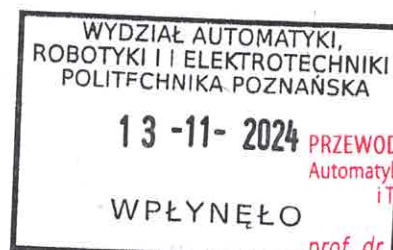


Szczecin, 5.11.2024 r.

prof. dr hab. inż. Krzysztof Okarma
Katedra Przetwarzania Sygnałów i Inżynierii Multimedialnej
Wydział Elektryczny
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie



PRZEWODNICZĄCY RADY DISCYPLIN
Automatyka, Elektronika, Elektrotechnika
i Technologie Kosmiczne
prof. dr hab. inż. Wojciech Szela

RECENZJA ROZPRAWY DOKTORSKIEJ
dla Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne
Politechniki Poznańskiej

opracowana na podstawie uchwały nr 02/2024-2028 Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Poznańskiej z dnia 17.09.2024 r. oraz pisma Przewodniczącego Rady prof. dr. hab. inż. Wojciecha Szela z dnia 19.09.2024 r.

Tytuł rozprawy: **Metody doboru struktury oprogramowania i optymalizacji parametrów w zadaniu lokalizacji wizyjnej na podstawie danych RGB-D**

Autor rozprawy: **mgr inż. Aleksander Kostusiak**

Dyscyplina naukowa: **automatyka, elektronika, elektrotechnika i technologie kosmiczne**

Promotor: **prof. dr hab. inż. Piotr Skrzypczyński**

I. TEMATYKA, TEZA NAUKOWA I CEL ROZPRAWY

Rozprawa doktorska mgr. inż. Aleksandra Kostusiaka pt. „Metody doboru struktury oprogramowania i optymalizacji parametrów w zadaniu lokalizacji wizyjnej na podstawie danych RGB-D” dotyczy zagadnień związanych z robotyką mobilną, widzeniem maszynowym oraz sztuczną inteligencją. Bardzo dobrze wpisuje się zatem, zarówno w dawną dyscyplinę naukową *automatyka i robotyka*, w której otwarty został przewód doktorski, którego jest zwieńczeniem, jak również w aktualnie obowiązującą dyscyplinę *automatyka, elektronika, elektrotechnika i technologie kosmiczne*. Zawiera ona również pewne elementy interdyscyplinarności związane z *informatyką techniczną i telekomunikacją*, co stanowi jej dodatkową zaletę oraz potwierdzenie bliskości obu tych dyscyplin, szczególnie w zakresie zastosowań metod analizy obrazów w robotyce czy też w zagadnieniach dotyczących rozwijającej się dynamicznie branży *automotive*. Tematykę rozprawy bez wątpienia uznać należy za aktualną, ściśle związaną z rozwojem nauki i techniki w tym zakresie. Postawiona hipoteza badawcza, wraz z dwiema hipotezami pomocniczymi, dobrze określa problem naukowy, którego rozwiązania podjął się Kandydat. Wyprzedzając nieco konkluzję, stwierdzić można, iż zostały one w pracy potwierdzone a cel pracy, który nie został co prawda jawnie sformułowany we wprowadzeniu, został zrealizowany.

Motywacja do wyboru tematyki pracy przedstawiona przez Doktoranta nie budzi zastrzeżeń. Słusznie podkreśla On mnogość rozwiązań technicznych, czy też różnego typu sensorów, które mogą być zastosowane w zagadnieniu lokalizacji wizyjnej robotów mobilnych, skupiając się w pracy głównie na systemach korzystających z odometrii wizyjnej (ang. *Visual Odometry* – VO) oraz jednoczesnej lokalizacji i mapowania przestrzeni (ang. *Simultaneous Localization and Mapping* – SLAM). Tematyka ta jest również w pełni zgodna z dorobkiem oraz zainteresowaniami naukowymi Promotora rozprawy.

Przedstawiona do recenzji rozprawa została poparta dwoma wymienionymi w bibliografii publikacjami naukowymi [163, 164], jednak zerknięcie choćby do bazy dblp pozwala znaleźć 4 kolejne pozycje autorstwa lub współautorstwa Doktoranta, ściśle związane z tematyką rozprawy. W bazie Scopus publikacji takich jest 8, co dobrze świadczy zarówno o wyborze tematyki badawczej, jak również o właściwym podejściu do publikowania na bieżąco cząstkowych wyników badań, m.in. w celu poddania ich niezależnej weryfikacji przez recenzentów zgłaszanych prac.

II. ZAWARTOŚĆ MERYTORYCZNA ROZPRAWY

Recenzowana rozprawa została podzielona na 7 zasadniczych rozdziałów, zawiera również spis symboli i skrótów stosowanych w pracy, a także bogatą bibliografię liczącą 248 pozycji, wśród których można jednakże zauważyć aż 89 źródeł w postaci stron internetowych, które nie zostały w żaden sposób opisane – Doktorant ograniczył się jedynie do podania adresów poszczególnych stron. Dla niektórych źródeł stanowiących publikacje naukowe, w ich opisie podane zostały zarówno odnośniki w postaci numerów DOI, jak również adresy URL, co wydaje się zbędne. Tym niemniej zauważyć należy, iż pozycje bibliograficzne zostały dobrane trafnie i są one ściśle związane z tematyką rozprawy.

We wstępnym rozdziale rozprawy Doktorant przedstawił motywację do podjęcia badań dotyczących systemu lokalizacji wizyjnej na podstawie danych RGB oraz informacji o głębokości. Sformułował również hipotezy badawcze oraz przedstawił plan badań.

W kolejnym rozdziale przedstawiony został przegląd rozwiązań z zakresu lokalizacji wizyjnej w robotyce, w tym metod pomiaru odległości z użyciem pasywnych oraz aktywnych sensorów wraz z zagadnieniami kalibracji kamer oraz technik stereoskopowych. Przedstawiono metody lokalizacji na podstawie znaczników oraz pełnej mapy otoczenia trójwymiarowego, a także możliwości zastosowania metod głębokiego uczenia. Następnie Doktorant omówił metody odometrii wizyjnej oraz różne warianty SLAM stosowane w przypadku braku znajomości mapy otoczenia 3D. Ich opis jest bardziej szczegółowy, co w oczywisty sposób wynika z tematyki pracy. W końcowej części rozdziału Kandydat przedstawił dwa sposoby uzupełniania brakujących danych (*inpainting*) w kontekście ich zastosowania do map głębokości (algorytmy Telei oraz Naviera-Stokesa), jak również podejście RANSAC, typowo wykorzystywane m.in. w stereowizji.

Rozdział trzeci dotyczy opisu systemu odometrii wizyjnej bazującego na danych pomiarowych uzyskanych z sensorów Kinect. Przedstawiono w nim przegląd metod detekcji cech oraz klasycznych deskryptorów (SURF, ORB, BRISK, KAZE, AKAZE), a także sieć SuperPoint, wraz z omówieniem metod dopasowania par deskryptorów oraz filtracji odstających cech stanowiących nieprawidłowe dopasowania. Rozdział ten kończy omówienie algorytmów służących do estymacji trajektorii ruchu robota oraz opis struktury opracowanego systemu odometrii wizyjnej.

Następna część rozprawy dotyczy metod optymalizacyjnych zastosowanych do poszukiwania parametrów systemu lokalizacji robota, w szczególności algorytmu ewolucyjnego oraz metody roju cząstek

(ang. *Particle Swarm Optimization* - PSO), a także ich analizy porównawczej. Ostatnie dwa podrozdziały są jednak bardzo krótkie, co budzi pewien niedosyt, zwłaszcza iż miary dopasowania (błędy bezwzględne ATE oraz względne RPE) zostały szczegółowo opisane w podrozdziale 6.5. W tym kontekście eksponowanie tego krótkiego podrozdziału w tym miejscu pracy budzi wątpliwości.

Rozdział 5. stanowi z kolei omówienie wybranych metod uczenia maszynowego, które Autor zastosował w swoich badaniach. Pierwsza część tego rozdziału przedstawia najważniejsze pojęcia z tego zakresu, jednak została ona podzielona na aż 14, w większości bardzo krótkich, sekcji. Kolejny podrozdział, dotyczący analizy dostępnych platform programistycznych oraz wyboru najbardziej adekwatnej spośród nich, zawiera z kolei 9 krótkich sekcji. Trzecia część tego rozdziału, traktująca o zastosowaniu sieci neuronowych w celu estymacji głębi, została z kolei podzielona na część ogólną (bez numeru) oraz wydzieloną sekcję dotyczącą jednej z sieci, mianowicie Monodepth. W tym miejscu zabrakło niestety porównania z innymi, nieco nowszymi rozwiązaniami, które mogłyby być zastosowane w tym samym celu, jak chociażby dwa modele udostępnione w ubiegłym roku na licencji open source przez firmę Intel – VI-Depth oraz MiDaS 3.1. Doktorant wskazuje co prawda na dostępny przegląd metod w publikacjach [158, 175], jednak są to publikacje z roku 2020 oraz 2022, a postęp technologiczny w tym zakresie jest bardzo szybki i nowe rozwiązania pojawiają się z miesiąca na miesiąc.

Kolejny rozdział pracy dotyczy przeprowadzonych badań eksperymentalnych. Kandydat przedstawił w nim zarówno metodologię trenowania sieci neuronowej przy użyciu platformy programistycznej fastai (zgodnie z oryginalną notacją nazwa ta powinna być pisana wyłącznie małymi literami, chociaż Doktorant używa w rozprawie różnych wersji notacji), jak też sposób przygotowania zestawów danych do jej uczenia wraz z uwagami na temat augmentacji danych. Przedstawione zostały także publicznie dostępne zestawy danych użyte w trakcie eksperymentów, takie jak TUM RGB-D przygotowany w Technische Universität München, czy też PUTKK (PUT Kinect 1 & Kinect 2 data set) oraz Messor II pochodzące z macierzystego wydziału Politechniki Poznańskiej (przygotowane przez zespół kierowany przez Promotora rozprawy). Kandydat omówił także użyte miary błędów (choć jak wspomniano, równie dobrze mogłyby one być bardziej szczegółowo przedstawione wcześniej, tj. w końcowej części rozdziału 4.). Podrozdział 6.6, o niezbyt trafnie dobranym tytule „Badania”, zawiera opisy eksperymentów przeprowadzonych z użyciem wcześniej opisanych testowych sekwencji wideo oraz ich wyniki. Dotyczyły one porównania par detektor-deskryptor, metod n -punktowych, a także analizy metod RGB-D SLAM. Ponadto w podrozdziale tym znalazły się opisy badań dotyczących użycia metod populacyjnych do poszukiwania optymalnych parametrów prostego systemu VO RGB-D. Badania dotyczyły zarówno parametrów algorytmu RANSAC, jak też zastosowanego w tym systemie detektora AKAZE w zagadnieniu offline. Dopełnieniem tej części tego obszernego podrozdziału (od str. 132 aż do str. 214), który równie dobrze mógłby stanowić odrębny rozdział rozprawy, jest analiza statystyczna bazująca na histogramach błędów translacji i rotacji. Następnie Kandydat przedstawił badania dotyczące analogicznego poszukiwania optymalnych parametrów systemu VO w trybie online, a także ciekawą koncepcję użycia metod *inpaintingu* w odniesieniu do mapy głębi – zarówno w odniesieniu do metod klasycznych, jak również z użyciem sztucznej sieci neuronowej. Przedstawione w tej części pracy wyniki oraz ich analizę uznaję za najbardziej wartościowy merytorycznie element rozprawy. W szczególności dotyczy to obserwacji związanych z uzupełnianiem danych różnymi metodami, gdzie można było zaobserwować efekt określany często jako halucynacja sieci neuronowej, w tym przypadku objawiający się np. pojawianiem się dodatkowych obiektów (np. nieistniejących nóg krzesła) utrudniających znalezienie właściwych punktów kluczowych niezbędnych do poprawnego wyznaczenia roto-translacji. Na końcu rozdziału szóstego – najobszerniejszego z całej rozprawy – Kandydat umieścił krótki podrozdział dotyczący oceny czasu wykonywanych obliczeń algorytmów ewolucyjnych oraz uzupełniania map głębi.

Ostatni rozdział rozprawy zawiera podsumowanie przeprowadzonych badań eksperymentalnych oraz wnioski z nich płynące. Przedstawiono także możliwości dalszych badań związanych z przedstawioną koncepcją. Nakreślone kierunki należy uznać za właściwe a recenzowaną rozprawę za opracowanie podsumowujące pewien ich etap. Biorąc pod uwagę ciągły rozwój metod uczenia maszynowego, pojawiające się wciąż nowe modele sieci neuronowych, a także zmieniające się platformy programistyczne, tego rodzaju badania prowadzone w ramach przewodu doktorskiego w naturalny sposób powinny się zakończyć w pewnym momencie, co w żaden sposób nie wyklucza ich kontynuacji i rozwijania opracowanego systemu po obronie rozprawy doktorskiej.

III. OGÓLNA OCENA ROZPRAWY I UWAGI DYSKUSYJNE

Rozprawa doktorska mgr. inż. Aleksandra Kostusiaka stanowi interesujące opracowanie zagadnienia o znacznym potencjale aplikacyjnym z zakresu dynamicznie rozwijającego się obszaru robotyki mobilnej oraz systemów wizyjnych. Wiele z przedstawionych w rozprawie koncepcji oraz metod stanowią znane algorytmy i rozwiązania, często posiadające dostępne implementacje, jednak nie obniża to wartości samej pracy. Jej Autor wykazał się bardzo dobrą orientacją w tematyce rozprawy, jak również znajomością dostępnych rozwiązań, ich specyfiki, uwarunkowań i możliwości zastosowania. W umiejętny sposób dobrał dostępne rozwiązania software'owe, uwzględniając potrzeby i ograniczenia, także sprzętowe, rozważanego systemu lokalizacji wizyjnej na podstawie danych RGB-D. Niewątpliwie integracja różnych metod oraz implementacja ich we wspólnym systemie wymagała nie tylko dobrej orientacji w niuansach z nimi związanych, ale także znacznego nakładu pracy. Uzyskane wyniki oraz wynikające z nich wnioski są interesujące i nie zawsze oczywiste. Zaproponowane metody doboru struktury oprogramowania, jak również optymalizacji parametrów systemu lokalizacji wizyjnej, zgodnie z tytułem rozprawy, stanowią o jej wartości merytorycznej. Dodatkowym atutem pracy jest fakt zweryfikowania wszystkich zaproponowanych metod oraz algorytmów z użyciem rzeczywistych danych benchmarkowych uzyskanych na podstawie zapisu z kamer zamontowanych na robotach mobilnych, ewentualnie z kamery RGB-D trzymanej w ręku.

Do najważniejszych osiągnięć Doktoranta zaliczyć można dokonanie rzetelnej analizy dostępnych metod oraz możliwości ich zastosowania w postawionym zadaniu, dobór odpowiednich narzędzi oraz rozwiązań, jak również dokonanie optymalizacji parametrów systemu odometrii wizyjnej oraz weryfikację jego działania dla różnych zestawów testowych sekwencji wideo wraz z dostępnymi mapami głębi. Ważnym osiągnięciem jest zbadanie możliwości użycia metod populacyjnych w celu optymalizacji parametrów metody RANSAC oraz detektora AKAZE, zarówno w zagadnieniu offline, jak też online. W mojej opinii na szczególną uwagę zasługują także przeprowadzone badania metod uzupełniania mapy głębi oraz ich wpływu na dokładność lokalizacji robota mobilnego. Dodatkowym potwierdzeniem wartości merytorycznej uzyskanych rezultatów są publikacje w czasopiśmie naukowych oraz w materiałach konferencyjnych indeksowanych w bazach Scopus, a także dblp, w których publikowane były częściowe wyniki badań eksperymentalnych.

Jak wspomniano, szczególnie ciekawym elementem pracy jest koncepcja użycia metod uzupełniania brakujących danych (*inpaintingu*) w odniesieniu do mapy głębi oraz związane z nią badania eksperymentalne, z których wynika interesujący wniosek dotyczący – często negatywnego – wpływu sieci neuronowej na uzyskiwane wyniki, utrudniające późniejsze dopasowywanie cech wynikające z obecności fałszywych obiektów (zjawisko tzw. halucynacji). Czy w zaistniałej sytuacji nie byłoby celowe zastanowienie



się nad możliwością zastosowania innych metod *inpaintingu* lub ewentualnie opracowania autorskiego rozwiązania dedykowanego dla map głębi?

W trakcie lektury rozprawy pojawiły się również pytania o charakterze dyskusyjnym oraz wątpliwości, których wyjaśnienie uważam za celowe. Część z nich zapewne wynika w dużej mierze z pewnej niestaranności przy pisaniu pracy, co prawdopodobnie jest efektem zbliżającego się ustawowego terminu zamknięcia przewodów doktorskich. Tym niemniej, wielu usterek można byłoby uniknąć, poświęcając nieco dodatkowego czasu na uważną ponowną lekturę rozprawy, zaś w przypadku konieczności naniesienia poprawek uniemożliwiających jej złożenie i obronę w wymaganym terminie, istniałaby przecież możliwość wszczęcia postępowania doktorskiego w „nowym” trybie przewidzianym w aktualnie obowiązującej Ustawie *Prawo o szkolnictwie wyższym i nauce*.

Do uwag o charakterze krytycznym i dyskusyjnym zaliczyłbym następujące spostrzeżenia:

- nie jest jasne stwierdzenie na str. 13 „System GPS zapewnia wysoką precyzję”, zwłaszcza w kontekście robotyki mobilnej, dla której zwykle oczekiwana jest większa dokładność lokalizacji aniżeli zapewniona przez systemy nawigacji satelitarnej; Autor wskazuje co prawda wady systemów GPS, ale nie precyzuje stwierdzenia dotyczącego błędów lokalizacji; podobne stwierdzenie („dokładne ustalanie pozycji za pomocą systemu GPS”) znajduje się na str. 31 – należałoby w tym kontekście podać tę dokładność;
- stwierdzenie ze str. 14 „podjęto próbę wykorzystania nowego, lepszego sensora do ulepszenia danych otrzymanych z starszego sensora za pomocą klasycznych metod oraz sztucznych sieci neuronowych” zawiera pewien skrót myślowy, który wymaga doprecyzowania – można się jedynie domyślić, że dane z nowszego sensora Kinect v2 mają charakter referencyjny;
- wątpliwości budzi stwierdzenie na str. 30 „Taki sposób działania został zapewne nie tylko dogłębnie sprawdzony przez naturę, ale i we wszelaki sposób podważany i kwestionowany poprzez alternatywne podejścia postrzegania rzeczywistości”;
- wyjaśnienia wymaga stwierdzenie dotyczące podziału robotów mobilnych na autonomiczne oraz nieautonomiczne na str. 30, gdzie Doktorant stwierdził, iż roboty autonomiczne cechują się brakiem ograniczenia ruchu, podczas gdy typowa definicja autonomiczności wiąże się z możliwością samodzielnego sterowania swoim położeniem bez ingerencji człowieka a nie tylko brakiem obecności przewodów zasilających czy sterujących (bezprzewodowość to nie zawsze autonomiczność, o czym świadczą chociażby drony sterowane przez człowieka za pomocą bezprzewodowego kontrolera);
- nie jest jasne, co Doktorant rozumie przez „rozmywanie o zmiennym przewodzeniu” – sformułowanie takie znajduje się na str. 69 w opisie deskryptora KAZE;
- w pracy Doktorant stosuje nieprawidłowe nazewnictwo dekompozycji SVD (str. 82) – nie jest to rozkład na wartości własne, ale rozkład względem wartości szczególnych;
- wątpliwości budzi stwierdzenie zawarte na str. 88 „Po znalezieniu najodpowiedniej nastawy, pozwalającego uzyskać najlepsze wyniki, takie poszukiwania mogą być dalej poprawiane w jego okolicy” – pomijając niespójność gramatyczną, czy oznacza to, że uzyskane wyniki jednak nie były najlepsze?
- stosowane w sieciach neuronowych pojęcie „momentum” kojarzy się m.in. z fizycznym pojęciem pędu – z lakonicznego opisu nie wynika, czy można stosować taką interpretację także w tym przypadku;
- nie wiadomo, co oznacza użycie wartości 10-20 mniejszej (str. 101) – podobnie „2-3 mniejsza” – nie jest jasne czy chodzi o wartość 10-krotnie mniejszą, czy mniejszą o 10 jednostek;

- wzór (5.6) jest błędny – w przypadku wskaźnika SSIM obliczany powinien być iloczyn trzech czynników określających składowe dotyczące zgodności luminancji, kontrastu oraz struktury, podczas gdy Kandydat podaje, iż są one sumowane; konieczne jest wyjaśnienie, w jaki sposób zaimplementowany został ten wskaźnik w procesie trenowania sieci i czy w tym kontekście uzyskiwane wyniki potwierdzają krytykę wskaźnika zawartą w sekcji 5.1.13 – zamieszczona uwaga dotycząca kierowania uczenia sieci w złym kierunku może przecież dotyczyć każdego wskaźnika, choć z wieloma argumentami nie sposób się nie zgodzić; nieuzasadnione jest także stwierdzenie, iż indeks przyjmuje wartości z zakresu (0, 1], choć zwykle faktycznie tak jest (ostatni czynnik może przecież przyjmować wartości ujemne);
- trudno zgodzić się z uwagą dotyczącą wskaźnika SSIM dotyczącą konieczności konwersji obrazu kolorowego do skali szarości – z jednej strony istnieje wiele modyfikacji wskaźnika SSIM, także dla obrazów kolorowych, a z drugiej strony wiele innych wskaźników jakości obrazu też korzysta przede wszystkim, a często nawet wyłącznie, z informacji o jasności pikseli bez użycia koloru;
- w pracy brak odniesienia się do modelu MiDaS udostępnionego przez firmę Intel w 2023 r., który służyć może do podobnych celów jak Monodepth – celowe wydaje się dokonania analizy możliwości zastosowania tego rozwiązania, zamiast lub jako uzupełnienie dotychczasowego;
- nie jest jasne, co oznacza stwierdzenie na str. 119 „W badaniach dokonano ręcznego podziału warstw sieci Monodepth na grupy, tak że każda grupa zawierała jedną warstwę”;
- zasadne byłoby wskazanie motywacji podziału zbioru danych w proporcji 90%/10% dla sieci neuronowej, wskazane na str. 122, jak też analiza przypadku innego podziału np. bardziej typowego 80%/20%;
- wniosek dotyczący sieci Monodepth zawarty na str. 192 związany z koniecznością douczania na docelowym zestawie danych jest dość typowy dla rozwiązań bazujących na CNN, z kolei na str. 205 Autor wskazuje na przeuczenie sieci, a na str. 214 stwierdza, iż „Być może zwiększenie modelu mogłoby [...] pomóc” – czy przy zwiększeniu modelu nie wystąpi w tej sytuacji niebezpieczeństwo zbyt dużego dopasowania do danych uczących (czyli właśnie przeuczenie)? Na str. 219 stwierdzono z kolei, iż „Rozważenie większego modelu sieci neuronowej mogłoby zapobiec nadmiernemu dopasowaniu do środowiska uczącego”, co miałoby na celu zwiększenie uniwersalności modelu i lepszą generalizację. Warto byłoby pokusić się o wyjaśnienie tego zagadnienia, także w kontekście rozwiązań z zakresu wyjaśnialnej sztucznej inteligencji (XAI), pożądaną także w robotyce mobilnej.

W końcowej części rozprawy można znaleźć trafną analizę popartą wysnutymi wnioskami, gdzie Kandydat dość „ostrożnie” stwierdził, że wnioski częściowe „sugerują prawdziwość tezy”, iż „zastosowanie populacyjnych metod optymalizacji oraz wybranych algorytmów uczenia maszynowego pozwala zoptymalizować strukturę i parametry systemu lokalizacji RGB-D, ze szczególnym uwzględnieniem lokalnego ruchu sensora”. Zapewne jest to spowodowane wątpliwościami, czy znaleziona struktura oraz parametry systemu będą faktycznie optymalne, choć niewątpliwie nastąpi ich poprawa. W mojej opinii, można jednak stwierdzić, iż wykonane badania eksperymentalne tę tezę potwierdzają, nawet jeśli w przyszłości pojawią się inne rozwiązania pozwalające na uzyskanie jeszcze lepszych rezultatów (biorąc pod uwagę postęp techniczny, będzie to raczej dotyczyło innego systemu lokalizacji RGB-D, także o innych uwarunkowaniach sprzętowych).

Całościową ocenę rozprawy obniża niestety znaczna liczba błędów edytorskich, typograficznych, językowych, brak spójności oznaczeń, liczne błędy interpunkcyjne, a nawet pojawiające się pojedyncze błędy ortograficzne. Niektóre spośród najbardziej rażących zostały wymienione w kolejnej części recenzji.

IV. UWAGI SZCZEGÓŁOWE

W pracy znaleźć można wyjątkowo dużą liczbę usterek edytorskich, czy też językowych. Poniżej wymieniono tylko niektóre spośród zauważonych, w tym te o rażącym charakterze:

- hipoteza badawcza na str. 15 jest zamieszczona dwukrotnie;
- zbyt często używana jest liczba mnoga w odniesieniu do wykonanych badań (np. „w naszych badaniach [...] skupiliśmy się” na str. 87, „nacisk położyliśmy” na str. 132, czy też „w naszych eksperymentach” na str. 145 i w wielu innych miejscach) – praca doktorska powinna być co do zasady pracą samodzielną, choć wykonywaną pod opieką promotora, a zatem znacznie lepszym rozwiązaniem jest użycie strony biernej (zresztą prawidłowo używanej w części tekstu pracy);
- w tekście rozprawy stanowczo zbyt często pojawiają się sformułowania o charakterze przypuszczającym, których powinno się unikać w pracy o charakterze naukowym; przykłady można znaleźć m.in. na str. 205, 210, 214 czy 215 („to co powoduje znaczne problemy, to prawdopodobnie maksymalne błędy RPE”, „może to wynikać”, „mogłoby w tym wypadku pomóc”, „zdaje się mieć przewagę”, „może to wynikać z faktu”), ponadto: „choć zdaje się, że i tak był on spowolniony” oraz „zdaje się, że ORB nie znajduje najlepszych punktów” na str. 138, słowo „ponoć” na str. 113, czy też wnioski na str. 145; podobna uwaga dotyczy sformułowań typu „udało się” (np. na str. 46 czy też str. 178), sugerujących przypadkowość osiągnięcia określonego celu;
- nie wszystkie rysunki zawierające anglojęzyczne elementy mają podane źródło bibliograficzne (np. rys. 2.13 lub 3.1) – skoro praca napisana jest w języku polskim, należy przypuszczać, iż pochodzą one z innych źródeł; z kolei na rys. 6.20 brak opisu znaczenia kolorów pomimo pozostawionej dużej ilości wolnego miejsca na str. 164 (Autor odsyła do głównego tekstu, gdzie wyjaśnienia tych oznaczeń trzeba szukać dwie strony wcześniej);
- opis deskryptora SURF w sekcji 3.2.1 pozostawia sporo do życzenia i zawiera dość liczne usterki, m.in. zapis „Hessianu” nie jest prawidłowy (w języku polskim powinno używać się sformułowań „hesjanu” lub „macierzy Hessego”);
- liczne potknięcia językowe dotyczące niewłaściwych końcówek wyrazów, prawdopodobnie powstałe częściowo podczas przeredagowywania niektórych fragmentów tekstu, ale też dotyczące nieprawidłowej odmiany wyrazów, np. nagminne stosowanie słowa „tą” zamiast „tę”: „tą fale” (str. 21), „tą linię” (str. 27), „tą samą scenę”, „tą zależność” (str. 29), czy „tą metodę” (str. 32), choć np. na str. 85 jest prawidłowo („tę operację”); a także inne błędy językowe np. na str. 169 („opartej o miarę” zamiast „opartej na mierze”);
- niewłaściwe rozpoczynanie zdań np. od słowa „Natomiast” (str. 14), „Czyli” (str. 55) lub „Przy czym” (str. 103), stosowanie pojęcia „ilość” zamiast „liczba” w odniesieniu do elementów policzalnych (np. na str. 88, 103, czy 210) lub odwrotnie (str. 113 – „spora liczba nowego kodu”);
- błędy ortograficzne, które w żadnym wypadku nie powinny występować w pracach doktorskich („macież” – str. 81, „zwarzone” – str. 105 oraz 106);
- liczne usterki typograficzne np. *Squar*, *StM* (str. 11), *dodatkowo* (str. 13), *technologi* (str. 23), *kamey* (str. 26), *metrologi* (str. 29), *opcjonanq* (str. 33), *Lowe’y* (str. 73), *benchamrkowe* (str. 85), *agorytmów* (str. 140), *ilnierów* (podpisy pod rysunkami 6.21–6.23) i inne; choć często są one trudne do wychwycenia przez słownik, np. gdy zamiast „m.in.” pojawia się „min.” (np. na str. 35);
- błędy o charakterze częściowo typograficznym a częściowo ortograficznym np. „dla tego trudno jest ocenić wpływ” (str. 14, podobnie na str. 23), „w okół punktu” (str. 55), „nie znanym obszarem” (str. 56), „NVidiom” (str. 109), „Wydłużyło by” (str. 121), czy „na przeciw” (str. 183);
- liczne błędy przestankowe, w szczególności brak przecinków lub ich lokowanie w niewłaściwych miejscach;
- niepotrzebne używanie łączników np. w wyrażeniach „radio-lokalizacji” (str. 19), „kilku-krotnie” (str. 102);
- stosowanie równoważników zdań, np. na początku str. 21, czy też na str. 72 (przedostatni akapit);
- zdarzają się zdania, w których pozostawiono zbędne fragmenty nie odpowiadające pozostałej ich części, np. „[...] , gdzie je dość ładnie i krótko opisano VO i SLAM” (str. 34) lub brakuje niektórych

wyrazów (np. „k losowo klatek” na str. 52); zdarzają się także niezręczne powtórzenia słów np. „Pełniejszy wgląd w wyniki daje wgląd w tab. 6.7” (str. 141);

- wzór, a właściwie zestaw wzorów (2.4), jest sformatowany w niezbyt czytelny sposób;
- pomiędzy wzorami na str. 28 warto byłoby podać także dość często używaną angielską nazwę macierzy zasadniczej (*essential matrix*);
- sformułowanie „metody douczania sieci pozwalają na jej efektywne uczenie do pracy w innym miejscu” (str. 33) nie jest właściwe, gdyż istotna jest w tym przypadku praca z innymi danymi lub ewentualnie w innym zastosowaniu – niekoniecznie w innym miejscu;
- zbyt częste użycie potocznych wyrazów, np. na str. 35 w jednym akapicie czterokrotnie użyto słów „zrobione/robi/zrobić”, także w odniesieniu do kamery czy biblioteki programistycznej, które wprost nie wykonują czynności „same z siebie”, czy też „robot będzie chciał dodać go do mapy” (na str. 40); podobnie np. sformułowanie „konieczność znania położenia” powinno być zastąpione frazą „konieczność znajomości położenia” (str. 37); razi też wyrażenie „uważany jest prawie że jako język sam w sobie” (str. 114);
- zdarzają się następujące po sobie zdania niezwiązane ze sobą w logiczny sposób, np. na str. 35 (ostatni akapit);
- w niektórych miejscach celowa byłaby zmiana kolejności cytowanych pozycji, np. zamiast [115,159,122] na str. 43 powinno być [115,122,159], podobna uwaga dotyczy np. str. 60;
- na rysunku 2.10 polskie znaki są nieprawidłowo kodowane, co skutkuje ich pominięciem („wierzchoków”, „mapy gbi”);
- sformułowanie „obraz czarno-biały” na str. 45 jest nieprecyzyjne, choć można domyślić się, iż chodzi o obraz monochromatyczny (w skali szarości), gdyż Doktorant pisze o jego wygładzeniu filtrem Gaussa; podobnym „skrótom myślowym” jest „losowanie za pomocą rozkładu normalnego” (str. 94), podczas gdy jest ono dokonywane za pomocą generatora o rozkładzie normalnym; podobna uwaga dotyczy „interpolacji najbliższego sąsiada” (str. 117);
- należy unikać sformułowań stanowiących jawne „kalki językowe”, jak np. „euklidesowy dystans” na str. 45 (powinno to być „odległość euklidesowa”), czy „dystans Hamminga” (str. 51), zwłaszcza jeśli istnieją znane polskie nazwy, podobnie zamiast „keypointa” (str. 73) można użyć polskiej nazwy „punkt kluczowy”, „splot” zamiast „konwolucja” (str. 117), czy „pionową i poziomą” zamiast „wertykalną i horyzontalną” (str. 125); niektóre angielskie nazwy nie są prawidłowo odmieniane (zamiast „Google’a” w pracy jest kilkakrotnie użyta odmiana „Googla”);
- nazewnictwo sekcji powinno być pełne, przykładowo nazwy sekcji 3.5.1, 3.5.2, 6.7.1, czy 6.7.2 mają sens jedynie w połączeniu z tytułami nadrzędnych podrozdziałów 3.5 oraz 6.7 – co prawda nie jest to rażące w spisie treści, ale utrudnia lekturę zasadniczej części pracy, gdy sekcja rozpoczyna się kilka stron później aniżeli podrozdział, którego jest częścią;
- w przypadku odnoszenia się do wzorów znajdujących się bezpośrednio za odnośnikiem wystarczające jest podanie numeru po prawej stronie wzoru (dotyczy to wielu wzorów m.in. (3.1)–(3.17), dużo naturalniej wygląda sposób umieszczenia w tekście wzoru (3.18));
- sformułowanie „pływ optyczny” (str. 79) nie jest raczej spotykane – pojęcie „*optical flow*” tłumaczy się zwykle jako „przepływ optyczny”; razi też sformułowanie „wyćwiczone sieci” (str. 109);
- w tekście na str. 82 kilka razy niepotrzebnie występuje słowo „punkt”;
- często zdarza się niekonsekwentna pisownia nazw własnych bez zachowania reguł dotyczących wielkich i małych liter (np. nazwy PyTorch, fastai, NVIDIA można zauważyć w pracy w różnych wersjach typograficznych, pomijając już „literówki” typu Pytocha – str. 111);
- ostatnie dwa zdania w sekcji 5.2.7 są napisane niejasnym, potocznym językiem, z kolei sformułowanie „co było przyczyną znaczących obszarach bez danych o odległość” (str. 127) jest przykładem całkowitego braku spójności gramatycznej, świadczącego niestety o braku staranności w redagowaniu tekstu rozprawy.

Usterki tego rodzaju w tekście rozprawy jest znacznie więcej, nie jest jednak celowe przytaczanie kolejnych; powyższa lista i tak jest stanowczo zbyt długa w przypadku rozprawy doktorskiej. Pewne uwagi można mieć również do formatowania tekstu, gdyż często zdarzają się pozostawione jednoliterowe spójniki

na końcach wierszy, niewłaściwie umiejscowione ilustracje, poniżej których znajduje się pojedynczy wiersz tekstu (widać to np. na str. 26, 27, 49, czy też str. 91), czy też brak dzielenia wyrazów skutkujący zaburzeniem linii prawego marginesu. Niewłaściwe jest także formatowanie tekstu na str. 176. Problemy te są możliwe do stosunkowo łatwego rozwiązania w systemie LaTeX, który został użyty do składu tekstu.

Szkoda, iż na niektórych rysunkach (np. rys. 6.34 na str. 187) Autor nie zastosował ujednocionej skali na wykresach, tak jak uczynił to na rysunku 6.33. Takie podejście, o ile jest możliwe, znacznie ułatwia wizualne porównanie uzyskanych wyników.

Edytorska strona pracy utrudnia niestety jej lekturę oraz znacząco obniża ogólną ocenę rozprawy. Jak wynika z jej lektury, Kandydat niestety nie pokusił się o uważne sprawdzenie rozprawy przed jej złożeniem, jednak z przedstawionych opisów wykonanych badań oraz analizy uzyskanych wyników wyłania się obraz ukazujący dużą ilość poświęconego czasu oraz znaczny wkład pracy w badania eksperymentalne. Ta merytoryczna, a także praktyczna, część pracy jest niewątpliwie najważniejszym jej elementem, choć dobra „wizytówka”, jaką stanowi pisemna rozprawa, również gra niebagatelną rolę. Tym niemniej, pomimo wskazanych wątpliwości i uwag, uznaję, iż założone cele pracy zostały osiągnięte a tezy udowodnione, co stanowi najważniejszy aspekt przedstawionej do oceny rozprawy doktorskiej.

V. WNIOSKI KOŃCOWE

Opiniowana rozprawa stanowi samodzielne rozwiązanie problemu naukowego, bez wątpienia mieszczącego się w zakresie dyscypliny naukowej *automatyka, elektronika, elektrotechnika i technologie kosmiczne*, zawierając także pewne elementy interdyscyplinarne z zakresu dyscypliny *informatyka techniczna i telekomunikacja*. W pracy Autor wykazał się znajomością zagadnień dotyczących robotyki mobilnej, analizy obrazów i widzenia maszynowego, jak również algorytmów sztucznej inteligencji. Wykazał również swoje umiejętności zastosowania i przystosowania metod widzenia maszynowego oraz algorytmów sztucznej inteligencji do wciąż pojawiających się nowych wyzwań współczesnej robotyki mobilnej, co potwierdza umiejętność samodzielnego rozwiązywania problemów naukowych, także na pograniczu obu ww. dyscyplin naukowych.

Stwierdzam, iż przedstawiona do recenzji rozprawa doktorska **mgr. inż. Aleksandra Kostusiaka** pt. ***Metody doboru struktury oprogramowania i optymalizacji parametrów w zadaniu lokalizacji wizyjnej na podstawie danych RGB-D***, której promotorem jest prof. dr hab. inż. Piotr Skrzypczyński, zrealizowana w ramach przewodu doktorskiego wszczętego w dniu 18.09.2018 r. w dziedzinie nauk technicznych w dyscyplinie *automatyka i robotyka*, zgodnie z Rozporządzeniem MNiSW z dnia 19.01.2018 r. (Dz. U. z 2018 r. poz. 261), pomimo uwag krytycznych wskazanych w treści recenzji, **spełnia wymagania** stawiane rozprawom doktorskim przez *Ustawę o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki* z dnia 14 marca 2003 roku (Dz. U. z 2017 r. poz. 1789), jak również przez aktualnie obowiązującą *ustawę Prawo o szkolnictwie wyższym i nauce* z dnia 20 lipca 2018 roku (tekst jednolity Dz. U. 2023 poz. 742, z późn. zm.), wraz z odpowiednimi przepisami wprowadzającymi oraz przejściowymi. **Wnioskuje zatem o jej przyjęcie i dopuszczenie do publicznej obrony.**