

I. IDENTIFIKAČNÍ ÚDAJE

Název práce:	Zrychlení lokalizace pomocí knihovny InLoc
Jméno autora:	Martin Sebera
Typ práce:	diplomová
Fakulta/ústav:	Fakulta elektrotechnická (FEL)
Katedra/ústav:	Katedra kybernetiky
Oponent práce:	Ing. Bc. Radim Špetlík
Pracoviště oponenta práce:	Katedra kybernetiky

II. HODNOCENÍ JEDNOTLIVÝCH KRITÉRIÍ

Zadání	lehčí
<i>Hodnocení náročnosti zadání závěrečné práce.</i>	
Hlavní téma práce, totiž optimalizace kódu knihovny InLoc, se specializace "Počítačové vidění a digitální obraz", kterou student absolvoval, týká jen okrajově, a svou podstatou je vhodná spíše pro kvalifikační práci studenta, který ukončuje bakalářské studium a obor "softwarové inženýrství".	

Splnění zadání	splněno s většími 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>	
Student se seznámil s lokalizační metodou InLoc, zřejmě provedl měření běhu optimalizovaných i neoptimalizovaných implementací, našel nejpomalejší části kódu a zrychlil je. Popis měření běhů je zmatené, autor sice uvádí v tabulkách 7.2-7.4 naměřené hodnoty, ovšem v textu práce jsem nenašel odkazy do těchto tabulek, takže se lze pouze dohadovat, co čísla v tabulkách znamenají. Dále např. na straně 73 autor píše, že „bylo nutné spustit celý algoritmus na jednom vlákně (tj. nahradit všechny paralelní cykly parfor za cykly for)“, přesto jsou v tabulce 7.2 názvy funkcí začínající „parfor_“. Dále autor používá nestandardní procesní diagramy (Obrázek 3.1 a Obrázek 4.1) a neodkazuje k nim v textu. Dále např. výsledky experimentu v Obrázku 7.1 zobrazuje nevhodným způsobem – v grafu je evidentně jedna hodnota, kterou lze označit za „outliera“, a vzhledem k tomu, že ji autor v datech ponechal a nepoužil vhodnější způsob zobrazení, je hodnota sdělení daného grafu téměř nulová.	

Zvolený postup řešení	částečně vhodný
<i>Posuďte, zda student zvolil správný postup nebo metody řešení.</i>	
Hlavním tématem práce je zrychlení metody InLoc. Z práce jsem pochopil, že autor si jako míru rychlosti běhu určil celkový čas běhu. Dle mého názoru se jedná o hrubou chybu, protože I/O dotazy, např. dotazy do databáze, nejsou pro zrychlení metody relevantní. Pouhé zmenšení počtu dotazů do databáze razantně snižuje celkovou dobu běhu, ale hlavně, počty I/O dotazů by měly být, dle mého názoru, identické pro zrychlenou i nezrychlenou verzi.	

Odborná úroveň	E - dostatečně
<i>Posuďte úroveň odbornosti závěrečné práce, využití znalostí získaných studiem a z odborné literatury, využití podkladů a dat získaných z praxe.</i>	
Vzhledem ke komentářům, které uvádím výše, hodnotím odbornou úroveň jako dostatečnou.	

Formální a jazyková úroveň, rozsah práce	E - dostatečně
<i>Posuďte správnost používání formálních zápisů obsažených v práci. Posuďte typografickou a jazykovou stránku.</i>	
Text práce je nakonec pochopitelný, ovšem chybí odkazy k Tabulkám a Obrázkům, takže čtenář musí odhadovat, jaký obrázek, či jakou tabulku, má autor na mysli, když v textu skutečnosti uvedené v tabulkách a obrázcích popisuje. Textu by pomohlo, kdyby jej autor před odevzdáním poskytl někomu, kdo dané problematice vůbec nerozumí – zjistil by, že takový člověk nemá šanci práci pochopit. Některé „chyby“ jsou spíše úsměvné, jako např. na straně 81, „Úprava implementace zahrnovala: 1. Oddělení role databázového a dotazového snímku: [...] 2. Není nutné ukládat mezivýsledky do souboru. [...] 3. Je možno vypnout možnost načítání mezivýsledku ze souboru [...]“, jiné, jako např. první věta na straně 4: „Lokalizace ze snímku má jen dva vstupy, tj. dotazový snímek a mapu prostředí reprezentovanou databází RGB-D snímku, pro než je póza kamery známá,“ čtenáře zcela dezorientují.	

Výběr zdrojů, korektnost citací

D - uspokojivě

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.

Autor sice cituje metody a postupy, které používá, ale např. algoritmus DEGENSAC, deskriptory DSIFT, či plně konvoluční síť NetVLAD zmiňuje již úvodní kapitole, kdežto citace se metody dočkají až mnohem později.

Další komentáře a hodnocení

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.

Autor na straně 34 uvádí, že CI/CD nepoužil, protože kód překládá MATLAB. MATLAB ale překládá pouze soubor _MEX.cpp. Ostatní je možné kompilovat a testovat v gitlabu v automatizační pipeline. Dále není kompilace MATLABem důvod pro vynechání statické analýzy kódu, též implementovatelné v GL pipeline, ani pro vynechání automatických integračních testů, o kterých víme, že je autor provádí, protože je popisuje na straně 35. Absence standardního kompilačního a testovacího prostředí naznačuje, že autor tyto nezná a používal pouze MS Visual Studio. Projekt neobsahuje standardní kompilační pomocníky, např. Makefile. Konkrétní návod, jak knihovnu zkompilovat, každopádně v repozitáři chybí pro jakýkoliv systém, i přes to, že obecný postup kompilace MEX je popsán na straně 27.

III. CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

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

Práci hodnotím jako dostatečnou zejména proto, že její odborná část se dle mého názoru nachází na hranici uznatelného. Autor si vybral čistě implementační zadání, které neodpovídá jeho specializaci, a vzhledem k absenci znalostí základů softwarového vývoje a jeho standardních nástrojů vytvořil práci, které je po odborné stránce na úrovni absolventa oboru softwarového vývoje bakalářského studijního programu. Autor optimalizoval čas běhu kódu, přičemž pro měření použil celkové časy běhu, tzn. časy včetně I/O dotazů – počítání I/O dotazů do výsledných časů je dle mého názoru hrubou chybou, protože pouhé zmenšení počtu dotazů razantně snižuje celkovou dobu běhu programu. O autorově odbornosti mě dále přesvědčuje absence standardní dokumentace kódu, která se ovšem částečně nachází rozházená v textu práce samotné, míchání norem pojmenovávání proměnných, absence standardních kompilačních a testovacích nástrojů a nevyužití automatizačních GitLab pipeline – ne nutně pro kompilaci MATLAB MEX, ale alespoň pro statickou analýzu a unit testy.

Otázky:

- 1) Je pravdou, že počet dotazů do databáze má zásadní vliv na rychlost běhu algoritmu. Ovšem míra zrychlení běhu snížením počtu I/O dotazů bude stejná u všech verzí algoritmu, a pokud se tedy v kapitole 7.4 zabýváte měřením času běhu jednotlivých funkcí, proč jste do času běhu programu započítal i I/O dotazy?
- 2) Proč jste v repozitáři kódu nepřipravil základní dokumentaci, např. v souboru README.md, a skript pro jednoduchou kompilaci pro MATLAB MEX?

Předloženou závěrečnou práci hodnotím klasifikačním stupněm **E - dostatečně**.

Datum: 5.1.2022

Podpis: