



Thesis title:	Investigating Visual Localization Based on Semi-Generalized Camera Pose Estimation
Author's name:	Alena Smutná
Type of thesis :	bachelor
Faculty/Institute:	Faculty of Electrical Engineering (FEE)
Department:	Department of Cybernetics
Thesis reviewer:	RNDr. Zuzana Kúkelová, Ph.D.
Reviewer's department:	Department of Cybernetics

The goal of this thesis was to investigate in detail how the accuracy of camera poses estimated from 2D-2D correspondences between images depends on the geometric relation between the images. This is an interesting and not well-understood problem with direct applications to visual localization.

Visual localization is the problem of estimating the position and orientation, i.e., the camera pose, from which an image was taken. Visual localization algorithms are important parts of applications such as augmented reality and autonomous robots, e.g., self-driving cars. Most commonly, visual localization algorithms represent the scene via a 3D model and estimate the camera pose of a query image from 2D-3D correspondences between pixels in the query image and 3D scene points. While these approaches are highly accurate, building and maintaining a 3D model of a scene can be challenging. An alternative to using a 3D model for pose estimation is to represent the scene as a set of images with known intrinsic and extrinsic parameters and to estimate the query image pose from 2D-2D matches between the query and database images. This approach avoids having to construct and maintain a 3D scene model. However, in our experience, this structure-less visual localization approach is less accurate than the classical structure-based approach described above.

As a step towards understanding this behavior, the goal of the thesis was to investigate how the camera pose accuracy depends on the geometric relationship between the query and the database images and between the database images themselves. To this end, the thesis of Alena Smutná, after explaining the concepts underlying such a localization strategy, provides a detailed experimental analysis of this behavior. In detail, the thesis makes the following contributions:

- The thesis reviews structure-based and structure-less localization pipelines and motivates the advantages of the latter with respect to the former. It further explains the individual parts of the structure-less pipeline in detail. Given that this topic is not taught as part of the Bachelor studies, Alena had to learn new concepts and techniques. Thus, this part of the thesis constitutes a contribution in itself.
- The thesis presents detailed experiments that analyze the behavior of the chosen structure-less method on multiple scenes of a standard dataset, exploring how the distances between query and database images and between database images impact pose accuracy. To be able to draw valid conclusions, Alena not only used real images but also created virtual images in order to have more control over the distances between the query and database images. She proposed and investigated three different ways to create such virtual views. The results presented in the thesis are informative for rendering-based approaches to the localization problem that can create such virtual views (e.g., to improve pose accuracy). The level of detail of the experimental evaluation is clearly beyond what one would expect for a Bachelor thesis and would be very good for a Master thesis as well.
- The results of the experiments are interesting and partially unexpected. For example, the availability of database images taken close to the query image (but not too close) leads to better pose accuracy (our initial hypothesis was that database images taken from farther away would be more useful). The results presented in this thesis and the conclusions drawn from them will certainly inspire our further research into this topic.



THESIS SUPERVISOR'S REPORT

Alena is a very good student. She worked very independently and showed great attention to detail in the way she set up the experiments, executed them, and analyzed the resulting data. She was able to quickly grasp new concepts and learn about new problems and their solutions. She was able to nicely describe the motivation of the thesis and the technical parts and to derive good conclusions from the results of her experiments. Even in the review of the state-of-the-art and in the technical part of the thesis, she was able to describe the content in her own words, rather than closely adapting text from papers when describing prior work. She fulfilled the goals of the thesis completely.

The grade that I award for the thesis is **A - excellent**.

Date: **14.6.2023**

Signature:

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Investigating Visual Localisation Based on Semi-Generalized Camera Pose Estimation
Jméno autora:	Alena Smutná
Typ práce:	bakalářská
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Ing. Čeněk Albl Ph.D.
Pracoviště oponenta práce:	Diminished Reality, Inc

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	náročnější
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
<p>Tématem práce jsou pokročilé metody počítačového vidění, konkrétně problematika vizuální lokalizace pomocí fotografií. Cílem byla analýza specifických scénářů a konfigurací existujících metod pro výpočet polohy a orientace kamery, která obraz zachytila, vzhledem ke 3D scéně, respektive referenčním kamerám, jejichž poloha a orientace oproti scéně je známá.</p> <p>Zadání hodnotím jako náročné, problematika patří spíše k těm složitějším. Vyžadováno bylo zorientování se v projektivní geometrii, algebraických metodách, robustním odhadování, optimalizaci a statistické analýzy. K práci bylo nutné použít několik specializovaných knihoven a transformovat data tak aby bylo možné jejich propojení. Experimenty musely být navrženy samotnou autorkou, tak aby ověřily zadané hypotézy. K vyvození závěrů byla potřeba důkladná analýza a zvážení mnoha faktorů, které na výsledky mají vliv.</p>	
Splnění zadání	splněno s menšími výhradami
<i>Posuďte, zda předložená závěrečná práce splňuje zadání. V komentáři případně uveďte body zadání, které nebyly zcela splněny, nebo zda je práce oproti zadání rozšířena. Nebylo-li zadání zcela splněno, pokuste se posoudit závažnost, dopady a případně i příčiny jednotlivých nedostatků.</i>	
<p>Zadání považuji z větší části za splněné. Prozkoumány byly vlivy dvou hlavních faktorů a to vzdálenost hledané kamery od kamer referenčních a vzájemná vzdálenost referenčních kamer. Parametry základních experimentů byly nastaveny dobře, výsledky důkladně zhodnoceny a v některých případech byly navrženy navazující experimenty lepšímu porozumění trendů, které se ve výsledcích objevily. Oceňuji zvláště badavý přístup autorky, která se snaží výsledkům opravdu porozumět, najít vysvětlení jak pro trendy, tak i pro anomálie ve výsledcích. Postupně eliminuje faktory ovlivňující výsledek, a tak izoluje pravou příčinu. Jde opravdu do hloubky, zvažuje možné závěry, ale také správně rozpoznává situace, kdy závěr nelze s určitostí udělat. Z tohoto hlediska považuji práci autorky za příkladnou a vědeckou.</p> <p>Za menší výhradu považuji několik technických nedostatků při návrhu experimentů, které vedly k nedostatečnému rozuzlení problému. Ačkoliv studován byl vliv polohy a orientace referenčních a hledaných kamer, výsledky byly ovlivněny i dalšími faktory. Jedním z nich je počet referenčních kamer s 2D korespondencemi k hledané kameře. Jelikož tento parametr nebyl konstantní, nebylo možné vznést závěry ohledně vlivu základních parametrů. Autorka si to alespoň v sekci 3.5.4 uvědomila a upravila experiment tak, aby vliv počtu kamer byl potlačen, nicméně nepodařilo se to úplně a navíc předchozí experimenty v sekci 3.3 takto upraveny nebyly. Dále zde velkou roli hraje faktor počtu korespondencí mezi hledanou kamerou a referenčními kamerami. Tento vliv byl zvažován pouze v sekci 3.4, která zkoumala vliv malých vzdáleností referenčních kamer za pomoci automatického generování syntetických kamer. Jsem toho názoru, že vliv těchto dvou zásadních faktorů se dal eliminovat ze všech experimentů, například omezením počtu korespondencí a referenčních kamer shora, tak aby v daném rozsahu zkoumaného parametru byly tyto počty identické. To by dalo vzneseným závěrům větší hodnotu a možná pomohlo vysvětlit více trendů, které se ve výsledcích objevily.</p>	
Zvolený postup řešení	správný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Autorka postupovala velmi metodicky a z velké většiny správně, viz předchozí sekce.	

Odborná úroveň	A - výborně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
K vypracování autorka použila znalosti algebry, projektivní geometrie, statistiky, optimalizace. Porozuměla velmi komplexní problematice minimálních řešení geometrických problémů a jejich úskalím. Naučila se používat rozsáhlé, state-of-the-art knihovny pro 3D rekonstrukci a geometrii COLMAP a PoseLib a jimi poskytované algoritmy. Zvládla pracovat s daty popisujícími 3D scénu a kamery, provádět prostorové transformace a simulovat projekci 3D do 2D. V neposlední řadě si osvojila robustní odhadování metodou RANSAC a jejími rozšířenými variantami.	
Formální a jazyková úroveň, rozsah práce	B - velmi dobře
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Práce je velmi dobře strukturovaná, obsahuje všechny náležité sekce a velmi dobře se čte. Experimenty jsou detailně popsány, grafická prezentace výsledků je na velmi dobré úrovni. Objevují se frekventované, ale nikoliv zásadní chyby v angličtině. Jako překážku hodnocení A vidím nedostatečný úvod do projektivní geometrie a modelu kamery, kdy od samotného začátku, až do sekce 3 autorka používá termíny a symboly pro střed kamery, orientaci kamery, 2D projekce 3D bodů a přitom první alespoň náznak popisu procesu projekce 3D bodu do kamery je až v obrázku 3.8.	
Výběr zdrojů, korektnost citací	A - výborně
<i>Vyjádřete se k aktivitě studenta při získávání a využívání studijních materiálů k řešení závěrečné práce. Charakterizujte výběr pramenů. Posuďte, zda student využil všechny relevantní zdroje. Ověřte, zda jsou všechny převzaté prvky řádně odlišeny od vlastních výsledků a úvah, zda nedošlo k porušení citační etiky a zda jsou bibliografické citace úplné a v souladu s citačními zvyklostmi a normami.</i>	
Zdroje jsou relevantní, vyčerpávající a jsou náležitě odkazovány v textu.	
Další komentáře a hodnocení	
<i>Vyjádřete se k úrovni dosažených hlavních výsledků závěrečné práce, např. k úrovni teoretických výsledků, nebo k úrovni a funkčnosti technického nebo programového vytvořeného řešení, publikačním výstupům, experimentální zručnosti apod.</i>	
Vložte komentář (nepovinné hodnocení).	

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Shrňte aspekty závěrečné práce, které nejvíce ovlivnily Vaše celkové hodnocení. Uvedte případné otázky, které by měl student zodpovědět při obhajobě závěrečné práce před komisí.

Práce je na velmi vysoké úrovni a v mnoha ohledech překonává rozsah bakalářské práce. Náročnou problematiku autorka zvládla s přehledem a prokázala schopnost proniknout do hloubky problému. K prozkoumání vlivu parametrů určených v zadání musela navrhnout vlastní experimenty a tohoto úkolu se zhostila velmi dobře. Oceňuji zejména autorky kritický přístup a analytické myšlení při hledání příčin trendů v datech. I přes výborný výkon nelze opomíjet některé nedostatky v návrhu experimentů a strukturní a jazykové stránce práce.

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **B - velmi dobře**.

Datum: 13.6.2023

Podpis: