

dr hab. inż. Andrzej Typiak, prof. WAT
Wojskowa Akademia Techniczna
Wydział Inżynierii Mechanicznej
Instytut Robotów i Konstrukcji Maszyn
ul. gen S. Kaliskiego 2
00-908 Warszawa
andrzej.typiak@wat.edu.pl

Warszawa 13 październik 2023 r.

S E K R E T A R I A T
Rady Dyscypliny AEEITK

Wpłynęło dnia ...20.10.2023...

Zarejestrowano pod nr

Podpis*Jm*.....

RECENZJA
rozprawy doktorskiej mgr. inż. Piotra RECZKA
na temat:

**„Nawe metody rozpoznawania gestów 3D bez użycia kamer,
w aplikacjach branży motoryzacyjnej”**

Promotor dr. hab. inż. Andrzej MŁYNIĘC, prof. AGH
Promotor pomocniczy dr hab. inż. Andrzej WETULA, prof. AGH

Podstawą opracowania recenzji rozprawy doktorskiej mgr. inż. Piotra Reczki było pismo Przewodniczącego Rady Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne Akademii Górniczo-Hutniczej Pana dr hab. inż. Ryszarda SROKI, prof. uczelni z dn. 06.07.2023 roku.

1. Ocena aktualności wybranego tematu

Komunikacja poprzez gesty stanowi znaczący obszar niewerbalnych sposobów porozumiewania się. Mowa ciała może czasami przekazywać informacje bardziej precyzyjnie niż komunikaty słowne. Przekazanie informacji na temat położenia okazuje się często łatwiejsze przy użyciu gestów niż słownie. Można założyć, że system wykrywający odpowiednie gesty może stanowić intuicyjny interfejs dla człowieka, w szczególności mając na uwadze rosnący stopień skomplikowania systemów teleinformatycznych otaczających ludzi w ich przestrzeni życiowej.

Największy odsetek systemów wykrywających gesty stanowią systemy wizyjne. Wydaje się to naturalne, ponieważ jest to sposób najbardziej zgodny z ludzkim sposobem rozpoznawania gestów. Jednak istotnym problemem systemu wizyjnego jest padanie światła słonecznego w obszar roboczy kamery. Powoduje to brak odpowiedniego poziomu kontrastu analizowanego obrazu.

Recenzowana rozprawa doktorska dotyczy opracowania nowatorskiego rozwiązania systemu rozpoznawania gestów w zamkniętej i ograniczonej przestrzeni wnętrza pojazdu bez użycia

kamer wideo. Opracowanie zostało zrealizowane na podstawie analizy istniejących w tej dziedzinie rozwiązań.

2. Cel i teza rozprawy doktorskiej

Cel i zakres pracy zostały opisane w rozdziale 3 rozprawy na stronie 46 i choć nie zostały one przedstawione w sposób bezpośredni to można jego brzmienie odczytać jako następujące: "Celem pracy były badania dotyczące poszukiwań i analizy nowych metod oraz systemów rozpoznających gesty, które mogłyby być zastosowane w branży motoryzacyjnej".

Doktorant podaje, że aby rozwiązanie mogło być stosowane w branży motoryzacyjnej to powinno ono spełnić szereg wymagań:

- poprawna praca systemu w zakresie temperatur od -40 do 85 °C;
- integracja systemu z istniejącymi elementami wnętrza pojazdu;
- koszt dodania systemu do rozpoznawania gestów nie powinien zwiększyć kosztu podstawowego modułu o więcej niż 100%;
- system powinien być możliwy do zintegrowania z produktami z portfolio partnera doktoratu – Firmy Merit, czyli przełącznikami i modułami sterującymi zlokalizowanymi wewnątrz samochodu.

Przyjęta teza pracy została podana również w rozdziale 3 na stronie 47 i brzmi następująco: "Fuzja danych z dwóch systemów pozwoli na uzyskanie lepszych rezultatów niż w przypadku wykorzystania każdego z nich osobno. Dotyczy to głównie skuteczności wykrywanych gestów, co pozwoliłoby na większą niezawodność systemu pozwalającą na zastosowanie w pojeździe. Dodatkowo postawiono tezę, że fuzja danych pozwoli na osiągnięcie lepszych rezultatów w przypadku większej liczby gestów."

Należy zaznaczyć, że podanie w rozprawie doktorskiej jej celu w sposób bezpośredni w sposób znaczący ułatwiłoby prowadzenie dyskusji nad recenzowaną rozprawą a także umożliwiłoby w dogłębniej ocenić osiągnięte rezultaty.

3. Analiza krytyczna treści rozprawy

Rozprawa doktorska napisana jest w języku polskim i zawarta jest na 117 stronach maszynopisu. Treść pracy jest zgodna z jej tytułem. Rozprawa składa się ze streszczenia w języku polskim i angielskim, podziękowań, oraz sześciu numerowanych rozdziałów, którymi są kolejno: wstęp, przegląd literatury oraz istniejących rozwiązań, cel i zakres pracy, system do rozpoznawania gestów oparty o czujniki pojemnościowe, system do rozpoznawania gestów oparty o radar, fuzja danych z systemów pojemnościowego oraz radarowego, podsumowanie oraz możliwe dalsze kierunki rozwoju. Zakończeniem pracy jest numerowany rozdział zawierający pozycje literaturowe. Struktura rozprawy jest przemyślana i logicznie uporządkowana. Pewne zastrzeżenia budzi użyte sformułowanie „oparty o”, choć językoznawcy uważają, że jest ono dopuszczalne to można użyć tytułu rozdziału „radarowy system do rozpoznawania gestów”.

W ramach realizacji pracy, dokonano przeglądu literatury oraz istniejących rozwiązań, które mogłyby zostać zastosowane do rozpoznawania gestów w środowisku wnętrza pojazdu.

Następnie przeanalizowano istniejące systemy w różnych kategoriach: Systemy wizyjne, Systemy wykorzystujące elektrody pojemnościowe, Systemy oparte o czujniki podczerwieni, Systemy radarowe, Systemy ultradźwiękowe, Systemy zlokalizowane na ręce. Każdą z zaproponowanych technologii poddano krytycznej analizie z perspektywy wymagań dla systemów do rozpoznawania gestów z potencjalnym zastosowaniem w branży motoryzacyjnej. Dodatkowo, każdą z nich przeanalizowano pod kątem wykorzystania do utworzenia nowego rozwiązania wykluczając z założenia systemy wizyjne. Na zakończenie przeglądu literatury omówiono fuzję danych oraz klasyfikację metod wykorzystywanych do jej przeprowadzania.

W rozdziałach „System do rozpoznawania gestów oparty o czujniki pojemnościowe” oraz „System do rozpoznawania gestów oparty o radar”, przedstawiono autorskie rozwiązania, w tym zastosowaną architekturę sprzętową, opracowane algorytmy oraz sposoby działania obu systemów. Rozwiązania te zostały opracowane w ramach realizacji projektu wdrożeniowego w firmie Merit. W rozdziale traktującym o systemie pojemnościowym zaproponowano również sposoby rozszerzania danych uczących dla przebiegów sygnału pojemnościowego z poszczególnych elektrod. W końcowej części rozdziału opisano proces konwersji modelu sieci neuronowej oraz uruchomienia go na mikrokontrolerze.

W kolejnym podrozdziale opisana została implementacja systemu działającego w oparciu o czujniki pojemnościowe, który został zamontowany w samochodzie testowym, co stanowiło wdrożeniowy charakter pracy. Na zakończenie rozdziału, przedstawiono nowatorską metodę automatyzacji akwizycji danych uczących dla potrzeb procesu treningu sieci neuronowej przy użyciu 4 osiowego manipulatora, który zastąpił ludzi wykonujących gesty. W rozdziale dotyczącym systemu używającego radaru fali ciągłej z modulacją częstotliwościową, oprócz opisu konstrukcji radaru oraz jego sprzętowej konfiguracji, przedstawiono zastosowane metody przetwarzania danych otrzymywanych bezpośrednio z radaru do postaci cech, które stanowiły wejście do sieci neuronowej klasyfikującej gesty.

W ostatnim rozdziale zostały przedstawione wyniki skuteczności rozpoznawania gestów dla obu opisanych rozwiązań, a także po dokonaniu fuzji danych pochodzących z każdego z nich. Sieć neuronowa wykorzystana do syntezy danych, przyjmowała rozszerzony wektor wejściowy obejmujący 10 cech w postaci 5 sygnałów z czujnika pojemnościowego, 2 przebiegów określających prędkość obiektu względem anten radaru, 2 przebiegów informujących o odległości wykrywanego obiektu od każdej z anten oraz położenia kąтового wykrywanego obiektu w płaszczyźnie wertykalnej radaru. Przeprowadzone eksperymenty wykazały, że modele sieci neuronowej rozpoznające zarówno 5 oraz 7 gestów, dzięki zastosowaniu fuzji danych, osiągnęły ponad 97 procentową skuteczność ich rozpoznawania, podczas gdy sieci neuronowe wykorzystujące dane tylko z jednego typu sensora (elektrod pojemnościowych i radaru) odpowiednio dla 5 gestów 93% i 94,2% oraz dla 7 gestów 87,8% i 92,5%. Otrzymane rezultaty dowiodły postawionej tezie, że fuzja danych z obu systemów spowoduje wzrost skuteczności wykrywanych gestów. W ostatnim rozdziale zaproponowano także obszary, których udoskonalenie powinno zostać rozważone pod kątem dalszego rozwoju przedstawionych systemów, zarówno w kontekście wykorzystania osobno jak i w kontekście ich fuzji.

4. Metoda badawcza i oryginalny dorobek Autora

Praca doktorska zatytułowana: „Nowe metody rozpoznawania gestów 3D bez użycia kamer, w aplikacjach w branży motoryzacyjnej”, składa się z części teoretyczno-analitycznej oraz praktyczno-wdrożeniowej. Pierwsza część, zawiera krytyczną analizę rozwiązań technicznych wykorzystywanych dla potrzeb tytułowej aplikacji oraz systematyzuje klasyfikację metod fuzji danych. Druga część pracy, obejmuje dokładny opis zbudowanych systemów do wykrywania gestów – pojemnościowego i radarowego oraz przeprowadzenie fuzji dla danych otrzymywanych z obu systemów. Praca realizowana była w ramach programu „Doktorat wdrożeniowy” we współpracy z firmą Merit Poland sp. z o.o. będącą jednym z ważniejszych dostawców rozwiązań mechatronicznych interfejsu użytkownika dla branży motoryzacyjnej.

W ramach prowadzonych badań, Doktorant przeprowadził dogłębną analizę istniejących technologii w kontekście ich wykorzystania do rozpoznawania gestów w branży motoryzacyjnej wraz klasyfikacją metod fuzji danych, definiując luki w wiedzy i technologii, niezbędne do wypełnienia w trakcie rozwoju nowego systemu do rozpoznawania gestów. Przeanalizował rozszerzony wektor wejściowy obejmujący 10 cech w postaci 5 sygnałów z czujnika pojemnościowego, 2 przebiegów określających prędkość obiektu względem anten radaru, 2 przebiegów informujących o odległości wykrywanego obiektu od każdej z anten oraz położenia kąтового wykrywanego obiektu w płaszczyźnie wertykalnej radaru. Przeprowadzone eksperymenty wykazały, że modele sieci neuronowej rozpoznające zarówno 5 oraz 7 gestów, dzięki zastosowaniu fuzji danych, osiągnęły ponad 97 procentową skuteczność ich rozpoznawania, podczas gdy sieci neuronowe wykorzystujące dane tylko z jednego typu sensora (elektrod pojemnościowych i radaru) odpowiednio dla 5 gestów 93% i 94,2% oraz dla 7 gestów 87,8% i 92,5%. Otrzymane rezultaty dowiodły postawionej tezie, że fuzja danych z obu systemów spowoduje wzrost skuteczności wykrywanych gestów. W ostatnim rozdziale przedstawiono obszary, których udoskonalenie powinno zostać rozważone pod kątem dalszego rozwoju przedstawionych systemów, zarówno w kontekście wykorzystania osobno jak i w kontekście ich fuzji.

W ramach części praktycznej opracował autorskie rozwiązania bazujące na czujnikach pojemnościowych oraz wykorzystujące radar wraz z systemem służącym do fuzji danych. Przeprowadził ponadto analizę porównawczą skuteczności rozpoznawania gestów przez opracowane systemy: wykorzystujący sensory pojemnościowy, radar oraz z zastosowaniem fuzji danych; wykazując wzrost skuteczności rozpoznawanych gestów dzięki użyciu fuzji danych. Rozpoznawanie gestów w każdym przypadku realizowała indywidualnie wytrenowana sieć neuronowa. Zadanie fuzji w przypadku wykorzystania połączonych systemów, również zostało dokonane za pomocą rozszerzonej sieci neuronowej. Dodatkowo w pracy doktorskiej przedstawiono nowatorską metodę uczenia sieci neuronowej dla systemu pojemnościowego z wykorzystaniem 4-osowego manipulatora, dzięki czemu możliwe było zrezygnowanie z grupy wolontariuszy wykonujących gesty.

Pod względem inżynierskim, praca wyjaśnia zasadę działania technologii stosowanych do rozpoznawania gestów, co stanowi praktyczne opracowanie ułatwiające wybór rozwiązania dla podobnych aplikacji w różnorodnych projektach inżynierskich. Ponadto w częściach

opisujących zrealizowane systemy, zawarto szczegółowy opis wykorzystanych zasobów sprzętowych, ich wzajemnej komunikacji oraz zastosowanych algorytmów. Umożliwia to wykorzystanie opisanych rozwiązań lub ich elementów w przyszłych projektach inżynierskich z tego obszaru w ramach aktywności firmy Merit Poland Sp. z o.o., partnera przemysłowego zrealizowanego doktoratu wdrożeniowego.

5. Niedociągnięcia dostrzeżone w rozprawie i problemy dyskusyjne

Rozprawa spełnia wszystkie kryteria formalne, stawiane pracom doktorskim, jednak zawiera także pewne niedociągnięcia. Autor w kilku miejscach nie uniknął niedopowiedzeń, powiązanych z nieustabilizowaną jeszcze terminologią oraz drobnych zawiłości stylistycznych.

Struktura rozprawy jest przemyślana i logicznie uporządkowana, jednak istotnym niedociągnięciem w strukturze pracy jest brak wykazu użytych oznaczeń i skrótów, który w sposób znaczący ułatwiłby czytanie rozprawy.

W pracy Autor nie ustrzegł się także błędów językowych i nieprecyzyjnych sformułowań i skrótów myślowych.

Str. 3: „Następnie przeanalizowano istniejące systemy w następujących kategoriach: Systemy wizyjne, Systemy wykorzystujące elektrody pojemnościowe, Systemy oparte o czujniki podczerwieni, Systemy radarowe, Systemy ultradźwiękowe, Systemy zlokalizowane na ręce.” – pomijając błędne pisanie słowa systemy (z wielkiej litery) to jakiego rodzaju systemami są Systemy zlokalizowane na ręce?;

Str. 14: Rysunek 2 – literą t oznaczamy czas, nie odległość;

Str. 16: „porównywanie ładunku zgromadzonego w kondensatorze referencyjnym oraz w wirtualnym – stworzonym przez ludzkie ciało.” – ludzkie ciało to nie jest kondensator wirtualny.

Str. 25: „Rysunek 9. Zastosowanie radarów w samochodzie” czy na rysunku przedstawiono zastosowanie radarów czy też obszary wokół samochodu które są „obserwowane” przez poszczególne radary?

Str. 55: W akapicie drugim podano „Na rysunku 29 przedstawiono schemat zastosowanej sieci neuronowej...”, w podpisie pod rysunkiem zamieszczono „Rysunek 29. Model zastosowanej sieci neuronowej dla rozpoznawania gestów z wykorzystaniem systemu wykorzystującego elektrody pojemnościowe” – natomiast zdaniem recenzenta jest to algorytm rozpoznawania gestów – podobnie rysunek 66.

Str. 58: Rysunki na str. 57 i 58 mają ten sam numer.

Str. 64.: Tabela 4 „Podobnie jak w przypadku Categorical Cross Entropy bazuje na logarytmie.” – o jakim logarytmie mowa?

Str.71: „Zagadnienie zbierania danych jest wąskim gardłem w procesach uczenia maszynowego w niemal wszystkich przypadkach” – slogan (wąskie gardło);

Str. 81.: „Należy zaznaczyć, że w dalszej części używana jest zgodnie z dokumentacją układu termin „częstotliwość pośrednia”, choć w praktyce nie mamy tu do czynienia z klasyczną demodulacją heterodynową. Termin ten w opisywanym kontekście oznacza częstotliwość różnicową, a nazwa wynika z faktu, że demodulacja nie jest ostatnim krokiem analizy.” W

jaki sposób określenie „częstotliwość pośrednia” odnosi się to faktu, że demodulacja nie jest ostatnim krokiem analizy?

Powyższe uwagi jak również znajdujące się w pracy drobne wady redakcyjne nie ujmują wartości merytorycznej rozprawy, szczególnie na tle przedstawianych w rozprawie interesujących i nowatorskich, możliwych do skomercjalizowania, wyników badań naukowo-technicznych.

Uwagi problemowe:

1. Na str. 26 przedstawiono „Rysunek 10. Wykres przedstawiający przesunięcie częstotliwościowe $f(t)$ między sygnałem wysłanym oraz odebrany [29]” – proszę przeanalizować czy na wykresie przedstawiono przesunięcie częstotliwościowe czy przesunięcie czasowe.
2. Na str. 48. Przedstawiono „Rysunek 22. Schemat systemu rozpoznającego gesty w oparciu o sygnał z elektrod pojemnościowych. ... „ – zamieszczony rysunek nie przedstawia schematu systemu tylko układy scalone użyte do budowy systemu. Proszę przedstawić schemat systemu z wyszczególnieniem poszczególnych układów i opisem ich roli w systemie. – podobnie rysunek 46.
3. Na str. 58 podano :” W aplikacji systemu pojemnościowego do zbierania danych wykorzystano pełny system przedstawiony na rysunku 22”. Czy można wykorzystywać nie pełny system? Jeżeli tak to proszę o rozwinięcie tego zagadnienia
4. Na str. 108 w tabeli 7 przedstawiono skuteczność rozpoznawania gestów dla systemu pojemnościowego, radarowego oraz po dokonaniu fuzji na podstawie danych z obu systemów. Dla 5 gestów system radarowy rejestrował poprawnie 93.0% gestów a pojemnościowy 94.2%. Dzięki fuzji danych z obu systemów poprawnie odczytano 97.7% gestów, czyli wzrost o 4,7 punktu procentowego względem systemu radarowego i 3,5 punktu procentowego względem systemu pojemnościowego. Natomiast dla 7 gestów system radarowy rejestrował poprawnie 87.8% gestów a pojemnościowy 92.5%. Dzięki fuzji danych z obu systemów poprawnie odczytano 97.1 % gestów, czyli wzrost o 10,7 punktu procentowego względem systemu radarowego i 4,6 punktu procentowego względem systemu pojemnościowego. Czy analizował Pan przyczynę wzrostu działania systemu wykorzystującego fuzję danych dla większej liczby gestów?
5. „Analizując otrzymane wyniki przez pryzmat fuzji danych, pierwszym aspektem, którego poprawa może przynieść jeszcze lepsze rezultaty, jest większa integracja obu systemów.” – w jaki sposób zamierza Pan zwiększyć integrację systemów?
6. Na str. 110 podano „Projekt został zakończony pomyślnie, ponieważ zaprojektowano i przetestowano rozwiązanie niewymagające użycia systemów wizyjnych, w postaci systemu pojemnościowego.” Czy wystarczy zaprojektować i przetestować układ by stwierdzić, że projekt został zakończony pomyślnie. Czy nie jest wymagane, aby porównać otrzymane rezultaty z przyjętymi na wstępie założeniami aby określić poziom realizacji?

6. Końcowa ocena rozprawy doktorskiej

Recenzowana rozprawa doktorska mgr inż. Piotra Reczka mieści się w dyscyplinie automatyka, elektronika, elektrotechnika i technologie kosmiczne. Moja ocena pracy jest pozytywna, pomimo wykazanych niedociągnięć. Praca zawiera szereg wyników poznawczych. Uzupełniają one w tym zakresie stan wiedzy z dynamicznie rozwijającej się tematyki pojazdów autonomicznych w branży motoryzacyjnej. Rozprawa ta przede wszystkim wnosi oryginalne wnioski w zastosowaniu do praktyki wdrożeniowej.

Na podstawie wnikliwej analizy przedstawionej do recenzji rozprawy doktorskiej potwierdzona została słuszność postawionej na wstępie pracy tezy twierdzącej, iż:

„Fuzja danych z dwóch systemów pozwoli na uzyskanie lepszych rezultatów niż w przypadku wykorzystania każdego z nich osobno. Dotyczy to głównie skuteczności wykrywanych gestów, co pozwoliłoby na większą niezawodność systemu pozwalającą na zastosowanie w pojeździe. Dodatkowo postawiono tezę, że fuzja danych pozwoli na osiągnięcie lepszych rezultatów w przypadku większej liczby gestów.”

Doktorant udowadniając tezę rozwiązał problemy badawcze w niej zawarte. Wykazał się także dobrym warsztatem naukowym, dobierając odpowiednie metody badawcze, a także udowodnił, że potrafi analizować i oceniać uzyskane rezultaty oraz formułować wnioski. Wykazał się przy tym bardzo dobrym przygotowaniem merytorycznym co stawia go w szeregu bardzo dobrych specjalistów w zakresie sterowania pojazdami autonomicznymi.

Uważam, że recenzowana rozprawa doktorska Pana mgr. inż. Piotra Reczka spełnia warunki określone przez Ustawę z dnia 20 lipca 2018 roku Prawo o szkolnictwie wyższym i nauce (Dz. U. 2023 poz. 742). Wnoszę zatem o dopuszczenie rozprawy doktorskiej mgr. Piotra Reczka, przez Radę Dyscypliny Automatyka, Elektronika, Elektrotechnika i Technologie Kosmiczne, Akademii Górniczo-Hutniczej, do dalszego procedowania.



