



Politechnika Łódzka
Instytut Informatyki

ROZPRAWA DOKTORSKA

Hybrydowa metoda śledzenia ruchu człowieka w czasie rzeczywistym

Wydział Fizyki Technicznej, Informatyki i Matematyki Stosowanej

Promotor: dr hab. inż. Adam Wojciechowski

Dyplomant: mgr inż. Grzegorz Glonek

Nr albumu: 800291

Łódź, 24 kwietnia 2017

Instytut Informatyki

90-924 Łódź, ul. Wólczańska 215, budynek B9

tel. 42 631 27 97, 42 632 97 57, fax 42 630 34 14, email: office@ics.p.lodz.pl



Pragnę serdecznie podziękować mojemu promotorowi dr. hab. inż. Adamowi Wojciechowskiemu oraz Pani dr hab. inż. Marii Pietruszce, za nieustanne motywowanie w pracy, życzliwą opiekę oraz pomoc w przygotowaniu niniejszej rozprawy doktorskiej.

Dziękuję również rodzinie i przyjaciołom za cierpliwość i wyrozumiałość.

Grzegorz Glonek

Michałowi i Magdzie

Spis treści

Lista symboli	7
Rozdział 1. Wstęp	11
1.1. Problematyka pracy	13
1.2. Cel i teza pracy	13
1.3. Struktura pracy	14
Rozdział 2. Przegląd metod śledzenia ruchu	17
2.1. Systemy śledzenia ruchu postaci	18
2.1.1. Optyczne systemy śledzenia ruchu	18
2.1.2. Systemy nieoptyczne	26
2.2. Komputerowy model postaci ludzkiej	34
2.2.1. Model kinematyczny szkieletu postaci	35
2.2.2. Model reprezentujcy kształt postaci	37
2.2.3. Wyznaczanie położenia stawów w modelu szkieletowym	39
2.2.4. Reprezentacja rotacji obiektu w przestrzeni	46
2.3. Charakterystyka wykorzystywanych urządzeń pomiarowych	51
2.3.1. Kontroler Microsoft Kinect v. 1	51
2.3.2. Urządzenia inercyjne	62
2.3.3. Ograniczenia w działaniu czujników inercyjnych	66
2.4. Systemy hybrydowe łączące optyczny system śledzenia ruchu bez markerów oraz inercyjny system śledzenia ruchu	70
Rozdział 3. Hybrydowa metoda śledzenia ruchu człowieka	79
3.1. Format danych używany przez urządzenia pomiarowe	82
3.1.1. Kontroler Kinect	82
3.1.2. Czujniki inercyjne	83
3.2. Kalibracja	85
3.3. Korekta danych z urządzeń pomiarowych	88
3.4. Synchronizacja czasowa	93
3.5. Łączenie danych	95
Rozdział 4. Eksperymenty badawcze	103
4.1. Scena	103
4.2. Badanie ruchu	106

Ćwiczenie 1 – Zgięcie ramienia w łokciu do góry	109
Ćwiczenia 2 i 3 – Zgięcie ramienia w łokciu do przodu oraz wyciągnięcie wyprostowanych ramion do przodu	112
Ćwiczenie 4 – Utrzymanie wyprostowanych ramion	116
4.3. Podsumowanie	118
Rozdział 5. Podsumowanie i wnioski	123

Dodatki

Dodatek A. Popularne metody filtracji i estymacji sygnałów pomiarowych	131
Liniowy filtr Kalmana (KF - <i>ang. Kalman filter</i>)	131
Rozszerzony filtr Kalmana (EKF - <i>ang. Extended Kalman Filter</i>)	133
Bezśadowy filtr Kalmana (UKF - <i>ang. Unscented Kalman Filter</i>)	134
Filtr komplementarny	136
Filtры Mahoney'a i Madgwicka	137
Dodatek B. Wariancja Allana	143
Bibliografia	147
Spis rysunków	159
Spis tablic	163

Lista symboli

- α kąt obrotu sylwetki użytkownika względem kontrolera Kinect
 $\bar{A} = [\bar{a}_x, \bar{a}_y, \bar{a}_z]$ uśrednione wartości pomiarów z akcelerometru
 $\bar{G} = [\bar{g}_x, \bar{g}_y, \bar{g}_z]$ uśrednione wartości pomiarów z żyroskopu
- β kąt zgięcia ręki mierzony w stawie łokciowym
- \blacksquare^F indeks oznaczający estymaty będące wynikiem fuzji danych
- \blacksquare^I indeks oznaczający dane uzyskane z modułów inercyjnych
- \blacksquare^K indeks oznaczający dane uzyskane z kontrolera Kinect
- \blacksquare_e indeks oznaczający staw łokciowy *ang. elbow*
- \blacksquare_j indeks oznaczający wybrany staw *ang. joint*
- \blacksquare_w indeks oznaczający staw nadgarstkowy *ang. wrist*
- \blacksquare_{j-1} indeks oznaczający staw nadrzędny w hierarchicznym modelu szkieletowym
- \blacksquare_{sh_L} indeks oznaczający staw barkowy lewy *ang. shoulder left*
- \blacksquare_{sh_R} indeks oznaczający staw barkowy prawy *ang. shoulder right*
- ϕ kąt Eulera określający obrót wokół osi X
- ψ kąt Eulera określający obrót wokół osi Z
- τ przesunięcie czasowe pomiędzy sygnałami
- τ_{max} przesunięcie czasowe pomiędzy sygnałami, dla którego korelacja pomiędzy nimi przyjmuje największą wartość
- θ kąt Eulera określający obrót wokół osi Y
- $\widetilde{\omega}$ średnia wartość szumu błądzenia (ARW - *ang. Angular Random Walk*) żyroskopu
- $A = [a_x, a_y, a_z]$ wartości pomiarów z akcelerometru
- A' skorygowany pomiar akcelerometru ze względu na temperaturę
- $A_0 = [a_x, a_y, a_z]_0$ oczekiwane wartości spoczynkowe pomiarów akcelerometru
- $A_{th} = [a_x, a_y, a_z]_{th}$ maksymalny dopuszczalny błąd spoczynkowych pomiarów akcelerometru
- $cor = [cor_A, cor_G]$ współczynniki korekty pomiarów modułów inercyjnych
- $cor_A = [c_{ax}, c_{ay}, c_{az}]$ współczynniki korekty pomiarów akcelerometru
- $cor_G = [c_{gx}, c_{gy}, c_{gz}]$ współczynniki korekty pomiarów żyroskopu

$d_e(P_1, P_2)$	Odległość euklidesowa pomiędzy dwoma dowolnymi punktami P_1, P_2 w przestrzeni
$E = [\phi, \theta, \psi]$	orientacja przestrzenna w postaci kątów Eulera
f_A	współczynnik konwersji bezpośrednich pomiarów akcelerometru do jednostek przyspieszenia grawitacyjnego
f_G	współczynnik konwersji bezpośrednich pomiarów żyroskopu do prędkości kątowej $^{\circ}/s$
f_m	współczynnik filtracji pomiarów z czujników inercyjnych, wykorzystywany w filtrze Madgwicka
f_T	współczynnik korekty temperatury
f_{LPF}	współczynnik filtracji pomiarów z kontrolera Kinect, wykorzystywany w filtrze LPF
$G = [g_x, g_y, g_z]$	wartości pomiarów z żyroskopu
g	jednostka przyspieszenia grawitacyjnego. W przybliżeniu $1g = 9.80665 m/s^2$
$G_0 = [g_x, g_y, g_z]_0$	oczekiwane wartości spoczynkowe pomiarów żyroskopu
$G_{th} = [g_x, g_y, g_z]_{th}$	maksymalny dopuszczalny błąd spoczynkowych pomiarów żyroskopu
$I * K$	korelacja wzajemna sygnałów o źródłach w module inercyjnym oraz kontrolerze Kinect
I	sygnał, którego źródłem jest moduł inercyjny
K	sygnał, którego źródłem jest Kinect
$m(A, G, f_m, \Delta t)$	formuła filtru Madgwicka
$P = [p_x, p_y, p_z]$	położenie przestrzenne stawu
P'	położenie przestrzenne stawu po filtracji
$Q = [q_w, q_x, q_y, q_z]$	orientacja przestrzenna w postaci kwaternionowej
T	pomiar temperatury pracy modułu inercyjnego
t	znacznik czasu otrzymanych pomiarów
T_0	neutralna temperatura pracy modułu inercyjnego wg specyfikacji
T_{deg}	pomiar temperatury przedstawiony w formie stopni Celsjusza wyznaczony na podstawie T_{raw}
T_{raw}	pomiar temperatury w formacie bezpośrednio odczytanym z modułu inercyjnego
t_{szum}	czas, w którym pomiary z kontrolera Kinect pozostają uznawane za niewiarygodne
w_ϕ	waga wartości kąta Eulera ϕ
w_ψ	waga wartości kąta Eulera ψ
w_θ	waga wartości kąta Eulera θ
EKF	<i>ang. Extended Kalman Filter</i> – Rozszerzony filtr Kalmana
GPS	<i>ang. Global Positioning System</i> – globalny system pozycjonowania

- IMU *ang. Inertial Measurement Unit* – inercyjne jednostki pomiarowe: akcelerometr i żyroskop
- KF Liniowy filtr Kalmana
- LPM *ang. Local Position Measurement System* – lokalny system pozycjonowania
- MARG *ang. Magnetic, Angular Rate, and Gravity* – inercyjne jednostki pomiarowe: akcelerometr i żyroskop wspierane przez magnetometr
- MoCap *ang. Motion Capture* – system śledzenia i rejestrowania ruchu
- UKF *ang. Unscented Kalman Filter* – Bezśladowy filtr Kalmana

Rozdział 1

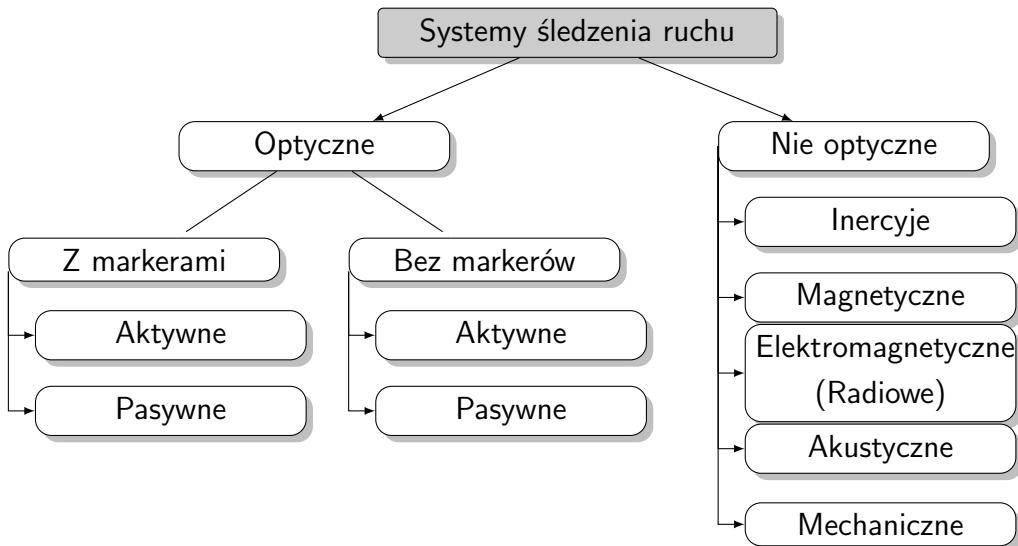
Wstęp

Śledzenie ruchu człowieka w systemach komputerowych jest rozumiane jako proces obliczeniowy, którego celem jest jednoznaczne i pozbawione znaczących opóźnień szacowanie położenia stawów układu szkieletowego człowieka w przestrzeni i w czasie. Dysponując położeniem stawów w czasie i modelem układu kostnego człowieka można oszacować kolejne pozy postaci oraz ich zmianę - ruch postaci.

Jednym z pionierów śledzenia ruchu człowieka był psycholog Gunnar Johansson. W roku 1973, w ramach eksperymentu o nazwie *Moving Light Display (MLD)*, stworzył on system śledzenia ruchu człowieka bazujący na kamerze video rejestrującej ruch odbijających światło markerów, przyczepionych do ciała śledzonej postaci [53]. System, zaproponowany wówczas na potrzeby eksperymentu prowadzonego przez Johanssoną, w formie udoskonalonej jest stosowany z powodzeniem współcześnie. Technologia pierwotnie wykorzystywana do badania i analizy ruchu kończyn, z czasem wyewoluowała w stronę zaawansowanych systemów używanych przez profesjonalistów do tworzenia animacji, czy analizy ruchu w zastosowaniach sportowych (trening sportowy) oraz medycznych (telerehabilitacja). Jest popularna również wśród amatorów, którzy w zaciszu swoich domów stosują ją do interakcji z komputerem, między innymi w grach oraz w środowiskach wirtualnych. Widoczny jest również trend wykorzystania systemów śledzenia ruchu do budowy Naturalnych Interfejsów Użytkownika[38] (NUI - ang. *Natural User Interface*).

Ponad 40 lat prowadzenia prac i badań związanych z **systemami śledzenia ruchu**, pozwoliło na opracowanie różnorodnych technik rejestracji danych dotyczących wykonywanego ruchu. Przełożyło się to na stworzenie wielu rodzajów systemów śledzenia ruchu począwszy od systemów wizyjnych wspieranych markerami oraz tzw. bezmarkerowych, systemów inercyjnych i radiowych, aż po zaawansowane egzo-szkielety śledzące ruch człowieka w sposób mechaniczny. Stosowana w systemach śledzenia ruchu technologia

obserwacji ruchu śledzonej osoby stanowi jedno z kryteriów podziału systemów śledzenia ruchu. Diagram przedstawiony na rysunku 1.1 przedstawia taksonomię systemów śledzenia ze względu na kryterium użytej technologii.



Rysunek 1.1: Podział systemów śledzenia ruchu ze względu na technologię pobierania danych (źródło: opracowanie własne)

Prowadzone są również badania nad systemami **hybrydowymi**, które łączą ze sobą kilka rodzajów urządzeń pomiarowych mogących się nawzajem uzupełniać w czasie śledzenia ruchu, co ma na celu zmniejszenie ograniczeń ich działania i zwiększenie dokładności oszacowania położenia poszczególnych stawów.

Niniejsza praca poświęcona została hybrydowemu systemowi śledzenia ruchu, łączącemu dane uzyskane za pomocą sensora głębi, umieszczonego w kontrolerze Microsoft Kinect, z danymi uzyskanymi z **autonomicznych czujników inercyjnych** działających w ramach samodzielnie stworzonych urządzeń pomiarowych, bazujących na platformie Arduino. Autonomicznymi czujnikami inercyjnymi nazywamy moduły mierzące, za pomocą wbudowanego żyroskopu, prędkość kątową, z jaką obraca się dany komponent oraz, za pomocą wbudowanego akcelerometru, siły (i przyspieszenia liniowe) na niego działające. Dzięki odniesieniu działających sił do siły grawitacji, można określić przybliżone przyspieszenie akcelerometru. Zazwyczaj akcelerometr i żyroskop połączone są ze sobą jako jedno urządzenie pomiarowe (moduł IMU – ang. *Inertial Measurement Unit*) umieszczone w niewielkim pudełku. W rozważanym systemie moduły IMU zostały umieszczone na ciele (kończynach) śledzonej postaci, jako urządzenia pomiarowe uzupełniające kontroler Kinect. W dalszej części pracy, moduły IMU, umieszczone na kończynach śledzonej postaci, będą nosić nazwę **modułów inercyjnych**.

1.1. Problematyka pracy

Głównym problemem z jakim zmagają się twórcy systemów śledzenia ruchu człowieka jest precyzyjne oszacowanie położenia stawów szkieletu postaci w przestrzeni. Precyza systemu śledzenia ruchu postaci uzależniona jest w dużej mierze od charakterystyk działania wykorzystywanych w nim urządzeń pomiarowych. W przypadku opisywanego w niniejszej pracy hybrydowego systemu śledzenia ruchu, wykorzystującego kontroler Kinect oraz czujniki inercyjne, problemami, na które należało zwrócić uwagę są między innymi: dokładność urządzeń pomiarowych oraz ograniczone możliwości śledzenia kończyn w przypadku przesłaniańia się części ciała - główny problem kontrolera Kinect. Dodatkowo istotnym problemem jest zasumienie pomiarów uzyskiwanych z czujników inercyjnych oraz łączenie danych uzyskanych ze wszystkich urządzeń pomiarowych w taki sposób, aby uzyskane oszacowanie położenia stawów w przestrzeni było jak najbardziej precyzyjne w odniesieniu do rzeczywistego ich położenia.

Jedną z inspiracji dla niniejszej pracy i jednocześnie przykładowym zastosowaniem tworzonego systemu była pomoc przy rehabilitacji ruchowej. Zadaniem systemu byłoby umożliwienie pacjentowi wykonywania ćwiczeń we własnym domu z zapewnieniem weryfikacji poprawności wykonywanych ruchów. Według Reinkensmeyera i Boningera, cechą charakterystyczną takich ćwiczeń jest ich wolne tempo, powiązane z rygorem dokładnego ich wykonywania [91]. Ci sami autorzy wskazują, że rehabilitacja wspomagana systemem śledzenia ruchu pozwala na zindywidualizowanie ćwiczeń dla konkretnego pacjenta. Z kolei Malouin i in. [79] wskazują, że systemy pozwalające na zautomatyzowany nadzór nad wykonywanymi ćwiczeniami rehabilitacyjnymi, pozwalają na ich zintensyfikowanie oraz na przyspieszenie ich wdrożenia w terapię, co ma pozytywny wpływ na powrót pacjenta do pełni zdrowia.

1.2. Cel i teza pracy

Celem niniejszej pracy było opracowanie metody połączenia sygnałów z wybranych urządzeń pomiarowych: kontrolera Kinect i czujników inercyjnych, która działa z częstotliwością porównywalną do częstotliwości pracy kontrolera Kinecta (30 Hz) i zapewnia wysoką, wyższą niż u innych autorów, dokładność oszacowania położenia śledzonych stawów układu kostnego człowieka. W konsekwencji śledzenie wykonywanych przez postać ruchów byłoby bardziej precyzyjne. Ruchy kończyn powinny być realizowane w tempie charakterystycznym dla ćwiczeń rehabilitacyjnych, aby zastosowanie zaproponowanych urządzeń pomiarowych było zasadne. Gdyby prędkość ruchu

kończyn przekroczyła zdolność rejestracji któregoś z wybranych urządzeń pomiarowych, należałyby zastąpić je urządzeniami o wyższej częstotliwości rejestracji pomiarów. Nie ogranicza to jednak zastosowania metody do łączenia danych pochodzących z innych urządzeń, pod warunkiem, że charakter zwracanych danych pomiarowych będzie analogiczny.

W pracy została postawiona następująca teza:

Zastosowanie autorskiej, hybrydowej metody śledzenia ruchu kończyn człowieka, łączącej dane pochodzące z sensora głębi i sensorów inercyjnych, uwzględniającej kontekstowe charakterystyki pracy urządzeń, pozwala na bardziej precyzyjne śledzenie ruchu.

W ramach pracy została opracowana i zaimplementowana autorska metoda przetwarzania i łączenia ze sobą sygnałów z inercyjnych czujników ruchu oraz kontrolera Kinect. Metoda wykorzystuje informacje o obrotach segmentów kończyn ciała aktora i, w połączeniu z oszacowanym modelem układu kości, pozwala na obliczenie pozycji stawów układu szkieletowego człowieka. Jako przykład został wybrany i zbadany ruch prawej ręki bez śledzenia ruchu dloni. Metoda ta została porównana z autorską implementacją hybrydowej metody śledzenia ruchu o najwyższej deklarowanej w literaturze dokładności, i wykorzystującej analogiczne urządzenia pomiarowe.

1.3. Struktura pracy

Niniejsza dysertacja została podzielona na 5 rozdziałów. Poza wstępem, kolejne rozdziały dotyczą: analizy aktualnego stanu zagadnienia, przedstawienia autorskiej hybrydowej metody śledzenia ruchu, a także weryfikacji skuteczności opracowanego rozwiązania. Całość pracy zwieńczona jest rozdziałem podsumowującym przedstawione zagadnienia.

Analiza aktualnego stanu zagadnienia (rozdział 2) zawiera omówienie istniejących systemów śledzenia ruchu (rozdział 2.1), sposobów reprezentacji ciała człowieka w systemach komputerowych (rozdział 2.2), analizę charakterystyk pracy i budowy kontrolera Microsoft Kinect oraz czujników inercyjnych (rozdział 2.3), a także analizę istniejących w literaturze realizacji hybrydowych systemów śledzenia ruchu człowieka, wykorzystujących te same lub analogiczne urządzenia pomiarowe (rozdział 2.4). Należy w tym miejscu zaznaczyć, że opisane w pracy charakterystyki kontrolera Kinect i czujników

inercyjnych, zostały w znacznej mierze wyznaczone na bazie własnych badań oraz na podstawie wiedzy dostępnej w literaturze.

Kolejny rozdział (rozdział 3), przedstawiający autorską hybrydową metodę śledzenia ruchu, wykorzystującą kontroler Microsoft Kinect i czujniki inercyjne, zawiera omówienie kolejnych etapów przetwarzania i łączenia danych uzyskanych z urządzeń pomiarowych. Kolejne podrozdziały zostały poświęcone akwizycji danych pomiarowych i ich formatom, dyskusji dotyczącej kwestii kalibracji urządzeń oraz etapom korekcji pozyskanych danych, ich synchronizacji czasowej oraz ich finalnemu łączeniu mającemu na celu oszacowanie położenia wybranych stawów ręki.

Rozdział 4 zawiera opis badań eksperymentalnych wraz z wynikami i ich omówieniem. Kolejny rozdział zawiera podsumowanie dysertacji wraz z analizą dalszych kierunków badań związanych z omawianą metodą.

Praca zawiera także 2 rozdziały dodatkowe, mające na celu przedstawienie informacji pomocnych w analizie przedstawionego tematu. Pierwszym z nich (dodatek A) jest opis podstawowych filtrów stosowanych do łączenia sygnałów, a wykorzystywanych w przytoczonych w niniejszej pracy artykułach (filtry Kalmana: liniowy [58], rozszerzony [108] i bezśladowy [55], a także filtr komplementarny [32] oraz filtry Mahoney'a [11] i Madgwicka [75]). Drugim (dodatek B) jest przybliżenie wariancji Allana [4] będącej narzędziem pozwalającym na analizę charakterystyki szumów, między innymi czujników inercyjnych.

Rozdział 2

Przegląd metod śledzenia ruchu

Śledzenie ruchu ma zastosowanie zarówno w przypadku istot żywych (ludzi i zwierząt), jak i w przypadku obiektów nieożywionych. W zależności od budowy śledzonego obiektu można stosować różne techniki i narzędzia. Za pomocą innych technik będzie śledzony ruch człowieka, który charakteryzuje się specyficzną dla ludzi kinematyką i dynamiką zmian ułożenia ciała, a inaczej będzie śledzony fizyczny przedmiot, który z jednej strony może podlegać fizycznie ograniczonemu ruchowi (na przykład ruch wystrzelonego pocisku w polu grawitacyjnym), a z drugiej strony może być immanentnie wyposażony w pomocnicze, autonomiczne urządzenia pomiarowe (na przykład rakieta wyposażona w GPS, akcelerometr i żyroskop).

W doborze techniki śledzenia ruchu istotną rolę odgrywa także określenie celu oraz pożąданie dokładności procesu śledzenia. Jeśli celem jest określenie położenia poszczególnych stawów ludzkiego ciała, wówczas niezbędne będzie zastosowanie systemów śledzenia ruchu postaci, które gwarantują dużą dokładność szacowania pozycji stawów. Jeśli istotne jest tylko zgrubne określenie położenia całej postaci w jej otoczeniu, wówczas wystarczy wykorzystanie systemu pozycjonowania GPS, przyczepionego do korpusu śledzonej osoby.

Innym istotnym aspektem, który determinuje zastosowanie techniki śledzenia ruchu, jest wykorzystanie konkretnych (dostępnych) urządzeń pomiarowych. Charakterystyka urządzeń pomiarowych wpływa na dobór metod analizy danych zwracanych przez urządzenia pomiarowe i końcową dokładność śledzenia zmian obiektu. Istotnym czynnikiem wpływającym na konstrukcję i oprogramowanie systemu śledzenia jest przewidywany zakres przedkości wykonywanego ruchu, który musi zostać zarejestrowany i zinterpretowany przez system pomiarowy. Przykładem ruchów człowieka o skrajnie różnej dynamice mogą być ćwiczenia rehabilitacyjne kończyn po urazach oraz dynamika ręki tenisisty podczas serwisu.

W literaturze przedmiotu można znaleźć wiele opisów prowadzonych prac badawczych, które bazując na charakterystyce śledzonego obiektu lub dostępnych urządzeniach pomiarowych próbują pokonać trudności związane

z niedostateczną dokładnością śledzenia ruchu, bądź niedostateczną czułością i bezwładnością układu pomiarowego w przypadku dużych prędkości zmian śledzonego obiektu. Istnieje też szereg prac, które stosując heterogeniczne systemy śledzenia, kompensują niedoskonałości składowych układów śledzących. Przegląd wybranych publikacji, sprofilowany głównie na śledzenie ruchu postaci ludzkiej, został przedstawiony w kolejnych podrozdziałach.

2.1. Systemy śledzenia ruchu postaci

Śledzenie ruchu postaci jest procesem obliczeniowym mającym na celu jednoznaczne określenie w czasie i w przestrzeni położenia i orientacji układu punktów (stawów) opisujących pozycję człowieka.

Jednym z kryteriów podziału systemów śledzenia ruchu człowieka jest jego zastosowanie. Głównymi obszarami zastosowań dyskutowanych systemów śledzenia są: nadzór, kontrola oraz analiza ruchu ludzkiego ciała [82]. Pod pojęciem nadzoru rozumiemy systemy śledzące zachowanie człowieka i na jego podstawie ocenяjące na przykład czy dana osoba potrzebuje pomocy [43, 61, 67]. Obszar zastosowań związanych z kontrolą odnosi się do interakcji człowiek–komputer, na przykład: w grach komputerowych, środowiskach wirtualnych, czy animacji [81]. Ostatni z wymienionych obszarów zastosowań - analiza, obejmuje takie zagadnienia jak diagnostyka i rehabilitacja medyczna^{1,2}, czy wspomaganie treningu sportowego³ [85, 87].

Innym kryterium podziału systemów śledzenia ruchu człowieka jest zastosowana w nich technologia, pozwalająca obserwować śledzoną osobę. Schematykę tej przedstawia diagram na rysunku 1.1 na stronie 12. Kolejne części niniejszego rozdziału zostały poświęcone przybliżeniu istotnych cech każdego z systemów śledzenia ruchu przedstawionych na tym diagramie.

2.1.1. Optyczne systemy śledzenia ruchu

Optyczny system śledzenia ruchu to system, w którym dane o ruchu obiektu pozyskuje się za pośrednictwem toru optycznego uzbrojonego w czujniki optyczne (na przykład: kamery video, sensory podczerwieni i in.). Jeśli informacje o ruchu można pozyskać na podstawie zbioru kontrolowanych

¹ Vicon Corporation, *Life science case study MOTEK medical + virtual rehabilitation*, <https://goo.gl/EIAoki>.

² Xsens Motion Technologies, *Interactive sonification of human movements for stroke rehabilitation*, <https://goo.gl/2CfIvn>.

³ Inmotio Corp, *Local Position Measurement (LPM) technology is the world's most precise sports tracking system*. <https://goo.gl/IVjiAQ>.

punktów (markerów) identyfikujących śledzony obiekt to takie systemy nazywamy systemami śledzenia wykorzystującymi markery (ang. *marker-based tracking system*). Przykładem tego typu systemów śledzenia ruchu mogą być produkty firm Vicon⁴, Qualisys⁵, czy Optitrack⁶. Istnieją również systemy bezmarkerowe (ang. *markerless tracking system*), które fizyczne markery umieszczone na śledzionym obiekcie, zastępują punktami wyznaczanymi na podstawie cech charakterystycznych danych obiektów (na przykład dla ciała ludzkiego mogą to być punkty reprezentujące stawy). W takim przypadku, strumień danych otrzymany z czujników zostaje poddany analizie i przetwarzaniu w celu wyznaczenia punktów, które pozwalają śledzić ruch obiektu. Przykładami bezmarkerowych optycznych systemów śledzenia ruchu mogą być komercyjne systemy śledzenia ruchu firmy Organic Motion⁷, czy systemy tworzone na uniwersytetach Stanforda⁸ i Maryland [112, 113].

Optyczne systemy śledzenia ruchu zbudowane są zazwyczaj z dwóch rodzajów urządzeń. Pierwszym z nich jest jedno lub wiele źródeł światła na przykład: kierunkowego, odbitego czy ustrukturyzowanego, a drugim jest jeden, lub wiele czujników optycznych. Najpopularniejszymi systemami optycznymi wykorzystywanymi do śledzenia ruchu człowieka są te, oparte o analizę obrazów rejestrowanych przez jedną [71, 99] lub układ wielu kamer^{4,9} [66, 113].

Optyczne systemy śledzenia ruchu wykorzystujące markery

Z punktu widzenia analizy optycznych systemów śledzenia ruchu wykorzystujących markery najistotniejszą cechą rozróżniającą stosowane markery jest to czy są one emiterem światła, czy jedynie je odbijają. Jeśli mamy do czynienia z markerami będącymi równocześnie źródłem światła wówczas nazywamy je markerami aktywnymi, natomiast markery odbijające światło z zewnętrznego źródła są nazywane markerami pasywnymi.

Markery pasywne, wykorzystywane w optycznych systemach śledzenia ruchu, mogą różnić się swoim kształtem, wielkością, czy budową, w zależności

⁴ Vicon Corporation, *Motion Capture Systems — VICON*, <https://goo.gl/DQ2dmr>

⁵ Qualisys Corporation, *Qualisys — Motion Capture Systems*, <https://goo.gl/nNzzHb>.

⁶ OptiTrack Corporation, *Motion capture systems - OptiTrack*, <https://goo.gl/1hEBtp>.

⁷ Organic Motion Corporation, *Markerless motion capture software by Organic Motion*, <https://goo.gl/jeA3t>.

⁸ Stanford University, *BioMotion Laboratory — Markerless Motion Capture System*, <https://goo.gl/zXUfpv>.

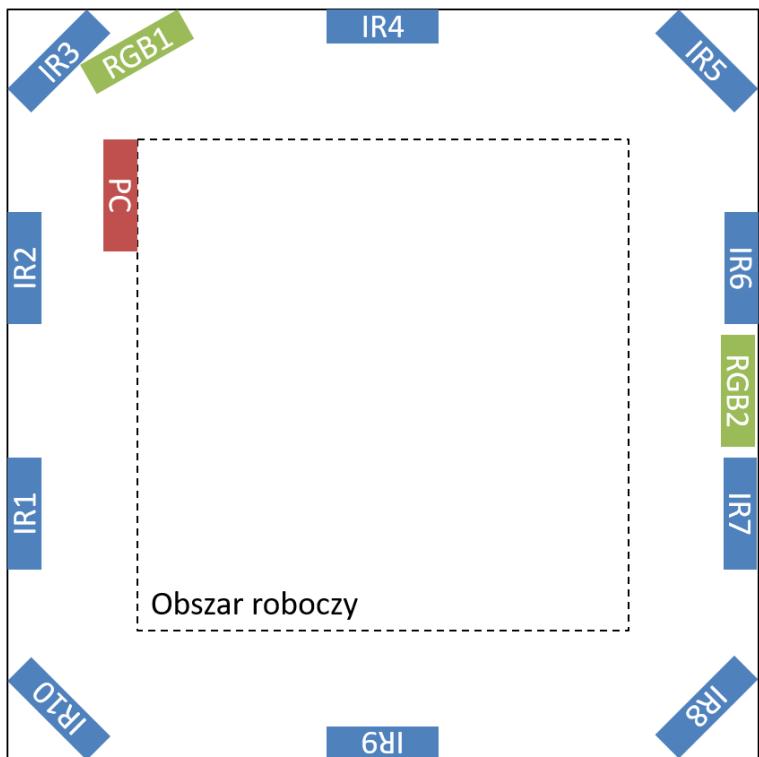
⁹ Qualisys Corporation, *Qualisys — Motion Capture Systems*, <https://goo.gl/nNzzHb>.

od ich przeznaczenia i oczekiwanej dokładności. Różnice w wielkości markerów można prześledzić na przykładzie markerów pasywnych, które wykorzystywane do śledzenia mimiki twarzy (problem wymagający dużej precyzji) mają średnicę od około 3 do około 5 mm, natomiast gdy mają być wykorzystywane do śledzenia położenia całej sylwetki (problem wymagający mniejszej precyzji, lecz rejestracji markera z większej odległości) ich średnica wynosi zazwyczaj około 25mm¹⁰. Systemy śledzenia ruchu oparte o markery pasywne opierają się dokładnie na pomyśle zaproponowanym przez G. Johanssona we wspomnianym wcześniej eksperymencie *MLD*. Jako współczesne realizacje optycznych systemów śledzenia ruchu stosujących markery pasywne można wymienić wspomniane już wielokamerowe profesjonalne systemy śledzenia firm Vicon, Qualisys, czy OptiTrack. Zamieszczony w dalszej części pracy opis budowy i zasad działania optycznego systemu śledzenia ruchu z pasywnymi markerami zostanie przedstawiony na podstawie systemu firmy Vicon¹¹ zainstalowanego na Politechnice Łódzkiej. Analizowany system firmy Vicon jest systemem wielokamerowym, który wykorzystuje wyspecjalizowane kamery cyfrowe video, o wysokiej rozdzielczości matrycy (4 MPix dla rozważanego systemu), wyposażone w filtr przepuszczający zakres światła podczerwonego. Każda z kamer wyposażona jest równocześnie w kierunkowe źródło światła emitujące fale z zakresu światła podczerwonego. System Vicon, dostępny na Politechnice Łódzkiej, zbudowany jest z 10 takich kamer umieszczonej dookoła pomieszczenia laboratoryjnego tak jak jest to pokazane na rysunku 2.1, gdzie elementy oznaczone IR1..10 to kamery wyposażone w filtr podczerwony, RGB1 i RGB2 to dwie kamery video pozwalające rejestrować tradycyjny obraz kolorowy oraz stanowisko komputerowe oznaczone jako PC. System posiada swoją funkcjonalność w tzw. przestrzeni roboczej śledzenia (ang. *tracking volume*). Jest to przestrzeń wyznaczona przez bryły widzenia kamer. Technicznie każdy marker w przestrzeni roboczej powinien być widoczny jednocześnie przez trzy kamery, aby możliwe było precyzyjne określenie jego pozycji. Ruchy wykonywane poza przestrzenią roboczą nie są śledzone przez system lub obarczone są istotnymi błędami pomiarowymi.

Kamery wykorzystane w systemie Vicon posiadają filtr przepuszczający jedynie fale z zakresu światła podczerwonego, co pozwala na stosunkowo łatwe wyodrębnienie z rejestrowanego obrazu markerów odbijających światło podczerwone. Markery są elementami pokrytymi farbą odbijającą padającą na nie światło w kierunku źródła emisji promieni. Ponieważ w przypadku zastosowanych kamer źródło światła jest umieszczone tuż obok obiektywu,

¹⁰ Vicon Corporation, *Markers and suits — VICON*, <https://goo.gl/Gw1P63>

¹¹ Wybór ten został podektywany dostępem do takowego systemu w Centrum Technologii Informatycznych Politechniki Łódzkiej.



Rysunek 2.1: Rysunek przedstawiający konfigurację sprzętową systemu Vicon dostępnego w Centrum Technologii Informatycznych Politechniki Łódzkiej. IR1..10 – kamery z filtrem podczerwonym, RGB1, RGB2 – kamery rejestrujące obraz kolorowy, PC – stanowisko komputerowe ze specjalistycznym oprogramowaniem do obsługi systemu śledzenia (źródło: opracowanie własne)

oświetlone markery stanowią najjaśniejsze punkty na rejestrowanym obrazie. Na rysunku 2.3 zaprezentowany został przykładowy zestaw markerów wykorzystywanych w omawianym systemie śledzenia ruchu, natomiast na rysunku 2.2 widać aktora w kombinezonie, przy którym owe markery są przymocowane. Markery wykorzystywane są do oznaczenia części ciała człowieka, których ruch będzie śledzony.

Oprócz kamer i markerów, istotnym elementem optycznego systemu śledzenia ruchu jest oprogramowanie obsługujące system. Jest ono odpowiedzialne za synchroniczną analizę sekwencji obrazów zarejestrowanych przez kamery tak, aby możliwe było określenie współrzędnych każdego z markerów w przestrzeni trójwymiarowej układu odniesienia. Aby możliwe było wyznaczenie położenia przestrzennego wybranego markera, musi być on widziany przez minimum trzy kamery w tym samym czasie. Pozwala to na precyzyjne

¹² Qualisys Corporation, *Super-spherical markers* — Qualisys, <https://goo.gl/aAUZYu>.



Rysunek 2.2: Zestaw znaczników pasywnych w systemie Vicon umieszczonych na ciele aktora¹⁰



Rysunek 2.3: Przykładowe pasywne markery wykorzystywane w optycznych systemach śledzenia ruchu¹²

wyznaczenie jego wszystkich współrzędnych za pomocą triangulacji. Jednakże zbudowanie profesjonalnego systemu śledzenia ruchu złożonego jedynie z trzech kamer byłoby niewystarczające ze względu na możliwość przysłonięcia markerów przez części ciała w trakcie wykonywania ruchu. Zwiększenie liczby kamer pozwala na obserwowanie sceny przez większą ich liczbę, co zwiększa prawdopodobieństwo, że w trakcie nawet złożonych ruchów przynajmniej trzy kamery zarejestrują każdy z markerów. Dodatkowo redundancja danych, dla lepiej widocznych markerów, sprzyja zwiększeniu dokładności wyznaczania ich pozycji. Oczywiście może się zdarzyć, że pomimo zastosowania wielu kamer marker zostanie zauważony przez mniej niż trzy kamery. W takiej sytuacji rolą oprogramowania jest oszacowanie przybliżonego położenia danego markera, na przykład na podstawie pozycji innych markerów i ich trajektorii ruchu.

Kolejnym ważnym elementem, za który odpowiada oprogramowanie systemu śledzenia ruchu jest etykietowanie (*ang. labelling*) śledzonych markerów. Dzięki temu możliwe jest określenie wzajemnych relacji pomiędzy poszczególnymi markerami i nadanie im dodatkowego znaczenia poza wyznaczeniem położenia w przestrzeni, na przykład: powiązania pomiędzy markerami, przypisanie kilku markerów do jednego obiektu na scenie. Proces ten, po wcześniejszym zdefiniowaniu wzajemnych relacji pomiędzy markerami, jest zazwyczaj automatyzowany i wymaga korekty ze strony operatora systemu jedynie w sytuacjach kiedy nie jest możliwa pełna rekonstrukcja sceny, na przykład gdy zbyt wiele markerów jest niewidocznych dla kamer i oprogramowanie nie jest w stanie oszacować ich położenia, a co za tym idzie niemożliwe jest prawidłowe przypisanie zbioru etykiet.

Rozpatrując śledzenie ruchu człowieka, istotne jest aby oprogramowanie było w stanie kompensować różnicę między ruchem markerów jaki został zarejestrowany, a ruchem postaci jaki został faktycznie wykonany. Szczególnie jest to widoczne w przypadku obrotów kończyn wokół własnej osi. Aby jak najlepiej odzwierciedlić taki ruch, markery powinny być umieszczone bezpośrednio na ciele w miejscach nie wrażliwych na przesunięcia względem układu kostnego człowieka - w pobliżu kości, w miejscach pozbawionych tkanki mięśniowej. Jednak i to nie gwarantuje, że markery umieszczone na powierzchni ciała wykonają taki sam ruch jak kończyna. Badania mające na celu określenie różnicy pomiędzy ruchem jaki wykonywany jest przez kość lub staw, a tym jak jest on widoczny na podstawie markerów umieszczonych na powierzchni skóry, prowadzone były już w latach 90-tych XX wieku [48, 92, 98]. Reinschmidt i in. [92] na przykładzie badania zgęcia kolana wykazał, że średnia różnica pomiędzy faktycznym ruchem kości, a ruchem zarejestrowanym przez markery może wynosić od 4.1° do 5.3° w zależności od osi, w której ruch ten się odbywał. Sati i in. [98] oraz Holden i in. [48] uzyskali zbliżone rezultaty dodatkowo zauważając, że markery mogą ulec przesunięciu w trakcie wykonywania ruchu nawet o $2cm$ [98], co również może wypaczyć uzyskane wyniki.

Optyczne systemy śledzenia ruchu z aktywnymi markerami działają i są zbudowane w analogiczny sposób do systemów z markerami pasywnymi. Podstawową różnicą pomiędzy tymi dwoma systemami śledzenia ruchu jest to, że aktywne markery są źródłem światła, co eliminuje konieczność wyposażenia całego systemu w dodatkowe urządzenia do oświetlania sceny. Przykładem realizacji takiego systemu może być wielokamerowy system Impulse X2 firmy PhaseSpace¹³.

Niewątpliwą zaletą optycznych systemów śledzenia z markerami jest ich wysoka dokładność szacowania pozycji śledzonych markerów przy jednoczesnej swobodzie wykonywania ruchów. Oficjalna specyfikacja systemu Vicon, a także niezależne badania wykorzystujące ten system, pokazują, że jest on w stanie oszacować położenie markerów z dokładnością $\pm 0.5mm$, a ich obrót z dokładnością $\pm 0.5^\circ$ ¹⁴ [125]. Dzięki tej precyzji wykorzystywane są one zarówno w przemyśle rozrywkowym, na przykład przy tworzeniu animacji postaci na potrzeby gier i filmów, jak i w analizie ruchu na potrzeby sportu, czy medycyny¹⁵.

¹³ PhaseSpace Corp, *PhaseSpace Motion Capture — Products — Impulse X2 Motion Capture Solution*, <https://goo.gl/8dHEur>.

¹⁴ Vicon Corporation, *Vantage Motion Capture Camera — VICON*, <https://goo.gl/NRBDC7>.

¹⁵ Vicon Corporation, *Life science case study MOTEK medical + virtual rehabilitation*, <https://goo.gl/EIAoki>.

Większość oferowanych obecnie komercyjnych optycznych systemów śledzenia ruchu pracuje standardowo z częstotliwością 100 Hz – 160 Hz, jednak możliwe jest także używanie ich w trybie 500Hz i więcej. Zwiększenie częstotliwości pomiarów wiąże się zazwyczaj z obniżeniem ich dokładności. Główną wadą omawianych systemów jest ich cena, która sprawia, że są one praktycznie niedostępne dla użytkowników chcących zbudować taki system śledzenia ruchu w warunkach domowych. Wystarczy wspomnieć, że koszt zakupu systemu śledzenia ruchu firmy Vicon złożonego z 10 kamer z filtrem podczerwonym, kupionego dla Centrum Technologii Informacyjnych Politechniki Łódzkiej, wyniósł w 2014 roku około 1000 000 zł.

Optyczne systemy śledzenia ruchu nie wymagające markerów

Drugim rodzajem optycznego systemu śledzenia ruchu są systemy nie wymagające zastosowania markerów. Bezmarkerowe systemy także możemy podzielić na systemy aktywne i pasywne [84]. Aktywne systemy stosują emisję światła (najczęściej ustrukturyzowanego), widzialnego bądź podczerwonego, oświetlającego śledzone obiekty, natomiast pasywne rejestrują jedynie obrazy z kamer wideo i rozpoznają ruch na podstawie analizy zawartości obrazów.

Systemy aktywne wykorzystują zróżnicowane techniki emisji światła do określenia odległości pomiędzy urządzeniem pomiarowym danego systemu a obserwowanym obiektem. Jedną z technik jest oświetlenie sceny ustrukturyzowaną mapą punktów świetlnych, której odkształcenie pozwala określić, w jakiej odległości przed kamerą znajdują się dane obiekty. Technika ta jest wykorzystywana między innymi przez kontroler Microsoft Kinect v1¹⁶, czy kamerę 3D systemu Intel RealSense¹⁷. Inną popularną techniką określania odległości, w jakiej znajduje się obserwowany obiekt, wykorzystywaną w kontrolerze Microsoft Kinect v2¹⁸, jest czas dotarcia wiązki światła do tego obiektu (*ang. Time-of-Flight*)[42].

Zarówno w systemach aktywnych jak i pasywnych kluczowym etapem przetwarzania danych jest wyodrębnienie pierwszego planu i oddzielenie go od tła. Metody wykorzystywane na tym etapie są często opisywane w literaturze [22, 41, 94, 114, 123]. Po separacji pierwszego planu zawierającego obraz śledzonej postaci od tła, kolejnym etapem jest estymacja modelu człowieka, który jest jego reprezentacją w systemie komputerowym.

Optycznym, bezmarkerowym i zarazem aktywnym systemem śledzenia ruchu, który niewątpliwie odniósł największy sukces komercyjny i posiada

¹⁶ J. L. Flatley, *Visualized: Kinect + night vision = lots and lots and lots of dots (video)*, <https://goo.gl/HuzK3>.

¹⁷ Intel Corporation, *RealSense - develop immersive experiences — Intel® Software*, <https://goo.gl/yuV8Yk>.

¹⁸ Microsoft Corporation, *Kinect hardware*, <https://goo.gl/zYa6Y0>.

największą rozpoznawalność jest system wbudowany w kontroler Microsoft Kinect. W dalszej części niniejszej pracy (rozdział 2.3) szczegółowo zostało opisane działanie kontrolera Microsoft Kinect v1 z uwagi na fakt, że opracowana autorska metoda śledzenia ruchu wykorzystuje to właśnie urządzenie.

Systemy śledzenia ruchu, zaimplementowane w każdej z dwóch wersji kontrolera Microsoft Kinect, zostały domyślnie zaprojektowane do śledzenia ruchów wykonywanych przez człowieka. Wykorzystywanie ich na przykład do śledzenia ruchu zwierząt lub przedmiotów wymaga ingerencji w proces przetwarzania sygnałów realizowany przez wspomniane kontrolery [86].

System śledzenia ruchu człowieka, zaimplementowany w kontrolerze Microsoft Kinect, w takiej konfiguracji w jakiej dostarcza go producent, jest systemem do zastosowań rozrywkowych dla użytkownika domowego. Istnieją jednak bezmarkerowe systemy śledzenia ruchu człowieka adresowane do zastosowań profesjonalnych. Jako przykład może posłużyć system *Organic Motion* zaprezentowany na konferencji SIGGRAPH w 2012 roku. Jest to system wielokamerowy (8–18 urządzeń), który rekonstruuje przesunięcia obiektu 3D w czasie na podstawie triangulacji wspólnych punktów rozpoznanych na zsynchronizowanych obrazach 2D [18].

W swoim artykule Brooks i Czarowicz [18] porównują podstawowe cechy systemów *Organic Motion* oraz systemu wbudowanego w kontroler Microsoft Kinect v1. W tabeli 2.1 znajduje się skrócone porównanie tych dwóch systemów.

Tablica 2.1: Porównanie cech charakterystycznych systemu *Organic Motion* oraz systemu wbudowanego w kontroler Microsoft Kinect v1 (na podstawie [18])

	Kinect	Organic Motion
Częstotliwość	30Hz	30Hz – 120Hz
Wymiarowanie kości szkieletu	Szacowanie Zmienne w czasie	Dokładny pomiar Stał w czasie
Śledzenie stóp	Brak	Jest
Śledzenie palców	Brak	Brak
Śledzenie 360°	Brak	Jest
Opóźnienie	250ms	50–100ms
Rozdzielcość	1 kamera 640x480px	8 – 18 kamer 640x480px każda

Cechy przedstawione w tabeli 2.1 pozwalają zauważyc kilka istotnych różnic pomiędzy porównywanyymi systemami, które mają znaczący wpływ na

działanie i obszary możliwych zastosowań dla każdego z nich. Warto zauważyc, że system oparty o kontroler Microsoft Kinect v1, w konfiguracji zalecanej przez jego producenta, opiera się o pojedyncze urządzenie rejestrujące obraz z rozdzielcością 640x480px, co sprawia, że jesteśmy pozbawieni możliwości śledzenia małych ruchów wykonywanych na przykład przez palce, czy stopy. Są to ruchy zbyt szczegółowe by można było je skutecznie zarejestrować i wyodrębnić na obrazie o tak niskiej rozdzielcości. Dodatkowo, pojedyncze urządzenie sprawia, że ruch obserwowany jest tylko z jednej płaszczyzny, więc tracimy informację o ruchu jaki wykonywany jest na przykład za plecami użytkownika. W przypadku systemu *Organic Motion* wykorzystanie większej liczby kamer daje możliwość pełnej obserwacji dookoła użytkownika w 360° co oznacza, że model postaci otrzymany w systemie komputerowym, przedstawia sylwetkę obserwowaną z dowolnego miejsca w przestrzeni. Obraz uzyskany z wielu kamer daje sumarycznie większą rozdzielcość co pozwala na zarejestrowanie szczegółowych ruchów. Daje także możliwość dokładnego, stałego w czasie, wymiarowania poszczególnych elementów modelu ciała człowieka (w szczególności długości jego kości) zamiast przybliżonego ich szacowania na podstawie chwilowych danych, co ma wpływ na częstą zmienność tych pomiarów. Wszystko to pokazuje, że o ile w systemach rozrywkowych, gdzie nie ma konieczności uzyskiwania stabilnych i dokładnych pomiarów oraz śledzenia ruchu, kontroler Microsoft Kinect jest wystarczający, to aplikacje związane na przykład z biomechaniką wymagają rozbudowanych systemów wielokamerowych.

Przedstawione powyżej systemy wizyjne, zarówno te wykorzystujące markery jak i bezmarkerowe, osiągają wysoką dokładność śledzenia ruchu w konfiguracjach wielokamerowych umieszczonych wewnętrz pomieszczenia, gdzie system nie jest podatny na działanie światła słonecznego. W przypadku śledzenia ruchu wykonywanego w innych warunkach niż studyjne, stosuje się systemy oparte o inne techniki niż optyczne, na przykład rejestruje się prędkości śledzonych części ciała w trakcie ruchu za pomocą czujników inercyjnych.

2.1.2. Systemy nieoptyczne

Systemy inercyjne i magnetyczne

Systemy inercyjne oraz systemy magnetyczne opierają swoje działanie na pomiarach wielkości fizycznych, takich jak przyspieszenie czy prędkość, które występują w trakcie wykonywania ruchu. Wykorzystują one moduły zbudowane z czujników pozwalających na pomiar wspomnianych wielkości, jakie występują w momencie wykonywania ruchu. Systemy inercyjne wykorzystują

dwa czujniki bezwładnościowe: akcelerometr oraz żyroskop. Każdy z nich może być jedno, dwu lub trójosiowy, co przekłada się na liczbę kierunków ruchu, w którym może być dokonany pomiar.

Akcelerometry są to czujniki mierzące siłę oddziałyującą liniowo wzdłuż poszczególnych osi. Pomiar oddziaływającej siły wyrażany jest w odniesieniu do siły grawitacji, co teoretycznie pozwala na określenie z jakim przyspieszeniem, w każdej z mierzonych osi, porusza się dany czujnik. W praktyce, precyzyjne określenie przyspieszeń w każdej osi, a następnie wyznaczenie na tej podstawie odległości na jaką przesunął się czujnik, jest utrudnione ze względu na duże zaszumienie danych. Możliwe jest natomiast przybliżone oszacowanie kąta obrotu danego czujnika względem siły grawitacji.

Drugim z czujników inercyjnych jest żyroskop, który pozwala zmierzyć prędkość kątową z jaką obraca się czujnik wokół każdej z osi układu odniesienia. W przypadku pozostawania czujnika bez ruchu, pomiary dla każdej z osi powinny wynosić 0. Podobnie jak w przypadku akcelerometru, zaszumienie danych znacząco utrudnia precyzyjne określenie o jaki kąt obrócił się czujnik.

Czujnikiem magnetycznym, jaki wykorzystywany jest w magnetycznych systemach śledzenia ruchu, jest magnetometr. Czujnik ten pozwala na określenie jak jest on obrócony względem pola magnetycznego ziemi.

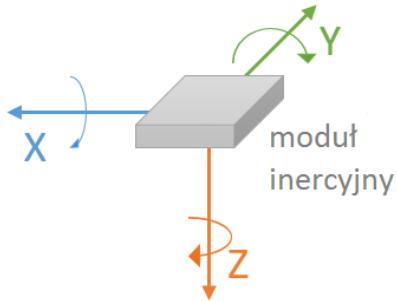
Czujniki te najczęściej wykorzystywane są w konfiguracji akcelerometr–żyroskop lub akcelerometr–żyroskop–magnetometr. Para akcelerometr–żyroskop występuje w literaturze anglojęzycznej pod nazwą modułu inercyjnego (*ang. Inertial Measurement Unit – IMU*) natomiast trójka wspomnianych czujników nazywana jest MARG od pierwszych liter nazw sensorów *Magnetic, Angular Rate, and Gravity*.

Wykorzystywanie tych czujników w parze lub trójce ma na celu połączenie ze sobą ich sygnałów i dzięki temu zmniejszenie wpływu szumów na pomiary uzyskiwane przez każdy z nich z osobna. Metody łączenia ze sobą sygnałów z czujników inercyjnych oraz magnetometru, a także identyfikacja szumów, jakie na nie działają, od wielu lat są przedmiotem badań uczonych ze wszystkich ośrodków badawczych, zajmujących się śledzeniem, jak również analizą ruchu, i są wciąż kontynuowane. Jako przykłady prac związanych z tym zagadnieniem można przytoczyć chociażby propozycję wykorzystania transformaty Fouriera do badania szumu magnetometrów w zakresie niskich częstotliwości ($0\text{Hz}, 5\text{Hz}$)[20]. Badania te były prowadzone przez zespół włoskich naukowców pod przewodnictwem Maurizio Candidiego już w latach 70-tych XX wieku na potrzeby przemysłu kosmicznego. W latach 90-tych zaszumienie danych z czujników inercyjnych: akcelerometru i żyroskopu oraz metody kompensacji tych szumów były obiektami badań przedstawionych

przez S. Woolvena i D. B. Reida z Uniwersytetu w Ontario [126]. Autorzy w swoim artykule zaproponowali łączenie danych z dwóch wspomnianych czujników inercyjnych za pomocą filtra Kalmana jako jedną z metod kompensacji szumów. Wśród współczesnych badań znaleźć można propozycję zastosowania falek do odszumiania pomiarów czujników inercyjnych (El-Sheemy [102]). Zastosowanie ich w kontekście systemu nawigacji inercyjnej (INS – ang. *Inertial Navigation System*) pozwoliło zmniejszyć niemal dziesięciokrotnie wariancję wyznaczonego kierunku poruszania się nawigowanego obiektu oraz trzykrotnie przyspieszyć ustabilizowanie się estymacji obranego kierunku ruchu. W dotychczas dyskutowanych publikacjach, metodą łączenia danych z wykorzystywanych czujników był filtr Kalmana. Jego główną wadą jest jednak wysoka złożoność obliczeniowa spowodowana wykorzystaniem rachunku macierzowego. W 2015 roku Pasquale Daponte wraz ze swoim zespołem przedstawili alternatywną, do filtru Kalmana, metodę łączenia danych z czujników inercyjnych i magnetycznych [25]. Obiektem ich badań była metoda zmniejszającej się gradientu wykorzystywana wcześniej przez Roberta Mahony'ego [78] oraz Sebastiana Madgwicka [76]. Autorzy dyskutowanego artykułu wskazali trzy główne zalety badanej przez nich metody w porównaniu do wykorzystawanego powszechnie filtru Kalmana: mniejsza złożoność obliczeniowa, szybsza stabilizacja uzyskiwanych wyników i ich lepsza stabilność w czasie. Dodatkowo, zwróciли oni uwagę, że dzięki zastosowaniu kwaternionów do reprezentacji obrotów, ułatwiono wykorzystanie otrzymywanych wyników w dalszych obliczeniach, a także uniknęto niejednoznaczności, które występują w przypadku innych metod reprezentowania rotacji.

Warto zauważyć, że zarówno moduł zbudowany z pary czujników inercyjnych jak i moduł zawierający akcelerometr, żyroskop oraz magnetometr określają jak względem poszczególnych osi obrócony jest dany moduł. W przypadku wykorzystania tylko czujników inercyjnych, możliwe jest określenie obrotu wokół dwóch osi prostopadłych do osi wyznaczonej przez siłę grawitacji. Na rysunku 2.4 są to osie X i Y. W przypadku wykorzystania również magnetometru możliwe jest określenie obrotów wokół każdej z tych osi. Warto również zauważyć, że układ współrzędnych połączonych czujników, w jakim przedstawione są obroty wokół każdej z osi, odpowiada osiom w układzie współrzędnych Ziemi. Oznacza to, że w przypadku wykorzystania wszystkich trzech czujników, możliwe jest określenie orientacji modułu względem czterech stron świata (obrót wokół osi Z) oraz względem płaszczyzny Ziemi (obroty względem osi X i Y).

W swojej pracy Alexiev i Nikolova [2] wyróżniają 4 klasy czujników inercyjnych ze względu na stopień zaszumienia ich pomiarów: nawigacyjne (oryg. *navigation*), taktyczne (oryg. *tactical*), przemysłowe (oryg. *industrial*) oraz



Rysunek 2.4: Układ współrzędnych dla czujników inercyjnych i magnetycznych (źródło: opracowanie własne)

Tablica 2.2: Błąd estymacji położenia w zależności od błędu pomiarów akcelerometru[2].

Klasa czujnika	Błąd pomiarów [$10^{-3}g$]	Błąd szacowania położenia [m]			
		1 s	10 s	60 s	1 godz.
Nawigacyjne	0.025	0.00013	0.012	0.44	1600
Taktyczne	0.3	0.0015	0.15	5.3	19000
Przemysłowe	3	0.015	1.5	53	190000
Samochodowe	125	0.62	60	2200	7900000

samochodowe (oryg. *automotive*). Poziom zaszumienia pomiarów czujników inercyjnych ma znaczący wpływ na to, czy możliwe jest wystarczająco precyzyjne określenie położenia modułu inercyjnego w przestrzeni wyznaczając metodami zliczeniowymi jego chwilowe przemieszczenie jedynie na podstawie pomiarów z jego czujników. Spowodowane jest to szybkim kumulowaniem się błędów, a co za tym idzie znacznym spadkiem dokładności w czasie. Tabela 2.2 prezentuje przykład zaczerpnięty z [2] jak szum pomiarów akcelerometru wpływa na dokładność oszacowania przemieszczenia się czujnika.

Jak widać z danych umieszczonych w tabeli 2.2, uzyskanie dużej dokładności szacowania położenia stawów szkieletu ludzkiego, za pomocą metod zliczeniowych, wykorzystując pomiary czujników inercyjnych, jest zadaniem trudnym, a w długim czasie śledzenia może okazać się wręcz niemożliwe. Fakt ten ma zatem wpływ na sposób budowania systemów śledzenia ruchu opartych na urządzeniach inercyjnych. W przypadku śledzenia ruchu człowieka wymagane jest uzupełnienie pomiarów poprzez zdefiniowanie hierarchicznego modelu szkieletu ludzkiego, zawierającego informację o długościach poszczególnych kości. Umożliwia to oszacowanie położenia poszczególnych stawów na podstawie orientacji czujników umieszczonych na poszczególnych częściach ciała. W związku z tym, duży wpływ na dokładność śledzenia ruchu

i szacowania pozycji poszczególnych stawów, w systemach wykorzystujących czujniki inercyjne i magnetyczne, ma dokładność pomiaru długości poszczególnych kości.

Przykładem komercyjnej implementacji takiego systemu może być rozwiązanie zaproponowane przez firmę Xsens, które z powodzeniem wykorzystywane jest zarówno do śledzenia ruchu na potrzeby animacji¹⁹, jak i do analizy ruchu u sportowców^{20,21}, czy w procesie rehabilitacji motorycznej²². Jednak na przykładzie zastosowań sportowych^{20,21} należy zauważyć, że układy IMU i MARG nie są wystarczające do zlokalizowania śledzonej postaci na scenie, a jedynie do odtworzenia jej pozy. Aby śledzić aktualną lokalizację postaci wykorzystywany jest dodatkowy system umożliwiający określenie położenie postaci w przestrzeni. W zależności od tego na jak rozległej przestrzeni odbywa się ruch, może być to system GPS lub systemy oparte o technologię LPM (*ang. local position measurement*). GPS przeznaczony jest do lokalizowania obiektów, wyposażonych w odpowiedni odbiornik, na bardzo dużej otwartej przestrzeni (śledzony obiekt nie powinien być wewnątrz budynku), przy dokładności określenia położenia rzędu kilku metrów²³. Technologia LPM może być wykorzystywana zarówno wewnątrz budynków, jak i na zewnątrz, wymaga jednak dodatkowych urządzeń obserwujących określony obszar, na którym odbywa się ruch. Jako przykład realizacji systemu LPM może posłużyć system firmy Inmotio²⁴ wykorzystujący technologię radiową RFID (identyfikacja radiowa, *ang. Radio-frequency identification*), która pozwala określić położenie śledzonej postaci na podstawie triangulacji pomiaru mocy sygnału pomiędzy stacjami bazowymi a odbiornikiem. Przykładowa konfiguracja takiego systemu widoczna jest na rysunku 2.5). Szczegółowy opis działania systemu Xsens można znaleźć w pracy Roetenberg i in. [93].

¹⁹ Xsens Motion Technologies, *Ted 2: Using Xsens for the grand dance opening scene*, <https://goo.gl/1xQaqi>.

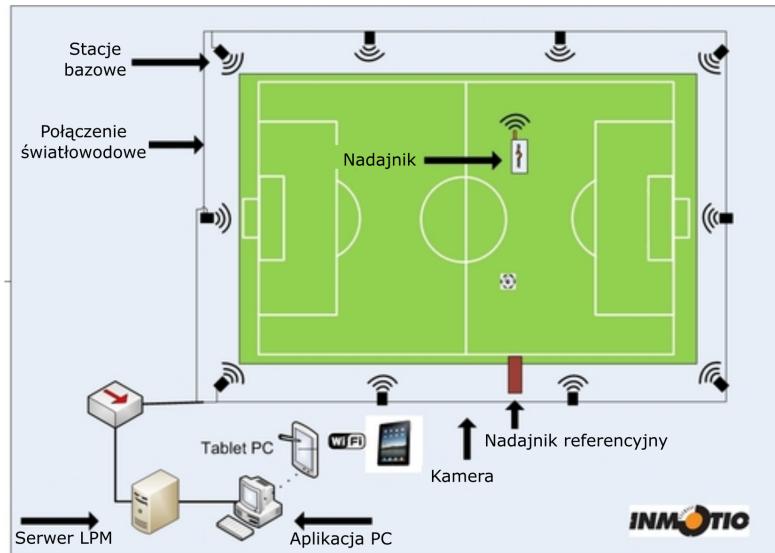
²⁰ Xsens Motion Technologies, *Rowing with Xsens MVN BIOMECH*, <https://goo.gl/5dYRVQ>

²¹ Xsens Motion Technologies, *Full body motion capture of an entire ski jump*, <https://goo.gl/pPd8bt>

²² Xsens Motion Technologies, *Interactive sonification of human movements for stroke rehabilitation*, <https://goo.gl/2CfIvn>.

²³ Według danych zamieszczonych na <http://www.gps.gov/> dokładność systemu GPS może wynieść około 5m dla odbiorników wykorzystywanych na przykład w telefonach komórkowych. Według wersji systemu, którego odbiorniki wykorzystują komunikację na dwóch częstotliwościach równocześnie (na przykład w przyrządach geodezyjnych), uzyskiwana dokładność położenia jest rzędu kilku milimetrów

²⁴ Inmotio Corp, *Local Position Measurement (LPM) technology is the world's most precise sports tracking system*. <https://goo.gl/IVjiAQ>

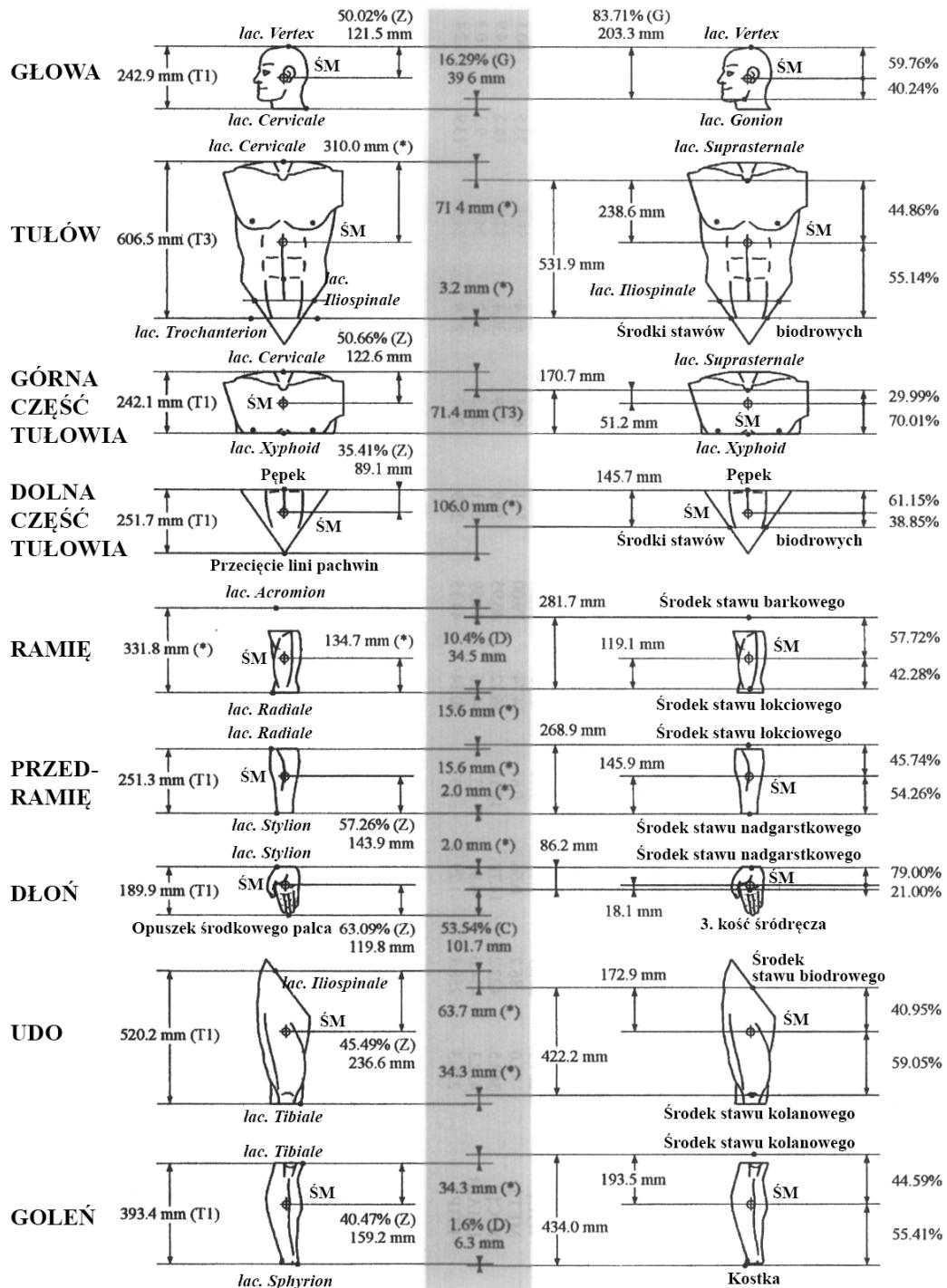


Rysunek 2.5: Przykładowa konfiguracja systemu LPM firmy Inmotio²⁴

Zmniejszenie wpływu zaszumienia pomiarów czujników inercyjnych w celu poprawy dokładności szacowania pozycji stawów w inercyjnych i magnetycznych systemach śledzenia ruchu jest często spotykanym zagadnieniem w literaturze naukowej. Wspomniani już Alexieva i Nikolovej [2], zaproponowali statystyczną metodę detekcji kiedy występuje prawdziwy ruch czujnika. Dzięki temu możliwe jest ograniczenie wpływu pomiarów złożonych jedynie z szumu, a to z kolei pozwala na spowolnienie spadku dokładności szacowania pozycji w czasie. Innym przykładem może być metoda przedstawiona przez Zhou i Hu [131, 132]. Zaproponowali oni, na podstawie badań opartych o śledzenia ruchu stawów ręki, wykorzystanie algorytmu symulowanego wyżarzania w celu zmniejszenia błędów pomiarowych czujników oraz zastosowali rozszerzony filtr Kalmana do połączenia sygnałów z sensorów. Zaproponowaną przez siebie metodę, Zhou i Hu przetestowali wykorzystując komercyjny, inercyjny system śledzenia ruchu firmy Xsens, natomiast dane referencyjne uzyskali za pomocą wizyjnego systemu śledzenia ruchu firmy Vicon. Zaproponowana przez nich metoda pozwoliła na poprawienie dokładności śledzenia ruchu stawów ręki o 26% z około 2.3cm do 1.7cm.

W przypadku systemów śledzenia ruchu człowieka opartych o IMU i MARG wpływ na dokładność wyznaczenia położenia poszczególnych stawów ma umiejscowienie czujników na ciele. Badania na ten temat prowadzili między innymi Vanegas i Stirling [117]. Badania te wykazały, że optymalnym miejscem umieszczenia tego typu sensorów są okolice środka masy danej części ciała lub obiektu. W przypadku umieszczenia czujników na ciele człowieka pomocne okazują się być wyniki badań prowadzonych w połowie lat 90-tych przez Paolo de Leva z Uniwersytetu stanu Indiana. De Leva w swojej pracy,

dotyczcej badań nad środkami mas poszczególnych części ciała [70], rozwinał wcześniejsze badania Zatsiorskiego i Seluyanova [129] i wyznaczył położenie poszczególnych środków mas kości. Rysunek 2.6 przedstawia diagram rozmieszczenia środków mas dla każdej z części ciała według badań de Leva.



Rysunek 2.6: Rozmieszczenie środków mas (ŚM) poszczególnych części ciała człowieka według propozycji Zatsiorskiego i Seluyanova (lewo) oraz de Leva (prawo) [70]

Systemy elektromagnetyczne (radiowe)

Elektromagnetyczne systemy śledzenia ruchu opierają się o pomiar strumienia elektromagnetycznego wytworzzonego między nadajnikiem umieszczonym na ciele postaci, której ruch jest śledzony, a odbiornikami otaczającymi obszar, na którym odbywa się śledzenie ruchu. Zarówno nadajnik jak i odbiorniki zbudowane są z kilku prostopadłych cewek, co pozwala na śledzenie ruchu w sześciu stopniach swobody (położenie oraz obroty). Dokładne położenie nadajnika w przestrzeni wyznaczone jest za pomocą triangulacji uzyskiwanych pomiarów. Elektromagnetyczne systemy śledzenia ruchu są podatne na zakłócenia wynikające z obecności elementów metalowych lub magnetycznych w pobliżu obszaru, na którym odbywa się śledzenie. W szczególności źródłem takiego zakłócenia mogą być stalowe elementy konstrukcyjne budynku, w którym znajduje się system śledzenia (np. pręty zbrojeniowe w stropie czy ścianach). Przykładem systemu śledzenia ruchu opartego o tę technologię może być produkt firmy Polhemus o nazwie Patriot²⁵.

Systemy akustyczne

Akustyczne systemy śledzenia opierają się na pomiarze czasu przemieszczania się fali dźwiękowej między nadajnikami umieszczonymi na ciele osoby, której ruch był śledzony, a odbiornikami umieszczonymi wokół obszaru, na którym ruch się odbywał. Systemy te również określają położenie nadajników na podstawie triangulacji. Niestety, są one podatne na wiele czynników naturalnych, które mogą wpływać na dokładność uzyskanych wyników. Do takich czynników należą między innymi: temperatura powietrza, wilgotność, ciśnienie atmosferyczne. Wymienione czynniki zewnętrzne powodują zmianę szybkości przemieszczania się fali dźwiękowej i w konsekwencji, zaburzenia rozchodzenia się dźwięku. Akustyczne systemy śledzenia ruchu zazwyczaj wykorzystują ultradźwięki, więc są niesłyszalne dla człowieka. Jednym z producentów systemów śledzenia opartych o ultradźwięki jest firma Nexonar²⁶, której produkty zostały wykorzystane między innymi w systemie wspomagającym naukę gry w golfa: Science and Motion²⁷.

Systemy mechaniczne

Mechaniczne systemy śledzenia ruchu różnią się znacznie od poprzednio opisanych systemów choćby ze względu na ich ingerencję w swobodę ruchu, jaki może wykonać śledzona osoba. Podstawowym urządzeniem wykorzystywanym w mechanicznych systemach śledzenia ruchu, jest strój nazywany

²⁵ Polhemus Corporation, *Polhemus Patriot*, <https://goo.gl/Gf5X0M>.

²⁶ Nexonar Corporation, *nexonar – next generation tracking technology*, <https://goo.gl/Z6jddH>.

²⁷ Science and Motion: <http://www.scienceandmotion.com/>

egzoszkieletem, zbudowany z szeregu czujników mechanicznych i elektromechanicznych, na przykład potencjometrów, których zadaniem jest odzwierciedlenie pozy jaką przyjmuje śledzona osoba. Przykładem egzoszkieletu może być układ mechaniczny będący elementem systemu śledzenia ruchu Gypsy²⁸. Przykładowy egzoszkielet został zaprezentowany na rysunku 2.7.



Rysunek 2.7: Egzoszkielet dla systemu śledzenia ruchu Gpsy²⁸

Niewątpliwą zaletą mechanicznych systemów śledzenia ruchu jest to, że głównym urządzeniem z jakiego się składa taki system to wspomniany egzoszkielet. Nie jest on podatny na zakłócenia swojego działania przez elementy znajdujące się w otoczeniu, czy przesłanianie się części ciała w trakcie ruchu. Mechaniczne systemy śledzenia ruchu zazwyczaj są w stanie pracować w czasie rzeczywistym. Niestety, wadą jest, wspomniane już wcześniej, ograniczenie swobody wykonywanego ruchu, co powoduje, że nie każdy rodzaj ruchu (na przykład dynamiczne ćwiczenia gimnastyczne) może być śledzony za pomocą układów mechanicznych.

2.2. Komputerowy model postaci ludzkiej

Systemy śledzenia ruchu kończyn człowieka opierają się często na komputerowym modelu postaci, który opisuje właściwości kinematyczne modelu szkieletu oraz, jeśli to możliwe, również kształt ciała śledzonej postaci. Reprezentacja geometryczna opisująca zewnętrzny wygląd postaci jest mapowana na jego układ szkieletowy, dzięki czemu zmiany szkieletu mogą być automatycznie przekładane na zewnętrzną pozę postaci, natomiast na podstawie kształtu i ustawienia ciała człowieka można wnioskować jaka jest konfiguracja szkieletu. Nie istnieje jeden standardowy model postaci, gdyż w zależności od

²⁸ Meta Motion Corp, *Inertial/Mechanical mocap technology*, <https://goo.gl/X4bk11>

zastosowań przyjmowane są różne uproszczenia. O ile szkielet kostny dorosłego człowieka składa się z 206 kości [16], o tyle w modelach komputerowych, na potrzeby rejestracji ruchu, stosuje się ich zazwyczaj kilkadziesiąt. Na przykład, model szkieletowy zastosowany w kontrolerze Kinect składa się z 20 stawów i 19 kości²⁹, zaś model szkieletowy systemu Optitrack zbudowany jest również z 20 stawów, ale połączonych ze sobą za pomocą 25 kości³⁰ (zwiększoną liczbą kości w obrębie klatki piersiowej oraz stóp). Przyjęcie właściwego modelu reprezentacji postaci wpływa na swobodę ruchów (liczba stopni swobody) modelowanego układu kostnego, co z kolei przekłada się na skuteczność i dokładność śledzenia kończyn człowieka.

2.2.1. Model kinematyczny szkieletu postaci

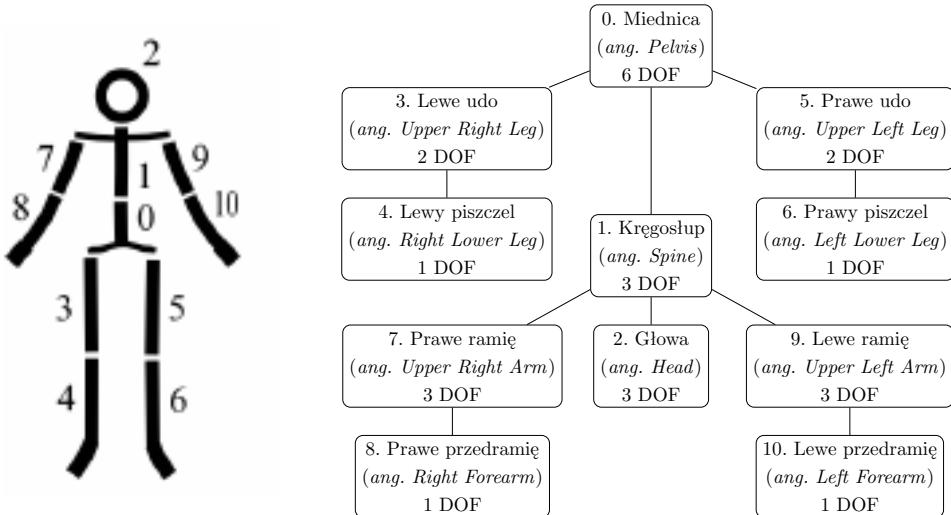
Model kinematyczny szkieletu postaci opisany jest za pomocą struktury przedstawiającej relacje między poszczególnymi stawami połączonymi ze sobą segmentami reprezentującymi kości. Każdy ze stawów opisany jest przez stopnie swobody definiujące w jaki sposób segmenty (kości) połączone danym stawem mogą się wzajemnie poruszać. Teoretycznie maksymalna liczba stopni swobody danego stawu może wynosić 6, co oznacza, że staw może przesunąć się wzdłuż każdej z trzech osi X , Y i Z układu współrzędnych oraz obracać się wokół każdej z nich, jednak w praktyce zakres ruchomości stawów człowieka jest ograniczony i osobniczo różny. W związku z tym pozycja każdego ze stawów jest determinowana stopniem swobody jego rodzica. Wiedząc zatem, że staw łokciowy ma jeden stopień swobody, ruch kości przedramienia jest ograniczony tylko do jednej płaszczyzny, a względna możliwa pozycja stawu nadgarstkowego znajduje się na łuku o promieniu równym długości kości przedramienia.

Model kinematyczny jest zazwyczaj definiowany jako hierarchiczny układ stawów co oznacza, że wielkości opisujące dany staw są wielkościami względnymi w odniesieniu do stawu bezpośrednio go poprzedzającego (nadzawanego) w hierarchii. Aby wyznaczyć wartości absolutne danego stawu, niezbędne jest przyrostowe złożenie informacji ze wszystkich stawów w łańcuchu kinematycznym, począwszy od stawu głównego (korzenia). Korzeniem jest jeden wyszczególniony staw modelu hierarchicznego, do którego doczepione są łańcuchy kinematyczne wszystkich pozostałych stawów. Korzeń może, ale nie musi, mieć swojego odpowiednika w stawach szkieletu ludzkiego. Jako przykład można przytoczyć propozycję uproszczonego modelu szkieletowego

²⁹ Microsoft Corporation, *Tracking users with Kinect skeletal tracking*, <https://goo.gl/iWkg0g>.

³⁰ OptiTrack Corporation, *Skeleton tracking - NaturalPoint product documentation*, <https://goo.gl/5L6MTv>.

człowieka wraz z przedstawieniem hierarchii poszczególnych segmentów oraz ich stopni swobody zaproponowaną przez Kwolek i in.[68], przedstawioną na rys. 2.8. W modelu kinematycznym, w którym określona jest hierarchia jego stawów, układ współrzędnych, w którym opisane są obroty i położenie danego stawu, ma swój początek w stawie poprzedzającym go (staw rodzica). Sprawia to, że wartości przypisane do poszczególnych stawów są wielkościami względnymi.



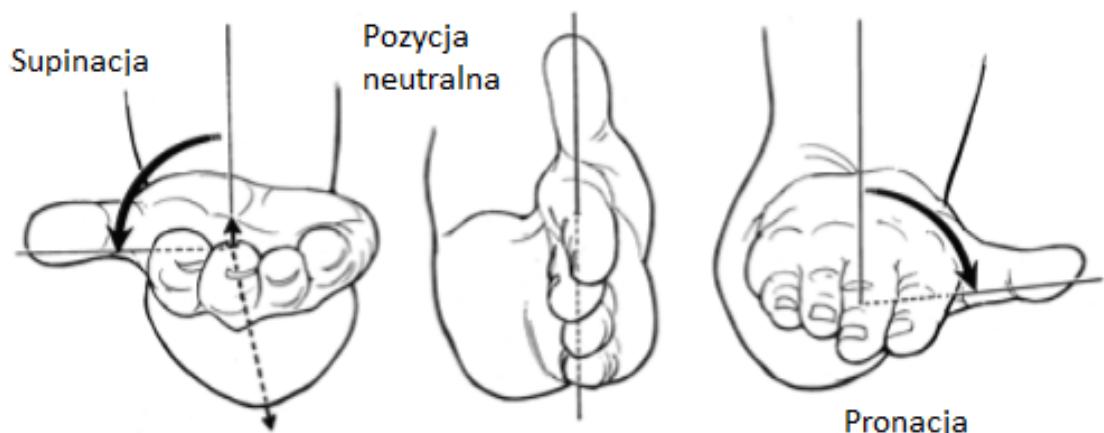
Rysunek 2.8: Uproszczony hierarchiczny model szkieletowy człowieka (lewy) wraz z diagramem przedstawiającym hierarchię poszczególnych elementów(prawy) [68]

Oprócz stopni swobody pojedynczego stawu, definiuje się również stopień swobody całego modelu kinematycznego. Taka wielkość rozumiana jest jako suma stopni swobody wszystkich jego elementów. W przypadku pełnego opisu modelu postaci możemy mieć zatem do czynienia z modelem kinematycznym, który posiada zdefiniowanych ponad 50 stopni swobody [1]. Nie zawsze jednak potrzebne są modele zawierające pełen opis stopni swobody. Okazuje się, że do śledzenia większości ruchów człowieka wystarczający jest model uproszczony definiujący około 30 stopni swobody [69, 107].

O ile model kinematyczny szkieletu jest w stanie jednoznacznie opisać pozę, w jakiej znajduje się w danej chwili śledzona postać, o tyle może okazać się niewystarczający w przypadku śledzenia ruchu w środowiskach wirtualnych, w których istotna jest interakcja postaci z otoczeniem. W takim przypadku niezbędne staje się określenie dodatkowo modelu kształtu ciała śledzonej postaci.

2.2.2. Model reprezentujcy kształt postaci

Modele kształtu ciała, podobnie jak modele kinematyczne, możemy określać w przestrzeni dwu- lub trójwymiarowej. Zwykle wystarczające jest przybliżone odwzorowanie kształtu poszczególnych części ciała za pomocą figur geometrycznych (modele 2D) lub brył (modele 3D). Umożliwia to zwizualizowanie budowy ciała postaci, której ruch jest śledzony, bez zbytniego obciążania procesora komputera, na którym odbywa się przetwarzanie danych pozyskanych z systemu śledzenia. Dokładne odwzorowanie budowy ciała postaci jest możliwe dzięki wykorzystaniu popularnych przestrzennych reprezentacji geometrycznych (np. siatka trójkątna 3D), jednak może to wpływać negatywnie na wydajność systemu. Niezależnie od tego, jakie reprezentacje geometryczne wykorzystujemy, istotne jest takie przedstawienie kształtu ciała, aby odpowiednio zwizualizować przynajmniej śledzone ruchy. Na przykład, jeśli wykorzystujemy trójwymiarowy uproszczony model zbudowany z kul i cylindrów, trudne będzie zwizualizowanie niektórych obrotów ciała, takich jak rotacja dłoni i nadgarstka jak na rys. 2.9.



Rysunek 2.9: Rotacja wewnętrzna (pronacja) i zewnętrzna (supinacja) dłoni³¹

W praktyce to w jaki sposób aktor będzie reprezentowany w systemie komputerowym ściśle powiązane jest z przeznaczeniem danego systemu. Inna reprezentacja będzie konieczna dla śledzenia i rozpoznawania pozy w jakiej znajduje się aktor, a inna jeśli śledzimy jego położenie w przestrzeni. I tak możemy wyróżnić następujące sposoby reprezentowania postaci:

- **punkt** – postać reprezentowana jest przez pojedynczy punkt (zazwyczaj środek ciężkości) [118] lub kilka punktów określających punkty

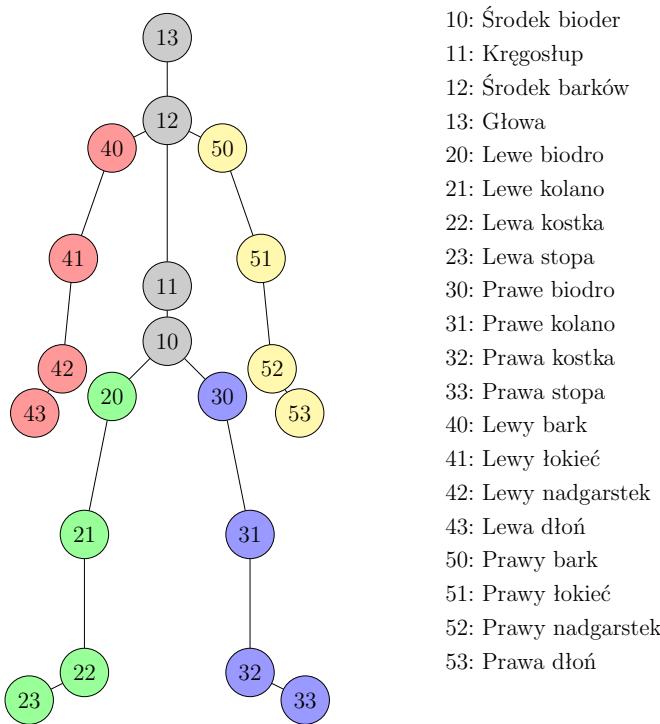
³¹ HookSlide, *Tigers starting pitchers' mechanics FAQ: Part 1*, <https://goo.gl/bJTi4J>

charakterystyczne na ciele [101]. Reprezentacja taka jest zazwyczaj użyteczna jeśli śledzony obiekt zajmuje niewielki fragment obszaru, na którym odbywa się śledzenie;

- **figury i bryły geometryczne** – śledzona postać może być reprezentowana przez pojedynczą figurę/bryłę [21] lub kilka brył odpowiadających za każdą część ciała. W tej reprezentacji wykorzystuje się proste figury/bryły, które mogą jednoznacznie zobrazować śledzone ciało. Do takich możemy zaliczyć między innymi: prostokąty, elipsy, cylindry, prostopadłościany;
- **sylwetka** – w przypadku tej reprezentacji możliwe są dwa warianty realizacji: kontur pusty lub kontur z wypełnieniem. Dodatkowo, sam obrys może być pełny, czyli zrealizowany jako figura zamknięta zaznaczona linią ciągłą, lub przedstawiony jedynie jako zbiór punktów kluczowych na obrębie sylwetki. Taka reprezentacja została zaproponowana między innymi w [128], do śledzenia interakcji człowieka z obiek-tami o nieregularnych kształtach;
- **szkielet** – model szkieletowy jest niejako graficzną reprezentacją modelu kinematycznego. Jest to hierarchiczna struktura zbudowana z prostych figur/brył geometrycznych i swoim kształtem oddaje pozycję, w jakiej znajduje się śledzona postać. Szkielet taki powinien być w stanie w pełni zawrzeć się w obrębie postaci [3]. Chociaż zwyczajowo punkty, z których zbudowany jest model szkieletowy nazywa się stawami (*ang. joints*), nie muszą one odpowiadać stawom w szkielecie biologicznym. Na przykład, często występujący w modelu szkieletowym staw reprezentujący kość miedniczą lub głowę, w rzeczywistości nie istnieje;

W niniejszej pracy wykorzystywana jest trójwymiarowa reprezentacja szkieletowa modelu 3D śledzonej postaci złożona z 20 stawów i 19 segmentów kości. Model ten przedstawiony jest na rysunku 2.10. W zależności od wykorzystywanego systemu śledzenia ruchu, inna jest metoda wyznaczania położenia i obrotów poszczególnych stawów. W dalszej części zamieszczony został opis metod związanych z technikami śledzenia wykorzystywany mi w niniejszej pracy. Omówiony został optyczny system śledzenia ruchu z markerami pasywnymi, inercyjny system śledzenia ruchu oraz optyczny system śledzenia ruchu bez markerów.

³² Microsoft Corporation, *Tracking users with Kinect skeletal tracking*, <https://goo.gl/iWkg0g>.



Rysunek 2.10: Model szkieletowy człowieka wykorzystywany w kontrolerze Microsoft Kinect³²

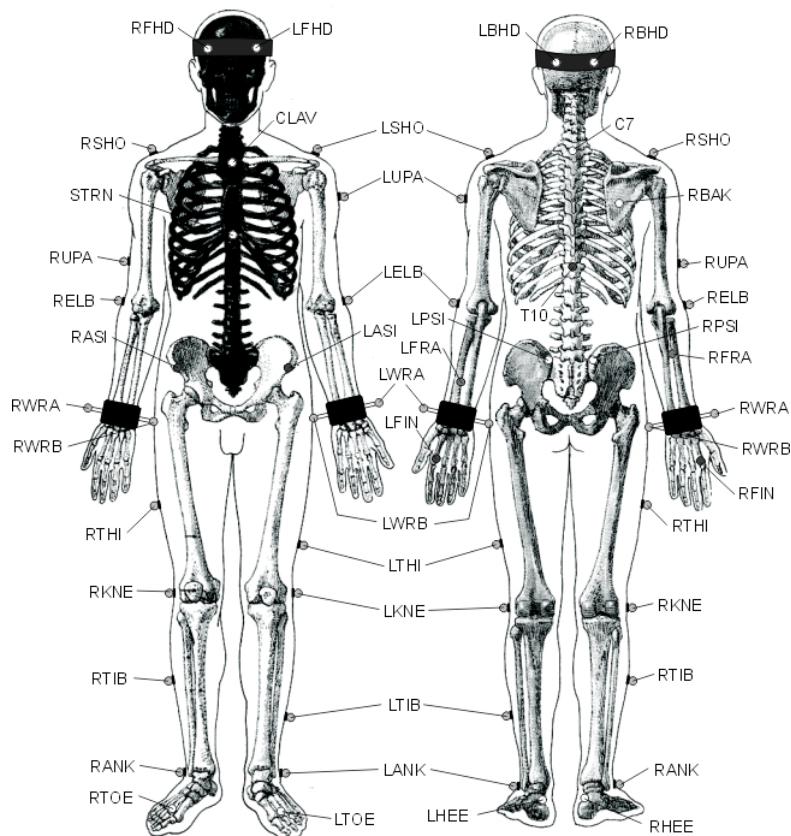
2.2.3. Wyznaczanie położenia stawów w modelu szkieletowym

Metody wyznaczania położenia oraz obrotu stawów w modelu szkieletowym mogą się znacząco od siebie różnić w zależności od zastosowanego systemu śledzenia ruchu. W przypadku systemów wykorzystujących markery możemy przyjąć, że podstawową różnicą (pomijając wykorzystywanie różnych zjawisk fizycznych, takich jak fala elektromagnetyczna, czy fala akustyczna) jest to, czy oszacowane położenie danego stawu jest wynikiem jego bezpośredniego śledzenia (bądź śledzenia markerów przymocowanych w sąsiedztwie tego stawu), czy jest to efekt przekształceń geometrycznych wynikających ze śledzenia ruchu markerów umieszczonych blisko środków ciężkości kości, pomiędzy dwoma kolejnymi stawami. W przypadku systemów śledzenia ruchu niewykorzystujących markerów, wyznaczenie położenia stawów odbywa się przy wykorzystaniu analizy i rozpoznawania obrazów postaci, na podstawie których wnioskuje się położenie poszczególnych stawów.

Wyznaczanie położenia stawów w optycznych systemach śledzenia ruchu z markerami

Optyczne systemy śledzenia ruchu kończyn stosujące markery są przykładem systemów, w których określenie położenia stawów odbywa się przez śledzenie ruchu markerów bezpośrednio przymocowanych do ciała śledzonej

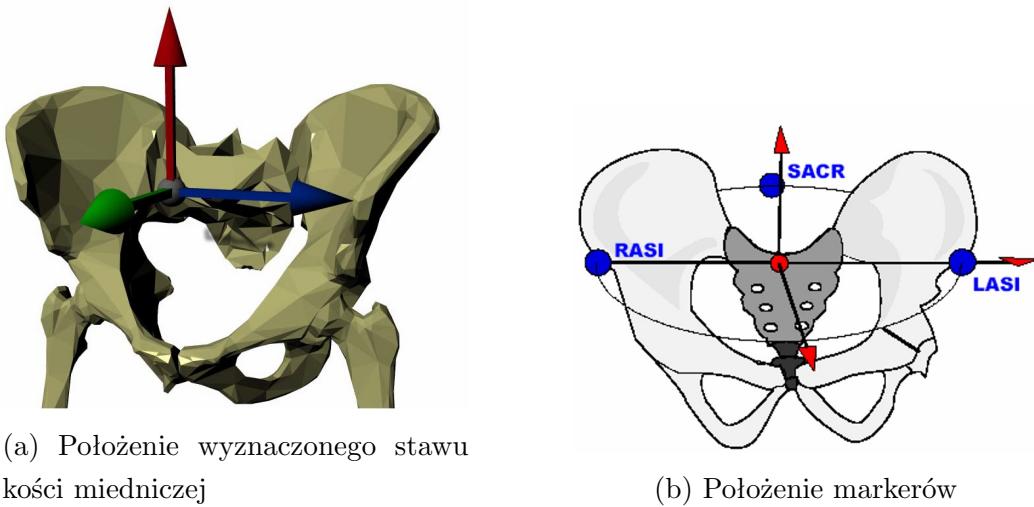
postaci. Aby możliwe było określenie położenia oraz obrotu danego stawu, każdy staw musi być oznaczony przez minimum dwa markery. Komercyjnie dostępne, optyczne systemy śledzenia ruchu stosujące markery posiadają zazwyczaj zdefiniowane schematy rozmieszczenia markerów na ciele człowieka, tak aby móc na ich podstawie automatycznie określić niezbędne parametry dla każdego ze stawów. Rysunek 2.11 przedstawia przykładowy schemat rozmieszczenia markerów w systemie śledzenia firmy Vicon. Szczegółowy opis akronimów użytych do oznaczenia markerów można znaleźć w dokumentacji systemu Vicon [120].



Rysunek 2.11: Schemat umieszczania markerów na ciele człowieka w systemie śledzenia ruchu firmy Vicon [120]

Kolejnym krokiem, po wyznaczeniu położenia poszczególnych markerów na podstawie obrazów z kamer, jest interpretacja tych danych w taki sposób, aby uzyskać opis stanu stawów, które są reprezentowane graficznie przez pojedyncze punkty posiadające informacje nie tylko o swoim położeniu, ale także o orientacji przestrzennej. Interpretacja ta odbywa się na podstawie obliczeń matematycznych, wynikających wprost z położenia przestrzennego wyznaczonych punktów i relacji między nimi, lub na podstawie przyjętego modelu biomechanicznego ciała ludzkiego.

Przykładem sposobu wyznaczania położenia stawu na podstawie wzajemnej relacji pomiędzy stawami, może być proces wyznaczania położenia kości miedniczej (*ang. pelvis*). Rysunek 2.12 pokazuje umiejscowienie markerów wykorzystywanych do wyznaczenia pozycji stawu reprezentującego kość miedniczą. Według dokumentacji [119] położenie kości miedniczej jest szacowane jako średnia arytmetyczna punktów RASI i LASI (rys. 2.12b). Jeśli jeden z tych dwóch punktów jest niewidoczny, wówczas położenie stawu reprezentującego kość miedniczą jest wyznaczane jako wierzchołek trójkąta prostokątnego, którego przeciwprostokątną wyznaczają punkty SACR oraz jeden z widocznych RASI lub LASI.



Rysunek 2.12: Rysunek przedstawiający wyznaczanie położenia kości miedniczej (a) na potrzeby modelu szkieletowego, na podstawie położenia markerów (b) w systemie Vicon [119]

Przykładem wyznaczania położenia stawu modelu szkieletowego na podstawie biomechanicznego modelu szkieletu ludzkiego, może być schemat wyznaczania położenia stawów biodrowych. Opierając się na pracy Davies i in. [26] do wyznaczenia położenia stawów zostały wykorzystane informacje o szerokości miednicy, długości kości udowej postaci, której ruch jest śledzony, oraz współczynników skalujących i stałych zdefiniowanych we wspomnianej pracy. Równania 2.1a–2.1c przedstawiają przykład wyznaczania współrzędnych położenia x, y, z dla prawego biodra:

$$X = f_C * \cos(\gamma) * \sin(v) - (f_h + \phi_m) * \cos(v) \quad (2.1a)$$

$$Y = -(f_C * \sin(\gamma) - f_a) \quad (2.1b)$$

$$Z = -f_C * \cos(\gamma) * \cos(v) - (f_h + \phi_m) * \sin(v) \quad (2.1c)$$

Przyjmując, że l_l i l_p to długości kości udowych obu nóg, współczynnik f_C zdefiniowany jest empirycznie według wzoru 2.2.

$$f_C = \frac{l_l + l_p}{2} * 0.115 - 15.3 \quad (2.2)$$

Współczynnik f_h wyznaczany jest na podstawie wartości l_l albo l_p w zależności od tego, czy wyznaczane jest położenie biodra dla lewej czy dla prawej strony. Współczynnik f_h zdefiniowany został według wzoru 2.3.

$$f_h = 0.1288 * l_p - 48.56 \quad (2.3)$$

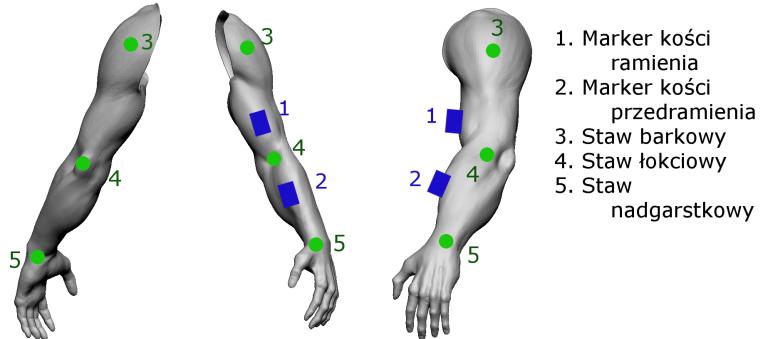
Współczynnik f_a , wykorzystany we wzorze 2.1b, wyznaczony jest jako połowa odległości pomiędzy stawem kości miedniczej a położeniem markera LASI albo RASI, natomiast współczynnik \varnothing_m , użyty we wzorach 2.1a i 2.1c, określa średnicę markerów wykorzystywanych w trakcie śledzenia. Kąty γ oraz v przyjmują stałe wartości odpowiednio 0.5 rad oraz 0.314 rad . W dokumentacji systemu śledzenia Vicon [119] można również znaleźć opisy procedur wyznaczania położenia dla pozostałych stawów.

Wyznaczanie położenia stawów w inercyjnych i magnetycznych systemach śledzenia ruchu

Inercyjne i magnetyczne systemy śledzenia ruchu wymagają nieco innego podejścia do wyznaczania położenia poszczególnych stawów, ponieważ nie udostępniają one bezpośrednio informacji dotyczących ich położenia w przestrzeni, a jedynie orientację oraz wartości sił działających na poszczególne czujniki. Pozwala to określić przyspieszenie i kierunek z jakim porusza się czujnik (akcelerometr) oraz prędkość z jaką się obraca (żyroskop). Dzięki wykorzystaniu pomiarów poszczególnych czujników inercyjnych, możliwe jest zastosowanie metod zliczeniowych do określenia przyblizonego położenia wynikającego z przemieszczenia się obiektu, do którego dany czujnik jest przy mocowany [52, 83]. Uzyskanie dokładnego położenia danego czujnika wymaga wykorzystania drogich komponentów, w których zasumienie danych pomiarowych jest stosunkowo niewielkie. W przypadku powszechnie dostępnych czujników inercyjnych wykorzystywanych na przykład w telefonach, czy w inercyjnych systemach śledzenia ruchu, takich jak XSense, moduły te wykorzystywane są do określenia orientacji przestrzennej takiego modułu, czy chociażby kształtu pozy jaką przyjęła w danej chwili śledzona osoba. Konsekwencją takiego podejścia jest to, że korzeń modelu szkieletowego w systemie śledzenia ruchu opartym o czujniki inercyjne i magnetyczne pozostaje zazwyczaj nieruchomy.

Możliwe jest zdefiniowanie tylko fragmentu modelu szkieletowego, zawierającego na przykład ramię i przedramię prawej ręki, którego korzeniem będzie staw barkowy. Wówczas cały model zbudowany będzie z trzech stawów, jednego nieruchomego odpowiadającego stawowi barkowemu i dwóch ruchomych: łokciowego i nadgarstkowego.

W inercyjnych i magnetycznych systemach śledzenia ruchu markery wyposażone w odpowiednie czujniki umieszczane są na powierzchni ciała osoby (wzdłuż kości), której ruch jest śledzony. Czujniki inercyjne działają najlepiej jeśli umieszczone są blisko środka masy poruszającego się obiektu. Wyznaczeniem środków mas kości człowieka zajmują się między innymi badacze z zakresu biomechaniki, a przykładem efektów tych prac może być diagram wyznaczający środki mas kości zaprezentowany w pracy Paolo de Leva [70], przedstawiony na rysunku 2.6.



Rysunek 2.13: Przykładowy schemat umieszczenia markerów inercyjnych na potrzeby śledzenia ruchu ręki. Niebieskie prostokąty oznaczają schematyczne umieszczenie czujników inercyjnych, zaś zielone kropki oznaczają symbolicznie pozycję stawów szkieletu ramienia (źródło: opracowanie własne)

Rysunek 2.13 pokazuje przykładowy sposób umieszczenia markerów inercyjnego systemu śledzenia ruchu na powierzchni ręki, tak aby ich położenie możliwe pokrywało się ze środkiem mas poszczególnych segmentów kończyny. Moduły inercyjne, dostarczające informacje o swojej orientacji w przestrzeni, nie przekazują wprost informacji o położeniu konkretnych stawów, a jedynie określają jaki obrót, względem przyjętego stanu początkowego, wykonała kończyna, do której te moduły zostały przymocowane. Posługując się przykładowym schematem z rysunku 2.13, przyjmijmy staw barkowy (punkt 3 z rys. 2.13) jako korzeń. Chcąc określić położenie stawu łokciowego (punkt 4 z rys. 2.13), niezbędna jest informacja o orientacji przestrzennej modułu inercyjnego umieszczonego na ramieniu (prostokąt 1 z rys. 2.13), oraz długość kości ramieniowej, która łączy stawy barkowy i łokciowy. Mając te dwie informacje należy dokonać obrotu wektora o długości takiej, jak

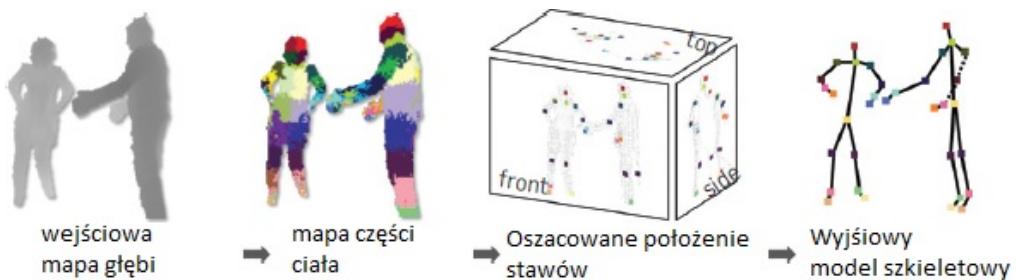
długość kości ramieniowej, do orientacji, w jakiej znajduje się rozważany moduł (prostokąt 1 z rys. 2.13). Wówczas punkt początkowy hierarchicznego modelu szkieletowego ręki reprezentuje staw barkowy i znajduje się w umownym punkcie $(0, 0, 0)$, natomiast punkt końcowy reprezentuje staw łokciowy, a jego współrzędne (X, Y, Z) odpowiadają relatywnemu położeniu tego stawu w przestrzeni. Wyznaczenie położenia stawu nadgarstkowego (punkt 5 z rys. 2.13) odbywa się analogicznie, przy czym punktem nadzewnętrznym jest położenie stawu łokciowego (punkt 4 z rys. 2.13), a modułem, którego orientacja jest brana pod uwagę, jest sensor umieszczony na przedramieniu (prostokąt 2 z rys. 2.13). W przypadku systemów śledzenia ruchu wykorzystujących moduły inercyjne, istotne jest zachowanie hierarchii stawów wykorzystywanego modelu szkieletowego i wykonywanie obliczeń zgodnie z tą hierarchią.

Wyznaczanie położenia stawów modelu szkieletowego w optycznym, bezmarkerowym systemie śledzenia ruchu

Optyczne, bezmarkerowe systemy śledzenia ruchu, w głównej mierze, opierają się na analizie i rozpoznawaniu rejestrowanego obrazu w celu wyodrębnienia elementów, których ruch będzie śledzony. Elementy takie można wówczas interpretować jako wirtualne znaczniki, bo w istocie system śledzenia jedynie na ich podstawie odtwarza model szkieletu postaci. Na przykładzie kontrolera Microsoft Kinect można przedstawić jeden z możliwych sposobów wyznaczania modelu postaci przez oprogramowanie wykorzystywane w systemie śledzenia ruchu nie bazującym na markerach, lecz stosującym algorytmy rozpoznawania obrazów. Algorytm zastosowany w kontrolerze Kinect jest w stanie rozpoznać tylko szkielet postaci ludzkiej. W zależności od trybu pracy szkielet ten przedstawia całą sylwetkę albo jedynie jej górną część, na którą składają się głowa, ręce oraz barki. Należy zwrócić uwagę, że model szkieletowy przedstawiający pełną sylwetkę, jest jedynie modelem uproszczonym zawierającym 20 stawów (rys. 2.10).

Kontroler Kinect wykorzystuje mechanizm uczenia maszynowego, dzięki któremu możliwe jest rozpoznawanie poszczególnych części ciała śledzonej postaci na podstawie obrazu mapy głębi. Zastosowana w kontrolerze metoda uczenia maszynowego opiera się na algorytmie losowych lasów decyzyjnych [23], który podlegał wcześniejszemu procesowi uczenia. W celu nauczenia algorytmów decyzyjnych, zastosowanych w kontrolerze Kinect, przygotowano bazę ponad miliona wzorców różnych póż, w jakich może się znajdować śledzona postać wraz z oznaczonymi częściami ciała. Zestaw zdjęć zastosowany

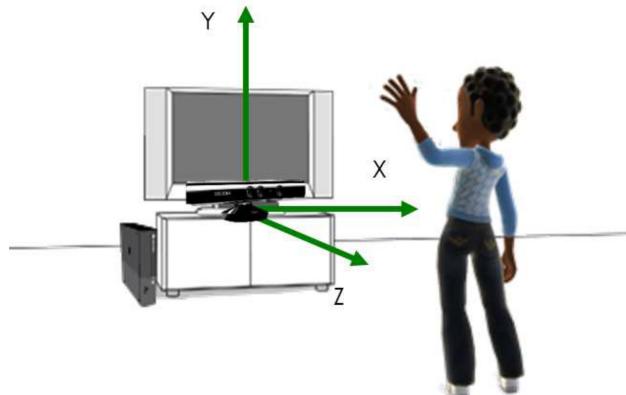
do uczenia algorytmu składał się z blisko stu tysięcy zdjęć póż rzeczywistych postaci, uzupełnionych pomiarami z systemu śledzenia ruchu o wysokiej dokładności oraz z kilkuset tysięcy wzorcowych zdjęć cyfrowych postaci wygenerowanych (wyrenderowanych) komputerowo[74]. Odpowiednio wytrenowany algorytm zyskał gotowość do rozpoznawania i klasyfikowania elementów przedstawionego obrazu jako części ciała człowieka. Wykorzystując między innymi wzajemne położenie elementów (pojedyncze punkty lub grupy punktów – obszary) występujących na mapie głębi, algorytm szacuje, które z nich należą do tej samej części ciała. Efektem działania algorytmu losowych lasów decyzyjnych jest mapa części ciała (*ang. body parts map*), której przykład przedstawiony jest jako drugi krok na schemacie widocznym na rysunku 2.14.



Rysunek 2.14: Schemat przedstawiający kolejne kroki wyznaczania modelu szkieletowego na podstawie mapy głębi (źródło: opracowanie własne na podstawie [74])

Mapa części ciała stanowi podstawę dla kolejnego kroku wyznaczania modelu szkieletowego, jakim jest oszacowanie położenia stawów w przestrzeni trójwymiarowej. Segmenty, z których składa się mapa części ciała, przetworzone są algorytmem *mean shift* [21], co w efekcie daje oszacowanie położenia poszczególnych stawów. Połączenie ich zgodnie ze zdefiniowaną hierarchią wyznacza model szkieletowy śledzonej postaci. Warto zaznaczyć, że większość tych operacji zaimplementowana została jako oprogramowanie wbudowane procesora firmy Prime Sense. Hipotetycznie, w celu wykorzystania Kinecta do podobnego rozpoznawania innych obiektów, czy zwierząt, należałoby zaimplementować przedstawiony powyżej proces, opierając się na specjalnie przygotowanych do tego danych.

Warto zwrócić uwagę, że proces wyznaczania mapy głębi i reprezentacji szkieletowej śledzonego użytkownika opiera się na danych pozyskanych tylko z jednego układu, jaki stanowią kamera i projektor światła podczerwonego (IR - *ang. infra red*). Kamera RGB nie jest w ogóle do tego zadania wykorzystywana, co z kolei skutkuje tym, że urządzenie jest w stanie działać w zaciemnionych pomieszczeniach.



Rysunek 2.15: Układ współrzędnych wykorzystywany w kontrolerze Microsoft Kinect³³

Kinect wykorzystuje reprezentację ciała użytkownika w postaci szkieletowej. Składa się ona z 20 jednopunktowych węzłów, które w dużej mierze pokrywają się ze stawami ludzkiego szkieletu (rys. 2.10). Każdy z węzłów posiada swoje współrzędne w przestrzeni trójwymiarowej związanej z kontrolerem, wyrażone w metrach, gdzie punktem (0,0,0) jest projektor podczerwieni w kontrolerze Kinect. Rysunek 2.15 przedstawia układ współrzędnych wykorzystywany w Kineckie do określenie położenia stawów.

2.2.4. Reprezentacja rotacji obiektu w przestrzeni

Istnieje wiele metod reprezentowania orientacji obiektu w przestrzeni, i każda z tych metod wyróżnia się charakterystycznymi dla siebie cechami, które mogą wpływać na łatwość obliczeń kosztem czytelności zapisu bądź na optymalizację użycia pamięci przy ich implementacji komputerowej. Do najpopularniejszych można zaliczyć następujące metody:

- Kąty Eulera,
- Macierze rotacji,
- Pary Osie-Kąt (*ang. Axis-Angle*),
- Kwaterniony.

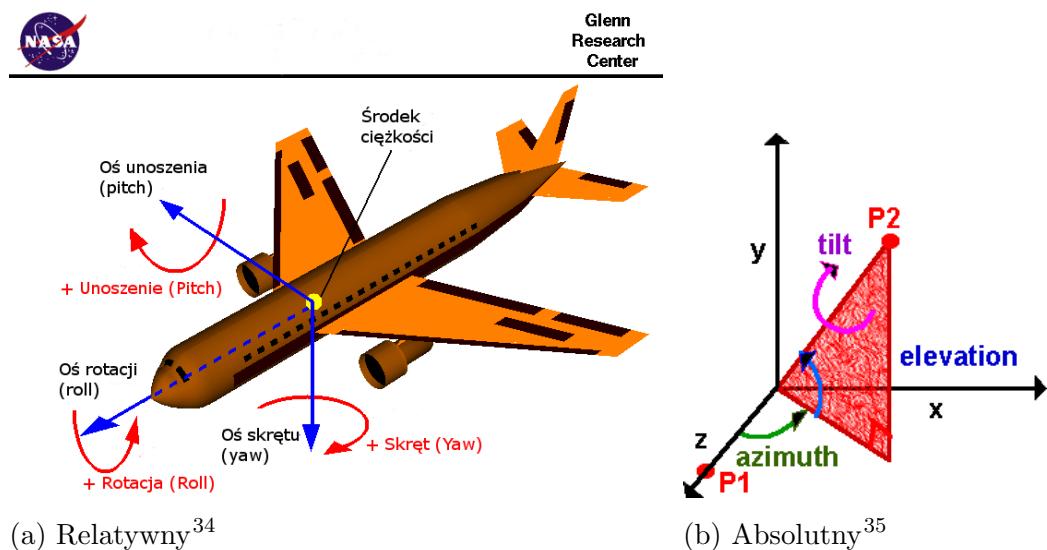
Autorska metoda łączenia danych opisywana w niniejszej pracy wykorzystuje reprezentację obrotów w formie kątów Eulera oraz kwaternionów, w związku z tym, te dwie formy zostaną przybliżone w dalszej części pracy.

Kąty Eulera

Reprezentacja orientacji obiektu za pomocą kątów Eulera jest najbardziej intuicyjną z powszechnie używanych reprezentacji, ponieważ przedstawia ona

³³ Microsoft Corporation, *Kinect for Windows SDK 1.8—Coordinate spaces*, <https://goo.gl/WMqfx0>.

trzy wartości kątów, o jakie nastąpił obrót względem odpowiednich osi układu współrzędnych. Należy jednak pamiętać, że konwencje nazewnicze odnoszące się do poszczególnych kątów i osi różnią się w zależności od dziedziny, w jakiej określana jest orientacja. Na przykład innych oznaczeń używa się w aviacji, a innych w przypadku określenia obrotów kończyn w systemach śledzenia. Jest to związane z układem odniesienia, jaki został przyjęty do określania obrotów. Możemy wyróżnić relatywny układ odniesienia, w którym obrót mierzony jest względem arbitralnie wybranego układu (*ang. body-fixed coordinate system*, rys.2.16a), lub układ absolutny – względem kierunków ziemi (*ang. world coordinate system*, rys.2.16b). Tabela 2.3 zawiera powszechnie przyjęte konwencje oznaczania obrotów. Nazwy osi układu współrzędnych w tabeli 2.3 są zgodne z osiami układu współrzędnych zaprezentowanym na rysunku 2.16b.



Rysunek 2.16: Katy Eulera wzgledem przyjetygo ukладu wspolrzednych: relatywnego (a) oraz absolutnego (b)

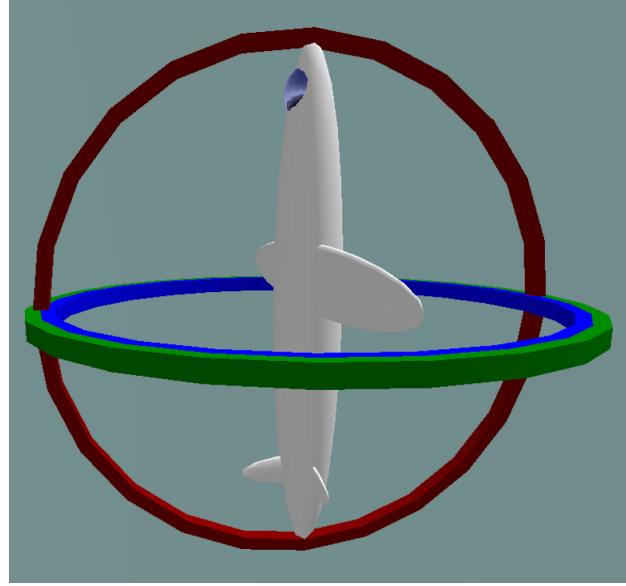
Tablica 2.3: Konwencje nazewnicze kątów Eulera

Oś	Układ Relatywny	Układ Absolutny	Symbol	Prędkość kątowa
X	<i>bank</i>	<i>tilt</i>	ϕ	<i>roll</i>
Y	<i>heading</i>	<i>azimuth</i>	ψ	<i>yaw</i>
Z	<i>attitude</i>	<i>elevation</i>	θ	<i>pitch</i>

³⁴ National Aeronautics and Space Administration, *Aircraft rotations*, <https:// goo.gl/wVLUqm>.

³⁵ M. Baker, *Maths - Euler angles - Martin Baker*, <https:// goo.gl/xh96W8>.

Chcąc prawidłowo interpretować podane wartości trzeba pamiętać, że definiują one zawsze obroty względem poszczególnych osi w ścisłe określonej kolejności. Spośród 27 możliwych kombinacji³⁶ określających kolejność wykonywania rotacji wokół osi, jedynie 12 może być wykorzystane do zdefiniowania obrotów w przestrzeni 3D [29]. Ograniczenie liczby możliwych do wykorzystania sekwencji obrotów, wynika z założenia, że dwa następujące po sobie obroty nie mogą być wykonane wokół tej samej osi. Dokładne opisy sekwencji wykonywanych obrotów można znaleźć w pracach poświęconych zagadnieniu rotacji obiektów [29, 90]. Innym utrudnieniem, które trzeba brać pod uwagę w przypadku stosowania kątów Eulera wyrażonych w relatywnym układzie współrzędnych, jest występowanie, podczas obrotu obiektu będącego w orientacji pionowej ($\phi = \pm 90^\circ$, rys. 2.17), niejednoznaczności, określanej jako blokada przegubu (*ang. gimbal lock*). Przykład przedstawiony na rysunku 2.17 pokazuje sytuację, w której w wyniku skierowania obiektu do orientacji pionowej, dwie z trzech osi obrotu stały się do siebie równoległe (osie rotacji i skrętu). Skutkuje to tym, że obracając obiekt wokół własnej osi (dokonując rotacji), nie ma pewności czy obrót został wykonany wokół osi rotacji czy skrętu. Może to prowadzić do poważnych błędów interpretacyjnych rzutujących na dalsze obliczenia.



Rysunek 2.17: Obiekt znajdujący się w orientacji z blokadą przegubu. Pierścień zewnętrzny (czerwony) - unoszenie, środkowy (zielony) - rotacja, wewnętrzny (niebieski) - skręt

³⁶ osie w sekwencji mogą się powtarzać, więc na podstawie wzoru wariancji z powtórzeniami otrzymujemy $3^3 = 27$

Kwaterniony

Reprezentacja orientacji w postaci kątów Eulera jest intuicyjna z punktu widzenia interpretacji przez człowieka, ale w systemach komputerowych, ze względu na chociażby problem blokady przegubu, wymaga implementacji dodatkowych mechanizmów zabezpieczających przed niepoprawną interpretacją danych. W związku z tym, zasadnym jest wykorzystywanie w systemach komputerowych innych sposobów reprezentacji obrotów, które nie są obarczone ewentualnymi błędami interpretacyjnymi. Jednym z możliwych rozwiązań jest wykorzystanie kwaternionów (zbiór kwaternionów oznaczany jest w matematyce jako \mathbb{H}) – liczb z ciała liczb zespolonych, składających się z części rzeczywistej oraz części urojonej. Postać analityczną kwaternionu $q \in \mathbb{H}$ można przedstawić jako $q = q_0 + q_1i + q_2j + q_3k : q_0, q_1, q_2, q_3 \in \mathbb{R}$, gdzie $i^2 = j^2 = k^2 = -1$ [50]. Część urojona reprezentowana jest przez trzy liczby określające wektor obrotu (q_1, q_2, q_3) , zaś część rzeczywista składa się z jednej liczby q_0 określającej kąt obrotu wokół danego wektora obrotu (q_1, q_2, q_3) . Zaletą kwaternionów jest łatwość i zwięzłość zapisu operacji obrotu obiektu oraz składania sekwencji obrotów. Obarczone jest to jednak utratą możliwości łatwego odczytania przez człowieka wartości obrotów wokół każdej z osi układu współrzędnych. Kwaterniony przedstawione w postaci analitycznej rzadko są wykorzystywane w implementacjach komputerowych. W systemach komputerowych, często spotykana jest reprezentacja kwaternionów w postaci wektorów tak jak we wzorze 2.4:

$$\mathbf{q} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} q_w \\ q_x \\ q_y \\ q_z \end{bmatrix} = \begin{bmatrix} \cos(\frac{\alpha}{2}) \\ e_1 \sin(\frac{\alpha}{2}) \\ e_2 \sin(\frac{\alpha}{2}) \\ e_3 \sin(\frac{\alpha}{2}) \end{bmatrix} \quad (2.4)$$

gdzie $q_w, q_x, q_y, q_z, e_1, e_2, e_3 \in \mathbb{R}, \alpha \in [-\pi, \pi], \sqrt{e_1^2 + e_2^2 + e_3^2} = 1$.

Wartości kąta α oraz wektor $[e_1, e_2, e_3]$ są tożsame z reprezentacją w postaci Osie-Kąt.

Wzór 2.4 nie wyczerpuje jednak wszystkich reprezentacji kwaternionów jakie występują w literaturze. Wybierając notację dla zapisu kwaternionów należy szczególnie zadbać o określenie, która z wartości w wektorze odpowiada za część rzeczywistą, a które za część urojoną. W niniejszej pracy część rzeczywista reprezentowana jest przez element wektora oznaczony indeksem 0 lub w . Wzór 2.4 pokazuje też związek pomiędzy kwaternionem, a reprezentacją w postaci pary Osie-Kąt.

Aby skutecznie dokonywać obliczeń za pomocą kwaternionów muszą być one znormalizowane, czyli spełniać zależność $\|\mathbf{q}\| = 1$.

Opis operacji matematycznych oraz geometrycznych jakie wykonywane mogą być na kwaterionach można znaleźć w szeregu publikacji (na przykład [24]), a także w podręcznikach akademickich dotyczących matematyki.

Konwersje

Dla każdej z powyższych reprezentacji orientacji istnieje możliwość dokonania konwersji na dowolną inną. Taka operacja może być przydatna jeśli chcemy, prezentować aktualną orientację obiektu użytkownikowi, a postać, która jest używana przy obliczeniach nie jest intuicyjnie interpretowalna przez człowieka. Z uwagi na to, że w niniejszej pracy, do wyrażenia orientacji przestrzennej kości, wykorzystane były dwie formy reprezentacji: kąty Eulera oraz kwateriony, konwersja pomiędzy tymi dwoma reprezentacjami została przedstawiona w niniejszym rozdziale.

Przyjmując, że kąty Eulera wyrażone są przez sekwencję obrotów $\psi - \theta - \phi$ (ang. *Tait–Bryan angles*) oraz $[\cos(\frac{\psi}{2}), 0, 0, \sin(\frac{\psi}{2})]^T$, $[\cos(\frac{\theta}{2}), 0, \sin(\frac{\theta}{2}), 0]^T$, $[\cos(\frac{\phi}{2}), \sin(\frac{\phi}{2}), 0, 0]^T$ to kwateriony określające obroty o zadany kąt wokół każdej z osi układu współrzędnych indywidualnie, wówczas konwersja z reprezentacji za pomocą kątów Eulera do reprezentacji w postaci kwaterionów wyrażona jest wzorem 2.5:

$$q = \begin{bmatrix} q_w \\ q_x \\ q_y \\ q_z \end{bmatrix} = \begin{bmatrix} \cos(\frac{\psi}{2}) \\ 0 \\ 0 \\ \sin(\frac{\psi}{2}) \end{bmatrix} \begin{bmatrix} \cos(\frac{\theta}{2}) \\ 0 \\ \sin(\frac{\theta}{2}) \\ 0 \end{bmatrix} \begin{bmatrix} \cos(\frac{\phi}{2}) \\ \sin(\frac{\phi}{2}) \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \cos \frac{\phi}{2} \cos \frac{\theta}{2} \cos \frac{\psi}{2} + \sin \frac{\phi}{2} \sin \frac{\theta}{2} \sin \frac{\psi}{2} \\ \sin \frac{\phi}{2} \cos \frac{\theta}{2} \cos \frac{\psi}{2} - \cos \frac{\phi}{2} \sin \frac{\theta}{2} \sin \frac{\psi}{2} \\ \cos \frac{\phi}{2} \sin \frac{\theta}{2} \cos \frac{\psi}{2} + \sin \frac{\phi}{2} \cos \frac{\theta}{2} \sin \frac{\psi}{2} \\ \cos \frac{\phi}{2} \cos \frac{\theta}{2} \sin \frac{\psi}{2} - \sin \frac{\phi}{2} \sin \frac{\theta}{2} \cos \frac{\psi}{2} \end{bmatrix} \quad (2.5)$$

Zmiana sekwencji obrotów w postaci kątów Eulera wpływa na kolejność wykonywanych mnożeń kwaterionów w czasie konwersji.

Przy tożsamej sekwencji kątów Eulera, tak jak przy poprzednio przedstawionym przekształceniu, konwersja z postaci kwaterionowej ($q = [q_w, q_x, q_y, q_z]^T$) do kątów Eulera ($[\phi, \theta, \psi]^T$) wyrażona jest wzorem 2.6:

$$\begin{bmatrix} \theta \\ \phi \\ \psi \end{bmatrix} = \begin{bmatrix} \arcsin(2(q_w q_y - q_x q_z)) \\ \text{arctan2}(2(q_y q_z + q_w q_x), 1 - 2(q_x^2 + q_y^2)) \\ \text{arctan2}(2(q_x q_y + q_w q_z), 1 - 2(q_y^2 + q_z^2)) \end{bmatrix} \quad (2.6)$$

2.3. Charakterystyka wykorzystywanych urządzeń pomiarowych

W niniejszej pracy zostały wykorzystane dwa typy urządzeń: kamera RGB-D w postaci kontrolera ruchu Microsoft Kinect dla konsoli gier Microsoft XBox 360 oraz inercyjne czujniki ruchu (*ang. Inertial Measurement Unit - IMU*): akcelerometr i żyroskop. Ich powszechna dostępność sprawiła, że zagadnienie śledzenia ruchu ciała ludzkiego przestało być zarezerwowane jedynie dla profesjonalistów dysponujących drogim i bardzo rozbudowanym systemem śledzenia ruchu (*ang. Motion Capture*). Korzystając z obu typów urządzeń należy zwrócić szczególną uwagę na charakterystykę ich działania oraz ograniczenia takie jak wrażliwość na okluzje sensora głębi, niedokładność pomiarów, czy wręcz ich brak, dla niektórych stopni swobody czujników inercyjnych. Świadomość ta jest niezbędna przy budowaniu algorytmów śledzenia ruchu w oparciu o wymienione urządzenia.

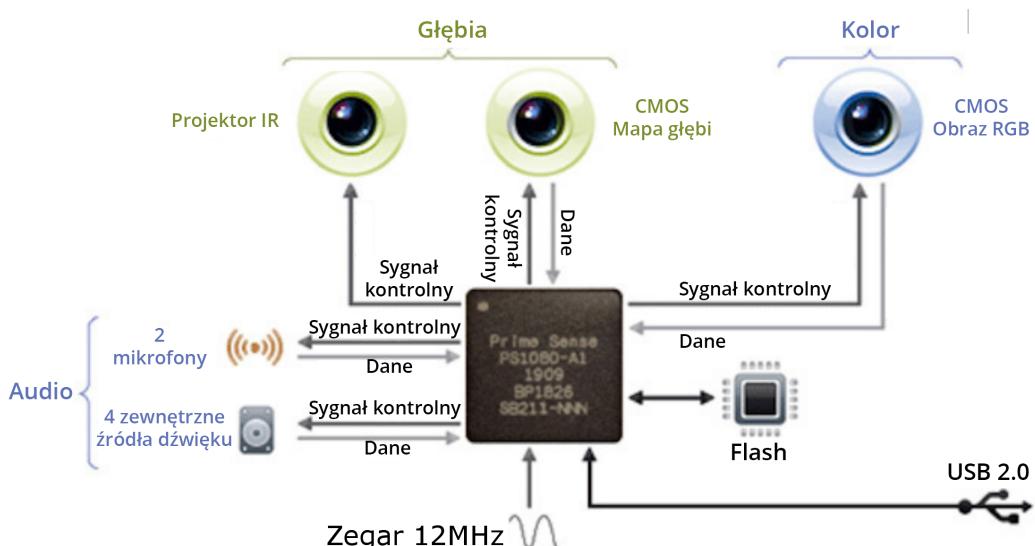
2.3.1. Kontroler Microsoft Kinect v. 1

Pierwsza wersja kontrolera ruchu Microsoft Kinect została oficjalnie udostępniona do sprzedaży rynkowej w 2010 roku (wcześniej prezentowano prototypy pod nazwą kodową "Project Natal"). Urządzenie to stało się pierwszym, powszechnie dostępnym, kontrolerem śledzącym ruch użytkownika bez wykorzystywania markerów, który odniósł faktyczny sukces komercyjny³⁷.

Budowa kontrolera

Śledzenie ruchu za pomocą urządzenia Kinect oparte jest o system wizyjny zbudowany z dwóch kamer z matrycami CMOS oraz z projektora światła podczerwonego. Jedna z kamer odpowiada za rejestrowanie tradycyjnego obrazu RGB, natomiast druga, wraz z projektorem, tworzy sensor głębi, który potrafi określić, w jakiej odległości od Kinecta znajdują się obiekty na scenie. Obie kamery działają z częstotliwością nie przekraczającą 30 Hz (*fps – ang. frames per second*). Ponadto urządzenie posiada cztery mikrofony, a głównym komponentem odpowiedzialnym za przetwarzanie i interpretację zarejestrowanych obrazów wejściowych jest dedykowany procesor opracowany i wyprodukowany przez izraelską firmę Prime Sense. Firma ta jest także autorem algorytmu rozpoznającego ruchy i gesty użytkownika na podstawie zarejestrowanych przez kamery sygnałów wejściowych. Rysunek 2.18 przedstawia uproszczony schemat budowy tego kontrolera.

³⁷ W lutym 2013 łączna sprzedaż kontrolera Microsoft Kinect osiągnęła 24 miliony sztuk (Z. Epstein, *Microsoft says Xbox 360 sales have surpassed 76 million units, Kinect sales top 24 million*, <https://goo.gl/SVfI3>)



Rysunek 2.18: Uproszczony schemat budowy kontrolera Microsoft Kinect v.1³⁸

Według oficjalnej dokumentacji³⁹ Kinect pracuje poprawnie, gdy obiekt (postać) znajduje się w odległości od 0.8m do 4m od urządzenia oraz zawiera się w polu widzenia urządzenia: 57° w orientacji poziomej i 43° w orientacji pionowej. Oficjalny zakres pracy urządzenia przedstawia rysunek 2.19. Dla odległości mniejszej niż 0.4m i większej niż 8m Kinect nie zwraca żadnych wyników pozwalających określić czy jakiś obiekt znajduje się przed nim, czy nie. Oficjalna dokumentacja nie zawiera jednak informacji o ewentualnej niejednorodności precyzyji działania w obszarze pracy urządzenia. Można więc przypuszczać, że urządzenie działa jednakowo precyzyjnie w dowolnym miejscu znajdującym się w zakresie roboczym urządzenia⁴⁰. Specyfikacja udostępniona na stronach MSDN⁴¹ przedstawia prawdopodobnie pewien uśredniony lub minimalny zakres pracy, ponieważ zgodnie z obserwacjami użytkowników⁴² oraz badaniami porównawczymi [30] okazuje się, że poszczególne serie produkcyjne urządzenia mogą mieć nieco inne zakresy pracy. Występowanie różnic w zakresie działania pomiędzy poszczególnymi egzemplarzami kontrolera Kinect zostało także przewidziane w oficjalnym SDK przygotowanym

³⁸ IFixIt, Xbox 360 Kinect teardown, <https://goo.gl/L9DxEu>.

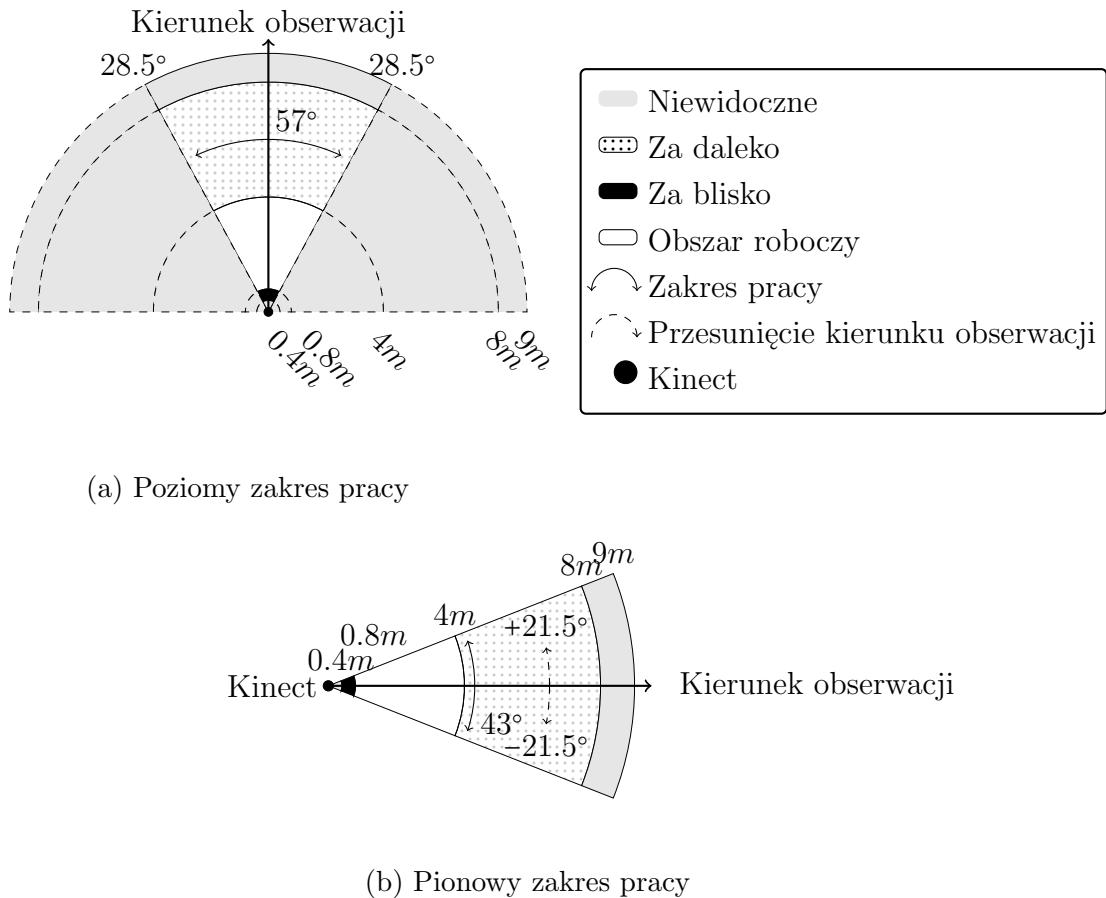
³⁹ Microsoft Corporation, *Kinect for Windows sensor components and specifications*, <https://goo.gl/08MjRG>

⁴⁰ Eksperymenty badawcze weryfikujące to przypuszczenie zostały zawarte w dalszej części niniejszego rozdziału

⁴¹ MSDN – *Microsoft Developer Network*, portal internetowy zawierający dokumentacje i poradniki głównie dla programistów korzystających z produktów i narzędzi firmy Microsoft. Dostępny pod adresem [http://https://msdn.microsoft.com](https://msdn.microsoft.com)

⁴² Stackoverflow Community, *Precision of the kinect depth camera*, <https://goo.gl/x4jWoX>.

przez firmę Microsoft⁴³. Analizując kod źródłowy SDK urządzenia można zauważać rozbieżność pomiędzy dokumentacją dotyczącą maksymalnych zakresów działania, a faktyczną implementacją. Jako przykład można podać wartość kąta widzenia kamery urządzenia w pionie, która według dokumentacji wynosi 43° , natomiast faktyczna implementacja podaje różne wartości dla kamer głębi i RGB, a wynoszą one odpowiednio 45.6° i 48.6° .



Rysunek 2.19: Poziomy (a) i pionowy (b) zakres pracy kontrolera Microsoft Kinect v.1 (źródło: opracowanie własne na podstawie³⁹⁾)

Wyznaczanie mapy głębi

Sposób w jaki Kinect określa odległość, w jakiej umieszczone są przed nim obiekty, opiera się o rozpoznawanie zniekształceń ściśle zdefiniowanego wzorca światlnego, którym została oświetlona scena (*ang. structured light*). W tym procesie biorą udział jedynie projektor oraz kamera światła podczerwonego, natomiast obraz z kamery RGB nie jest wykorzystywany na żadnym etapie określania odległości. Technika ta jest wykorzystywana powszechnie

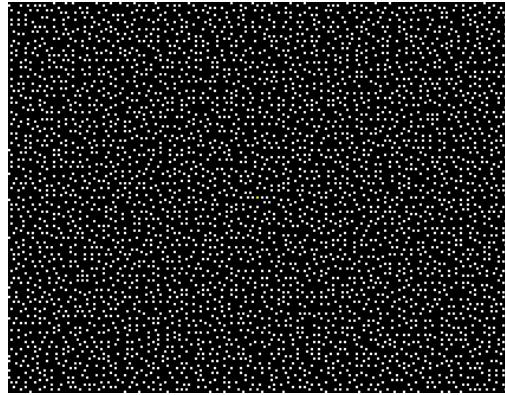
⁴³ Microsoft Corporation, *Kinect for Windows SDK 1.8—Constants*, <https://goo.gl/L3W5Pf>.

w skanerach 3D^{44,45}. Dokładne działanie kontrolera Kinect nie zostało oficjalnie ujawnione publicznie przez firmę Microsoft. Możliwe jest jednakże przedstawienie przybliżonego działania tego kontrolera dzięki niezależnym badaniom (na przykład MacCormick[74]) oraz analizie zgłoszeń patentowych[34, 105, 106] należących do producenta procesora wykorzystanego do budowy kontrolera Kinect – firmy Prime Sense.

W pierwszym kroku cała scena zostaje oświetlona zbiorem punktów za pomocą projektora światła podczerwonego (rys. 2.20) zgodnie ze ścisłe określonym wzorcem (rys. 2.21). Wzorzec ten nie został nigdy oficjalnie opublikowany przez firmy Prime Sense, czy Microsoft, natomiast wzorzec zamieszczony na rysunku 2.21 został wyznaczony przez Andreasa Reichingera, doktoranta na Uniwersytecie Technicznym w Wiedniu. Na swoim blogu opublikował on dokładny opis sposobu wyznaczania zamieszczonego wzorca⁴⁶.



Rysunek 2.20: Scena oświetlona promieniami IR⁴⁷



Rysunek 2.21: Wzorzec oświetlenia sceny przez kontroler Kinect⁴⁶

Pełen wzorzec zbudowany jest z dziewięciu powtarzających się regionów w układzie 3x3 o rozmiarze 211x165 punktów, co daje pełen rozmiar wzorca równy 633x495 punktów. Jasne punkty wyznaczone przez wiązki światła podczerwonego stanowią około 11% powierzchni tego wzorca, przy czym ich jasność nie jest jednorodna. Wyraźnie można zauważyć, że część punktów jest jaśniejsza od pozostałych. Ponadto w punkcie centralnym każdego z regionów (jak i całego wzorca) znajduje się najjaśniejszy z punktów. Dodatkowo wzorzec ten jest skośnie symetryczny, co oznacza niewrażliwość na obrócenie urządzenia o 180°.

⁴⁴ DAVID Vision Systems GmbH, *DAVID - 3D scanner*, <https://goo.gl/ohPxk2>.

⁴⁵ LMI Technologies, *Home—LMI Technologies*, <https://goo.gl/hyxbiI>.

⁴⁶ A. Reichinger, *Kinect pattern uncovered*, <https://goo.gl/D0qRxZ>

⁴⁷ J. L. Flatley, *Visualized: Kinect + night vision = lots and lots and lots of dots (video)*, <https://goo.gl/HuzK3>.

Następnym krokiem jest wyznaczenie mapy głębi na podstawie analizy zniekształcenia wyświetlonego wzorca. Opierając się na badaniach prowadzonych przez MacCormicka [74], analiza zniekształcenia wzorca bazuje na dwóch technikach:

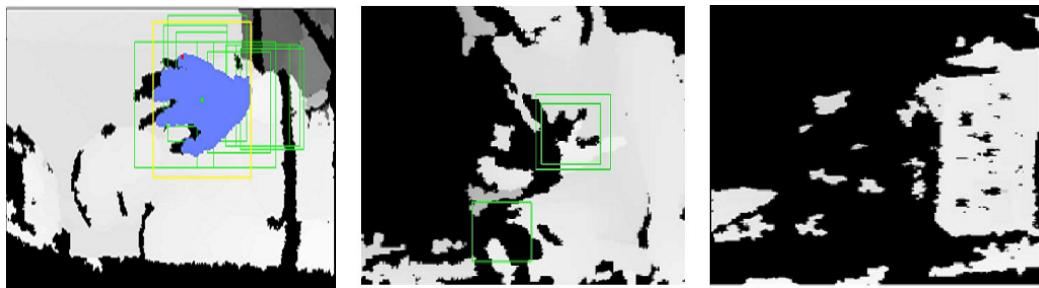
- wyznaczenie głębi na podstawie ostrości punktów,
- wyznaczenie głębi na podstawie zjawiska paralaksy.

Pierwsza z technik opiera się na fakcie, że obraz obiektów umieszczonych bliżej soczewki jest mniej wyostrzony od tych znajdujących się dalej. Dodatkowo soczewki umieszczone w kamerze IR Kinecta posiadają różną ogniskową dla osi "X" i "Y", co sprawia, że okrągłe punkty stają się eliptyczne z pochyleniem zależnym od głębokości w jakiej znajduje się dany obiekt. Druga z zastosowanych technik opiera się na zjawisku paralaksy, które występuje w wyniku przesunięcia pomiędzy projektorem światła podczerwonego a kamerą wyposażoną w filtr światła podczerwonego. Szczegółowe opisy działania obu technik można znaleźć w wielu publikacjach naukowych m.in.: Rzeszotarski, Strumiłło i in. [95], czy Fofi, Sliwa i Voisin [33]. Informacje uzyskane na podstawie obu powyższych technik są następnie ze sobą łączone w celu uzyskania mapy głębi.

Ograniczenia w działaniu kontrolera Kinect

Kinect przez ponad pięć lat dostępności na rynku jest obiektem badań związanych zarówno z jego zastosowaniem i poprawą działania, jak i dokładnym opisaniem jego cech, oszacowaniem dokładności oraz zdefiniowaniem ograniczeń. Jednym z podstawowych ograniczeń jakie występują w tym urządzeniu jest wrażliwość na światło słoneczne. Jak to zostało wcześniej opisane, jedynym źródłem danych wykorzystywanym do budowy mapy głębi, a także do stworzenia modelu szkieletowego jest kamera z filtrem podczerwieni oraz wzorzec punktów podczerwonych emitowanych przez wbudowany projektor. Światło słoneczne natomiast składa się z pełnego spektrum barw, zarówno z zakresu widzialnego jak i niewidzialnego dla człowieka. Oznacza to, że promienie słoneczne mogą wprowadzić dodatkowe punkty na scenie, które nie należą do wzorca, a będą zarejestrowane przez kamerę. Zarejestrowanie przez kamerę podczerwoną kontrolera Kinect punktów spoza przyjętego wzorca prowadzi do niepoprawnego zbudowania mapy głębi wykorzystywanej w procesie wyznaczania modelu szkieletowego. Rysunek 2.22 przedstawia mapę głębi zbudowaną wewnętrz pomieszczenia 2.22a oraz na zewnątrz w dwóch przypadkach, w miejscu zacienionym 2.22b oraz w pełnym słońcu 2.22c.

Jak widać na przykładzie z rysunku 2.22, Kinect nie był w stanie zbudować prawidłowej mapy w przypadku oświetlenia sceny światłem słonecznym.



(a) Wewnątrz pomieszczenia
 (b) Na zewnątrz pomieszczenia w cieniu
 (c) Na zewnątrz pomieszczenia w świetle

Rysunek 2.22: Mapa głębi ręki w różnych warunkach oświetleniowych[111].

Z podobnych względów niezalecane jest używanie wielu Kinectów do śledzenia ruchu na jednej scenie. Co prawda SDK wspiera wykorzystanie aż czterech urządzeń równocześnie, jednak zgodnie z oficjalną dokumentacją może to negatywnie wpływać na dokładność rozpoznawania części ciała, a przez to wyznaczenia modelu szkieletowego i śledzenia ruchu stawów⁴⁸. Scenariusz wykorzystania wielu kontrolerów Microsoft Kinect w przypadku jednej sceny, jest jednym z obszarów badawczych wśród naukowców. Po zastosowaniu odpowiednich algorytmów wspomagających proces kalibracji i synchronizacji [63], czy wprowadzających elementy uczenia maszynowego (rozpoznawanie szkieletu z wykorzystaniem klasyfikatora HCRF - ang. *Hidden-state Conditional Random Fields*) [62], badacze osiągnęli możliwość stabilnej pracy układu kilku kontrolerów Kinect z dokładnością wyższą niż w przypadku pojedynczego kontrolera. Badając możliwości wykorzystania kilku kontrolerów Kinect do obserwacji jednej sceny badacze byli w stanie poprawić także działanie systemu śledzenia z wykorzystaniem jednego urządzenia. Na przykład, Astriadi [8] oraz Baek [10] opisali metody wykorzystania dwóch kontrolerów Kinect, dzięki którym zmniejszyli oni podatność prezentowanego systemu śledzenia na występowanie okluzji. To z kolei przekłada się na dokładność oszacowania położenia stawów. Shroeder [100] zaproponował natomiast metodę poprawiającą jakość wyznaczonej mapy głębi. Według opublikowanych danych zmniejszył on liczbę pikseli niepoprawnie określających głębię z około 20% do 1%.

Kolejnym z ograniczeń, które ma znaczący wpływ na możliwości systemu opartego o kontroler Microsoft Kinect, jest niepełna informacja o wykonywanych obrotach. Kontroler Kinect, na podstawie swoich pomiarów, estymuje

⁴⁸ Microsoft Corporation, *Skeleton tracking with multiple Kinect sensors*, <https://goo.gl/hk1Cdt>.

trójwymiarowy model szkieletowy śledzonej postaci, w którym pozycja każdego stawu opisana jest przez trójkę liczb pozwalającą określić jednoznacznie położenie tego stawu. Twórcy tego urządzenia przyjęli, że z każdym stawem powiązany jest tylko jeden punkt i jest on powiązany ze swoim sąsiednim stawem definiując w ten sposób kość. Na podstawie informacji udostępnionych przez kontroler Kinect, możliwe jest określenie jaki obrót wokół osi układu współrzędnych został wykonany przez śledzoną postać. Jednak informacja o obrotach ograniczona jest jedynie do dwóch kątów. Przyjmując notację kątów Eulera zaprezentowaną na rysunku 2.16a, możliwe jest określenie obrotów wokół osi skrętu i unoszenia, natomiast brakuje informacji dotyczącej obrotu wokół osi rotacji.

Innym ograniczeniem, jakie należy wziąć pod uwagę, jest zasłanianie stawów śledzonej postaci. Okluzja może wystąpić z dwóch powodów:

- przysłonięcie przez obiekty znajdujące się na scenie, na przykład meble,
- przysłonięcie przez inne części ciała.

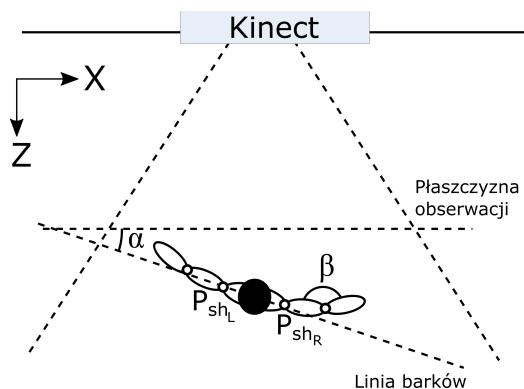
W tym miejscu należy wspomnieć o informacji przypisanej do każdego ze stawów, określającej stan jego śledzenia. Informacja ta może przyjąć jedną z trzech wartości:

- *Tracked* – wartość przyjmowana kiedy dany staw jest w pełni widoczny i śledzony bez żadnych znaczących zakłóceń;
- *Interferred* – wartość przyjmowana kiedy Kinect nie jest w stanie śledzić danego stawu, ale na podstawie innych wartości jest w stanie oszacować jego położenie. Algorytm jakim w tej sytuacji posługuje się Kinect do szacowania położenia nie został opublikowany;
- *NotTracked* – wartość przyjmowana kiedy dany staw jest zupełnie niewidoczny i nie ma możliwości podania nawet jego przybliżonego położenia;

W praktyce, położenie stawów jest najczęściej oznaczone jako *Interferred*, co oznacza, że zanim otrzymane dane zostaną użyte, należy dodatkowo weryfikować, na przykład, czy położenie danego stawu pomiędzy kolejnymi pomiarami jest prawdopodobne, lub wystąpił jednak błąd szacowania. Sposobem na taką weryfikację może być wyznaczenie stopnia przesunięcia danego stawu pomiędzy kolejnymi pomiarami. Przy częstotliwości akwizycji danych przez kontroler Kinect, która wynosi 30 klatek na sekundę, czas jaki mija pomiędzy dwiema kolejnymi klatkami to około 30ms. Tak krótki czas sprawia, że przesunięcia stawów przy normalnym ruchu postaci są realnie niewielkie. Jeśli zatem wyznaczone przesunięcie będzie wynosiło na przykład 1 metr, istnieje duże prawdopodobieństwo, że położenie stawu zostało oszacowane błędnie.

Określając czy dany pomiar Kinecta można uznać za prawidłowy albo błędny, warto rozważyć go w pewnym kontekście dotyczącym wykonywanych ruchów. Jako przykład można tutaj podać sytuację, w której ruch był wykonywany raczej płynnie z niewielkimi przesunięciami stawów i nagle z otrzymanych pomiarów wynika, że nastąpiło krótkotrwałe, znaczco większe przesunięcie położenia stawu, kontynuowane ponownie niewielkimi przesunięciami. W takim wypadku, skokowej zmiany oszacowania położenia, również można przypuszczać, że nastąpiło błędne oszacowanie położenia stawów.

Szczególne przypadki okluzji występują, gdy użytkownik skieruje wyprostowaną kończynę w kierunku kamery Kinecta lub obróci się cała jego sylwetka względem płaszczyzny obserwacji kontrolera Kinect. Analizując pierwszy scenariusz najłatwiej przedstawić go na przykładzie ruchu ręki. Trzymając ją na wprost kamery kontrolera Kinect, wyprostowaną w łokciu, w pełni widoczny jest tylko staw nadgarstkowy, a także dłoń. Praktycznie całkowitemu przysłonięciu ulegają zaś łokieć i staw barkowy.



Rysunek 2.23: Kąt obrotu (α) użytkownika pomiędzy linią barków a płaszczyzną obserwacji Kinecta

W trakcie badań własnych przeprowadzony został eksperyment, którego celem było oszacowanie przy jakim kącie obrotu ciała śledzonej postaci względem kontrolera Kinect następuje praktyczna utrata zdolności śledzenia wybranych stawów przez kontroler. Eksperyment polegał na obracaniu się postaci wokół własnej osi w tak zwanej pozycji *T-pose*, co oznacza, obie ręce wyprostowane w łokciach i wyciągnięte w bok, tak że sylwetka przypomina literę „T”. W trakcie wykonywania ruchu mierzony był kąt obrotu α pomiędzy linią barków (P_{Sh_L} , P_{Sh_R}), a płaszczyzną obserwacji (rys. 2.23). Kąt ten jest wyznaczony w przestrzeni dwuwymiarowej i wyrażony jest wzorem 2.7.

$$\alpha = \begin{cases} \text{atan}\left(\frac{|P_{Sh_R,Z}^K - P_{Sh_L,Z}^K|}{|P_{Sh_R,X}^K - P_{Sh_L,X}^K|}\right) & , |P_{Sh_R,X}^K - P_{Sh_L,X}^K| \neq 0 \\ \frac{\pi}{2} & , |P_{Sh_R,X}^K - P_{Sh_L,X}^K| = 0 \end{cases} \quad (2.7)$$

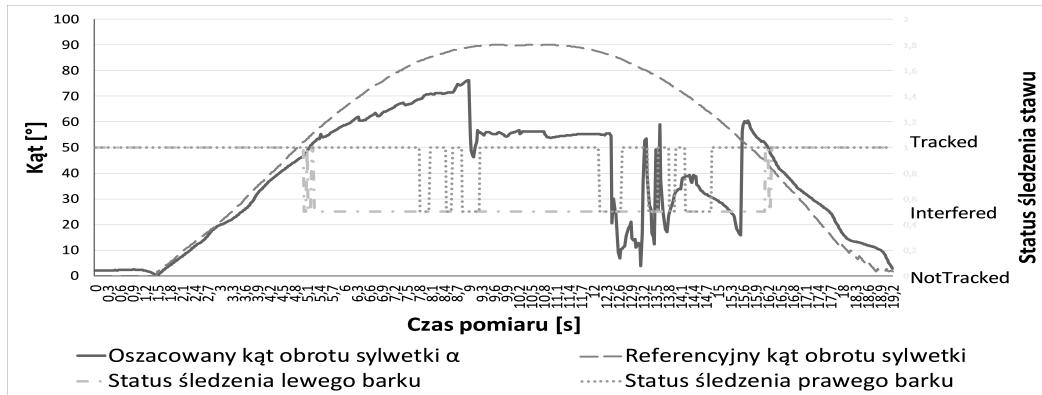
gdzie: $P_{Sh_L,X}^K, P_{Sh_L,Z}^K$ to współrzędne pobrane z kontrolera Kinect, określające położenie lewego barku użytkownika w osiach X i Z, natomiast $P_{Sh_R,X}^K, P_{Sh_R,Z}^K$ są współrzędnymi określającymi położenie prawego barku użytkownika. Wszystkie współrzędne użyte w tym wzorze wyrażone są w układzie współrzędnych Kinecta (rys. 2.15).

W wyniku eksperymentu okazało się, że przy kącie obrotu α powyżej 50° kontroler Kinect traci z pola widzenia blisko połowę stawów użytkownika, a to z kolei powoduje, że pomiar wszystkich pozostałych stawów staje się niewiarygodny. Objawia się to dużymi różnicami w oszacowaniu pozycji stawów pomiędzy dwoma kolejnymi pomiarami. Co ciekawe, różnice te są także zauważalne dla stawów oznaczonych jako *Tracked*, czyli teoretycznie takich, które są w pełni widoczne dla kontrolera. Na tej podstawie można wysunąć dwa wnioski. Po pierwsze wyznaczenie pozycji poszczególnych stawów odbywa się z wykorzystaniem informacji o pozostałych stawach. W innym przypadku stawy oznaczone jako *Tracked* zawsze byłyby pozycjonowane w poprawny sposób. Drugi wniosek jaki się nasuwa dotyczy pól wykorzystywanych jako zestaw uczący dla algorytmu losowych lasów decyzyjnych. Pozy wykorzystywane do nauczenia Kinecta prawidłowego rozpoznawania części ciała człowieka przedstawiają prawdopodobnie postać stojącą na wprost, względnie obróconą w niewielkim stopniu do płaszczyzny obserwacji. Stąd problemy kontrolera Kinect w określaniu położenia nawet widocznych stawów w sytuacji, gdy śledzona postać jest obrócona.

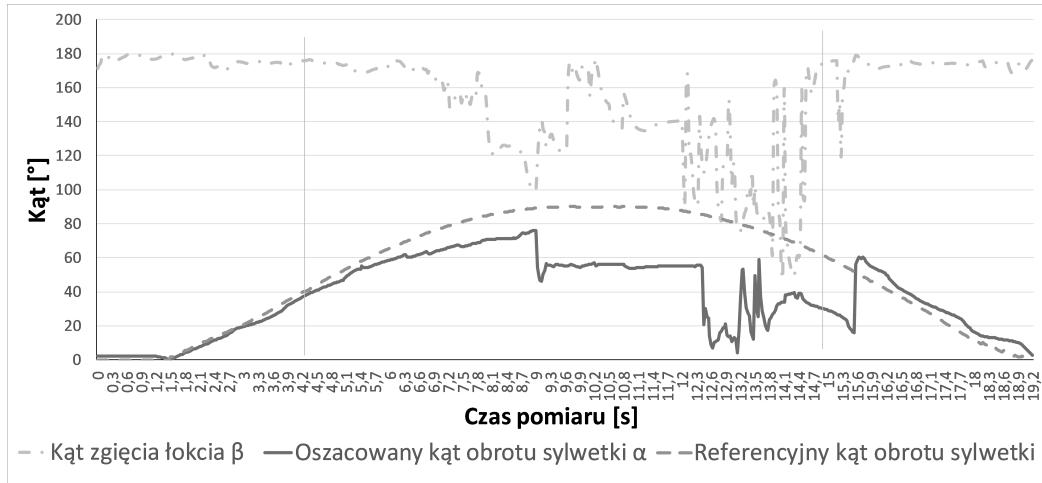
Wykres na rysunku 2.24a przedstawia status śledzenia lewego i prawego barku (linie przerywane) w zależności od kąta obrotu α wyznaczonego na podstawie pomiarów położenia stawów barkowych przez kontroler Kinect (linia ciągła). Stan śledzenia stawu: *Tracked*, *Interferred* oraz *NotTracked* oznaczony jest na prawej osi pionowej. Obrót wykonywany był w kierunku przeciwnym do ruchu wskaźówek zegara co oznacza, że prawa strona ciała była cały czas widoczna dla urządzenia śledzącego ruch. Jak widać na tym wykresie, przy kącie zbliżonym do 50° status śledzenia prawego stawu łokciowego zaczyna zmieniać się gwałtownie, aby ustalić swoją wartość na *Interferred* po przekroczeniu rzeczonego kąta. Warto również zauważyć, że w przedziale od 50° do 90° występowały duże wahania oszacowania wartości kąta obrotu α , co jest spowodowane błędnym szacowaniem położenia lewego barku.

Wykres przedstawiony na rysunku 2.24b pokazuje również zmianę wartości kąta mierzonego w łokciu prawej ręki β (rys. 2.23) w trakcie wykonywania obrotu. Ręka przez cały czas trwania obrotu była wyprostowana w łokciu i stanowiła przedłużenie linii pomiędzy barkami (*T-pose*). Jak widać na wykresie, pomimo pełnej widoczności prawej ręki przez kamery Kinecta przy obrocie w zakresie $50^\circ < \alpha \leq 90^\circ$, pomiar kąta β jest niestabilny, a jego

wartość nie jest nawet zbliżona do wartości kąta wyprostowanej w łokciu ręki ($170^\circ - 180^\circ$).



(a) Stan śledzenia stawów podczas obrotu ciała (*T-pose*) względem kontrolera



(b) Kąt w łokciu prawej ręki podczas obrotu ciała (*T-pose*) względem kontrolera

Rysunek 2.24: Pomiary zależne od kąta obrotu względem Kinecta (źródło: badania własne)

Szacowanie odległości pomiędzy kontrolerem Kinect, a śledzona postacią

Kolejną istotną cechą kontrolera Kinect jest zmieniająca się dokładność szacowania głębi w zależności od odległości, w jakiej znajduje się śledzona postać. Specyfikacja urządzenia nie zawiera informacji o niejednorodności dokładności pomiaru, a jedynie określa zakres w jakim działa urządzenie. W trakcie własnych badań, prowadzonych nad charakterystyką kontrolera Kinect, zaobserwowano zmianę dokładności szacowania głębi wraz z oddalaniem się użytkownika od urządzenia. Eksperyment, dzięki któremu udało się zaobserwować omawianą zmianę dokładności szacowania odległości, polegał na płynnym oddalaniu się od kontrolera Kinect, rejestrując uzyskany dzięki

niemu pomiar odległości stawu barkowego oraz równoczesny pomiar tej samej odległości za pomocą systemu śledzenia Vicon. Wykres z rysunku 2.25 przedstawia różnicę pomiędzy pomiarami uzyskanymi za pomocą kontrolera Kinect oraz tymi z systemu Vicon. Pomiary z systemu Vicon posłużyły jako dane referencyjne dla odległości pomiędzy postacią a kontrolerem Kinect oraz zostały wykorzystane jako dane na osi odciętych. Jak widać na wykresie z rysunku 2.25 zmiana dokładności oszacowania odległości, w jakiej znajduje się śledzona osoba od kontrolera Kinect, nie jest funkcją prostoliniową i wraz z oddalaniem się postaci od urządzenia pomiarowego zmienia się od niedoszacowania do przeszacowania wartości. Funkcja reprezentująca przybliżony model dokładności oszacowania odległości przez kontroler Kinect, ma postać wielomianu 3-ego rzędu ($y = a_0 + a_1x + a_2x^2 + a_3x^3$).

$$X = \begin{bmatrix} x_1^0 & x_1^1 & x_1^2 & x_1^3 \\ x_2^0 & x_2^1 & x_2^2 & x_2^3 \\ x_3^0 & x_3^1 & x_3^2 & x_3^3 \\ \dots \\ x_n^0 & x_n^1 & x_n^2 & x_n^3 \end{bmatrix}, A = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}, Y = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix} \quad (2.8)$$

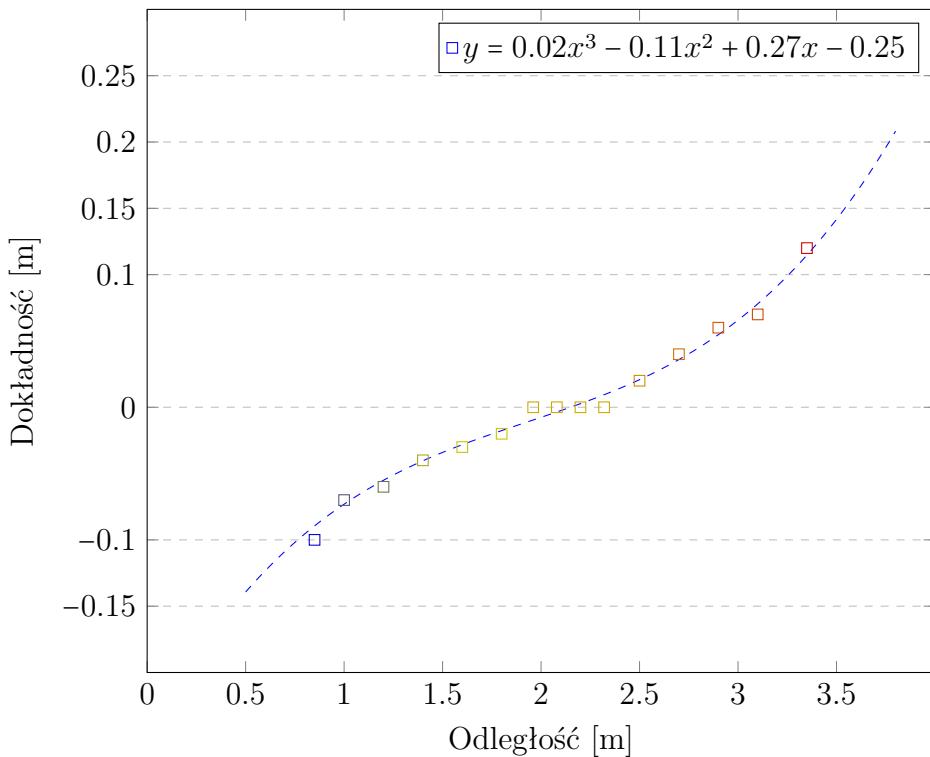
$$X^T X A = X^T Y$$

Wzór 2.8 przedstawia układ równań, w postaci macierzowej, pozwalający wyznaczyć współczynniki a_0, a_1, a_2, a_3 szukanego wielomianu na podstawie n próbek pomiarowych. Macierze X i Y wykorzystane we wzorze 2.8 zawierają odpowiednio argumenty i wartości dla tych argumentów, na podstawie których możliwe będzie wyznaczenie współczynników wielomianu zawartych w macierzy A . W przypadku omawianego eksperymentu argumentami są wybrane odległości od Kinecta, w jakich znajdowała się śledzona postać, natomiast wartościami są różnice pomiędzy oszacowaniem Kinecta, a odlegością uzyskaną dzięki pomiarom systemu Vicon we wspomnianych punktach (argumentach). Na podstawie pomiarów uzyskanych w trakcie eksperymentu opisanego powyżej, został wyznaczony następujący wektor współczynników:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} -0.25 \\ 0.27 \\ -0.11 \\ 0.02 \end{bmatrix} \quad (2.9)$$

Należy zauważyć, że zaprezentowana na wykresie z rysunku 2.25 charakterystyka oszacowania odległości pomiędzy kontrolerem Kinect a użytkownikiem dotyczy obszaru roboczego tego kontrolera (rys. 2.19). W związku

z tym, dokładność oszacowania pozycji poszczególnych stawów w modelu szkieletowym użytkownika jest uzależniona od tego, gdzie on się znajduje i należy to uwzględnić w prowadzonych obliczeniach, aby wykazane niedokładności skorygować. Jest to o tyle istotne, że na podstawie uzyskanych wyników, przestrzeń, w której szacowanie odległości pomiędzy użytkownikiem a Kinectem jest zbliżone do faktycznej odległości, wynosi jedynie $0.3m$ pomiędzy $2m$, a $2.3m$ od kamery.



Rysunek 2.25: Dokładność pomiaru głębi w zależności od odległości śledzonej postaci od kontrolera Kinect (źródło: badania własne)

2.3.2. Urządzenia inercyjne

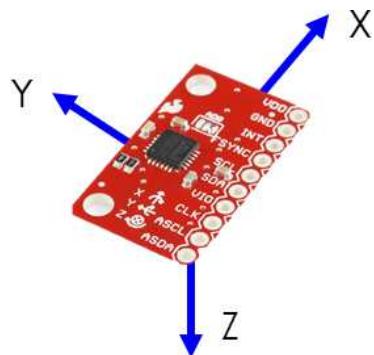
Urządzenia inercyjne są czujnikami, których działanie opiera się na mierzeniu wielkości fizycznych, jakie działają na te czujniki w trakcie ruchu oraz w spoczynku. Wyróżniamy dwa podstawowe czujniki inercyjne: akcelerometr mierzący przyspieszenia liniowe oraz żyroskop mierzący prędkości kątowe. Choć urządzenia te są znane i stosowane od dawna, dopiero z chwilą ich implementacji w formie mikroukładów elektromechanicznych (MEMS – *microelectromechanical systems*) stały się one powszechnie dostępne i stosowane na przykład w telefonach, czy systemach nawigacji inercyjnej (INS – *indoor*

navigation systems). Przykład zastosowania mikroukładów inercyjnych w nawigacji został zaprezentowany w pracy [12, 13]. W badaniach przedstawionych w niniejszej pracy wykorzystany został układ o symbolu MPU-6050 firmy InvenSense integrujący oba te czujniki (tabela 2.4 zawiera podstawowe parametry pracy tego układu). Układ ten oprócz wcześniej wspomnianych czujników posiada wbudowany termometr oraz procesor ruchu pozwalający na przetworzenie danych pomiarowych i udostępnienie ich w postaci kwaternionu reprezentującego orientację układu w przestrzeni. Algorytm wyznaczania orientacji układu, zaimplementowany w tym procesorze, nie został jednak opublikowany w żadnej dokumentacji, w związku z czym funkcjonalność ta nie była wykorzystywana w badaniach.

Tablica 2.4: Zestawienie podstawowych parametrów pracy układu MPU-6050⁴⁹

Osie pomiaru:	X, Y, Z (układ współrzędnych widoczny na rys. 2.26)
Interfejs komunikacyjny:	I2C (TWI) - 400 kHz
Rozdzielcość danych:	16-bitów dla każdej osi
Zakres pomiarowy akcelerometru:	programowalny: $\pm 2g$, $\pm 4g$, $\pm 8g$, $\pm 16g$
Zakres pomiarowy żyroskopu:	programowalny: $\pm 250^\circ/s$, $\pm 500^\circ/s$, $\pm 1000^\circ/s$, $\pm 2000^\circ/s$

We wszystkich badaniach w niniejszej pracy, w których wykorzystywane są czujniki inercyjne, są one skonfigurowane do pracy w zakresie $\pm 4g$ dla akcelerometru oraz $\pm 500^\circ/s$ dla żyroskopu. Zakresy te zostały wybrane na podstawie specyfikacji komercyjnie dostępnego urządzenia pozwalającego na śledzenie ruchu – Nintendo WiiRemote.



Rysunek 2.26: Układ współrzędnych wykorzystywany w czujnikach MPU-6050 (źródło: opracowanie własne)

Akcelerometr

Akcelerometr mierzy siłę działającą na niego w dowolnym z trzech kierunków, wyrażoną w odniesieniu do siły grawitacji w jednostkach "g" ("g" jest to jednostka standardowego przyspieszenia ziemskiego. $1g = 9.80665 \text{ m/s}^2$). Pozwala to na szacunkowe określenie z jakim przyspieszeniem wzdłuż każdej osi porusza się dany czujnik. Ze względu na to, że siła grawitacji działa na wszystkie obiekty nieustannie, akcelerometr leżący horyzontalnie na płaskiej powierzchni i nie poruszający się, będzie wskazywał wartość $1g$ wzdłuż osi równoległej do wektora grawitacji i $0g$ w pozostałych dwóch osiach. Pomijając wpływ szumów na odczyty pomiarów z akcelerometru, jeśli dla nieporuszającego się czujnika wartości pomiaru będą inne niż podane powyżej, oznacza to, że czujnik jednak nie leży w pozycji horyzontalnej. Pomiar wynoszący $0g$ we wszystkich trzech osiach oznacza, że dany obiekt spada swobodnie w orientacji idealnie poziomej.

Zastosowany w tym układzie akcelerometr jest czujnikiem typu pojemościowego. Oznacza to, że jego wewnętrzna struktura opiera się na układzie ruchomych półprzewodnikowych belek ułożonych pomiędzy analogicznymi, nieruchomymi belkami, które tym samym tworzą układ kondensatorów (rys. 2.27). Pomiędzy dwiema sąsiednimi nieruchomoymi belkami znajduje się dokładnie jedna belka ruchoma tworząc w ten sposób parę dwóch kondensatorów o zmiennych pojemnościach C_1 i C_2 . Siła działająca na czujnik powoduje przesunięcie się ruchomej belki w kierunku przyłożonej siły o dystans d . Przyjmując dystans pomiędzy nieruchomoymi belkami a belką ruchomą w czasie spoczynku jako d_0 , to przy powierzchni s belek i stałej dielektrycznej ϵ materiału, z którego wykonane są belki, pojemności C_1 i C_2 określone są wzorami 2.10a i 2.10b:

$$C_1 = \frac{\epsilon s}{d_0 + d} \quad (2.10a)$$

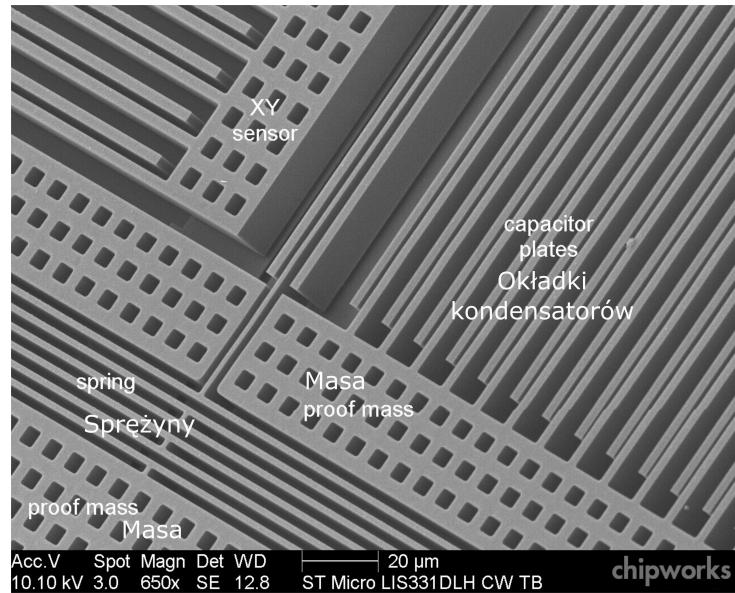
$$C_2 = \frac{\epsilon s}{d_0 - d} \quad (2.10b)$$

Następnie, na podstawie wartości C_1 i C_2 oraz wartości napięcia elektrycznego jaki zasila układ akcelerometru, możliwe jest wynaczenie wartości napięcia wypływającego z takiej pary kondensatorów V_C przy podanym napięciu wejściowym V_0 za pomocą wzoru 2.11:

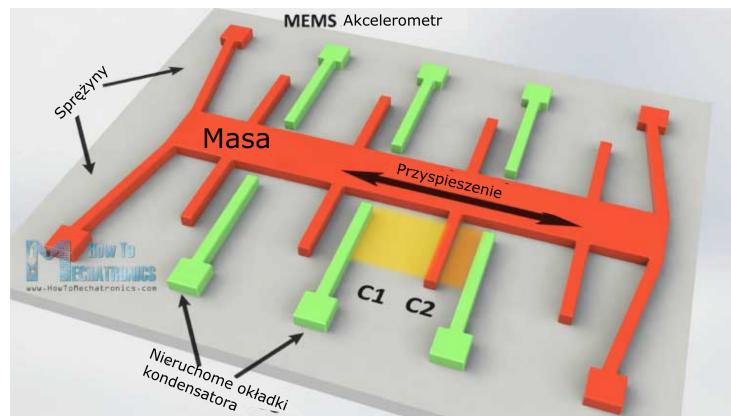
$$V_C = V_0 \frac{C_2 - C_1}{C_2 + C_1} = V_0 \frac{d}{d_0} \quad (2.11)$$

⁵⁰ S. Dixon-Warren, *Motion sensing in the iPhone 4: MEMS accelerometer*, <https://goo.gl/6m06U8>.

⁵¹ D. Nedelkovski, *MEMS accelerometer gyroscope magnetometer and Arduino*, <https://goo.gl/JDYruw>



(a) Struktura wewnętrzna⁵⁰



(b) Wizualizacja budowy⁵¹

Rysunek 2.27: Struktura wewnętrzna akcelerometru w technologii MEMS

Ponieważ wartość napięcia V_C jest zależna od wartości przemieszczenia d ruchomej belki w układzie dwóch kondensatorów, wartość ta odzwierciedla siłę jaka działa na dany czujnik. Bezpośredni odczyt wartości V_C musi być jednak poddany dodatkowemu przetwarzaniu w celu określenia kierunku w jakim zadziałała siła. Kierunek działania siły na czujnik, odzwierciedlony jest w postaci znaku + albo - przy wartości V_C . Następnie możliwe jest wyznaczenie wartości przyspieszenia a działającego na dany czujnik. W tym celu należy wykorzystać fakt, że siła sprężystości F_s dla małych odkształceń, na podstawie prawa Hooke'a, jest wprost proporcjonalna do odkształcenia sprężyny (w tym wypadku do przemieszczenia ruchomej belki d). Oznaczając

przez f_s współczynnik sprężystości charakterystyczny dla materiału, z którego zbudowana jest dana sprężyna, siłę sprężystości można zdefiniować wzorem 2.12.

$$F_s = f_s d \quad (2.12)$$

Wykorzystując drugie prawo Newtona możemy przedstawić związek pomiędzy siłą (F_s) działającą na ciało, jego masą m oraz jego przyspieszeniem a . Na tej podstawie wzór 2.12 można przekształcić do postaci określonej wzorem 2.13.

$$F_s = ma = f_s d \quad (2.13)$$

Następnie łącząc ze sobą wzory 2.11 oraz 2.13 można wyznaczyć wartość przyspieszenia zgodnie ze wzorem 2.14.

$$ma = f_s d \Rightarrow a = \frac{f_s}{m} d = \frac{f_s}{m} \frac{d_0 V_C}{V_0} \quad (2.14)$$

Szczegółowy opis budowy i działania akcelerometrów zbudowanych w architekturze MEMS można znaleźć w pracy Mateja Andrejasica z Uniwersytetu w Lublanie [6].

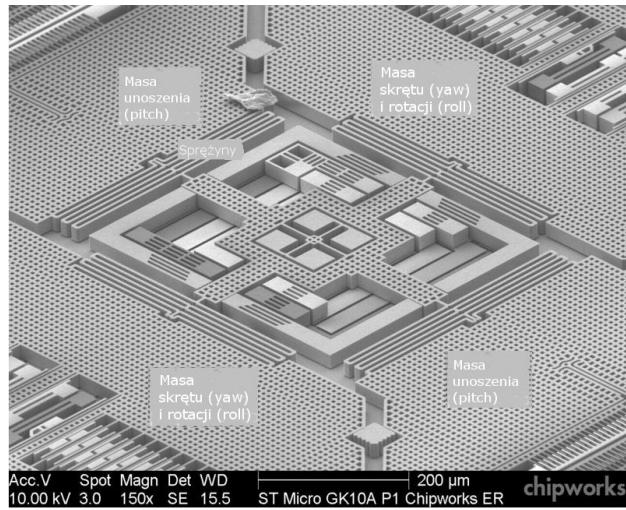
Żyroskop

Budowa wykorzystanego żyroskopu, podobnie jak akcelerometru, opiera się na układzie kondensatorów o zmiennej pojemności, jednak w tym przypadku są one dodatkowo wprawione w drgania. Działanie takiego żyroskopu opiera się na efekcie Coriolisa⁵¹. Budowę i uproszczony sposób działania czujnika przedstawia rysunek 2.28. Wykorzystane urządzenie mierzy prędkość kątową wokół trzech osi, wyrażoną w stopniach na sekundę ($^{\circ}/s$), odpowiadającą konwencji nazewnicy związaną z reprezentacją obrotów w postaci kątów Eulera (rozdział 2.2.4).

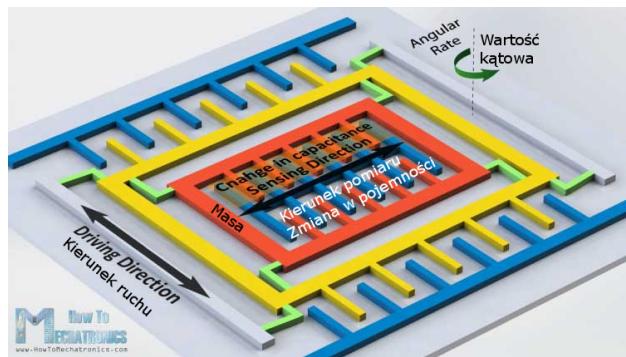
2.3.3. Ograniczenia w działaniu czujników inercyjnych

Ograniczenia w działaniu obu czujników inercyjnych można rozdzielić na te dotyczące ogólnie urządzeń stworzonych w architekturze MEMS, a tym samym wynikające z ich budowy i właściwości fizycznych, oraz na te charakterystyczne dla poszczególnych rodzajów czujników związane z ich działaniem, na przykład: szum danych, czy wpływ temperatury urządzenia na uzyskiwane pomiary.

⁵² S. Dixon-Warren, *Motion sensing in the iPhone 4: MEMS gyroscope*, <https://goo.gl/KKysQr>.



(a) Struktura wewnętrzna⁵²



(b) Wizualizacja budowy⁵¹

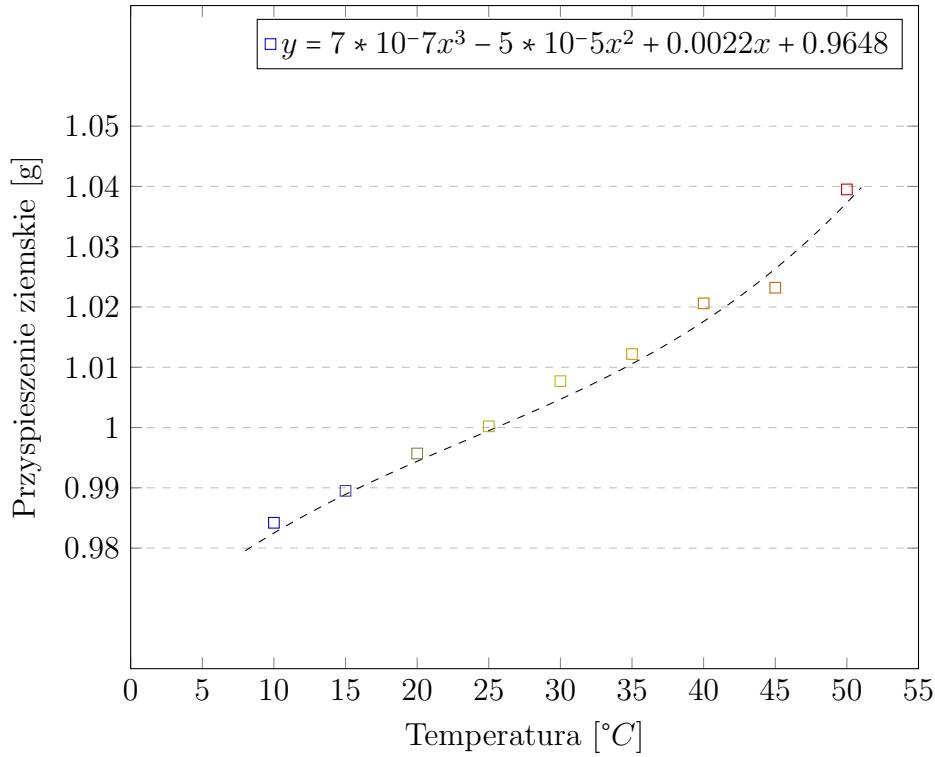
Rysunek 2.28: Struktura wewnętrzna żyroskopu w technologii MEMS

Pomiary dostarczane przez oba czujniki zawierają szum, który w znaczący sposób wpływa na dalsze obliczenia. Do określenia charakteru zakłóceń pomiaru najczęściej wykorzystuje się wariancje oraz odchylenie Allana [4], która to metoda została przyjęta przez IEEE jako standardowa do określania szumu w IMU [17]. Z uzyskanych za pomocą tej metody wyników można odczytać, że dwoma typami szumów, które przeważają w sygnale uzyskanym z żyroskopu i akcelerometru są: błędzenie losowe związane z mierzonymi wielkościami (*ang. angular/velocity random walk*), a także szum o wysokiej częstotliwości oraz niestabilność błędu systematycznego (*ang. bias instability*) o niskiej częstotliwości. Fakt występowania obu szumów w czujnikach inercyjnych jest dodatkowo powiązany z różnymi charakterystykami częstotliwościowymi właściwych sygnałów. Dla akcelerometru, mierzącego działające na niego siły w odniesieniu do siły grawitacji, sygnał prawidłowy jest o niskiej częstotliwości, więc szczególnego znaczenia nabiera zakłócenie i zmienność wartości

błędu systematycznego. Z kolei sygnał zarejestrowany przez żyroskop jest wysokoczęstotliwościowy, więc błędzenie losowe wpływa na uzyskane wyniki.

Jednym z czynników wpływających na jakość pomiarów uzyskiwanych za pomocą czujników tworzonych w architekturze MEMS jest temperatura pracy tych urządzeń. Jej wpływ na uzyskiwane wyniki, jak w swojej analizie wskazują Liu i in. [72, 73], mają trzy czynniki: moduł Younga, deformacja materiału pod wpływem temperatury oraz naprężenia materiału. Każdy z tych czynników wpływa na możliwości przepływu prądu w omawianych układach, co bezpośrednio wpływa na jakość interpretacji zgromadzonych ładunków w kondensatorach układów MEMS. Zgodnie ze specyfikacją układu MPU-6050, temperatura neutralna dla ich pracy to 25°C , a błąd związany z temperaturą urządzenia zmienia się nieliniowo. Zmiana temperatury układów MEMS wynika z temperatury otoczenia, w której pracują oraz z naturalnego wydzielania się ciepła podczas pracy. W wyniku własnych eksperymentów udało się zaobserwować wpływ zmiany temperatury na uzyskane pomiary przyspieszenia ziemskiego w zakresie temperatur $10\text{--}50^{\circ}\text{C}$ (rys. 2.29). Ponieważ temperatura pracy urządzeń umieszczonych przy ciele stabilizuje się w okolicach 32°C , czynnik ten musi być wzięty pod uwagę i skorygowany przed obliczeniami bazującymi na danych akcelerometru. Podobnej zależności nie zaobserwowano w przypadku używanych żyroskopów. Prawdopodobnie temperatura wpływa na sygnały o niskiej częstotliwości, które są odfiltrowane w trakcie odczytu pomiarów. Wynika to z faktu, że prawidłowy sygnał, jaki rejestrowany jest przez żyroskop, powinien mieć wysoką częstotliwość, a w związku z tym cały sygnał przetwarzany jest filtrem górnoprzepustowym. W takim przypadku szum niskiej częstotliwości jest w dużym stopniu usunięty z uzyskiwanych pomiarów.

Do precyzyjnego określenia orientacji, w przestrzeni urządzenia pomiarowego złożonego z czujników inercyjnych i magnetycznych, nie powinno się bazować jedynie na pojedynczym czujniku (na przykład wyłącznie na żyroskopie), ale należy połączyć ze sobą sygnały (*ang. data fusion*) z kilku różnych czujników na przykład: akcelerometru i żyroskopu, czy żyroskopu i magnetometru. Wynika to z szumów jakie występują w mierzonych sygnałach i konieczności ich odfiltrowania. Ponieważ charakterystyki szumów i sygnałów mierzonych przez każdy z czujników różnią się od siebie, to właśnie zsynchronizowane łączenie danych daje rezultaty lepsze niż na przykład zastosowanie filtracji na pojedynczym sygnale. Jakość uzyskanych wyników takiego łączenia danych jest zależna od wybranej metody (dodatek A) oraz posiadanej wiedzy dotyczącej charakterystyk urządzeń. W przypadku zastosowania pary akcelerometr – żyroskop, pomimo tego, że każdy z nich jest urządzeniem trzyosiowym, możliwe jest wyznaczenie orientacji jedynie względem dwóch osi.



Rysunek 2.29: Pomiar przyspieszenia ziemskiego mierzonego przez moduł inercyjny w przedziale temperatur 10-50°C (źródło: badania własne)

Ograniczenie to wynika z pomiarów, jakie można uzyskać z akcelerometru. Aby móc połączyć ze sobą dane z obu czujników, zarówno akcelerometr jak i żyroskop muszą udostępnić dane dotyczące obrotów wokół tych samych osi. O ile żyroskop faktycznie dokonuje pomiaru prędkości kątowych wokół wszystkich trzech osi, o tyle akcelerometr podaje informację o tym jak czujnik jest obrócony tylko względem dwóch osi. Dla akcelerometru, obrót wokół osi działania siły grawitacji (obrót względem osi Z na rys. 2.26) jest niesignalizowany, ponieważ obrót taki nie zmienia wartości z jaką ta siła działa na czujnik w danej chwili. Z tego zaś powodu, orientacja względem osi działania siły grawitacji nie jest wyznaczona drogą łączenia danych z akcelerometrem i żyroskopem, a jedynie może być wyznaczona na podstawie całkowania prędkości kątowej zmierzanej przez sam żyroskop, co jest obarczone znaczącymi błędami z uwagi na zaszumienie danych. W związku z tym nie bierze się tej wartości pod uwagę w obliczeniach i przyjmuje się, że urządzenia oparte na parze czujników inercyjnych określają orientację względem dwóch osi.

Opisane w niniejszym rozdziale urządzenia pomiarowe (kontroler Microsoft Kinect oraz akcelerometr i żyroskop) są wykorzystywane do budowy dwóch różnych systemów śledzenia ruchu. Wykorzystując kontroler Microsoft Kinect można zbudować optyczny system śledzenia ruchu niewykorzystujący

markerów (roz. 2.1.1). Z kolei akcelerometry i żyroskopy stanowią podstawę dla nieoptycznych inercyjnych systemów śledzenia ruchu (roz. 2.1.2). Badacze zwróciли uwagę na możliwość tworzenia hybrydowych systemów śledzenia ruchu, które opierają się na kontrolerze Kinect, akcelerometrach oraz żyroskopach. Tworzenie takich hybrydowych systemów śledzenia ruchu możliwe jest dzięki temu, że urządzenia pomiarowe, które wykorzystują, mogą się wzajemnie uzupełniać w działaniu.

2.4. Systemy hybrydowe łączące optyczny system śledzenia ruchu bez markerów oraz inercyjny system śledzenia ruchu

Poszukiwanie rozwiązań pozwalających na łączenie ze sobą optycznych systemów śledzenia ruchu bez markerów oraz inercyjnych systemów śledzenia jest zagadnieniem stosunkowo często spotykany w literaturze. Wiele ośrodków badawczych na całym świecie prowadzi badania w tym obszarze mające na celu zwiększenie dokładności pozycjonowania stawów i śledzenia ich ruchu. Obecnie, dzięki powszechniej dostępności urządzeń pomiarowych takich jak kontroler Kinect oraz akcelerometry i żyroskopy, jak również dzięki badaniom związanym z ich połączeniem w jeden system, można budować hybrydowe systemy śledzenia ruchu o wysokiej dokładności, dostępne do użytku domowego. Otwiera to z kolei wiele możliwości w obszarze zastosowań takich systemów, jak chociażby zaawansowane systemy do telerehabilitacji, czy programy do precyzyjnych treningów sportowych. W dalszej części pracy przedstawione zostaną te opracowania, które wykorzystują kamery RGB-D dostępne na rynku konsumenckim (w szczególności kontroler Kinect) oraz urządzenia inercyjne lub magnetyczne.

W 2014 roku Destelle i in. [27] zaproponował połączenie ze sobą kontrolera Kinect oraz systemu inercyjnego, w celu stworzenia kompletnego systemu śledząco–pozycjonującego, nie korzystającego z sygnału GPS oraz zaawansowanych systemów LPM (lokalny system pozycjonowania – *ang. Local Position Measurement System*) w celu określenia położenia użytkownika na scenie. W rozwiązaniu autorów Kinect został wykorzystany na dwa sposoby:

1. Jako system definiujący model szkieletowy.
2. Jako system odpowiedzialny za śledzenie położenia aktora na scenie.

W omawianym rozwiązaniu autorzy zdecydowali się na wykorzystanie urządzeń inercyjnych stworzonych przez firmę X-IO Technologies⁵³. Urządzenia inercyjne firmy X-IO Technologies to układy zbudowane na bazie czujników inercyjnych albo czujników inercyjnych wspieranych przez czujnik magnetyczny. Łączenie danych z wymienionych czujników, mające na celu określenie orientacji przestrzennej urządzenia, stosuje autorską metodę założyciela tej firmy, nazwaną jego nazwiskiem – metodę Madgwicka.

Aby uniknąć konieczności budowania modelu szkieletowego na podstawie własnych pomiarów, na potrzeby systemu inercyjnego, autorzy postanowili wykorzystać estymację szkieletu dokonywaną przez Kinecta. Aby uniknąć problemów związanych ze zmiennością szacowania długości poszczególnych kości, a co za tym idzie zmiennością proporcji modelu szkieletowego, niezbędne pomiary pozwalające wyznaczyć model szkieletowy dokonywane są tylko raz w czasie inicjalizacji systemu i zostają przyjęte jako stałe na cały czas trwania śledzenia ruchu. W trakcie śledzenia ruchu rolą urządzeń inercyjnych było określenie orientacji przestrzennej tych kości, do których takie moduły były przymocowane, a rolą Kinecta było określenie położenia całego modelu postaci na scenie. Pozycjonowanie modelu na scenie zostało zrealizowane poprzez śledzenie przez kontroler Kinect położenia pojedynczego punktu znajdującego się na korpusie obserwowanej postaci. Przemieszczenie tego pojedynczego punktu traktowane jest jako przemieszczenie całego wyznaczonego modelu szkieletowego. W omawianym artykule, Destelle wskazuje średniokwadratowy błąd wyznaczania kątów zgięcia kończyn w stawach na $4^\circ - 14^\circ$ w zależności od stawu, wobec błędu średniokwadratowego w przedziale $10^\circ - 30^\circ$ dla tych samych stawów w przypadku pomiaru kątów wyłącznie za pomocą samego kontrolera Kinect. Warto jednak zauważyć, że w przypadku poniższych obarczonych największym błędem średniokwadratowym następuowała czasowa utrata śledzenia poszczególnych stawów. Autor nie zamieścił natomiast żadnych danych pozwalających na oszacowanie dokładności wyznaczenia pozycji poszczególnych stawów. Łatwo można zauważyć, że system zaproponowany przez Destelle nie wykorzystuje w pełni możliwości łączenia danych z obu urządzeń, gdyż wykorzystywane są one w odrębnych obszarach i mają minimalny wpływ na siebie. Stąd, porównanie dokładności samego Kinecta z dokładnością systemu hybrydowego w zakresie mierzenia zgięcia kończyn w stawach jest de facto porównaniem dokładności tego pierwszego z urządzeniem Madgwicka. Z drugiej zaś strony, omawiany system dobrze pokazuje zasadność wykorzystania w badaniach łatwo dostępnych i tanich urządzeń inercyjnych, które pozwalają poprawić wyniki kontrolera Kinect.

⁵³ X-io Technologies, *x-IMU our original versatile imu board*, <https://goo.gl/eAkVba>.

Bo i in. [15] w 2011 roku zaproponowali metodę szacowania kąta stawu kolanowego wykorzystując układ czujników inercyjnych o pięciu stopniach swobody (dwa stopnie swobody dla żyroskopu i trzy stopnie swobody dla akcelerometru) oraz Kinecta. Metoda ma niejako dwa tryby działania: dla sytuacji kiedy dane z Kinecta są dostępne i wtedy kiedy ich brakuje. W obu przypadkach to czujniki inercyjne są wykorzystywane do oszacowania kąta zgięcia kończyny w obserwowanym stawie, przy czym wartości te obliczane są osobno dla żyroskopu (Θ_G) i akcelerometru (Θ_A). Oszacowanie wartości kąta, wykorzystuje całkowanie pomiarów żyroskopu ($\tilde{\omega}_y$) zgodnie ze wzorem 2.15.

$$\Theta_G = \int \tilde{\omega}_y dt \quad (2.15)$$

oraz bazuje na obliczeniach trygonometrycznych uwzględniających stosunek siły działającej na akcelerometr w wybranych osiach (f_x, f_z) do siły grawitacji g według wzoru 2.16.

$$\Theta_A = \alpha \arccos \frac{f_x}{\|g\|} + (1 - \alpha) \arcsin \frac{f_z}{\|g\|} \quad (2.16)$$

Następnie wartości te są ze sobą łączone z wykorzystaniem liniowego filtra Kalmana. W przypadku dostępności danych z Kinecta, dodatkowym krokiem jest korekta pomiarów na podstawie danych z tego urządzenia. Obliczana jest wówczas różnica pomiędzy wartościami wyznaczonymi z czujników inercyjnych i z Kinecta, a wartość różnicy uwzględniana jest jako korekta estymacji kąta dokonanej jedynie na podstawie pomiarów akcelerometru.

Metoda zaproponowana przez Bo i in. opiera się na założeniu, że Kinect jest wystarczająco dokładnym urządzeniem, aby jego pomiary przyjąć jako referencyjne i na ich podstawie dokonywać korekty innych pomiarów. Dokładna analiza zakresów działania Kinecta (dokładność oszacowania odległości osoby śledzonej od kontrolera) przedstawiona w rozdziale 2.3 pokazuje, że takie założenie jest prawdziwe dla ruchów, które odbywają się bez istotnej zmiany odległości pomiędzy kontrolerem Kinect a postacią. Problemem może być jednak śledzenie ruchu rąk przesuwanych do przodu, ponieważ nawet stojąc w optymalnej odległości od kontrolera (około 2m), ręka w trakcie ruchu najprawdopodobniej nie będzie znajdowała się obszarze, w którym Kinect szacuje położenie stawów z dokładnością zbliżoną do pomiarów rzeczywistych. Dodatkowo, przyglądając się badaniom widać, że metoda była testowana w taki sposób, że staw kolanowy, którego kąt zgięcia podlegał oszacowaniu, był dobrze widoczny i znajdował się w stałej odległości od Kinecta. Osoba, której ruch był śledzony, była obrócona bokiem do kontrolera

i na przemian wstawała i siadała na krzesle. Dzięki temu uniknięto problemów związanych z okluzją oraz zmianą dokładności śledzenia wynikającej ze zmian odległości postaci od kontrolera Kinect.

Wyniki zamieszczone w omawianym artykule autorsta Bo i in.[15] nie pozwalają na ocenę dokładności proponowanej metody względem wartości rzeczywistych. Pokazują one natomiast relację oszacowań będących wynikiem proponowanej metody do pomiarów Kinecta, które zostały przyjęte przez autorów jako referencyjne. Pokazują również, że dzięki wykorzystaniu łączenia danych z żyroskopu i akcelerometru udało się uzyskać stabilizację pomiarów kąta za pomocą czujników inercyjnych wobec braku stabilności pomiarów każdego z czujników z osobna. Zamieszczone w artykule wykresy wyraźnie pokazują negatywny wpływ dryfu na wyniki uzyskane za pomocą żyroskopu, objawiające się stopniowym pogarszaniem jakości pomiarów. Dryf ten jest wyraźny i na podstawie opublikowanych przez autorów wykresów, można go ocenić na $2^{\circ}/s$. Tendencja ta nie jest widoczna w przypadku pomiarów akcelerometru, jednak zaobserwować można ciągłe, małe odchylenia od pomiaru referencyjnego. Widać także, że wykorzystanie filtra Kalmana w liniowej postaci jest w stanie oba te problemy znacznie ograniczyć.

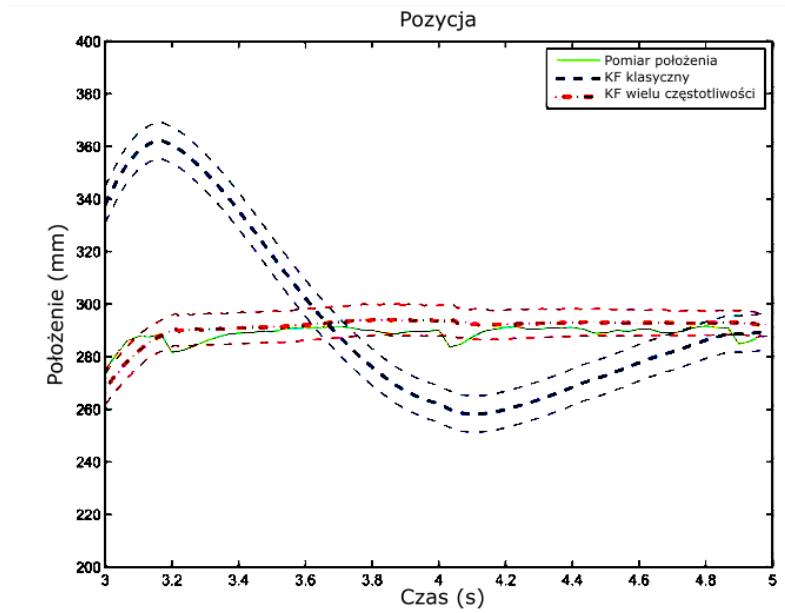
W 2015 roku Tian i in. [116] zaproponował łączenie danych z czujników inercyjnych za pomocą bezśadowego filtra Kalmana (UKF - ang. *Unscendent Kalman Filter*). W prezentowanej metodzie autorzy zdecydowali się na korektę uzyskanych wyników za pomocą z góry narzuconych ograniczeń geometrycznych, wynikających z modelu biomechanicznego szkieletu ludzkiego. Początkowo autorzy w swojej metodzie wyznaczają orientację każdej z kości, których ruch jest śledzony za pomocą modułów inercyjnych. Orientacje te są wyznaczone bez uwzględniania ich poprawności, to znaczy bez sprawdzenia czy w ogóle możliwe jest, aby człowiek wykonał taki ruch na jaki wskazują pomiary. Dopiero w następnym kroku następuje weryfikacja uzyskanych oszacowań ze względu na ograniczenia wynikające z mechaniki ruchu ciała ludzkiego. Pozwoliło to wyeliminować oszacowania obrotów, które w sposób oczywisty są niepoprawne oraz ograniczyć wpływ dryfu żyroskopu na ostateczny wynik. Przykładem takiego ograniczenia, dzięki któremu można zweryfikować poprawność oszacowania orientacji kości, jest maksymalny kąt, jaki można uzyskać w łokciu przy wyprostowanej ręce, wynoszący około 180° (w zależności od występowania anomalii osobniczych kąt ten może być nieco większy w przypadku przeprostów lub nieco mniejszy w przypadku występowania na przykład przykurczy mięśniowych). Jeśli zatem oszacowanie, na podstawie danych z czujników, wskazywałoby na przykład kąt zbliżony do 220° to można uznać, że jest to oszacowanie błędne. Po zakończeniu fuzji

danych z czujników inercyjnych przeprowadzano weryfikację dostępności pomiarów z kontrolera Kinect. Jeśli były one dostępne i były dostatecznie dobrzej jakości, zostały one połączone z wielkościami uzyskanymi z IMU za pomocą filtra UKF. Autorzy nie podali wprost kryterium jakości pomiarów Kinecta, ale z treści artykułu można wywnioskować, że takim kryterium jest wariancja położenia konkretnych stawów. Jeśli dane są niewiarygodne, wartość ta jest bardzo duża w każdym kolejnym przedziale czasowym. Jeśli dane z Kinecta nie mogą być użyte do ostatecznego wyznaczenia pozycji stawów, algorytm bazuje tylko na pomiarach z czujników inercyjnych i odpowiednio aktualizuje położenie stawów na bazie wyznaczonego oszacowania. Porównanie uzyskanych wyników z wynikami referencyjnymi uzyskanymi z systemu wizyjnego z markerami, pozwoliło autorom oszacować dokładność ich metody śledzenia poniżej 20° dla kąta zgięcia łokcia.

Cechą wspólną powyższych metod jest to, że podstawą ich działania były pomiary z czujników inercyjnych, a Kinect stanowił uzupełnienie lub korektę uzyskanych pomiarów. W przypadku Bo i in. oraz Destelle i in. dane z Kinecta nie były w żaden sposób poddane weryfikacji co do ich wiarygodności. Podejście takie było obarczone dużym ryzykiem wprowadzenia dodatkowych błędów do uzyskanego wyniku, które mogło całkowicie wypaczyć wyniki. Metoda zaprezentowana przez Tian i in. wprowadziła krok weryfikujący poprawność danych Kinecta i wraz z ograniczeniami wynikającymi z biomechaniki minimalizowała wpływ zaszumienia danych na ostateczny rezultat. Nie wiadomo jednak jak długo dyskutowana metoda jest w stanie, w poprawny sposób, działać bez pomiarów Kinecta. Taki eksperyment nie został opisany w artykule.

Kolejną grupę hybrydowych metod łączących sygnały inercyjne z sygnałami kontrolera Kinect stanowią metody, które wykorzystują w sposób ciągły oba te źródła. Wymagają one ciągłego analizowania obu sygnałów i na podstawie ich jakości decydują z jakim stopniem istotności je połączyć. W 2014 roku Feng i Murray-Smith [104] zaprezentowali metodę łączenia danych pochodzących z kontrolera Kinect i czujników inercyjnych za pomocą zmodyfikowanego filtra Kalmana potrafiącego działać na sygnałach o różnej częstotliwości (*ang. multi-rate Kalman Filter* [28]). Wybór takiego filtra został podeyktywany tym, że częstotliwość pracy Kinecta to 30Hz, a użytego układu MARG to 90Hz i połączenie ich sygnałów klasyczną wersją filtru wymagałaby wyrównania tych częstotliwości. Feng i Murray-Smith wykorzystali w swojej pracy wariant filtru liniowego, natomiast Armesto i Smyth dla podobnego zadania wykorzystali warianty odpowiednio EKF[7] i UKF[109]. Zastosowanie modyfikacji metody łączenia danych, uwzględniającej różne częstotliwości łączonych sygnałów, pozwoliło uzyskać lepszą reaktywność na zmiany

zachodzące w śledzonym ruchu. Różnicę w czasie reakcji na zakończenie ruchu widać na rysunku 2.30. Moment zakończenia ruchu uwzględniony został w przedziale czasowym pomiędzy $3s$ a $3.4s$. Niebieski wykres na rysunku 2.30 pokazuje oszacowanie położenia wykorzystujące klasyczny filtr Kalmana, natomiast czerwony wykres przedstawia oszacowanie położenia z wykorzystaniem zmodyfikowanego filtru.



Rysunek 2.30: Szacowanie położenia stawu z użyciem klasycznego filtru Kalmana (niebieski) i filtru Kalmana dostosowanego do różnych częstotliwości (czerwony) [104]

Na podstawie wykresów zamieszczonych w artykule Fanga i Murray-Smitha można oszacować, że zaproponowana przez nich metoda określa pozycję stawów z dokładnością $1.5cm - 2cm$. Opublikowane diagramy pokazują jedynie wykresy wygenerowane przez zaledwie $5s$, więc problematyczne staje się oszacowanie jak zachowa się ta metoda w dłuższym okresie. Widać także, że metoda wykorzystująca modyfikację filtru Kalmana i uwzględniająca zróżnicowanie częstotliwości łączonych sygnałów, szybciej reaguje na zmiany w wykonywanym ruchu, niż metoda oparta o klasyczną implementację filtru Kalmana (ma mniejszą bezwładność).

W 2013 roku Helten i in. [44] zaproponował połączenie ze sobą sygnałów z IMU oraz mapy głębi wyznaczonej przez kontroler Kinect. Dane uzyskane z obu źródeł pozwalają na zbudowanie mapy widoczności poszczególnych części ciała. Informacje zawarte w tej mapie pozwalają z kolei wygenerować

model ciała odzwierciedlający pozę, w jakiej znajduje się śledzona postać. Wygenerowanie modelu ciała śledzonej postaci pozwala na przeprowadzenie klasyfikacji tego modelu na podstawie wcześniej zdefiniowanego treningowego zbioru zawierającego modele ciała w określonych, nazwanych pozach. Istotną cechą metody polegającej na klasyfikacji poz, w jakiej znajduje się śledzona postać, był brak możliwości określenia położenia poszczególnych stawów, a jedynie monitorowanie aktualnej aktywności śledzonej postaci. Dzięki temu można było określić czy dana osoba stoi, chodzi, czy siedzi, a to z kolei zapewniło, że metoda zaproponowana przez zespół Thomasa Heltena mogła być z powodzeniem zastosowana w systemach nadzorujących na przykład zachowanie osób starszych. W związku z tym metoda ta nie może być porównywana z takimi metodami jak choćby opisana powyżej metoda zaproponowana przez Bo i in. [15], natomiast autorzy porównali ją z analogicznymi metodami rozpoznającymi pozę człowieka na przykład Ganapathi i in. [36] czy Baak i in. [9]. Według opisu zamieszczonego w omawianym artykule, autorzy wskazali, że uzyskane przez nich rezultaty prawidłowego rozpoznania póż oraz wyznaczenia na ich podstawie modelu szkieletowego były nie gorsze niż innych, podobnych metod opisanych w literaturze dla wybranego zestawu sześciu póż, wśród których były między innymi obroty czy kopnięcia. Zdefiniowano błąd wyznaczenia wybranych szesnastu stawów uproszczonego modelu szkieletowego, jaki udało się wyznaczyć na podstawie rozpoznanej poz, względem modelu szkieletowego otrzymanego w wyniku śledzenia póż optycznym systemem śledzenia ruchu z markerami firmy PhaseSpace. Średni błąd wyznaczania stawów przez metodę proponowaną przez zespół Thomasa Heltena wyniósł około $75mm$, co stanowiło poprawę o blisko 50% względem metody o największym średnim błędzie pośród metod porównywanych w omawianym artykule.

Ostatnią z przytoczonych metod, ujętych w literaturze i dyskutowanych w niniejszej pracy jest metoda zaproponowana w artykule Kalkbrenner i in. [57]. Autorzy w swojej metodzie wykorzystali dwa filtry łączące dane uzyskane za pomocą kontrolera Kinect oraz urządzeń pomiarowych opartych o czujniki inercyjne. Pierwszym z nich jest filtr zaproponowany przez Sebastiana Madgwicka [57], drugim – liniowy filtr Kalmana. Filtr Madgwicka (odatek A) pozwala na wyznaczanie orientacji urządzeń pomiarowych w dwóch osiach dla czujników inercyjnych: akcelerometru i żyroskopu oraz w trzech dla czujników inercyjnych wspartych przez magnetometr. Implementacja filtra przygotowana przez jego twórcę operowała na kwaterionach, co jest nie-wątpliwym ułatwieniem dla dalszych przekształceń. Jest to o tyle istotne, że metoda zaproponowana przez Kalkbrennera wymaga, na podstawie wyniku

uzyskanego za pomocą filtra Madgwicka oraz modelu długości kości, wyznaczenia pozycji kolejnych stawów w analogiczny sposób, jaki ma miejsce przy budowaniu hierarchicznego modelu szkieletowego ciała człowieka. Po wyznaczeniu pozycji stawów (w przypadku dyskutowanego artykułu wyznaczane są jedynie stawy jednej ręki: barkowy, jako korzeń, łokciowy oraz nadgarstkowy) były one łączone z analogcznymi pozycjami stawów otrzymanymi z Kinecta za pomocą liniowego filtru Kalmana.

Warto tutaj dodać, że autorzy uzależnili wartość współczynnika Kalmana K (*ang. Kalman gain*) od wartości przemieszczenia stawów δs w modelu szkieletowym Kinecta. Jeśli przemieszczenie to pomiędzy kolejnymi pomiarami ($\delta s = [\delta s_x, \delta s_y, \delta s_z]$) było zbyt duże, oznaczało to utratę śledzenia i brak wiarygodności pomiarów, a co za tym idzie obniżenie istotności tych danych w trakcie łączenia. Jako graniczną wartość zbyt dużego przemieszczenia się stawu, po której następował ponowne wyznaczenie współczynnika Kalmana, autorzy przyjęli wartość $15cm$. W sytuacji kiedy przemieszczenie się stawu pomiędzy kolejnymi pomiarami przekroczyło tę wartość graniczną następuwało ponowne wyznaczenie najpierw macierzy kowariancji kontrolera Kinect R według wzoru 2.17.

$$R = \begin{pmatrix} R_x & 0 & 0 \\ 0 & R_y & 0 \\ 0 & 0 & R_z \end{pmatrix} + \begin{pmatrix} \delta s_x^2 & 0 & 0 \\ 0 & \delta s_y^2 & 0 \\ 0 & 0 & \delta s_z^2 \end{pmatrix} * \kappa \quad (2.17)$$

a następnie współczynnika Kalmana K wykorzystując wzór 2.18:

$$K = P * H^T * (H * P * H^T + R)^{-1} \quad (2.18)$$

gdzie: H jest macierzą definiującą powiązanie pomiędzy dwoma kolejnymi pomiarami kontrolera Kinect (autorzy zdefiniowali ją jako macierz jednostkową), R to macierz kowariancji kontrolera Kinect wyznaczona według wzoru 2.17, a P to macierz kowariancji czujników inercyjnych. Wartość współczynnika κ we wzorze 2.17 została przez autorów artykułu wyznaczona empirycznie na 15.

Twórcy omawianej metody podali wartość średniego odchylenia wyznaczenia położenia wybranych stawów na $\pm 2.2cm$ wobec $\pm 5.5cm$ dla kontrolera Kinect pracującego samodzielnie. Z opisu zaproponowanej metody wynika, że dla wyznaczenia pozycji stawów na podstawie IMU, model długości poszczególnych kości pobrany był wprost z danych Kinecta. Z jednej strony ułatwiło to późniejsze złączenie ze sobą pozycji stawów wyznaczonych przez Kinecta oraz wyznaczonych na podstawie pomiarów czujników inercyjnych, ponieważ jedyna różnica pomiędzy rzeczywistymi stawami wynika z orientacji kości łączących śledzone stawy. Z drugiej zaś strony, oszacowanie długości

kości przez kontroler Kinect jest niedokładne i zmienia się niemal w każdej klatce pomiarowej, więc uniemożliwiało to wyznaczenie prawdziwego modelu szkieletowego, w którym długości poszczególnych kości odpowiadałyby prawdziwym długościom.

Na tej podstawie zasadnym wydaje się wprowadzenie modelu szkieletowego, w którym długości poszczególnych kości miałyby stałe wartości. Wykazałoby to jednak wcześniejszego dokonania pomiarów poszczególnych kości w celu zbudowania takiego modelu. Eliminując zaszumienie ostatecznych wyników przez zmienność długości kości, łączenie danych z urządzenia Kinect i czujników inercyjnych powinno odbywać się z wykorzystaniem informacji o tym, w jakiej orientacji znajdują się poszczególne kości. Wartość taka dostępna jest jako jedna z danych udostępnianych przez oprogramowanie kontrolera Kinect, a także jest dostępna w wyniku działania urządzeń pomiarowych opartych o czujniki inercyjne. Następnie wykorzystanie informacji o obrotach, z przygotowanym wcześniej modelem szkieletowym, pozwala wyznaczyć położenie poszczególnych stawów. Można przypuszczać, że wprowadzenie modelu szkieletowego o stałych długościach kości i wykorzystanie informacji o ich obrocie da większą dokładność wyznaczania pozycji stawów niż bezpośrednie łączenie ze sobą pozycji stawów wyznaczonych w dwóch modelach szkieletowych, dla każdego z urządzeń pomiarowych mających zmienne długości kości. Powyższe spostrzeżenia stanowią podstawową przesłankę do budowy nowej, autorskiej hybrydowej metody śledzenia ruchu kończyn. Wstępna propozycja takiej metody została przedstawiona przez autora niniejszej dysertacji podczas konferencji *International Conference Information Technologies in Biomedicine (ITIB)* 2016 [39], a następnie jej wersja rozszerzona o synchronizację czasową sygnałów została zaprezentowana w trakcie konferencji *International Conference on Computer Vision and Graphics (ICCVG)* 2016 [40].

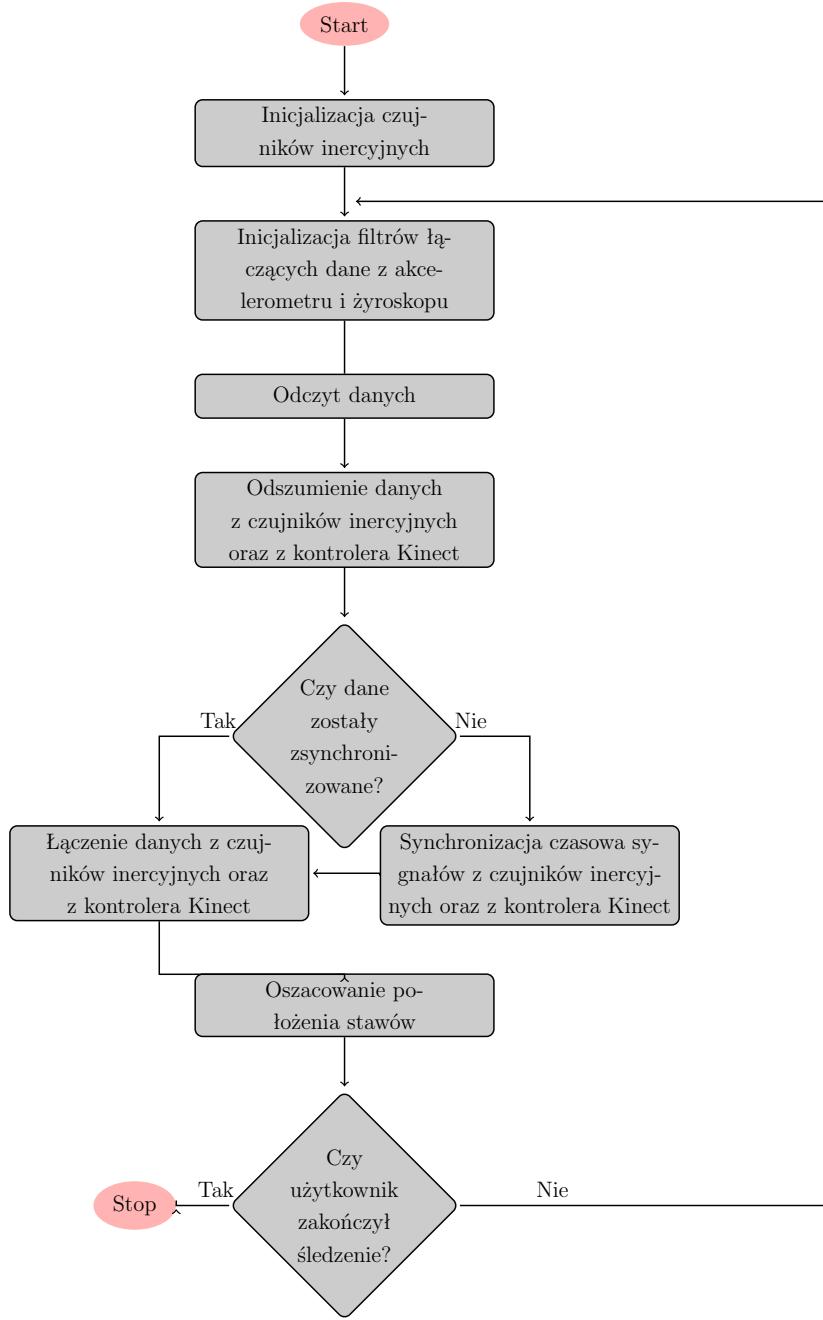
Rozdział 3

Hybrydowa metoda śledzenia ruchu człowieka

Prezentowana w niniejszej dysertacji hybrydowa metoda śledzenia ruchu człowieka, wykorzystuje dwa rodzaje urządzeń pomiarowych: kontroler Microsoft Kinect oraz czujniki inercyjne - akcelerometr i żyroskop, których charakterystyki zostały przedstawione w rozdziale 2.3. Mając na uwadze dokładność danych przesyłanych przez każde z zastosowanych urządzeń pomiarowych oraz ograniczenia w ich działaniu, takie jak wrażliwość na występowanie okluzji sensora głębi, czy ograniczona liczba stopni swobody układów inercyjnych, opracowana przez autora hybrydowa metoda śledzenia ruchu człowieka umożliwia ograniczenie wpływu niedoskonałości każdego ze składowych urządzeń. Autorskie połączenie danych (*ang. data fusion*) z kontrolera Kinect i czujników inercyjnych poprawia precyzję szacowania położenia stawów, a co za tym idzie wpływa korzystnie na precyzję śledzenia wykonywanego przez użytkownika ruchu. Ograniczenie wpływu niedoskonałości każdego z wykorzystanych urządzeń pomiarowych na dokładność śledzenia ruchu możliwe jest dzięki temu, że są one względem siebie w pewnym stopniu komplementarne to znaczy, że niedoskonałości występujące w jednym urządzeniu i wpływające, na przykład, na przerwy w śledzeniu ruchu w jednej z płaszczyzn, mogą zostać ograniczone przez wykorzystanie możliwości śledzenia ruchu w tej płaszczyźnie przez drugie z urządzeń.

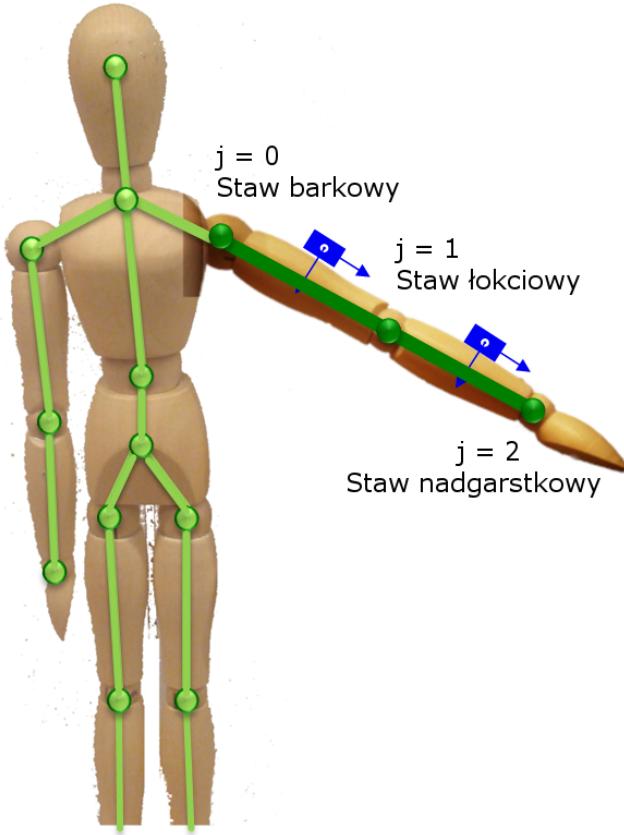
Proces łączenia danych uzyskanych z kontrolera Kinect i z czujników inercyjnych podzielony został na kilka etapów, z których część jest wykonana jedynie na samym początku, a część jest powtarzana w trakcie śledzenia ruchu. Schemat kolejnych kroków przetwarzania danych prezentuje diagram na rysunku 3.1. Kroki te zostały omówione w dalszej części niniejszego rozdziału.

Zadaniem, omawianej w niniejszym rozdziale, hybrydowej metody śledzenia ruchu człowieka jest oszacowanie w chwili czasu t położenia wybranego stawu j szkieletu postaci ($P_{j,t}^F = [p_{j,x}^F, p_{j,y}^F, p_{j,z}^F]_t$, F – *ang. Fused*) w przestrzeni trójwymiarowej opisanej w referencyjnym układzie współrzędnych



Rysunek 3.1: Algorytm przetwarzania danych w autorskiej metodzie łączenia danych (źródło: badania własne)

XYZ , tożsamym z układem współrzędnych kontrolera Kinect. W zakresie pojedynczego modułu inercyjnego szacowanie położenia stawu odbywa się na podstawie pomiaru sił działających na akcelerometr w każdej z trzech osi ($A = [a_x, a_y, a_z]$), pomiaru prędkości kątowych, z jakimi porusza się żyroskop względem każdej z osi ($G = [g_x, g_y, g_z]$) oraz aktualnej temperatury pracy czujników inercyjnych (T). Pomiary uzyskane z akcelerometru oraz żyroskopu wyrażone są w układzie współrzędnych modułu inercyjnego (rys. 2.26).



Rysunek 3.2: Przykład hierarchii stawów dla modelu szkieletowego ręki (źródło: opracowanie własne)

W przypadku kontrolera Kinect pobierane są pozycje dwóch stawów: stawu j , którego pozycja będzie szacowana ($P_j^K = [p_{j,x}^K, p_{j,y}^K, p_{j,z}^K]$, K -Kinect) oraz stawu $j - 1$ ($P_{j-1}^K = [p_{j-1,x}^K, p_{j-1,y}^K, p_{j-1,z}^K]$), będącego jego rodzicem w modelu szkieletowym Kinecta (rys. 3.2). Fuzja danych z czujników inercyjnych (A oraz G) oraz z kontrolera Kinecta (P_j^K) pozwala oszacować pozycję stawu j w danej chwili t ($P_{j,t}^F$) i odbywa się z pewnym interwałem czasowym Δt , określającym czas pomiędzy kolejnymi aktualizacjami pomiarów kontrolera Kinect, a będącym jednocześnie interwałem czasowym dla procesu decygnacji sygnału urządzeń inercyjnych. Dodatkowymi parametrami wykorzystywanymi w procesie łączenia danych z modułu inercyjnego z pomiarami Kinecta są pozycje obu stawów barkowych uzyskane za pomocą kontrolera Kinecta ($P_{sh_L,t}^K, P_{sh_R,t}^K$). Pozwalają one oszacować:

1. jak śledzona postać jest obrócona względem Kinecta,
2. pozycję stawu rodzica ($P_{j-1,t}^F$) w globalnym układzie współrzędnych, będącą wynikiem osobnej fuzji danych,
3. długość kości l , na której umieszczony jest moduł inercyjny.

Biorąc pod uwagę powyższe, oszacowanie pozycji wybranego stawu j , w hybrydowej metodzie śledzenia ruchu człowieka, można wyrazić za pomocą uogólnionego wzoru 3.1, którego szczegółowy kształt zostanie przedstawiony w dalszej części pracy.

$$P_{j,t}^F = [p_{j,x}^F, p_{j,y}^F, p_{j,z}^F]_t = f(A, G, T, P_{j-1,t}^K, P_{j,t}^K, \Delta t, P_{j-1,t}^F, P_{sh_L,t}^K, P_{sh_R,t}^K, l) \quad (3.1)$$

Wzór 3.1 odnosi się do każdego modułu inercyjnego zbudowanego z pary czujników: akcelerometru i żyroskopu, umieszczonego na ciele pomiędzy dwoma kolejnymi stawami, których pozycje oszacowane przez kontroler Kinect są argumentami metody. W przypadku śledzenia ruchu więcej niż jednego stawu, powyższą metodę należy wykonać zgodnie z ustaloną hierarchią stawów rozpoczynając od przyjętego korzenia. Oszacowana, w wyniku fuzji, pozycja stawu $P_{j,t}^F$ określa jego położenie w globalnym układzie współrzędnych prezentowanego systemu śledzenia ruchu, który pokrywa się z układem współrzędnych kontrolera Kinect.

3.1. Format danych używany przez urządzenia pomiarowe

3.1.1. Kontroler Kinect

Kontroler Kinect w ramach pakietu pomiarów, rejestrowanego z częstotliwością $30Hz$, udostępnia 3 kategorie danych dotyczących obserwowanej sceny: obraz z kamery RGB, mapę głębi sceny oraz model szkieletowy postaci widocznej na scenie. Kontroler Kinect może jednocześnie zrekonstruować modele szkieletowe dwóch postaci i śledzić ich kolejne pozy, rejestrując tym samym ich ruch. W zależności od potrzeb, aplikacja wykorzystująca kontroler Kinect może uzyskać wszystkie 3 kategorie danych jednocześnie lub wybrać tylko te z nich, które potrzebuje do swojego działania. Opisywana w niniejszej dysertacji hybrydowa metoda śledzenia ruchu kończyn człowieka bazuje na informacji dotyczącej modelu szkieletowego śledzonej postaci wybierając z pełnego modelu tylko informacje związane z wybranymi stawami reprezentującymi śledzoną kończynę oraz te, na podstawie których można określić wiarygodność pomiarów uzyskanych za pomocą kontrolera Kinect. Na przykład przy śledzeniu ruchu prawej ręki są to oba stawy barkowe, prawy staw łokciowy oraz prawy staw nadgarstkowy. Każdy staw j reprezentowany jest przez trzy współrzędne ($P_j^K = [p_{j,x}^K, p_{j,y}^K, p_{j,z}^K]$) określające jego położenie w przestrzeni w układzie współrzędnych kontrolera Kinect oraz

informacje o stanie śledzenia tego stawu: *Tracked*, *Interferred* albo *NotTracked*. Na podstawie współrzędnych stawów barkowych postaci obliczany jest kąt obrotu sylwetki względem Kinecta (wzór 2.7). Dodatkowo każdy pakiet danych zawiera informacje o swoim numerze, znacznik czasu wygenerowany przez kontroler Kinect oraz godzinę odebrania danych na komputerze, do którego podłączone jest to urządzenie.

3.1.2. Czujniki inercyjne

Moduły inercyjne (IMU1 i IMU2) zbudowane zostały z czujników inercyjnych: akcelerometru i żyroskopu, podłączonych bezpośrednio do jednostki centralnej opartej na platformie Arduino. Częstotliwość odczytu danych z czujników kontrolowana jest przez zewnętrzny moduł zegara (RTC DS3230), który również został podłączony bezpośrednio do Arduino⁵⁴. Zadaniem jednostki centralnej jest odczyt danych z czujników, sformatowanie ich i wysłanie z wykorzystaniem technologii Bluetooth (BT HC-05) do komputera PC, gdzie następuje dalsze ich przetwarzanie (uproszczony schemat urządzenia do pomiaru danych inercyjnych⁵⁵ przedstawia rys. 3.3). Odczyt pakietu danych z pojedynczego modułu obejmuje:

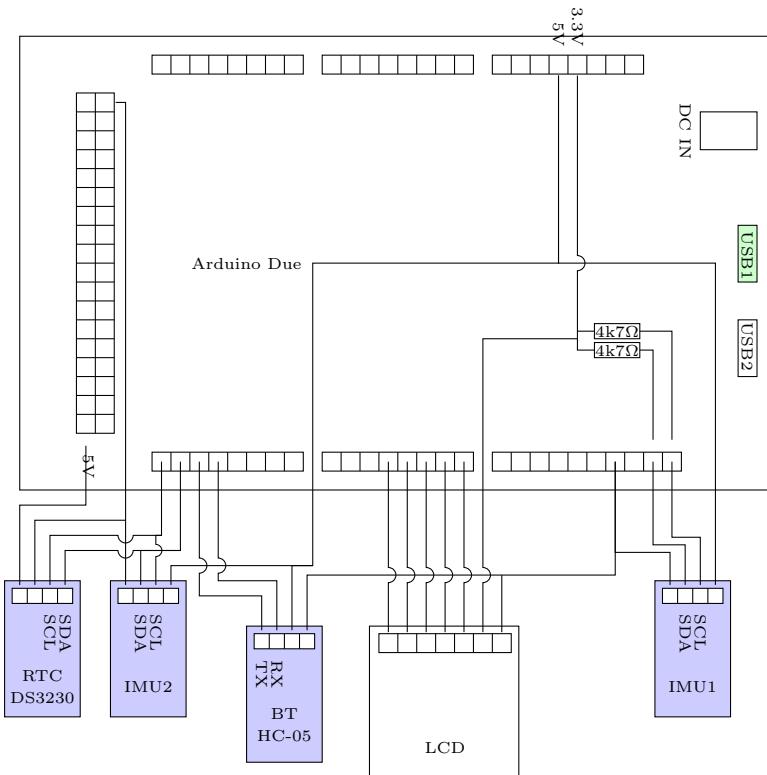
1. zestaw pomiarów udostępnionych przez żyroskop (trzy wartości obrotu $G = [g_x, g_y, g_z]$, po jednej dla każdej z osi),
2. zestaw pomiarów z akcelerometru (trzy wartości przyspieszenia $A = [a_x, a_y, a_z]$),
3. pomiar temperatury czujników T .

Wszystkie pomiary przesyłane są jako 16-bitowe liczby całkowite ze znakiem i pomiary ze wszystkich podłączonych modułów wysyłane są do komputera PC jednocześnie. Przesyłane wartości muszą być następnie skonwertowane do jednostek odpowiadających wielkościom fizycznym mierzonym przez poszczególne czujniki. Dla żyroskopu są to stopnie na sekundę ($^{\circ}/s$), dla akcelerometru – jednostki przyspieszenia grawitacyjnego (g), które można również przeliczyć na jednostki przyspieszenia m/s^2 , a dla temperatury stopnie Celsjusza ($^{\circ}C$). W przypadku danych z czujników inercyjnych konwersja polega na podzieleniu wartości bezpośredniego odczytu z czujników przez współczynniki odpowiednie dla zakresu pracy tych czujników. Tabele 3.1a oraz 3.1b przedstawiają zestawienie współczynników konwersji f_A i f_G wykorzystywanych do konwersji pomiarów odpowiednio z akcelerometru i żyroskopu. Jak

⁵⁴ Arduino – platforma programistyczna dla systemów wbudowanych, www.arduino.cc

⁵⁵ Urządzenie do pomiaru danych inercyjnych jest urządzeniem zaprojektowanym i wykonanym przez autora niniejszej dysertacji.

już zostało to zaznaczone wcześniej, w trakcie prowadzonych eksperymentów badawczych, zakres pracy akcelerometru został ustalony na $\pm 4g$, natomiast żyroskopu na $\pm 500^\circ/s$. Jako zakresy pracy czujników modułów inercyjnych przyjęto podobne wartości do tych zastosowanych w kontrolerze ruchu Nintendo Wii Remote Plus.



Rysunek 3.3: Schemat budowy urządzenia pomiarowego opartego o komputer Arduino (źródło: badania własne). IMU1 i IMU2 – Moduły inercyjne, RTC – moduł zegara oparty o komponent DS3230, BT – moduł Bluetooth oparty o komponent HC-05, LCD – ekran ciekłokrystaliczny

Tablica 3.1: Współczynniki konwersji bezpośrednich pomiarów akcelerometru (a) i żyroskopu (b) w zależności od przyjętego zakresu pracy⁵⁶

(a) Akcelerometr

Zakres pomiaru	Współczynnik
$\pm 2g$	16384
$\pm 4g$	8192
$\pm 8g$	4096
$\pm 16g$	2048

(b) Żyroskop

Zakres pomiaru	Współczynnik
$\pm 250 \frac{^\circ}{s}$	131.0
$\pm 500^\circ/s$	65.5
$\pm 1000 \frac{^\circ}{s}$	32.8
$\pm 2000 \frac{^\circ}{s}$	16.4

Konwersja pomiaru temperatury czujników (T_{raw}) do wartości wyrażonej w $^{\circ}C$ (T_{deg}) wymaga zastosowania wzoru 3.2

$$T_{deg} = 36.53 + T_{raw}/340.0 \quad (3.2)$$

wynikającego ze specyfikacji zastosowanego układu elektronicznego InvenSense MPU-6050. Wartości stałe zastosowane w tym wzorze związane są z rozdzielczością pomiarów wykorzystanego w tym układzie elektronicznym termometru (wartość = 340.0) oraz ze stałego przesunięcia dla temperatury $0^{\circ}C$ (wartość = 36.53)

Do danych pobranych ze wszystkich podłączonych modułów inercyjnych, podobnie jak w przypadku danych z Kinecta, dołączone są znaczniki czasowe zarówno ustawiane przez urządzenie, jak i przez komputer PC w momencie odbioru danych oraz numeracja kolejnych pakietów danych. Odczyt danych z modułów inercyjnych odbywa się z częstotliwością $100Hz$.

3.2. Kalibracja

W proponowanym systemie czujniki inercyjne wymagają kalibracji przy każdym uruchomieniu. Kalibracja Kinecta nie jest wymagana, a jedyne co należy zrobić to upewnić się, że kąt nachylenia kamery pozwala na obserwację śledzonej postaci.

Kalibracja czujników inercyjnych przebiega dwuetapowo w następującej kolejności:

1. Wyznaczenie współczynników korekty odchylenia pomiarów czujników inercyjnych dla wartości spoczynkowych.
2. Inicjalizacja filtru Madgwicka wyznaczającego orientację sensora na podstawie danych z czujników inercyjnych.

Krok pierwszy służy do wyznaczenia macierzy współczynników korekty ($cor = [cor_A \ cor_G]^T = [[c_{AX}, c_{AY}, c_{AZ}] \ [c_{GX}, c_{GY}, c_{GZ}]]^T$) pomiarów spoczynkowych (G, A) dla każdego modułu inercyjnego indywidualnie. W trakcie tego kroku kalibrowane czujniki należy umieścić możliwie jak najbardziej poziomo tak, żeby os Z sensora była równoległa do kierunku działania siły grawitacji. Dla czujników znajdujących się w takim właśnie położeniu można określić jakich wartości pomiarów należy oczekiwać ($A_0 = [0, 0, 1]$ dla akcelerometru oraz $G_0 = [0, 0, 0]$ dla żyroskopu). W praktyce jednak, ze względu na zaszumienie danych i brak możliwości całkowitego ich oczyszczenia, wartości idealne nie są osiągalne. W związku

⁵⁶ InvenSense, *MPU-6050 — InvenSense*, <https://goo.gl/vXcRhE>.

z tym, algorytm odpowiedzialny za kalibrację działa iteracyjnie (s – indeks iteracji) tak długo, dopóki jakikolwiek element macierzy średnich błędów pomiarów czujników inercyjnych ($[\bar{A} \quad \bar{G}]^T = [[\bar{a}_x, \bar{a}_y, \bar{a}_z] \quad [\bar{g}_x, \bar{g}_y, \bar{g}_z]]^T$) jest większy od elementu o tym samym indeksie w macierzy akceptowalnych błędów pomiaru ($[A_{th} \quad G_{th}]^T = [[a_{x,th}, a_{y,th}, a_{z,th}] \quad [g_{x,th}, g_{y,th}, g_{z,th}]]^T$). Stop następuje, gdy prawdziwy jest następujący warunek:

$$[\bar{A} \quad \bar{G}] \leq [A_{th} \quad G_{th}] \Leftrightarrow \bar{a}_x \leq a_{x,th} \wedge \bar{a}_y \leq a_{y,th} \wedge \bar{a}_z \leq a_{z,th} \wedge \bar{g}_x \leq g_{x,th} \wedge \bar{g}_y \leq g_{y,th} \wedge \bar{g}_z \leq g_{z,th}.$$

W rezultacie elementy macierzy cor pozwalają uzyskać pomiary w każdej z osi z założoną dokładnością. Przyjmując, że macierz $[A_0 \quad G_0]$ zawiera oczekiwane wartości pomiarów dla akcelerometru i żyroskopu, a $[A \quad G]_{s,i}$ to i -ty z n kolejnych pomiarów obu czujników dla iteracji s , średnie błędy pomiarów wyznaczone są zgodnie ze wzorem 3.3.

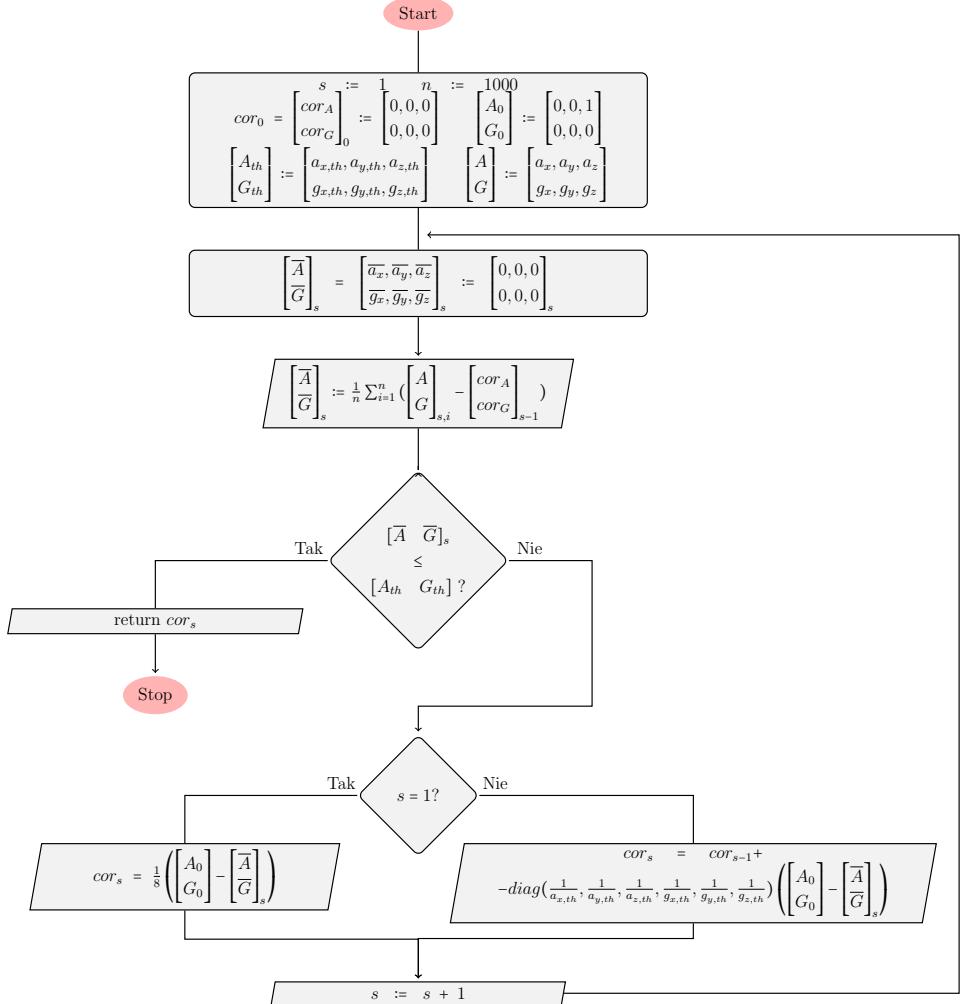
$$\begin{bmatrix} \bar{A} \\ \bar{G} \end{bmatrix}_s = \begin{cases} \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} A \\ G \end{bmatrix}_{s,i} & s = 1 \\ \frac{1}{n} \sum_{i=1}^n \begin{bmatrix} A \\ G \end{bmatrix}_{s,i} - \begin{bmatrix} cor_A \\ cor_G \end{bmatrix}_{s-1} & s > 1 \end{cases} \quad (3.3)$$

natomiast macierz współczynników korekty cor wyznaczona jest na podstawie wzoru 3.4.

$$\begin{aligned} cor_s = \begin{bmatrix} cor_A \\ cor_G \end{bmatrix}_s &= \begin{bmatrix} c_{AX} & c_{AY} & c_{AZ} \\ c_{GX} & c_{GY} & c_{GZ} \end{bmatrix}_s = \\ &\begin{cases} \frac{1}{8} \left(\begin{bmatrix} A_0 \\ G_0 \end{bmatrix} - \begin{bmatrix} \bar{A} \\ \bar{G} \end{bmatrix}_1 \right) & s = 1 \\ cor_{s-1} - diag\left(\frac{1}{a_{x,th}}, \frac{1}{a_{y,th}}, \frac{1}{a_{z,th}}, \frac{1}{g_{x,th}}, \frac{1}{g_{y,th}}, \frac{1}{g_{z,th}}\right) \left(\begin{bmatrix} A_0 \\ G_0 \end{bmatrix} - \begin{bmatrix} \bar{A} \\ \bar{G} \end{bmatrix}_n \right) & s > 1 \end{cases} \quad (3.4) \end{aligned}$$

Wyznaczone w ten sposób współczynniki korekty: c_{AX} , c_{AY} , c_{AZ} , c_{GX} , c_{GY} , c_{GZ} wykorzystywane są następnie do ciągłej korekty pomiarów uzyskiwanych z akcelerometrów i żyroskopów. Kalibracja modułu zajmuje zwykle około 10 iteracji, natomiast zdarzały się sytuacje, w których trwało to ponad 2 razy dłużej. Było to zazwyczaj spowodowane poruszeniem modułu w trakcie kalibracji. Schemat na rysunku 3.4 przedstawia schemat blokowy procedury kalibracji czujników inercyjnych.

Kolejnym krokiem jest inicjalizacja filtru Madgwicka łączącego dane z czujników inercyjnych w celu wyznaczenia ich orientacji w przestrzeni. Łączenie to ma na celu ograniczenie dryfu żyroskopu za pomocą pomiarów akcelerometru, a następnie wyznaczenie orientacji przestrzennej modułu inercyjnego. W literaturze najczęściej spotykanym filtrem wykorzystywanym



Rysunek 3.4: Diagram przedstawiający kalibrację czujników inercyjnych (źródło: badania własne)

do tego celu jest filtr Kalmana [80, 96, 97, 130] (szczegółowy opis filtrów Kalmana jak i filtru Madgwicka znajduje się w dodatku A). W roku 2010 Sebastian Madgwick opublikował raport ze swoich badań, w którym przedstawił wzór autorskiego filtru łączącego dane z akcelerometru i żyroskopu oraz wariant uzupełniony o dane z magnetometru, przedstawiający orientację urządzenia pomiarowego w chwili t , zbudowanego z tych czujników, w postaci kwaternionów [75, 76] ($Q_t^I = [q_w, q_x, q_y, q_z]$). W przytoczonych pracach opublikowane zostały także wyniki testów porównujących zaproponowany filtr z filtrem Kalmana, z których wynika, że filtr Madgwicka osiąga lepsze wyniki niż filtr Kalmana w wyznaczaniu orientacji przestrzennej. W opisywanej w tym rozdziale autorskiej hybrydowej metodzie śledzenia ruchu kończyn człowieka, wykorzystany został filtr Madgwicka ($m(\dots)$) opisany wzorem 3.5

$$Q_t^I = m(Q_{t-1}^I, A, G, \Delta t, f_m) \quad (3.5)$$

gdzie: Q_t^I reprezentuje kwaternion reprezentujący orientację czujnika inercyjnego w przestrzeni, A jest wektorem reprezentującym pomiar z akcelerometru ($A = [a_x, a_y, a_z]$), G – wektor reprezentujący pomiar z żyroskopu ($G = [g_x, g_y, g_z]$), Δt to czas pomiędzy kolejnymi pomiarami wyrażony w sekundach, a f_m jest współczynnikiem filtracji filtru Madgwicka.

Inicjalizacja filtru Madgwicka odbywa się poprzez jego wielokrotne zastosowanie dla uśrednionych pomiarów z czujników inercyjnych znajdujących się i dokonujących pomiary w stanie spoczynku. W trakcie inicjalizacji filtru Madgwicka powinien być on zastosowany taką liczbę razy jakiej liczności była próbka do wyznaczenia uśrednionych wartości pomiarów czujników. Na przykład, jeśli średnie wartości pomiarów były wyznaczone z próbki zawierającej 1000 wartości, to w ramach inicjalizacji filtr Madgwicka powinien być zastosowany 1000-krotnie dla uśrednionych danych.

Zgodnie z zaleceniami twórcy tej metody, w trakcie inicjalizacji wartość współczynnika filtracji f_m powinna być relatywnie wysoka. W opisywanej w niniejszej dysertacji metodzie, w momencie inicjalizacji filtru $f_m = 2$. Po zakończeniu inicjalizacji współczynnik f_m zostaje wyznaczony według wzoru 3.6.

$$f_m = \sqrt{\frac{3}{4}\tilde{\omega}} \quad (3.6)$$

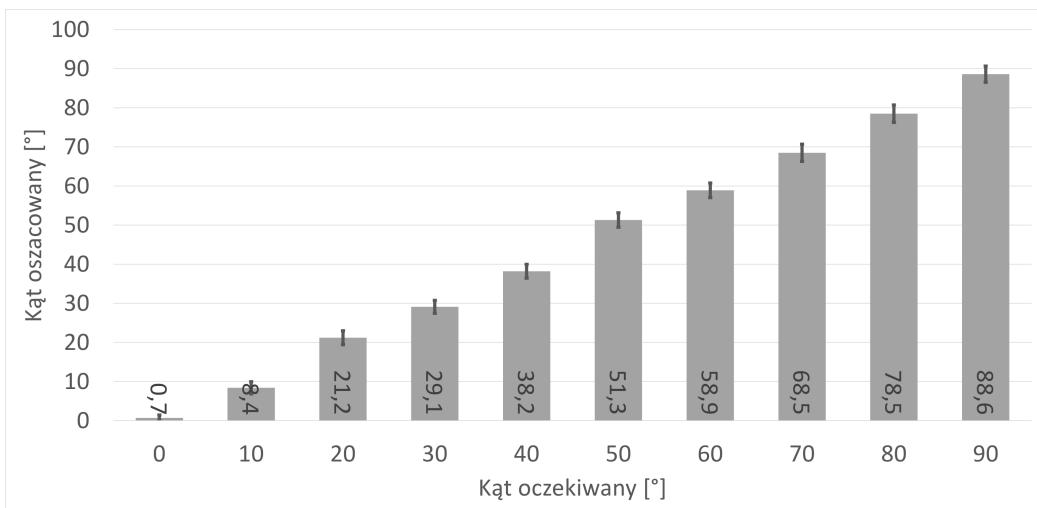
Jest on zależny od średniej wartości szumu błądzenia $\tilde{\omega}$ (błąd ARW – *ang. Angle Random Walk*) wyznaczonego dla żyroskopu na podstawie wariancji Allana (dodatek B). W przypadku wykorzystanych czujników wartość ARW wynosi $\tilde{\omega} = 0.009$, więc dla wykorzystywanych urządzeń przyjęto współczynnik $f_m = 0.082$.

3.3. Korekta danych z urządzeń pomiarowych

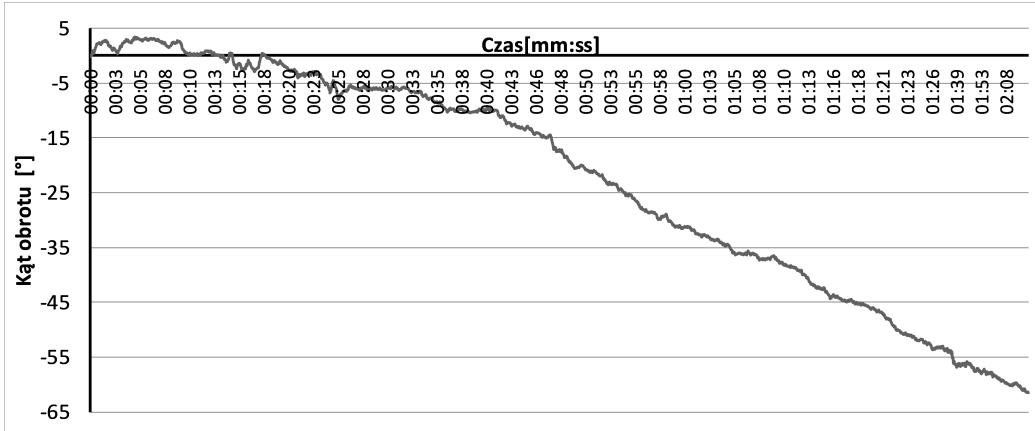
Dane pobrane z czujników inercyjnych oraz z kontrolera Kinect muszą zostać poddane filtracji, aby ograniczyć wpływ szumów na uzyskane pomiary. W przypadku czujników inercyjnych, wpływ szumów wynikający z błądzenia losowego zostaje skutecznie ograniczony w trakcie łączenia danych pobranych z akcelerometru z danymi pobranymi z żyroskopu za pomocą filtru Madgwicka, którego wynikiem jest oszacowana orientacja, w jakiej znajduje się dany moduł inercyjny. Ograniczenie wpływu szumów obecnych w danych pomiarowych ma bezpośrednie przełożenie na stabilność i dokładność oszacowanej orientacji. Dla osi X i Y , oszacowana wartość orientacji jest stabilna w czasie, a dokładność w stosunku do prawdziwej

orientacji modułu inercyjnego, zweryfikowana w trakcie własnych eksperymentów, wyniosła około $\pm 2^\circ$ (rys. 3.5). Oszacowanie orientacji w osi Z odbywa się jedynie na podstawie danych pomiarowych żyroskopu, co powoduje, że uzyskiwany wynik nie jest stabilny w czasie. Analizując wykres z rysunku 3.6, przedstawiający wartość oszacowania orientacji w osi Z w czasie dla modułu inercyjnego będącego w spoczynku, można zaobserwować ciągłą zmianę wartości tak jakby urządzenie pomiarowe wciąż się obracało. W trakcie przeprowadzanych badań zarejestrowano dryf dochodzący do około 60° w przeciągu 2 minut pomiaru. W związku z tym oszacowanie obrotu wokół tej osi jest całkowicie niewiarygodne i nie nadaje się do dalszego wykorzystania w obliczeniach.

Wyznaczając orientację przestrenną modułu inercyjnego zbudowanego jedynie z akcelerometru i żyroskopu, należy mieć na uwadze układ odniesienia dla otrzymywanych wartości. Oszacowanie będące wynikiem działania filtra Madgwicka dla osi X i Y wykorzystuje jako swój punkt odniesienia wektor siły grawitacji, co oznacza że te dwie wartości będą wskazywały orientację względem grawitacji ziemskiej. Orientacja względem osi Z , oszacowana jedynie na podstawie pomiarów żyroskopu, co do zasady wyraża obrót w stosunku do pozycji początkowej, w jakiej znajdował się moduł w trakcie inicjalizacji filtra Madgwicka. Sposobem na ustabilizowanie oszacowania orientacji wokół osi Z i odniesienia go względem grawitacji Ziemi byłoby uzupełnienie urządzenia pomiarowego o magnetometr.



Rysunek 3.5: Wykres przedstawiający uśrednione wartości oszacowania kąta obrotu wokół osi X i Y za pomocą modułu inercyjnego (źródło: badania własne)



Rysunek 3.6: Wykres przedstawiający dryf oszacowania kąta obrotu modułu inercyjnego wokół osi Z (źródło: badania własne)

Zanim jednak dane z akcelerometru i żyroskopu zostaną połączone ze sobą z wykorzystaniem filtru Madgwicka, pomiary uzyskane za pomocą akcelerometru poddane są korekcje ze względu na aktualną temperaturę pracy sensora.

$$A' = \frac{A}{1 + f_T(T - T_0)} \quad (3.7)$$

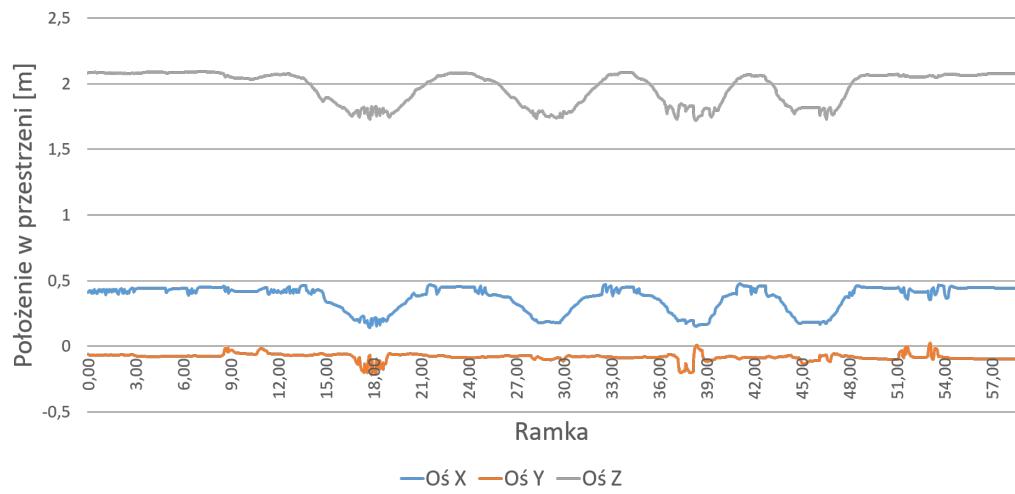
Wzór 3.7 przedstawia sposób korekty pomiaru z akcelerometru A w określonej temperaturze T . Temperatura T_0 to temperatura neutralna, która według specyfikacji wykorzystanego układu elektronicznego oraz na podstawie własnych badań dotyczących wpływu temperatury na dokładność pomiarów czujników inercyjnych, wynosi $25^\circ C$ (wykres na rys. 2.29). W znanej mnie literaturze przedmiotu, korekta taka nie była uwzględniana lub autorzy tych publikacji nie zamieścili stosownej informacji o tym fakcie.

Również dane dotyczące położenia w przestrzeni wybranych stawów szkieletu, pobrane z kontrolera Kinect, muszą zostać poddane filtracji. Umożliwia to zmniejszenie efektu "drgania" położenia stawów, których ruch jest śledzony oraz eliminuje krótkotrwałe zaburzenia śledzenia mające swój efekt w postaci znaczących różnic w położeniu danego stawu pomiędzy dwoma kolejnymi pomiarami. Na przykład przy częstotliwości pracy kontrolera Kinect wynoszącej 30 Hz przemieszczenie się stawu o kilkanaście centymetrów, pomiędzy dwoma kolejnymi oszacowaniami, może wskazywać błąd wyznaczania pozycji lub błąd pomiaru. Aby ograniczyć wpływ opisanych szumów na dalsze etapy działania opracowanej przez autora hybrydowej metody śledzenia ruchu, pozycje wybranych stawów, których ruch jest śledzony, zostały poddane filtracji dolnoprzepustowej za pomocą filtru wykładniczego I-go rzędu. Jego działanie opiera się na komplementarnym łączeniu ze sobą zaszumionych danych dotyczących pozycji P stawu j pobranych z kontrolera Kinect w chwili

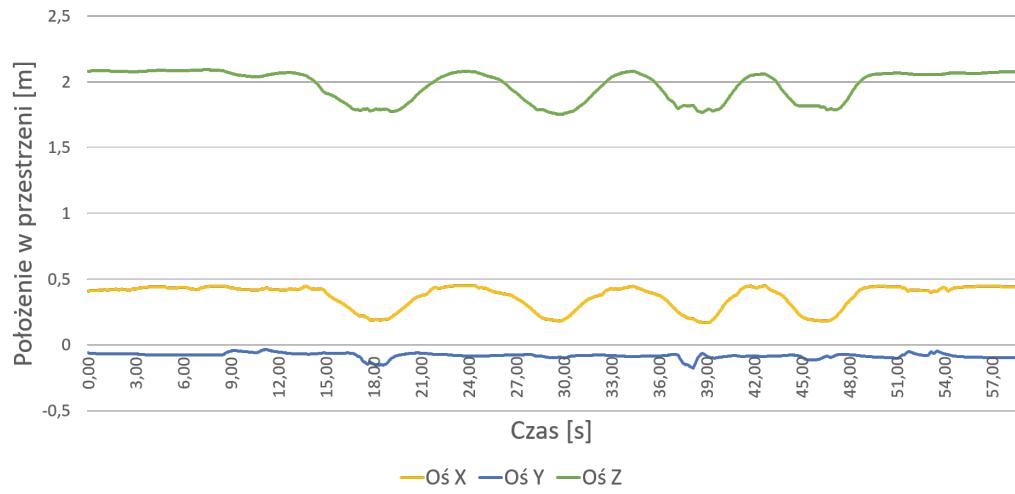
t ($P_{j,t}^K$) z wynikiem działania filtra w chwili $t - 1$ ($P_{j,t-1}^{tK}$) przy zastosowaniu współczynnika filtracji f_{LPF} zgodnie ze wzorem 3.8.

$$P_{j,t}^{tK} = f_{LPF} P_{j,t}^K + (1 - f_{LPF}) P_{j,t-1}^{tK} \quad (3.8)$$

Przykład danych dotyczący położenia stawu łokciowego przed i po zastosowaniu omówionego filtra wykładniczego I-go rzędu, przedstawiają wykresy z rysunków 3.7a oraz 3.7b.



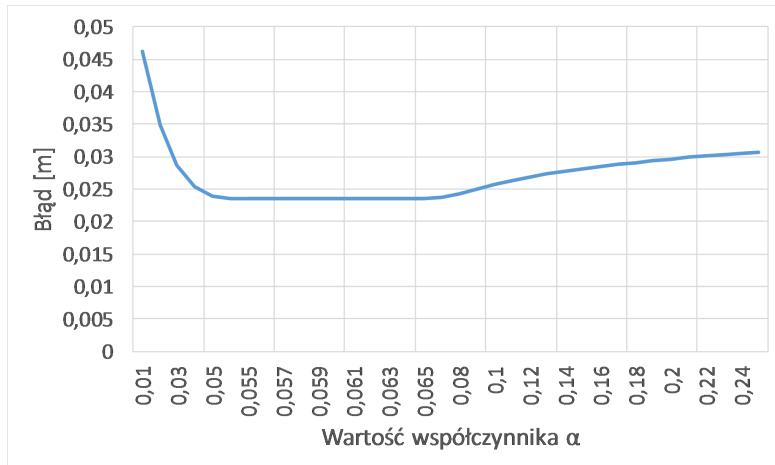
(a) Pomiar bezpośredni – zaszumiony



(b) Pomiar odszumiony

Rysunek 3.7: Wykres przedstawiający oszacowanie położenia stawu łokciowego za pomocą kontrolera Kinect na podstawie sygnału zaszumionego (a) oraz odszumionego (b) filtrem wyrażonym wzorem 3.8 (źródło: badania własne)

Wartość współczynnika filtracji f_{LPF} została wyznaczona w drodze prowadzonych badań własnych i została ustalona na 0.065. Przeprowadzone przez mnie badanie wpływu wartości współczynnika filtracji na dokładność oszacowania pozycji stawów, za pomocą kontrolera Kinect, polegało na wyznaczeniu średniego błędu oszacowania pozycji stawów między wskazaniami kontrolera Kinect, a danymi referencyjnymi uzyskanymi z systemu śledzenia ruchu firmy Vicon. W trakcie badania były wykonywane ruchy samych kończyn jak i całego ciała, śledzone równocześnie przez system Vicon jak i kontroler Kinect. Następnie wyznaczany był średni błąd oszacowania pozycji dla wybranych stawów, z wykorzystaniem filtra oraz bez niego, a także z różnymi wartościami współczynnika f_{LPF} . Rysunek 3.8 przedstawia wykres zależności uzyskanego średniego błędu oszacowania pozycji stawów w zależności od wartości współczynnika filtracji f_{LPF} .



Rysunek 3.8: Wykres przedstawiający średni błąd pozycjonowania stawów za pomocą kontrolera Kinect w zależności od współczynnika filtracji α (źródło: badania własne)

Korekcie muszą zostać poddane także oszacowania odległości od kontrolera Kinect, w jakiej znajdują się wybrane stawy. Krok ten nie był uwzględniany w dotychczas opublikowanej, znanej autorowi literaturze, a ze względu na zauważalną różnicę pomiędzy oszacowaniami Kinecta a rzeczywistą odlegością, w jakiej stoi użytkownik, ma on wpływ na błędy w dalszych obliczeniach. Zgodnie ze wzorami 2.8 i 2.9, definiującymi model błędu szacowania odległości pomiędzy osobą, której ruch jest śledzony a Kinectem, funkcja korygująca ten błąd przybiera postać wzoru 3.9.

$$f(z) = z' = -0.02z^3 + 0.11z^2 - 0.27z + 0.25 \quad (3.9)$$

gdzie wartość z to oszacowanie odległości podlegające korekcie, a z' to wartość skorygowana. Formuła korygująca oszacowanie odległości wykonane

przez kontroler Kinect jest efektem własnego eksperymentu opisanego szerzej w rozdziale 2.3.1.

3.4. Synchronizacja czasowa

Aby dane reprezentujące położenie badanego stawu, na podstawie czujników inercyjnych oraz kontrolera Kinect, mogły być poprawnie złączone, powinny zostać najpierw zsynchronizowane w czasie. Jest to proces, który powinien zostać przeprowadzony przed każdą sesją śledzenia ruchu, na przykład przed konkretną sekwencją ćwiczeń do wykonania. Synchronizacja czasowa ma na celu określenie przesunięcia czasowego, jakie występuje pomiędzy oszacowaniami położenia tych samych stawów w trakcie tego samego ruchu za pomocą każdego z urządzeń pomiarowych osobno – czujników inercyjnych i kontrolera Kinect. Informacja dotycząca kroku związanego z synchronizacją sygnałów rozważanych urządzeń nie znalazła się w żadnej ze znanych autorowi pozycji bibliograficznych dotyczących omawianego przedmiotu, co może wskazywać na jego pominięcie w procesie przetwarzania i łączenia sygnałów źródłowych. To z kolei może skutkować powstaniem błędów na skutek łączenia nie skorelowanych ze sobą danych.

Do wyznaczenia przesunięcia czasowego pomiędzy sygnałem uzyskiwanym z czujników inercyjnych ($I[t]$), a tym z kontrolera Kinect ($K[t]$) wykorzystany został algorytm korelacji wzajemnej (*ang. cross-correlation*) określony wzorem 3.10.

$$(I * K)(\tau) = \int_{-\infty}^{+\infty} I[t]K[t + \tau]dt \quad (3.10)$$

Parametr τ odpowiada za opóźnienie jednego sygnału względem drugiego. Aby wyznaczyć jakie jest przesunięcie czasowe (τ_{max}) pomiędzy dwoma sygnałami należy znaleźć taką wartość argumentu τ dla której wartość funkcji korelacji wzajemnej pomiędzy badanymi sygnałami jest największa, co przedstawia wzór 3.11.

$$\tau_{max} = argmax_{\tau}((I * K)(\tau)) \quad (3.11)$$

Aby można było skutecznie wyznaczyć tę wartość, oba sygnały muszą posiadać tę samą częstotliwość próbkowania. Spełnienie tego warunku uzyskano poprzez obniżenie częstotliwości próbkowania sygnału uzyskanego z czujników inercyjnych ze 100Hz do 30Hz, to znaczy została ona zrównana do tej, z jaką pracuje Kinect. Pomiary z czujników inercyjnych są tymczasowo buforowane i gdy zostanie pobrany pakiet danych z Kinecta, pomiary z modułów inercyjnych zgromadzone w tymczasowym buforze zostają poddane filtrowi

medianowemu, aby wyeliminować ewentualne krótkotrwałe zaburzenia pomiarów.

W znanej autorowi niniejszej dysertacji literaturze przedmiotu, korekta taka nie była uwzględniania lub autorzy tych publikacji nie zamieścili stosownej informacji o tym fakcie.

Wyznaczona wartość przesunięcia w czasie pomiędzy sygnałami (τ_{max}) wykorzystywana jest następnie do modyfikacji wartości znacznika określającego czas odebrania danych z modułu inercyjnego przez komputer PC. Tak zmodyfikowany znacznik czasu wykorzystywany jest do wyznaczenia pomiarów z czujników inercyjnych odpowiadających pomiarom uzyskanym z kontrolera Kinect. Metoda określająca odpowiadające sobie pomiary z kilku sygnałów, na podstawie czasów otrzymania wartości każdego z badanych sygnałów, nazywana jest decymacją [47]. Przyjmując, że t_K jest znacznikiem czasowym otrzymania na komputerze PC danych z kontrolera Kinect, a $t_I[]$ to zbiór znaczników czasowych próbek otrzymanych z czujników inercyjnych i umieszczonych w buforze, wówczas wybór próbki danych pomiarowych z czujników inercyjnych skorelowanych z danymi z kontrolera Kinect wyrażony jest wzorem 3.12

$$i_{res} = \underset{i}{\operatorname{argmin}}(|t_K - (t_I[i] + \tau_{max})|) \quad (3.12)$$

gdzie i_{res} to numer próbki umieszczonej w buforze danych odpowiadającej danym otrzymanym z Kinecta ze znacznikiem czasowym t_K .

W trakcie prowadzonych autorskich prac badawczych, oprócz wyznaczania przesunięcia czasowego pomiędzy czujnikami inercyjnymi, a kontrolerem Kinect, zostało zbadane przesunięcie czasowe pomiędzy tymi urządzeniami pomiarowymi, a systemem śledzenia ruchu Vicon. Przyjmując, że Vicon działa w czasie rzeczywistym, przesunięcia te określają de facto jakie jest opóźnienie między IMU i Kinectem, a rzeczywistym ruchem. I tak, średnie przesunięcie czasowe pomiędzy czujnikami inercyjnymi a systemem Vicon wyniosło około $0.04s$, natomiast pomiędzy Kinectem a systemem Vicon to około $0.09s$. Korelacja wyznaczona między czujnikami inercyjnymi a kontrolerem Kinect wskazuje, że średnie przesunięcie szacowania pozycji wynosi około $0.06s$. Z uwagi na to, że dane z czujników inercyjnych są łączone z pomiarami uzyskanymi z kontrolera Kinect, zgodnie z częstotliwością pracy Kinecta, zasadnym jest wskazanie, że opisywana przez autora metoda łączenia danych ze wspomnianych urządzeń posiada opóźnienie $0.09s$ względem rzeczywistego ruchu.

3.5. Łączenie danych

Strumienie danych pochodzące z poszczególnych sensorów, zsynchronizowane czasowo oraz poddane filtracji mogą zostać połączone celem uzyskania wypadkowych wartości, lepiej odzwierciedlających układ szkieletu śledzonej postaci. Na tym etapie prezentowana metoda opiera się na informacji o orientacji przestrzennej poszczególnych kości zamiast na położeniu konkretnych stawów. Jest to nowe podejście do łączenia sygnałów z czujników inercyjnych z sygnałami z kontrolera Kinect, które stanowi jedno z głównych osiągnięć prezentowanej pracy. Podejście takie podyktowane jest faktem, że w przypadku czujników inercyjnych, wyznaczenie położenia stawów wymaga połączenia informacji o orientacji przestrzennej czujników z modelem szkieletowym zawierającym informację o długościach poszczególnych kości. Z uwagi, że wykorzystując definicję modelu szkieletowego, definicja długości jego segmentów jest stała, należy zwrócić szczególną uwagę na dokładność pomiarów tych długości. Wszelkie błędy i niedokładności popełnione na tym etapie przygotowywania modelu mogą skutkować pojawiением się istotnych błędów w oszacowaniu położenia poszczególnych stawów. Co więcej, ze względu na hierarchiczną budowę modelu szkieletowego, ewentualne błędy wynikające z braku dokładnych pomiarów długości poszczególnych kości szkieletu będą się kumulowały wraz z odległością danego segmentu od korzenia.

Wykorzystanie położenia stawów szkieletu, w procesie łączenia sygnałów z obu źródeł, dodatkowo utrudnia fakt niedokładności w oszacowaniu modelu szkieletowego za pomocą kontrolera Kinect. Model ten nie ma zdefiniowanych długości poszczególnych kości na stałe (choćby na czas pojedynczej sesji śledzenia) i są one determinowane przez oszacowanie położenia kolejnych stawów, obarczonych zauważalną zmiennością w czasie. To z kolei skutkuje tym, że długości poszczególnych kości modelu szkieletowego, oszacowanego przez kontroler Kinect, różnią się nawet o kilka *cm* pomiędzy kolejnymi pomiarami.

Dane z obu źródeł (kontroler Kinect i moduł inercyjny) zawierają informacje o orientacji przestrzennej wyrażone w kwaterionach, jednak dalsze obliczenia, mające na celu połączenie tych informacji, są bardziej intuicyjne przy reprezentacji w postaci kątów Eulera. Konwersja taka przeprowadzona jest zgodnie ze wzorem 2.6, a w jej wyniku wartości obrotów wokół każdej z osi układu współrzędnych danego urządzenia pomiarowego, które składają się na orientację przestrzenną danej kości, są wyrażone bezpośrednio. Reprezentacja w postaci kątów Eulera ułatwia także konwersję informacji o obrotach do wspólnego, globalnego układu współrzędnych, co jest niezbędne, żeby można było je w dalszych krokach łączyć. Informacje o orientacji kości, wyznaczone na podstawie danych

pomiarowych z czujników inercyjnych oraz kontrolera Kinect, są od siebie niezależne i uzupełniają się nawzajem. Wzajemna niezależność i uzupełnianie się danych umożliwia ich połączenie w sposób komplementarny. Łącząc dane w sposób komplementarny, każdej z wartości podlegającej łączeniu przypisane zostają wagi, których suma wynosi 1, a następnie są one ze sobą sumowane z uwzględnieniem wag określających poziom istotności poszczególnych wartości kątów. Wynik sumowania reprezentuje wartość złączonych danych wejściowych.

Prezentowana w niniejszej dysertacji hybrydowa metoda śledzenia ruchu kończyn człowieka łączy ze sobą dane pochodzące z dwóch źródeł. Dane te są w postaci trójelementowych wektorów orientacji $E^i = [\phi^i \ \theta^i \ \psi^i]$ ($i = \{I, K\}$, I – IMU; K – Kinect), gdzie każdy z elementów oznacza obrót wokół pojedynczej osi układu współrzędnych – odpowiednio X Y Z . Wektor ten stanowi reprezentację orientacji przestrzennej segmentu szkieletu w postaci kątów Eulera, wyznaczoną przez konwersję kwaternionów otrzymanych na podstawie pomiarów czujników inercyjnych (E^I) lub kontrolera Kinect(E^K). Wagi wykorzystywane w procesie łączenia danych są wyznaczane indywidualnie dla każdego z elementów wektora. Przyjmując, że wagi w_ϕ, w_θ, w_ψ oznaczają poziom istotności informacji otrzymanych na podstawie pomiarów czujników inercyjnych ($E^I = [\phi^I \ \theta^I \ \psi^I]$), wówczas komplementarne łączenie ich z danymi otrzymanymi z kontrolera Kinect ($E^K = [\phi^K \ \theta^K \ \psi^K]$) dla danej chwili t pozwala oszacować złączone kąty Eulera E^F (F – złączone, ang. *Fused*) według wzoru 3.13

$$E_t^F = \begin{bmatrix} \phi^F \\ \theta^F \\ \psi^F \end{bmatrix}_t = \begin{bmatrix} w_\phi * \phi^I \\ w_\theta * \theta^I \\ w_\psi * \psi^I \end{bmatrix}_t + \begin{bmatrix} (1 - w_\phi) * \phi^K \\ (1 - w_\theta) * \theta^K \\ (1 - w_\psi) * \psi^K \end{bmatrix}_t \quad (3.13)$$

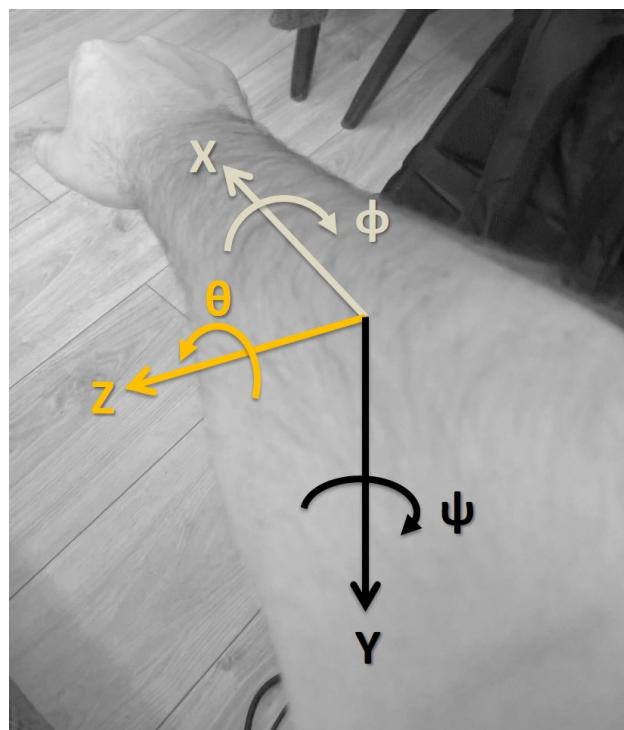
Wzór 3.13 jest wykorzystywany zawsze wtedy, kiedy dane podlegające łączeniu są dobrej jakości i możemy je uznać za prawidłowe. W takiej sytuacji współczynniki wag w_ϕ, w_θ, w_ψ są wartościami stałymi dobranymi eksperymentalnie przez autora i wynoszącymi odpowiednio 0.98, 0.05, 0.65.

Waga w_ϕ odpowiada za obrót wokół osi przechodzącej wzdłuż ciała (oś X , rys. 2.15 i 3.9). Wysoka wartość tej wagi wynika z faktu, że wartość obrotu wokół tej osi jest mierzalna jedynie przez czujniki inercyjne. Dzięki wadze zbliżonej do wartości 1, pomiary uzyskane z Kinecta są niemal ignorowane. w_θ jest wagą określającą istotność informacji o obrocie wokół osi zgodnej z kierunkiem działania grawitacji (oś Y , rys. 2.15 i 3.9). Wartość tego obrotu jest bardziej wiarygodnie i stabilnie szacowana w czasie przez kontroler Kinect. Niska wartość wagi w_θ zmniejsza wpływ danych wyznaczonych na

podstawie pomiarów czujników inercyjnych (obarczonych znaczącym dryfem) na wartości wypadkowe będące wynikiem łączenia danych.

Waga w_ψ określa poziom istotności informacji o obrocie wokół osi skierowanej w stronę obserwacji kontrolera Kinect (rys. 2.15). Obrót ten jest szacowany zarówno przez kontroler Kinect jak i na podstawie pomiarów z czujników inercyjnych. Oszacowania te mają jednak różną dokładność i różnica ta jest odzwierciedlona w przyjętej empirycznie wartości wagi. Według badań opublikowanych przez twórcę filtru Madgwicka [75], wykorzystywanego w niniejszej pracy, jego metoda wyznaczania orientacji przestrzennej, na podstawie pomiarów z czujników inercyjnych, zapewnia dokładność około $\pm 2^\circ$. W badaniach własnych autora dokładność ta była zbliżona do $\pm 3^\circ$. Z kolei dokładność oszacowania obrotu wokół osi Z kontrolera Kinect (rys. 2.15, rys. 3.9) przez to właśnie urządzenie, określona na podstawie badań własnych jak i opisów dostępnych w literaturze [49], wynosi około $\pm 6^\circ$. Niemal dwukrotnie mniejszy błąd oszacowania obrotu wokół osi Z , na podstawie pomiarów z czujników inercyjnych, ma swoje przełożenie na wartość wagi w_ψ .

Rysunek 3.9 przedstawia układ współrzędnych wraz z obrotami odpowiadającymi każdej z osi tego układu, w odniesieniu do ręki człowieka.



Rysunek 3.9: Oznaczenie układu współrzędnych XYZ wraz z obrotami ϕ, θ, ψ w odniesieniu do ręki człowieka (źródło: opracowanie własne)

Określenie czy oszacowania kątów, uzyskane za pomocą kontrolera Kinect, są wystarczającej jakości do łączenia ich z oszacowaniami kątów uzyskanymi

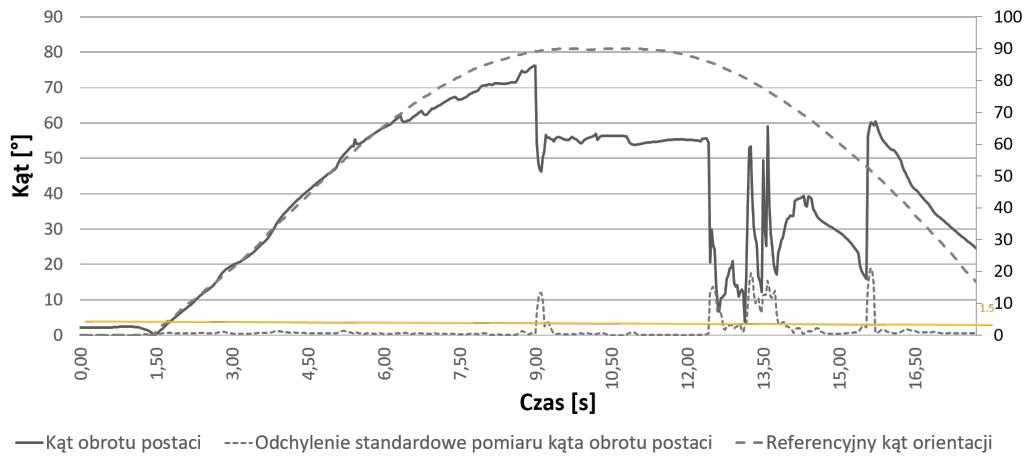
z czujników inercyjnych, wykonywane jest każdorazowo na podstawie kryteriów związań z realizowanym ruchem.

Pierwszym kryterium branym pod uwagę jest obrót całej sylwetki względem kamery, liczony jako kąt α pomiędzy linią barków a kamerą w płaszczyźnie $X-Z$ (rys. 2.23) zgodnie ze wzorem 2.7. Na podstawie zbadanych wcześniej charakterystyk obu urządzeń wiadomo, że dane które można wiarygodnie wykorzystać do łączenia są możliwe do uzyskania tylko gdy sylwetka nie jest obrócona o więcej niż 50° . Oprócz wartości kąta obrotu sylwetki brane jest także pod uwagę odchylenie standardowe estymacji tego obrotu. Przyjmując, że $\alpha[]$ to zbiór wartości oszacowań kąta obrotu sylwetki z ostatnich $3s$ (ok. 90 elementów w zbiorze), a $\bar{\alpha}$ to średnia arytmetyczna elementów w tym zbiorze, wówczas odchylenie standardowe σ_α definiuje wzór 3.14.

$$\sigma_\alpha = \sqrt{\frac{1}{n} \sum_{i=1}^n (\alpha[i] - \bar{\alpha})^2} \quad (3.14)$$

Branie pod uwagę zarówno wartości estymacji obrotu sylwetki wraz z odchyleniem standardowym uzyskiwanych oszacowań spowodowane jest faktem, że w momencie wystąpienia okluzji jednego lub obu stawów barkowych i błędów oszacowania ich położenia, następuje błędne oszacowanie wartości obrotu. W praktyce oznacza to, że wyznaczony kąt obrotu sylwetki może przyjąć dowolną wartość. Przykład szacowania kąta obrotu sylwetki względem kontrolera Kinect przedstawia wykres z rysunku 3.10. W trakcie tego eksperymentu badawczego wykonywany był obrót postaci względem kontrolera Kinect do kąta 90° i powrót do pozycji równoległej do urządzenia pomiarowego, czyli do kąta 0° . Przy obrocie sylwetki nie przekraczającym 50° odchylenie standardowe oszacowań z przyjętego okresu czasu nie przekroczało wartości 1.5° . W związku z tym w opisywanej w niniejszej dysertacji hybrydowej metodzie śledzenia ruchu kończyn człowieka, autor przyjął, że pomiar obrotu postaci w stosunku do kontrolera Kinect jest wiarygodny, jeśli oszacowanie jego kąta obrotu jest $\alpha \leq 50^\circ$ przy odchyleniu standardowym $\sigma_\alpha \leq 1.5^\circ$.

Oprócz obrotu całej postaci, ocenie podlega także stabilność pomiarów samych stawów. Pierwszym parametrem, jaki brany jest pod uwagę, jest stan śledzenia poszczególnych stawów, który może przyjąć jedną z trzech wartości opisanych szerzej w podrozdziale 2.3.1. Wartość *NotTracked* jest od razu traktowana jako wskazanie, że dane związane z danym stawem (jego pozycja, ale także orientacja kości z nim związanych) są niewiarygodne. W przypadku pozostałych dwóch możliwych stanów śledzenia stawu, sprawdzany jest poziom szumu za pomocą filtra górnoprzepustowego w postaci wzoru 3.15.

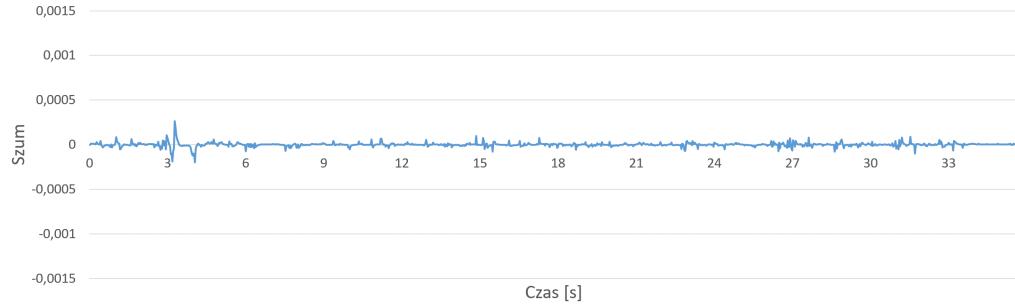


Rysunek 3.10: Odchylenie standardowe estymacji kąta obrotu postaci względem Kinecta (źródło: badania własne)

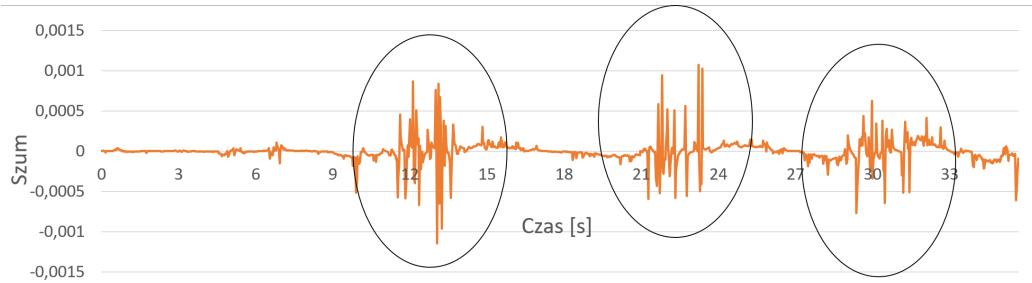
$$n_{j,t}^K = f_{HPF} * (n_{j,t-1}^K + P_{j,t}^K - P_{j,t-1}^K) \quad (3.15)$$

gdzie: $n_{j,t}^K$ to poziom szumu n w aktualnym oszacowaniu położenia danego stawu j w czasie t dla Kinecta, $n_{j,t-1}^K$ – poziom szumu n w poprzednim oszacowaniu położenia danego stawu j w czasie $t-1$ dla Kinecta, $P_{j,t}^K$ jest położeniem P stawu j w aktualnym oszacowaniu t dla Kinecta, $P_{j,t-1}^K$ to położenie P stawu j w poprzednim oszacowaniu $t-1$ dla Kinecta, natomiast f_{HPF} jest współczynnikiem filtracji $f_{HPF} = 0.01$.

Wykresy z rysunków 3.11a i 3.11b przedstawiają wynik działania filtru górnoprzepustowego dla oszacowania położenia stawu nadgarstkowego w przypadku kiedy staw cały czas posiadał status śledzenia *Tracked* oraz kiedy w trakcie wykonywanego ruchu status zmieniał się na *Interferred*. Dla zachowania przejrzystości, wykresy przedstawiają wspomniany wynik dla położenia stawu wyznaczonego wzdłuż jednej z osi. Na wykresie z rysunku 3.11b zostały zaznaczone ovalami te miejsca, w których nastąpiła zmiana statusu śledzenia. Na wspomnianych wykresach widać, że poziom szumu pozostaje niski dla oszacowania położenia stawu, kiedy status jego śledzenia przyjmuje wartość *Tracked*. W innym przypadku poziom szumu wzrasta wielokrotnie i utrzymuje się do momentu ponownego przyjęcia statusu *Tracked*. Zauważać można wówczas częste zmiany oszacowania pozycji danego stawu przez kontroler Kinect, co wiąże się z widocznym efektem drgania. Na podstawie przeprowadzonych badań własnych, w metodzie śledzenia ruchu, jako graniczną wartość zaszumienia danych pomiarowych kontrolera Kinect (n^K), dla których możliwe jest ich łączenie z danymi uzyskanymi z czujników inercyjnych na podstawie wzoru 3.13, autor przyjął $|n^K| \leq 0.0004$.



(a) Staw śledzony ze statusem *Tracked*



(b) Staw śledzony ze statusami *Tracked* i *Interferred* (oznaczone owalem)

Rysunek 3.11: Wykres wartości szumu występującego w oszacowaniach do położenia wybranych stawów w przestrzeni w osi Z dla stawu śledzonego ze stałym statusem *Tracked* (a) oraz zmiennym statusem *Tracked* i *Interferred* (b) (źródło: badania własne)

W przypadku, kiedy położenie stawów, wyznaczone przez kontroler Kinect, nie może być uznane za wiarygodne (zazwyczaj spowodowane jest to okluzją), równanie 3.13 ulega modyfikacji i przyjmuje postać jak w równaniu 3.16.

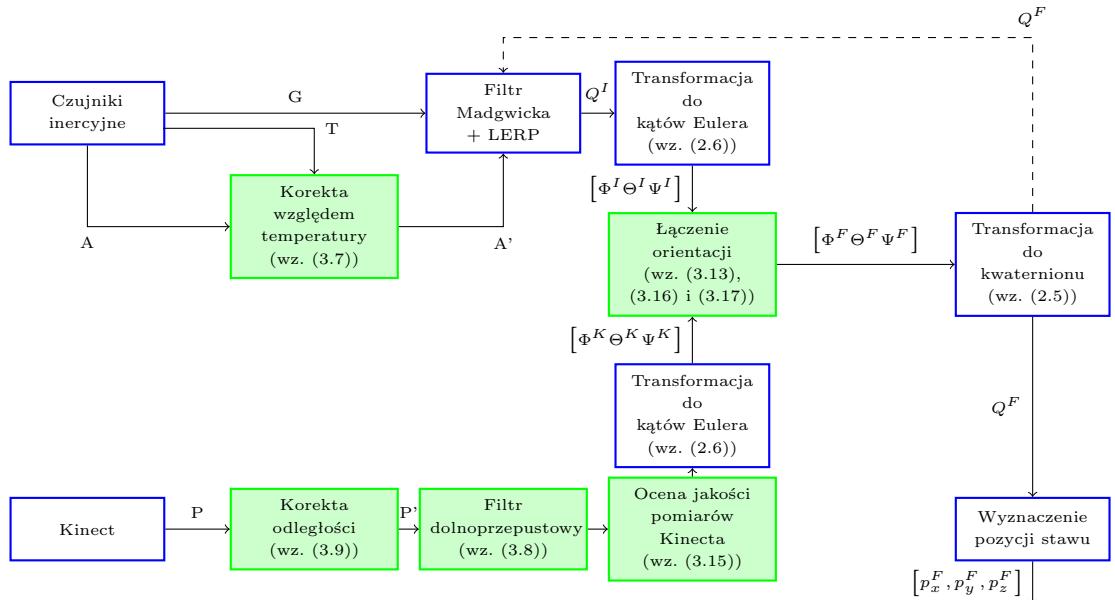
$$E_t^F = \begin{bmatrix} \phi^F \\ \theta^F \\ \psi^F \end{bmatrix}_t = \begin{bmatrix} \phi^F \\ \theta^F \\ \psi^F \end{bmatrix}_{t-1} + \text{diag}(w_\phi, w_\theta, w_\psi) \left(\begin{bmatrix} \phi^I \\ \theta^I \\ \psi^I \end{bmatrix}_t - \begin{bmatrix} \phi^I \\ \theta^I \\ \psi^I \end{bmatrix}_{t-1} \right) \quad (3.16)$$

W tym przypadku zmianie ulegają wartości części wag, z jakimi zostają połączone dane. O ile wartość w_ϕ nie ulega zmianie i wynosi 0.98, o tyle pozostałe dwie wagi są zależne od czasu i wyrażają się wzorem 3.17.

$$w_\theta = w_\psi = f(t_{szum}) = (1 - \frac{t_{szum}}{10}) * 0.65 \quad (3.17)$$

Wartość t_{szum} oznacza czas trwania okresu niepewności pomiarów wyrażony w sekundach. Jeśli czas, w którym utrzymuje się niepewność pomiarów kontrolera Kinect, przekracza 10s wówczas wartość t_{szum} przyjmuje wartość

10s, a wagi w_θ i w_ψ utrzymują wartość 0. Maksymalny czas kiedy szacowanie położenia stawów może odbywać się tylko na podstawie modułów inercyjnych uzależniony jest od stabości ich pomiarów w czasie i tego jaki jest maksymalny dopuszczony błąd szacowania położenia stawów dla projektowanego systemu. Jako przykład może tutaj posłużyć dryf pomiarów jaki występuje przy szacowaniu obrotu wokół osi równoległej do wektora grawitacji (kąt θ). Jeśli dryf ten wynosi około $0.5^\circ/s$ dla wykorzystanych modułów inercyjnych wówczas w ciągu 10s musimy liczyć się z błędem szacowania obrotu dochodzącym do 5° co przekłada się na błąd szacowania pozycji stawu dochodzący do 2.5cm dla kości o długości 30cm . Wobec uzyskiwanych dokładności szacowania pozycji stawów, w systemach śledzenia łączących dane z Kinecta z danymi z czujników inercyjnych, które wynoszą $2.5 - 3\text{cm}$ oznaczałoby to dwukrotny spadek dokładności szacowania w ciągu 10s. W związku z tym, w zaproponowanej metodzie, po utracie możliwości łączenia danych z obu źródeł, szacunkowa pozycja stawu ((3.16) i (3.17)) jest aktualizowania właśnie przez 10s. Po tym czasie musi nastąpić możliwość korekty oszacowania położenia stawów na podstawie danych z obu urządzeń. W przeciwnym razie użytkownik mógłby być wprowadzany w błąd.



Rysunek 3.12: Diagram przedstawiający kolejne etapy działania autorskiej metody łączenia danych IMU i Kinecta (źródło: badania własne)

Poszczególne etapy metody opisanej powyżej przedstawione są na diagramie z rysunku 3.12 i odzwierciedlają kroki opisane w niniejszym rozdziale. Kolorem zielonym zostały oznaczone etapy opisywanej metody, które stanowią autorskie dokonanie autora niniejszej pracy. Zostały one samodzielnie opracowane przez autora (wz. (3.9), (3.13), (3.16)), kryteria określenia

użyteczności danych Kinecta dla procesu łączenia ich z danymi z czujników inercyjnych, m.in. (3.15)) lub po raz pierwszy, kompleksowo uwzględnione w procesie łączenia danych z kontrolera ruchu Kinect i sensorów inercyjnych (wz. (3.7) i (3.8)). Elementy te miały zauważalny wpływ na jakość uzyskiwanych wyników.

Rozdział 4

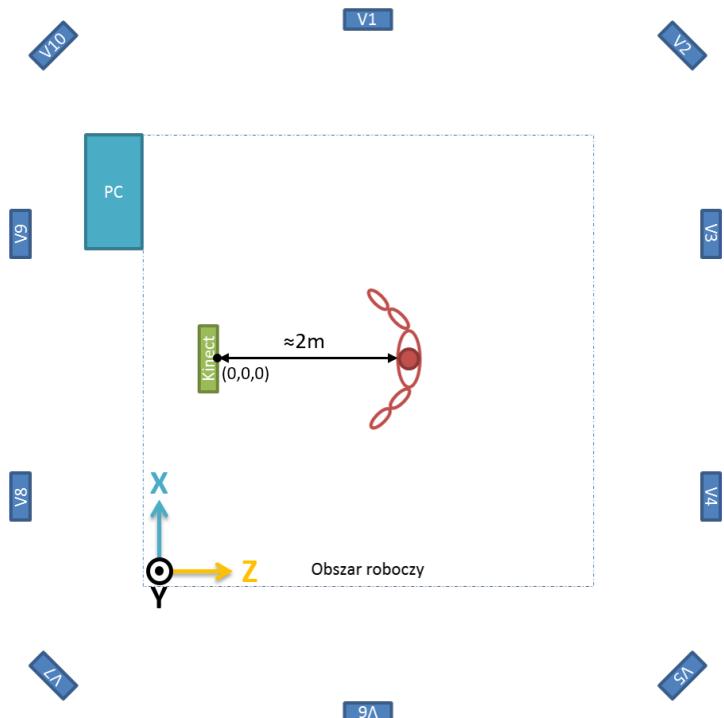
Eksperymenty badawcze

Eksperymenty badawcze wyryfikujące charakterystyki sensorów, jak również dokładności opracowanej autorsko hybrydowej metody śledzenia ruchu kończyn, zostały wykonane w laboratorium technik multimedialnych w Centrum Technologii Informacyjnych Politechniki Łódzkiej. Jako źródło danych referencyjnych wykorzystany został zainstalowany tam system śledzenia ruchu firmy Vicon zbudowany na bazie 10 kamer Vicon T-40S⁵⁷ rejestrujących ruch markerów. System ten został skonfigurowany do działania z częstotliwością 100 klatek na sekundę, a dane przetwarzane były przez dedykowaną aplikację Nexus w wersji 2.0. Dane testowe niezbędne do zweryfikowania omawianej metody zostały zgromadzone przy użyciu kontrolera Microsoft Kinect w wersji dla konsoli XBox 360 model 1414 oraz przez urządzenie rejestrujące dane z czujników inercyjnych zaprojektowane i wykonane samodzielnie w oparciu o platformę programistyczną dla systemów wbudowanych Arduino. Szczegółowy opis przeprowadzonych eksperymentów badawczych znajduje się w dalszej części niniejszego rozdziału.

4.1. Scena

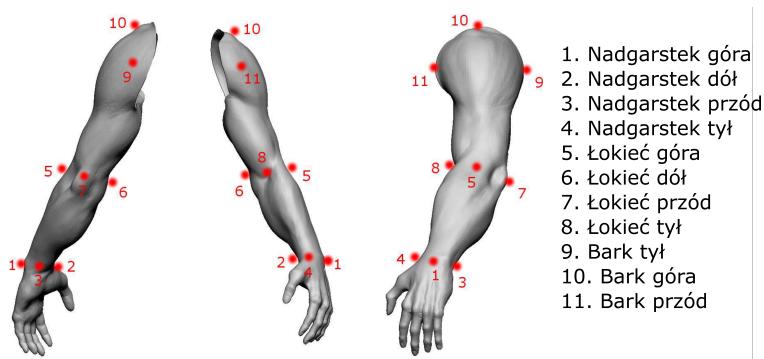
Obszar roboczy, na którym dokonywane były pomiary ruchu, został zamknięty w przestrzeni prostopadłościanu, obserwowanej przez kamery systemu Vicon, której obrys podstawy zbliżony był do kwadratu $4m \times 4m$. W trakcie kalibracji systemu referencyjnego punkt początkowy $(0, 0, 0)$ został zdefiniowany w miejscu, w którym został umieszczony kontroler Kinect. W trakcie trwania eksperymentu, badany użytkownik stał w odległości ok. $2m$ od Kinecta, czyli w odległości, w której działa on najbardziej dokładnie. Pełne rozmieszczenie urządzeń na scenie prezentuje rysunek 4.1. Prostokąty dookoła obszaru roboczego oznaczone V1 – V10 to kamery systemu Vicon.

⁵⁷ Vicon Corporation, *Vantage Motion Capture Camera — VICON*, <https://goo.gl/NRBDC7>.



Rysunek 4.1: Schemat rozmieszczenia elementów sceny w laboratorium (źródło: opracowanie własne)

Eksperymenty badawcze, opisane w niniejszej pracy, zostały przeprowadzone na przykładzie ruchu ręki, a oszacowaniu położenia w przestrzeni podlegał staw łokciowy oraz nadgarstkowy prawej ręki. Nie ogranicza to jednak możliwości zastosowania opracowanej w pracy metody do śledzenia pozostałych kończyn. Markery niezbędne do prawidłowego działania systemu śledzenia ruchu firmy Vicon zostały umieszczone według schematu zamieszczonego na rysunku 4.2.



Rysunek 4.2: Schemat umieszczenia markerów na ręce zgodny z zaleceniami systemu Vicon (źródło: opracowanie własne)

Sposób umieszczenia markerów został zaczerpnięty ze schematu ich rozmieszczenia zawartego w instrukcji do systemu Vicon, jednak poszerzony

został o dodatkowe markery tak, aby nadgarstek i łokieć śledzony był za pomocą czterech znaczników, a ramię za pomocą trzech. Aby móc skutecznie porównać ze sobą położenie stawów, wyznaczonych na podstawie pomiarów z kontrolera Kinect i z czujników inercyjnych, z tymi wyznaczonymi przez system śledzenia ruchu Vicon, współrzędne śledzonych markerów, reprezentujące wybrany staw, muszą być uśrednione, tak żeby na ich podstawie obliczyć współrzędne stawów analizowanej kończyny. Do dokładnego wyznaczenia pojedynczego punktu reprezentującego staw, pomocne jest oszacowanie położenia trzech markerów. Dzięki zastosowaniu uzupełniających markerów do układu rekomendowanego przez firmę Vicon, udało się znacznie ograniczyć liczbę pomiarów, na których były widoczne mniej niż trzy markery, przez co precyzja śledzenia ruchu kończyn była większa. System śledzenia ruchu Vicon, w wykorzystanej konfiguracji, dostarczał pomiary co 10 ms – działał z częstotliwością 100 Hz.

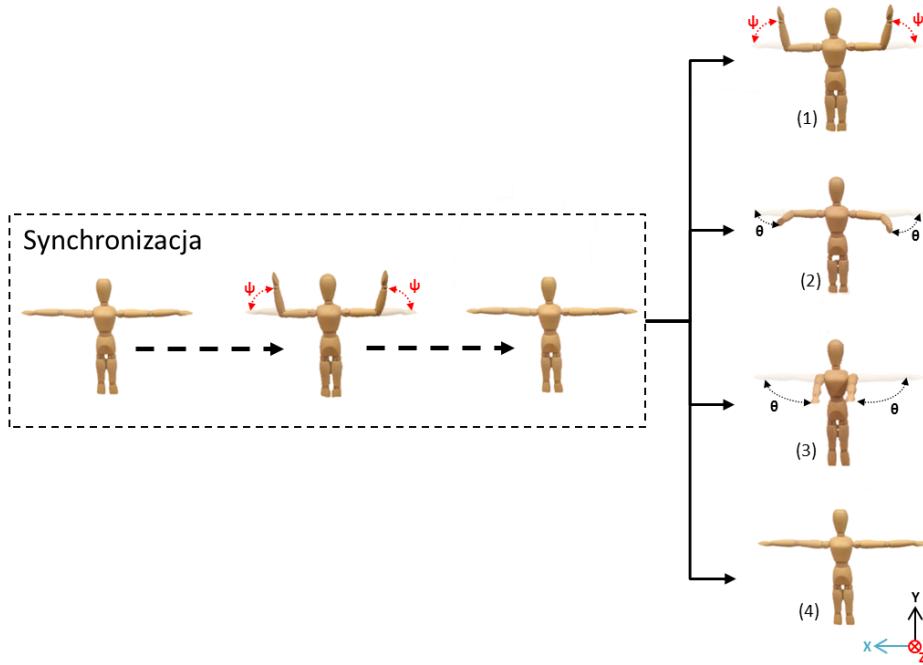
Oprócz markerów niezbędnych dla poprawnego działania systemu Vicon na przedramieniu i ramieniu, w połowie ich długości, zostały umieszczone czujniki inercyjne. Czujniki te, aby możliwie wiernie rejestrować ruchy kończyn zostały przymocowane do skóry za pomocą taśmy klejącej oraz dodatkowo docisnięte elastycznymi opaskami. Dzięki temu były one w stanie wykonywać obroty wraz z kończynami minimalizując względne przesunięcia. Zostały one umieszczone na wierzchniej stronie ręki, tam gdzie jest relatywnie cienka warstwa mięśni i ścięgien tak jak na rysunku 4.3. Dzięki temu czujniki w większym stopniu reagowały na obroty kości i były mniej wrażliwe na pracę otaczających je tkanek (wpływ ruchu między innymi skóry, ścięgien i mięśni na badanie ruchu kości można znaleźć w [92, 98]).



Rysunek 4.3: Moduł inercyjny umieszczony na przedramieniu autora(źródło: badania własne)

4.2. Badanie ruchu

Każdy z ruchów ręki był zrealizowany przez trzech użytkowników i wykonywany w dwóch seriach po pięć powtórzeń. Każda seria stanowiła osobną sesję śledzenia co oznacza, że każdy ruch został zarejestrowany w sześciu sesjach (trzy osoby po dwie sesje śledzenia na każde ćwiczenie). Podzielenie całego ćwiczenia na serie z mniejszą liczbą powtórzeń związane było z efektem zmęczenia aktorów, jakie było widoczne wraz z kolejnymi powtórzeniami. Ćwiczenia jakie należało wykonać zostały dobrane w taki sposób, aby sprawdzić opisaną metodę zarówno w sytuacjach, w których kontroler Kinect i czujniki inercyjne działają w uprzywilejowanym dla nich zakresie pracy jak i w takich, w których jedno z urządzeń posiada istotne ograniczenia w działaniu. Wszystkie sekwencje rozpoczynały się od pozycji wyjściowej z rękoma rozłożonymi na boki w tak zwanej pozycji T (*ang. T-pose*). Następnie, wykonywano zgięcie rąk w łokciu ku górze (w płaszczyźnie XY) i ich wyprostowanie w celu zsynchronizowania pomiarów Kinecta i czujników inercyjnych w czasie. Po sekwencji synchronizującej następowały właściwe ćwiczenia, w których każde powtόrzenie zaczynało i kończyło się pozycją T (*ang. T-pose*). Wykonywane ćwiczenia prezentuje rysunek 4.4.



Rysunek 4.4: Ćwiczenia wykonywane w ramach eksperymentu (źródło: badania własne)

Lista wykonywanych ćwiczeń obejmuje kolejno:

1. Zgięcie ręki w łokciu do góry (zgięcie w płaszczyźnie XY układu współrzędnych Kinecta);
2. Zgięcie ręki w łokciu do przodu (zgięcie w płaszczyźnie XZ układu współrzędnych Kinecta);
3. Wyciągnięcie wyprostowanych ramion do przodu (ruch w płaszczyźnie XZ układu współrzędnych Kinecta);
4. Utrzymanie wyprostowanych ramion (pozycja T) w bezruchu przez czas około 60s;

Ćwiczenie 1 wykonywane było w płaszczyźnie, w której czujniki inercyjne są w stanie dokonywać dokładnych i stabilnych w czasie pomiarów, natomiast dla Kinecta nie następuje okluzja stawów. Ćwiczenia 2 i 3 wymagają wykonania ruchu w płaszczyźnie, w której pomiary czujników inercyjnych są błędne i nie mogą być traktowane jako wiarygodne. Dodatkowo, w trakcie wykonywania ćwiczenia, kontroler Kinect traci możliwość obserwacji stawów ręki z powodu ich przesłaniania się. Ostatnie ćwiczenie z kolei sprawdza stabilność pomiaru w sytuacji kiedy nie jest wykonywany żaden ruch. Pomiar taki obarczony jest dryfem występującym w danych uzyskanych za pomocą akcelerometru i żyroskopu.

Wszystkie pomiary zostały porównane z pomiarami referencyjnymi zarejestrowanymi w systemie Vicon w celu określenia dokładności wyznaczania pozycji stawów i kąta zgięcia w łokciu. Układy współrzędnych systemu śledzenia Vicon i systemu opartego na autorskiej metodzie śledzenia ruchu zostały ujednolicone tak aby można było łatwo określić precyzję szacowania położenia stawów. W trakcie kalibracji systemu Vicon, początek jego układu współrzędnych (punkt $(0, 0, 0)$) został określony na podłodze, dokładnie pod kontrolerem Kinect, który został także oznaczony markerami w celu ustalenia jego położenia. Dzięki temu możliwe było skorelowanie układów współrzędnych Kinecta i systemu Vicon, a w związku z tym także porównanie ze sobą oszacowań uzyskanych za pomocą obu systemów.

Błąd estymacji pozycji stawów, w pojedynczej sesji wykonywanego ruchu, został zdefiniowany jako pierwiastek kwadratowy z błędu średniokwadratowego (RMSE – *ang. root mean square error*) odległości euklidesowej ($d_e()$) pomiędzy referencyjną pozycją danego stawu (j) zmierzoną w systemie Vicon P_j^V , a pozycją oszacowaną (P_j^F) i jest wyrażony wzorem 4.1

$$RMSE_j^F = \sqrt{\frac{1}{n} \sum_{i=1}^n d_e(P_{j,i}^V, P_{j,i}^F)^2} \quad (4.1)$$

Następnie została obliczona średnia arytmetyczna z wyznaczonych błędów średniokwadratowych każdej z sześciu sesji śledzenia według wzoru 4.2

$$\overline{RMSE_j^F} = \frac{\sum_{s=1}^6 RMSE_{j,s}^F}{6} \quad (4.2)$$

w celu wyznaczenia dokładności szacowania pozycji stawów w każdym z wykonanych ćwiczeń. Podobną metodologię określania błędu oszacowania badanej metody zastosowali w swoich pracach między innymi Kassim i in. [14] oraz Tadano i in. [115]. W analogiczny sposób (wzór 4.1 w którym pójcja stawu została zastapiona kątem zgięcia kończyn w stawie) został oszacowany błąd wyznaczania kąta zgięcia stawu łokciowego β .

Błędy szacowania pozycji stawów oraz kąta zgięcia stawu łokciowego zostały wyznaczone dla opisywanej w niniejszej dysertacji metody łączenia danych, opartej na obrotach kości $(P_j^{F,O}, \beta^{F,O})$ oraz dla analogicznych oszacowań uzyskanych za pomocą metody o najniższym opisanym w literaturze błędzie szacowania położenia stawów – metodzie Kalkbrennera i in. [57]. Metoda Kalkbrennera wykorzystuje dane o pozycjach poszczególnych stawów, zarejestrowanych niezależnie przez kontroler Kinect i czujniki inercyjne, które są łączone celem zmniejszenia niedokładności oszacowania $(P_j^{F,P}, \beta^{F,P})$.

Każda z sesji śledzenia konkretnego ruchu składała się z 5 nieprzerywanych powtórzeń tego samego ruchu. Dla każdej z takich sesji zostały wyliczone wartości błędu średniokwadratowego (wzór 4.1) każdego z wyżej wymienionych oszacowań (pozycje stawów łokciowego ($P_E^F, j = E$ ang. Elbow) i nadgarstkowego ($P_W^F, j = W$ ang. Wrist) oraz kąt zgięcia łokcia β^F). Całkowity błąd szacowania metody został określony jako średnia arytmetyczna błędów średniokwadratowych (\overline{RMSE}) ze wszystkich sześciu sesji śledzenia ruchu (indeks sesji śledzenia – s) dla danego stawu (łokcia E lub nadgarstka W) lub kąta β (wz. 4.2).

Wyniki zostały zebrane w tabelach 4.1, 4.2, 4.3 i 4.4. Porównanie pomiędzy wynikami uzyskanymi za pomocą metody Kalkbrennera i metody łączenia danych opartej na obrotach kości zostało wyznaczone jako stosunek procentowy (r) różnicy pomiędzy uśrednionym błędem średniokwadratowym dla metody Kalkbrennera (\overline{RMSE}_j^P) a uśrednionym błędem średniokwadratowym dla metody opracowanej przez autora niniejszej dysertacji (\overline{RMSE}_j^O) do \overline{RMSE}_j^P , gdzie j oznacza analizowany staw szkieletu ($j = E, W$) lub kąt β (wz. 4.3).

$$r = \frac{\overline{RMSE}_j^P - \overline{RMSE}_j^O}{\overline{RMSE}_j^P} * 100\% \quad (4.3)$$

Dodatnia wartość r oznacza, że uśredniony błąd średniokwadratowy wyników uzyskanych za pomocą metody opracowanej przez autora jest mniejszy niż uśredniony błąd średniokwadratowy wyników uzyskanych metodą Kalkbrennera, co oznacza, że nastąpiła poprawa oszacowań. W przypadku wartości ujemnej mamy do czynienia z pogorszeniem uzyskiwanych wyników. W obu przypadkach wartość r określa o ile procent zmieniły się wyniki. Porównanie uśrednionych błędów średniokwadratowych dla oszacowań:

- położenia stawu łokciowego,
- położenia stawu nadgarstkowego,
- kąta zgięcia stawu łokciowego (β),

wykonanych za pomocą obu metod, zostało zaprezentowane na wykresach znajdujących się w dalszej części niniejszego rozdziału.

Ćwiczenie 1 – Zgięcie ramienia w łokciu do góry

Pierwsze ćwiczenie, jakie zostało wykonane z użyciem omawianej metody, to zgięcie ręki w łokciu do góry (płaszczyzna XY układu odniesienia). Jest to ruch łatwy do śledzenia dla obu urządzeń pomiarowych ze względu na brak występowania okluzji stawów w płaszczyźnie ruchu prostopadłej do kierunku obserwacji Kinecta oraz ze względu na to, że nie odbywa się on wokół wektora grawitacji. Wyniki uzyskane w tym ćwiczeniu prezentuje tabela 4.1.

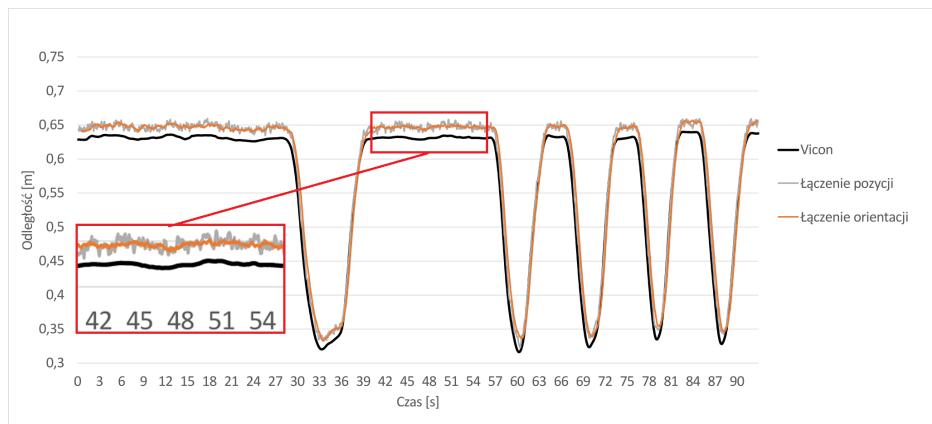
Tablica 4.1: Średni błąd szacowania \overline{RMSE} (wz. 4.3) dla ćwiczenia nr 1 (źródło: badania własne)

Użytkownik	Sesja śledzenia	Metoda autorska			Metoda Kalkbrennera		
		$RMSE_E^O$ [cm]	$RMSE_W^O$ [cm]	$RMSE_\beta^O$ [°]	$RMSE_E^P$ [cm]	$RMSE_W^P$ [cm]	$RMSE_\beta^P$ [°]
1	1	2.25	2.75	3.24	2.65	3.09	3.44
1	2	2.26	2.76	3.33	2.66	3.10	3.77
2	3	2.23	2.74	3.15	2.55	3.02	3.42
2	4	2.26	2.76	3.18	2.56	3.07	3.48
3	5	2.06	2.78	3.36	2.49	2.94	3.51
3	6	2.26	2.77	3.52	2.56	2.99	3.63
Średnia $RMSE$		2.22	2.76	3.30	2.58	3.04	3.54
Odchylenie		0.07	0.01	0.12	0.06	0.06	0.12

Ćwiczenie polegające na wielokrotnym zgięciu ręki w łokciu do góry, zgodnie z przewidywaniami, dla obu urządzeń pomiarowych nie było trudne. Przez cały czas trwania ćwiczenia stawy były poprawnie identyfikowane i w pełni śledzone przez kontroler Kinect. Także czujniki inercyjne były w stanie zarejestrować w pełni wykonywany ruch. Obie porównywane ze

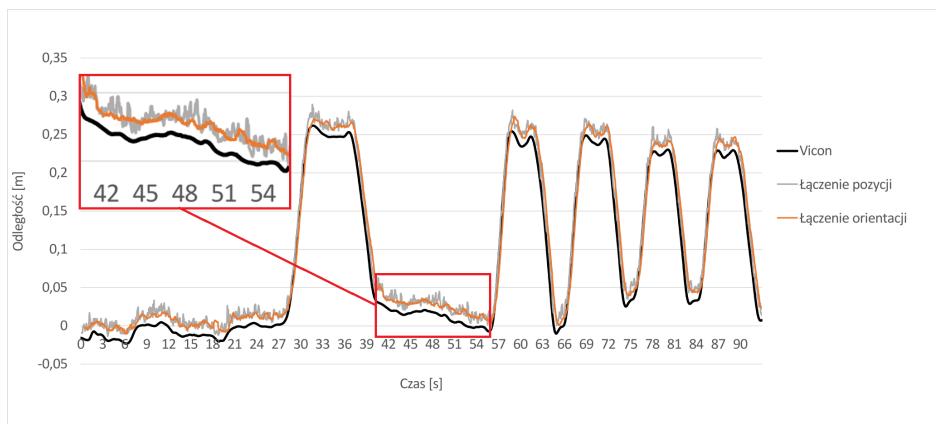
sobą metody uzyskiwały podobne wyniki, a zakres szacowanego przez nie ruchu był zbliżony do tego wykonywanego w rzeczywistości. Widoczne było natomiast zróżnicowanie w dopasowaniu szkieletu do ciała ze względu na przyjęty model długości kości. W metodzie Kalkbrennera miała miejsce częsta zmiana tego oszacowania, co wpływało na to, że śledzony staw potrafił zmienić swoje położenie podczas gdy postać pozostawała w bezruchu, a także potrafił znaleźć się poza konturem postaci. W przypadku metody autora niniejszej rozprawy, zauważalne były częste, jednak niewielkie, fluktuacje wartości oszacowanych pozycji stawów. Były one jednak na tyle niewielkie, że nie wpływały w istotny sposób na wartość błędu średniokwadratowego uzyskanego szacowania, w przeciwieństwie do wpływu na ten błąd zmiennej długości segmentów reprezentujących kości wykorzystywanych w metodzie Kalkbrennera. Wydaje się również, że to właśnie brak stałego modelu długości kości w metodzie Kalkbrennera ma największy wpływ na różnicę uzyskiwanych wyników w ćwiczeniu 1.

W omawianym ćwiczeniu ruch wykonywany był głównie przez przedramię, czyli istotne i zamierzone zmiany położenia były widoczne na stawie nadgarstkowym. Wykresy z rysunków 4.5, 4.6, 4.7 pokazują położenie stawu nadgarstkowego w każdej z trzech osi w trakcie wykonywania ruchu. Widać na nich wyraźną różnicę w stabilności pomiarów objawiającą się ciągłymi drganiemiami widocznymi na wykresie łączenia danych za pomocą metody Kalkbrennera.

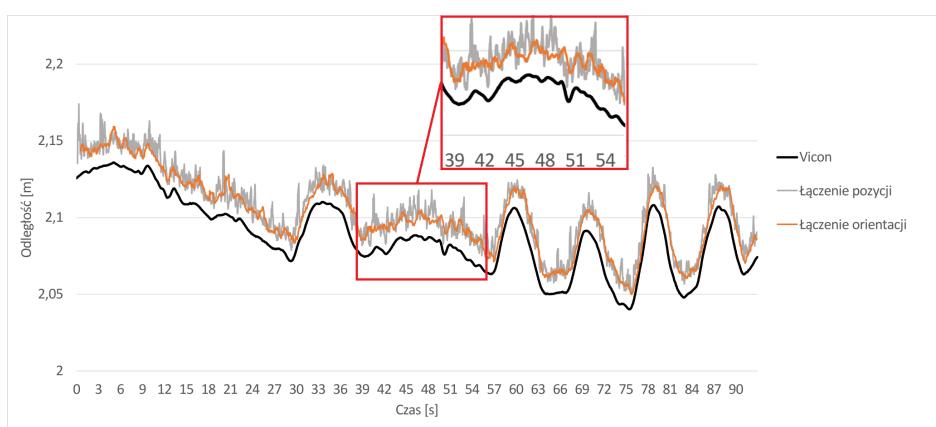


Rysunek 4.5: Wykres przedstawiający położenie stawu nadgarstkowego w osi X w ćwiczeniu 1 (źródło: badania własne)

Wykres na rysunku 4.8 przedstawia oszacowanie kąta zgięcia ręki w łokciu (β). Widać na nim, wyraźną różnicę w dokładności i stabilności szacowania wartości w trakcie wykonywania ruchu oraz w położeniach skrajnych (wyprostowana ręka lub zgięta). W położeniu skrajnym widoczne są fluktuacje

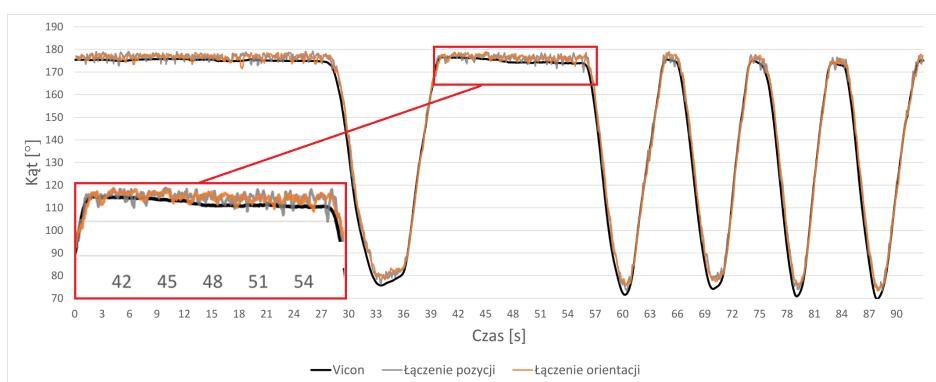


Rysunek 4.6: Wykres przedstawiający położenie stawu nadgarstkowego w osi Y w ćwiczeniu 1 (źródło: badania własne)



Rysunek 4.7: Wykres przedstawiający położenie stawu nadgarstkowego w osi Z w ćwiczeniu 1 (źródło: badania własne)

oszacowań zarówno w przypadku metody Kalkbrennera jak i metody autorskiej. Jednakże fluktuacje obecne w oszacowaniach metody, łączącej ze sobą obroty są, widocznie mniejsze niż te występujące w metodzie opartej o łączenie pozycji stawów.



Rysunek 4.8: Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 1 (źródło: badania własne)

Rysunki 4.9a oraz 4.9b wizualizują ruch ręki podczas tego ćwiczenia. W przypadku obu metod łączenia danych widać poprawę pozycjonowania stawów względem danych otrzymanych bezpośrednio z urządzeń pomiarowych.



(a) Wizualizacja ruchu bezpośrednio
z urządzeń pomiarowych

(b) Wizualizacja ruchu po złączeniu da-
nych z urządzeń pomiarowych

Rysunek 4.9: Wizualizacja ruchu ręki w ćwiczeniu 1. Kolory: czarny – Vicon, niebieski – czujniki inercyjne, zielony – Kinect, szary – metoda Kalkbrennera, pomarańczowy – metoda autorska (źródło: badania własne)

Ćwiczenia 2 i 3 – Zgięcie ramienia w łokciu do przodu oraz wyciągnięcie wyprostowanych ramion do przodu

Ćwiczenia 2 i 3 przedstawiają ruchy o zwiększym stopniu trudności śledzenia dla obu urządzeń pomiarowych. W tym przypadku następuje stopniowe przysłanianie jednego ze stawów co oznacza utratę możliwości dokładnego śledzenia ruchu przez kontroler Kinect. Dodatkowo ruch odbywał się w płaszczyźnie trudnej z punktu widzenia czujników inercyjnych, ponieważ obrót dokonywany był wokół osi równoległej do wektora grawitacji. Tabela 4.2 przedstawia średnie dokładności wykonywanego ruchu samego przedramienia, natomiast tabela 4.3 przedstawia wyniki dla ruchu całej wyprostowanej ręki.

W obu ćwiczeniach, w trakcie wykonywania ruchu, następował moment, w którym jeden lub dwa stawy stawały się niewidoczne dla kontrolera Kinect, a co za tym idzie, obie metody musiały bazować tylko na jednym źródle danych. Metoda opracowana przez autora niniejszej pracy nie wykorzystywała w trakcie łączenia danych pomiarów uzyskanych z niepewnego źródła i stopniowo wstrzymywała aktualizację estymowanego położenia stawu. Natomiast metoda Kalkbrennera wykorzystywała w dalszym ciągu dane z kontrolera Kinect zmieniając jedynie ich stopień istotności (rys. 4.10).

Tablica 4.2: Średni błąd szacowania \overline{RMSE} (wz. 4.3) dla ćwiczenia nr 2 (źródło: badania własne)

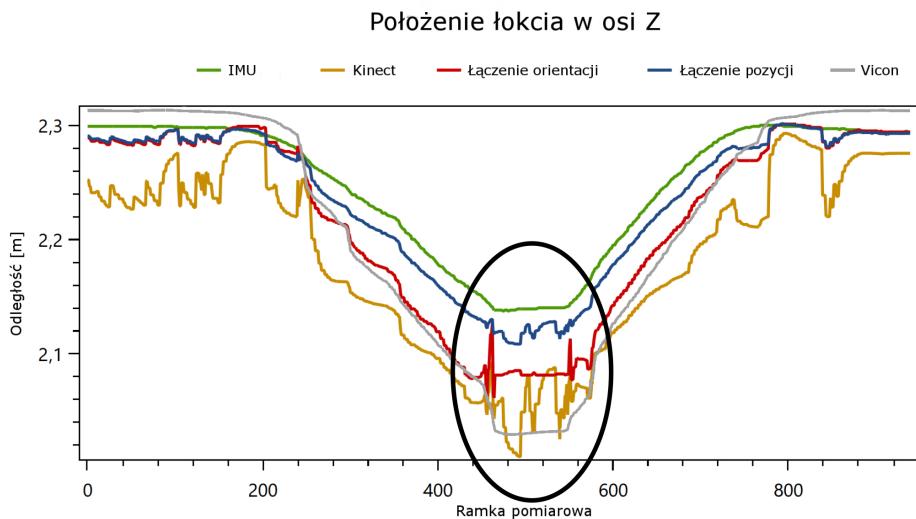
Użytkownik	Sesja śledzenia	Metoda autorska			Metoda Kalkbrennera		
		$RMSE_E^O$ [cm]	$RMSE_W^O$ [cm]	$RMSE_\beta^O$ [°]	$RMSE_E^P$ [cm]	$RMSE_W^P$ [cm]	$RMSE_\beta^P$ [°]
1	1	2.67	3.02	14.46	3.08	3.48	15.01
1	2	2.50	2.85	12.72	3.07	3.51	13.07
2	3	2.36	2.74	12.78	2.86	3.24	14.50
2	4	2.41	2.93	13.38	3.14	3.53	15.28
3	5	2.45	2.88	13.41	3.16	3.47	14.61
3	6	2.51	2.79	13.34	2.95	3.50	14.50
Średnia $RMSE$		2.48	2.87	13.35	3.06	3.46	14.50
Odchylenie		0.10	0.09	0.62	0.11	0.10	0.76

Tablica 4.3: Średni błąd szacowania \overline{RMSE} (wz. 4.3) dla ćwiczenia nr 3 (źródło: badania własne)

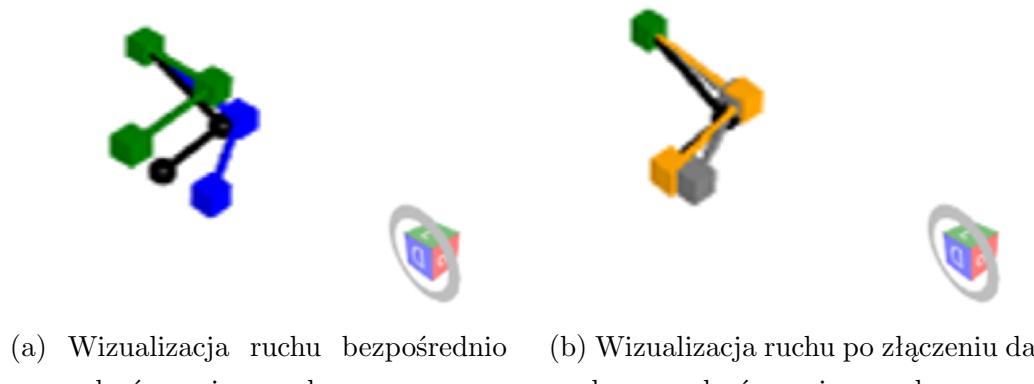
Użytkownik	Sesja śledzenia	Metoda autorska			Metoda Kalkbrennera		
		$RMSE_E^O$ [cm]	$RMSE_W^O$ [cm]	$RMSE_\beta^O$ [°]	$RMSE_E^P$ [cm]	$RMSE_W^P$ [cm]	$RMSE_\beta^P$ [°]
1	1	2.60	2.97	13.32	3.13	3.56	15.75
1	2	2.24	2.53	12.37	2.99	3.33	13.12
2	3	2.76	3.14	13.18	2.70	3.14	12.54
2	4	2.73	3.03	14.78	3.17	3.51	16.41
3	5	2.22	2.58	16.77	2.74	3.07	15.75
3	6	2.20	2.69	13.41	2.75	2.90	14.66
Średnia $RMSE$		2.46	2.82	13.97	2.91	3.25	14.70
Odchylenie		0.25	0.23	1.44	0.19	0.24	1.43

Ze względu na to, że kontroler Kinect, pomimo utraty możliwości śledzenia wybranego stawu, wciąż estymuje jego położenie, zasadne jest sprawdzenie czy fuzja pomiarów dostarczonych przez kontroler Kinect z pomiarami urządzeń inercyjnych nie sprawi, że ostateczna estymacja położenia stawu stanie się zbyt niedokładna. Metoda Kalkbrennera co prawda dokonuje takiego sprawdzenia w trakcie stosowania filtra Kalmana, jednak widoczne było podążanie szacunkowych wartości położenia stawu za skokami danych z Kinecta. Oznacza to, że wpływ pomiarów kontrolera Kinect na ostateczną estymację jest znaczący i zastosowana w metodzie Kalkbrennera zmiana stopnia ich istotności, w procesie łączenia danych, jest niewystarczająca na ograniczenie negatywnego wpływu niedokładności kontrolera Kinect.

Rysunki 4.11a i 4.12a przedstawiają wizualizację położenia stawów ręki, których położenie wyestymowano na podstawie pomiarów uzyskanych bezpośrednio z urządzeń pomiarowych. Widoczne na nich jest niedokładne szacowanie położenia stawu łokciowego przez kontroler Kinect w momencie jego przysłonięcia, a z tym związane jest także niedokładne szacowanie położenia nadgarstka, pomimo jego pełnej widoczności. Widoczne jest także błędne szacowanie obrotu zarówno kości przedramienia jak i ramienia wokół osi grawitacji przez urządzenia inercyjne. Wizualizuje to w pełni ograniczenia w działaniu obu urządzeń pomiarowych, które zostały opisane w rozdziale 2.3. W przypadku wizualizacji pozycjonowania stawów



Rysunek 4.10: Wykres przedstawiający podążanie wyznaczania pozycji stawu za pomiarami Kinecta w ćwiczeniu 3 (źródło: badania własne)



Rysunek 4.11: Wizualizacja ruchu ręki w ćwiczeniu 2. Kolory: czarny – Vicon, niebieski – czujniki inercyjne, zielony – Kinect, szary – metoda Kalkbrennera, pomarańczowy – metoda autorska (źródło: badania własne)

łokciowego i nadgarstkowego, za pomocą metody Kalkbrennera oraz metody autora niniejszej rozprawy (rys. 4.11b oraz 4.12b), w obu przypadkach widoczne jest zmniejszenie wpływu występowania okluzji stawów jak i błędów pomiaru obrotu wokół osi grawitacji na ostateczne szacowanie położenia stawów. Widać równocześnie wierniejsze odwzorowanie ruchu w przypadku metody autora.

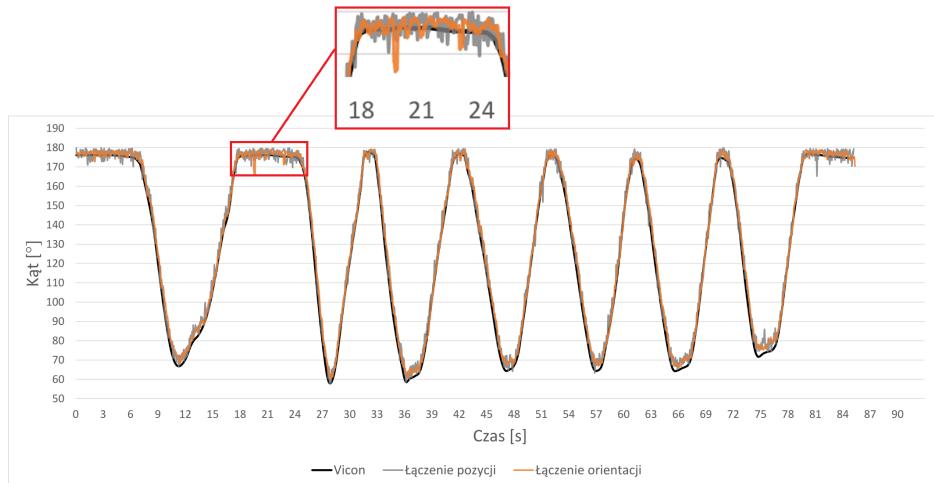
Wykresy z rysunków 4.13 i 4.14 przedstawiają oszacowania wartości kąta β w trakcie wykonywania ćwiczeń 2 i 3. Porównując ze sobą oba te wykresy, można zauważać znaczząco większe fluktuacje uzyskanych wyników szacowania wartości kąta β w ćwiczeniu 3 niż w ćwiczeniu 2. W obu tych ćwiczeniach



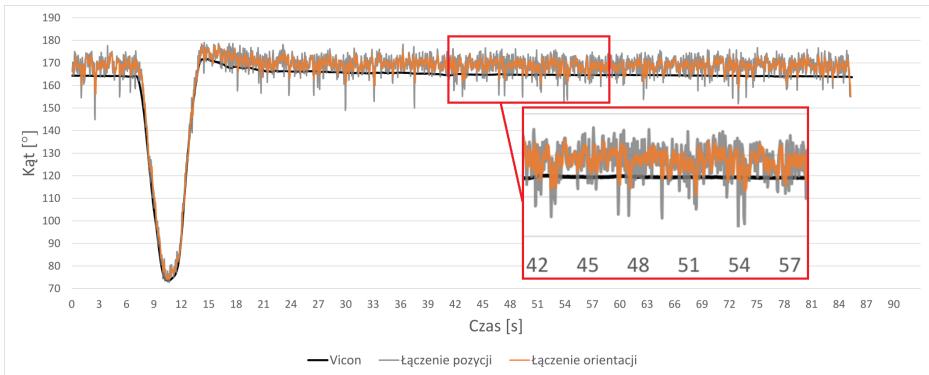
(a) Wizualizacja ruchu bezpośrednio
z urządzeń pomiarowych (b) Wizualizacja ruchu po złączeniu da-
nych z urządzeń pomiarowych

Rysunek 4.12: Wizualizacja ruchu ręki w ćwiczeniu 3. Kolory: czarny – Vicon, niebieski – czujniki inercyjne, zielony – Kinect, szary – metoda Kalkbrennera, pomarańczowy – metoda autorska (źródło: badania własne)

dochodziło do czasowego przysłonięcia stawów ręki, w ćwiczeniu 2 tylko łokcia, natomiast w ćwiczeniu 3 zarówno łokcia jak i barku. Skutkowało to pogorszeniem jakości pomiarów Kinecta co miało bezpośredni wpływ na działanie obu metod łączenia danych. W obu przypadkach można zauważyć, że fluktuacje wyników uzyskanych za pomocą metody łączącej informację o obrotach (metody autora pracy) były mniejsze od tych uzyskanych za pomocą metody opartej o łączenie pozycji stawów.



Rysunek 4.13: Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 2 (źródło: badania własne)



Rysunek 4.14: Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 3 (źródło: badania własne)

Ćwiczenie 4 – Utrzymanie wyprostowanych ramion

Celem ostatniego ćwiczenia było sprawdzenie stabilności pomiarów w przypadku długiego braku ruchu. W tym wypadku urządzeniem, którego dane pomiarowe mogą szczególnie ulec pogorszeniu w trakcie działania jest urządzenie zbudowane z czujników inercyjnych. Tabela 4.4 zawiera zestawienie dokładności pomiarów dla tego eksperymentu.

Tablica 4.4: Średni błąd szacowania \overline{RMSE} (wz. 4.3) dla ćwiczenia nr 4 (źródło: badania własne)

Użytkownik	Sesja śledzenia	Metoda autorska			Metoda Kalkbrennera		
		$RMSE_E^O$ [cm]	$RMSE_W^O$ [cm]	$RMSE_\beta^O$ [°]	$RMSE_E^P$ [cm]	$RMSE_W^P$ [cm]	$RMSE_\beta^P$ [°]
1	1	2.12	2.39	4.55	2.40	2.73	4.86
1	2	2.26	2.62	3.86	2.44	2.76	5.10
2	3	2.16	2.50	3.70	2.58	2.85	4.24
2	4	2.30	2.58	4.23	2.58	2.85	4.61
3	5	2.58	2.47	4.32	2.47	2.79	4.85
3	6	2.18	2.48	4.02	2.36	2.75	4.27
Średnia $RMSE$		2.27	2.51	4.11	2.47	2.79	4.65
Odchylenie		0.15	0.08	0.28	0.08	0.05	0.32

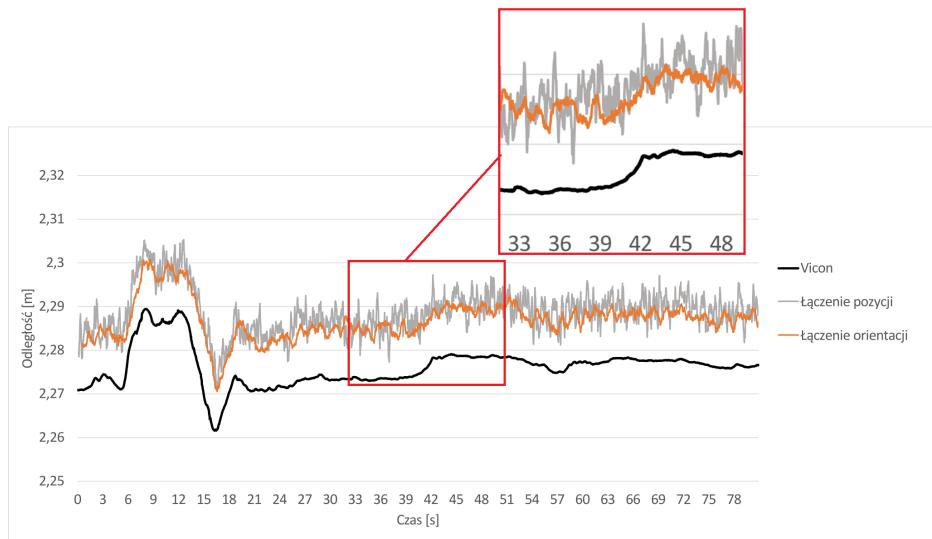
Podobnie jak w ćwiczeniu 1, tak i w ćwiczeniu 4, oba urządzenia były w stanie śledzić ruchy wykonywane przez śledzonego aktora przez cały czas trwania eksperymentu. Jednak, podobnie jak we wspomnianym ćwiczeniu, zmianie ulegały długości segmentów kości i miało to przełożenie na działanie metody Kalkbrennera. Wizualizacje łańcuchów kinematycznych przedstawione na rysunku 4.15 pokazują, że Kinect, mimo pełnej widoczności całej ręki, miał trudność z poprawnym określeniem jej pozycji w osi pionowej. Mimo to obie metody były w stanie wyznaczyć pozycje stawów z lepszą dokładnością niż każde z urządzeń pomiarowych osobno.



(a) Wizualizacja ruchu bezpośrednio z urządzeń pomiarowych (b) Wizualizacja ruchu po złączeniu danych z urządzeń pomiarowych

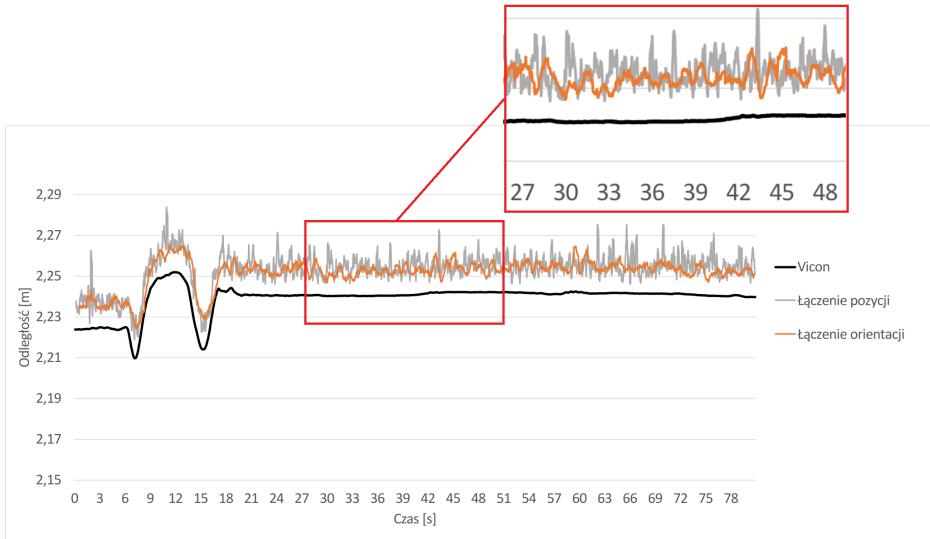
Rysunek 4.15: Wizualizacja ruchu ręki w ćwiczeniu 4. Kolory: czarny – Vicon, niebieski – czujniki inercyjne, zielony – Kinect, szary – metoda Kalkbrennera, pomarańczowy – metoda autorska (źródło: badania własne)

Wykresy estymacji położenia stawu łokciowego (rys. 4.16) oraz nadgarstekowego (rys. 4.17) pozwalają zaobserwować wpływ stabilności pomiarów jednego stawu na drugi. Jest to szczególnie widoczne w wykresie metody Kalkbrennera, gdzie amplituda drgań pozycji nadgarstka jest widocznie większa niż w przypadku drgań łokcia.



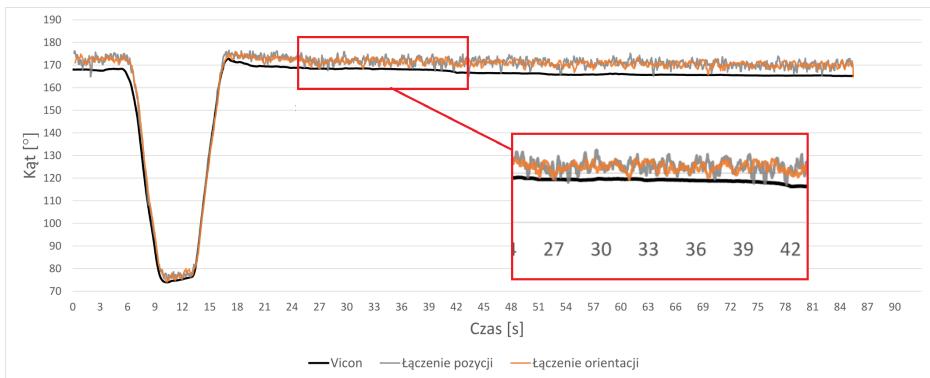
Rysunek 4.16: Wykres położenia stawu łokciowego w osi Z w ćwiczeniu 4 (źródło: badania własne)

Wykres estymacji kąta β w trakcie ćwiczenia 4, zaprezentowany na rysunku 4.18, pozwala zauważać podobne zachowanie estymowanych wartości kąta do tego zaobserwowanego w ćwiczeniu 1. W trakcie wykonywania ruchu (w tym ćwiczeniu zamierzona zmiana kąta w łokciu następowała tylko podczas początkowej synchronizacji) widać estymację kąta za pomocą obu



Rysunek 4.17: Wykres położenia stawu nadgarstkowego w osi Z w ćwiczeniu 4 (źródło: badania własne)

metod zbliżoną do danych referencyjnych. Różnica jest natomiast dobrze widoczna podczas braku ruchu. Wówczas widoczne są wyraźne fluktuacje w oszacowaniach wartości kąta, które wpływają na wartość uzyskanego błędu średniokadratowego.



Rysunek 4.18: Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 4 (źródło: badania własne)

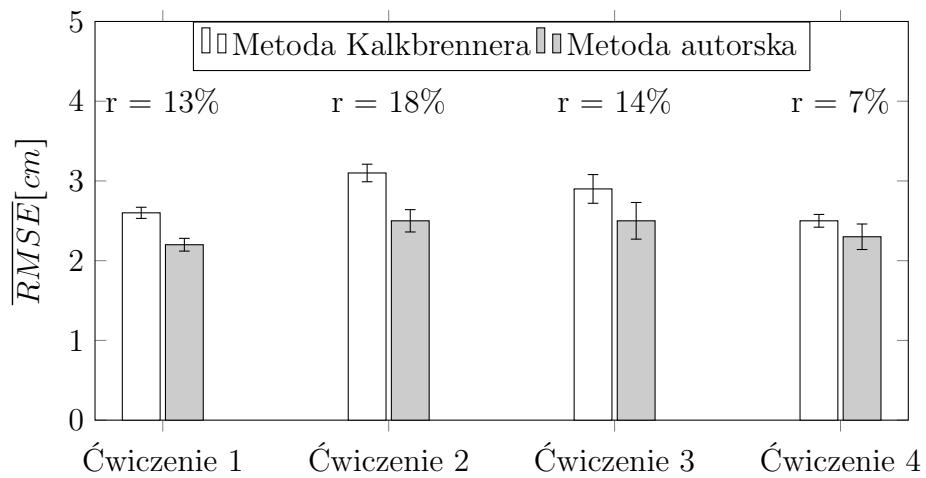
4.3. Podsumowanie

Ćwiczenia wykonywane w ramach eksperymentów miały za zadanie zweryfikować dokładność metody autora niniejszej pracy i porównać ją z metodą Kalkbrennera w sytuacji, gdy pomiary z czujników przestają być wiarygodne

oraz kiedy są one degradowane w czasie. Dla kontrastu weryfikacja dokładności porównywanych metod, dokonana została także na bazie ćwiczenia, które dla obu urządzeń pomiarowych może być nazywane neutralnym (ćwiczenie 1).

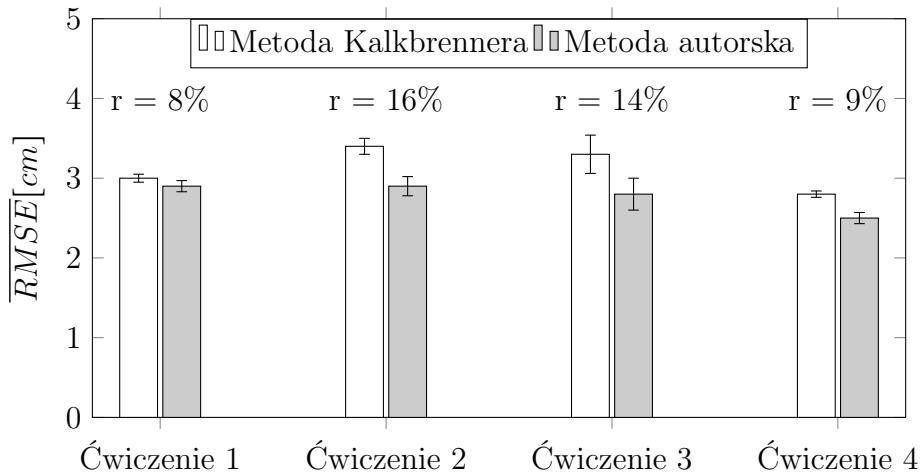
Porównane ze sobą zostały wartości trzech parametrów wybranego układu szkieletowego, wyznaczone przez dwie metody: metodę Kalkbrennera i metodę łączącą dane opartą o obroty kości. Porównanymi parametrami kończyny górnej były: położenie stawu łokciowego, położenie stawu nadgarstkowego oraz kąt zgięcia ręki w łokciu (β). Błąd oszacowania pozycji stawów w przestrzeni (wz. 4.2) może być wykorzystany jako wskaźnik użyteczności danej metody przy wykrywaniu ewentualnej interakcji pomiędzy estymowanym modelem a otoczeniem. Jako przykład można podać środowisko wirtualne, w którym znajdują się obiekty, którymi można manipulować. Każdy taki obiekt ma swoje ścisłe określone położenie w przestrzeni, a estymacja pozycji stawów jest niezbędna żeby określić czy doszło do interakcji z takim obiektem czy nie.

Trzeci parametr układu szkieletowego ręki – kąt zgięcia ręki w łokciu (β) – jest uniezależniony od przyjętych długości kości i ewentualnych błędów w pomiarach tych długości. Określa on wzajemną orientację, pomiędzy kością przedramienia a ramieniem, w zakresie stopnia swobody stawu łokciowego. Pozwala też na zweryfikowanie śledzenia zakresu ruchu stawu łokciowego na potrzeby ruchu, co może być przydatne w procesie rehabilitacji ruchowej.



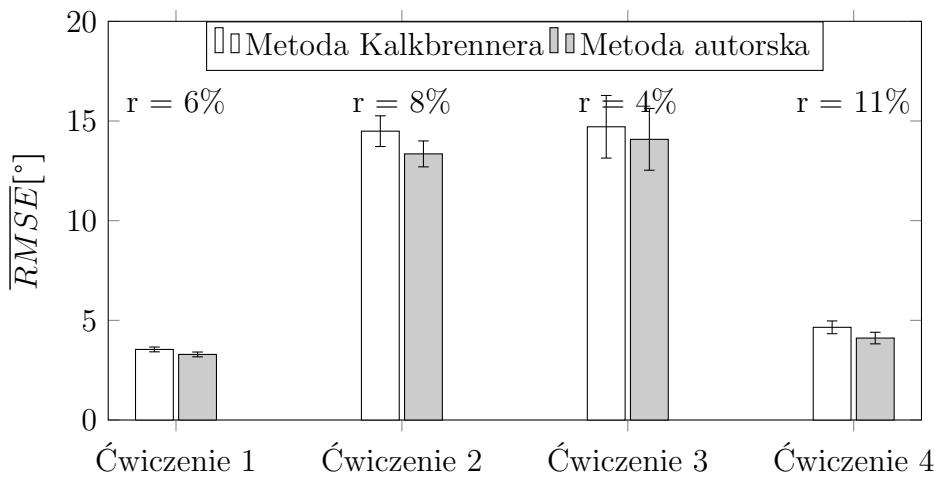
Rysunek 4.19: Średni błąd średniokwadratowy wyznaczania pozycji łokcia (źródło: badania własne)

Wykresy z rysunków 4.19, 4.20 oraz 4.21 pokazują podsumowanie porównania uśrednionych wartości błędu uzyskanych oszacowań (\overline{RMSE}) za pomocą metody Kalkbrennera oraz metody autora tej pracy. Widać na nich poprawę w wyznaczaniu położenia stawu łokciowego do 18%, nadgarstkowego



Rysunek 4.20: Średni błąd średniokwadratowy wyznaczania pozycji nadgarstka (źródło: badania własne)

do 16% oraz zdolności szacowania kąta zgięcia ręki w łokciu do 11%. Z uzytkowanych rezultatów eksperymentów wynika większa poprawa w wyznaczaniu pozycji łokcia niż w przypadku nadgarstka. Spowodowane jest to kumulowaniem się błędów własnych estymacji położenia nadgarstka, jak i błędów estymacji położenia łokcia. Jest to cecha charakterystyczna stosowania hierarchicznego modelu ciała człowieka. Dla przyjętej hierarchicznej reprezentacji układu kostnego postaci, czym badany staw jest bardziej odległy od punktu początkowego (korzenia) tym dokładność estymacji jego położenia jest mniejsza i mniejsze są możliwości jej poprawy.



Rysunek 4.21: Średni błąd średniokwadratowy wyznaczania kąta zgięcia ręki w łokciu β (źródło: badania własne)

Również na wykresie z rysunku 4.21, reprezentującym wyniki szacowania kąta zgięcia ręki w łokciu za pomocą obu metod, widać zmniejszenie

błędu wyników uzyskiwanych przez metodę opisywaną w niniejszej dysertacji. Sugeruje to, że łącząc ze sobą informacje o orientacji kości, możliwe jest dokładniejsze odtworzenie pozy, w jakiej znajduje się śledzona postać, niż ma to miejsce w przypadku bezpośredniego łączenia informacji o pozycjach stawów. Precyzyjne oszacowanie wzajemnej orientacji pomiędzy poszczególnymi kośćmi, z jednoczesnym wykorzystaniem modelu szkieletowego o stałych wartościach długości poszczególnych segmentów, ma bezpośrednie przełożenie na wielkość uzyskiwanego błędu. Można przyjąć, że kombinacja tych dwóch cech: łączenie ze sobą informacji o obrotach kości oraz niezmiennej, oszacowanej podczas kalibracji metody, model długości kości, mają decydujący wpływ na błąd szacowania pozycji stawów. Jednocześnie, poprawienie szacowania wartości kąta zgięcia ręki wpływa na zmniejszenie ilości występujących drgań poszczególnych segmentów modelu szkieletowego wyznaczonego przez metodę autora tej pracy.

Rozdział 5

Podsumowanie i wnioski

Zaprezentowana w pracy metoda zawiera nowe podejście do łączenia ze sobą danych z czujników inercyjnych i opisu modelu szkieletowego udostępnianego przez kontroler Kinect. W wyniku hybrydowego łączenia danych o orientacji kości układu szkieletowego, poprawiona została dokładność śledzenia ruchu stawów oraz możliwa była skuteczna aproksymacja informacji o pozycji stawów szkieletu, których nie można było precyzyjnie wyznaczyć bezpośrednio z wykorzystanych urządzeń pomiarowych. Względem metod opisanych w dostępnej i znanej autorowi literaturze unikalnymi elementami zaproponowanej metody są:

- łączenie informacji uzyskanych na podstawie danych z obu czujników, na etapie szacowania orientacji kości i szacowanie pozycji stawów na podstawie wypadkowej wartości ich orientacji,
- kompleksowa kompensacja mierzalnych błędów pomiarowych obu urządzeń,
- zróżnicowanie sposobu łączenia ze sobą danych w zależności od jakości uzyskanych sygnałów.

Zgodnie z przyjętą tezą pracy mówiącą, że

zastosowanie autorskiej, hybrydowej metody śledzenia ruchu kończyn człowieka, łączącej dane pochodzące z sensora głębi i sensorów inercyjnych, uwzględniającej kontekstowe charakterystyki pracy urządzeń, pozwala na bardziej precyzyjne śledzenie ruchu

oraz celami badań, metoda opisana w niniejszej pracy, łącząca ze sobą sygnały z ogólnodostępnych urządzeń pomiarowych, jakimi są kontroler Kinect i zintegrowane moduły inercyjne pokazuje, że odejście od łączenia informacji dotyczących położenia wybranych stawów na rzecz połączenia informacji o tym jak obrócone są związane z nimi kości, zmniejsza błąd oszacowania położenia tychże stawów. Spowodowane jest to tym, że łączenie danych o pozycjach śledzonych stawów uwypukla błędy związane z określeniem długości

poszczególnych kości modeli szkieletowych postaci wykorzystywanych przez oba te urządzenia. Jednocześnie komplementarne uwzględnianie sygnałów z obu urządzeń niweluje wady sygnałów każdego z kontrolerów. W przypadku modułów inercyjnych umieszczonych na ciele, informacja o ich orientacji przestrzennej (a co za tym idzie orientacji przestrzennej kości, do której są przyczepione) jest niejako domyślna i wymaga mniej złożonego przetwarzania danych. Wyznaczenie położenia stawów powiązanych z daną kośćią wymaga, oprócz wyznaczenia orientacji kości, dokładnego manualnego pomiaru jej długości tak, żeby wykorzystując przekształcenia geometryczne można było podać pozycję stawów. W związku z tym błąd szacowania położenia stawu wynika z niedokładności określania orientacji kości jak i błędów pomiaru długości kości. W przypadku kontrolera Kinect, model szkieletowy śledzonej postaci zawiera immanentnie informacje o długościach poszczególnych segmentów, jednak są to informacje szacowane na bieżąco i długość tej samej kości w tym modelu może się różnić nawet o kilka centymetrów pomiędzy następującymi po sobie pomiarami. Jednocześnie, jeśli w trakcie śledzenia ruchu, za pomocą Kinecta nie występuje okluzja, wówczas oszacowania orientacji poszczególnych kości są bardziej stabilne niż pozycji stawów. W związku z powyższym, łączenie sygnałów z dwóch urządzeń pomiarowych: Kinecta i modułów inercyjnych, wykorzystując informację o obrotach, niweluje istotne źródło błędu.

Uwzględnienie charakterystyk obu urządzeń pomiarowych oraz kompleksowa kompensacja mierzalnych błędów, występujących w ich pomiarach, jest także elementem wyróżniającym proponowaną metodę na tle innych, opisanych w literaturze i pozwala istotnie poprawić dane uzyskane z każdego z urządzeń przed ich złączeniem. Istotne jest także to, że wiele z tych charakterystyk jest związanych ściśle z kontekstem wykonywanego przez użytkownika ruchu. Na przykład, charakterystyczna zmienność błędu szacowania odległości między kamerą a użytkownikiem, w zależności gdzie się on znajduje, czy też to jak kąt obrotu sylwetki użytkownika względem płaszczyzny obserwacji Kinecta wpływa na stabilność jego pomiarów. Celem prowadzonych badań było opracowanie hybrydowej metody śledzenia ruchu kończyn człowieka, która, uwzględniając powyższe aspekty, w istotny sposób zmniejsza błąd oszacowania pozycji śledzonych stawów modelu szkieletowego człowieka. Zaproponowana metoda, przetestowana na przykładzie kończyn górnych, pozwoliła zmniejszyć błąd pozycjonowania stawów średnio do 18% dla stawu łokciowego i do 16% dla stawu nadgarstkowego względem wyników

uzyskiwanych za pomocą metody o nominalnie największej deklarowanej dokładności, oraz do 11% mniejszy błąd szacowania kąta zgięcia ręki w łokciu (β). Uzyskane wyniki charakteryzujące się mniejszym błędem szacowania pozycji stawów, niż metody opisane w literaturze, udowadniają stawianą tezę, że **zastosowanie autorskiej, hybrydowej metody śledzenia ruchu kończyn człowieka, łączącej dane pochodzące z sensora głębi i sensorów inercyjnych, uwzględniającej kontekstowe charakterystyki pracy urządzeń, pozwala na bardziej precyzyjne śledzenie ich ruchu.**

Opisywana metoda może być w dalszym ciągu rozwijana i modyfikowana. Celowe wydaje się sprawdzenie i ewentualne dostosowanie omawianej metody do obsługi kontrolera Kinect w wersji 2.0 (wersja dla konsoli Xbox One) działającego w oparciu o inną technologię, niż urządzenie użyte w niniejszej pracy. Odmiennosć tego urządzenia może sprawić, że część aktualnie wykonywanych kroków może być pominięta lub zastąpiona innymi. Innym możliwym kierunkiem badań jest poszerzenie spektrum ruchów, charakterystycznych dla dyscyplin sportowych odznaczających się dużą dynamiką zmian, co z kolei może mieć wpływ na dobór wartości wykorzystywanych jako kryterium sposobu łączenia sygnałów z urządzeń pomiarowych. Na przykład większa fragmentacja danych, przy szybkim ruchu kończyny, może mieć wpływ na wartość wariancji charakterystycznej dla prawidłowych pomiarów w określonym przedziale czasu.

Duże znaczenie dla przeprowadzonych badań miały warunki, zwłaszcza oświetleniowe, jakie panowały w laboratorium. Zdarzyło się kilkukrotnie, że ćwiczenia musiały zostać przerwane, bądź to ze względu na unoszenie się zanieczyszczenia niewiadomego pochodzenia w powietrzu, które odbijało światło w sposób zakłócający działanie Kinecta, bądź też ze względu na zakłócenia jakim poddana była transmisja danych za pomocą technologii komunikacji bezprzewodowej Bluetooth. Objawiało się to tym, że do komputera, który odbierał i rejestrował wszystkie dane, pakiety z pomiarami IMU trafiały niekompletne lub uszkodzone. Elementami, które wymagać mogą więc dalszego rozwoju jest na przykład poprawa formatu danych na taki, który byłby w stanie je odtworzyć w przypadku niewielkiego uszkodzenia w trakcie transmisji. Ostatnim z rozważanych przeze mnie kierunków dalszych badań jest modyfikacja urządzenia agregującego sygnały z czujników inercyjnych. Modyfikacja ta polegałaby na zamianie obecnie wykorzystywanego modelu opartego o jeden moduł centralny z podłączonymi do niego modułami inercyjnymi i zapewniającego komunikację z komputerem PC (rys. 3.3), na rzecz modelu opartego o wiele niezależnych, autonomicznych modułów inercyjnych, z których każdy komunikuje się niezależnie z komputerem PC, na

którym odbywa się łączenie wszystkich zaglebowanych pomiarów. Głównym powodem rozważanej modyfikacji jest przede wszystkim wygoda użytkowania systemu poprzez wyeliminowanie przewodów łączących moduły inercyjne z modułem centralnym. Z drugiej jednak strony, taka zmiana spowodowałaby potencjalne problemy synchronizacji sygnałów pomiędzy nawet kilkunastoma urządzeniami pomiarowymi, co mogłoby wprowadzić istotne błędy w śledzeniu ruchu postaci.

Dodatki

Dodatek A

Popularne metody filtracji i estymacji sygnałów pomiarowych

Istotnym etapem przetwarzania sygnałów jest ich filtracja. Zazwyczaj celem tej operacji jest usunięcie niepożądanego szumu z uzyskanych pomiarów. Znanych jest wiele metod filtracji, które weszły do powszechnego użytku, jednak szczególną grupą filtrów są filtry Kalmana. Przez ponad 50 lat znalazły one zastosowanie w przemyśle, a także stały się obiektem wielu⁵⁸ badań i opracowań naukowych [31, 51, 56, 88, 124].

Liniowy filtr Kalmana (KF - ang. *Kalman filter*)

Liniowy filtr Kalmana jest oryginalną metodą filtracji sygnału opracowaną i opisaną przez Rudolpha E. Kalmana w 1960 [58] i 1961 [59]. Metoda ta pozwala na oszacowanie wartości kolejnych pomiarów tylko na podstawie bieżących danych oraz modelu matematycznego obserwowanego zjawiska. Działanie filtra możemy podzielić na dwie następujące bezpośrednio po sobie fazy:

1. Predykacji,
2. Korekcji.

Faza pierwsza polega na oszacowaniu aktualnej wartości (stanu) pomiaru na podstawie ostatniego znanego pomiaru oraz modelu przejścia pomiędzy kolejnymi stanami. Jeśli w układzie występuje źródło sygnału kontrolnego, pomiary z niego uzyskane można wykorzystać w tej fazie do wymuszenia określonego stanu. Oszacowanie aktualnego pomiaru uwzględnia także występowanie szumu, który z założenia jest szumem białym. Równania A.1a oraz A.1b przedstawiają matematyczny zapis fazy predykacji nazywanej także aktualizacją czasową.

$$\widehat{x}_k = A\widehat{x}_{k-1} + Bu_{k-1} + w_{k-1} \quad (\text{A.1a})$$

$$P_k^- = AP_{k-1}A^T + Q \quad (\text{A.1b})$$

⁵⁸ W samej tylko bibliotece cyfrowej IEEE Xplore (<http://ieeexplore.ieee.org>) wyszukiwanie frazy "Kalman filter" zwraca około 25000 wyników.

gdzie:

- \widehat{x}_k – prognozowana wartość stanu w kroku k ,
- \widehat{x}_{k-1} – wartość stanu oszacowana w kroku $k-1$,
- P_k^- – prognozowana wartość kowariancji estymacji,
- P_{k-1} – faktyczna wartość kowariancji estymacji wykonana w kroku $k-1$,
- A – stała macierz definiująca przejście pomiędzy dwoma, następującymi po sobie stanami,
- B – stała macierz definiująca powiązanie pomiędzy aktualnym stanem/pomiarem, a pomiarem sygnału sterującego,
- w – zmienna losowa o rozkładzie normalnym będąca szumem procesu,
- Q – kowariancja szumu procesu.

Efektem działania tej fazy jest oszacowana wartość aktualnego stanu obserwowanego źródła wraz z kowariancją tego oszacowania.

Drugim etapem jaki możemy wyszczególnić w trakcie działania omawianego filtru jest faza korekcji lub inaczej aktualizacji pomiarowej. Matematycznie fazę tę można zapisać za pomocą równań A.2a, A.2b, A.2c oraz A.2d.

$$z_k = H * \widehat{x}_k + v_k \quad (\text{A.2a})$$

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (\text{A.2b})$$

$$\widehat{x}_k = \widehat{x}_{k-1} + K_k (z_k - H \widehat{x}_{k-1}) \quad (\text{A.2c})$$

$$P_k = (I - K_k H) P_{k-1}^- \quad (\text{A.2d})$$

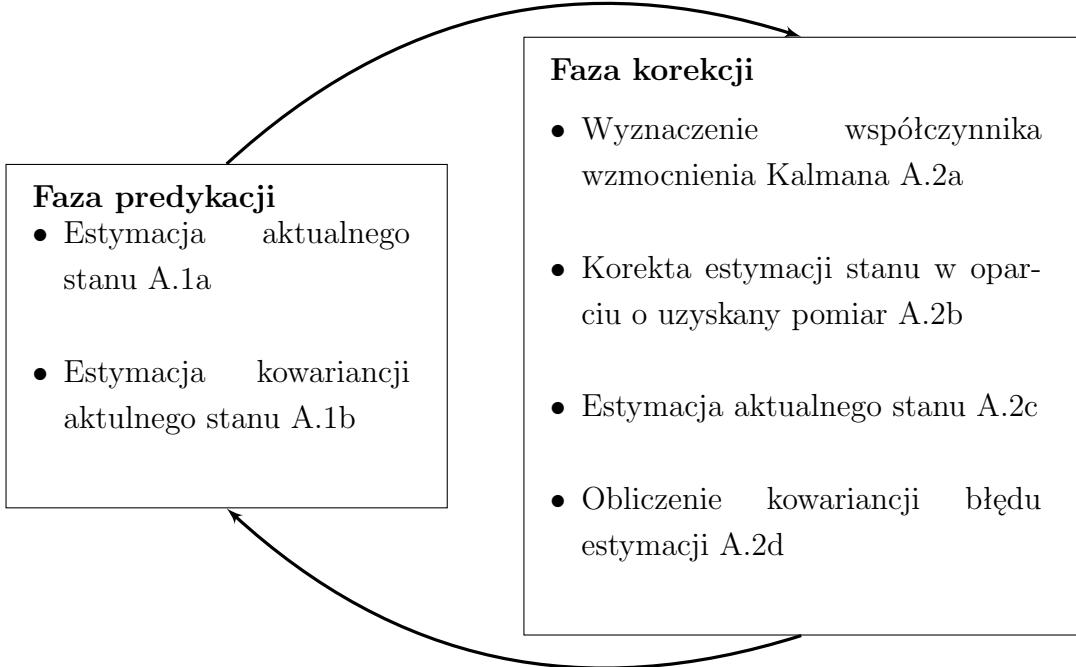
gdzie:

- z_k – aktualny pomiar uzyskany ze źródła w kroku k ,
- H – macierz definiująca powiązanie aktualnego pomiaru z oszacowanym stanem,
- v_k – szum pomiaru,
- K_k – wzmacnienie Kalmana (*ang. Kalman gain*),
- R – macierz wariancji uzyskanego pomiaru,
- I – macierz jednostkowa.

Wartością, którą filtr zwraca jako wynik swojego działania jest \widehat{x}_k .

Obie fazy wykonywane są cyklicznie, a kompletny algorytm przedstawia schemat z rysunku A.1.

Filtr Kalmana, dzięki swojej budowie, może być używany nie tylko w procesie odszumiania sygnału. Drugim ważnym i popularnym jego zastosowaniem jest łączenie ze sobą sygnałów z różnych źródeł. Wówczas wartości jednego z nich wykorzystywane są w fazie predykacji, a drugiego w fazie korekcji.



Rysunek A.1: Schemat działania filtru Kalmana (źródło: opracowanie własne)

Rozszerzony filtr Kalmana (EKF - ang. *Extended Kalman Filter*)

Filtr Kalmana w swojej pierwotnej formie stanowi optymalny estymator dla procesów o charakterze liniowym. Filtracja sygnałów nieliniowych wymagała modyfikacji oryginalnej formuły i prace takie zostały podjęte niemal natychmiast po opublikowaniu prac Rudolpha Kalmana. W 1962 Stanley F. Schmidt, jeden z dyrektorów NASA zaangażowany w program Apollo, wraz z zespołem opublikował pracę zawierającą zmodyfikowaną formułę liniowego filtru Kalmana przetwarzającą dane nieliniowe [108]. Działanie filtru opiera się na linearyzacji przetwarzanego sygnału za pomocą szeregu Taylora. Schemat działania jest analogiczny do tego przedstawionego na rysunku A.1, jednak zmianie ulegają poszczególne równania. Faza estymacji opiera się na równaniach A.3a i A.3b.

$$\widehat{x}_k^- = f(\widehat{x}_{k-1}, u_{k-1}, w_{k-1}) \quad (\text{A.3a})$$

$$P_k^- = A^J P_{k-1} A^{J^T} + W^J Q W^{J^T} \quad (\text{A.3b})$$

gdzie:

- $f()$ – funkcja nieliniowa określająca przejście pomiędzy kolejnymi stanami,

- A^J, W^J – jakobiany funkcji $f()$.

Równania A.4a, A.4b, A.4c oraz A.4d definiują natomiast fazę korekcji.

$$z_k = h(\widehat{x}_k, v_k) \quad (\text{A.4a})$$

$$K_k = P_k^- H^{JT} (H^J P_k^- H^{JT} + V^J R V^{JT})^{-1} \quad (\text{A.4b})$$

$$\widehat{x}_k = \widehat{x}_k + K_k(z_k - h(\widehat{x}_k, v_k)) \quad (\text{A.4c})$$

$$P_k = (I - K_k H^J) P_k^- \quad (\text{A.4d})$$

gdzie:

- h – funkcja nieliniowa definiująca relację pomiędzy estymowanym stanem a faktycznym pomiarem,
- H^J, V^J – jakobiany funkcji h .

Sposób wyznaczenia użytych jakobianów prezentują równania A.5a, A.5b, A.5c oraz A.5d.

$$A_{[i,j]}^J = \frac{\partial f_{[i]}}{\partial x_{[j]}}(\widehat{x}_{k-1}, u_{k-1}, w_{k-1}) \quad (\text{A.5a})$$

$$W_{[i,j]}^J = \frac{\partial f_{[i]}}{\partial w_{[j]}}(\widehat{x}_{k-1}, u_{k-1}, w_{k-1}) \quad (\text{A.5b})$$

$$H_{[i,j]}^J = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\widehat{x}_k, v_k) \quad (\text{A.5c})$$

$$V_{[i,j]}^J = \frac{\partial h_{[i]}}{\partial v_{[j]}}(\widehat{x}_k, v_k) \quad (\text{A.5d})$$

Filtr EKF okazał się skuteczny dla filtracji sygnałów nieliniowych jednak konieczność obliczania jakobianów zwiększa koszty obliczeniowe tego rozwiązania, co jest szczególnie widoczne w przypadku częstego dokonywania pomiarów.

Bezśadowy filtr Kalmana (UKF - ang. *Unscented Kalman Filter*)

W 1995 w wyniku prac prowadzonych przez S. Juliera, J. Uhlmanną i H. Durrant-Whyte'a została opublikowana [55] metoda filtrowania sygnałów nieliniowych oparta o analizę danych statystycznych filtrowanego sygnału. Realizowana jest ona poprzez zastosowanie przekształcenia bezśadowego operującego na rozkładzie prawdopodobieństwa filtrowanego sygnału, ponieważ mniej kosztowne, a jednocześnie dokładniejsze jest aproksymowanie

takich danych, niż funkcji nieliniowych [54]. Względem EKF, w przypadku tego algorytmu, występuje dodatkowy krok – wyznaczenie punktów sigma (równanie A.6).

$$\chi_{k-1}^{\alpha} = [\widehat{x}_{k-1}^{\alpha} \widehat{x}_{k-1}^{\alpha} \pm \sqrt{(L + \lambda) P_{k-1}^{\alpha}}] \quad (\text{A.6})$$

Kolejnym krokiem jest faza predykacji wyrażona za pomocą wzorów A.7a - A.7e

$$\chi_k^x = F[\chi_{k-1}^x, \chi_{k-1}^v] \quad (\text{A.7a})$$

$$\widehat{x}_k = \sum_{i=0}^{2L} W_i^{(m)} \chi_{i,k}^x \quad (\text{A.7b})$$

$$P_k^- = \sum_{i=0}^{2L} W_i^{(c)} [\chi_{i,k}^x - \widehat{x}_k] [\chi_{i,k}^x - \widehat{x}_k]^T \quad (\text{A.7c})$$

$$\gamma_k = H[\chi_k^x, \chi_{k-1}^n] \quad (\text{A.7d})$$

$$\widehat{y}_k = \sum_{i=0}^{2L} W_i^{(m)} \gamma_{i,k} \quad (\text{A.7e})$$

Faza korekcji została wyrażona wzorami A.8a - A.8e

$$P_{\widehat{y}_k \widehat{y}_k} = \sum_{i=0}^{2L} W_i^{(c)} [\gamma_{i,k} - \widehat{y}_k] [\gamma_{i,k} - \widehat{y}_k]^T \quad (\text{A.8a})$$

$$P_{\widehat{x}_k \widehat{y}_k} = \sum_{i=0}^{2L} W_i^{(c)} [\chi_{i,k}^x - \widehat{x}_k] [\gamma_{i,k} - \widehat{y}_k]^T \quad (\text{A.8b})$$

$$K_k = P_{\widehat{x}_k \widehat{y}_k} P_{\widehat{y}_k \widehat{y}_k}^{-1} \quad (\text{A.8c})$$

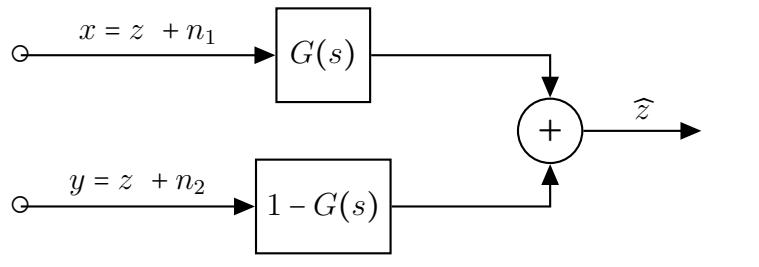
$$\widehat{x}_k = \widehat{x}_k + K_k (y_k - \widehat{y}_k) \quad (\text{A.8d})$$

$$P_k = P_k^- - K_k P_{\widehat{y}_k \widehat{y}_k} K_k^T \quad (\text{A.8e})$$

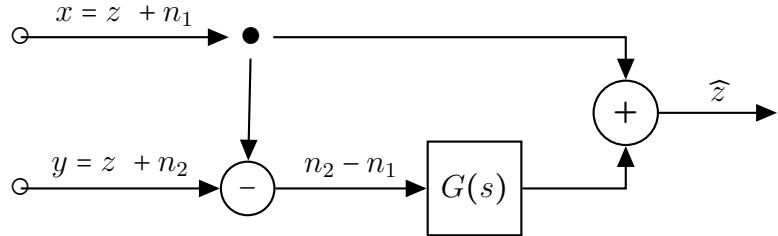
Dokładny opis bezśadowego filtru Kalmana można znaleźć między innymi w pracach E. Wana oraz R. van der Merwe [121, 122], którzy pracowali nad rozwinięciem tej metody. Wspomniani autorzy, w tej samej pracy, porównali ze sobą dokładność szacowania dla EKF i UKF wykorzystując sygnał wygenerowany na podstawie równania Mackey-Glassa [37] zaburzonego białym szumem. Zamieszczone wyniki jednoznacznie pokazują, że filtr UKF był w stanie oszacować odfiltrowany sygnał z mniejszym błędem niż EKF. Podobne badania, porównujące oba te filtry, były prowadzone także w polskich ośrodkach badawczych między innymi w Wojskowej Akademii Technicznej. Także i w tym przypadku, opublikowane wyniki wskazywały na większą dokładność filtru UKF [64, 65].

Filtr komplementarny

Filtr Kalmana w swojej podstawowej postaci, jak i w wersji rozszerzonej został od razu wykorzystany w pozaakademickim środowisku i stał się niejako standardowym narzędziem filtracji i fuzji sygnałów. Jednak w przypadku kiedy stan obserwowanego systemu jest wyrażony przez więcej niż jedną wartość liczbową, implementacja filtru Kalmana wymaga wykorzystania algebry macierzowej, która jest złożona obliczeniowo [60]. Świadomość tego ograniczenia skłoniła badaczy do poszukiwania alternatywnego sposobu na filtrację sygnałów przy zachowaniu zbliżonej dokładności. Przykład efektów takich badań możemy zobaczyć w pracy Higginsa [46], w której prezentuje on postać filtru komplementarnego oraz wykazuje związek takiej metody z filtrem Kalmana. Typowym zastosowaniem filtru komplementarnego jest złączenie ze sobą sygnałów opisujących to samo zjawisko, ale odznaczających się różną częstotliwością szumu (wysoką i niską). Ogólny schemat podstawowego filtru komplementarnego zaprezentowany został na rysunku A.2a.



(a) Wariant podstawowy



(b) Wariant pracujący na szumie

Rysunek A.2: Schemat działania filtru komplementarnego (źródło: opracowanie własne)

gdzie:

- x, y – zaszumione sygnały wejściowe,
- z – sygnał nieposiadający szumu (idealny),
- n_1, n_2 – szum jakiemu poddany jest sygnał z ,
- $G(s)$ – filtr dolnoprzepustowy; $G(s) = \frac{1}{\tau s + 1}$, $\tau \in \mathbb{R}$,
- $1 - G(s)$ – filtr górnoprzepustowy $1 - G(s) = 1 - \frac{1}{\tau s + 1} = \frac{\tau s}{\tau s + 1}$, $\tau \in \mathbb{R}$,

- \widehat{z} – sygnał wyjściowy będący odfiltrowanym złączeniem sygnałów x i y .

Zarówno wspomniany już Higgins, jak i Brown [19] w swoich artykułach pokazali związek pomiędzy filtrem Kalmana i filtrem komplementarnym. Wykazali tym samym, że może on stanowić mniej złożoną i mniej kosztowną obliczeniowo alternatywę dla filtra Kalmana bez znaczącej utraty jakości uzyskanego wyniku. Zagadnienie, w kontekście którego odbywały się badania nad tym filtrem, obejmowało problem integracji sygnałów z urządzeń pozwalających wyznaczyć prędkość oraz kierunek ruchu obiektu (w tym przypadku chodziło o poruszanie się pojazdów kosmicznych).

Filtry Mahoney'a i Madgwicka

Podobnie jak w przypadku filtru Kalmana, istnieje także wariant filtru komplementarnego dla sygnałów nieliniowych. W podobnym kontekście prace nad opracowaniem takiego filtra prowadził między innymi Robert Mahoney, który zaprezentował serię artykułów przedstawiających zarówno sam filtr, jak i sposób jego projektowania oraz zastosowanie [11, 32, 77, 78]. Na bazie tych prac w 2010 roku Sebastian Madgwick opracował autorską metodę [75, 76] integracji danych z czujników inercyjnych i magnetycznych w celu wyznaczenia orientacji w przestrzeni obserwowanego obiektu. Algorytm ten, według opublikowanych wyników eksperymentów, wykazuje większą dokładność niż algorytm oparty o filtr Kalmana. Pierwszym krokiem wykonywanym przez ten filtr jest wyznaczenie orientacji modułu inercyjnego na podstawie odczytów z żyroskopu oraz akcelerometru, a także wykorzystując dotychczasową estymację tego filtra. Przyjmując oznaczenia zaproponowane przez Madgwicka, ${}^S\omega = [0, \omega_x, \omega_y, \omega_z]$ wyraża wartości pomiarów otrzymanych z żyroskopu, natomiast pochodna kwaternionu opisującego zmianę orientacji czujnika inercyjnego względem Ziemi wyrażona jest wzorem A.9:

$${}^S_E\dot{q} = \frac{1}{2} {}^S_E\hat{\mathbf{q}} \otimes {}^S\omega \quad (\text{A.9})$$

gdzie ${}^S_E\hat{\mathbf{q}} = [q_1, q_2, q_3, q_4]$ ⁵⁹ oznacza dotychczasową orientację czujnika. Wykorzystując wzór A.9, możliwe jest określenie bieżącej orientacji czujnika względem Ziemi w chwili czasowej t (${}^S_E\mathbf{q}_{\omega,t}$) za pomocą całkowania numerycznego pochodnej ${}^S_E\dot{\mathbf{q}}_{\omega,t}$ wyznaczonej na podstawie aktualnego odczytu danych żyroskopu i estymacji orientacji czujnika w chwili $t - 1$ (${}^S_E\hat{q}_{est,t-1}$) zgodnie ze wzorem A.10.

$${}^S_E\dot{\mathbf{q}}_{\omega,t} = \frac{1}{2} {}^S_E\hat{q}_{est,t-1} \otimes {}^S\omega_t \quad (\text{A.10})$$

⁵⁹ q_4 w tym przypadku oznacza część rzeczywistą kwaternionu

Przyjmując, że Δt oznacza okres próbkowania, wówczas aktualna orientacja czujnika, wyznaczona na podstawie odczytów z żyroskopu, wyrażona jest wzorem A.11:

$$\overset{S}{E}\mathbf{q}_{\omega,t} = \overset{S}{E}\hat{q}_{est,t-1} \otimes \overset{S}{E}\dot{\mathbf{q}}_{\omega,t} \Delta t \quad (\text{A.11})$$

Na podstawie pomiarów czujnika ${}^S\hat{\mathbf{s}} = [0, s_x, s_y, s_z]$ określającego kierunki sił działających w polu grawitacyjnym (akcelerometr) lub magnetycznym (magnetometr) możliwe jest wyznaczenie orientacji czujnika w tym polu. W związku z tym, definiując referencyjny kierunek działania sił ${}^E\hat{\mathbf{d}} = [0, d_x, d_y, d_z]$, jednoznaczne wyznaczenie orientacji przestrzennej można przedstawić jako problem optymalizacji $\min_{\overset{S}{E}\hat{q} \in \Re^4} \mathbf{f}({}^S\hat{q}, {}^E\hat{\mathbf{d}}, {}^S\hat{\mathbf{s}})$ gdzie funkcja $\mathbf{f}()$ wyrażona wzorem A.12.

$$\begin{aligned} \mathbf{f}({}^S\hat{q}, {}^E\hat{\mathbf{d}}, {}^S\hat{\mathbf{s}}) &= {}^S\hat{\mathbf{q}}^* \otimes {}^E\hat{\mathbf{d}} \otimes {}^S\hat{\mathbf{q}} - {}^S\hat{\mathbf{s}} = \\ &\left[\begin{array}{l} 2d_x(1/2 - q_3^2 - q_4^2) + 2d_y(q_1q_4 + q_2q_3) + 2d_z(q_2q_4 - q_1q_3) - s_x \\ 2d_x(q_2q_3 - q_1q_4) + 2d_y(1/2 - q_2^2 - q_4^2) + 2d_z(q_1q_2 + q_3q_4) - s_y \\ 2d_x(q_1q_3 + q_2q_4) + 2d_y(q_3q_4 + q_1q_2) + 2d_z(1/2 - q_2^2 - q_3^2) - s_z \end{array} \right] \end{aligned} \quad (\text{A.12})$$

Wzór A.12 pozwala na wyznaczenie różnicy pomiędzy pomiarami czujnika, a wartościami referencyjnymi. Symbol ${}^S\hat{\mathbf{q}}^*$ oznacza koniugację kwaternionu ${}^S\hat{\mathbf{q}}$. W przypadku omawianego filtru Madgwicka, metoda optymalizująca funkcję $\mathbf{f}()$ została zaimplementowana w formie iteracyjnej metody malejących gradientów zgodnie ze wzorem A.13:

$$\overset{S}{E}\hat{q}_{k+1} = \overset{S}{E}\hat{q}_k - \mu \frac{\nabla \mathbf{f}({}^S\hat{q}, {}^E\hat{\mathbf{d}}, {}^S\hat{\mathbf{s}})}{\|\nabla \mathbf{f}({}^S\hat{q}, {}^E\hat{\mathbf{d}}, {}^S\hat{\mathbf{s}})\|}, k = 0, 1, 2 \dots n \quad (\text{A.13})$$

gdzie μ oznacza rozmiar kroku przesunięcia danych pomiędzy iteracjami, natomiast gradient funkcji $\mathbf{f}()$ (wz. A.12) przedstawia wzór A.14

$$\nabla \mathbf{f}({}^S\hat{q}, {}^E\hat{\mathbf{d}}, {}^S\hat{\mathbf{s}}) = \mathbf{J}^T({}^S\hat{q}, {}^E\hat{\mathbf{d}}) \mathbf{f}({}^S\hat{q}, {}^E\hat{\mathbf{d}}, {}^S\hat{\mathbf{s}}) \quad (\text{A.14})$$

gdzie $\mathbf{J}({}^S\hat{q}, {}^E\hat{\mathbf{d}})$ przedstawia Jakobian wyrażony wzorem A.15.

$$\begin{aligned} \mathbf{J}({}^S\hat{q}, {}^E\hat{\mathbf{d}}) &= \\ &\left[\begin{array}{cccc} 2d_yq_4 - 2d_zq_3 & 2d_yq_3 + 2d_zq_4 & -4d_xq_3 + 2d_yq_2 - 2d_zq_1 & -4d_xq_4 + 2d_yq_1 + 2d_zq_2 \\ -2d_xq_4 + 2d_zq_2 & 2d_xq_3 - 4d_yq_2 + 2d_zq_1 & 2d_xq_2 + 2d_zq_4 & -2d_xq_1 - 4d_yq_4 + 2d_zq_3 \\ 2d_xq_3 - 2d_yq_2 & 2d_xq_4 - 2d_yq_1 - 4d_zq_2 & 2d_xq_1 + 2d_yq_4 - 4d_zq_3 & 2d_xq_2 + 2d_yq_3 \end{array} \right] \end{aligned} \quad (\text{A.15})$$

Przyjmując, że wykorzystywanym czujnikiem jest akcelerometr (${}^S\hat{\mathbf{a}} = [0, a_x, a_y, a_z]$) działający w polu grawitacyjnym Ziemi ${}^E\hat{\mathbf{g}} = [0, 0, 0, 1]$, wówczas funkcja ze wzoru A.12 przyjmuje postać jak we wzorze A.16

$$\mathbf{f}({}_E^S \hat{q}, {}^S \hat{\mathbf{a}}) = \begin{bmatrix} 2(q_2 q_4 - q_1 q_3) - a_x \\ 2(q_1 q_2 + q_3 q_4) - a_y \\ 2(1/2 - q_2^2 + q_3^2) - a_z \end{bmatrix} \quad (\text{A.16})$$

natomiaszt Jakobian przyjmuje postać przedstawioną we wzorze A.17

$$\mathbf{J}_g({}_E^S \hat{q}) = \begin{bmatrix} -2q_3 & 2q_4 & -2q_1 & 2q_2 \\ 2q_2 & 2q_1 & 2q_4 & 2q_3 \\ 0 & -4q_2 & -4q_3 & 0 \end{bmatrix} \quad (\text{A.17})$$

Na podstawie wzorów A.13, A.16, A.17, orientacja przestrzenna akcelerometru w polu grawitacyjnym Ziemi w chwili t wyrażona jest wzorem A.18

$${}_E^S q_{\nabla, t} = {}_E^S \hat{q}_{est, t-1} - \mu_t \frac{\nabla \mathbf{f}}{\|\mathbf{f}\|} \quad (\text{A.18})$$

gdzie $\nabla \mathbf{f} = \mathbf{J}_g^T({}_E^S \hat{q}_{est, t-1}) \mathbf{f}_g({}_E^S \hat{q}_{est, t-1}, {}^S \hat{\mathbf{a}}_t)$.

Kolejnym krokiem po wyznaczeniu orientacji na podstawie żyroskopu oraz akcelerometru jest fuzja tych dwóch wartości. Madgwick przyjął jako metodę fuzji orientacji z obu źródeł filtr komplementarny zdefiniowany zgodnie ze wzorem A.19

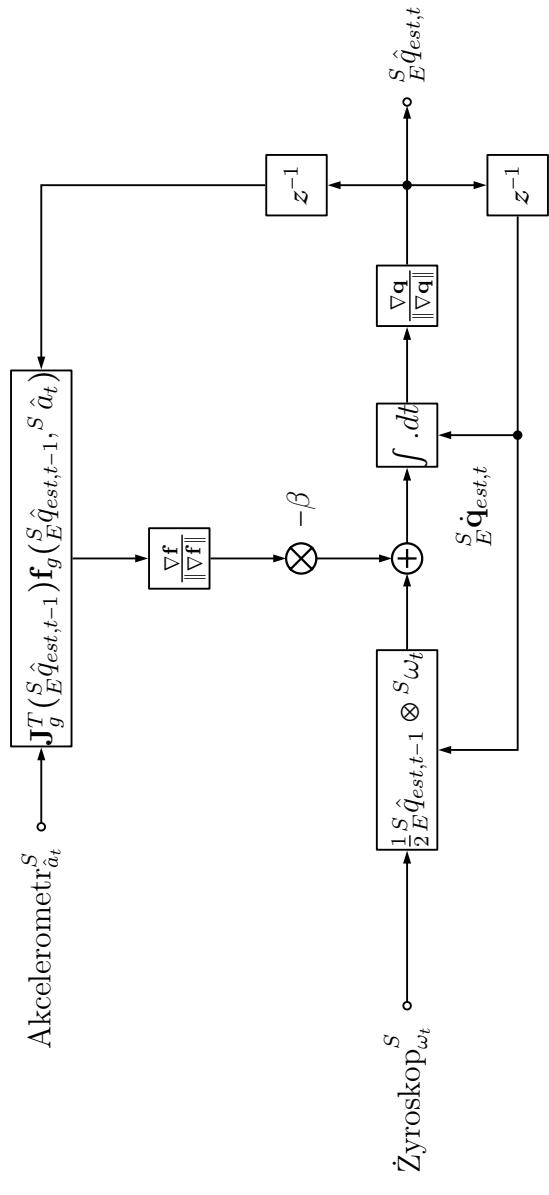
$${}_E^S \hat{q}_{est, t} = \alpha_t {}_E^S q_{\nabla, t} + (1 - \alpha_t) {}_E^S q_{\omega, t}, \quad 0 \leq \alpha \leq 1 \quad (\text{A.19})$$

gdzie α_t określa poziom istotności orientacji wyznaczonej na podstawie akcelerometru w danej chwili t . Wartość wagi α_t jest zmienna w czasie i jest wyznaczona na podstawie wzoru A.20.

$$\alpha_t = \frac{\beta}{\mu_t / \Delta t + \beta} \quad (\text{A.20})$$

Współczynnik β jest określany na podstawie błędu pomiarów żyroskopu wyznaczonego za pomocą Wariancji Allana.

Rysunek A.3 przedstawia schemat blokowy filtra Madgwicka. Warto zauważyć analogię w jego budowie do schematu filtra komplementarnego przedstawionego na rysunku A.2a.



Rysunek A.3: Diagram blokowy przedstawiający filtr Madgwicka dla czujników inercyjnych (źródło: opracowanie własne)

Dodatek B

Wariancja Allana

Wariancja Allana jest metodą wykorzystywaną do określenia charakterystyki szumów zawartych w sygnałach mierzonych przez układy oparte o oscylatory. Metoda ta została opisana w [4] jako część pracy magisterskiej w latach 60-tych XX wieku i od tamtego czasu została przyjęta jako standardowa do wyznaczania charakterystyk tego typu układów. W literaturze można znaleźć między innymi prace opisujące analizę sygnałów układów inercyjnych [35, 103], czy układów GPS [127]. Wariancja Allana ($\sigma_{\Omega}^2(\Delta t)$) wyznaczana jest dla sygnału Ω wyrażonego w domenie czasu i podzielonego na N przedziałów o czasie trwania Δt każdy. Przyjmując, że n oznacza numer przedziału badanego sygnału, a $\Omega_n(\Delta t)$ oznacza średnią wartość sygnału w przedziale o numerze n o długości czasu Δt , wariancję Allana możemy wyrazić za pomocą równania B.1[5, 110].

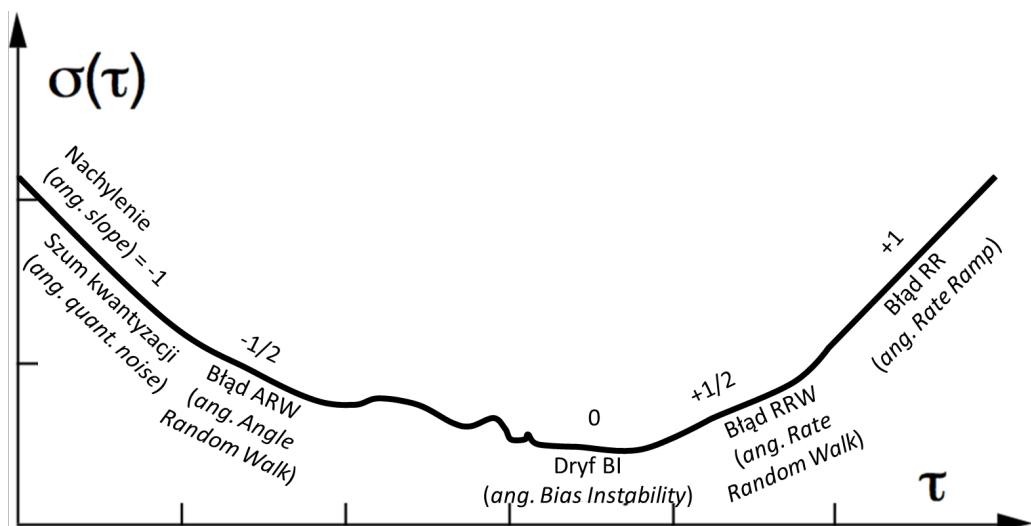
$$\sigma_{\Omega}^2(\Delta t) = \frac{1}{2(N-1)} \sum_{k=1}^{N-1} [\Omega_{n+1}(\Delta t) - \Omega_n(\Delta t)]^2 \quad (\text{B.1})$$

Aby na podstawie wyznaczonej wariancji określić charakterystyki szumów należy obliczyć odchylenie Allana $\sigma_{\Omega}(\Delta t) = \sqrt{\sigma^2(\Delta t)}$ i przeanalizować przebieg tej funkcji na wykresie w skali logarytmicznej (przykładowy wykres odchylenia Allana dla żyroskopu przedstawia wykres z rysunku B.1). Wykres ten zawiera informacje o 5 różnych rodzajach szumu jakie są zawarte w badanym sygnale (nazwy szumów na podstawie [89]):

1. Błąd kwantyzacji – szum wprowadzony w trakcie konwersji sygnału analogowego na cyfrowy i spowodowany przez różnice pomiędzy faktyczną wartością sygnału, a wartością wynikającą z punktu, w którym nastąpiło próbkowanie.
2. Błąd ARW – szum wysokiej częstotliwości obserwowalny nawet w czasie krótkich pomiarów. Jego występowanie znaczco obniża dokładność prowadzonych obliczeń.
3. Błąd BI – szum niskiej częstotliwości związany ze zmianą właściwości fizycznych materiałów, z których zbudowane są czujniki.

4. Błąd RRW – szum wysokiej częstotliwości pojawiający się w pomiarach w dłuższym okresie czasu. Jego pochodzenie nie jest do końca znane.
5. Błąd RR – systematycznie pojawiający się błąd pomiarów, którego wartość zależna jest od czasu działania. Pomiary zakłócone tego rodzaju szumem nie są w stanie powrócić do wartości początkowych.

Powyższe zakłócenia pomiarów zawarte są na wykresie odchylenia jako obszary, w których kąt nachylenia wykresu do osi X jest zbliżony do kąta nachylenia prostej (*ang. slope*) o współczynniku kierunkowym wynoszącym odpowiednio: $-1, -0.5, 0, 0.5, 1$. Obszary te są widoczne na wykresie z rysunku B.1.



Rysunek B.1: Obszary szumów na wykresie odchylenia Allana [45]

Nie zawsze wykres funkcji $\sigma(\tau)$ musi zawierać wszystkie 5 obszarów co oznacza, że niektóre rodzaje szumów mogą nie występować.

Bibliografia

- [1] A. Agarwal i B. Triggs, “Recovering 3D human pose from monocular images”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, s. 44–58, 2006. DOI: 10.1109/TPAMI.2006.21.
- [2] K. Alexiev i I. Nikolova, “An algorithm for error reducing in IMU”, *2013 IEEE International Symposium on Innovations in Intelligent Systems and Applications*, 2013. DOI: 10.1109/INISTA.2013.6577663.
- [3] A. Ali i J. K. Aggarwal, “Segmentation and recognition of continuous human activity”, *Proceedings - IEEE Workshop on Detection and Recognition of Events in Video, EVENT 2001*, s. 28–35, 2001. DOI: 10.1109/EVENT.2001.938863.
- [4] D. W. Allan, “Statistics of atomic frequency standards”, *Proceedings of the IEEE*, s. 221–230, 1966. DOI: 10.1109/PROC.1966.4634.
- [5] D. W. Allan, “Should the classical variance be used as a basic measure in standards metrology?”, *IEEE Transactions on Instrumentation and Measurement*, s. 646–654, 1987, ISSN: 0018-9456. DOI: 10.1109/TIM.1987.6312761.
- [6] M. Andrejasic, “Mems accelerometers”, Seminar, University of Ljubljana, Ljubljana, s. 1–17, 2008. adr.: <https://goo.gl/eg27U9>.
- [7] L. Armesto, J. Tornero i M. Vincze, “Fast Ego-motion estimation with multi-rate fusion of inertial and vision”, *The International journal of Robotics Research*, s. 577–589, 2007. DOI: 10.1177/0278364907079283.
- [8] S. Asteriadis, A. Chatzitofis, D. Zarpalas, D. S. Alexiadis i P. Daras, “Estimating human motion from multiple Kinect sensors”, *Proceedings of the 6th International Conference on Computer Vision / Computer Graphics Collaboration Techniques and Applications - MIRAGE '13*, s. 1–6, 2013. DOI: 10.1145/2466715.2466727.
- [9] A. Baak, M. Muller, G. Bharaj, H. P. Seidel i C. Theobalt, “A data-driven approach for real-time full body pose reconstruction from a depth camera”, *Proceedings of the IEEE International Conference*

- on Computer Vision*, s. 1092–1099, 2011. DOI: 10.1109/ICCV.2011.6126356.
- [10] S. Baek i M. Kim, “Real-time performance capture using multiple Kinects”, *2014 International Conference on Information and Communication Technology Convergence (ICTC)*, s. 647–648, 2014. DOI: 10.1109/ICTC.2014.6983241.
 - [11] G. Baldwin, R. Mahony, J. Trumpf, T. Hamel i T. Cheviron, “Complementary filter design on the Special Euclidean group SE(3)”, *Proc. of the European Control Conference*, s. 3763–3770, 2007.
 - [12] P. Barański i P. Strumiłło, “Układ nawigacji zliczeniowej do korekcji odczytów GPS w terenie miejskim”, *Przegląd Elektrotechniczny*, s. 160–163, 2009.
 - [13] P. Barański, P. Strumiłło i P. Polańczyk, “Fusion of data from inertial sensors, raster maps and GPS for estimation of pedestrian geographic location in urban terrain”, *Metrology and Measurement Systems*, s. 145–158, 2011. DOI: 10.2478/v10178-011-0014-3.
 - [14] Bingbing Ni, A. Kassim i S. Winkler, “A hybrid framework for 3-d human motion tracking”, *IEEE Transactions on Circuits and Systems for Video Technology*, s. 1075–1084, 2008. DOI: 10.1109/TCSVT.2008.927108.
 - [15] A. P. L. Bo, M. Hayashibe i P. Poignet, “Joint angle estimation in rehabilitation with inertial sensors and its integration with Kinect”, *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, s. 3479–3483, 2011. DOI: 10.1109/IEMBS.2011.6090940.
 - [16] A. Bochenek i M. Reicher, “Kości, stawy i więzadła”, w *Anatomia człowieka Tom I. Anatomia ogólna, kości, stawy i więzadła, mięśnie*, W. Łasiński, red., XI (VII), Warszawa: Państwowy zakład wydawnictw lekarskich, s. 187–219. 1999.
 - [17] “IEEE standard specification format guide and test procedure for linear, single-axis, non-gyroscopic accelerometers”, *IEEE Std 1293-1998 (R2008)*, K. Breitfekder, red., s. 1–249, 2011. DOI: 10.1109/IEEESTD.2011.5960745.
 - [18] A. L. Brooks, A. Czarowicz i N. Bohrs, “Markerless motion tracking : MS Kinect and Organic Motion OpenStage”, *Proceedings of the 9th Intl Conference on Disability, Virtual Reality and Associated Technologies*, s. 10–12, 2012.
 - [19] R. G. BROWN, “Integrated navigation systems and kalman filtering: a perspective”, *Navigation*, s. 355–362, 1972. DOI: 10.1002/j.2161-4296.1972.tb01706.x.

- [20] M. Candidi, R. Orfei, F. Palutan i G. Vannaroni, “FFT analysis of a space magnetometer noise”, *IEEE Transactions on Geoscience Electronics*, s. 23–28, 1974. DOI: 10.1109/TGE.1974.294327.
- [21] D. Comaniciu, V. Ramesh i P. Meer, “Kernel-based object tracking”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, s. 564–577, 2003. DOI: 10.1109/TPAMI.2003.1195991.
- [22] S. Corazza, L. Mundermann, a. M. Chaudhari, T. Demattio, C. Cobelli i T. P. Andriacchi, “A markerless motion capture system to study musculoskeletal biomechanics: visual hull and simulated annealing approach.”, *Annals of Biomedical Engineering*, s. 1019–29, 2006. DOI: 10.1007/s10439-006-9122-8.
- [23] A. Criminisi, J. Shotton i E. Konukoglu, “Decision forests for classification , regression , density estimation , manifold learning and semi-supervised learning”, *Learning*, s. 81–227, 2011. DOI: 10.1561/0600000035.
- [24] N. Dantam, “Quaternion computation”, Georgia Institute of Technology, spraw. tech., s. 1–9, 2014.
- [25] P. Daponte, L. De Vito, S. Rapuano, M. Riccio i F. Picariello, “Compensating magnetic disturbances on MARG units by means of a low complexity data fusion algorithm”, *2015 IEEE International Symposium on Medical Measurements and Applications (MeMeA) Proceedings*, s. 157–162, 2015. DOI: 10.1109/MeMeA.2015.7145191.
- [26] R. B. Davis, S. Ounpuu, D. Tyburski i J. R. Gage, “A gait analysis data collection and reduction technique”, *Human Movement Science*, s. 575–587, 1991. DOI: 10.1016/0167-9457(91)90046-Z. arXiv: 101.
- [27] F. Destelle, A. Ahmadi, N. E. O’Connor, K. Moran, A. Chatzitofis, D. Zarpalas i P. Daras, “Low-cost accurate skeleton tracking based on fusion of kinect and wearable inertial sensors”, *European Signal Processing Conference*, s. 371–375, 2014.
- [28] R. Dhuli, “Multirate Kalman filter for sensor data fusion”, *Proceedings of the National Conference on Communications (NCC)*, s. 2005–2008, 2009.
- [29] J. Diebel, “Representing attitude: Euler angles, unit quaternions, and rotation vectors”, *Journal of Experimental Botany*, s. 4019–4026, 2007. DOI: 10.1093/jxb/erm298.
- [30] N. M. DiFilippo i M. K. Jouaneh, “Characterization of different Microsoft Kinect sensor models”, *IEEE Sensors Journal*, s. 4554–4564, 2015. DOI: 10.1109/JSEN.2015.2422611.

- [31] L. Drolet, F. Michaud i J. Cote, “Adaptable sensor fusion using multiple Kalman filters”, *Proceedings. 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2000) (Cat. No.00CH37113)*, s. 1434–1439, 2000. DOI: 10.1109/IROS.2000.893222.
- [32] M. Euston, P. Coote, R. Mahony, T. Hamel, J. Kim i T. Hamel, “A complementary filter for attitude estimation of a fixed-wing UAV”, *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, s. 340–345, 2008. DOI: 10.1109/IROS.2008.4650766.
- [33] D. Fofi, T. Sliwa i Y. Voisin, “A comparative survey on invisible structured light”, *SPIE Machine Vision Applications in Industrial Inspection XII*, J. R. Price i F. Meriaudeau, red., s. 1–90, 2004. DOI: 10.1117/12.525369.
- [34] B. Freedman, A. Shpunt, M. Machline i A. Yoel, “Depth mapping using projected patterns”, 2010. adr.: <https://goo.gl/RRA8zF>.
- [35] Freescale Semiconductor, Inc., “Allan variance: noise analysis for gyroscopes”, s. 1–9, 2015. adr.: <https://goo.gl/qUtvON>.
- [36] V. Ganapathi, C. Plagemann, D. Koller i S. Thrun, “Real time motion capture using a single time-of-flight camera”, *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2010)*, s. 755–762, 2010. DOI: 10.1109/CVPR.2010.5540141.
- [37] L. Glass i M. Mackey, “Mackey-Glass equation”, *Scholarpedia*, s. 6908, 2010. DOI: 10.4249/scholarpedia.6908. adr.: <https://goo.gl/70SxNy>.
- [38] G. Glonek i M. Pietruszka, “Natural user interfaces (NUI): review”, *Journal of Applied Computer Science*, s. 27–45, 2012. adr.: <https:// goo.g1/1rDwzJ>.
- [39] G. Glonek i A. Wojciechowski, “Hybrid method of human limb joints positioning—hand movement case study”, w *Information Technologies in Medicine, 2016, vol. 2*, Springer International Publishing, s. 307–320. 2016. DOI: 10.1007/978-3-319-39904-1_28.
- [40] G. Glonek i A. Wojciechowski, “Kinect and IMU sensors imprecisions compensation method for human limbs tracking.”, w *Computer Vision and Graphics. International Conference, ICCVG 2016, Warsaw, Poland, September 19-21, 2016. Proceedings*, Springer International Publishing, s. 316–328. 2016. DOI: 10.1007/978-3-319-46418-3_28.
- [41] P. Guan, A. Weiss, A. O. Balan i M. J. Black, “Estimating human shape and pose from a single image”, *2009 IEEE 12th International Conference on Computer Vision*, s. 1381–1388, 2009.

- [42] M. Hansard, S. Lee, O. Choi i R. Horaud, “Characterization of Time-of-Flight data”, w *Time-of-Flight cameras: principles, methods and applications*, London: Springer London, s. 1–28. 2013, ISBN: 978-1-4471-4658-2. DOI: 10.1007/978-1-4471-4658-2_1.
- [43] I. Haritaoglu, D. Harwood i L. S. Davis, “W4S: A real-time system for detecting and tracking people in 2 1/2D”, *Computer Vision - ECCV'98, 5th European Conference on Computer Vision*, H. Burkhardt i B. Neumann, red., s. 877–892, 1998. DOI: 10.1007/BFb0055710.
- [44] T. Helten, M. Muller, H.-P. Seidel i C. Theobalt, “Real-time body tracking with one depth camera and inertial sensors”, *2013 IEEE International Conference on Computer Vision*, s. 1105–1112, 2013. DOI: 10.1109/ICCV.2013.141.
- [45] J. Hidalgo, P. Poulakis, J. Köhler, J. Del-Cerro i A. Barrientos, “Improving planetary rover attitude estimation via MEMS sensor characterization”, *Sensors*, s. 2219–2235, 2012. DOI: 10.3390/s120202219.
- [46] W. Higgins, “A comparison of complementary and Kalman filtering”, *IEEE Transactions on Aerospace and Electronic Systems*, s. 321–325, 1975. DOI: 10.1109/TAES.1975.308081.
- [47] O. Hinton, “Multirate digital signal processing”, w *EEE305*, *EEE801 Part A: Digital Signal Processing*, 2001 wyd., Newcastle: University of Newcastle upon Tyne, rozd. 9, s. 1–8. 2001. adr.: <https://goo.gl/fiuA6o>.
- [48] J. P. Holden, J. A. Orsini, K. L. Siegel, T. M. Kepple, L. H. Gerber i S. J. Stanhope, “Surface movement errors in shank kinematics and knee kinetics during gait”, *Gait and Posture*, s. 217–227, 1997. DOI: 10.1016/S0966-6362(96)01088-0.
- [49] M. E. Huber, A. L. Seitz, M. Leeser i D. Sternad, “Validity and reliability of Kinect skeleton for measuring shoulder joint angles: a feasibility study”, *Physiotherapy (United Kingdom)*, s. 389–393, 2015. DOI: 10.1016/j.physio.2015.02.002.
- [50] J. Huerta, “Introducing the quaternions”, Fullerton College, s. 1–25, 2010.
- [51] Y. Huo, Z. Cai, W. Gong i Q. Liu, “A new adaptive Kalman filter by combining evolutionary algorithm and fuzzy inference system”, *IEEE Congr. Evol. Comput.*, s. 2893–2900, 2014. DOI: 10.1109/CEC.2014.6900320.
- [52] D. Hyun, H. S. Yang, G. H. Yuk i H. S. Park, “A dead reckoning sensor system and a tracking algorithm for mobile robots”, *2009*

- IEEE International Conference on Mechatronics*, s. 1–6, 2009. DOI: 10.1109/ICMECH.2009.4957155.
- [53] G. Johansson, “Visual perception of biological motion and a model for its analysis”, *Perception and Psychophysics*, s. 201–211, 1973. DOI: 10.3758/BF03212378.
 - [54] S. J. Julier i J. K. Uhlmann, “New extension of the Kalman filter to nonlinear systems”, *Proceedings of SPIE - The International Society for Optical Engineering*, I. Kadar, red., s. 1–12, 1997. DOI: 10.1117/12.280797.
 - [55] S. J. Julier, J. K. Uhlmann i H. F. Durrant-Whyte, “A new approach for filtering nonlinear systems”, *Proceedings of 1995 American Control Conference - ACC'95*, s. 1–4, 1995. DOI: 10.1109/ACC.1995.529783.
 - [56] J. Kędzierski, “Filtr Kalmana - zastosowania w prostych układach sensorycznych”, *Koło Naukowe Robotyków KoNaR*, s. 1–33, 2007. DOI: 10.13140/RG.2.1.1336.4884.
 - [57] C. Kalkbrenner, S. Hacker, M.-e. Algorri i R. Blechschmidt-trapp, “Motion capturing with inertial measurement units and Kinect - tracking of limb movement using optical and orientation information”, *Proceedings of the International Conference on Biomedical Electronics and Devices*, s. 120–126, 2014. DOI: 10.5220/0004787601200126.
 - [58] R. E. Kalman, “A new approach to linear filtering and prediction problems”, *Journal of Basic Engineering*, s. 1–35, 1960. DOI: 10.1115/1.3662552.
 - [59] R. E. Kalman i R. S. Bucy, “New results in linear filtering and prediction theory”, *Journal of Basic Engineering*, s. 95–108, 1961. DOI: 10.1115/1.3658902.
 - [60] E. L. Kaltofen i A. Storjohann, “Complexity of computational problems in exact linear algebra”, w *Encyclopedia of Applied and Computational Mathematics*, Berlin, Heidelberg: Springer Berlin Heidelberg, s. 227–233. 2015. DOI: 10.1007/978-3-540-70529-1_173.
 - [61] M. Kepski i B. Kwolek, “Fall detection using body-worn accelerometer and depth maps acquired by active camera”, *Hybrid Artificial Intelligence Systems*, s. 414–426, 2016. DOI: 10.1007/978-3-319-32034-2_35.
 - [62] A. Kitsikidis, K. Dimitropoulos, S. Douka i N. Grammalidis, “Dance analysis using multiple Kinect sensors”, *2014 International Conference on Computer Vision Theory and Applications (VISAPP)*, s. 789–795, 2014.
 - [63] Y. Kohno, M. Maeda, T. Hamamura i B. Irie, “The measurement of carried cartons using multiple Kinect sensors”, *MVA2013 IAPR*

- International Conference on Machine Vision Applications*, s. 367–370, 2013.
- [64] S. Konatowski, “Dokładność procesu estymacji w wybranych filtrach Kalmana”, *Biuletyn Wojskowej Akademii Technicznej*, s. 233–242, 2007.
 - [65] S. Konatowski i B. Pudlak, “Ocena dokładności filtracji nieliniowej w procesie określania położenia”, *Biuletyn Wojskowej Akademii Technicznej*, s. 243–253, 2007.
 - [66] T. Krzeszowski, “Śledzenie ruchu postaci ludzkiej w systemie wielokamerowym”, prac. dokt., Politechnika Śląska, s. 1–113, 2013. adr.: <https://goo.gl/2lpUf7>.
 - [67] B. Kwolek i M. Kepski, “Fall detection using Kinect sensor and fall energy image”, *Hybrid Artificial Intelligent Systems*, 2013.
 - [68] B. Kwolek, T. Krzeszowski, A. Michalcuk i H. Josinski, “3D gait recognition using spatio-temporal motion descriptors”, *6th Asian Conference on Intelligent Information and Database Systems 2014: Intelligent Information and Database Systems*, Lecture Notes in Computer Science, N. T. Nguyen, M. T. Le i J. Świątek, red., s. 595–604, 2014. DOI: [10.1007/978-3-319-05458-2_61](https://doi.org/10.1007/978-3-319-05458-2_61).
 - [69] B. Kwolek, T. Krzeszowski i K. Wojciechowski, “Swarm intelligence based searching schemes for articulated 3D body motion tracking”, w *Advanced Concepts for Intelligent Vision Systems: 13th International Conference, ACIVS 2011, Ghent, Belgium, August 22–25, 2011. Proceedings*, Berlin, Heidelberg: Springer Berlin Heidelberg, s. 115–126. 2011, ISBN: 978-3-642-23687-7. DOI: [10.1007/978-3-642-23687-7_11](https://doi.org/10.1007/978-3-642-23687-7_11).
 - [70] P. de Leva, “Adjustments to Zatsiorsky-Seluyanov’s segment inertia parameters”, *Journal of Biomechanics*, s. 1223–1230, 1996. DOI: [10.1016/0021-9290\(95\)00178-6](https://doi.org/10.1016/0021-9290(95)00178-6).
 - [71] R. Li, M.-H. Yang, S. Sclaroff i T.-P. Tian, “Monocular tracking of 3D human motion with a coordinated mixture of factor analyzers”, *Computer Vision – ECCV 2006: 9th European Conference on Computer Vision, Graz, Austria, May 7–13, 2006. Proceedings, Part II*, s. 137–150, 2006. DOI: [10.1007/11744047_11](https://doi.org/10.1007/11744047_11).
 - [72] G. Liu, A. Wang, T. Jiang, J. Jiao i J.-B. Jang, “Effects of environmental temperature on the performance of a micromachined gyroscope”, *Microsystem Technologies*, s. 199–204, 2007. DOI: [10.1007/s00542-007-0411-4](https://doi.org/10.1007/s00542-007-0411-4).

- [73] G. Liu, F. Yang, X. Bao i T. Jiang, “Robust optimization of a MEMS accelerometer considering temperature variations”, *Sensors*, s. 6342–6359, 2015. DOI: 10.3390/s150306342.
- [74] J. MacCormick, “How does the Kinect work?”, Dickinson College, s. 1–52, 2011. adr.: <https://goo.gl/5rcbZS>.
- [75] S. O. H. Madgwick, A. J. L. Harrison i R. Vaidyanathan, “An efficient orientation filter for inertial and inertialmagnetic sensor arrays”, *Internal Report*, s. 1–7, 2010.
- [76] S. O. H. Madgwick, A. J. L. Harrison i R. Vaidyanathan, “Estimation of IMU and MARG orientation using a gradient descent algorithm”, *IEEE International Conference on Rehabilitation Robotics*, s. 179–185, 2011. DOI: 10.1109/ICORR.2011.5975346.
- [77] R. Mahony, T. Hamel i J. M. Pflimlin, “Nonlinear complementary filters on the special orthogonal group”, *IEEE Trans. Automat. Contr.*, s. 1203–1218, 2008. DOI: 10.1109/TAC.2008.923738.
- [78] R. Mahony, T. Hamel i J.-M. Pflimlin, “Complementary filter design on the special orthogonal group SO(3)”, *Proceedings of the 44th IEEE Conference on Decision and Control*, s. 1477–1484, 2005. DOI: 10.1109/CDC.2005.1582367.
- [79] F. Malouin, C. L. Richards, B. McFadyen i J. Doyon, “To walk again after a stroke: new perspectives of locomotor rehabilitation”, *medecine/sciences*, s. 994–998, 2003. DOI: 10.1051/medsci/20031910994.
- [80] S. Mau, “What is the Kalman filter and How can it be used for Data Fusion?”, Carnegie Mellon University, s. 1–10, 2005.
- [81] T. B. Moeslund, “Interacting with a virtual world through motion capture”, w *Virtual Interaction: Interaction in Virtual Inhabited 3D Worlds*, L. Qvortrup, red., London: Springer London, 2001, s. 221–234, ISBN: 978-1-4471-3698-9. DOI: 10.1007/978-1-4471-3698-9_11.
- [82] T. B. Moeslund i E. Granum, “A survey of computer vision-based human motion capture”, *Computer Vision and Image Understanding*, s. 231–268, 2001. DOI: 10.1006/cviu.2000.0897.
- [83] F. Montorsi, F. Pancaldi i G. M. Vitetta, “Design and implementation of an inertial navigation system for pedestrians based on a low-cost MEMS IMU”, *2013 IEEE International Conference on Communications Workshops (ICC)*, s. 57–61, 2013. DOI: 10.1109/ICCW.2013.6649201.
- [84] L. Mündermann, S. Corazza i T. P. Andriacchi, “The evolution of methods for the capture of human movement leading to markerless

- motion capture for biomechanical applications”, *Journaltitle of NeuroEngineering and Rehabilitation*, s. 1–6, 2006. DOI: 10.1186/1743-0003-3-6.
- [85] J. Neville, A. Wixted, D. Rowlands i D. James, “Accelerometers: an underutilized resource in sports monitoring”, *2010 Sixth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, s. 287–290, 2010. DOI: 10.1109/ISSNIP.2010.5706766.
 - [86] S. Nirjon i J. A. Stankovic, “Kinsight: localizing and tracking household objects using depth-camera sensors”, *2012 IEEE 8th International Conference on Distributed Computing in Sensor Systems*, s. 67–74, 2012. DOI: 10.1109/DCOSS.2012.27.
 - [87] S. Noiumkar i S. Tirakoat, “Use of optical motion capture in sports science: a case study of golf swing”, *2013 International Conference on Informatics and Creative Multimedia*, s. 310–313, 2013. DOI: 10.1109/ICICM.2013.58.
 - [88] C. Pandey i S. Nemade, “Enhancement of the speech quality by the implementation of second order fast adaptive Kalman filter algorithm”, *2014 Annu. IEEE India Conf.*, s. 1–6, 2014. DOI: 10.1109/INDICON.2014.7030515.
 - [89] J. Paszek, “Analiza błędów losowych czujników bezwładnościowych przy pomocy metod wariancyjnych”, *Przegląd Elektrotechniczny*, s. 64–69, 2016. DOI: 10.15199/48.2016.01.14.
 - [90] R. L. Pio, “Euler angle transformations”, *IEEE Transactions on Automatic Control*, s. 707–715, 1966. DOI: 10.1109/TAC.1966.1098430.
 - [91] D. J. Reinkensmeyer i M. L. Boninger, “Technologies and combination therapies for enhancing movement training for people with a disability”, *Journal of NeuroEngineering and Rehabilitation*, s. 1–17, 2012. DOI: 10.1186/1743-0003-9-17.
 - [92] C. Reinschmidt, A. J. Van Den Bogert, B. M. Nigg, A. Lundberg i N. Murphy, “Effect of skin movement on the analysis of skeletal knee joint motion during running”, *Journal of Biomechanics*, s. 729–732, 1997. DOI: 10.1016/S0021-9290(97)00001-8.
 - [93] D. Roetenberg, H. Luinge i P. Slycke, “Xsens MVN: full 6DOF human motion tracking using miniature inertial sensors”, Xsens Motion Technologies BV, s. 1–9, 2013. adr.: <https://goo.gl/GAmb9b>.
 - [94] B. Rosenhahn, C. Schmalz, T. Brox, J. Weickert, D. Cremers i H.-P. Seidel, “Markerless motion capture of man-machine interaction”, *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, s. 1–8, 2008.

- [95] D. Rzeszotarski, P. Strumiłło, P. Pełczyński, B. Więcek i A. Lorenc, “System obrazowania stereoskopowego sekwencji scen trójwymiarowych”, *Elektronika : prace naukowe*, s. 165–184, 2005.
- [96] A. M. Sabatini, “Kalman-Filter-based orientation determination using inertial/magnetic sensors: observability analysis and performance evaluation”, *Sensors*, s. 9182–9206, 2011. DOI: [10.3390/s111009182](https://doi.org/10.3390/s111009182).
- [97] J. Sasiadek i P. Hartana, “Sensor data fusion using Kalman filter”, *Information Fusion, 2000. FUSION 2000. Proceedings of the Third International Conference on*, s. 941–952, 2000. DOI: [10.1109/IFIC.2000.859866](https://doi.org/10.1109/IFIC.2000.859866).
- [98] M. Sati, J. de Guise, S. Larouche i G. Drouin, “Quantitative assessment of skin-bone movement at the knee”, *The Knee*, s. 121–138, 1996. DOI: [10.1016/0968-0160\(96\)00210-4](https://doi.org/10.1016/0968-0160(96)00210-4).
- [99] J. Schmidt, J. Fritsch i B. Kwolek, “Kernel particle filter for real-time 3D body tracking in monocular color images”, *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, s. 567–572, 2006. DOI: [10.1109/FGR.2006.69](https://doi.org/10.1109/FGR.2006.69).
- [100] Y. Schroder, A. Scholz, K. Berger, K. Ruhl, S. Guthe i M. Magnor, “Multiple Kinect studies”, Technische Universität Braunschweig, Braunschweig, s. 1–30, 2011.
- [101] D. Serby, E. K. Meier i L. V. Gool, “Probabilistic object tracking using multiple features”, *Proceedings of the 17th International Conference on Pattern Recognition*, s. 1–10, 2004. DOI: [10.1109/ICPR.2004.1334091](https://doi.org/10.1109/ICPR.2004.1334091).
- [102] N. El-Sheimy, S. Nassar i A. Noureldin, “Wavelet de-noising for IMU alignment”, *IEEE Aerospace and Electronic Systems Magazine*, s. 32–39, 2004. DOI: [10.1109/MAES.2004.1365016](https://doi.org/10.1109/MAES.2004.1365016).
- [103] N. El-Sheimy, H. Hou i X. Niu, “Analysis and modeling of inertial sensors using Allan variance”, *IEEE Transactions on Instrumentation and Measurement*, s. 140–149, 2008. DOI: [10.1109/TIM.2007.908635](https://doi.org/10.1109/TIM.2007.908635).
- [104] Shimin Feng i R. Murray-Smith, “Fusing Kinect sensor and inertial sensors with multi-rate Kalman filter”, *IET Conference on Data Fusion and Target Tracking 2014: Algorithms and Applications*, s. 1–9, 2014. DOI: [10.1049/cp.2014.0527](https://doi.org/10.1049/cp.2014.0527).
- [105] A. Shpunt, “Depth mapping using multi-beam illumination”, U.S. Patent: 20 100 020 078, 2010. adr.: <https://goo.gl/ViFu8B>.
- [106] A. Shpunt, T. Petah, Z. Zalevsky i H. Rosh, “Depth-varying light fields for three dimensional sensing”, U.S. Patent: 20 080 106 746, 2008. adr.: <https://goo.gl/RVwHfU>.

- [107] L. Sigal, A. O. Balan i M. J. Black, “HumanEva: synchronized video and motion capture dataset and baseline algorithm for evaluation of articulated human motion”, *International Journal of Computer Vision*, s. 4–27, 2010. DOI: 10.1007/s11263-009-0273-6.
- [108] G. L. Smith, S. F. Schmidt i L. A. McGee, “Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle”, National Aeronautics i Space Administration, s. 1–27, 1962.
- [109] A. Smyth i M. Wu, “Multi-rate Kalman filtering for the data fusion of displacement and acceleration response measurements in dynamic system monitoring”, *Mechanical Systems and Signal Processing*, s. 706–723, 2007. DOI: 10.1016/j.ymssp.2006.03.005.
- [110] D. Sochacka, “Zastosowanie wariancji Allana do oceny wyników pomiarów napięcia wzorców napięcia z diodami Zenera”, *Materyły XXXVI Międzyuczelnianej Konferencji Metrologów MKM'04*, s. 309–316, 2004.
- [111] J. Suarez i R. R. Murphy, “Using the Kinect for search and rescue robotics”, *2012 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, s. 1–2, 2012. DOI: 10.1109/SSRR.2012.6523918.
- [112] A. Sundaresan, “Towards markerless motion capture: model estimation, initialization and tracking”, prac. dokt., University of Maryland, s. 1–127, 2007. adr.: <https://goo.gl/8Ym5bs>.
- [113] A. Sundaresan i R. Chellappa, “Markerless motion capture using multiple cameras”, *Computer Vision for Interactive and Intelligent Environment (CVIIE'05)*, s. 15–26, 2005. DOI: 10.1109/CVIIE.2005.13.
- [114] E. Surer, A. Cereatti, E. Grossi i U. D. Croce, “A markerless estimation of the ankle–foot complex 2D kinematics during stance”, *Gait and Posture*, s. 532–537, 2011. DOI: 10.1016/j.gaitpost.2011.01.003.
- [115] S. Tadano, R. Takeda i H. Miyagawa, “Three dimensional gait analysis using wearable acceleration and gyro sensors based on quaternion calculations”, s. 9321–9343, 2013. DOI: 10.3390/s130709321.
- [116] Y. Tian, X. Meng, D. Tao, D. Liu i C. Feng, “Upper limb motion tracking with the integration of IMU and Kinect”, *Neurocomputing*, s. 207–218, 2015. DOI: 10.1016/j.neucom.2015.01.071.
- [117] M. Vanegas i L. Stirling, “Characterization of inertial measurement unit placement on the human body upon repeated donnings”, *2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks, BSN 2015*, 2015. DOI: 10.1109/BSN.2015.7299398.

- [118] C. Veenman, M. Reinders i E. Backer, “Resolving motion correspondence for densely moving points”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, s. 54–72, 2001. DOI: [10.1109/34.899946](https://doi.org/10.1109/34.899946).
- [119] Vicon Corporation, “Plug-in-Gait modelling instructions. ”, *Vicon{R} Manual*, s. 1–612, 2002.
- [120] Vicon Corporation, “Plug-in-Gait Marker Placement”, *Mocap*, s. 1–4, 2006.
- [121] E. A. Wan i R. Van Der Merwe, “The unscented Kalman filter for nonlinear estimation”, *Proc. IEEE 2000 Adapt. Syst. Signal Process. Commun. Control Symp.*, s. 153–158, 2000. DOI: [10.1109/ASSPCC.2000.882463](https://doi.org/10.1109/ASSPCC.2000.882463).
- [122] E. A. Wan i R. van der Merwe, “The unsented Kalman filter”, w *Kalman Filtering and Neural Networks*, Wiley, rozd. 7, s. 221–269. 2001.
- [123] L. Wang, W. Hu i T. Tan, “Recent developments in human motion analysis”, *Pattern Recognition*, s. 585–601, 2003. DOI: [10.1016/S0031-3203\(02\)00100-0](https://doi.org/10.1016/S0031-3203(02)00100-0).
- [124] G. Welch i G. Bishop, “An Introduction to the Kalman Filter”, *SIGGRAPH 2001 Course Pack*, s. 1–82, 2001. DOI: [10.1.1.117.6808](https://doi.org/10.1.1.117.6808).
- [125] M. Windolf, N. Gotzen i M. Morlock, “Systematic accuracy and precision analysis of video motion capturing systems-exemplified on the Vicon-460 system.”, *Journal of Biomechanics*, s. 2776–80, 2008. DOI: [10.1016/j.jbiomech.2008.06.024](https://doi.org/10.1016/j.jbiomech.2008.06.024).
- [126] S. Woolven i D. Reid, “IMU noise evaluation for attitude determination and stabilization in ground and airborne applications”, *Proceedings of 1994 IEEE Position, Location and Navigation Symposium - PLANS'94*, s. 817–822, 1994. DOI: [10.1109/PLANS.1994.303396](https://doi.org/10.1109/PLANS.1994.303396).
- [127] J. R. Wright, “GPS composite clock analysis”, *Proceedings of the IEEE International Frequency Control Symposium and Exposition*, s. 523–528, 2007. DOI: [10.1109/FREQ.2007.4319128](https://doi.org/10.1109/FREQ.2007.4319128).
- [128] A. Yilmaz, X. Li i M. Shah, “Contour-based object tracking with occlusion handling in video acquired using mobile cameras”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, s. 1531–1536, 2004. DOI: [10.1109/TPAMI.2004.96](https://doi.org/10.1109/TPAMI.2004.96).
- [129] V. M. Zatsiorsky i V. Seluyanov, “The mass and inertia characteristics of the main segments of the human body”, *Biomechanics VIII-B: Proceedings of the Eighth International Congress of Biomechanics*, H. Matsui i K. Kobayashi, red., s. 1152–1159, 1983.

- [130] S. Zhongbin, Y. Lei i K. Qingming, “Quaternion Kalman filter design based on mems sensors”, *The 7th International Conference on Control and Automation (CA 2014)*, A. Stoica, B.-H. Kang i T.-h. Kim, red., s. 93–97, 2014. DOI: [10.14257/astl.2014.76.20](https://doi.org/10.14257/astl.2014.76.20).
- [131] H. Zhou, H. Hu i N. Harris, “Wearable inertial sensors for arm motion tracking in home-based rehabilitation”, *Proceedings of the 9th International Conference on Intelligent Autonomous Systems*, s. 930–937, 2006.
- [132] H. Zhou i H. Hu, “Inertial motion tracking of human arm movements in stroke rehabilitation”, *2005 IEEE International Conference Mechatronics and Automation*, s. 1306–1311, 2005. DOI: [10.1109/ICMA.2005.1626742](https://doi.org/10.1109/ICMA.2005.1626742).

Spis rysunków

1.1	Podział systemów śledzenia ruchu ze względu na technologię pobierania danych (źródło: opracowanie własne)	12
2.1	Rysunek przedstawiający konfigurację sprzętową systemu Vicon dostępnego w Centrum Technologii Informatycznych Politechniki Łódzkiej	21
2.2	Zestaw markerów pasywnych w systemie Vicon umieszczonych na ciele aktora w trakcie sesji śledzenia ruchu	22
2.3	Przykładowe pasywne markery wykorzystywane w optycznych systemach śledzenia ruchu	22
2.4	Układ współrzędnych dla czujników inercyjnych i magnetycznych . . .	29
2.5	Przykładowa konfiguracja systemu LPM firmy Inmotio	31
2.6	Rozmieszczenie środków mas poszczególnych części ciała człowieka według propozycji Zatsiorskiego i Seluyanova oraz de Leva	32
2.7	Egzoszkielet dla systemu śledzenia ruchu Gypsy	34
2.8	Uproszczony hierarchiczny model szkieletowy człowieka wraz z diagramem przedstawiającym hierarchię poszczególnych elementów .	36
2.9	Rotacja wewnętrzna (pronacja) i zewnętrzna (supinacja) dloni	37
2.10	Model szkieletowy człowieka wykorzystywany w kontrolerze Microsoft Kinect	39
2.11	Schemat umieszczania markerów na ciele człowieka w systemie śledzenia ruchu firmy Vicon	40
2.12	Wyznaczenie położenia kości miedniczej na potrzeby modelu szkieletowego, na podstawie położenia markerów w systemie Vicon . .	41
2.13	Przykładowy schemat umieszczania markerów inercyjnych na potrzeby śledzenia ruchu ręki	43
2.14	Schemat przedstawiający kolejne kroki wyznaczania modelu szkieletowego na podstawie mapy głębi	45
2.15	Układ współrzędnych wykorzystywany w kontrolerze Microsoft Kinect	46
2.16	Kąty Eulera względem przyjętego układu współrzędnych: relatywnego (a) oraz absolutnego (b)	47
2.17	Obiekt znajdujący się w orientacji z blokadą przegubu	48
2.18	Uproszczony schemat budowy kontrolera Microsoft Kinect v.1	52
2.19	Poziomy i pionowy zakres pracy kontrolera Microsoft Kinect v.1	53

2.20	Scena oświetlona promieniami podczerwonymi	54
2.21	Wzorzec oświetlenia sceny przez kontroler Kinect	54
2.22	Mapa głębi ręki w różnych warunkach oświetleniowych	56
2.23	Kąt obrotu (α) użytkownika pomiędzy linią barków a płaszczyzną obserwacji Kinecta	58
2.24	Pomiary zależne od kąta obrotu względem Kinecta (źródło: badania własne)	60
2.25	Dokładność pomiaru głębi w zależności od odległości śledzonej postaci od kontrolera Kinect	62
2.26	Układ współrzędnych wykorzystywany w czujnikach MPU-6050 (źródło: opracowanie własne)	63
2.27	Struktura wewnętrzna akcelerometru w technologii MEMS	65
2.28	Struktura wewnętrzna żyroskopu w technologii MEMS	67
2.29	Pomiar przyspieszenia ziemskiego w przedziale temperatur 10-50°C	69
2.30	Szacowanie położenia stawu z użyciem klasycznego filtra Kalmana i filtru Kalmana dostosowanego do różnych częstotliwości	75
3.1	Algorytm przetwarzania danych w autorskiej metodzie łączenia danych (źródło: badania własne)	80
3.2	Przykład hierarchii stawów dla modelu szkieletowego ręki (źródło: opracowanie własne)	81
3.3	Schemat budowy urządzenia pomiarowego opartego o komputer Arduino (źródło: badania własne). IMU1 i IMU2 – Moduły inercyjne, RTC – moduł zagara oparty o komponent DS3230, BT – moduł Bluetooth oparty o komponent HC-05, LCD – ekran ciekłokrystaliczny	84
3.4	Diagram przedstawiający kalibrację czujników inercyjnych (źródło: badania własne)	87
3.5	Wykres przedstawiający uśrednione wartości oszacowania kąta obrotu wokół osi X i Y za pomocą modułu inercyjnego (źródło: badania własne)	89
3.6	Wykres przedstawiający dryf oszacowania kąta obrotu modułu inercyjnego wokół osi Z (źródło: badania własne)	90
3.7	Wykres przedstawiający oszacowanie położenia stawu łokciowego za pomocą kontrolera Kinect na podstawie sygnału zaszumionego oraz odszumionego	91
3.8	Wykres przedstawiający średni błąd pozycjonowania stawów za pomocą kontrolera Kinect w zależności od współczynnika filtracji α (źródło: badania własne)	92
3.9	Oznaczenie układu współrzędnych XYZ wraz z obrotami ϕ, θ, ψ w odniesieniu do ręki człowieka (źródło: opracowanie własne)	97
3.10	Odchylenie standardowe estymacji kąta obrotu postaci względem Kinecta (źródło: badania własne)	99
3.11	Wykres wartości szumu w oszacowaniach kontrolera Kinect	100

3.12	Diagram przedstawiający kolejne etapy działania autorskiej metody łączenia danych IMU i Kinecta (źródło: badania własne)	101
4.1	Schemat rozmieszczenia elementów sceny w laboratorium (źródło: opracowanie własne)	104
4.2	Schemat romieszczenia markerów na ręce zgodny z zaleceniami systemu Vicon (źródło: opracowanie własne)	104
4.3	Moduł inercyjny umieszczony na przedramieniu autora(źródło: badania własne)	105
4.4	Ćwiczenia wykonywane w ramach eksperymentu (źródło: badania własne)	106
4.5	Wykres przedstawiający położenie stawu nadgarstkowego w osi X w ćwiczeniu 1 (źródło: badania własne)	110
4.6	Wykres przedstawiający położenie stawu nadgarstkowego w osi Y w ćwiczeniu 1 (źródło: badania własne)	111
4.7	Wykres przedstawiający położenie stawu nadgarstkowego w osi Z w ćwiczeniu 1 (źródło: badania własne)	111
4.8	Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 1 (źródło: badania własne)	111
4.9	Wizualizacja ruchu ręki w ćwiczeniu 1	112
4.10	Wykres przedstawiający podążanie wyznaczania pozycji stawu za pomiarami Kinecta w ćwiczeniu 3 (źródło: badania własne)	114
4.11	Wizualizacja ruchu ręki w ćwiczeniu 2	114
4.12	Wizualizacja ruchu ręki w ćwiczeniu 3	115
4.13	Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 2 (źródło: badania własne)	115
4.14	Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 3 (źródło: badania własne)	116
4.15	Wizualizacja ruchu ręki w ćwiczeniu 4	117
4.16	Wykres położenia stawu łokciowego w osi Z w ćwiczeniu 4	117
4.17	Wykres położenia stawu nadgarstkowego w osi Z w ćwiczeniu 4	118
4.18	Wykres przedstawiający oszacowanie wartości kąta zgięcia ręki w łokciu (β) w ćwiczeniu 4 (źródło: badania własne)	118
4.19	Średni błąd średniokwadratowy wyznaczania pozycji łokcia (źródło: badania własne)	119
4.20	Średni błąd średniokwadratowy wyznaczania pozycji nadgarstka (źródło: badania własne)	120
4.21	Średni błąd średniokwadratowy wyznaczania kąta zgięcia ręki w łokciu β (źródło: badania własne)	120
A.1	Schemat działania filtra Kalmana (źródło: opracowanie własne)	133
A.2	Schemat działania filtra komplementarnego (źródło: opracowanie własne)	136

A.3	Diagram blokowy przedstawiający filtr Madgwicka dla czujników inercyjnych (źródło: opracowanie własne)	140
B.1	Obszary szumów na wykresie odchylenia Allana	144

Spis tablic

2.1	Porównanie cech charakterystycznych systemu <i>Organic Motion</i> oraz systemu wbudowanego w kontroler Microsoft Kinect v1	25
2.2	Błąd estymacji położenia w zależności od błędu pomiarów akcelerometru	29
2.3	Konwencje nazewnicze kątów Eulera	47
2.4	Zestawienie podstawowych parametrów pracy układu MPU-6050	63
3.1	Współczynniki konwersji bezpośrednich pomiarów akcelerometru (a) i żyroskopu (b) w zależności od przyjętego zakresu pracy	84
4.1	Średni błąd szacowania \overline{RMSE} dla ćwiczenia nr 1	109
4.2	Średni błąd szacowania \overline{RMSE} dla ćwiczenia nr 2	113
4.3	Średni błąd szacowania \overline{RMSE} dla ćwiczenia nr 3	113
4.4	Średni błąd szacowania \overline{RMSE} dla ćwiczenia nr 4	116