

# Oponentský posudek na diplomovou práci Jakuba Žemličky:

## Recognition of 3D Objects with Uniform Surface Reflectance

Posudek je napsán i jako zpětná vazba pro diplomanta. Při obhajobě, pro účely hodnocení, lze přeskočit k otázkám do diskuse a k celkovému zhodnocení.

### Téma a cíle práce

Práce se zabývá rozpoznáváním objektů v obrazech. Objekt je reprezentován kolekcí malých částí obrazů. Úloha nalezení (rozpoznání) objektu v obraze je pak formulována jako nalezení dostatečného počtu *korespondujících* částí obrazu. Tato úloha je obvykle nazývána *Image matching* a je velmi atraktivní v oboru počítačového vidění a rozpoznávání. Nehledě na významný pokrok v oblasti v uplynulé dekádě, úloha ještě není zcela vyřešena. Problematické je zejména nalézání objektů s rozsáhlými homogenními oblastmi bez textury. Diplomová práce se zabývá právě touto problémovou oblastí.

Cílem práce bylo navrhnout nový detektor obrazových částí pro objekty s málo členitou texturou a ověřit jeho funkčnost na problému klasifikace automobilů.

### Splnění cílů práce – postup řešení, výsledky

Diplomant navrhl detektor polygonálních oblastí. Což považují za rozumnou volbu vzhledem k častému výskytu polygonů na objektech „vyrobených“ člověkem (man-made objects).

Metoda začíná nalezením hranových pixelů ve vstupním intenzitním obrázku pomocí Cannyho detektoru. Pro robustní detekci přímých hranových úseků navrhl diplomant originální fit-split algoritmus.

Pro detekci společných bodů přímek navrhl diplomant poměrně překvapivý, z mého pohledu elegantní, algoritmus, který prohledává vylepšený rastrový obrázek. Tedy nikoli analytické řešení, které by se po nalezení parametrů přímých úseků nabízelo. Volba algoritmu je konzistentně vysvětlena a algoritmus sám je dobře vysvětlen.

Z takto zpracovaného obrazu je vytvořen graf, kde nalezeným přímým úsekům odpovídají vrcholy, které jsou spojeny hranou, pokud mají společný bod. Vrcholy odpovídající kolineárním úsekům jsou sloučeny v rámci dalšího zpracování.

Polygony v obraze pak odpovídají cyklům v grafu. Jelikož počet všech cyklů v grafu může být příliš velký, navrhl diplomant algoritmus, který hledá pouze speciální cykly, které jsou vhodné pro danou úlohu. Jde o prohledávání do hloubky, kde prohledávání je zastaveno v okamžiku, kdy je porušena některá z podmínek kladená na vhodný cyklus (polygon) např. konvexita.

Těžiště experimentů je v nalézání registračních značek na automobilech. Metoda našla 84% registračních značek v kolekci 3688 obrazů obsahujících 3329 čitelných značek. Druhý experiment, který testuje opakovatelnost detekce je popsán nedostatečně. Implementace algoritmu je dostatečně rychlá, běží téměř v reálném čase.

### Otázky do diskuse

Algoritmus má poměrně velké množství parametrů. Z experimentů není jasné, jakým způsobem byly nalezeny použité parametry. Byla použita křížová validace na trénovací množině? Jak jsou dané parametry stabilní? Neboli jak moc je potřeba je přenastavit pro použití s jinými daty?

Ze zhodnocení experimentů mi také není jasné, jestli algoritmus dosáhl požadované úspěšnosti nebo ne.

## Text práce

Práce je napsána konzistentně. Oceňuji, že se autor soustředí na hlavní téma a zbytečně nepopisuje triviální kroky a známé postupy. Jím navržené algoritmy jsou vždy popsány v pseudokódu, což by mělo umožnit snadnou reimplementaci, Vysoko hodnotím obsažný a kriticky zpracovaný přehled existujících přístupů, především v části 2.1.2.

Experimenty jsou asi nejslabší částí práce. Experiment na opakovatelnost detekce je popsán zcela nedostatečně.

Titul neodpovídá zcela obsahu práce. Doporučil bych nějaký více konkrétní, např. „Detection of polygonal features in images“ a pod. V práci není v podstatě nic o „Recognition“, tedy rozpoznávání 3D objektů.

Text je psán dobrou angličtinou a je srozumitelný. Jen pár drobných výtek:

- Vychýlená hodnota je „outlier“, nikoli „outlayer“.
- Pro zobrazení hran a hranových úseků, obr. 2.6, bych doporučil použít negativní zobrazení, které by bylo mnohem přehlednější.
- U některých algoritmů či postupů chybí citace, např. u Cannyho detektoru nebo Bresenhamova algoritmu.
- Množné číslo od „maximum“ je „maxima“.
- Tvar neurčitého členu „a, an“ závisí na výslovnosti slova nikoli počátečním písmenu samotnému. Tedy „a unique“, nikoli „an unique“ a podobně.
- Grafy na obrázku 3.4 bych doporučil rozdělit podle rozlišení vstupních obrazů. Takto je bimodalita histogramů poněkud matoucí.