the first and fourth point of the polyline  $P$ .

In a special case, when the straight line on which the point  $T$  lies passes through the first and fourth point of the polyline  $P$ , the two segments in  $P_G$  are on the same straight line and the point  $T$  can be removed from  $P_G$ , i.e.  $P_G = \{S_1(T_1, T_4)\}$ .

## 6 The Application of Function $e$ to Cartographic Line Generalization

The defined function  $e$  can be used for cartographic line generalization. In order to define a complete procedure for cartographic line generalization, it is necessary to introduce additional definitions and constraints, primarily:

1. the constraints for selecting three consecutive segments, i.e. four consecutive points in polyline which is going to be transformed through function  $e$ ,
2. the constraint for selecting the position of point  $T$  and
3. the method and order in which the polyline with more than 4 points is going to be treated by function  $e$ .

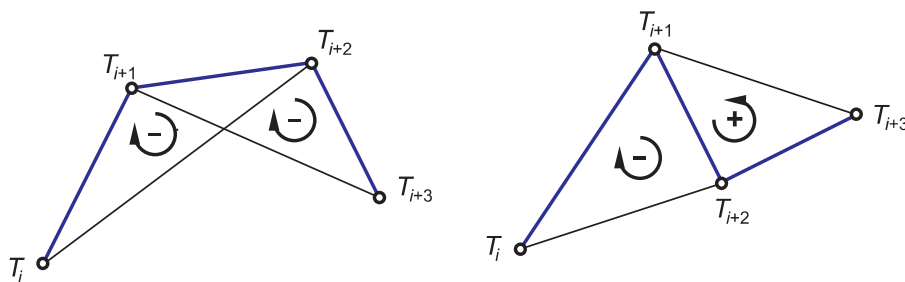


Fig. 6. The same and different orientation of triangles made of four consecutive points in polyline

Slika 6. Trokuti jednake i različite orijentacije, sastavljeni od četiri uzastopne točke u polilinjiji

Ako je  $x_1 = x_4$  i  $y_1 = y_4$  tada su prva i četvrta točka jednake i riječ je o zatvorenoj polaznoj polilinjiji  $P$ , a u tom slučaju kako je već definirano  $P_G = P$ .

Koeficijent smjera pravca na kojem leži točka  $T$  određen je izrazom  $\frac{y_4 - y_1}{x_4 - x_1}$ . Dakle pravac na kojem leži točka  $T$  paralelan je s pravcem kroz prvu i četvrtu točku polilinjije  $P$ .

U specijalnom slučaju kada pravac na kojem leži točka  $T$  prolazi kroz prvu i četvrtu točku segmenti u  $P_G$  leže na istom pravcu pa je moguće izbacivanje točke  $T$  iz  $P_G$ , odnosno  $P_G = \{S_1(T_1, T_4)\}$ .

### 6. Primjena funkcije generalizacije na kartografsku generalizaciju linija

Gore navedenu funkciju može se primijeniti i za kartografsku generalizaciju linija. Da bi definirali postupak generalizacije gornjom funkcijom potrebno je uvesti dodatne uvjete i definicije koje se prvenstveno odnose na:

1. uvjete kada niz od tri uzastopna segmenta odnosno 4 vrha u polilinjiji treba generalizirati definiranom funkcijom,
2. izbor položaja točke  $T$  i
3. način i redoslijed na koji će se složena polilinjija obuhvatiti funkcijom za generalizaciju  $e$ .

### 7. Jedan algoritam za ekvivalentno pojednostavljenje polilinjije

Na temelju prethodnih definicija i rasprave u nastavku će se definirati jedan konkretan algoritam za ekvivalentnu generalizaciju polilinjija. Da bi se to ostvarilo potrebno je konkretizirati uvjete iz prethodnog odjeljka.

#### 7.1. Uvjet kada niz od tri uzastopna segmenta odnosno 4 vrha u polilinjiji treba generalizirati definiranom funkcijom $e$

Moguće je zamisliti mnogo uvjeta kada niz od tri uzastopna segmenta odnosno 4 vrha u polilinjiji treba

generalizirati definiranom funkcijom  $e$ . U skladu s motivacijom za ovaj članak uvjet se može definirati na sljedeći način:

1. dio polilinjije od tri uzastopna segmenta će se generalizirati ako je kut između prvog i drugog segmenta i kut između drugog i trećeg segmenta mjereni s iste strane polilinjije jednom veći, a drugi put manji od ispruženog kuta i
2. duljina srednjeg segmenta mora biti manja od neke unaprijed zadane konstante.

Kako bismo utvrdili prvi dio uvjeta poslužiti ćemo se orijentacijom trokuta. Trokut je pozitivno orijentiran ako je on, obilazeći njegov rub od vrha  $T_1$  preko  $T_2, T_3$  i ponovo  $T_1$  u smjeru suprotnom od smjera kazaljke na satu, uvijek s lijeve strane. Trokut je negativno orijentiran ako nije pozitivno orijentiran (Lapaine i Frančula, 2001) (slika 5).

Lapaine i Frančula definiraju preslikavanje  $F : S \rightarrow R$  na sljedeći način:

$$F(T_1, T_2, T_3) = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

i pokazuju 7 svojstava tog preslikavanja, a dva ovdje važna svojstva su:

1.  $F(T_1, T_2, T_3) > 0$  ako je trokut pozitivno orijentiran i
2.  $F(T_1, T_2, T_3) < 0$  ako je trokut negativno orijentiran.

Prvi dio gore navedenog uvjeta za niz točaka  $T_i, T_{i+1}, T_{i+2}$  i  $T_{i+3}$ , sada se može izraziti na sljedeći način:

$$\text{sgn} F(T_i, T_{i+1}, T_{i+2}) \neq \text{sgn} F(T_{i+1}, T_{i+2}, T_{i+3}),$$

tj. ako je orijentacija trokuta  $\Delta T_i T_{i+1} T_{i+2}$  različita od orijentacije trokuta  $\Delta T_{i+1} T_{i+2} T_{i+3}$  tada je kut između prvog i drugog segmenta i kut između drugog i trećeg segmenta mjereni s iste strane polilinjije jednom veći, a drugi put manji od ispruženog kuta (slika 6).

Kada bi prethodni uvjet za generalizaciju niza od 4 uzastopne točke funkcijom  $e$  bio jedini linija bi se generalizirala sve dok svi trokuti sastavljeni od tri uzastopne točke u polilinjiji ne bi bili jednake orijentacije. Kako bi se



## 7 Algorithm for Equiareal Line Generalization

Based on the previous discussion and definitions, one exact algorithm for equiareal line generalization is going to be defined. In order to do this, it is necessary to specify the constraints from the previous section.

### 7.1 The constraint for selecting three consecutive segments in polyline for generalization with function $e$

It is possible to define many different constraints for selecting three consecutive segments in polyline which should be generalized with function  $e$ . According to the motivation for this paper, the constraint can be defined as follows:

1. the part of the polyline consisting of three consecutive segments is going to be generalized based on the value of angles between the first and second segment, and the second and the third segment. If one of the angles is greater than  $180^\circ$  and the other is less than  $180^\circ$ , then the three consecutive segments are going to be generalized by function  $e$  and
2. the length of the middle segment is less than a given constant value.

In order to determine if segments satisfy the first part of the constraint, we are going to use triangle orientation. A triangle is positively oriented if it lies on left side while touring its boundary from the vertex  $T_1$  over  $T_2, T_3$  and again to  $T_1$  in the counterclockwise direction. The triangle is negatively oriented if it is not positively oriented (Lapaine and Frančula 2001) (Fig. 5).

Lapaine and Frančula define the mapping  $F: S \rightarrow R$

$$\text{as } F(T_1, T_2, T_3) = \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

and show the seven properties of such mapping. Two properties of such mapping important for triangle orientation are:

1.  $F(T_1, T_2, T_3) > 0$  if the triangle is positively oriented and
2.  $F(T_1, T_2, T_3) < 0$  if the triangle is negatively oriented.

The first part of the constraint for four consecutive points  $T_i, T_{i+1}, T_{i+2}$  and  $T_{i+3}$  can be expressed as:

$$\text{sgn}F(T_i, T_{i+1}, T_{i+2}) \neq \text{sgn}F(T_{i+1}, T_{i+2}, T_{i+3}),$$

i.e. if orientation of the triangle  $\Delta T_i T_{i+1} T_{i+2}$  is different from the orientation of the triangle  $\Delta T_{i+1} T_{i+2} T_{i+3}$ , then the angle between the first and the second segment is less than  $180^\circ$  and the angle between the second and the third segment is greater than  $180^\circ$  or vice versa (Fig. 6).

If the constraint consisted of this part only, generalization with function  $e$  would result in a polyline with the same orientation of all triangles made of three

consecutive points. In order to remove only a certain amount of zigzags in the polyline, the length of the middle segment has to be shorter than a given value  $\varepsilon$ , i.e.

$$d(T_2, T_3) < \varepsilon.$$

It is useful to relate the value  $\varepsilon$  with other parameters, such as map scale, media for map presentation, line width, etc.

### 7.2 Point $T$ position selection

There is infinite number of possible positions for point  $T$  since it lies on the line. Some useful constraints are:

1. Point  $T$  is chosen so that the angle between segments  $\overline{T_1T}$  and  $\overline{TT_2}$  from  $P_G$  is the largest.
2. The intersection area of the regions  $A$  and  $A_G$  is maximal, i.e. for the point  $T$ :

$$T \rightarrow \max(A \cap A_G).$$

If the region overlap is maximal, polyline closeness is indirectly expressed this way.

3. Point  $T$  lies on the segment  $\overline{T_2T_3}$ .

Of the constraints listed, the first is going to be used for the purpose of the new algorithm, i.e. the point  $T$  is chosen to maximise the angle between segments  $\overline{T_1T}$  and  $\overline{TT_2}$  from  $P_G$ . In order to find the coordinates of the point  $T$ , we have to express this constraint analytically. The angle  $\angle T_1TT_2$  is shown in Fig. 6. It is going to be maximal when

$$p = q = \frac{b}{2}.$$

This can be easily proven in the following way. The sum of angles in a triangle is  $180^\circ$  and

$$\gamma = 180^\circ - \alpha - \beta.$$

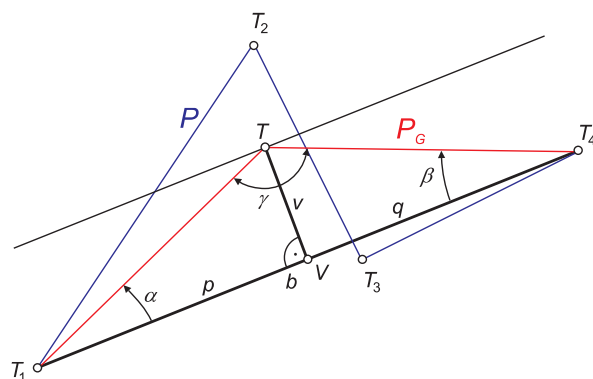


Fig. 7. Illustration for point  $T$  selection when the angle between segments  $\overline{T_1T}$  and  $\overline{TT_2}$  from  $P_G$  is maximal

Slika 7. Uz izbor točke  $T$  kada je kut između segmenata  $\overline{T_1T}$  i  $\overline{TT_2}$  iz  $P_G$  najtuplji

vijugavost polilinije uklonila samo do određene razine uvodi se dodatni uvjet duljine srednjeg segmenta, tj.:

$$d(T_2, T_3) < \varepsilon \text{ gdje je}$$

$\varepsilon$  veličina koju zadaje korisnik. Poželjno je dovesti veličinu  $\varepsilon$  u vezu s nekim drugim veličinama, npr. mjerilom karte, medijem prikaza karte, debljinom linije itd.

## 7.2. Izbor položaja točke $T$

Moguće je zamisliti beskonačno mnogo uvjeta za izbor položaja točke  $T$ . Neke mogućnosti su:

1. Točka  $T$  bira se tako da je kut između segmenata  $\overline{T_1T}$  i  $\overline{TT_2}$  iz  $P_G$  najtuplji.
2. Presjek područja  $A$  i  $A_G$  je maksimalan, odnosno točku  $T$  izabrat će se tako da vrijedi:

$$T \rightarrow \max(A \cap A_G).$$

Tim uvjetom zadaje se da je površina područja koje nije zajedničko najmanja i na posredan način zadana je bliskost izvorne i generalizirane polilinije.

3. Točka  $T$  leži na segmentu  $\overline{T_2T_3}$ .

Od nabrojanih uvjeta za konkretan algoritam u ovom radu upotrijebit će se prvi navedeni uvjet odnosno točka  $T$  bira se tako da je kut između segmenata  $\overline{T_1T}$  i  $\overline{TT_2}$  iz  $P_G$  najtuplji. Kako bismo odredili koordinate točke  $T$  potrebno je matematički izraziti navedeni uvjet. Na slici 7. prikazan je kut  $\angle T_1TT_2$ . On će poprimiti maksimalni iznos kada je

$$p = q = \frac{b}{2}.$$

Ta se tvrdnja može lako pokazati na sljedeći način. Zbroj kutova u trokutu je  $180^\circ$  pa vrijedi:

$$\gamma = 180^\circ - \alpha - \beta.$$

Iz pravokutnih trokuta  $\Delta T_1TV$  i  $\Delta TT_2V$  dobije se:

$$\tan \alpha = \frac{v}{p} \text{ i } \tan \beta = \frac{v}{q}. \text{ Sada imamo:}$$

$$\gamma = 180^\circ - \arctan \frac{v}{p} - \arctan \frac{v}{q} \text{ i } p = b - q.$$

Nađimo prvu derivaciju od  $\gamma$  po  $q$  te nakon izjednačavanja s nulom i sređivanja uz poznati izraz

$$\frac{d}{dx} \arctan x = \frac{1}{1+x^2} \text{ dobije se } q = \frac{b}{2}.$$

Drugim riječima točka  $T$  leži na presjeku pravca koji je rješenje funkcije za generalizaciju  $e$  i okomice na pravac kroz  $T_1$  i  $T_2$  položene kroz polovište dužine  $\overline{T_1T_2}$ .

## 7.3. Metoda obuhvaćanja složene polilinije funkcijom za generalizaciju $e$

Općenito, polilinija se sastoji od većeg broja točaka, a funkcija za generalizaciju  $e$  kao argument prima niz od 4 uzastopne točke. Dakle, potrebno je zadati način na

koji će se polazna linija obuhvatiti funkcijom za generalizaciju. Moguće je opet zamisliti veći broj načina za tu operaciju. Nakon nekih početnih testiranja različitih pristupa uočeno je da redosljed primjene funkcije generalizacije na dijelove polazne polilinije može dati dosta različite rezultate. Zbog toga smatramo da se ovaj dio algoritma treba posebno i detaljno istražiti u budućnosti.

U ovom radu odabran je sljedeći način:

1. brojač točaka postavlja se na 1,
2. točka koju zadaje brojač i sljedeće tri točke polazne polilinije šalju se u funkciju za generalizaciju  $e$ ,
3. ako je došlo do promjene, tj. vraćene su tri točke iz funkcije za generalizaciju, brojač točaka se umanjuje za 2 (ako je veći od 3), četiri točke u polaznoj poliliniji zamijenjuju se s tri točke dobivene kao rezultat generalizacije, poziva se korak 2,
4. ako nije došlo do promjene, brojač točaka se uvećava za 1 (ako je za 4 i više manji od broja točaka u poliliniji), poziva se korak 2.

Ovako zadanim načinom polazna polilinija se funkcijom za generalizaciju obilazi od početka do kraja samo jednom. Moguće je da se brojač točaka i smanjuje i povećava, tj. da se polilinija generalizira lokalno sve dok više nisu zadovoljeni uvjeti za generalizaciju i tek tada brojač točaka napreduje prema kraju polilinije. Nakon svakog zadovoljenog uvjeta za generalizaciju polazna polilinija se modificira i kao takva ulazi u ponovno razmatranje.

Kod zatvorenih polilinija treba voditi računa o tome da se na kraj popisa točaka dodaju prve tri točke iz popisa kako bi funkcija za generalizaciju mogla eventualno modificirati i te točke.

## 8. Primjena i testiranje algoritma

Za implementaciju gore opisanog algoritma izabran je GRASS GIS (URL3). GRASS GIS je program otvorenog kôda i već unutar modula *v.generalize* ima definiran čitav niz algoritama za pojednostavljenje i izgladivanje linija ili mreža. Za pojednostavljenje ugrađeni su algoritmi Douglas-Peuckera, Langa, Reumann-Witkama te algoritmi za filtriranje točaka na temelju jednostavnih uvjeta. Za izgladivanje tu su algoritmi Boylea, Snakesa, Chaikena, dva McMasterova algoritma i hermitska interpolacija.

Izvorni kôd modula *v.generalize* izmijenjen je tako da implementira ovdje opisanu metodu radnog naziva "equi-areaal". Izvorni kôd potrebno je prevesti u izvršni nakon čega je metodu moguće upotrebljavati unutar okruženja GRASS GIS-a što ima čitav niz prednosti od učitavanja podataka do njihovog ispisa.

Za testiranje algoritma upotrijebljena je obalna crta otoka Raba. Prikaz obalne crte otoka Raba na kartama različitih mjerila istražen je u radovima (Vučetić i dr. 2006) i (Tutić i dr. 2007). Na temelju tih istraživanja može se zaključiti da prikaz otoka Raba na kartama različitih mjerila ima približno jednake površine, a da se duljina obalne crte značajno mijenja.

From rectangular triangles  $\Delta T_1TV$  and  $\Delta TT_4V$ , we have:

$\tan \alpha = \frac{v}{p}$  and  $\tan \beta = \frac{v}{q}$ . Now:

$$\gamma = 180^\circ - \arctan \frac{v}{p} - \arctan \frac{v}{q} \text{ and } p = b - q.$$

Let us find the first derivative of  $\gamma$  by  $q$ , using expression  $\frac{d}{dx} \arctan x = \frac{1}{1+x^2}$ , and after solving it we get

$$q = \frac{b}{2}.$$

In other words, point  $T$  lies at the intersection of the straight line which is the solution of the function  $e$  and the straight line which passes through the middle point of segment  $\overline{T_1T_4}$  and is perpendicular to it.

### 7.3 Method in which the polyline with more than four points is going to be treated by function $e$

Generally, a polyline consists of a number of points. The generalization function  $e$  takes as argument only polylines with four points. The order of how the given polyline is going to be treated by the generalization function should be specified. Again, it is possible to define a number of such methods. After initial experimenting with different methods, it was found that this can have a substantial effect on the results. Therefore, it is necessary to further investigate this part of the algorithm.

For the purpose of the new algorithm the following method is used:

1. point counter is set to 1 (first point in polyline)
2. the point indicated by the counter and next three points are input in the generalization function  $e$ ,
3. if there is a change, i.e. the result has three points, the point counter is decreased by 2 (if it is larger than 3), the four points in the given polyline are replaced with three points resulting from the generalization function  $e$ , step 2 is invoked,
4. if there was no change, the point counter is increased by 1 if it is less than the number of points in polyline minus 3, step 2 is called; otherwise, the process ends.

This way, the given polyline is traversed only once from the beginning to the end by the generalization function. It is possible the point counter decreases and increases back and forth, i.e. the polyline is going to be treated locally until no generalization constraint is satisfied, which is when the point counter increases towards the end of the polyline. After each satisfied generalization constraint, the given polyline is updated and then enters the next consideration.

When closed polylines are treated, the first three points should be added to the end of the point list in order to consider all segments for generalization.

## 8 Application and Testing of the Algorithm

GRASS GIS was chosen for implementation of the described algorithm (URL3). GRASS GIS is an open source software which already has the module *v.generalize* with a number of implemented algorithms for simplification and smoothing of lines and networks. For simplification, there are algorithms of Douglas-Peucker, Lang, Reumann-Witkam and simple filtering algorithms. For smoothing, there are algorithms of Boyle, Snakes, Chaiken, two McMaster algorithms and the hermit interpolation.

The source code of the module *v.generalize* is modified to implement the described algorithm with the working title *equiareal*. The source code needs to be compiled in order to use the algorithm inside the GRASS GIS environment. This has an advantage when importing, manipulating and printing data inside the same environment.

The coastline of the Island of Rab was used to test the algorithm. Representation of this coastline on maps at different scales was explored in (Vučetić et al. 2006) and (Tutić et al. 2007). Based on that research, it can be concluded the island has approximately the same area in all representations, while the coastline length changes substantially.

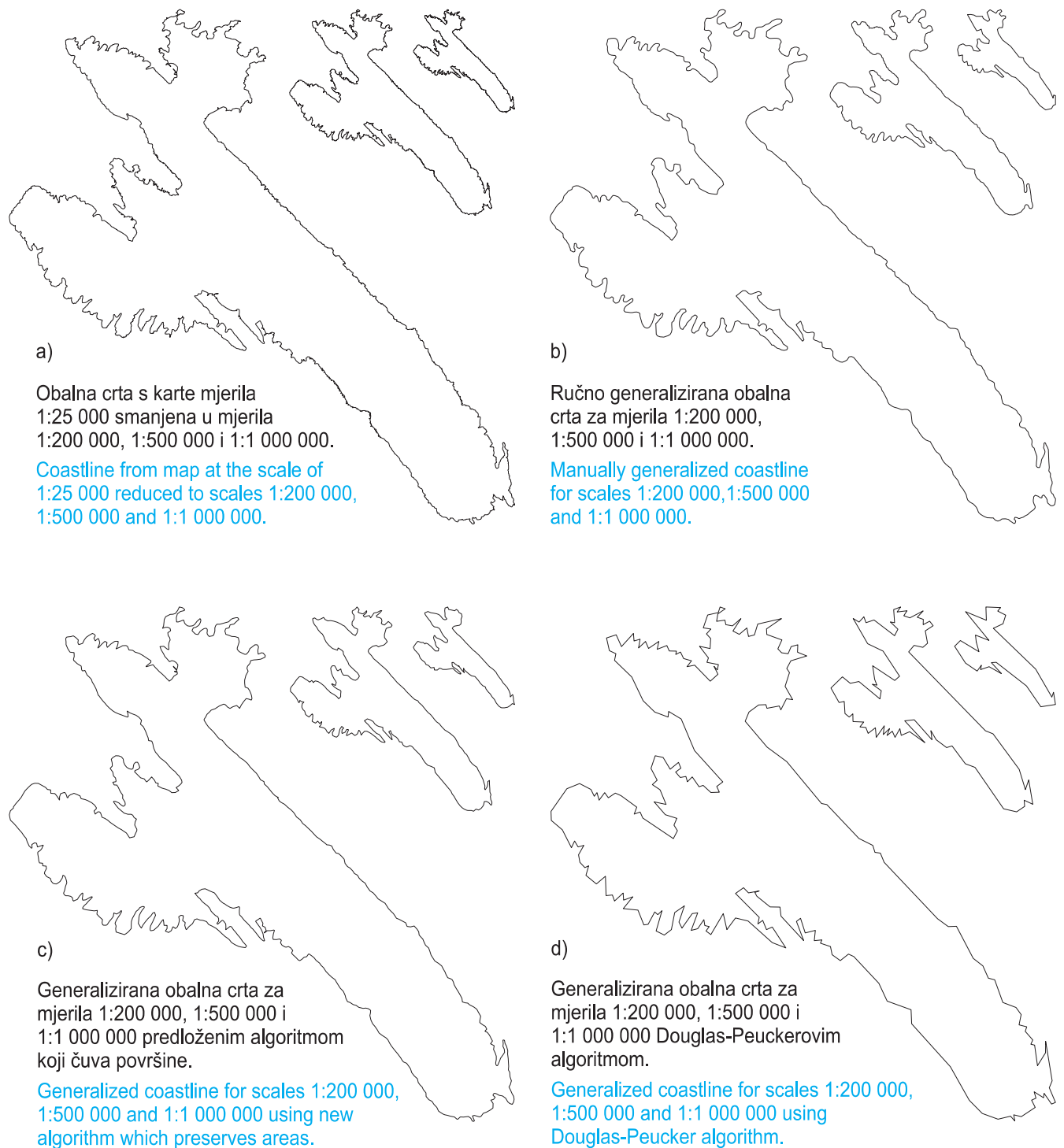
For the input polyline, the one from the map at the scale of 1:25 000 was chosen and it was generalized for the scales 1:200 000, 1:500 000 and 1:1 000 000. The generalized lines with selected algorithms were constrained to have the same length as the manually generalized coastlines.

Fig. 8a shows the input polyline reduced to the scales 1:200 000, 1:500 000 and 1:1 000 000, 8b shows the manually generalized polyline on official topographic maps, 8c shows generalized polyline using the new algorithm *equiareal* and 8d shows the polyline generalized with the Douglas-Peucker algorithm.

Because the new algorithm can also be considered as the smoothing algorithm, the following comparison was performed. The coastline was smoothed with the Boyle algorithm in order to gain the same length of the generalized coastline as of those generalized manually. The number of points in such polylines is the same as of the input polyline. The next step involved the application of the Douglas-Peucker algorithm in order to obtain the same number of points as in those polylines obtained by the new algorithm *equiareal*. The result of these two operations is shown in Fig. 9.

The length, number of points and area of all coastlines in Fig. 8 and 9 are given in Table 1.

Based on data in Table 1, some observations can be given. First, it should be noted the new algorithm always yields polylines with the same area as the input polyline which results from the given properties. The largest change in area results from manual generalization and

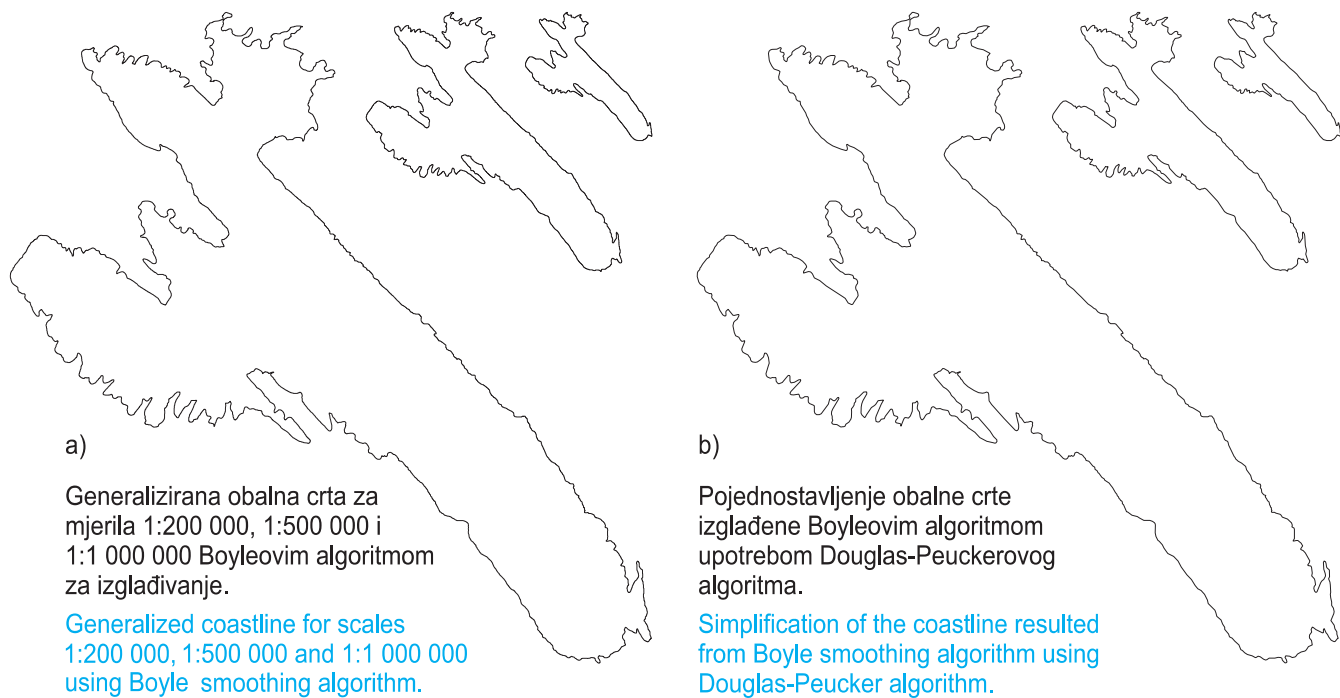


Parametar predloženog algoritma koji čuva površine i Douglas-Peuckerovog algoritma izabran je tako da duljina linije odgovara duljini linije dobivenoj ručnom generalizacijom za pojedino mjerilo karte.

Parameter of the new algorithm which preserves areas and Douglas-Peucker algorithm is chosen to give lengths of coastline which are (almost) equal to the lengths of manually generalized coastline in given map scale.

Fig. 8. Coastline of the Island of Rab at the scales of 1:200 000, 1:500 000 i 1:1 000 000: a) photographically reduced coastline, b) manually generalised coastline, c) coastline generalised using new algorithm, d) coastline simplified using Douglas-Peucker algorithm

Slika 8. Obalna crta otoka Raba u mjerilima 1:200 000, 1:500 000 i 1:1 000 000: a) fotografski smanjeni izvorni prikaz, b) izvorni prikaz ručno generaliziran, c) prikazi generalizirani novim algoritmom opisanim u ovom radu, d) prikazi pojednostavljeni Douglas-Peuckerovim algoritmom



Parametar Boyleovog algoritma za izgladivanje izabran je tako da duljina linije odgovara duljini linije dobivenoj ručnom generalizacijom za pojedino mjerilo karte. Parametar Douglas-Peuckerovog pojednostavljenja izabran je tako da broj točaka odgovara broju točaka u liniji dobivenoj predloženim algoritmom koji čuva površine za pojedino mjerilo karte.

Parameter of the Boyle smoothing algorithm is chosen to give lengths of coastline which are (almost) equal to the lengths of manually generalized coastline at the given map scale. Parameter of the Douglas-Peucker algorithm is chosen to give coastline with the same number of points as coastline resulted from new algorithm which preserves areas at the given map scale.

Fig. 9. Coastline at the scales of 1:200 000, 1:500 000, 1:1 000 000 smoothed using the Boyle algorithm (a) and then simplified with the Douglas-Peucker algorithm (b)

Slika 9. Prikaz obalne crte u mjerilima 1:200 000, 1:500 000, 1:1 000 000 izgladene Boyleovim algoritmom (a) te potom pojednostavljene Douglas-Peuckerovim algoritmom (b)

the Douglas-Peucker algorithm, and a relatively small change using Boyle smoothing. When the length of a generalized polyline is given in advance, the new algorithm resulted with polylines with five to eight times more points than using the Douglas-Peucker algorithm. If we look at the generalized coastlines, it can be concluded the new algorithm yields satisfactory results for cartographic line generalization application.

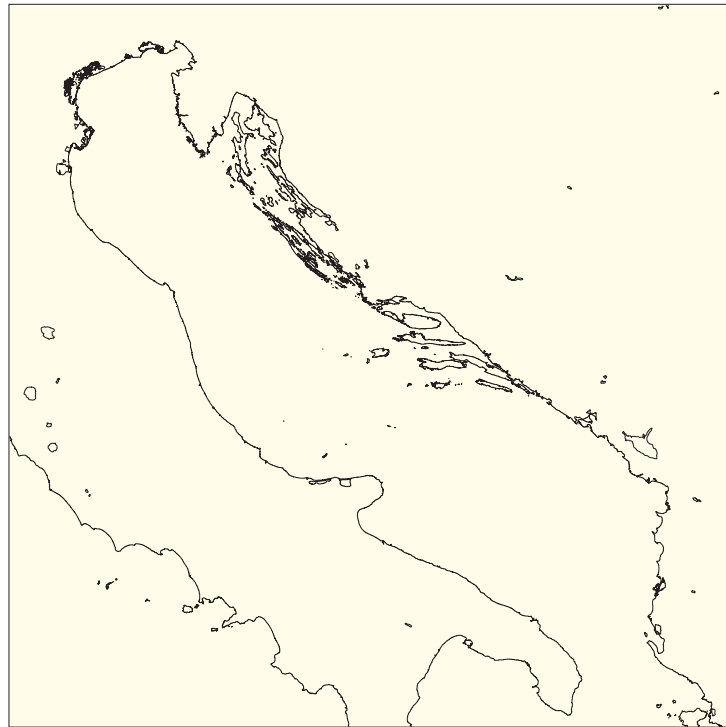
## 9 Selection of the Parameter Value for Cartographic Purposes

Each polyline on a map at a given scale is drawn using a defined width. The map scale and the line width are just some of the parameters which can be used for estimating the initial parameter value of the algorithm for generalization. The next important factor is the media used for map presentation and its resolution. After the initial value is chosen and applied, the cartographer needs

to analyze the results and potentially redefine the parameter. In the case of automated systems for generalization, it is important to have good initial values of the parameter because there is no interactive correction.

Based on the values in Table 1, estimation can be attempted of the parameter value as a function of the map scale for the new algorithm. All lines are drawn on paper, with width of 0.13 mm and these two factors are not going to be accounted for now. The ratio of the scale denominator and the value of the parameter range from 2300 to 2900, increasing with the scale denominator. Let us take the value 2500 as a rule of thumb. A better relation of the scale denominator and other factors to the parameter value of the new algorithm deserves further research including a larger set of test data.

Fig. 10a shows the coastlines of the mainland and islands of the Adriatic Sea at the scale 1:625 000 as it is defined at the most detailed level of GSHHS data (URL 4). Fig. 10b shows the generalized coastlines using the



*Fig. 10. Shorelines of Adriatic Sea represented in GSHHS (up) and generalised representation using new algorithm and filtering of small areas (down)*  
*Slika 10. Prikaz obalne crte kopna i otoka Jadranskog mora: prikaz podataka GSHHS-a (gore) i prikaz dobiven generalizacijom s pomoću novog algoritma i filtriranjem malih površina (dolje)*

Table 1. The length, number of points and area of all coastlines of the island of Rab in Fig. 8a-8d and 9a-9b  
 Tablica 1. Podaci o duljini, broju točaka i površini svih obalnih crta Raba na slikama 8a-8d i 9a-9b

<b>Input polyline</b> Izvorna polilinija				
Scale Mjerilo karte	Length [km] Duljina [km]	Area [km <sup>2</sup> ] Površina [km <sup>2</sup> ]	Number of points Broj točaka	Parameter Parametar
1:25 000	127.9	86.3	15 678	
<b>Manual generalization</b> Ručna generalizacija				
1:200 000	103.2	86.5		
1:500 000	90.5	87.7		
1:1 000 000	76.9	85.4		
<b>New algorithm equiareal</b> Novi algoritam equiareal				
1:200 000	103.2	86.3	1372	85
1:500 000	90.5	86.3	515	176
1:1 000 000	76.8	86.3	217	343
<b>Douglas-Peucker algorithm</b> Douglas-Peuckerov algoritam				
1:200 000	103.2	85.9	251	92
1:500 000	90.5	85.2	82	258
1:1 000 000	76.8	85.3	29	650
<b>Boyle algorithm</b> Boyleov algoritam				
1:200 000	103.8	86.3	15 678	8
1:500 000	90.3	86.2	15 678	21
1:1 000 000	76.9	86.1	15 678	58
<b>Combination Boyle &amp; Douglas-Peucker</b> Kombinacija Boyle i Douglas-Peucker				
1:200 000	102.8	86.3	1372	*
1:500 000	89.3	86.2	515	*
1:1 000 000	76.0	86.1	217	*

\* A variant of Douglas-Peucker algorithm was used in which a percentage of points is given in the generalized line in relation to the input line

\* Upotrijebljena je varijanta Douglas-Peuckerova algoritma u kojemu se zadaje postotak točaka u generaliziranoj liniji u odnosu na polaznu liniju.

Kao izvorna (polazna) polilinijska linija izabrana je ona s karte u mjerilu 1:25 000, a generalizirana je za mjerila 1:200 000, 1:500 000 i 1:1 000 000, ali tako da duljina generalizirane obalne crte odgovara duljini obalne crte dobivene ručnom generalizacijom.

Slika 8a. prikazuje izvornu polilinijsku fotografiju smanjenu u mjerila 1:200 000, 1:500 000 i 1:1 000 000, 8b. prikazuje ručno generalizirane obalne crte na službenim topografskim kartama, 8c. generaliziranu obalnu crtu predloženim algoritmom "equiareal" i 8d. generaliziranu crtu Douglas-Peuckerovim algoritmom.

Kako predloženi algoritam spada i u algoritme za izgladivanje napravljena je i sljedeća usporedba. Obalna crta je najprije izgladivljena Boyleovim algoritmom tako da se dobije duljina obalne crte koja odgovara duljini obalne crte dobivene ručnom generalizacijom. Tako izgladivljena linija ima jednak broj točaka kao i polazna. U sljedećem koraku tako izgladivljena linija pojednostavljena je Douglas-Peuckerovim algoritmom uz uvjet da je broj točaka jednak broju točaka u obalnoj crti dobivenoj novim algoritmom koji čuva površine. Rezultat se vidi na slici 9.

Podaci o duljini, broju točaka i površini svih obalnih crta Raba na slikama 8a-8d i 9e-9f dani su u tablici 1.

Na temelju podataka iz tablice 1. mogu se donijeti određene prosudbe. Najprije valja uočiti da novi algoritam za sve generalizirane linije ima jednaku površinu kao i polazna linija, što proizlazi iz postavljenih uvjeta. Najveću promjenu površine daje ručna generalizacija i Douglas-Peuckerov algoritam, a relativno malu Boyleov algoritam za izgladivanje. Novi algoritam za unaprijed zadanu duljinu polilinijske linije (uzetu ne temelju ručne generalizacije) daje generalizirane linije koje imaju 5-8 puta više točaka od linije dobivene Douglas-Peuckerovim algoritmom. Pogledaju li se prikazi dobivenih generaliziranih linija lako se uočava da u smislu kartografske generalizacije novi algoritam daje prihvatljive rezultate.

## 9. O izboru vrijednosti parametra za potrebe kartografije

Svaka polilinijska linija na karti zadanog mjerila crta se unaprijed zadanom debljinom. Mjerilo karte i debljina linije samo su neki podaci koji mogu poslužiti za izbor početne vrijednosti parametra algoritma za generalizaciju. Sljedeća važna stvar je medij na kojemu se karta prikazuje i razlučivost tog medija. Nakon početnog izbora vrijednosti parametra i uvida u rezultat, kartograf treba donijeti odluku o njezinoj korekciji. Kod automatiziranih sustava za generalizaciju potrebno je unaprijed imati dovoljno dobru vrijednost parametra jer nema mogućnosti interaktivne korekcije.

Na temelju vrijednosti iz tablice 1 može se pokušati donijeti gruba procjena vrijednosti parametra za novi

algoritam na temelju mjerila karte. Sve crte prikazane su debljinom 0.13 mm na papiru tako da se te dvije veličine u ovom trenutku neće razmatrati. Kvocijent nazivnika mjerila i vrijednosti parametra kreće se od otprilike 2300 do 2900 i raste s nazivnikom mjerila karte. Uzmimo grubu procjenu tog kvocijenta kao konstante jednake 2500. Bolja funkcijska veza između vrijednosti parametra i mjerila karte zaslužuje daljnja istraživanja koja će obuhvatiti i veći skup testnih podataka.

Na slici 10a. prikazana je obalna crta kopna i otoka Jadranskog mora iz podataka GSHHS, razine s najviše detalja, (URL4) u mjerilu 1: 625 000. Slika 10b. prikazuje generaliziranu obalu Jadranskog mora i otoka predloženim algoritmom s vrijednosti parametra 2500 (6250000/2500). Osim toga provedeno je i filtriranje (izbacivanje) malih površina koje su manje od 0.8 mm<sup>2</sup> u mjerilu karte. Takvo filtriranje može se provesti prije ili poslije generalizacije s obzirom na svojstvo čuvanja površina.

## 10. Zaključak

U radu je opisana metoda generalizacije linije koja čuva površine područja omeđenih polilinijskim linijama. Svojstvo očuvanja površina poželjno je iz dva razloga. Prvi je da i ručno generalizirane linije imaju tendenciju očuvanja površine područja omeđena tim linijama (Vučetić i dr. 2006). Drugi razlog je primjena u GIS-u i bazama prostornih podataka kada se vrijednosti atributa odnose na površine objekata. Primjerice filtriranje po vrijednosti površine može se obaviti prije i poslije generalizacije ovakvom metodom.

Opisani algoritam temelji se na funkciji za generalizaciju koja dio polilinijske linije od 3 uzastopna segmenta transformira u polilinijsku s 2 uzastopna segmenta i opisana je detaljno u §5. Moguć je veliki broj različitih načina primjene takve funkcije na generalizaciju linije od kojih će svaki rezultirati očuvanim površinama.

Prikazani način daje dobre rezultate prilikom primjene za kartografsku generalizaciju. Maksimalno generalizirane zatvorene polilinijske linije dovode do konveksnog oblika, što je u analogiji s oblikom područja prilikom ručne generalizacije kad se približavamo minimalnim veličinama.

Ovim radom otvara se mogućnost daljnjih istraživanja. U ovom radu izabrani uvjeti za generalizaciju, izbor položaja nove točke te pristupa složenim polilinijskim linijama su samo jedna mogućnost. Preostaje ispitati i druge mogućnosti. Moguće je definirati i inverznu funkciju koja će za polazna 2 uzastopna segmenta u polilinijskim rezultatima s 3 segmenta uz isti uvjet očuvanja površina, te na taj način omogućiti uklanjanje oštrog prijeloma u polilinijskim linijama. Problem samosječanja u polaznoj i rezultirajućoj polilinijskoj liniji također treba dodatno istražiti.



new algorithm with the parameter value 2500 (6250000/2500). Filtering (deletion) of small areas less than 0.8 mm<sup>2</sup> on the map was also performed. Such filtering can be performed before or after generalization due to area preservation.

## 10 Conclusions

This paper describes a new method for line generalization which has the property of area preservation of regions bounded by polylines. The area preservation property is desirable for two reasons. The first is that manually generalized regions also have a tendency to preserve areas (Vučetić et al . 2006). The second reason is the application in GIS and spatial databases when feature attributes are related to object areas. As an example, filtering based on area can be performed before or after generalization using this method.

The algorithm described is based on the generalization function which, for the input polyline of three segments, yields a polyline with two segments and is described in §5 in detail. A large number of different approaches for applying this function to line generalization is possible, each of them resulting in preserved areas.

The method described yields satisfactory results when used for cartographic line generalization. The maximal polygon generalization with this method results in convex polygon which is in accordance with the shape of the polygons when approaching minimal dimensions.

This paper opens possibilities for further research. The additional constraints selected represent just one possible solution. Other possibilities are left for future work. It is possible to define the inverse of the generalization function which is going to transform two consecutive input segments in the polyline into three new segments with the constraint of area preservation. Such an inverse function can be used to refine pikes in lines. Self-intersection in input and output polylines also needs to be investigated and implemented in such algorithms.

## References / Literatura

- Braden, B. (1986): The Surveyor's Area Formula. *The College Mathematics Journal*, Vol. 17, No. 4, pp. 326-337.
- Lapaine, M., Frančula, N. (2001): Točka u poligonu. *Geodetski list*, Vol. 55(78), No.3, pp. 207-222.
- Macarol, S. (1978): *Praktična geodezija*. 3. izdanje, Tehnička knjiga, Zagreb.
- Mackaness, W. A., Ruas, A and Sarjakoski, L. T. (eds.) (2007): *Generalisation of Geographic Information: Cartographic Modelling and Applications*. Elsevier, Amsterdam.
- McMaster, R. B. (1986): A Statistical Analysis of Mathematical Measures for Linear Simplification. *The American Cartographer*, Vol. 13, No. 2, pp. 103-116.
- McMaster, R. B. (1987): Automated Line Generalization. *Cartographica*, Vol. 24, No. 2, pp.74-111.
- Thapa, K. (1988): A Review of Critical Points Detection and Line Generalization Algorithms. *Surveying and Mapping*, Vol. 48, No. 3, pp. 185-205.
- Tutić, D., Lapaine, M., Vučetić, N. (2007): The Length of Coastline of the Island of Rab. XXIII International Cartographic Conference, August 4-10, 2007, Moscow, Abstracts of Papers, Moskva, Roskartografija, pp. 205-205.
- Vučetić, N. Frančula, N., Šimat, I. (2006): The Effect of Map Scale on the Determination of the Coastline Length and the Area of Islands in the Adriatic Sea - the Example of the Island of Rab, *Kartografija i geoinformacije*, Vol. 5, No. 6, pp. 68-75.

URL1: ICA Commission on Generalisation and Multiple Representation - Bibliography

<http://aci.ign.fr/bibliography/all.php> (1. 6. 2009.)

URL2: ESRI Training and Education

<http://training.esri.com/campus/library/index.cfm> (1. 6. 2009.)

URL3: GRASS GIS - The World Leading Free Software GIS

<http://grass.itc.it/> (1. 6. 2009.)

URL4: GSHHS - A Global Self-consistent, Hierarchical, High-resolution Shoreline Database

<http://www.soest.hawaii.edu/wessel/gshhs/gshhs.html> (1. 6. 2009.)