

# GEOGRAFICKÝ ČASOPIS

46

1994

1

*Karol Husár\**

## VEKTOROVÁ DIGITALIZÁCIA AREÁLOVÝCH OBJEKTOV A ICH PRIESTOROVÁ ANALÝZA V PODMIENKACH GEOGRAFICKÉHO ÚSTAVU SAV

Karol Husár: Vector digitalization of areal objects and their spatial analysis in the conditions of the Institute of Geography SAS. Geogr. čas., 46, 1994, 1, 4 figs., 16 refs.

Presented work contains the basic principles and techniques of the area of vector data digitalization and their primary processing oriented to those that were applied in programme system APTAB and separately in the programme system DIGEDIT, in both cases products of the Institute of Geography of the SAS. The work therefore represents a certain methodological salient point for digitized processing of any analog form.

Key words: vector digitalization, morphometric analysis.

### ÚVOD

K ťažiskovým problémom spracovania priestorových dát počítačovými systémami patrí vstup údajov do počítačového prostredia. Na rozdiel od zariadení, ktoré priestorové údaje získavajú v rastrovom móde (skenerové digitalizačné zariadenia, multispektrálne skenery atď.), stále značná časť údajov existuje v analógovej forme a musí byť predmetom konverzie do digitálnej, počítačovo akceptovateľnej formy vektorovými digitalizačnými

---

\* Geofyzikálny ústav SAV, Dúbravská cesta 9, Bratislava

zariadeniami. Tento proces označujeme ako poloautomatická alebo manuálna *vektorová digitalizácia*.

Vektorová digitalizácia reprezentuje jednu z možných foriem *vstupu priestorových dát do prostredia GIS-u*, obsahujúceho priestorové, deskriptívne a časové aspekty objektov (javov), pričom je možné vyčleniť štyri úrovne vstupu dát (Husár, manuscript):

- a) vstup priestorovo-lokalizovaných dát (digitalizácia),
- b) vstup deskriptívnych (nie priestorovo-lokalizovaných) dát,
- c) vstup časovo-orientovaných dát,
- d) spájanie priestorových, deskriptívnych a časových dát.

Ťažiskovým bodom tejto práce je digitalizácia vektorových objektov ako súčasť a) a náčrt možností ich priestorovej analýzy v podmienkach Geografického ústavu SAV (GÚ SAV).

Uprednostnenie vektorovej pred rastrovou digitalizáciou je podmienené jej vysokou presnosťou, ktorá môže spĺňať aj náročné kartografické kritériá. Navyše, druhým závažným dôvodom, ktorý hovorí v prospech vektorovej digitalizácie pred rastrovou digitalizáciou skenerovým zariadením, sú v prípade rastrovej digitalizácie skenerom obrovské nároky na pamäťové médiá.

Na druhej strane, k nedostatkom vektorovej digitalizácie patrí jej časová náročnosť a náchylnosti k chybám spôsobených ľudským faktorom.

## 1 DIGITALIZÁCIA V BÝVALEJ ČSFR

Tak ako v iných oblastiach, digitalizácia u nás zaostáva za vývojom vo svete, a to aj napriek asi 15 až 20-ročnej snahe niektorých organizácií urobiť v tejto oblasti isté kroky. Aktivity jednotlivých organizácií v tomto smere sa vyznačovali vzájomnou nekoordinovanosťou krokov, metodickou, ako aj aplikačnou redundanciou.

Doterajší výskum v oblasti digitalizácie bol u nás sústredený najmä na digitalizáciu niektorých *základných mapových prvkov*, a to najmä: administratívne hranice, vodstvo a reliéf. Menej už na komunikácie, sídla, hranice lesa a tematické prvky mapy. Riešiteľmi úloh týkajúcich sa digitalizácie boli predovšetkým pracoviská FMNO, ČÚGK, SÚGK, SVŠT (terajšia STÚ) Bratislava, Geografický ústav ČSAV Brno, Terplan Praha, Lesprojekt Zvolen a Brandýs nad Labem, Urbion Bratislava (resp. SÚÚRA Bratislava), Geografický ústav SAV Bratislava a iné. V poslednom období sa vďaka existencii výkonného výpočtového systému ARC/INFO na niektorých okresných úradoch životného prostredia začalo s digitalizačnými prácami na príslušnej regionálnej úrovni.

V rezorte FMNO bývalej ČSFR sa pomerne dlhé časové obdobie sústredili na teoreticko-metodologické aspekty digitalizácie, najmä z oblasti vodstva a reliéfu. Z nich osobitne problematika digitalizácie *liniových hierarchicko-sieťových entít* (vodstva) bola podrobne rozpracovaná v práci Moravca (1986). V realizačnej fáze bola táto metodika uplatnená pri digitalizácii niekoľkých mapových listov z územia ČR. Z oblasti digitalizácie prvku reliéf, rezort FMNO vlastní zdigitalizované nadmorské výšky z územia ČSFR, a to vo forme štvorcovej mriežky so stranou 1 km a 100 m. Nedostatkom týchto dát je relatívne vysoký výskyt chýb. Okrem toho FMNO spravovalo digitálne údaje všetkých základných mapových prvkov topografickej mapy 1:200 000. Ide o jediné digitálne mapové dielo bývalej ČSFR.

Na Geografickom ústave ČSAV (t.č. už zrušené) bolo roku 1987 vytvorené u nás ojedinelé dielo "Atlas obyvateľstva ČSSR", reprezentujúce údaje z censusu. Islo vtedy o prvé mapové dielo v ČSFR, ktorého časť kartografického spracovania tematického obsahu bola vytvorená automatizovaným spôsobom na báze topologicky zdigitalizovaných administratívnych hraníc po úroveň okresov v mierke 1:750 000 a digitalizovaných údajov referenčných bodov základných sídelných jednotiek. Tematický obsah atlasu bol vyjadrený šrafovaním a kartodiagramami. Bývalý GÚ ČSAV okrem toho zdigitalizoval administratívne hranice základnej mapy ČSFR po úroveň katastrálneho územia obce v mierke 1:50 000.

SÚÚRA rovnako vlastní topologické digitálne údaje administratívnych hraníc katastrálnych území obcí v mierke 1:50 000, avšak iba z územia SR. Navyše spravuje digitálne údaje 75. funkčných typov zastavaných plôch vybraných sídiel Slovenska získaných zo základnej mapy mierky 1:10 000.

Okrem uvedených pracovísk má Geografický ústav SAV zdigitalizované areály geomorfologických jednotiek Slovenska (Mazúr a Lukniš, 1980) v mierke 1:200 000. GÚ SAV vlastní tiež digitálne údaje Interpretáčnej schémy foriem využitia krajiny juhozápadnej časti Slovenska v mierke 1:200 000 (Feranec a Otáhel, 1992) a formy využitia krajiny okolia Šurian v mierke 1:50 000 (Feranec et al. v tlači). Uvedené typy údajov sú digitalizované v topologickej štruktúre.

Z iných pracovísk, ktoré vlastní digitálne údaje mapových alebo tematických prvkov, spomeňme Geofond Bratislava, ktorý spravuje digitálne údaje zosuvov a iných svahových deformácií a digitálne údaje skládok odpadov tuhého komunálneho odpadu v mierke 1:25 000. Lesprojekt Zvolen t.č. dokončuje digitalizáciu máp lesných hospodárskych celkov Slovenska v mierke 1:25 000. Rovnako Lesprojekt Brandýs nad Labem digitalizuje lesnícke mapy mierky 1:25 000. V rezorte SÚGK boli digitalizované sídla, komunikácie, vodstvo, hranice a výškopis pre medzinárodnú mapu Európy v mierke 1:2 500 000, ako aj základné mapové prvky z oblasti Vysokých Tatier získaných digitalizáciou interpretačných schém systémom Gradis.

Z rodiaciach sa projektov na poli digitalizácie priestorových dát je potrebné spomenúť najmä rozsiahly projekt Zeměměřického ústavu v Prahe, na základe ktorého sa od 1.1.1993 začali realizovať práce na tvorbe kartografickej bázy dát skenerovou digitalizáciou základných máp v mierke 1:10 000 v súradnicovom systéme JTSK. Pre ilustráciu ide o 4,5 tisíc mapových listov, z ktorých každý bude rozložený na vrstvy jednotlivých mapových prvkov. Projekt, z ktorého iba hodnota hardverových a softverových komponentov bola odhadnutá na čiastku 40 miliónov Kčs, je postavený na rastrovej digitalizácii, jej vektorizácii a topologizácii. Ťažiskom tohto projektu bude vektorizácia, následná identifikácia objektov a topologizácia.

Podobný projekt na Slovensku plánovala realizovať SÚGK (t.č. ÚGKK), avšak práce nie sú v takom štádiu prípravy, ako je to v Českej republike. T.č. sa pod patronátom ÚGKK vyvíja koncepcia takejto digitalizácie a hľadajú sa zdroje na jej realizáciu.

## 2 PRINCÍP VEKTROVEJ TOPOLOGICKEJ DIGITALIZÁCIE

Väčšina už uvedených disponibilných digitálnych údajov v bývalej ČSFR bola získaná vektorovou topologickou digitalizáciou. Na ilustráciu predností topologických dátových

štruktúr pred jednoduchými vektorovými štruktúrami vychádzajúcimi z princípu spaghetti modelu (Pequet 1984, Smith et al. 1987, Husár 1992) uveďme nevýhody týchto jednoduchých polygónových štruktúr (Burrough 1986):

- a) hraničné segmenty medzi susednými polygónmi sú digitalizované dvakrát, čo nevyhnutne vedie k získaniu dvoch rôznych hraničných čiar medzi susednými polygónmi;
- b) absencia akýchkoľvek topologických informácií;
- c) nemožnosť akceptovať tzv. ostrov, t.j. polygón obsiahnutý v inom polygón;
- d) zložitejší spôsob identifikácie a opravy chýb vyplývajúcich z digitalizácie.

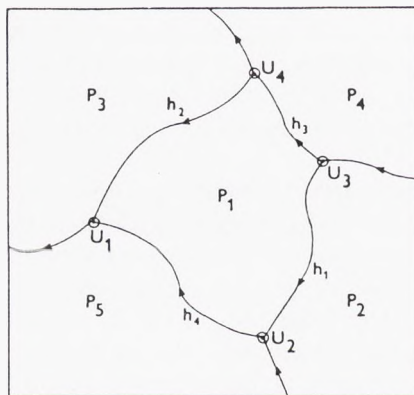
Závažným nedostatkom jednoduchých polygónových štruktúr je, ako bolo uvedené v bode c), že nedokážu pracovať s ostrovmi. Tento problém je síce v niektorých špeciálnych prípadoch riešiteľný aj v rámci spaghetti modelu, nie je to však riešenie všeobecne použiteľné pri rôznych typoch aplikačných úloh priestorovej analýzy. Uspokojivé vyriešenie tohto problému je možné iba s pomocou explicitne vyjadrených topologických vzťahov, teda na úrovni vektorovej topologickej digitalizácie.

Keďže ďalšia časť práce sa týka vektorovej topologickej digitalizácie, stručne uvedieme základné pojmy danej problematiky.

Elementárnou logickou jednotkou vo vektorovom modeli je *bod a čiara* (vektor), ktorá v geografickom kontexte môže korešpondovať s cestnou komunikáciou, riekou alebo s hraničným segmentom polygónu<sup>1</sup>.

Pod *topológiou* rozumieme explicitné vyjadrenie vzťahov medzi entitami v priestore.

Na obr. 1, znázorňujúcom základné topologické elementy  $u_1$ ,  $u_2$ ,  $u_3$  a  $u_4$  reprezentujú *uzlové body* (spravidla tzv. *trojné body poľa*), teda miesto styku *hraničných segmentov*



Obr. 1. Základné topologické prvky a ich analógové vzťahy.

(arcs), ktoré sú na obrázku znázomené ako  $h_1$ ,  $h_2$ ,  $h_3$  a  $h_4$ . Hraničné segmenty zoradené vo zvolenom smere vytvárajú *polygóny*, ktoré sú na obrázku reprezentované ako  $P_1$ ,  $P_2$ ,  $P_3$ ,  $P_4$  a  $P_5$ . Šípky na hraničných segmentoch naznačujú, že ide o orientované hrany.

Pre techniku topologickej digitalizácie je nevyhnutné definovať topologické parametre, napr. názov polygónu po ľavej strane a názov polygónu po pravej strane daného hraničného

<sup>1</sup> Termín "polygón" je používaný v oblasti vektorovej počítačovej grafiky. Jeho pendantom na poli geografie je pojem "areál" alebo "región". V práci tieto termíny používame v synonymnom význame.

segmentu. (Smer vľavo a vpravo pritom chápeme relatívne, v závislosti od smeru postupu digitalizácie). Sú možné 2 riešenia definovania topológie, a to počas vlastnej digitalizácie, alebo po nej (Chrisman 1987). Prikláňame sa k prvému spôsobu, pričom tento spôsob bol použitý aj v nižšie uvedenom programe APTAB.

Aby sme mohli uvažovať o priestorovej analýze nad dátami vo vektorovom formáte, je potrebné, aby priestorové dáta boli topologicky konzistentné.

Pod *topológiou* rozumieme explicitné vyjadrenie vzťahov medzi entitami (bodmi, hranami) v priestore.

Vymedziť pojem "topologický" možno na príklade deformácie mapy. Deformovaním mapy sa menia niektoré jej atribúty, a to dĺžky, uhly a plošné rozsahy. Iné atribúty zostávajú konštantné. Sú to napr. "susednosť", priestorové vzájomné vzťahy ako "nachádza sa v (vo)", "pretína" atď. Okrem toho deformovaním mapy typu priestorových objektov zostávajú nezmenené, t.j. body zostávajú bodmi, čiary čiarami, areály areálmi. Topologické vlastnosti sú potom všetky tie vlastnosti, ktoré po deformácii mapy zostanú konštantné.

Predpokladom uplatnenia topológie pri digitalizácii je splnenie nasledovných podmienok (Husár, manuscript):

i) každý bod príslušného analógového podkladu je súčasťou ("leží", "nachádza sa v(vo)") práve jednej entity (okrem hraničného segmentu),

ii) areály pokrývajú celý rozsah územia,

iii) každý hraničný segment oddeľuje práve dva areály,

iiii) areály sa neprekrývajú.

Tvrdenie v iii) nie je nezávislé od tvrdenia v i). Uviedli sme ho iba ako názorný dôsledok tvrdenia v i) a z dôvodov, že na rozdiel napr. od Haggetta (1965), ktorý túto situáciu uvádza pod názvom "prekrývajúce hranice", my pri vektorovej topologickej digitalizácii prekrývajúce areály neberieme do úvahy.

Na základe uvedeného možno tvrdiť, že predmetom vektorovej topologickej digitalizácie je územie spĺňajúce podmienky i), ii), a iii).

### 3 DIGITALIZÁCIA NA GEOGRAFICKOM ÚSTAVE SAV

V súčasnosti sú z hľadiska softveru i hardveru na Geografickom ústave SAV vytvorené dobré podmienky na digitalizáciu analógových dát, ako aj ich manipuláciu a výstup. Z oblasti hardveru ústav vlastní poloaautomatické zariadenie Tektronix 4958 formátu A0, s rozlišovacou schopnosťou 0.025 mm a ploter Hewlett-Packard formátu A0, ktoré sú pripojené k počítaču IBM PC/386. Zo softverových produktov ústav prevádzkuje systém IDRISI (Clark University), Design CAD 2-D (American Small Business Computers), program TYDIG ako relatívne samostatnú časť systému SPANS 5.0 (Intera Tydac) a vlastný softver na vektorovú digitalizáciu APTAB a DIGEDIT, určeného na prvé spracovanie a priestorovú analýzu dát.

Ak vychádzame z požiadavky topologickej digitalizácie, potom systém IDRISI a Design CAD 2-D nie sú určené na takýto druh digitalizácie.

Pre potreby topologickej vektorovej digitalizácie slúži program TYDIG a APTAB. Komerčný systém SPANS 5.0 (INTERA TYDAC, 1991), ktorého súčasťou je aj program



TYDIG, patrí medzi výkonné GIS-y, operujúci so všetkými v súčasnosti bežnými základnými dátovými modelmi, teda: s vektorom, mriežkou a štvorstromom (quadtree). Akokoľvek, SPANS narába s vektorovými štruktúrami do tej miery, pokiaľ sú súčasťou TYDIG-u. Väčšina ostatných operácií a aplikácií po importovaní zdigitalizovaných údajov z TYDIG-u do SPANS-u a konvertovaní do mriežky sa dejú na úrovni štvorstromu. Z toho je zjavné, že SPANS z hľadiska priestorových analýz nie je vektorovo orientovaným GIS-om. Podstata problému spočíva v transformácii vektorových dát do mriežky, pri ktorej dochádza k istému úbytku informačného potenciálu. Následné priestorové analýzy v SPANS-e sú takto zaťažené istou chybou vyplývajúcou z danej konverzie.

Z uvedených dôvodov na účely vektorovej topologickej digitalizácie na GÚ SAV preferujeme program APTAB (FORTRAN 77 pod MS-DOS), ktorý umožňuje všetky základné operácie potrebné pri topologickej digitalizácii polygónov.

Postup, ktorým sa uskutočňuje transformácia analógových dát do počítačového prostredia a ktorého súčasťou je aj vektorová topologická digitalizácia programom APTAB, možno rozdeliť do 3 základných krokov (Husár, manuscript):

1. príprava grafického podkladu na digitalizáciu,
2. vlastná digitalizácia programom APTAB,
3. kontrola a opravy zdigitalizovaných údajov.

1. Pred vlastnou digitalizáciou je vhodné uskutočniť prípravu (adjustáciu) grafického podkladu, a to buď priamo na jeho originále alebo na jeho kópiách. Kvôli zachovaniu originálu však doporučujeme prípravu realizovať na jeho kópiách, napr. fóliách, výhodou ktorých (v porovnaní s papierovými médiami) o.i. je ich stabilnosť s ohľadom na kartografické deformácie.

Pod prípravou grafického podkladu rozumieme:

- mierkové transformácie (spravidla v smere zhotovenia zväčšeniny) - v prípade, ak sa čiarové prvky mapového podkladu, ktoré budú predmetom digitalizácie, vyznačujú veľkou krivosťou alebo hustotou na jednotku plochy alebo pri existencii väčšieho počtu relatívne malých areálových jednotiek (rádovo do niekoľko  $\text{mm}^2$ ). Pritom je zásadne dôležité dbať na to, aby nedošlo ku kartografickým deformáciám;
- určenie a vyznačenie identických bodov potrebných na výpočet transformačných koeficientov na prevod súradníc z lokálneho súradnicového systému do požadovaného globálneho;
- separácia, vyznačenie ( a identifikácia) prvkov mapového podkladu, ktoré budú predmetom digitalizácie, zvlášť v prípade obsahovo preťaženej a neprehľadnej mapy, obsahujúcej aj také prvky, ktoré nie sú predmetom digitalizácie;
- určenie poradia digitalizácie prvkov, prípadne logických celkov postupu digitalizácie;
- vyznačenie tzv. štartovacích bodov, t.j. bodov začiatku a konca digitalizácie jednoduchých ostrovov. Možno odporučiť, aby štartovací bod bol na hraničnom segmente ostrova zvolený v tých miestach, kde je krivosť najmenšia;
- fixovanie mapového podkladu a pomocnej "menu" tabuľky (v prípade jej existencie - s ohľadom na použitý systém) na účinnú snímáciu plochu digitizéra.

Nie všetky z uvedených krokov prípravy mapového podkladu sú rovnako dôležité, pričom niektoré z nich môžu byť používateľom ignorované. Praktické skúsenosti autora z prípravy grafického podkladu, ako aj z vlastného procesu digitalizácie však ukazujú, že ich precízna aplikácia zásadnou mierou ovplyvňuje korektnosť postupu a presnosť digitalizácie, teda aj kvalitu získaného dátového fondu, ako aj rozsah opravných procesov vedúcich k odstráneniu vzniknutých chýb.

2. Elementárnymi prvkami vlastnej digitalizácie sú súradnice polohy bodu, ktoré pri digitalizácii sú dané polohou snímačej hlavice digitizéra (kurzora, myši), resp. ortogonálneho priemetu stredu jej osového križa v čase stlačenia príslušného tlačidla. To znamená, že pri digitalizácii bodu snímacia hlavica sa stredom osového križa presunie kolmo nad daný bod a stlačí sa tlačidlo "Z" na snímačej hlavici digitizéra TEKTRONIX 4958.

Digitalizované súradnice bodov sú primárne zaznamenané v lokálnom ľavotočivom kartézskom súradnicovom systéme so začiatkom v ľavom dolnom rohu účinnej snímačej plochy digitizéra.

Používateľ programového systému APTAB po jeho odštartovaní musí najskôr v dialógovom režime s počítačom zadať niekoľko údajov, tvoriacich tzv. *úvodnú vetu* (digitálneho) súboru, obsahujúcu:

- zvolenú rozlišovaciu schopnosť digitizéra (jednu z hodnôt 0.025, 0.0254, 0.05, 0.0508, 0.0635, 0.1, 0.127 alebo 0.254 mm). Pre bežné druhy digitalizácií na poli geografie postačuje voľba menšej rozlišovacej schopnosti (napr. 0.1 mm kvôli jednoduchému prevodu medzi relatívnymi a požadovanými globálnymi súradnicami v prípade kontroly a odstraňovania chýb);

- minimálne 4 a maximálne 20 identických bodov na výpočet transformačných koeficientov, pri výpočte ktorých APTAB používa algoritmus afinnej transformácie. Vypočítané transformačné koeficienty slúžia najskôr na kontrolu korektnosti identických bodov, t.j., z lokálnych súradníc identických bodov sa transformačnými rovnicami vypočítajú globálne súradnice a tie sa porovnávajú so zadanými. V prípade, že rozdiel medzi vypočítanými a zadanými globálnymi súradnicami identického bodu je väčší ako stanovený parameter  $d$ , tento bod je vylúčený zo zoznamu identických bodov a opakuje sa celá procedúra afinnej transformácie, a to až dovtedy, kým rozdiely medzi zadanými a vypočítanými súradnicami všetkých identických bodov v globálnej súradnicovej sústave budú  $\leq d$  (napr. 5m v APTAB-e);

- zadanie rozsahu digitalizovaného územia ( 4 body);

- určenie polohy "menu" tabuľky (pomocou 2 bodov), ktorá slúži na operatívne zmeny funkcií počas digitalizácie.

Nebudeme sa tu zaoberať všetkými možnými variantmi snímania<sup>2</sup>, ktoré je možné programom APTAB aplikovať. Spomenieme iba, že program umožňuje digitalizáciu v 4 základných *režimoch snímania*, t.j. v režime "BODY" (poloha individ. bodu), v režime "HRANA" (poloha a topológia), v režime "POLYGÓN" (poloha uzatvoreného polygónu) a v režime "SKALÁR" (poloha a skalár). Okrem týchto základných funkcií program umožňuje elimináciu invalidných bodov, resp. úsekov hraničných segmentov v každom z uvedených režimov snímania.

Z hľadiska cieľa práce sa digitalizácia areálových objektov v APTAB-e realizuje snímaním v režime "HRANA", umožňujúcim topologicky orientovanú digitalizáciu. Postup je daný voľbou logickej sekvencie krokov A až E (Husár, manuscript):

- A; v tabuľke "menu" sa zvolí režim snímania "HRANA", a číselné identifikátory objektov nachádzajúcich sa **vľavo** a **vpravo** od snímačej (orientovanej) hrany a kód hrany, definujúci jej špecifikáciu s ohľadom na jej polohu na okraji mapového listu;

- B; Digitalizácia priebehu podrobných bodov hraničného segmentu sa uskutočňuje vo zvolenom smere od uzlového bodu poľa po uzlový bod tlačidlom "Z" snímačej hlavice digitizéra, s hustotou (presnosťou) závislou od

<sup>2</sup> Termíny "snímanie" a "snímať" používame ako synonymné ekvivalenty k termínom "digitalizácia" a "digitalizovať".

účelu a voľby používateľa, resp. operátora. Vo všeobecnosti, hustota voľby podrobných bodov hraničného segmentu je priamo závislá od jeho krivosti. Každý druhý a ďalší podrobný bod hrany je na monitore lineárne spájaný s predchádzajúcim zosnímaným bodom tej istej hrany a teda stáva sa novým vrcholom polygónu. V pravom hornom rohu monitora sa po zosnímaní ľubovoľného bodu, a teda aj podrobného bodu hrany, objaví transformáciou získané jeho digitálne globálne súradnice. Ukončenie snímania podrobných bodov príslušnej hrany je dané buď novou voľbou režimu snímania "HRANA" v tabuľke "menu" alebo tzv. opakovacím tlačidlom (pozri ďalej).

- C; Našou snahou bolo vyhnúť sa zbytočným presunom snímačej hlavice digitizéra (od posledného zdigitalizovaného bodu príslušnej hrany do tabuľky "menu" kvôli opakovanej voľbe režimu "HRANA"), pri ktorých často dochádza k strate kontextu postupu digitalizácie (a tým aj k časovým stratám a možným chybám). APTAB preto rozpoznáva tzv. opakovacie tlačidlo, nachádzajúce sa vpravo hore na snímačej hlavici digitizéra označené číslom "3", stlačenie ktorého má za následok nastavenie počítača na digitalizáciu novej hrany. Snímacia hlavica pritom zostáva na mieste posledne zosnímaného uzlového bodu, ktorý v konkrétnom prípade je pravdepodobným počiatočným bodom v poradí ďalšej digitalizovanej hrany. Funkcia tlačidla "3" je teda ekvivalentná funkcii v tabuľke kódov zosnímanému režimu "HRANA", avšak z uvedených dôvodov preferujeme využitie práve tohto tlačidla.

- D; Jednoduchý ostrov sa digitalizuje ako uzatvorená hrana so začiatkom a koncom v tom istom bode, pričom rozdielne súradnice počiatočného a koncového bodu sú nahradené priemerovaním vyrovnanými súradnicami. Pri digitalizácii hrán, tvoriacich jednoduchý ostrov je nevyhnutné konštantne definovať smer postupu digitalizácie týchto hrán. APTAB vyžaduje pravotočivý smer postupu digitalizácie ostrova, pri ktorom okolie ostrova sa nachádza vždy po ľavej strane ostrova a ľubovoľný vnútorný bod ostrova sa nachádza vždy po pravej strane od jeho hraničného segmentu. Dodržanie tohto pravidla je dôležité z hľadiska ďalších, topologicky orientovaných operácií. (Dodajme, že kód hrany jednoduchého ostrova je rovný 5.)

- E; Chybný podrobný bod hrany sa zo súboru vyhadzuje zvolením funkcie "Delete hrana (úsek)" v tabuľke "menu". N-tým opakovaním tejto funkcie je možné vyhodit' n zosnímaných bodov aktuálne snímanej hrany. Ich počet je daný počtom stlačení tlačidla "Z", pričom na obrazovke sa zruší čiara medzi 2 posledne zosnímanými podrobnými hranami.

Nová voľba režimu "HRANA" alebo voľba opakovacieho tlačidla "3" spôsobuje definitívnu akceptáciu práve zosnímanej hrany počítačom, a teda vylučuje možnosť uskutočniť akékoľvek opravy na úrovni tejto hrany v priebehu digitalizácie. Preto, vzhľadom na minimalizáciu chýb a z nich vyplývajúcich korekcií po digitalizácii, je vhodné pred voľbou novej hrany uskutočniť na monitore vizuálnu kontrolu digitalizovanej hrany.

3. Dôležitosť *vyhľadávania* (detekcie) a *opravy* (korekcie) kontroly a opravy zdigitalizovaných údajov a ich vplyv na validitu aplikačných výstupov je zrejmé. V tejto časti máme na mysli najmä takú kontrolu a opravu, ktorá nasleduje po procese vlastnej digitalizácie, v rámci programového systému DIGEDIT, ktorý okrem viacerých aplikačne zameraných procedúr slúži aj na primárnu detekciu chýb. DIGEDIT nebol stavaný na editovanie súborov, preto sa všetky diskrétné opravy v základnom digitálnom súbore uskutočňujú pomocou profesionálnych editovacích programov ako Multi-Edit (American Cybernetics) alebo staršieho editora Edit 3 (IBM).

Na elementárne spôsoby zisťovania chýb možno použiť okná (windows) v DIGEDIT, a to vizuálnou konfrontáciou originálu mapového podkladu a jej počítačovo vizualizovanej digitálnej reprezentácie.

Ťažisko detekcie invalidných dát sa uskutočňuje na úrovni procedúry spájania hrán v uzlovom bode a čiastočne na úrovni generovania polygónov.



Procedúra spájania hrán identifikuje všetky hrany, ktoré nemajú korektne zadané topologické charakteristiky. (K bežným chybám patrí napr. vzájomne vymenené poradie ľavého a pravého objektu.) Okrem toho, program apriori považuje všetky uzlové body, ktoré nie sú trojnými, za podozrivé. Zvlášť to platí o uzloch prvého a druhého stupňa. V prvom prípade ide o neukončený ("slepý") koniec hrany, v druhom buď chýba tretia incidujúca hrana, alebo to značí nadbytočnosť konštrukcie tohto uzlového bodu, a teda príslušné dve incidujúce hrany môžu byť zlúčené do jednej.

Konkrétny spôsob korektúry invalidného digitálneho dátového súboru vo všeobecnosti nie je možné vykonať na úrovni počítača (ak vylúčime riešenia na úrovni expertných systémov) a vo väčšej miere je v kompetencii používateľa.

Výstupom z programu APTAB okrem výstupu na monitor, ktorý nie je uchovávaný, sú dva súbory, a to: základný digitálny súbor v ASCII kóde a grafický metasúbor (metafile) slúžiaci predovšetkým na kontrolný výstup na tlačiarňu alebo ploter.

Jedným z výstupov programu APTAB je topologicky orientovaný súbor hraničných segmentov (v ASCII formáte), digitalizovaných v ľubovoľnom poradí a v ľubovoľnom smere. Tento súbor predstavuje vstup do programového systému DIGEDIT (Fortran 77 pod MS-DOS). Programový systém DIGEDIT obsahuje procedúry podriadené šiestim hlavným funkciám:

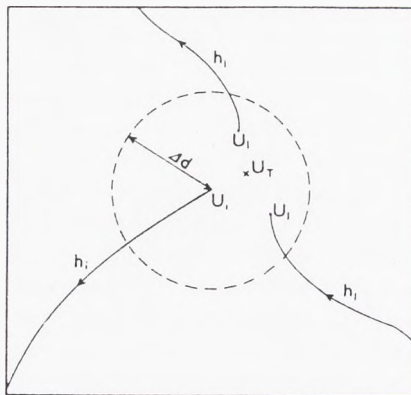
1. čiastočná kontrola a identifikácia metrických a topologických chýb;
2. spájanie hrán v uzlových bodoch a generovanie polygónov ako základný štandardný predpoklad vektorovej priestorovej analýzy;
3. vytvorenie grafických súborov (metafiles) a ich prezentácia na monitore, ihlicovej tlačiarňu alebo ploteri;
4. výpočet morfometrických parametrov regiónu;
5. výpočet závislosti kvalitatívnych javov;
6. export súborov do iných štandardných formátov.

Ďalej sa z hľadiska cieľa práce dotkneme princípu polygonalizácie (2.) a niektorých základných výpočtov morfometrických parametrov regiónu (4.) v prípade objektov reprezentovaných areálovými typmi údajov<sup>3</sup>.

Kardinálnou úlohou vektorového spracovania dát je *polygonalizácia*, t.j. proces spájania hraničných segmentov do siete polygónov. Tento proces spočíva v tom, že k testovanému koncovému bodu hraničného segmentu sú vypočítané vzdialenosti koncových bodov zvyšných hraničných segmentov. Z nich sú vybrané najviac 3 hrany s najmenšími vzdialenosťami od testovanej hrany a zároveň spĺňajúce podmienku, že sú menšie ako definované okolie  $\Delta d$ . Štandardnou hodnotou pre okolie  $\Delta d$  v programe DIGEDIT je 1.5 mm (obr.2).

V prípade existencie dvoch hraničných segmentov spĺňajúcich uvedeníu podmienku ide o klasický prípad trojného bodu poľa  $u_T$ , ktorý je konštruovaný spriemerovaním súradníc koncových bodov príslušných hrán. Následne, koncové súradnice hrán sú nahradené súradnicami vypočítaného trojného bodu poľa a každej z troch hrán sú priradené smerníky (adresy) zvyšných dvoch. Tým vznikne predpoklad na automatické generovanie polygónov.

<sup>3</sup> V súvislosti s výpočtom polygonalizácie a morfometrických parametrov regiónu, naše úvahy sa pohybuju v rámci euklidovského metrického priestoru  $\langle \Omega, d \rangle$ , kde  $\Omega$  je rovina a  $d$  je tzv. euklidovská metrika, pre ktorú platí  $d(x,y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}$  (Kučina, 1979).



Obr. 2. Príklad konštrukcie trojného bodu poľa.

V prípade, že v okolí koncového bodu testovanej hrany sa nachádzajú 3 hrany, je potrebné zistiť, či ide o uzlový bod štvorného bodu poľa, ktorý na rozdiel od trojného je zriedkavý (avšak reálne sa vyskytujúci) alebo či ide o trojný bod poľa, v okolí ktorého sa nachádza koncový bod inej hrany.

#### 4 PRVOTNÉ SPRACOVANIE VEKTOROVÝCH DÁT A ICH MORFOMETRICKÁ ANALÝZA

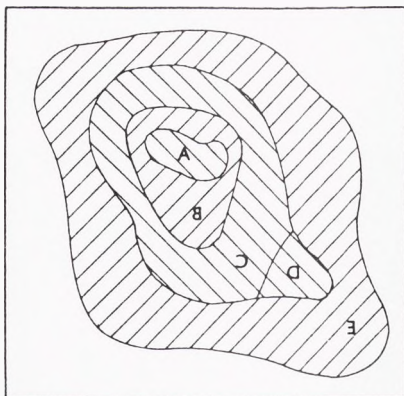
V práci pod označením "*morfometrická analýza areálov (regiónov)*" rozumieme také kvantitatívne metódy, ktoré manipulujú s *priestorovými* údajmi areálov regionálnej typizácie. Tieto metódy delíme do piatich skupín, z ktorých každá má svoje vyjadrenie v DIGEDIT:

- početnosť výskytu,
- plošný obsah a obvod,
- tvar (forma),
- priestorová orientácia regiónu,
- priestorové vzťahy (autokorelácia) medzi regiónmi.

Z hľadiska vymedzeného priestoru práce, z uvedených metód sa dotkneme princípov výpočtu plošného obsahu a obvodu a čiastočne priestorových vzťahov (autokorelácie) medzi regiónmi.

Prv než sa dotkneme niektorých *metód morfometrickej analýzy*, uvedme príklad, ktorý naznačí istý typ charakteristických problémov riešených na úrovni vektorových priestorových analýz.

Vychádzajúc z obr. 3, polygón E obsahuje ostrov, ktorý je tvorený polygónmi C a D, z nich polygón C obsahuje ostrov B a ten ostrov A. Vyriešenie prípadu vnorených ostrovov je dôležité z hľadiska následného výpočtu plošného obsahu polygónov, ako i šrafovania. S týmto problémom, ktorý, ako sme už uviedli, nie je riešiteľný na úrovni jednoduchých polygónových štruktúr, sa možno stretnúť najmä v prípade typologickej regionalizácie (príklad foriem krajinného krytu (Feranec et al. v tlači) atď), v ktorej výskyty ostrovov



Obr. 3. Příklad vnorených ostrovov.

nie sú zvláštnosťou. Ak napríklad polygón E je rovnakého typu ako polygón B, potom plochu polygónov E a B vypočítame zo vzťahu:

$$P(E+B) = P(E) - P(D) - P(C) + P(E) - P(A).$$

Z uvedeného je zrejmé, že počítač musí pri výpočte plochy polygónu (a rovnako i pri šrafovaní a iných operáciách) poznať všetky ostrovy daného polygónu ako aj hĺbku vnorenia prípadných ďalších polygónov - ostrovov. Keďže počítač nedisponuje zrakom, nie je preňho, na rozdiel od človeka, celkom triviálne zistiť, či testovaný polygón obsahuje alebo neobsahuje ostrov. Toto zistenie možno uskutočniť postupnosťou dvoch krokov:

1. testovanie, či polygón sa nachádza v okne testovaného polygónu (pomocou súradníc  $x_{\min}$ ,  $y_{\min}$ ,  $x_{\max}$  a  $y_{\max}$ ). Polygón spĺňajúci túto podmienku ešte nemusí byť ostrovom v testovanom polygóne. Preto je potrebný aj nasledujúci krok;

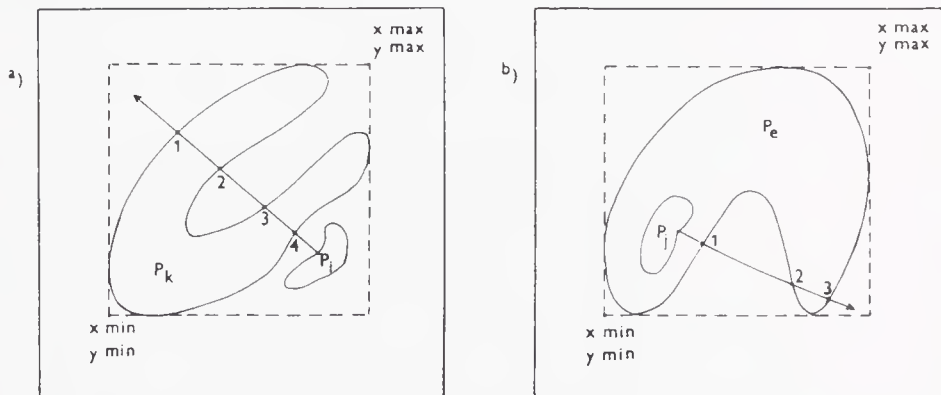
2. zistenie počtu priesečníkov orientovanej polpriamky ľubovoľného bodu polygónu a testovaného polygónu. V prípade že počet priesečníkov je párný, ide o polygón nachádzajúci sa v okne testovaného polygónu, avšak nejde o ostrov. V prípade, že počet priesečníkov je nepárny, ide o polygón nachádzajúci sa vnútri testovaného polygónu, teda ide o ostrov (obr. 4).

Naznačili sme základné logické kroky a rámcový algoritmus, ktorý predchádza vlastnému výpočtu plošného obsahu a niektorým iným metódam priestorovej analýzy nad vektorovým typom dát.

Z formálneho hľadiska pod plošným obsahom regiónu v dvojdimenzionálnom priestore rozumieme nezáporné reálne číslo, o ktorom predpokladáme, že má nasledujúce vlastnosti: ak región  $\alpha$  je podmnožinou  $\beta$  regiónu<sup>4</sup>, potom platí:

$$P(\alpha) \leq P(\beta),$$

kde  $P(\alpha)$  je plošný obsah regiónu  $\alpha$  (Horák a Niepel, 1982).



Obr. 4. Riešenie problému ostrova:

a) polygón  $P_i$  nie je ostrovom polygónu  $P_k$ ,

b) polygón  $P_j$  je ostrovom polygónu  $P_e$ .

V prípade, že  $\alpha_1, \alpha_2, \dots, \alpha_n$  sú rovinné regionálne typy tej istej kvality, potom pre plošný obsah ich zjednotenia platí:

$$P(\alpha_1, \alpha_2, \dots, \alpha_n) \leq P(\alpha_1) + P(\alpha_2) + \dots + P(\alpha_n) \quad (4.1)$$

Vzhľadom na to, že  $\alpha_1, \alpha_2, \dots, \alpha_n$  sú produktom regionálnej typizácie a sú navzájom disjunktné (podmienka i), resp. iii) v § 2), potom nerovnosť v (4.1) možno zmeniť na rovnosť.

Samotný výpočet plochy polygónu vypočítame z analytického vzťahu:

$$P = \sum_{i=1}^{n-1} \frac{(x_{(i+1)} - x_{(i)}) * (y_{(i+1)} + y_{(i)})}{2} \quad (4.2)$$

kde  $P$  - je plošný obsah polygónu,

$n$  - je celkový počet hraničných bodov, ktoré tvoria obvod "pravotočivo" orientovaného polygónu,

$x_{(i)}, y_{(i)}$  - sú súradnice priebežného hraničného bodu  $i$ .

Vzťah (4.2) platí iba pre tzv. "pravotočivo" orientované polygóny. Ďalej platí:

$$P(\alpha_{(L)}) < 0 \text{ a } P(\alpha_{(P)}) = -P(\alpha_{(L)}) = |P(\alpha_{(L)})|, \quad (4.3.)$$

kde  $P(\alpha_{(P)})$  je plošný obsah "pravotočivo" orientovaného polygónu  $\alpha$ ,

$P(\alpha_{(L)})$  je plošný obsah toho istého polygónu orientovaného "ľavotočivo".



V prípade výpočtu plošného obsahu polygónu, ktorý je tvorený hraničnými segmentmi navzájom rôzne orientovanými, je vhodné vyhnúť sa predpisovaniu "pravo-" alebo "ľavo-točivej" orientácie hraničných segmentov. Potom v zmysle (4.3.) pre plošný obsah platí

$$P = \sum_{i=1}^{n-1} \left| \frac{(x_{(i+1)} - x_{(i)}) * (y_{(i+1)} + y_{(i)})}{2} \right| \quad (4.4.)$$

S výpočtom plošného obsahu sa spravidla stretávame pri regionálnej typizácii, v ktorých nejde o výpočet plochy individuálneho polygónu, ale o výpočet celkovej plochy daného typu regiónu. Výpočet takejto plochy je potom v zmysle (4.1.) súčtom jednotlivých výskytov daného typu regiónu a platí

$$P = \sum_m \sum_{i=1}^{n-1} \left| \frac{(x_{(i+1)} - x_{(i)}) * (y_{(i+1)} + y_{(i)})}{2} \right| \quad (4.5.)$$

kde m - je počet výskytov daného typu regiónu.

Výstupom programového bloku výpočtu plošného obsahu areálov toho istého typu je tabuľka obsahujúca hodnoty:

- celkového plošného obsahu,
- najmenšieho plošného obsahu,
- najväčšieho plošného obsahu,
- mediánu.

Druhým dôležitým morfometrickým ukazovateľom je *obvod* (dĺžka hraníc) polygónov. V prípade výpočtu obvodu polygónu, tento je v konkrétnej programovej aplikácii podriadený štruktúre uloženia vektorového dátového súboru. Vo všeobecnosti, obvod regiónu je daný súčtom na troch úrovniach:

- výpočet dĺžky hraničných segmentov, ktorá je daná sumou euklidovských vzdialeností medzi jednotlivými hraničnými bodmi príslušného hraničného segmentu,
- sumou dĺžok tých hraničných segmentov, ktoré tvoria obvod regiónu,
- sumou dĺžok obvodov regiónov tvoriacich tú istú triedu.

Vzhľadom na uvedené, obvod polygónu je určený vzťahom:

$$O = \sum_k ( \sum_l ( \sum_m ( \sqrt{((x_{(i+1)} - x_{(i)})^2 + (y_{(i+1)} - y_{(i)})^2})) ) ) ) \quad (4.6)$$

kde  $x_{(i)}$ ,  $y_{(i)}$ ,  $x_{(i+1)}$ ,  $y_{(i+1)}$  - sú súradnice dvoch za sebou nasledujúcich bodov hraničného segmentu,

- k - je počet výskytov regiónu daného typu,
- l - je počet hraničných segmentov príslušného regiónu,
- m - je počet bodov príslušného hraničného segmentu i.

Výstupom programového bloku na výpočet obvodu polygónu je:

- celková dĺžka hraníc jednotlivých typov polygónov;
- podiel dĺžky hraníc na jednotku plošného obsahu;
- symetrická matica dĺžok MD<sub>A</sub>, v ktorej jednotlivé riadky (stĺpce) reprezentujú dĺžku spoločných hraníc istého typu polygónu so všetkými ostatnými typmi polygónov. Diago-

nála takejto matice je nulová a ostatné nulové prvky matice poukazujú na absenciu spoločnej hranice. Suma dĺžok v riadku (stĺpci) matice je celkovou dĺžkou hraníc (obvodom) polygónu daného typu;

d. podobná symetrická matica  $MD_R$ , ako bola uvedená v bode c, avšak prvky tejto matice sú tvorené percentuálnym podielom hraníc istého typu polygónu so všetkými ostatnými typmi polygónov. Celková suma hodnôt v riadku (stĺpci) sa rovná hodnote 100.

Matice  $MD_A$  a  $MD_R$  sú mocným zdrojom informačného potenciálu pri skúmaní priestorových vzťahov (autokorelácie) v rámci rôznych vektorových a areálových typoch dát ako: *vzdialenosť medzi regiónmi, susednosť, asociativnosť, adaptabilita* (Husár, manuscript) a inými charakteristikami, ktoré na základe geometrických atribútov referujú minimálne o dvoch priestorových areálových jednotkách.

S ohľadom na vymedzený rozsah našim cieľom nebolo poukázať na všetky možné typy priestorových analýz, ktoré umožňuje DIGEDIT.

Výstupy z programu DIGEDIT, ktorý okrem už opísaných procedúr obsahuje aj tvorbu okien (windiwing), tvorbu rôznych grafických súborov (napr. šrafovanie areálov), výpočet závislosti medzi kvalitatívnymi typmi dát, transformácie formátu APTAB do iných štandardných formátov a iné, poskytujú široký priestor na geografickú interpretáciu a zároveň môžu byť zdrojom dát na využitie pri výpočte ďalších typov priestorových analýz v iných výpočtových systémoch.

## ZÁVER

Cieľom práce bolo zhodnotiť stav digitalizácie vektorových polygónových dát na Geografickom ústave SAV v kontexte riešenia tejto problematiky v bývalej ČSFR a poukázať na základné princípy, techniky a metódy z oblasti vektorovej topologickej digitalizácie a niektorých vybraných priestorových analýz nad vektorovým typom dát, ktoré boli aplikované v programovom systéme APTAB a DIGEDIT.

Práca teda predstavuje istý metodický rámec na digitálne spracovanie ľubovoľnej predlohy analógovej formy, t.j. základnej mapy, topografickej mapy alebo interpretačnej schémy, vytvorenej pomocou techník diaľkového prieskumu Zeme.

## LITERATÚRA

- BURROUGH, P.A. (1986). *Principles of geographical information systems for land resources assessment*. Oxford (Clarendon).
- FERANEC, J., OŤAHEL, J. (1992). Land cover forms in Slovakia identified by application of colour infrared space photographs at the scale 1:500 000. *Geografický časopis*, 44, 120-126.
- FERANEC, J., OŤAHEL, J., PRAVDA, J., HUSÁR, K. (v tlači). Formy krajinného krytu identifikované v rámci projektu CORINE Land Cover. *Geografický časopis*.
- GOODCHILD, M.F. (1989). Toward enumeration and classification of GIS functions. In *The GIS Sourcebook*. Colorado (GIS World), pp. 22-26.
- HAGGETT, P. (1965). *Locational analysis in human geography*. London (Edward Arnold Ltd).
- HORÁK, P., NIEPEL, L. (1982). *Prehľad matematiky*. Bratislava (Alfa).
- HUSÁR, K. (1992). Základná taxonómia a porovnanie priestorových dátových modelov z hľadiska geografických informačných systémov. *Geografický časopis*, 44, 306-318.
- HUSÁR, K. (manuscript). *Vektorová digitalizácia a morfometrická analýza areálov na príklade foriem land cover juhozápadného Slovenska*. Geografický ústav SAV, Bratislava.

- CHRISMAN, N.R. (1987). Efficient digitizing through the combination of appropriate hardware and software for error detection and editing. *International Journal of Geographical Information System*, 1, 265-277.
- INTERA TYDAC (1991). *SPANS User Guide, Version 5.2*. Ontario, Canada (Intera Tydac technologies Inc.)
- KUŘINA, F. (1979). *Metrika a topologie*. Hradec Králové (Pedagogická fakulta).
- MAZÚR, E., LUKNIŠ, M. (1980). Geomorfologické jednotky, 1:500 000. In *Atlas SSR*. Bratislava (Slovenská akadémia vied a Slovenský úrad geodézie a kartografie).
- MORAVEC, D. (1986). *Modelování automatizované tvorby topografických map*. Dizertační práce, Výzkumné středisko 090, Praha.
- PEQUET, D.J. (1984). A conceptual framework and comparison of spatial data model. *Cartographica*, 21, 66-113.
- SMITH, T.R. et al. (1987). Requirements and principles of large-scale geographic information system. *International Journal of Geographical Information System*, 1, 13-31.

Karol Husár

## VECTOR DIGITALIZATION OF AREAL OBJECTS AND THEIR SPATIAL ANALYSIS IN THE CONDITIONS OF THE INSTITUTE OF GEOGRAPHY SAS

The work is, besides the introductory part to the subject, divided into 3 relatively independent units, the first of which is treating the state and evaluation of the data digitalization in the former CSFR.

Assuming that the work will be subject to the study of broader geographical community not familiar with the theme, the second part is dedicated to the basic principles of the area of vector topological digitalization of polygons.

The third part of the work deals with digitizing of vector polygon data of area type, by the APTAB programme system stressing the possibilities offered by this software to the user. We also touch briefly the principles of polygonalization.

The fourth part contains some methods of morphometric analysis of regions and provides a survey of these methods that form part of the DIGEDIT programme system. More space is dedicated to the methods used in the calculation of areas, polygons and lengths of the boundary lines of polygons. Besides, this part of work presents matrices of the lengths of the boundary segments  $MD_A$  and  $MD_R$  that represent an important information potential and at the same time a valuable source of input data for the calculation of further types of spatial analyses in the environment of DIGEDIT or other professional systems.

Fig. 1. Basic topological elements and their analog relations.

Fig. 2. Example of construction of triple point of field.

Fig. 3. Example of immersed islands.

Fig. 4. Solution of the problem of island:

- a) polygon  $P_i$  is not island of polygon  $P_k$
- b) polygon  $P_j$  is island of polygon  $P_e$ .

Translated by H. Contrerasová