

Detekce kartografického zobrazení z množiny bodů

Tomáš Bayer

Katedra aplikované geoinformatiky a kartografie, Přírodovědecká fakulta, Univerzita Karlova v Praze, Albertov 6, 120 78, Praha 2, Česká republika, bayertom@natur.cuni.cz

Abstrakt

Detekce kartografického zobrazení z množiny bodů o známých rovinných souřadnicích představuje poměrně zajímavý a nepříliš často řešený problém. Lze ho využít při analýze kartografických produktů, které jsou k dispozici pouze v analogové formě, a nenesou informace o kartografickém zobrazení, ve kterém byly vyhotoveny. Níže uvedený postup používá k detekci kartografického zobrazení srovnání množiny bodů v mapě s množinou korespondujících bodů ve známém kartografickém zobrazení. Nad oběma množinami jsou zkonstruovány Voronoiovy diagramy, parametry jednotlivých Voronoiovy buněk následně analyzujeme. Za podobné jsou prohlášeny takové množiny, které mají podobné Voronoiovy diagramy. Článek se snaží ukázat souvislost mezi matematickou kartografií a 2D geometrickými strukturami.

Klíčová slova: Matematická kartografie, kartografické zobrazení, Voronoi diagram.

1 Úvod

Kartografické zobrazení K zobrazuje referenční plochu R_1 na referenční plochu R_2 , lze ho popsat prostřednictvím zobrazovacích rovnic v explicitním vyjádření

$$x = f(\varphi, \lambda) \quad (1)$$

$$y = g(\varphi, \lambda) \quad (2)$$

O zobrazovacích funkcích f , g se předpokládá, že jsou spojité a snadno diferencovatelné, vzájemně nezávislé. Proměnné φ, λ jsou zeměpisné souřadnice bodu na referenční ploše R_1 , x, y pravoúhlé souřadnice téhož bodu na referenční ploše R_2 .

Voronoiova teselace VT přiřazuje každému bodu množiny $P = \{P_1, P_2, \dots, P_n\}$ uzavřenou či otevřenou oblast, tj. $V(P) = \{V(P_1), V(P_2), \dots, V(P_n)\}$, takovou, že libovolný bod $X \in V(P_i)$ je blíže k bodu P_i než k jakémukoliv bodu P_j . Platí $d(X, V(P_i)) \leq d(X, V(P_j))$. Uzavřenou oblast $V(P_i)$ nazýváme *Voronoiovou buňkou*.

Body ležící na společné hranici $V(P_i)$ a $V(P_j)$, ($i \neq j$) jsou od bodů P_i, P_j stejně vzdáleny. Platí (1) Voronoioův diagram $V(P)$ je planárním grafem (2) Pro každý bod P je $V(P)$ konvexní (3) Vrchol Voronoiovy buňky je průnikem tří hran, právě když jsou Voronoiovy diagramy (dále jen VD) nedegenerované. (4) Leží-li bod P_i na konvexní obálce, je $V(P_i)$ otevřená.

VD je poměrně *citlivý* k poloze bodů, změna polohy některého z bodů může způsobit přegenerování jeho tvaru. Této vlastnosti využijeme při detekci tvaru kartografického zobrazení.

Voronoioův diagram a kartografické zobrazení.

Množiny podobných bodů generují i podobné VD. Můžeme tvrdit, že mezi analyzovanou množinou bodů a množinou bodů v kartografickém zobrazení existuje podobnost, pokud jsou podobné i jejich VD. Cílem postupu je nalezení co nejpodobnější dvojice VD vygenerovaných nad množinou analyzovaných bodů a množinou jim odpovídajících bodů v nějakém kartografickém zobrazení. V této souvislosti vyvstává řada otázek:

- Jak zhodnotit podobnost dvou VD?
- Budou pro analýzu použity všechny $V(P_i)$ nebo jen některé?
- Jaké množství bodů použít ke generování VD?
- Jakou metodikou realizovat výběr těchto bodů?
- Která kartografická zobrazení pro porovnání zvolit a jaké mají být jejich kartografické parametry?

Komplexní odpověď na tyto otázky přesahuje rozsah tohoto článku a vyžádala by si provedení výzkumu s následným statistickým zhodnocením dosažených parametrů. Berme tento článek jako prvotní seznámení s popisovanou metodou.

2 Proces analýzy

První krok zahrnuje rozvahu týkající se výběru množiny vstupních bodů P , nad kterými budou generovány VD. Vzhledem k faktu, že kartografické dílo může být (a většinou i bývá) z hlediska obsahového poměrně obsáhlé, není

únosné při analýze pracovat se všemi jeho prvky (resp. s jejich číselnou reprezentací), ale pouze s podmnožinou těchto prvků. U některých prvků bývá upřednostňována geografická věrnost nad půdorysnou přesností. Tento fakt ovlivňuje úspěšnost analýzy.

2.1 Množina vstupních bodů na mapě

Při volbě množiny vstupních bodů na mapě mohou nastat dva případy:

- *Mapa neobsahuje zeměpisnou síť.*

Množina vstupních bodů P představuje výběrový soubor ze všech bodových, liniových či plošných kartografických znaků mapy. Pro každý bod P_i máme k dispozici rovinné kartézské souřadnice $[x_i, y_i]$ (lze je pořídit např. za použití kartometrické digitalizace) a zeměpisné souřadnice $[\varphi_i, \lambda_i]$ (z aktuálních mapových podkladů). Odpovídající množinu obrazů bodů P v rovině kartografického zobrazení označíme Q . Libovolný bod $Q_i = [x'_i, y'_i]$, $i = 1, \dots, N$, určíme výpočtem ze zobrazovacích rovnic (1), (2) pro hodnoty φ_i, λ_i . Množiny P a Q tedy představují *identické body*.

- *Mapa obsahuje zeměpisnou síť.*

V tomto případě je situace jednodušší. Lze použít body P_i ležící v uzlových bodech poledníků a rovnoběžek, či doplnit je případně dalšími body volnými a označenými dle výše uvedených zásad.

Aby byly výpočty nad oběma množinami dobře podmíněné, výpočty zobrazovacích rovnic je vhodné provádět nad jednotkovou koulí ($R = 1$). Uveďme také nejdůležitější parametry týkající se volby vstupní množiny bodů, které ovlivňují významným způsobem kvalitu detekce zobrazení:

- *Typ kartografického znaku.*

S ohledem na metodiku tvorby mapových podkladů se pro body P_i jeví jako nejvýhodnější bodové kartografické znaky použité pro reprezentaci sídel či liniové kartografické znaky využívané pro znázornění významných prvků polohopisu/výškopisu (soutoky řek). Za nevhodné lze považovat kartografické znaky plošné znázorňující zpravidla lesní nebo vodní plochy, u kterých existuje poměrně malá pravděpodobnost, že se znázorněný stav krajiny zachoval až do současnosti.

- *Počet bodů.*

Pro praktické posouzení není zřejmě třeba, aby byl počet N prvků výběrového souboru příliš vysoký. Pro testování byly použity množiny bodů s rozsahem $N \in (50, 1500)$. Zřejmě tedy neplatí, že kvalita detekce kartografického zobrazení je závislá pouze na hodnotě N .

- *Rozložení bodů.*

Důležitou roli hraje rozložení bodů množiny P . V analyzovaném kartografickém díle by měly být body P_i rozloženy co nejrovnoměrněji po celé ploše území, abychom se vyvarovali pouze lokální shody množin P a Q . Pokud jsou body P_i rozloženy pravidelně, lze dosažené výsledky zobecnit: důsledkem podobnosti množin P a Q je tedy i podobnost jejich základních souborů. V opačném případě takové zobecnění nemusí být platné a dosažené výsledky mohou být zkreslené. Není vhodné umísťovat analyzované body do obrazů singulárních bodů (tj. zeměpisných či kartografických pólů) a jejich blízkého okolí.

- *Velikost území.*

Velikost analyzovaného území hraje podstatnou roli. Čím větší je analyzované území, tím více se liší vzájemné polohy jednotlivých bodů množin P , Q a detekce kartografického zobrazení je úspěšnější. Pro malá území s plochou menší než 100km^2 není prakticky možné detekci provést. Polohy bodů obou množin se ve všech kartografických zobrazeních liší jen velmi málo.

- *Přesnost zákresu bodů množiny P .*

V případě, kdy je situace na analyzované mapě znázorněna s nízkou přesností, výrazně klesá účinnost procesu automatické analýzy kartografického zobrazení. Tento bod se týká mapových podkladů, které nebyly konstruovány na geodetickém základu, tedy převážně historických map. Pokud jsou jednotlivé body P_i zatíženy náhodnými chybami v poloze a následně zobrazeny v kartografickém zobrazení, jejich srovnání s množinou odpovídajících identických bodů ve stejném zobrazení nezátíženou náhodnými chybami (aktuální mapový podklad) vykáže výrazně nižší podobnost.

2.2 Parametry kartografického zobrazení

Důležitou roli při procesu detekce hraje volba kartografických zobrazení a jejich parametrů, pro které budou hodnoty Q určovány. Paleta kartografických zobrazení, se kterými budeme srovnání provádět, by měla být co nejširší. Je vhodné do ní začlenit jak zobrazení jednoduchá, tak zobrazení nepravá, polykónická či neklasifikovaná.

Zajímavěji se jeví volba parametrů těchto kartografických zobrazení. Postup jejich určování by měl být automatizovaný, aby bylo možno proces rozpoznávání kartografického zobrazení snadno algoritmizovat. V zásadě lze volbu kartografických parametrů provést pro dva typy zobrazení:

- Volba parametrů pro zobrazení v normální poloze
- Volba parametrů pro zobrazení v obecné poloze

Volba parametrů pro zobrazení v normální poloze. Tato varianta je poměrně jednoduchá. Provedeme analýzu vstupní množiny P a určíme její těžiště $[\varphi_t, \lambda_t]$. Zeměpisná délka základního poledníku může být volena tak, že $\lambda_0 = \lambda_{Grenw.}$ nebo $\lambda_0 = \lambda_t$. Zeměpisná šířka nezkreslené rovnoběžky může být volena jako $\varphi_0 = 0$ nebo $\varphi_0 = \varphi_t$.

Volba parametrů pro zobrazení v obecné poloze. Tato varianta je obtížnější, pro množinu bodů je zapotřebí určit několik parametrů. První parametr představuje stanovení převládajícího směru území, ve kterém se body množiny P nacházejí. Tento směr může být charakterizován min-max boxem natočeným o úhel ω vzhledem k ose x . Převládající směr území je dán delší ze spojnic středů protějších stran.

Otázkou je, jak takto definovaný min-max box náležet. Hledání je založeno na myšlence opakované konstrukce min-max boxů s různými úhly ω , které jsou množině P opsány. U každého min-max boxu je určena jeho plocha S_ω , (S je funkcí ω), platí $S_\omega > S_{min}$. Pokud $S_\omega^i < S_{min}$, pak $S_{min} = S_\omega$. Po nalezení obdélníku S_{min} delší ze spojnic středů jeho protějších stran představuje převládající směr území. Takto vzniklé body označme jako $K_1 = [\varphi_1^k, \lambda_1^k]$, $K_2 = [\varphi_2^k, \lambda_2^k]$.

V dalším kroku nalezneme souřadnice φ_Q, λ_Q kartografického pólu Q . U azimutálního zobrazení platí $\varphi_Q = \varphi_t$ a $\lambda_Q = \lambda_t$. Při hledání kartografického pólu pro válcová a kuželová zobrazení využijeme faktu, že body K_1 a K_2 mohou ležet na ortodromě. Platí tedy

$$\lambda_Q = \arctan \frac{-\tan \varphi_1^k \cos \lambda_2^k + \tan \varphi_2^k \cos \lambda_1^k}{\tan \varphi_1^k \sin \lambda_2^k - \tan \varphi_2^k \sin \lambda_1^k} \quad (3)$$

$$\varphi_Q = \arctan \frac{\cos(\lambda_Q - \lambda_1^k)}{\tan \varphi_1} \quad (4)$$

Kartografickou šířku s_0 základní rovnoběžky určíme ze vztahu

$$s_0 = \frac{\pi}{2} - \arccos(\sin \varphi_1^k \sin \varphi_Q + \cos \varphi_1^k \cos \varphi_Q \cos(\lambda_1^k - \lambda_Q)).$$

Ostatní parametry určíme ze známých vztahů matematické kartografie, které již nebudeme uvádět. Popsaný postup můžeme pro kuželová a válcová zobrazení zjednodušit. Prohlásíme, že nezkreslená rovnoběžka je $\varphi_0 = \varphi_t$, základní poledník je $\lambda_0 = \lambda_t$. Vzhledem k předpokladu zhruba čtvercového tvaru území tak provádíme test pouze v normální poloze.

2.3 Určení míry identity

Důležitým krokem je posouzení míry identity množin bodů P a Q . Mezi oběma množinami zřejmě existuje vztah (obecně nemusí být lineární) definovaný předpisem $\xi : P \leftrightarrow Q$.¹ Platí tedy, že se jedná o vztah dvou nestejnorodých souřadnicových systémů.

¹Pro malá území lze tento vztah považovat za lineární a místo níže uvedené kubické transformace použít transformaci lineární (Helmertova).

²Transformace vyšších stupňů než 3 se již zpravidla nepoužívají, nepřinášejí podstatnější efekt.

Cílem tohoto kroku je nalezení a odstranění identických bodů v množinách P a Q s "příliš" velkou polohovou odchylkou, které by mohly nevhodným způsobem ovlivnit tvar VD. Vzhledem k faktu, že používáme nadbytečný počet identických bodů, využijeme ke stanovení míry identity metodu nejmenších čtverců. Je-li ξ nelineární, jako vhodná se jeví transformace vyššího stupně, např. kubická.² Minimální počet identických bodů je $N = 10$. Kubická transformace je dána rovnicemi

$$x' = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5v^2 + a_6x^3 + a_7x^2y + a_8xy^2 + a_9y^3 \quad (5)$$

$$y' = b_0 + b_1x + b_2y + b_3x^2 + b_4xy + b_5v^2 + b_6x^3 + b_7x^2y + b_8xy^2 + b_9y^3. \quad (6)$$

Tato transformace zobrazí úsečky jako kubické křivky, je konformní, tj. zachovává úhel v bodě mezi těmito křivkami. Označíme-li matici oprav jako $v(2N, 2)$, matici vyrovnaných hodnot jako $h(2N, 2)$ platí

$$v = Ah - l$$

Prvky matice $A(2N, 8)$ můžeme určit ze vztahu

$$A = \begin{pmatrix} 1 & x_1 & y_1 & x_1^2 & x_1y_1 \\ 1 & x_2 & y_2 & x_2^2 & x_2y_2 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_N & y_N & x_N^2 & x_Ny_N \\ y_1^2 & x_1^3 & x_1^2y_1 & x_1y_1^2 & y_1^3 \\ y_2^2 & x_2^3 & x_2^2y_2 & x_2y_2^2 & y_2^3 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ y_N^2 & x_N^3 & x_N^2y_N & x_Ny_N^2 & y_N^3 \end{pmatrix}. \quad (7)$$

Matice $l(2N, 2)$ má tvar

$$l = \begin{pmatrix} x'_1 & y'_1 \\ x'_2 & y'_2 \\ \dots & \dots \\ \dots & \dots \\ x'_N & y'_N \end{pmatrix}. \quad (8)$$

Podmínku MNC lze zapsat ve tvaru

$$h = (A^T A)^{-1} A^T l.$$

Dosazením (7) a (8) do (9) získáme hodnoty oprav v_x, v_y jednotlivých souřadnic, ze kterých lze určit v_{xy}

$$v_{xy} = \sqrt{v_x^2 + v_y^2}.$$

Opravu v_{xy} lze použít jako charakteristiku přesnosti popisující míru ztotožnění obou množin. Alternativní charakteristiku přesnosti může představovat výpočet poklesu oprav. Z matic A, l opakovaně odstraňujeme j -tý a $(j + N)$ -tý řádek, $j = 1, \dots, N$. Takto vzniklé matice označíme $A'(2(N-1), 8), l'(2(N-1), 2)$. Získáme celkem N takových matic. Platí

$$h' = (A'^T A')^{-1} A'^T l' \quad (9)$$

Opravy $v'(2(N-1), 2)$ určíme z rovnice

$$v' = A' h' - l'.$$

Vlastní pokles oprav Δv_{xy} lze vypočítat jako rozdíl čtverců oprav v a v'

$$\Delta v_{xy} = v^T v - v'^T v'$$

Z dalšího výpočtu budou vyloučeny body s výrazně vyšší hodnotou v_{xy} resp. Δv_{xy} (např. body, u kterých přesahuje $v_{xy} > 2\sigma_v$ resp. $\Delta v_{xy} > 2\sigma_{\Delta v}$).

2.4 Konstrukce Voronoiova diagramu

Ke konstrukci VD lze využít řadu postupů, autorem byl v práci použit inkrementální algoritmus založený na postupném přidávání bodů do již hotové Voronoiovy tesselace. Složitost algoritmu je $O(N^2)$. Autor využil vlastní implementaci v jazyce C++ provádějící export hran VD do formátu DXF, který lze načíst ve většině geoinformačticky orientovaného softwaru.

Před konstrukcí VD je nutno z množin P a Q odstranit duplicitní body. Týká se to zejména obrazů singulárních bodů v takových kartografických zobrazeních, které zobrazují singulární body opět jako body (kuželová, azimutální zobrazení).

Pro detekci bodů nacházejících se na konvexní obálce ($V(P_i)$ je otevřená oblast) bylo použito Jarvisovo skeonování, při kterém je posuzován úhel ω mezi poslední stranou konvexní obálky tvořenou body P_{i-1}, P_i a úsečkou $P_i - P_{akt}$. Hledá se takový bod P_{akt} , pro nějž je tento úhel maximální (tj. $\omega = \max$). Složitost algoritmu je $O(N \log N)$. Uvedený postup je výhodný pro množiny $N < 1000$.

Pro efektivnější konstrukci je možné použít např. algoritmus qhull využívající adaptivní matematiku (<http://www.qhull.org>), který je implementován např. v softwaru MATLAB. VD byly generovány na oblasti vymezené min-max boxem (rozměry $\Delta x_{mm}, \Delta y_{mm}$) s 25% přesahem. Upravený min-max box je dán vrcholy $V_1 - V_4$.

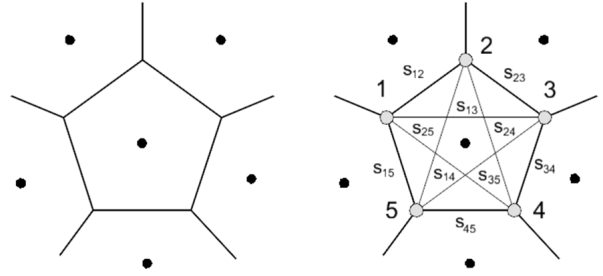
2.5 Podobnost $V(P)$ a $V(Q)$

Posouzení podobnosti $V(P)$ a $V(Q)$ je klíčovým krokem analýzy. Vzhledem k faktu, že VD jsou značně citlivé k poloze generátorů, lze tvrdit: $V(P)$ je podobné $V(Q)$ právě když jsou podobné i jejich geometrické parametry.

Pokud nalezneme takové zobrazení ξ , které představuje lineární operátor (tj. vyjadřuje existenci lineárního vztahu mezi $V(P)$ a $V(Q)$), jsou obě množiny P a Q podobné (tj. vzájemně posunuté, otočené, zvětšené či zmenšené).

Základním zprostředkujícím parametrem sloužícím pro výpočet vzájemného vztahu obou množin bude vztah vzdálenosti s mezi dvojicí bodů náležejících $V(P_i)$ a jí odpovídající vzdáleností s' mezi dvojicí bodů náležejících $V(Q_i)$. Konkrétněji se bude jednat o dvojice bodů ležících na hraně libovolné Voronoiovy buňky. Takových dvojic lze pro Voronoiovu buňku s n uzly nalézt $\binom{n}{2}$.

Pro tento účel převedeme Voronoiovu buňku na strukturu představovanou úplným neorientovaným grafem, viz obr. 1.



Obrázek 1: Převod Voronoiovy buňky na úplný neorientovaný graf.

Úplný graf. Graf $G = (H, U, \rho)$, mezi jehož každými dvěma uzly $u_i, u_j \in U$ existuje právě jedna hrana, nazýváme úplným grafem. Prvky množiny H představují hrany, prvky množiny U uzly grafu G , ρ pak incidenci grafu. Každá hrana $h_{ij} \in H$ má přiřazeno ohodnocení $w_{ij} \in \mathbb{R}$.

Matic W . Ohodnocení hran grafu budou v tomto případě představovat euklidovské vzdálenosti s . Graf G lze popsat čtvercovou maticí W řádu n , jejíž prvky w_{ij} jsou definovány následujícím vztahem.

$$w_{ij} = \begin{cases} 0, & \text{pro } i = j, \\ s(u_i, u_j), & \text{pro } i \neq j. \end{cases} \quad (10)$$

Matici W nad množinou $V(P_i)$ označíme W_{P_i} , $u_i \approx P_i$,

$$W_P = \begin{pmatrix} s_{11} & s_{12} & \dots & s_{1n} \\ s_{12} & s_{22} & \dots & s_{2n} \\ \dots & \dots & \dots & \dots \\ s_{1n} & s_{2n} & \dots & s_{nn} \end{pmatrix},$$

matici W nad množinou $V(Q_i)$ označíme W_{Q_i} , $u_i \approx Q_i$.

$$W_Q = \begin{pmatrix} s'_{11} & s'_{12} & \dots & s'_{1n} \\ s'_{12} & s'_{22} & \dots & s'_{2n} \\ \dots & \dots & \dots & \dots \\ s'_{1n} & s'_{2n} & \dots & s'_{nn} \end{pmatrix}.$$

Je-li ξ lineární operátor, zřejmě platí

$$\xi(W_{P_i}) = W_{Q_i}.$$

Charakteristická rovnice ξ . Vyjdeme z charakteristické rovnice lineárního operátoru ξ . Skalár λ představuje vlastní číslo lineárního operátoru ξ , pokud existuje vektor $v, v \neq 0$, pro který platí

$$\xi(v) = \lambda v.$$

Říkáme, že v je vlastní vektor operátoru ξ . Charakteristickou rovnici lze zapsat ve tvaru

$$\det(\lambda E - \xi) = 0.$$

Pro výpočet je praktičtější tvar

$$\begin{vmatrix} \lambda - \xi_{11} & -\xi_{12} & \dots & -\xi_{1n} \\ -\xi_{21} & \lambda - \xi_{22} & \dots & -\xi_{2n} \\ \dots & \dots & \dots & \dots \\ -\xi_{n1} & -\xi_{n2} & \dots & \lambda - \xi_{nn} \end{vmatrix} = 0$$

Kořeny rovnice s neznámou λ označíme λ_i ,

$$(\lambda - \lambda_1)(\lambda - \lambda_2)\dots(\lambda - \lambda_n) = 0$$

představují vlastní čísla lineárního operátoru ξ . Lze je určit např. s použitím QR rozkladu.

Podobnost matic a vlastní čísla. Analýza je založena na faktu, že podobné matice mají stejná vlastní čísla. Čtvercové matice B, C nazýváme podobnými, existuje-li taková regulární matice M , aby platilo

$$C = M^{-1}BM.$$

Sloupce matice M pak tvoří vlastní vektory. Pro vlastní čísla matic B, C platí

$$\lambda_B = \lambda_C.$$

Jestliže pro $\alpha \in \mathbb{R}$

$$C = \alpha M^{-1}BM,$$

pro vlastní čísla matic B, C platí

$$\lambda_C = \alpha \lambda_B.$$

Matice W je invariantní vůči úhlu natočení množin P a Q . V tomto případě je $M = E$, matice W_{Q_i} je α -násobkem matice W_{P_i} . Zřejmě platí

$$W_{Q_i} = \alpha W_{P_i} \quad (11)$$

$$\lambda_{Q_i} = \alpha \lambda_{P_i}. \quad (12)$$

Pro stopy matic lze psát

$$tr(W_{Q_i}) = \sum_{i=1}^n \lambda_{Q_i},$$

$$tr(W_{P_i}) = \sum_{i=1}^n \lambda_{P_i}.$$

Pak platí

$$tr(W_{Q_i}) = \alpha tr(W_{P_i}).$$

Jako testovací kritérium detekující, zda W_{Q_i} je α -násobkem W_{P_i} , použijeme

$$t_{Q_i} = \sum_{i=1}^n \lambda_{Q_i}^2,$$

$$t_{P_i} = \sum_{i=1}^n \lambda_{P_i}^2.$$

Pro vztah mezi t_{Q_i} a t_{P_i} platí

$$t_{Q_i} = \alpha^2 t_{P_i}, \alpha \in \mathbb{R}. \quad (13)$$

2.6 Selekcce $V(P)$ a $V(Q)$

Před vlastním výpočtem testovacího kritéria je nutno provést selekci množin $V(P)$ a $V(Q)$. Podobnost nebudeme posuzovat mezi všemi dvojicemi buněk $V(P_i)$ a $V(Q_i)$, množinu dvojic omezíme. Autor navrhuje provést selekci (z důvodu jednoduchosti výpočtu) na základě geometrických parametrů.

Geometrické parametry je nutno zvolit tak, aby reprezentativním způsobem charakterizovaly každou buňku, a současně byl jejich výpočet snadný. Autor použil následující trojici geometrických parametrů: (1) počet hran h , (2) obvod O , (3) obsah S . Postup lze formálně shrnout do dvou bodů:

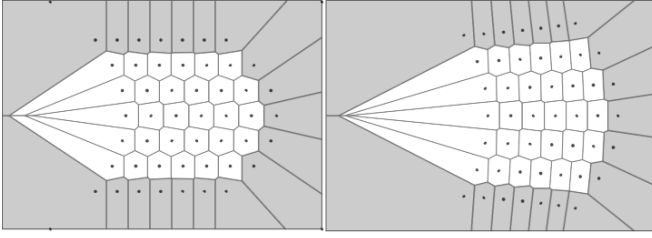
1. Pro každé $V(P_i)$ byly následně určeny jeho geometrické parametry: (1) počet hran h_P^i , (2) obvod O_P^i , (3) obsah S_P^i . Tyto údaje byly uloženy do dynamické datové struktury realizované maticí s rozměry $(3, N)$. Pro $V(P_j)$ ležící na konvexní obálce nebylo možné tyto hodnoty určit, a proto jsme pro další výpočty definovali $h_P^j = 0$, $O_P^j = 0$, $S_P^j = 0$. Stejným způsobem byly nad $V(Q_i)$ spočteny parametry h_Q^i , O_Q^i , S_Q^i .

Ke každé trojici geometrických parametrů nad buňkou $V(P_i)$ tedy existuje korespondující trojice parametrů nad buňkou $V(Q_i)$. Na základě výše uvedeného postupu vyvstává otázka, zda do porovnání nezahrnout i jiné geometrické parametry VD. Tímto problémem se však publikovaný článek podrobněji nezabývá.

2. V dalším kroku spočteme poměry mezi vzájemně odpovídajícími dvojicemi parametrů

$$k_h^i = \frac{h_P^i}{h_Q^i}, \quad k_O^i = \frac{O_P^i}{O_Q^i}, \quad k_S^i = \sqrt{\frac{S_P^i}{S_Q^i}}$$

a jim příslušející směrodatné odchylky σ_{kh} , σ_{kO} , σ_{kS} . Z výpočtu vyloučíme takové buňky $V(P_i)$ a $V(Q_i)$:



Obrázek 2: Selektce $V(P_i)$ a $V(Q_i)$, u kterých alespoň jeden z generátorů leží na konvexní obálce.

- (a) u kterých je počet hran různý. Neexistuje-li žádná dvojice buněk $V(P_i)$ a $V(Q_i)$ se stejným počtem hran, je hodnota testovacího kritéria $m_\alpha = \infty$.
- (b) alespoň jeden z dvojice parametrů se vztahuje ke generátoru P_i nebo Q_i ležícímu na konvexní obálce. Těmto geometrickým parametrům byla přiřazena hodnota 0. Uvedeným krokem odstraníme všechny neuzavřené buňky obou množin.
- (c) velikosti opravy $v^i = \bar{k}^i - k^i > 2\sigma$. Tímto krokem odstraníme buňky s výrazně odlišnými charakteristikami (tj. tvarově odlehle buňky) v obou množinách. Typickým příkladem jsou buňky nacházející se na okrajích množin, jejichž tvary a velikosti bývají značně rozdílné.

Vlastní detekci podobnosti obou množin P a Q kartografického zobrazení je možno provést, pokud počet zbylých buněk $V(P_i)$ a $V(Q_i)$ vyjádřený relativním poměrem splňuje podmínku $\frac{N'}{N} > 0.2$. V opačném případě by nebylo množství zbylých buněk statisticky reprezentativní. Hodnota 0.2 byla empiricky stanovena autorem. Na první pohled se jeví jako poměrně “malá”. Je však nutné vzít v potaz, že výše uvedenými kroky dojde k odstranění poměrně velkého množství buněk $V(P_i)$ a $V(Q_i)$. Se zvyšujícími se hodnotami N a snižujícími se hodnotami kroků $\Delta\varphi, \Delta\lambda$ se budou zřejmě hodnoty tohoto poměru zvětšovat.

2.7 Výpočet testovacího kritéria.

Po selekci buněk VD můžeme provést vlastní výpočet testovacího kritéria indikujícího splnění (11). Počet takto vzniklých Voronoiových buněk označme N' . Pro zbylé buňky $V(P_i)$ a $V(Q_i)$ naplníme matice $t_Q(1, N')$ a $t_P(1, N')$. V dalším kroku určíme prvky matice $\alpha(1, N')$

$$\alpha = \sqrt{\frac{t_Q}{t_P}}.$$

Reference

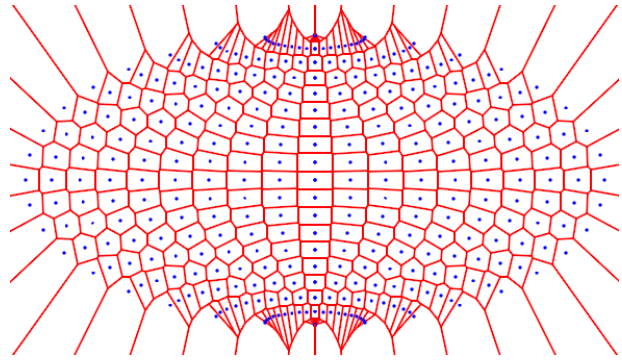
- [1] Hebák P.: Vícerozměrné statistické metody, Informatorium, 2005, Praha.

Pokud ξ představuje lineární operátor, pak $\alpha_{1i} = \alpha_{1j}$ pro libovolné $i \neq j$. V praxi díky náhodným či systematickým chybám nebudou hodnoty α_{1i} stejné, budou se lišit. Jako kritérium přesnosti a tím pádem i hodnotu ilustrující podobnost $V(P)$ a $V(Q)$ resp. množin P a Q lze použít směrodatnou odchylku

$$m_\alpha = \sqrt{\frac{\sum v_\alpha v_\alpha}{N' - 1}}. \quad (14)$$

Čím je hodnota m_α nižší, tím podobnější jsou si množiny P a Q . Provádíme-li srovnání $P^{(k)}$ s více množinami $Q^{(k,j)}$ reprezentujícími obrazy bodů v různých kartografických zobrazeních, nejpodobnější bude množina s m_α^{min} .

$$m_\alpha^{min} = \min(m_{\alpha_1}, m_{\alpha_2}, \dots, m_{\alpha_{N'}}) \quad (15)$$



Obrázek 3: Voronoiový diagram nad množinou P Hamme-rova zobrazení.

2.8 Závěr

Tento příspěvek se snažil nastínit možnost detekce kartografického zobrazení z množiny bodů srovnáním s množinou korespondujících bodů ve známém kartografickém zobrazení. Článek poukazuje na souvislost matematické kartografie, výpočetní geometrie a teorie grafů. Podrobnější analýzu dosažených výsledků při praktických testech se pokusíme přinést v některém z dalších čísel.

3 Poděkování

Článek vznikl za podpory projektu grantu GAČR č. 205/04/088 s názvem “Kartometrická a semiotická analýza a vizualizace starých map českých zemí z období 1518-1720”.

- [2] Buchar P., Hojovec V.: Matematická kartografie, Vydavatelství ČVUT, 1996, Praha.
- [3] Rourke O. J.: Computational Geometry in C, Cambridge University Press, 2005.