

# Posudek oponenta bakalářské práce

Autor/autorka práce: Ondřej KAAS

Název práce: Využití shlukování pro GIS aplikace

## Obsah práce

Autor kvalifikační práce se na potřebné úrovni seznámil s teorií shlukovacích metod a systematicky přehledně naznačil přednosti a vhodnost jejich použití především z pohledu objemu dat a kapacity výpočetní techniky. Při využití již existující shlukovací knihovny vytvořené na KIV bylo doplněno uživatelské prostředí pro práci s touto knihovnou, popsány prvky ovládání a ověřena funkcionální knihovny hierarchického shlukování na uměle vytvořených 2D datech s rozdílnou RGB informací. Výsledky dosažených nezávislých experimentů na uměle vytvořených rastrových datech prokázaly, že shlukovací knihovna pracuje spolehlivě.

Specifickou oblastí pro testování shlukovací knihovny následně byla geodata diferencí zpřesňujících transformací v procesu digitalizace katastrálních map. U tohoto typu dat je variantnost zadání velice široká (prostorová poloha, limity rozsahu území, difference transformačního přiřazení apod), a proto je nutné operativně volit nastavení parametrů shlukování včetně váhových koeficientů. Toto byl v minulosti problém často limitující exaktní hodnocení výsledků zpřesňujících transformací v procesu tvorby katastrálních map digitalizovaných z analogových map sáhových měřítek.

## Kvalita řešení a dosažených výsledků

Ne vždy se však jedná o 4-dimenzionální vektor  $t=(x_t, y_t, X_t, Y_t)$  s interpretací, která je popisována v kap. 3.3. Příkladem jsou data globálního transformačního klíče pro prostor Čech, kde je požadavek shlukovat vektor  $t=(x_t, y_t, X_t, Y_t, dx, dy)$  a kde  $x_t, y_t$  jsou souřadnice polohy zdrojové soustavy,  $X_t, Y_t$  jsou souřadnice polohy cílové soustavy a shluky jsou vytvářeny podle parametrů  $dx, dy$  a nikoli dle vzorců 3.1 a 3.2 jak je uvedeno v práci. Již správně je na tuto skutečnost reagováno v kap. 4.3.2 volbou váhových koeficientů. Jako oponent oceňuji ochotu autora práce zabývat se poněkud nestandardními datovými sadami. To je také jistě důvod poněkud problematické odborné terminologie z oblasti geodat, která se v práci objevuje...

## Formální úroveň

Po formální stránce je práce vyhotovena na odpovídající a požadované úrovni.

## Práce s literaturou

Protože v původním zadání práce bylo uvedeno, že seznam odborné literatury dodá vedoucí bakalářské práce, je literatura převážně z elektronických zdrojů a z provenience KIV asi na místě. Osobně bych uvítal, kdyby autor práce se alespoň okrajově zmínil o zdrojích datových sad (globální transformační klíče pro převod souřadnicových systémů stabilního katastru do S-JTSK, nebo problematika zpřesňujících transformací v procesu digitalizace analogových katastrálních map).

### Splnění zadání

Práce zadané téma ověření knihovny KIV pro hierarchické clusterování data streamů a její využití pro shlukování geodat podle odlišných atributů včetně volby různých vah splnila.

### Doplňující informace k práci

Vysoce kladně hodnotím ochotu autora aplikovat problematiku hierarchického clusterování na problematiku datových sad geodat.

### Dotazy k práci

- 1) Jak vysvětlíte, že při zmenšujícím se parametru *MaxDistance* se sníží počet shluků (viz kap. 4.3.2)?
- 2) Popište, na základě jakých vlastností byly testovány „hrubé chyby“ v datech (viz kap. 4.3.2)?

Navrhuji hodnocení známkou **velmi dobře** a práci doporučuji k obhajobě.



V Plzni 28.5.2013

Doc. Ing. Václav Čada, CSc.