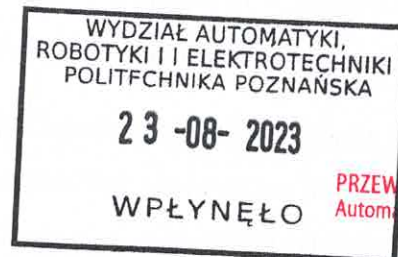


Warszawa, 21.08.2023

dr hab. Barbara Siemiątkowska, prof. PW
Politechnika Warszawska
Wydział Mechatroniki
e-mail Barbara.Siemiątkowska@pw.edu.pl



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

RECENZJA ROZPRAWY DOKTORSKIEJ

Mgr inż. Marty Rostkowskiej

pt. „Nowe metody pasywnej percepcji wizyjnej w zadaniu nawigacji autonomicznego robota mobilnego”

Promotor: prof. dr hab. inż. Piotr Skrzypczyński

1. Wybór tematu badawczego

Obecnie obserwujemy gwałtowny rozwój metod sztucznej inteligencji i jednocześnie wzrost zainteresowania robotyką mobilną. Roboty mobilne zaczynają być stosowane nie tylko w przemyśle, ale także w życiu codziennym. Przykładem tego typu rozwiązań są autonomiczne odkurzacze, kosiarki, wózki pracujące w szpitalach, biurach, supermarketach. Rozwija się także rynek samochodów autonomicznych.

Robot nie tylko powinien skutecznie wykonywać powierzone zadanie np. odkurzanie, mycie powierzchni, transport, ale jednocześnie pracować w sposób bezpieczny dla otoczenia. W przypadku robotów o wysokim stopniu autonomii istotna jest też umiejętność określenia położenia. Możemy wyróżnić lokalizację metryczną oraz semantyczną. W przypadku lokalizacji metrycznej obliczane są współrzędne i orientacja robota w przyjętym układzie współrzędnych. W przypadku lokalizacji semantycznej określamy typ miejsca w którym znajduje się robot.

W proponowanej rozprawie przedstawiono koncepcję, wzorowanego na organizmach żywych sensora wizyjnego, który umożliwia obserwowanie otoczenia w zakresie 360° .

Zaprezentowano też metody lokalizacji metrycznej, topologicznej i częściowo semantycznej, na podstawie sztucznych znaczników i naturalnych cech otoczenia. Istotną cechą rozprawy jest to, że obliczenia mogą być realizowane w czasie rzeczywistym.

Przedstawione wyniki badań odpowiadają najnowszym trendom badawczym robotyki mobilnej i stwarzają szerokie możliwości aplikacyjne. Wybór tematyki pracy doktorskiej uważam za uzasadniony, a wyniki badań stwarzają możliwości aplikacyjne.

2. Charakterystyka pracy

Recenzowana praca została wykonana w Instytucie Robotyki i Inteligencji Maszynowej, na Wydziale Automatyki, Robotyki i Elektrotechniki Politechniki Poznańskiej pod kierunkiem prof. dr hab. inż. Piotr Skrzypczyńskiego.

Rozprawa doktorska została wydana w formie rozprawy naukowej. W wyniku prowadzonych badań powstało 12 publikacji, w 11 mgr inż. Marta Rostkowska jest głównym autorem.

Przedstawiona do recenzji praca składa się z 8 rozdziałów. We wstępie Autorka rozprawy przedstawiła motywację do podjęcia prac dotyczących potrzeby budowy systemów wizyjnych o niekonwencjonalnej strukturze. Przedstawiono sformułowanie problemu naukowego i tezę pracy:

„Wykorzystanie inspirowanej biologicznie koncepcji akwizycji obrazów oraz ich równoległego przetwarzania pozwala realizować w czasie rzeczywistym wybrane funkcje nawigacyjne robota mobilnego w systemie wbudowanym o ograniczonych zasobach obliczeniowych”.

Ponieważ teza jest ogólna sformułowano trzy hipotezy pomocnicze, które pozwalają zawęzić zakres pracy. Tezy te są związane z kalibracją układu kamer, realizacją obliczeń w czasie rzeczywistym. Najważniejszą tezą jest wykazanie, że hybrydowy system wizyjny umożliwia lokalizację robota (na podstawie znaczników sztucznych i naturalnych) w szerokim zakresie odległości i kątów między sensorem, a znacznikami.

W Rozdz. 2 przedstawiono stan wiedzy w dziedzinie systemów wizyjnych i ich wykorzystania w robotyce mobilnej. Omówiono:

- Wykorzystanie systemów wizyjnych w robotach mobilnych pracujących w różnych warunkach i wykonujących różne zadania.

- Przedstawiono inspiracje biologiczne, omówiono narządy wzroku różnego rodzaju zwierząt i związek między wzrokiem, a działaniem organizmów żywych.
- Podział i charakterystyką systemów wizyjnych wykorzystywanych w lokalizacji robota mobilnego oraz przegląd metod lokalizacji.
- Przedstawiono podział systemów lokalizacji robota. Autorka skupiła się głównie na sieciach konwolucyjnych, które były wykorzystywane w prowadzonych badaniach.

W kolejnym rozdziale szczegółowo przedstawiono zastosowanie sieci neuronowych w zagadnieniach klasyfikacji obiektów występujących na obrazie i segmentacji obrazów.

Dalsza część pracy stanowi omówienie autorskich rozwiązań mgr inż. Marty Rostkowskiej. Rozdział 4 poświęcono omówieniu koncepcji proponowanego hybrydowego systemu wizyjnego, który składa się z dwóch typów systemów wizyjnych: klasycznej kamery i kamery dookólnej. Dzięki temu rozwiązaniu kąt widzenia wynosi 360° .

W rozdz. 5, autorka rozprawy omówiła metody kalibracji kamery klasycznej, dookólnej i hybrydowego układu stereowizyjnego składającego się z kamer o istotnie różnej geometrii. Metoda kalibracji układu hybrydowego jest oryginalnym rozwiązaniem mgr inż. Marty Rostkowskiej.

W rozdz. 6 przedstawiono metody lokalizacji robota na podstawie sztucznych znaczników. Przedstawione rozwiązanie może być stosowane w przypadku robotów usługowych pracujących w znanym środowisku. Przeprowadzono wyniki przeprowadzonych testów. Określono błąd określenia orientacji robota (4°), i błąd określenia położenia robota, który wynosi 3cm w osi OX i 5cm w osi OY. Autorka skupiła się także na czasie wykonywania obliczeń. Porównano czas wykrywania znaczników za pomocą sieci YOLO realizowanej klasycznie i w technologii CUDA.

W kolejnym rozdziale została opisana metoda lokalizacji robota na podstawie naturalnych cech otoczenia. Przetestowana metodę oceny odległości wykorzystując klasyczne deskryptory obrazu, cechy SURF, ORB i SIFT. Oceniono błąd pomiaru, oraz czas wykrywania cech wykorzystując CPU i GPU. W rozdz. 7.2 przedstawiono metodę lokalizacji topologicznej w której sieci konwolucyjne służą do wykrywania charakterystycznych cech, a metoda k-najbliższych sąsiadów (zaimplementowana w optymalny sposób) umożliwia dokonanie klasyfikacji topologicznej. Testowano różnego rodzaju sieci. Eksperymentalnie dobrano warstwę sieci (cech), która jest najbardziej odporna na zmiany oświetlenia. Istotne jest to, że zadbane o zbalansowanie zbioru danych. Eksperymenty przeprowadzono w budynku Centrum

Mechatroniki. Dodatkowo przeprowadzono eksperymenty wykorzystując publiczną bazę danych.

Na końcu pracy przedstawiono podsumowanie wyników badań, propozycje przyszłych prac oraz bibliografię. Bibliografia zawiera 318, aktualnych dobrze dobranych pozycji literaturowych.

Recenzowana rozprawa doktorska jest pracą teoretyczno–doświadczalną. Napisana w języku polskim rozprawa została starannie przygotowana, zawiera nieliczne literówki. Przygotowane rysunki, tabele i wykresy nie budzą zastrzeżeń. Autorka przedstawiła propozycję oryginalnego hybrydowego systemu wizyjnego i autorskich algorytmów, które umożliwiają lokalizację robota czasie rzeczywistym.

W opracowanych algorytmach zastosowano nowoczesne narzędzia sztucznej inteligencji i metody równoległego przetwarzania danych. Wszystkie opracowane algorytmy zostały starannie przetestowane. Określono błędy określenia położenia robota i czas przeprowadzonych obliczeń.

W mojej ocenie praca jest wykonana samodzielnie i stanowi oryginalne opracowanie postawionego problemu. Umiejętność ta świadczy o dojrzałości Doktorantki do samodzielnej pracy naukowo-badawczej.

Do najważniejszych osiągnięć pracy, będących w mojej ocenie elementami nowości naukowej, należy zaliczyć:

1. System wizyjny, który łączy kamerę dookólną i klasyczną kamerę.
2. Metoda kalibracji sensora stereowizyjnego zbudowanego z kamer o zupełnie innych cechach.
3. Lokalizacja robota mobilnego na podstawie sztucznych znaczników i hybrydowego systemu wizyjnego.
4. Metoda lokalizacji topologicznej, na podstawie danych pochodzących z kamery dookólnej z wykorzystaniem sieci konwolucyjnej i mapy globalnych deskryptorów.

Podsumowując wykonane w pracy badania należy podkreślić, że Autorka przedstawiła oryginalną koncepcję systemu wizyjnego wzorowanego na sposobie widzenia organizmów żywych. Omówione zostały metody wykorzystania systemu w zagadnieniach lokalizacji metrycznej, topologicznej i semantycznej (dla publicznej bazy danych), wykorzystując znaczniki naturalne i sztuczne. Przedstawiono oryginalną metodę lokalizacji podstawie

danych pochodzących z kamery dookólnej z wykorzystaniem sieci konwolucyjnej i mapy globalnych deskryptorów. Na szczególną uwagę zasługują liczne testy. Sprawdzono nie tylko dokładność określenia położenia robota, ale także czas obliczeń, wykorzystując CPU i GPU. Opracowana metoda ma oprócz znaczenia naukowego, aspekt aplikacyjny. Po niewielkich modyfikacjach może być wykorzystana w robotach usługowych.

3. Uwagi i zapytania

Mam następujące pytania i uwagi:

- W streszczeniu pracy znajduje się sformułowanie: „podstawowym zadaniem robota mobilnego jest znalezienie odpowiedzi na pytanie, gdzie jestem?”, według mnie nie jest prawdą. Robot mobilny powinien realizować różnego rodzaju zadania, np. transportowe, ratownicze itd. Lokalizacja jest potrzebna, aby te zadania mogły być zrealizowane w sposób prawidłowy.
- We wstępie stwierdzono, że w przypadku organizmów żywych, analizę obrazów przeprowadza mózg, pominięto rolę siatkówki, która jest odpowiedzialna za wstępnie przetwarzanie. Rolę tą opisano dopiero w kolejnym rozdziale.
- We wstępie znajdujemy zdanie, że fiksacja wzroku umożliwia wykrywanie ruchu. Warto ten temat omówić nieco szerzej.
- Wstęp jest nieco chaotyczny brakuje spójnego przekazu. Według mnie w sposób niedostateczny uwypuklono podstawową zaletę pracy tzn. połączenie dwóch sposobów zbierania informacji o otoczeniu.
- Rozdział 2.1 zbyt ogólny, zawiera dużo powszechnie znanych informacji dotyczących zastosowania robotów mobilnych. Brakuje szerszej informacji dotyczącej zalet sensorów wizyjnych, ich ograniczeń.
- Rozdział 2.2 ma podobnie jak rozdz. 2.1 charakter popularno-naukowy, opisuje zmysł wzroku różnego rodzaju zwierząt. Autorka wskazuje zależność między sposobem widzenia, a sposobem zachowania i środowiskiem w którym zwierzę się znajduje. Rozdz. 2.2 jest bardzo dobrze napisany, ale warto rozwinąć omawianą właściwość w kierunku zastosowania w robotyce.
- Rys.2.6 jest bardzo ważny - zawiera schemat drogi wzrokowej, nie został jednak szerzej opisany w tekście.

- W rozdz. 2.3 opisano systemy wizyjne w nawigacji. Rozdział zaczyna się zdaniem „uzyskanie podstawowych informacji o otoczeniu jest podstawowym zadaniem robota mobilnego”, podobnie jak w przypadku streszczenia nie zgadzam się z tym zdaniem.
- Chmura punktów przechowuje informacje o otoczeniu w n-wymiarowej tablicy (niekoniecznie $n=2$). Nieuporządkowana chmura punktów niekoniecznie powstaje w wyniku składania chmur punktów, może być wynikiem sposobu zbierania danych.
- Zdanie: montaż kamer w środowisku pracy jest bardzo często używany w systemach wielorobotowych ..., nie jest jasne. (str. 27)
- Tytuł rozdziału 3 jest zbyt szeroki. Autorka opisała wąski fragment metod SI - sieci neuronowe. W rozdziale znajdują się literówki.
- Pisząc o przewadze sieci konwolucyjnych nad metodami wykorzystującymi deskryptory obrazu Doktorantka, napisała, że CNN jest bardziej odporna na zmiany oświetlenia. I w jednym i w drugim przypadku wiele zależy od zbioru uczącego. Według mnie podstawową zaletą sieci konwolucyjnych jest automatyczne generowanie zbioru cech. Nie omówiono jednak problemu, który występuje w przypadku sieci konwolucyjnych – wpływu kontekstu na wynik klasyfikacji, istnieją sytuacje, że przypadkowe elementy tła mogą powodować trudne do przewidzenia błędy klasyfikacji.
- W rozdz. 3.2 przedstawiono architektury sieci CNN. Skupiono się na omówieniu zalet, pominięto jednak istotną wadę – obszar w którym znajduje się obiekt jest prostokątem o bokach równoległych do krawędzi obrazu. Wartość parametru IoU, zależy nie tylko od jakości dopasowania, ale także od orientacji obiektu w przestrzeni.
- Sformułowania: pełność i precyzja oferują kompromis (str. 60) i przyśpieszono niemaksymalne tłumienie, nie są jasne.
- Omawiając obliczenia równoległe skupiono się głównie na technologii CUDA i rozwiązaniach hardwarowych. Pominięto istotne narzędzia takie jak TensorFlow i PyTorch.
- Co to są piksele całkowito-liczbowe?
- Zdanie: bardzo ważnym aspektem jest, aby kamery były ustawione ... co nie jest możliwe w praktyce. (str. 84), jest nieprecyzyjne, tu mamy do czynienia z rozbieżnością między teorią a praktyką. Warto szerzej rozwinąć ten temat.
- Na stronie 108 Autorka pisze o zniekształconym środowisku, według mnie to obraz jest zniekształcony, a nie środowisko

- Pisząc o kamerze dookólnej Doktorantka napisała, że linie proste, na obrazie to łuki, a obrazem kwadratu jest prostokąt. Zdanie to nie jest prawdziwe, obraz prostej zależy od orientacji prostej. A obrazem kwadratu nie jest prostokąt.
- Nie wiem dlaczego w przypadku kamery dookólnej poszukiwanie sztucznych znaczników odbywa się na obrazie przetworzonym, a nie oryginalnym? Według mnie warto wykorzystać informację, że na obrazie krawędzie pionowe są w przybliżeniu odcinkami radialnymi. Analizując, między innymi kąt między liniami radialnymi można wykryć znaczniki (białe kwadraty), bez konieczności dokonywania transformacji.
- Zwykle robot określa swoje położenie w czasie jazdy, w jakich warunkach prowadzone były eksperymenty? Jak prędkość robota wpływa na dokładność lokalizacji? Czy zaobserwowano drgania układu optycznego w czasie jazdy.
- Str. 116 – do obliczenia macierzy zasadniczej zastosowano algorytm Ransac. Wybór algorytmu należy szerzej uzasadnić. Jaki jest czas generowania macierzy.
- Jeśli popatrzymy na rys. 7.4, to możemy zauważyć, że zależność między rzeczywistą odległością, a oszacowaną na podstawie ORB jest mniejsza niż zmierzona (o ok. 10%). Błąd nie ma więc charakteru losowego i można go próbować zredukować. Będę wdzięczna za komentarz.
- W pracy warto określić błąd względny określania odległości do znacznika (procentowy, zależnie od odległości), a nie błąd bezwzględny.
- Badając lokalizację robota na podstawie sztucznych znaczników, nie określono wpływu liczby znaczników na dokładność lokalizacji.
- Metoda lokalizacji topologicznej jest opisana dość jasno, nie jest jednak dla mnie jasne co to jest średni błąd odległości euklidesowej (str. 128). Jeśli jest to błąd określenia położenia robota, to jak to położenie jest obliczane?
- W podpisie pod tabelą 7.7 jest błąd ortograficzny.

4. Ocena końcowa

Praca doktorska Pani mgr inż. Marty Rostkowskiej jest pełnym, wyczerpującym opracowaniem tytułowego zagadnienia. Doktoranta wykazał się znajomością problematyki związanej z tematem pracy. Praca stanowi oryginalne, samodzielne rozwiązanie problemu naukowego i potwierdza nie tylko wiedzę teoretyczną Doktorantki, ale także umiejętności

zastosowania wiedzy w praktyce. Przedstawione wyniki badań odpowiadają najnowszym trendom badawczym robotyki mobilnej. Wybór tematyki pracy doktorskiej uważam za uzasadniony, a wyniki badań stwarzają możliwości aplikacyjne.

Teza pracy: „Wykorzystanie inspirowanej biologicznie koncepcji akwizycji obrazów oraz ich równoległego przetwarzania pozwala realizować w czasie rzeczywistym wybrane funkcje nawigacyjne robota mobilnego w systemie wbudowanym o ograniczonych zasobach obliczeniowych”, jest jasno określona. W pracy w sposób jednoznaczny wykazano jej prawdziwość.

Podsumowując, stwierdzam, że rozprawa doktorska mgr inż. Marty Rostkowskiej, złożona w dyscyplinie Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne spełnia wszystkie warunki przewidziane dla prac doktorskich, zawarte w art. 13 ust.1 ustawy z dnia 14 marca 2003 r. o stopniach i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (tekst jednolity Dz.U. z 2014 r. poz.1852, oraz z 2015r. poz. 249 i 1767) – wnoszę zatem o przyjęcie rozprawy doktorskiej i dopuszczenie mgr inż. Marty Rostkowskiej do dalszego postępowania w przewodzie doktorskim, zgodnie z obowiązującymi przepisami.

B. G.