

Metodika vyhledávání mapových značek na digitalizovaných starých mapách

Jan Kotera, Milan Talich

Realizováno z programového projektu DF11P01OVV021: Program aplikovaného výzkumu a vývoje národní a kulturní identity financovaného MK ČR v rámci projektu

„Kartografické zdroje jako kulturní dědictví. Výzkum nových metodik a technologií digitalizace, zpřístupnění a využití starých map, plánů, atlasů a glóbulů.“

Výzkumný ústav geodetický, topografický a kartografický, v.v.i.

říjen 2013

Obsah

1. Předmět metodiky	3
2. Struktura metodiky	4
3. Úprava kontrastu mapy	5
4. První fáze prohledání mapy	7
5. Druhá fáze prohledání.....	10
6. Třetí fáze prohledání	12
7. Závěr	15
8. Seznam použitých zdrojů	16

1. Předmět metodiky

Cílem metodiky je popsat vytvořený algoritmus pro automatické vyhledávání bodových mapových značek na digitalizovaných starých mapách, který je výsledkem aplikovaného výzkumu uskutečněného v rámci projektu DF11P01OVV021 „Kartografické zdroje jako kulturní dědictví. Výzkum nových metodik a technologií digitalizace, zpřístupnění a využití starých map, plánů, atlasů a glóbů.“. Tento algoritmus tvoří základ systému pro uvedený účel.

System umožní snazší práci s mapu, což přispěje k jejímu porozumění a užitku. Na základě vzorových exemplářů značky a nastavení vyhledávacích parametrů pro konkrétní druh mapy software vyhledá značky na mapě a zaznamená jejich polohu do databáze. Předpokládá se použití pro jednorázové dávkové zpracování jednotlivých listů mapy a uchování výsledků v databázi, aby pak mohly být kdykoliv vyvolány bez nutnosti opakovaného vyhledávání. System je primárně vyvíjen pro mapy III. vojenského mapování. Hlavními nároky na vyvíjený system jsou rozumná rychlost, efektivita a nízké nároky na objem vstupních dat, tedy zejména počet vzorových značek. Největšími problémy, kterým bylo nutné vývoj a postup přizpůsobit, jsou velká diverzita nakreslení jednotlivých exemplářů značky na mapách a častý překryv značky a zbytku kresby na mapě.

2. Struktura metodiky

System vyhledává vždy jen jediný druh mapové značky, pro vyhledání několika druhů značek na jednom mapovém listu je třeba provést vyhledání opakovaně s různými vstupními daty, která odpovídají jednotlivým značkám. Metodika vyhledávání tak bude popsána na příkladu vyhledávání konkrétní mapové značky.

Prohledávání mapy probíhá v několika fázích:

1. Úprava kontrastu mapy, převod z barevného na šedotónový obrázek. V případě vyhledávání barevných značek na barevné mapě je převod automaticky udělán tak, aby zvýraznil dominantní barvu značku. V případě map III. vojenského mapování, na které je systém primárně vyvíjen, jsou ale značky pouze tmavé na světlém pozadí.
2. Prohledání celého mapového listu a porovnání s jedním (hlavním) exemplářem hledané značky. Prohledávání probíhá v několika geometrických transformacích hledané značky (úhly natočení zvětšení/zmenšení). Pro posouzení, zda by se na daném místě mapy mohla nacházet značka, slouží normalizovaná korelace zahrnující masku okolí značky (vysvětleno dále). Tato fáze slouží k omezení dalšího pomalejšího hledání jen na malou část mapy.
3. Pouze v místech, která jsou pozitivně identifikována v první fázi, je provedeno další prohledání pomocí korelace s maskou, tentokrát ale pro všechny vstupní exempláře značky a pro jemnější odstupňování geometrických transformací značky. Tato fáze umožní přesnější stanovení polohy značky a umožňuje tak použít tvrdší práh pro pozitivní detekci než v předchozí fázi.
4. V poslední fázi je v každém výřezu, který je podezřelý na výskyt značky, nejprve provedeno odstranění pozadí, které není součástí značky, ale zbytku mapy. Toto pozadí činí identifikaci obtížnější a je tak žádoucí se ho zbavit. Nakonec je provedeno konečné rozhodnutí, zda se jedná o hledanou značku. K tomuto lze použít například Diceův koeficient podobnosti (DSC) vůči vzorovým exemplářům značky nebo porovnání příznakového vektoru histogramu orientace gradientů (HOG) vůči vzorovým exemplářům s prahem na maximální odchylku. Místo HOG je možné použít jiný typ příznaků, je vhodné experimentálně určit, který funguje nejlépe.

Uživatelský vstup má 2 hlavní části:

1. Vzorové exempláře hledané značky (vyjmuté z mapy) spolu s jejich „maskami“, tedy označením pixelů, které tvoří značku a nikoliv pozadí. První z těchto vzorových značek je považována za „hlavní“ a je použita v prvním prohledání pomocí korelace.
2. Parametry hledání, což jsou hlavně rozmezí geometrických transformací, které jsou prohledávány, a prahové hodnoty (na korelaci, DSC, popř. příznakové blízkosti) pro rozhodnutí zda daný výřez postupuje do další fáze vyhledávání nebo bude v konečné fázi označen za nalezenou značku.

V dalším textu budou tyto části podrobně popsány.

3. Úprava kontrastu mapy

Na obrazu mapy je nejprve provedeno lokální roztažení histogramu jednotlivých kanálů a převod na šedotónový obrázek. Tato úprava je nutná vzhledem k tomu, že jednotlivé části mapy jsou jinak tmavé, především kvůli nekvalitě originální papírové mapy.

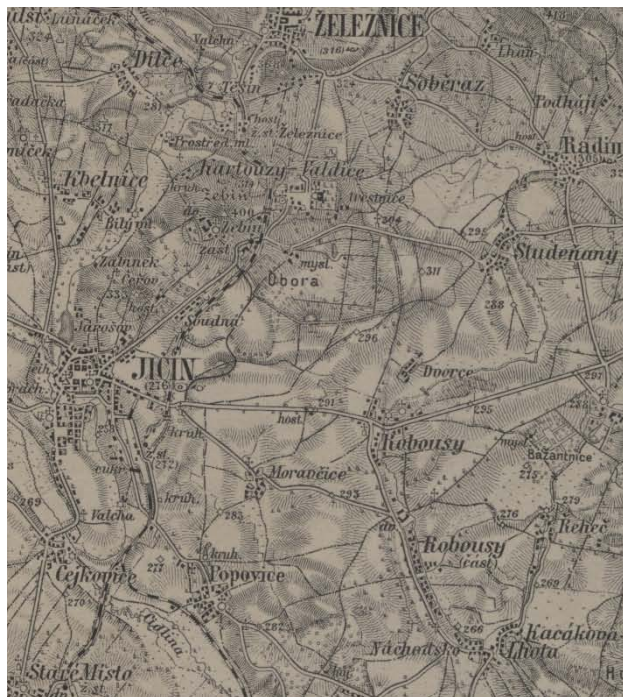
Mapa je rozdělena na nepřekrývající se dlaždice stanovené velikosti (pro testované mapy je např. použito 500x500px) a v každé dlaždici jsou spočítány konstanty a a b pro lineární transformaci intenzit jednotlivých kanálů podle

$$out_i = a(in_i - b),$$

kde in_i resp. out_i je i -tý vstupní resp. výstupní intenzitní level. Výpočet je proveden tak, aby dané procento pixelů (použito je 1%) mělo maximální resp. minimální hodnotu. Hodnoty těchto konstant jsou takto stanoveny pro středové pixely jednotlivých oken, pro zbytek obrázku jsou lineárně interpolovány. Každý barevný kanál je pak (zvlášť) převeden touto transformací a jednotlivé kanály jsou zprůměrovány pro získání šedotónového obrázku.

Pro shrnutí, úprava kontrastu a převod na šedotónový obrázek je provedena následujícím postupem:

1. Mapa je rozdělena do nepřekrývajících se dlaždic dané velikosti.
2. V každé dlaždici a pro každý barevný kanál jsou spočteny parametry takové lineární transformace intenzity, aby dané procento pixelů mělo po transformaci minimální/maximální hodnotu.
3. Parametry z bodu 2 jsou uvažovány pro středové pixely dlaždic z bodu 1, pro ostatní pixely jsou bilineárně interpolovány.
4. Intenzitní hodnoty v každém kanálu jsou převedeny stanovenou transformací.
5. Mapa je převedena na šedotónový obrázek zprůměrováním intenzit barevných kanálů.



Ukázka úpravy kontrastu a transformace barev. Vlevo výřez vstupní mapy, vpravo odpovídající výřez po provedené transformaci.

4. První fáze prohledání mapy

Tato fáze provede prohledání celé vstupní mapy. Vstupní data tvoří jeden vzorový exemplář vstupní značky, uvažované úhly otočení (diskrétní hodnoty, parametr `angles1`) a uvažovaná zvětšení/zmenšení (diskrétní hodnoty, mohou být izotropní nebo v každém směru jiné, parametr `scales1`). Prohledávání mapy pak probíhá tak, že pro každý obdélníkový výřez (okno) z mapy velikosti vzorové značky je proveden výpočet korelace se vzorovou značkou v rámci její masky (viz dále).

Nejprve je stanovena maska, která určuje, jaká část okna bude pro výpočet korelace relevantní. Značka totiž nezabírá celé obdélníkové okno a jeho část je tak zaplněna jiným obsahem mapy, který ztěžuje porovnávání. Současně je však třeba uvažovat pro porovnání alespoň malé těsné okolí značky, jinak by vznikalo mnoho falešných detekcí (značka by například zdánlivě existovala v každé homogenní oblasti atd.). Součástí vstupních dat je i maska značky (bez pozadí), tato maska je mírně zvětšena dilatací, aby obsahovala nejen přímo značku, ale i její těsné okolí. Míru zvětšení udává parametr `neighborhood`.

Je provedena transformace vzorové značky a vypočtené masky pro všechna zvětšení/zmenšení a všechny úhly (tím jsou myšleny všechny kombinace). Pro každou takto vzniklou značku je spočítána normalizovaná korelace s každým možným umístěním značky na mapu. Do výpočtu korelace je zahrnuta maska značky a jejího těsného okolí, takže pixely mimo tuto masku nejsou v korelaci uvažovány. V daném okně vstupního obrázku je tak normalizovaná korelace spočítána podle vzorce

$$cc = \frac{\sum_i (f_i - \bar{f})(t_i - \bar{t}) m_i}{\sum_j (f_j - \bar{f})^2 m_j \sum_j (t_j - \bar{t})^2 m_j},$$

kde f jsou pixely okna obrázku (mapy), t jsou pixely šablony, m jsou pixely masky ($m_i \in \{0, 1\}$), indexy i, j probíhají v prohledávaném okně a hodnoty \bar{f} resp. \bar{t} značí průměr okna obrázku resp. šablony v rámci masky, tedy

$$\bar{f} = \frac{1}{N} \sum_i f_i m_i, \quad \bar{t} = \frac{1}{N} \sum_i t_i m_i, \quad N = \sum_i m_i,$$

kde indexy i probíhají přes okno obrázku resp. šablony, N je počet nenulových prvků v masce. Maximalizace takovéto korelace je ekvivalentní minimalizaci kvadrátu odchylek intenzity okna obrázku od intenzity šablony značky v rámci masky, je-li předem provedena normalizace intenzit okna obrázku na afinní transformaci.

Tato hodnota korelace je spočítána pro každý bod obrázku mapy, což odpovídá všem možným umístěním značky na mapu. V hodnotách korelace jsou nalezena lokální maxima a tyto hodnoty jsou dále oprahovány stanoveným prahem (hodnota `cc011` ve vstupních parametrech). Takto nalezené body na mapě jsou označeny za kandidáty na značku a oblast mapy s malým okolím okolo takového bodu postupuje do další fáze prohledávání. Protože byla korelace počítána pro několik natočení a zvětšení vzorové značky, může být na mapě několikrát označeno „stejně“ místo jako kandidát na značku, proto jsou nejprve jednotlivé nálezy seskupeny, pokud se odpovídající obrazy značek na daném místě mapy překrývají, a jako kandidát postupuje do další zpracování odpovídající část mapy jen jednou.

Zpracovávané mapy jsou ručně kreslené a jednotlivé exempláře mapových značek nejsou zdaleka všechny stejné, od jakéhosi „ideálu“ se odlišují mnohými deformacemi, jako je natočení, změna velikosti, zkosení a jiné. Vzorové značky jsou tak pro účely prohledání mírně otáčeny a zvětšeny/zmenšeny, aby bylo možné dosáhnout lepší shody s deformovaným exemplářem na mapě. Ve druhé fázi prohledání je rozsah uvažovaných deformací ještě větší a jemnější.

Výsledkem první části prohledání jsou tedy malé výřezy na mapě, které je nutné dále prohledat (neboť se nemusí skutečně jednat o značku). Parametry by měly být nastaveny tak, aby existovalo minimum skutečných značek, které první fází neprojdou (takové už nemohou být dále „nalezeny“). Předpokládá se, že budou vybrány i části mapy, které skutečně značku neobsahují. Nemělo by jich jen být příliš mnoho, to by mělo negativní vliv na rychlost druhé fáze.

Pro shrnutí, první fáze prohledání pomocí korelace je provedena následujícím postupem:

1. Stanoví se maska značky + malého okolí (dilatací masky ze vstupních dat). Tato maska je použita i pro korelaci ve druhé fázi prohledání a částečně pro filtraci pozadí ve třetí fázi, maska se tedy stanoví pro všechny vzorové značky, ačkoliv v první fázi korelace je použita pouze první.
2. Pro každou dvojici úhel-zvětšení je šablona značky a maska transformována a proveden výpočet normalizované korelace v rámci masky pro všechna možná umístění šablony na mapu.
3. V hodnotách korelace z bodu 2 jsou nalezena lokální maxima vyšší než práh dle parametrů. Pro každé takové maximum jsou na mapě označeny pixely, které odpovídají umístění značky na mapu v místě, kde bylo dosaženo maxima.
4. Je provedeno seskupení nálezů (označených pixelů) z bodu 3 – pokud se obrys značky (ve smyslu masky ze vzorových dat) překrývá pro dva a více nálezů (což se snadno stane v rámci podobných úhlů a zvětšení), jsou tyto nálezy dány do jedné skupiny a do další fáze postupují jako jediný nález.
5. Kolem každého nálezu (po seskupení z bodu 4) je udělán výřez ze vstupní mapy, v další fázi již není prohledávána celá mapa, ale jen dílčí výřezy okolo podezřelých míst nalezených v první fázi.



Ukázka výsledku první fáze prohledání při vyhledávání značky „kostel“. Je označeno mnoho falešných detekcí, které však budou eliminovány v dalších fázích. Většina plochy mapy je tím z dalšího prohledávání vyloučena.

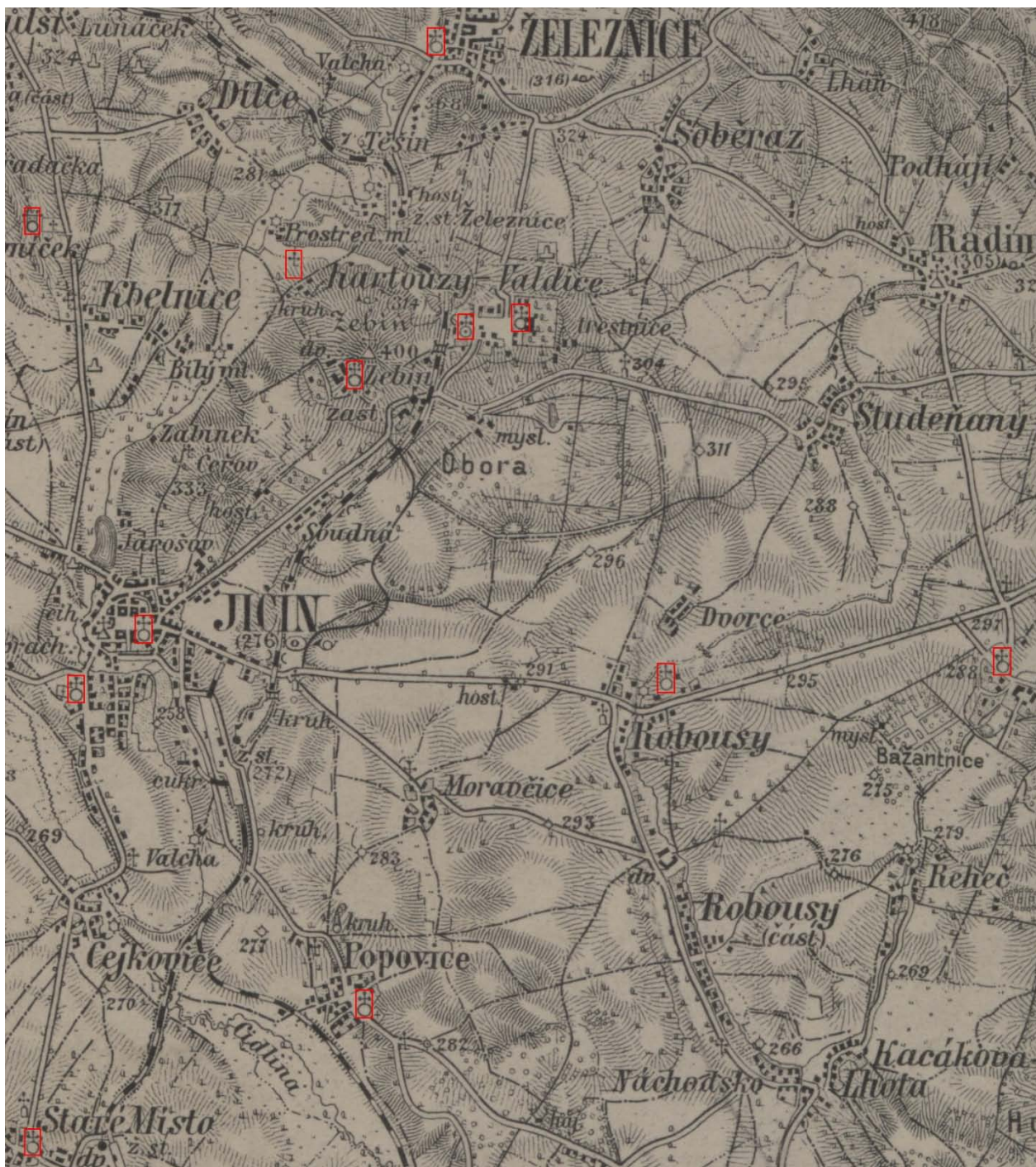
5. Druhá fáze prohledání

V této fázi jsou prohledávány už jen malé výřezy vstupního obrázku v okolí bodů, které prošly prahováním v první fázi. Porovnání opět probíhá pomocí korelace v rámci masky, stejně jako v předchozí fázi. Rozdíl je v tom, že je použito mnohem jemnější dělení úhlů (parametr `angles2`) a zvětšení značky (parametr `scales2`), což by v první fázi nebylo možné kvůli časové náročnosti, a pro prohledání jsou použity všechny vzorové exempláře hledané značky, nikoliv jen jediný, jako tomu bylo v předchozím případě. Cílem je stanovit přesněji geometrickou transformaci značky na mapě proti vzorové šabloně, což umožní použití silnějšího prahu na korelaci a efektivnější filtraci pozadí, která usnadní konečnou identifikaci značky ve třetí fázi.

Pro každý prohledávaný výřez obrázku je zvolena a zaznamenána šablona a parametry transformace (úhel, zvětšení), které dávají nejvyšší hodnotu korelace. Pokud tato hodnota překračuje práh pro druhou fázi korelace (parametr `cctol2`), pokračuje výřez do dalšího zpracování. Na výřez obrázku je aplikována inverzní geometrická transformace (tzn. otočení a změna velikosti o opačnou hodnotu, než jaká je detekována). Takto znormalizovaný výřez postupuje do poslední rozhodovací fáze. Protože v poslední fázi se za pomoci prahování odstraňuje pozadí výřezu, je vhodné, aby výřez co nejlépe „pasoval“ na značku, čehož se dosáhne právě důkladnějším prohledáním pomocí korelace a inverzí detekované geometrické transformace. Smyslem tedy je provést co nejpřesnější registraci výřezu mapy a vzorové značky (s uvažováním chyby pouze v malém okolí šablony značky, aby to odpovídalo masce v korelaci) – je tak možné použít jiný registrační algoritmus, je však třeba ohlídat, aby nezdivergoval do nesmyslné polohy.

Pro shrnutí, druhá fáze prohledání je provedena následujícím postupem:

1. Pro každou dvojici úhel-zvětšení jsou všechny vzorové šablony značky a jejich masky transformovány a proveden výpočet normalizované korelace v rámci masky na všech výřezech z první fáze (pro všechna možná umístění šablony na daný výřez).
2. U každého výřezu jsou zaznamenány parametry (úhel, zvětšení, zvolená vzorová šablona), jaké vedou k nejvyšší hodnotě korelaci v rámci daného výřezu. Pokud tato hodnota překračuje práh dle parametrů, postupuje výřez do dalšího zpracování.
3. Na základě parametrů (úhel otočení, míra zvětšení) stanovených v bodě 2 je provedena inverzní geometrická transformace malého výřezu okolo kandidáta na značku. Transformovaný výřez má pak stejnou velikost jako vzorová šablona. Tímto dojde k narovnání mírné deformace nakreslené značky do „standardní“ polohy.



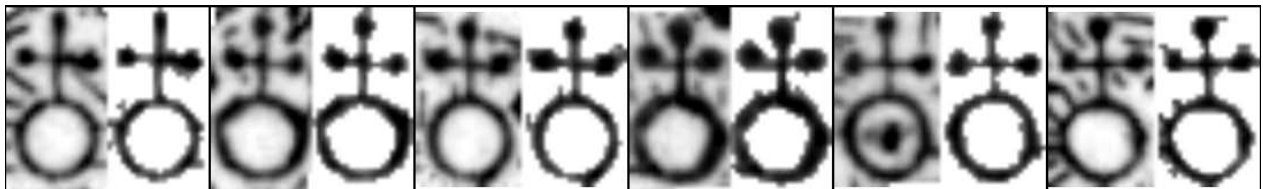
Výsledek druhé fáze prohledání na stejném výřezu mapy. Většina falešných detekcí byla odstraněna již nyní, všechny označené výřezy však ještě vstupují do třetí rozhodovací fáze.

6. Třetí fáze prohledání

V této fázi se již neprohledává mapa nebo její část, cílem je projít všechny normalizované výřezy z předchozí části a u každého rozhodnout, zda představuje značku či nikoliv. Před tímto rozhodnutím je nejprve za pomoci prahování odstraněno pozadí výřezu, které nesouvisí se značkou. Myšlenka prahování je následující (pro připomenutí, značka je po transformaci kontrastu/barev znázorněna tmavou na světlém pozadí): Filtrování pozadí pomocí prahování je provedeno ve 2 fázích, nejprve jsou zvolena jen silně tmavá místa (se silným prahem), poté obecně tmavá místa (se slabším prahem), která jsou však připojena, jen pokud se dotýkají některého silně tmavého místa (jedná se o podobný princip jako je například u Cannyho hranového detektoru). Aby skutečně došlo k zachování značky a určitému odfiltrování pozadí, jsou zahrnuty ještě další prvky. Za prvé musí prvotní silně tmavé části ležet jen v místě očekávané značky (takto silná preference hledaného tvaru je nebezpečná v tom, že pak jakýkoliv obrázek začne vypadat jako hledaný tvar – je to však vykompenzováno následným napojováním), za druhé je kolem obrysu vyhledávané značky vytvořena jakási hranice, přes kterou je spojitost uvažována jen přes silně tmavé pixely, navíc je vyžadována spojitost přes alespoň tři pixely, aby se vyloučil jen slabý dotyk nakreslených symbolů na mapě. Smysl takového postupu je za prvé takový, aby byly odstraněny tmavé izolované objekty ležící zcela mimo značku a aby se za „izolované“ považovaly i objekty, které jsou se značkou spojeny jen slabě (kresba starých map je velmi špatná a obsahuje mnohé šmouhy, které komplikují identifikaci značky). Postup prahování výřezu za účelem odstranění nežádoucího pozadí je tedy následující:

1. Nalezení silně tmavých částí výřezu (s intenzitou menší než stanovený nízký práh), jsou ponechány jen ty části, které leží v masce značky.
2. Nalezení obrysu značku, tzn. hranice masky značky – ta je udělána jako logický rozdíl mezi zvětšenou maskou značky (z první fáze prohledání) a maskou ze vzorových vstupních dat.
3. Nalezení obecně tmavých částí celého výřezu (je použit slabší práh než v bodě 1)
4. Filtrace pixelů z bodu 3: jsou ponechány jen oblasti, které jsou napojeny na některou z oblastí vybranou v bodě 1, přičemž konektivita přes pixely „hranice“ stanovené v bodě 2 je uvažována, jen pokud je přes alespoň 3 silně tmavé pixely (ve smyslu prahu použitého v bodě 1).
5. Z výřezu jsou ponechány (tzn. například v případě prahování pro DSC nastaveny na černou) ty pixely, které byly vybrány v bodě 1 a prošly filtrováním v bodě 4.

Tímto procesem je částečně odfiltrováno pozadí, které nesouvisí se symbolem značky. Vzhledem ke kvalitě starých map se to většinou však povede jen částečně.



Ukázka výsledku odstranění pozadí. V každé dvojici je vždy levý obrázek identifikovaný výřez mapy, pravý obrázek výsledek filtrace pozadí.

Konečné rozhodovací kritérium je výpočet vhodného příznaku na odfiltrovaném výřezu obrázku a porovnání, zda je získaná hodnota dostatečně blízká k hodnotě stejného příznaku vypočteného u vzorových exemplářů značky. Jedním z implementovaných kritérií je tzv. Diceův koeficient podobnosti, který je vypočten mezi oprahovaným výřezem a maskou vzorové značky, která byla nejlepší korelací určena ve fázi 2. Koeficient je určen vztahem

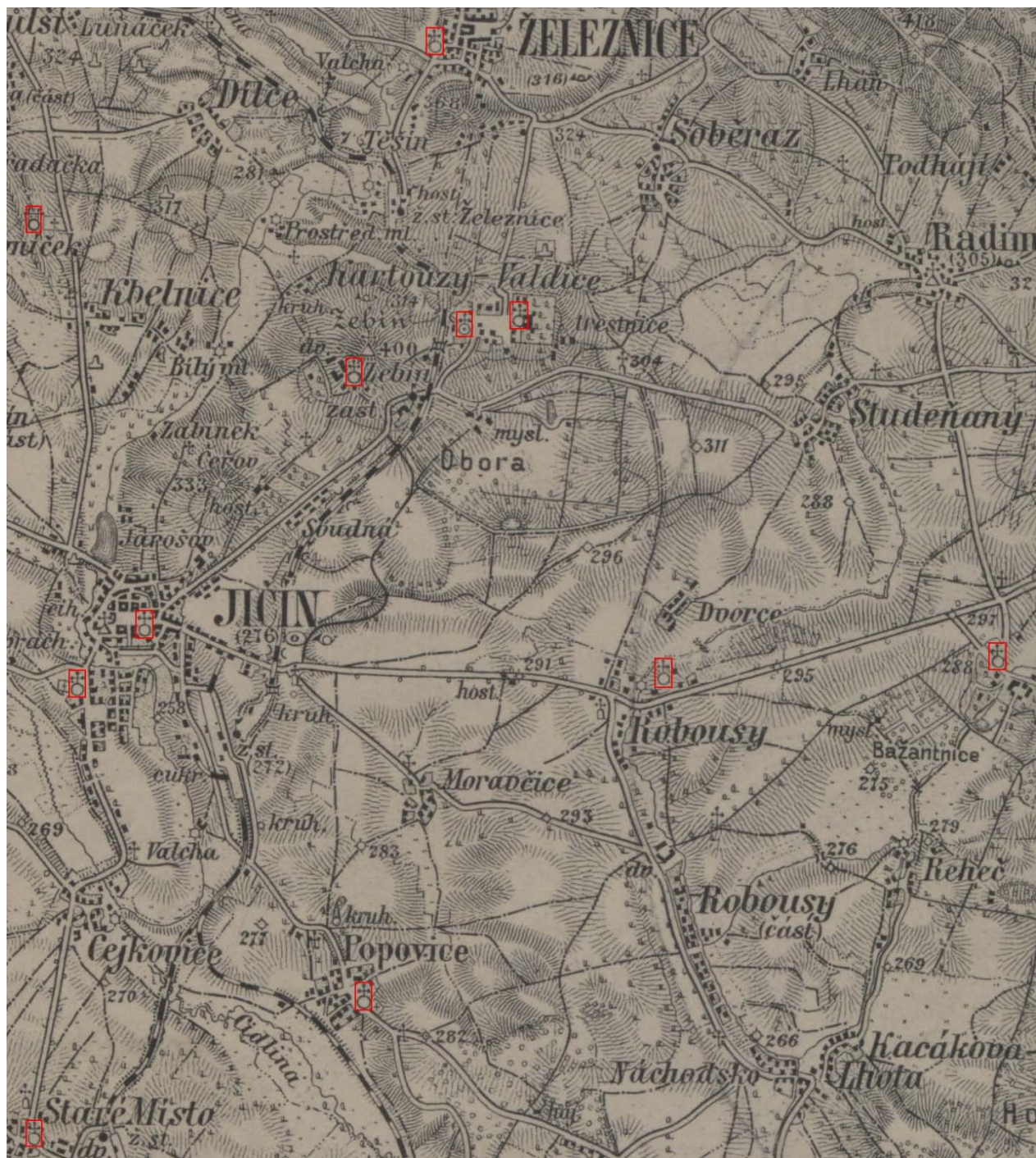
$$DSC(A, B) = \frac{2|A \cap B|}{|A| + |B|}.$$

Pokud je hodnota vyšší než stanovený práh (parametr `dsc_tol`), je výřez označen jako hledaná značka. Další vhodnou možností je příznakový vektor histogramu orientace gradientů (HOG), spočítaný pro vyfiltrovaný výřez, přičemž je porovnává blízkost (v příznakovém prostoru, udává parametr `hog_tol`) k některému ze vzorových exemplářů.

Výřezy mapy, které pozitivně projdou tímto posledním testem, jsou označeny jako hledaná značka a poloha je zaznamenána do databáze.

Postup konečné rozhodovací fáze je tedy následující:

1. Na oprahovaných jen vyčištěných (viz předchozí část) výřezech je spočítán příznakový vektor dle výběru (DSC proti masce zvolené vzorové značky, nebo HOG, popř. jiný příznak).
2. Pokud je příznaková hodnota podobnosti vyšší než stanovený práh (v případě DSC), nebo je příznakový vektor dostatečně blízký k stejně spočítanému příznakovému vektoru pro některou ze vzorových značek, je daný výřez označen jako exemplář hledané značky a jeho poloha na mapě zaznamenána do databáze. V opačném případě je zahozen a zpracování pokračuje pro další výřez.



Finální výsledek vyhledávání. Všechny značky byly nalezeny, dále byla nesprávně označena značka vlevo nahoře, která sice představuje kostel, nicméně zakreslený jinou (velmi podobnou) značkou.

7. Závěr

Uvedená metodika popisuje algoritmus vyhledávání bodových mapových značek na digitalizovaných starých mapách. Ten může tvořit základ jakéhokoliv specializovaného systému pro řešení uvedené úlohy. Umožní se tím snazší práce s mapu, což přispěje k jejímu rychlejšímu porozumění a většímu užítku.

Algoritmus byl vyvinut primárně pro mapy III. vojenského mapování, jejichž mapový klíč je velmi bohatý a které jsou díky své obsahové bohatosti a velkému časovému období, po které byly používány, nejžádanějšími starými mapami v ČR. Systém je vyvíjen v MATLABu. Podle doposud provedených experimentů je úspěšnost optimálně nastaveného rozpoznávacího systému přibližně 90%, míra falešných odezev je nižší než 10%. Stávající rychlost výpočtu je v řádu minut na jeden mapový list velikosti přibližně 60 megapixelů. Tato rychlost však velmi záleží na parametrech vyhledávání. U kvalitních map (tj. s malou diverzitou nakreslení jednotlivých exemplářů značky a s malým překryvem značek a zbytků kresby) stačí první část prohledání dělat pro málo hodnot transformačních parametrů a také jednotlivé prahy je možné nastavit přísnější (vyšší), rychlost zpracování je tak možné zlepšit až řádově. U méně kvalitních map je to samozřejmě také možné, pokud je rychlost prioritní, je to však na úkor kvality.

8. Seznam použitých zdrojů

- [1] GUYOMARD, J.; THOME, N.; CORD, M.; ARTIERES, T. (2012). Contextual detection of drawn symbols in old maps. ICIIP 2012, s. 837-840, ISSN 1522-4880
- [2] VIOLA, P.; JONES, M., (2001). Rapid object detection using a boosted cascade of simple features. CVPR 2001, Vol. 1, s. 511-518, ISSN 1063-6919
- [3] DALAL, N.; TRIGGS, B. (2005). Histograms of oriented gradients for human detection. CVPR 2005, Vol. 1, s. 886-893, ISSN 1063-6919
- [4] BISHOP, C. M. (2006). Pattern recognition and machine learning. ISBN 978-0-387-31073-2.
- [5] BOGUSZAK, František a CÍSAŘ, Jan. Vývoj mapového zobrazení území Československé socialistické republiky. 3. díl, Mapování a měření českých zemí od poloviny 18. století do počátku 20. století. 1. vyd., Ústř. správa geodézie a kartografie, Praha 1961