

AKADEMIA GÓRNICZO-HUTNICZA IM. STANISŁAWA STASZICA



Wydział Informatyki, Elektroniki i Telekomunikacji

AUTOMATIC ANALYSIS OF TECHNIQUES AND BODY MOTION
PATTERNS IN SPORT

AUTOMATYCZNA ANALIZA TECHNIK ORAZ FORM RUCHU CIAŁA
W SPORCIE

Autoreferat rozprawy doktorskiej

Filip Malawski

Promotor:
dr hab. inż. Bogdan Kwolek, prof. AGH

Kraków, 2019

1. Wprowadzenie

Technologia odgrywa istotną rolę we współczesnym sporcie. Elektroniczne pomiary czasu w biegach sprinterskich, weryfikacja uderzenia piłki w boisko na podstawie nagrań wideo w siatkówce, czy wykorzystanie pozycjonowania GPS do śledzenia trasy kolarzy są obecnie powszechnie stosowane w praktyce. Ponadto, w ostatnim dziesięcioleciu rozwinięto bardziej zaawansowane, nowoczesne systemy do analizy ruchu w sporcie, które pozwalają na wykonywanie dokładnych pomiarów parametrów ruchu oraz dostarczanie istotnych informacji zwrotnych zarówno dla sportowców jak i trenerów [47].

Analiza ruchu w sporcie różni się zasadniczo od ogólnej analizy ruchu w kilku istotnych aspektach. Po pierwsze, akcje sportowców są bardzo specyficzne i często nie mają swoich odpowiedników w codziennych czynnościach, co utrudnia wykorzystanie ogólnych metod analizy ruchu. Po drugie, charakteryzują się one większą szybkością i dynamiką wykonania. Ponadto, w sporcie istotne jest nie tylko rozpoznawanie akcji, ale też ich analiza jakościowa, która jest konieczna, aby dostarczyć sportowcom użyteczne informacje zwrotne. Zasadniczą cechą akcji sportowych jest również zróżnicowanie pomiędzy dyscyplinami sportowymi, czego efektem jest potrzeba adaptacji metod analizy ruchu do każdej dyscypliny z osobna. W związku ze wspomnianymi aspektami, automatyczna analiza ruchu w sporcie stanowi znaczne wyzwanie. Motywacją dla badań przeprowadzonych w tej rozprawie był brak dostępnych w literaturze rozwiązań ukierunkowanych na powyższe zagadnienia.

1.1 Teza rozprawy

Celem niniejszej pracy było opracowanie metod pozwalających na automatyczną analizę ruchu w sporcie, które mogłyby wspomagać zawodników oraz trenerów. Na tej podstawie postawiono następującą tezę:

Automatyczna analiza techniki oraz ruchu ciała sportowca generuje kanał zwrotny, użyteczny na potrzeby doskonalenia umiejętności sportowych

Rozprawa dotyczy wybranych problemów związanych z automatyczną analizą ruchu w sporcie. Po pierwsze, poruszane jest zagadnienie rozpoznawania akcji specyficznych dla sportu, które mają podobne trajektorie ruchu, ale inną dynamikę wykonania, co przekłada się na różne techniki sportowe. Po drugie, omawiane są zagadnienia związane z segmentacją czasową ciągłego ruchu oraz analizą jakościową wyodrębnionych akcji. Parametry jakościowe wykonywanego ruchu stanowią istotną informację zwrotną dla sportowców. Po trzecie, prezentowane są analizy i rozwiązania w zakresie metod przekazywania informacji zwrotnej w sposób intuicyjny i jednocześnie w czasie rzeczywistym. Wynikiem końcowym są autorskie metody przebadane w rzeczywistych scenariuszach dla każdego z powyższych zagadnień badawczych.

1.2 Osiągnięcia rozprawy

Metody opracowane w niniejszej pracy zostały zweryfikowane na przykładzie szermierki, która jest bardzo technicznym sportem. Do najważniejszych osiągnięć rozprawy można zaliczyć:

- wyczerpujący przegląd literatury w zakresie metod rozpoznawania ruchu, w szczególności ruchu w sporcie, w oparciu o różne modalności sygnałów
- nowe metody klasyfikacji podobnych akcji sportowych w oparciu o dynamikę ruchu, w tym nowe deskryptory, oparte o wizyjne oraz inercyjne dane, jak również nowe metody selekcji oraz fuzji cech
- dedykowany, multimodalny zbiór danych obrazujący pracę nóg szermierzy, udostępniony w Internecie
- ewaluacja opracowanych metod klasyfikacji na dwóch zbiorach danych: własnym oraz ogólnie dostępnym zbiorze UTD-MHAD, na obu zbiorach uzyskano wyniki lepsze niż metody prezentowane w literaturze
- nowe metody filtracji sygnałów inercyjnych oraz szkieletowych, pozwalające na detekcję akcji wypadków w ciągłej pracy nóg szermierza, jak również metody analizy jakościowej wyodrębnionych akcji
- implementacja systemu do detekcji i analizy akcji w pracy nóg szermierza w czasie rzeczywistym
- ewaluacja opracowanego systemu na danych z dwóch dedykowanych zbiorów danych
- nowe metody umożliwiające śledzenie broni, uczenie modelu ruchu oraz weryfikację poprawności akcji wykonywanych bronią, a także wykorzystanie rozszerzonej rzeczywistości jako intuicyjnego kanału dostarczania informacji zwrotnej
- implementacja systemu opartego o rozszerzoną rzeczywistość wspomagającego ćwiczenia z bronią w szermierce
- ewaluacja opracowanych metod w kontekście wspomagania ćwiczeń z bronią w szermierce

2. Obszar badań oraz przegląd literatury

Celem rozwoju metod automatycznej analizy ruchu w sporcie jest wspomaganie procesu doskonalenia umiejętności sportowych. Umiejętności posiadają dwie składowe - sprawność oraz technikę [57]. Cechy sprawnościowe, takie jak siła, elastyczność czy koordynacja definiują maksymalny zakres oraz dokładność wykonywanego ruchu. Technika jest natomiast cechą znaną wykonywania akcji motorycznej. Definiuje jak powinna być wykonana akcja sportowa, np. przyjęcie piłki w siatkówce lub ruch zasłony bronią w szermierce. Analiza poprawności wykonywanych technik jest więc podstawą dla procesu wspomagania doskonalenia umiejętności. Zadaniem tym zajmuje się trener, jednakże trenerzy muszą dzielić uwagę pomiędzy wielu zawodników. Celem badań w niniejszej pracy było opracowanie metod analizy ruchu, które zapewnią użyteczne informacje zwrotne dla ćwiczeń wykonywanych bez udziału trenera.

2.1 Automatyczna analiza ruchu człowieka

Metody pozwalające na automatyczną analizę ruchu człowieka są rozwijane od wielu lat, ponieważ znajdują one zastosowanie w wielu dziedzinach, takich jak interfejsy człowiek-komputer [45], rozpoznawanie chodu [30], rozpoznawanie upadku [31] czy rozpoznawanie języka miganego [28]. Do analizy ruchu wykorzystywane są różnorodne modalności

danych. Najczęściej są to zarejestrowane sekwencje obrazów z kamer RGB, mapy głębi uzyskiwane za pomocą sensorów głębi takich jak Kinect [55], oraz przyspieszenie, prędkość kątowna i wartość pola magnetycznego, odczytywane przez sensory inercyjne. Proponowane w literaturze metody są ukierunkowane na szereg zagadnień, takich jak wariacje w wykonywaniu ruchu przez różne osoby, wpływ zmian oświetlenia oraz przesłaniania obiektów w nagraniach z kamer RGB, ograniczony zasięg sensorów głębi, czy też akumulacja błędów w odczytach z sensorów inercyjnych [60].

Przegląd prac w zakresie analizy ruchu na podstawie obrazów RGB można znaleźć w pracach [10, 64]. W tym obszarze można wyróżnić kilka podejść. Pierwsze z nich to metody oparte o detekcję osoby i analizę sylwetki, opcjonalnie z segmentacją poszczególnych części ciała. Obejmuje to m.in. obrazy energii i historii ruchu [9] oraz histogramy zorientowanych gradientów [19]. W drugim podejściu znajdują się metody pozwalające na detekcję przestrzenno-czasowych punktów kluczowych [32] oraz ich opis za pomocą dedykowanych deskryptorów [34]. Lokalne deskryptory mogą być również wyznaczane w siatce punktów, co wykorzystywane jest z powodzeniem w metodzie lokalnych wzorców binarnych [49] oraz w metodzie przepływu optycznego [7]. Dużą popularnością cieszą się ostatnio głębokie sieci neuronowe [26], które zapewniają wysoką skuteczność rozpoznawania nawet przy bardzo dużej liczbie klas [29]. Metody analizy ruchu w oparciu o dane z sensorów głębi stały się popularne głównie dzięki powszechnej dostępności sensora Kinect [16]. W tym obszarze możemy wyróżnić metody wykorzystujące mapy głębi [50], oraz metody wykorzystujące wyodrębniony automatycznie szkielet postaci [52]. Sensory inercyjne pozwalają na uzyskanie danych o przyspieszeniu, prędkości kątowej, polu magnetycznym, a pośrednio również orientacji [3]. Opis ruchu na podstawie takich danych polega zazwyczaj na określeniu cech statystycznych w dziedzinie czasu lub cech w dziedzinie częstotliwości przez wyznaczenie szybkiej transformaty Fouriera [6]. Łączenie wielu sensorów jest również popularnym podejściem, zarówno w przypadku sensorów tego samego typu [56], jak i takich, które dostarczają dane różnych modalności [37].

2.2 Automatyczna analiza ruchu w sporcie

Analiza ruchu w sporcie różni się od ogólnej analizy ruchu. Akcje sportowe charakteryzują się specyficznymi formami ruchu, w których mamy do czynienia ze zmienną dynamiką wykonania. Segmentacja czasowa ruchu oraz analiza jakościowa wykonywanych akcji są podstawą uzyskania użytecznych dla sportowców informacji. Sposób przekazywania tych informacji ma istotny wpływ na przydatność systemów wspomagających ćwiczenia. Akcje sportowe są również mocno zróżnicowane pomiędzy dyscyplinami sportowymi. W związku z tym analiza ruchu w sporcie wiąże się z rozwojem specyficznych metod, często ukierunkowanych na konkretną dyscyplinę.

Wiele prac w literaturze zorientowanych jest na sporty zespołowe. Poruszane zagadnienia dotyczą wykrywania obszaru należącego do boiska [43], śledzenia graczy [46] oraz identyfikacji graczy [22]. W dyscyplinach wykorzystujących piłkę istotnym problemem jest jej śledzenie [15]. W sportach indywidualnych metody są dużo bardziej zróżnicowane i specyficzne dla danej dyscypliny. W tenisie ziemnym śledzona jest piłka oraz gracze [53], w golfie analizowany jest ruch kija [63], w karate klasyfikowane są wykonane techniki [24]. Można również znaleźć prace dotyczące klasyfikacji uderzeń w boksie [27], czy też analizy cykliczności ruchu w pływaniu [51]. Niniejszy przegląd pokazuje jak bardzo zróżnicowanym oraz złożonym problemem jest analiza ruchu w sporcie. Warto również wspomnieć, że o ile dla ogólnych akcji istnieje wiele ogólnie dostępnych baz da-

nych [62], o tyle w odniesieniu do sportu ich liczba jest bardzo ograniczona i najczęściej dotyczą one problemu klasyfikacji dyscyplin, a nie analizy technik, a w szczególności ruchu sportowca [48, 54].

Na podstawie dogłębnego przeglądu literatury zidentyfikowano istotne, nierozwiązane dotychczas problemy związane z analizą ruchu w sporcie: rozpoznawanie podobnych akcji w oparciu o analizę dynamiki ruchu, segmentacja czasowa ciągłego ruchu, analiza jakościowa wykrytych akcji oraz dostarczanie informacji zwrotnej w czasie rzeczywistym. Na potrzeby przeprowadzonych badań wybrano dyscyplinę sportową, w której wszystkie te zagadnienia pełnią istotną rolę, to znaczy szermierkę. Dwa podstawowe elementy składowe szermierki to praca nóg oraz działania z bronią. Klasyfikacja wypadów w szermierczej pracy nóg wymaga analizy dynamiki wykonania. Detekcja wypadów w ciągłej pracy nóg jest podstawą do określenia ich parametrów jakościowych, takich jak szybkość, czy też długość. Zagadnienie dostarczania informacji zwrotnej w czasie rzeczywistym jest szczególnie istotne w przypadku ćwiczeń z bronią, które wymagają dużej precyzji i powtarzalności ruchu.

Prace dotyczące analizy ruchu szermierza oraz szeroko rozumianej szermierki są stosunkowo ograniczone. Analiza wypadów została przeprowadzona z wykorzystaniem stereofotogrametrii [23] oraz elektromiografii [11]. Porównanie dynamiki wypadu oraz rzutu szermierczego przedstawiono w pracy [8]. Klasyfikacja akcji broni została wykonana z użyciem wielokamerowego systemu przechwytywania ruchu [44]. Rozpoznawanie poprawnych i niepoprawnych wykonania akcji bronią na podstawie sygnałów inercyjnych opisano w pracy [12]. Przebadano wpływ różnych rękoności na mięśnie nadgarstka [33]. Została również wykazana większa dynamiczna kontrola równowagi wśród szermierzy niż wśród sportowców ćwiczących sporty statyczne [25]. Pomimo zróżnicowanych podejść, przedstawione prace nie odnoszą się do wspomnianych wyżej zagadnień.

Niniejsza rozprawa podzielona została na trzy główne części, dotyczące zagadnień z zakresu analizy ruchu w sporcie. W Rozdziale 3. zaproponowano oraz przebadano metody, które umożliwiają bardziej efektywną klasyfikację różnych typów wypadów w szermierce, na podstawie danych głębi oraz inercyjnych. W Rozdziale 4. zaprezentowano opracowane i przebadane metody adaptacyjnego filtrowania w oparciu o model ruchu, które umożliwiają bardziej niezawodne wykrywanie akcji wypadów w ciągłej pracy nóg szermierza. Parametry jakościowe wykrytych akcji są wyznaczane oraz przesyłane do ćwiczącego szermierza w czasie rzeczywistym. W Rozdziale 5. zaproponowano oraz przebadano metody do śledzenia ruchu broni, uczenia modelu poprawnych akcji broni oraz ewaluacji ćwiczeń z bronią. Ponadto zastosowano okulary do rozszerzonej rzeczywistości, które tworzą mieszany widok świata wirtualnego i rzeczywistego, w którym wyświetlane są modelowe oraz aktualne trajektorie ruchu. Zapewnia to intuicyjny oraz innowacyjny kanał dostarczania informacji zwrotnej w czasie rzeczywistym.

3. Rozpoznawanie dynamiki akcji w pracy nóg w szermierce

Rozpoznawanie akcji jest jednym z najpopularniejszych zagadnień w obszarze automatycznej analizy ruchu. Zadanie to realizowane jest przez ekstrakcję cech opisujących ruch w danych wejściowych, opcjonalnie wybór cech o największej sile dyskryminacyjnej, oraz uczenie klasyfikatorów [58]. Zagadnienie wyboru zestawu danych oraz cech, które w naj-

wierniejszy sposób opisać wykonywany ruch jest jednym z kluczowych czynników. Z przeglądu literatury wynika, że obecnie wiele zbiorów danych jest ogólnie dostępnych, jednakże zawierają one akcje które w znaczący sposób różnią się między sobą [62]. W sporcie występują często podobne akcje, które ze względu na inną dynamikę wykonania przekładają się na różne techniki sportowe. Przykładem takich akcji są różne typy wypadów szermierczych. W niniejszym rozdziale przedstawiono opracowane metody, dzięki którym uzyskano skuteczne rozpoznawanie wypadów szermierczych w oparciu o analizę dynamiki ruchu. Wykonane eksperymenty pokazały, że również na ogólnie dostępnym zbiorze danych UTD-MHAD [14], metody te pozwalają uzyskać skuteczność lepszą niż metody znane z literatury. Przeprowadzone badania opisano w trzech pracach [38, 40, 41].

3.1 Zbiór danych z szermierczą pracą nóg

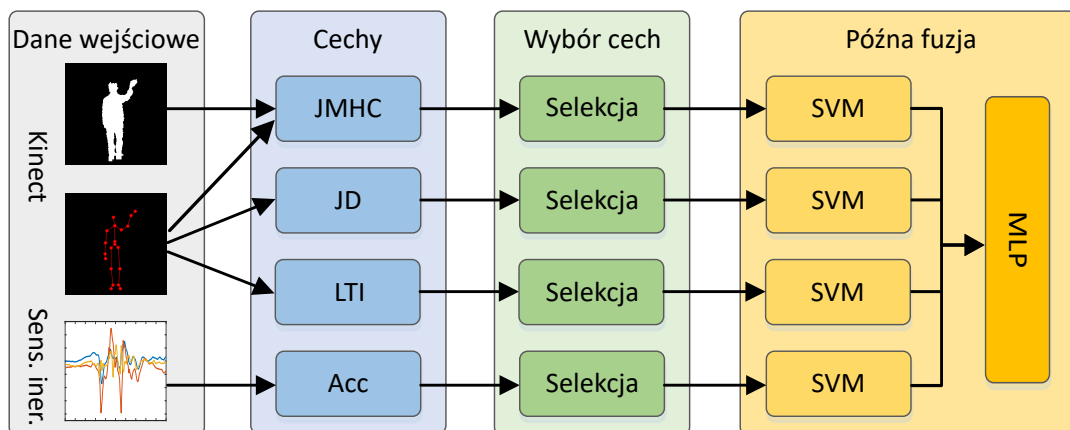
Mając na względzie to, że brak jest ogólnie dostępnych zbiorów danych obrazujące szermierczą pracę nóg, nagrano własny zbiór, nazwany *Fencing Footwork Dataset* (FFD). Obejmuje on 6 podstawowych akcji: krok w przód, krok w tył oraz cztery typy wypadów - szybkościowo-zrywowy, z narastającą szybkością, z wyczekiwaniem oraz skocznoślizgowy. Każdy z omawianych typów wypadu przekłada się na element taktyki zawodnika, np. atak z zaskoczenia lub sprowokowanie określonego działania przeciwnika. Wspomniane wyżej akcje wybrane zostały na podstawie analizy prac prof. Czajkowskiego [18] oraz konsultacji z trenerami szermierki. Akwizycja danych przeprowadzona została z użyciem sensora Kinect, dostarczającego dane RGB, mapy głębi, oraz szkielet wykrytej postaci, jak również wykorzystano sensor inercyjny, zamontowany na kolanie, który dostarczał informacje o przyspieszeniu, prędkości kątowej oraz polu magnetycznym. Sesje nagraniowe przeprowadzone zostały z udziałem 10 szermierzy ze Szkoły Fechtunku Aramis [4]. Każda akcja powtarzana była 10 lub 11 razy. Zarejestrowany zbiór danych został udostępniony w Internecie [35].

3.2 Zaproponowane metody

Na potrzeby rozpoznawania podstawowych akcji w szermierczej pracy nóg zaproponowano nowe metody do ekstrakcji, selekcji oraz fuzji cech. Ekstrakcja obejmuje cechy uzyskane z sygnału przyspieszenia (ang. *Accelerometric* (Acc)), dwa zestawy cech wyznaczanych na podstawie danych szkieletowych - dynamika stawów (ang. *Joint Dynamics* (JD)) oraz lokalne obrazy śladów (ang. *Local Trace Images* (LTI)), a także cechy oparte o dane szkieletowe i dane głębi, które nazywane są kontekstem historii ruchu stawów (ang. *Joint Motion History Context* (JMHC)). Dla każdego zestawu cech realizowana jest selekcja cech o największej sile dyskryminacyjnej za pomocą opracowanej metody wyboru cech, a następnie dla każdego z nich przeprowadzana jest klasyfikacja z użyciem maszyn wektorów nośnych (ang. *Support Vector Machine* (SVM)). Uzyskane prawdopodobieństwa klas są potem łączone we wspólny wektor, który jest klasyfikowany przez perceptron wielowarstwowy (ang. *Multilayer Perceptron* (MLP)). Zarys zaproponowanego podejścia przedstawiono na Rys. 3.1.

Cechy Acc wyznaczone zostały w następujący sposób. Sygnał przyspieszenia poddano filtracji górnoprzepustowej, a następnie wyznaczano pierwszą pochodną przefiltrowanego sygnału oraz różnicę między przefiltrowanym oraz oryginalnym sygnałem. Uzyskane przebiegi zostały poddane interpolacji w celu ujednoczenia ich długości pomiędzy na-

granicami. Następnie zastosowano podział na 7 okien czasowych z 50% nakładaniem się okien. W wspomnianych oknach wyznaczano szereg cech statystycznych w dziedzinie czasu, jak również inne cechy w dziedzinach częstotliwościowej i falkowej. Ostatecznie zastosowano cechy statystyczne w dziedzinie czasu, ponieważ pozwalały one uzyskać lepszą skuteczność klasyfikacji niż pozostałe cechy.



Rysunek 3.1: Zarys zaproponowanego podejścia.

Dla cech JD wykorzystane zostały pozycje stawów szkieletu, dostarczane przez sensor Kinect. Na podstawie pozycji stawów wyznaczone zostały prędkości oraz przyspieszenia. Otrzymane szeregi czasowe poddano interpolacji celem uzyskania ujednocionej dla wszystkich nagrań długości wektora danych, a następnie zastosowano podział na okna czasowe. Okna obejmowały trzy poziomy, przy czym w kolejnych poziomach długość okna była o połowę mniejsza. Poczynając od okna o długości równej długości wektora danych, przy 50% nakładaniu się okien, dało to w sumie 11 okien. W każdym z nich wyznaczona została krótkookresowa transformata Fouriera. Pierwsze trzy współczynniki transformaty, wyznaczone dla każdego kanału danych oraz w każdym oknie zostały wykorzystane jako cechy.

Kolejne zaproponowane cechy, to znaczy LTI, które oparte zostały również o dane szkieletowe, dostarczają innego typu informacji. Dla każdego stawu szkieletu tworzony jest obraz śladu ilustrujący ruch tego stawu w czasie. Dla każdej klatki nagrania zostają wygenerowane na czarnym tle piksele o jasności wynikającej z dwuwymiarowego rozkładu Gaussa zlokalizowanego w miejscu położenia danego stawu szkieletowego. Zastosowanie rozkładu Gaussa wprowadza generalizację, która jest niezbędna dla porównania różnych wykonania danego ruchu. Jasność uzyskanych obrazów jest normalizowana względem ich liczby. Następnie zostają one zsumowane dla każdego stawu z osobna. Ostateczny wektor cech składa się z połączonych obrazów generowanych dla wszystkich użytych stawów.

Cechy JMHC wykorzystują zarówno dane szkieletowe jak i mapy głębi. Ruch opisywany jest poprzez logarytmiczno-polarne histogramy umieszczone na stawach w obrazach będących różnicą sylwetki między aktualną, a wcześniejszą klatką. Zaproponowany deskryptor wykorzystuje sumę ważoną histogramów wyznaczanych dla różnic między aktualną oraz trzema wcześniejszymi klatkami. Podobnie jak w przypadku cech Acc oraz JD, dane są interpolowane oraz dzielone na okna czasowe, głównie w celu uzyskania ostatecznego wektora cech obejmującego całe nagranie. Zastosowano 7 okien z 50% nakładaniem się oraz wyznaczono cechy statystyczne w każdym z nich.

Na potrzeby selekcji cech opracowany został nowy algorytm, należący do klasy metod filtrujących [13]. Na początku tworzona jest macierz histogramów z rozkładem wartości każdej cechy dla każdej klasy. Następnie, na podstawie odległości między histogramami, wyznaczana jest macierz opisująca siłę dyskryminacyjną każdej cechy do odróżniania klas w każdej możliwej parze klas. Wybór cech odbywa się iteracyjnie na podstawie wartości wynikających z macierzy opisującej siłę dyskryminacyjną, wraz z uwzględnieniem korelacji między aktualnie rozważaną cechą, a cechami wybranymi wcześniej.

W kontekście klasyfikacji akcji przebadano zarówno wczesną fuzję, na poziomie cech, jak również późną fuzję, na poziomie decyzji. Wczesna fuzja zrealizowana została przez połączenie wektorów cech po selekcji. Późna fuzja obejmowała klasyfikację każdego zestawu cech przez osobny klasyfikator SVM, a następnie przekazanie wyznaczonych prawdopodobieństw klas w formie wspólnego wektora do klasyfikatora MLP.

3.3 Wyniki badań eksperymentalnych

Zaproponowane metody zostały zweryfikowane eksperymentalnie na dwóch zbiorach danych. Pierwszym z nich był własny zbiór FFD, a drugim ogólnie dostępny zbiór UTD-MHAD [14]. Zawiera on dane RGB, głębi, szkieletowe oraz inercyjne dla 27 akcji, wykonywanych przez 8 osób i powtarzanych 4-krotnie. Dla zbioru FFD cechy wyznaczane były jedynie dla 8 dolnych stawów szkieletowych w jednej płaszczyźnie ruchu (zgodnej z płaszczyzną obrazu sensora Kinect). Wynika to ze specyfiki pracy nóg w szermierce, w której brak jest ruchu w dodatkowych płaszczyznach, a ruch górnych części ciała może być niezależny od ruchu nóg. Dla zbioru UTD-MHAD wykorzystano wszystkie stawy szkieletu w trzech ortogonalnych płaszczyznach.

Badania eksperymentalne dla zbioru FFD przeprowadzono dla danych uczących i testowych tej samej osoby (ang. *Person Dependent* (PD)) z użyciem 5-krotnej walidacji krzyżowej, jak również dla przypadku uczenia na innych osobach niż osoba testowana (ang. *Person Independent* (PI)), z użyciem walidacji krzyżowej typu *leave-one-out*. Scenariusz PD służył do weryfikacji spójności wykonania akcji przez pojedyncze osoby, zaś scenariusz PI do zbadania możliwości generalizacji rozpoznawania. Pierwszym krokiem było porównanie skuteczności klasyfikatora SVM oraz lasów drzew losowych (ang. *Random Forest* (RF)) dla wszystkich typów cech. Klasyfikator SVM okazał się lepszy i dlatego wykorzystano go w dalszych eksperymentach. Następnie wykonano ręczną selekcję wykorzystywanych stawów szkieletowych, co pokazało, że dla cech JD oraz LTI należało usunąć cechy związane ze stawami kostek. W kolejnym kroku przebadano skuteczność zaproponowanej metody selekcji dla poszczególnych zestawów cech z osobna, jak również dla podejścia z wczesną i późną fuzją danych. Zaproponowana metoda selekcji okazała się lepsza niż użyte dla porównania metody Lasso [59], AdaBoost [42] oraz PCA [1]. Przebadano również zastosowanie klasyfikatora SVM z jądrem opartym o radialne funkcje bazowe (ang. *Radial Basis Function* (RBF)), ale nie poprawiło to wyników. Ostatecznie udało się uzyskać wyniki klasyfikacji 98.67% dla scenariusza PD oraz 86.31% dla scenariusza PI. Dokładność ta jest znacznie wyższa niż dokładność metod znanych z literatury, które uzyskiwały 94.55% dla scenariusza PD oraz 76.14% w scenariuszu PI.

Badania eksperymentalne dla zbioru UTD-MHAD wykonano zgodnie z zalecanym przez autorów zbioru protokołem. W omawianym protokole zbiór uczący obejmuje osoby 1, 3, 5 i 7, zaś zbiór testowy osoby 2, 4, 6 i 8. Przeprowadzono eksperymenty weryfikujące

skuteczność zaproponowanej metody selekcji na każdym zestawie cech osobno, jak również przy wykorzystaniu późnej fuzji. Na pojedynczych zestawach cech zaproponowana metoda nie zawsze dawała najlepsze rezultaty, niemniej jednak przy zastosowaniu fuzji okazała się skuteczniejsza niż metody użyte do porównania - Lasso, AdaBoost oraz PCA. Zastosowanie jądra RBF dla klasyfikatora SVM poprawiło wynik klasyfikacji. Opracowana metoda umożliwia uzyskanie 94.91% skuteczności klasyfikacji, który to wynik jest lepszy od wyników uzyskiwanych przez metody znane z literatury, które są oparte o te same modalności danych.

4. Detekcja oraz analiza akcji w pracy nóg w szermierce w czasie rzeczywistym

Większość prac w literaturze poświęconych analizie ruchu wykorzystuje posegmentowane zbiory danych, w których każda akcja rozpatrywana jest jako osobny przykład [62]. W analizie ruchu w sporcie etap segmentacji czasowej jest bardzo istotny, szczególnie w kontekście analizy danych w czasie rzeczywistym. W sporcie ważne jest również wyznaczenie parametrów jakościowych akcji, które odzwierciedlają poprawność wykonywanego ruchu. W niniejszym rozdziale zaproponowano oraz przebadano metody adaptacyjnego filtrowania danych szkieletowych oraz inercyjnych w oparciu o model ruchu, pozwalające na detekcję segmentów czasowych zawierających akcję wypadu w ciągłej szermierczej pracy nóg. Wykryte akcje są następnie analizowane pod kątem jakościowym, a informacje zwrotne przekazywane są do ćwiczących szermierzy w czasie rzeczywistym, umożliwiając poprawę ruchu w trakcie ćwiczenia. Wyniki badań eksperymentalnych omówiono w pracy [39].

4.1 Akwizycja danych

Na potrzeby opracowania metod segmentacji czasowej oraz analizy jakościowej akcji nagrano dwa dedykowane zbiory danych obrazujące ciągłą szermierczą pracę nóg. Pierwszy zbiór został pozyskany przy użyciu sensora Kinect, który dostarczał danych głębi oraz szkieletowych. W sesjach nagranych uczestniczyło 8 szermierzy o średnim i zaawansowanym poziomie umiejętności. Każdy z nich proszony był o wykonanie pokazowych sekwencji akcji z pracą nóg, polegających na przemieszczaniu się w przód i w tył krokami szermierczymi oraz wykonywaniu wypadów w wybranym przez siebie momencie. Zbiór danych obejmuje od 2 do 5 nagrań dla każdej osoby, a każde z nich zawiera od 4 do 6 akcji wypadu. W sumie zbiór obejmuje 149 wypadów szermierczych.

Drugi zbiór danych został nagrany z wykorzystaniem zarówno sensora Kinect, jak i własnoręcznie zbudowanego systemu złożonego z dwóch sensorów inercyjnych podłączonych do mikrokontrolera. Sensory były zamontowane na klatce piersiowej oraz na łokciu zawodnika. Podobnie jak dla pierwszego zbioru, zarejestrowano ciągłą pracę nóg. W drugim scenariuszu zebrano dane dla 9 osób, po 3 do 4 nagrań na osobę, od 5 do 6 wypadów na nagranie. Łącznie drugi zbiór danych obejmuje 162 akcje wypadu szermierczego.

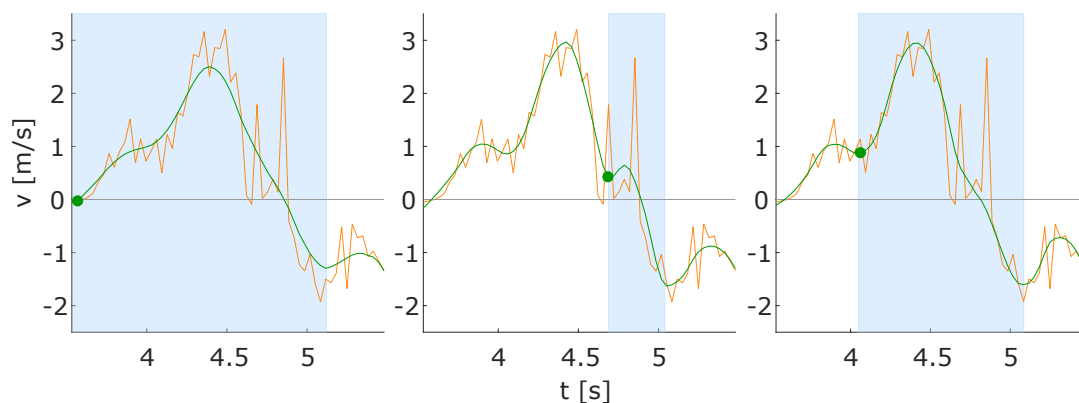
4.2 Zaproponowane metody

Opracowane podejście do analizy szermierczej pracy nóg składa się z następujących etapów. Na początku sygnał wejściowy jest filtrowany w celu wygładzenia przebiegu.

Następnie identyfikowane są segmenty czasowe potencjalnie zawierające akcje wypadu. Segmenty te klasyfikowane są w celu znalezienia rzeczywistych wypadów. Kluczowe momenty są wykrywane w każdym segmencie z wypadem, co umożliwia analizę jakościową. Informacja zwrotna przekazywana jest w czasie rzeczywistym zawodnikowi oraz opcjonalnie trenerowi. Zaproponowane metody zostały pierwotnie zaprojektowane dla danych szkieletowych, a następnie zaadaptowane dla danych inercyjnych.

Sensor Kinect wykrywa pozycje 20 stawów szkieletu postaci ludzkiej. W celu analizy ciągłej pracy nóg wykorzystano prędkość stawu zlokalizowanego przy podstawie kręgosłupa, który jest śledzony najbardziej stabilnie. Szermierze poruszają się w pozycji bocznej, z ręką zbrojną w kierunku przeciwnika i drugą ręką ułożoną za plecami. Akcja wypadu polega na wyproście ręki zbrojnej, następnie wyproście nogi przedniej i mocnym wybiciu się z nogi tylnej. Na końcu następuje powrót do pozycji podstawowej. W sygnale prędkości wybranego stawu przekłada się to na występowanie lokalnego maksimum większego od zera, po którym pojawia się lokalne minimum mniejsze od zera. Taka para lokalnych ekstremów wskazuje na potencjalną akcję wypadu, może jednak być pomyłona z sekwencją kroków w przód oraz w tył. W związku z powyższym potrzebna jest klasyfikacja potencjalnych akcji wypadu. Rozpoznawanie akcji wypadu wykonywane jest w oparciu o prędkości wszystkich stawów szkieletu. Dane ilustrujące prędkości są interpolowane do wspólnej długości, a następnie podzielone na okna czasowe, w których wyznaczana jest szybka transformata Fouriera. Pierwsze trzy współczynniki transformaty wykorzystano do uczenia klasyfikatora SVM.

W celu wiarygodnego wykrycia ekstremów w przebiegu prędkości wybranego stawu wymagana jest filtracja. Początkowo przebadano użyteczność średniej ruchomej (ang. *Moving Average* (MA)) oraz filtru Loess [17] opartego o wielomiany drugiego stopnia. Pojawił się jednak problem z określeniem początku akcji wypadu, który wskazywany jest przez dodatkowe lokalne minimum, występujące przed opisaną wcześniej sekwencją ekstremów. Filtr MA zbyt wygładzał przebieg na zboczu rosnącym, co powodowało sporadycznie pominięcie lokalnego minimum, natomiast filtr Loess sporadycznie generował dodatkowe, nieprawidłowe lokalne minimum na zboczu opadającym (zob. Rys. 4.1). Wspomniane zjawiska były motywacją do zaproponowania własnej, adaptacyjnej metody filtracji opartej o model ruchu. Metoda ta filtruje sygnał w oparciu o sumę ważoną filtrów MA oraz Loess, zaś wagi zależą od kierunku zbocza (rosnące lub opadające). Pozwoliło to na skuteczne wykrywanie ekstremów w przebiegu prędkości (zob. Rys. 4.1).



Rysunek 4.1: Detekcja segmentu z wypadem. Filtry MA (po lewej) oraz Loess (na środku) mają trudności ze znalezieniem początku wypadu. Zaproponowana metoda (po prawej) poprawnie znajduje początek wypadu.

Analiza jakościowa wykrytej akcji wykorzystuje określone przez trenerów szermierki parametry. Na wstępie wyznaczana jest różnica między momentem wyprostowania ręki (wykrywanym na podstawie kąta między stawami w barku, łokciu i nadgarstku) oraz rozpoczęcia ruchu wypadu. Następnie wyznaczana jest długość oraz czas trwania wypadu. Podawane są również parametry dynamiczne - średnia i maksymalna prędkość oraz przyspieszenie. Analiza danych wykonywana jest w czasie rzeczywistym. Informacje zwrotne przekazywane są na bieżąco, przez wysłanie ich do smartfona, który jest trzymany w ręce przez osobę ćwiczącą. Dzięki omawianemu rozwiązaniu możliwa jest korekta wykonywanego ruchu w trakcie treningu.

Podobną funkcjonalność uzyskano w oparciu o sensory inercyjne. Analiza przebiegu sygnału wartości pola magnetycznego uzyskanego z sensora zamontowanego przy łokciu umożliwiła skuteczne wykrycie momentu wyprostowania ręki. Analiza przebiegu sygnału przyspieszenia z sensora zamontowanego na klatce piersiowej pozwoliła z kolei na znalezienie początku i końca ruchu wypadu. Na tej podstawie uzyskano podobne parametry jakościowe wykrytych akcji wypadu jak dla sensora Kinect, z wyjątkiem długości oraz prędkości. Informacja zwrotna przekazywana była w czasie rzeczywistym.

4.3 Wyniki badań eksperymentalnych

Zaproponowane metody przebadano na obu zarejestrowanych zbiorach danych. W oparciu o dane z sensora Kinect dla obu zbiorów przygotowano ręcznie opis danych, uwzględniając dla każdego wypadu jego początek i koniec oraz moment wyprostowania ręki. Wspomniana adnotacja danych umożliwiła ewaluację omawianych metod. Dla wszystkich rozważanych filtrów (MA, Loess oraz zaproponowany), z różnymi długościami okna filtrowania, wyznaczono parametry ilościowe oraz jakościowe określające skuteczność detekcji oraz analizy akcji. Wyznaczono liczbę poprawnie wykrytych wypadów (ang. *True Positive* (TP)), liczbę mylnie wykrytych wypadów (ang. *False Positive* (FP)) oraz dwie miary skuteczności klasyfikacji - czułość (ang. *recall*) i precyzję (ang. *precision*). Czułość określa stosunek poprawnie wykrytych wypadów do wszystkich wypadów w zbiorze danych, natomiast precyzja określa stosunek poprawnie wykrytych wypadów do wszystkich wykrytych wypadów. Dla pierwszego zbioru danych jedynie zaprojektowana metoda filtracji uzyskuje jednocześnie 100% czułości oraz precyzji. Metoda ta dawała również najlepsze wyniki w zakresie dokładności określenia czasu trwania i długości wypadu, jak również różnicy między momentem wyprostowania ręki, a rozpoczęciem wypadu.

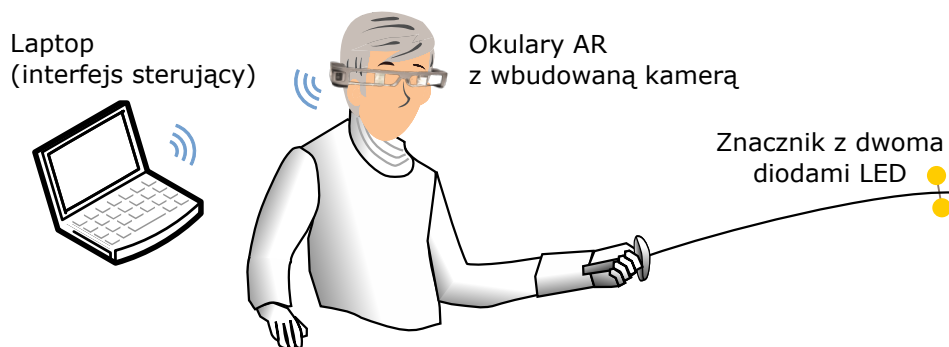
Dla drugiego zbioru danych eksperymenty przeprowadzono zarówno dla danych z sensora Kinect jak i dla danych inercyjnych. Wspomniane eksperymenty realizowane były celem porównania użyteczności obydwu metod. Dla danych z sensora Kinect uzyskano ponownie 100% czułości oraz precyzji, tym razem dla różnych metod filtracji, aczkolwiek zaproponowana metoda pozwalała na największy zakres długości okna filtrowania. Dla danych inercyjnych najlepszy wynik uzyskano dla filtrów MA oraz zaproponowanego. Dla obu filtrów były to czułość 99.38% oraz precyzja 98.77%. Analiza dokładności wykrywania czasu trwania wypadu, a także różnicy czasu wyprostowania ręki oraz rozpoczęcia wypadu pokazała, że metody oparte o dane inercyjne uzyskują mniejszą dokładność. Niemniej jednak jest ona wystarczająca, aby zapewnić użyteczne informacje zwrotne dla ćwiczących szermierzy. Trenerzy szermierki pozytywnie ocenili prototyp systemu w kontekście wspomagania doskonalenia ruchu wypadu.

5. Immersyjny kanał zwrotny dla ćwiczeń z bronią w szermierce z użyciem rozszerzonej rzeczywistości

Dostarczanie informacji zwrotnej dla sportowców, która jest prezentowana oraz interpretowana intuicyjnie oraz w czasie rzeczywistym jest zagadnieniem rzadko poruszonym w literaturze. W nielicznych pracach autorzy proponują wizualizacje specyficzne dla danych dyscyplin sportowych, takich jak tenis stołowy czy biathlon [5]. Brakuje jednak tego typu rozwiązań, które wspierałyby ćwiczenia z bronią w szermierce. W niniejszym rozdziale opisano opracowany system do wspomagania ćwiczeń z bronią, który zarazem oferuje immersyjny kanał zwrotny przez wykorzystanie rozszerzonej rzeczywistości (ang. *Augmented Reality* (AR)). System został opisany w pracy [36].

5.1 Architektura systemu

Trening z bronią w szermierce wymaga wiernego odtwarzania ruchu zaprezentowanego przez trenera, najpierw samodzielnie, a następnie w ćwiczeniach z partnerem. Jednym z najbardziej intuicyjnych sposobów rozumienia ruchu broni jest percepcja i postrzeganie trajektorii jej końcówki. Celem badań i prac eksperymentalnych było opracowanie kanału do dostarczenia użytecznej informacji zwrotnej w oparciu o trajektorie broni. W tym celu wykorzystano okulary do rozszerzonej rzeczywistości, które tworzą mieszany widok świata wirtualnego i rzeczywistego. Wspomniana operacja realizowana jest poprzez nakładanie generowanych obiektów na widok otoczenia, dzięki zastosowaniu półprzezroczystych wyświetlaczy. Śledzenie ruchu broni zrealizowane jest za pomocą kamery wbudowanej w okulary oraz znacznika z diodami LED, który jest zamontowany na końcówce klingi. Dzięki opracowanym metodom do mapowania współrzędnych między światem rzeczywistym oraz wirtualnym możliwe jest zobrazowanie na wyświetlaczach okularów trajektorii, które są dopasowane do fizycznego ruchu broni. Dodatkowo zaproponowano metody uczenia modeli akcji wykonywanych bronią, co umożliwi prowadzenie ćwiczeń poprzez wyświetlanie w czasie rzeczywistym trajektorii modelowych oraz aktualnych. Zapewnia to intuicyjny, immersyjny kanał dostarczania informacji zwrotnej w czasie treningu. Sterowanie systemem realizowane jest w oparciu o aplikację na laptopie, który komunikuje się bezprzewodowo z okularami. Architektura systemu zaprezentowana została na Rys. 5.1.



Rysunek 5.1: Architektura zaproponowanego systemu (materiały graficzne: [21, 61]).

5.2 Zaproponowane metody

Automatyczna analiza ruchu broni w szermierce jest złożonym zagadnieniem. Klinga porusza się bardzo szybko oraz jest cienka i wykonana ze stali, która odbija światło. Utrudnia to znacznie wykorzystanie typowych algorytmów śledzenia obiektów, szczególnie w oparciu o obrazy pobierane z kamery wbudowanej w okulary AR, która dostarcza strumienia wideo w stosunkowo niskiej jakości. Z tego powodu zdecydowano się wykorzystać znacznik z diodami LED zamontowany na końcu broni. Dzięki zastosowaniu dwóch diod LED, umieszczonych prostopadle do klingi, możliwe jest również oszacowanie głębi oraz rotacji klingi. W pierwszej kolejności na obrazach pobranych z kamery realizowana jest operacja progowania, której efektem są binarne obrazy, na których jasne punkty odpowiadające diodom LED są zaznaczone na białą, a tło pozostaje czarne. Granica progowania znajduje się automatycznie w oparciu o nieskomplikowaną procedurę kalibracyjną.

W kolejnym kroku realizowana jest klasteryzacja białych pikseli. Za pomocą rekursywnej metody rozrastania obszarów [2] znajdują się klastry białych pikseli, które odpowiadają dwóm diodom LED. Środki tych klastrów nazywane są w niniejszej pracy punktami kluczowymi. Wspomniane punkty kluczowe wykorzystywane są do wyznaczenia położenia, głębi oraz rotacji klingi. Położenie końcówki broni wyznaczane jest jako punkt środkowy pomiędzy punktami kluczowymi. Głębina szacowana jest w oparciu o pomiar odległości między tymi punktami. Im bliżej kamery, tym ta odległość jest większa. Rotacja klingi określona jest przez kąt wyznaczony między linią łączącą punkty kluczowe oraz poziomą osią kamery.

Uzyskane w ten sposób położenia końca broni w kolejnych klatkach stanowią zapis trajektorii ruchu. Trajektorie te są interpolowane do wspólnej długości z użyciem funkcji sklepanych, co umożliwi ich późniejsze porównywanie na podstawie modeli ruchu. Wymagany jest jednak również wspólny punkt startowy. Tworzenie modelu ruchu akcji bronią zaczyna się od ustalenia punktu startowego w przestrzeni. W omawianej procedurze trener umieszcza broń w wybranej pozycji i zapisuje położenie punktu w aplikacji sterującej. Następnie wykonuje on szereg powtórzeń akcji, zaczynając zawsze w punkcie startowym, który wyświetlany jest w okularach AR. Trajektorie z kolejnych powtórzeń akcji zostają uśrednione i stanowią model, z którym będą porównywane ćwiczenia wykonywane przez zawodników. Na potrzeby ewaluacji ćwiczeń, model przechowuje również odchylenie standardowe wyznaczone w każdym punkcie trajektorii. W czasie treningu trajektoria modelowa oraz aktualna porównywane są punkt po punkcie w oparciu o analizę odchylenia standardowego. Oznacza to, że po każdym wykonaniu ćwiczenia wyświetlana jest procentowa średnia zgodność z modelem akcji. Model przechowuje również średnie wartości i odchylenia standardowe kątów obrotu w każdym punkcie. Oszacowanie zgodności kątów przeprowadzane jest analogicznie.

Kolejne powtórzenia wykonania akcji, zarówno na etapie uczenia modeli jak i ewaluacji poprawności wykonania ćwiczeń, wykrywane są automatycznie przez maszynę stanów. W fazie początkowej system jest w stanie detekcji znaczników. Po najejchaniu końcem broni na punkt startowy następuje uruchomienie licznika, który po jednej sekundzie zmienia stan na gotowy do rozpoczęcia akcji. Po wykryciu rozpoczęcia ruchu klingi, stan maszyny zmienia się na stan nagrywania akcji. Powrót do pierwotnego stanu następuje po zatrzymaniu końca broni.

W zaproponowanym systemie wykorzystano urządzenie Epson Moverio BT-300, które zawiera okulary AR oraz jednostkę przetwarzającą, pracującą pod kontrolą systemu Android. Okulary te mogą pracować w trybie 3D. Dzięki wyświetlaniu przesuniętego obrazu dla lewego i prawego oka, uzyskiwany jest efekt stereoskopowy, co pozwala postrzegać głębię wyświetlanych obiektów. Zasadniczym problemem do rozwiązania było opracowanie metody pozwalającej na dopasowanie wirtualnych trajektorii do fizycznego położenia końca broni w mieszanym widoku rzeczywistości rozszerzonej. W tym celu opracowana została metoda kalibracji mapowania rzeczywistych współrzędnych znacznika LED widzianego w obrazie z kamery na wirtualne współrzędne obiektów wyświetlanych na okularach. W czasie kalibracji użytkownik umieszcza koniec broni tak, aby pokrywał się z wyświetlanymi na okularach znacznikami. Pozwala to zebrać szereg punktów kalibracyjnych, na podstawie których wyznaczane są zależności opisujące mapowanie współrzędnych między światem rzeczywistym oraz wirtualnym.

5.3 Wyniki badań eksperymentalnych

Zaproponowane metody zostały przebadane w kilku eksperymentach. Na początku zweryfikowano poprawność śledzenia położenia oraz rotacji klingi. W tym celu zarejestrowano, z użyciem kamery w okularach AR, krótkie sekwencje ćwiczeń z bronią, wykonywane w trzech różnych lokalizacjach, co przekładało się na różne warunki oświetleniowe. Rozdzielczość obrazu wynosiła 640 x 480. We wszystkich nagraniach, na potrzeby ewaluacji dokładności śledzenia oznaczono ręcznie położenie znacznika z diodami LED. Uzyskano następujące wyniki. Diody LED były poprawnie wykrywane w 99.32% klatek. Położenie końca broni śledzone było z średnią dokładnością 1.17 ± 0.42 piksela. Odległość między diodami LED była wyznaczana ze średnią dokładnością 0.64 ± 0.99 piksela, zaś rotacja ze średnią dokładnością 0.92 ± 0.96 stopnia. Stwierdzono również, że wysoką dokładność śledzenia można uzyskać jedynie, jeśli w polu widzenia kamery nie występują inne źródła światła lub odbłaski. Wspomniane wyniki są zadowalające, co wskazuje na to, że opracowane metody do śledzenia położenia oraz rotacji klingi pozwalają uzyskać użyteczną informację zwrotną.

Metoda szacowania głębi na podstawie odległości między diodami LED w obrazie z kamery została zweryfikowana przez umieszczenie broni w różnych odległościach od kamery (co 10 cm) i zmierzenie odległości między diodami w zarejestrowanych obrazach. Okazało się, że w zakresie odpowiadającym akcjom wykonywanym z bronią (130 do 170 cm) głębia mogła być oszacowana z użyciem zależności liniowej. Zmierzono również średni czas przetwarzania jednej klatki, który wynosił 36 ms dla rozdzielczości obrazu 640 x 480 oraz 13 ms dla rozdzielczości obrazu 320 x 240. Mając na względzie powyższe, zasadne jest uznanie, że zaimplementowany system jest w stanie działać w czasie rzeczywistym.

Weryfikacja metody do porównywania trajektorii została przeprowadzona przez wygenerowanie przykładowej uśrednionej trajektorii ruchu broni oraz porównanie jej z szeregiem trajektorii odbiegających od niej w różnym stopniu. Eksperyment pokazał, że poziom zgodności trajektorii wyznaczony przez system jest znacząco zgodny z oceną człowieka. Zweryfikowano również mechanizm automatycznego wykrywania kolejnych powtórzeń akcji. Omawiana weryfikacja była realizowana na podstawie opinii użytkowników, którzy stwierdzili, że jest on wygodny w obsłudze i pozwala na ewaluację ćwiczeń bez użycia dodatkowego interfejsu sterującego.

W końcowej części badań przeprowadzono eksperymenty z immersyjnym kanałem dostarczania informacji zwrotnej przez okulary AR. W badaniach uczestniczyli szermierze, trenerzy szermierki oraz osoby bez doświadczenia w tym sporcie. Na początku wykonano kalibrację z udziałem trzech osób, co pozwoliło stwierdzić, że system poprawnie dopasowuje wirtualne trajektorie do fizycznego położenia broni. Następnie nagrano z trenerami modele trzech akcji, z spośród których dwie wykorzystano do eksperymentów. Trzecia akcja zawierała ruch w przód, którego ocena okazała się mało miarodajna, ze względu na niewystarczające postrzeganie głębi wygenerowanych trajektorii przez osoby ćwiczące. Szermierze byli w stanie powtórzyć wybrane dwie akcje z dość dobrą powtarzalnością. Osoby bez doświadczenia szermierczego uzyskiwały zróżnicowane wyniki. Trenerzy zauważyli, że system pozwala osobom początkującym wykonywać ćwiczenia bardziej poprawnie, niż przy tradycyjnym treningu, co potwierdza jego użyteczność na potrzeby wspomagania ćwiczeń z bronią.

6. Podsumowanie

W rozprawie zaprezentowano zagadnienia związane z automatyczną analizą ruchu w sporcie. W tym celu zaproponowano, przebadano oraz zaimplementowano szereg metod, które w efekcie umożliwiają wspomaganie doskonalenie umiejętności sportowych. Badania zrealizowano w trzech etapach. W etapie pierwszym zaproponowano metody pozwalające na rozpoznawanie specyficznych akcji sportowych, w oparciu o analizę dynamiki ruchu, na bazie multimodalnych danych. Opracowane algorytmy okazały się skuteczne dla zarejestrowanego zbioru danych obrazującego szermierczą pracę nóg, jak również dla bardziej ogólnych akcji z ogólnie dostępnego zbioru UTD-MHAD. W drugim etapie badania dotyczyły segmentacji czasowej ciągłego ruchu oraz analizy jakościowej wyodrębnionych akcji. Zaproponowane metody pozwalają na skuteczną detekcję akcji wypadków w szermierczej pracy nóg w czasie rzeczywistym, jak również wyznaczenie ich parametrów jakościowych oraz przesłanie ich do ćwiczących szermierzy w trakcie treningu. W trzecim etapie opracowano system do wspomagania ćwiczeń z bronią w szermierce, który wykorzystuje okulary do rozszerzonej rzeczywistości. Zaproponowano metody do śledzenia ruchu klingi, tworzenia modeli akcji wykonywanych bronią, jak również procedurę kalibracji mapowania współrzędnych rzeczywistych na wirtualne. Metody te złożyły się na innowacyjne rozwiązanie, w którym immersyjny kanał dostarcza informacji zwrotnej w postaci mieszanego widoku wirtualnych trajektorii dopasowanych do fizycznego położenia broni.

Opracowane metody rozwiązują istotne problemy w obszarze automatycznej analizy ruchu w sporcie oraz pokazują, że z ich pomocą możliwe jest wspomaganie procesu doskonalenia umiejętności sportowych, nawet bez nadzoru trenera. Metody opisane zostały w pracach konferencyjnych oraz w czasopiśmie recenzowanych [37–41]. Dalsze prace skupiają się głównie nad rozwojem metod segmentacji czasowej oraz analizy jakościowej akcji w szermierczej pracy nóg, w szczególności celem wykrywania innych typów akcji oraz wyznaczania dodatkowych parametrów jakościowych. Aktualnie prace realizowane są we współpracy ze szkołą szermierki Delta Fencing Center [20], zlokalizowaną w Kalifornii, w Stanach Zjednoczonych.

Bibliografia

- [1] Abdi H., Williams L.J.: Principal component analysis. In: *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2(4), pp. 433–459, 2010.
- [2] Adams R., Bischof L.: Seeded region growing. In: *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16(6), pp. 641–647, 1994.
- [3] Altun K., Barshan B., Tunçel O.: Comparative study on classifying human activities with miniature inertial and magnetic sensors. In: *Pattern Recognition*, vol. 43(10), pp. 3605–3620, 2010.
- [4] Aramis Fencing School. <https://aramis.pl/>. Last access on Jan 2019.
- [5] Baca A., Kornfeind P.: Rapid feedback systems for elite sports training. In: *IEEE Pervasive Computing*, vol. 5(4), pp. 70–76, 2006.
- [6] Bao L., Intille S.: Activity recognition from user-annotated acceleration data. In: *Pervasive Computing*, pp. 1–17, 2004.
- [7] Beauchemin S.S., Barron J.L.: The computation of optical flow. In: *ACM Computing Surveys (CSUR)*, vol. 27(3), pp. 433–466, 1995.
- [8] Bober T., Rutkowska-Kucharska A., Jaroszczuk S., Barabasz M., Woźnica W.: Kinematic characterisation of the lunge and the fleche in epee fencing: two case studies. In: *Polish J. of Sport and Tourism*, vol. 23(4), pp. 181–185, 2016.
- [9] Bobick A.F., Davis J.W.: The recognition of human movement using temporal templates. In: *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23(3), pp. 257–267, 2001.
- [10] Borges P.V., Conci N., Cavallaro A.: Video-based human behavior understanding: A survey. In: *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 23(11), pp. 1993–2008, 2013.
- [11] Borysiuk Z., Piechota K., Minkiewicz T.: Analysis of performance of the fencing lunge with regard to the difficulty level of a technical-tactical task. In: *J. of Combat Sports and Martial Arts*, vol. 4(2), pp. 135–139, 2013.
- [12] Campaniço A.T., Valente A., Seródio R., Escalera S.: Data’s hidden data: qualitative revelations of sports efficiency analysis brought by neural network performance metrics. In: *Motricidade*, vol. 14(4), pp. 94–102, 2018.
- [13] Chandrashekar G., Sahin F.: A survey on feature selection methods. In: *Computers & Electrical Engineering*, vol. 40(1), pp. 16–28, 2014.
- [14] Chen C., Jafari R., Kehtarnavaz N.: UTD-MHAD: A multimodal dataset for human action recognition utilizing a depth camera and a wearable inertial sensor. In: *IEEE Int. Conf. on Image Processing (ICIP)*, pp. 168–172. IEEE, 2015.
- [15] Chen H.T., Tsai W.J., Lee S.Y., Yu J.Y.: Ball tracking and 3D trajectory approximation with applications to tactics analysis from single-camera volleyball sequences. In: *Multimedia Tools and Applications*, vol. 60(3), pp. 641–667, 2012.
- [16] Chen L., Wei H., Ferryman J.: A survey of human motion analysis using depth imagery. In: *Pattern Recognition Letters*, vol. 34(15), pp. 1995–2006, 2013.

- [17] Cleveland W.S.: Robust locally weighted regression and smoothing scatterplots. In: *J. of the American Statistical Association*, vol. 74(368), pp. 829–836, 1979.
- [18] Czajkowski Z.: *Understanding Fencing. The Unity of Theory and Practice*. SKA Swordplay Books, 2005.
- [19] Dalal N., Triggs B.: Histograms of oriented gradients for human detection. In: *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, pp. 886–893. IEEE, 2005.
- [20] Delta Fencing Center. <http://www.deltafencingcenter.com/>. Last access on Jan 2019.
- [21] Freepik vectorpocket. <https://www.freepik.com/vectorpocket>. Last access on Jan 2019.
- [22] Gerke S., Linnemann A., Müller K.: Soccer player recognition using spatial constellation features and jersey number recognition. In: *Computer Vision and Image Understanding*, 2017.
- [23] Gholipour M., Tabrizi A., Farahmand F.: Kinematics analysis of lunge fencing using stereophotogrametry. In: *World J. of Sport Sciences*, vol. 1(1), pp. 32–37, 2008.
- [24] Hachaj T., Ogiela M.R., Koptyra K.: Effectiveness comparison of Kinect and Kinect 2 for recognition of Oyama karate techniques. In: *18th Int. Conf. on Network-Based Information Systems (NBIS)*, pp. 332–337. IEEE, 2015.
- [25] Herpin G., Gauchard G.C., Lion A., Collet P., Keller D., Perrin P.P.: Sensorimotor specificities in balance control of expert fencers and pistol shooters. In: *J. of Electromyography and Kinesiology*, vol. 20(1), pp. 162–169, 2010.
- [26] Hinton G.E., Osindero S., Teh Y.W.: A fast learning algorithm for deep belief nets. In: *Neural Computation*, vol. 18(7), pp. 1527–1554, 2006.
- [27] Kasiri S., Fookes C., Sridharan S., Morgan S.: Fine-grained action recognition of boxing punches from depth imagery. In: *Computer Vision and Image Understanding*, 2017.
- [28] Kasprzak W., Wilkowski A., Czapnik K.: Hand gesture recognition based on free-form contours and probabilistic inference. In: *Int. J. of Applied Mathematics and Computer Science*, vol. 22(2), pp. 437–448, 2012.
- [29] Krizhevsky A., Sutskever I., Hinton G.E.: Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, pp. 1097–1105. 2012.
- [30] Krzeszowski T., Kwolek B., Michalczuk A., Świtoński A., Josiński H.: View independent human gait recognition using markerless 3D human motion capture. In: *Int. Conf. on Computer Vision and Graphics*, pp. 491–500. Springer, 2012.
- [31] Kwolek B., Kepski M.: Improving fall detection by the use of depth sensor and accelerometer. In: *Neurocomputing*, vol. 168, pp. 637–645, 2015.
- [32] Laptev I.: On space-time interest points. In: *Int. J. of Computer Vision*, vol. 64(2-3), pp. 107–123, 2005.

- [33] Lin F., Chang C., Jou Y., Pan H., Hsu T.: The study of influence of fencing handle type and handle angle on wrist for a fencing game. In: *17th Int. Conf. on Industrial Engineering and Engineering Management (IE&EM)*, pp. 1624–1627. IEEE, 2010.
- [34] Lowe D.G.: Distinctive image features from scale-invariant keypoints. In: *Int. J. of Computer Vision*, vol. 60(2), pp. 91–110, 2004.
- [35] Malawski F.: Fencing Footwork Dataset (FFD). <http://home.agh.edu.pl/~fmal/ffd/>. Last access on Jan 2019.
- [36] Malawski F.: Real-time first person perspective tracking and feedback system for weapon practice support in fencing. In: *Applications of Intelligent Systems (AP-*PIS*)*, vol. 310, pp. 79 – 88, 2018.
- [37] Malawski F., Gałka J.: System for multimodal data acquisition for human action recognition. In: *Multimedia Tools and Applications*, vol. 7(18), pp. 1–26, Springer, 2018.
- [38] Malawski F., Kwolek B.: Classification of basic footwork in fencing using accelerometer. In: *Signal Processing: Algorithms, Architectures, Arrangements, and Applications (SPA)*, vol. 6, p. 20. 2016.
- [39] Malawski F., Kwolek B.: Real-time action detection and analysis in fencing footwork. In: *40th IEEE Int. Conf. on Telecommunications and Signal Processing (TSP)*, pp. 520–523. IEEE, 2017.
- [40] Malawski F., Kwolek B.: Improving multimodal action representation with joint motion history context. In: *Resubmitted with minor revision to Journal of Visual Communication and Image Representation*, 2019.
- [41] Malawski F., Kwolek B.: Recognition of action dynamics in fencing using multimodal cues. In: *Image and Vision Computing*, vol. 75, pp. 1–10, Elsevier, 2018.
- [42] Malawski F., Kwolek B., Sako S.: Using Kinect for facial expression recognition under varying poses and illumination. In: *Int. Conf. on Active Media Technology (AMT)*, pp. 395–406. Springer, 2014.
- [43] Manaffard M., Ebadi H., Moghaddam H.A.: A survey on player tracking in soccer videos. In: *Computer Vision and Image Understanding*, 2017.
- [44] Mantovani G., Ravaschio A., Piaggi P., Landi A.: Fine classification of complex motion pattern in fencing. In: *Procedia Engineering*, vol. 2(2), pp. 3423–3428, 2010.
- [45] Maqueda A.I., del Blanco C.R., Jaureguizar F., García N.: Human-computer interaction based on visual hand-gesture recognition using volumetric spatiograms of local binary patterns. In: *Computer Vision and Image Understanding*, vol. 141, pp. 126–137, 2015.
- [46] Mauthner T., Koch C., Tilp M., Bischof H.: Visual tracking of athletes in beach volleyball using a single camera. In: *Int. J. of Computer Science in Sport*, vol. 6(2), pp. 21–34, 2007.
- [47] Moeslund T.B., Thomas G., Hilton A.: *Computer vision in sports*. Springer, 2014.
- [48] Niebles J.C., Chen C.W., Fei-Fei L.: Modeling temporal structure of decomposable motion segments for activity classification. In: *European Conf. on Computer Vision (ECCV)*, pp. 392–405. Springer, 2010.

- [49] Ojala T., Pietikainen M., Maenpaa T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. In: *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 24(7), pp. 971–987, 2002.
- [50] Oreifej O., Liu Z.: HON4D: Histogram of oriented 4D normals for activity recognition from depth sequences. In: *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, pp. 716–723. 2013.
- [51] Pan M.S., Huang K.C., Lu T.H., Lin Z.Y.: Using accelerometer for counting and identifying swimming strokes. In: *Pervasive and Mobile Computing*, vol. 31, pp. 37–49, 2016.
- [52] Presti L.L., La Cascia M.: 3D skeleton-based human action classification: A survey. In: *Pattern Recognition*, vol. 53, pp. 130–147, 2016.
- [53] Reno V., Mosca N., Nitti M., D’Orazio T., Guaragnella C., Campagnoli D., Prati A., Stella E.: A technology platform for automatic high-level tennis game analysis. In: *Computer Vision and Image Understanding*, 2017.
- [54] Safdarnejad S.M., Liu X., Udpa L., Andrus B., Wood J., Craven D.: Sports videos in the wild (SVW): A video dataset for sports analysis. In: *11th IEEE Int. Conf. and Workshops on Automatic Face and Gesture Recognition (FG)*, vol. 1, pp. 1–7. IEEE, 2015.
- [55] Sarbolandi H., Lefloch D., Kolb A.: Kinect range sensing: Structured-light versus time-of-flight Kinect. In: *Computer Vision and Image Understanding*, vol. 139, pp. 1–20, 2015.
- [56] Sharma A., Agarwal M., Sharma A., Dhuria P.: Motion capture process, techniques and applications. In: *Int. J. on Recent and Innovation Trends in Computing and Communication*, vol. 1(4), pp. 251–257, 2013.
- [57] Stallings L.M.: *Motor Learning: From Theory to Practice*. Mosby Inc., 1982.
- [58] Stapor K.: *Metody Klasyfikacji Obiektów w Wizji Komputerowej*. Wydawnictwo Naukowe PWN, 2011.
- [59] Tibshirani R.: Regression shrinkage and selection via the lasso. In: *J. of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- [60] Weinland D., Ronfard R., Boyer E.: A survey of vision-based methods for action representation, segmentation and recognition. In: *Computer Vision and Image Understanding*, vol. 115(2), pp. 224–241, 2011.
- [61] Wikipedia: Fencing foil valid surfaces. https://commons.wikimedia.org/wiki/File:Fencing_foil_valid_surfaces_2009.svg/. Last access on Jan 2019.
- [62] Zhang J., Li W., Ogunbona P.O., Wang P., Tang C.: RGB-D-based action recognition datasets: A survey. In: *Pattern Recognition*, vol. 60, pp. 86–105, 2016.
- [63] Zhang L., Hsieh J.C., Ting T.T., Huang Y.C., Ho Y.C., Ku L.K.: A Kinect based golf swing score and grade system using GMM and SVM. In: *5th Int. Congress on Image and Signal Processing (CISP)*, pp. 711–715. IEEE, 2012.
- [64] Zhu F., Shao L., Xie J., Fang Y.: From handcrafted to learned representations for human action recognition: A survey. In: *Image and Vision Computing*, vol. 55, pp. 42–52, 2016.