

# Posudek oponenta závěrečné práce

Název práce: Strojové učení pro robotickou exploraci  
Jméno autora: Jáchym Staněk  
Typ práce: Bakalářská  
Ústav: ČVUT v Praze  
Fakulta: Fakulta elektrotechnická  
Katedra: Katedra kybernetiky  
Oponent práce: Ing. Pavel Petráček, ČVUT v Praze, Katedra kybernetiky

## Náročnost zadání

Zadání popisuje problém korekce hloubkových dat jakožto podproblém přesné lokalizace a mapování při robotické exploraci. Zadání navrhuje detailní postup k metodologickému přístupu k řešení problému. Zadání neobsahuje požadavek na rešerši odborné literatury a hodnotím jej i proto jako **průměrně náročné**.

## Splnění zadání

Zadání hodnotím jako **splněné**.

## Zvolený postup řešení

Postup řešení byl již velmi detailně specifikován v zadání práce. Studentovo řešení tento postup velmi úzce sledovalo a považuji jej tak za **správný**, avšak dle mého názoru mohlo obsahovat více vlastní realizace studenta. Prostor pro takovou realizaci by se našel například v rozšíření modelu o penalizaci paprskových artefaktů v hloubkových datech, jejichž filtrace je velmi krátce zmíněna v kapitole C.1 a je vyhodnocena jako hlavní faktor zhoršující přesnost lokalizace při využití modelu k doplnění dat.

## Formální a jazyková úroveň, rozsah práce

Práce je psána čitelnou angličtinou, avšak obsahuje velký počet chyb, které místami brání porozumění textu. Anglicky psaný text obsahuje systematické chyby. Příkladem je psaní *"e.g. text"* a *"i.e. text"*, jež se správně píše s čárkou jako *"e.g., text"* a *"i.e., text"*.

Všechny rovnice jsou psány špatně. Nejsou psány jako součástí věty a neobsahují číslování. Většina obrázků je formálně sazena správně, avšak některé nesprávně obsahují redundantní titul (Obr. 1.1, Obr. 2.3, Obr. 4.2, všechny obrázky v Dop. B), často s výrazně nepasující velikostí použitého fontu. Tabulky s výsledky v Kap. 5 mi přijdou nešťastné a velmi matoucí. Doporučil bych tabulky rozdělit na jednotlivé části a tyto části lépe popsat v popisku i v textu. Data v grafech v Dop. B bych doporučil oříznout pouze na jejich doménu bez zbytečného přesahu vně.

Referencování obrázků, tabulek a sekcí je nekonzistentní a ve většině části práce špatně. Jedním z mnoha příkladů je referencování obrázků na str. 14 (*"from the pictures C.6A and C.6b"* namísto *"in Figure C.6"* případně *"in Fig. C.6"*). Dalším příkladem jsou chybějící názvy reference na str. 15 (*"using the MSE loss (see 4.2.1)"* namísto *"using the MSE loss (see Section 4.2.1)"*) a nebo na str. 17 (*"in the pipeline C.2"* namísto *"in the pipeline shown in Figure C.2"*). Obrázek 4.3 není v textu referencován vůbec.

Struktura práce je místami matoucí. Kapitola 4. obsahuje jak popis datasetů, použité metriky a balíčky třetích stran, tak ale i popis navržených experimentů a parametrů použitých metod. Navržení experimentů

pasuje mimo tuto kapitolu a zasloužilo by si vlastní kapitolu, jelikož se takto velmi dobře schovalo mezi zbytek textu. Zbytek kapitoly by pak patřil do úvodu pod kapitolu běžně známou jako *Preliminaries*. Sazení grafů do Doplnkových kapitol mi přijde taktéž nešťastné. Za lepší řešení bych považoval jejich přesun do příslušných kapitol a jejich patřičné referencování a komentování.

Rozsah práce je pro bakalářskou práci v pořádku. Kritérium formální a jazykové úrovně hodnotím stupněm **D - uspokojivě**.

## Odborná úroveň

Student naimplementoval jak specifikovalo zadání. Práce využívá publikovaný model pro doplňování hloubkových dat a tento model trénuje na veřejně přístupném datasetu pomocí dvou ztrátových funkcí, z nichž *Chamfer loss* je pravděpodobně pro tuto úlohu použita poprvé. Výstup modelu poté student využívá jako vstup do vedoucím-poskytnutého SLAM přístupu a porovnává jeho výstup s ground truth daty. Z kódu ani z historie veřejného git repozitáře však není přesně jasné, kterou část práce udělal student sám a co je prací vedoucího práce. K odborné úrovni vyčísluji své poznámky níže.

1. Popis algoritmu ICP na straně 7 ukazuje na nepochopení algoritmu samotného. Takto psáno nemůže algoritmus fungovat, jelikož je popsána pouze jedna iterace a popis na řádcích 2-5 (navíc s neinicializovanými  $R'$  a  $t'$  uvnitř sumace na řádku 4.) nedokverguje, pokud se tedy optimum nenachází v nulovém posunu a jednotkové rotaci, kde pak ale využití ICP postrádá smysl. V návaznosti na špatný popis chybí pro ICP typické ukončovací kritérium konvergence.
2. Práce debatuje nad využitím strojové učení v doplnění hloubkových dat a jejich následné využití v úloze SLAM. Dosažené výsledky v Kap. 5.2 ukazují, že implementované řešení dosahuje horších výsledků, nežli využití pouze řídkých dat. V Kap. 6 je diskutováno, že tento výsledek je pravděpodobně způsoben modelem generujícím nedeterministické paprsky (viz. Obr. C.4), jež lze považovat za šum. Pokus o jejich filtraci pomocí bilaterálního filtru a nebo oříznutí vzhledem ke vzdálenosti od zdroje paprsku je neúčinné. Na základě Obr. C.4 však pochybuji o správnosti využití bilaterálního filtru při filtraci. I přesto, že pravá část obrázku je nezarovnána a chybí popis barevného schéma (domýšlím si, že zbarvení naznačuje vzdálenost od zdroje ve vertikální ose), je v pravé části obrázku jasně vidět, že použitý filtr data přidává, namísto aby je odebíral nebo vyhlazoval. Tato skutečnost mi napovídá o špatné aplikaci/implementaci filtru jako takového. Jako specifikovaný zdroj špatných výsledků bych čekal širší analýzu řešení tohoto problému, například pomocí lokálního průměrování a nebo využití minmax filtrace.
3. V práci mi chybí porovnání metody založené na strojovém učení s dobře prostudovanými metodami pro interpolaci dat. Velkým přínosem práce by bylo využití interpolačních technik (lineární, kvadratické, atd.) pro doplnění dat a jejich porovnání s navrženým přístupem. Obzvláště, když výstup generativního modelu je zašuměný, jak je prezentováno.
4. Délka učení v Dop. B je krátká a z grafů lze vyčíst, že hodnota trénovací ztrátové funkce v případě *Chamfer loss* ještě nedokvergovala. Jaký byl důvod k zastavení učení?
5. V číslem nespecifikované rovnici na straně 22 je definována chyba lokalizace jako  $L = d + \alpha_x$ , kde  $d$  je posun v pozici a  $\alpha_x$  je jakýsi úhel rotace. Jednotky nejsou specifikovány a domýšlím si tak, že se v rámci této metriky sčítají metry a radiány. Taktéž není definováno o jaký úhel rotace se jedná. Velmi pochybuji o správnosti této funkce při zpětném učení modelu strojového učení.
6. Dosažené výsledky v Kap. 5.2 vykazují velké rozdíly při využití plných a řídkých dat. Doporučil bych navržené experimentů a nebo využití normalizačních technik tak, aby šly použité metody porovnat. V prezentovaném kontextu jsou výsledky přístupů neporovnatelné, jelikož produkují podezřele stejnou *localization accuracy* (viz. mé pochybnosti v předchozím bodě) při velmi rozdílných MAE a RMSE.

Porovnání s publikovanými daty v Tab. 5.3 je irelevantní, jelikož proběhlo na jiných datech, jak je zmíněno v textu.

Odbornou úroveň práce hodnotím stupněm **D - uspokojivě**.

## Výběr zdrojů, korektnost citací

Relevantní práce v oblasti jsou citovány a výběr zdrojů je korektní. Více než polovina referencí je však citováno špatně — 8 referencí vůbec neobsahuje informaci o tom, kde byl článek publikován. Některé časopisecké citace obsahují stránky, jiné zase ne. Stejně tak u konferenčních příspěvků. Online reference neobsahují informaci o tom, že jsou online a ani o datu návštěvy webu. Výběr zdrojů a korektnost citací hodnotím stupněm **E - dostatečně**.

## CELKOVÉ HODNOCENÍ, OTÁZKY K OBHAJOBĚ, NÁVRH KLASIFIKACE

Předloženou práci hodnotím stupněm **D - uspokojivě**.

## Doplňující otázky

1. Vysvětlete rozdíly a přínosy metod *differential SLAM* oproti běžným přístupům k problému *SLAM*.
2. Popište princip algoritmu *Adam Optimizer* a vysvětlete rozdíl varianty *AdamW Optimizer* oproti variantě *Adam*. Variantu *AdamW* jste použil ve svém navrženém řešení. Jakým způsobem jste volil hodnoty hyperparametrů *learning rate* a *weight decay*?
3. Popište princip základní verze algoritmu *Iterative Closest Point*.

V Praze 30. května 2022

Pavel Petráček