

Recenzja Rozprawy Doktorskiej
mgr inż. Filipa Malawskiego
przygotowana dla Rady Wydziału Elektroniki, Informatyki i Telekomunikacji
Akademii Górniczo-Hutniczej w Krakowie

Tytuł rozprawy: **Automatic analysis of techniques and body motion patterns in sport**
(Automatyczna analiza technik oraz form ruchu ciała w sporcie)

Autor rozprawy: **mgr inż. Filip Malawski**

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy /teza rozprawy/ i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Rozprawa mgr inż. Filipa Malawskiego dotyczy wybranych problemów analizy danych pochodzących z inercyjnych czujników nasobnych człowieka lub zadanych znaczników umieszczonych na ostrzu szermierza i danych z kamery RGB-D, związanych z automatyczną analizą ruchu w sporcie. Ze względu na swoją tematykę, charakter rozwiązań i typ analizowanych danych rozprawę można zaliczyć do obszarów badawczych informatyki, w szczególności obejmujących metodyki „Sztucznej Inteligencji” i „Rozpoznawania Obrazów” służące do „wizualnej analizy dynamicznych scen” i „interpretacji danych o ruchu człowieka”.

Teza rozprawy, czyli zagadnienie naukowe (zadanie badawcze) rozpatrywane w pracy została sformułowana dostatecznie jasno, jednak w mojej ocenie zbyt ogólnie. Teza brzmi następująco: *„Automatyczna analiza techniki oraz ruchu ciała sportowca generuje kanał zwrotny, użyteczny na potrzeby doskonalenia umiejętności sportowych.”*

Oczekiwałbym przynajmniej wskazania charakteru zastosowanych lub nowo proponowane metodyk analizy strumieni danych, charakteru analizowanych danych (sygnały z czujników inercyjnych, sekwencje przetworzonych obrazów z Kinect-a – mapa głębi i szkielet złączy) i określenia stopnia uniwersalności bądź specjalizacji proponowanych technik dla wybranej dyscypliny sportowej.

Rozprawa posiada jasno określony cel praktyczny, jakim jest opracowanie metod pozwalających na automatyczną analizę ruchu w sporcie, które mogłyby wspomagać zawodników oraz trenerów w procesie doskonalenia umiejętności sportowych. W szczególności, celem badań i rozwiązań przedstawionych w rozprawie było zapewnienie użytecznej informacji zwrotnej dla zawodnika co do poprawności wykonywanych technik, także bez nadzoru trenera.

Rozprawa ma dominujący charakter eksperymentalny i wzbogacona jest o dyskusję naukową dotyczącą opracowania i wyboru algorytmów. Oprócz wstępu i podsumowania, rozprawa zawiera cztery rozdziały, które odpowiadają czterem grupom zagadnień i są kluczowe dla oceny

rozprawy. Najpierw Autor dokonuje przeglądu istniejących rozwiązań, algorytmów i technik analizy danych o ruchu człowieka oraz skupia się na wyjaśnieniu problemu automatycznej analizy ruchu stosowanej w celu wspomagania treningu sportowca. Następnie kolejno przedstawiane są wyniki prac badawczych zebrane w trzech rozdziałach. Każde z nich obejmuje analizę problemu, ogólny opis autorskiego rozwiązania i wyniki jego eksperymentalnej weryfikacji w rzeczywistych scenariuszach analizy ruchu nóg lub narzędzia szermierza w systemie wspomagania treningu szermierki.

- Najpierw omawiane jest zagadnienie modelowania akcji specyficznych dla sportu szermierczego, które mają podobne trajektorie ruchu, ale inną dynamikę wykonania, co przekłada się na różne techniki sportowe. Proponowany i implementowany jest system rozpoznawania akcji nóg szermierza, obejmujący akwizycję i analizę pochodzących z Kinect-a i czujników inercyjnych, ekstrakcję sekwencji cech i ich klasyfikację.
 - Następnie omawiane są zagadnienia segmentacji czasowej procesów ciągłego ruchu oraz oceny jakości wykonania przez szermierza akcji „wypadu”. Proponowane rozwiązanie obejmuje detekcję i lokalizację segmentów w sekwencji czasowej danych pochodzących z Kinect-a i czujnika inercyjnego a także metodę jakościowej oceny poprawności wykonanej akcji.
 - Na koniec analizowane są sposoby akwizycji i analizy danych o ruchu ostrza. Przedstawiono metody śledzenia punktów ostrza w czasie i uczenia się ich poprawnych trajektorii. Wyniki analizy ruchu narzędzia szermierza przekazywane są zwrotnie do sportowca w sposób intuicyjny i w czasie rzeczywistym, z wykorzystaniem nowoczesnej technologii obrazowania na pół-przezroczystych okularach.
2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle / świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Rozprawa liczy 97 stron właściwego tekstu, podzielonego na 6 rozdziałów. W ocenie rozmiaru należy uwzględnić fakt, że tekst pisany jest małą czcionką (10 pt. (?)) z pojedynczym odstępem. Praca zawiera także 22 strony bibliografii (314 pozycji literatury), spis treści, spis 25 tabel i 49 rysunków.

Już na podstawie samej liczby pozycji literatury można przypuszczać, że Autor dokonuje bardzo szczegółowego i obszernego przeglądu literatury adekwatnej do tematyki rozprawy. Sam Autor za jedno z zasadniczych osiągnięć rozprawy uznaje „wyczerpujący przegląd literatury w zakresie metod rozpoznawania ruchu, w szczególności ruchu w sporcie, w oparciu o różne modalności sygnałów”. Mogę potwierdzić, że tak jest w istocie. Przegląd literatury został podzielony na dwie zasadnicze części: metody analizy danych o ruchu człowieka i zastosowania metod analizy ruchu w sporcie.

- Przegląd metod jest bardzo dobrze uporządkowany w oparciu o właściwie kryterium rodzaju analizowanych danych. Autor omawia analizę danych o postaci: sekwencji czasowych obrazów RGB (dane kolorowego wideo), sekwencji map głębi i obrazów RGB-D oraz sygnałów pochodzących z czujników inercyjnych. Dokonuje On także przeglądu rozwiązań łączących ze sobą analizę danych tych trzech zasadniczych typów. Autor odwołuje się do wielu artykułów naukowych (16) mających charakter przeglądowy, zasadniczo opublikowanych w okresie lat 2008-2018. W przeglądzie nie

zabrakło istotnego odwołania się do sztucznych sieci neuronowych, stanowiących nowoczesne i burzliwie rozwijające się metodyki stosowane m.in. do rozpoznawania obrazów. Są to poprawnie przez Autora zidentyfikowane sieci głębokie: CNN (warto tu także dodać R-CNN), RNN czy LSTM.

- Bardzo bogaty jest także przegląd literatury dotyczący zastosowania metod analizy ruchu w sporcie. W ostatnich latach systemy wizyjne, służące do monitorowania zawodów sportowych i powtarzania ujęć oraz akcji pojawiają się w kolejnych dyscyplinach sportowych. Jednak rozprawa dotyczy zagadnień daleko wykraczających poza te technologie. Jej celem jest automatyczna analiza ruchu umożliwiająca klasyfikację akcji i ocenę poprawności wykonania akcji przez sportowca. Przegląd literatury na temat zastosowań w sporcie skupia się niejako na trzech obszarach. Po pierwsze, omawiane są techniki analizy ruchu proponowane dla różnych dyscyplin sportowych. Po drugie, podane są wybrane i ogólnie dostępne zbiory danych o ruchu człowieka wykonywanego w ramach różnych życiowych aktywności oraz takie zbiory, które są specyficzne dla akcji sportowca w różnych dyscyplinach sportowych. Po trzecie, omawiane są dotychczasowe rozwiązania dotyczące analizy ruchu na potrzeby szermierki.

Niewątpliwie tak obszerny przegląd literatury światowej w zakresie stanu wiedzy o metodach automatycznej analizy ruchu i ich praktycznych zastosowaniach świadczący o dużej wiedzy Autora co pozwoli mu zrealizować bardzo dojrzały proces projektowania i weryfikacji własnych rozwiązań mających na celu dla wykazanie zasadności zadanej w rozprawie tezy badawczej.

Mankamentem jest jednak brak podsumowania większości cząstkowych przeglądów literatury i wskazania wniosków dla dalszego postępowania Autora, wynikających z przeglądów źródeł.

Należy zauważyć, że Autor odwołuje się w swojej rozprawie do 13 pozycji, w których jest jedynym lub pierwszym autorem, w tym do dwóch artykułów opublikowanych lub procedowanych w renomowanych czasopismach. Ta obserwacja a także wykonana w dalszej części recenzji analiza całego dorobku publikacyjnego Autora, pozwalają stwierdzić, iż dorobek ten jest bardzo znaczący i całkowicie spójny z tematyką rozprawy.

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Cele szczegółowe określono jako rozwiązanie trzech zagadnień:

1. Rozpoznawanie dynamiki akcji w pracy nóg w szermierce;
2. Detekcja oraz analiza akcji w pracy nóg w szermierce w czasie rzeczywistym;
3. Intuicyjny kanał zwrotny dla ćwiczeń z bronią w szermierce z użyciem rozszerzonej rzeczywistości.

W ich ramach w rozprawie zaproponowano, zaimplementowano i zweryfikowano eksperymentalnie szereg metod, które w efekcie umożliwiają wspomaganie doskonalenia umiejętności sportowych.

Zagadnienie 1: rozpoznawanie dynamiki akcji w pracy nóg w szermierce

W rozdziale 3. zaproponowano i zweryfikowano metody analizy dynamiki ruchu, które umożliwiają efektywną klasyfikację różnych typów wypadów w szermierce, na podstawie multimodalnych danych - map głębi i sygnałów inercyjnych.

Metody zweryfikowano na zarejestrowanym przez Autora zbiorze danych obrazującym pracę nóg szermierza, nazwanym *Fencing Footwork Dataset* (FFD), jak również dla bardziej ogólnych akcji z ogólnie dostępnego zbioru UTD-MHAD.

Na podstawie analizy prac prof. Czajkowskiego i w wyniku konsultacji z trenerami szermierki wybrano 6 podstawowych akcji: krok w przód, krok w tył oraz cztery typy wypadów - zrywowy, z narastającą szybkością, z wyczekiwaniem i skoczno-ślizgowy.

Autor zaproponował kilka sposobów ekstrakcji informacji o ruchu, które dają silnie ze sobą skorelowane zestawy cech i w zasadzie klasyfikacja mogłaby być oparta osobno na każdym zestawie. Ostatecznie tworzone są cztery zestawy cech.

1. Cechy uzyskane z sygnału przyspieszenia (ang. *Accelerometric (Acc)*) - po 24 cechy w dziedzinie czasu (typu wartość średnia i rms) dla 3 sygnałów wtórnych podzielonych na 7 okien czasowych (łącznie 168 cech dla bazy FFD).
- 2-3. Dwa zestawy cech wyznaczane są na podstawie trajektorii danych szkieletowych (po 64 punkty każda) pochodzących z Kinect-a, przy czym korzysta się jedynie z 8 złączy dla dolnych partii ciała:
2. Jeden zestaw reprezentuje „dynamikę złączy” (ang. *Joint Dynamics (JD)*) - po wyznaczeniu z trajektorii 2 sygnałów prędkości i 2 sygnałów przyspieszeń w płaszczyźnie x-y, podzieleniu sygnałów na 7 okien o 3 rozmiarach i obliczeniu 3 pierwszych współczynników amplitudowych Fouriera dla każdego okna powstaje łącznie wektor o 1056 elementach;
3. Kolejny zestaw stanowią „lokalne obrazy śladów” (ang. *Local Trace Images (LTI)*) – dodawane są do siebie mapy położenia złączy w kolejnych chwilach czasu (przy czym każde położenie jest rozmywane funkcją Gaussa dla właściwego wyrażenia ruchu złącza) – po normalizacji rozmiaru mapy ruchu dla każdego złącza do 16 x 16 i konkatencji 8 map powstaje obraz o 2048 elementach.
4. Cechy oparte o dane szkieletowe o złączach (stawach sylwetki człowieka) i binarne obrazy (maski) zmian sylwetki uzyskiwane z map głębi; dla każdej maski i każdego złącza tworzony jest (znormalizowany do sumy 1) histogram 2D wyrażający odległości „pikseli zmienionych” w otoczeniu każdego złącza we współrzędnych biegunowych; po wyznaczeniu trzech zestawów histogramów dla trzech kolejnych ramek następujących po ramce referencyjnej tworzona jest ważona suma histogramów każdego złącza; po konkatencji histogramów dla każdej ramki referencyjnej powstaje wektor cech nazywany „kontekstem historii ruchu stawów” (ang. *Joint Motion History Context (JMHC)*) o 480 elementach.

Niewątpliwie każdy zestaw cech dla reprezentacji ruchu posiada swoje uzasadnienie i oparty jest na wiedzy eksperckiej z zakresu analizy sekwencji obrazów i szeregów czasowych. Ewentualnie można oczekiwać głębszego uzasadnienia dla zestawu 2 i 3 obok podanego dla zestawu 1, że próbowano tworzyć alternatywne zestawy cech (w dziedzinie transformaty Fouriera czy falkowej dla danych inercyjnych), ale w wyniku wstępnych testów wybrano te właściwe.

Następnie Autor przechodzi do etapu stosowania technik analizy („drażenia”) danych i klasyfikacji cech. Każdy zestaw cech jest z osobna poddawany redukcji własnym algorytmem Autora opartym o znane metody selekcji cech (AdaBoost, Lasso, PCA) i kryterium dekorelacji. Dla każdej pary klas realizowane jest uporządkowanie cech według ich siły dyskryminacyjnej a następnie do najlepszych cech dobierane są zdekorowane z nimi cechy.

Na koniec badane są dwa podejścia do fuzji zestawów cech i klasyfikacji cech. W pierwszym z nich – fuzja odbywa się na etapie tworzenia wektora cech i klasyfikacja jest jedno-etapowa. Wyselekcjonowane z zestawów podzbiory cech łączone są w jeden wektor cech i podlegają one klasyfikacji z użyciem maszyn wektorów nośnych (ang. *Support Vector Machine*, SVM). W drugim podejściu fuzja dotyczy wyników odrębnych klasyfikacji i proces decyzyjny jest dwuetapowy. Wyselekcjonowany podzbiór cech z każdego zestawu jest pre-klasyfikowany z użyciem maszyn wektorów nośnych (ang. *Support Vector Machine* (SVM)). Uzyskane prawdopodobieństwa klas są łączone we wspólny wektor, który jest klasyfikowany przez perceptron wielowarstwowy (ang. *Multilayer Perceptron* (MLP)).

W rozprawie zawarto wyniki zbiorcze wielu testów porównawczych wykonanych przez Autora. Najpierw, w oparciu o swoją bazę danych, porównał on skuteczność odrębnej klasyfikacji poszczególnych 4 zestawów cech bez ich selekcji czy fuzji, stosując dwa scenariusze – z modelem zależnym od sprawdzanego użytkownika (PD) lub niezależnym (PI) od niego. W pierwszym przypadku możliwe było uzyskanie skuteczności klasyfikacji powyżej 90% w „łatwiejszym” scenariuszu PD ale jedynie poniżej 80% w wariancie PI. Następnie Autor porównał jakość swojej metody selekcji cech z metodami PCA, AdaBoost i Lasso oraz z sytuacją braku selekcji cech. Wykazał on wyższą skuteczność swojej metody nad resztą dla 3 zestawów z 4 – w scenariuszu PD i 2 z 4 w scenariuszu PI. W dalszej kolejności Autor podał wyniki porównania skuteczności klasyfikacji w oparciu o powyższe różne metody selekcji cech po wykonaniu fuzji cech dwoma sposobami – fuzji na poziomie cech lub na poziomie wyników pre-klasyfikacji. Można zauważyć, że fuzja na poziomie cech znacząco poprawia skuteczność klasyfikacji (z ok. 94% do ok. 98%) ale fuzja na poziomie pre-klasyfikacji nieznacznie obniża skuteczność (o ok. 2 %) w scenariuszu PD. Z kolei dla scenariusza PI dla obu typów fuzji skuteczność się poprawia (z ok. 80 do ok. 83 lub ok. 86%) - nawet jest wyższa dla przypadku fuzji na poziomie wyników pre-klasyfikacji. Na koniec testów w oparciu o autorski zestaw danych FFD, Autor zamieścił wyniki porównania skuteczności najlepszej wersji całego swojego rozwiązania z kilkoma znanymi rozwiązaniami do rozpoznawania sekwencji cech o uniwersalnym charakterze. Uzyskane wyniki Autora są znacząco lepsze od wyników dla innych metod. W drugiej części eksperymentów stosowano ogólnodostępną bazę akcji UTD-MHAD klasyfikacji sekwencji czasowych. Dla tego zbioru wykorzystano wszystkie stawy szkieletu w trzech ortogonalnych płaszczyznach, zwiększając odpowiednio rozmiar wektorów cech. Nie podano tego jawnie, ale należy założyć, że testowano jedynie scenariusz PI. Również w tym przypadku autorska metoda selekcji cech konkuruje skutecznie z innymi testowanymi metodami (PCA, AdaBoost, Lasso) a po wykonaniu fuzji cech na poziomie pre-klasyfikacji, przewyższa ona jakością wszystkie pozostałe metody. Inaczej niż dla zbioru FFD, pre-klasyfikacja za pomocą SVM z jądrem RBF prowadzi do wyższej skuteczności niż przy użyciu liniowej SVM. Na koniec Autor w oparciu o bazę UTD-MHAD porównał całościowe rozwiązanie z szeregiem rozwiązań znanych z literatury wykazując, że jego rozwiązanie jest jedynie nieznacznie gorsze od metody VGG-16, które dodatkowo korzysta z obrazu RGB, ale nie posiada informacji z czujników inercyjnych.

Zagadnienie 2. Detekcja i analiza akcji w pracy nóg w szermierce w czasie rzeczywistym

W rozdziale 4. zaprezentowano i przebadano własną metodę adaptacyjnego filtrowania sygnału próbującego wartość prędkości zmierzoną dla szermierza w kolejnych chwilach czasu w oparciu o dane z Kinect-a lub z przetworzonego sygnału pochodzącego z akcelerometru. W porównaniu do uniwersalnych filtrów uśredniających (filtra średniej kroczącej lub typu „loess”)

Autor skorzystał ze znanego modelu ruchu szermierza, co umożliwiło bardziej niezawodne wykrywanie ekstremalnych wartości w sygnale odpowiadających akcjom wypadów lub sekwencji ruchu wprzód – w tył. Niestety opis następującego po filtracji algorytmu segmentacji sygnałów, czyli wiedza o tym, które ekstremum sygnału jakiego złącza jest wykrywane, jest enigmatyczny i trudno zrozumiały. Bardzo pomocne byłoby zastosowanie precyzyjnej notacji, w tym pseudokodu, dla jego przedstawienia.

Dla sygnałów prędkości złącz szkieletu pozyskiwanego z obrazu Kinect-a odpowiadających poszczególnych czasom trwania segmentów wyznaczone zostają cechy (znacznie uproszczone wobec poprzednio badanych) – 120 współczynników Fouriera, po 3 dla każdego złącza i składowej ruchu x lub y (dla 20 złączy) Cechy są ostatecznie klasyfikowane za pomocą liniowego SVM. Jednocześnie analizowane są wybrane własności wyodrębnionych segmentów dla wykrytego typu akcji takie, jak: szybkość wyrzutu ręki, czas trwania wypadu, długość wypadku, średnie i maksymalne przyspieszenie, średnia i maksymalna prędkość.

Wykryte własności (7 liczb) wysyłane są zwrótnie do smartfonu szermierza poprzez sieć bezprzewodową. Przez czas rzeczywisty całego procesu Autor rozumie ograniczenie opóźnienia otrzymania przez szermierza wyniku analizy względem czasu wykonania (zakończenia) jego akcji. Czas ten szacowany jest na 300 ms i w zasadzie zdeterminowany jest rozmiarem filtra uśredniającego sygnał o prędkości złącza. Samo przetwarzanie danych na komputerze wymaga mniej niż 1 ms czasu.

Na potrzeby opracowania i weryfikacji metod segmentacji czasowej oraz analizy jakościowej akcji nagrano dwa zbiory danych obrazujące pracę nóg szermierza. Pierwszy zbiór został pozyskany przy użyciu sensora Kinect, który dostarczał danych głębi i szkieletowych.

Drugi zbiór danych obejmuje zarówno dane z sensora Kinect, jak i z własnoręcznie zbudowanego systemu złożonego z dwóch sensorów inercyjnych i mikrokontrolera. Sensory były zamontowane na klatce piersiowej oraz na łokciu zawodnika. Analiza sygnału uzyskanego z sensora zamontowanego przy łokciu umożliwia wykrycie momentu wyprostowania ręki. Analiza przebiegu sygnału przyspieszenia z sensora zamontowanego na klatce piersiowej pozwala z kolei na znalezienie początku i końca ruchu wypadu.

Zagadnienie 3. Intuicyjny kanał zwrotny dla ćwiczeń z bronią w szermierce z użyciem rozszerzonej rzeczywistości.

W rozdziale 5. zaproponowano oraz przebadano metody do śledzenia ruchu broni, uczenia modelu poprawnych akcji broni oraz ewaluacji ćwiczeń z bronią. Wykorzystano dwa markery (świecące LED-y) umieszczone na końcu ostrza do łatwej detekcji pozycji ostrza w układzie współrzędnych kamery umieszczonej w półprzezroczystych okularach wirtualnej rzeczywistości zakładanych przez szermierza.

Dzięki kalibracji kamery i łatwej detekcji markerów w obrazie możliwy jest pomiar głębi położenia markerów i kąta nachylenia ostrza oraz odwzorowanie współrzędnych rzeczywistych na współrzędne wirtualne. Należy zauważyć, że algorytm śledzenia markerów został opisany jedynie w postaci tekstowej i wydaje się być prostą detekcją położenia markerów w kolejnych chwilach czasu. Na okularach szermierza wyświetlane są modelowe oraz aktualne trajektorie ruchu ostrza.

Główne cele rozprawy, rozumiane jako rozwiązanie podanych wyżej trzech zagadnień, zostały osiągnięte a postawiona teza została pozytywnie zweryfikowana.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Zasadnicza oryginalność rozprawy polega na unikalnej dziedzinie zastosowania metod analizy ruchu w multi-modalnych sekwencjach danych (obrazach z Kinect-a i sygnałach z czujników inercyjnych). Proponowane rozwiązanie, opisane tezą pracy, stanowi samodzielny i oryginalny dorobek Autora. Dzięki zastosowaniu nowoczesnych technik akwizycji danych i prezentacji wyników, propozycji własnych algorytmów dla cząstkowych rozwiązań, praca reprezentuje wysoki poziom w skali międzynarodowej i wnosi istotny wkład do literatury światowej dotyczącej jej tematyki – czyli metod automatycznej analizy ruchu w zastosowaniu do wspomaganie treningu sportowca.

Do najważniejszych osiągnięć rozprawy zaliczam nowo proponowane metody analizy sekwencji multi-modalnych danych. Są nimi

- autorska metoda klasyfikacji zbliżonych do siebie akcji sportowych w oparciu o dynamikę ruchu;
- nowe cechy charakteryzujące obrazy i sygnały inercyjne, powstałe poprzez nowe metody selekcji oraz fuzji deskryptorów;
- autorska metoda filtracji sygnałów prędkości uzyskiwanych czujników inercyjnych i obrazów szkieletowych z Kinect-a oraz jej zastosowanie do detekcji akcji wypadów pracy nóg szermierza;
- analiza jakościowa akcji wypadu poprzez wyznaczanie własności dla wykrytych segmentów czasowych trwania akcji takich jak: czas trwania akcji, długość wypadu, szybkość ruchu ręki.

Podkreślić należy także duży zakres stosowanych metod i mnogość eksperymentów pokazujących skuteczność proponowanych rozwiązań i ich przewagi wobec wcześniejszych metod znanych z literatury. Z uznaniem należy odnotować zebranie dużego zbioru multi-modalnych danych obrazujących pracę nóg szermierzy i jego udostępnienie w Internecie.

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników /zwięzłość, jasność, poprawność redakcyjna rozprawy/?

Rozprawa jest napisana poprawnym językiem angielskim i jej treść jest w pełni zrozumiała. Struktura pracy właściwie oddaje wymagania wobec rozpraw doktorskich i właściwie grupuje w rozdziały i podrozdziały wykonane przez Doktoranta badania oraz osiągnięte wyniki. Treść podawana jest w zwięzły sposób, ale jest ona nadmiernie zwięzła, gdyż w wielu miejscach mogłaby być wzbogacona o autorskie podsumowanie, refleksje lub szczegóły stosowanych algorytmów. Brakuje mi podsumowania przeglądu literatury i wniosków, oraz szczegółowych opisów proponowanych i zrealizowanych algorytmów.

Nie mam zastrzeżeń redakcyjnych do ostatecznej wersji rozprawy. Większość zagadnień jest prezentowana w poglądowy sposób i ilustrowana graficznie. Omawiane zagadnienia są przedstawione poprawnie pod względem merytorycznym.

6. Jakie są słabe strony rozprawy i jej główne wady?

Rozprawa stanowi propozycję innowacyjnego rozwiązania opracowanego dla rozwiązania trzech zasadniczych zagadnień występujących w specyficznym obszarze zastosowania. Praca ma charakter eksperymentalny - dla weryfikacji rozwiązania korzysta z dostępnych technologii pozyskiwania danych o ruchu człowieka i obiektu oraz tworzenia wirtualnej rzeczywistości. Samo rozwiązanie polega na opracowaniu i zastosowaniu nowych algorytmów, zmodyfikowaniu istniejących bądź zastosowaniu uznanych technik analizy obrazów, sekwencji obrazów i szeregów czasowych. Biorąc po uwagę powyższy charakter pracy, nie dostrzegam żadnych zasadniczych wad rozprawy.

Nieliczne słabe strony rozprawy postrzegam w nadmiernej zwięzłości prezentacji (jedynie w postaci tekstowej) kilku algorytmów, a także w sformułowaniu jednej, dość ogólnej tezy. Można też sformułować dyskusyjne w swoim charakterze uwagi i pytania dotyczące braku bardziej teoretycznego wyjaśnienia przyczyny wyboru proponowanych algorytmów.

1. Rozprawa doktorska (nie mająca charakteru doktoratu wdrożeniowego) powinna wносить ważny wkład w rozwój wiedzy i siłą rzeczy sięgać bardzo głęboko w specjalistyczny obszar metod czy zastosowań. Dlatego w tezie rozprawy (str. 2) oczekiwałbym przynajmniej wskazania charakteru zastosowanych lub nowo proponowane metodyk analizy strumieni danych, charakteru analizowanych danych (sygnały z czujników inercyjnych, sekwencje obrazów z Kinect-a, ale bez obrazu RGB) i określenia stopnia uniwersalności bądź specjalizacji proponowanych technik dla wybranej dyscypliny sportowej.
2. Można się domyślać, że przez „klasyfikator SVM” (str. 43, 47-48) Autor rozumie „wieloklasowe SVM”, które jest zespołem binarnych klasyfikatorów SVM. Jest kilka sposobów agregacji binarnych klasyfikatorów w jeden klasyfikator wieloklasowy. Pomimo zastosowania gotowej implementacji z pakietu Weka warto było wyjaśnić przyjętą zasadę wieloklasowego SVM.
3. Czy wobec faktu, że metoda VGG-16 dla bazy danych UTD-MHAD wykazała lepszą skuteczność klasyfikacji akcji aniżeli metoda Autora, dla uzyskania większej uniwersalności rozwiązania, nie należałoby uzupełnić zestaw cech o informację pozyskaną z obrazu kolorowego, który to obraz jest dostępny w Kinect-cie?
4. Nie negując potrzeby zastosowanych w pracy klasycznych technik analizy obrazów i sygnałów powstaje pytanie: czy w czasach burzliwego rozwoju stosowania sieci neuronowych w analizie obrazów, w tym nowych typów sieci takich jak CNN, R-CNN czy LSTM, które szybko zastępują tradycyjne techniki segmentacji i pozyskiwania cech, lokalizacji obiektów i ich klasyfikacji, a także klasyfikacji sekwencji wzorców w czasie, czy jest jeszcze racja bytu dla klasycznych technik „drażenia danych” – wyznaczania cech, selekcji i fuzji cech? Jakie wnioski Autor wyciąga z przeglądu literatury dotyczącej sztucznych sieci neuronowych?

7. Jaka jest przydatność rozprawy dla nauk technicznych?

Opracowane metody służą do automatycznej analizy ruchu sportowca i innych obiektów w sporcie co może stanowić istotną pomoc w procesie doskonalenia technik sportowca, nawet bez bezpośredniego nadzoru trenera. Na podkreślenie zasługuje fakt, że w poszukiwaniu

praktycznego zastosowania swojego rozwiązania Autor nawiązał współpracę ze szkołami szermierki w Polsce (szkoła Aramis) i USA („Delta Fencing Center”).

Tematyka rozprawy wpisuje się doskonale w aktualne światowe trendy badawcze dotyczące rozwoju metod i technik pozwalających na wykorzystanie gwałtownego rozwoju technologii sensorowej, w tym czujników obrazowania wbudowanych w kamery cyfrowe i czujników elektronicznych zakładanych jako element ubrania (tzw. czujniki nasobne).

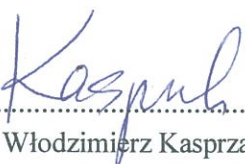
Podjęmowane na dużą skalę prace badawczo-rozwojowe w tym obszarze stają się motorem napędowym szeregu zjawisk społecznych. Pozyskiwane duże zbiory behawioralnych, psychofizycznych danych i opracowywane algorytmy pozwalają na optymalizację życia i działalności człowieka w różnych jego przejawach. Przykłady zastosowań to nie tylko optymalizowanie treningu i technik w sporcie, ale także wspomaganie służby ochrony zdrowia, w tym medycyny sportowej, i przeciwdziałanie wykluczeniu społecznemu wynikającemu ze starzenia się i braku pełnej sprawności człowieka.

Aktualność tematyki badawczej, szeroki zakres stosowanych metod i technik (analiza danych z sensorów, rozpoznawanie dynamicznych wzorców i synteza wirtualnej rzeczywistości) oraz weryfikacja na rzeczywistych danych w przyjętym scenariuszu zastosowania, pozwalają uznać, że jest to bardzo wartościowa praca dla rozwoju nauk technicznych i jej zastosowań.

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

Praca z nadmiarem spełnia wszelkie wymagania stawiane rozprawom doktorskim w dziedzinie nauk technicznych. Autor stosuje poprawną metodykę badawczą, weryfikuje swoje rozwiązanie w sposób eksperymentalny i porównuje różne techniki realizacji stosując obiektywne kryteria. Praca wyróżnia się unikalną tematyką i innowacyjnym rozwiązaniem w skali międzynarodowej. Tematycznie powiązany z rozprawą dorobek publikacyjny Autora jest rozpoznawalny w skali międzynarodowej. Mgr inż. Filip Malawski posiada znaczący dorobek publikacyjny, obejmujący 21 istotnych publikacji, w tym 5 artykułów w czasopismach z listy A (w tym jedna po recenzji), 8 artykułów w czasopismach z listy B MNiSW oraz 8 publikacji konferencyjnych. Pod względem liczby i jakości publikacji ten dorobek znacznie przewyższa oczekiwania w przewodach doktorskich.

Rozprawę zaliczam do najwyższej kategorii „wybitnie dobra, zasługująca na wyróżnienie”.


.....
/ Włodzimierz Kasprzak /