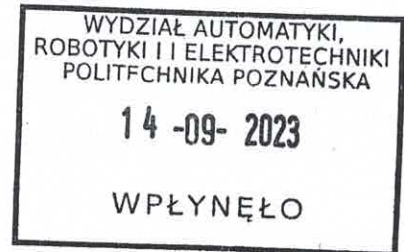


PRZEWODNICZĄCY RADY DYSCYPLINY
Automatyka, Elektronika, Elektrotechnika
i Technologie Kosmiczne
prof. dr hab. inż. Wojciech Szelaąg

Szczecin, 6.09.2023 r.

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



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

*opracowana na podstawie uchwały nr 41/2022-2023 Rady Dyscypliny Automatyka,
Elektronika, Elektrotechnika i Technologie Kosmiczne Politechniki Poznańskiej z dnia 21.06.2023 r.
oraz pisma Przewodniczącego Rady prof. dr. hab. inż. Wojciecha Szelaąga z dnia 6.07.2023 r.*

Tytuł rozprawy: **Nowe metody pasywnej percepcji wizyjnej w zadaniu nawigacji autonomicznego robota mobilnego**

Autor rozprawy: **mgr inż. Marta Rostkowska**

Dyscyplina naukowa: **automatyka i robotyka**

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

I. TEMATYKA, TEZA NAUKOWA I CEL ROZPRAWY

Widzenie maszynowe i komputerowe staje się w ostatnich latach jedną z najdynamiczniej rozwijających się gałęzi nowoczesnej automatyki i robotyki, integrującą także elementy informatyki technicznej oraz elektroniki. Sprzyja temu rosnąca dostępność wydajnego sprzętu obliczeniowego wraz z jego miniaturyzacją, powszechność i relatywnie niska cena kamer wysokiej rozdzielczości, a także rozwój platform i środowisk programistycznych, w szczególności wsparcie dla głębokich sieci neuronowych. Dynamiczny rozwój metod sztucznej inteligencji, zwłaszcza uczenia maszynowego, w zastosowaniach związanych z analizą obrazu spowodował także wzrost zainteresowania tematyką dotyczącą przemysłowych zastosowań widzenia maszynowego i komputerowego, w tym związanych z robotyką mobilną.

Recenzowana rozprawa bardzo dobrze wpisuje się w te trendy, jak również w nową dyscyplinę naukową *automatyka, elektronika, elektrotechnika i technologie kosmiczne*, w której stanowić ma ona podstawę nadania stopnia doktora nauk inżynierijno-technicznych. Podjęta w niej tematyka jest niewątpliwie ważna i aktualna, a postawiona teza naukowa dotycząca celowości wykorzystania inspirowanej biologicznie koncepcji akwizycji obrazów oraz ich równoległego przetwarzania do

realizacji w czasie rzeczywistym wybranych funkcji nawigacyjnych robota mobilnego w systemie wbudowanym o ograniczonych zasobach obliczeniowych jest w mojej opinii trafna, co znajduje zresztą potwierdzenie w treści rozprawy. Jest ona sformułowana w precyzyjny sposób i została udowodniona w przedstawionej do oceny rozprawie, której cele dotyczyły głównie zbudowania stanowiska eksperymentalnego, opracowania metod przetwarzania i analizy obrazów pochodzących z pary różnych kamer („klasycznej” perspektywicznej oraz dookólnej kamery katadioptrycznej), jak też weryfikacji eksperymentalnej postawionej hipotezy wraz z opracowaniem bazy danych testowych niezbędnej zarówno do uczenia sieci neuronowej, jak również weryfikacji dokładności samolokalizacji robota mobilnego na podstawie opracowanych metod. Cele te zostały w rozprawie osiągnięte, udowodnione zostały również sformułowane przez Doktorantkę hipotezy pomocnicze dotyczące możliwości zastosowania zasady współpracy widzenia peryferyjnego i centralnego znanej z przyrody, a także poprawy dokładności szacowania odległości od wybranych obiektów (dzięki odpowiedniej kalibracji kamer i rektyfikacji obrazu) oraz możliwości uzyskiwania obrazów wynikowych i ich deskryptorów w czasie rzeczywistym (dzięki zrównoleżonemu przetwarzaniu danych z kamery dookólnej).

W mojej opinii osiągnięcia opisane w dysertacji potwierdzają osiągnięcie założonego celu oraz pozytywną weryfikację hipotez badawczych przedstawionych we wstępie rozprawy.

II. ZAWARTOŚĆ MERYTORYCZNA ROZPRAWY

Rozprawa doktorska pt. *„Nowe metody pasywnej percepcji wizyjnej w zadaniu nawigacji autonomicznego robota mobilnego”* została podzielona na 8 zasadniczych rozdziałów, przy czym pierwszy z nich stanowi wprowadzenie a ostatni zawiera podsumowanie i wnioski. We wstępnej części rozprawy Doktorantka przedstawiła motywację do podjęcia badań nad proponowanym systemem wizyjnym, a także sformułowała problem naukowy i hipotezę badawczą oraz przybliżyła strukturę pracy. W rozdziale 2. Doktorantka przedstawiła stan wiedzy dotyczący pasywnych systemów wizyjnych używanych w robotyce mobilnej, natomiast w kolejnym rozdziale skupiła się na przedstawieniu aktualnego stanu rozwoju zastosowań uczenia maszynowego do analizy danych pozyskiwanych z kamer. W ramach analizy zastosowań systemów wizyjnych do celów samolokalizacji i nawigacji robotów mobilnych Autorka przedstawiła procesy percepcji obrazu u zwierząt i ludzi, które stanowiły inspirację do rozwoju systemu zaproponowanego w dysertacji. Ponadto dokonany został przegląd istniejących systemów oraz metod wizyjnego pomiaru odległości oraz samolokalizacji robotów mobilnych. Zagadnienia dotyczące inspiracji biologicznych zostały przedstawione w bardzo szczegółowy sposób, co umożliwiło przedstawienie analogii do proponowanego rozwiązania. W podrozdziale 2.4 omówione zostały różne metody lokalizacji bazujące na danych pozyskanych z pasywnych sensorów wizyjnych. Autorka omówiła lokalizację na podstawie naturalnych cech otoczenia oraz znaczników, jak też lokalizację topologiczną, wspominając także o metodach lokalizacji grupy robotów.

Rozdział 3 rozpoczyna się opisem klasycznych sieci neuronowych oraz metod ich uczenia. Następnie przedstawione zostały architektury sieci spłotowych używanych w zadaniach klasyfikacji ze szczególnym uwzględnieniem sieci głębokich. Na uwagę zasługuje rzetelne porównanie różnych rodzajów sieci oraz czytelny sposób przedstawienia ich rozwoju, a także ograniczeń i zalet poszczególnych architektur sieci. W kolejnym podrozdziale omówione zostały architektury sieci CNN

używanych do lokalizacji i detekcji obiektów ze szczególnym uwzględnieniem algorytmu YOLO. Następnie omówione zostały zagadnienia zrównoleglenia metod przetwarzania obrazu pod kątem robotyki, głównie z zastosowaniem architektury CUDA. Omówiona część pracy ma charakter opisowy, bazując na rzetelnym przeglądzie literaturowym zawierającym wiele pozycji z ostatnich lat. Warto zauważyć, iż bibliografia liczy aż 318 pozycji w znakomitej większości stanowiących publikacje konferencyjne oraz artykuły opublikowane w uznanych czasopismach naukowych ściśle związane z tematyką rozprawy.

W rozdziale 4 przedstawiono koncepcję systemu wizyjnego o hybrydowym polu widzenia opartego na podstawach biologicznych. System ten bazuje na założeniu wykorzystywania kamery dookólnej (peryferyjnej) o szerokim polu widzenia jednocześnie z kamerą perspektywiczną (widzenie centralne), zapewniającą lepsze odwzorowanie szczegółów dla wybranego fragmentu obserwowanej sceny. Znaczącym utrudnieniem implementacji proponowanego rozwiązania jest różna geometria obu kamer. Zastosowanie znanych algorytmów stereoskopowych wymaga w takim przypadku odpowiedniego przetworzenia obrazu uzyskanego z kamery katadioptrycznej, co stanowi jeden z najistotniejszych elementów ocenianej rozprawy. Przedstawiona motywacja zastosowanego rozwiązania jest przekonująca, zwłaszcza biorąc pod uwagę analizę możliwości sprzętowych. W rozdziale tym przedstawiono opis zależności geometrycznych w obu rozważanych rodzajach kamer, a także klasyczny układ stereoskopowy oraz metody rektyfikacji obrazu.

Ze względu na fakt, iż dokładność analizy obrazów stereoskopowych, a zwłaszcza wyznaczania odległości od obiektów, jest uzależniona od występowania zniekształceń geometrycznych, rozdział 5. rozprawy poświęcony został omówieniu metod kalibracji systemu wizyjnego. Przedstawiono zarówno proces kalibracji kamery perspektywicznej, jak również katadioptrycznej, a także całego dwukamerowego układu stereowizyjnego. Rozdział ten również bazuje na znanych z literatury metodach, jednak stanowi niezbędny element pracy pozwalający na całościową analizę rozprawy i proponowanego w niej rozwiązania.

W rozdziale 6. przedstawiono metody lokalizacji na podstawie sztucznych znaczników głównie wykorzystujące binarne kody matrycowe, takiej jak np. powszechnie znane kody QR. Warto zauważyć, iż Autorka dokonała porównania wyników stosowania różnych znaczników, a także określiła błędy pomiaru odległości oraz kąta widzenia znacznika na obrazach uzyskanych z kamery perspektywicznej. Przedstawiono także metodę samolokalizacji robota na podstawie znaczników z kodem QR przy pomocy hybrydowego systemu wizyjnego. Szczególną uwagę Doktorantka zwróciła na zagadnienia wyszukiwania znaczników na obrazach uzyskanych z kamery dookólnej, a także na możliwości zastosowania w tym celu konwolucyjnej sieci YOLO. Przedstawione zostały także wyniki dodatkowych eksperymentów przeprowadzonych w typowym środowisku domowym z użyciem znaczników zawierających kody QR.

Za najbardziej wartościową pod względem merytorycznym część pracy uznać można rozdział 7. dotyczący samolokalizacji robota na podstawie naturalnych cech otoczenia z użyciem proponowanego hybrydowego układu stereowizyjnego. Stanowi on swoiste zwięźcenie rozprawy, w którym przedstawiono wyniki eksperymentalne dotyczące dopasowania cech obrazów z użyciem różnych deskryptorów, a także analizy uzyskanej dokładności pomiaru odległości oraz czasu przetwarzania z użyciem procesora graficznego GPGPU oraz standardowej jednostki CPU. Następnie

przedstawiono metody lokalizacji oparte na spłotowych sieciach neuronowych z użyciem obrazów pozyskanych z kamery katadioptrycznej. Weryfikacja koncepcji dokonana została na podstawie dwóch baz testowych, z których jedna stworzona została w budynku Centrum Mechatroniki Politechniki Poznańskiej w ramach realizacji ocenianej rozprawy. Ważnym elementem tego rozdziału jest badanie własności proponowanej metody lokalizacji na obu zbiorach danych (drugi zbiór stanowi baza zarejestrowana w budynku uniwersytetu we Freiburgu).

Ostatni rozdział zawiera podsumowanie rezultatów badań, wnioski, a także koncepcję dalszego rozwoju zaproponowanego w rozprawie rozwiązania.

III. OGÓLNA OCENA ROZPRAWY I UWAGI DYSKUSYJNE

Przedstawiona do oceny rozprawa doktorska mgr inż. Marty Rostkowskiej zawiera interesujące wyniki teoretyczne oraz eksperymentalne z zakresu zastosowania systemów wizyjnych w robotyce mobilnej. Została ona zredagowana dość starannie, zawiera na ogół wysokiej jakości ilustracje oraz bogatą bibliografię. Zawiera ona trafnie dobrane pozycje, w zdecydowanej większości z okresu ostatnich kilku bądź kilkunastu lat, będące w ścisłym związku z tematyką dysertacji. Niestety doktorantka nie ustrzegła się dość licznych usterek typograficznych interpunkcyjnych, a także pewnej liczby potknięć językowych. Obniżają one nieco ogólną ocenę rozprawy, choć nie wpływają w zasadniczy sposób na jej wartość merytoryczną, którą ocenić należy wysoko. Niewątpliwą zaletę rozprawy stanowi fakt udanej implementacji rozwiązań znanych ze świata przyrody w systemach robotycznych, a także szczegółowy przegląd stosowanych metod i technik oraz możliwości ich wykorzystania w proponowanym rozwiązaniu. Doktorantka w przekonujący sposób przedstawia motywację zastosowania proponowanych metod, jasno wskazując własne osiągnięcia, a także porównując je z aktualnym stanem wiedzy w zakresie odpowiadającym tematyce doktoratu. Potwierdza to zarówno dobrą jego znajomość, jak również rzetelność jego analizy przeprowadzoną przez Kandydatkę.

Przedstawiona analiza pasywnej percepcji wizyjnej w zadaniach nawigacji robotów mobilnych w szczególności ich samolokalizacji, jak również inspiracji biologicznych, w połączeniu z propozycją hybrydowego systemu wizyjnego zaimplementowanego na rzeczywistej platformie robota mobilnego stanowią oryginalne osiągnięcie naukowe, wymagające szerokiej wiedzy zarówno z zakresu automatyki i robotyki, jak również technik programistycznych, w szczególności dotyczących metod sztucznej inteligencji oraz widzenia maszynowego i komputerowego.

Do oryginalnych osiągnięć naukowych Doktorantki, trafnie wskazanych w rozdziale 8., zaliczyć należy:

- opracowanie procedury szybkiej transformacji obrazu pozyskanego z kamery dookólnej do obrazu panoramicznego zgodnego z obrazem uzyskiwanym z kamery perspektywicznej pod względem parametrów geometrycznych,
- opracowanie metody kalibracji sensora stereoskopowego składającego się z kamery perspektywicznej oraz kamery wirtualnej,
- opracowanie procedury tworzenia obrazu z kamery wirtualnej oraz koncepcji użycia tego obrazu jako elementu składowego pary stereowizyjnej,

- propozycja metody detekcji sztucznych znaczników z kodami QR na obrazach 360° oraz ich użycia do lokalizacji robota mobilnego,
- zaproponowanie techniki szybkiej detekcji tych wskaźników na obrazach 360° z użyciem sieci neuronowej YOLOv3,
- koncepcja metody lokalizacji topologicznej robota mobilnego z użyciem mapy globalnej deskryptorów dla poszczególnych lokacji oraz sieci neuronowej.

Pewien niedosyt pozostawia niepełne porównanie uzyskanych wyników z alternatywnymi rozwiązaniami znanymi z literatury. W szczególności uwaga ta dotyczy braku porównania wyników samolokalizacji robota uzyskanych przy użyciu proponowanego w rozprawie hybrydowego systemu wizyjnego w porównaniu z systemem bazującym wyłącznie na kamerze perspektywicznej. Niestety Tabela 6.3 zawiera wyłącznie wyniki uzyskane dla systemu hybrydowego, co utrudnia nieco jednoznaczną ocenę przydatności i zasadności stosowania dodatkowej kamery katadioptrycznej, przykładowo zamiast drugiej kamery perspektywicznej, co uprościłoby implementację pod względem programistycznym.

Pewne zastrzeżenia budzić może brak ciągłości wywodu w niektórych miejscach pracy, jak chociażby na stronach 16-17 w podrozdziale 2.2. Można także odnieść wrażenie, iż opis inspiracji biologicznych jest nieco zbyt obszerny w porównaniu z pozostałymi rozdziałami pracy, która przecież przedstawiona jest jako rozprawa doktorska w dyscyplinie automatyka i robotyka. Z drugiej jednak strony pewne cechy interdyscyplinarności pracy mogą (i powinny) być postrzegane jako zalety świadczące o szerokiej wiedzy Doktorantki nie ograniczającej się tylko do jednej dyscypliny naukowej.

Wątpliwość budzi także sformułowanie zawarte na stronie 28 rozprawy, gdzie Doktorantka stwierdziła, iż aktywne systemy wizyjne nie mogą być stosowane na zewnątrz budynków. w mojej ocenie stwierdzenie to jest zbyt uogólnione, ponieważ rozwiązania takie istnieją, choć faktycznie ich realizacja jest znacznie trudniejsza aniżeli systemów pracujących w budynkach. Omówione w rozdziale 6 sztuczne znaczniki wykorzystujące kody matrycowe nie uwzględniają dość popularnego rozwiązania w robotyce, którym są tzw. znaczniki AprilTags. Zostały one w rozprawie pominięte, jednak warto byłoby określić tego przyczynę lub przedstawić wyniki badań porównawczych z kodami QR. Szkoda także, iż na str. 76 nie umieszczono bardziej szczegółowej informacji dotyczącej wielomianowego modelu zniekształceń obrazu uzyskiwanego z kamery, dla którego stosowane są pierwsze wyrazy szeregu Taylora.

Na str. 104 Doktorantka stwierdziła, iż w pierwszym kroku przetwarzania obraz znacznika poddawany jest procesowi przekształcenia morfologicznego w celu usunięcia szumów z obrazu. Niestety nie sprecyzowała, o jakie konkretnie przekształcenie morfologiczne chodzi. Jak można domyślić się z rysunku 6.7, jest to erozja (wskazana zresztą na schemacie w języku angielskim /Erode/ zamiast polskim). Niejasne jest również sformułowanie znajdujące się na str. 108 dotyczące zamiany kolorowego obrazu z kamery katadioptrycznej na obraz czarno-biały. Pojęcie to jest mylące, gdyż kojarzy się z binaryzacją obrazu (nie wiadomo jaką metodą), podczas gdy w rzeczywistości chodzi zapewne o konwersję obrazu do skali szarości, jak wynika z rysunku 6.11. Nie jest również jasne, która z metod konwersji została wykorzystana i czy zbadany został wpływ metody konwersji obrazu kolorowego do skali szarości na uzyskiwane wyniki.

Bliższego omówienia wymagałoby także stwierdzenie zawarte na str. 57 „To równe skalowanie wszystkich warstw nie jest optymalne”, zwłaszcza iż Autorka nie powołuje się przy nim na żadną pozycję bibliograficzną. Nie jest zatem jasne, czym jest spowodowane takie stwierdzenie i względem jakiego kryterium wskazana optymalność nie występuje.

Jak wskazuje Autorka na str. 64, w chwili pisania pracy najnowszą wersją sieci YOLO jest jej 8. wersja, jednak w pracy użyto sieci YOLOv3. Wybór ten nie został jednak w wystarczająco przekonujący sposób umotywowany w rozprawie.

Doprecyzowania wymagałoby również stwierdzenie ze str. 82 dotyczące wyznaczania parametrów modelu za pomocą metody Levenberga-Marquardta. Czy zdaniem Autorki porównywalne lub nawet lepsze wyniki mogłoby zostać uzyskane innymi metodami optymalizacyjnymi? Co Autorka rozumie przez „dokładne odległości” określone również na str. 82?

Na str. 102 Autorka wskazała zastosowanie zmodyfikowanej gradientowej transformaty Hougha, jednak nie zostało wyjaśnione, na czym ta modyfikacja polega. Nie wskazano tam także, jakie algorytmy znajdowania konturów zostały użyte.

Dyskusyjne jest również stwierdzenie ze str. 7 „Powszechność kamer w życiu codziennym spowodowała ich rozwój, miniaturyzację, energooszczędność i ciągłe dążenie do otrzymywania obrazów o jak największej rozdzielczości”, gdyż równie dobrze można stwierdzić, iż to właśnie ich rozwój i miniaturyzacja wpłynęły na wzrost ich dostępności oraz powszechności.

Użycie określenia „zniekształcenia radialne i tangensowe” (po raz pierwszy na str. 31) powinno być raczej zastąpione częściej używanym (zresztą także przez Doktorantkę w dalszej części rozprawy) sformułowaniem „zniekształcenia radialne i styczne”, ewentualne zamienne użycie określenia „zniekształcenia tangensowe” zamiast „styczne” powinno być wyjaśnione przy pierwszym jego wystąpieniu, tak aby nie było mylące dla czytelników mniej zorientowanych w zagadnieniach kalibracji kamer. Nieco niefortunne jest także zdanie „Usunięcie zniekształceń na obrazie odbywa się na podstawie modelu matematycznego kamery w procesie kalibracji” (str. 31) – podczas kalibracji w zasadzie jest uzyskiwany model oraz wyznaczane są parametry kamer a sam proces usuwania zniekształceń następuje później. Podobna uwaga dotyczy sformułowania na str. 73 „kamera perspektywiczna, za pomocą której mam miejsce szczegółowa analiza sceny” (zasadniczo jest ona przeprowadzana za pomocą procesora).

Sformułowanie zawarte na str. 35 „gdy obiekt oddala się od kamery, rozmiar obrazu zmniejsza się” nie wydaje się prawidłowe – zapewne Autorka miała na myśli zmniejszający się rozmiar obiektu widocznego na obrazie a nie samego obrazu. Znajdujące się na str. 45 określenie „do momentu uzyskania zbieżności wartości wyjściowej z rzeczywistą wartością” jest niejasne – zapewne Doktorantce chodziło o „zgodność” a nie „zbieżność”, kojarzącą się raczej z procesem zmierzania do określonej wartości w czasie (a nie konkretnym momentem). Podobnie niejasna jest fraza „transfer uczenia” na str. 49 – właściwe tłumaczenie to „uczenie transferowe”, zresztą używane prawidłowo w dalszej części rozprawy. Kolejne tego typu uwagi dotyczą niezbyt precyzyjnych sformułowań: „ponieważ mniejszy rozmiar nie może uchwycić informacji lewo-prawo i góra-dół” (str. 51), czy też „Wszystko co robi, to używa konwolucji” (str. 53), jak również „dwa eksperymenty uzyskały najlepsze wyniki” (str. 131).

W pracy pojawiają się także nieliczne sformułowania o charakterze potocznym, czy nawet nieco slangowym, np. „za pomocą niedrogiej konwolucji” oraz „konwolucja 1×1 ścisła sieć” (str. 56), „obrazów z datasetu” (str. 136). Nie jest także jasne, co Autorka rozumie przez „uzyskanie idealnego obrazu poziomego” (str. 74).

Interpretacja wykresów zamieszczonych na rysunkach 7.19 oraz 7.20 byłaby zapewne nieco łatwiejsza gdyby słupki reprezentujące poszczególne wartości były pogrupowane inaczej, tzn. wyniki uzyskane dla trzech różnych metod uczenia pokazane na słupkach obok siebie na jednym wykresie; wówczas każdy wykres dotyczyłby osobno wyników uzyskanych w dni pochmurne, słoneczne oraz w nocy (reprezentowanych przez poszczególne kolory na każdym z wykresów).

IV. UWAGI SZCZEGÓŁOWE

Podczas lektury rozprawy zauważono szereg usterek mniejszej wagi, które wpływają na ogólny odbiór pracy, jednak nie mają wpływu na jej wysoką ocenę pod względem merytorycznym. Niektóre spośród częściej powtarzających się błędów i usterek zostały wymienione poniżej:

- na stronie tytułowej rozprawy przyimek „w” nie powinien występować na końcu wiersza lecz być przeniesiony do nowego wiersza a tytuł dysertacji nie powinien kończyć się kropką;
- w pracy pojawiają się dość liczne usterki typograficzne, głównie polegające na braku znaków diakrytycznych oraz niewłaściwej końcówce wyrazu, lub kropek po „np.” (dwukrotnie na str. 14 i dalej), a także na zamianie kolejności liter (tzw. „czeskie błędy”), np. „z jednego oko” (str. 25), „punty” (str. 27), „kat widzenia” (str. 29), „przykłady takie systemu” (str. 37), „uzyskania” (str. 41), „komonikacji” (str. 43), „warst”, „wartwy” oraz „wartswa” (str. 46), „wszytkie” (str. 47), „skupiowo” (str. 49), „Schamt” i „przekzywany” (str. 52), „wartswy” i „wsadowową” (str. 53), „ostania” (str. 55), „dokładnością” (str. 56), „kowolucyjne” (str. 57), „mairze” (str. 58), „podobieństwa” (str. 59), „wykywania” (str. 61), „przetrzeni” (2× na str. 73), „opicanych” (str. 85), „występje” (str. 88), „katadioprycznej” i „zwierciadłem” (str. 92), „wyniki” (str. 94), „szchownic” (2× na str. 95), „wirtualną” (str. 96), „użciem” (str. 101), „przedstawionowego” (str. 107), „Jeston” (2× na str. 117), „zadnym” (3× na str. 126) i sporo temu podobnych;
- w pracy często pojawiają się błędy interpunkcyjne, zwykle w postaci brakujących przecinków, np. przed wyrażeniami „przez którą”, „a także”, „aby”, „co umożliwia”, „gdzie”, „czyli”, itp.;
- w podrozdziale 1.3 brak odniesienia do zawartości rozdziału 3;
- kilkakrotnie zdarza się użycia sformułowania „oparty o” zamiast „oparty na”, np. str. 12, 40, 56, 98, 127, 130;
- kilkakrotnie występuje sytuacja, w której zdanie rozpoczyna się słowem „natomiast” lub „ale”, np. str. 10, 14, 41, 74, 98, 100, 107, 113, 125, 128;
- brak oddzielenia wartości od jednostek np. „440Hz” oraz „400nm” na str. 16, czy też „265Hz” i „53Hz” na str. 18;
- większość wyliczeń, które powinny stanowić jedno zdanie, zawiera elementy rozpoczynające się wielką literą zakończone kropką (np. str. 37, 49, 92, 97, 99, 136);
- występujące błędy ortograficzne polegające na użyciu końcówki „-ą” zamiast „-om” spowodowane być może użyciem funkcji dyktafonu, np. „określonym zachowaniom i interakcją

- są przypisane ..." (str. 16), „ciągłym zmianą” (str. 38), „ponadto jednostką wykonawczym GPU zezwolono ..." (str. 67), „poddawany operacją” (str. 109), a także „porządane” zamiast „pożądane” na str. 55 oraz „wyekstrachowanymi” zamiast „wyekstrahowanymi” na str. 68;
- akronim ToF (Time of Flight) użyty po raz pierwszy na stronie 28 powinien być tam wyjaśniony, podobnie jak MLP na str. 53, z kolei dwukrotne rozwinięcie akronimu CUDA (str. 73) jest zbędne;
 - określenia „dense” i „sparse” w odniesieniu do stereowizji na str. 31 powinny być (przynajmniej przy pierwszym wystąpieniu) zastąpione – używanymi zresztą w dalszej części pracy – polskimi odpowiednikami;
 - słowa „mikrosekundowej” (str. 33), „indoor” (str. 41), czy też „jednordzeniowy” (str. 69), powinny być pisane bez myślników, słowa „na przemian” (str. 99) oddzielnie, zaś słowo „megapikseli” (str. 68) powinno być pisane łącznie;
 - błędy typograficzne w angielskich słowach „Threshold” (dwukrotnie na str. 43), „Rectified” (str. 45) oraz „transfer” i „tuning” (str. 47), czy też „matching” (str. 69);
 - na str. 42 fraza „niezmienna na zmiany oświetlenia” powinna być zastąpiona inną np. „niezmienna względem zmian oświetlenia” lub „odporna na zmiany oświetlenia”; podobnie fraza „występowania w wideo” powinna być zastąpiona przez „występowania w sekwencji wideo”;
 - sformułowanie „piksel wykryje zmianę jasności sceny” nie jest zbyt fortunne, gdyż piksel jedynie reprezentuje jasność, niczego nie wykrywając; podobnie nieprecyzyjne określenie występuje na str. 77 („każdy piksel na zarejestrowanym obrazie mierzy natężenie światła”);
 - słowo „znajdywanie” powinno być zastąpione przez „znajdowanie” (str. 52), „50\$” przez „50%” (str. 68), zaś „uproszczają” przez „upraszczają” (str. 97);
 - określenie „za pomocą nieliniowej funkcji $f()$ ” na str. 78 nie pozwala określić, o jaką funkcję chodzi, zapewne jest to funkcja określona wzorem (4.9), jednak nie wynika to wprost z tekstu;
 - w kilku miejscach pracy nieprawidłowo użyto słowa „liczba” zamiast „ilość” w odniesieniu do rzeczowników niepoliczalnych (dane, zasoby, informacje), np. na str. 61;
 - w wielu miejscach pracy po wzorach pojawiają się zbędne wcięcia tekstu, zwykle przed słowem „gdzie” (np. w rozdziale 4. począwszy od str. 78, czy też w rozdziałach 5.-7.);
 - określenie „osie optyczne obu kamer są równoległe i przecinają się w nieskończoności” (str. 84) nie jest zbyt fortunne – w rzeczywistości te osie po prostu się nie przecinają;
 - fraza „ostatnim krokiem, gdy już zostały wyznaczone wszystkie współczynniki ma miejsce ich optymalizacja” powinna zostać przereklamowana;
 - na str. 103 obszar zainteresowania został omyłkowo oznaczony jako IoU zamiast RoI;
 - kolorystyka wykresów na rysunku 6.9 jest mało intuicyjna i nie sprzyja ich czytelności; z kolei czytelność wykresu na rysunku 7.4a poprawiłoby umieszczenie linii ukośnej ilustrującą wzajemnie równe odległości na obu osiach;
 - sformułowanie „Pozycja kamery składa się z 6 stopni swobody” (str. 106) nie jest precyzyjne;
 - niektóre pozycje bibliograficzne nie są sformatowane prawidłowo, występują braki tytułów monografii, w których opublikowano rozdziały np. [23, 259, 263], jak również brak użycia wielkich liter w nazwach własnych konferencji i czasopism (powodujące niespójność opisów bibliograficznych dla różnych artykułów nawet z tej samej konferencji), np. CVPR [2,7, 94, 279], ICRA [11, 318], PAMI [56, 314], ICCV [93], ICUAS [84], ICAIM [269], jak też w tytułach publikacji (np. Lucas-Kanade w [29], SLAM [151]), itp.;



Powyższa lista usterek nie jest pełna (nie wszystkie zostały wymienione w recenzji), jednak warto zauważyć, iż pomimo ich obecności praca napisana jest zrozumiałym i na ogół poprawnym językiem. Można uznać, iż w zdecydowanej większości nie są to błędy szczególnie rażące, zwłaszcza iż często zdarza się kilkukrotne wystąpienie usterek tego samego rodzaju, w szczególności interpunkcyjnych, czy typograficznych. Można je uznać za usterki drobne, nie utrudniające w istotnym stopniu lektury i zrozumienia zagadnień przedstawianych w rozprawie. Z tego względu nie umniejszają one mojej pozytywnej oceny rozprawy.

V. WNIOSKI KOŃCOWE

Opiniowana rozprawa stanowi oryginalne rozwiązanie problemu naukowego mieszczącego się w zakresie dyscypliny naukowej *automatyka i robotyka*, w której otwarty został przewód doktorski. Zawiera ona pewne elementy interdyscyplinarne, związane częściowo z zagadnieniami informatyki technicznej. Tematyka i zakres dysertacji bardzo dobrze wpisuje się w zakres aktualnej dyscypliny *automatyka, elektronika, elektrotechnika i technologie kosmiczne*. Jej Autorka wykazała się biegłą znajomością zagadnień związanych zarówno z technikami widzenia maszynowego i komputerowego, jak również z metodami nawigacji i samolokalizacji robotów mobilnych, potwierdzając umiejętność rozwiązywania problemów naukowych z tego zakresu oraz ogólną wiedzę teoretyczną właściwą dla dyscypliny naukowej.

Stwierdzam, iż przedstawiona do recenzji rozprawa doktorska **mgr inż. Marty Rostkowskiej** pt. ***Nowe metody pasywnej percepcji wizyjnej w zadaniu nawigacji autonomicznego robota mobilnego***, której promotorem jest prof. dr hab. inż. Piotr Skrzypczyński, zrealizowana w ramach przewodu doktorskiego wszczętego w dniu 20.09.2016 r. w dyscyplinie automatyka i robotyka w dziedzinie nauk technicznych, zgodnie z Rozporządzeniem MNiSW z dnia 30.10.2015 r. (Dz. U. z 2015 r. poz. 1842) spełnia wszystkie 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 2014 r. poz. 1852, oraz z 2015 r. poza. 249 i 1767), jak również przez aktualnie obowiązującą *Ustawę Prawo o szkolnictwie wyższym i nauce* wraz z odpowiednimi przepisami wprowadzającymi oraz przejściowymi. **Wniosuję o jej przyjęcie i dopuszczenie do publicznej obrony.**

