



godina  
years

Geodetskog fakulteta  
of Faculty of Geodesy

Zagreb, 26.-27. rujna 2002.

N. Vučetić<sup>1</sup>, S. Petrović<sup>2</sup>, M. Lapaine<sup>1</sup>, N. Frančula<sup>1</sup>

## POJEDNOSTAVLJENJE LINIJA NA TEMELJU KOEFIČIJENTA LINEARNE KORELACIJE

**SAŽETAK.** Cilj pojednostavljenja linija u digitalnoj kartografiji je smanjenje informacijskog sadržaja, pri čemu pojednostavljena linija treba biti slična izvornoj, tj. da na promatrača ostavlja sličan vizualni dojam kao izvorna te da metrička odstupanja ostanu unutar unaprijed zadane granice. Iskustva iz teorije prepoznavanja uzoraka, ali i iz graničnih područja geodezije i geofizike ukazuju na jaku povezanost sličnosti vizualnih dojmova i sličnosti oblika. Ako se odustane od u statistici uobičajenog ograničenja da se koeficijent linearne korelacije primjenjuje samo na stohastičke veličine i primijeni ga se na veličine koje definiraju geometrijske konfiguracije (npr. na koordinate), dolazi se do jedne nove interpretacije toga koeficijenta. Time on postaje mjera sličnosti oblika. U ovom radu dana su načela pojednostavljenja pojedinačnih linija na temelju maksimalne korelacije između izvorne i pojednostavljene linije. Mogućnost primjene postupka ilustrirana je primjerima. Pritom se pokazalo da ta metoda mnogo bolje zadržava karakteristični oblik linije nego postupci koji se temelje samo na metričkim kriterijima.

**Ključne riječi:** kartografska generalizacija, pojednostavljenje linija, koeficijent linearne korelacije, sličnost oblika.

## LINE SIMPLIFICATION BASED ON THE COEFFICIENT OF LINEAR CORRELATION

**ABSTRACT.** The intention of line simplification in digital mapping is the reduction of information contents in such a way that the simplified line be similar to the original (it is, it should produce a similar visual impression on an observer as the original one) and that the metric deviations remain inside the predefined limit. Experiences from the theory of pattern recognition, but also from bordering regions of geodesy and geophysics point to a strong relation between the similarity of visual impressions and the similarity of forms. Giving up the limitation which is usual in statistics, that the coefficient of linear correlation can be applied

---

<sup>1</sup>Dr.sc. Nada Vučetić, prof.dr.sc. Miljenko Lapaine, prof.dr.sc. Nedjeljko Frančula, Sveučilište u Zagrebu, Geodetski fakultet, Kačićeva ul. 26, 10000 Zagreb, Hrvatska.

<sup>2</sup>Dr.sc. Svetozar Petrović, Technische Universität Berlin, Institut für Geodäsie und Geoinformationstechnik, Sekretariat H12, D-10623 Berlin, Straße des 17. Juni 135.

only to stochastic quantities and applying it to quantities which define geometric configurations (like coordinates), a new interpretation of this coefficient is obtained. In such a way it becomes a measure of the similarity of forms. This paper presents the fundamentals of the simplification of individual lines based on maximum correlation between the original and the simplified line. The scope of the application of the procedure is illustrated on examples. It came out that the new method retains the characteristic form of a line to a much higher extent than the procedures based on metric criteria alone.

**Key words:** cartographic generalization, line simplification, coefficient of linear correlation, similarity of forms.

## 1. UVOD

Jedan od najvećih izazova kartografskom radu u digitalnom okruženju je automatizacija kartografske generalizacije (McMaster i Shea 1992, João 1998), tako da se već tridesetak godina u pojedinim znanstvenim institucijama u svijetu provode istraživanja na tom polju. Kako su 80% elemenata karte linijski elementi (Thapa 1988, Müller i dr. 1995), većina istraživača bavi se problemom automatizacije generalizacije linija. Istraživanja automatizacije generalizacije linija provode se i u Zavodu za kartografiju Geodetskog fakulteta Sveučilišta u Zagrebu (Frančula 2000).

U postojećoj literaturi nailazimo na raznorazne algoritme za automatsku generalizaciju linija. Najveći broj tih postupaka temelji se na izravnom smanjenju broja točaka linije (pojednostavljenju) (Ivanov 1965, Lang 1969, Ramer 1972, Douglas i Peucker 1973, Reumann i Witkam 1974, Cromley 1991, Visvalingam i Whyatt 1993, Wang i Müller 1998, ...). Najznačajniji i najrašireniji predstavnik te vrste algoritama, koji je ugrađen kao osnova u mnogim komercijalnim programskim paketima, npr. ArcInfu (Li i Openshaw 1992, Hassen i Beard 1998), AutoCAD Mapu (Vučetić 2000), MGE Map Generalizeru (Lee 1995, Lechthaler i Kasyk 1999) je Douglas-Peuckerov algoritam (Douglas i Peucker 1973).

Ispitivanjima na nekoliko algoritama koji najviše obećavaju pokazalo se da algoritmi koji se temelje na smanjenju broja točaka imaju dobrih i loših osobina, s time da nedostaci ne proizlaze iz redukcije broja točaka nego iz činjenice da se smanjenje broja temelji isključivo na udaljenostima od izvorne linije, pri čemu se udaljenosti mjere euklidskim udaljenostima, površinama ili nekim drugim metričkim značajkama. S druge pak strane pokušaji razrade algoritama koji nastoje u nekom smislu biti utemeljeni na zadržavanju bitnih karakteristika, tj. zapravo na sličnosti oblika, zasnivaju se u dosadašnjim realizacijama na tome da se karakteristični oblici također pokušavaju opisati na metrički način, kao npr. u radu Wanga i Müllera (1998), gdje se sličnost dvaju zavoja definira kao jednakost njihovih površina dodatno pomnoženih recipročnim vrijednostima indeksa kompaktnosti. Princip je dakle opet uklanjanje na temelju odstupanja od osnovne linije zavoja, koje se u navedenom primjeru mjeri površinama.

Metoda, čija se načela daju u ovome rada, pokušava kombinirati dosadašnje pristupe tako što se broj točaka smanjuje, ali se to smanjenje želi zasnovati na sličnosti oblika između izvorne i pojednostavljene linije koje se ne mjeri tolerancijama odstupanja ili površina nego iznosima koeficijenta linearne korelacije.

## 2. POJEDNOSTAVLJENJE LINIJA I KOEFICIJENT LINEARNE KORELACIJE

Glavna značajka pojednostavljena linija je, pored smanjenja informacijskog sadržaja izvorne linije, sličnost pojednostavljene linije izvornoj u smislu da na promatrača ostavlja sličan vizualni dojam kao izvorna.

Iskustva iz teorije prepoznavanja uzoraka, ali i iz graničnih područja geodezije i geofizike (gdje se uspoređuju različite vrste geodetskih i geofizičkih podataka) podsjećaju nas na jaku povezanost sličnosti vizualnih dojmova i sličnosti oblika. Sličnost oblika mjeri se pak koeficijentom linearne korelacije (Petrović i dr. 1987, Petrović i dr. 1988, Petrović 1991, Vučetić 2001).

Do koeficijenta linearne korelacije kao mjere sličnosti dolazi se drugačijom geometrijskom interpretacijom nego što je ona koja se može pronaći u standardnoj literaturi matematičke statistike. Interpretacija se temelji na razmatranju geometrijskih odnosa među dvjema konfiguracijama točaka u ravnini, odnosno prostoru. Naime, ako je apsolutna vrijednost koeficijenta linearne korelacije  $r$  jednak 1 tada se geometrijski gledano jedna konfiguracija točaka može dobiti rastezanjem druge. U slučaju kada je ona različita od 1 nije moguće dobiti jednu konfiguraciju točaka rastezanjem druge, već neke njoj "više" ili "manje" bliske. Koliko je rastegnuta slika jedne konfiguracije "bliska" drugoj konfiguraciji govori iznos koeficijenta linearne korelacije.

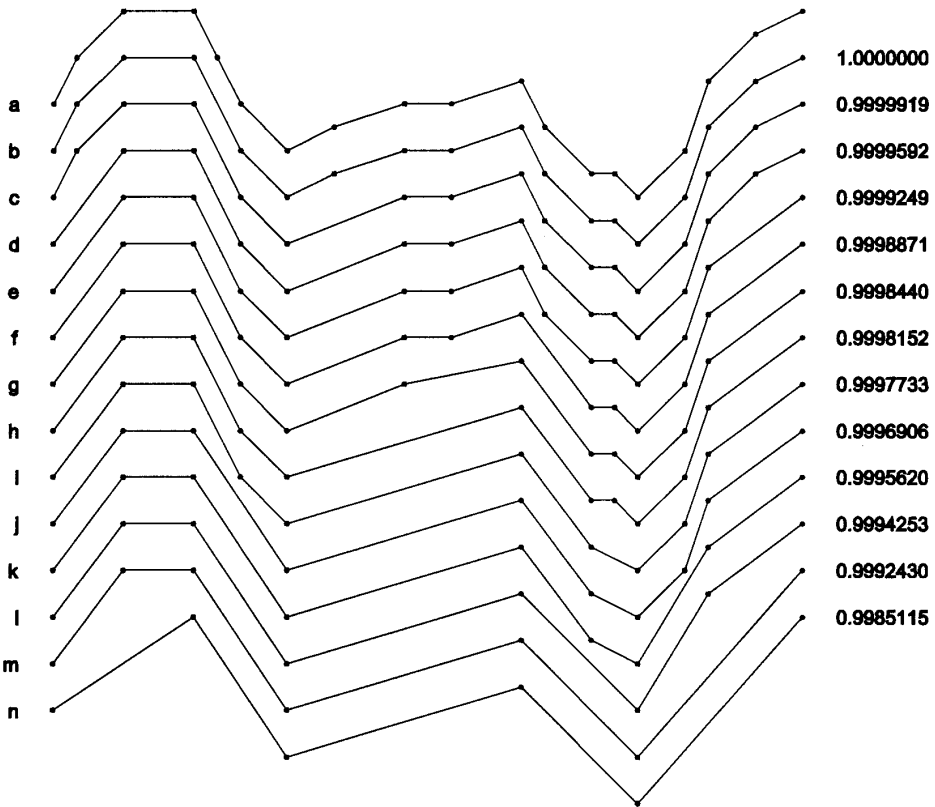
Pod pojmom sličnosti u ovom se radu misli na sličnost oblika, koja se ne smije miješati sa sličnosti u matematičkom smislu.

Sličnost je geometrijsko svojstvo. Dvije konfiguracije su u matematičkom smislu slične ako se pomoću transformacije sličnosti, tj. kompozicije translacije, rotacije i promjene mjerila, mogu prevesti jedna na drugu. Dva lika u matematičkom smislu mogu biti ili slični ili ne.

Pod pojmom sličnosti oblika podrazumijeva se svojstvo geometrijskih konfiguracija koje poprima numeričku vrijednost između 0 i 1, a mjeri se apsolutnom vrijednošću koeficijenta linearne korelacije.

### 2.1. *Pojednostavljenje linija na temelju kriterija maksimalne sličnosti*

Algoritam za pojednostavljenje linija na temelju maksimalne sličnosti može se ukratko opisati na sljedeći način. Tražimo onu točku izvorne linije čije će izbacivanje najmanje utjecati na sličnost oblika izvorne i pojednostavljene linije. Dakle, koeficijent linearne korelacije između izvorne linije i linije reducirane za jednu točku treba biti maksimalan. Pronađenu točku izbacimo. Među preostalim točkama tražimo opet onu čije izbacivanje najmanje snizuje koeficijent korelacije i izbacujemo je. Postupak se ponavlja dok se ne zadovolji unaprijed zadani kriterij za završetak postupka (npr. broj izbačenih točaka izvorne linije u skladu s općepoznatim Töpferovim zakonom drugog korijena (Töpfer i Pillewizer 1966, Töpfer 1974)).



Slika 1. Izvorna linija (a) i linije koje nastaju iz nje uzastopnim izbacivanjem po jedne točke

Na sl. 1 prikazana je najprije izvorna linija (a), a nakon toga rezultat svakog pojedinačnog koraka. Dakle, slijedi prikaz linije nakon izbacivanja jedne točke (b), zatim nakon izbacivanja ukupno dviju točaka (c), itd. Pored svake krivulje iskazan je i pripadni koeficijent linearne korelacije između izvorne linije i njenog pojednostavljenog prikaza. Na slici je vidljivo da koeficijent korelacije zaista dobro odražava vizualnu sličnost.

### 3. PRIMJENA POSTUPKA ZA PRIKAZE DIJELA OBALNE LINIJE REPUBLIKE HRVATSKE

Za ilustraciju pojednostavljenja linija opisanim postupkom izabrana je obalna linija Istre i Kvarnera. Upotrijebljeni podaci dobiveni su digitalizacijom u lokalnom sustavu digitalizatora, te su kasnije Helmertovom transformacijom koordinata transformirani u koordinatni sustav Gauss-Krügerove projekcije. Digitalizirana je karta teritorijalne podjele Republike Hrvatske u mjerilu 1:1 000 000, izrađena u Zavodu za kartografiju Geodetskog fakulteta 1979. godine u Gauss-Krügerovoj projekciji sa srednjim meridijanom  $\lambda_0 = 16^\circ 30'$ . Na sl. 2 dan je izvorni prikaz linije koja sadrži ukupno 568 točaka.



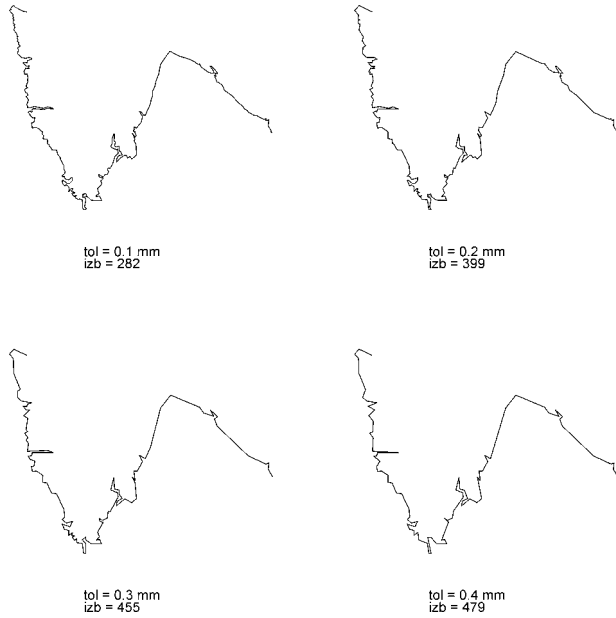
*Slika 2. Obalna linija Istre i Kvarnera u mjerilu 1:1 000 000*

Kako je u literaturi najčešći način ispitivanja primjenjivosti nekog algoritma za pojednostavljenje linija usporedba s prikazima dobivenim Douglas-Peuckerovim algoritmom te ručno generaliziranim prikazom, ovdje su dani i takvi prikazi. Osim toga, radi jednostavnije vizualne usporedbe dani su i izvorni prikazi linije smanjeni u mjerilo za koje se pojednostavljuje (vidi sliku 3).

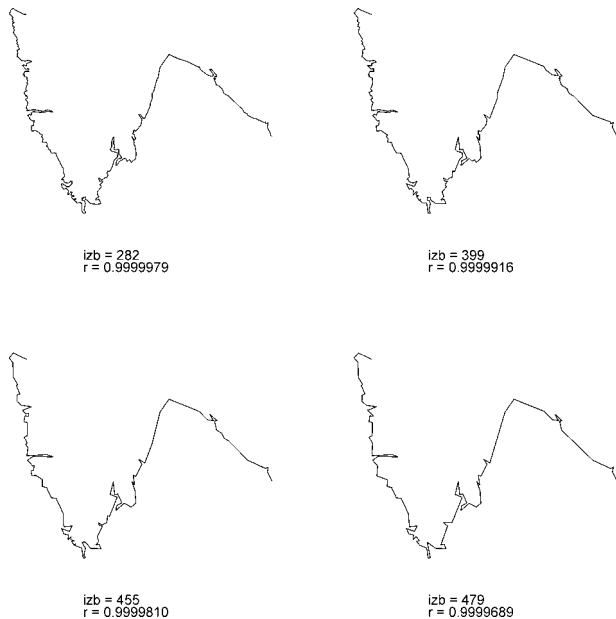


*Slika 3. Izvorni prikaz obalne linije Istre i Kvarnera smanjen u mjerilo 1:2 000 000*

Da bi se lakše uočile razlike između prikaza linija pojednostavljenih predloženim i Douglas-Peuckerovim algoritmom, na sl. 4 dani su prikazi za tolerancije 0,1 - 0,4 mm u mjerilu pojednostavljenog prikaza. Broj točaka linije koje se izbacuju na sl. 5 određen je na temelju pojednostavljenja Douglas-Peuckerovim algoritmom uz zadanu toleranciju.



*Slika 4. Prikaz obalne linije Istre i Kvarnera u mjerilu 1:2 000 000 pojednostavljen Douglas-Peuckerovim algoritmom*

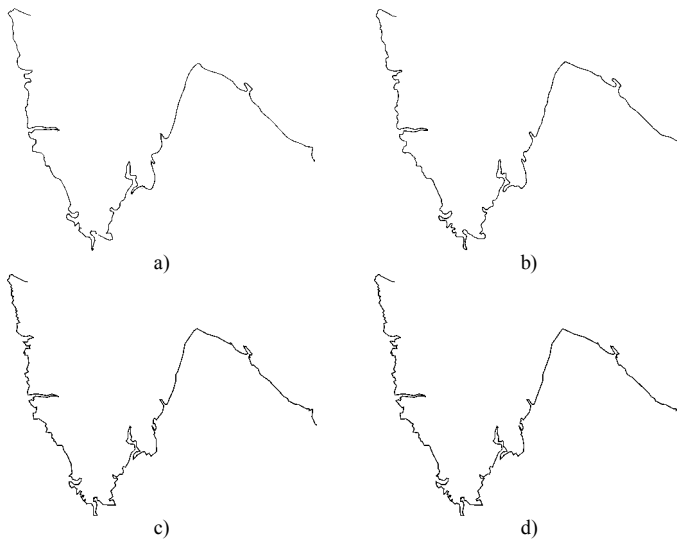


*Slika 5. Prikaz obalne linije Istre i Kvarnera u mjerilu 1:2 000 000 pojednostavljen algoritmom koji se predlaže u ovom radu*

Iako je na dvama posljednjim prikazima na sl. 4 i 5 ( $tol = 0,3$  i  $tol = 0,4$ , odnosno  $izb = 455$  i  $izb = 479$ ) prevelik stupanj pojednostavljenja za mjerilo karte 1:2 000 000 oni su ovdje ipak dani radi usporedbe prikaza dobivenih pomoću ta dva algoritma.

Uspoređujući prikaze na sl. 4 i sl. 5 sa sl. 3, može se zaključiti da se smanjivanjem broja točaka linije prikazi pojednostavljeni algoritmom danim u ovom radu sve više razlikuju od prikaza dobivenih Douglas-Peuckerovim algoritmom, te da predloženi algoritam više čuva karakteristične dijelove linije, tj. da su prikazi sličniji polaznoj, izvornoj liniji (vidi npr. Bakarski zaljev i najjužniji dio Istre).

Ostaje nam da vidimo koliko se slažu prikazi po jednoj i drugoj metodi s prikazom ručno pojednostavljene linije (sl. 6 i sl. 7). Prikazi dobiveni ručnom generalizacijom (slike 6a i b, 7b i c) preuzeti su iz rada N. Vučetić (2000). U tom radu upotrijebljeni su isti izvorni podaci.

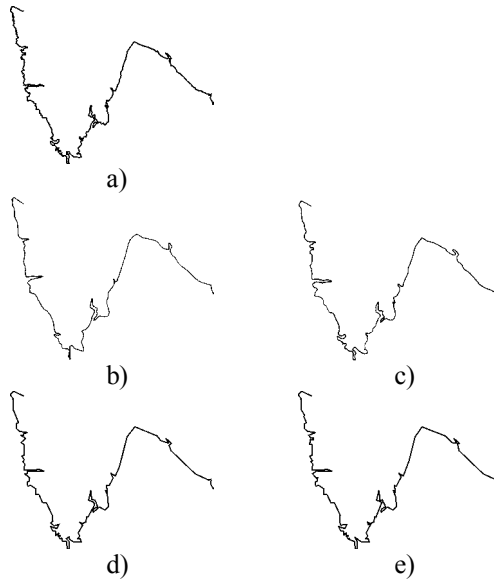


*Slika 6. Obalna linija Istre i Kvarnera u mjerilu 1:2 000 000, ručna generalizacija:*

- a) izvorni prikaz smanjen pa generaliziran,*
- b) izvorni prikaz generaliziran za mjerilo 1:2 000 000, potom smanjen fotografski u mjerilo 1:2 000 000,*
- c) prikaz pojednostavljen Douglas-Peuckerovom metodom (izbačeno 327 točaka),*
- d) prikaz pojednostavljen predloženom metodom (izbačeno 327 točaka)*

Ako pogledamo sl. 6a i sl. 7b, uočava se da je preveliki stupanj generalizacije. To je najvjerojatnije zato što to nije uobičajen način ručne generalizacije. Naime, ovdje je prvo izvorni prikaz smanjen u mjerilo izvedene karte te potom generaliziran. Na sl. 6b i sl. 7c dan je klasičan način ručne generalizacije (izvorni prikaz generaliziran za željeno mjerilo izvedene karte te potom fotografski smanjen u to mjerilo). Na tim prikazima vidljivo je da se izgubio karakteristični oblik Bakarskog zaljeva. Treba napomenuti da je obje generalizacije proveo isti kartograf.

Ako se izuzme prikaz Bakarskog zaljeva, vidljivo je da se prikaz dobiven danim algoritmom vrlo dobro slaže s prikazom ručne generalizacije. Posljednja dva prikaza linije na sl. 6 i sl. 7 sadrže broj točaka koji je u skladu s Töpferovim zakonom drugog korijena.



*Slika 7. Obalna linija Istre i Kvarnera u mjerilu 1:4 000 000,*

*a) smanjeni izvorni prikaz, ručna generalizacija;*

*b) izvorni prikaz smanjen pa generaliziran,*

*c) izvorni prikaz generaliziran za mjerilo 1:4 000 000 i potom smanjen fotografski u mjerilo 1:4 000 000,*

*d) prikaz pojednostavljen Douglas-Peuckerovim algoritmom (izbačeno 460 točaka),*

*e) prikaz pojednostavljen predloženim algoritmom (izbačeno 460 točaka)*

#### 4. ZAKLJUČAK

Postupci za automatsko pojednostavljenje linija koje se može naći u literaturi postižu smanjenje informacijskog sadržaja primjenom kriterija koji se temelje na metričkim pokazateljima kao što su linearno odstupanje, površina ili promjena smjera. Time se postiže željena točnost prikaza, ali se ne uspijeva najbolje zadržati karakteristični oblik linije, što proturječi osnovnoj ideji kartografske generalizacije.

Pojam karakterističnog oblika linije, koji se nastoji očuvati pri ručnoj generalizaciji, usko je povezan s vizualnom sličnošću (između izvorne i pojednostavljene linije). Kako se vizualna sličnost u dobroj mjeri može opisati iznosom koeficijenta linearne korelacije, u ovom su radu dana načela postupka za pojednostavljenje linija koji se temelji na maksimalnoj sličnosti, dakle na maksimalnom koeficijentu linearne korelacije između izvorne i pojednostavljene linije.



Pokazalo se da predloženi postupak daje prihvatljivije rezultate nego poznati postupak koji kao kriterij upotrebljava metričke značajke. Novim se postupkom bolje zadržavaju karakteristični detalji linije i dobivaju rezultati koji su bliži rezultatima ručne kartografske generalizacije.

## LITERATURA

- Cromley, R. G. (1991): Hierarchical methods of line simplification, *Cartography and Geographic Information Systems*, Vol. 18, No. 2, 125-131.
- Douglas, D. H., Peucker, T. K. (1973): Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, *The Canadian Cartographer*, Vol. 10, No. 2, 112-122.
- Frančula, N. (2000): Kartografska generalizacija, Sveučilište u Zagrebu, Geodetski fakultet.
- Hassen, K., Beard, K. (1998): Visual evaluation of GIS algorithms using a reference grid, *Cartography and Geographic Information Systems*, Vol. 25, No. 1, 42-50.
- Ivanov, V. V. (1965): O nekih mogućnostima automatizacije topografskih kart, *Geodezija i kartografija*, 1, 62-66.
- João, E. M. (1998): Causes and consequences of map generalisation, *Research Monographs in Geographical Information Systems*, Taylor & Francis, London.
- Lang, T. (1969): Rules for robot draughtsmen, *Geographical Magazine*, 1, 50-51.
- Lechthaler, M., Kasyk, S. (1999): Systemunterstützte kartographische Generalisierung linearer Kartenobjekte, *KN*, 2, 52-59.
- Lee, D. (1995): Experiment on formalizing the generalization process, U: Müller, J. C., Lagrange, J. P., Weibel, R. (urednici): *GIS and generalization, Methodology and practice, GISDATA I*, Taylor & Francis, London, 219-234.
- Li, Z., Openshaw, S. (1992): Algorithms for automated line generalization based on a natural principle of objective generalization, *International Journal of Geographical Information Systems*, Vol. 6, No. 5, 373-389.
- McMaster, R. B., Shea, K. S. (1992): *Generalization in digital cartography*, Association of American Geographers, Washington.
- Müller, J. C., Weibel, R., Lagrange, J. P., Salgé, F. (1995): Generalization: state of the art and issues, U: Müller, J. C., Lagrange, J. P., Weibel, R. (urednici): *GIS and generalization, Methodology and practice, GISDATA I*, Taylor & Francis, London, 3-17.
- Petrović, S. (1991): Geometry of the correlation coefficient and its application in geodesy, *Mitteilungen der geodätischen Institute der Technischen Universität Graz*, Folge 71.
- Petrović, S., Lapaine, M., Vučetić, N. (1988): Some new interpretations of the linear correlation coefficient, 6<sup>th</sup> Seminary in applied mathematics, *Tara* 31. 8.-3. 9. 1988.
- Petrović, S., Vučetić, N., Čolić, K. (1987): Jedna nova interpretacija koeficijenta linearne korelacije pri obradi podataka mjerenja, *Geodetski list*, 10-12, 281-295.
- Ramer, U. (1972): An iterative procedure for the polygonal approximation of plane curves, *Computer Graphics and Image Processing*, 1, 244-256.

- Reumann, K., Witkam, A. P. M. (1974): Optimizing curve segmentation in computer graphics, International Computing Symposium, North-Holland Publishing Company, Amsterdam, 467-472.
- Thapa, K. (1988): A review of critical points detection and line generalization algorithms, Surveying and Mapping, Vol. 36, No. 3, 185-205.
- Töpfer, F. (1974): Kartographische Generalisierung, VEB Hermann Haack, Geographisch-Kartographische Anstalt Gotha, Leipzig.
- Töpfer, F., Pillewizer, W. (1966): The principles of selection, The Cartographic Journal, Vol. 3, No. 1, 10-16.
- Visvalingam, M., Whyatt, J. D. (1993): Line generalization by repeated elimination of points, The Cartographical Journal, Vol. 30, No. 1, 46-51.
- Vučetić, N. (2000): Kartografska generalizacija i AutoCAD Map, Seminarski rad, Sveučilište u Zagrebu, Geodetski fakultet.
- Vučetić, N. (2001): Generalizacija linijskih elemenata karte po kriteriju maksimalne sličnosti, Doktorska disertacija, Sveučilište u Zagrebu, Geodetski fakultet.
- Wang, Z., Müller, J.-C. (1998): Line generalization based on analysis of shape characteristics, Cartography and Geographic Information Systems, Vol. 25, No. 1, 3-15.