

# Posudek závěrečné práce Tomáše Krupky

oponent: Karel Zimmermann

Problém bezznačkového snímání pohybu artikulovaných objektů (např. lidí nebo zvířat) z několika statických kamer je nedílnou součástí mnoha úloh od virtuální reality až po analýzu pohybu pro sportovní a lékařské účely. Tomáš Krupka vychází ze současné metody [1] jak bylo doporučeno školitelem a objekt modeluje jako součet normálních rozdělení vázaných artikulovanou strukturou (kapitola 2). Odhad parametrů modelu je založen na gradientní minimalizaci podobnosti průmětu modelu s kamerovými daty (kapitola 3). Implementace založená na metodě "automatické diferenciace" dosahuje rychlosti zpracování v reálném čase (kapitola 5). Ve speciální kapitole řeší problém semi-autonomní inicializace, který považuji za zajímavý a z praktického pohledu velmi užitečný, hlavně pro účely rychlého manuálního opravování. Celkově je práce dobře čitelná a pochopitelná i pro člověka neznalého v počítačovém vidění a výsledný algoritmus je ověřen na několika nasnímaných sekvencích. Pro budoucí korektní vyhodnocení implementace a experimentů doporučuji použít nějakou ze standardních testovacích sad.

V zadání bylo dále požadováno: (i) návrh metody pro adaptaci modelu na změnu osvětlení, (ii) ověření robustnosti na složitých scénách s několika pohybujícími se objekty. I přesto, že se nepovedlo úspěšně naimplementovat (i) a nenašel jsem žádné výsledky zahrnující (ii), považuji způsob napsání práce i implementaci za nadprůměrnou a hodnotím ji známkou **A - výborně**.

## Literatura:

- [1] Stoll et al.: Fast Articulated Motion Tracking using a Sums of Gaussians Body Model, Proceedings of IEEE International Conference on Computer Vision, pp. 951-958, 2011.
- [2] R.Plankers and P.Fua: Articulated Soft Objects for Multi-View Shape and Motion Capture, Transactions on Pattern Analysis and Machine Intelligence, 2003.
- [3] D.Piponi: Automatic differentiation, C++ templates, and photogrammetry. Journal of graphics, GPU, and game tools, 9(4):41-55, 2004. **Otázky do diskuze:**

1. Plankers a Fua [2] dosáhli impresivních výsledků kombinací siluet a hloubky ze stera. Domníváte se, že současný rozvoj RGB-D senzorů (jako je Kinect nebo RealSense) bude mít zásadní vliv na metody bezznačkového snímání pohybu? Umožňuje váš přístup jednoduše zahrnout hloubková

data?

2. Objekt je modelován jako součet normálních rozdělení, jejichž podobnost je měřena např. vztahem (2.11). Zvážíte i jiné míry pro výpočet podobnosti pravděpodobnostních rozdělení, napr. Kullback–Leibler divergenci?
3. Je optimalizační problém konvexní?