

Extrakce digitálních prostorových dat z historických map metodami segmentace obrazu

Karel JELÍNEK
Geografický ústav
Přírodovědecká fakulta, Masarykova univerzita
Kotlářská 2
611 37 Brno
E mail:
k.a.jelinek@centrum.cz

Abstract

The possibility of automatic extraction of digital spatial data using methods of image segmentation is tested in chosen historical maps in this thesis. Acceptable parameters of segmentation and classification were projected with using eCognition Professional 4.0 software. On the basis of classification, vector maps of land use were designed and consequently compared with the maps of the same area created by analog interpretation. From results it follows that automatic classification in eCognition is closely associated with used cartographic symbology on individual maps.

Abstrakt

Práce testuje možnosti automatické extrakce digitálních prostorových dat metodou segmentace obrazu na vybraných historických mapách. Prostřednictvím programu eCognition Professional 4.0 byly navrženy vhodné parametry segmentace a klasifikace. Na základě klasifikace byly sestaveny vektorové mapy land use, které byly porovnány s mapami téhož území získanými analogovou interpretací. Z výsledků vyplývá, že přesnost automatické klasifikace v programu eCognition je úzce spjata s použitými kartografickými vyjadřovacími prostředky zpracovávaných map.

Úvod

Kromě své kartografické a umělecké hodnoty poskytují mapy vojenských mapování z 18. a 19. století unikátní cenné informace o minulosti krajiny zahrnující i dnešní Českou republiku. Jejich studiem lze vysvětlit zákonitosti vývoje krajiny i jejího osídlení i v dnešní době. Jsou důležitým zdrojem dat pro řadu odborníků nejen z oblasti geografických věd, ale i z oblastí krajinné ekologie, historie, archeologie, botaniky, geodézie, kartografie a mnohých dalších [5].

Lepší možnosti využití prostorových dat z historických map, jejich uchovávání i dostupnost poskytuje digitální forma těchto dat. [6] Pro některé úlohy řešené v prostředí GIS je výhodné (nebo i zcela nutné) tato data aplikovat ve formě vektoru. K získání těchto dat se nabízí použití nástrojů, postupů a technik geoinformatiky a zpracování obrazu [1].

Cílem této práce bylo na vybraných historických mapách testovat možnosti automatické extrakce dat pomocí segmentace obrazu. Na základě následně provedené klasifikace pak bylo cílem vytvořit vektorovou mapu základních druhů povrchů. Ta byla posléze porovnána s mapou vzniklou analogovou interpretací téhož území.

Zdrojová data

Vstupní data pro tuto práci poskytlo brněnské středisko AOPK ČR v digitální podobě ve formátu TIFF. Jedná se o naskenované listy map čtyř mapování, na nichž je zobrazeno zájmové území katastr Kohoutovice (část města Brna). Tato data jsou společně charakterizovaná rozlišením 400 dpi a bitovou hloubkou 8, případně 24 bit. Konkrétně byla použita následující data:

- I. vojenské mapování (josefské): 1764 – 1768 a 1780 – 1783 (rektifikace), měřítko 1: 28 800, část Morava, list 77, velikost 12441 x 9756 pixelů,
- II. vojenské mapování (Františkovo): 1836 – 1852, měřítko 1: 28 800, část Morava, list O-10-I, velikost 11811 x 9752 pixelů,
- III. vojenské mapování (Františko-josefské): 1876 – 1878 (Morava a Slezsko), měřítko 1: 25 000, list 4357-1, velikost 12528 x 9880 pixelů,

- vojenská topografická mapa : 1990, měřítko 1:25 000, list M-33-106-A-c, velikost 6985 x 7401 pixelů.

Postup zpracování

Vlastní zpracování se skládalo z dílčích úloh týkajících se jednak úpravy vstupních dat a postupů specifických při klasifikaci „per-object“ vedoucích až k tvorbě vektorové mapy základních druhů povrchů. Jedná se zejména o následující úlohy:

- **pořízení výřezu zájmového území,**
- **proces segmentace obrazu,**
- **proces klasifikace,**
- **export výsledku klasifikace do vektorového formátu.**

Pořízení výřezu zájmového území

Použití výřezu zájmového území z listu skenované mapy do značné míry zjednodušuje a zrychluje zpracování ve všech následujících krocích, zejména z důvodu vysoké paměťové náročnosti obrazových dat celého listu.



Obr. 1. Výřez zájmového území (katastr Kohoutovice) z mapy II. vojenského mapování

Proces segmentace obrazu

V procesu segmentace [3] je obraz rozčleněn dle zadaných kritérií do jednotlivých nepřekrývajících se segmentů (objektů). Těmito kritérii jsou jednak obdobné spektrální vlastnosti (v našem případě jsou DN hodnoty dané použitou barvou při zhotovení mapy a následnou digitalizací) a požadavky na tvar objektu. Při procesu segmentace je přiřazeno každému objektu množství atributových údajů, jež se týkají zejména spektrálních vlastností, tvarových vlastností, textury a kontextu s ostatními prvky. O tom, jak bude vypadat výsledek segmentace, tedy kolik, jak velkých a jak tvarovaných segmentů vznikne, rozhoduje míra heterogenity (respektive homogenity), která do procesu vstupuje na začátku segmentace při zadání parametrů:

SCALE parameter - ovlivňuje velikost výsledných segmentů,

SHAPE factor - udává váhu tvarové heterogenity oproti spektrální,

COMPACTNESS - udává váhu kritéria kompaktnosti oproti hladkosti.

Hledání vhodných parametrů je do značné míry otázkou zkušenosti. Je třeba testovat nastavení těchto parametrů, dokud není dosaženo vhodného výsledku procesu. Za ten je možné považovat situaci, kdy vzniklé segmenty respektují hranice mezi třídami, jež je třeba určit klasifikací. Tímto způsobem byly nalezeny parametry pro zpracovávané mapy. Pro příklad mapy II. vojenského mapování jsou popsány v tabulce 1.

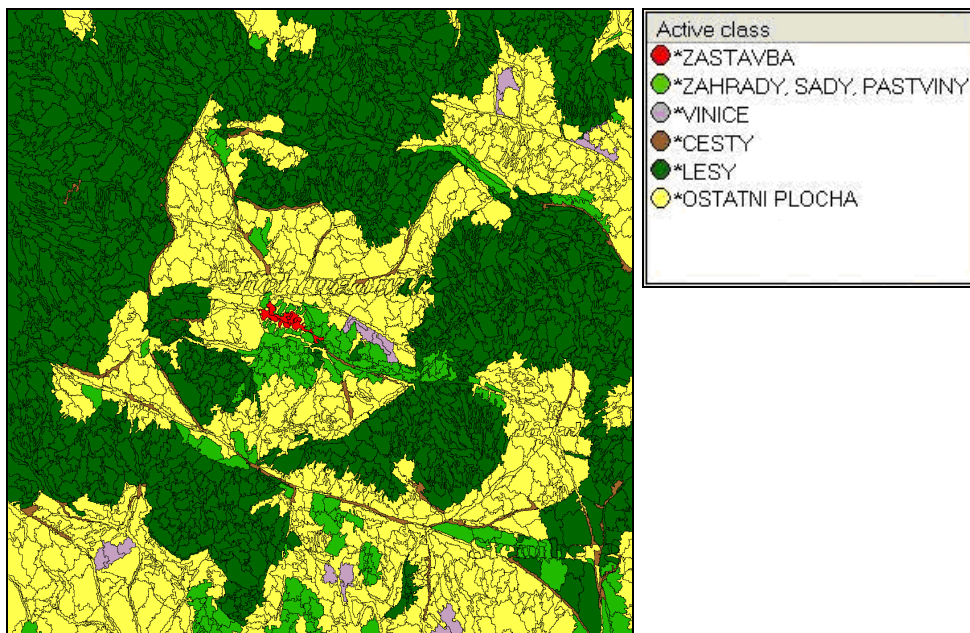
Tab. 1. Vhodné nastavení parametrů segmentace mapy II. vojenského mapování

Úroveň	Parametr		
	Scale	Shape	Compactness
			Smoothness
1	10	0,05	0,7
			0,3
2	40	0,3	0,7
			0,3
3	150	0,3	0,7
			0,3

Jak vyplývá z tabulky 1, bylo použito více úrovní segmentace. To umožnilo při pozdější klasifikaci uvažovat více atributů (zejména týkajících se topologie) a tím dosáhnout přesnějších výsledků.

Proces klasifikace

Klasifikace „per-object“ probíhá na úrovni segmentů vzniklých v předchozím procesu [2]. V práci bylo použito klasifikace založené na *principu neostrých množin* (fuzzy logic). Příslušnost jednotlivých segmentů ke každé třídě je v této klasifikaci hodnocena prostřednictvím funkce příslušnosti (membership function) pro každý z uvažovaný *atributů*. Nejprve však bylo vždy třeba stanovit klasifikační schéma. V případě mapy II. vojenského mapování je klasifikační schéma patrné z obr. 2., kde je zachycen i výsledek klasifikace, na němž jsou jednotlivé segmenty obrazu znázorněny černým obrysem.



Obr. 2. Výsledek

klasifikace mapy II. vojenského mapování

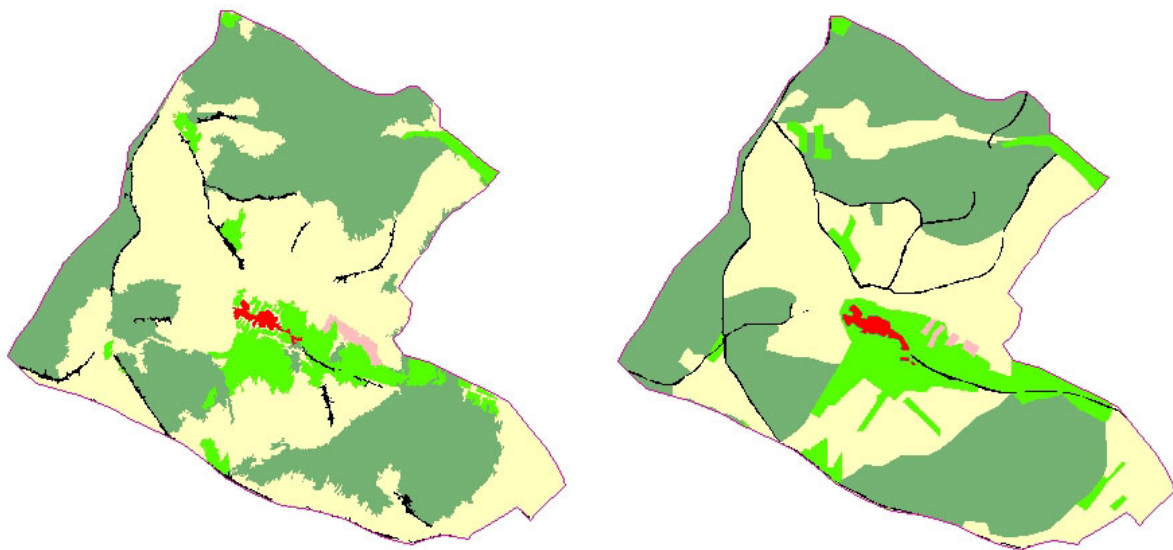
Export výsledku do vektorového formátu

V této práci byl proveden konkrétně export všech obrazových objektů (segmentů) do jednoho souboru SHP, a příslušnost k daným třídám byla automaticky zaznamenána v atributové tabulce.

Porovnání výsledku klasifikace s analogovou interpretací

Pro korektní zjištění přesnosti automatické klasifikace se využilo porovnání mapy vzniklé pomocí výše zmiňovaných postupů s mapou téhož území vzniklou analogovou interpretací v programu ARC VIEW 3.2. Takto vzniklá mapa je tvořena stejnými třídami z klasifikačního schématu (má stejnou legendu) a je považována za referenční. Porovnáním lze přesně určit kolik pixelů z referenčních dat (čili pořízených analogovou interpretací) je zařazeno v té které třídě dat získaných automatickou klasifikací v programu eCognition a naopak. Pro zhodnocení takto získané informace byla sestavena chybová matice [4] a spočteny ukazatele přesnosti klasifikace, jež jsou patrné z tabulky 2.

- Hranice katastru
- Zastaba
- Cesty
- Vinice
- Zahrady sady, pastviny
- Ostatní plocha
- Lesy



Obr. 3. Vizualizace vektorů získaných automatickou klasifikací v programu eCognition (vlevo) a analogovou interpretací (vpravo)

Tab. 2. Chybová matice a ukazatele přesnosti klasifikace pro II. vojenské mapování

II. VOJENSKÉ MAPOVÁNÍ									
ANALOGOVÁ INTERPRETACE									
eCognition	třída	Zástavba	Zahrady	Vinice	Cesty	Lesy	Pole	SUMA	PU [%]
	Zástavba	2829	96	0	0	0	0	2925	96,72
	Zahrady	520	46788	458	496	611	3641	52514	89,10
	Vinice	0	896	1304	0	0	1308	3507	37,17
	Cesty	0	805	0	5421	2347	3655	12228	44,33
	Lesy	199	9007	0	2960	264282	18456	294904	89,62
	Pole	846	16583	589	4644	47034	254431	324127	78,50
	SUMA	4394	74175	2350	13521	314273	281491	690204	
CHO [%]	35,61	36,92	44,54	59,91	15,91	10,64			
CHZ [%]	2,18	7,72	93,76	50,34	9,74	24,76			
PZ [%]	64,39	63,08	55,46	40,09	84,09	90,39			
PP [%]	83,32				κ	0,7244			

PP.....průměrná přesnost
 PU.....přesnost z hlediska uživatele
 PZ.....přesnost z hlediska zpracovatele
 CHO...chyba z opomenutí
 CHZ...chyba z nesprávného zařazení
 κ.....Kappa koeficient

Závěr

Pro zájmové území byly pořízeny výřezy z vojenské topografické mapy (1990) a ze starých map I. II. a III. vojenského mapování. Tyto mapy se od sebe podstatně lišily kartografickými vyjadřovacími prostředky. V programu eCognition byly testovány a posléze nalezeny vhodné parametry procesu segmentace (Scale parametr, Shape Factor a Compactness) pro každý typ mapy. Tak vzniklo množství vzájemně se nepřekrývajících segmentů, na jejichž úrovni se odehrává klasifikace „per-object.“ Pomocí klasifikace založené na principu neostrých množin byl všem vzniklým segmentům přiřazen význam z předem stanoveného klasifikačního schématu. Takováto klasifikace byla exportována do vektorového formátu SHP. Za účelem zhodnocení přesnosti této klasifikace byly také vytvořeny mapy analogovou interpretací. Klasifikací bylo tedy docíleno průměrné přesnosti 87,04 % u vojenské topografické mapy, 79,54 % u mapy I. vojenského mapování, 83,32 % u mapy II. vojenského mapování a 75,78 % u mapy III. vojenského mapování. Přičemž rozdíly v přesnosti mezi jednotlivými typy map jsou dány zejména použitými kartografickými vyjadřovacími prostředky jednotlivých map a geografických objektů a jevů na nich znázorněných.

Tento příspěvek je shrnutím bakalářské práce Karla Jelínka "Extrakce digitálních prostorových dat z historických map metodami segmentace obrazu." aktuálně zpracovávané na Přírodovědecké fakultě Masarykovy univerzity.

Literatura

[1] BRŮNA, V. – UHLÍŘOVÁ, L.: *Metodika přístupu k interpretaci historických map se zvláštním zřetelem k udržení a obnově ekologické stability krajiny* [online]. 2000 [cit. 2006-03-02]. Dostupný na WWW: <<http://projekty.geolab.cz/files/Metodika.pdf>>.

[2] DOBROVOLNÝ, P.: *Dálkový průzkum Země Digitální zpracování obrazu*. 1. vyd. Brno : Masarykova univerzita v Brně, 1998. 210 s. ISBN 80-210-1812-7.

[3] eCognition downloads [online]. Definiens imaging, Mnichov, 2006. [cit. 2006-02-02] Dostupný na WWW: <<http://www.definiens-imaging.com/down/>>.

[4] LILLESAND, T. M. – KIEFFER, R. W.: *Remote Sensing and Image Interpretation*. 4. vyd., New York: John Willey & Sons, 1999, 724 s. ISBN 0-471-25515-7.

[5] LIPSKÝ, Z.: *Sledování změn v kulturní krajině: učební text z předmětu Krajinná ekologie*. Kostelec nad Černými lesy: Česká zemědělská univerzita Praha v nakladatelství Lesnická práce, s. r. o, 2000. 71 s. ISBN 80-213-0643-2.

[6] *Prezentace starých mapových děl z území Čech, Moravy a Slezska* [online]. Laboratoř geoinformatiky UJEP. Ústí nad Labem, 2005. [cit 2005-11-11]. Dostupný na WWW: <<http://oldmaps.geolab.cz/>>.