

POROVNÁVÁNÍ BODOVÝCH MNOŽIN

Point Sets Matching

Jiří Rošický

Abstrakt: Práce popisuje zefektivnění metody pro porovnávání 2D bodových množin. Prvky bodových množin jsou páry souřadnic. Metoda je založena na popisu bodových množin, který je invariantní vůči čtyřem typům transformací: posunutí, otočení, šum v souřadnicích a ztráta nebo přidání bodů.

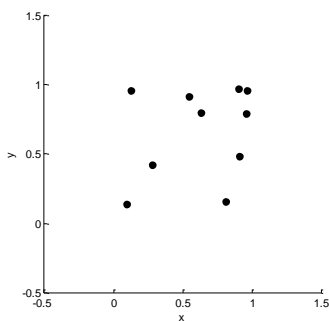
Key words: bodová množina, porovnávání, analýza dat, point set, matching, data analysis.

1. Úvod

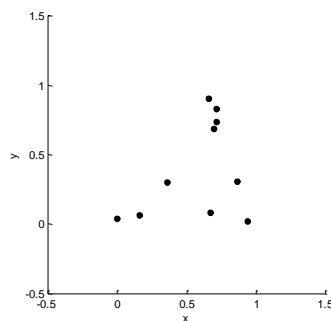
Vývoj metody pro porovnávání množin bodů je motivován problematikou identifikace fyzických objektů. Existuje řada způsobů, jak objekty značit a identifikovat. V případech, kde lze objekty reprezentovat 2D množinami bodů, se nabízejí některé zajímavé aplikace (např. z hlediska bezpečnosti).

2. Porovnávání bodových množin

Cílem práce bylo optimalizovat metodu navrženou v [1]. Konkrétním cílem bylo urychlení vlastního procesu porovnávání. Připomeňme zadání úlohy. Jsou dány dvě množiny bodů, A a B (**Obrázek 1**, **Obrázek 2**) a chceme vyhodnotit míru jejich podobnosti.

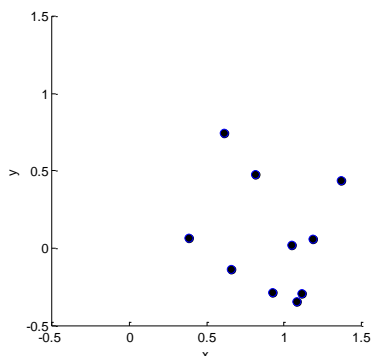


Obrázek 1: Množina bodů A .

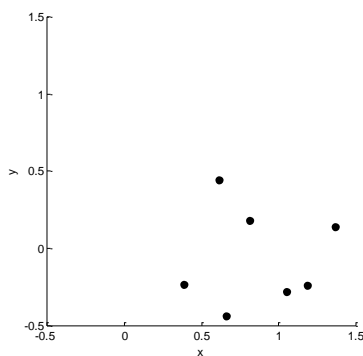


Obrázek 2: Množina bodů B .

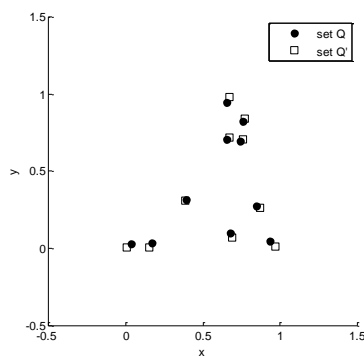
Uskupení bodů může být v rámci množiny posunuto a otočeno (**Obrázek 3**), v množině mohou chybět některé body (**Obrázek 4**), případně mohou být souřadnice bodů zašuměny (**Obrázek 5**). Neuvažujeme změnu měřítka.



Obrázek 3: Množina bodů A , posunutá a otočená.



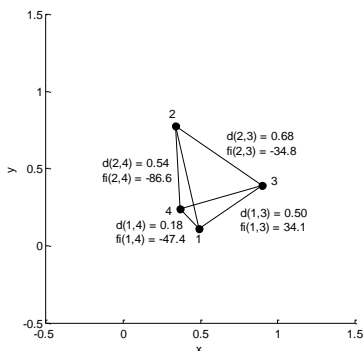
Obrázek 4: Množina bodů A , posunutá, otočená a bez tří bodů.



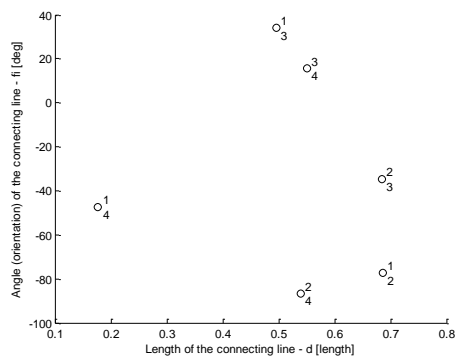
Obrázek 5: Dvě realizace jedné množiny bodů, souřadnice bodů jsou zašuměné.

3. Přehled současných řešení

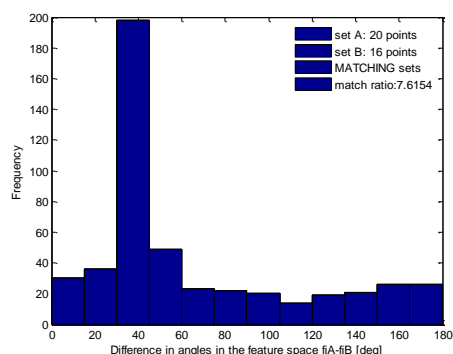
Možná řešení byla navržena např. v [2], [3], [4] a [1]. Vzhledem k cíli práce zmíníme přístup podle [1]. Množina bodů je zde popsána množinou příznaků, jejímiž prvky jsou páry: délka spojnice bodů d a úhel φ , který daná spojnice svírá s horizontálou (Obrázek 6). Příznaky jsou zaneseny do příznakového prostoru (Obrázek 7), který se dále analyzuje. Výsledkem je konstrukce veličiny, v jejímž rozdělení četnosti se vyhodnocuje přítomnost významného maxima, což je indikátor míry podobnosti porovnávaných množin (Obrázek 8 a Obrázek 9). Prohlédává se okolí každého příznaku, což je výpočetně náročné. Pro množiny s počtem bodů ~ 50 trvá jejich porovnání řádově sekundy (2,5 s [1]).



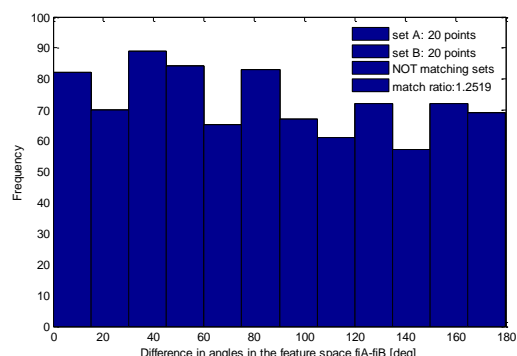
Obrázek 6: Popis množiny bodů – délky spojníc d a jejich orientace φ vzhledem k horizontále.



Obrázek 7: Příznakový prostor délek d a úhlů φ množiny bodů podle Obr. 6.



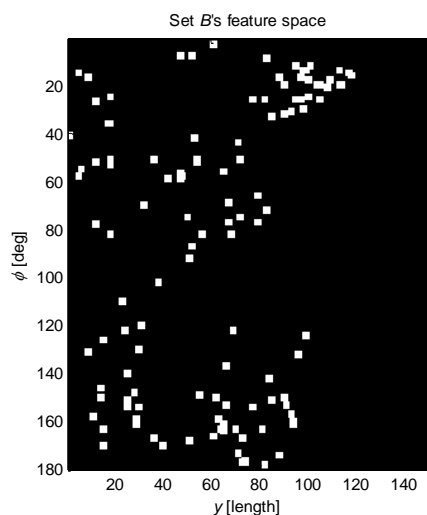
Obrázek 8. Rozdělení úhlových rozdílů.
Přítomnost výrazného maxima indikuje souhlas množin.



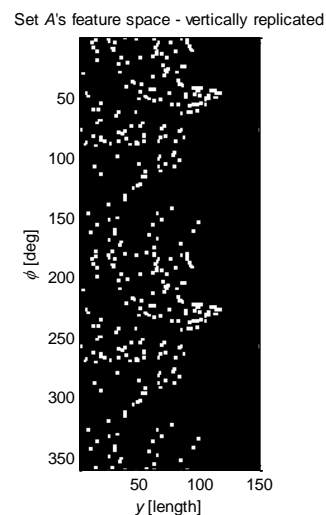
Obrázek 9. Rozdělení úhlových rozdílů.
„Ploché“ rozdělení indikuje nesouhlas množin.

4. Zrychlení analýzy příznakového prostoru

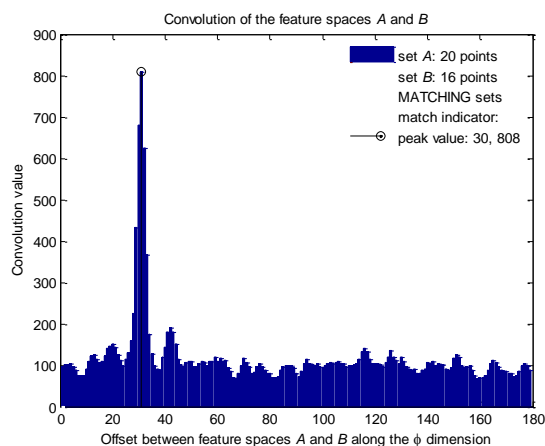
Zrychlení analýzy příznakového prostoru lze dosáhnout použitím konvoluce. Příznakové prostory obou porovnávaných množin se převedou na „binární obrazy“ s jedničkami v oblastech příznaků, s nulami na ostatních místech (**Obrázek 10**). Obrazová verze jedné porovnávané množiny se položí dvakrát „nad sebe“ ve směru úhlové dimenze φ (**Obrázek 11**). Obraz druhé množiny je přikládán postupně na (duplikovaný) obraz první množiny a pro všechny polohy podél úhlové dimenze φ se spočte suma součinů prvek po prvku obou obrazů. Výsledkem je funkce, která indikuje míru souhlasu porovnávaných množin (**Obrázek 12**, **Obrázek 13**).



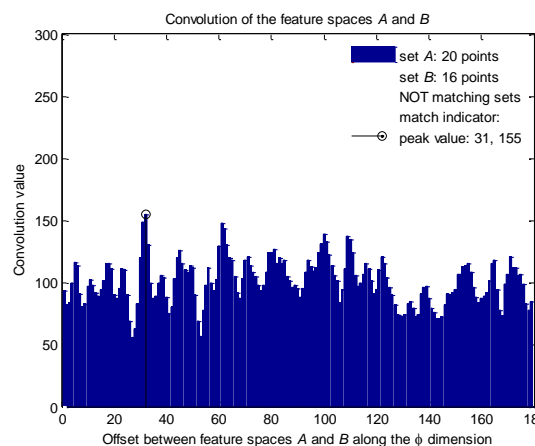
Obrázek 10. „Binární obraz“
příznakového prostoru.



Obrázek 11. Duplikovaný „binární
obraz“ příznakového prostoru.



Obrázek 12. Konvoluce příznakových prostorů – *přítomnost* výrazného maxima indikuje *souhlas* množin.



Obrázek 13. Konvoluce příznakových prostorů – *nepřítomnost* výrazného maxima indikuje *nesouhlas* množin.

5. Implementace

Navržená úprava byla implementována na systému IBM T60 notebook (2CPU, 1.8GHz, 2.5GB RAM), MATLAB R2009. Porovnání libovolně velkých bodových množin trvá přibližně 0,05 s pro příznakový prostor s rozlišením 150 bodů v délkové dimenzi y a 180 bodů v úhlové dimenzi φ (Obrázek 10).

6. Shrnutí

Popsanou úpravou bylo dosaženo přibližně konstantní doby potřebné pro porovnání bodových množin bez ohledu na množství bodů. Porovnávací doba 0,05s je ve srovnání s výsledky původní metody [1] (10 bodů \rightarrow 0,02s; 20 bodů \rightarrow 0,1s; 50 bodů \rightarrow 2,5s) významné zefektivnění.

Příští práce bude zaměřena na další optimalizaci. Možné zrychlení nabízí použití vlastností Fourierových obrazů konvolovaných funkcí. Dalšího zefektivnění bude možné dosáhnout omezením velikosti těchto funkcí např. vážením jejich okrajových oblastí (tzn. oblastí, kde příznak y nabývá extrémních hodnot). Zkvalitnění porovnávání bude pravděpodobně možné dosáhnout vyšší strukturací obrazů příznakových prostorů. Současné dvě úrovně příznaků (0/1) by bylo možné rozšířit na více úrovní a implementovat tak vážení.

Poděkování

Projekt byl podpořen interním grantem ČVUT SGS10/051/OHK2/1T/12.

Literatura

- [1] Rošický, Jiří: Structure Description and Matching. In. *Technical Computing Prague*. 2009. Praha : Humusoft, 2009, s. 87. ISBN 978-80-7080-733-0.
- [2] Murtagh F.: A New Approach to Point Pattern Matching. *Publications of the Astronomical Society of the Pacific*, 1992, in press.
- [3] Mathworks: Procrustes. *Matlab Documentation*. Available on-line: [http://www.mathworks.com/access/helpdesk/help/toolbox/stats/procrustes.html]
- [4] Klaus Kriegel et al.: An alternative approach to deal with geometric uncertainties in computer analysis of two-dimensional electrophoresis gels. *Electrophoresis* 2000, 21, 2637-2640