

Autoreferat

1. Imię i nazwisko.

Dr inż. Tomasz Grzejszczak

2. Posiadane dyplomy, stopnie naukowe lub artystyczne – z podaniem podmiotu nadającego stopień, roku ich uzyskania oraz tytułu rozprawy doktorskiej.

Stopień naukowy doktora - Politechnika Śląska w Gliwicach, Wydział: Automatyki, Elektroniki i Informatyki, Dyscyplina: Automatyka i Robotyka, specjalność: Wizja Komputerowa i Przetwarzanie Obrazów, Tytuł pracy: „Wykrywanie i lokalizacja punktów charakterystycznych dłoni w obrazach cyfrowych” Promotorzy: dr hab. inż. Adam Gałuszka, dr inż. Michał Kawulok. Data podjęcia uchwały o nadaniu stopnia doktora: 28.04.2015.

Stopień naukowy magistra inżyniera - Politechnika Śląska w Gliwicach, Wydział: Automatyki, Elektroniki i Informatyki, Kierunek: Makrokierunek (studia stacjonarne, jednolite, w języku angielskim) Specjalizacja: Automatyka i Robotyka, Studia zakończone obroną pracy magisterskiej pt. „Semantyczna reprezentacja Świata Klocków w zastosowaniu do sterowania manipulatorem ARM1”.

3. Informacja o dotychczasowym zatrudnieniu w jednostkach naukowych lub artystycznych.

a) Zatrudnienie w jednostkach naukowych

2015 - nadal Adiunkt (przed 2019 – asystent), Politechnika Śląska w Gliwicach, Wydział Automatyki, Elektroniki i Informatyki Politechniki Śląskiej, Katedra Automatyki i Robotyki (dawniej Zakład Sterowania i Robotyki)

b) Staże naukowe i dydaktyczne w jednostkach naukowych

2017 3 miesięczny staż naukowy w Niemczech na: University of Wuppertal, School of Electrical, Information, Media and Engineering, 42119 Wuppertal, Germany, Department of Automation/Computer Science, Rainer-Gruenter-Str. 21

2019 1 tygodniowy staż dydaktyczny w Rosji na: Uniwersytet: ITMO, Faculty of Control Systems and Robotics, Rosja, St. Petersburg, Ulitsa Lomonosova, 9, 1417

2025 1 miesięczny staż dydaktyczny w Chinach na: Yanshan University, Silesian College of Intelligent Science and Engineering, Qinhuangdao, Hebei.

c) Zatrudnienie w projektach naukowo badawczych (B+R)

2024 - nadal Konsultant B+R w projekcie SMART, Tritem P.S.A., Ligocka 103.11 Katowice

2022 - 2023 Ekspert ds. Sztucznej Inteligencji w projekcie GPW-PCOL finansowanym z NCBiR. Projekt pt. „Opracowanie Polskiego Cyfrowego Operatora Logistycznego – PCOL” POIR.01.01.01-00-0142/21

2021 - 2023 Kierownik zespołu B+R w projekcie NCBiR: „Modułowy, wielofunkcyjny, zdalnie sterowany pojazd podwodny typu Bottom Crawler Work Class” realizowanego na zlecenie SR Robotics sp. z o.o. POIR.01.01.01-00-1023/21

2020 - 2022 Kierownik zespołu B+R finansowanego z NCBiR dla Infotower Technologies sp. z o.o. „Opracowanie zaawansowanego algorytmu do selekcji danych multimedialnych oraz innowacyjnej metody wizualizacji tych danych na potrzeby platformy wspierającej obsługę klienta w branży turystycznej” POIR.01.01.01-00-1111/19

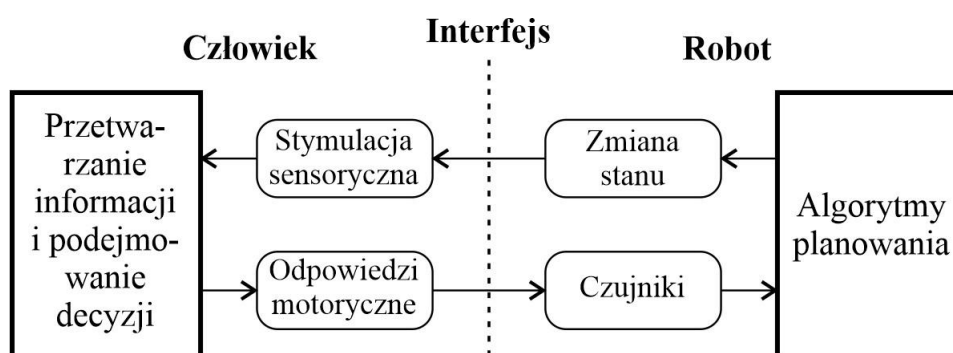
- 2019 - 2021 Data Scientist w projekcie GPW finansowanym z NCBiR w trakcie realizacji 2 etapów projektu: „Platforma GPW Data jako innowacyjny system wykorzystujący techniki sztucznej inteligencji celem wspierania decyzji inwestycyjnych” POIR.01.01.01-00-0162/19
- 2019 - 2021 Wykonawca w projekcie badawczym: „Inteligentny, efektywny system prowadzenia specjalistycznych prac podwodnych” realizowanego na zlecenie SR Robotics sp. z o.o. POIR.01.01.01-00-0266/18
- 2017 Kierownik projektu: „Wyjazd stażowy w celu prowadzenia badań i poszerzania wiedzy z zakresu wykrywania gestów” finansowanego przez NCN – MINIATURA1
- 2015-2019 Kilukrotne zlecenie przez Menard Polska S. z o.o. na wykonywanie skryptów i programów na potrzeby badawczo rozwojowe.

d) Zatrudnienie w instytucjach oceniających wnioski o finansowanie badań

- 2017 - 2019 Ekspert NCBiR – umowa na zlecenie, ocena wniosków w ramach szybkiej ścieżki POIR.01.01.01 (ocena 4 wniosków, z czego 1 raz w roli eksperta wiodącego)
- 2023 Ekspert NCBiR – ocena 1 wniosku w ramach INFOSTRATEG IV

4. Omówienie osiągnięć, o których mowa w art. 219 ust. 1 pkt. 2 ustawy z dnia 20 lipca 2018 r. Prawo o szkolnictwie wyższym i nauce (Dz. U. z 2021 r. poz. 478 z późn. zm.). Omówienie to winno dotyczyć merytorycznego ujęcia przedmiotowych osiągnięć, jak i w sposób precyzyjny określać indywidualny wkład w ich powstanie, w przypadku, gdy dane osiągnięcie jest dziełem współautorskim, z uwzględnieniem możliwości wskazywania dorobku z okresu całej kariery zawodowej.

W ramach prac prowadzonych po doktoracie, badania skupiały się nad dziedziną interakcji człowieka z robotem (HRI – ang. Human Robot Interaction), a dokładniej nad procesiem przetwarzania danych pozyskanych z różnych czujników stanowiących system percepcyjny robota.



Rysunek 1 Pętla interakcji Człowieka z Robotem (HRI)

Interakcję człowieka z robotem lub inną maszyną można przedstawić jako pętlę. Zgodnie z rys. 1, najpierw decyzja jest formułowana w ludzkim mózgu i wyrażana za pomocą układu nerwowego i mięśni, tworząc reakcje motoryczne. Reakcje te są notowane i przetwarzane przez czujniki, w które robot jest wyposażony. Następnie podejmowane są decyzje przez algorytm planowania. Wynikiem algorytmu planowania jest wykonanie akcji za pomocą dostępnych siłowników, które przyczyniają się do zmiany stanu maszyny i środowiska. Zmiany stanu można również wizualizować za pomocą

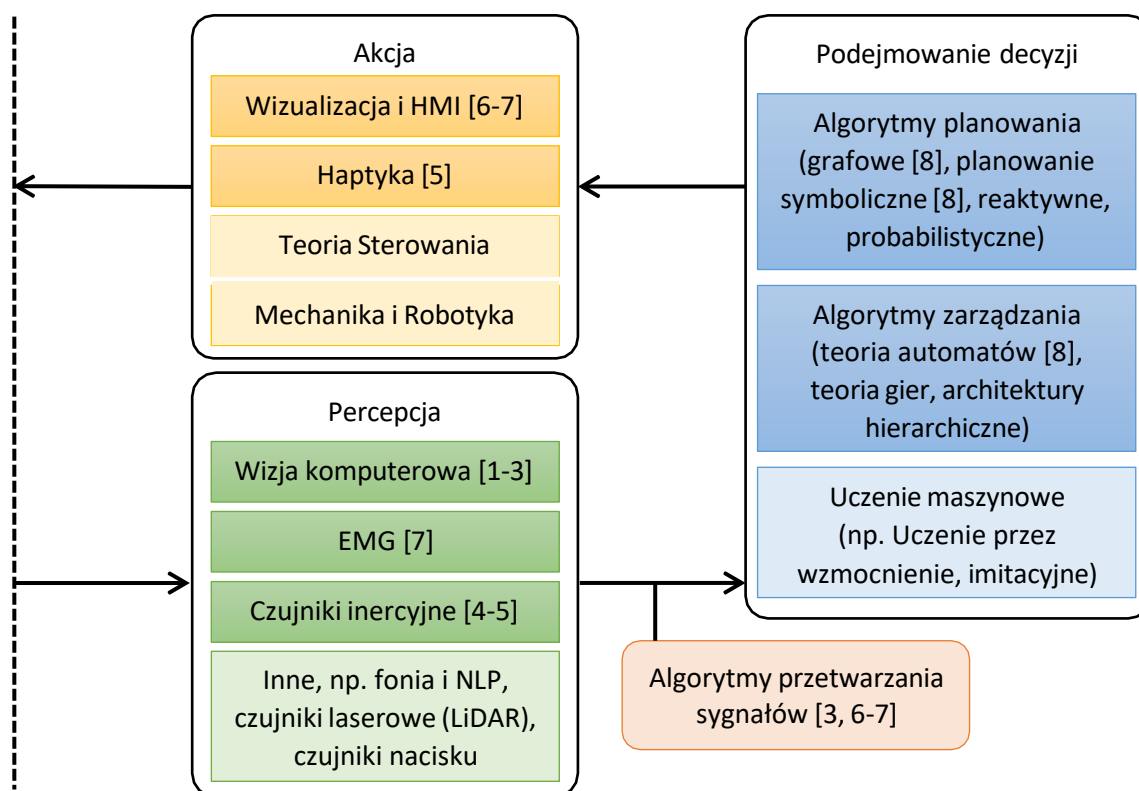
wyświetlaczy w postaci wartości, wykresów lub symulacji. Dane są odbierane przez ludzkie zmysły, czyli neuronalne bodźce sensoryczne, takie jak wzrok lub receptory skóry. Możliwość obserwacji stanu maszyny zachęca człowieka do przetwarzania informacji i podejmowania nowych decyzji, co zamyka pętlę interakcji. Kluczowym elementem łańcucha interakcji człowieka z robotem jest interfejs, czyli sposób połączenia i przekazywania danych (miejsce przecięcia z linią przerywaną na rys. 1), składający się z 2 aspektów:

- Odpowiednie dostosowanie zestawu czujników robota do możliwości motorycznych ludzkiego ciała (ergonomia) oraz właściwe ich przetwarzanie tak aby zapewnić najlepszą wiedzę o stanie środowiska dla algorytmów planowania,
- Odpowiednie wykonywanie ruchów i wyświetlanie stanu, tak aby dostarczyć użytkownikowi najlepszej wiedzy o działaniach robota (User Interface) w celu zachęcenia i wzmocnienia interakcji oraz wywołania zjawiska immersji.

Efektem badań w dziedzinie interakcji człowieka z robotem (HRI) jest poznanie sposobu działania ludzkiego ciała (psychologia) tak aby tworzyć maszyny najbardziej komunikatywne i kompatybilne z potrzebami ludzi (robotyka). **W ramach badań własnych wnioskodawcy sprawdzano metody akwizycji danych analizujących ruchy użytkownika pobrane z użyciem różnych czujników, rozwijając metody przetwarzania danych w postaci szeregów czasowych, w celu rozwoju algorytmów planowania i podejmowania decyzji w robotyce.** Prawie każda z 54 publikacji zawiera element badający aspekty pętli HRI przedstawionej na rys. 1 - 2. Z załączonych publikacji wybrano najbardziej znaczące artykuły naukowe o największym wkładzie wnioskodawcy, przedstawiające powiązany tematycznie zbiór stanowiących znaczny wkład w rozwój określonej wyżej dyscypliny. Cykl artykułów naukowych to:

- [1] **Grzejszczak Tomasz**, Kawulok Michał, Gałuszka Adam: Hand landmarks detection and localization in color images, Multimedia Tools and Applications, Springer, vol. 75, nr 23, 2016, s. 16363-16387, DOI:10.1007/s11042-015-2934-5
- [2] **Grzejszczak Tomasz**, Molle R., Roth R.: Tracking of dynamic gesture fingertips position in video sequence, Archives of Control Sciences, Sciendo, vol. 30, nr 1, 2020, s. 101-122, DOI:10.24425/acs.2020.132587
- [3] **Grzejszczak Tomasz**, Roth R., Moller R.: Hand gesture classification with use of convolutional neural networks, W: 32nd Annual European Simulation and Modelling Conference : ESM 2018, Ghent, Belgium, 24-26 October 2018, 2018, EUROSIS, ISBN 978-9492859-05-1, s. 49-53
- [4] Babiartz Artur, Bieda Robert, **Grzejszczak Tomasz**, Jaskot Krzysztof, Kozyra Andrzej, Ściegienka Piotr: Underwater human arm manipulator - system calibration, Advances in Science and Technology Research Journal, Society of Polish Mechanical Engineers and Technicians (SIMP), vol. 18, nr 2, 2024, s. 36-46, DOI:10.12913/22998624/183569
- [5] Babiartz Artur, Bieda Robert, Borowik Tomasz, **Grzejszczak Tomasz**, Hartwig Tomasz, Jaskot Krzysztof, Kozyra Andrzej, Ściegienka Piotr: Underwater manipulator that imitates the movements of the human arm, Frontiers in Marine Science, Frontiers Media S.A., vol. 10, 2023, Numer artykułu: 1271185, s. 1-10, DOI:10.3389/fmars.2023.1271185
- [6] **Grzejszczak Tomasz**, Rokseła Anna, Poświata Anna, Siemianowicz Anna, Kiełboń Agnieszka, Mikulski Michał: Surface electromyography data analysis for evaluation of physical exercise habits between athletes and non-athletes during indoor rowing, Sensors, Multidisciplinary Digital Publishing Institute, vol. 24, nr 6, 2024, Numer artykułu: 1964, s. 1-16, DOI:10.3390/s24061964
- [7] **Grzejszczak Tomasz**, Probiez Eryka, Gałuszka Adam, Simek Krzysztof, Jędrasiak Karol, Wiśniewski Tomasz: Dynamic time warping in financial data - modification of algorithm in context of stock market similarity analysis, European Research Studies Journal, University of Piraeus, International Strategic Management Association, vol. 25, nr 1, 2022, s. 967-979, DOI:10.35808/ersj/2897
- [8] **Grzejszczak Tomasz**, Nocoń Witold: Cooperative and individual planning strategies for autonomous robots in warehouse parcel transportation, Advances in Science and Technology Research Journal, Society of Polish Mechanical Engineers and Technicians (SIMP), vol. 19, nr 1, 2025, s. 406-418, DOI:10.12913/22998624/195453

Cykl publikacji podzielono na 5 aspektów, które przedstawiono w następnych podrozdziałach. **Rozdziały 4.1 - 4.3** przedstawiają metody przetwarzania danych pozyskanych z różnych czujników, w tym kamery [1-3] (**rozdz. 4.1**), czujników inercyjnych montowanych na rękę [4-5] (**rozdz. 4.2**) oraz czujników elektromiograficznych EMG [7] (**rozdz. 4.3**). Przetwarzane dane to gesty użytkowników. Akwizycja danych w czasie sprawia, że przetwarzanymi danymi są szeregi czasowe. W **Rozdziale 4.4** przedstawiono jak metody używane do przetwarzania szeregów czasowych mogą znaleźć zastosowanie w innych problemach naukowych, w tym przypadku do przetwarzania danych giełdowych [6]. Następnym krokiem w pętli HRI jest algorytm planowania. **Rozdział 4.5** przedstawia koncepcje planowania w robotyce z użyciem informacji pozyskanych z czujników percepcyjnych [8]. Powiązanie tematyczne wybranych artykułów naukowych przedstawiono na rys. 2.



Rysunek 2 Fragment Pętli interakcji Człowieka z Robotem (HRI) z opisem algorytmów

4.1. Rozpoznawanie gestów dłoni przetwarzając obraz cyfrowy

[1] Hand landmarks detection and localization in color images – to pierwsza kluczowa publikacja, która ukazała się po obronie doktoratu. W publikacji przedstawiono rozwinięcie badań z doktoratu i dodatkowe testy algorytmu lokalizacji punktów charakterystycznych dłoni w obrazach cyfrowych.

OŚWIADCZENIE O WKŁADZIE:

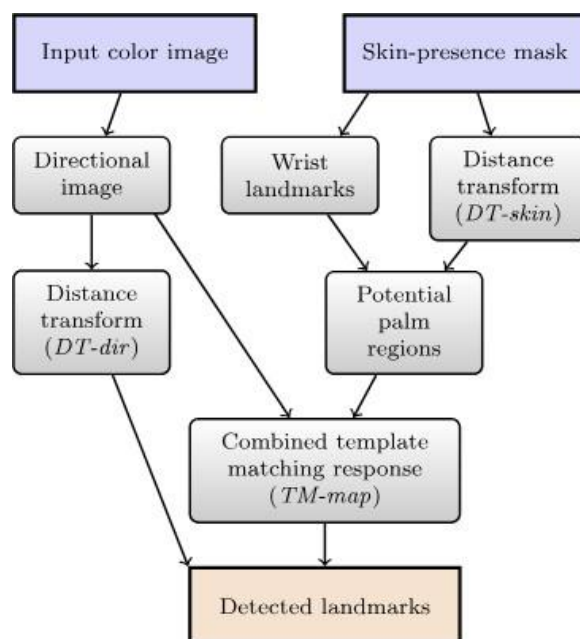
Tomasz Grzejszczak, jako pierwszy autor, miał największy wkład w proces powstawania artykułu. Artykuł powstał bazując na jego koncepcji, opracował metodologię, utworzył oprogramowanie, przeprowadził analizę i walidację wyników oraz napisał treść wraz z wizualizacją wyników. Główny wkład pozostałych autorów to nadzór merytoryczny, recenzja, gromadzenie danych testowych, konsultacja w sprawie proponowanej koncepcji, metodologii i oprogramowania



W roku 2016 popularny trend rozwoju technik interakcji człowiek-komputer (HCI) miał na celu przede wszystkim zwiększenie intuicyjności sterowania. Ekrany dotykowe stały się bardzo popularne, i pojawiały się pierwsze bezdotykowe interfejsy oparte na gestach jako rozwiązania komercyjne. Wprowadzenie sensora Kinect znacznie poprawiło możliwości rozpoznawania gestów, jednak w celu zwiększenia dostępności interfejsów bezdotykowych, bardziej efektywne rozwiązania oparte na wizji komputerowej pomogłyby wykorzystać dużą popularność i niską cenę charakterystyczną dla standardowych kamer 2D. Systemy oparte na wizji wciąż jednak nie pozwalały na rozpoznawanie gestów z wystarczającą dokładnością, co rodziło potrzebę opracowania nowych algorytmów analizy obrazu do rozpoznawania gestów. Jednym z głównych problemów, poruszonych w niniejszym artykule, jest wykrywanie i lokalizowanie punktów charakterystycznych dłoni, takich jak opuszki palców, stawy lub paliczki.

W ramach prac związanych z doktoratem opracowano nowatorski algorytm rozpoznawania gestów, który najpierw przekształca obraz wejściowy dłoni w obraz kierunkowy, a następnie wykonuje transformację odległości na podstawie wyrenderowanego obrazu kierunkowego i konturu dłoni. Lokalne maksima na mapie odległości tworzą zestaw kandydatów na punkty charakterystyczne, który jest filtrowany w oparciu o dane wyodrębnione za pomocą dopasowania szablonu. Na koniec oznaczane są punkty charakterystyczne za pomocą zestawu reguł heurystycznych. Wykorzystując przekształcenie odległości wykonane na obrazie kierunkowym w celu wykrycia punktów charakterystycznych dłoni, co umożliwia ich znalezienie niezależnie od tego, czy znajdują się na konturze dłoni, czy wewnątrz sylwetki dłoni. Kontrastuje to z alternatywnymi istniejącymi technikami, które koncentrują się wyłącznie na konturze i nie wykrywają punktów charakterystycznych zgiętych palców, ponieważ najczęściej bazują na konturze. Wyniki obszernego badania eksperymentalnego wyraźnie pokazują zalety proponowanej metody w porównaniu z szeregiem alternatywnych, najnowocześniejszych podejść. Ponadto wyniki zidentyfikowały różne aspekty analizowanych algorytmów, które mogą być pomocne w projektowaniu nowej metody hybrydowej.

Ogólny schemat blokowy algorytmu został przedstawiony na Rys. 3, a jego kolejne kroki wraz z przykładowymi wynikami zostały zwizualizowane na Rys. 4. Kolorowy obraz wejściowy (rys. 4a) wraz z maską regionu skóry (rys. 4b) są przetwarzane w trzech łańcuchach w celu: (i) zlokalizowania linii nadgarstka (czerwone strzałki na rys. 4), (ii) znalezienia potencjalnych obszarów dłoni (zielone strzałki) oraz (iii) wykrycia kandydatów na punkty charakterystyczne (niebieskie i fioletowe strzałki)

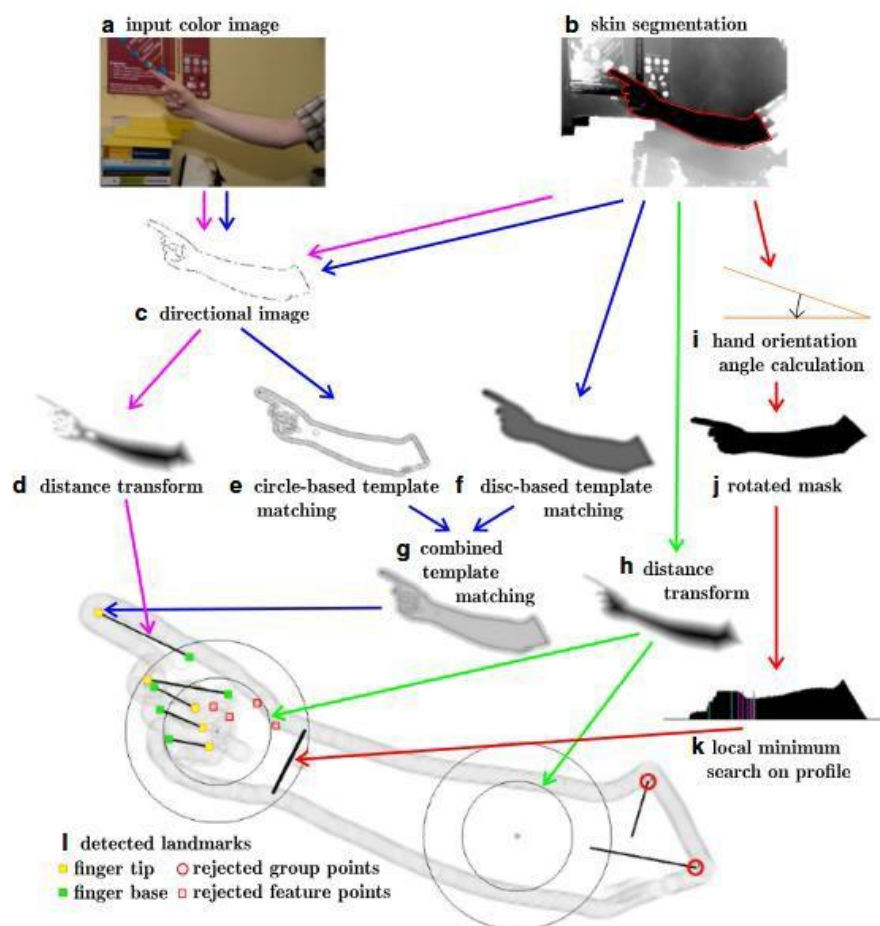


Rysunek 3 Schemat proponowanego algorytmu lokalizacji punktów charakterystycznych dłoni [1]

W celu zlokalizowania linii nadgarstka (tj. punktów określanych jako W 1 i W 2), najpierw szacowany jest kąt orientacji dłoni (rys. 4i). Dla maski obróconej o kąt orientacji wyznaczany jest poziomy profil (rys. 4j), który jest analizowany w celu określenia pozycji nadgarstka (rys. 4k) metodą poszukiwania minimum z użyciem ekspansji i kontrakcji przedziału. Wykryta linia nadgarstka dzieli całą maskę na dwie części, a dla każdej części znajduje się globalne maksimum na mapie odległości uzyskanej po zastosowaniu transformacji odległości do konturu dłoni (rys. 4h). Każde maksimum (określane jako P1 i P2) jest uważane za środek potencjalnego obszaru dłoni o promieniu wyznaczonym z wartości globalnego maksimum transformacji odległościowej.

Kandydaci na punkty charakterystyczne są określani jako lokalne krzywizny wykryte na obrazie kierunkowym dłoni (rys. 4c), a także w konturze maski dłoni. Pierwszy z nich podlega dopasowaniu szablonu (Template matching) opartego na okręgu (rys. 4e), podczas gdy drugi jest przetwarzany za pomocą dopasowania szablonu opartego na dysku (rys. 4f), a te dwie odpowiedzi dopasowania szablonu są łączone ze sobą (rys. 4g). Obraz kierunkowy przedstawia kontur dłoni wraz z wewnętrznymi konturami cyfr wewnątrz maski skóry pozyskany z wizualizacji lokalnego gradientu spadku wartości pikseli na obrazie źródłowym. Szablon w kształcie koła zastosowany do obrazu kierunkowego umożliwia zlokalizowanie opuszków palców zgiętych występujących wewnątrz konturu dłoni, podczas gdy mapa dopasowania szablonu w kształcie dysku zastosowana do maski dłoni podkreśla opuszki palców wyprostowanych palców. Promień szablonu jest ustawiany proporcjonalnie na podstawie wykrytego rozmiaru obszaru dłoni. Następnie każde lokalne maksimum w połączonej odpowiedzi dopasowania szablonu jest traktowane jako kandydat na punkt charakterystyczny. Korzystając z mapy transformacji odległości (rys. 4d) uzyskanej z obrazu kierunkowego, kandydaci na punkty charakterystyczne są weryfikowani na podstawie potencjalnych obszarów dłoni.

Obszar dłoni, w którym wykryto więcej punktów charakterystycznych, jest uznawany za prawidłowy, podczas gdy drugi jest odrzucany. Kandydaci na punkty charakterystyczne powiązane z zaakceptowanym obszarem dłoni są uznawane za ostatecznie wykryte punkty charakterystyczne (rys. 4l).



Rysunek 4 Przykłady wyników uzyskanych na kolejnych etapach wykrywania punktów charakterystycznych

Przedstawione punkty algorytmu: (c-l) są autorstwa wnioskodawcy. Samą metodę wyliczenia wartości gradientów lokalnych i wizualizacji w formie obrazu kierunkowego (c) opracował współautor, jednak sposób użycia i dalszego przetworzenia (d-e) tego obrazu należy do wnioskodawcy. Dokładny sposób implementacji i metodę działania można znaleźć w artykule.

Ponadto, w ramach publikacji przedstawiono obszerne badania metody z porównaniem do istniejących wiodących na tamten czas rozwiązań. W przypadku badanych zestawów można zauważyć, że proponowana metoda pozwala wykryć więcej punktów charakterystycznych niż istniejące podejścia.

[2] Tracking of dynamic gesture fingertips position in video sequence –

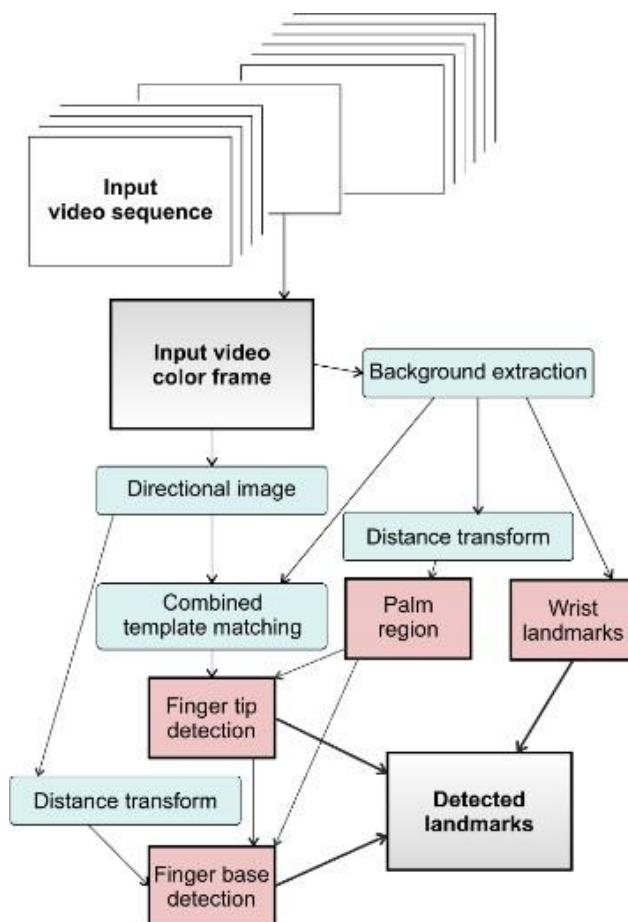
kolejnym etapem prac nad algorytmami rozpoznawania gestów była analiza lokalizacji wykrytych punktów charakterystycznych w czasie. Pozwoliło to na rozpoznawanie gestów dynamicznych. Badania zakładały przedstawienie zestawu wykrytych punktów danej klasy jako jeden zbiór podlegający regresji liniowej.

OŚWIADCZENIE O WKŁADZIE:

Tomasz Grzeszczak, jako pierwszy autor, miał największy wkład w proces powstawania artykułu. Artykuł powstał bazując na jego koncepcji, opracował metodologię, utworzył oprogramowanie, przeprowadził analizę i walidację wyników oraz napisał treść wraz z wizualizacją wyników. Główny wkład pozostałych autorów to nadzór merytoryczny, recenzja, pomoc przy gromadzeniu danych testowych, konsultacja w sprawie proponowanej koncepcji, metodologii i oprogramowania.

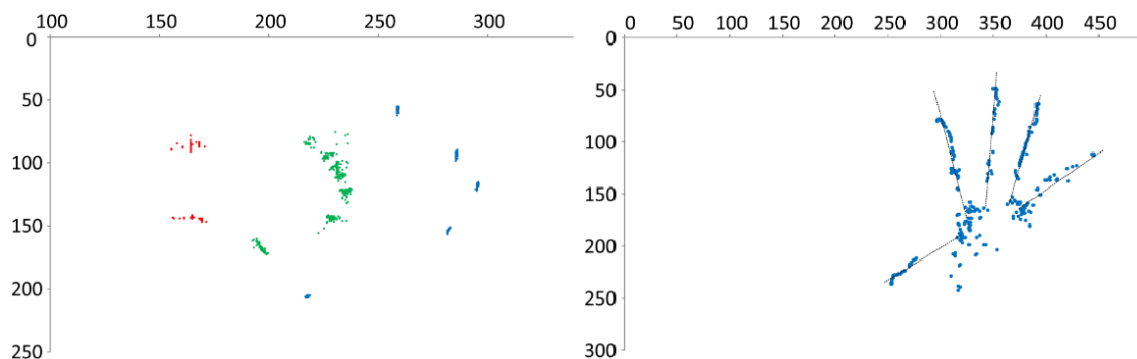


Wraz z rozwojem dziedziny interakcji człowieka z komputerem powstawały dedykowane czujniki, takie jak kontroler ruchu Kinect czy LeapMotion – do rozpoznawania ułożenia dłoni. Czujniki te były to dedykowane rozwiązania bazujące na mapach głębi generowanych za pomocą zjawiska Time of Flight. Celem niniejszych badań było prześledzenie możliwości rozwiązań opartych na czujnikach głębokości, ale z wykorzystaniem standardowych obrazów z aparatów cyfrowych, które są łatwo dostępne i powszechnie stosowane w wielu komputerach lub smartfonach. Dlatego też, spośród wszystkich różnych podejść wizyjnych, badania przedstawione w niniejszym artykule przeprowadzono z wykorzystaniem systemu wizyjnego przetwarzającego kolejne obrazy 2D zapisane jako sekwencja wideo. Dla każdej klatki wideo zbiór wykrytych punktów charakterystycznych jest dodawany do danych szeregu czasowego zawierającego całą ścieżkę ruchu punktów charakterystycznych. Wkład w rozwój dziedziny polegał na zastosowaniu reguł heurystycznych dotyczących obrazu kierunkowego i transformacji odległości do śledzenia wykrytych punktów charakterystycznych dłoni w sekwencji wideo. Techniki zostały opracowane i porównane ze stanem techniki w poprzednich badaniach nad detekcją i lokalizacją punktów charakterystycznych dłoni na stacjonarnych obrazach kolorowych. W niniejszym artykule wejściem jest sekwencja wideo, więc metoda została zmodyfikowana w celu ciągłego śledzenia punktów charakterystycznych dłoni. W niniejszym badaniu zastosowano metodę opartą na obrazie kierunkowym, ponieważ alternatywne techniki koncentrują się wyłącznie na konturze i nie wykrywają punktów charakterystycznych palców zgiętych. Wyniki badań eksperymentalnych jednoznacznie wskazują na zalety i możliwości ciągłego śledzenia punktów charakterystycznych dłoni, zarówno w przypadku palców wyraźnie uniesionych, jak i zgiętych. Modyfikacja algorytmu przedstawiona jest na rysunku 5.



Rysunek 5 Schemat blokowy wykrywania i lokalizacji punktów charakterystycznych dłoni [2]

W tamtych czasach wiodące metody bazowały na analizie konturów, która zawodziła w przypadku, gdy punkty charakterystyczne znajdowały się wewnątrz konturu dłoni. Najbardziej charakterystyczną cechą tego podejścia jest możliwość wykrywania opuszków palców zgiętych. Przedstawione eksperymenty mają na celu weryfikację dokładności śledzenia palców w sekwencjach, w których palce są zgięte, zachodzą na siebie lub są wykrywane wewnątrz konturu dłoni. Zaproponowano trzy testy, aby w pełni potwierdzić i udowodnić tę koncepcję. Podczas każdego testu prezentowany jest jeden dynamiczny gest, w którym opuszki palców poruszają się wzdłuż wyznaczonej linii. Nawet jeśli niektóre gesty są prezentowane jako nieruchome, stają się dynamiczne poprzez przesuwanie ich wzdłuż ścieżki. Dla współrzędnych wykrytych punktów charakterystycznych obliczana jest odległość od linii powstałej z wyznaczenia regresji wszystkich punktów jednej klasy w danym eksperymencie.



Rysunek 6 Wyniki działania algorytmu śledzenia punktów charakterystycznych dla gestu statycznego i dynamicznego. Dla gestu dynamicznego przedstawiono wizualizację regresji liniowej tworzącej ścieżkę ruchu.

Dla wykrytego zbioru punktów charakterystycznych wyznaczono trajektorię liniową po której punkt powinien się przemieszczać, a następnie wyznaczono błąd metody d_i z użyciem wzoru

$$d_i = \frac{1}{N\sqrt{\bar{a}_i^2 + \bar{b}_i^2}} \sum_{j=1}^N |\bar{a}_i p_x^{[i,j]} + \bar{b}_i p_y^{[i,j]} + \bar{c}_i|. \quad (1)$$

Błąd liczony jest z uwzględnieniem liczby wykrytych punktów N , wartości linii trajektorii a, b i c , oraz współrzędnych wykrytego punktu p .

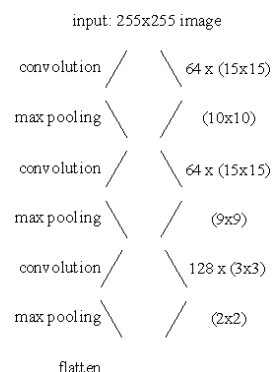
Przedstawione porównanie pokazuje, że opisanie podejście oparte na wizji i prostej kamerze światła widzialnego dorównuje rozwiązaniu używającego dedykowanych czujników głębi, takich jak kontroler Leap Motion w przypadku gestów dynamicznych. W przypadku powtarzalności obliczanej na podstawie gestu statycznego, błąd jest znacznie wyższy w porównaniu z Leap Motion. Prawdopodobnym wpływem jest to, że Leap Motion był testowany z manipulatorem, natomiast ludzka ręka często drga wykonując mikroruchy, niewystępujące w manipulatorach przemysłowych. Co więcej, w przedstawionych testach błąd statyczny wynosił średnio 1 piksel, co jest skalą pomiaru, dlatego nie można było uzyskać lepszych wyników w tym środowisku. Podsumowując metoda nie odbiegała znacząco jakością od metod używających dedykowanych kontrolerów, bazując na klasycznym obrazie z ogólnodostępnych kamer.

[3] Hand gesture classification with use of convolutional neural networks

– W czasie prowadzenia badań nad algorytmami wykrywania gestów z użyciem obrazów cyfrowych, coraz bardziej popularne stawały się metody bazujące na głębokim uczeniu z użyciem splotowych sieci neuronowych. Ten artykuł był odgałęzieniem od dziedziny klasycznego rozpoznawania obrazów w kierunku metod głębokiego uczenia.

OŚWIADCZENIE O WKŁADZIE:

Tomasz Grzejszczak, jako pierwszy autor, miał największy wkład w proces powstawania artykułu. Artykuł powstał bazując na jego koncepcji, opracował metodologię, utworzył oprogramowanie, przeprowadził analizę i walidację wyników oraz napisał treść wraz z wizualizacją wyników. Główny wkład pozostałych autorów to nadzór merytoryczny, recenzja, pomoc przy gromadzeniu danych testowych, konsultacja w sprawie proponowanej koncepcji, metodologii i oprogramowania.



Badania prowadzone w ramach [1] i [2] wykorzystywały klasyczne podejście przetwarzania obrazów cyfrowych, które polegało na tworzeniu zbioru reguł decyzyjnych. Nowe podejście oparte na sieciach neuronowych zakładało, że utworzona sieć neuronowa sama tworzy reguły decyzyjne dostrajając wagi warstw w procesie trenowania. Wadą tego rozwiązania jest konieczność posiadania zbioru treningowego. Każdy obraz w bazie danych jest wyposażony w opis i lokalizację punktów charakterystycznych oraz maskę obszaru skóry. Informacje te nie są jednak przydatne w podejściu głębokiego uczenia. W ramach badań [1-2] powstała baza obrazów zawierająca oznaczone punkty charakterystyczne oraz klasę wskazanego gestu. Zdjęcia zostały wykonane w różnych warunkach, tj. z kontrolowanym oświetleniem i tłem oraz bez kontroli tła i oświetlenia. Na potrzeby tworzenia zbioru, osoby wykonywały ten sam gest kilkakrotnie, co zapewnia różnorodność gestów. Każdy obraz jest opisany na podstawie identyfikatora osoby, znaczenia gestu i skojarzenia gestów (polski lub amerykański język migowy). W niniejszym artykule wykorzystano fragmenty tego zbioru danych. Każdy test opisuje rozmiar i typ użytego zbioru obrazów. Główną wadą tego zbioru jest to, że jest on stosunkowo mały, jak na standardy wielkości zbiorów uczących w podejściu głębokiego uczenia.

Konwolucyjna sieć neuronowa wykorzystuje model sekwencyjny. Warstwa wejściowa zawiera 255x255 neuronów (Obraz jest przeskalowany do takiej rozdzielczości), na które podawane są kolejne wartości pikseli obrazu. Następnie występuje warstwa konwolucyjna, aktywacji typu ReLU (rectified linear unit) i warstwy uśredniania maksymalnego uśredniania, która jest powtarzana 3 razy z różnymi parametrami. Następnie ostatnie warstwy są kombinacją spłaszczenia, zagęszczenia i porzucenia. Końcowa klasyfikacja jest wykonywana przy użyciu klasyfikatora softmax. Sieć neuronowa jest trenowana przy użyciu binarnej funkcji straty crossentropii.

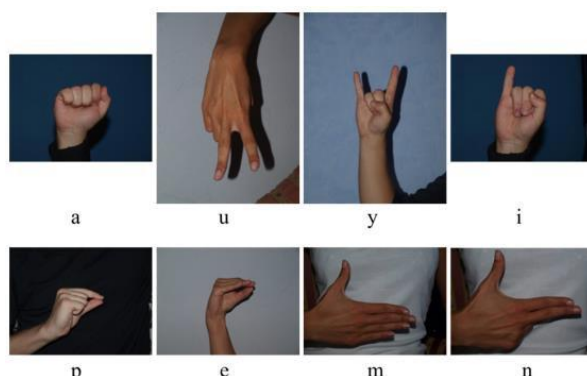
W ramach badań wykonano 3 testy, o różnym stopniu trudności, czyli z różną liczbą klas wyjściowych oraz zbiorem gestów łatwych do rozróżnienia oraz podobnych do siebie. Pierwszy test był najłatwiejszy i zawierał 6 typów gestów. W zależności od liczby uniesionych palców, prezentowane gesty odpowiadały gestom oznaczającym cyfry 0–6 z amerykańskiego języka migowego. Liczba uniesionych palców bezpośrednio wskazuje na typ gestu, dlatego test ten nazywany jest najłatwiejszym. Zadanie liczenia uniesionych palców było wielokrotnie badane z wykorzystaniem różnych podejść. Co więcej, aby było jeszcze łatwiejsze, przyjmuje się stałe tło. W tym teście 18 osób prezentuje 6 gestów. Każdy obraz jest powtarzany 10 razy z małą transformacją w celu zwiększenia zbioru treningowego, co daje łącznie 1080 próbek, po 180 dla każdej klasy. Obrazy są dzielone na zbiory treningowe, walidacyjne i testowe w proporcji 13:4:1. Wynik testu przedstawiono w Tabeli 1. Każdy wiersz przedstawia prawdopodobieństwo przypisania obrazu do jednej z 6 klas. Nazwy klas podano w pierwszym wierszu i pierwszej kolumnie. Intensywność koloru tła wskazuje na wartość pewności.

Tabela 1 Macierz pomyłek dla testu z 6 gestami [3]

	0	1	2	3	4	5
0	1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1	0.0018	0.9891	0.0058	0.0018	0.0014	0.0000
2	0.0020	0.0632	0.9170	0.0072	0.0106	0.0000
3	0.0000	0.0000	0.0001	0.9575	0.0000	0.0423
4	0.0039	0.0178	0.1123	0.0115	0.8393	0.0152
5	0.0000	0.0000	0.0000	0.0001	0.0004	0.9995

Z Tabeli 1 można odczytać, że każdy gest jest poprawnie sklasyfikowany. Najlepsza klasyfikacja, ze 100% dokładnością, dotyczyła gestu „0” bez pokazanych palców. Najgorsza klasyfikacja dotyczy gestu „4”, który można pomylić z gestem „2”. Powodem takiego wyniku jest to, że osoba wykonała ten gest nieprecyzyjnie, a palce się stykały. Ogólna pewność tego testu wynosi około 95%, co jest bardzo wysokim wynikiem. Dowodzi to, że ta metoda ma potencjał i może być stosowana wśród innych metod opartych na liczeniu palców uniesionych.

Kolejny test jest bardziej złożony. Tym razem wykorzystano zestaw gestów prezentujący liczby od 0 do 9. Problem polega na tym, że gesty 6-9 są podobne. Klasyfikacji nie można przeprowadzić poprzez liczenie uniesionych palców, ponieważ w każdym teście uniesione są 3 palce, a różnica polega na tym, czy kciuk dotyka palca wskazującego, serdecznego, czy małego. Gesty 6-9 są nie tylko podobne do siebie, ale także do gestów 1-5, co sprawiło, że w eksperymencie nie uzyskano satysfakcjonujących wyników świadczących o poprawnej klasyfikacji. Bardzo trudno jest określić położenie palców. Mimo braku powodzenia, test pokazuje najważniejsze problemy klasyfikacji podobnych gestów.



Rysunek 7 Fragment zbioru HGR prezentujący gesty statyczne z podpisem znaczenia prezentowanego gestu [3]

W trzeciej próbie, test przeprowadzany jest na pełnym zbiorze danych, czyli zbiorze wszystkich stacjonarnych gestów z polskiego języka migowego. 12 osób prezentuje 27 różnych gestów. Czasami jedna osoba prezentuje gest kilka razy, co daje łącznie 33 obrazy na jeden gest. Obrazy są podzielone na klasy treningowe, walidacyjne i testowe w proporcji 27:5:1.

Rys. 7 przedstawia część zbioru testowego. Można zauważyć, że w każdej klasie znajdują się losowe osoby. Tło jest raczej jednolite, ale nie ściśle kontrolowane. Patrząc na rys. 7, można zauważyć pewne podobieństwa niektórych gestów i unikalność innych. Obserwację tę potwierdzają wyniki testów przedstawione w tabeli 2. Gesty „a”, „u” i „y” są unikatowe. Poprzednie testy wykazały, że gest zaciśniętej pięści („0” lub „a”) jest najbardziej unikatowy. Tutaj również „u” i „y” są unikatowe, ponieważ jeden jest jedynym gestem odwróconym do góry nogami, a drugi wymaga szerokiego odstępu między palcem wskazującym a małym. „l” zostało sklasyfikowane jako marginalnie poprawne i zgodnie z macierzą pomyłek można je pomylić z „a” lub „y”. Gesty te są podobne i różnią się uniesieniem palca

wskazującego lub krocza. Inne podobne pary gestów to „p” – „e” i „m” – „n”. Gesty te różnią się jedynie drobnymi szczegółami, a te właściwości można zaobserwować w macierzy pomyłek.

Ogólna liczba pozytywnie sklasyfikowanych gestów wynosi 10 na 27 (37%). Jeśli przyjmiemy, że odpowiedź jest alternatywą dwóch najlepszych wyborów (na przykład potraktujemy odpowiedź „gest „e” to „p” lub „e” jako poprawną), to ogólny wynik wyniesie 16 na 27 (60%). Jest to wciąż niewielka liczba, jednak wystarczająco duża, aby udowodnić słuszność tego podejścia.

Warto również zauważyć, że gesty „p” i „e” różnią się pod względem tła i położenia łokcia, a ich podobieństwa dotyczą jedynie ułożenia palców. Idea splotowych sieci neuronowych gwarantuje pominięcie cech nieistotnych. W tym przypadku, gdy różne osoby są losowo przydzielane do różnych klas ze względu na różne gesty, takie indywidualne cechy, jak orientacja dłoni, kolor tła czy kolor koszuli, nie są brane pod uwagę.

Podsumowując badanie, prosta sieć typu CNN została wytrenowana na bazie, która zawierała bardzo dużo klas (w ostatnim teście było to 27 różnych gestów). Każdy gest prezentowany był przez zaledwie 18 zdjęć, z czego 13 użyto do trenowania, 4 do walidacji a 1 do testów. Wyniki tak relatywnie małego zbioru były zaskakująco wysokie.

Tabela 2 Macierz pomyłek dla zbioru 27 gestów. Prezentowane liczby to % pewności klasyfikacji. [3].

	1	2	3	4	5	a	b	c	d	e	f	g	h	i	k	l	m	n	o	p	r	s	t	u	w	y	z
1	28	09	04	00	00	00	00	01	01	00	00	00	00	00	03	32	04	02	01	00	10	01	01	00	00	00	03
2	01	01	01	04	01	02	00	03	03	07	00	00	00	01	00	03	34	27	00	08	01	00	00	00	00	01	00
3	28	30	30	01	04	00	00	00	00	00	00	00	00	00	00	04	02	01	00	00	00	00	00	00	00	00	00
4	01	03	03	48	05	00	00	02	01	03	00	00	00	00	00	01	18	09	01	03	00	00	00	00	00	00	00
5	02	01	01	08	03	04	03	08	07	05	06	08	03	03	01	02	02	02	08	05	01	06	05	00	01	02	02
a	00	00	00	00	00	98	00	00	00	01	00	00	00	01	00	00	00	00	00	00	00	00	00	00	00	00	00
b	00	00	00	00	00	00	24	00	00	00	00	00	01	05	37	03	00	00	00	00	24	00	00	00	01	00	04
c	00	00	00	01	00	15	00	09	04	29	03	02	01	03	00	00	02	04	01	20	01	00	01	01	00	02	00
d	00	00	00	00	00	00	00	45	07	13	13	06	00	00	00	00	00	00	04	08	00	00	03	00	00	00	00
e	00	00	00	00	00	02	00	12	08	36	01	07	00	00	00	00	00	00	00	33	00	00	00	00	00	00	00
f	00	00	00	00	00	00	00	02	01	00	18	07	00	00	00	00	00	00	27	00	00	20	25	00	00	00	00
g	00	00	00	00	00	00	00	07	02	00	10	44	01	01	00	00	00	00	05	00	00	06	22	00	00	00	01
h	00	00	00	00	00	00	01	11	03	01	04	06	16	07	01	01	00	00	01	00	01	02	21	00	03	17	04
i	00	00	00	00	00	27	03	02	02	04	02	05	04	28	01	00	00	00	00	01	01	00	01	00	01	17	01
k	00	00	00	00	00	02	02	18	09	04	17	17	05	04	01	00	00	00	01	04	01	01	09	00	00	02	02
l	00	00	00	00	00	01	04	01	01	00	01	00	01	06	15	21	00	00	05	00	10	07	05	02	02	00	17
m	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	84	16	00	00	00	00	00	00	00	00	00
n	00	00	00	01	00	00	00	00	00	00	00	00	00	00	00	00	84	15	00	00	00	00	00	00	00	00	00
o	00	00	00	00	00	00	03	05	03	01	15	06	04	02	03	01	00	00	12	01	02	13	22	00	02	00	05
p	00	00	00	00	00	06	00	04	02	74	00	00	00	01	00	00	00	00	00	12	00	00	00	00	00	01	00
r	00	00	00	00	00	00	01	00	00	00	00	00	06	01	24	15	00	00	00	00	13	03	07	00	03	00	26
s	00	00	00	00	00	00	01	01	01	00	04	04	10	03	03	00	00	00	01	00	01	04	41	00	13	02	11
t	00	00	00	00	00	00	02	01	01	00	07	03	11	06	03	00	00	00	00	00	03	01	30	00	07	02	22
u	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	99	00	00	00
w	00	00	00	00	00	00	10	00	00	00	00	00	12	05	28	01	00	00	00	00	08	00	02	00	20	04	08
y	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	00	99	00	00
z	00	00	00	00	00	04	03	19	11	05	09	02	02	05	01	01	00	00	08	04	03	02	11	04	00	01	03

Publikacja stanowiła wczesne i istotne rozpoznanie potencjału głębokiego uczenia, w szczególności konwolucyjnych sieci neuronowych (CNN), w dziedzinie rozpoznawania gestów. W momencie jej powstawania w 2019 roku, dominujące podejścia opierały się na klasycznych metodach przetwarzania obrazów, a zastosowania CNN w tym obszarze nie były jeszcze powszechne. Praca pełniła zatem rolę pionierskiego dowodu koncepcji, wykazując, że nawet prosta architektura sieci, zasilona stosunkowo niewielkim zbiorem danych, jest w stanie osiągnąć satysfakcjonującą i obiecującą skuteczność.

Choć bezpośredni wpływ artykułu miał charakter raczej przyczynowy niż przełomowy, to jego trwały wkład w rozwój dyscypliny znajduje potwierdzenie w notowanej liczbie cytowań. Publikacja wpisała się w rodzący się nurt badawczy, którego pełny rozkwit obserwujemy obecnie, gdy zaawansowane rozwiązania używające głębokiego uczenia, takie jak MediaPipe, stały się standardem w efektywnym rozwiązywaniu złożonych problemów wizyjnego rozpoznawania gestów i śledzenia punktów charakterystycznych. W ten sposób praca przyczyniła się do stopniowej zmiany paradygmatu, torując drogę dla nowoczesnych, opartych na sieciach neuronowych rozwiązań w tej dynamicznie rozwijającej się dziedzinie.

4.2. Przetwarzanie danych z czujników inercyjnych dla HRI

[4] Underwater human arm manipulator - system calibration

[5] Underwater manipulator that imitates the movements of the human arm

– Efektem prowadzonych badań z zakresu HRI było zaproszenie do zespołu badawczego we współpracy z firmą SR-Robotics w ramach projektu finansowanego z NCBiR. Efektem zrealizowanych prac były dwie publikacje w czasopiśmie. Celem prac było kreowanie interakcji człowiek – robot.

OŚWIADCZENIE O WKŁADZIE:

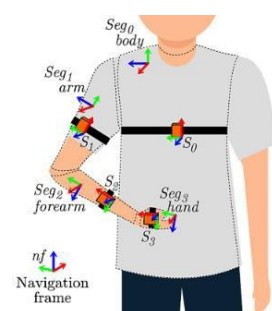
Publikacje są ze sobą powiązane, ponieważ stanowią efekt projektu naukowego. Wkład Tomasza Grzejszczaka w procesie powstawania tych publikacji jest następujący:

- Uczestniczył w procesie opracowywania metodologii, głównie skupiając się na aspektach interakcji Człowieka z Robotem, procesem pobierania i analizy danych z czujników inercyjnych oraz sterowania manipulatorem pojazdu podwodnego. Przygotowywał oprogramowanie do wyżej wymienionych aspektów
- Uczestniczył w procesie powstawania treści artykułu „Underwater human arm manipulator - system calibration” oraz uczestniczył w procesie recenzji obu artykułów.

- Przygotowywał dane testowe oraz prowadził eksperymenty walidacyjne

Znaczący wkład pozostałych autorów w procesie publikacji:

- Opracowanie koncepcji, metodologii, testów i oprogramowania związanych z: równaniami kinematyki, procesem kalibracji, urządzeniami haptycznymi, aspektami konstrukcyjnymi.



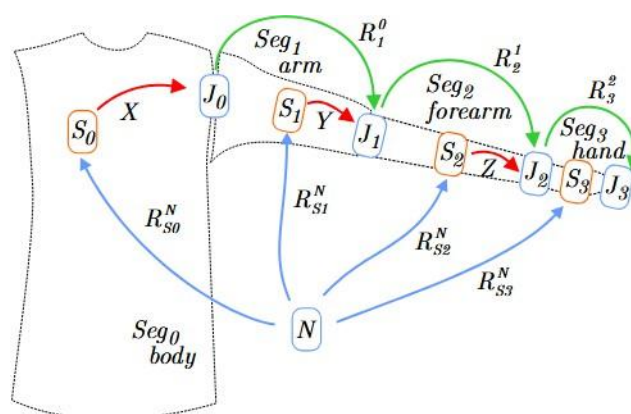
W dziedzinie robotyki metody sterowania ramieniem robota pozostają niezmiennie od dziesięcioleci. Głównym zastosowaniem ramion robota są powtarzalne, ciągłe operacje, dlatego główne metody programowania dzielą się na planowanie on-line i off-line. Rzadszym, ale nie mniej ważnym zastosowaniem manipulatorów jest teleoperacja. W tym przypadku ramię robota jest sterowane w czasie rzeczywistym przez teleoperatorów w celu wykonywania trudnych lub precyzyjnych zadań. Teleoperacja jest wykorzystywana w trudnych środowiskach np. w medycynie, rozbrajaniu ładunków wybuchowych, eksploracji kosmosu lub w zdalnym sterowaniu pojazdami mobilnymi w miejscach trudno dostępnych.

Rozwój urządzeń ułatwiających sterowanie manipulatorami znacznie odciążył operatorów i zmniejszył występowanie błędów. Literatura potwierdza, że wdrożenie urządzeń wspomagających, takich jak systemy dotykowego sprzężenia zwrotnego, interfejsy rzeczywistości rozszerzonej i zaawansowane algorytmy automatyzacji, znacznie usprawniło sterowanie robotami podwodnymi. Tworząc zdalnie sterowany manipulator, twórcy stają przed problemem stworzenia interfejsu człowiek-robot (HRI), czyli zestawu elementów sterujących i wyświetlaczy stanu, który ułatwia proces manipulacji. Większość gotowych rozwiązań, zaczerpniętych z programowania online, jest dobra, ale brakuje im intuicyjności i wymagają przeszkolenia operatorów. W przypadku zdalnie sterowanych robotów podwodnych (ROV) zazwyczaj jest co najmniej dwóch operatorów, z których jeden odpowiada

za stabilność platformy, a drugi za obsługę manipulatora. Operator manipulatora korzysta z kontrolera, który jest modelem manipulatora i imituje jego kształt. Operator może zmieniać położenie elementów modelu. Ruchy te są przekładane na ruch rzeczywistej maszyny. W tym przypadku system nazywa się sterowaniem master-slave, w którym manipulator jest urządzeniem podrzędnym, podążającym za ruchem urządzenia nadrzędnego (kontrolera). To rozwiązanie nadal jest mało intuicyjne i wymaga przeszkolenia. Z drugiej strony, przemysł filmowy i animacyjny powszechnie wykorzystuje technologię przechwytywania ruchu (motion capture), aby dokładnie rejestrować i śledzić ruchy ciała.

Celem tych badań było wykorzystanie inercyjnych jednostek pomiarowych IMU do śledzenia ruchu rąk i wykorzystania danych sekwencyjnych do sterowania manipulatorem w ROV. W takim podejściu operator może manipulować ramieniem robota w intuicyjny sposób poruszając własną ręką. W tamtych czasach podobne rozwiązania zazwyczaj ograniczały się do bezpośredniego mapowania przetwarzając jeden kanał IMU na jeden przegub robota lub sterowania za pomocą predefiniowanych poleceń. Wkładem tej pracy jest procedura kalibracji, która służy do mapowania macierzy kątów obrotu z 4 jednostek IMU na podstawie położenia ręki operatora jako kątów zgięcia nadgarstka, łokcia i ramienia. Celem jest zaproponowanie systemu sterowania manipulatorem okrętu podwodnego, który imituje ruchy ramienia człowieka na podstawie sygnałów z czujników IMU. Podejście to nie wymaga stosowania zaawansowanych algorytmów planowania trajektorii, takich jak RRT, PRM, RRT* lub podobnych. Dzięki zastosowaniu prezentowanego podejścia możliwe jest bardzo łatwe sterowanie ruchem robota bez żadnych ograniczeń poza tymi wynikającymi z naturalnych ograniczeń ruchu ramienia człowieka. Jedynym wymogiem dla proponowanego rozwiązania jest kalibracja systemu czujników.

Realizacja przedstawionych prac badawczych pozwoliła na opracowanie innowacyjnego systemu sterowania manipulatorem podwodnym, który w sposób intuicyjny odwzorowuje ruchy kończyny górnej operatora. Kluczowym elementem dorobku jest opracowanie metodologii interakcji Człowieka z Robotem (HRI), bazującej na wykorzystaniu bezprzewodowych czujników inercyjnych (IMU) systemu Xsens. Indywidualny wkład w ten obszar obejmował zaprojektowanie procesu pobierania i wielokanałowej analizy danych z sensorów rozmieszczonych na ciele operatora (S_0 – S_3), ze szczególnym uwzględnieniem eliminacji wpływu ruchów całego ciała poprzez zastosowanie czujnika referencyjnego S_0 . Takie podejście umożliwiło uzyskanie stabilnych odczytów orientacji członów operatora nawet w trudnych warunkach pracy na kołyszącej się jednostce pływającej.



Rysunek 8 Schemat połączenia czujników pomiarowych IMU z elementami łańcucha kinematycznego ramienia operatora. [4]

W warstwie technologicznej osiągnięcie obejmuje przygotowanie autorskiego oprogramowania w języku C#, odpowiedzialnego za akwizycję danych pomiarowych w czasie rzeczywistym oraz ich przetwarzanie na parametry sterujące. Opracowane algorytmy integrowały odczyty z sensorów i przesyłały wypracowane nastawy dla przegubów manipulatora do sterownika PLC (Wago) za

pośrednictwem protokołu Modbus. Należy jednak precyzyjnie wskazać, że teoretyczne wyprowadzenie matematycznych równań kinematyki prostej i odwrotnej manipulatora oraz szczegółowe rozwiązanie problemu Wahby w procesie kalibracji nie stanowiły przedmiotu mojego wkładu, lecz były wynikiem prac pozostałych współautorów. Moja rola koncentrowała się na implementacji oprogramowania sterującego wyżej wymienionymi aspektami oraz ich merytorycznemu opisowi.

W celu pomiaru ruchów ręki operatora niezależnie od ruchów jego ciała, zastosowano system czterech bezwładnościowych czujników IMU (S0–S3), zamontowanych na ciele (Seg0), ramieniu (Seg1), przedramieniu (Seg2) i dłoni (Seg3). Czujnik referencyjny S0 stanowi ruchomą bazę układu odniesienia, eliminując wpływ przemieszczeń operatora (np. na statku) na pomiar względnego ruchu łańcucha kinematycznego ręki. Orientację czujników opisano przy użyciu macierzy rotacji, wykorzystując sekwencję kątów Eulera XYZ

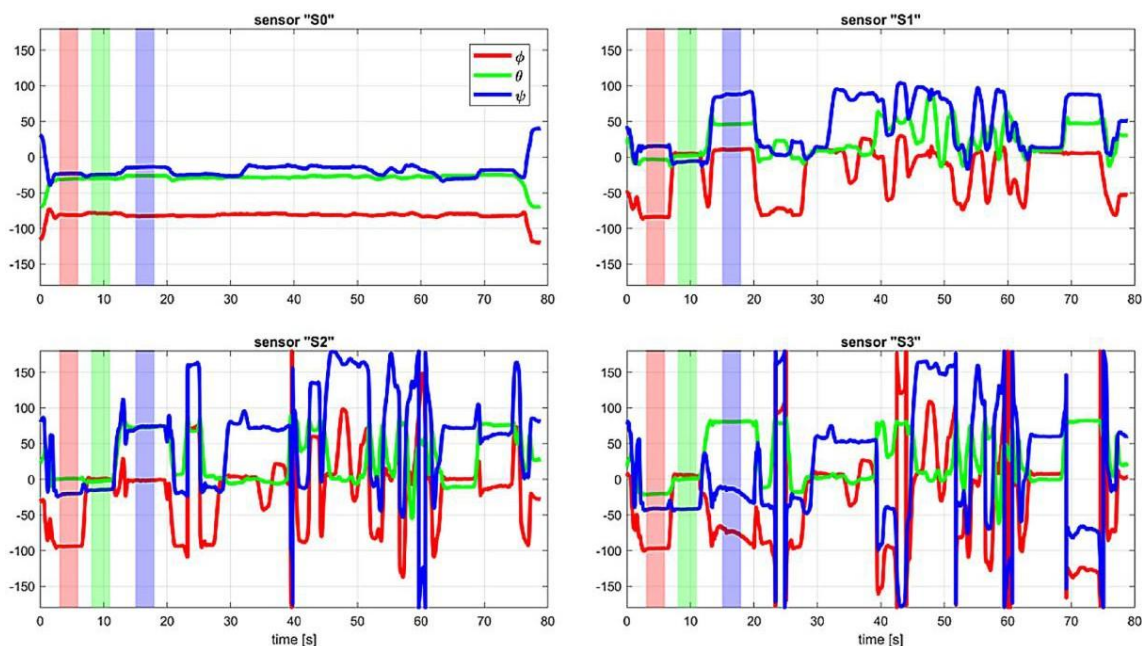
$$R = Rot_Z(\psi)Rot_Y(\theta)Rot_X(\varphi) = \begin{bmatrix} c_\psi & -s_\psi & 0 \\ s_\psi & c_\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} c_\theta & 0 & s_\theta \\ 0 & 1 & 0 \\ -s_\theta & 0 & c_\theta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & c_\phi & -s_\phi \\ 0 & s_\phi & c_\phi \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \quad (2)$$

o określonych zakresach zmienności ($\phi \sim \pm 180^\circ$, $\theta \sim \pm 90^\circ$, $\psi \sim \pm 180^\circ$). Kalibracja systemu polega na wyznaczeniu nieznanych macierzy obrotu (X, Y, Z, V) spełniających układ równań

$$\begin{cases} R_{S_0}^N X R_1^0 = R_{S_1}^N Y \\ R_{S_0}^N X R_2^0 = R_{S_2}^N Z \\ R_{S_0}^N X R_3^0 = R_{S_3}^N V \end{cases} \quad (3)$$

gdzie: $R_1^0 = R_0^0 R_1^1$, $R_2^0 = R_0^0 R_1^1 R_2^2$. Proces kalibracji wykorzystuje trzy statyczne pozycje operatora: „I” (ramię wzdłuż tułowia), „T” (ramię w bok) oraz „F” (ramię w przód). Pozwala to uprościć układ równań do postaci: $R_1^0 = R_0^0 = R_2^0 = R_{pose}$, gdzie R_{pose} to macierz rotacji modelu dla danej pozycji (RI, RT, RF).

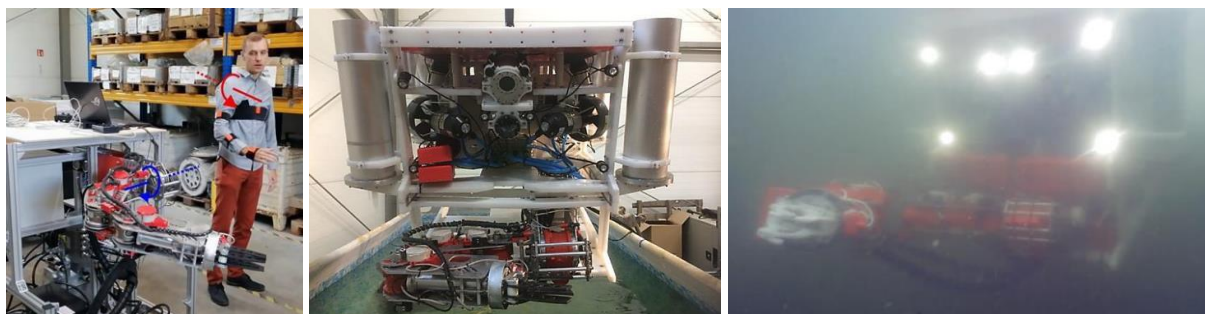
Kalibrację należy przeprowadzać przy każdym założeniu zestawu czujników.



Rysunek 9 Przebiegi kątów obrotu poszczególnych czujników pomiarowych w eksperymencie kalibracyjnym „IMU-arm”[4]

Wynikiem kalibracji jest możliwość wyznaczenia kątów obrotu (ϕ , θ , ψ) dla par segmentów: ciało–ramię, ramię–przedramię, przedramię–dłoń (wykres na rys. 9). Mimo różnic anatomicznych, procedura umożliwiła określenie współczynników przeliczeniowych ruchu ludzkiej ręki na kąty stawów manipulatora. Dzięki referencji do ciała (S0), operator może swobodnie się przemieszczać, utrzymując chwytak w stałej pozycji, co realizuje intuicyjne, aczkolwiek nieco opóźnione, sterowanie manipulatorem.

Proces weryfikacji naukowej obejmował przygotowanie danych testowych oraz prowadzenie wieloetapowych eksperymentów walidacyjnych, w tym testów w warunkach basenowych oraz badań w środowisku naturalnym na Morzu Bałtyckim. Eksperymenty te pozwoliły na ocenę poprawności odwzorowania trajektorii ruchu oraz analizę opóźnień sterowania wynikających z dynamiki układu hydraulicznego. Warto zaznaczyć, że aspekty konstrukcyjne 5-osiowego manipulatora podwodnego oraz projekt i budowa urządzeń haptycznych (rękawica z driverami DRV2605L) pozostawały poza zakresem mojego indywidualnego wkładu. Sumarycznie, opisane osiągnięcia stanowią spójne rozwiązanie zwiększające intuicyjność teleoperacji w robotyce podwodnej.



Rysunek 10 Wnioskodawca w trakcie badania ruchu w górę i w dół w laboratorium sterując ramieniem pojazdu podwodnego. Widok pojazdu ROV z przymocowanym manipulatorem w laboratorium i pod wodą w trakcie pracy [5]

Te dwie publikacje są przykładem praktycznego zastosowania metod przetwarzania sygnałów zmiennych w czasie w zastosowaniu do HRI. Niewątpliwy wkład w rozwój dziedziny nauki został pozytywnie oceniony przez ekspertów NCBiR, zarówno w procesie naboru jak i rozliczania wniosku. Pojazd wraz z działającą innowacyjną metodą teleportacji wszedł w ofertę wnioskodawcy projektu.

4.3. Powierzchniowe czujniki elektromiograficzne - sEMG

[7] Surface electromyography data analysis for evaluation of physical exercise habits between athletes and non-athletes during indoor rowing –

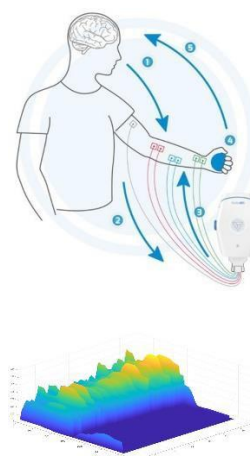
Praktycznym zastosowaniem algorytmów interakcji człowieka z robotem jest robotyka rehabilitacyjna. Podobnie jak w przypadku sygnałów czujników inercyjnych, akwizycja danych z czujników elektromiograficznych również prowadzi do powstania szeregów czasowych wymagających obróbki.

OŚWIADCZENIE O WKŁADZIE:

Artykuł zawiera sekcję **Author Contributions**, która precyzyjnie rozdziela wkład autorów. Tomasz Grzejszczak był zaangażowany w: opracowanie Metodologii, Oprogramowanie, Walidację oraz proces pisanie treści artykułu. Główny wkład to przetwarzanie danych pomiarowych, dalsza ich obróbka, automatyczne wykrywanie wzorców, klasteryzacja, opracowanie i wizualizacja danych statystycznych.

Znaczący wkład pozostałych autorów publikacji:

Przygotowanie środowiska pomiarowego, filtracja sygnału EMG, przygotowanie testów i zbieranie pomiarów, nadzór merytoryczny i finansowy.

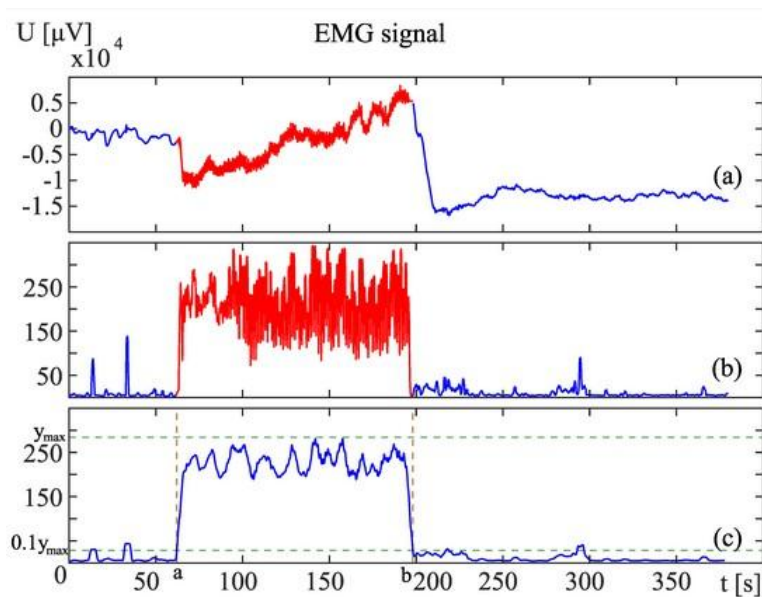


W artykule podjęto problem klasyfikacji i oceny jakości treningu bazując na czujnikach elektromiograficznych zbierających dane w trakcie pracy z ergometrem wioślarskim. W ostatnich latach można zauważyć zwiększoną chęć ludzi do dbania o swój rozwój fizyczny i uprawiania sportów amatorskich. Ponadto, amatorzy często korzystają z różnego rodzaju sprzętu i gadżetów, aby ułatwić sobie ćwiczenia. Urządzenia noszone na ciele mierzące pewne wielkości fizyczne lub fizjologiczne mogą zapewnić lepsze porady trenerskie. Zastosowanie nauki i technologii może zapewnić znaczącą przewagę konkurencyjną. Urządzenia technologiczne noszone na ciele mają ogromny potencjał w niedalekiej przyszłości, umożliwiając spłaszczenie krzywej kosztów opieki zdrowotnej.

Wioślarstwo stacjonarne aktywuje wiele mięśni i części ciała, zwłaszcza ramiona (biceps), nogi (mięsień czworogłowy uda) i mięśnie brzucha. Zazwyczaj wioślarz stacjonarny (ergometr wioślarski) jest wyposażony w mechanizm oporu powietrza i czujnik, który przelicza przyłożoną siłę ciągu. W tych badaniach mierzony i przetwarzany jest czas aktywacji mięśni. Czas aktywacji mięśni jest często wykorzystywany do określania poprawności wykonywanych zadań lub do identyfikacji ograniczeń. Badanie aktywacji mięśni może służyć do określania zmęczenia mięśni, przeciążenia lub być podłożem do diagnozy schorzeń. W tych badaniach pomiarów dokonywano za pomocą elektromiografii powierzchniowej (sEMG)

Z użyciem sEMG można określić czas i siłę skurczu mięśni w celach diagnostycznych i rehabilitacyjnych. Oprócz celów medycznych i rehabilitacyjnych, sEMG jest wykorzystywane do badania nowych podejść do interakcji człowiek-robot (HRI). W literaturze można odnaleźć metody wykrywania skurczów mięśni za pomocą sEMG, wykorzystane do rozpoznawania gestów dłoni lub do bezpośredniego sterowania manipulatorem.

Wkładem niniejszego artykułu jest metoda przetwarzania sygnału sEMG uzyskanego z bicepsów i mięśni czworogłowych uda podczas ćwiczeń wioślarskich na ergometrze wioślarskim. Uzyskano zestaw parametrów w celu rozróżnienia wyników poszczególnych osób i znalezienia typowych błędów w ćwiczeniach. Parametry osób niebędących sportowcami wyraźnie wskazują na typowe błędy w porównaniu z profesjonalnymi sportowcami. Ponadto, przedstawiona metoda pozwala odróżnić osoby z potencjałem sportowym od osób niebędących sportowcami.



Rysunek 11 Przykładowy sygnał sEMG zarejestrowany podczas ćwiczenia. Ćwiczenie trwało od 62 do 198 sekund, a ten obszar zaznaczono na czerwono. Wykresy przedstawiają (a) sygnał surowy, (b) sygnał przekształcony metodą RMS i (c) przefiltrowany sygnał RMS. [7]

W niniejszym badaniu jako metodę pomiaru zastosowano wyłącznie sEMG z trzema kanałami: bicepsa, czworogłowego i referencyjnego. Po uzyskaniu sygnału RMS, kolejnym krokiem jest określenie czasu trwania ćwiczenia. Ćwiczenie rozpoczyna się od pierwszego skurczu mięśnia i kończy po ostatnim rozluźnieniu. Charakteryzuje się ono harmonicznymi i szybkimi zmianami sygnału RMS na rysunku 11b. Można łatwo zaobserwować, że ćwiczenie na rysunku 11b trwa od 62 s do 198 s.

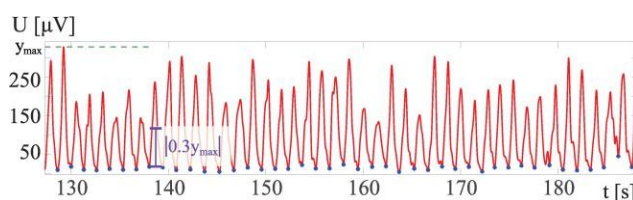
Rejestracja rozpoczyna się po podłączeniu wszystkich elektrod, ale przed rozpoczęciem ćwiczeń. Również po ćwiczeniach sygnał jest rejestrowany przez pewien czas. W tym czasie przed i po ćwiczeniach mogą wystąpić pojedyncze skurcze mięśni. Te pojedyncze impulsy można zaobserwować w przykładzie na rysunku 11b w 16, 36 i 296 sekundzie. Co więcej, podczas ćwiczeń rozluźnienie mięśni może sprowadzić sygnał do wartości bliskich zeru. Te właściwości uniemożliwiają zastosowanie prostego progowania.

Aby automatycznie określić czas trwania ćwiczenia, sygnał RMS na rysunku 11b jest filtrowany w celu uzyskania sygnału na rysunku 11c. Filtracja ma na celu zmniejszenie liczby pojedynczych skurczów przed i po ćwiczeniu oraz zapobieżenie zbyt niemu spadkowi sygnału w trakcie eksperymentu. Sygnał (rysunek 11c)

$$y(n) = \frac{1}{N_w} \sum_{i=0}^{N-1} x(n-i) \quad (4)$$

jest uzyskiwany poprzez filtrację uśredniającą ruchomą z użyciem próbek N_w w oknie filtru z sygnału RMS. Aby wygładzić sygnał, rozmiar okna N_w jest ustawiony na co najmniej równy $2t_c \cdot fs$, gdzie t_c to średni czas cyklu, natomiast fs to częstotliwość próbkowania.

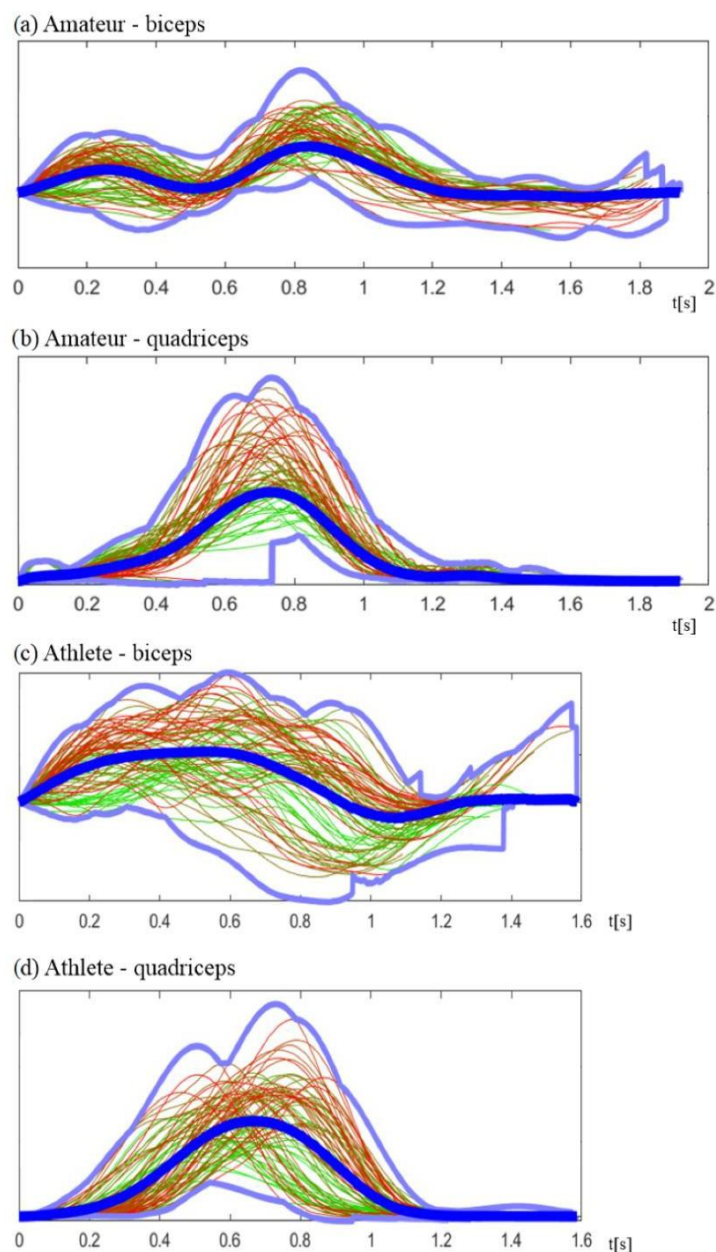
Ćwiczenie polega na powtarzających się skurczach i rozluźnieniach mięśni, które tworzą cykl. Sygnał jest dzielony na segmenty zgodnie z minimami lokalnymi odnoszącymi się do momentu największego rozluźnienia mięśni. Prawidłowo podzielone sygnały powinny mieć czas cyklu $t_c \approx t_{\hat{e}}$. Prawidłowo podzielone sygnały powinny charakteryzować się czasem cyklu z małym odchyleniem standardowym $\sigma(t_c)$. Wartości $t_c \approx 2t_{\hat{e}}, 3t_{\hat{e}}, 0.5t_{\hat{e}}$, zazwyczaj wynikają z nieprawidłowego położenia minimów i podziału sygnału. Aby znaleźć prawidłowe minima lokalne, brana jest pod uwagę minimalna wartość wzniesienia otaczających maksimów. Wartość głębokości minimum określa, jak bardzo wartość wyróżnia się ze względu na swoją wewnętrzną głębokość i położenie względem innych minimów. Podczas eksperymentów wartość 0,3 zapewnia najmniejszą akceptowalną długość segmentów sygnału. Każde minimum jest oznaczane, gdy otaczające je maksima przekraczają wartość $0,3 y_{max}$. Fragment sygnału z rysunku 10c z widocznymi cyklami, wykrytymi minimami lokalnymi oraz wizualizacją obliczonego minimalnego wzniesienia przedstawiono na rysunku 12.



Rysunek 12 Fragment przykładowego sygnału sEMG RMS z rysunku 10b z oznaczeniami minimów lokalnych, które dzielą sygnał na cykle. Minimalna wzniesienie jest zilustrowane fioletowym odcinkiem linii o długości $0,3 y_{max}$. [7]

Odpowiednie podzielenie sygnału na cykle pozwala na precyzyjną analizę oraz wyznaczenie dodatkowych parametrów ćwiczenia, takich jak liczba cykli, średni czas aktywacji, lokalne minima i maksima cyklu oraz długość aktywacji mięśnia. Każdy z tych parametrów może być użyty do klasyfikacji i oceny jakości wykonywanego ćwiczenia.

W prawidłowo wykonanym ćwiczeniu każda faza powinna być wykonywana po kolei. Oznacza to, że bicepsy powinny być aktywowane po mięśniach czworogłowych. Średni wykres dla profesjonalnych sportowców (Rys. 13c, d) ilustruje prawidłowe wykonanie ćwiczenia. Po pozycji wyjściowej z fazy 1, w czasie $t=0,2 - 0,8$ sekundy, aktywowane są mięśnie czworogłowe (faza 2). Podczas prostowania nogi, w czasie $t=0,2 - 0,9$, aktywowane są bicepsy (faza 3). Podczas regeneracji przy $t>1,0$ (faza 4), mięśnie nie są aktywne. W przykładzie na Rys. 13 a-b, amator aktywował bicepsy zbyt wcześnie, rozpoczynając fazę 3 przed zakończeniem fazy 2. Maksimum lokalne wykresu na Rys. 13a znajduje się przed maksimum na Rys. 13b. Co więcej, ten błąd spowodował potrzebę drugiej aktywacji bicepsów przy $t=0,8$ i drugim maksimum lokalnym, dalej skutkując krótszą i mniej efektywną fazą końcową. Jedną ze zmiennych pozwalających na ocenę poprawności wykonania ćwiczenia jest liczba maksimumów lokalnych i odpowiadających im pozycji.

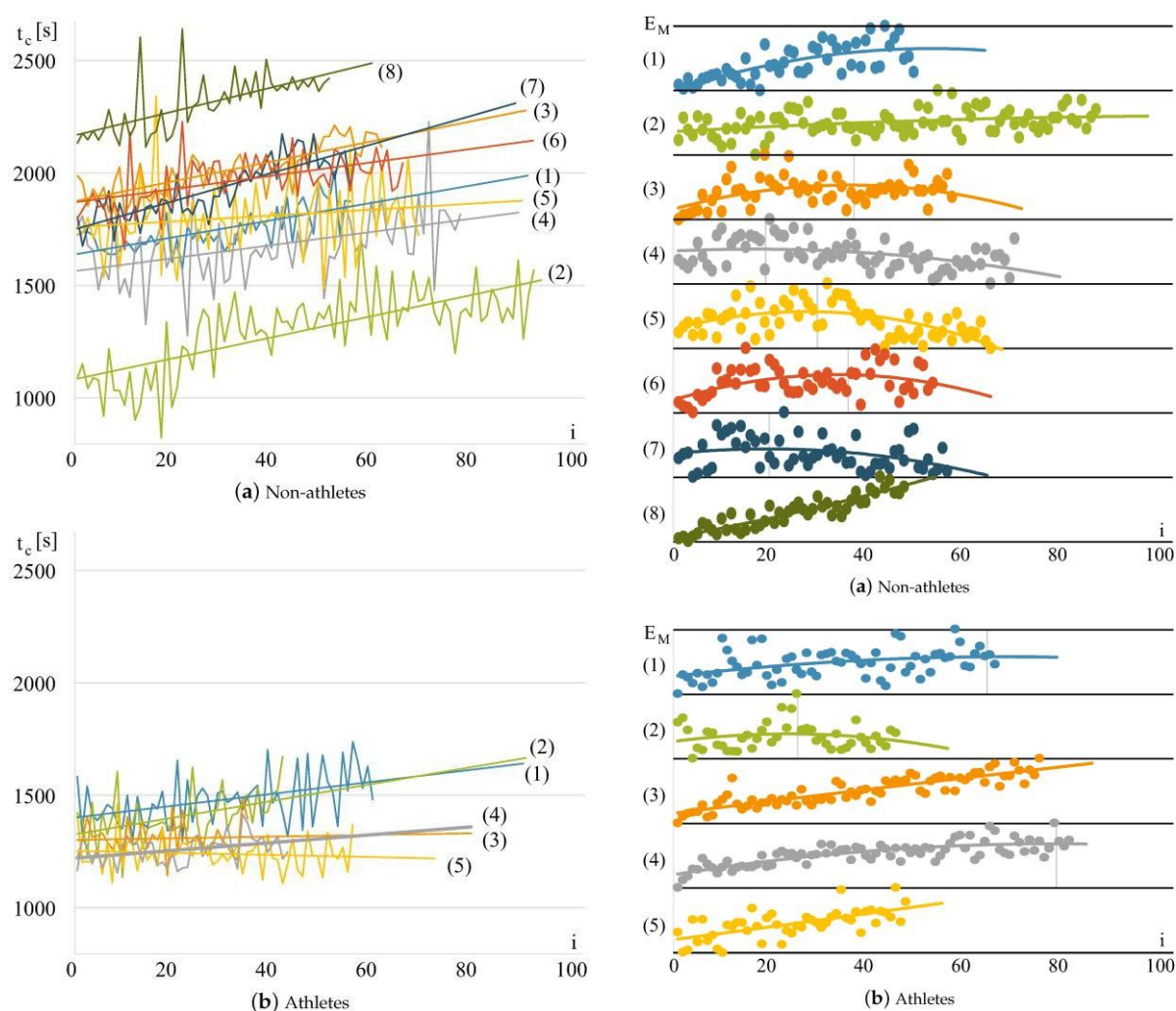


Rysunek 13 Wykresy połączonych cykli treningowych N_c , których kolejność można rozróżnić po gradiencie kolorów. Pierwszy cykl ćwiczenia jest przedstawiony na zielono, a ostatni na czerwono. Wykresy przedstawiają klasę: (a, b) Amatorzy i (c, d) sportowcy zawodowi. (a, c) Mięśnie czworogłowe i (b, d) Bicepsy. Niebieskie linie oznaczają minimalną, maksymalną i średnią aktywację mięśni. Wartości są znormalizowane (brak wartości na pionowej osi), ponieważ nie są porównywalne między poszczególnymi osobami. [7]

W artykule przedstawiono metody obliczania zaawansowanych wskaźników, takich jak:

- Liczba cykli ćwiczenia N_c
- Czas trwania ćwiczenia t_c oraz średni czas trwania cyklu $t_c^{[i]}$
- Odchylenie standardowe czasu cyklu oraz wartość trendu czasu cyklu a świadcząca o wytrzymałości
- Siła mięśni rozumiana jako całkowita wykryta siła neuronalna działająca na mięśnie w każdym cyklu
- Punkt zmęczenia, jako punkt maksimum regresji wielomianowej zmiany siły mięśni E_M

Wybrane wskaźniki przedstawiono na rysunku 14, natomiast dokładne metody ich wyliczania można odnaleźć w artykule źródłowym [7].



Rysunek 14 Po lewej: Zmiana długości cyklu w trakcie ćwiczenia wraz z wyliczonym trendem. Po prawej: Zmiana siły mięśni wraz z regresją wielomianową. Podział na grupy: a- amatorzy, b- sportowcy. [7]

Warto zauważyć, że w miarę trwania ćwiczenia, jak osoba staje się coraz bardziej zmęczona, trzeba aktywować więcej włókien mięśniowych. Wyliczona za pomocą całki moc mięśni

$$E_M[i] = \int_0^{t_c} y_c^{[i]}(n) \, dn, \quad (5)$$

wzrasta w początkowej fazie ćwiczenia, jednak w pewnym momencie wartość ta osiąga maksimum a ćwiczący nie jest już w stanie aktywować więcej włókien mięśniowych i energia mięśniowa zaczyna w kolejnych cyklach maleć. Ten punkt maksymalnego wydatkowania energii mięśniowej nazywany jest punktem zmęczenia. W niektórych przypadkach nie wyznaczono maksimum ponieważ całe ćwiczenie zostało wykonane bez nadmiernego zmęczenia, co świadczy o tym, że ćwiczący był w stanie zrobić więcej – oznacza to, że nie przyłożył się należycie do treningu.

Tabela 3 Dane z testu ćwiczeniowego. Wartości dla każdego koloru zmiennej zgodnie z formatowaniem warunkowym. [7]

	t_t	N_c	σ	$t^{[i]}$	Trend a	Zmęczenie
Amatorzy:	92,45	53	0,092	1,74	4,025	0%
	114,45	88	0,179	1,30	4,955	0%
	119,61	59	0,102	2,03	4,713	41%
	124,08	74	0,151	1,68	3,063	61%
	156,62	66	0,157	1,80	1,368	45%
	104,19	54	0,126	1,93	6,674	15%
	120,93	63	0,110	1,92	3,162	46%
	124,36	49	0,124	2,31	5,731	0%
Sportowcy:	94,37	64	0,087	1,47	2,708	2%
	64,51	46	0,066	1,40	3,779	43%
	98,39	74	0,075	1,34	0,312	3%
	98,39	80	0,072	1,23	0,454	0%
	58,85	47	0,062	1,25	1,731	0%

Dla ośmiu osób niebędących sportowcami i pięciu sportowców, ćwiczenie zostało ocenione, biorąc pod uwagę wiedzę na temat sześciu kluczowych czynników, które przedstawiono w Tabeli 3. Wartości są warunkowo pokolorowane: najgorsze wartości na czerwono, a najlepsze na zielono. Ćwiczenie powinno zakończyć się jak najszybciej. Liczba segmentów cyklu N_c to liczba wykonanych faz ćwiczenia. Przy mniejszej liczbie cykli, energia mięśni musi być wyższa, aby pokonać zadany dystans. Zakończenie ćwiczenia z mniejszą liczbą cykli może wskazywać na większą moc mięśni, ale nie jest to kluczowy czynnik. Odchylenie standardowe σ ($t[i]c$) odpowiada podobieństwu długości segmentu w sekundach. Wyższa wartość σ wskazuje, że dana osoba miała trudności z utrzymaniem pożądanego rytmu cykli. Średni czas cyklu $t[i]$ to średni czas całej fazy ćwiczeń. Obserwacje dotyczące średniego czasu cyklu są podobne do tych dotyczących liczby segmentów cyklu. Czas cyklu zmienia się w trakcie ćwiczeń. W procesie oceny nachylenie tych zmian oblicza się jako $\hat{a}$, które jest nachyleniem linii regresji zmieniającego się czasu cyklu. Zdolność do utrzymania stałego rytmu cyklu jest wskazywana małą wartością parametru $\hat{a}$. Podczas ćwiczeń występuje punkt, w którym osoba jest zmęczona i wykonuje ostatnie cykle w stanie wyczerpania. W Tabeli 3 zmęczenie wskazuje, jaka część ćwiczenia w procentach została wykonana w stanie zmęczenia.

Ponadto dane z Tabeli 3 wykorzystano do obliczenia istotności statystycznej (Tabela 4). Dla każdej zmiennej w Tabeli 3 i dla każdej grupy obliczono wartość średnią i odchylenie standardowe (SD), aby uzyskać wartość t i wartość p zgodnie z testem t dla niezależnych prób. W rezultacie można zauważyć, że według konwencjonalnych kryteriów prawie wszystkie sugerowane zmienne zapewniają różnicę, którą uważa się za bardzo istotną statystycznie. Jedynym wyjątkiem jest N_c , która nie jest istotna statystycznie. W artykule zaprezentowano wyjaśnienia, dlaczego N_c samo nie może być klasyfikatorem.

Tabela 4 Dane testów statystycznych [7]

Zmienna	Sportowcy Średnia	SD	Amatorzy Średnia	SD	Wartość t	Wartość p
t_t	119.5863	18.5784	82.9020	19.5450	3.3983	0.0059
N_c	63.25	12.80	62.20	15.43	0.1333	0.8964
σ	0.13013	0.02988	0.07240	0.00961	4.1275	0.0017
$t^{[i]}$	1.8388	0.2928	1.3380	0.1008	3.6388	0.0039
Trend a	4.21137	1.67834	1.79680	1.48083	2.6318	0.0233

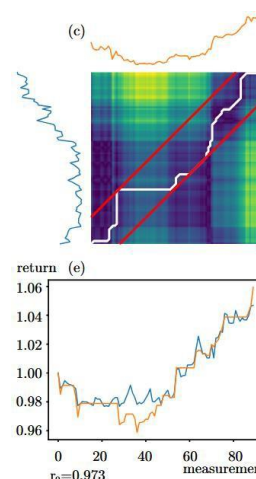
Dogłębna analiza, dokładniej opisana w artykule [7] oraz zestawienie wyników badań (tabela 3) pozwoliły na wyciągnięcie szeregu wniosków. Zaproponowano sposób pomiaru indywidualnej aktywności sportowej. Podczas eksperymentu uzyskano szereg oczywistych wniosków: profesjonalści wykonują ćwiczenie lepiej niż amatorzy, a zmęczenie pojawia się w czasie podczas ćwiczeń, powodując mniejszą siłę mięśni. Z drugiej strony, najważniejsze jest to, że niniejsze badanie zapewnia sposób pomiaru tych podstawowych zjawisk. Dane przedstawione i wyjaśnione w artykule umożliwiają pomiar wyników ćwiczenia jednostki na podstawie czasów aktywacji mięśni i lokalnych pozycji ekstremów. Przedstawiono sposób wskazywania zmęczenia, które można zauważyć na podstawie zmian amplitudy i czasu cyklu, co stanowi interesującą alternatywę dla mediany częstotliwości i jej nachylenia podczas wysiłku, często wykorzystywanej podczas oceny zmęczenia za pomocą sEMG. Wreszcie, faza wyczerpania, która może sugerować moment zakończenia wysiłku, może być zauważona wraz ze spadkiem amplitudy cyklu. Analiza poszczególnych wyników przedstawionych w Tabeli 3 może przynieść cenne uwagi na temat oceny wykonywania ćwiczenia, które mogą stanowić zestaw rekomendacji podczas treningu.

4.4. Inne zastosowanie metod przetwarzania szeregów czasowych – Dane giełdowe

[6] Dynamic time warping in financial data - modification of algorithm in context of stock market similarity analysis - Opracowywane algorytmy przetwarzania i klasyfikacji szeregów czasowych zaczerpnięte z dziedziny HRI znalazły również zastosowania w innych projektach i innych dziedzinach nauk niezwiązanych z robotyką, takich jak odnajdywanie wzorców i podobieństwa w danych giełdowych

OŚWIADCZENIE O WKŁADZIE:

Tomasz Grzejszczak, jako pierwszy autor, miał największy wkład w proces powstawania artykułu. Artykuł powstał bazując na jego koncepcji, opracował metodologię, utworzył oprogramowanie, przeprowadził analizę i walidację wyników oraz napisał znaczącą treść wraz z wizualizacją wyników. Główny wkład pozostałych autorów to nadzór merytoryczny, opracowanie treści, recenzja, gromadzenie danych testowych, konsultacja w sprawie proponowanej koncepcji, metodologii i oprogramowania.



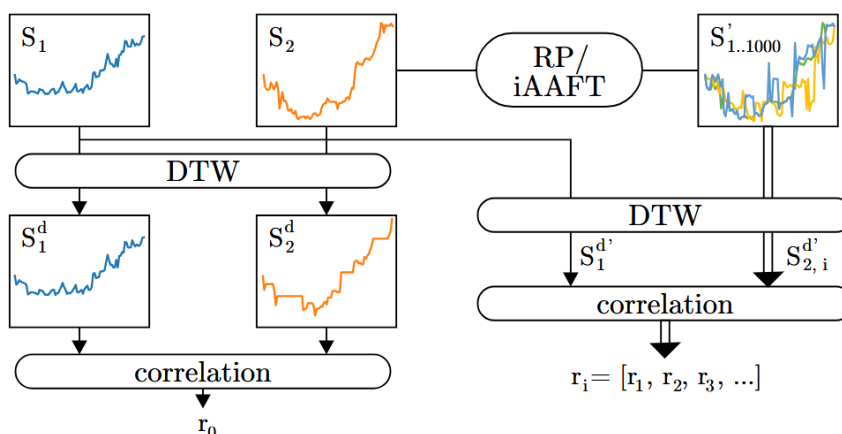
Podobnie jak w przypadku prac związanych z IMU, również tutaj doświadczenie w dziedzinie przetwarzania sygnałów poskutkowało rozpoczęciem prac finansowanych z NRBIR. Pomimo tego, że tematyka odbiega od HRI, to możliwość zastosowania metod przetwarzania i klasyfikacji sygnałów w postaci szeregów czasowych ukazuje interdyscyplinarność i szeroki pozytywny wpływ na rozwój wielu dziedzin nauki.

Głównym algorytmem rozwijanym w trakcie tych badań jest algorytm Dynamic Time Warping (DTW), który stosowany jest do analizy szeregów czasowych, takich jak te wykorzystywane w dotychczasowych badaniach wnioskodawcy do kreowania interakcji człowieka z robotem. Wybór

algorytmu DTW został podyktowany jego szerokim zastosowaniem w innych i pokrewnych dziedzinach oraz możliwością jego rozszerzenia o kolejne modyfikacje, co pozwala na szerokie zastosowanie tego narzędzia w danych finansowych. Należy jednak pamiętać, że DTW jest również szeroko stosowany w innych obszarach, takich jak analiza danych z czujników wykorzystywanych do wczesnego wykrywania usterek i działań sprzętu w pomiarach cyklu, rozpoznawanie gestów na podstawie analizy trajektorii ruchu w obrazie czy analiza i rozpoznawanie mowy. Szeroki zakres zastosowań wyraźnie wskazuje, że obecnie istnieje wiele modyfikacji i udoskonaleń DTW, które mogą zapewnić lepsze parametry dla konkretnych danych.

Miary podobieństwa stanowią istotny element analizy wskaźników giełdowych. Umiejętność prawidłowego odczytywania informacji z danych zawartych na wykresach lub tabelach jest kluczową cechą udanego inwestora. Wskaźniki giełdowe zakładają istnienie pewnych powtarzalnych wzorców. Założenie powtarzalności i wzajemnych powiązań między przyczyną a skutkiem rynku jest często niewypowiedzianą, ale jednocześnie niezbędną podstawą do przeprowadzania analiz związanych z badaniem podobieństwa pewnych wzorców lub ich przewidywaniem na podstawie dostępnych danych. Celem niniejszego artykułu jest zastosowanie metody Dynamic Time Warping (DTW) do analizy danych Giełdy Papierów Wartościowych w Warszawie w celu zbadania skuteczności tego algorytmu na posiadanych danych oraz porównania uzyskanych wyników z metodami klasycznie stosowanymi. Takie rozwiązanie ma szansę pozwolić w przyszłości na swobodne porównywanie dwóch spółek w różnych przedziałach czasowych w celu wykrycia i przeanalizowania wzorców, które wystąpiły już w przeszłości, a także późniejszych konsekwencji, jakie wystąpiły dla danej spółki. Algorytm ten jest w stanie określić nie tylko, czy benchmarki obu spółek są podobne, ale także stopień podobieństwa. Miara ta może stanowić ważny czynnik w podejmowaniu kolejnych decyzji inwestycyjnych. W niniejszym artykule przedstawiamy klasyczne zastosowanie DTW, oraz metodę sprawdzania zgodności poprzez wygenerowanie zbioru podobnych szeregów czasowych z wykorzystaniem losowych permutacji lub transformacji Fouriera.

Celem przeprowadzenia badań jest opracowanie prostej i szybkiej metody wyszukiwania podobnych akcji. Aby znaleźć podobieństwo między dwoma szeregami czasowymi (S_1, S_2), można obliczyć korelację krzyżową $r(S_1, S_2)$. Jednak w cenach akcji dwie spółki mogą wykazywać ten sam trend, ale jedna z nich może być opóźniona w czasie. Aby pominąć efekt opóźnienia, można przeprowadzić analizę DTW. Niestety, badania pokazują, że obliczanie korelacji po DTW może być zawodne. Charakter DTW sprawia, że wynikowe szeregi czasowe (S^d) wykazują większe podobieństwo widoczne przez zwiększenie korelacji $r(S_1^d, S_2^d) > r(S_1, S_2)$. W artykule zaproponowano zestaw testów statystycznych, aby dodatkowo sprawdzić wynik korelacji po DTW. Koncepcja testu statystycznego została przedstawiona na rys. 14.



Rysunek 15 Koncepcja metodologii badań

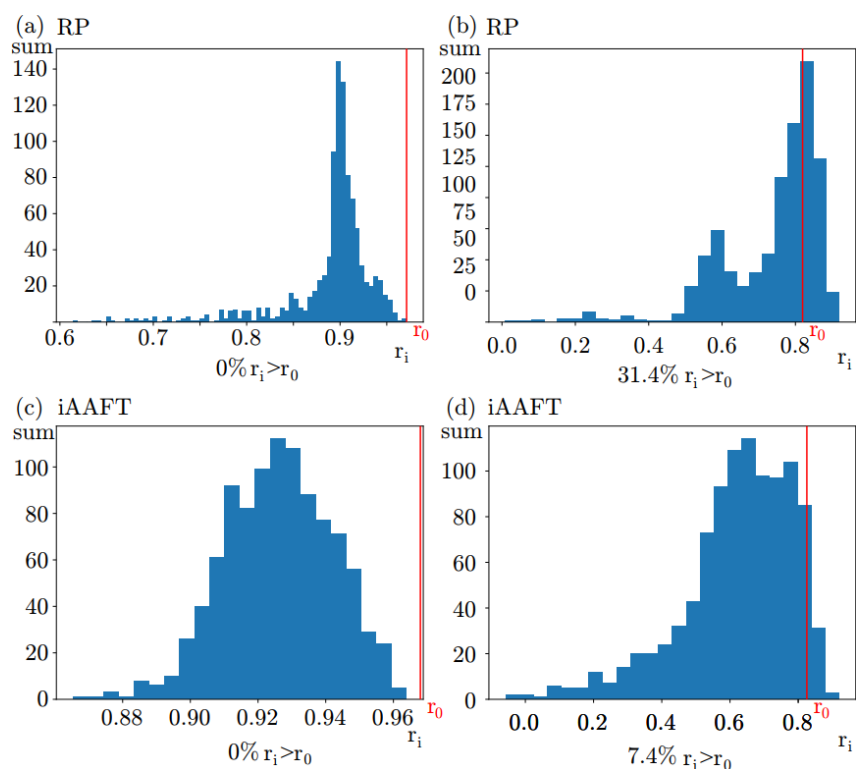
Celem testu statystycznego jest przeprowadzenie testu istotności po losowej zmianie kolejności występowania danych w szeregu, stawiając hipotezę

$$H_0: \rho(S_1^d, S_2^d) = 0 \quad (6)$$

stanowiącą, że para szeregów nie jest podobna. Zatem, aby udowodnić podobieństwo, należy odrzucić tę hipotezę. Aby przetestować tę hipotezę, konieczne jest wygenerowanie losowych kopii szeregów $(S_1^{d'}, S_2^{d'})$ z szeregów (S_1^d, S_2^d) , które są zgodne z H_0 . W tym podejściu wygenerowano 1000 kopii szeregów wejściowych, wykorzystując 2 algorytmy:

- **RP** – losowa permutacja, która losowo tasuje próbki, zachowując ten sam rozkład brzegowy, co szereg oryginalny
- **iAAFT** – iterowana transformata Fouriera z dostosowaną amplitudą, która zachowuje oryginalne widmo mocy. Surogaty szeregów oryginalnych są tworzone za pomocą biblioteki pyunicorn.

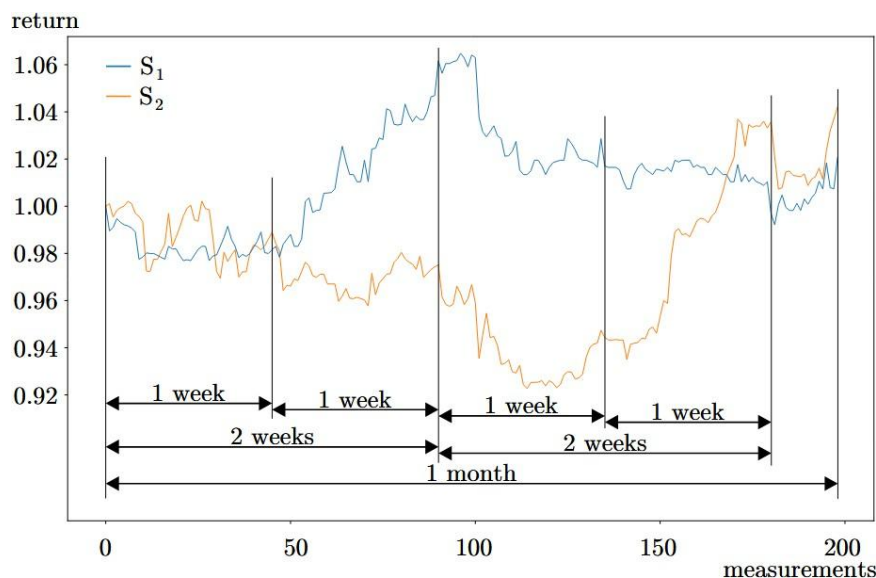
Celem testu jest porównanie oryginalnej korelacji krzyżowej r_0 (S_1^d, S_2^d) z korelacją obliczoną na podstawie sztucznie wygenerowanych szeregów r_i ($S_1^{d'}, S_2^{d'}$), $i=1,2,\dots,1000$. Hipoteza H_0 zostaje odrzucona, jeśli r_0 leży na końcach rozkładu próby o losowym uporządkowaniu danych szeregu czasowego (przetasowaniu). Przykłady wyników testów statystycznych szeregów zgodnych z metodyką z rys. 15, przedstawiono na rys. 16.



Rysunek 16 Histogramy korelacji po DTW sztucznie wygenerowanych próbek r_i (niebieskie) utworzonych przez RP (a-b) i iAAFT (c-d) w porównaniu do oryginalnej korelacji r_0 (czerwone).

Porównanie przeprowadza się między dwoma szeregami czasowymi o tej samej długości. W celu normalizacji danych, wartości akcji w czasie są dzielone przez pierwszą wartość, tworząc w ten sposób wykres stopy zwrotu z akcji. W zależności od preferencji użytkownika, długość szeregu i próbkowanie są zmienne. W przeprowadzonym teście próbkowanie jest ustawione na odczyt ceny akcji co 1 godzinę.

Zestaw eksperymentów przeprowadzono dla 1 tygodnia, 2 tygodni, 1 miesiąca i kwartału. Przykładowe dwa wykresy stopy zwrotu z akcji przedstawiono na rys. 16.

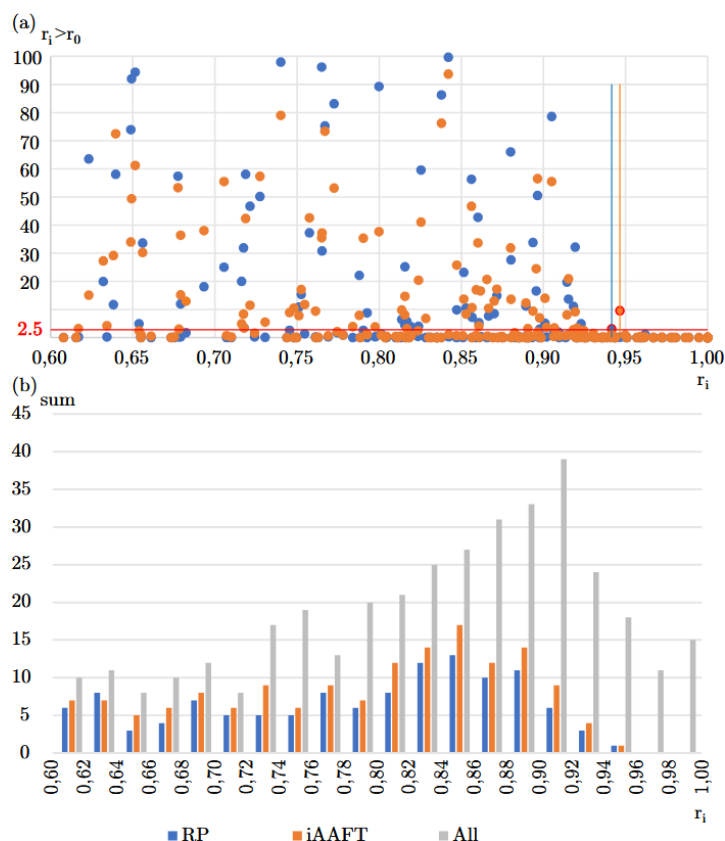


Rysunek 17 Przykład stopy zwrotu dwóch różnych akcji. W zależności od wybranego zakresu, szereg czasowy może mieć różną długość.

Celem testu jest znalezienie wartości progowej korelacji krzyżowej θ która gwarantuje, że wszystkie testy powyżej tej korelacji odrzucają H_0 , co oznacza, że testowana para jest faktycznie podobna. To zapotrzebowanie wynika z praktycznego zastosowania, w którym użytkownik chciałby uzyskać listę podobnych spółek i ich akcji. W zależności od preferencji użytkownika, istnieje możliwość dostosowania pewnych parametrów, które wpływają na czas obliczeń i wyniki. W metodzie DTW istnieją dwa główne parametry: liczba próbek i ograniczenie DTW Sakoe-Chiby. W zależności od preferencji użytkownika, jeden może być zainteresowany długoterminowymi zyskami dziennymi z jednodniowym próbkowaniem, inny może być zainteresowany krótkoterminowym handlem intraday z częstym próbkowaniem. Jednak z punktu widzenia algorytmu, wszystkie te parametry wpływają w rzeczywistości tylko na jedną ostatnią zmienną, czyli liczbę próbek.

Kluczową wizualizację wyników przeprowadzonych testów statystycznych przedstawia rys. 18, który pozwala na wyznaczenie obiektywnych kryteriów podobieństwa szeregów czasowych. Wykres punktowy (rys. 18 a) pokazuje procentowy udział przypadków, w których korelacja szeregów wygenerowanych sztucznie była większa od korelacji oryginalnej pary szeregów czasowych. Czerwona linia na poziomie 2,5% wyznacza próg odrzucenia hipotezy zerowej. Punkty znajdujące się poniżej tej linii oznaczają pary, które zostały uznane za statystycznie podobne. Poniżej, na rys. 18 b, znajduje się histogram wyników. Szare słupki reprezentują całkowitą liczbę badanych par w danym przedziale korelacji, natomiast kolorowe słupki (dla różnych algorytmów losowego generowania danych, czyli niebieskie dla RP i pomarańczowe dla iAAFT) pokazują liczbę przypadków, w których nie udało się odrzucić hipotezy (czyli mimo wysokiej korelacji, testy nie potwierdziły istotnego statystycznie podobieństwa).

Badania dowodzą istnienie wartości progowej możliwej do użycia w wyszukiwaniu spółek o podobnej zmienności kursu. Wykazano, że istnieje konkretna wartość korelacji, powyżej której hipoteza o braku podobieństwa jest zawsze odrzucana. Dla danych z tego eksperymentu progi te wyniosły 0,94 dla metody RP oraz 0,95 dla metody iAAFT.



Rysunek 18 Wynik testów statystycznych. (a) Procentowe umiejscowienie r_0 w rozkładzie r_i w zależności od korelacji krzyżowej r_0 . (b) Histogram wyników testów, gdzie szary kolor oznacza całkowitą liczbę testów w zakresie korelacji, a kolorowe słupki odpowiadają zaakceptowanym H_0 .

Efektom przeprowadzonych prac było wyznaczenie wartości progowych korelacji, które gwarantują odrzucenie hipotezy o braku podobieństwa dwóch szeregów czasowych. Wykazano, że możliwe jest skuteczne wykrywanie wysokiego podobieństwa między wykresami giełdowymi przy użyciu DTW i jego obiektywna weryfikacja za pomocą testów statystycznych.

Badania te wnoszą istotny wkład w rozwój finansów ilościowych oraz informatyki ekonomicznej, przesuując punkt ciężkości z tradycyjnej analizy fundamentalnej w stronę zaawansowanych algorytmów uczenia maszynowego i analizy szeregów czasowych, wspierając interdyscyplinarność. Wprowadzenie testów statystycznych do analizy wyników DTW w kontekście danych giełdowych jest rozwiązaniem nowatorskim, które pozwala na określenie prawdopodobieństwa wystąpienia danego wzorca. Ponadto, prace te były realizowane i finansowane w ramach projektu NCBiR (01.01.01-00-0162/19). Projekt ten, został oceniony przez instytucję finansującą jako innowacyjny. Badania opisane w artykule stanowiły istotną część merytoryczną tego przedsięwzięcia, a pozytywne przyjęcie sprawozdania końcowego oraz ostateczne rozliczenie projektu potwierdzają wartość naukową i wdrożeniową uzyskanych rezultatów.

Zaproponowana metoda statystycznej weryfikacji istotności podobieństwa sygnałów w postaci szeregów czasowych za pomocą algorytmu DTW znajduje bezpośrednie zastosowanie w HRI, umożliwiając precyzyjne rozpoznawanie gestów oraz trajektorii ruchu człowieka o zmiennej dynamice czasowej. Dzięki wprowadzeniu obiektywnych wartości progowych korelacji, systemy sterowania mogą skuteczniej odróżniać celowe sygnały i wzorce zachowań użytkownika od przypadkowych ruchów, co znacząco podnosi niezawodność i bezpieczeństwo autonomicznych systemów robotycznych.

4.5. Modelowanie stanu maszyny

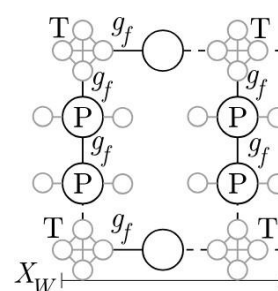
[8] Cooperative and individual planning strategies for autonomous robots in warehouse parcel transportation – Następnym istotnym krokiem w interakcji człowieka z robotem jest modelowanie stanu maszyny. W tym celu prowadzono badania w dziedzinie planowania bazującego na danych zmiennych w czasie w celu kooperacyjnego sterowania robotami

OŚWIADCZENIE O WKŁADZIE:

Artykuł naukowy zawiera dwa aspekty:

- 1) Symulacja środowiska bazująca na automatach komórkowych – Witold Nocoń
- 2) Koncepcja algorytmów kooperacyjnych – wkład: Tomasz Grzeszczak

Obaj autorzy ściśle współpracowali w każdym z etapów powstawania publikacji i każdy autor posiada równy wkład (50%) w każdy aspekt publikacji.



Podczas gdy wcześniejsze prace w obszarze interakcji człowiek-robot (HRI) koncentrowały się głównie na etapie akwizycji i przetwarzania danych sensorycznych (szeregów czasowych), niniejszy artykuł przesuwa punkt ciężkości na kolejny, kluczowy etap: planowanie i podejmowanie decyzji. Kontekstem badawczym jest kooperacyjne planowanie tras dla zespołu autonomicznych robotów jeżdżących w środowisku magazynowym. Poruszając się myślą zarysowaną we wstępie wniosku i rys. 1 - 2, obecny artykuł porusza tematykę zawartą w etapie związanym z planowaniem symbolicznym.

Środowisko magazynowe oraz zadania logistyczne zostały sformalizowane w języku opisowym PDDL (Planning Domain Definition Language). Notacja ta pozwoliła na precyzyjne zdefiniowanie przestrzeni stanów oraz zbioru możliwych akcji dla robotów. Na tym fundamencie zbudowano i zbadano algorytmy planowania, których celem było porównanie efektywności strategii czysto indywidualnych (gdzie każdy robot wykonuje swoje zadanie niezależnie) oraz kooperacyjnych (w których roboty aktywnie koordynują swoje działania, aby uniknąć konfliktów i osiągnąć globalnie lepsze rezultaty). Głównym wkładem pracy jest więc analiza i ocena, w jaki sposób formalne modele planowania oraz mechanizmy współpracy mogą zwiększyć wydajność i niezawodność zautomatyzowanych systemów magazynowych. Dodatkowo, warto zauważyć, że w klasycznym podejściu uczenia maszynowego (nie głębokiego), poprawne zdefiniowanie przestrzeni akcji, stanu i funkcji celu jest pierwszym krokiem do przeprowadzenia algorytmu uczenia ze wzmocnieniem (np. Q-learning)

Algorytm kooperacyjny opisany w niniejszym artykule opiera się na podejściu „Bucket Brigade”, czyli koncepcji zaczerpniętej z efektywnej walki z pożarem polegającej na podawaniu wiader kolejnej osobie z rzędu. To podejście jest implementowane głównie w harmonogramowaniu wieloagentowym i jest również wykorzystywane w innych dziedzinach badań do optymalizacji zużycia energii. Ponadto opracowane algorytmy opierają się na teorii automatów komórkowych do implementacji ruchu, modelu topologicznym do obliczania kosztów w reprezentacji mapy świata oraz algorytmach STRIPS lub PDDL do implementacji algorytmów planowania.

Innowacyjność zaproponowanego rozwiązania opiera się na wprowadzeniu strategii kooperacyjnej umożliwiającej pozostawianie paczek w punktach pośrednich trasy, co stanowi istotne novum w stosunku do klasycznych podejść logistycznych. Przeprowadzona w artykule analiza literatury, obejmująca m.in. warianty problemu multi-agent pickup and delivery (MAPD), wykazała, że żadne z dostępnych rozwiązań nie zakłada możliwości przerwania realizacji zamówienia w trakcie jego trwania w celu przekazania ładunku innemu agentowi; wszystkie zidentyfikowane metody opierają się na sztywnym modelu „od pobrania do dostarczenia” (pickup-and-delivery) przez tę samą jednostkę. Moje badania wypełniają tę lukę, wprowadzając zdecentralizowany algorytm, który pozwala robotom na

autonomiczne porzucanie i przejmowanie zadań w zależności od aktualnej konfiguracji otoczenia, co jak wykazano w testach, pozwala na znaczną optymalizację czasu pracy i redukcję pustych przebiegów bez potrzeby stosowania centralnego nadzoru.

Główną jednostką robotyczną lub agentem zautomatyzowanego magazynu jest AVG (Automated Guided Vehicle). Definicja systemu wieloagentowego wymaga, aby agenci byli autonomiczni i pracowali w sposób zdecentralizowany. Co więcej, w środowisku wieloagentowym zachowanie agenta można określić jako kooperacyjne, jeśli wspólna użyteczność jest świadoma i chętnie zwiększana w porównaniu z użytecznością referencyjną lub kooperatywne wykonywanie tego samego algorytmu.

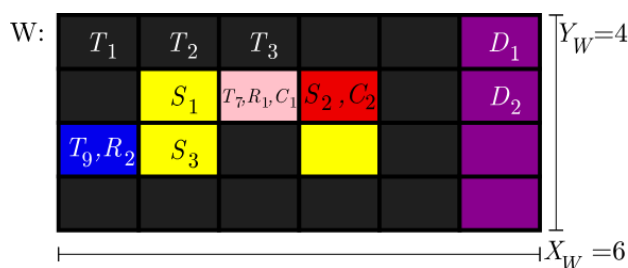
W artykule przedstawiono dwa podejścia do planowania robotów w obszarze zdecentralizowanych robotów autonomicznych. W pierwszym podejściu każdy robot pobiera ładunek i dostarcza go bezpośrednio do miejsca zrzutu. Taka strategia jest w niniejszym artykule nazywana strategią indywidualną. W strategii kooperacyjnej roboty starają się przekazywać ładunek innym robotom, a także odbierać ładunek pozostawiony przez inne roboty. Fizyczna konstrukcja magazynu ogranicza ruch robotów, co powoduje kolizje i korki, które należy obsłużyć. Dodatkowo, ze względu na charakter modelu fizycznego, na którym oparta jest prezentowana symulacja, ruch po linii prostej jest znacznie szybszy niż skręcanie.

Notacja formalna pozwoliła na opracowanie i zaimplementowanie dwóch algorytmów autonomicznego sterowania robotami. Symulacja jest zdefiniowana za pomocą formalnej notacji krotki $\langle P, O, W \rangle$ podobnej do notacji PDDL lub STRIPS, zawierającej zbiór predykatów P , operatorów O oraz świat W , który jest modelem magazynu.

Model użyty do symulacji opisany został zgodnie z podejściem PDDL. W symulacji istnieje magazyn $W[X_W, Y_W]$ o zadanej szerokości X_W i wysokości Y_W . Magazyn zawiera zestaw robotów R , torów T , regałów magazynowych S , ładunków C i miejsca docelowego D , więc $W=\{R, T, S, C, D\}$. Magazyn jest wizualizowany na rysunku 19. W zbiorze ładunków znajduje się N_C ładunków, z których każdy ma swoją pozycję w magazynie, więc

$$C = \{C_1(x_{C1}, y_{C1}), C_2(x_{C2}, y_{C2}), \dots, C_{N_C}(x_{CNC}, y_{CNC})\} \quad (7)$$

oraz $x_C \in X_W, y_C \in Y_W$. Podobne definicje sformułowano dla T, S, D i R .



Rysunek 19 Model przykładowego magazynu $W=\{R, T, S, C, D\}$

Tak rozpisana notacja formalna posłużyła do zdefiniowania stanu magazynu z użyciem listy predykatów oraz możliwych do wykonania akcji, które modyfikują zestaw predykatów. Robot może zostać zaprogramowany do stosowania jednej z dwóch strategii planowania: indywidualnej (kod na rys.20) lub kooperacyjnej (kod na rys 21). Oba algorytmy zostały przedstawione za pomocą pseudokodu. Najważniejszą częścią każdego algorytmu jest część planowania z algorytmem wyszukiwania ścieżki, którego wynikiem jest ścieżka P . Jego celem jest znalezienie zestawu współrzędnych kolejnych punktów w ścieżce P , które łączą miejsce odbioru paczki z miejscem docelowym.

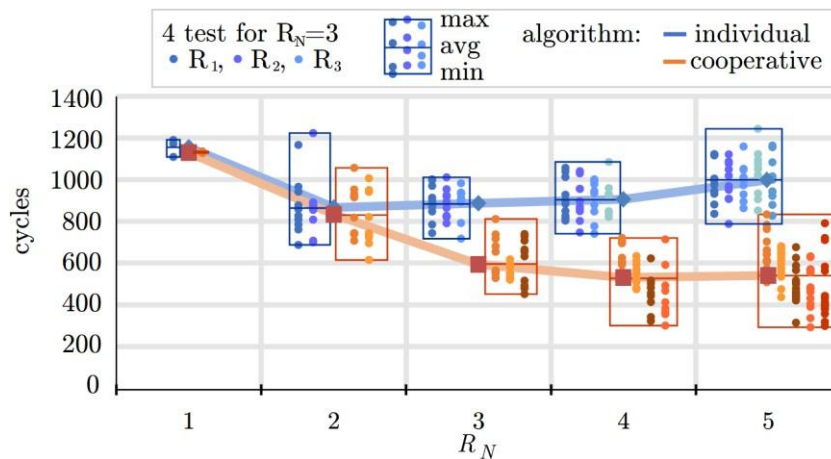
Require: $R_n(x_{Rn}, y_{Rn}), C_m(x_{Cm}, y_{Cm}), D_m(x_{Dm}, y_{Dm})$	
Ensure: $(x_{Cm}, y_{Cm}) = (x_{Dm}, y_{Dm})$	
1:	For given n find m that indicates C_m closest to R_n
2:	$GOTO_PICK(R_n, C_m)$: Find shortest path P_1 from (x_{Rn}, y_{Rn}) to $(x_{Cm} + 1, y_{Cm})$
3:	for each element i of P_1 do
4:	if $T_{P1i} \neq \text{empty}$ then
5:	wait (random(1–5) cycles)
6:	Go to: 1
7:	else
8:	$MOVE(R_n, T(x_{Rn}, y_{Rn}), T(P_{1i}))$
9:	end if
10:	end for
11:	$PICK_CARGO(R_n, C_m)$
12:	$GOTO_DROP(R_n, D_m)$: Find shortest path P_2 from (x_{Rn}, y_{Rn}) to $(x_{Dm} - 1, y_{Dm})$
13:	for each element i of P_2 do
14:	if $T_{P2i} \neq \text{empty}$ then
15:	wait (random(5–10) cycles)
16:	Go to: 12
17:	else
18:	$MOVE(R_n, T(x_{Rn}, y_{Rn}), T(P_{2i}))$
19:	end if
20:	end for
21:	$DROP_CARGO(R_n, D_m)$

Rysunek 20 Algorytm indywidualny

Require: $R_n(x_{Rn}, y_{Rn}), C_m(x_{Cm}, y_{Cm}), D_m(x_{Dm}, y_{Dm})$	
Ensure: $(x_{Cm}, y_{Cm}) = (x_{Dm}, y_{Dm})$	
1:	Find alley x_a with $\max(\sum ON_SHELF(C_a, S_{(xa-1, y)}) - \sum ON_SHELF(C_a, S_{(xa+1, y)}))$
2:	Find m that indicates first C_m with $x_{Cm} = x_a - 1$
3:	$GOTO_PICK(R_n, C_m)$: Find shortest path P_1 from (x_{Rn}, y_{Rn}) to $(x_{Cm} + 1, y_{Cm})$
4:	for each element i of P_1 do
5:	if $TP_{1i} \neq \text{empty}$ then
6:	wait (random(1–5) cycles)
7:	Go to: 1
8:	else
9:	$MOVE(R_n, T(x_{Rn}, y_{Rn}), T(P_{1i}))$
10:	end if
11:	end for
12:	$PICK_CARGO(R_n, C_m)$
13:	$TRY_DROP_RIGHT(R_n)$: Find y_e with $\neg ON_SHELF(C(x_a + 1, y_e), S(x_a + 1, y_e))$
14:	if $TRY_DROP_RIGHT(R_n)$ failed then
15:	repeat with $x_a = x_a + 2$
16:	end if
17:	$GOTO_DROP(R_n, D_m)$: Find shortest path P_2 from (x_{Rn}, y_{Rn}) to (x_a, y_e)
18:	for each element i of P_2 do
19:	if $T_{P2i} \neq \text{empty}$ then
20:	wait (random(5–10) cycles)
21:	Go to: 17
22:	else
23:	$MOVE(R_n, T(x_{Rn}, y_{Rn}), T(P_{2i}))$
24:	end if
25:	end for
26:	$DROP_CARGO(R_n, (x_a, y_e))$

Rysunek 21 Algorytm kooperacyjny

Opracowana symulacja magazynu umożliwiła dowolną modyfikację środowiska testowego. Oba algorytmy przetestowano w 2 różnych scenariuszach zakładających różną wielkość magazynu. Testy wykonano dla różnej liczby robotów. Sprawdzone potrzebną liczbę cykli do zrealizowania wszystkich zadań, całkowite zużycie energii oraz zliczenie pustych przebiegów. Dane obszernie przedstawiono i skomentowano w artykule. Jeden przykładowy wykres prezentujący dane przedstawiono na rys. 22.



Rysunek 22 Średnia liczba cykli dla grupy robotów realizujących strategię kooperacyjną lub indywidualną

W związku z tym, że scenariusze testowe zawierały elementy losowości, każda konfiguracja poważana była kilkakrotnie, dlatego występujące na rys. 22 punkty wykresu powtarzają się kilkakrotnie w danej kolumnie. Wraz ze wzrostem liczby robotów R_N , wzrastała liczba kolumn z punktami w danej serii, a wynik każdego robota przedstawiono w innym odcieniu. Ostatecznie punkty uśrednione zostały zwizualizowane łamaną linią trendu, która ukazuje zdecydowaną przewagę rozwiązania kooperacyjnego, który posiada średnio prawie dwukrotnie szybsze ukończenie zadania niż podejście indywidualne.

Dodatkową zaletą przedstawionej strategii kooperacyjnej jest to, że dobierana jest optymalna liczba współpracujących robotów. Selekcja ta odbywa się naturalnie w oparciu o mechanizm rezerwacji globalnej listy zamówień. Gdy wszystkie korytarze są zarezerwowane, dodatkowy robot wraca na swoje miejsce parkingowe lub miejsce ładowania.

Główna wartość naukowa polega na opracowaniu i formalnej analizie porównawczej strategii indywidualnego oraz kooperacyjnego planowania tras dla floty robotów magazynowych. Wprowadzenie formalnego opisu dziedziny w języku PDDL umożliwiło precyzyjne modelowanie problemu i stworzyło solidne podstawy dla oceny algorytmów. W rezultacie praca dostarczyła konkretnych dowodów oraz analiz, które przyczyniły się do rozwoju wiedzy na temat efektywności planowania kooperacyjnego, oferując ważny punkt odniesienia i metodologiczny fundament dla dalszych badań nad inteligentną koordynacją systemów wieloagentowych w złożonych środowiskach logistycznych.

Istotnym wkładem prowadzonych badań jest opracowanie zdecentralizowanego algorytmu planowania, który przekształca globalny problem optymalizacji w lokalne decyzje poszczególnych agentów. Udowodniono, że inteligentne zachowania stadne (kooperacja bez komunikacji) mogą skutecznie kompensować ograniczenia fizyczne robotów i mieć zastosowanie w magazynach, w których większość rozwiązań wciąż bazuje na przydzielaniu pojedynczych zadań typu „od pobrania do dostarczenia”. Badania wpisują się w spójny cykl publikacji, ponieważ rozwijają metody zdecentralizowanego sterowania, które są fundamentem bezpiecznej pracy robotów w środowiskach współdzielonych z ludźmi, gdzie brak centralnego nadzoru nad każdym ruchem zwiększa elastyczność systemu.

5. Informacja o wykazywaniu się istotną aktywnością naukową albo artystyczną realizowaną w więcej niż jednej uczelni, instytucji naukowej lub instytucji kultury, w szczególności zagranicznej.

Główną afiliacją wnioskodawcy jest Politechnika Śląska w Gliwicach i to tam realizowano istotną aktywność naukową opisaną w niniejszym wniosku. Najważniejszym efektem prowadzonej latami pracy naukowej była 4-krotna zespołowa nagroda Rektora stopnia I (2016 r., 2017 r.) i II (2023 r., 2024 r.).

Odnosząc się do innych, w szczególności zagranicznych instytucji naukowych, w roku 2017 odbyło 3 miesięczny staż naukowy w Niemczech na: University of Wuppertal, School of Electrical, Information, Media and Engineering, 42119 Wuppertal, Germany, Department of Automation/Computer Science, Rainer-Gruenter-Str. 21. Staż zrealizowano w ramach projektu: „Wyjazd stażowy w celu prowadzenia badań i poszerzania wiedzy z zakresu wykrywania gestów” finansowanego przez NCN – MINIATURA1

5.1. Opis celu działania naukowego

W ramach prowadzonych prac rozwijany jest algorytm analizy gestów rąk i śledzenia punktów charakterystycznych podlegający pod dziedzinę interfejs człowiek-maszyna. Celem działania naukowego było badanie kolejnego etapu, po etapie opracowania algorytmu wykrywania i lokalizacji punktów charakterystycznych dłoni, czyli sformułowanie reguł decyzyjnych i konstrukcja klasyfikatora w celu rozpoznawania i wykrywania gestów. W trakcie stażu badane były metody klasyfikacji w zastosowaniu do rozpoznawania gestów bazując na zbiorze punktów charakterystycznych.

5.2. Opis realizacji działania naukowego

W ramach projektu opłacona została delegacja na wyjazd stażowy, czyli dieta, nocleg oraz transport. Pozyskane środki pozwoliły na trzymiesięczną pracę na Uniwersytecie w Wuppertalu, wydziale Inżynierii Elektrycznej, Informacyjnej i Mediów, na stanowisku wolontariusza. Staż rozpoczął się dnia 14 września a zakończył 13 grudnia 2017 roku.

Głównym tematem badań było odnalezienie metod klasyfikacji zbioru wykrytych punktów charakterystycznych. Po przedstawieniu tematyki na zebraniu zakładu wywiązała się dyskusja, której konkluzją była sugestia zastosowania Konwolucyjnych Sieci Neuronowych (ang. Convolutional Neural Networks), czyli metody Głębokiego Uczenia (ang. Deep Learning).

Opracowywany algorytm wykorzystuje trójetapowe podejście polegające na segmentacji, przetwarzaniu oraz klasyfikacji. W klasycznym podejściu, programista tworzy reguły, które implementuje w postaci algorytmu. W zasugerowanym podejściu głębokiego uczenia tworzony jest tylko klasyfikator w postaci dużej, konwolucyjnej sieci neuronowej, do której dostarczane są surowe, nieprzetworzone dane. Na etapie trenowania sieci neuronowej klasyfikator samemu dobiera odpowiednie cechy, które w najlepszy sposób rozróżniają klasyfikowane dane, wzmacniając odpowiednie połączenia neuronowe. Reguły eksperckie nie są wymagane.

Zasugerowane podejście podważa dotychczasowy sposób realizacji i podejście do tworzenia algorytmu analizy gestów rąk i śledzenia punktów charakterystycznych. Dlatego, w trakcie stażu skupiono się głównie na sprawdzeniu tego podejścia. Ostatecznie, w ramach stażu, prowadzono jednocześnie:

- badania dokładności śledzenia wykrytych punktów charakterystycznych przy wykonywaniu gestów dynamicznych,
- naukę o Konwolucyjnych Sieciach Neuronowych, jako metody klasyfikacji gestów statycznych,
- badaniu możliwości praktycznego zastosowania metod wykrywania gestów jako część interfejsu człowiek-komputer.

Ponadto, w ramach współpracy i wymiany wiedzy, pomagałem przy realizacji prac, które związane były z osobistym rozwojem oraz nawiązywaniem kontaktów. Wymienione niżej prace miały między innymi na celu przyszłe współautorskie publikacje.

- Konsultacja przy projekcie rekonstrukcji, kategoryzacji odłamków antycznej ceramiki.
- Konsultacja przy projekcie symulacji i wizualizacji 3D pola elektromagnetycznego.
- Konsultacja przy pracy inżynierskiej o tematyce projekcji obrazu i danych w rzeczywistości rozszerzonej.
- Pomoc przy egzaminie i prowadzenie wykładu o tematyce metody przetwarzania obrazu w interfejsach człowiek-komputer, w ramach przedmiotu Multimodalne interfejsy człowiek-maszyna.
- Wstępny etap konsultacji nad opracowaniem projektu dotyczącego zastosowania algorytmu wykrywania gestów.

5.3. Opis efektów realizacji działania naukowego

Prowadzone prace były podzielone na 2 tematy, związane z wykrywaniem gestów dynamicznych oraz zastosowaniem rozwijającej się w tamtym okresie dziedziny głębokiego uczenia do rozpoznawania gestów

5.3.1. Śledzenie punktów charakterystycznych w celu analizy gestów dynamicznych

W ramach badania dokładności śledzenia wykrytych punktów charakterystycznych przy wykonywaniu gestów dynamicznych przebudowano istniejący algorytm. W badaniach skupiono się na sprawdzeniu dokładności detekcji punktów przy ich śledzeniu oraz na powtarzalności przy wykrywaniu gestu statycznego. Badania porównano do wyników odnalezionych w przeglądzie literatury. Badania wykazały, że opracowywany algorytm bazujący na analizie sekwencji wideo jest niewiele gorszy niż badany, obecnie najczęściej stosowany czujnik Leap Motion, bazujący na detekcji głębi w refleksji światła podczerwonego, dedykowany do wykrywania ruchów dłoni. Przewagą opracowywanego algorytmu jest możliwość pracy na każdym sprzęcie z ogólnodostępną kamerą i nie wymaga dedykowanego sprzętu. W ramach podsumowania tej pracy powstała publikacja, która po poprawkach została opublikowana w czasopiśmie Archives of Control Sciences [2] (Opisana dokładnie w rozdziale 4.2 wniosku)

5.3.2. Badanie zastosowania sieci konwolucyjnych do rozpoznawania gestów

Możliwość nauki o Konwolucyjnych Sieciach Neuronowych, jako metody klasyfikacji gestów statycznych, była najbardziej przełomową częścią stażu, wpływającą na dalszy, osobisty rozwój kariery naukowej. W ramach tego etapu poznałem podstawy i możliwości metod głębokiego uczenia. Nauczyłem się programowania w języku Python z wykorzystaniem biblioteki Keras i TensorFlow oraz wykorzystania procesora graficznego wspomagającego CUDA do obliczeń wielowątkowych. Stworzyłem program do klasyfikacji gestów bazując na posiadanej wcześniej niewielkiej, jak na potrzeby trenowania konwolucyjnych sieci neuronowych, bazie danych. Publikacja, po dodaniu wyników, została wygłoszona na konferencji European Simulation and Modelling Conference 2018. [3] (Opisana dokładnie w rozdziale 4.3 wniosku)

Pozornie, zastosowanie głębokiego uczenia podważa stosowność prowadzenia prac wspomnianych w punkcie pierwszym, jednak jest kilka cech przemawiających za przewagą klasycznego podejścia. Związane są one z faktem, że do trenowania dużej sieci neuronowej wymagana jest duża baza obrazów oraz duża moc obliczeniowa. Proste reguły decyzyjne przenoszą się na szybkość działania algorytmu. Jednak aby to zweryfikować, należy zmodyfikować funkcje celu przy uczeniu sieci neuronowej tak, aby

celem uczenia było jak najlepsze dopasowanie punktów charakterystycznych dłoni. Będzie to tematem przyszłych badań.

5.3.3. Inne efekty

Celem badania możliwości praktycznego zastosowania metod wykrywania gestów jako część interfejsu człowiek-komputer było odnalezienie zapotrzebowania rynkowego. Wynikiem tych rozmów było wskazanie potencjalnego odbiorcy. Wskazano, że rozwiązanie może być wykorzystywane przy stanowiskach pracy manualnej do nauki lub oceny wykonania precyzyjnego ruchu dłoni.

Efektom prac realizowanych w ramach wymiany wiedzy była jedna współautorska publikacja: „Adapting strategies to display simulations of electric fields spawned by point charges for augmented reality application”, która odbiega od tematyki wniosku i nie została w nim opisana.

6. Informacja o osiągnięciach dydaktycznych, organizacyjnych oraz popularyzujących naukę lub sztukę.

Uczestnictwo w akcjach popularnonaukowych:

- Targi edukacyjne (Reprezentacja kierunku Automatyka i Robotyka, Pol.Śl.): 4
- Dni otwarte Politechniki Śląskiej: 5
- Noc Naukowców (Politechnika Śląska): 4
- Śląski Festiwal Nauki (Katowice): 2
- Zawód od kuchni (MOK w Zabrze): 4
- Zajęcia popularnonaukowe dla Zabrzeńskich szkół i przedszkoli w tematyce robotyki i sztucznej inteligencji (IX LO, 1 Tech., SP 43, SP 42, Przedsz. 34): 7

Promotor: **14** prac magisterskich (w tym 4 w języku angielskim)

29 prac inżynierskich (w tym 7 w języku angielskim)

3 prac inżynierskich prowadzonych na Yanshan University (YSU), Qinhuangdao, Chiny

promotor pomocniczy: **11** prac inżynierskich prowadzonych na Yanshan University (YSU), Chiny

Opiekun koła naukowego: MKN Social Robots (2020r.-2023r.),
SpaceCoffee (2023r.-obecnie)

Opiekun naukowy projektów studenckich: Project Based Learning (PBL): **6** projektów,
Inicjatywa Doskonałości – Uczelnia Badawcza (IDUB): **5** projektów.

Mentor w programie mentorskim „Rozwiń Skrzydła”: dla 2 osób (2019r. – 2023r. oraz 2023r. - obecnie)

7. Inne informacje, ważne z punktu widzenia wnioskodawcy, dotyczące kariery zawodowej.

Brak

.....
(podpis wnioskodawcy)