



**GDAŃSK UNIVERSITY  
OF TECHNOLOGY**

FACULTY OF ELECTRONICS, TELECOMMUNICATIONS  
AND INFORMATICS



Imię i nazwisko autora rozprawy: Tomasz Merta  
Dyscyplina naukowa: automatyka i robotyka

## **ROZPRAWA DOKTORSKA**

Tytuł rozprawy w języku polskim: Odwzorowanie obiektów ograniczonego środowiska na trójwymiarowej mapie cyfrowej z wykorzystaniem robotów mobilnych zaopatrzonych w stereowizję

Tytuł rozprawy w języku angielskim: Mapping of objects of a limited environment on a three-dimensional digital map with the use of mobile robots equipped with stereovision

Promotor

*podpis*

prof. dr hab. inż. Zdzisław Kowalczyk

Gdańsk, rok 2019





POLITECHNIKA GDAŃSKA

WYDZIAŁ ELEKTRONIKI  
TELEKOMUNIKACJI I INFORMATYKI



---

*Katedra Systemów Decyzyjnych i Robotyki*

**Tomasz Merta**

---

ODWZOROWANIE OBIEKTÓW OGRANICZONEGO  
ŚRODOWISKA NA TRÓJWYMIAROWEJ MAPIE  
CYFROWEJ Z WYKORZYSTANIEM ROBOTÓW  
MOBILNYCH ZAOPATRZONYCH W STEREOWIZJĘ

---

Rozprawa Doktorska

Promotor:  
prof. dr hab. inż. Zdzisław  
Kowalczyk  
Wydział Elektroniki  
Telekomunikacji  
i Informatyki  
Politechnika Gdańska

Gdańsk, 10 lipca 2019 r.





## *Streszczenie*

Problem podjęty w rozprawie dotyczy trójwymiarowego odwzorowania otoczenia przez roboty mobilne. Wpisuje się on w zagadnienie SLAM – jednoczesnego określania położenia i tworzenia mapy. W pracy skupiono się na wybranej realizacji mapowania opartej na stereowizji (konkretnego zagadnienia SLAM). Duży nacisk położono także na warunki konieczne do zapewnienia prawidłowej stereowizji oraz na sprzętowy układ eksperymentu, który umożliwia odwzorowanie ograniczonego otoczenia.

Stereowizja dostarcza danych przestrzennych na podstawie obrazów z dwóch kamer. W typowym systemie stereowizyjnym zależności między kamerami są stałe. Natomiast w pracy założono zmienną bazę stereo, więc parę obrazów uzyskuje się, gdy robot znajduje się w różnych pozycjach. Mechanizm ten można zaadaptować także dla grupy robotów, wówczas każdy obraz jest uzyskiwany z kamery innego robota. W obu przypadkach każda para obrazów ma inną bazę stereowizyjną w wyniku akwizycji danych z kamer znajdujących się w różnych pozycjach w przestrzeni.

Ponadto systemy wizyjne z dwiema kamerami są odpowiednie do wyświetlania obrazu stereoskopowego operatorowi robota (człowiekowi). Wprowadzając odpowiednie ograniczenie różnic między obrazami (które mogą powstać przy dużych odległościach między kamerami) umożliwiono wizualizację poprawnego obrazu 3D za pomocą dostępnej techniki stereoskopowej.

Istotnym problemem jest także dokładna estymacja położenia i ułożenia robota, ponieważ pozycja kamery zmienia się wraz z bazą. Ze względu na trójwymiarową reprezentację środowiska, położenie i ułożenie wyznaczone są także w trzech wymiarach.

Estymacja położenia i ułożenia robota zaproponowana w pracy opiera się na fuzji danych inercyjnych i filtracji Kalmana. Odpowiednia filtracja pozwala estymować pozycję robota dzięki zidentyfikowanemu modelowi dynamiki robota i modelowi akcelerometru, a także dzięki danym akcelerometrycznym i sygnałowi sterującemu robotem. Estymacja jest użyteczna mimo istnienia złożonych i silnych szumów pomiarowych. Metoda została zweryfikowana w warunkach rzeczywistych przy użyciu laboratoryjnej platformy sprzętowej. Dodatkowo parametry modelu testowej platformy robota wyznaczono autorską metodą estymacji opartą na przetwarzaniu danych z kamery CMOS (z wykrywaniem kolorowego znacznika).

Na podstawie rozważań analitycznych wyznaczono wrażliwość stereowizji na błędy położenia i ułożenia robota, co pozwoliło określić, w jaki sposób błędy te propagują się na głębię widzenia. Ewaluację wrażliwości stereowizji przeprowadzono w środowisku wirtualnym i rzeczywistym uwzględniając dokładność mapowania ze zmienną bazą stereo. Wyznaczono wirtualną mapę głębi (jako mapę odniesienia) dla wybranych 30 wirtualnych scen a także odrębne cztery mapy głębi dla wybranych baz. Dzięki porównaniu tych map z mierzonymi mapami głębi określono błąd odwzorowania. Na tej podstawie zweryfikowano ilościowo, w jakim stopniu wykorzystanie zmiennej bazy poprawia docelowe odwzorowanie otoczenia. Zrealizowane testy rzeczywiste wykazały, że opracowany system ze zmienną bazą jest użyteczny przy mapowaniu otoczenia robota.

## ***Abstract***

The dissertation concerns the problem of three-dimensional mapping of the environment by mobile robots. It is a part of the SLAM issue, which simultaneously determines the robots location and creates a map. The work is related to selected implementations of mapping that are based on stereovision (an issue of SLAM). We place a large emphasis on the necessary conditions to ensure correct stereovision and the hardware experimental setup that is sufficient for the mapping process of a limited environment.

Stereovision provides spatial data by processing images from two cameras. In a typical stereovision system relations between cameras are fixed. In this work we consider the method of variable baseline, so a pair of images is determined when the robot takes different positions. Such a mechanism can also be implemented by a group of robots, when each image is obtained from the camera of another robot. In both cases, each pair of images has a different stereovision baseline, as a result of acquiring data from cameras which are in different positions (in space).

Moreover, vision systems with two cameras are suitable for showing the stereoscopic image to a robot operator (human). Introducing appropriate limitation of differences between images (which can arise for long distances between cameras), it is possible to properly visualize 3D images using the available stereoscopic technique.

Precise estimation of the robot's position and pose is also an important problem, because each new position of the camera changes the baseline. Due to the three-dimensional representation of environment, the position and pose can also be determined in three dimensions.

Estimation of the position and pose of the robot is obtained by the fusion of inertial data and Kalman filtering. Appropriate filtration allows us to estimate the robot's position using both identified dynamic model of the robot and the accelerometer model, as well as accelerometer data and a robot control signal. The estimation algorithm works properly despite the existence of complex and strong measurement noise. The method has been verified in real conditions using a laboratory hardware platform. Moreover, the parameters of the model of the hardware robot platform have been determined using an original estimation method founded on processing data from the CMOS camera (by detecting a colored marker).

On the basis of analytical considerations, we have determined the sensitivity of stereovision to the pose and position errors. As a consequence, we have found, how these errors propagate to depth of vision. Further evaluation of stereovision sensitivity was carried out in virtual and real environments, including the accuracy of mapping with variable baseline. We determined a virtual (reference) depth map for selected 30 virtual scenes, and other four depth maps taken at selected baselines. The comparison of these maps with the measured depth maps gave us the value of mapping error. Therefore, we also verified the improvement of the environment mapping by using the variable baseline. The real tests have showed that the developed system with a variable baseline is useful for mapping the robotic environment.



*”Żaden robot nie został zaprojektowany, aby być świadomy tego, co robi; ale w większości przypadków, my też nie jesteśmy tego świadomi.”*

John McCarthy





---

# SPIS TREŚCI

<b>1</b>	<b>Wprowadzenie</b>	<b>1</b>
1.1	Cel pracy . . . . .	2
1.2	Motywacja i dorobek autora . . . . .	3
1.3	Tezy . . . . .	4
1.4	Omówienie zakresu tematyki rozdziałów . . . . .	5
<b>2</b>	<b>Przegląd rozwiązań zrobotyzowanych</b>	<b>7</b>
2.1	Rodzaje robotów . . . . .	7
2.2	Roboty mobilne . . . . .	8
2.2.1	Roboty latające . . . . .	9
2.2.2	Roboty wodne . . . . .	12
2.2.3	Roboty lądowe . . . . .	13
2.3	Wyposażenie robotów mobilnych . . . . .	17
2.3.1	Napęd robota kołowego . . . . .	17
2.3.2	Sensory . . . . .	19
2.4	Mapowanie . . . . .	24
2.4.1	Podział zagadnień SLAM . . . . .	24
2.4.2	Mapowanie oparte na wizji . . . . .	26
2.4.3	Wybrane realizacje mapowania 3D realizowane za pomocą robotów mobilnych . . . . .	28
2.5	Podsumowanie . . . . .	30
<b>3</b>	<b>Teoretyczne podstawy pomiarów</b>	<b>31</b>
3.1	Sensory inercyjne i ich charakterystyka . . . . .	32
3.1.1	Pomiar ułożenia . . . . .	32
3.1.2	Pomiar położenia . . . . .	33
3.2	Transformacja układu odniesienia . . . . .	33
3.2.1	Rotacja i macierze obrotu . . . . .	34
3.2.2	Kwaterniony . . . . .	36
3.2.3	Szum sensorów inercyjnych . . . . .	37
3.3	Model stanowy i estymacja danych . . . . .	38
3.3.1	Model stanowy . . . . .	39
3.3.2	Cechy układu . . . . .	40

3.3.3	Obserwacja stanu . . . . .	41
3.3.4	Filtracja Kalmana . . . . .	42
3.3.5	Algorytm SLAM oparty na filtracji Kalmana . . . . .	44
3.4	Stereowizja . . . . .	46
3.4.1	Współrzędne jednorodne . . . . .	46
3.4.2	Modele kamery . . . . .	47
3.4.3	Parametry kamer . . . . .	49
3.4.4	Geometria epipolarna . . . . .	51
3.4.5	Wyznaczanie głębi . . . . .	52
3.4.6	Mapowanie 3D . . . . .	54
3.5	Podsumowanie . . . . .	55
<b>4</b>	<b>System VisRobot</b>	
	<b>ze zmienną bazą stereowizyjną</b>	<b>57</b>
4.1	Ogólna charakterystyka systemu . . . . .	57
4.2	Wyznaczanie ułożenia robota . . . . .	59
4.2.1	Kalibracja sensorów inercyjnych . . . . .	60
4.2.2	Fuzja danych inercyjnych . . . . .	61
4.3	Wyznaczanie położenia robota . . . . .	63
4.3.1	Model akcelerometru . . . . .	64
4.3.2	Model kinematyki robota mobilnego . . . . .	66
4.3.3	Model dynamiki robota mobilnego . . . . .	69
4.4	Wykorzystanie zmiennej bazy stereoskopowej . . . . .	72
4.4.1	Wyznaczanie dysparycji . . . . .	73
4.4.2	Błąd mapy dysparycji . . . . .	74
4.4.3	Fuzja map dysparycji i ich wiarygodność . . . . .	78
4.5	Wizualizacja i ocena głębi . . . . .	82
4.5.1	Chmura punktów . . . . .	83
4.5.2	Obraz 3D . . . . .	83
4.5.3	Ocena jakości głębi obrazu . . . . .	85
4.6	Podsumowanie . . . . .	86
<b>5</b>	<b>Symulacje i eksperymenty</b>	<b>87</b>
5.1	Analiza wrażliwości danych akcelerometru . . . . .	88
5.1.1	Symulacja ruchu robota i filtracja pasmowa . . . . .	88
5.1.2	Szum akcelerometru . . . . .	91
5.1.3	Wrażliwość pomiaru położenia . . . . .	93
5.2	Wyznaczanie trajektorii robota . . . . .	95
5.2.1	Testy symulacyjne . . . . .	95
5.2.2	Testy doświadczalne . . . . .	99
5.3	Analiza dokładności mapowania . . . . .	107
5.3.1	Wrażliwość systemu stereowizyjnego dla wirtualnej sceny . . . . .	107
5.3.2	Wrażliwość systemu stereowizyjnego dla sceny rzeczywistej . . . . .	113
5.3.3	Dokładność mapowania wirtualnego otoczenia . . . . .	118
5.3.4	Mapowanie wybranego otoczenia rzeczywistego . . . . .	121
5.4	Podsumowanie . . . . .	126

<b>6</b>	<b>Zakończenie</b>	<b>129</b>
6.1	Zrealizowane założenia . . . . .	129
6.2	Odniesienie do tez pracy . . . . .	130
6.3	Główne osiągnięcia pracy . . . . .	132
6.4	Krytyczna analiza pracy . . . . .	132
	<b>Spis rysunków</b>	<b>135</b>
	<b>Spis tablic</b>	<b>139</b>
	<b>Wykaz ważniejszych skrótów i oznaczeń</b>	<b>141</b>
	<b>Bibliografia</b>	<b>145</b>



---

## WPROWADZENIE

Mapowanie jest naturalnym procesem, który jest nieustannie realizowany przez człowieka i większość innych organizmów żywych. W trakcie codziennych zachowań każdy organizm musi orientować się, gdzie aktualnie się znajduje (np. jak daleko znajduje się od domu lub innego bezpiecznego miejsca). Zapamiętując nowe elementy znajdujące się w otoczeniu lub nowe nieznanne tereny, zwykle nieświadomie, tworzymy mapę. Jednakże proces ten nie jest prosty i wymaga odpowiednich sensorów, złożonej analizy danych, postrzegania obiektów, zestawienia ich z informacją zapisaną w pamięci.

Przez tysiące lat organizmy dostosowywały się do otaczającego je środowiska. Zmienne i zróżnicowane warunki życia sprawiły, że rozmaite zwierzęta i rośliny na drodze ewolucji wykształciły mechanizmy umożliwiające zaawansowaną eksplorację otoczenia. Zwierzęta łowne skutecznie polują dzięki bardzo czułemu powonieniu, ryby mają linię boczną czułą nawet na bardzo małe zmiany ciśnienia, nietoperze doskonale poruszają się w jaskiniach, gdzie nie ma dostępu światła. Człowiek, podobnie jak wiele zwierząt, wykształcił zmysły ułatwiające egzystencję w środowisku. Błądźnik i zmysł równowagi pozwala określić, czy obracamy się lub zmieniamy swoje ułożenia względem ziemi. Słuch, węch czy dotyk umożliwiają odbieranie bodźców z otoczenia. Jednakże to wzrok dostarcza największej ilości informacji o otaczającym świecie i pozwala na dokładne obserwowanie obiektów znajdujących się w środowisku. Dlatego właśnie mapowanie oparte na wizji jest skutecznie realizowane przez wiele organizmów żywych.

Bezpieczne poruszanie się oraz unikanie przeszkód wymagają zatem odpowiedniego orientowania się w terenie. Problem sprowadzić można do określenia, jakie jest położenie ruchomego obiektu (np. człowieka) w relacji z innymi charakterystycznymi obiektami. W tym celu wykonuje się mapowanie, które generalnie polega na określeniu swojego położenia oraz wszystkich charakterystycznych obiektów znajdujących się w otoczeniu. Dopiero, gdy znane są relacje przestrzenne między obiektami można realizować bardziej złożone zadania. Zalicza się do nich obserwowanie i analizowanie otoczenia, wchodzenie w interakcję z nieznanymi obiektami, eksplorowanie nowych obszarów, utrzymywanie odpowiedniego dystansu od obiektów niebezpiecznych, itp.

Robot mobilny, podobnie jak zwierzęta i człowiek, powinien poruszać się bezpiecznie w swoim środowisku, aby zrealizować zamierzoną pracę. Jeżeli robot jest dedykowany do prostego zadania w ściśle ograniczonym środowisku, najczęściej wystarczą proste mechanizmy ruchu (w ograniczonym zakresie). Przykładem takiego ograniczenia może być

przewodnica, po której może poruszać się robot. W przypadku złożonych zadań robot musi być zaopatrzony w odpowiednie sensory i system, który określa pozycję<sup>1</sup> robota w przestrzeni. Implementacja takiego systemu wymaga rozwiązania licznych problemów i zaprojektowania zaawansowanych algorytmów.

## 1.1 Cel pracy

Zamiarem jest opracowanie metody odwzorowania obiektów ograniczonego środowiska na trójwymiarowej mapie cyfrowej, z wykorzystaniem robotów mobilnych zaopatrzonych w stereowizję. Zadanie to sprowadza się do mapowania ograniczonego otoczenia robota mobilnego. Podstawą do tworzenia mapy 3D jest stereowizja, która oparta jest na parze obrazów, które są rejestrowane przez kamery umieszczone na robotach mobilnych.

Przedstawione zagadnienie nazywa się jednoczesnym określaniem położenia i tworzeniem mapy SLAM (ang. *Simultaneous Localization and Mapping*). Problem SLAM zakłada, że budowa mapy i określanie położenia realizowane są jednocześnie, gdyż aktualizowana mapa jest tworzona z wykorzystaniem aktualnego położenia robota, a zmiana położenia jest określana względem mapy. Mimo bardzo ogólnych założeń konkretna realizacja SLAM zależy od rodzaju robota, ograniczeń ruchu, dostępnych sensorów, itp. Ponadto przy rzeczywistych pomiarach realizacja SLAM wymaga rozwiązania dodatkowych zadań, np. fuzji i filtracji danych z różnych sensorów, wyznaczenia modelu ruchu i modelu zbierania pomiarów, łączenia nowych pomiarów do mapy znanej przed pomiarem, zamykanie pętli itp.

W niniejszej pracy skupiamy się na wybranej realizacji mapowania opartej na stereowizji, co jest konkretnym zagadnieniem SLAM. Zatem duży nacisk będzie położony na realizację rozwiązania, które spełnia warunki konieczne do prawidłowej stereowizji w celu odwzorowania ograniczonego otoczenia. Ważnym aspektem przyjętego podejścia jest uwzględnienie sprzętowego układu eksperymentu. Dlatego zakłada się weryfikację rozwiązania nie tylko w środowisku symulacyjnym ale także w warunkach rzeczywistych.

Stereowizja dostarcza danych przestrzennych, podobnie jak para oczu u człowieka. Przesunięcie między parą oczu czy parą kamer skutkuje powstaniem różnic między obrazami. Te zaś pozwalają na wyznaczenie głębi, czyli odległości kamery od obiektu widocznego w kadrze. Odległość między oczami (tzw. baza stereoskopowa) jest (w ogólności) stała. Natomiast odległość między kamerami może być dowolna. Ponadto system wizyjny składający się z 2 kamer pozwala na wyświetlanie obrazu stereoskopowego operatorowi robota (człowiekowi). Przy odpowiednim ograniczeniu różnic między obrazami (które mogą powstać przy dużych odległościach między kamerami) operator powinien ujrzyć poprawny obraz 3D.

W pracy założono, że na stereowizyjne wejście można podać dwa obrazy uzyskane przez robota, znajdującego się w różnych pozycjach. Należy zatem zapewnić, aby położenie i ułożenie robota w każdym przypadku było znane z zadaną dokładnością. Mechanizm ten można zaadaptować także dla grupy robotów, wówczas każdy obraz może pochodzić z kamery innego robota. W obu przypadkach na skutek akwizycji danych z kamer, które znajdują się w różnych pozycjach w przestrzeni, każda para obrazów ma inną bazę ste-

<sup>1</sup>Przez pozycjonowanie (lub lokalizowanie) należy rozumieć określenie położenia i ułożenia w przestrzeni. Opisując położenie podaje się współrzędne danego punktu (np. środka platformy robota), natomiast ułożenie oznacza orientację, jak skierowany jest robot (za pomocą współrzędnych kątowych).

reowizyjną. Zmiana bazy powoduje zmianę mapowanego zakresu, dlatego w pracy rozważamy też metodę wykorzystania zmiennej bazy stereowizyjnej<sup>2</sup> w procesie mapowania ograniczonego środowiska.

Dla prawidłowego działania stereowizji i mapowania otoczenia szczególnie istotne jest dokładne określenie położenia i ułożenia kamery. Ze względu na trójwymiarowe odwzorowanie otoczenia, położenie i ułożenie należy wyznaczyć także w trzech wymiarach. Istniejące systemy oparte są na informacji z czujników odometrycznych, odległościowych, radarowych lub wizyjnych. Dane o względnym położeniu robota obarczone są kumulującymi się błędami. Bezwzględny pomiar położenia z zewnętrznego globalnego systemu zapewnia stały błąd (stałą dokładność), niemniej dane z globalnego systemu nie zawsze są dostępne a ich dokładność wystarczająca. Określenie ułożenia w istniejących rozwiązaniach oparte jest na zewnętrznych znacznikach umieszczonych w otoczeniu lub na sensorach inercyjnych, które nie wymagają ingerowania w otoczenie. W obu przypadkach do określenia ułożenia konieczny jest dedykowany system. Z powyższego wynika, że autonomiczne nawigowanie robota oraz odwzorowanie otoczenia na mapie jest zadaniem trudnym z inżynierskiego punktu widzenia.

Podsumowując, problem odwzorowania otoczenia na mapie 3D z wykorzystaniem systemu wizyjnego robota opartego na układzie kamer ze zmienną szerokością bazy, powinien uwzględniać wiele osobnych problemów badawczych. Należą do nich: określanie położenia i ułożenia robota w przestrzeni, głębi w obrazie oraz wpływu zmienności bazy na wyznaczanie głębi. Dodatkowo wymagane jest opracowanie odpowiedniego algorytmu SLAM, w tym fuzji danych z wielu źródeł. Ze względu na złożoność powyższych problemów oraz w celu zachowania spójności pracy, konieczne jest odpowiednie ograniczenie zakresu badań.

## ***1.2 Motywacja i dorobek autora***

Tematyka pracy ściśle związana jest z zainteresowaniami autora ulokowanymi w dziedzinie przetwarzania obrazu. Wcześniejsze doświadczenia obejmują rozpoznawanie obiektów za pomocą algorytmów przetwarzania obrazu, gdzie różnorodne podejścia nie zaowocowały oczekiwaną wysoką skutecznością rozpoznawania obiektów. Obiekty ograniczone pod względem obrotu i położenia nietrudno jest rozpoznać. Jednak w praktyce obiekty nie zawsze są widoczne w tym samym miejscu w kadrze kamery i nie ustawiają się w pożądanym sposobie. Kamera rejestruje płaski obraz, który jest tylko rzutem przestrzennego obiektu na płaszczyznę. Stąd bierze się potrzeba wspomaganie rozpoznawania płaskiego obrazu za pomocą danych przestrzennych.

Problemy związane z rozpoznawaniem obiektów związane są też z zagadnieniem uczenia. Aby system poprawnie rozpoznawał dany obiekt zwykle potrzebne są liczne zdjęcia (dane uczące) wielu obiektów tego samego typu z różnych stron, przy różnych warunkach oświetleniowych, itp. Robot mobilny pozwala na automatyczne generowanie kolejnych danych uczących podczas obserwacji obiektu z różnych stron. Warunkiem jest zastosowanie odpowiedniej funkcji SLAM, aby robot mógł bezpiecznie obserwować otoczenie. Można zatem uczyć robota rozpoznawania wybranych obiektów bez ręcznego zapewniania dużych baz uczących. Dotyczy to zarówno uczenia i rozpoznawania nowych obiektów, jak i dodawania nowych danych o znanym obiekcie. Po zaopatrzeniu robota w odpowiednie

---

<sup>2</sup>W kontekście odwzorowania otoczenia, pojęcia kamery stereoskopowej i stereowizyjnej można używać zamiennie.

człony wykonawcze jest szansa, że taki robot będzie mógł zbierać dane o świecie poprzez obserwację kolejnych obiektów i interakcję z nimi, podobnie jak robią to organizmy żywe.

Doświadczenie związane z przetwarzaniem obrazów przyczyniło się do wyboru tematyki, związanej z uwzględnieniem głębi w przetwarzaniu płaskich obrazów. Problem rozmiaru bazy uczącej przy rozpoznawaniu obiektu stanowił motywację do skupienia się na odwzorowaniu otoczenia, aby w przyszłości można było zautomatyzować proces poszerzania bazy uczącej. Dlatego opracowując zarysy pracy uwzględniono zagadnienie pozyskiwania danych przestrzennych o otoczeniu.

W badaniach nad metodami określania pozycji robota szczególną rolę odgrywa zagadnienie całkowania danych w kontekście testowania danych z akcelerometru. Ponieważ lektura literatury naukowej nie dała jednoznacznej odpowiedzi, podjęto badania w celu wyjaśnienia, czy dane te są użyteczne w procesie wyznaczania położenia.

W ramach badań własnych zrealizowano kilka ważnych zadań, które opisano w publikacjach w materiałach indeksowanych oraz zaprezentowano na uznanych konferencjach międzynarodowych:

- Kowalczyk Z., Merta T. (2013): Stereovision system for a mobile robot, 18th International Conference on Methods and Models in Automation and Robotics, pp 794-799, Międzyzdroje, Polska.
- Kowalczyk Z., Merta T. (2014): Wizualizacja obrazu stereowizyjnego w systemie VISROBOT. *Pomiary, Automatyka, Kontrola*, vol. 10, no. 1., pp 803-808.
- Kowalczyk Z., Merta T. (2015): Evaluation of position estimation based on accelerometer data. *International Workshop on Robot Motion and Control*, pp 246-251, Poznań, Polska.
- Kowalczyk Z., Merta T. (2016): Evaluating the Position of a Mobile Robot Using Accelerometer Data. *Advanced and Intelligent Computations in Diagnosis and Control*, vol. 386, pp 131-143, Springer International Publishing, Londyn, UK.
- Kowalczyk Z., Merta T. (2016): Three-dimensional mapping for data collected using variable stereo baseline. *21st International Conference on Methods and Models in Automation and Robotics*, pp 1082-1087, Międzyzdroje, Polska.
- Kowalczyk Z., Merta T. (2017): Estimation of DC motor parameters using a simple CMOS camera. *22nd International Conference on Methods and Models in Automation and Robotics*, pp 872-877, Międzyzdroje, Polska.

### 1.3 Tezy

Dla postawienia wymienionych wyżej celów pracy oraz krótkiej charakterystyki tematyki rozprawy doktorskiej zdefiniowano główną tezę badawczą oraz dwie tezy pomocnicze.

**Teza:** Odpowiednie przetworzenie informacji z sensorów inercyjnych oraz kamer systemu stereoskopowej wizji, umieszczonych na robotach mobilnych, pozwala adekwatnie odwzorować otoczenie w postaci mapy 3D

**Subteza 1:** Istotnym czynnikiem poprawy jakości w systemie stereowizyjnym jest wykorzystanie obrazów z kamer przy zmiennej bazie stereoskopowej

**Subteza 2:** Istnieje optymalne ustawienie bazy obrazów stereoskopowych pod względem jakości oceny głębi obrazu

Z tezą główną wiąże się zamiar weryfikacji koncepcji użycia kamer stereoskopowych montowanych na robotach mobilnych, tj. sprawdzić, racjonalność takiego podejścia. Przy



czym 'adekwatnie' - oznacza odwzorowanie, które może być wykorzystane przez robota (lub człowieka) poruszającego się w przestrzeni rzeczywistej, wirtualnej, rozszerzonej. Występujące w tezie głównej sensory inercyjne służą wyznaczaniu położenia i ułożenia. W literaturze można spotkać rozwiązania, w których docelowa nawigacja oparta jest na stereowizji, jednak dotyczą one przede wszystkim aplikacji SLAM (Shimadu *et al.*, 2012; Bai *et al.*, 2010; Mokri i Yasamin, 2005; Lu-Fang *et al.*, 2007) dostosowanych do specjalistycznych zastosowań, wykorzystujących stałą bazę (oraz fabrycznie skalibrowane kamery). W odróżnieniu od powyższych rozwiązań, metoda oparta na sensorach inercyjnych prezentuje nowatorskie podejście. Kolejną innowacją tej dysertacji jest zastosowanie zmiennej bazy, czego dotyczą tezy pomocnicze.

Subteza 1 ma na celu ewaluację metody zmiennej bazy przy mapowaniu opartym na stereowizji. W literaturze można znaleźć nieliczne przykłady użycia zmiennej bazy do odwzorowania otoczenia (Gallup *et al.*, 2008; Jeon *et al.*, 2001). Dotyczą one jednak generalnie stereowizji oraz nie rozważają kompletnego zagadnienia mapowania w oparciu o możliwości robotyki mobilnej z uwzględnieniem błędów położenia i ułożenia. Można głównie spotkać wyniki badań konkretnych algorytmów analizy pary obrazów.

Subteza 2 pogłębia zagadnienie zmiennej bazy w stereowizji przez określenie, która z map głębi jest optymalna<sup>3</sup> (najlepsza jakościowo). Mapa taka powinna zawierać możliwie najwięcej użytecznych pikseli<sup>4</sup>, czyli nieść najwięcej informacji o głębi. Optymalna mapa głębi uzyskana przy danej szerokości bazy jest najbardziej odpowiednia dla operatora ludzkiego, gdyż zawiera minimalną liczbę pikseli z nadprogową dysparycją. Taką głębię można zobrazować w postaci obrazu stereoskopowego. W literaturze istnieje wiele prac dotyczących wyświetlania obrazu stereoskopowego, związanych przede wszystkim z produkcją filmów 3D (Zilly *et al.*, 2011). Jednakże przy ustawianiu kamer to operator ma zadbać, aby obraz stereoskopowy był dostosowany do układu widzenia człowieka. Podczas przeglądu literatury autor nie napotkał prac na temat doboru optymalnej bazy i poprawnej oceny głębi. W przeanalizowanych pracach skupiano się na poszukiwaniach automatycznej oceny odległych obiektów oraz filtracji obszarów, które są niepoprawne dla ludzkiej percepcji. Zatem określenie optymalnej bazy i odpowiednia wizualizacja obrazów stereoskopowych ludzkiemu operatorowi również stanowi nowatorski element niniejszej pracy.

## **1.4 Omówienie zakresu tematyki rozdziałów**

Niniejsza rozprawa doktorska podzielona jest na 6 rozdziałów przy czym cztery rozdziały stanowią główną część tej pracy. We wprowadzeniu przedstawiono cel pracy, tezy oraz podstawowe problemy rozważane w jej ramach.

Rozdziały 2 i 3 dotyczą zagadnień teoretycznych. Pierwszy z rozdziałów jest przeglądem rozwiązań z zakresu robotyki i prezentuje rodzaje robotów, ich wyposażenie oraz podejścia do problemu mapowania. Ze względu na różnorodność rozwiązań dla istniejących rodzajów robotów, przegląd wyposażenia jest ograniczony do dostępnych rozwiązań sprzętowych. Drugi rozdział teoretyczny opisuje podstawy matematyczne, metody pomiarowe i algorytmy ważne dla zrozumienia problemu i realizacji celu pracy. Należy tu

<sup>3</sup>Definicje funkcji celu zagadnienia optymalizacji przedstawiono w dalszej części pracy.

<sup>4</sup>Czyli najmniej pikseli nieużytecznych, którym nie można przypisać akceptowalnej dysparycji

wyróżnić zagadnienia związane z sensorami, transformacją współrzędnych, estymacją danych, a także stereowizją i przetwarzaniem obrazów.

Rozdział 4 przedstawia eksperymentalną implementację, która w praktyczny sposób realizuje cel pracy. W początkowej części rozdziału prezentowana jest koncepcja ogólna, a następnie opisane są szczegółowo kolejne realizowane elementy. Szczególny nacisk położono na problem określania położenia i ułożenia kamer, wyznaczania dysparycji dla zmiennej bazy i wyświetlania punktów 3D.

W rozdziale 5 opisano eksperymenty związane z precyzją mapowania, wyznaczeniem trajektorii robota a także wyniki analizy wrażliwości systemu stereowyjnego i jego kluczowych sensorów. Eksperymenty obejmują badania symulacyjne oraz laboratoryjne oparte na zaprojektowanym i zbudowanym robocie mobilnym. Symulacje pozwalają określić skuteczność mapowania przy dowolnej precyzji pomiarowej, natomiast eksperymenty rzeczywiste ilustrują jakość mapowania uzyskaną przy ograniczonej dokładności pomiarowej, związanej z dostępnymi zasobami sprzętowymi.

Zakończenie podsumowuje najważniejsze wyniki pracy oraz odnosi się do celów pracy i postawionych tez. W rozdziale tym zawarto także krytyczną analizę wyników i możliwe kierunki dalszych badań.

---

## PRZEGLĄD ROZWIĄZAŃ ZROBOTYZOWANYCH

Obecnie roboty są powszechnie używane w wielu dziedzinach życia. Ich głównym celem jest realizowanie określonych zadań w sposób automatyczny. W przemyśle roboty pozwalają na produkcję określonych dóbr w sposób szybki, precyzyjny i powtarzalny. W różnych dziedzinach roboty używane są zwykle, aby usprawnić określone czynności wykonywane przez człowieka albo zastąpić człowieka w miejscach niebezpiecznych lub niedostępnych.

Zadania wykonywane przez robota są tym prostsze im bardziej ograniczone jest ich środowisko pracy. Dlatego stacjonarne roboty produkcyjne, odpowiednio zainstalowane, skutecznie umieszczają określone podzespoły na taśmie produkcyjnej powtarzając proste rozkazy. Roboty, które mogą zmieniać swoje położenie w przestrzeni, która nie jest ściśle ograniczona, wymagają rozwiązania wielu dodatkowych problemów, aby w ogóle możliwe było wykonanie zwykle złożonego zadania. Skuteczność wykonywania takiego zadania jest silnie zależna od tego, czy robot jest w stanie poprawnie nawigować w danym środowisku. Jednym z kluczowych problemów związanych z nawigacją jest mapowanie, które jest głównym zagadnieniem omawianym w niniejszej pracy. Zagadnienie to dotyczy wybranego typu robota poruszającego się w środowisku opisanym w przestrzeni trójwymiarowej. Dane służące do określenia tego środowiska są dostarczane i przetwarzane przez system wizyjny.

Opis najważniejszych problemów poruszanych w tej pracy wymaga omówienia i przeglądu stosowanych rozwiązań technicznych i technologicznych. Należy do nich zaliczyć ogólny podział robotów oraz ich budowy, ze szczególnym uwzględnieniem robotów mobilnych. Ponadto niezbędne jest omówienie problemu sterowania robotem i przetwarzania danych dostarczanych przez sensory umieszczone na robotach. Po opisie możliwości, jakie dają wybrane roboty, skupimy się na zagadnieniu mapowania i lokalizacji robota w przestrzeni, aby przejść do wybranych problemów dotyczących przetwarzania obrazu w celu mapowania środowiska.

### *2.1 Rodzaje robotów*

Podjęcie się zadania zaprojektowania i zbudowania robota jest zwykle motywowane konkretnym zadaniem, które robot ma realizować. Typ, możliwości funkcjonalne, czy podzespoły są dobierane tak, aby robot mógł jak najlepiej wykonać zamierzone funkcje. Tym można tłumaczyć dużą różnorodność robotów występującą obecnie. W literaturze najczę-



ściej można spotkać ogólny podział robotów na stacjonarne i mobilne. Przez stacjonarne należy rozumieć te, które nie mogą się przemieszczać, a ruch sprowadza się do przemieszczania manipulatora lub chwytaka. Roboty mobilne nie mają powyższego ograniczenia i w ogólności mogą zmieniać swoje położenie. Stąd zdecydowanie większa różnorodność i liczność podziałów dotyczy robotów mobilnych. Typowe kryteria podziału robotów to (Siegwart i Nourbakhsh, 2004):

- branża (przemysłowe, usługowe)
- budowa robota (monolityczne, modułowe, pseudomodułowe)
- środowisko pracy (naziemne, wodne, powietrzne)
- rodzaj napędu (elektryczny, hydrauliczny, pneumatyczny)
- sposób poruszania się (kołowe, gąsienicowe, kroczące, pełzające, skaczące, toczące się, hybrydowe, oparte na wysuwaniu elementów, zmianie kształtu, śrubie, adhezji, przywieraniu, podciśnieniu)
- funkcjonalność/dziedzina zastosowań (eksperymentalne i badawcze, wojskowe i policyjne, domowe, publiczne, eksploracyjne i kosmiczne, medyczne, rozrywkowe i sportowe, transportowe).

Podział ten nie jest jednoznaczny i spójny zwłaszcza, że kryteria te nie są niezależne. Co więcej szczegółowy podział robotów zwykle wynika z podejścia do danego problemu, jaki przyświeca konstruktorowi. Jeżeli chce on zbudować robota, który wyróżnia się określoną cechą dokonuje podziału robotów względem tej cechy. Dlatego w szczególnych przypadkach można znaleźć różne podziały m.in. ze względu na (Peng *et al.*, 2003; Levander, 2017; Tzafestas, 2014; Kowalczyk i Czubenko, 2015):

- strukturę kinematyczną (otwarty/zamknięty łańcuch kinematyczny)
- sposób komunikacji (przewodowa/bezprzewodowa)
- zasięg działania (maksymalna odległość, pułap, itp.)
- manewrowość (np. minimalny promień skrętu)
- odporność na warunki zewnętrzne (temperatura, promieniowanie, uderzenia itp.)
- dynamikę ruchu (np. maksymalne przyspieszenie)
- czas działania (czas bez ładowania baterii)
- rodzaj nawigacji (oparta o dane inercyjne, znaczniki, zewnętrzny/globalny system pozycjonowania)
- algorytmy uczenia (z nauczycielem, bez nauczyciela).

Ze względu na mnogość rozwiązań, w pracy omówione zostaną tylko wybrane roboty mobilne, które są najbardziej użyteczne w kontekście tematu pracy. Przez użyteczne należy rozumieć tylko te platformy robota, które umożliwiają swobodne przemieszczanie się i mapowanie otoczenia z zastosowaniem stereowizji.

## 2.2 Roboty mobilne

Specyfika pracy robota mobilnego jest inna w przypadku środowiska wodnego, powietrznego oraz lądowego. Różnice dotyczą nie tylko konstrukcji i napędu robota ale sposobu pozyskania danych o jego otoczeniu. Dlatego poniższy przegląd rozwiązań oparty jest na podziale względem środowiska, w którym robot ma się poruszać.



Rysunek 2.1 Bezzałogowy samolot Ikhana (źródło: NASA licencja: Public Domain).

### 2.2.1. Roboty latające

Roboty latające wyróżnia wysokie zapotrzebowanie na energię podczas wznoszenia, a także w trakcie manewru zatrzymania w locie. Dlatego przy projektowaniu robotów latających należy zadbać, aby robot był lekki, miał niskie zużycie energii i mógł bezpiecznie wylądować przed rozładowaniem źródła zasilania. Roboty latające, mimo ustawienia zadanej trajektorii, mogą zmieniać położenie i orientację na skutek działania warunków zewnętrznych, takich jak wiatr i prądy powietrzne. Do robotów latających zalicza się jednostki bezzałogowe UAV (ang. *Unmanned Aircraft Vehicle*) sterowane zdalnie lub działające autonomicznie. Jednostki te wyróżniają się względnie dużymi rozmiarami. Wśród mniejszych jednostek coraz częściej spotykane są małe drony SUAV (ang. *Small Unmanned Aircraft Vehicle*). Samoloty bezzałogowe UAV latają na dużych wysokościach (rzędu setek metrów) z prędkością rzędu kilkuset  $km/h$  i wykorzystywane są przede wszystkim w celach wojskowych.

Przykładem samolotu bezzałogowego jest Ikhana - konstrukcja NASA zobrazowana na rys. 2.1, która została szczegółowo opisana w pracy (Merlin, 2009), a której rozwój związany jest ze zbieraniem danych o pożarach na terenach dzikich. W tym celu na pokładzie samolotu instaluje się specjalne skanery do termicznego mapowania terenu, co umożliwi odwzorowanie powierzchni Ziemi na mapie termograficznej. Obrazowanie przestrzenne w tym przypadku jest niepotrzebne.

Innym autonomicznym (bezzałogowym) samolotem jest Predator zbudowany dla Departamentu Obrony USA (rys. 2.2). Jest on przeznaczony do typowych zadań wojskowych związanych z mapowaniem, rozpoznaniem i śledzeniem. Proces mapowania zrealizowano za pomocą specjalnego radaru z syntetyczną aperturą SAR (ang. *Synthetic Aperture Radar*), który pozwala odwzorować dany obszar na mapie z dokładnością bliską  $0.1m$  przy założeniu, że położenie jednostki jest znane z dużą dokładnością. W swojej historii robot był użyty także do zniszczenia wykrytych celów. Opis tego samolotu oraz jednostek podobnych pod względem rozmiarów, budowy i zastosowań są dostępne w pracy (Beard i McLain, 2012; Austin, 2011), zaś rozwój i perspektywy dla bezzałogowych samolotów opisano w pracy (Weatherington i Deputy, 2005).

Przelet dużych bezzałogowych robotów podlega licznym regulacjom prawnym związanym z bezpieczeństwem zbliżonym do przepisów dotyczących samolotów cywilnych.



Rysunek 2.2 Bezzałogowy samolot RQ-1 Predator (*źródło*: Departament Obrony USA *licencja*: Public Domain).

Mniejsze jednostki zwykle wymagają zgody na lot nad terenami miejskimi, niemniej regulacje prawne dotyczące dronów dopiero powstają (EASA, 2015). Rozmiar jednostek latających jest istotnym ograniczeniem także pod kątem instalacji sensorów oraz komputerów do przetwarzania danych. Podzespoły mogą być zbyt duże i zbyt ciężkie dla małej jednostki, (np. drona). Generalnie tylko duże jednostki są w stanie udźwignąć wysokiej jakości sprzęt do nawigacji i mapowania.

Mapowanie realizowane przez jednostki latające może dotyczyć zarówno określania położenia innych obiektów latających, jak i obiektów znajdujących się na powierzchni ziemi. W środowisku powietrznym występuje niewiele obiektów latających, zwłaszcza takich, które są niezidentyfikowane przez lotnictwo lub wojsko. Dlatego jednostki latające, takie jak zwykle samoloty, z reguły nie mają urządzeń do detekcji innych obiektów poruszających się w powietrzu. Informacje dotyczące położenia i kursu innych samolotów dostarczane są z odpowiednich systemów radarowych instalowanych na powierzchni ziemi. Natomiast proces mapowania terenu, nad którym przelatuje robot zwykle dotyczy odwzorowania na płaskiej mapie obiektów znajdujących się na powierzchni ziemi. Mapowanie 3D ma większy sens dla dużych obiektów, np. budynków mieszkalnych. Ze względu na duże odległości między samolotem a mapowanym terenem, obiekty widoczne z góry są odwzorowane w zakresie elementów pionowych ze znacznie mniejszą dokładnością niż elementy poziome. Co więcej systemy wizyjne dedykowane do mapowania z dużych wysokości wymagają bardzo drogich kamer i układów optycznych (Beard i McLain, 2012). Podobne problemy nie występują dla systemów, które mapują otoczenie z powierzchni ziemi. Niestety, zmniejszenie pułapu lotu samolotu bezzałogowego do kilkudziesięciu metrów (wysokości porównywalnych z wysokością budynków) jest niemożliwe ze względów bezpieczeństwa. Mimo to samoloty bezzałogowe są bardzo użyteczne w wielu zastosowaniach dotyczących monitorowania wybranego obszaru dla potrzeb wojskowych.

Drony (SUAV) wyróżniają się małymi rozmiarami i nieco innym obszarem zastosowań. Roboty tego typu są budowane z napędem śmigłowym ustawionym poziomo. Silniki wytwarzają ciąg w kierunku pionowym, przez co nie jest potrzebna minimalna prędkość w kierunku poziomym konieczna, aby zapewnić unoszenie drona w powietrzu. Zasięg i pułap drona jest silnie ograniczony pojemnością źródła zasilania, która wystarcza zwykle na kilkanaście minut lotu. Małe rozmiary ograniczają wagę sensorów, które można zain-



Rysunek 2.3 Mały dron Phantom 4 (źródło: DJI licencja: Creative Commons 0).

stalować na dronie. Zainstalowanie dodatkowych sensorów i procesorów również wpływa negatywnie na czas lotu i dynamikę sterowania. Dlatego duże systemy radarowe, skanery laserowe, czy złożone układy optyczne nie mogą być instalowane na pokładzie małych dronów. Niemniej typowy dron o rozpiętości ramion mniejszej niż  $1m$  jest w stanie udźwignąć mały system wizyjny zawierający kilka mini-kamer i kilka małych sensorów inercyjnych lub sonarów. Co więcej dron może przemieszczać się w pobliżu wybranego obiektu, a nawet utrzymać swoje położenie (gdy występują warunki bezwietrzne). Dlatego dron w ogólności umożliwia mapowanie nie tylko powierzchni ziemi, ale także mapowanie przestrzenne, np. obszarów miejskich.

Przykładem autonomicznego drona jest konstrukcja firmy DJI o nazwie Phantom 4 zaprezentowana na rys. 2.3. Według specyfikacji dron ten umożliwia śledzenie obiektu i omijanie przeszkód w sposób automatyczny. Zadania te są realizowane poprzez system składający się z sonarów i czterech kamer z dedykowanymi jednostkami do przetwarzania obrazu (Sagitov i Gerasimov, 2017).

Inne implementacje dronów opisywane w literaturze są częściowo autonomiczne i umożliwiają mapowanie. W pracy Angrisaniego (2015) autorzy skupiają się na ładowaniu baterii drona w sposób zautomatyzowany. Praca McAree *et al.* (2016) opisuje system do weryfikacji bezpieczeństwa drona, gdzie założono, że robot jest pół-autonomiczny. Podejście opisane przez Esrafiliana i Taghirada (2016) przedstawia system mapowania i śledzenia obiektów oparty na danych wizyjnych pobranych z kwadrokoptera. Jednakże robot z prezentowanym systemem testowany był na materiale offline (zapisanym wcześniej na karcie pamięci) oraz przemieszczał się w bardzo ograniczonej przestrzeni.

Obecnie dla małych dronów można znaleźć liczne podejścia do mapowania i zwiększania autonomiczności, jednak nie są to konkretne implementacje uwzględniające ograniczenia wagi podzespołów elektronicznych, mocy procesora, czy zapewniające, że algorytmy będą działały w różnych warunkach w czasie rzeczywistym. Ponadto drony są stosunkowo podatne na uszkodzenia i kolizje podczas eksperymentów. Roboty SUAV mogą być także niebezpieczne dla osób znajdujących się w ich pobliżu - w szczególności, gdy system jest w fazie wstępnych testów. W ogólności mapowanie przestrzenne realizowane przez roboty latające jest problemem trudnym i problematycznym w przypadku małych dronów. Dla większych jednostek latających na dużych wysokościach mapowanie sprowadza się do odwzorowania 2D dużych obiektów.

### 2.2.2. Roboty wodne

Zagadnienie mapowania w przypadku robotów pracujących w środowisku wodnym wymaga rozwiązania problemów innego rodzaju. Statki pływające na wodzie w ogólności mają bardzo ubogie środowisko do mapowania. Podczas rejsu statku nad wodą zwykle widoczne są pojedyncze jednostki pływające. Jedynie przy nabrzeżu występują tereny bardziej urozmaicone, których mapowanie może być potrzebne. Co więcej określenie położenia na wodzie nie jest łatwe, gdyż obiekty zanurzone w wodzie poddają się ruchom wody. Dlatego autonomiczne statki wymagają zewnętrznego globalnego systemu pozycjonowania. Przykłady autonomicznych robotów pływających można znaleźć w pracach Levandera (2017), Johansena i Pereza (2016). Ze względu na duże rozmiary statków do ustalania pozycji używane są systemy GPS i radiolatarnie, zaś do detekcji innych obiektów wykorzystuje się radary i dedykowane systemy informacji, np. AIS (ang. *Automatic Identification System*). Dla bezpieczeństwa statku badana jest także głębokość za pomocą echosond. Mapowanie przestrzenne dla statku nie jest w ogólności potrzebne. Do nawigacji wystarczająca jest mapa płaska, która jest znana *a priori* oraz system pozycjonowania.

Środowisko znajdujące się pod powierzchnią wody jest mniej interesujące dla człowieka. Mimo to istnieje wiele robotów dedykowanych do zadań podwodnych. Wśród podwodnych pojazdów (robotów) bezzałogowych UUV (ang. *Underwater Unmanned Vehicle*) wyróżnia się roboty zdalnie sterowane ROV (ang. *Remotely Operated Vehicle*) oraz roboty autonomiczne AUV (ang. *Autonomous Underwater Vehicle*). Roboty zdalnie sterowane są używane szczególnie w pracach, których człowiek nie może wykonać, np. ze względu na wysokie ciśnienie występujące na dużych głębokościach. Typowe zadania wykonywane z pomocą ROV to prace w przemyśle petrochemicznym przy sprawdzaniu odwiertów, naprawie, inspekcji statków, czy eksploracji wraków statków (Bono *et al.*, 1999; Nakajoh *et al.*, 2012). Roboty autonomiczne AUV są przydatne w wielu prostszych zadaniach, których realizacja wymaga długiego czasu pracy. Do takich zadań należy zaliczyć mapowanie dna morskiego, szukanie i rozbijanie min lub innych obiektów niebezpiecznych (Soberi i Zakaria, 2016; Vikranth *et al.*, 2016). Oba typy robotów poruszają się dzięki obrotom śruby napędowej i zbiornikom wypornościowym.

W środowisku wodnym występują istotne problemy z propagacją sygnałów. Fale radiowe są silnie tłumione przez wodę, natomiast fale mechaniczne zmieniają swoje właściwości w zależności od temperatury i zasolenia wody. Co więcej, fale akustyczne odbijają się od dna i innych obiektów zanurzonych w wodzie, przez co komunikacja na drodze akustycznej jest możliwa tylko w bardzo ograniczonym zakresie. Dlatego też roboty ROV najczęściej komunikują się z użytkownikiem za pomocą łączności przewodowej. Roboty autonomiczne podczas wykonywania zadania mogą nie mieć komunikacji z operatorem. Natomiast dane zebrane podczas pracy można odczytać dopiero po wynurzeniu się robota nad powierzchnię wody. Przykładem okrętu podwodnego jest robot Boeinga - Echo Voyager, który może pracować autonomicznie przez wiele tygodni (rys. 2.4). Jest to jednostka wielozadaniowa, której zadania to między innymi zbieranie danych o organizmach podwodnych lub położeniu okrętów, czy pilnowanie szlaków przed piratami. W literaturze opisano rozwiązania, które zbierają i przetwarzają dane wizyjne, aby wykrywać obiekty znajdujące się w wodzie. Jednak wzburzona woda oraz wzniesiony piasek w obszarach przydennych danego akwenu, tworzą zniekształcenia obrazu. Stąd akwizycja i przetwarzanie obrazu w przypadku robotów pracujących pod wodą mogą być nieskuteczne w warunkach rzeczywistych (Oleari *et al.*, 2016; Huang *et al.*, 2012).





Rysunek 2.4 Autonomiczny robot podwodny Echo Voyager (*źródło: DARPA licencja: Public Domain*).

### 2.2.3. Roboty lądowe

Ostatnia i jednocześnie najbardziej liczna grupa to roboty poruszające się w środowisku lądowym. Jest ono bardzo bogate w obiekty, które na co dzień są interesujące dla człowieka. Co więcej ruch odbywa się najczęściej na stabilnym podłożu, przez co położenie robota w ogólności nie ulega zmianie, gdy napęd jest wyłączony. Dlatego szczegółowe mapowanie i lokalizowanie obiektów na lądzie jest realizowane stosunkowo często w porównaniu ze środowiskiem wodnym i powietrznym.

Do mobilnych robotów lądowych należą roboty kołowe. Roboty te są bardzo powszechne, gdyż pojazd kołowy na płaskim podłożu może poruszać się bardzo wydajnie. Typowy robot kołowy może się poruszać bez większych ograniczeń do przodu i do tyłu. Ograniczenia występują dla manewru skręcania. Robot nieholonomiczny posiada istotne ograniczenia ruchu, np. może skręcać tylko po łuku z pewnym minimalnym promieniem skreću albo zablokować ruch poprzez odpowiednie ustawienie koła. Przykładem pojazdu nieholonomicznego jest samochód. Robot holonomiczny nie ma takich ograniczeń i może wykonywać ruchy w dowolnym kierunku bez wcześniejszej zmiany ułożenia. W przypadku prostych kół, możliwy jest tylko taki ruch robota, który jest konsekwencją obrotu koła zamocowanego na wale napędowym. Dodatkowo do pary kół stałych umieszczonych w przedniej części robota może być koło obrotowe instalowane w tylnej części robota. W efekcie robot może obracać się w miejscu i poruszać się do przodu lub tyłu po linii prostej albo po łuku. Większe możliwości mają koła szwedzkie (wszechkierunkowe), które pozwalają na ruch robota w lewo lub w prawo bez konieczności obracania robota.

Opis przykładowego robota z kołami szwedzkimi oraz implementację dedykowanego sterownika znaleźć można w pracach Hana (2016) i Pawitana (2016). Warto nadmienić, że brak ograniczeń robota (holonomicznego), zbudowanego np. z 3 kół szwedzkich, może być szkodliwy w użytkowaniu (bez odpowiednich hamulców robot może się samoczynnie przemieszczać w dowolnym kierunku na pochyłym podłożu).

W przypadku gdy robot mobilny musi poruszać się po nierównym, górzystym terenie często stosuje się napęd gąsienicowy. Taki pojazd może przejechać nawet przez przeszkody, które są niewielkie w porównaniu z rozmiarami robota. Gąsienice pozwalają także na jazdę po zboczach o dużym nachyleniu nawet na podłożu w ogólności nieutwardzonym. Niestety

każdy obrót robota wiąże się z dużym tarciem gąsienic o podłoże, przez co ruch takiego robota może być mało efektywny energetycznie. Przykładem robota gąsienicowego jest robot Talon przedstawiony na rys. 2.5 używany przez Departament Obrony USA oraz agencję DARPA (ang. *Defense Advanced Research Projects Agency*). Roboty gąsienicowe o podobnej budowie produkowane są przez PIAP (Przemysłowy Instytut Automatyki i Pomiarów) opisane, np. przez Trojneckiego (2015). Konstrukcje te są używane zwykle do zadań związanych z bezpieczeństwem. Dzięki zaopatrzeniu w specjalny manipulator na długim ramieniu robot może przenosić ładunki wybuchowe w celu ich neutralizacji. Manipulator można zamienić na inne elementy wykonawcze lub sensory, np. czujnik badający poziom promieniowania.



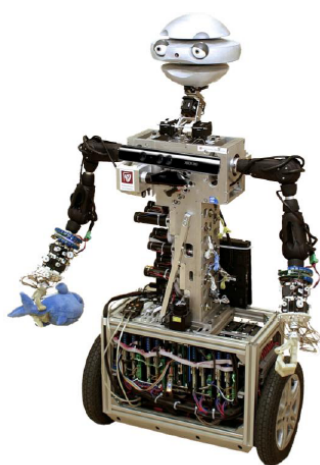
Rysunek 2.5 Robot gąsienicowy Talon (źródło: Departament Obrony USA licencja: Public Domain).

Zarówno roboty kołowe, jak i gąsienicowe można wyposażyć w odpowiednie sensory i układy sterowania pozwalające na mapowanie i analizę otoczenia. Komunikacja z robotem lądowym w ogólności jest możliwa nie tylko za pomocą przewodów, ale także poprzez rozwiązania bezprzewodowe. W otoczeniu robotów lądowych, w odróżnieniu od robotów środowiska wodnego, nie występują istotne przeszkody dla poprawnej pracy sensorów robota, w szczególności dla systemu wizyjnego.

Wśród lądowych robotów mobilnych można wyróżnić także takie roboty, których mechanizm poruszania się jest inspirowany sposobem poruszania się organizmów żywych. Należą do nich roboty kroczące, skaczące, pełzające, itp. Roboty kroczące poruszają się przez ruch odnóży i odpowiednie przenoszenie środka ciężkości. Jest to dużo trudniejszy sposób poruszania się w porównaniu z ruchem kołowym czy gąsienicowym (Pfeifer *et al.*, 2007). Robot posiadający jedną parę odnóży wymaga odpowiedniego sterowania, które zapewnia równowagę robota i chroni przed upadkiem podczas ruchu a także podczas stania. Do takich robotów należą roboty humanoidalne, które budową i wyglądem przypominają człowieka. Definicja ta nie jest ścisła, dlatego również niektóre roboty jeżdżące zalicza się do humanoidów (Tzafestas, 2014; Rodriguez, 2011). Roboty humanoidalne tworzone są głównie z myślą, aby były towarzyszami człowieka. Co więcej robot powinien być na tyle zaawansowany, aby można było się z nim komunikować, a także by mógł wykonywać interakcje z okolicznymi obiektami. Ze względu na złożoność realizowanych zadań, humanoid jest zaopatrzone w liczne sensory oraz kamery i sonary. Niemniej liczba senso-

rów, ich ciężar i rozmieszczenie, wpływają na utrzymywanie równowagi, a zabezpieczenie robota przed upadkiem wymaga dodania nowych sensorów. Ponadto w obecnych rozwiązaniach problemem jest akwizycja danych z wielu sensorów jednocześnie i przetworzenie ich w czasie rzeczywistym przez energooszczędny procesor robota. Aktualnie istniejące rozwiązania nie pozwalają na zbudowanie humanoida, który byłby w stanie analizować otoczenie kontekstowo i wykonywać bardzo złożone zadania. Niemniej, nie ma przeszkód, aby humanoid był wyposażony w podzespoły pozwalające na mapowanie otoczenia.

Przykładem robota humanoidalnego jest FLASH widoczny na rys. 2.6 zrealizowany na Politechnice Wrocławskiej. Jest to robot posiadający obrotowy element przypominający głowę oraz dwa ruchome ramiona przypominające ręce a jego napęd jest oparty na dwóch kołach. Zatem sterownik robota musi zapewniać odpowiedni ruch kołami, aby umożliwić utrzymanie równowagi (Kędziński *et al.*, 2015).



Rysunek 2.6 Robot humanoidalny FLASH (źródło: lirec.eu, prawa autorskie: Lirec Group).

Przykład innej, mniejszej konstrukcji, która wizualnie bardziej przypomina człowieka to robot Nao przedstawiony na rys. 2.7. Konstrukcja zawiera elementy przypominające głowę, ręce oraz nogi. Roboty humanoidalne, w przeciwieństwie do innych rodzajów robotów, służą m.i. badaniom nad wyrażaniem emocji, analizą emocji, postrzeganiem i zachowaniami ludzkimi (Czubenko *et al.*, 2015; Kowalczyk i Merta, 2016b). Coraz częściej roboty humanoidalne można spotkać w leczeniu autyzmu czy w zwykłej rozrywce, gdzie robot jest postrzegany jako zabawka.

Roboty kroczące zbudowane z więcej niż jednej pary odnóży są spotykane również w nauce i przemyśle (mają jednak inne zastosowania niż roboty humanoidalne). Roboty czteronożne można spotkać w zastosowaniach wojskowych. Przykładem takiego robota jest Big Dog widoczny na rys. 2.8, którego zadaniem jest przemierzanie trudnego terenu. Robot jest w stanie przenosić ciężkie ładunki wspierając małe grupy żołnierzy w terenie, gdzie pojazdy kołowe nie mają dostępu. Rozmiary i kształt robota pozwalają na zaopatrzenie go w sensory do nawigowania w terenie (Meng *et al.*, 2016).

Inne roboty lądowe wykorzystujące nietypowe rodzaje ruchu, np. pełzanie, tworzone są w specyficznych konkretnych celach. Robot pełzający inspirowany ruchem węża pozwala na dotarcie do zamkniętych wąskich obszarów, takich jak rury. Wykorzystując możliwość



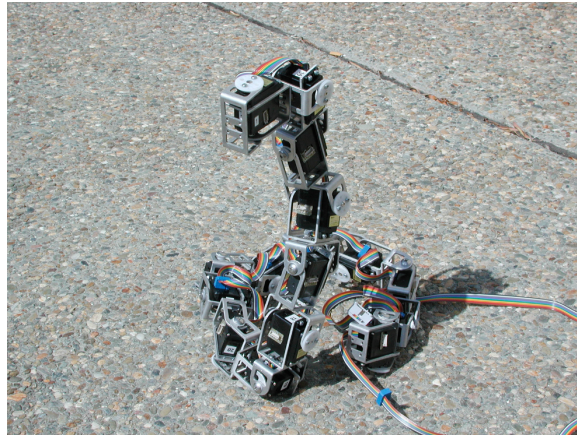
Rysunek 2.7 Robot humanoidalny Nao (źródło: Ubahnverleih wikipedia licencja: Creative Commons 0).

zaciskania wokół kolumny robot ma możliwość wspięcia się po kolumnie (do góry). Przykładem robota pełzającego jest SnakeBot widoczny na rys. 2.9. Jest on modułowy, przez co zniszczenie danego modułu nie skutkuje awarią całego robota. Robot pozwala także na przemierzanie stromego skalistego terenu. Niemniej prędkość przemieszczania się robota jest mała. Ze względu na rozmiary i budowę robota nie można zainstalować na robocie wielu sensorów przydatnych w mapowaniu otoczenia (Dennis i Edwards, 2013).

Wśród robotów lądowych istnieją także rozwiązania hybrydowe, które łączą kilka typów mechanizmów poruszania się robota mobilnego. Do nietypowych rozwiązań można także zaliczyć ruch polegający na przemieszczaniu się wielu robotów (systemy agentowe). Niemniej na potrzeby niniejszej pracy omówione zostały tylko konstrukcje robotów, które można w łatwy sposób zaopatrzyć w system wizyjny oraz inne sensory potrzebne do mapowania przestrzennego. Różnice w mapowaniu otoczenia w przypadku środowiska wodnego, powietrznego i lądowego, możliwość prowadzenia bezpiecznych eksperymentów



Rysunek 2.8 Robot czworonożny BigDog (źródło: DARPA licencja: Public Domain).



Rysunek 2.9 Robot SnakeBot (źródło: NASA licencja: Public Domain).

a także dostępność podzespołów powodują, że konieczne jest skupienie się na jednym konkretnym typie robota. Wobec powyższego w dalszej części pracy pod pojęciem robot mobilny będzie rozumiany robot kołowy, poruszający się po lądzie. Szczegóły dotyczące nawigacji takiego robota w terenie zostaną omówione w dalszej części rozdziału.

## 2.3 Wyposażenie robotów mobilnych

Mapowanie obiektów znajdujących się w środowisku lądowym w ogólności wymaga rekonstruksji po zadanym obszarze. Dlatego sposób, w jaki robot mobilny porusza się istotnie wpływa na proces mapowania. Szczególnie ważną rolę odgrywają tu napęd robota oraz sensory. Do napędu należy zaliczyć wszystkie elementy, które biorą udział w zamianie określonego rodzaju energii, np. elektrycznej lub pochodzącej ze spalania paliwa, na ruch robota. Natomiast sensory dostarczają dane pomiarowe związane ze środowiskiem zewnętrznym a także ze stanem robota.

### 2.3.1. Napęd robota kołowego

Elementem napędowym, który zamienia zmagazynowaną energię bezpośrednio na energię mechaniczną, jest silnik. Kołowe roboty mobilne są zaopatrzone najczęściej w silnik elektryczny. Silnik spalinowy używany jest w dużych jednostkach o dużej bezwładności, wymagających dynamicznego ruchu (jak autonomiczne samochody). Niemniej ze względów na negatywny wpływ na środowisko, silniki spalinowe są coraz częściej zastępowane przez silniki elektryczne (Cairns i Albertus, 2010).

Mobilne roboty kołowe budowane są przede wszystkim w oparciu o silniki prądu stałego DC (ang. *Direct Current*), prądu przemiennego, oraz silniki krokowe. Występuje wiele realizacji silników prądu stałego i przemiennego (szczotkowe i bezszczotkowe), które różnią się sposobem wzbudzenia pola magnetycznego. Niemniej mają wspólną cechę, po przyłożeniu określonego napięcia do zacisków silnika pracującego pod obciążeniem, silnik będzie przyspieszał, aż do uzyskania stałej prędkości obrotowej. Silnik taki w zasadzie można łatwoysterować, niemniej precyzyjne określenie drogi, którą przejechał robot nie jest łatwe i wymaga odpowiednich pomiarów. Silniki krokowe pozwalają na dokładniejsze określenie dystansu, jaki pokonał robot, gdyż silnik jest sterowany impulsowo a pojedynczy impuls powoduje obrót wału silnika o pewien określony mały kąt. Sterowa-

nie silnikiem krokowym jest jednak trudniejsze. Wymuszenie gwałtownego przyspieszenia bądź hamowania, bądź nadmierne obciążenie silnika może powodować gubienie kroków. Silnik krokowy ma większy pobór prądu w porównaniu ze zwykłym silnikiem DC. Dlatego silniki krokowe są szczególnie użyteczne w zastosowaniach, gdzie obciążenie silnika jest generalnie małe i stałe, a prędkość obrotowa nieduża (Holdaway, 2000; Krishnamurthy i Khorrami, 2003). Projekt kołowego robota mobilnego, którego napęd jest oparty na silniku krokowym ma istotne ograniczenia.

W napędzie robota występuje także element, który zamienia sygnał małej mocy sterujący robotem (wymuszenie) na sygnał dużej mocy. Układ ten ma ograniczenia związane z wydajnością prądową a także wprowadza zakłócenia w obwodzie zasilania. Stąd nie każdy sygnał sterujący może być poprawnie przetworzony przez układ wzmacniania mocy (ang. *Motor Driver*)<sup>1</sup>. Wpływ tego układu należy uwzględnić przy modelu opisującym ruch robota szczególnie, gdy sygnał sterujący robotem jest w postaci sygnału PWM (ang. *Pulse Width Modulation*).

Napęd robota wprowadza ograniczenia na ruchy jakie może wykonywać robot. W zależności od rozmieszczenia silników na platformie robota oraz użytego typu kół konieczne są różne podejścia do sterowania robotem. Szczegółowy podział można znaleźć w pracy Rodriguez (2011), w rozdziale dot. robotów mobilnych. Natomiast sposób sterowania różnymi rodzajami robotów kołowych można znaleźć w pracy Abdolahi (2012).

Wśród rozwiązań napędu dla robotów wielokołowych można wyróżnić sterowanie różnicowe dla robotów czterośladowych, sterowanie dla robotów trzyśladowych (z kołem obrotowym), czy sterowanie dla napędu zawierającego cztery koła szwedzkie (omni). Szczegóły opisu możliwych realizacji nie są tu ważne, natomiast z punktu widzenia niniejszej pracy kluczowe jest, aby dla wybranego robota mobilnego i konkretnego napędu można było podać model matematyczny, który w możliwie prosty sposób opisze istotne ograniczenia dynamiki ruchu robota. Robot wykorzystywany do mapowania powinien swobodnie poruszać się po zadanym terenie bez większych ograniczeń. Najlepiej zatem wybrać rozwiązanie, które ma nieznaczne ograniczenia ruchu i pozwala na przemieszczanie się po zadanej trajektorii w możliwie prosty sposób.

Model robota jest istotny przede wszystkim dla zapewnienia właściwego sterowania robotem. Znając model można przewidzieć z pewną dokładnością, jaki będzie ruch robota w każdej dopuszczalnej sytuacji. Wiarygodność modelu w zasadniczy sposób zależy od dopuszczonych układów eksperymentu, gdyż nawet bardzo skomplikowany model zwykle nie uwzględnia wszystkich możliwych sytuacji rzeczywistych. Dlatego, nawet zakładając brak szumów w układzie sterowania w praktyce występują błędy różnego rodzaju. Co więcej na dokładność matematycznego odwzorowania wpływają nie tylko ograniczenia danego typu robota i struktura modelu, ale także precyzja, z jaką wyznacza się parametry modelu w przypadku konkretnego robota.

W teorii sterowania układ napędowy robota reprezentowany musi być za pomocą modelu, który wystarczająco dobrze opisuje dynamikę ruchu robota. Dla zadanych sygnałów wejściowych (wymuszeń), przy znanych parametrach modelu oraz mierzonych wielkościach fizycznych istotnych dla modelu, można estymować położenie robota. Znajomość położenia jest istotna w procesie mapowania, gdyż względem aktualnego położenia

<sup>1</sup>Określenie angielskie „motor driver”, które powinno być bezpośrednio tłumaczone jako „sterownik silnika” (nie należy mylić z układem sterującym czyli „sterownikiem”) w tekście będzie zwykle nazywane „układem mocy”.

zbierane są pomiary otoczenia. Szum pomiarowy może powodować istotne błędy estymowanego położenia. Jednak estymacja położenia realizowana w oparciu o obserwatory może skutecznie odfiltrować błędy pomiarowe. Innymi słowy model analizowanego zjawiska dynamicznego ogranicza dynamikę jego sygnałów, w tym zmian jego położenia (Kowalczyk i Merta, 2015). Dobór sensorów używanych przy określaniu położenia jest zatem bardzo ważny dla poprawnego mapowania otoczenia. Oczywiście proces mapowania może być właściwie realizowany tylko w oparciu o sensory o odpowiedniej precyzji. Dlatego poniżej omówimy dostępne sensory oraz sposoby pomiaru położenia i ułożenia robota w przestrzeni, a także mapowania przestrzennego, które umożliwi lokalizację obiektów znajdujących się w otoczeniu.

### **2.3.2. Sensory**

Sensory spełniają podobną rolę jak zmysły organizmów żywych. Mobilny robot kołowy może być zaopatrzone praktycznie w dowolny sensor pod warunkiem, że platforma robota będzie w stanie bez przeszkód udźwignąć i odpowiednio zasilić go. Zależnie od zastosowania można znaleźć roboty z sensorami mierzącymi temperaturę, wilgotność, stężenie promieniowania i inne wielkości związane z zastosowaniem jednostki. Dzięki temu robot stanowi mobilną stację pomiarową. Wśród różnych sensorów istotną rolę odgrywają czujniki dedykowane nawigacji robota, które często stanowią podstawę pracy autonomicznej.

W literaturze spotyka się różne podziały sensorów związanych z nawigacją ze względu na szczególne zastosowanie, rodzaj pomiaru, oraz typ mierzonej wielkości fizycznej (Everett, 1995). Siegwart i Nourbakhsh (2004) w rozdziale dotyczącym percepcji sygnałów proponują podział inspirowany układami czucia organizmów żywych (proprioceptywny i eksteroceptywny). Do czujników proprioceptywnych można zaliczyć sensory wewnętrzne mierzące, np. prędkość obrotową silnika, obciążenie, poziom naładowania baterii, położenie manipulatora. Sensory eksteroceptywne zbierają dane zewnętrzne pochodzące ze środowiska – typu odległość do celu, natężenie pola magnetycznego, natężenie światła, temperatura powietrza, itp.

Sensory dzieli się także na pasywne i aktywne. Pasywne mierzą energię sygnałów naturalnie generowanych przez środowisko. Czujniki aktywne dodatkowo emitują określony sygnał do środowiska, a następnie mierzą i analizują odbierany sygnał (zmodyfikowany przez środowisko). Dlatego sensory aktywne w ogólności dostarczają danych lepszych jakościowo. Jednak odbywa się to kosztem pewnej ingerencji w środowisko. Na danym ograniczonym obszarze aktywne sensory mogą nie tylko zakłócać się wzajemnie, ale także szkodzić organizmom żywym. W pracy przyjęto ograniczającą zasadę dostosowania podziału czujników do możliwości ich aplikacji na małym kołowym robocie mobilnym działającym w ograniczonym środowisku.

#### *Nawigacja zliczeniowa*

Nawigacja wymaga danych sensorycznych, którymi mogą być pomiary pochodzące z dedykowanych czujników oraz odpowiednio przetworzone rozkazy sterujące. W przypadku robota kroczącego dane związane z ruchem członów wykonawczych można otrzymać przez zliczanie kolejnych kroków. Taki rodzaj nawigacji określa się jako nawigację zliczeniową. Jest to najstarszy i generalnie mało dokładny typ nawigacji. Dla robota kołowego nawigacja zliczeniowa polega na obliczeniu aktualnej pozycji na podstawie drogi przebytej przez

robota od ostatniej obserwowanej pozycji. Metoda ta od dawna była wykorzystywana w nawigacji morskiej, gdzie na podstawie kursu i szybkości statku na mapie regularnie zaznaczano przebywany odcinek drogi. Jest ona niedokładna, gdyż elementy takie jak dryf, prądy morskie, czy błąd pomiaru, kumulują się wraz z przebytym dystansem. W przypadku robota kołowego potrzebny jest sensor, który pozwala na określenie liczby obrotów koła. Na błąd pozycji wpływa zatem dokładność przyrostowego wyznaczania drogi a także inne czynniki, takie jak ślizg kół.

W nawigacji zliczeniowej robota kołowego wykorzystywane są sensory odometryczne, które zamieniają obroty koła na impulsy elektryczne. Typowe rozwiązania oparte są na enkoderach optycznych, magnetycznych, pojemnościowych, indukcyjnych, potencjometrach (Everett, 1995). Impulsy generowane są w oparciu o tarczę zainstalowaną na wale silnika, która składa się zwykle z kilku do kilkudziesięciu obszarów rozróżnianych przez detektor. Przykładowo dla enkodera magnetycznego tarcza składa się z domen magnetycznych ułożonych naprzemiennie. Warto zaznaczyć, że dokładność wyznaczania drogi za pomocą enkoderów jest zależna nie tylko od liczby obszarów, na które podzielona jest tarcza, ale także od przekładni silnika w przypadku, gdy tarcza jest umieszczona na głównym wale silnika. Enkodery pozwalają nie tylko na wyznaczenie odległości, jaką przebył robot kołowy. Analizując impulsy z enkoderów z co najmniej dwóch kół, określić można, w jakim stopniu robot skręca na płaszczyźnie ruchu. Nawigacja zliczeniowa mimo licznych wad jest wciąż używana na krótkich dystansach oraz w sytuacji, gdy nie można określić położenia względem innych znanych punktów – sygnalizatorów znajdujących się w środowisku (Siegwart i Nourbakhsh, 2004).

### *Sensory inercyjne*

Inna grupa czujników, które w pewnym stopniu można wykorzystać do nawigacji zliczeniowej to sensory inercyjne. Nawigacja ta, określana jest jako nawigacja bezwładnościowa, polega na pomiarze przyspieszeń działających na robota, które są efektem zmiany jego położenia. Ponadto sensory inercyjne umożliwiają wyznaczenie ułożenia robota (Titterton i Weston, 2005). W obu pomiarach zliczanie polega na akumulacji danych wraz z upływem czasu. Ułożenie (położenie kątowe) oblicza się całkując dane z żyroskopu, zaś położenie (liniowe) obliczane jest poprzez podwójne całkowanie danych z akcelerometru (Hardt von der et al., 1996). Ze względu na szумы sensorów w mierzonym ułożeniu można zaobserwować dryf. Natomiast określenie położenia jest dodatkowo utrudnione przez obecność przyspieszenia ziemskiego. Do sensorów inercyjnych zalicza się także magnetometr mierzący natężenie pola magnetycznego (Bekir, 2007). Magnetometr po odpowiedniej kalibracji pozwala na określenie ułożenia robota względem północy magnetycznej. Często przy sensorach inercyjnych można spotkać wysokościomierz, który mierzy wysokość na podstawie ciśnienia powietrza. Jest on użyteczny w samolotach latających na dużych wysokościach, gdzie ciśnienie powietrza zmienia się wraz z wysokością. W przypadku obu sensorów podczas pomiarów nie jest potrzebne zliczanie, przez co nie występuje dryf.

Przedstawione sensory inercyjne zwykle dostępne są w postaci jednostki do nawigacji inercyjnej IMU (ang. *Inertial Measurement Unit*). Jednostka ta zawiera 3-osiowy akcelerometr, 3-osiowy żyroskop, 3-osiowy magnetometr oraz wysokościomierz, które obecnie produkowane są najczęściej w technologii MEMS (ang. *Micro Electro Mechanical Systems*) (Albarbar et al., 2009; Ang i Khoo, 2004; Pang i Liu, 2001). Jednostkę IMU spotyka





się przede wszystkim w zastosowaniach, które wymagają dokładnego pomiaru ułożenia. Mimo, że dane z magnetometru pozwalają na określenie ułożenia, szumy sensora i wpływ ferromagnetyków powodują istotne błędy pomiarowe. Z drugiej strony akcelerometr pozwala na określenie odchylenia robota od pionu przy założeniu, że na robota mogą działać tylko inne małe siły w porównaniu z siłą grawitacji. Dlatego ułożenie robota obliczane jest na podstawie fuzji danych z jednostki IMU, co ma miejsce, np. w systemie AHRS (ang. *Attitude and Heading Reference System*) używanym w lotnictwie do nawigacji bezwładnościowej (Li *et al.*, 2006).

### *Pomiar odległości*

Nawigacja zliczeniowa jest wykorzystywana także obecnie w sytuacjach, gdzie naturalne zewnętrzne sygnały lub punkty odniesienia nie występują, oraz żaden sztuczny sygnał odniesienia nie jest dostępny (lub wystarczający do zrealizowania określonego zadania). W pozostałych przypadkach do zadań nawigacyjnych można użyć sensorów wykrywających zewnętrzne sygnały lub punkty w sposób aktywny bądź pasywny. W tej grupie sensorów można wyróżnić czujniki odległości oraz dotyku. Czujniki odległości generują sygnał, który odbija się od powierzchni danego obiektu. Powracający sygnał dostarcza danych dotyczących odległości na podstawie pomiaru czasu przelotu, amplitudy lub fazy sygnału. W zależności od budowy sensora generowane są sygnały optyczne, akustyczne, radiowe, magnetyczne, indukcyjne, pojemnościowe, itp. (Everett, 1995). Czujniki dotyku mogą być użyteczne, gdy przeszkoda znajduje się bezpośrednio przy sensorze (np. przy unikaniu kolizji).

Sensory odległości zwykle dostarczają pomiaru dotyczącego jednego punktu bądź otoczenia danego punktu. Pomiar taki może być wystarczający do oceny położenia obiektu względem robota, tylko jeśli porusza się on po linii prostej. Bardziej użyteczne są dane z czujników, które mierzą odległość względem wielu punktów. Do takich sensorów zalicza się radary i sonary, które wysyłają i odbierają wiele sygnałów w stosunkowo krótkim czasie. Sensory te wysyłają sygnał sondujący o stosunkowo wąskiej wiązce w danym kierunku a następnie odbierają i analizują sygnał odbity. Sonary i radary fizycznie obracają się lub odpowiednio przestrajają generowany sygnał, przez co zbierane są odbicia z szerokiego przedziału kąтового. Dzięki temu w krótkim czasie opomiarowane jest otoczenie robota. Radary wysyłają fale radiowe w zakresie mikrofal, zaś sonary wysyłają fale akustyczne, w szczególności ultradźwiękowe (Borenstein *et al.*, 1997). Fala dźwiękowa porusza się w powietrzu blisko  $10^6$  razy wolniej niż fala radiowa i jest bardziej tłumiona, stąd radar może w krótkim czasie odebrać sygnał odbity od dalekich obiektów. Ważne jest, aby obiekty te dobrze odbijały dany zakres fal. Sonar działa na dużo krótszym dystansie i pozwala wykrywać takie powierzchnie obiektów, które dobrze odbijają fale mechaniczne (dźwiękowe lub ultradźwiękowe).

Bardziej zaawansowanym sensorem mierzącym odległość jest lidar (ang. *Light Detection and Ranging*), który działa podobnie jak radar, ale sygnał sondujący składa się z impulsów światła laserowego (Bekir, 2007). Sygnał odbity jest wykrywany za pomocą układu optycznego z fotodiodą i kamerą, a następnie odebrany sygnał podawany jest dalszej analizie komputerowej. Lidar to złożony system pomiarowy zawierający wewnętrzny system wizyjny. Dla zapewnienia dokładności pomiarów lidar zwykle wzbogacony jest o dodatkowe sensory określające ułożenie i położenie. Typowy radar i sonar pozwala

na wyznaczenie odległości i położenia obiektu (na płaszczyźnie) we współrzędnych biegunowych, jednakże nie jest znana współrzędna wysokościowa obiektu. Natomiast lidar zmieniając kierunek emisji i detekcji światła laserowego pozwala na pomiar położenia przeszkody w 3 wymiarach (względem sensora). Warto zaznaczyć, że lidar podobnie jak inne sensory odległości jest czujnikiem aktywnym emitującym sygnał, którego moc może być stosunkowo duża. W przypadku, gdy wiele takich samych aktywnych sensorów pracuje w tym samym otoczeniu należy zadbać, aby urządzenia nie zakłócały się wzajemnie, np. zapewniając różne sygnały sondujące (z różnymi długościami fali).

### *Sensory wizyjne*

Problem wzajemnego zakłócania się sensorów nie występuje w przypadku systemu składającego się z pasywnych kamer. Sensory te pozwalają na zobrazowanie elementów otoczenia (lądowego i powietrznego) przy odpowiednich warunkach oświetleniowych. Światło słoneczne, które w czasie dnia występuje naturalnie w środowisku, oświetla obiekty znajdujące się w otoczeniu. Światło odbite jest rejestrowane przez matrycę kamery. Obraz dwuwymiarowy rejestrowany przez kamerę jest projekcją perspektywiczną świata trójwymiarowego (Ma *et al.*, 2006). Ponadto obraz z kamery pozwala na określenie cech obiektów statycznych i ruchomych. Jednak analiza kontekstowa obrazu z kamery jest bardzo złożona zwłaszcza, że kamera nie dostarcza danych przestrzennych o obiektach.

Kamera w najprostszym przypadku jest używana, aby wyświetlić użytkownikowi obraz do analizy. Bardziej złożony pomiar z użyciem kamery pozwala na automatyczną detekcję krawędzi, obszarów o określonej jasności lub kolorze, segmentację, itp. Złożone algorytmy przetwarzania obrazu pozwalają na wykrywanie, rozpoznawanie i śledzenie wybranych cech (Ben-Tzvi i Xu, 2010). Do takich rozwiązań należy zaliczyć także odometrię wizyjną, która pozwala na nawigację zliczeniową na podstawie danych wizyjnych.

Systemy wizyjne często wykorzystują wiele kamer, co pozwala na określenie odległości między kamerą a wybranym obszarem widocznym w obrazie. W szczególności do wyznaczania głębi w obrazie dedykowana jest stereowizja, która umożliwia znalezienie zależności między punktami odpowiadającymi danemu obiektowi w przestrzeni. Inne podejście zakłada śledzenie określonych punktów w kolejnych klatkach obrazu, gdy kamera jest w trakcie ruchu. W podejściu zwanym SfM (ang. *Structure from Motion*) możliwe jest uzyskanie danych przestrzennych dla śledzonych punktów (Hartley i Zisserman, 2003).

Problem analizy zawartości obrazu z kamery jest w ogólności trudny. Różnorodność i brak pełnej powtarzalności rejestrowanej sytuacji powoduje, że wiele systemów wizyjnych upraszcza zadanie przez wykorzystanie dodatkowych elementów, takich jak oświetlacze czy znaczniki. Dzięki temu problem detekcji związany z niedoświetlonym lub gładkim obiektem, który występuje w stereowizji, jest eliminowany przez dodanie światła o określonej charakterystyce przypominającej szum (Jia *et al.*, 2003; Zhao *et al.*, 2010). Podejście to określane jest jako aktywna stereowizja i działa tylko dla obiektów znajdujących się blisko sensora. Dodatkowe oświetlenie wykorzystywane jest także w jednokamerowych systemach wizyjnych takich jak Microsoft Kinect w celu uzyskania odległości obiektów widocznych w obrazie. Oświetlając scenę światłem strukturalnym, które zawiera specjalny wzór, określić można odległość poprzez analizę deformacji tego wzoru na obiekcie. Można także w prostszy sposób określać położenie obiektu względem znacznika, dodając sztuczne znaczniki do środowiska (Ahrary i Ishikawa, 2006).

Sensory instalowane na robocie umożliwiają nawigację opierając się na położeniu obiektu względem położenia robota. Określenie pozycji w zewnętrznym niezależnym układzie odniesienia wymaga systemu, który obejmuje całe środowisko pracy robota i umożliwia lokalizację. Atutem zewnętrznego rozwiązania jest stały poziom błędu niezależnie od drogi, którą przebył robot.

System GPS (ang. *Global Positioning System*) jest powszechnym systemem globalnego pozycjonowania (Li *et al.*, 2006; Todescato *et al.*, 2016). System ten przy określaniu pozycji wymaga jedynie zainstalowania odbiornika na robocie i zapewnienia dostępności sygnału GPS. Pomiar jest realizowany za pomocą trilateracji, która mierzy pseudoodległość (czasy odbioru odpowiednich kodów zawartych w sygnale) z co najmniej 4 satelitów. Każdy emituje sygnał radiowy zawierający zakodowane dane dotyczące położenia i prędkości satelity oraz czasu, w którym zbocze sygnału jest wysyłane. System ten wymaga precyzyjnego pomiaru czasu w urządzeniach nadawczych i odbiorczych. Po stronie nadawczej instalowane są zegary atomowe, natomiast miniaturowe odbiorniki mają zegary kwarcowe o niskiej precyzji. Dlatego dostępne są dodatkowe systemy takie jak DGPS, WAAS czy EGNOS, które przesyłają poprawki do odbiorników GPS (Dubrovin i Scherbatyuk, 2016). Dokładność systemu GPS dla zastosowań cywilnych z uwzględnieniem poprawek szacuje się na około 3m dla pojedynczego pomiaru. Uśrednienie sygnału pozwala na zmniejszenie błędu poniżej 1m. Taka precyzja występuje przy założeniu, że robot z zainstalowanym odbiornikiem odbiera sygnały z co najmniej 6 satelitów jednocześnie. Podobne zasady działania cechują inne globalne systemy pozycjonowania takie jak rosyjski GLONASS, czy europejski Galileo (Tabibi *et al.*, 2017).

System GPS działa poprawnie, gdy odbiornik znajduje się w otwartej przestrzeni, z dala od wysokich budynków. W pozostałych przypadkach jest on często bezużyteczny. Jego dokładność jest zwykle wystarczająca przy określaniu pozycji dużych obiektów takich jak statek, samolot, kiedy błąd położenia rzędu 1m nie powoduje istotnych konsekwencji w działaniu.

Określanie pozycji w zewnętrznym układzie odniesienia realizowane jest także w sensie lokalnym. Rozwiązania takie oparte są na sygnalizatorach nawigacyjnych (ang. *Navigation Beacons*). Do pasywnych sygnalizatorów można zaliczyć znaczniki naturalne takie jak gwiazdy, charakterystyczne elementy krajobrazu oraz znaczniki sztuczne, np. nabieżniki morskie. Rozwiązania aktywne to radiolatarnie, sygnalizatory świetlne, a także stacje bazowe sieci komórkowej, lub punkty dostępu do sieci bezprzewodowej WiFi. Wewnątrz budynku do określania pozycji może służyć kamera ustawiona tak, aby rejestrowała wybrany obszar z góry. Podobnie odpowiednio ustawiony system Motion Capture może rejestrować położenie i ułożenie dedykowanego znacznika w specjalnie przygotowanym pomieszczeniu (Lee *et al.*, 2016).

Określanie pozycji za pomocą sygnalizatorów nawigacyjnych jest bardzo wygodne. Niemniej wymaga pewnej znajomości środowiska, poznania naturalnych sygnalizatorów lub umieszczenia w środowisku sztucznych sygnalizatorów. W specyficznych sytuacjach sygnał pochodzący z sygnalizatorów nie zawsze jest dostępny (podobnie jak sygnał GPS). W takich wypadkach konieczne jest korzystanie z sensorów pomiarowych określających położenie w układzie odniesienia robota.



## 2.4 Mapowanie

Mapowanie jest procesem, który ma na celu wyznaczenie pozycji elementów znajdujących się w otoczeniu (uwzględniając także robota mapującego) i zaznaczenie ich na mapie. Jednakże przed analizą konkretnej realizacji mapowania należy rozważyć problem w szerszym kontekście. Aby robot mógł stworzyć mapę, konieczna jest znajomość jego pozycji na mapie, zaś wyznaczenie pozycji robota wymaga znajomości mapy. Przez pozycję należy rozumieć położenie i ułożenie robota.

Zagadnienie to jest przykładem problemu „jajka i kury” i jest definiowane w robotyce jako SLAM – jednoczesna lokalizacja i mapowanie (ang. *Simultaneous Localization and Mapping*). Znane jest też jego wcześniejsze określenie jako CML (ang. *Concurrent Mapping and Localization*) (Bailey i Durrant-Whyte, 2006a; Martinelli *et al.*, 2003). Uznaje się, że problem SLAM zdefiniowano wraz z opisem probabilistycznym w 1986 roku na konferencji IEEE w San Francisco dotyczącej automatyki i robotyki. Metody teoretyczne związane z estymacją (szacowaniem) lokalizacji oraz mapy przedstawiali Jim Crowley, Hugh Durrant-Whyte. SLAM ostatecznie zostało uznane za fundamentalny problem dla autonomiczności robota (Bailey i Durrant-Whyte, 2006a).

Określanie położenia i ułożenia może być oparte na danych z czujników omówionych w poprzedniej sekcji. Jest to stosunkowo proste, gdy mapa jest już znana (*a priori*), zaś lokalizacja jest realizowana za pomocą globalnego systemu pozycjonowania (względem zewnętrznego układu odniesienia). Taka sytuacja ma miejsce, np. przy nawigacji samochodowej opartej na systemie GPS zapewniającej stały błąd położenia.

Brak możliwości bezwzględnego określania położenia robota powoduje, że błąd lokalizacji nie jest stały. Ponadto położenie i ułożenie robota przed rozpoczęciem ruchu musi być znane, zaś w celu ograniczenia błędów położenia, należy sprawdzić, czy mapowane elementy (w danej chwili) pokrywają się z elementami istniejącymi na mapie. Podobny problem występuje w sytuacji, gdy *a priori* nie jest dostępna mapa. Tworzenie nowego fragmentu mapy jest obarczone zmiennym błędem, gdyż bazuje na ostatnim położeniu i ułożeniu robota. Dlatego w takim przypadku konieczne jest zarówno określanie położenia na mapie i tworzenie nowego fragmentu mapy.

### 2.4.1. Podział zagadnień SLAM

Od pierwszej definicji zagadnienia SLAM przez lata stworzono wiele podziałów i podejść. Do najważniejszych elementów, które wyróżnia się w rozważaniach dotyczących SLAM, można zaliczyć:

- dostępność pomiarów
- reprezentację niepewności (modeli danych pomiarowych, ruchu i mapowania)
- obecność znaczników na mapie
- zamykanie pętli
- estymację położenia robota i obiektów na mapie.

Ze względu na dostępność danych wyróżnia się *offline* SLAM (*full* SLAM) i *online* SLAM. Podejście pierwsze zakłada, że generowanie mapy oraz całej ścieżki jaką przejechał robot, jest realizowane dopiero po zakończeniu ruchu. Stąd mapa oraz trajektoria robota obliczane są na podstawie pełnych danych pomiarowych (*full* SLAM). Takie podejście jest szczególnie przydatne przy analizie i testowaniu skuteczności algorytmów SLAM w takich samych warunkach. Metoda *online* jest trudniejsza, gdyż dane do mapo-

wania pojawiają się na bieżąco, a mapa tworzona jest na podstawie niepełnych danych o otoczeniu. Algorytm musi wyznaczać aktualny fragment ścieżki, który będzie obciążony większym błędem niż ten sam fragment ścieżki wyznaczany w warunkach *full* SLAM. Podejście *online* SLAM jest ściśle związane z autonomicznością robota. Jego implementacja na platformie robota wymaga działania w czasie rzeczywistym (Goel *et al.*, 1999; Bailey i Durrant-Whyte, 2006a).

Przy opisie poszczególnych elementów obejmujących problem SLAM używa się modeli matematycznych opartych na prawdopodobieństwie, co w literaturze określa się hasłem SLAM stochastyczny (Bailey i Durrant-Whyte, 2006b; Chandra *et al.*, 2010; Martinelli *et al.*, 2003). Opis taki zawiera modele obserwacji, ruchu oraz mapy. Model obserwacji opisuje prawdopodobieństwo dokonania danego pomiaru (obserwacji), gdy znane są lokalizacje robota i znaczników. Model ruchu określa, jak stan robota w poprzedniej chwili czasu wpływa na stan obecny, przy znanym wymuszeniu (sygnale sterującym) robota. Wykorzystując oba modele wyznacza się model mapy, który pozwala na identyfikację położenia danego znacznika (w postaci obszaru o zadanym przedziale prawdopodobieństwa). Dzięki temu można określić, czy znacznik znajduje się w danym obszarze z określonym prawdopodobieństwem (np. nie mniejszym niż 90%).

Ze względu na sposób reprezentacji obiektów jako zgrupowanie znaczników na mapie, wyróżnić można SLAM metryczny i topologiczny. SLAM metryczny wymaga, aby odległości między znacznikami oraz położenie robota na mapie były zgodne z rzeczywistymi odległościami, natomiast SLAM z mapą topologiczną przedstawia obszar w postaci ścieżek między obiektami, co można opisać w postaci grafu (Estrada *et al.*, 2005).

Rozwiązania metryczne realizuje się w postaci obszarów zajętości opisanych równomierną siatką z oczkami w postaci kwadratów lub sześciokątów (na płaszczyźnie). Dyskretyzacja mapy metrycznej pozwala na dodatkowe oznaczenie obszarów wolnych lub zajętych przez przeszkodę. Taka mapa umożliwia odwzorowanie względnie małego otoczenia (Santos *et al.*, 2013).

Mapowanie dużych obszarów wymaga bardzo dużych nakładów obliczeniowych, dlatego też w takim przypadku przydatna jest mapa topologiczna. W rozwiązanych topologicznych węzłach grafu reprezentuje konkretną lokację reprezentującą obiekt (np. przeszkodę), zaś krawędzie określają relacje między innymi węzłami (Bailey i Durrant-Whyte, 2006a; Kuzmin, 2018). Na podstawie takiej mapy można opisać duże obszary oraz łatwo planować przejścia między wybranymi węzłami, natomiast sama mapa topologiczna nie reprezentuje dystansów między tymi węzłami (można je oczywiście odczytać z parametrów krawędzi).

W celu zmniejszenia wymagań na moc obliczeniową tworzy się również mapy oparte na cechach obiektów (ang. *Feature Based Maps*). Algorytmy detekcji wykrywają cechy (naturalne sygnalizatory), które można umieszczać na mapie (Santos *et al.*, 2013).

Reprezentacja obiektów realizowana jest także przy użyciu map hybrydowych, które łączą cechy map metrycznych i map topologicznych. W szczególnych przypadkach, gdy dane charakteryzują się dużą niepewnością pomiaru, można spotkać rozwiązania oparte na mapach hierarchicznych. Podejście hierarchiczne zakłada wyznaczenie lokalnych map zajętości, które służą do odpowiedniej filtracji danych mapowania (Kitajima *et al.*, 2010). Warto zaznaczyć, że w zależności od potrzeb opis otoczenia na mapie jest realizowany na płaszczyźnie (SLAM 2D) lub w przestrzeni trójwymiarowej (SLAM 3D).

Błąd pomiaru położenia zwykle jest mały na początku trajektorii ruchu robota i rośnie wraz z odległością od punktu początkowego. Zamykanie pętli polega na ponownym przejechaniu robota przez punkt startowy lub inny odwiedzony punkt. W chwili zamykania pętli w określonym punkcie, estymowane położenie na mapie jest obciążone istotnym błędem. Przyrównując położenie estymowane z położeniem punktu (startowego) wykrytego podczas zamykania pętli, można w pewnym stopniu skorygować nie tylko chwilowe położenie, ale także całą (wcześniejszą) trajektorię robota. Zamykaniu pętli towarzyszą liczne problemy takie, jak skuteczność i dokładność wykrywania zamknięcia pętli lub obszaru przeszukiwania mapy oraz porównywania i skojarzenia dwóch lokalnych map (czy analizowany fragment mapy był odwiedzony wcześniej), jak również wpływ dokładności ułożenia na detekcję zamykania pętli, itp. W literaturze istnieją różne podejścia mające na celu zwiększenie pewności detekcji zamykania pętli, jednak najnowsze rozwiązania oparte są przede wszystkim na analizie obrazu z kamer.

#### 2.4.2. Mapowanie oparte na wizji

Odwzorowanie otoczenia na mapie istotnie różni się dla przypadku danych 2D i 3D. Mapę dwuwymiarową można tworzyć na podstawie wielu względnie prostych sensorów. Natomiast mapa trójwymiarowa wymaga użycia sensorów takich jak lidar czy kamera, a także fuzji danych z wielu sensorów.

Z powyższego w literaturze można znaleźć głównie algorytmy dedykowane do problemu SLAM zakładające użycie konkretnego sensora, zaś pozostała istotna część dotycząca platformy sprzętowej nie jest określona. Na podstawie przeglądu literatury można dokonać porównania konkretnych metod (Labbé i Michaud, 2014; Angeli *et al.*, 2008; Mahon *et al.*, 2008). Jednak osiągalne opisy konkretnych modeli robotów pozwalają określić co najwyżej, czy dostępne sensory pozwalają na implementację konkretnej metody. Co więcej, wyniki działania metod są przedstawiane na przykładach indywidualnych eksperymentów, co nie pozwala na ilościową ocenę dokładności odwzorowania. Wyjątek stanowią serwisy, np. Middelbury (Scharstein *et al.*, 2014), który ocenia algorytmy stereowizyjne na wybranych parach obrazów. Przegląd rozwiązań zostanie zawężony do metod związanych z problematyką systemów SLAM dedykowanych dla układów kamer.

Obrazy z kamer pozwalają na określenie trójwymiarowych danych opisujących otoczenie oraz przypisanie koloru do danego punktu mapy 3D. Rozwiązania oparte na wizji dotyczące SLAM obejmują systemy:

- jednokamerowe
- stereowizyjne (dwukamerowe)
- wielokamerowe.

Porównanie algorytmów przedstawiono w tab. 2.1, gdzie oprócz liczby kamer uwzględniono, czy algorytm działa w czasie rzeczywistym (przy implementacji na procesor ogólnego przeznaczenia CPU lub procesor graficzny GPU), przetwarza chmurę punktów (ang. *Point Cloud*), a także czy wymaga ciągłości pomiarów.

System jednokamerowy MonoSLAM działający w czasie rzeczywistym został przedstawiony przez Davidsona (2007). Rozwiązanie to zakłada, że w otoczeniu występują naturalne nieruchome znaczniki (obiekty), których odpowiednikami w dziedzinie przetwarzania obrazu są przede wszystkim narożniki (ang. *Corners*). Położenie wykrytych punktów jest śledzone w kolejnych klatkach obrazu. Położenie znaczników w obrazie w ogólności zmienia się wraz ze zmianą położenia kamery. Analiza tych zmian w kolejnych klatkach



Tablica 2.1 Porównanie algorytmów SLAM opartych na danych wizyjnych.

Algorytm	liczba kamer	wymaga ciągłości pomiarów	działa w czasie rzeczywistym	przetwarza chmurę punktów
MonoSLAM (Sparse)	jedna	nie	tak (CPU/GPU)	nie
MonoSLAM (Dense)	jedna	nie	tak (GPU)	nie
Stereo (Sparse)	dwie	tak	tak (CPU/GPU)	nie
Stereo BM (Dense)	dwie	tak	tak (CPU/GPU)	nie
Stereo SGBM (Dense)	dwie	tak	tak (GPU)	nie
EdgeGraph3D	wiele	tak	-	nie
MVS Liu	wiele	tak	nie	nie
MVS Wei	wiele	tak	nie	nie
Multi-view Stereopsis	wiele	tak	nie	tak

obrazu pozwala określić położenie znaczników względem kamery. Technika ta jest znana w postaci SfM, która nie jest bezpośrednio związana z problematyką SLAM (Harris i Pike, 1988). Pierwotny MonoSLAM oparty był na Rozszerzonym Filtrze Kalmana EKF (ang. *Extended Kalman Filter*) i zakładał ruch kamery opisany modelem gaussowskim.

Aby system działał poprawnie konieczna jest wcześniejsza kalibracja kamery. Decyzje o dodaniu nowego znacznika lub usunięciu starego podejmowane są w taki sposób, aby utrzymać poprawne znaczniki w wybranych obszarach i jednocześnie nie przeciążać procesora przetwarzaniem zbyt wielu znaczników. System wymaga także, aby znaczniki były widoczne w kolejnych klatkach obrazu. W przypadku braku ciągłości śledzenia znaczników, algorytm nie pozwala na wyznaczenie położenia wcześniej śledzonych znaczników względem aktualnego położenia kamery. Rozwiązanie to zalicza się do problemu SLAM opartego na cechach (ang. *Feature Based SLAM*), gdzie algorytm detekcji porównuje tylko wybrane cechy w sąsiednich klatkach obrazu (ang. *Sparse Detection Algorithm*). Inne podejścia opierają się na algorytmach detekcji, które porównują wszystkie piksele w sąsiednich klatkach (ang. *Dense Detection Algorithm*) i wymagają dedykowanego sprzętu. Nieznaczące różnice dotyczą głównie koncepcji SLAM, metod detekcji narożników, oraz modelu ruchu kamery, itp. (Stasse *et al.*, 2006; Lategahn *et al.*, 2011).

Systemy stereowizyjne używane w SLAM nie wymagają, aby znaczniki były widoczne w sąsiednich klatkach. Stereowizja zakłada obraz z dwóch kamer przesuniętych względem siebie o stałą odległość, więc w każdym przypadku znaczniki są w obu obrazach. W typowej stereowizji położenie i ułożenie kamer względem siebie nie ulega zmianie, więc analiza przesunięć obiektów widocznych w obu obrazach pozwala na tworzenie obrazu 3D. W literaturze istnieją liczne lokalne i globalne podejścia do wyznaczania głębi w obrazie stereowizyjnym. Wśród lokalnych można wyróżnić proste porównywanie ob-



szarów BM (ang. *Block Matching*) oraz wersje wzbogacone o liczne algorytmy poprawy głębi (Hirschmuller, 2008). Wśród globalnych lub semi-globalnych można znaleźć metody oparte na analizie informacji wzajemnej (SGBM (ang. *Semi-Global Block Matching*)), energii, zmiennych okien analizy itp. (Futami *et al.*, 2002; Stefano *et al.*, 2004; Diebel i Reutersward, 2004; Hartley i Zisserman, 2003; Ma *et al.*, 2006).

Rozwiązania wielokamerowe, w szczególności składające się z trzech kamer, pozwalają na uzyskanie lepszej głębi niż w typowym systemie dwukamerowym, ponieważ dodatkowa kamera może rozwiązać problemy z niejednoznacznością związaną z parą stereo. Wyróżnić można systemy ze statycznymi kamerami, gdzie położenie kamer dobrane jest tak, aby kolejne pary kamer obserwowały obiekty z innej strony. Dzięki temu elementy niewidoczne dla danej pary kamer mogą być obserwowane przez inną parę. W podejściu tym korzysta się z filtracji i łączenia chmur punktów uzyskanych ze stereowizji a pochodzących z kolejnych par kamer (Jia *et al.*, 2003; Chrysostomou *et al.*, 2010). Rozważa się też rozwiązania z wieloma ruchomymi kamerami generującymi chmury punktów zgodnie z zasadą MonoSLAM. Różnica wynika z tego, że MonoSLAM analizuje klatki z jednej kamery w kolejnych chwilach, zaś w rozwiązaniu wielokamerowym analizie poddawane są klatki z różnych kamer w tej samej chwili (Qian i Lo, 2015). Podejście wielokamerowe łączy zatem analizę MonoSLAM i stereowizję (Martinez-Gomez *et al.*, 2014), czego przykładem są metody Multi-view Stereopsis (Furukawa i Ponce, 2007), MVS Liu (Liu *et al.*, 2017), MVS Wei (Wei *et al.*, 2016). W analizowanych przykładach, w oparciu o zestawienie z tab. 2.1 można zauważyć, że wraz ze wzrostem liczby kamer tworzone są bardziej złożone algorytmy, które dodatkowo przetwarzają chmurę punktów i nie działają w czasie rzeczywistym nawet na procesorze GPU.

Inny problem SLAM związany jest z zadaniem porównania różnych chmur punktów. Problem ten dotyczy nie tylko kamer ale także innych sensorów, które dostarczają dane mapowania otoczenia. Aby uzyskać spójną mapę należy zapewnić, aby dwie chmury punktów uzyskane dla różnych położzeń sensora zamocowanego na robocie, odpowiednio się nakładały. W szczególności należy wyznaczyć niezbędne transformacje polegające na przesunięciu i obrocie pomiędzy dwiema chmurami punktów, których większość opisuje te same elementy przestrzeni. Wśród ogólnych podejść do opisu kształtów opisanych w przestrzeni trójwymiarowej wyróżnia się porównywanie zbiorów punktów, krawędzi (Bignoli *et al.*, 2018) linii, linii łączonych (ang. *Polylines*), trójkątów i innych krzywych parametrycznych itp. (Besl i McKay, 1992). Aktualne rozwiązania są oparte na algorytmie iteracyjnego najbliższego punktu (ang. *Iterative Closest Point*) bądź jego modyfikacji (Rusinkiewicz i Levoy, 2001; Tomono, 2009). Metoda ta pozwala zminimalizować różnice między chmurami punktów. Jednak dla chmur zawierających dużą liczbę 'odstających' błędnych punktów (ang. *Outliers*), metoda może prowadzić do grubych błędów (Gelfand *et al.*, 2003; Xu *et al.*, 2016).

#### **2.4.3. Wybrane realizacje mapowania 3D realizowane za pomocą robotów mobilnych**

W literaturze można znaleźć realizacje złożonych systemów mapowania 3D dla konkretnych robotów mobilnych. Różnorodność typów robotów, napędów, sensorów, itd., powoduje, że zwykle trudno jest przeprowadzić prostą klasyfikację. Dlatego poniżej przedstawimy tylko kilka wybranych rozwiązań.



Względnie prosty kompletny system mapowania przy użyciu robota mobilnego przedstawiono w pracy (Georgiades *et al.*, 2004). Robot amfibia o nazwie AQUA wykonuje zadanie SLAM podczas poruszania się po łodzi oraz w wodzie. Określanie lokalizacji oparto przede wszystkim na kombinacji danych z GPS oraz IMU, przy czym dla ruchu pod wodą wykorzystuje się także sensory akustyczne. Dodatkowo robota zaopatrzone w pierwszym podejściu w pojedynczą kamerę, a następnie w trzy kamery. Stosując model pola losowego Markova (ang. *Markov Random Field*) do interpretacji danych wizyjnych wyznaczana jest głębina. Po czym uzyskane chmury punktów ('niezachodzące' bądź 'zachodzące' na siebie) łączone są na podstawie danych inercyjnych z użyciem algorytmu dopasowującego *Bundle Adjustment*. Założona możliwość zmian warunków pracy prowadzi do zastosowania odmiennej lokalizacji w wodzie i na łodzi.

Jedno z bardziej interesujących rozwiązań kompletnego systemu mapowania zostało zaproponowane przez Sim *et al.* (2006). System ten wykorzystuje stereowizję, zasadzając się na filtrze cząsteczkowym, opartym na twierdzeniu Rao-Blackwella, który znany jest też pod nazwą FastSLAM. Wyznaczanie głębi realizowane jest przez przetwarzanie punktów charakterystycznych w obrazie za pomocą deskryptora SIFT (ang. *Scale-Invariant Feature Transform*), który realizuje podejście selektywne (ang. *Sparse*) wykrywające znaczniki. Rozwiązanie posiada liczne opcje związane z przetwarzaniem obrazu, np. mapowanie z użyciem jednej kamery (odpowiednik monoSLAM), dwóch kamer (stereo) lub estymację ruchu własnego (odometrię wizyjną). W pracy tej użyto także rozszerzonego filtra Kalmana (EKF) do śledzenia indywidualnych (zaakceptowanych) znaczników. W celu potwierdzenia skuteczności działania przeprowadzono eksperymenty na robocie RWI B14 oraz skalibrowanej fabrycznie komercyjnej kamerze stereo BumbleBee.

Nieco inne podejście prezentuje się w pracy (Belter *et al.*, 2012), gdzie proponowany system dostosowany jest do robota nożnego, który sposobem poruszania się przypomina pająka. Do odwzorowania terenu wykorzystuje się dane ze skanera laserowego URG-04LX oraz skalibrowaną kamerę stereowizyjną STOC firmy Videre Design. Praca uwzględnia niepewność związaną z pomiarami stereowizyjnymi (o stałej bazie). Finalna mapa odwzorowująca otoczenie oparta jest na (dwuwymiarowej) siatce zajętości oraz rozszerzeniu tej siatki o trzeci wymiar, wysokość (ang. *Elevation Map*). Estymacja wysokości opiera się na filtracji Kalmana. Dodatkowe metody poprawiające odwzorowanie, np. do aproksymacji/interpolacji niezdefiniowanych obszarów wykorzystuje algorytm rojowy. Metodę sprawdzono poprzez symulację ruchu mechanizmu kroczącego z wykorzystaniem robota przemysłowego. Wykonano także eksperyment w warunkach zewnętrznych.

Najnowsze badania obejmują sensory RGBD do mapowania opartego na danych sensorycznych, które zawierają obraz kolorowy (RGB) oraz głębiny (ang. *Depth*). Do najczęściej spotykanych systemów RGBD należy zaliczyć MS Kinect, Intel RealSense, Primesense Carmine, ASUS Xtion (Zollhöfer *et al.*, 2018). System przedstawiony w pracy (Belter *et al.*, 2016) stosuje dwa takie sensory oraz konfigurację z jednym sensorem RGBD wraz z systemem stereowizyjnym. Podobnie jak w powyżej przedstawianym przykładzie (Belter *et al.*, 2012), również ten system dokonuje mapowania w oparciu o odrębny pomiar podłoża oraz obiektów znajdujących się w otoczeniu. Przeprowadzone testy wykazały, że odpowiednio przetworzone dane z sensorów Asus Xtion i PrimeSense Carmine prowadzą do lepszych wyników niż system (Belter *et al.*, 2012). Dobrą stroną sensorów RGBD jest możliwość łączenia ich pomiarów, co ukazano w pracy (Walas *et al.*, 2016), gdzie użyto kamery TOF (ang. *Time of Flight*) oraz ASUS Xtion. Ze względu na różne zakresy mapo-

wania oraz inny charakter niepewności pomiarowej, odpowiednia fuzja danych umożliwiła poprawę odwzorowania otoczenia.

Opisy kompletnych systemów często nie uwypuklają problemów związanych z ograniczeniami sprzętowymi. Jak widać z dokonanego przeglądu rozwiązań, istnieje wiele metod opartych na wizji, które odwzorowują otoczenie. Wśród nich można znaleźć także kompletne systemy odwzorowujące otoczenie z użyciem stereowizji. Nie natknięto się jednak na kompletne rozwiązanie oparte na stereowizji ze zmienną bazą. Koncept zmiennej bazy znaleźć można co prawda w istniejących rozwiązaniach, ale najwyraźniej nie są one dedykowane dla robotów mobilnych. Przykład zastosowania zmiennej bazy stereo dotyczy suwnicy (Nakabo *et al.*, 2005), bądź bliżej nieokreślonego mechanizmu przesuwającego kamerę (Gallup *et al.*, 2008). W niniejszej pracy zmienna baza uzyskiwana jest dzięki przemieszczaniu się robota, zaś jednoczesna estymacja położenia i ułożenia robota stanowi drugi istotny element proponowanego rozwiązania. Ponadto nowym elementem (niespotykanym w literaturze) jest wykorzystanie danych przyspieszenia do estymowania położenia kamery.

## **2.5 Podsumowanie**

W rozdziale przedstawiono przegląd rozwiązań dotyczących najważniejszych zagadnień związanych z tematyką poruszaną w pracy. Scharakteryzowano roboty używane do mapowania otoczenia, wyróżniono cechy robota istotne z perspektywy implementacji i możliwości testowania. Opisano napęd robota kołowego i jego wpływ na możliwości mapowania. Następnie omówiono sensory potrzebne do nawigowania robotem, które określają położenie i ułożenie robota w przestrzeni oraz są dedykowane operacji mapowania (łącznie z sensorami wizyjnymi). Ponadto omówiono zagadnienie mapowania skupiając się na problemie SLAM z wyróżnieniem rozwiązań opartych na systemach wizyjnych.

W dalszej części pracy nacisk położymy na charakterystykę problemów związanych z możliwościami i dostępnością platform robotów (jednostek nośnych) oraz sensorów. Rozważane będą zadania determinacji położenia i ułożenia kołowego robota mobilnego w oparciu o sensory inercyjne oraz mapowania opartego na stereowizji.

---

## TEORETYCZNE PODSTAWY POMIARÓW

Problem mapowania środowiska przez robota mobilnego jest ściśle powiązany z określeniem położenia i ułożenia robota. Sposób, w jaki wyznaczana jest pozycja robota, dokładność pomiaru, wpływa na precyzję odwzorowania obiektów na mapie. Proces jest zależny od typu i jakości sensorów oraz zastosowanej metody pomiaru. Wybór konkretnego rozwiązania podyktowany jest najczęściej celem, jaki ma realizować robot, dostępnością czujników oraz ograniczeniami, jakie dany robot ma spełniać. Dlatego na wczesnym etapie projektowania konieczny jest wybór konkretnego podejścia do określania położenia i ułożenia, które będzie spełniało założone ograniczenia projektowe.

Analizując dostępne rozwiązania dotyczące określania pozycji robota na potrzeby mapowania zauważyć można różnorodność rozwiązań nawet dla konkretnego typu robota (Nez *et al.*, 2016; Czubenko *et al.*, 2015; Joubert *et al.*, 2016). Jednostki latające i pływające wymagają odmiennego podejścia niż jednostki jeżdżące bądź kroczące. Duże roboty pracujące w warunkach zewnętrznych mogą korzystać z satelitarnych (globalnych) systemów pozycjonowania, gdzie błąd położenia jest mniejszy niż rozmiar robota. Z kolei małe jednostki zwykle wymagają innego podejścia, gdyż błąd ich położenia w systemie GPS może być większy niż ich wymiary (Tabibi *et al.*, 2017). Co więcej, ocena poprawności i skuteczności mapowania dużych jednostek jest utrudniona, gdyż wymaga też dużych przestrzeni laboratoryjnych. Małe jednostki można testować w stosunkowo niewielkiej przestrzeni w powtarzalnych warunkach, bez dużych nakładów finansowych.

Biorąc pod uwagę powyższe ograniczenia, w niniejszej pracy problem określania położenia i ułożenia rozważany będzie przy użyciu sensorów inercyjnych. Generalnie czujniki te pozwalają na stosunkowo dokładne określenie ułożenia robota w przestrzeni (Tzafestas, 2014). Ostatecznie testy dokładności wykażą, jaka jest precyzja wyznaczania ułożenia, a także pozwolą określić, czy jest ona wystarczająca dla potrzeb mapowania. Natomiast pomiar położenia z zastosowaniem sensorów inercyjnych wymaga znajomości ułożenia. Jest on w ogólności mało precyzyjny, gdyż opiera się na nawigacji zliczeniowej. Niemniej dla małych odległości metoda ta może dać dobre wyniki. Warto ją zbadać także dlatego, że istnieją źródła literaturowe, które ogólnie negują przydatność czujników inercyjnych w określaniu położenia (Siegwart i Nourbakhsh, 2004; Tzafestas, 2014) oraz liczne źródła, które potwierdzają ich przydatność (Albarbar *et al.*, 2009; Pang i Liu, 2001; Ang i Khoo, 2004; Gilbert *et al.*, 2010).

### 3.1 Sensory inercyjne i ich charakterystyka

Sensory inercyjne mierzą wielkości fizyczne związane z bezwładnością robota. W sekcji 2.3.2 opisano istniejące rozwiązania nawigacji robota z użyciem sensorów inercyjnych. Czujniki te są całkowicie pasywne, czym wyróżniają się na tle innych rozwiązań, zaś w przeciwieństwie do silników krokowych i enkoderów są one wrażliwe na przemieszczenie także w sytuacji, gdy koło traci kontakt z podłożem.

Do sensorów inercyjnych generalnie należy zaliczyć akcelerometr, magnetometr, żyroskop oraz altymetr (wysokościomierz). Czujniki te są dostępne w postaci jednostki IMU, która jest dedykowana przede wszystkim do określania ułożenia. W zależności od tego, czy pomiar ułożenia jest potrzebny do utrzymania równowagi lub stałego kierunku, czy odwzorowania dynamiki obrotu, zwykle jeden z wymienionych sensorów odgrywa kluczową rolę (Borenstein *et al.*, 1997; Grygiel i Bieda, 2014).

#### 3.1.1. Pomiar ułożenia

Pojedynczy akcelerometr mierzy przyspieszenie obiektu w jednym określonym kierunku. Typowa jednostka IMU zawiera 3-osiowy akcelerometr, który pozwala na pomiar przyspieszeń w przestrzeni kartezjańskiej. Akcelerometr mierzy przyspieszenia związane z siłami, które działają na obiekt, w tym także przyspieszenie ziemskie (Joubert *et al.*, 2016). Zakładając, że ruch jednostki ma charakter czysto obrotowy, pomiar przyspieszenia ziemskiego pozwala ustalić odchylenie obiektu od pionu. Dlatego określanie ułożenia na potrzeby zadań, takich jak utrzymanie równowagi przez robota, opiera się głównie na danych z akcelerometru (Godfrey *et al.*, 2007).

Magnetometr mierzy natężenie pola magnetycznego w danym kierunku. Aby określić ułożenie obiektu potrzebny jest 3-osiowy magnetometr, który mierzy natężenie pola wzdłuż 3 prostopadłych osi w przestrzeni kartezjańskiej. W idealnym przypadku, z dala od źródeł pola magnetycznego i ferromagnetyków, magnetometr pozwala na określenie ułożenia względem magnetycznej północy, która dla danej lokalizacji jest generalnie stała (Titterton i Weston, 2005). W przypadku obiektów pokonujących duże odległości (rzędu tysięcy kilometrów), jak samoloty czy statki, należy uwzględnić deklinację magnetyczną, która dla danej lokalizacji określa kąt między południkiem magnetycznym a geograficznym. Jest to istotne także, gdy tworzona jest mapa kartograficzna (odwzorowana względem północy geograficznej). Na potrzeby tworzenia lokalnej mapy zwykle można pominąć deklinację magnetyczną. W przypadku, gdy w otoczeniu magnetometru występują ferromagnetyki nieruchome względem czujnika, konieczne jest uwzględnienie dewiacji kompasu, która wyraża różnicę między kursem kompasowym (mierzonym) a magnetycznym. Po uwzględnieniu opisanych problemów, magnetometr pozwala opisać aktualne ułożenie względem północy magnetycznej. Jednakże pomiar można łatwo zakłócić ruchomym silnym ferromagnetykiem. Ponadto błędy obliczeniowe istotnie wpływają na dokładność wyznaczania północy (kompasowej), która zmienia się w zależności od ułożenia obiektu.

Żyroskop dostarcza danych dotyczących prędkości kątowej jednostki. W szczególności żyroskop 3-osiowy pozwala określić zmianę ułożenia w trakcie dowolnego obrotu. Dlatego żyroskop jest szczególnie przydatny, gdy istotne jest określenie dynamiki obrotu jednostki. Jednakże szum danych z żyroskopu powoduje, że prędkość obrotowa nieruchomego robota jest niezerowa, co można zobrazować jako drgania jednostki. Wyznaczenie aktualnego ułożenia (położenia kątowego) robota na podstawie danych o prędkości kątowej (obrotowej)

wymaga całkowania danych. Proces ten powoduje, że przy danych obciążonych szumem pojawia się dryf. Dlatego żyroskop bez danych np. z magnetometru jest użyteczny przy opisie dynamiki pojedynczych, stosunkowo krótkich, obrotów (Edwan *et al.*, 2012).

Wysokościomierz wchodzi w skład typowego IMU, jednak nie jest to czujnik ściśle związany z inercją jednostki. Interpretacja wysokości oparta jest na pomiarze ciśnienia, dlatego w przypadku lądowego robota mobilnego wysokościomierz zwykle nie jest przydatny przy wyznaczaniu pozycji (Bekir, 2007).

Dane z pojedynczego sensora wchodzącego w skład IMU pozwalają na określenie ułożenia, jednak precyzja takiego pomiaru może być mała. Nawet dla jednostki nieruchomej dane z magnetometru mogą sugerować, że obiekt jest w ruchu (drga). Również podczas płynnego ruchu obiektu mogą one wskazywać postęp skokowy. Żyroskop sugeruje małe zmiany ułożenia robota nieruchomego ze względu na dryft danych, przez co z upływem czasu niepewność pomiaru wzrasta. Akcelerometr jest czuły na odchylenie obiektu od pionu, nie dostarcza jednak danych związanych z obrotem obiektu względem osi pionowej. Co więcej, gdy jednostka nośna podczas obrotu ulega dodatkowemu przesunięciu, wyznaczone chwilowe ułożenie będzie obciążone błędem tym większym, im większe jest przyspieszenie powodujące przesunięcie. Wykorzystanie jednocześnie danych z magnetometru, żyroskopa i akcelerometru pozwala na otrzymanie dokładniejszych estymat ułożenia obiektu, niemniej tego rodzaju dane wymagają odpowiedniej filtracji i fuzji (Li *et al.*, 2006; Bekir, 2007).

### **3.1.2. Pomiar położenia**

Wyznaczenie położenia na podstawie mierzonego przyspieszenia oprzeć można na całkowaniu sygnału przyspieszenia. Pojedyncze całkowanie przyspieszenia pozwala na ocenę prędkości, zaś podwójne całkowanie daje w wyniku położenie. Z perspektywy określania położenia sytuację komplikuje fakt, że akcelerometr mierzy nie tylko rzeczywiste przyspieszenie obiektu, ale także przyspieszenie ziemskie. Dlatego też przed całkowaniem konieczne jest usunięcie wpływu przyspieszenia grawitacyjnego, które jest stałe w ziemskim (zewnętrznym) układzie odniesienia. Innymi słowy, aby wyznaczyć położenie w oparciu o akcelerometr, trzeba najpierw wyznaczyć dokładnie ułożenie obiektu, aby na tej podstawie usunąć wpływ wektora grawitacji (Siegwart i Nourbakhsh, 2004). Magnetometr i żyroskop nie są jednak bezpośrednio przydatne przy wyznaczaniu przemieszczenia robota. Potrzebne są inne sensory, pełniące kluczową rolę w określaniu ułożenia, bez którego generalnie nie jest możliwe mapowanie oraz usuwanie wpływu przyspieszenia ziemskiego. Dalsza analiza wyznaczania położenia i ułożenia jednostki na podstawie danych z sensorów IMU wymaga matematycznego opisu transformacji obiektu w przestrzeni.

## **3.2 Transformacja układu odniesienia**

Określenie ułożenia obiektu w przestrzeni jest związane z występowaniem oddziaływań o charakterze wewnętrznym (lokalnym), które są stałe w układzie odniesienia robota oraz oddziaływań zewnętrznych, które pozostają niezmiennie w zewnętrznym układzie odniesienia. Grawitacja i północ magnetyczna mają charakter zewnętrzny, natomiast sensory inercyjne zainstalowane na jednostce nośnej wykonują pomiar w wewnętrznym układzie odniesienia (robota).

W układzie współrzędnych kartezjańskich transformację między zewnętrznym (globalnym) układem współrzędnych (ang. *World Coordinates*) a lokalnym układem odniesienia (ang. *Local Coordinates*) można zapisać w następującej postaci:

$$\begin{bmatrix} p_x^w \\ p_y^w \\ p_z^w \end{bmatrix} = \mathcal{T}_l^w \begin{bmatrix} p_x^l \\ p_y^l \\ p_z^l \end{bmatrix} \quad (3.1)$$

gdzie  $p_x^w$  oznacza położenie względem osi  $x$  we współrzędnych zewnętrznych  $w$ , a  $p_x^l$  oznacza położenie względem osi  $x$  we współrzędnych obiektu lokalnych  $l$ . Położenie względem osi  $y$  i  $z$  zapisano w sposób analogiczny. Transformacja  $\mathcal{T}_l^w$  jest funkcją przejścia między lokalnym układem odniesienia  $l$  (np. robota) a układem zewnętrznym  $w$ .

Ogólną afiniczną transformację punktu między wewnętrznym a zewnętrznym układem odniesienia można zapisać w postaci złożenia rotacji wokół osi  $x$ ,  $y$ ,  $z$  i translacji względem tych osi. Translacja punktu  $p$  we współrzędnych lokalnych a następnie obrót tego punktu można zapisać w postaci:

$$\begin{bmatrix} p_x^w \\ p_y^w \\ p_z^w \end{bmatrix} = \mathbf{R} \begin{bmatrix} p_x^l + t_x \\ p_y^l + t_y \\ p_z^l + t_z \end{bmatrix} \quad (3.2)$$

gdzie  $\mathbf{R}$  jest macierzą rotacji zaś  $t_x, t_y, t_z$  to składowe wektora translacji  $\mathbf{T}$ . W powyższej formule najpierw wykonywana jest translacja później rotacja. Jednak kolejność wykonywania operacji, w których skład może wchodzić wiele rotacji i translacji, nie ma znaczenia pod warunkiem, że wybrany opis matematyczny uwzględnia transformacje w zewnętrznym układzie odniesienia. Dlatego też ważny jest wybór właściwego opisu matematycznego, w szczególności opisu rotacji (Siegwart i Nourbakhsh, 2004; Titterton i Weston, 2005).

### 3.2.1. Rotacja i macierze obrotu

W układzie kartezjańskim rotacja obiektu jest przedstawiana w postaci pojedynczych obrotów  $\mathbf{R}_x, \mathbf{R}_y, \mathbf{R}_z$  wokół osi obiektu odpowiednio  $x, y, z$ . Każda z rotacji jest niezależna, niemniej kolejność wykonania poszczególnych obrotów w takim zapisie wpływa na końcowe położenie obiektu. Dla prawoskrętnego układu współrzędnych rotacja  $\phi$  względem osi  $x$  jest przechyleniem (ang. *roll*). Obrót  $\theta$  względem osi  $y$  jest to pochylenie (ang. *pitch*), zaś rotacja  $\psi$  względem osi  $z$  to myśkowanie (ang. *yaw*). Obroty te można zapisać w postaci następujących macierzy obrotu:

$$\mathbf{R}_x(\phi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \quad (3.3)$$



$$\mathbf{R}_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \quad (3.4)$$

$$\mathbf{R}_z(\psi) = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (3.5)$$

Złożenie rotacji  $\mathbf{R}_x$ ,  $\mathbf{R}_y$ ,  $\mathbf{R}_z$  w jedną macierz można zapisać zgodnie z wyrażeniem (3.6). Dla uproszczenia zapisu, funkcje trygonometryczne zostały podane w skróconej postaci ( $s\alpha = \sin \alpha$ ,  $c\alpha = \cos \alpha$ ).

$$\begin{aligned} \mathbf{R}_{xyz} &= \mathbf{R}_x(\phi)\mathbf{R}_y(\theta)\mathbf{R}_z(\psi) = \\ &= \begin{bmatrix} c\theta c\psi & c\theta s\psi & -s\theta \\ -c\phi s\psi + s\phi s\theta s\psi & c\phi c\psi + s\phi s\theta s\psi & s\phi c\theta \\ s\phi s\psi + c\phi s\theta c\psi & -s\phi c\psi + c\phi s\theta s\psi & c\phi c\theta \end{bmatrix} \end{aligned} \quad (3.6)$$

Macierze powyższe są unitarne, dzięki czemu odwrotna rotacja jest opisana transpozycją macierzy pierwotnej rotacji. Ponadto macierze te należą do tzw. specjalnej grupy obrotu  $SO(3)$ . Rotację odwrotną można zapisać jako:

$$\mathbf{R}_{xyz}(\phi, \theta, \psi)^{-1} = \mathbf{R}_z(\psi)^{-1}\mathbf{R}_y(\theta)^{-1}\mathbf{R}_x(\phi)^{-1} \quad (3.7)$$

gdzie kolejność poszczególnych rotacji jest odwrócona. Powyższe macierze reprezentują obroty według kątów Eulera, przez co zapis jest prosty i czytelny. Przy założeniu, że robot kołowy porusza się co najwyżej po pagórkowatej powierzchni i nie wykonuje myśkowania (obrotów względem osi pionowej) przekraczającego zakres  $(-180^\circ, +180^\circ)$ , opis ten może być wystarczający. Warto jednak zaznaczyć, że przy używaniu kątów Eulera w przypadku robota, który może dowolnie zmieniać swoje ułożenie, napotkać można na problem blokady przegubu (ang. *Gimbal Lock*). Obroty względem osi można interpretować jako ruch trzech niezależnych przegubów. Przeguby te pokrywają się z płaszczyznami  $xz$ ,  $zy$ ,  $yx$  i w ogólności mogą się obracać niezależnie. Jednak na skutek obrotu obiektu o  $90^\circ$  lub  $270^\circ$  względem danej osi, dwie płaszczyzny przegubów pokrywają się i przeguby się blokują. Kolejny obrót jednego z zablokowanych przegubów powoduje także ruch drugiego przegubu skutkiem czego następuje utrata jednego stopnia swobody podczas obrotu (Titterton i Weston, 2005). Z perspektywy opisu matematycznego problem zablokowania przegubu jest widoczny w postaci osobliwości przy wyznaczaniu kątów w oparciu o macierz  $\mathbf{R}_{xyz}$  (3.6).

Aby zapobiec problemowi blokady przegubu należy użyć innego bardziej zaawansowanego opisu matematycznego. Wśród dostępnych rozwiązań można spotkać zapis obrotów w postaci kwaternionów lub macierzy kosinusów kierunkowych DCM (ang. *Direct Cosine Matrix*). Ze względu na szybkość obliczeń w grafice 3D obroty często realizuje się przy użyciu kwaternionów (Bekir, 2007), zatem w projektowanym rozwiązaniu należy użyć zapisu kwaternionowego.

### 3.2.2. Kwaterniony

Kwaterniony są uogólnieniem liczb zespolonych. Umożliwiają globalną parametryzację specjalnej grupy obrotu  $SO(3)$  zapisanej za pomocą czterech elementów: jednostki skalarnej „1” i trzech liczb zespolonych  $i, j, k$  (Kozłowski *et al.*, 2012; Titterton i Weston, 2005), które spełniają warunek:

$$i^2 = j^2 = k^2 = ijk = -1 \quad (3.8)$$

Pojedynczy element ze zbioru kwaternionów jest definiowany w postaci:

$$\mathbf{q} = (q_0, \mathbf{q}) = (q_0, q_1, q_2, q_3) = q_0 + q_1i + q_2j + q_3k \quad (3.9)$$

gdzie  $q_0 \in \mathbb{R}$  jest skalar, a  $\mathbf{q} = (q_1, q_2, q_3) \in \mathbb{R}^3$  jest wektorem urojonym tworzącym kwaternion. Użycie liczby zespolonej do opisu obrotu względem danej osi jest wygodne, gdyż jej argument zmienia się w przedziale  $(-\pi, \pi)$  w sposób ciągły (nie występują skoki dla skrajnych wartości, np. między  $360^\circ$  a  $0^\circ$ ).

Dla dwóch kwaternionów  $\mathbf{q} = (q_0, \mathbf{q})$  i  $\mathbf{p} = (p_0, \mathbf{p})$  możliwe są przekształcenia algebraiczne, takie jak dodawanie lub odejmowanie:

$$\mathbf{q} \pm \mathbf{p} = (q_0 \pm p_0, \mathbf{q} \pm \mathbf{p}) \quad (3.10)$$

oraz operacje mnożenia, które nie jest przemienne, podobnie jak kolejność mnożenia macierzy rotacji (3.7). Operacja mnożenia ma postać:

$$\mathbf{q} \cdot \mathbf{p} = (q_0p_0 - \mathbf{q} \bullet \mathbf{p}, q_0\mathbf{p} + p_0\mathbf{q} + \mathbf{q} \times \mathbf{p}) \quad (3.11)$$

gdzie operacje ‘ $\bullet$ ’ i ‘ $\times$ ’ oznaczają odpowiednio iloczyn skalarny i wektorowy. Kwaternion odwrotny, używany do określenia rotacji odwrotnej, opisany jest następująco:

$$\mathbf{q}^{-1} = \mathbf{q}^* / \|\mathbf{q}\|^2 \quad (3.12)$$

$$\|\mathbf{q}\| = \sqrt{\mathbf{q}\mathbf{q}^*} = \sqrt{q_0^2 + q_1^2 + q_2^2 + q_3^2}$$

gdzie  $\mathbf{q}^* = (q_0, -\mathbf{q})$  jest kwaternionem sprzężonym, a  $\|\mathbf{q}\|$  jest normą kwaternionu  $\mathbf{q}$ .

Zdefiniowany powyżej kwaternion może być użyty do opisu dowolnej rotacji w przestrzeni. Wcześniejszy opis rotacji z wykorzystaniem macierzy rotacji (3.6) wymagał złożenia 3 obrotów wokół ortogonalnych osi  $x, y, z$ . Tę samą operację polegającą na pojedynczym obrocie o kąt  $\alpha$  wokół jednej osi zdefiniowanej za pomocą wektora  $\mathbf{u}$  można opisać za pomocą kwaternionu:

$$\mathbf{q} = (q_0, \mathbf{q}) = \left( \cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right) \mathbf{u} \right) \quad (3.13)$$

Na podstawie powyższego, rotację wokół osi  $x$  można zapisać jako

$$\mathbf{q}_x(\alpha) = \left( \cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right), 0, 0 \right)_{|\mathbf{u}=(1, 0, 0)} \quad (3.14)$$

Analogicznie obrót wokół osi  $y$  można wyrazić za pomocą kwaternionu

$$\mathbf{q}_y(\alpha) = \left( \cos\left(\frac{\alpha}{2}\right), 0, \sin\left(\frac{\alpha}{2}\right), 0 \right)_{|\mathbf{u}=(0, 1, 0)} \quad (3.15)$$



zaś obrót wokół osi  $z$  zrealizować można używając kwaternionu

$$\mathbf{q}_z(\alpha) = \left( \cos\left(\frac{\alpha}{2}\right), 0, 0, \sin\left(\frac{\alpha}{2}\right) \right)_{|\mathbf{u}=(0, 0, 1)} \quad (3.16)$$

Bekir (2007) szerzej opisuje własności kwaternionów, dowody wyprowadzeń a także konwersje między różnymi formami opisu rotacji.

### 3.2.3. Szum sensorów inercyjnych

Dysponując sensorami inercyjnymi oraz odpowiednim aparatem matematycznym, który pozwala opisać rotację względem różnych układów odniesienia, wyznaczyć można ułożenie i położenie obiektu. Niemniej precyzja ułożenia i położenia jest zależna od typu sensora oraz jakości jego wykonania. Obecnie spotkać można niszowe rozwiązania wysokiej jakości dla zastosowań militarnych lub lotniczych, które oparte są na układach optycznych, systemach kalibracji temperaturowej oraz elementach tłumiących zakłócenia magnetyczne (Shortelle *et al.*, 1998; Sheard *et al.*, 2008). Rozwiązania te charakteryzują się stosunkowo niewielkim szumem, są drogie i trudno dostępne.

Ze względu na miniaturowe rozmiary oraz niskie koszty produkcji, sensory MEMS stały się powszechnie dostępne. Są masowo używane w telefonach, tabletach, aparatach fotograficznych. Charakteryzują się dużo większym szumem w porównaniu z drogimi rozwiązaniami używanymi w lotnictwie i militariach. Aby upewnić się, czy precyzja danego czujnika jest wystarczająca do wyznaczania ułożenia i położenia na potrzeby mapowania, konieczne jest omówienie parametrów opisujących szum sensorów MEMS oraz wpływu szumu danego sensora na dryft.

Istnieje wiele czynników, które wpływają na szum sensorów MEMS (Mohd-Yasin *et al.*, 2009). Szczególnie istotne, choć mało problematyczne, są współczynnik skalowania (ang. *Scale Factor*) oraz przesunięcia, czyli uchybu (ang. *Bias*). Współczynnik skalujący, związany z czułością sensora, określanym jest jako stosunek zmian na wyjściu czujnika do zmian na jego wejściu, które ma być mierzone. W konsekwencji, zakres sygnału mierzonego przez sensor oraz poziom szumu są zmienne, zależne od współczynnika skalującego. Problem przesunięcia można opisać jako niezerową wartość, o jaką różni się wyjście czujnika od oczekiwanej wartości rzeczywistej. W praktyce uchyb ten mierzy się przez pewien okres czasu dla zadanych warunków pracy i aproksymuje się go jedną stałą wartością. Współczynnik skalujący i przesunięcie są związane z czynnikami materiałowymi i konstrukcyjnymi sensora, np. efektami międzyosiowymi (ang. *Cross-Axis Effects*), nieliniowością i histerezą. Problemy sensorów MEMS są także związane z warunkami pracy, które dotyczą głównie temperatury i ciśnienia (Albarbar *et al.*, 2009). Można tu wyróżnić takie parametry, jak zmienność czułości pod wpływem temperatury (ang. *Sensitivity Change vs. Temperature*) i zmianę poziomu zerowego lub obrotów zerowych pod wpływem temperatury (ang. *Zero-G/Zero-Rate Level Change vs. Temperature*). W ogólności wpływ opisanych czynników można zmniejszyć przez odpowiednią kalibrację. Co więcej dodatkowe układy mierzące i stabilizujące warunki zewnętrzne pozwalają na wprowadzenie odpowiedniej korekty do danych wyjściowych, co jest stosowane w znacznie droższych rozwiązaniach komercyjnych.

Źródłem szumu w czujnikach MEMS są różne efekty elektromechaniczne (Agrawal *et al.*, 2013). Dlatego też trudno jest opisać wszystkie istotne cechy szumu. Niemniej można wyróżnić dwa ważne rodzaje szumu, które wynikają z własności układów scalo-

nych, wzmacniaczy operacyjnych, rezystorów, itp. Są to szum termomechaniczny (ang. *Thermo Mechanical Noise*) oraz szum migoczący (ang. *Flicker Noise*). Szum termomechaniczny uwidacznia się w postaci sygnału wysokiej częstotliwości, która jest zdecydowanie wyższa od częstotliwości próbkowania sensora (wyższa od największej możliwej częstotliwości obecnej w sygnale mierzonym). W efekcie szum daje zakłócenie danych występujące dla wysokich częstotliwości o charakterze zbliżonym do białego szumu gaussowskiego z zerową wartością średnią (Woodman, 2007). Natomiast szum migoczący interpretowany jest jako szum różowy lub  $1/f$  ze względu na to, że wprowadza fluktuacje przesunięcia sygnału (Beeby *et al.*, 2004). Oba rodzaje szumu można zintegrować do jednego sygnału, który zazwyczaj modeluje się w postaci szumu zwanego błędzeniem losowym (ang. *Random Walk*). Warto zaznaczyć, że przypadkowe zmiany sygnału (zbliżone do białego szumu) po scałkowaniu także powodują szum typu błędzenie losowe. W przypadku żyroskopu przypadkowe zmiany przesunięcia powodują błędzenie losowe w danych ułożenia, przez co niepewność rośnie proporcjonalnie do  $t^{3/2}$ . Natomiast dla akcelerometru przypadkowe zmiany przyspieszenia wprowadzają błędzenie losowe dla prędkości (niepewność proporcjonalna do  $t^{3/2}$ ), oraz błędzenie losowe trzeciego rzędu dla położenia (niepewność proporcjonalna do  $t^{5/2}$ ). W konsekwencji podwójne całkowanie danych z akcelerometru MEMS w ogólności powoduje błąd położenia, który rośnie kwadratowo z czasem (Pang i Liu, 2001; Gilbert *et al.*, 2010).

Przedstawione niepewności pomiaru powodowane przez szum sensorów MEMS są szczególnie istotne przy wyznaczaniu położenia i ułożenia obiektu. Dlatego też należy uwzględnić, w jakim stopniu szum czujnika pogarsza precyzję określania pozycji robota, oraz sprawdzić czy precyzja ta jest wystarczająca dla potrzeb mapowania 3D. Jak już wspomniano, aby określić ułożenie możliwie dokładnie, konieczna jest odpowiednia filtracja i fuzja danych z magnetometru, żyroskopu i akcelerometru. Oczywiście dane te są nadmiarowe i każdy sensor dostarcza danych dotyczących ułożenia z nieco innej perspektywy. Dlatego algorytm fuzji danych powinien filtrować dane obciążone dużym szumem i wyznaczać ułożenie w oparciu o informacje o wyższej precyzji (Pang i Liu, 2001).

### 3.3 Model stanowy i estymacja danych

Problem wyznaczania nieznannej wielkości modelowanej na podstawie danych pomiarowych sprowadza się do problemu estymacji. Dane pomiarowe mogą pochodzić z wielu czujników, z których każdy podaje wartość niepewną (niedokładną) również ze względu na obecność szumu. Estymacja pozwala zatem szacować wartości nieznanymi parametrów modelu na podstawie obserwacji pewnej próby losowej. Taka obserwacja odnoszona może być do nieznannej postaci modelu danych pomiarowych (rozkładu zmiennej losowej). Generalnie można przyjąć, że znając model opisujący dane i szum, można oszacować prawdziwe wartości mierzonego procesu. Jeżeli dodatkowo proces ten można opisać za pomocą modelu to estymacja może wyeliminować nawet bardzo duże błędy pomiarowe (Kowalczyk i Domżański, 2013).

Do opisu dynamiki bądź kinematyki robota mobilnego należy racjonalnie dobrać model matematyczny, który pozwoli na poprawną estymację/predykcję pozycji robota. Ponieważ założono użycie robota kołowego, który ma 2 koła stałe z napędem umieszczone w przedniej części platformy robota oraz koło obrotowe (bez napędu) ułożone w tylnej części platformy robota, właściwy model kinematyki takiego robota opisują odpowied-



nie równania nieliniowe, które wyrażają istotne ograniczenia ruchu. Dokładne omówienie nieliniowego modelu robota z uwzględnieniem ograniczeń związanych z zamocowaniem i rodzajem kół, stopnia mobilności i sterowności, a także redukowalności, stabilizowalności i sterowalności, jak również możliwości linearyzacji znacznie wykracza poza zakres tej pracy. Zastosowany opis kinematyki robota oraz najistotniejszych elementów z nim związanych zostanie przedstawiony w rozdziale czwartym. Dla uproszczenia dyskusji i omówienia najważniejszych zagadnień w dalszej części rozdziału zakłada się uproszczenie modelu robota do postaci liniowej.

### 3.3.1. Model stanowy

Problem estymacji danych wymaga odpowiedniego opisu matematycznego (modelu) obiektu oraz danych. W automatyce i robotyce obserwowane procesy ciągłe można wyrazić w postaci równań różniczkowych, gdzie zmienne mają określoną interpretację fizyczną. Dlatego też większość systemów sterowania opiera się na układach równań stanu.

Układy równań różniczkowych wysokiego rzędu można sprowadzić do postaci stanowej, która reprezentuje równania różniczkowe pierwszego rzędu. Model stanowy zatem można zredukować do zależności wejściowo-wyjściowej oraz operatorowego opisu transmitancyjnego otrzymywanego poprzez przekształcenie Laplace'a lub wywodzące się z tego przekształcenia metody częstotliwościowe. Ogólny model stanu dla liniowego układu, ciągłego i niezmienniczego opisać można w postaci równania stanu i równania wyjścia

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)\end{aligned}\tag{3.17}$$

gdzie  $\mathbf{x}(t) \in \mathbb{R}^n$  to wektor zmiennych stanu,  $\dot{\mathbf{x}}(t)$  to pochodna wektora stanu po czasie,  $\mathbf{u}(t) \in \mathbb{R}^m$  wektor zmiennych wejściowych (wymuszenie),  $\mathbf{y}(t) \in \mathbb{R}^p$  wektor zmiennych wyjściowych, zaś  $\mathbf{A} \in \mathbb{R}^{qn}$  jest macierzą tranzycji stanu,  $\mathbf{B} \in \mathbb{R}^{qm}$  macierzą wejść,  $\mathbf{C} \in \mathbb{R}^{pn}$  macierzą wyjść i  $\mathbf{D} \in \mathbb{R}^{pm}$  macierzą przejścia wejście-wyjście. Macierz stanu  $\mathbf{A}$  opisuje dynamikę procesu. W przypadku problemu sterowania robotem i określania jego pozycji, macierz  $\mathbf{A}$  reprezentuje modelowaną dynamikę<sup>1</sup> robota, która wiąże zmienne ruchu robota. Macierz  $\mathbf{B}$  określa w jaki sposób wymuszenie (sygnał sterujący robotem) wpływa na zmienne stanu  $\mathbf{x}(t)$ . Macierz  $\mathbf{C}$  reprezentuje sposób obserwacji zmiennych stanu  $\mathbf{x}(t)$  poprzez sensory. Macierz  $\mathbf{D}$  zwykle jest zerowa (w przypadkach SISO, gdzie stopień licznika transmitancji jest mniejszy niż stopień jej mianownika).

Układ opisany modelem stanowym (3.17) ma swoją reprezentację w dziedzinie częstotliwości w postaci transmitancyjnej:

$$\mathbf{G}(s) = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B} + \mathbf{D}\tag{3.18}$$

Warto podkreślić, że opis stanowy nie jest jednoznaczny, tzn. istnieje nieskończenie wiele realizacji w postaci czwórki macierzy  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$ , które reprezentują daną transmitancję  $\mathbf{G}(s)$ . Dlatego w literaturze można spotkać wiele użytecznych postaci stanowych o różnych cechach implementacyjnych realizujących transmitancję  $\mathbf{G}(s)$  (Kowalczyk, 1992).

Układy rzeczywiste (fizyczne) modeluje się jako procesy ciągłoczasowe. Ponieważ pomiary (za pomocą układów MEMS) odbywają się w dyskretnych momentach czasu,

<sup>1</sup>Zgodnie z poczynionym założeniem przyjęto uproszczenie do liniowego modelu robota

w praktyce przydatne są modele stanowe z czasem dyskretnym. Do modelu ciągłego stosuje się transformatę Laplace'a, zaś do postaci dyskretniej transformatę  $\mathcal{Z}$ . W związku z tym należy pamiętać, że wartości parametrów transmitancji danego układu w postaci ciągłej i dyskretniej będą różne.

Układ dyskretny, liniowy, niezmienniczy (inwariantny) zapisać można w postaci następującego modelu stanowego:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k) \end{aligned} \quad (3.19)$$

gdzie  $k = 0, 1, 2, \dots$  jest numerem próbki oznaczającym dyskretny czas (przesunięcie).

Oprócz przedstawionych modeli stanowych LNP (liniowy, niezmienniczy, przyczynowy), ciągłego (3.17) i dyskretnego (3.19) istnieją oczywiście bardziej złożone postaci stosowne dla wariantnych układów (niestacjonarnych), nieliniowych, jednak w niniejszej pracy rozpatrywana będzie głównie stanowa postać odpowiednia dla układów LNP.

### 3.3.2. Cechy układu

Generalnie układy automatycznego sterowania posiadają cechy, które warunkują ich użyteczność do pracy autonomicznej. Konieczną cechą układu sterowania jest stabilność. Jest to cecha układu polegająca na powracaniu do stanu równowagi, bądź do niezaburzonej ewolucji po zaniku czynników zakłócających, które wytrąciły układ z równowagi (Kowalczyk, 1992). W szczegółowych rozważaniach wyróżnia się stabilność asymptotyczną oraz lokalną i globalną, które określają powrót do stanu równowagi odpowiednio: po nieskończonym czasie oraz dla małych i dowolnych wychyleń. Często wyróżnia się także stabilność w sensie BIBO (ang. *Bounded Input Bounded Output*), gdzie dla ograniczonego sygnału wejściowego  $\mathbf{u}(t)$  układ odpowiada ograniczonym sygnałem wyjściowym  $\mathbf{y}(t)$ . W praktyce istotne jest też, aby układ zamknięty (ze sterownikiem i ujemną pętlą sprzężenia zwrotnego) był tak zaprojektowany, aby nie dopuścić do samoistnego wzbudzenia się niezależnie od wartości sygnału wejściowego.

Z perspektywy opisu stanowego w wersji ciągłej układ jest stabilny, gdy wszystkie bieguny transmitancji układu zamkniętego znajdują się w lewej półpłaszczyźnie zespolonego operatora  $\mathcal{S}$ , co oznacza, że część rzeczywista wszystkich pierwiastków równania charakterystycznego (utworzonego, np. przez przyrównanie do zera mianownika transmitancji układu zamkniętego) jest ujemna. W przypadku układu dyskretnego stabilność wymaga, aby wszystkie bieguny transmitancji znajdowały się w okręgu jednostkowym na płaszczyźnie operatora zespolonej częstotliwości  $\mathcal{Z}$ .

Inne kluczowe cechy układu, to sterowalność i obserwowalność. Dany układ jest sterowalny, jeżeli dla dowolnego stanu początkowego  $\mathbf{x}(0)$  można zastosować sterowanie  $\mathbf{u}(t)$ , które spowoduje sprowadzenie układu do zadanego stanu ustalonego (zerowego)  $\mathbf{x}(t_f) = 0$ . Można spotkać wiele sposobów na sprawdzenie warunku sterowalności. Jednym z nich jest kryterium rzędu macierzy sterowalności, które wymaga, aby macierz  $\mathbf{R}$  była pełnego rzędu:

$$\text{rank}(\mathbf{R}) = \text{rank} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{n-1}\mathbf{B} \end{bmatrix} = n \quad (3.20)$$

gdzie  $\mathbf{A}$  i  $\mathbf{B}$  to odpowiednio macierz tranzycji stanu o wymiarze  $n \times n$  i macierz sterowania (wejść) o wymiarze  $n \times r$ , zaś *rank* to rząd macierzy. Jeżeli układ nie jest w pełni sterowalny, istnieją stany, których nie można zrelaksować (wysterować do zera). Jeśli są one stabilne, a pozostałe stany (w tym niestabilne) są sterowalne to w ogólności układem wciąż można sterować. W tym kontekście mówi się o układzie stabilizowalnym (którego wszystkie stany niesterowalne są stabilne). W zadaniu sterowania robotem mobilnym kluczowe jest, aby dla wybranego układu sterowania można było tak sterować robotem, aby osiągnąć pożądaną zmianę położenia z zadaną dokładnością. Kiedy docelowy stan nie jest zerowy, wymaga się o osiągalności (pełnej lub częściowej, jak w przypadku sterowalności). Przy nieosobliwej macierzy tranzycji stanu  $\mathbf{A}$  pojęcia sterowalności i osiągalności utożsamiają się i można ich używać wymiennie (Nise, 2015).

Z sensorami obecnymi w układzie wiąże się pojęcie obserwowalności. Jeżeli odczyt z sensorów w dowolnym czasie pozwala na określenie wartości wszystkich zmiennych stanu, mówi się o obserwowalności i odtwarzalności. Formalna definicja określa układ obserwowalny/odtworzalny jako układ, w którym dla dowolnego stanu  $\mathbf{x}(t)$  i sterowania  $\mathbf{u}(t)$  można określić stan początkowy/aktualny na podstawie obserwacji wyjścia  $\mathbf{y}(t)$  (w skończonym czasie). Kryterium obserwowalności jest związane z rzędem macierzy obserwowalności zdefiniowane jako

$$\text{rank}(O) = \text{rank} \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{n-1} \end{bmatrix} = n \quad (3.21)$$

gdzie  $\mathbf{A}$  i  $\mathbf{C}$  to odpowiednio macierz stanu o wymiarze  $q \times n$  i macierz wyjść o wymiarze  $p \times n$  (*rank* - rząd macierzy). Aby układ był obserwowalny macierz obserwowalności musi być pełnego rzędu. Podobnie jak przy sterowaniu, w kontekście obserwowalności/odtworzalności wyróżnia się pojęcie wykrywalności, kiedy wszystkie stany nieobserwowalne są stabilne (Kreindler i Sarachik, 1964) – w analogii do pojęcia stabilizowalności. W przypadku fuzji danych z sensorów inercyjnych brak obserwowalności/odtworzalności oznacza, że nie można określić wektora stanu w oparciu o określony model dynamiczny. Jeżeli na przykład model pozwala tylko na określenie prędkości obiektu, położenie robota może być nieobserwowalne. W takim przypadku konieczny jest dodatkowy pomiar albo zastosowanie innego sposobu obliczenia położenia.

### 3.3.3. Obserwacja stanu

Z problemem obserwowalności/odtworzalności (utożsamianym w przypadku nieosobliwości macierzy  $\mathbf{A}$ ) wiąże się pojęcie obserwatora stanu. Jest to pewien układ oparty na modelu układu rzeczywistego, który dostarcza estymaty (oszacowania) stanu układu wykorzystując obserwacje (pomiar) wejścia i wyjścia. Obserwator stanu jest szczególnie przydatny w sytuacji, gdy za pomocą bezpośredniego pomiaru nie można określić stanu układu (Luenberger, 1964). Układ musi być obserwowalny, aby za pomocą obserwatora estymować wszystkie stany. Obserwator Leuenbergera wykorzystuje wzmocnienie w postaci macierzy (lub wektora)  $\mathbf{L}$ . Macierz tę dobiera się tak, aby obserwator był stabilny

i wystarczająco szybki. Dla układu opisanego deterministycznym modelem (3.17), przy zerowej macierzy  $\mathbf{D}$ , stan estymuje się za pomocą obserwatora o następującej postaci:

$$\dot{\hat{\mathbf{x}}}(t) = \mathbf{A}\hat{\mathbf{x}}(t) + \mathbf{B}\mathbf{u}(t) + \mathbf{L}(\mathbf{y}(t) - \mathbf{C}\hat{\mathbf{x}}(t)) \quad (3.22)$$

gdzie  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  są macierzami modelu stanowego,  $\hat{\mathbf{x}}(t)$  jest estymowanym stanem układu,  $\mathbf{u}(t)$  jest sygnałem sterującym, a  $\mathbf{y}(t)$  jest sygnałem wyjściowym układu. Błąd obserwatora  $\mathbf{e}(t) = \hat{\mathbf{x}}(t) - \mathbf{x}(t)$  opisuje równanie

$$\dot{\mathbf{e}}(t) = (\mathbf{A} - \mathbf{L}\mathbf{C}) \mathbf{e}(t). \quad (3.23)$$

Obserwator Leuenbegera pozwala na efektywną estymację stanu układu, przy czym poprawność estymacji zależy silnie od doboru wzmocnienia  $\mathbf{L}$ . Aby obserwator był dostatecznie szybki należy odpowiednio dobrać jego bieguny (np. obserwator powinien być 10-krotnie szybszy niż układ). Wysoka dynamika układu obserwatora wymaga dużego wzmocnienia, które niestety sprzyja destabilizacji obserwowanego procesu. Obserwator daje wyniki poprawne asymptotycznie, dlatego w początkowym okresie może dawać błąd relatywnie duży (względem obserwowanej wartości zmiennej stanu). Dlatego też przy estymacji stanu konieczne jest odpowiednie dopasowanie wzmocnienia (i innych parametrów systemu) do dynamiki sygnałów.

### 3.3.4. Filtracja Kalmana

Bardziej zaawansowanym obserwatorem stanu jest filtr Kalmana. Jest to filtr optymalny w sensie średniokwadratowym dla liniowego układu dynamicznego modelowanego z określoną niepewnością, gdzie szумы związane z procesem i pomiarem mają charakter przypadkowy (Kalman, 1959). Filtracja ta w szczególności uwzględnia szum o rozkładzie gaussowskim. Filtr Kalmana działa zarówno dla układów stacjonarnych jak i niestacjonarnych oraz znajduje powszechne zastosowanie w wielu dziedzinach (Kushner, 1962). Równania stanu opisujące niepewny probabilistyczny proces w wersji dyskretnej mają postać:

$$\begin{aligned} \mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{w}(k) \end{aligned} \quad (3.24)$$

gdzie  $k$  jest przesunięciem (numerem próbki),  $\mathbf{v}(k)$  jest szumem procesowym (przetwarzania),  $\mathbf{w}(k)$  szumem pomiarowym,  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  są macierzami modelu stanowego (macierz  $\mathbf{D}$  jest nieużywana), natomiast  $\mathbf{x}(k)$ ,  $\mathbf{u}(k)$  i  $\mathbf{y}(k)$  to odpowiednio stan, sterowanie i pomiar (obserwacja). W modelu tym szum sparametryzowany jest za pomocą wartości średniej i wariancji (tutaj są to macierze kowariancji  $\mathbf{Q}$  i  $\mathbf{R}$ ):

$$\begin{aligned} \mathbf{v}(k) &\sim N(0, \mathbf{Q}) \\ \mathbf{w}(k) &\sim N(0, \mathbf{R}) \end{aligned} \quad (3.25)$$

Oba rodzaje szumów (procesowy i pomiarowy) charakteryzują się zerową wartością średnią i daną wariancją, które są stałe w czasie.

W filtracji Kalmana wyróżnia się etapy predykcji i korekcji. Równania opisujące predykcję nazywane są także równaniami aktualizacji czasu, gdyż predykcja wybranej



zmiennej dla aktualnego kroku  $k$  wyznaczana jest na podstawie oszacowania tej zmiennej w kroku poprzednim  $k - 1$ . Owe równania predyktora mają następującą postać:

$$\begin{aligned}\hat{\mathbf{x}}(k|k-1) &= \mathbf{A}\hat{\mathbf{x}}(k-1|k-1) + \mathbf{B}\mathbf{u}(k-1) \\ \mathbf{P}(k|k-1) &= \mathbf{A}\mathbf{P}(k-1|k-1)\mathbf{A}^T + \mathbf{Q}\end{aligned}\quad (3.26)$$

gdzie  $\hat{\mathbf{x}}(k|k-1)$  jest estymatą stanu w chwili  $k$  wyznaczoną na podstawie stanu  $\mathbf{x}(k-1)$  z chwili poprzedniej,  $\mathbf{P}(k|k-1)$  jest estymatą macierzy kowariancji błędu w chwili  $k$  obliczoną w oparciu o macierz kowariancji błędu z chwili poprzedniej. Oba wyrażenia zawierają wartości *a priori*, które obliczone są przed uwzględnieniem aktualnego pomiaru (wyjścia). Natomiast wyrażenia  $\hat{\mathbf{x}}(k-1|k-1)$  oraz  $\mathbf{P}(k-1|k-1)$  opisują wielkości *a posteriori* (uwzględniają wartości mierzone w poprzedniej chwili). Macierze  $\mathbf{A}$ ,  $\mathbf{B}$  i  $\mathbf{Q}$  to odpowiednio macierze: przejścia/systemowa, sterowania/wejścia i kowariancji szumu pomiarowego. Macierze kowariancji błędu *a priori* i *a posteriori* posiadają następującą interpretację:

$$\begin{aligned}\mathbf{P}(k|k-1) &= E[\mathbf{e}(k|k-1)\mathbf{e}^T(k|k-1)] \\ \mathbf{P}(k-1|k-1) &= E[\mathbf{e}(k-1|k-1)\mathbf{e}^T(k-1|k-1)]\end{aligned}\quad (3.27)$$

gdzie  $\mathbf{e}$  oznacza odpowiednio *aprioryczny* i *aposteryoryczny* błąd estymacji stanu w postaci różnicy między wartością rzeczywistą wektora stanu  $\mathbf{x}$  a jego wartością estymowaną  $\hat{\mathbf{x}}$ , co można zapisać jako

$$\begin{aligned}\mathbf{e}(k|k-1) &= \mathbf{x}(k|k-1) - \hat{\mathbf{x}}(k|k-1) \\ \mathbf{e}(k-1|k-1) &= \mathbf{x}(k-1|k-1) - \hat{\mathbf{x}}(k-1|k-1)\end{aligned}\quad (3.28)$$

Podczas predykcji przewiduje się estymatę stanu  $\hat{\mathbf{x}}$  oraz kowariancję błędu  $\mathbf{P}$  dla aktualnego kroku  $k$  (na estymowaną wartość pomiar nie wpływa). Procedura korekcji związana jest jeszcze z aktualizacją pomiarów. Na podstawie wyników predykcji korektor wyznacza zatem wzmocnienie filtra  $\mathbf{K}(k)$ , względem wartości *a posteriori* wektora stanu  $\hat{\mathbf{x}}(k|k)$  i macierzy kowariancji błędu  $\mathbf{P}(k|k)$  dla próbki  $k$ :

$$\begin{aligned}\mathbf{K}(k) &= \mathbf{P}(k|k-1)\mathbf{C}^T (\mathbf{C}\mathbf{P}(k|k-1)\mathbf{C}^T + \mathbf{R})^{-1} \\ \hat{\mathbf{x}}(k|k) &= \hat{\mathbf{x}}(k|k-1) + \mathbf{K}(k) (\mathbf{y}(k) - \mathbf{C}\hat{\mathbf{x}}(k|k-1)) \\ \mathbf{P}(k|k) &= (\mathbf{I} - \mathbf{K}(k)\mathbf{C}) \mathbf{P}(k|k-1)\end{aligned}\quad (3.29)$$

gdzie  $\mathbf{K}(k)$  nosi nazwę wzmocnienia Kalmana,  $\mathbf{C}$  jest macierzą wyjść/obserwacji, zaś  $\mathbf{I}$  to macierz jednostkowa.

Wzmocnienie  $\mathbf{K}$  reguluje wpływ (wiarygodność) estymacji i pomiaru. Jeżeli szum pomiarowy jest mały, pomiar jest bardziej wiarygodny niż estymacja, przez co pomiar ma większy wpływ na estymowany stan *a posteriori* niż estymacja stanu *a priori*. W przeciwnym przypadku, jeśli estymowany stan *a priori* nie jest obciążony dużym błędem to pomiar ma niewielki wpływ na estymowany stan *a posteriori*. Przy zerowej macierzy kowariancji szumu pomiarowego  $\mathbf{R}$  stan *a posteriori* odpowiada pomiarowi ( $\mathbf{K} = \mathbf{C}^{-1}$ ).

Przy zerowym błędzie estymacji  $\mathbf{P} = 0$ , wzmocnienie  $\mathbf{K} = 0$ , stan *a posteriori* odpowiada estymacji *a priori*, zaś wpływ pomiaru jest zerowy.

Filtr Kalmana jest algorytmem rekursywnym, w którym realizowane są na zmianę predykcja i korekcja. Przed uruchomieniem algorytmu należy przyjąć estymowany stan początkowy  $\hat{\mathbf{x}}(0|0)$  i macierz kowariancji błędu  $\mathbf{P}(0|0)$ . Predyktor wyznacza *aprioryczne* wartości stanu i kowariancji błędu estymacji, które przekazywane są do korektora. W oparciu o te dane oraz pomiar poza instrumentalnym wzmocnieniem, korektor wyznacza *aposterioryczne* wartości stanu i kowariancji błędu. Aby filtr był skuteczny, należy upewnić się, że szum sensoryczny ma charakter gaussowski, a następnie dobrać macierz  $\mathbf{Q}$  w oparciu o wariancję szumu sensora. Określenie macierzy opisującej szum procesowy  $\mathbf{R}$  oraz innych wartości początkowych jest jeszcze trudniejszym zadaniem. Zły dobór parametrów może powodować niedokładną lub powolną estymację, szczególnie na początku działania filtru. Złe wartości parametrów dla układów z biegunami położonymi blisko granicy stabilności mogą wiązać się z utratą stabilności obserwatora.

Przedstawiony filtr Kalmana dopasowany jest do dyskretnych układów liniowych. W literaturze można spotkać modyfikacje filtru Kalmana, np. filtr Kalmana z szumem niestacjonarnym, rozszerzony filtr Kalmana EKF, bezśladowy filtr Kalmana UKF (ang. *Unscented Kalman Filter*), które używane są dla układów o silnej nieliniowości (Albarbar *et al.*, 2009; Julier i Uhlmann, 2004). Mimo istnienia licznych odmian, w praktyce często stosuje się prosty filtr Kalmana, ponieważ występują istotne trudności w doborze parametrów w rozwiązaniach złożonych typu EKF i UKF. Okazuje się również, że rozwiązania problemu fuzji danych i sterowania robotem mobilnym oraz zadań SLAM mogą być oparte na prostym filtrze Kalmana (Sinopoli *et al.*, 2004).

### 3.3.5. Algorytm SLAM oparty na filtracji Kalmana

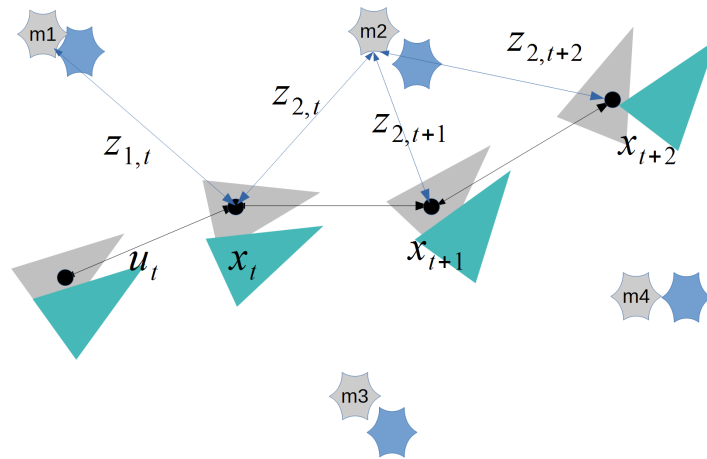
Algorytm dedykowany do jednoczesnego określania lokalizacji i mapowania (SLAM) obejmuje szereg elementów charakterystycznych dla problemów estymacji i sterowania. Konkretnie rozwiązania SLAM oparte są na filtrach cząsteczkowych lub filtracji Kalmana, którą opisano w poprzednim rozdziale. Zasada działania algorytmu SLAM została zilustrowana na rys. 3.1

Typowy system wykonujący zadanie SLAM zakłada przemieszczanie robota po całym badanym obszarze. Jest to algorytm iteracyjny, w którym każda iteracja to realizacja mapowania aktualnego otoczenia i przemieszczenie robota. Pojedynczy pomiar polega na określeniu położenia obiektów znajdujących się w bliskim otoczeniu robota i zaznaczeniu ich na mapie. Następnie robot przemieszcza się o zadany krok w kierunku obszaru, który nie był wcześniej eksplorowany. Mapowanie obiektów położonych w dalszym otoczeniu robota wymaga wielu iteracji algorytmu, co wiąże się z wieloma pomiarami i zmianami położenia wykonywanymi naprzemiennie.

Realizacja lokalizacji i mapowania wymaga pomiaru położenia znaczników (ang. *Landmarks*)  $m_1 - m_4$  znajdujących się w otoczeniu<sup>2</sup>. Pomiar  $z_{1,t}$  dostarcza danych o położeniu obiektu  $m_1$  w chwili czasu  $t$ . W przedstawionym przykładzie, w kolejnych dwóch chwilach czasu  $t + 1$  i  $t + 2$ , nie jest dostępny pomiar położenia znacznika  $m_1$ . Natomiast położenie znacznika  $m_2$  jest mierzone zarówno w chwili czasu  $t$ ,  $t + 1$ , jak i  $t + 2$  w postaci

<sup>2</sup>Znacznik jest określany także jako punkt orientacyjny. W praktyce może być to dowolny obiekt rozróżnialny od innych obiektów znajdujących się w otoczeniu.





Rysunek 3.1 Schemat działania SLAM.

pomiarów odpowiednio  $z_{2,t}$ ,  $z_{2,t+1}$  i  $z_{2,t+2}$ . W chwili czasu  $t$  wektor stanu robota zaznaczony jest w postaci  $\mathbf{x}_t$ . Wektor stanu zawiera zmienne stanu istotne dla dynamiki robota (oraz sterowania) takie jak, położenie, prędkość, ułożenie robota itp. W wektorze stanu uwzględnia się także wszystkie analizowane znaczniki. Na stan robota w chwili  $t$  wpływa oczywiście wymuszenie  $\mathbf{u}_t$ , które jest sygnałem sterującym (napędzającym) ruch robota.

Rzeczywiste położenie robota oraz znaczników (zaznaczone kolorem szarym) zwykle nie jest znane. Posługując się pomiarami można określić jedynie wartości estymowane (zaznaczone kolorem niebieskim). W trakcie przemieszczania robot mierzy położenie tych samych znaczników w kolejnych chwilach czasu. Zatem uśredniając dane pomiarowe związane z tym samym znacznikiem błąd estymacji położenia znaczników w ogólności powinien maleć, gdyż położenie znaczników jest stałe. Położenie znaczników wyznaczone jest względem pozycji robota, więc na pomiar położenia znacznika wpływa także błąd położenia i ułożenia robota. Jednakże przy naprzemiennym (jednoczesnym) wyznaczaniu położenia znaczników i robota wraz ze wzrostem liczby pomiarów błąd położenia powinien maleć. Pewność pomiaru może być utożsamiana z korelacją położenia robota względem znaczników (Bailey i Durrant-Whyte, 2006a,b; Hong, 2012).

Aby użyć filtracji Kalmana w zadaniu SLAM, konieczne jest wcześniejsze zdefiniowanie modelu dynamiki robota. Zatem należy odpowiednio zdefiniować elementy wektora stanu  $\mathbf{x}$  oraz parametry macierzy  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$  opisujące zachowanie robota oraz pozwalające na obserwowanie jego pozycji. Następnie model stanowy należy rozszerzyć o współrzędne znaczników (w wektorze stanu) oraz uzupełnić macierze  $\mathbf{A}$  i  $\mathbf{B}$  o odpowiednie zerowe wiersze oraz kolumny, a także rozszerzyć macierz  $\mathbf{C}$  tak, aby uwzględniała pomiary położenia znaczników. Dodatkowo należy uwzględnić szum związany z pomiarem położenia znaczników w postaci odpowiednich odchyłeń standardowych w wektorze szumu pomiarowego  $\mathbf{w}$ .

W przypadku dwuwymiarowym przykładowy wektor stanu przedstawić można w następującej postaci:

$$\mathbf{x} = [x, y, \psi, m_{1x}, m_{1y}, m_{2x}, m_{2y}, \dots]^T \quad (3.30)$$

gdzie  $x, y$  są współrzędnymi położenia robota, kąt  $\psi$  określa ułożenie robota na płaszczyźnie (w tym przypadku składowe ułożenia  $\phi$  oraz  $\theta$  są pomijane), natomiast pary  $(m_{1x}, m_{1y})$  i  $(m_{2x}, m_{2y})$  stanowią współrzędne znaczników, odpowiednio  $m_1$  i  $m_2$ . Jeżeli w danej chwili zostanie wykryty nowy znacznik, konieczne jest odpowiednie rozszerzenie wektora stanu i macierzy stanu.

Na początku działania estymowane położenie znaczników może być niedokładne, gdyż opiera się na małej liczbie danych pomiarowych. Ponadto początkowy duży błąd skutkuje dużym wzmocnieniem Kalmana, które generuje względnie dużą wariancję zmiennych. Po pewnym czasie, wzmocnienie Kalmana zmniejsza się a zatem i wahania wartości zmiennych. Dopóki ruch robota (modelowanego z użyciem macierzy  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ ) bliski jest założonemu modelowi liniowemu, zmienne stanu mają względnie małą wariancję.

Warto zaznaczyć, że filtracja Kalmana stosowana w SLAM ma stosunkowo dużą złożoność obliczeniową nie przekraczającą  $O(n^3)$ , gdzie  $n$  jest iloczynem liczby stanów i punktów pomiarowych (Samsuri *et al.*, 2014). Zatem przy dużej liczbie punktów pomiarowych algorytm może nie spełniać warunków czasu rzeczywistego. Dlatego w literaturze można spotkać rozwiązania oparte na ograniczaniu pomiarów i przyspieszaniu działania algorytmu SLAM (Montemerlo i Thrun, 2007).

### 3.4 Stereowizja

Stereowizja, podobnie jak procedury zaszyte w płacie potylicznym mózgu człowieka, generuje wrażenie przestrzenności na podstawie płaskiego obrazu. Naśladując system widzenia człowieka, w pewnych warunkach umożliwia się wizualizację obrazu stereoskopowego, np. z wykorzystaniem okularów 3D, co może być niezwykle użyteczne dla takiego systemu (Kowalczyk i Merta, 2013). Dwie kamery, tzw. stereopara (analogicznie jak para oczu) dostarczają obrazy, które przedstawiają obiekty z dwóch perspektyw. Pojedyncza kamera daje obraz, który jest jedynie rzutem obiektów widocznych w przestrzeni trójwymiarowej na płaszczyznę dwuwymiarową. Na podstawie obrazu z pojedynczej kamery nie można wnioskować o głębi tj. odległości danego punktu od kamery. Natomiast rzut tego samego punktu w przestrzeni na dwie oddalone od siebie płaszczyzny kamer pozwala na określenie głębi. (Kumar, 2009). Opis zależności geometrycznych związanych ze stereowizją wymaga użycia współrzędnych jednorodnych (ang. *Homogeneous Coordinates*) i przedstawienia modelu oraz parametrów kamery.

#### 3.4.1. Współrzędne jednorodne

Współrzędne jednorodne pozwalają na dogodną reprezentację transformacji w przestrzeni  $n$ -wymiarowej za pomocą  $n+1$  współrzędnych. Jest to szczególnie stosowny zapis dla zależności łączących współrzędne dwuwymiarowe i trójwymiarowe. Przy transformacji współrzędnych forma ta pozwala na zapis rotacji i translacji w postaci pojedynczej operacji, co w przypadku równań (3.2) nie jest możliwe. Współrzędne jednorodne pozwalają na wygodne wyrażenie także operacji typu skalowanie, rotacja, pochylenie, translacja itp. w postaci pojedynczej macierzy (Choi i Ha, 2010). Ponadto współrzędne jednorodne pozwalają na zwężony opis rzutu perspektywicznego przestrzeni  $n$ -wymiarowej na płaszczyznę  $(n-1)$ -wymiarową. W szczególności konwencja ta pozwala na opis rzutu przestrzeni trójwymiarowej na dwuwymiarową płaszczyznę kamery.

Dla współrzędnych trójwymiarowych dowolną afiniczną transformację (rotację i translację) zapisać można w następującej postaci:

$$\begin{bmatrix} x' \\ y' \\ z' \\ w' \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ w \end{bmatrix} \quad (3.31)$$

gdzie punkt o współrzędnych  $(x, y, z)$  po transformacji ma współrzędne  $(x', y', z')$ ,  $w$  i  $w'$  to współczynniki skalujące,  $t_x, t_y, t_z$  to kolejne elementy wektora translacji, zaś  $r_{11}, \dots, r_{33}$  to poszczególne elementy macierzy rotacji  $\mathbf{R}$  zdefiniowanej np. za pomocą wzoru (3.6).

Z perspektywy kamery transformację między współrzędnymi w przestrzeni trójwymiarowej a ich dwuwymiarowym rzutem na płaszczyznę zapisać można jako:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \mathcal{T} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (3.32)$$

gdzie  $(x, y, z)$  to współrzędne punktu w przestrzeni,  $(x', y', w')$  przy  $w' = 1$  to wektor współrzędnych rzutu tego punktu, zaś macierz  $\mathcal{T}$  jest macierzą przekształcenia.

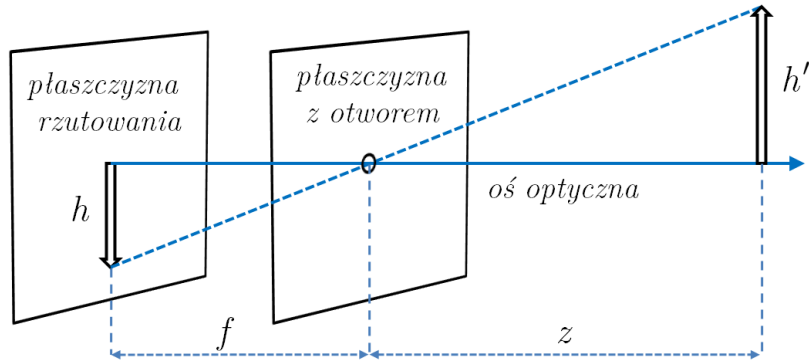
### 3.4.2. Modele kamery

Aby zrozumieć istotę rzutowania elementów świata zewnętrznego na płaszczyznę kamery, należy najpierw opracować model kamery. Proces akwizycji obrazu z kamery tłumaczy się drogą jaką pokonuje promień światła. Źródło światła w postaci słońca lub żarówki emituje promienie świetlne, które docierają do obiektów znajdujących się w otoczeniu. Pewna część energii świetlnej zostaje pochłonięta przez taki obiekt, natomiast pozostała część pozostaje odbita, przez co staje się on wtórnym źródłem światła. Promień świetlny wpadający do kamery załamuje się na soczewkach obiektywu, a następnie pada na materiał światłoczuły. W kamerach cyfrowych materiałem tym jest matryca małych elementów światłoczułych typu CCD (ang. *Charge-Coupled Device*) lub CMOS (ang. *Complementary Metal-Oxide-Semiconductor*), które na ekranie są reprezentowane w postaci pikseli.

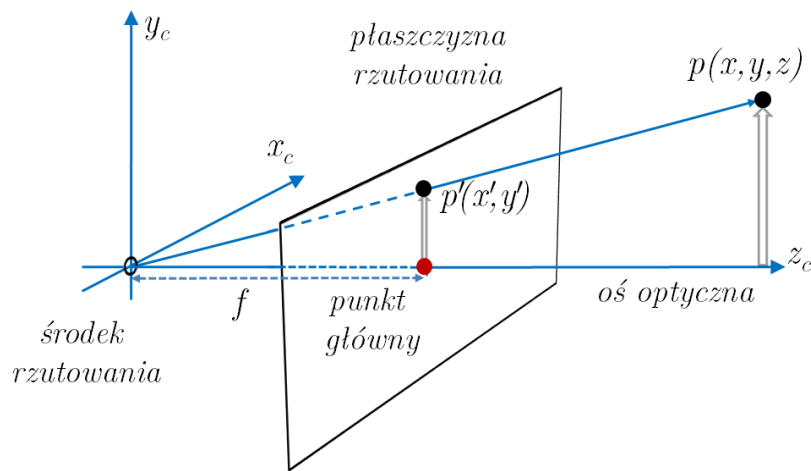
Najprostszy model kamery przedstawiony na rys. 3.2, znany jako *camera obscura* lub aparat otworkowy (ang. *Pinhole Camera*) jest oparty na obiektywie w postaci małego otworu. Otwór (zaznaczony na rys. 3.2) powoduje, że do płaszczyzny obrazu docierają tylko te promienie, które przecinają określony punkt w przestrzeni (Bradski i Kaehler, 2013). Promienie światła odbite od obiektu oddalonego od kamery na odległość  $z$  przechodzą przez otwór i docierają do płaszczyzny kamery oddalonej od otworu na odległość  $f$ , która jest ogniskową obiektywu. Rzut obiektu o wysokości  $h$  na płaszczyznę kamery ma wysokość  $h'$ . Z twierdzenia Talesa otrzymujemy zależność:

$$h' = f \frac{h}{z} \quad (3.33)$$

Równoważny model można uzyskać poprzez modyfikację kamery *obscura* tak, aby zarówno obiekt, jak i płaszczyzna rzutowania położone były po tej samej stronie otworu,



Rysunek 3.2 Model otworkowy kamery.



Rysunek 3.3 Równoważny model otworkowy kamery.

jak pokazano na rys. 3.3. Otwór w zmodyfikowanym modelu interpretuje się jako środek rzutowania (ang. *Center of Projection*), zaś punkt przecięcia osi optycznej z płaszczyzną rzutowania (kamery) to tzw. punkt główny (ang. *Principal Point*). Odcinek łączący środek rzutowania z punktem głównym reprezentuje ogniskową  $f$ . Model ten umożliwia prostszy opis zależności między rzeczywistym punktem  $p(x, y, z)$  w przestrzeni a jego obrazem  $p'(x', y')$  powstałym na płaszczyźnie kamery.

W przypadku idealnej kamery oś optyczna powinna przechodzić przez środek płaszczyzny kamery. W rzeczywistej kamerze oś optyczna jest nieznacznie odchyłona, przez co punkt główny nie znajduje się w środku płaszczyzny kamery. W komercyjnych kamerach (szczególnie tanich) występują piksele prostokątne, co oznacza inne ogniskowe w kierunku poziomym i pionowym. Po uwzględnieniu przemieszczenia punktu głównego oraz ogniskowych, dwuwymiarowy opis obrazu (3.33) można wyrazić w postaci:

$$\begin{aligned} x' &= f_x \frac{x}{z} + c_x \\ y' &= f_y \frac{y}{z} + c_y \end{aligned} \tag{3.34}$$

gdzie  $(x, y, z)$  są współrzędnymi punktu  $p$  w przestrzeni,  $(x', y')$  to współrzędne rzutu  $p'$  tego punktu na płaszczyznę kamery, zaś  $c_x$  i  $c_y$  reprezentują współrzędne punktu głównego,  $f_x$  i  $f_y$  są ogniskowymi w kierunkach odpowiednio  $x$  i  $y$ . Oczywiście parametry  $f_x$  i  $f_y$  reprezentują jedną fizyczną ogniskową  $f$ , którą przeliczamy na składowe.

Przedstawiony model i parametry pozwalają na wyznaczenie współrzędnych w płaszczyźnie kamery na podstawie współrzędnych punktu w przestrzeni. Odwrotny proces oparty na pojedynczej kamerze w ogólności nie jest możliwy. Model kamery przedstawiony powyżej jest wystarczający do omówienia najważniejszych parametrów kamery.

### 3.4.3. Parametry kamer

Właściwe ustawienie typowej kamery cyfrowej do pracy polega na dostrajaniu wielu różnych nastaw. Parametry mogą dotyczyć optyki (np. jasność obiektywu, ogniskowa i jej zakres zmienności, czy autofocus) lub elektroniki matrycy (np. typ i wymiary matrycy, maksymalna rozdzielczość, liczba klatek na sekundę, rozdzielczość przetwornika lub format zapisu). Dodatkowo istotne są parametry takie, jak kąt widzenia AOV (ang. *Angle of View*), zakres dynamiki WDR (ang. *Wide Dynamic Range*), kompensacja światła tylnego/przedniego BLC/HLC (ang. *Backlight/Highlight Compensation*), które nie są związane bezpośrednio z matrycą lub optyką i zależą od funkcjonalności i oprogramowania kamery.

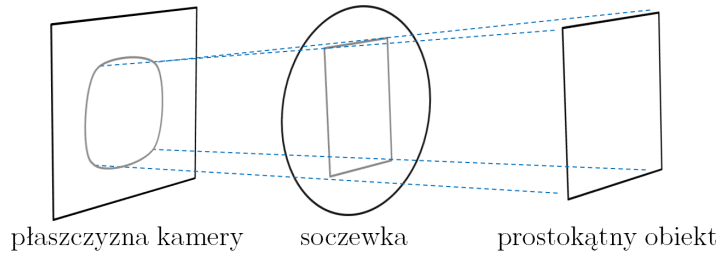
Z perspektywy stereowizji najistotniejsze są parametry kamery związane z dokładnością odwzorowania przestrzeni na płaszczyźnie kamery. Należą do nich parametry wewnętrzne (ang. *Intrinsic Parameters*) i zewnętrzne (ang. *Extrinsic Parameters*) (Bradski i Kaehler, 2013). Parametry wewnętrzne kamer łączą współrzędne punktu na płaszczyźnie obrazu (w pikselach) z odpowiadającymi im współrzędnymi w przestrzeni (w metrach), które opisane są w układzie odniesienia kamery. Natomiast parametry zewnętrzne definiują położenie i orientację (ułożenie) we współrzędnych zewnętrznych.

Część parametrów wewnętrznych kamery omówiono w opisie modelu kamery (3.34). W zagadnieniach przetwarzania obrazu parametry te przedstawia się w postaci jednej macierzy parametrów wewnętrznych  $\mathbf{M}$  (ang. *Camera Intrinsic Matrix*), która pozwala na zwiezły opis przejścia między współrzędnymi punktu trójwymiarowego  $p(x, y, z)$  w układzie odniesienia kamery a współrzędnymi tego punktu na obrazie (rzut  $p'(x', y', 1)$  we współrzędnych jednorodnych):

$$p' = \mathbf{M} \cdot p \quad (3.35)$$

$$\mathbf{M} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

Powyzsze wyrażenie przedstawione jest w postaci jednorodnej, gdzie  $c_x$  i  $c_y$  oznaczają współrzędne punktu głównego,  $f_x$  i  $f_y$  reprezentują ogniskową w kierunkach odpowiednio  $x$  i  $y$ . Parametry wewnętrzne poprawnie opisują rzut danego punktu na płaszczyznę obrazu tylko dla idealnego obiektywu. W przypadku rzeczywistej kamery zespół soczewek wprowadza istotne zniekształcenia (Hartley i Zisserman, 2003). Dla typowych kamer i obiektywów model zniekształceń obejmuje zniekształcenia radialne związane ze sferycznym kształtem soczewek i zniekształcenia styczne, które powstają podczas montażu kamery na skutek niedokładnego rozmieszczenia matrycy względem obiektywu (Ma *et al.*, 2006; Nedeveschi *et al.*, 2002).



Rysunek 3.4 Zniekształcenia radialne.

Zniekształcenia radialne są widoczne w postaci wyoblenia (zaokrąglenia) konturów obiektów widocznych w obrazie kamery, jak to zilustrowano na rys. 3.4. Zniekształcenia te są stosunkowo niewielkie w centralnej części obrazu i rosną wraz z odległością od jego środka. Dla typowego obiektywu opisuje się je za pomocą trzech parametrów  $k_1$ ,  $k_2$  i  $k_3$  (Bradski i Kaehler, 2013; Junchuan *et al.*, 2006):

$$\begin{aligned} x_k &= x(1 + k_1r^2 + k_2r^4 + k_3r^6) \\ y_k &= y(1 + k_1r^2 + k_2r^4 + k_3r^6) \end{aligned} \quad (3.36)$$

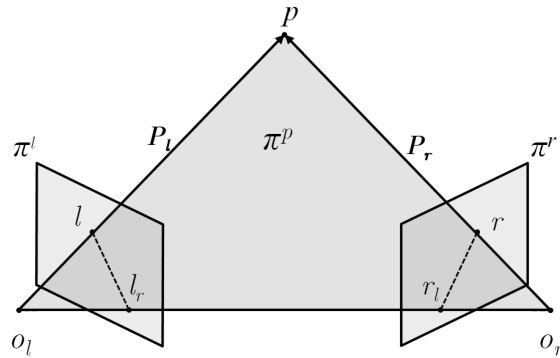
gdzie  $x_k$  i  $y_k$  są współrzędnymi obrazu  $x$ ,  $y$  po korekcji zniekształcenia, współczynnik  $k_1$  jest związany z niedużymi zniekształceniami, które rosną kwadratowo wraz z odległością  $r$  od środka optycznego,  $k_2$  odpowiada zniekształceniom rosnącym według  $r^4$ , zaś  $k_3$  odpowiada największym zniekształceniom, które rosną jak  $r^6$ . W praktyce dla typowej kamery parametr  $k_3$  przyjmuje wartość 0, gdyż dotyczy on zniekształceń, które pojawiają się tylko dla obiektywów szerokokątnych typu ‘rybie oko’.

Drugi rodzaj zniekształceń (styczne) pojawia się przede wszystkim w tanich kamerach, gdzie niedokładny montaż powoduje, że płaszczyzna matrycy kamery nie jest równoległa do płaszczyzny obrazowania (płaszczyzny, przez którą przechodzi otwór w modelu *camera obscura* pokazanej na rys. 3.2). Zniekształcenia styczne opisuje się za pomocą dwóch parametrów  $p_1$  i  $p_2$ :

$$\begin{aligned} x_k &= x + [2p_1y + p_2(r^2 + 2x^2)] \\ y_k &= y + [p_1(r^2 + 2y^2) + 2p_2x] \end{aligned} \quad (3.37)$$

gdzie  $x_k$  i  $y_k$  są współrzędnymi obrazu  $x$  i  $y$  po korekcji zniekształcenia, zaś  $r$  jest odległością od środka obrazu. Parametry radialne i styczne są często zapisywane w postaci wektora zniekształceń  $[k_1 \ k_2 \ p_1 \ p_2 \ k_3]^T$  (Bradski i Kaehler, 2013).

Macierz parametrów wewnętrznych wraz z wektorem zniekształceń pozwalają na geometryczne odwzorowanie danego punktu w przestrzeni na płaszczyznę kamery. Natomiast parametry zewnętrzne umożliwiają określenie położenia i ułożenia kamery względem danego punktu w przestrzeni. W przypadku systemu kamer parametry zewnętrzne i wewnętrzne pozwalają ocenić, gdzie na płaszczyznach kamer będzie zobrazowany dany punkt z przestrzeni. W skład parametrów zewnętrznych wchodzi macierz rotacji  $\mathbf{R}$  oraz wektor translacji  $\mathbf{T}$ . Parametry zewnętrzne reprezentowane są w pojedynczej macierzy we współrzędnych jednorodnych (Bradski i Kaehler, 2013).



Rysunek 3.5 Geometria epipolarna.

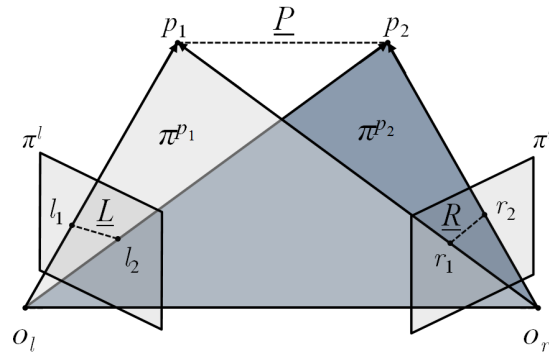
Parametry wewnętrzne i zewnętrzne ustala się w wyniku procedury kalibracji, która oparta jest na detekcji znanego wzorca z wielu ujęć. Typowym wzorcem jest fragment szachownicy. Zadaniem algorytmu kalibracji jest wykrycie charakterystycznych (dystyngtywnych) punktów w obrazie zniekształconej szachownicy. Następnie iteracyjnie dobierane są parametry wewnętrzne oraz wektor zniekształceń tak, aby zminimalizować błąd odpowiadający zniekształceniu szachownicy. Typowy średni błąd kalibracji nie powinien przekraczać wartości 0.25 piksela. Właściwa kalibracja jest niezbędna dla prawidłowego działania stereowizji.

#### 3.4.4. Geometria epipolarna

Poprawne odwzorowanie przestrzeni sprowadza się do zadania opisu tego samego punktu w obrazach pozyskiwanych z różnych kamer. Szczególnie istotne są zależności geometryczne pomiędzy punktem przestrzeni i jego obrazem w dwóch kamerach. Zagadnienie to jest rozważane w geometrii epipolarnej (Hirschmuller, 2008; Zhao *et al.*, 2010).

Szczegółowe rozważania matematyczne związane z odwzorowaniem przestrzeni opartym na systemie jedno-, dwu- i trzykamerowym opisał Hartley (2003). Wariant dwukamerowy (lub dwuobrazowy) zilustrowano na rys. 3.5, gdzie matryce kamer przedstawiono w postaci płaszczyzn obrazów  $\pi^l$  i  $\pi^r$  związanych odpowiednio z lewą i prawą kamerą. Na każdą z płaszczyzn obrazów rzutowany jest punkt  $p$  znajdujący się w przestrzeni. Przypominając, że  $o_l$  i  $o_r$  oznaczają środki rzutowania (odpowiednio dla lewej i prawej kamery), zdefiniujemy płaszczyznę  $\pi^p$  przechodzącą przez punkt  $p$  oraz oba środki rzutowania ( $o_l$  i  $o_r$ ) jako płaszczyznę epipolarną. Jej przecięcie z obrazami  $\pi^l$ ,  $\pi^r$  na matrycy każdej kamery tworzy sprzężone linie oraz odcinki epipolarne, odpowiednio  $(l, l_r)$  i  $(r, r_l)$ , zaś obraz  $r_l$  środka rzutowania  $o_l$  w prawej kamerze oraz obraz  $l_r$  środka rzutowania  $o_r$  w lewej kamerze stanowią środki epipolarne. Odcinek łączący środki rzutowania  $o_l$  i  $o_r$  jest bazą stereowizyjną zwaną także bazą stereoskopową, bądź krótko bazą stereo (ang. *Stereo Baseline*). Choć położenie i ułożenie kamer w przestrzeni jest nieznanne, dany punkt  $p$  w przestrzeni opisany jest wektorami  $P_l$  i  $P_r$  zaczepionymi w środkach rzutowania  $o_l$  i  $o_r$  w każdej z kamer (Kowalczyk i Merta, 2014; Davison *et al.*, 2007). Dokładne powiązanie tych wektorów jest możliwe po określeniu parametrów zewnętrznych kamer.

Położenie i ułożenie prawej kamery względem lewej można opisać za pomocą wektora translacji  $T$  oraz macierzy rotacji  $R$  (Ma *et al.*, 2006). Relacja między wektorami  $P_l$  i  $P_r$



Rysunek 3.6 Rzutowanie odcinka na płaszczyzny nieskalibrowanych kamer.

uzyskuje wtedy oczywistą postać:

$$\mathbf{P}_r = \mathbf{R}(\mathbf{P}_l - \mathbf{T}) \quad (3.38)$$

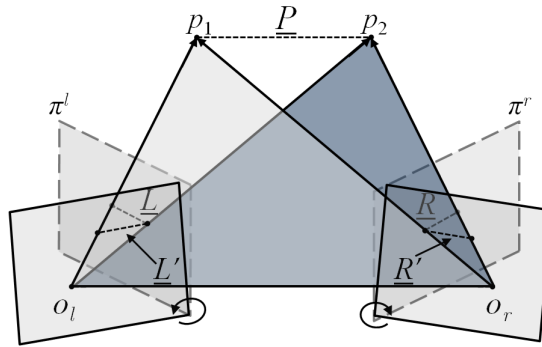
Zauważmy, że wektor  $\mathbf{T}$  i macierz  $\mathbf{R}$  można wyznaczyć podczas kalibracji kamer i są one niezbędne do dalszej analizy. W ogólności rzutowanie punktu 3D na płaszczyzny  $\pi^l$ ,  $\pi^r$  nie jest zadaniem trywialnym. Dodanie drugiego punktu w przestrzeni nieco komplikuje analizę, ale pozwala lepiej opisać jak dana linia jest rzutowana na płaszczyzny obrazów kamer. Na rys. 3.6 przedstawiono dany odcinek rzeczywisty  $\underline{P}(p_1, p_2)$  równoległy do układu kamer  $(o_l, o_r)$ , który jest rzutowany na płaszczyzny obrazów  $\pi^l$  i  $\pi^r$  czego skutkiem są powstałe odcinki odpowiednio  $\underline{L}(l_1, l_2)$  i  $\underline{R}(r_1, r_2)$ . Aby wyznaczyć odcinek rzeczywisty  $\underline{P}$  na podstawie obrazu z dwóch kamer konieczna jest odpowiednia analiza przesunięcia jego rzutu  $\underline{R}$  względem  $\underline{L}$  (zwykle współrzędne lewej kamery stanowią początek układu odniesienia). W ogólności odcinki  $\underline{L}$  i  $\underline{R}$  nie są względem siebie równoległe, co więcej nie pokrywają się z poziomymi liniami na płaszczyznach kamer ( $\pi^l$ ,  $\pi^r$ ). Dlatego przed analizą odcinków należy dokonać takiej transformacji obrazów z obu kamer, aby  $\underline{L}$  i  $\underline{R}$  były równoległe względem osi poziomej każdej z kamer.

### 3.4.5. Wyznaczanie głębi

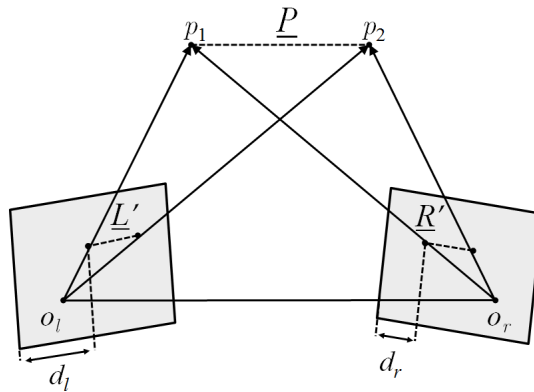
Aby ułatwić analizę rzutów punktu w przestrzeni na płaszczyzny obrazów dwóch kamer wykonuje się rektyfikację obrazów. Jest to transformacja perspektywiczna każdego z obrazów, która skutkuje przekształceniem płaszczyzn obrazów  $\pi^l$  i  $\pi^r$ . Po poprawnej rektyfikacji sprzężone odcinki  $\underline{L}$  i  $\underline{R}$  uzyskują postać odcinków  $\underline{L}'$  i  $\underline{R}'$ , które mieszczą się na tej samej (np. 5 od góry) linii poziomej na obrazach  $\pi^l$  i  $\pi^r$  (rys. 3.7). Ponadto parametry zewnętrzne kamer w typowym systemie stereowizyjnym opisane relacją (3.38) charakteryzują się brakiem rotacji i jedynie poziomą translacją, która jest równa długości bazy stereo  $B = (o_l, o_r)$ . Proces rektyfikacji ułatwia wyznaczenie głębi w obrazie, gdyż odpowiednie przetwarzanie sprowadza się do analizy tej samej pojedynczej linii poziomej w obrazach z kamer. W praktyce konieczna jest analiza obszaru obejmującego kilka sąsiednich linii (mimo rektyfikacji i wcześniejszej kalibracji kamer).

Zrektyfikowany obraz w kolejnym etapie przetwarzania pozwala na wyznaczenie dysparycji, która określa przesunięcie  $d = (d_l - d_r)$  w pikselach obrazu danego punktu 3D na płaszczyznach  $\pi^l$  i  $\pi^r$  jak to zilustrowano na rys. 3.8. Przesunięcie to obliczone dla wszystkich punktów widocznych w obrazie wyraża się mapą dysparycji. W ogólności zna-





Rysunek 3.7 Rektyfikacja obrazów z kamer.



Rysunek 3.8 Dysparycja obrazów po rektyfikacji obrazów z kamer.

jąc bazę  $B$  (odległość między środkami rzutowania kamer  $o_l$  i  $o_r$ ), punktom na mapie dysparycji można przypisać odległości w przestrzeni, tj. wyznaczyć mapę głębi jako:

$$z = \frac{Bf}{d} \quad (3.39)$$

gdzie  $z$  jest głębią obrazu w metrach,  $B$  jest szerokością bazy stereo w metrach,  $f$  jest ogniskową w pikselach,  $d$  zaś – dysparycją w pikselach. Błąd wyznaczania głębi obrazu  $\Delta z$  (w metrach) uwarunkowany jest błędem dysparycji  $\Delta d$  (w pikselach):

$$\Delta z = \frac{z^2}{Bf} \Delta d \quad (3.40)$$

Wyznaczana w ten sposób głębia  $z$  jest zależna od ustalonych parametrów ogniskowej  $f$ , dysparycji  $d$  oraz długości bazy stereo  $B$ . Im większa jest dysparycja tym dany punkt znajduje się bliżej układu kamer. Jak widać błąd pomiaru głębi rośnie wraz z kwadratem odległości od układu kamer. Dlatego zachowanie stałej dokładności pomiaru wymaga liniowego zwiększenia jednocześnie parametrów  $B$  i  $f$  wraz z liniowym wzrostem  $z$  (Gallup *et al.*, 2008). Warto zauważyć, że błąd (3.40) uwzględnia wielkości fizyczne w postaci ciągłej, natomiast dysparycja jest dostępna tylko w postaci dyskretniej. Dlatego w dalszej części rozdziału przeprowadzimy dokładniejszą analizę błędu map dysparycji i głębi.

Przy tworzeniu mapy dysparycji należy pamiętać, że wartości parametrów wyznaczone poprzez analizę obrazu wyrażone są w pikselach. Według (3.39) dysparycja  $d$  i ogniskowa

$f$  (lub baza stereoskopowa  $B$ ) są wyznaczone w pikselach. Wówczas baza  $B$  (lub ogniskowa) oraz głębia mogą być podane w metrach. Można też wyznaczyć rozmiar pojedynczego piksela, co można określić w procesie kalibracji mierząc długość określonego wzorca (np. bok pola szachownicy) lub obliczyć na podstawie wymiarów matrycy podanych przez producenta. Wartość ogniskowej w metrach, na podstawie wartości w pikselach, można wyznaczyć za pomocą prostego wzoru:

$$f_m = f_p \Theta \quad (3.41)$$

gdzie  $f_m$  jest ogniskową w metrach,  $f_p$  jest ogniskową w pikselach, zaś  $\Theta$  jest rozmiarem właściwym pojedynczego piksela (w  $m/piksel$ ). Warto zauważyć, że rozmiar właściwy piksela  $\Theta$  zwykle ma różne wartości w kierunkach  $x$  i  $y$ , dlatego niektóre parametry rozpatruje się oddzielnie w kierunku poziomym i pionowym, czego przykładem jest zapis ogniskowej  $f_x$  i  $f_y$ . Po wyznaczeniu mapy dysparycji (w pikselach) oraz relacji między wartością danego parametru w pikselach i metrach, wyznaczyć można odległość kamery od obiektów widocznych w obu kamerach.

Ponadto z ogniskową wiąże się zagadnienie głębi ostrości. Przy długiej ogniskowej obraz jest ostry tylko dla wybranego zakresu odległości od kamery. Dlatego do zastosowań mapowania warto wybrać krótką ogniskową, dzięki czemu obraz można ustawić ostry obraz dla punktów położonych dalej niż pewna wartość progowa.

### 3.4.6. Mapowanie 3D

Mapa trójwymiarowa odwzorowuje otoczenie w postaci chmury punktów. Każdy punkt opisany jest za pomocą współrzędnych  $(x, y, z)$ , co można zobrazować w odpowiednim środowisku symulacyjnym. W stereowizji każdemu pikselowi obrazu odpowiada konkretna dysparycja i odpowiadająca jej głębia, dlatego współrzędna  $z$  wszystkich punktów na mapie 3D (na podstawie wzoru (3.39)) jest w zasadzie znana. W rzeczywistości jednak współrzędna  $z$  jest ograniczona przez maksymalną dysparycję, której odpowiada punkt znajdujący się blisko kamery, oraz zerową dysparycję, której odpowiada punkt znajdujący się w nieskończonej odległości od kamery. Natomiast współrzędne  $x$  i  $y$  punktów na mapie 3D nie są znane. Dlatego należy je wyznaczyć w oparciu o mapę dysparycji i parametry kamery.

Transformację punktów widocznych w kamerze odniesienia (zwykle lewej) o zadanej dysparycji do punktów punktów 3D można zapisać za pomocą wzoru:

$$\begin{bmatrix} x' & y' & z' & w' \end{bmatrix}^T = \mathbf{Q} \begin{bmatrix} x & y & d(x, y) & 1 \end{bmatrix}^T \quad (3.42)$$

gdzie  $x', y', z', w'$  są współrzędnymi jednorodnymi punktu 3D,  $x$  i  $y$  są współrzędnymi pikselowymi (na matrycy kamery odniesienia),  $d(x, y)$  to dysparycja w punkcie o współrzędnych pikselowych  $(x, y)$ , macierz  $\mathbf{Q}$  jest definiowana jako macierz reprojekcji lub odwzorowania (ang. *Reprojection Matrix*) i stanowi pewne rozszerzenie macierzy (3.31). Zwykle macierz  $\mathbf{Q}$  zawiera parametry wyznaczone podczas kalibracji kamer (Hartley i

Zisserman, 2003):

$$Q = \begin{bmatrix} 1 & 0 & 0 & -c_x \\ 0 & 1 & 0 & -c_y \\ 0 & 0 & 0 & f \\ 0 & 0 & \frac{1}{B} & (c_{x_L} - c_{x_R})/B \end{bmatrix} \quad (3.43)$$

gdzie  $c_x$  i  $c_y$  są współrzędnymi punktu głównego,  $f$  jest ogniskową<sup>3</sup> obiektywu,  $B$  to baza stereo zaś wartość  $(c_{x_L} - c_{x_R})$  stanowi przesunięcie współrzędnej  $x$  punktu głównego między kamerami.

Współrzędne wyznaczanych punktów 3D opisuje się w postaci jednorodnej, zatem ostatecznie te współrzędne mają postać  $(x'/w', y'/w', z'/w')$ . Współczynnik  $w'$  wyraża efekt skalowania i wpływa na wszystkie współrzędne punktu 3D. Zgodnie z (3.42) współczynnik ten można zapisać w następującej postaci:

$$w' = \frac{d + (c_{x_L} - c_{x_R})}{B} \quad (3.44)$$

Nietrudno zauważyć, że  $w'$  zależy od parametrów stałych ( $c_{x_L}$ ,  $c_{x_R}$  oraz  $B$ ) oraz od dysparycji  $d$ , której błąd nie jest stały. Oznacza to, że błąd dysparycji wpływa nie tylko na odległość punktu od kamery (3.40), ale także na błąd położenia punktu w pionie i poziomie (Kowalczyk i Merta, 2016b).

### 3.5 Podsumowanie

W rozdziale omówiono najważniejsze podstawy teoretyczne, których znajomość jest wymagana do zrozumienia problemów związanych z mapowaniem 3D opartym na stereowizji. Podstawy te dotyczą wyznaczania położenia i ułożenia robota oraz modelu matematycznego umożliwiającego opis pozycji w danym układzie odniesienia oraz transformację współrzędnych. W dalszej części opisano zagadnienie estymacji danych w oparciu o model stanowy i obserwator stanu. Omówiono estymację opartą na filtracji Kalmana. Ponadto przedstawiono model kamery oraz zależności geometryczne istotne dla stereowizji, a także sposób odtworzenia chmury punktów 3D na podstawie mapy dysparycji.

W kolejnym rozdziale pracy zaprezentowana zostanie koncepcja systemu odwzorowania otoczenia na mapie cyfrowej. W oparciu o tę koncepcję oraz opisane podstawy teoretyczne zaproponujemy konkretną realizację odwzorowania otoczenia opartego na robocie kołowym. Realizacja uwzględni wybrane zagadnienie stereowizji w celu odwzorowania otoczenia. Natomiast proces odwzorowania realizowany będzie na podstawie estymacji położenia i ułożenia kołowego robota mobilnego. Estymacja ta uwzględni pomiary z sensorów inercyjnych oraz autorski model kołowego robota mobilnego.

<sup>3</sup>Parametry te opisane są jako parametry wewnętrzne kamer (3.35).



---

## SYSTEM VISROBOT

### ZE ZMIENNĄ BAZĄ STEREOWIZYJNĄ

Stereowizja umożliwia trójwymiarowe odwzorowanie otoczenia, gdyż po odpowiednim ustawieniu pary kamer można wyznaczyć chmurę punktów 3D. Takie odwzorowanie realizuje się w oparciu o dwukamerowy system wizyjny zakładając przy tym, że para obrazów pochodzi z jednej kamery, której położenie zmienia się w czasie.

Odwzorowanie otoczenia także wpisuje się w tematykę SLAM, która obejmuje szeroki zakres zagadnień dotyczących autonomicznej nawigacji. W tym kontekście nie skupiając się na typowym rozwiązaniu *online* lub *offline* SLAM, proponujemy innowacyjne podejście do SLAM (stereowizji i odwzorowania otoczenia), w którym system wizyjny oparty jest na zmiennej bazie stereo.

W przypadku robota mobilnego mapującego teren, oprócz właściwego systemu wizyjnego należy uwzględnić odpowiednie procedury określania położenia i ułożenia robota. Metoda wyznaczania pozycji robota zostanie dopasowana do możliwości kołowego robota mobilnego, którego napęd stanowią silniki prądu stałego DC. Należy zaznaczyć, że użycie innej elektromechanicznej konstrukcji platformy robota mogłoby wymagać opracowania zupełnie innego podejścia do zadania determinacji pozycji robota.

Przy realizacji konkretnego systemu wizyjnego robota należy opracować i zaimplementować eksperymentalne środowisko testowe, które pozwoli sprawdzić wykonalność prezentowanego podejścia. Po wykonaniu odpowiednich testów można także ocenić jakościowo system, w szczególności błąd oceny pozycji robota oraz błąd odwzorowania otoczenia.

#### **4.1 Ogólna charakterystyka systemu**

Opracowany system stereowizyjny ze zmienną bazą nazwano systemem VisRobot, który zakłada eksplorację wewnętrznego statycznego środowiska. Przy takim podejściu zmienną bazę stereo można również uzyskać w sytuacji, gdy para obrazów jest rejestrowana nawet przez 1 kamerę w różnych chwilach. W koncepcji tej zakłada się użycie robota kołowego o typowej konstrukcji zawierającej silniki DC, z kołami w części przedniej oraz tylnym kołem, które porusza się swobodnie w każdą stronę. Taka konstrukcja jest dogodna w sterowaniu i nie wymaga uwzględniania mechanizmu różnicowego lub ograniczeń przy skręcaniu. Zakładamy też, że robot jest zaopatrzony w dwie statyczne (nieobrotowe) kamery oraz sensory inercyjne. Trajektoria robota jest tak projektowana, aby robot poruszał się

względnie prostopadłe do odwzorowywanego kierunku (w lewo lub prawo) na stosunkowo niewielką odległość. Dzięki temu odległość między kamerami się zmienia i odpowiednie pary obrazów tworzą zmienną bazę stereoskopową. Ponadto prezentowane rozwiązanie umożliwia także zobrazowanie stereoskopowe wybranej pary obrazów stereo, tak aby możliwe było wyświetlanie jej ludzkiemu operatorowi.

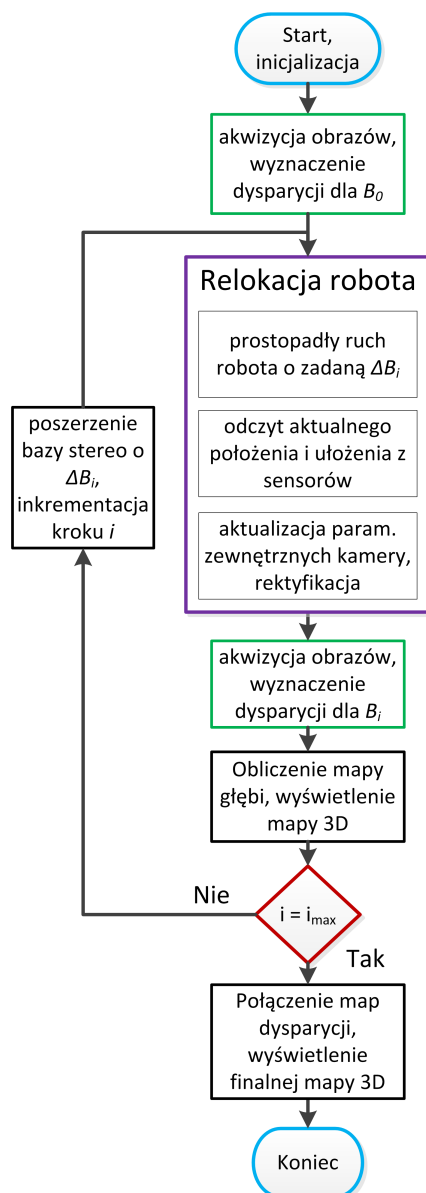
W ramach systemu VisRobot założono wykonanie pojedynczego złożonego pomiaru obiektów widocznych w kadrze kamery. Zmiana położenia robota w systemie ma na celu zwiększenie rozdzielczości głębinowej, czyli możliwości dokładniejszego określenia położenia obiektów widocznych w kadrze, w szczególności tych które znajdują się daleko od robota.

Do poprawnej pracy eksperymentalny system wymaga wyznaczenia ułożenia systemu kamer, co przy nieruchomych kamerach sprowadza się do określenia ułożenia całego robota. Wyznaczanie ułożenia oparto na danych pochodzących z sensorów zawartych w IMU. Ponadto system wymaga pomiaru odległości, którą robot pokonuje podczas przemieszczania. Dzięki wykorzystaniu akcelerometru, sterowanie położeniem robota realizowane może być w pętli otwartej. Podejście takie podyktowane jest stosunkowo niedużymi odległościami, które robot ma pokonywać, a także próbą zweryfikowania niejednoznacznych opinii w literaturze dotyczących użyteczności akcelerometru do wyznaczenia położenia<sup>1</sup>. Warto zaznaczyć, że dane wizyjne używane przez system nie wymagają pracy kamer w trybie ciągłym, gdyż klatki stanowiące stereoparę są zapisywane tylko na żądanie.

Ogólny schemat działania systemu VisRobot jest przedstawiony na rys. 4.1. Po uruchomieniu systemu ustawiane są wszystkie istotne parametry, w szczególności stałe parametry wewnętrzne kamer i początkowe wartości parametrów zewnętrznych, liczba par stereo (liczba kroków)  $i$ , krok przemieszczenia robota  $\Delta B_i$ , który związany jest z wartościami poszczególnych baz  $B_1, B_2, \dots, B_i$ . Ponadto po inicjalizacji wykonywany jest pierwszy pomiar w postaci akwizycji pierwszej stereopary, na podstawie której wyznaczana jest początkowa mapa dysparycji z bazą  $B_0$ . Następnie wykonywana jest pętla programu, w której pojedyncza iteracja ma na celu wyznaczenie mapy dysparycji dla danej bazy  $B_i$ . Na początku każdej iteracji ma miejsce przemieszczenie robota stosownie do szerokości bazy. Robot porusza się w lewo (bądź prawo) o zadaną odległość. Gdy punkt docelowy jest osiągnięty robot pozostaje nieruchomy do końca aktualnej iteracji. Na podstawie danych inercyjnych wyznaczane są aktualne ułożenie i położenie robota, które zapisywane są w postaci parametrów zewnętrznych kamer, co można zapisać w postaci wektora  $[R \ T]$ .

Dla aktualnego położenia robot wykonuje akwizycję obrazu z kamery (lewej lub prawej) stosownie do ustawień początkowych. Przy ustalonej szerokości bazy  $B_i$ , wyznaczana jest para obrazów stereo, w której skład wchodzi aktualna ( $i$ -ta) klatka oraz pierwsza zapisana klatka przy bazie  $B_0$ . Oba obrazy są poddawane rektyfikacji, po czym system wyznacza i zapisuje mapę dysparycji. Opcjonalnie na tym etapie mapa dysparycji może być przekształcona na mapę głębi i wyświetlona w postaci chmury punktów lub obrazu 3D z użyciem danej techniki stereoskopowej. Jeżeli aktualna iteracja nie jest ostatnia, algorytm aktualizuje indeks  $i$  oraz powtarza cały cykl dla nowej bazy  $B_i$ . Proces ten kończy się, gdy  $i$  osiągnie wartość  $i_{max}$  zadaną wcześniej przez operatora. Po zakończeniu pętli wszystkie zapisane mapy dysparycji są odpowiednio łączone tak, aby uzyskać ulepszoną mapę głębi o większej rozdzielczości głębinowej. Finalna mapa głębi jest przekazywana do

<sup>1</sup>Źródła literaturowe oraz uzasadnienie użycia akcelerometru podano na początku rozdziału „Teoretyczne podstawy pomiarów”.



Rysunek 4.1 Schemat działania systemu VisRobot ze zmienną bazą.

modułu wizualizacji, który umożliwia konwersję mapy głębi do chmury punktów 3D lub odpowiedniego obrazu 3D.

Opisany schemat postępowania w praktyczny sposób rozwiązuje omawiany problem odwzorowania ograniczonego środowiska na mapie cyfrowej z wykorzystaniem robota mobilnego zaopatrzonego w stereowizję. Szczegółowe mechanizmy tego rozwiązania, dotyczące wyznaczania położenia i ułożenia robota, sterowania, wyznaczania mapy dysparycji oraz precyzji odwzorowania obiektów na mapie wymagają dalszej analizy.

## 4.2 Wyznaczanie ułożenia robota

Problem ułożenia robota i kamer zainstalowanych na robocie opiera się na pomiarach z sensorów wchodzących w skład jednostki IMU. Najważniejsze cechy oraz specyfika

miaru wielkości fizycznych, które są mierzone przez akcelerometr, magnetometr oraz żyroskop, zostały opisane w sekcjach 2.3.2 oraz 3.1. Konieczność wyznaczenia ułożenia jest związana z dwoma aspektami. Z perspektywy mapowania istnieje bowiem potrzeba, aby wyznaczać aktualne ułożenie kamer, znajdujących się na platformie robota. Drugim aspektem jest wyznaczenie położenia w oparciu o dane z akcelerometru, które wymaga wcześniejszego wyznaczenia ułożenia w celu usunięcia składowej grawitacyjnej. Jeżeli ułożenie nie zostanie wyznaczone wystarczająco dokładnie to również estymowana składowa grawitacji będzie odchyłona od grawitacji rzeczywistej. Wówczas pomimo usunięcia mierzonego przyspieszenia ziemskiego według danych akcelerometrycznych, pozostanie pewne resztowe przyspieszenie, co istotnie wpłynie na błąd położenia.

Ułożenie kamery w oparciu o dane inercyjne jest możliwe z wykorzystaniem pojedynczego 3-osowego magnetometru. Jednak pomiar ten może być niewystarczający szczególnie przy wyznaczaniu położenia. Dlatego też w opracowanym systemie określenie ułożenia oparto na fuzji danych z magnetometru, żyroskopu i akcelerometru. Należy zaznaczyć, że dokładność pomiaru ułożenia zależy w dużym stopniu od jakości sensorów i systemów niwelujących wpływ warunków zewnętrznych. Szczegóły dotyczące najważniejszych czynników, które wpływają na poziom szumu sensora zostały przedstawione w sekcji 3.2.3. Oczywiście sensory muszą być odpowiednio skalibrowane przed fuzją danych.

#### 4.2.1. Kalibracja sensorów inercyjnych

Celem kalibracji jest ustawienie sensora tak, aby pomiar był zgodny z daną wielkością fizyczną mierzoną w określonym miejscu, przy określonych warunkach lokalnych. Akcelerometr mierzy przyspieszenie obiektu, w tym także przyspieszenie ziemskie. Dlatego kalibracja akcelerometru wymaga wykonania pomiaru przy sensorze ułożonym tak, aby wektor grawitacji był równoległy do jednej z osi akcelerometru. Następnie należy obracać sensor o wielokrotność  $90^\circ$  wokół każdej osi i dokonywać kolejnych pomiarów. Jeżeli przy żadnym ułożeniu pomiary nie pokazują pełnego przyspieszenia ziemskiego ( $\pm 1g$ ), konieczne jest wprowadzenie odpowiedniej korekty. W praktyce po zebraniu pomiarów kalibracyjnych dopasowuje się współczynnik skalujący i odchyłkę (offset) dla każdej z trzech osi tak, aby zminimalizować błąd przyspieszenia (Sipos *et al.*, 2012).

Kalibracja akcelerometru zwykle realizowana jest tylko raz, ponieważ przyspieszenie ziemskie w różnych miejscach na Ziemi różni się nieznacznie. Nieduże zmiany dotyczą wartości (modułu) wektora przyspieszenia (Harms, 2015), natomiast kierunek tego wektora jest generalnie stały, skierowany do środka masy Ziemi. Dlatego inicjalna kalibracja wykonywana przez producenta powinna dawać poprawne wyniki.

Dane z żyroskopu nie są zależne od zewnętrznych wielkości fizycznych, zatem kalibrację żyroskopu wykonuje się zwykle tylko raz. Proces kalibracji można zrealizować w oparciu dane referencyjne pochodzące ze stanowiska laboratoryjnego, które pozwala na kontrolowany obrót sensora o zadany kąt względem wybranej osi. Niemniej można spotkać rozwiązania oparte na danych referencyjnych z innego sensora, np. akcelerometru. Przy kalibracji kluczowe są 2 parametry. Należy ustawić właściwą wartość zerowej prędkości obrotowej dla każdej osi, dzięki czemu sensor wskaże poprawną zerową prędkość obrotową (oscylacje względem zera) przy braku takiego ruchu. Drugim parametrem jest czułość, która wpływa na wartość prędkości obrotowej. Niezależnie od danych referencyjnych kalibrację żyroskopu można oprzeć na odrębnym pomiarze prędkości obrotowej,



co w ogólności wymaga specjalnego sprzętu do stosunkowo trudnego pomiaru, lub na całkowaniu prędkości obrotowej podczas krótkiego ruchu (Wu i Pei, 2017).

Magnetometr, w przeciwieństwie do akcelerometru i żyroskopu, wymaga częstej kalibracji. Sensor pozwala określić ułożenie na podstawie pomiaru pola magnetycznego Ziemi. Elementy takie jak dewiacja i deklinacja magnetyczna, które wpływają na pomiar, zostały omówione w sekcji 3.1.1. Opracowany system zakłada, że sensor jest używany lokalnie, więc korygowanie danych z magnetometru jest związane z obecnością ferromagnetyków w bliskim otoczeniu sensora. Dokładne określenie, jak ferromagnetyki wpływają na pomiar magnetometru, jest zadaniem trudnym, ponieważ zależy od ich materiału, rozmiaru, kształtu, ułożenia względem pola magnetycznego Ziemi, itp. Niemniej można wyznaczyć składową stałą natężenia pola magnetycznego (dla każdej osi) dla ferromagnetyków umieszczonych na robocie. Natomiast inne silne ferromagnetyki należy umieścić możliwie daleko od robota biorąc pod uwagę fakt, że natężenie pola magnetycznego maleje wraz z kwadratem odległości (Hu i You, 2010).

Procedura kalibracji magnetometru polega na obracaniu sensora w taki sposób, aby otrzymać możliwie dużo różnych ułożeń sensora. Odczytane wartości dla danej osi zawierają natężenie pola magnetycznego Ziemi (przy danym ułożeniu) wraz ze składową stałą, pochodzącą od ferromagnetyków. Dla trzech osi zaszumione dane można przedstawić w postaci punktów na powierzchni elipsoidy, której współrzędne poszczególnych środków odpowiadają kolejnym składowym stałym związanym z obecnością ferromagnetyków. Kalibracja polega na znalezieniu takich dwóch parametrów dla każdej osi, które pozwolą na transformację elipsoidy (rys. 4.2a) do kuli o zadanym promieniu i środku położonym w początku układu współrzędnych (rys. 4.2b). W najprostszym przypadku można oprzeć się na wartości maksymalnej i minimalnej dla każdej osi. Jednak obecność szumu (którego skutki widać na rys. 4.2) może powodować błędy kalibracji. Dlatego w systemie VisRobot zastosowano algorytm kalibracji oparty na metodzie najmniejszych kwadratów. Metoda w kolejnych krokach ocenia, w jakim stopniu dane są dopasowane do sfery. Błąd kalibracji wyznaczamy jako średnie niedopasowanie wszystkich mierzonych punktów do sfery.

#### ***4.2.2. Fuzja danych inercyjnych***

Poprawnie skalibrowane sensory inercyjne dostarczają danych wejściowych dla algorytmu fuzji danych. Zadaniem takiego algorytmu w omawianym przypadku jest takie złożenie danych pochodzących z różnych sensorów, aby otrzymać dokładniejsze ułożenie obiektu w porównaniu z ułożeniem wyznaczonym za pomocą pojedynczego sensora.

Każdy sensor dostarcza indywidualnych danych związanych z ułożeniem. Zatem problem fuzji sprowadza się do odpowiedniej filtracji, która usuwa dane obniżające dokładność pomiaru oraz estymacji zmiennych. W systemie VisRobot zaimplementowano algorytm, który opiera się na całkowaniu danych z żyroskopu w układzie odniesienia definiowanym przez północ magnetyczną (magnetometr) i wektor grawitacji (akcelerometr). Ze względu na możliwość zajścia zjawiska blokady przegubu, wewnętrzne operacje obrotu oparte są na kwaternionach. Dane wyjściowe przedstawione mogą być w postaci kwaternionów lub kątów Eulera. Schemat działania programu przedstawiono na rys. 4.3. Po jego uruchomieniu, następuje inicjalizacja ustawień IMU, takich jak czułość danego sensora, parametrów transmisji danych, częstotliwości. W kolejnym etapie ładowane są dane kali-

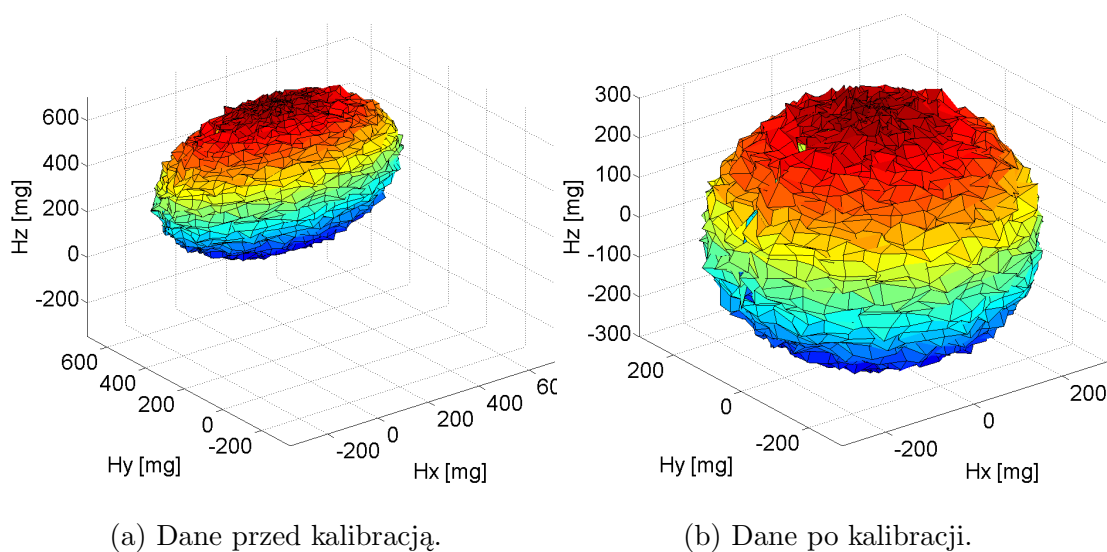
bracyjne. Następnie realizowana jest główna pętla programu, w której pojedyncza iteracja musi być wykonana w czasie  $1/f_s$ , gdzie  $f_s$  jest częstotliwością próbkowania.

W pojedynczej iteracji program odczytuje dane z akcelerometru, żyroskopu i magnetometru. Bezpośrednio po odczycie dane z magnetometru są korygowane zgodnie z załadowanymi danymi kalibracyjnymi. Następnie program sprawdza, czy dane z akcelerometru mogą być użyte do określania ułożenia. Jeżeli występuje stosunkowo gwałtowne przemieszczenie sensora to kierunek wektora grawitacji wyznaczany za pomocą akcelerometru jest obciążony błędem. Jeżeli norma wektora przyspieszenia  $|a_0|$  przekracza wartość progową  $a_p$  to akcelerometr nie może być użyty do określania ułożenia w aktualnej iteracji (kiedy  $|a_0|$  obliczone jest przez odjęcie wartości przyspieszenia ziemskiego  $1g$  od mierzonego przyspieszenia).

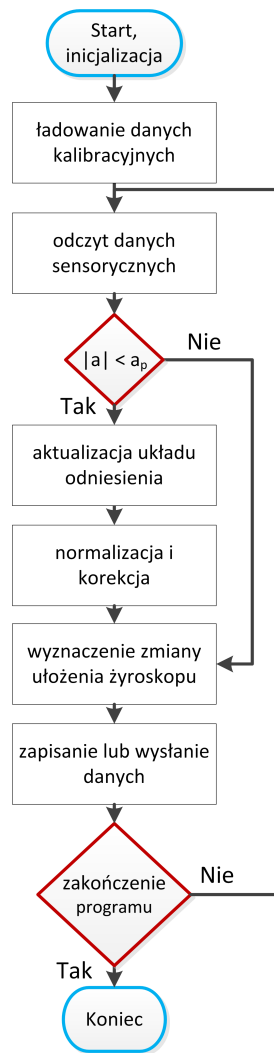
Jeżeli przyspieszenie  $|a_0|$  jest zbyt duże to wyznaczanie ułożenia opiera się tylko na podstawie całkowanych danych z żyroskopu. W przeciwnym wypadku układ odniesienia aktualizowany jest na podstawie danych magnetometru (północ magnetyczna), akcelerometru (pion), zaś kierunek poziomy należy obliczyć w oparciu o trzecią oś prostopadłą do dwóch pozostałych. Ponadto dane związane z układem odniesienia są odpowiednio korygowane i normalizowane. Dzięki temu otrzymuje się ortogonalny układ odniesienia (kierunki) z zaszumionych pomiarów magnetometru i akcelerometru.

Dalsza część programu wykonywana jest niezależnie od spełnienia warunku  $|a_0| < a_p$ . Dane z żyroskopu są całkowane w czasie  $T_g = 1/f_s$ , po czym układ odniesienia jest obracany o wyznaczoną drogę kątową. Podczas aktualizacji układu odniesienia, a także po scałkowaniu danych żyroskopu, obroty są realizowane w oparciu o rachunek kwaternionowy. Jest to szczególnie istotne, gdy sensor jest przechylany lub pochylany o kąt bliski  $90^\circ$  względem początkowego ułożenia, aby uniknąć zjawiska blokady przegubu.

Przedstawione rozwiązanie pozwala na stosunkowo dokładne wyznaczenie ułożenia obiektu. Zastosowanie filtracji Kalmana dla danych z żyroskopu oraz ułożenia układu odniesienia (na podstawie danych z magnetometru i akcelerometru) pozwala na mini-



Rysunek 4.2 Kalibracja magnetometru.



Rysunek 4.3 Schemat algorytmu wyznaczania ułożenia oparty na danych inercyjnych.

malizację skutków szumu w danych ułożenia (Gui *et al.*, 2015; Madgwick *et al.*, 2011). Filtr pozwala uzyskiwać mały błąd, gdy obrót obiektu odbywa się w sposób płynny. Przy gwałtownych zmianach ułożenia, filtr może dawać większy błąd niż rozwiązanie przedstawione na rys. 4.3, co jest zależne od przyjętej wartości wariancji szumu pomiarowego i procesowego. Podczas mapowania realizowanego przez robota kołowego nie występują gwałtowne obroty, dlatego warto zastosować filtrację, która wygładza dane pomiarowe. Wyznaczane ułożenie ma istotny wpływ na położenia kamery i dalej na mapowanie otoczenia. Dlatego w części symulacyjnej i doświadczalnej należy określić, w jakim stopniu błąd ułożenia (wg użytej metody) wpływa na błąd położenia i błąd mapowania. Co więcej w eksperymentach warto zdefiniować, jakie wymagania należy spełnić, aby mapowanie było realizowane z góry zadaną dokładnością.

### 4.3 Wyznaczanie położenia robota

Podejścia do problemu wyznaczania położenia obiektu opisano w sekcji 3.1, a także na początku rozdziału 3. Wyznaczanie położenia na podstawie danych inercyjnych wymaga

podwójnego całkowania danych przyspieszenia pochodzących z akcelerometru. Ze względu na obecność grawitacji akcelerometr mierzy zarówno rzeczywiste przyspieszenie obiektu, jak i przyspieszenie ziemskie, co w konwencji wektorowej można zapisać wzorem:

$$\mathbf{a}_m = \mathbf{a}_o + \mathbf{g} \quad (4.1)$$

gdzie  $\mathbf{a}_m$  jest przyspieszeniem mierzonym przez akcelerometr,  $\mathbf{a}_o$  jest rzeczywistym przyspieszeniem obiektu, a  $\mathbf{g}$  to przyspieszenie ziemskie. Aby określić położenie na podstawie przyspieszenia obiektu  $\mathbf{a}_o$ , należy usunąć wpływ grawitacji, która w ogólności zmienia się wraz ze zmianą ułożenia obiektu. We współrzędnych zewnętrznego świata wektor grawitacji jest stały, jedynie zmienia się ułożenie obiektu. Dlatego też wyznaczanie położenia wymaga znajomości ułożenia obiektu w każdej chwili, w której pomiar jest dokonywany. Ponadto podczas obrotu obiektu występują dodatkowe przyspieszenia. Aby właściwie opisać zależności występujące między przyspieszeniami należy zdefiniować możliwie prosty model akcelerometru.

#### 4.3.1. Model akcelerometru

Model akcelerometru opracowany na potrzeby pracy uwzględnia wpływ wszystkich istotnych sił, które powstają zarówno w zewnętrznym układzie odniesienia, jak i w układzie odniesienia obiektu. Spójny opis wymaga transformacji odpowiednich przyspieszeń do wspólnego układu odniesienia zgodnie ze wzorem (3.1). Dla zachowania czytelności zapisu zmienne kartezyjskie w zewnętrznym układzie współrzędnych będą wyrażone w postaci  $a_x, a_y, a_z$ , natomiast zmienne w układzie odniesienia obiektu w postaci  $a_l, a_s, a_t$ , gdzie  $l$  jest składową wzdłużną (ang. *Longitudinal Component*),  $s$  jest składową boczną (ang. *Side Component*), zaś  $t$  to składowa poprzeczna (ang. *Transversal Component*).

W sytuacji statycznej, gdy obiekt się nie porusza, przyspieszenia działające na obiekt można zapisać za pomocą następującego równania:

$$\begin{bmatrix} \tilde{a}_x \\ \tilde{a}_y \\ \tilde{a}_z \end{bmatrix} = \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} \quad (4.2)$$

gdzie  $\tilde{a}_x, \tilde{a}_y, \tilde{a}_z$  to składowe wektora mierzonego przyspieszenia w zewnętrznym układzie odniesienia. Właściwy pomiar ma miejsce w układzie odniesienia obiektu, zatem (4.2) można zapisać w bardziej użytecznej formie przedstawionej poniżej:

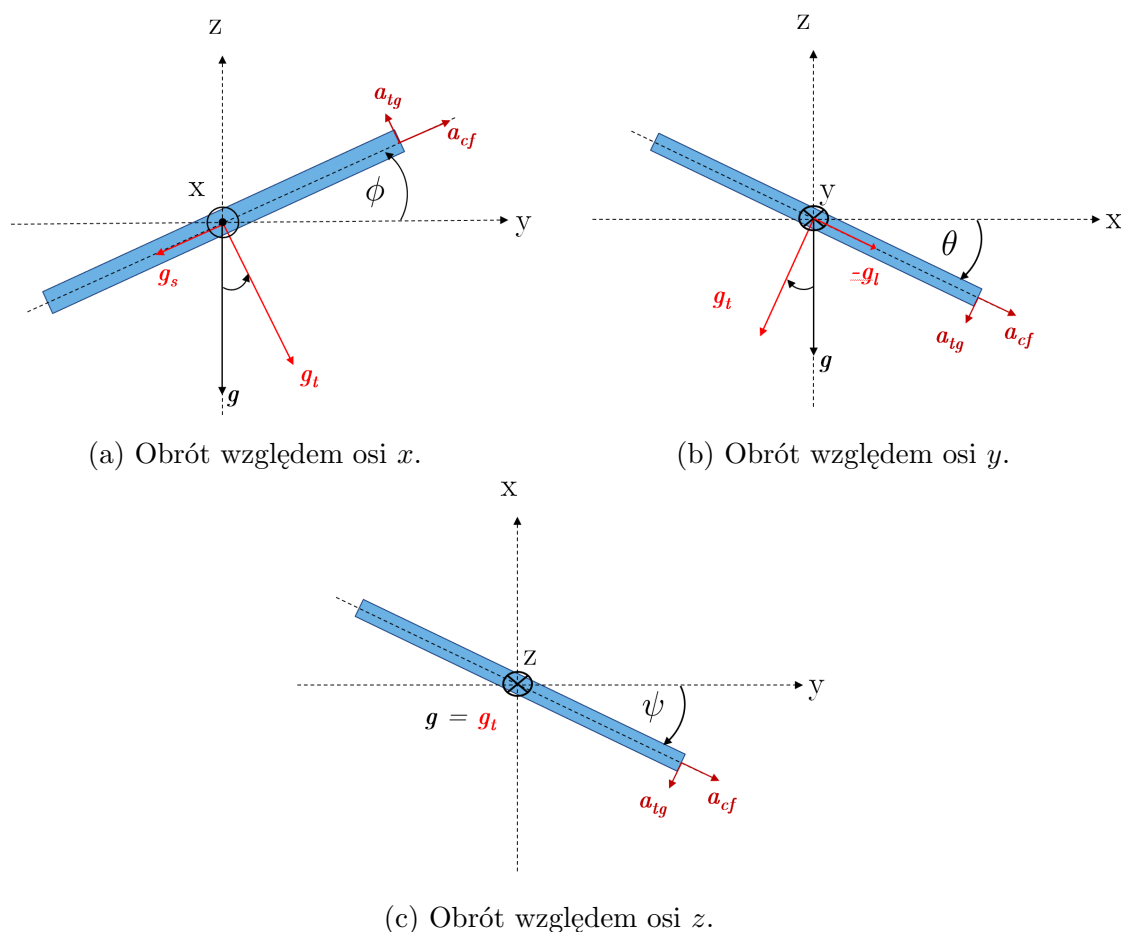
$$\begin{bmatrix} \tilde{a}_l \\ \tilde{a}_s \\ \tilde{a}_t \end{bmatrix} = \mathbf{R}_{xyz} \begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} + \mathbf{R}_{xyz} \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} = \begin{bmatrix} a_l \\ a_s \\ a_t \end{bmatrix} + g \begin{bmatrix} -\sin\theta \\ \sin\phi \cos\theta \\ \cos\phi \cos\theta \end{bmatrix} \quad (4.3)$$

gdzie  $\mathbf{R}_{xyz}$  jest rotacją opisującą obrót współrzędnych zewnętrznych względem współrzędnych obiektu zgodnie z wyrażeniem (3.6).

W przypadku, gdy na obiekt działa dodatkowa siła, która powoduje obrót obiektu należy uwzględnić przyspieszenie odśrodkowe  $a_{cf}$  (ang. *Centrifugal Acceleration*) oraz styczne  $a_{tg}$  (ang. *Tangential Acceleration*). W zależności od osi obrotu ( $x, y, z$ ) pojawiają



się istotne różnice odnośnie kierunków przyspieszeń, co zostało przedstawione odpowiednio na rys. 4.4a, rys. 4.4b, rys. 4.4c.



Rysunek 4.4 Przyspieszenia podczas obrotu sensora względem osi  $x$ ,  $y$ ,  $z$ .

Warto zauważyć, że obrót względem osi  $z$  nie wpływa na składowe wektora grawitacji co oznacza, że dane z akcelerometru nie pozwalają określić zmiany ułożenia związanej z myśkowaniem. Ponadto, przyspieszenie odśrodkowe  $a_{cf}$  jest (oczywiście) prostopadłe do kierunku obrotu, zaś przyspieszenie styczne  $a_{tg}$  jest zgodne z kierunkiem obrotu. Oba przyspieszenia są opisane względem układu odniesienia obiektu i są zależne od odległości sensora od środka obrotu  $r$ . W praktyce promień obrotu  $r$  nie zawsze jest stały dla wszystkich (trzech) osi obrotu, dlatego należy uwzględnić składowe  $r_l$ ,  $r_s$ ,  $r_t$  promienia.

Przyspieszenie odśrodkowe  $a_{cf}$  występuje, gdy prędkość kątowna  $(\dot{\phi}, \dot{\theta}, \dot{\psi})$  jest niezerowa, natomiast przyspieszenie styczne  $a_{tg}$  przyjmuje niezerowe wartości, gdy przyspieszenie kątowne  $(\ddot{\phi}, \ddot{\theta}, \ddot{\psi})$  jest niezerowe. W przypadku, gdy obiekt jest jednocześnie przesuwany i obracany mierzone przyspieszenie można zapisać w postaci:

$$\begin{bmatrix} \tilde{a}_l \\ \tilde{a}_s \\ \tilde{a}_t \end{bmatrix} = \begin{bmatrix} a_l \\ a_s \\ a_t \end{bmatrix} + g \begin{bmatrix} -\sin\theta \\ \sin\phi \cos\theta \\ \cos\phi \cos\theta \end{bmatrix} + \begin{bmatrix} (\dot{\theta}^2 - \ddot{\psi})r_l \\ (\dot{\phi}^2 + \dot{\psi}^2)r_s \\ (\ddot{\phi} - \ddot{\theta})r_t \end{bmatrix} \quad (4.4)$$



Poprzez analizę rys. 4.4 wyznaczyć można ostatni składnik związany z przyspieszeniem stycznym i kątowym. Zakładamy, że przed wystąpieniem obrotu osie zewnętrznego układu współrzędnych  $x, y, z$  pokrywają się z osiami sensora odpowiednio  $a_l, a_s, a_t$ . W takim przypadku przy obrocie względem osi  $x$  pojawia się przyspieszenie kątowe w kierunku  $a_s$  oraz przyspieszenie styczne w kierunku  $a_t$ . Analogicznie, przy obrocie względem osi  $y$  przyspieszenie kątowe występuje w kierunku  $a_l$ , styczne w kierunku  $-a_t$ , zaś przy obrocie względem osi  $z$  przyspieszenie kątowe wpływa na kierunek  $a_s$ , a styczne na kierunek  $-a_l$ .

Oprócz przyspieszeń związanych z obracaniem obiektu można wyróżnić także zjawisko precesji i siły Coriolisa, które mają wpływ na wskazania sensora IMU i są zależne od prędkości obrotowej (Dauwalter i Ha, 2005; Almabrouk *et al.*, 2016). Jednak w rozważanym zastosowaniu kołowego robota mobilnego wpływ precesji i siły Coriolisa będzie pominięty, gdyż robot wykonuje tylko małe i wolne obroty podczas skręcania. Szum w modelu akcelerometru musi być uwzględniony, gdyż ma on istotny wpływ na mierzone przyspieszenia. Ostateczna postać modelu akcelerometru wyraża się za pomocą następującego wyrażenia:

$$\begin{bmatrix} \tilde{a}_l \\ \tilde{a}_s \\ \tilde{a}_t \end{bmatrix} = \begin{bmatrix} a_l \\ a_s \\ a_t \end{bmatrix} + g \begin{bmatrix} -\sin\theta \\ \sin\phi \cos\theta \\ \cos\phi \cos\theta \end{bmatrix} + \begin{bmatrix} (\dot{\theta}^2 - \ddot{\psi})r_l \\ (\dot{\phi}^2 + \dot{\psi}^2)r_s \\ (\ddot{\phi} - \ddot{\theta})r_t \end{bmatrix} + \begin{bmatrix} \eta_l \\ \eta_s \\ \eta_t \end{bmatrix} \quad (4.5)$$

gdzie  $\eta_l, \eta_s, \eta_t$  reprezentuje szum sensora w kierunkach odpowiednio wzdłużnym, bocznym i poprzecznym. Powyższy model pozwala ocenić, jak poszczególne elementy wpływają na przyspieszenie mierzone przez sensor. W tym celu należy przeprowadzić odpowiednią analizę pozwalającą określić poziom błędów położenia przy zadanych wartościach parametrów.

Prosta analiza modelu pozwala ocenić, jak dany parametr wpływa na mierzone przyspieszenie. Oczywiście w takim przypadku zakłada się niezmiennosc pozostałych parametrów (*ceteris paribus*). Przykładowo, jeżeli wektor grawitacji równoległy do  $a_t$  zostanie odchyłony o  $1^\circ$  względem  $a_l$  lub  $a_s$  to pomiary akcelerometru wskażą dodatkowe przyspieszenie około  $17\text{cm/s}^2$  w kierunkach odpowiednio  $a_s$  lub  $a_l$ .

Za pomocą modelu można także oszacować wpływ określonego parametru na położenie. Przy estymacji położenia opartej tylko na danych z akcelerometru można stwierdzić przykładowo, że szum sensora  $1\text{cm/s}^2$  powoduje zwiększenie błędów prędkości o  $1\text{cm}$  z każdą sekundą oraz błąd położenia  $50\text{cm}$  po czasie  $10\text{s}$ .

Oszacowanie wpływu błędnego ułożenia podczas obrotu o dużej dynamice, a także wpływ wielu elementów jednocześnie wymaga dokładnego opisu obrotu, co najprościej można wyznaczyć dopiero na drodze symulacji. Dlatego w dalszej części pracy opisano wyniki symulacji wybranych sygnałów związanych z ruchem robota mobilnego. W szczególności na drodze symulacyjnej określono warunki graniczne, przy których występuje maksymalny błąd położenia akceptowalny w systemie.

#### 4.3.2. Model kinematyki robota mobilnego

Rozważany w pracy robot mobilny jest nieholonomiczny i ma 2 koła stałe na wspólnej osi oraz jedno koło obrotowe w tylnej części platformy robota. Tylne koło (samonastawne) umieszczone jest na wzdłużnej osi symetrii robota. Ograniczenia kół stałych uniemożliwiają ruch robota bezpośrednio<sup>2</sup> w lewo/prawo. Zaletą tego ograniczenia jest możliwość

<sup>2</sup>prostokątne do płaszczyzny kół, oczywiście bez wcześniejszego obrotu w miejscu.

Tablica 4.1 Ogólne parametry kinematyczne opisujące rozważanego robota mobilnego.

parametr	$\alpha$	$\beta$	$l$
koło 1f	0	0	L
koło 2f	$\pi$	0	L
koło 3oc	$3/2\pi$	-	L

utrzymywania pozycji robota w czasie postoju na lekko pochyłej powierzchni (w lewą lub prawą stronę). Robot o takich właściwościach należy do klasy Typu (2,0), czyli ze stopniem mobilności i sterowności równymi odpowiednio 2 i 0 (Siciliano i Khatib, 2008).

Zwykle kinematykę robota opisuje się na płaszczyźnie. Wobec czego pozycję robota (położenie i ułożenie) określa wektor:

$$\xi = [x \ y \ \psi]^T \quad (4.6)$$

gdzie  $(x, y)$  opisuje wybrany punkt platformy robota (np. środek geometryczny) będący jednocześnie punktem referencyjnym  $P$  w zewnętrznym układzie współrzędnych, zaś  $\psi$  jest ułożeniem robota na płaszczyźnie (obrót realizuje się zatem w oparciu o jedną macierz rotacji). Aby opisać pozycję kół zamocowanych na platformie robota należy zdefiniować następujące parametry:

- $l$  - odległość zamocowania koła od punktu referencyjnego  $P$  (promień we współrzędnych biegunowych),
- $\alpha$  - ułożenie koła względem zewnętrznego układu współrzędnych (kąt we współrzędnych biegunowych),
- $\beta$  - ułożenie płaszczyzny koła względem odcinka  $l$ ,
- $d_{oc}$  - odległość środka koła samonastawnego od punktu mocowania koła,
- $r_w$  - promień koła (stałego i samonastawnego).

W oparciu o powyższe parametry definiuje się ograniczenia w postaci dwóch równań, które opisują ograniczenie wzdłuż i w poprzek płaszczyzny koła (Campion *et al.*, 2006). Ogólny układ takich równań można przedstawić w postaci:

$$\begin{aligned} \mathbf{J}_1(\beta_c, \beta_{oc})\mathbf{R}(\psi)\dot{\xi} + \mathbf{J}_2\dot{\varrho} &= 0 \\ \mathbf{C}_1(\beta_c, \beta_{oc})\mathbf{R}(\psi)\dot{\xi} + \mathbf{C}_2\dot{\beta}_{oc} &= 0 \end{aligned} \quad (4.7)$$

gdzie  $\mathbf{J}_1$  jest wyrażeniem opisującym ograniczenia wzdłużne (dla każdego typu koła),  $\mathbf{J}_2$  stanowi współczynnik skalujący wyrażający wpływ rozmiaru kół (tutaj promień koła jest stały i równy  $r_w$ ),  $\mathbf{C}_1$  opisuje ograniczenia poprzeczne związane z typami kół,  $\mathbf{C}_2$  uwzględnia ograniczenia poprzeczne wynikające ze stałego przesunięcia koła samonastawnego  $d_{oc}$ ,  $\beta_c$  jest parametrem koła obrotowego wycelowanego (nie uwzględniane

w pracy),  $\beta_{oc}$  stanowi parametr koła samonastawnego (tylnego),  $\varrho$  to kąt obrotu koła samonastawnego,  $\mathbf{R}$  jest macierzą rotacji<sup>3</sup>.

Parametry związane z kinematyką robota z 2 kołami stałymi (1f, 2f) oraz kołem samonastawnym zamocowanym niecentralnie (3oc) podano w tab. 4.1 (za długość  $l$  przyjęto wartość  $L$ ). Uwzględniając te parametry, pełny model kinematyczny uzyskuje postać powyższej (4.7) pary równań, gdzie

$$\mathbf{J}_1 = \begin{bmatrix} 0 & 1 & L \\ 1 & -1 & L \\ \cos(\beta_{oc}) & \sin(\beta_{oc}) & L\cos(\beta_{oc}) \end{bmatrix} \quad \mathbf{J}_2 = r_w$$

$$\mathbf{C}_1 = \begin{bmatrix} 1 & 0 & 0 \\ -1 & 0 & 0 \\ \sin(\beta_{oc}) & -\cos(\beta_{oc}) & \varrho + L\sin(\beta_{oc}) \end{bmatrix} \quad \mathbf{C}_2 = \begin{bmatrix} 0 \\ 0 \\ d_{oc} \end{bmatrix} \quad (4.8)$$

Postać ta opisuje wszystkie istotne ograniczenia ruchu robota. Jednak nie reprezentuje bezpośrednio modelu stanowego. Pełny opis konwersji powyższej postaci do tzw. modelu kinematycznego postawy (ang. *Posture Kinematic Model*) można znaleźć w (Siciliano i Khatib, 2008). Generalnie, w przypadku rozważanego w pracy robota (z kołami stałymi oraz kołem samonastawnym), okazuje się, że prędkość jest ograniczona do postaci

$$\dot{\xi} = \mathbf{R}^T(\psi) \Sigma \mathbf{u} \quad (4.9)$$

gdzie zbiór wektorów rozpinających tworzy bazę  $\Sigma = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^T$  w przestrzeni sterowań  $\mathbf{u}$  (wymuszenie). W oparciu o powyższe równanie, model kinematyczny postawy ze sterowaniem różnicowym dwóch kół stałych i trzecim kołem (niesterowanym i nie blokującym ruchu), zapisać można w postaci

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} -\sin(\psi) & 0 \\ \cos(\psi) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}. \quad (4.10)$$

Warto zaznaczyć, że w przypadku innych robotów mobilnych zachodzić mogą dodatkowe komplikacje. Przykładowo, jeśli robot jest Typu (2, $x$ ), gdzie  $x > 0$  (tj. stopień sterowności jest niezerowy), baza  $\Sigma$  będzie uzależniona od parametru  $\beta_c$ <sup>4</sup>. Formę modelu kinematycznego postawy można przekształcić także do modelu dynamicznego konfiguracji (ang. *Configuration Dynamic Model*) poprzez formalizm Lagrange'a, jednak taki opis nie jest przedmiotowy w kontekście niniejszej pracy.

Model kinematyczny postawy (holonomicznego) robota Typu (3,0) jest całkowicie sterowalny oraz linearyzowalny. Procedurę linearyzacji można przeprowadzić korzystając ze sztywnego sprzężenia od stanu (ang. *Static State Feedback*).

<sup>3</sup>Dla zachowania spójności opisu  $\mathbf{R}(\psi) = \mathbf{R}_x(\psi)$ ,  $\mathbf{R}_x$  opisano za pomocą wzoru (3.3)

<sup>4</sup>Parametr zdefiniowano wcześniej w opisie (4.7)



Natomiast model robota Typu (2,0) analizowany w niniejszej rozprawie formalnie jest niesterowalny i nielinearyzowalny. Tym niemniej pełna linearyzacja może zostać osiągnięta poprzez rozważanie ogólniejszej postaci z dynamicznym sprzężeniem zwrotnym od stanu oraz rozszerzenie stanu o dodatkowe wejście (Isidori, 1995).

Warto zaznaczyć, że model kinematyczny postawy robota Typu (2,0) jest stabilizowalny przez sztywne sprzężenie od stanu ale poprzez wprowadzenie wariantności (zmienności w czasie). W podsumowaniu można zatem przyjąć, że robot może być przemieszczony do dowolnej pozycji w skończonym czasie poprzez zastosowanie odpowiedniej modyfikacji wejścia prędkościowego (Siciliano i Khatib, 2008).

W dysertacji założono realizację prostych ruchów polegających na obrocie w miejscu oraz jeździe do przodu. Założenie takie oznacza, że uproszczone sterowanie robota może być oparte na skalarnym sygnale sterującym.

#### **4.3.3. Model dynamiki robota mobilnego**

Model akcelerometru przedstawiony w sekcji 4.3.1 pozwala ocenić dokładność pomiaru położenia. Jednakże bezpośrednio mierzoną wielkością jest przyspieszenie. Jest ono obserwowalne podobnie jak wyższe pochodne przyspieszenia. Prędkość i położenie można obliczyć całkując dane przyspieszenia, natomiast z perspektywy zależności dynamicznych wielkości te nie są w ogólności obserwowalne (Kowalczyk i Merta, 2015). Estymując położenia tylko w oparciu o pomiar przyspieszenia, nie uwzględnia się ograniczeń ruchu robota związanych z napędem. Aby zminimalizować skumulowane błędy pomiarowe związane z całkowaniem zaszumionych danych przyspieszenia, należy zdefiniować i zastosować odpowiedni model dynamiki robota. W zastosowanym układzie eksperymentu, wymagającym realizacji prostych ruchów (ruch do przodu oraz obroty w miejscu), do syntezy układu sterowania wystarczy uwzględnić uproszczony model kinematyczny robota.

Robot napędzany jest silnikami prądu stałego (DC), więc model dynamiki robota wykorzystuje parametry silnika prądu stałego. Model zakłada, że robot wyposażony jest w dwa jednakowe silniki, gdzie główny wał napędowy każdego silnika wprawia w ruch jedno koło. Założono także, że masa platformy oraz wszystkich elementów umieszczonych na niej jest rozłożona równomiernie między dwa koła z napędem. Ponadto, zgodnie z przyjętymi założeniami dotyczącymi modelu kinematycznego, wpływ tylnego koła na dynamikę ruchu należy uznać za znikomy (ze względu na zamocowanie go na wzdłużnej osi symetrii).

#### *Model silnika prądu stałego*

Właściwy model robota wymaga odpowiedniego opisu silnika prądu stałego. W literaturze typowy silnik DC jest modelowany za pomocą układu drugiego rzędu (Wolm *et al.*, 2008; Kowalczyk i Merta, 2016a). W pewnych zastosowaniach można uprościć model do postaci układu pierwszego rzędu, czego konsekwencją może być znaczący błąd podczas przyspieszania robota. Dlatego w opracowanym systemie model opiera się na równaniu różniczkowym drugiego rzędu sprowadzonym do modelu stanowego, gdzie wymuszeniem jest napięcie  $V$ , zaś wektor stanu zawiera zmienne prądu  $i$  płynącego przez zwoje silnika, oraz prędkości obrotowej  $\omega$  głównego wału silnika:

$$\begin{aligned} \begin{bmatrix} \dot{i} \\ \dot{\omega} \end{bmatrix} &= \begin{bmatrix} -\frac{R_r}{L_r} & -\frac{k_e}{L_r} \\ \frac{k_m}{J} & -\frac{B_v}{J} \end{bmatrix} \cdot \begin{bmatrix} i \\ \omega \end{bmatrix} + \begin{bmatrix} \frac{1}{L_r} \\ 0 \end{bmatrix} \cdot V \\ y &= \begin{bmatrix} 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} i \\ \omega \end{bmatrix} + \begin{bmatrix} 0 \end{bmatrix} \end{aligned} \quad (4.11)$$

gdzie:

$k_m$  – stała mechaniczna

$k_e$  – stała elektryczna

$J$  – moment bezwładności stojana

$B_v$  – współczynnik tarcia

$R_r$  – rezystancja silnika

$L_r$  – indukcyjność silnika.

Model ten w postaci transmitancyjnej można wyrazić jako:

$$G_s = \frac{k_m}{JL_r s^2 + (R_r J + B_v L_r) s + (k_m k_e + R_r B_v)} \quad (4.12)$$

Przedstawiona postać nie jest związana bezpośrednio z przyspieszeniem. Dlatego, przy założeniu obserwowalności systemu, pierwotna postać stanowa przekształcona jest do odpowiedniej postaci kanonicznej, gdzie wektor stanu zawiera wymagane zmienne stanu, czyli prędkość kątową i przyspieszenie kątowe. Konwersja ta wymaga pewnych uproszczeń. Mianowicie, przy zmodyfikowanym modelu zakłada się, że  $R_r B_v \ll 1$ ,  $B_v L_r \ll R_r J$ ,  $L_r \ll \frac{J R_r^2}{k_m k_e}$ , co jest w ogólności słuszne dla silników prądu stałego. Można także przyjąć, że parametry  $k_m$  oraz  $k_e$  mimo różnej interpretacji fizycznej mają takie same wartości ( $k = k_e = k_m$ ), co w przypadku opracowywanego systemu nie jest konieczne (Das *et al.*, 2013; Wolm *et al.*, 2008).

Przy małej induktancji  $L_r$  w porównaniu do wartości bezwładności  $J$  oraz bardzo małym współczynnikowi tarcia  $B_v$ , transmitancję (4.12) można uprościć do użytecznej postaci ( $\underline{\omega}$  i  $\underline{V}$  to odpowiednio prędkość kątowa i wymuszenie w dziedzinie Laplace'a):

$$G_s = \frac{\underline{\omega}(s)}{\underline{V}(s)} = \frac{\frac{1}{k_e}}{\frac{JL_r}{k_m k_e} s^2 + \frac{R_r J}{k_m k_e} s + 1} \quad (4.13)$$

Powyższą postać można interpretować w kontekście stałych czasowych silnika. Definiując odpowiednio wzmocnienie  $K = \frac{1}{k_e}$ , mechaniczną stałą czasową  $T_m = \frac{R_r J + B_v L_r}{k_m k_e} \simeq \frac{R_r J}{k_m k_e}$ , oraz elektryczną stałą czasową  $T_e = \frac{L_r}{R_r}$ , transmitancję modelu zapisać można w uproszczonej postaci

$$G_s = \frac{K}{T_m T_e s^2 + (T_e + T_m) s + 1} \simeq \frac{K}{T_m T_e s^2 + (T_m) s + 1} \quad (4.14)$$

słusznej dla  $T_e \ll T_m$ .

Przedstawiony model silnika DC oraz jego uproszczone wersje stanowią podstawę budowy systemu pomiarowego robota mobilnego dostosowanego do danych pochodzących z akcelerometru.

### Model robota uwzględniający uproszczony model silnika DC

Proponowany opis dynamiki robota oparty jest na uproszczonym modelu drugiego rzędu danym transmitancją (4.13). Ciągłoczasowy model stanowy systemu można przedstawić w następującej postaci:

$$\begin{bmatrix} \dot{\omega} \\ \dot{\epsilon} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k_m k_e}{JL_r} & -\frac{R_r}{L_r} \end{bmatrix} \cdot \begin{bmatrix} \omega \\ \epsilon \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{k_m}{JL_r} \end{bmatrix} \cdot V \quad (4.15)$$
$$y = \begin{bmatrix} 0 & \frac{k_m}{JL_r} \end{bmatrix} \cdot \begin{bmatrix} \omega \\ \epsilon \end{bmatrix} + \begin{bmatrix} 0 \end{bmatrix}$$

Wektor stanu w prezentowanym rozwiązaniu zawiera prędkość kątową  $\omega$  oraz przyspieszenie kątowe  $\epsilon$ , dzięki czemu można powiązać pomiary akcelerometru z przyspieszeniem poprzez macierz wyjść  $\mathbf{C}$ . Oczywiście należy uwzględnić też przełożenie przekładni *ratio* (stosunek prędkości obrotowej głównego wału silnika do prędkości obrotowej koła) i promień koła  $r_w$ . Są to parametry, które umożliwiają przeliczenie wartości kątowych związanych z ruchem głównego wału silnika na wartości liniowe związane z ruchem robota. Wyznaczenie położenia wymaga pojedynczego całkowania prędkości. Położenie  $w_1$  wyznaczone na podstawie drugiej zmiennej stanu (prędkości kątowej  $\omega$ ) w przedziale czasu  $\langle t_0; t_1 \rangle$  ma postać:

$$w_1 = \frac{k_m}{JL_r} \cdot r_w \cdot ratio \cdot \int_{t_0}^{t_1} \omega(t) dt \quad (4.16)$$

Przedstawiony model stanowi ograniczenie dynamiki ruchu robota ze względu na wymuszenie  $V$  oraz przyspieszenie wzdłużne  $a_l$  oparte na danych z akcelerometru. Dane przyspieszenia są korygowane zgodnie z aktualnym ułożeniem  $(\phi, \theta, \psi)$  w celu usunięcia wpływu grawitacji. Przyspieszenie poprzeczne  $a_t$  związane z ruchem pionowym robota nie jest istotne dla robota poruszającego się po powierzchni. Przyspieszenie boczne  $a_s$  może pojawić się przy (dynamicznym) ruchu robota po łuku. Aby uniknąć niepotrzebnego komplikowania modelu dla potrzeb mapowania założono, że robot wykonuje obroty w miejscu.

### Filtracja Kalmana z wykorzystaniem modelu robota

Opracowany model robota jest podstawą do określania położenia robota oraz kamer zainstalowanych na jego platformie. Estymacja położenia zakłada użycie filtru Kalmana. Należy zaznaczyć, że w celu uproszczenia eksperymentu użyto jednowymiarowego sygnału sterującego.

W procesie estymacji położenia zastosowano filtrację Kalmana do wyznaczenia drogi w jednowymiarowej przestrzeni, jaką przejechał robot w oparciu o wiedzę o przyspieszeniu wzdłużnym  $a_l$ . Estymacja położenia w trzech wymiarach nie jest realizowana (amplituda przyspieszenia bocznego  $a_s$  jest zbyt mała, przez co dane nie są użyteczne; podobnie jak dane przyspieszenia poprzecznego/pionowego  $a_t$ ). Wektor stanu zawiera tylko prędkość kątową i przyspieszenie kątowe. Droga jest wyznaczona przez całkowanie estymowanej prędkości (4.16) z uwzględnieniem promienia koła  $r_w$ . Trajektoria robota określana jest

Tablica 4.2 Skrócony zapis rozkazów sterowania w systemie VisRobot przy 4 bazach.

$V_{pwm} \langle 0; 100 \rangle$	0	0	100	0	0	0	100	0	0	0	100	0	0
obrót ( $^{\circ}$ )	0	90	0	-90	0	90	0	-90	0	90	0	-90	0

na podstawie estymowanego położenia i ułożenia. W analizowanym systemie VisRobot przebyta droga pełni kluczową funkcję. To, że robot wykonuje obrót w miejscu (zmienia kierunek jazdy) i porusza się do przodu, pozwala na prostą estymację trajektorii ruchu polegającą na poruszaniu się w prawo względem mapowanego obszaru, w celu osiągnięcia nowego położenia determinującego szerokość bazy. Wektor sterowań rozumiany jest jako skończona sekwencja wartości (skalarne) sygnału sterującego, co schematycznie przedstawiono w tab. 4.2, gdzie  $V_{pwm}$  jest sygnałem PWM z zakresu  $\langle 0; 100 \rangle$ . Akwizycja danych z kamer i ich przetwarzanie realizowane jest po zatrzymaniu robota, co opisuje zestaw zerowych współrzędnych wyrażających ruch (wymuszenia i obrotu).

W filtracji Kalmana wymagane jest ustawienie wartości odpowiednich parametrów początkowych, w tym parametryzacja modelu szumu procesowego i pomiarowego. Jak wynika z przeprowadzonych eksperymentów oraz analizy szumu sensorów inercyjnych MEMS (sekcja 3.2.3) przyjęto, że błędy pomiarowe mają charakter gaussowski. W wykorzystanym modelu dynamiki robota przyjęto, że szum procesowy charakteryzuje się parametrem podobnym do szumu pomiarowego.

Dokładne wartości parametrów związane z konkretną realizacją fizyczną robota zostaną przedstawione w części doświadczalnej pracy (sekcja 5.2.2). Ponadto, jak wykazano w poprzednich rozdziałach, w celu ewaluacji dokładności mapowania należy w miarę precyzyjnie określić wpływ błędu położenia i ułożenia na błąd mapowania 3D - wpływają one bowiem bezpośrednio na wyznaczone parametry zewnętrzne kolejnych par stereo w procesie mapowania.

#### 4.4 Wykorzystanie zmiennej bazy stereoskopowej

W omawianym systemie problem mapowania opiera się na parach obrazów stereo uzyskanych przy różnych bazach stereo. Zmienną bazę otrzymuje się poprzez przemieszczenie robota o zadaną odległość zgodnie ze schematem przedstawionym na rys. 4.1.

Pojedyncza para obrazów jest przetwarzana przez algorytm stereowizyjny, który wyznacza mapę dysparycji. Teoretyczny opis zagadnienia przedstawiono w sekcjach 3.4.4 i 3.4.5. Parę obrazów należy odpowiednio skalibrować i dokonać rektyfikacji, aby ułatwić analizę porównawczą. Algorytm stereowizyjny szuka obszarów w danym (np. prawym) obrazie, które odpowiadają obszarom w obrazie referencyjnym (np. lewym). Przesunięcie pomiędzy tymi obszarami w kategoriach współrzędnych obrazu jest określane jako dysparycja. W praktyce analiza nawet poprawnie zrektfikowanej pary obrazów może nie być łatwa, gdyż występuje szereg problemów powodujących niejednoznaczności (Szeliski i Scharstein, 2004; Scharstein i Szeliski, 2002), przez co mapy dysparycji oraz głębi nie odzwierciedlają w pełni rzeczywistych odległości.



#### 4.4.1. Wyznaczanie dysparycji

System VisRobot ma na celu przede wszystkim uzyskanie ulepszonej mapy 3D, której rozdzielczość głębia będzie większa w porównaniu z wynikami klasycznej stereowizji. Do wyznaczenia mapy dysparycji wybrano typowy algorytm oparty na dopasowaniu obszarów (ang. *Block-Matching*). Ponieważ proste porównywanie obszarów charakteryzuje się względnie niską skutecznością, implementację należało wzbogacić o szereg dodatkowych procedur, które zmniejszają wrażliwość algorytmu na zmiany oświetlenia (Hirschmuller, 2008). Ze względu na złożoność implementacyjną w prezentowanym rozwiązaniu zastosowano metodę SGBM (ang. *Semi-Global Block Matching*) dostępną w bibliotece OpenCV. Metoda ta charakteryzuje się dobrą skutecznością (Scharstein *et al.*, 2014) i może być wykonywana bardzo szybko, dzięki optymalizacji, na przykład dla technologii OpenCL oraz OpenMP. Po wstępnej ocenie trudności implementacyjnej okazało się, że samodzielna implementacja takiego zoptymalizowanego algorytmu byłaby niezwykle czasochłonna.

Algorytm oparty na dopasowaniu obszarów można podzielić na 4 etapy:

- wyznaczenie funkcji kosztu
- agregacja kosztu
- wyznaczenie dysparycji
- filtracja końcowa.

Funkcja kosztu jest obliczana w celu oceny podobieństwa pomiędzy obszarem wokół danego punktu lewego obrazu a odpowiednim obszarem wokół punktu prawego obrazu. Obszar jest kwadratem, którego rozmiar jest zależny od nastaw programu (domyślna długość boku jest równa 9 pikseli). Dla wybranych obszarów obliczona jest suma bezwzględnych różnic:

$$\sum_{u,v} |I_L(x+u, y+v) - I_R(x+u+d, y+v)| \quad (4.17)$$

gdzie  $I_L$  i  $I_R$  to luminancja obrazu lewego i prawego dla piksela o określonych współrzędnych  $(x, y)$ , parametry  $u$  i  $v$  określają zakres badanego obszaru wokół tego punktu, natomiast  $d$  to potencjalna dysparycja. Zakres analizy potencjalnych dysparycji  $d_{max_{alg}}$  jest ograniczony typowo do 160 pikseli, aby zapewnić szybkie działanie algorytmu.

Koszt wyznacza się dla konkretnej ograniczonej wartości dysparycji  $d$ . Natomiast agregacja realizowana jest poprzez sumowanie kosztu w oknie analizy określonym parametrami  $(u, v)$ . Minimalna wartość zagregowanego kosztu dla konkretnego analizowanego obszaru determinuje dysparycję. Po wyznaczeniu najmniejszego zagregowanego kosztu dla wszystkich analizowanych obszarów (w lewym obrazie) otrzymujemy pierwotną (początkową) mapę dysparycji.

Pierwotna mapa dysparycji zawiera zwykle liczne błędy, które mają bardzo różnorodny podłoże (Ma *et al.*, 2006; Stefano *et al.*, 2004). Do najczęściej spotykanych problemów, które negatywnie wpływają na wiarygodność mapy dysparycji można zaliczyć:

- zniekształcenia fotometryczne i perspektywiczne
- powierzchnie odbijające światło
- obszary gładkie
- powtarzające się wzorce
- okluzję
- nieciągłość.

Są to problemy trudne, a szczegółowy ich opis oraz eliminacja zdecydowanie wykraczają poza zakres niniejszej pracy. O rozmiarze problemu może świadczyć liczba algorytmów wyznaczania dysparycji w serwisie Middelbury (Scharstein *et al.*, 2014).

Wobec powyższego wymagana filtracja jest zwykle bardzo złożona i realizowana poprzez rozmaite opcjonalne algorytmy. W naszym rozwiązaniu implementacja zawiera:

- dodatkową funkcję kosztu opartą na rozkładzie prawdopodobieństwa luminancji
- wygładzenie gaussowskie rozkładów prawdopodobieństwa
- poszerzenie okna analizy o 5 kierunków (uodpornienie na niewielką rotację obrazu)
- wyznaczenie wstępnej mapy dysparycji dla prawego obrazu
- uspoźnienie dwóch wstępnych map dysparycji
- funkcję oceny okluzji.

Powyższa naszkicowana procedura jest podstawą do wyznaczenia mapy dysparycji oraz głębi dla pojedynczej pary obrazów. Problem fuzji map dysparycji dla wielu par obrazów wymaga określenia, w jakim stopniu wiarygodna jest dana mapa dysparycji, a także złożenia tych map w oparciu o pewne ujęcie ich wiarygodności.

W dalszej części pracy algorytm, który wyznacza mapę dysparycji i głębi będzie określany jako algorytm SGBM, stereowizyjny bądź stereo.

#### 4.4.2. Błąd mapy dysparycji

Zmienna baza stereowizyjna modyfikuje zakres dostępnych odległości mierzony przez opracowany algorytm stereowizyjny. Szerokość bazy  $B$  wpływa liniowo na odległość  $z$ , jednak po uwzględnieniu dyskretnego charakteru dysparycji  $d$  oraz zależności wykładniczej między dysparycją ( $d$ ) a odległością ( $z$ ), wpływ zmiennej bazy  $B$  na wyznaczone położenie punktu  $(x, y, z)$  w przestrzeni wymaga dalszej analizy.

Oznaczając początkową bazę stereowizyjną jako  $B_0$ , równanie (3.39) można zapisać w nowej postaci:

$$\Delta z = z_{i+1} - z_i = B_0 f \cdot (1/d_2 - 1/d_1) \quad (4.18)$$

gdzie błąd  $\Delta z$  jest różnicą dwóch odległości  $z_{i+1}$  i  $z_i$ .

Przy założeniu, że wartości dysparycji są sąsiadujące i spełniają warunek  $d_{i+1} - d_i = 1$ , to błąd oceny odległości można zapisać w postaci następującego równania:

$$\Delta z = B_0 f \cdot (1/d_{i+1} - 1/d_i) = \frac{B_0 f}{d_{i+1} d_i} \quad (4.19)$$

W analizowanym przypadku błąd  $\Delta z$  zależy od odwrotności iloczynu sąsiednich dysparycji, w szczególności dla dysparycji  $d_{i+1} = 1$  i  $d_i = 0$  błąd jest nieskończony. Dlatego szerokość dobieranych baz powinna prowadzić do możliwie dużych wartości dysparycji.

Powyższy błąd uwzględnia tylko ograniczenia związane z dyskretyzacją dysparycji i dotyczy kierunku  $z$ . Aby określić wpływ zmiennej bazy na ogólną dokładność mapowania, należy uwzględnić błąd we wszystkich kierunkach ( $x$ ,  $y$  i  $z$ ) dotyczący położenia i ułożenia kamer oraz błąd metody związany z dyskretyzacją dysparycji. Błąd położenia i ułożenia wynika z problemu niedokładności pomiaru opisanego w 4.3 i 4.2. Błąd położenia rozumiany jest jako fizyczne przesunięcie kamery o pewną odległość (w kierunkach  $x$ ,  $y$  i  $z$ ), co można zapisać w postaci  $\Delta t_x$ ,  $\Delta t_y$ ,  $\Delta t_z$ , natomiast błąd rotacji można zapisać poprzez zmianę ułożenia kamery względem osi  $x$ ,  $y$  i  $z$  w postaci  $\Delta \phi$ ,  $\Delta \theta$ ,  $\Delta \psi$ . Warto dodać, że w stereowizji błąd położenia i ułożenia związany z kierunkiem  $x$ , wpływa także

na szerokości bazy  $B$ . Dla dalszej analizy błędu należy odpowiednio opisać zależności między punktem na obrazie a punktem w przestrzeni dla pojedynczej kamery oraz dla pary kamer.

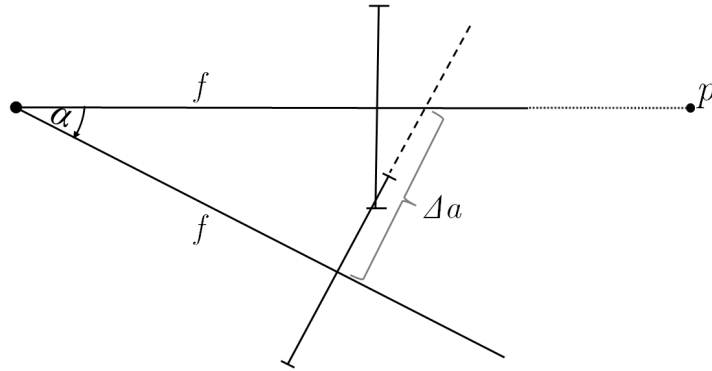
Z analizy modelu otworkowego kamery wynika, że punkt w przestrzeni oddalony o odległość  $h$  od osi optycznej, po uprzednim przeskalowaniu względem odległości od kamery  $z$  odpowiada punktowi na obrazie oddalonemu od punktu głównego o  $h'$  przeskalowanemu względem ogniskowej  $f$ . W związku z tym określenie, jak błąd położenia w przestrzeni wpływa na położenie na obrazie kamery jest proporcjonalne do współczynnika  $f/z$ . Natomiast dla pary kamer znana jest zależność (3.39), w której wyrażenie  $f/z$  jest proporcjonalne do stosunku dysparycji  $d$  i szerokości bazy  $B$ . Łatwo zauważyć analogię do pojedynczej kamery, gdyż szerokość bazy  $B$  związana z błędem położenia jest skalowana względem odległości od kamer  $z$ , co odpowiada różnicy w położeniach danego punktu (dysparycji  $d$ ) przeskalowanej względem ogniskowej  $f$ . Zatem błąd związany z szerokością bazy wpływa na zależności między punktami na ekranie także w sposób proporcjonalny do współczynnika  $f/z$ . Wykorzystując powyższy efekt skalowania można w prosty sposób wyznaczyć, jakie będą zależności geometryczne na obrazie na podstawie zależności geometrycznych w przestrzeni i odwrotnie. Ostatecznie błąd można przedstawić we współrzędnych obrazu oraz wyznaczyć dysparycję  $d$ , a następnie wyznaczyć błąd mapowania w oparciu o równanie (3.42).

Błąd związany z przesunięciem w kierunku  $x$  bezpośrednio wpływa na szerokość bazy  $B$ . Warto zaznaczyć, że baza  $B$  zdefiniowana jest jako odległość między kamerami w równoległym systemie kamer. Błędy położenia i ułożenia kamery w ogólności skutkują niespełnieniem warunku równoległości systemu kamer. Dlatego założono, że przesunięcia w kierunkach  $y$  oraz  $z$  nie wpływają na szerokość bazy  $B$ , a przede wszystkim przyczyniają się one do niewyznaczenia dysparycji dla pewnych obszarów obrazu. Dzieje się tak dlatego, że taka para obrazów nie jest zrektyfikowana, a problem jest tym większy, im bliżej układu kamer znajduje się dany obiekt. Ponadto przesunięcie w kierunku  $z$  powoduje błąd mapowania, który rośnie nieliniowo wraz z odległością od linii głównej danej kamery. Natomiast przesunięcie w kierunkach  $x$  i  $y$  generuje błąd, który jest stały dla całej płaszczyzny obrazu.

Błędy ułożenia są związane z rotacją kamery względem środka rzutowania kamery. W tym przypadku baza  $B$  pozostaje niezmienna, ale położenie punktów 3D ulega zmianie na skutek zmiany dysparycji  $d$ . W szczególności rotacja względem osi pionowej  $y$  wpływa na głębie, natomiast pozostałe obroty mogą przyczynić się do niewyznaczenia dysparycji.

Z powyższych rozważań wynika, że błąd mapowania jest zależny od wielu czynników, w tym od wyniku algorytmu wyznaczania dysparycji. Zatem dokładne wyznaczanie błędu mapowania na drodze analitycznej jest trudne i warunkowane znalezieniem przez algorytm punktów będących rzutami punktu 3D na płaszczyzny obrazów. Niemniej można w pewnym uproszczeniu określić wpływ błędu położenia i ułożenia na mapowanie.

Na rys. 4.5 można zaobserwować, jak obrót kamery o zadany kąt względem osi  $x$  lub  $y$  wpływa na przesunięcie współrzędnych obrazu. Dany punkt  $p$  w przestrzeni jest widoczny na płaszczyźnie kamery o ogniskowej  $f$  w jej punkcie głównym. Wskutek obrotu kamery o kąt  $\alpha$ , rzut punktu  $p$  przesunie się o odległość  $\Delta a$  na płaszczyźnie obrazu kamery. Na tej



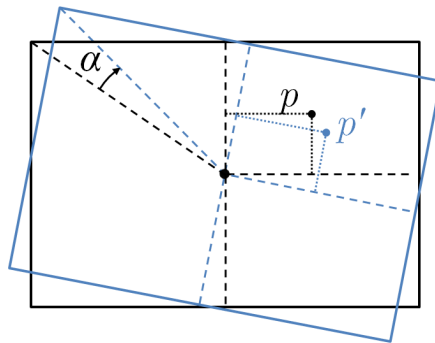
Rysunek 4.5 Błąd związany z obrotem kamery względem osi  $x$  oraz  $y$

podstawie można zapisać, że  $\Delta a = f \operatorname{tg} \alpha$  więc dla obrotów względem osi  $x$  i  $y$  mamy:

$$\begin{aligned} \Delta x &= f \operatorname{tg} \Delta \phi \\ \Delta y &= f \operatorname{tg} \Delta \theta \end{aligned} \quad (4.20)$$

gdzie  $\Delta x$  i  $\Delta y$  są przesunięciami na płaszczyźnie kamery w kierunku odpowiednio  $x$  i  $y$  na skutek obrotów  $\Delta \phi$  względem osi  $x$  i  $\Delta \theta$  względem osi  $y$ .

Obrót względem osi  $z$  przedstawiono na rys. 4.6. Przy obrocie kamery o kąt  $\alpha$ , rzut



Rysunek 4.6 Błąd związany z obrotem kamery względem osi  $z$ .

punktu  $p$  zostanie przesunięty zgodnie z macierzą rotacji (3.5) względem punktu głównego kamery. Rzut punktu  $p$  na płaszczyznę  $xy$  znajduje się w stałym miejscu przed obrotem i po obrocie. Punkt  $p'$  reprezentuje na obróconym obrazie współrzędne płaszczyzny kamery, w których przed obrotem znajdował się punktu  $p$ . Zatem we współrzędnych kamery punkty  $p$  oraz  $p'$  mają takie same współrzędne. Jeśli założymy, że początkiem układu płaszczyzny kamery jest jej punkt główny widoczny na środku, to wskutek obrotu o mały kąt  $\alpha$ , punkt główny nie zmieni się. Współrzędne poziome, a zatem także dysparycja  $d$ , zmniejszą się w II i III oraz zwiększą się w I i IV ćwiartce. Natomiast współrzędne pionowe zmniejszą się w I i II oraz zwiększą się w III i IV ćwiartce układu. Obrót spowoduje zmianę współrzędnych poziomych oraz dysparycji  $d$  o  $\pm x \cos \alpha$ , a także



przesunięcie w pionie o  $\pm x \sin\alpha$ , co może przyczynić się do niewyznaczenia dysparycji. W opisanym przypadku można stwierdzić, że z perspektywy mapowania błąd związany z obrotem kamery względem osi  $z$  ma postać:

$$\Delta x = x \sin\Delta\psi \quad (4.21)$$

gdzie  $x$  jest współrzędną poziomą danego punktu na płaszczyźnie kamery,  $\Delta\psi$  jest obrotem względem osi  $z$ .

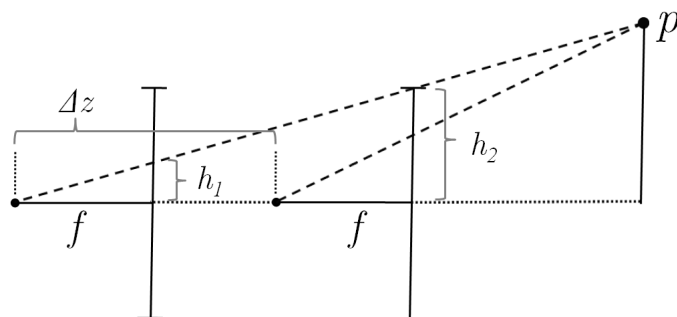
Podobnie błąd związany z dodatkowym przesunięciem można w pewnym uproszczeniu opisać analitycznie. Przesunięcie w kierunku  $x$  nie zmienia warunku równoległości kamer, więc przesunięcie interpretuje się jako zmianę szerokości bazy  $B$  o błędzie  $\Delta t_x$ :

$$B_m = B_{org} + \Delta t_x \quad (4.22)$$

gdzie  $B_{org}$  jest prawdziwą szerokością bazy, zaś  $B_m$  to mierzona szerokość bazy.

W przypadku przesunięcia w kierunku  $y$  założono, że szerokość bazy  $B$  oraz dysparycja  $d$  nie ulegają zmianie. Natomiast przesunięcie przyczynia się do niewyznaczenia dysparycji przez algorytm SGBM, analogicznie jak w przypadku zmian współrzędnych pionowych przy obrocie względem osi  $z$ .

Przesunięcie w kierunku  $z$  powoduje efekt podobny do przybliżania (ang. *zoom in*), gdzie punkty przesuwają się w kierunku zewnętrznych krawędzi kadru. Skutkuje to w ogólności przesunięciem punktów na płaszczyźnie kamery do zewnętrznych obszarów kadru. Wpływ opisywanego przesunięcia na położenie rzutu danego punktu 3D na płaszczyznę kamery przedstawiono na rys. 4.7. Punkt  $p$  jest rzutowany na płaszczyznę kamery



Rysunek 4.7 Błąd związany z przesunięciem kamery w kierunku  $z$ .

o ogniskowej  $f$ . Przed przesunięciem kamery rzut punktu  $p$  na jej płaszczyznę znajduje się w odległości  $h_1$  od linii głównej kamery. Na skutek przesunięcia kamery w kierunku  $z$  (w prawo) rzut punktu  $p$  na płaszczyznę kamery oddala się od linii głównej kamery o odległość  $h_2$ . Z perspektywy mapowania istotna jest zmiana odległości  $h_2$  i  $h_1$ . Korzystając z podobieństwa trójkątów można zapisać, że

$$\frac{h_2}{h_1} = \frac{\Delta t_z + f}{f} \quad (4.23)$$

zaś po obliczeniu  $h_2$  wyznaczyć można

$$\Delta h = h_2 - h_1 = h_1 \left( \frac{\Delta t_z + f}{f} - 1 \right) = h_1 \frac{\Delta t_z}{f} \quad (4.24)$$

Parametr  $\Delta h$  dotyczy dowolnego punktu na płaszczyźnie kamery. Zakładając, że punkt główny kamery jest początkiem układu współrzędnych płaszczyzny kamery to przesunięcie w kierunku  $z$  spowoduje przesunięcie wszystkich punktów na płaszczyźnie kamery, które są równo odległe od punktu  $(0, 0)$ . Dowolne przesunięcie na płaszczyźnie kamery można zapisać jako superpozycję przesunięcia we współrzędnych poziomych i pionowych. Dlatego też zależność (4.24) można zapisać oddzielnie dla współrzędnych poziomych i pionowych danego punktu w płaszczyźnie kamery. Jest to istotne, gdyż przesunięcie współrzędnych poziomych obrazu powoduje zmianę dysparycji, zaś przesunięcie współrzędnych pionowych przyczynia się do niewyznaczenia dysparycji.

Po uwzględnieniu opisanych zależności związanych z wpływem błędów translacji i rotacji kamery na dokładność mapowania podać można uproszczony całkowity błąd, który jest spowodowany błędem dysparycji  $\Delta d$  oraz błędem szerokości bazy  $\Delta B$ . Przy założeniu, że algorytm SGBM wyznaczył dysparycję, oraz że dyskretyzacja dysparycji daje błąd opisany równaniem (4.19), uzyskujemy

$$\begin{aligned}\Delta d &= f \cdot \operatorname{tg} \Delta \phi + x \left( \sin \Delta \psi + \frac{\Delta t_z}{f} \right) \\ \Delta B &= \Delta t_x\end{aligned}\tag{4.25}$$

gdzie  $x$  jest współrzędną poziomą płaszczyzny kamery,  $f$  jest jej ogniskową,  $\Delta \phi$  i  $\Delta \psi$  to błąd ułożenia względem osi odpowiednio  $x$  i  $z$ , zaś  $\Delta t_x$  i  $\Delta t_z$  to błąd położenia odpowiednio wzdłuż osi  $x$  i  $z$ . Błąd ułożenia i położenia względem osi  $y$  może powodować niewyznaczenie głębi, natomiast nie wpływa bezpośrednio na jej dokładność. Ponadto można przyjąć, że błąd dysparycji jest nie większy niż

$$\Delta d_{max} = f \cdot \operatorname{tg} \Delta \phi + \frac{1}{2} \rho_x / \left( \cos \Delta \psi + \frac{\Delta t_z}{f} \right)\tag{4.26}$$

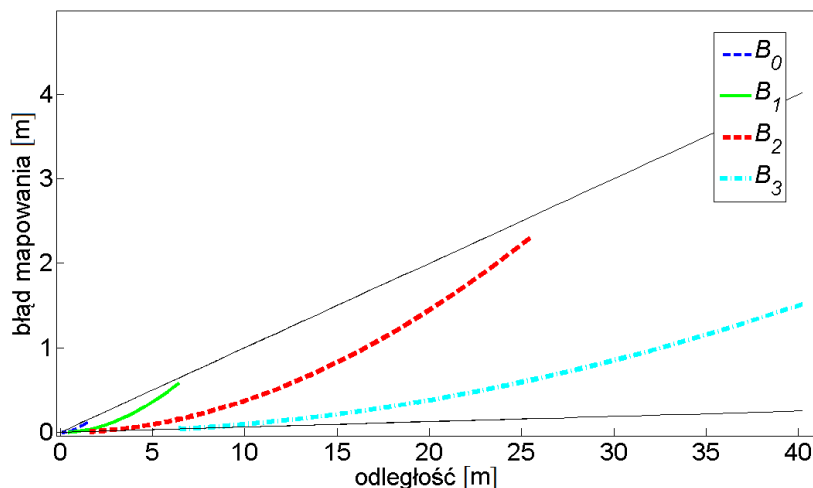
gdzie  $\rho_x$  jest rozdzielczością poziomą kamery, zaś  $\rho_x/2$  odpowiada punktowi płaszczyzny kamery, który jest najbardziej oddalony od punktu głównego. Błąd (4.26) jest to maksymalny możliwy błąd, który pojawia się tylko dla punktów płaszczyzny kamery znajdujących się na krawędzi kadru. Dla pozostałych punktów błąd jest mniejszy, stosownie do wartości  $x$  w równaniu (4.25).

Z powyższej analizy wynika, że błąd położenia i ułożenia ma istotny wpływ na dysparycję oraz jest zależny od współrzędnych obrazu. Uwzględniając zmienność bazy w systemie VisRobot należy także sprawdzić, jak błąd wpływa na mapowanie przy wybranych szerokościach baz, aby dobrać bazę możliwie najlepiej do danego zakresu głębi.

#### 4.4.3. Fuzja map dysparycji i ich wiarygodność

Opracowany system ze zmienną bazą zakłada, że pierwsza mapa dysparycji jest wyznaczona przy ustalonej szerokości bazy. Następne mapy dysparycji uzyskane przy szerszych bazach stereo mają na celu poprawienie jakości pierwotnej mapy. Aby określić, które obszary w dostępnych mapach dysparycji są najdokładniejsze, należy wyznaczyć ich wiarygodność na podstawie błędu (4.25). Założono przy tym, że wiarygodność jest odwrotnie proporcjonalna do funkcji błędu (Kowalczyk i Merta, 2016b). Zatem funkcja wiarygodności jest odwróconą i znormalizowaną funkcją błędu. Aby ocenić zależności między wiarygodnością map dysparycji dla różnych baz, posłużyć się można przykładowymi

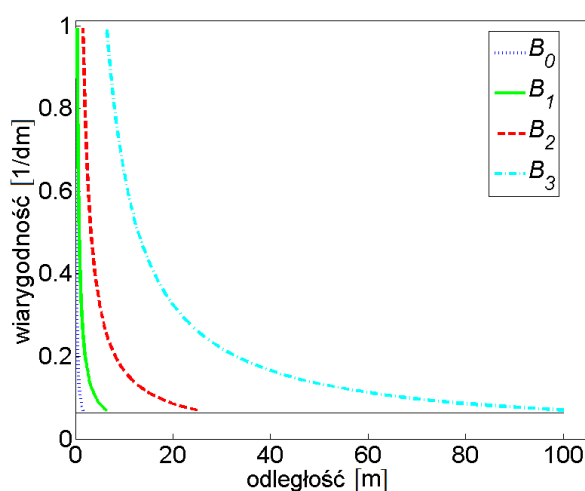




Rysunek 4.8 Błąd głębi dla różnych baz stereo.

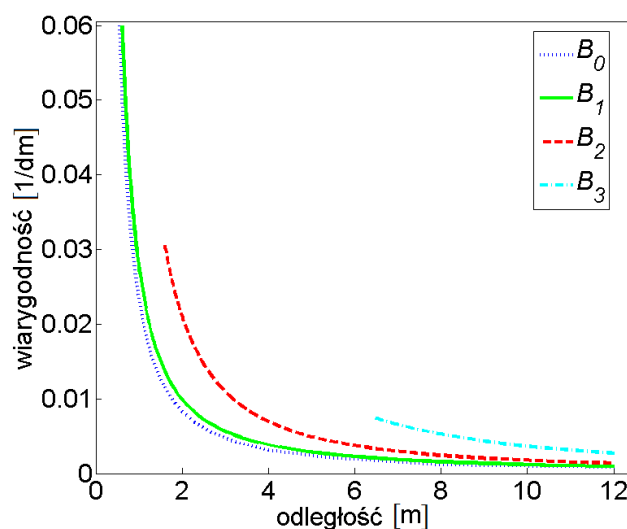
danymi. Ze względu na istnienie błędów związanych z dyskretyzacją dysparycji (4.18) na rys. 4.8 przedstawiono zależność pomiędzy głębią a tym błędem dla 4 różnych baz stereowizyjnych:  $B_1 = 0.01m$  (niebieski),  $B_2 = 0.04m$  (zielony),  $B_3 = 0.16m$  (czerwony),  $B_4 = 0.64m$  (turkusowy).

Na wykresie można zaobserwować, że wraz ze wzrostem szerokości bazy, zwiększa się zakres mierzonych odległości. Ponadto dla zakresu, który jest dostępny dla kilku baz, najmniejszy błąd występuje dla najszerszej bazy. Przy braku dodatkowych błędów związanych z parametrami zewnętrznymi kamery, na rys. 4.9 zobrazowano pierwotną funkcję wiarygodności, która uwzględnia jedynie błąd dyskretyzacji dysparycji.



Rysunek 4.9 Teoretyczny przebieg funkcji wiarygodności w zależności od odległości.

Ze względu na rodzaj pomiaru położenia i ułożenia kamery przy różnych bazach, wymagane jest określenie, jak błąd położenia zmienia się wraz ze wzrostem odległości. Ogólnie można wyróżnić systemy, które zapewniają stały błąd niezależnie od odległości oraz systemy, w których błąd rośnie wraz odległością w sposób liniowy bądź nieliniowy. W omawianej aplikacji błąd ułożenia powinien być stały, co jest związane z dokładnością IMU. Natomiast błąd położenia może rosnać wraz odległością. Dlatego też rozważa się 2 przypadki. Na rys. 4.10 - 4.11 zobrazowano funkcję wiarygodności dla 4 baz w przypadku błędu stałego (dla czterech baz) i kumulującego się (dla baz  $B_1$ ,  $B_2$ ,  $B_3$ ). Baza  $B_0$  w obu przypadkach jest tworzona w oparciu o parę kamer zainstalowanych na platformie robota, co daje stały błąd położenia i ułożenia kamer, który wynika z dokładności kalibracji kamer (Bier i Luchowski, 2009; Nedevschi *et al.*, 2002; Min *et al.*, 2007; Hansen *et al.*, 2012).

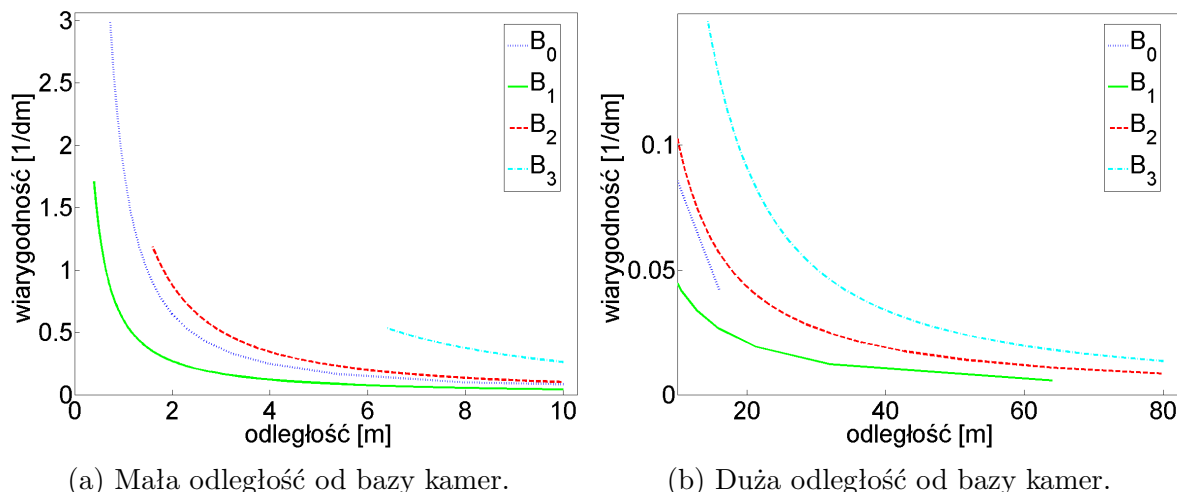


Rysunek 4.10 Funkcja wiarygodności z uwzględnieniem stałego błędu położenia i ułożenia.

Porównując rys. 4.10 i rys. 4.11 zauważyć można, że błąd położenia i ułożenia skorelowany z większymi odległościami wpływa negatywnie na wiarygodność pomiaru. Jednak dla każdej szerszej bazy zwiększa się zakres odległości. Ponadto dla szerszej bazy przy tej samej odległości od kamery wiarygodność jest wyższa. Wyjątkiem jest przypadek baz  $B_0$  oraz  $B_1$ , ponieważ baza  $B_0$  dla mniejszego zakresu odległości przewyższa  $B_1$  dzięki stałemu błędowi położenia (związanemu z kalibracją kamer). Zatem dla tak dobranych szerokości baz wyniki związane z bazą  $B_1$  można pominąć, ponieważ nie zwiększają one wiarygodność mapowania, a zakres odległości mapowany przy bazie  $B_1$  jest pokryty przez bazy  $B_0$  i  $B_2$ .

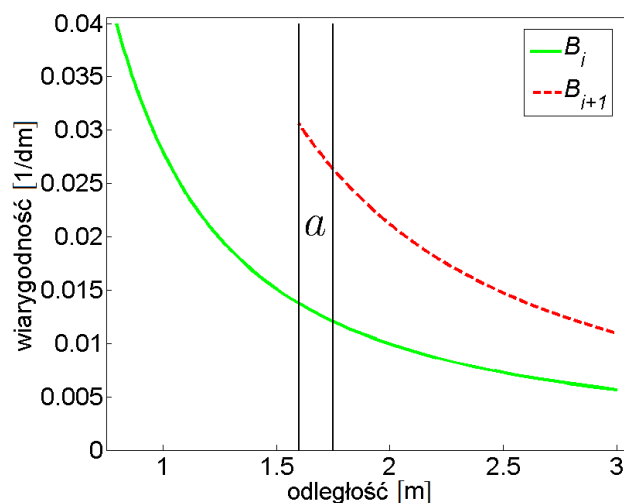
W związku z tym podczas tworzenia mapy, dla danego zakresu odległości, należy wybrać pomiar związany z bazą, która ma największą wiarygodność. Powyższe obserwacje dowodzą, że przy opisanych założeniach system zmiennej bazy pozwala na polepszenie procesu mapowania opartego na stereowizji.

W rozwiązaniu VisRobot występuje problem tworzenia mapy ze składowych map uzyskanych przy różnych bazach. Konsekwencją jest możliwe występowanie nieciągłości na krańcach map. Poprawna mapa powinna cechować się nie tylko założoną dokładnością,



Rysunek 4.11 Funkcja wiarygodności uwzględniająca stały błąd ułożenia oraz zmienny błąd położenia baz  $B_1$ - $B_3$ .

ale także spójnością. Dlatego też zastosowano tu uśrednianie obszarów blisko granicy nieciągłości. Obszar uśredniania jest zależny od parametru  $\eta$  określającego zakres odległości poddawany uśrednianiu. Jest on wspólny dla dwóch sąsiednich baz, co zaznaczono na rys. 4.12 w postaci przykładowego obszaru  $a$ .



Rysunek 4.12 Obszar uśredniania głębi na granicach zakresu sąsiadujących baz.

Przy fuzji danych głębi pojawia się kilka problemów technicznych. Każdą mapę głębi można odwzorować w chmurę punktów 3D. Chmury takie można agregować (w chmurę o większym zakresie odległości). Ponieważ błąd położenia robota przekłada się na błąd szerokości bazy, pojawia się niepewność (błąd) odwzorowania głębi (w kierunku  $z$ ). Błąd bazy propaguje się na pozostałe kierunki  $x$  i  $y$  poprzez prawy, dolny element macierzy  $Q$  (3.43). Błąd ten zatem powoduje nie tylko przesunięcie i rotację chmury punktów, ale także efektywne skalowanie we wszystkich trzech kierunkach.

Zauważmy, że łączenie przeskalowanych chmur punktów jest bardzo uciążliwe zwłaszcza, że dostępne metody kodowania (rejestracji) chmur punktów uwzględniają tylko<sup>5</sup> rotację i przesunięcie. Dlatego postanowiono dokonać rejestracji jeszcze na poziomie dysparycji, dzięki czemu niepewność głębi (na tym etapie) nie wpływa błędy w kierunkach  $x$  i  $y$ .

W celu połączenia map dysparycji, wymagane jest unormowanie każdej mapy względem największej bazy, tj. przeskalowanie każdej wartości (piksel) mapy dysparycji przez  $B_N/B_i$ , gdzie  $B_i$  to szerokość analizowanej bazy, a  $B_N$  to szerokość bazy największej. Konsekwencją takiego skalowania są mapy dysparycji o takiej samej skali jak mapa dysparycji związana z najszerszą bazą. Następnie, dla każdego elementu wynikowej<sup>6</sup> mapy dysparycji wybierana jest wartość o największej wiarygodności (jak to opisano powyżej). Często w wynikowej mapie dysparycji wyizolować można odrębne obszary związane z konkretnymi bazami. W przypadku nachodzenia na siebie obszarów (w 2D) przypisanych dwóm bazom, wystąpić może nieciągłość wgłębna (w trzecim wymiarze) tj. skoki w dysparycji. Aby wygładzić obszary graniczne zdefiniowano parametr  $a$  określający (wgłębny) zakres uśredniania, co zaznaczono na rys. 4.12. W tym zakresie wynikowa dysparycja jest wartością średnią dysparycji z dwóch sąsiednich baz (na rys. 4.12 zaznaczono je symbolami  $B_i$  i  $B_{i+1}$ ).

Jeżeli na jednej z pierwotnych map dysparycji dany punkt nie ma odpowiednika na drugiej mapie, to wynikowa mapa zawiera tylko istniejący punkt. Wynikowa mapa dysparycji uwzględniająca pomiary przy wszystkich założonych szerokościach baz jest podstawą do wyznaczenia współrzędnych 3D w oparciu o macierz  $Q$  (3.43) i równanie (3.42).

## 4.5 Wizualizacja i ocena głębi

Według klasycznej definicji, trójwymiarowa mapa cyfrowa jest zbiorem punktów o trzech współrzędnych. Bez odpowiedniego przetworzenia i zobrazowania zwykle taki zbiór jest mało użyteczny. Dla weryfikacji poprawności odwzorowania, analizy otoczenia wybranego punktu czy też obserwacji wybranych obszarów mapy, konieczny jest mechanizm wizualizujący mapę. Zobrazowanie takie realizuje program poprzez generowanie wirtualnej przestrzeni wyświetlanej na płaskim monitorze, ale także odpowiednie wyświetlenie tej przestrzeni w trzech wymiarach.

Implementacja eksperymentalnego systemu VisRobot dotyczy wąskiego rozumienia procesu mapowania, kiedy nacisk kładzie się na proces odwzorowania otoczenia na mapie 3D. W związku z tym rozważana jest jedna mapa utworzona na podstawie map dysparycji uzyskanych przy różnych bazach. W kompletnym systemie nawigacji robota mobilnego istotne jest rozgraniczenie mapy na wewnętrzną i zewnętrzną. Mapa wewnętrzna dotyczy tylko bliskiego otoczenia robota i ma na celu spójne złożenie danych z wielu sensorów nawigacyjnych. Celem mapy zewnętrznej jest wyznaczenie celu podróży oraz ścieżki, po której ma się poruszać robot. W rozwiązaniu VisRobot założono operowanie na jednej ogólnej mapie, która odwzorowuje wybraną część otoczenia. Ponadto rozwiązanie to nie reprezentuje typowego podejścia do SLAM, gdyż fragmenty mapy nie są dodawane w kolejnych krokach algorytmu, ale jeden fragment mapy jest składany z wielu par stereo ze zmienną bazą.

<sup>5</sup>np. algorytm IPC (ang. *Iterative Point Cloud*)

<sup>6</sup>zbiorowej, uzyskanej poprzez nałożenie na siebie wszystkich unormowanych map dysparycji

Wyznaczane punkty 3D przedstawia się w postaci chmury punktów w danym środowisku graficznym (w którym każdy punkt ma określony rozmiar). Aby otrzymać powierzchnie, sąsiednie punkty można łączyć, uzyskując w ten sposób siatkę trójkątów. Obrazy stereowizyjne umożliwiają także wizualizację obrazu bezpośrednio dla użytkownika za pomocą wybranej techniki stereoskopowej. Dzięki temu użytkownik może obserwować obraz (w początkowym etapie przetwarzania). Jedynym wymaganiem jest zapewnienie, aby dana para obrazów stereo była dostosowana do narządu wzroku człowieka.

#### 4.5.1. Chmura punktów

Implementację systemu VisRobot oparto na zewnętrznej bibliotece OpenCV i technologii OpenGL. Autorski program o nazwie VIS pozwala na zarządzanie parametrami wirtualnej kamery, której obraz jest uzyskany poprzez renderowany rzut chmury punktów na płaszczyznę kamery. Ponadto program ten umożliwia zaznaczanie punktów, wyświetlanie informacji dotyczących punktów i odległości między nimi, zmianę rozmiaru punktów, itp. Wizualizacja w tym przypadku jest realizowana na potrzeby eksperymentów w celu ewaluacji systemu.

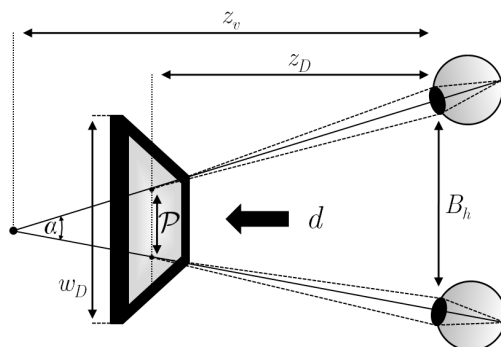
#### 4.5.2. Obraz 3D

Dane z kamer w systemie stereowizyjnym mogą służyć do odwzorowania otoczenia w formie trójwymiarowej. Wówczas otoczenie można przedstawić użytkownikowi, wyświetlając parę obrazów odpowiednio dla lewego i prawego oka, co wykorzystano w systemie VIS. Efektywną wizualizację pary obrazów realizuje się z wykorzystaniem okularów z techniką anaglifową, polaryzacyjną lub migawkową. Alternatywą są wyświetlacze autostereoskopowe, które nie wymagają okularów. Ze względu na specyfikę obrazu generowanego przez oko ludzkie oraz ograniczenia wyświetlaczy 3D, obraz z kamer stereo musi być tak przetworzony, aby zapewnić prawidłową fuzję dwóch obrazów (stereopary) w mózgu obserwatora (Siegel i Nagata, 2000).

System kamery dedykowany do stereowizji w idealnym przypadku jest równoległym układem kamer. Natomiast wzrok człowieka ma parę oczu w układzie wpatrzonym (ang. *Toad-In*). Wzrok zdrowego osobnika skupiony jest w jednym wybranym punkcie (skutkiem czego osie optyczne oczu spotykają się w tym punkcie). W praktyce ludzkiemu operatorowi należy przedstawić obraz (odpowiednio odległy) przed oczami tak, aby wybrana część trafiła do lewego lub prawego oka.

Do poprawnej syntezy obrazu stereoskopowego potrzebny jest opis zależności głębi obrazu od parametrów wyświetlacza i warunków wyświetlania. Dysparycja  $d$  odzwierciedla różnicę między obrazami (na matrycach kamer), zaś paralaksa  $\mathcal{P}$  oddaje różnicę między obrazami na wyświetlaczu 3D oraz występuje, gdy obserwacja obrazu na wyświetlaczu 3D jest realizowana za pomocą układu wpatzonego. Można zatem w pewnych warunkach utożsamiać te dwa pojęcia ze sobą. Okazuje się mianowicie, że jeśli obraz o szerokości  $w_S$  z obu kamer zostanie zrektyfikowany i przeskalowany adekwatnie do wymiarów wyświetlacza o szerokości  $w_D$  to można utożsamiać paralaksę  $\mathcal{P}$  z dysparycją  $d$  zgodnie ze wzorem:

$$\mathcal{P} = \frac{w_D}{w_S} d \quad (4.27)$$



Rysunek 4.13 Postrzeganie głębi w obrazie stereoskopowym.

Ujemna paralaksa ma miejsce, gdy wzrok jest skupiony w punkcie, który znajduje się przed ekranem. Dodatnia paralaksa świadczy o skupieniu wzroku za ekranem. Przy zerowej paralaksie wzrok naturalnie skupia się na ekranie.

Znając odległość obserwatora od ekranu  $z_D$ , szerokość bazy stereo  $B_h$  obserwatora (średnio  $0.064m$ ) oraz paralaksę  $\mathcal{P}$ , możliwa jest rekonstrukcja głębi  $z_v$  jako (4.28).

$$z_v = \frac{z_D B_h}{B_h - \mathcal{P}} \quad (4.28)$$

Powyższe zależności przedstawiono na rys. 4.13, gdzie można dodatkowo zaobserwować konsekwencje niezerowej paralaksy. Dla dowolnego rzeczywistego punktu 3D narząd wzroku dostosowuje ogniskową i kąt patrzenia. Występują tu mechanizmy akomodacji i konwergencji. Akomodacja narządu wzroku w danej chwili ustawia ogniskową stosownie do odległości  $z_D$ , aby obiekt w obrazie był ostry, zaś konwergencja ustawia kąt patrzenia każdego z oczu stosownie do  $z_v$ , aby obiekt ustawiony był na środku obu obrazów stereo. Z powodu konfliktu akomodacji i konwergencji różnica między  $z_D$  i  $z_v$  nie może być duża. Na drodze empirycznej wyznaczono zakres zmian kąta konwergencji  $\alpha$ , który nie zakłóca poprawnego złożenia obrazów stereoskopowych przez człowieka (Jones *et al.*, 2001). Różnica między wartością maksymalną i minimalną wynosi  $\Delta\alpha = 0.02[\text{rad}]$ , co ogranicza zakres paralaksy zgodnie ze wzorem:

$$\Delta\mathcal{P} = z_D \Delta\alpha \quad (4.29)$$

W literaturze dotyczącej zasad tworzenia nagrań 3D istnieje tzw. reguła 1/30, która opiera się na zaprezentowanym parametrze  $\Delta\alpha$  oraz proporcji między szerokością ekranu a odległością ekranu od widza  $z_d/w_D = 1.67$  (Zilly *et al.*, 2011). Zgodnie z regułą 1/30 wartość paralaksy (dodatniej lub ujemnej) nie powinna być większa niż 1/30 szerokości ekranu. Przyjmując maksymalną paralaksę  $\mathcal{P}_{max} = 1/30 w_D$  i uwzględniając równanie (4.27) można wyznaczyć dopuszczalny zakres dysparycji  $d_{max} = w_s/30$ . Istnieją oczywiście inne elementy przyczyniające się do niepoprawnego wyświetlania obrazu stereoskopowego, jednak ich wpływ jest niewielki (Zilly *et al.*, 2011).

Na podstawie reguły 1/30 w systemie VIS założono, że obraz wyświetlany użytkownikowi musi mieć ograniczoną dysparycję. W związku z tym system maskuje obszary o zbyt dużej dysparycji, przez co użytkownik może obserwować tylko te obszary głębi,



które spełniają zależność:

$$z \leq \frac{B}{d_{max}} f \quad (4.30)$$

Zgodnie z powyższymi ograniczeniami w ramach prowadzonych badań opracowano metodę opartą na stereowizji, która odwzorowuje otoczenie nie tylko za pomocą chmury punktów, ale także pozwala na bezpośrednie trójwymiarowe przedstawienie otoczenia ludzkiemu operatorowi.

Lewy i prawy obraz z kamery, po dostosowaniu do wzroku użytkownika, mogą być wyświetlone za pomocą wybranej techniki stereoskopowej. Operator, który nie dysponuje monitorem lub rzutnikiem dostosowanym do techniki polaryzacyjnej lub migawkowej, może użyć tanich okularów anaglifowych. Zaprezentowane rozwiązanie pozwala użytkownikowi na obserwację głębi mapowanej przestrzeni. Co więcej, przy użyciu odpowiednio dużej bazy można otrzymać efekt ‘pogłębienia przestrzennego’ powodujący zwiększenie rozdzielczości głębi (Zilly *et al.*, 2011). Dzięki temu, przy dużych szerokościach bazy obserwator może dostrzec zależności przestrzenne między odległymi obiektami (co wynika z dużej dysparycji). Dla małych baz zależności są niezauważalne, gdyż dysparycja jest bliska zeru.

#### 4.5.3. Ocena jakości głębi obrazu

W ramach postawionej subtezy 2 zdefiniowano zagadnienie optymalnego wyboru bazy obrazów stereoskopowych pod względem jakości głębi obrazu. Taki wybór pozwala ocenić, który pomiar jest optymalny. Ponadto taka optymalizacja umożliwia wyświetlanie ludzkiemu operatorowi pogłębionego obrazu stereoskopowego<sup>7</sup>.

Do zdefiniowania jakości głębi przyjęto, że najlepsza mapa głębi powinna zawierać najwięcej użytecznych pikseli, czyli nieść najwięcej informacji o głębi. Taka sytuacja zajdzie, kiedy suma pikseli o zbyt dużej dysparycji (zakrytych dla operatora) i liczba pikseli o zerowej dysparycji (tych najdalszych) jest najmniejsza. W tym celu z mapy dysparycji wyklucza się wartości przekraczające możliwości postrzegania (zgodnie z regułą 1/30) lub wynikające z ograniczeń algorytmu stereowizyjnego  $d_{max_{alg}}$  (w takim przypadku  $d_{max} \leq d_{max_{alg}}$  opisuje przekroczenie zakresu analizy).

W przypadku spełnienia założonego ograniczenia zakresu dysparycji, maksymalizowane kryterium przyjmuje postać sumy ‘użytecznych’ dysparycji. Funkcję celu można zapisać w postaci

$$F(IMG) = \frac{1}{x_{max}y_{max}} \sum_{x,y \in IMG} d^*(x, y) \quad (4.31)$$

gdzie  $x, y$  są współrzędnymi obrazu IMG (mapy głębi) uzyskanego przy danej bazie,  $x_{max}$  i  $y_{max}$  to odpowiednio rozdzielczość pozioma i pionowa obrazu IMG, zaś  $d^*$  jest efektywną (użyteczną) wartością dysparycji, która ma następującą postać:

$$d^*(x, y) = \begin{cases} 0 & d(x, y) = 0 \\ d(x, y) & 1 \leq d(x, y) \leq d_{max} \\ 0 & d(x, y) > d_{max} \end{cases} \quad (4.32)$$

<sup>7</sup>Umożliwia widzenie głębiej/dalej oczywiście z ograniczeniami.

Użyteczne wartości efektywnej dysparycji są niezerowe (w zakresie  $\langle 1; d_{max} \rangle$ ). Optymalna baza wynika z kryterium optymalizacyjnego:

$$B_{opt} = \arg \max_{i=1 \dots N} F(\text{IMG}_{B_i}) \quad (4.33)$$

gdzie  $i$  jest numerem bazy.

Funkcja celu (4.31) pozwala ilościowo ocenić jakość głębi i zostanie wykorzystana przy ocenie wyników eksperymentalnych.

## 4.6 Podsumowanie

W rozdziale przedstawiono opracowany eksperymentalny system, którego celem jest wykorzystanie robotów mobilnych w celu eksploracji wewnętrznego statycznego środowiska w oparciu o stereowizję ze zmienną bazą. Rozwiązanie to charakteryzuje się innowacyjnym podejściem do wykorzystania zmiennej bazy stereo, gdzie pary obrazów pochodzą z różnych lokalizacji robota. Tworzona mapa 3D jest złożona z wielu map głębi uzyskanych przy różnych bazach. Zobrazowanie otoczenia uwzględnia wyświetlanie chmury punktów oraz odpowiednio przygotowanych obrazów stereo.

Na dokładność odwzorowania silnie wpływa ocena położenia i ułożenia robota (z kamerą). Dlatego poświęcono wiele wysiłku na dokonanie dokładnych pomiarów tych zmiennych. Pomiar położenia i ułożenia oparto na danych inercyjnych, co jest nowatorskim podejściem w kontekście praktycznych metod mapowania i stereowizji. Do określania pozycji robota opracowano własny model robota, co znacznie zwiększyło praktyczność użycia akcelerometru przy wyznaczaniu przebytej drogi.

Analiza błędu dysparycji umożliwiła dowiedzenie na drodze analitycznej, że generalnie możliwa jest poprawa jakości procesu mapowania opartego na stereowizji przy wykorzystaniu zmiennej bazy. Jakość warunkowana jest dokładnością położenia i ułożenia kamer, a także wszystkich elementów, które mają wpływ na położenie i ułożenie kamer. Aby stwierdzić, w jakim stopniu prezentowany system VisRobot pozwala na poprawienie jakości odwzorowania otoczenia, konieczne są odpowiednie badania symulacyjne oraz eksperymenty, będące treścią kolejnego rozdziału.

---

## SYMULACJE I EKSPERYMENTY

W poprzedzających rozdziałach przedstawiono między innymi rozmaite rozwiązania dostępne w literaturze, które umożliwiają realizację zadania mapowania przestrzennego. Ze względu na różnorodność budowy robotów, rodzajów napędu, sensorów, a także mnogość praktycznych problemów występujących w zagadnieniach SLAM, podejście do odwzorowania otoczenia ograniczono do autorskiego rozwiązania VisRobot, które było przedmiotem rozdziału 4. Implementacja tego systemu oznacza realizację metody mapowania otoczenia za pomocą stereowizji ze zmienną bazą. Para obrazów stereo składa się z obrazu uzyskanego w początkowym położeniu robota oraz obrazu uzyskanego po przemieszczeniu robota (determinującym szerokość bazy stereoskopowej).

Do poprawnej pracy system VisRobot wymaga precyzyjnego pomiaru położenia kamery zainstalowanej na robocie, który oparto na dynamicznym modelu robota mobilnego i danych z akcelerometru. Ponadto w prezentowanym rozwiązaniu duże znaczenie ma pomiar ułożenia kamery uzyskany w wyniku fuzji danych inercyjnych. Natomiast stereowizja, która realizuje właściwe mapowanie wymaga poprawnej rektyfikacji obrazu z kamer. Charakteryzuje się ona rozdzielczością głębi, która maleje wraz ze wzrostem odległości od układu (bazy) kamer. Co więcej wyznaczanie głębi wymaga właściwych warunków oświetleniowych i działa tylko, gdy w obrazie obecne są dystynktywne elementy. Zależności dotyczące dokładności odwzorowania otoczenia przez system przedstawiono w sekcji 4.4.3. Zależności te są jednak prawdziwe, tylko jeżeli algorytm realizujący stereowizję znajdzie odpowiadające sobie punkty w lewym i prawym obrazie.

Ewaluacja systemu wymaga sprawdzenia poprawności działania oraz określenia precyzji wszystkich istotnych elementów, które wywierają wpływ na proces mapowania. Na drodze symulacji należy sprawdzić odporność (wrażliwość) systemu na błędy modelu akcelerometru i modelu robota mobilnego, które są kluczowe przy wyznaczaniu położenia. Ponadto na podstawie oceny dokładności wyznaczania punktów 3D w oparciu o stereowizję należy dokonać odpowiednich obliczeń (symulacji) błędu położenia punktu 3D dla konkretnych wartości szerokości baz, błędu położenia i ułożenia kamer. Wyniki takiej symulacji należy skonfrontować z eksperymentami rzeczywistymi w wykorzystaniu platformy (robota kołowego). W szczególności należy wyznaczyć błąd położenia i ułożenia dla zadanej trajektorii oraz błąd mapowania dla wybranych punktów w przestrzeni. Należy zaznaczyć, iż pomiary wybranych punktów 3D z wysoką dokładnością wymagają specjalistycznego certyfikowanego laboratorium, dlatego eksperymenty rzeczywiste będą

zrealizowane z uwzględnieniem najbardziej istotnych ograniczeń w charakterystyce przyrządów pomiarowych.

## 5.1 Analiza wrażliwości danych akcelerometru

Szczególnie istotnym elementem opracowanego systemu jest pomiar położenia robota, który oparto na danych pochodzących z akcelerometru. Motywacją do takiego podejścia jest brak jednoznacznej odpowiedzi, czy akcelerometr jest użyteczny do wyznaczania położenia, co zostało opisane na początku rozdziału 3. W kontekście systemu VisRobot dokładność położenia musi być wystarczająca do wyznaczenia szerokości bazy i mapowania 3D. Aby ocenić użyteczność akcelerometru należy przeprowadzić odpowiednią analizę wrażliwości i definitywnie określić czy (wymagające całkowania) dane z akcelerometru mogą być użyte do wyznaczania położenia w wybranych zastosowaniach. Szczegóły dotyczące badania wrażliwości typowego akcelerometru MEMS można znaleźć w pracy (Kowalczyk i Merta, 2015).

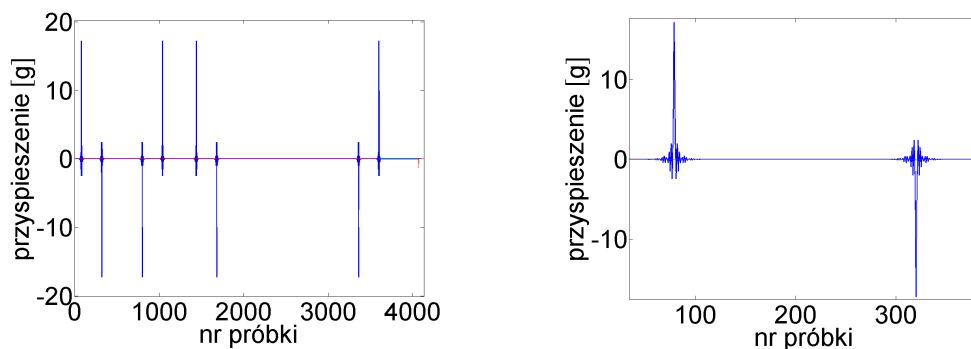
Analiza danych z akcelerometru jest oparta na modelu zaproponowanym w sekcji 4.3.1. Jest to model uwzględniający dynamikę ruchu w układzie odniesienia sensora oraz szum typowy dla stosowanych sensorów MEMS (4.5). Modelowanie z wykorzystaniem szumu wprowadza znaczne uproszczenie cech rzeczywistego szumu, który ogólnie ma bardzo złożony nieliniowy charakter. Niemniej nadal nie ma rozwiązań, które w sposób precyzyjny i użyteczny opisują szum sensorów MEMS i zwykle używa się w tym celu modelu błędzenia losowego (Agrawal *et al.*, 2013; Mohd-Yasin *et al.*, 2009; Wang *et al.*, 1998). Czynniki wolnozmiennne, takie jak dryft temperaturowy, nie są uwzględniane, gdyż czas przeprowadzanych eksperymentów jest stosunkowo krótki.

Wyznaczanie położenia opartego na pomiarze przyspieszenia sprowadza się do podwójnego całkowania danych po uprzednim usunięciu składowej grawitacji. Położenie zależne jest tylko od wolnozmiennnej składowej sygnału przyspieszenia (całkowanie jest operacją liniową).

Ze znanej właściwości całkowania wynika, że sygnał przyspieszenia ulegnie silnej filtracji dolnoprzepustowej. Na skutek filtracji wysokoczęstotliwościowe składowe szumu sensora w znacznej mierze zostaną usunięte. Jednak bardzo niskie częstotliwości oraz składowa stała sygnału pozostaną niezmiennione. Zatem na drodze symulacji należy przeprowadzić ewaluację proponowanego modelu akcelerometru, a także sprawdzić, w jakim stopniu dodatkowa filtracja pasmowa sygnału przyspieszenia może zmniejszyć błąd położenia (wyznaczany przy podwójnym całkowaniu sygnału przyspieszenia).

### 5.1.1. Symulacja ruchu robota i filtracja pasmowa

Pierwszy test zrealizowano dla symulowanego ruchu robota po płaskiej powierzchni. Ruch ten został zobrazowany na rys. 5.1 – 5.3. Właściwy sygnał sterujący silnikami robota jest powiązany z prędkością obrotową silnika, jednak na potrzeby dalszej analizy na rys. 5.1 zobrazowano odpowiadający mu sygnał przyspieszenia (kompletny oraz jego fragment). Sygnał ten wywołuje ruch robota do przodu lub do tyłu (w jego osi  $y$ ). W symulacji założono, że w trakcie ruchu robota jego ułożenie (oraz kamery) także zmienia się, co zobrazowano na rys. 5.2. Natomiast dodatkowe przyspieszenie spowodowane zmianą ułożenia (wpływ grawitacji) można zaobserwować na rys. 5.3.



(a) Pełny sygnał.

(b) Fragment sygnału.

Rysunek 5.1 Przyspieszenie robota względem osi  $y$ .

Ruch robota w symulacji charakteryzuje się dużą dynamiką. Na rys. 5.1a można zauważyć krótkie zmiany przyspieszenia (dodanie i ujemne), które dotyczą gwałtownych ruchów robota do przodu i hamowania oraz ruchu przy ustalonej prędkości. Po krótkim postoju schemat ruchu powtarza się w przeciwnym kierunku, przy czym występują inne przedziały czasowe między momentami przyspieszeń.

Aby poprawnie oddać dynamikę rzeczywistego ruchu, wszystkie sygnały zostały poddane niewielkiej filtracji dolnoprzepustowej o częstotliwości odcięcia 50Hz. Efekt filtracji można zaobserwować na rys. 5.1b, gdzie sygnały obrazujące skoki przyspieszenia mają kształt funkcji *sinc*.

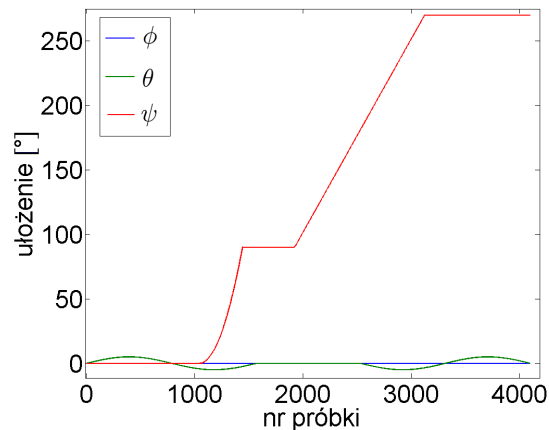
Prosty ruch robota wzbogacono o zmiany ułożenia (rys. 5.2). Ułożenie względem osi  $z$  jest zmienne i na początku ruchu wynosi  $0^\circ$ , zaś na koniec osiąga wartość  $270^\circ$ . Natomiast ułożenie względem osi  $y$  ulega niewielkim sinusoidalnym wahaniom, co symuluje ruch po względnie płaskiej powierzchni. Niezmienne jest ułożenie względem osi  $x$ .

Efektom zmian ułożenia jest dodatkowe niewielkie przyspieszenie robota zobrazowane na rys. 5.3. Warto zaznaczyć, że takiego przyspieszenia nie można pominąć, w kontekście całkowania danych. Zatem należy skupić się na dodatkowym przyspieszeniu związanym z błędami pomiarowymi.

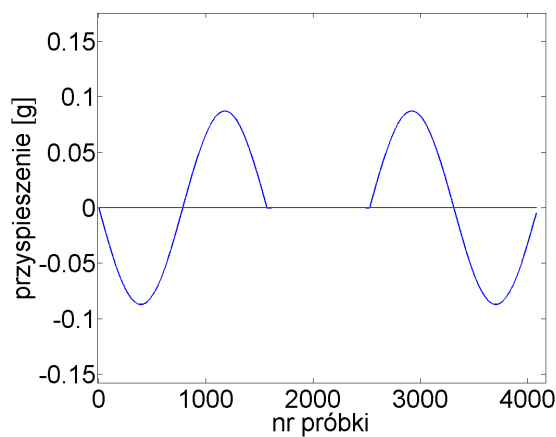
Szum akcelerometru oraz błędy związane z niedokładnością operacji usuwania wpływu przyspieszenia ziemskiego są niezmiernie istotne przy całkowaniu danych pomiarowych. W praktyce położenie wyznaczone przez podwójne całkowanie danych przyspieszenia bez filtracji daje błędy większe nawet niż pokonany dystans. Prosta filtracja dolnoprzepustowa mierzonego przyspieszenia nie doprowadziła do żadnej poprawy. Dlatego w ramach przeprowadzonych testów sprawdzono skuteczność filtracji pasmowej sygnału przyspieszenia poddawanego całkowaniu.

Dla opisanego powyżej symulowanego ruchu sprawdzono i zobrazowano różne przypadki filtracji pasmowej. Dla zapewnienia czytelności wykresów przedstawiono tylko wyniki filtracji dla reprezentatywnego przypadku, gdzie częstotliwość odcięcia  $f_{stop} = 0.5\text{Hz}$ . Warto również zaznaczyć, że ze względu na wysoki rząd filtru konieczne było odpowiednie wydłużenie sygnału wejściowego (wstawienie zer) i usunięcie opóźnienia wprowadzanego przez filtr.

Na rys. 5.4a można zaobserwować oryginalny symulowany sygnał przyspieszenia robota. Po silnej filtracji dolnoprzepustowej na wykresie 5.4b przyspieszenie robota zmieniło



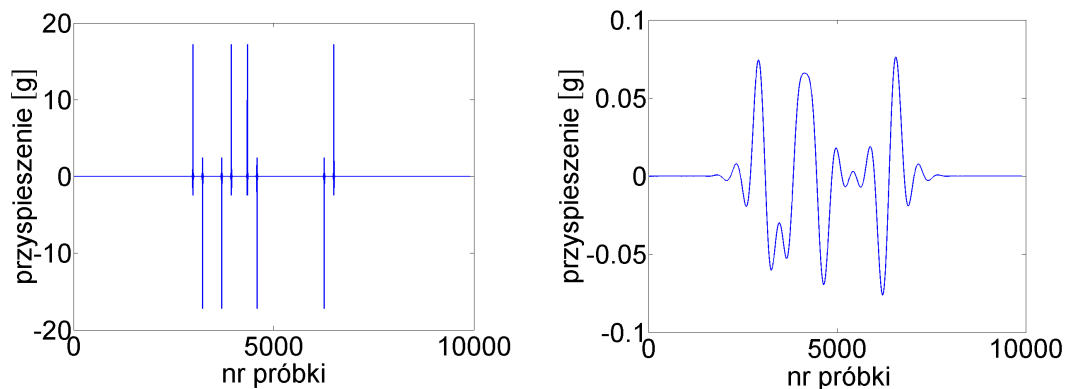
Rysunek 5.2 Ułożenie robota względem osi  $x$  (niebieski),  $y$  (zielony),  $z$  (czerwony).



Rysunek 5.3 Przyspieszenie robota wzdłuż osi  $x$  spowodowane zmianą ułożenia.

się w sposób zauważalny, natomiast widoczne jest wyraźne podobieństwo (pod względem kształtu) do oryginalnego przyspieszenia. Mimo zauważalnej różnicy wynikającej z filtracji sygnału przyspieszenia występuje tylko niewielka różnica w sygnale położenia, będąca skutkiem podwójnego całkowania obu sygnałów, widoczna odpowiednio na rys. 5.5a i rys. 5.5b.

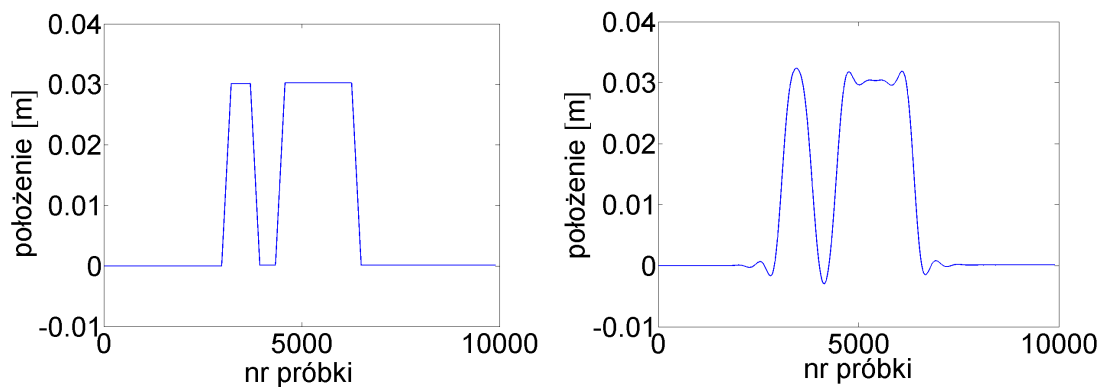
Można zatem przyjąć, że dla sygnału przyspieszenia analizowanego przy częstotliwości próbkowania  $f_s = 400\text{Hz}$ , filtracja dolnoprzepustowa o częstotliwości odcięcia  $f_{stop} = 0.5\text{Hz}$  nie zmienia w sposób istotny oceny położenia. W eksperymentach rzeczywistych nie projektuje się jednak tak stromych filtrów, gdyż opóźnienie filtrów o skończonej odpowiedzi impulsowej FIR (ang. *Finite Impulse Response*) jest bardzo duże, zaś filtry o nieskończonej odpowiedzi impulsowej IIR (ang. *Infinite Impulse Response*) wymagają gwarantowania stabilności. Można zatem wysnuć wniosek, że filtracja dolnopasmowa mająca na celu poprawę jakości sygnału z akcelerometru jest nieskuteczna. Należy pamiętać też o tym, że wszelka filtracja mająca na celu poprawę jakości sygnału z akcelerometru musi



(a) Oryginalny (rzeczywisty) sygnał.

(b) Sygnał przefiltrowany.

Rysunek 5.4 Przyspieszenie robota.



(a) Oryginalny (rzeczywisty) sygnał.

(b) Sygnał przefiltrowany.

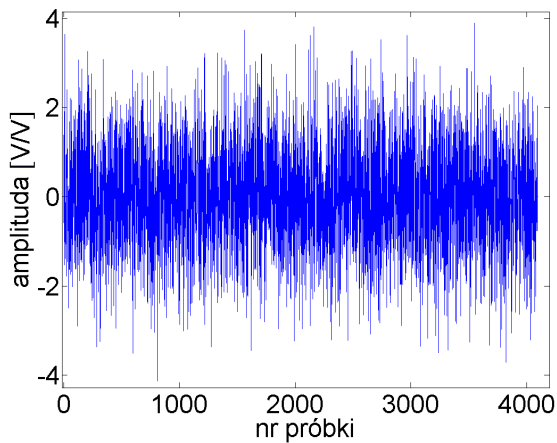
Rysunek 5.5 Położenie robota.

wiennie oddawać niskoczęstotliwościowe składowe, a przede wszystkim zachować wiernie składową stałą sygnału.

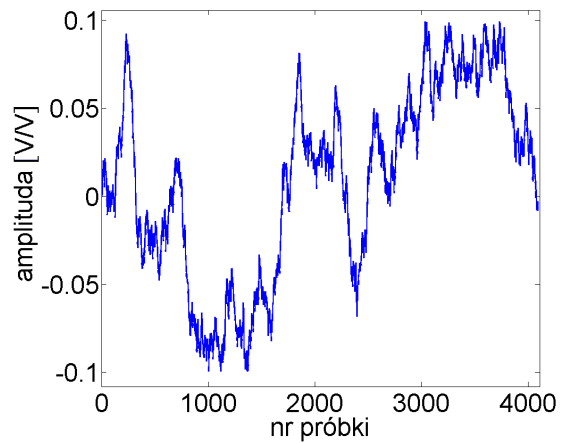
### 5.1.2. Szum akcelerometru

Aby określić, w jakim stopniu szum akcelerometru wpływa na estymację położenia robota, przeprowadzono kolejne symulacje. Jako podstawowy model szumu akcelerometru zastosowano biały szum, zobrazowany na rys. 5.6a. Wariancja generowanego szumu wynosi 1 [V/V], przy wartości średniej równej 0. Szum akcelerometru uwzględniany w badaniu przyjmuje postać losowego błędzenia, które wynika z całkowania białego szumu oraz odpowiedniego ograniczenia amplitudy sygnału do wartości  $\pm 0.1[V/V]$ , co zilustrowano na rys. 5.6b.

Aby zaobserwować, w jaki sposób błędzenie losowe jest filtrowane w wyniku całkowania, na rys. 5.7a przedstawiono wycinek szumowego sygnału typu błędzenie losowe przed filtracją, oraz na rys. 5.7b wynik jego filtracji dolnoprzepustowej z częstotliwością odcięcia 0.5Hz. Widać, że filtracja wyraźnie zmienia postać sygnału, jednak kształt i poziom amplitudy nie uległ wielkiej zmianie.

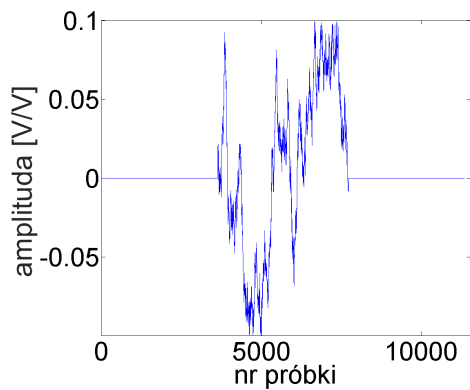


(a) Symulowany biały szum.

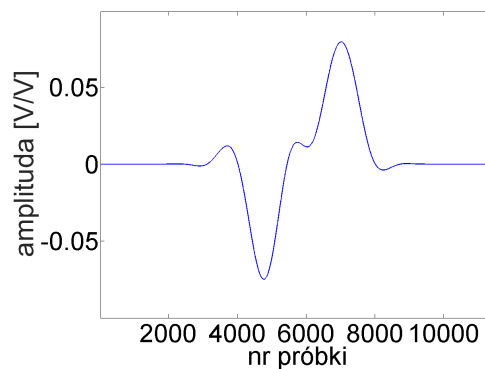


(b) Szum typu błądzenie losowe.

Rysunek 5.6 Generowane szумы użyte w symulacji.



(a) Oryginalna postać.



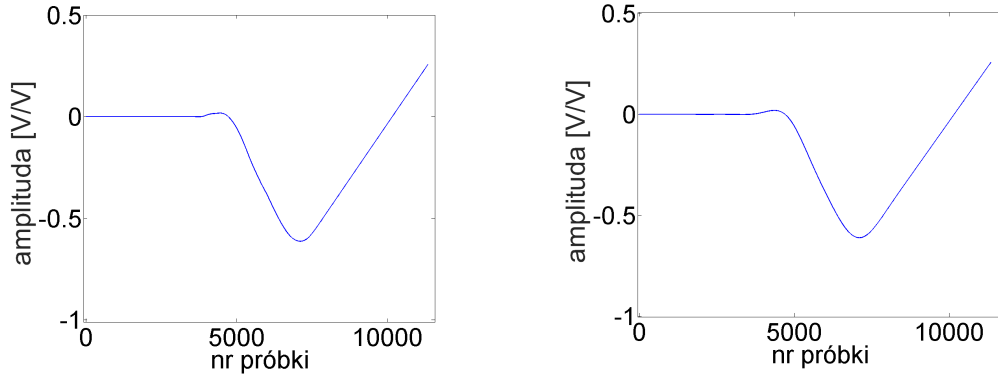
(b) Wynik filtracji dolnoprzepustowej.

Rysunek 5.7 Szum typu błądzenie losowe.

Wynik po podwójnym całkowaniu sygnału szumowego typu błądzenie losowe z rys. 5.7a (bez filtracji) oraz po filtracji (rys. 5.7b) można zaobserwować na rys. 5.8. Porównując oba wykresy trudno zauważyć różnicę. Potwierdza się zatem poprzednio wysnuty wniosek, że filtrowanie szumu akcelerometrycznego jest mało skuteczne. Ponadto nietrudno zauważyć narastanie negatywnego wpływu szumu wraz z czasem trwania ruchu.

Powyższe przykłady obrazują, że wyznaczenie poprawnego położenia robota w oparciu o całkowanie sygnału przyspieszenia jest zagadnieniem trudnym. Całkowanie silnie filtruje szum (z wyjątkiem jego składowej stałej). Analizując powyższe przykłady można wnioskować, że nie ma prostej metody poprawy sygnału przyspieszenia na drodze filtracji pasmowej. Dlatego warto ustalić, jaka jest niezbędna dokładność danych przyspieszenia oraz ułożenia, aby błąd położenia nie przekraczał określonej wartości przy zadanych warunkach ruchu.





(a) Sygnał bez filtracji.

(b) Sygnał po filtracji dolnoprzepustowej.

Rysunek 5.8 Dwukrotnie całkowany szum typu błądzenie losowe.

### 5.1.3. Wrażliwość pomiaru położenia

Wyznaczanie położenia poprzez całkowanie danych przyspieszenia jest wrażliwe na dane ułożenia i szum własny sensora. Błąd ułożenia powoduje, że pewna część przyspieszenia związanego z grawitacją dodaje się do mierzonego przyspieszenia. Wpływ grawitacji uwzględniony jest w modelu (4.5) w postaci drugiego składnika. Natomiast wpływ szumu sensora (rozważany w poprzedniej sekcji) wyraża się w ostatnim składniku równania (4.5).

Na potrzeby analizy założono, że zmienność szumu nie zależy od kierunku oraz że błąd ułożenia jest taki sam dla wszystkich trzech kątów obrotu. Zatem odchylenie standardowe szumu  $\Delta\eta$  oraz ułożenia  $\Delta\gamma$  można zapisać następująco:

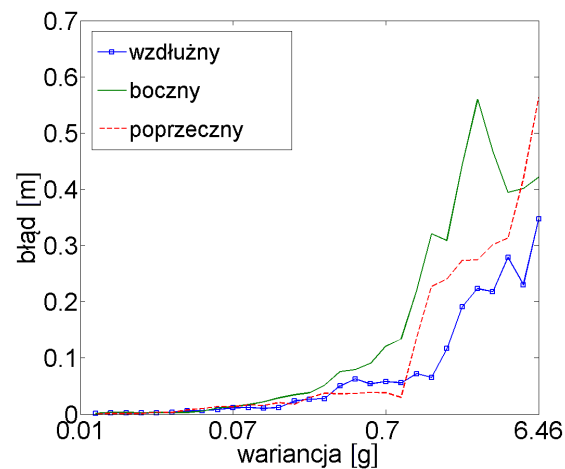
$$\begin{aligned} \eta_l = \eta_s = \eta_t = \Delta\eta \\ \phi = \theta = \psi = \Delta\gamma \end{aligned} \quad (5.1)$$

W badaniu symulacyjnym oceniano błąd położenia robota poruszającego się prostoliniowo po pierwszych trzech sekundach. Rzeczywisty dystans przejechany przez robota w symulacji to 1.2m (co dało 1200 próbek przy częstotliwości próbkowania 400Hz). Błąd wyznaczono po uśrednieniu ostatnich 40 próbek sygnału położenia (po zatrzymaniu robota). Aby skompensować wpływ konkretnych realizacji szumu, symulację wykonano dla 30 różnych realizacji szumu. Symulację przeprowadzono oddzielnie dla błędu przyspieszenia oraz błędu ułożenia.

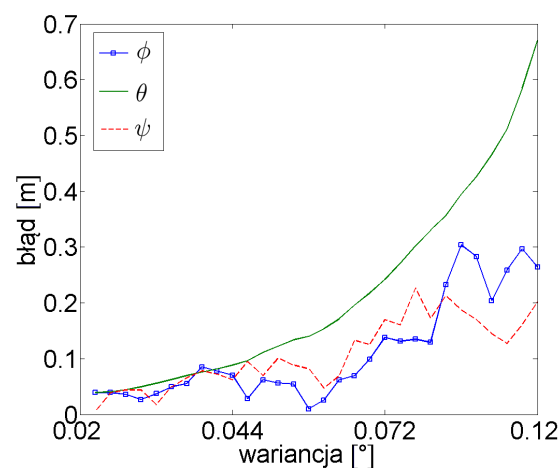
Na rys. 5.9 zobrazowano średni błąd położenia w zależności od wariancji szumu akcelerometru, przy zerowym błędzie ułożenia. Natomiast błąd położenia w zależności od wariancji błędu ułożenia, przy zerowym błędzie oceny przyspieszenia, można zaobserwować na rys. 5.10. Błąd ułożenia  $\Delta\gamma$  został uwzględniony także przy wyznaczaniu wartości przyspieszenia obrotowego i dośrodkowego, jak następuje:

$$\begin{bmatrix} \tilde{\Delta}a_l \\ \tilde{\Delta}a_s \\ \tilde{\Delta}a_t \end{bmatrix} = g \begin{bmatrix} -\sin\Delta\gamma \\ \sin\Delta\gamma \cos\Delta\gamma \\ \cos\Delta\gamma \cos\Delta\gamma \end{bmatrix} + \begin{bmatrix} (\dot{\Delta}\gamma^2 - \ddot{\Delta}\gamma)r_l \\ (\dot{\Delta}\gamma^2 + \ddot{\Delta}\gamma)r_s \\ (\ddot{\Delta}\gamma - \dot{\Delta}\gamma)r_t \end{bmatrix} + \begin{bmatrix} \Delta\eta \\ \Delta\eta \\ \Delta\eta \end{bmatrix} \quad (5.2)$$

gdzie  $\tilde{\Delta}a_l$ ,  $\tilde{\Delta}a_s$ ,  $\tilde{\Delta}a_t$  to wyznaczony błąd przyspieszenia, zaś  $r_l$ ,  $r_s$ ,  $r_t$  to odległości od środka obrotu odpowiednio w kierunku wzdłużnym, bocznym i poprzecznym.



Rysunek 5.9 Wpływ szumu akcelerometru na ocenę położenia robota.



Rysunek 5.10 Wpływ szumu ułożenia na ocenę położenia robota.

Na obu wykresach widoczny jest eksponencyjny wzrost błędu położenia ze wzrostem wariancji szumu. Błąd ten jest nieznacznie większy w kierunku bocznym, w porównaniu z kierunkiem wzdłużnym i poprzecznym robota. Ponadto błąd położenia jest bardziej wrażliwy na błędy ułożenia niż błędy oceny przyspieszenia. W kontekście postawionych celów pracy, na podstawie analizy dokładności mapowania zawartej w 4.4.3, można wstępnie przyjąć, że błąd położenia z pewnością nie powinien przekraczać wartości kilku centymetrów. Takie zaś założenie prowadzi do wymagania, aby wariancja błędu ułożenia była mniejsza niż  $0.04^\circ$ , zaś wariancja szumu akcelerometru była mniejsza niż  $0.1g$  dla ruchu trwającego  $3s$ . W przypadku krótszego czasu (całkowania) wymagana może być mniejsza precyzja pomiaru.

Na podstawie powyższej analizy wrażliwości całkowania danych z akcelerometru można wnioskować, że określanie położenia na podstawie przyspieszenia jest realizowalne. Takie podejście wymaga bardzo dokładnego pomiaru przyspieszenia i ułożenia nawet dla bardzo krótkiego ruchu robota (trwającego kilka sekund). Pasmowa filtracja sensorycznych sygnałów przyspieszenia o mniejszej precyzji jest generalnie nieskuteczna. Dlatego przy pomiarze przyspieszenia i ułożenia z mniejszą dokładnością należy dokonać obserwacji stanu po uprzednim zdefiniowaniu modelu ruchu robota. Taki model został wcześniej zaprezentowany w sekcji 4.3.3, zatem w oparciu o ustalone parametry tego modelu oraz modelu akcelerometru należy zweryfikować skuteczność obserwacji odpowiednio zamodelowanych zmiennych stanu oraz ostatecznej estymacji zadanej trajektorii robota odpowiednim studium symulacyjnym i eksperymentalnym.

## 5.2 Wyznaczanie trajektorii robota

Model robota mobilnego zrealizowany w ramach opracowanego systemu VisRobot pozwala przeprowadzić estymację położenia robota na podstawie zadanego wymuszenia i sygnałów pomiarowych przyspieszenia. Najistotniejszą cechą modelu matematycznego jest możliwość odwzorowania ograniczeń w dynamice ruchu, które również ograniczają błędy całkowania przyspieszenia opisane w sekcji 5.1. Oczywiście zastosowany model matematyczny dynamiki robota nie jest doskonały, ponieważ nie uwzględnia takich elementów, jak poślizg, upadek ze wzniesienia, uderzenie w przeszkodę, itp. Realizowany obserwatorowy pomiar przyspieszenia dostarcza też bieżącej oceny prędkości, choć bezpośrednia obserwacja położenia (z perspektywy układu dynamicznego) formalnie nie jest możliwa. Poprawność estymacji położenia robota zostanie zweryfikowana na drodze symulacji oraz eksperymentalnego badania rzeczywistego robota.

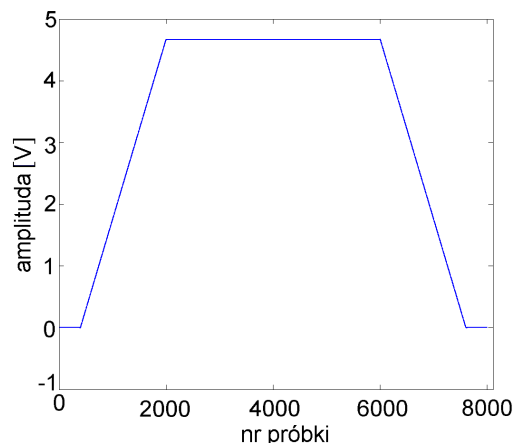
### 5.2.1. Testy symulacyjne

Testy estymacji położenia opartego na modelu dynamicznym platformy robota oraz danych z akcelerometru wymagają ustalenia parametrów eksperymentów (modelu oraz wymuszenia). Przy symulacji ruchu robota sprawdzono wrażliwość systemu z obserwatorem stanu na dane pomiarowe zawierające szum. Testy zrealizowano w oparciu o koncepcję obserwatora Leuenbergera oraz prostego filtra Kalmana (Kowalczyk i Merta, 2016a).

W badaniu symulacyjnym założono ruch niedużego robota na drodze 1.2m. Parametry symulacji dobrano tak, aby dynamika ruchu robota była zbliżona do ruchu rzeczywistego małego robota mobilnego używanego w badaniach doświadczalnych. Istotne parametry związane z modelem robota to promień koła  $r_w = 0.03m$  oraz przełożenie silnika  $ratio = 1:19$ . Ponadto część parametrów modelu robota (opisanego w sekcji 4.3.3) ustalono na drodze estymacji<sup>1</sup>.

Po ustaleniu wartości wszystkich parametrów, w ramach symulacji wygenerowano sygnał wymuszenia zilustrowany na rys. 5.11. Sygnał ma czasową postać napięciową o amplitudzie nie przekraczającej 5V i częstotliwościowe pasmo ograniczone za pomocą filtra dolnoprzepustowego (co ogranicza wielkość skoków pochodnej sygnału). Przyjęto także, że robot może się poruszać przy napięciu wejściowym o wartości co najmniej 0.5V. Na podane wymuszenie model robota odpowiada sygnałem położenia, co zobrazowano na rys. 5.12. Jest to odpowiedź systemu, w którym nie występuje szum. W rzeczywistym

<sup>1</sup>Autorską procedurę estymacji parametrów przedstawimy w sekcji 5.2.2

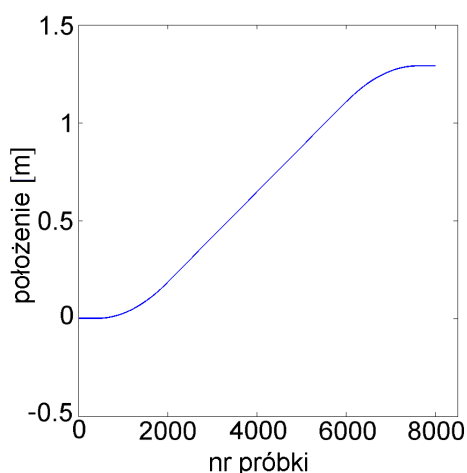


Rysunek 5.11 Symulowany sygnał wymuszenia robota.

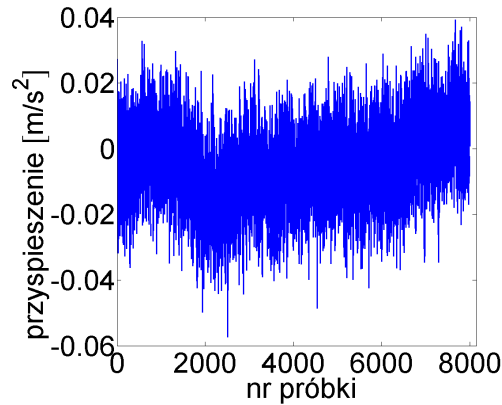
systemie występuje szum pomiarowy akcelerometru (i jego składowe). W symulacji przyjęto wartości szumu wynikające ze specyfikacji technicznej akcelerometru oraz wcześniejszej analizy. Proces szumowy typowy dla akcelerometru  $\eta_{acc}$  zamodelowano przy użyciu szumu typu błądzenie losowe  $\eta_d$  o następującym opisie:

$$\begin{aligned}\eta_d(n) &= \eta_d(n-1) + N(0, \sigma_d) \\ \eta_{acc}(n) &= \eta_d(n) + N(0, \sigma_n)\end{aligned}\tag{5.3}$$

gdzie  $N$  jest procesem gaussowskim, zaś  $\sigma_d$  i  $\sigma_n$  to odchylenia standardowe związane odpowiednio z błądzeniem losowym i białym szumem. Zgodnie ze specyfikacją maksymalna wartość szumu własnego akcelerometru wynosi 0.06g, zaś szum ułożenia nie przekracza  $0.7^\circ$ . W odróżnieniu od poprzednich badań symulacyjnych częstotliwość próbkowania zwiększono do wartości  $f_s=800\text{Hz}$ . Wizualizację symulowanego sygnału szumowego przedstawia rys. 5.13.

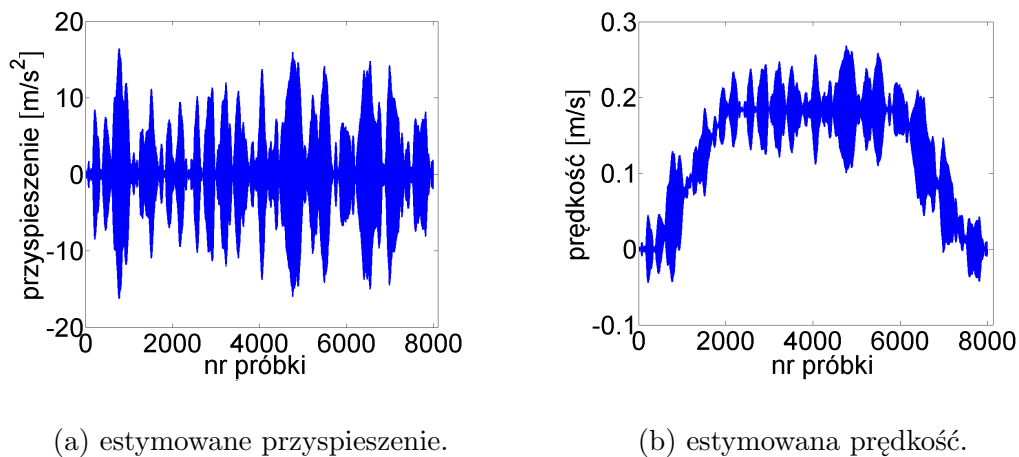


Rysunek 5.12 Symulowane położenie bez obecności szumu pomiarowego.



Rysunek 5.13 Symulowany szum akcelerometru.

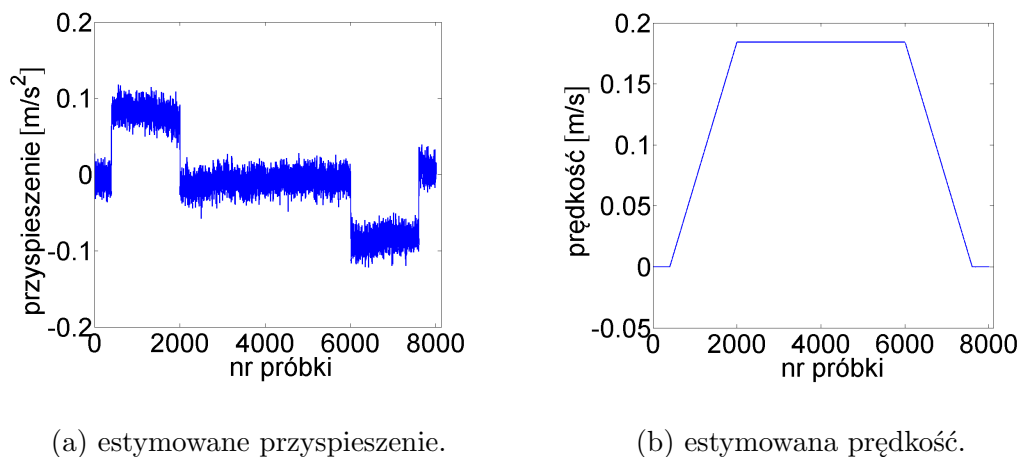
W pierwszym podejściu sprawdzono, jak szum akcelerometru wpływa na estymację stanu generowaną przez obserwator Leuenbergera przedstawiony w sekcji 3.3.3. Metodą lokowania biegunów zapewniono, że obserwator jest szybszy od badanego układu. Symulacja została zrealizowana dla sygnału wymuszenia zobrazonego na rys. 5.11. Wpływ szumu na obserwowane przyspieszenie oraz prędkość ilustruje rys. 5.14. Bez trudu zauważyć można, że estymowane przyspieszenie jest obarczone błędem, który jest dużo większy niż szum akcelerometru. Błędy prędkości są także istotne, jednak widać wyraźne podobieństwo estymowanego sygnału do prędkościowego wymuszenia zaprezentowanego na rys. 5.11. Natomiast uzyskano błąd estymowanego położenia (dla każdej próbki) mniejszy niż  $10^{-3}m$  (dlatego też pominięto wizualizację estymowanego położenia). Na tej podstawie można wnioskować, że nawet prosty obserwator stanu jest w stanie poprawnie estymować położenie robota przy obecności typowego szumu akcelerometru.



Rysunek 5.14 Estymacja stanu oparta na obserwatorze Leuenbergera dla typowego szumu akcelerometru.

Przewidując możliwość dalszej poprawy estymacji położenia, przeprowadzono symulacyjne badania testowe z użyciem filtracji Kalmana. Założono przy tym, że składowa szumu procesowego związana z przyspieszeniem jest o rząd wielkości większa od składowej związanej z prędkością.

Przebieg estymacji wektora stanu dla typowego szumu pochodzącego z akcelerometru można zaobserwować na rys. 5.15. Na wykresie estymowanej prędkości nie jest widoczny wpływ szumu pomiarowego, choć estymowane przyspieszenie zawiera istotny szum, którego amplituda nie przekracza  $0.05\text{m/s}^2$ . Jest to istotna poprawa względem estymacji stanu opartej na obserwatorze Leuenbergera. Niemniej w obu przypadkach błąd estymowanego położenia wyznaczonego przez całkowanie estymowanej prędkości jest pomijalnie mały.

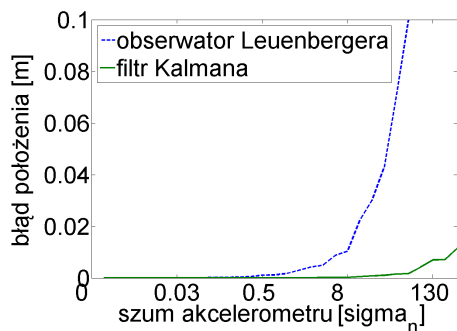


Rysunek 5.15 Estymacja stanu oparta na filtracji Kalmana dla typowego szumu z akcelerometru.

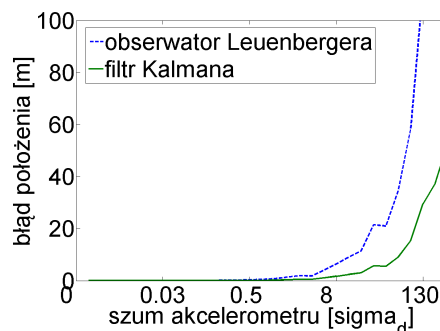
W ostatniej fazie badań symulacyjnych przeprowadzono analizę czułości estymacji położenia z użyciem obserwatora Leuenbergera i filtracji Kalmana. Błąd położenia został wyznaczony po 10s ruchu i uśredniony dla 20 różnych realizacji szumu. Obliczenia wykonano dla 30 różnych odchyłek standardowych szumu białego oraz szumu typu błądzenie losowe. Dla czytelności wykresu odchylenie standardowe szumu zwiększono wykładniczo w przybliżonym przedziale (0.001; 200). Wykresy czułości dla filtracji Kalmana i obserwatora Leuenbergera przedstawiono na rys. 5.16. Filtr Kalmana prowadzi do wyższej jakości estymacji w porównaniu do obserwatora Leuenbergera. Oczywiście przy małych wartościach wariancji szumu, oba obserwatory zachowują się podobnie (z małym błędem oceny położenia).

Na podstawie przeprowadzonych badań symulacyjnych można wnioskować, że estymacja położenia oparta na danych z akcelerometru, modelu robota i obserwatorze stanu może być skuteczna, albowiem błąd związany z całkowaniem danych jest ograniczony dzięki zastosowaniu dynamicznego modelu oraz obserwatora stanu. Filtr Kalmana daje lepsze wyniki estymacji w porównaniu z obserwatorem Leuenbergera. Warto też zaznaczyć, że w obecności dużego szumu pomiarowego, wynik estymacji zależy przede wszystkim od sygnału wymuszenia użytego w dynamicznym modelu robota, przy czym należy pamiętać o istotnej roli wzmocnienia Kalmana w procesie estymacji.

Warto także dodać, że dla estymacji położenia szum pomiarowy jest tym bardziej szkodliwy im mniejsze są stałe czasowe dynamiki robota. W związku z tym przyspieszenia chwilowe robota (w przód) powinny być możliwie duże. Natomiast ruch po łagodnym torze (łuku) daje małe chwilowe przyspieszenia, wówczas estymacja położenia robota w kierunku innym niż wzdłużny, wymagałaby silnego ograniczenia szumów pomiarowych.



(a) Biały szum.



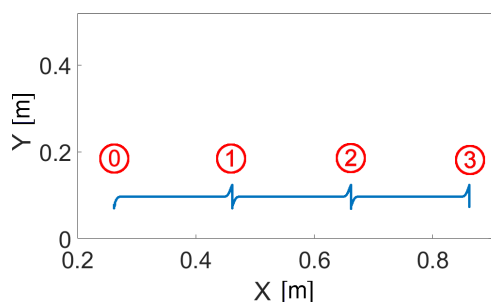
(b) Szum typu błądzenie losowe.

Rysunek 5.16 Studium czułości estymacji położenia dla obserwatora Leuenbergera (niebieski) i filtra Kalmana (zielony) w obecności szumu pomiarowego w postaci.

Biorąc pod uwagę powyższe własności, nie można zagwarantować, że przy rzeczywistym ruchu robota estymowany błąd nie będzie większy niż ten osiągnięty w studium symulacyjnym. Warto zatem przeprowadzić również testy doświadczalne w kontrolowanych laboratoryjnych warunkach, aby jednoznacznie określić, czy rozwiązanie jest skuteczne. Ze względu na charakter ruchu robota w opracowanym rozwiązaniu należy uwzględnić estymację położenia przede wszystkim w kierunku wzdłużnym robota.

### 5.2.2. Testy doświadczalne

Testy symulacyjne zaprezentowane w poprzedniej sekcji skupiały się na problemie estymacji położenia na podstawie danych akcelerometrycznych, modelu robota oraz sygnału wymuszenia. Wyniki symulacji przedstawiają położenie robota w postaci jednowymiarowej - w kierunku wzdłużnym robota. Zadanie mapowania w ogólności wymaga określenia położenia robota w trzech wymiarach. W przypadku systemu VisRobot, który służy stereowizji ze zmienną bazą, robot porusza się po trajektorii zilustrowanej na rys. 5.17. Po zapisie obrazów pierwotnej bazy (punkt 0), robot wykonuje obrót w miejscu o  $90^\circ$ ,

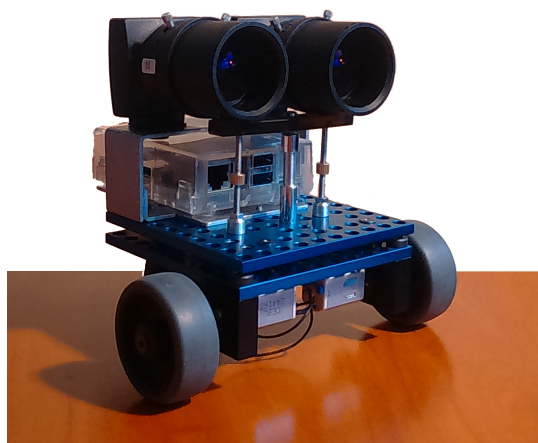


Rysunek 5.17 Trajektoria robota w systemie VisRobot.

ruchu po linii prostej (możliwie zbliżonej do linii prostej), obrót o  $-90^\circ$ . W tej pozycji następuje ewentualna korekta położenia (ruch do tyłu). Następnie zapisywany jest obraz

związany z pierwszą bazą (punkt 1), po czym schemat ten się powtarza (w przedstawionym przykładzie ostatnią pozycję robota zaznaczono punktem 3). Istotny jest tutaj dystans przejechany w danym kierunku (rozważany we wcześniejszych symulacjach) oraz kierunek, który wyznaczony będzie poprzez fuzję danych z jednostki IMU.

Testy doświadczalne realizowane w ramach pracy dotyczyły estymacji położenia wzdłużnego robota oraz kierunku ruchu. Określenie aktualnych wartości przesunięcia i kierunku ruchu robota sprowadzono do określenia jego trajektorii. Testy oparte były na rzeczywistym małym robocie kołowym widocznym na rys. 5.18. Aby w prosty sposób przeprowadzić pomiar i obserwację przejazdu robota założono, że robot ma przejechać drogę po obrysie prostokąta. Robot miał wykonać sekwencję ruchów do przodu, obrót w lewo (w miejscu), powtarzając to cztery razy, po czym wykonać pełny obrót w prawo. Aby ocenić poprawność estymacji ruchu i posiadać układ odniesienia konieczne było wykonanie dodatkowego zewnętrznego bezwzględnie pomiaru położenia robota. W celu zbadania położenia (w wybranych punktach) i dynamiki ruchu robota, zaprojektowano zatem układ odniesienia oparty na obrazie ze statycznej kamery, umieszczonej nad obszarem ruchu, oraz własnym algorytmie przetwarzania obrazu z wykorzystaniem kolorowego znacznika.



Rysunek 5.18 Platforma testowa robota mobilnego.

### *Bezwzględny pomiar trajektorii*

Bezwzględny pomiar trajektorii miał za zadanie określić położenie robota w oparciu o statyczną kamerę obserwującą robota z góry. Najpierw należało dobrać sygnał wymuszenia robota tak, aby trajektoria robota mieściła się w kadrze kamery i zajmowała znaczną jego część. Trajektoria dobrze widoczna w obrazie z kamery, skutkuje bowiem dużą liczbą pikseli opisujących ruch i wpływa pozytywnie na precyzję pomiaru. Ponadto przed właściwym pomiarem zawsze należy określić zależność pomiędzy odpowiednimi współrzędnymi obrazu i świata. Zależność ta wyznaczana była poprzez procedurę kalibracji kamery, która także eliminowała istotne zniekształcenia radialne obiektywu.

Bezwzględny pomiar położenia robota zrealizowano przez detekcję platformy robota w każdej klatce obrazu z kamery. W tym celu doświadczalnie zdefiniowano zakres barw w systemie HSV, który z zapasem obejmuje kolor platformy robota. Algorytm detekcji



wykrywa piksele zawierające się w zdefiniowanym zakresie, a następnie na masce binarnej z zaznaczonymi pikselami wykrywany jest największy kontur. Środek geometryczny tego konturu przyjmowany jest jako środek robota.

Algorytm ten jest prosty i skuteczny w wykrywaniu położenia robota dla poprawnie zdefiniowanego zakresu barw HSV, co zostało sprawdzone dla każdej klatki obrazu. Wcześniejsza implementacja oparta była na algorytmie MOG2, który wykorzystuje mikstury gaussa do usuwania tła (Zivkovic i der Heijden, 2006). Niestety maska obiektu wykrywana przez ten algorytm zmieniała się wskutek szumów kamery, co uniemożliwiało poprawne wykrycie środka obiektu podczas ruchu.

Bezwzględny pomiar trajektorii prowadził do stałego błędu położenia, który wynika z rozdzielczości kamery oraz kąta widzenia. Matryca użytej kamery, przy rozdzielczości 1280x1024 pikseli, była ustawiona równoległe do płaszczyzny ruchu robota. Algorytm działa z dokładnością do 1 piksela, co dla zastosowanego układu eksperymentu skutkowało dokładnością 0.12cm. Ze względu na małe zniekształcenia (pozostające po kalibracji) wartość ta wahała się w zakresie  $\pm 0.01\text{cm}$  dla brzegów kadru.

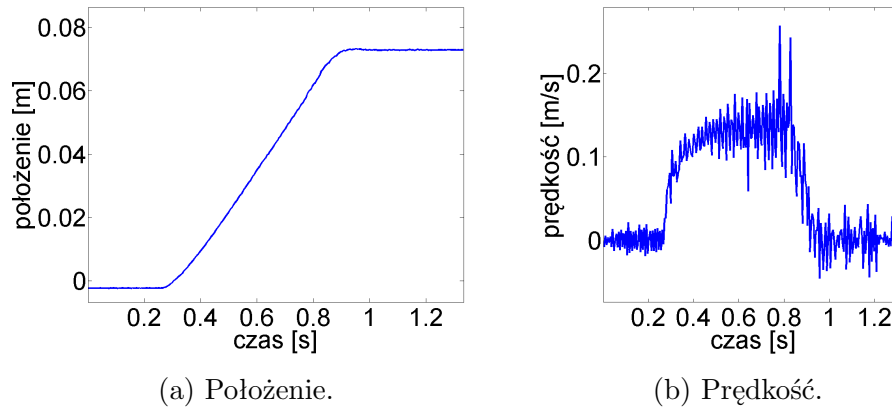
### *Estymacja parametrów modelu*

W przypadku oceny położenia robota mobilnego, czyli filtracji rzeczywistego sygnału przez obserwator, istotną rolę pełni także zagadnienie wyznaczania parametrów modelu robota. Model ten opiera się na opisie mechanizmów silnika prądu stałego. Parametry dużych silników z dedykowanym sterownikiem są zwykle podane przez producenta, natomiast dokumentacja małych (niedrogich) silników zawiera co najwyżej szczętkowe dane. W przypadku robota testowego użyto małych silników Pololu ze sterownikiem DRV8835, opartym na mostku H, których parametry są niedostępne. Pomiar parametrów jest ponadto czasochłonny i wymaga odpowiedniego stanowiska pomiarowego.

Aby wyznaczyć parametry modelu zastosowano autorską metodę estymacji parametrów silnika DC wykorzystującą obraz z kamery CMOS. Alternatywą mogłaby być estymacja oparta na danych z enkoderów, gdyby platforma była zaopatrzona w odpowiednie enkodery. Metoda autorska wykorzystuje własność kamery CMOS, która pozwala na zwiększenie liczby klatek na sekundę (FPS) kosztem rozdzielczości. Dzięki temu kamera CMOS o rozdzielczości 1280x1024 pikseli i 30 klatkach na sekundę umożliwia również pozyskiwanie obrazu o rozdzielczości 320x240 pikseli i blisko 200 FPS. Szczegółowy opis problemu został opisany w publikacji (Kowalczyk i Merta, 2017).

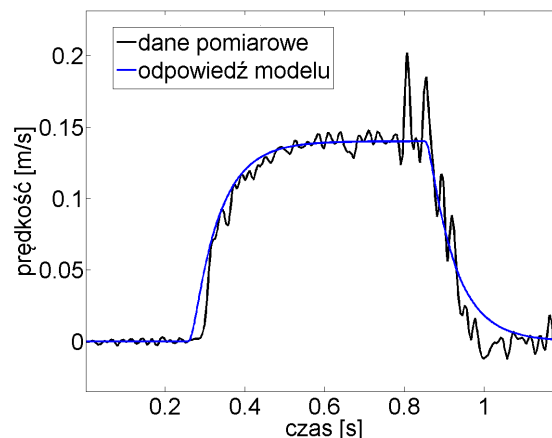
Algorytm estymacji parametrów polega na detekcji kolorowego znacznika zainstalowanego na robocie. Akwizycja obrazu obejmuje początek ruchu robota, przyspieszenie, jazdę z ustaloną prędkością i zatrzymanie robota. Ruch ten reprezentuje odpowiedź robota na sygnał wymuszenia w postaci prostokątnego skoku. Położenie znacznika w obrazie po odpowiednim przeskalowaniu określa rzeczywiste położenie robota. Właściwą odpowiedzią modelu jest prędkość, dlatego zarejestrowane zmienne położenie jest różniczkowane, a następnie poddane odpowiedniej filtracji dolnoprzepustowej. Częstotliwość odcięcia filtru oraz tłumienie zostało dobrane tak, aby usunąć tylko szum wysokoczęstotliwościowy. Mierzone położenie robota oraz prędkość, użyte do estymacji parametrów, zostały przedstawione na rys. 5.19.

Poprawność pomiaru jest uzależniona od stałej czasowej silnika DC pod obciążeniem. Przy szybkości 200 klatek na sekundę i założeniu, że potrzeba co najmniej 10 próbek



Rysunek 5.19 Mierzone położenie oraz prędkość.

reprezentujących odpowiedź skokową robota, proponowana metoda jest użyteczna, gdy dominująca stała czasowa robota jest nie mniejsza niż  $50ms$ . Dla bardziej dynamicznych i lekkich robotów kołowych może być potrzebna kamera umożliwiająca rejestrację większej liczby klatek na sekundę. Zauważmy jednak, że rejestracja znacznej liczby klatek na sekundę wiąże się z krótkim czasem otwarcia przysłony. Dlatego podczas takiego pomiaru należy zapewnić źródło światła, dla którego nie występuje efekt migotania.



Rysunek 5.20 Dopasowanie odpowiedzi modelu do danych pomiarowych (prędkości).

Po wyznaczeniu odpowiedzi układu (w postaci sygnału prędkości) realizowana jest procedura dopasowania parametrów modelu do zarejestrowanej odpowiedzi. W tym celu użyto nieliniowej metody najmniejszych kwadratów (dostępnej w programie Matlab). Po zapisaniu danych początkowych, algorytm iteracyjnie ze zmiennym krokiem zmienia wartości parametrów, po czym wyznacza aktualny błąd. Sygnał prędkości (rys. 5.19) wyznaczony w oparciu o dane z kamery został zatem poddany filtracji dolnoprzepustowej. Następnie za pomocą nieliniowej metody najmniejszych kwadratów dopasowano odpowiedź modelu robota do przefiltrowanych danych pomiarowych. Wyniki takiej identyfikacji przedstawiono na rys. 5.20, gdzie zobrazowano odpowiedź modelu dla zidentyfikowanych parametrów oraz mierzoną odpowiedź.

Tablica 5.1 Wartości parametrów modelu robota opisanego wzorem (4.15).

Parametr modelu	$k_m [\frac{Nm}{A}]$	$k_e [\frac{Vs}{rad}]$	$J[kg m^2]$	$B_v[\frac{kg}{m s}]$	$R_r [Ohm]$	$L_r [H]$
Wartość	0.31	0.31	$2.9 \cdot 10^{-5}$	$0.95 \cdot 10^{-4}$	2.1	$1.54 \cdot 10^{-3}$

Dokładność użytej metody jest związana z drogą jaką robot pokonuje od początku ruchu do uzyskania stałej prędkości. Szczególnie istotna jest liczba pikseli odpowiadająca tej drodze. W przedstawionym przypadku robot przejechał dystans  $6.8cm$ , co odpowiada liczbie 270 pikseli. Przy braku istotnych zniekształceń obrazu uzyskano błąd położenia  $1mm$  ( $\pm 0.5$  piksela), co jest w zupełności wystarczające dla potrzeb rozważanego eksperymentu identyfikacyjnego. Uzyskane w poniższym eksperymencie estymowane parametry fizycznego modelu platformy mobilnej przedstawiono w tab. 5.1.

#### Doświadczalna ocena położenia

Dysponując dodatkowym (bezwzględnym) pomiarem położenia oraz metodą pozwalającą na wyznaczenie parametrów modelu, można przystąpić do właściwego testu, który ocenia i wizualizuje estymowane położenie robota, w oparciu dane inercyjne i filtrację Kalmana (system VisRobot-KF). Trajektoria robota została dobrana tak, aby robot poruszał się po obszarze, który zajmuje możliwie dużą część obrazu z kamery. Dzięki temu precyzja bezwzględnego pomiaru położenia robota nie będzie niższa niż  $0.12cm$  (co opisano w sekcji 5.2.2).

Doświadczalnej oceny położenia dokonano dla robota kołowego, którego parametry modelu wyznaczono opisaną poprzednio autorską metodą estymacji opartą na obrazie z kamery CMOS. Przed właściwym doświadczeniem, po złożeniu wszystkich elementów składowych robota, skalibrowano sensory wchodzące w skład IMU, w szczególności magnetometr.

Zgodnie z przyjętym modelem dynamiki robota mobilnego (4.15) filtr Kalmana zastosowany został do estymacji drogi z uwzględnieniem wartości parametrów (wyestymowanych) przedstawionych w tab. 5.1. Po wyznaczeniu macierzy  $A, B, C, D$  i ich konwersji z wersji ciągłoczasowej na dyskretnoczasową przy częstotliwości próbkowania  $fs = 50Hz$ , równania stanu opisujące proces dyskretny podlegający filtracji Kalmana zapisać można następującym wyrażeniem:

$$\begin{aligned}
 \mathbf{x}(k+1) &= \begin{bmatrix} 0.7900 & 0.0025 \\ -11.1082 & -0.0337 \end{bmatrix} \mathbf{x}(k) + \begin{bmatrix} 0.000048 \\ 0.0025 \end{bmatrix} \mathbf{u}(k) + \mathbf{v}(k) \\
 \mathbf{y}(k) &= \begin{bmatrix} 0 \\ 4130 \end{bmatrix} \mathbf{x}(k) + \mathbf{w}(k)
 \end{aligned} \tag{5.4}$$

Wektory stanu (zawierający prędkość i przyspieszenie kątowe), szumu procesowego oraz pomiarowego mają następującą postać:

$$\mathbf{x} = \begin{bmatrix} \omega \\ \epsilon \end{bmatrix}$$

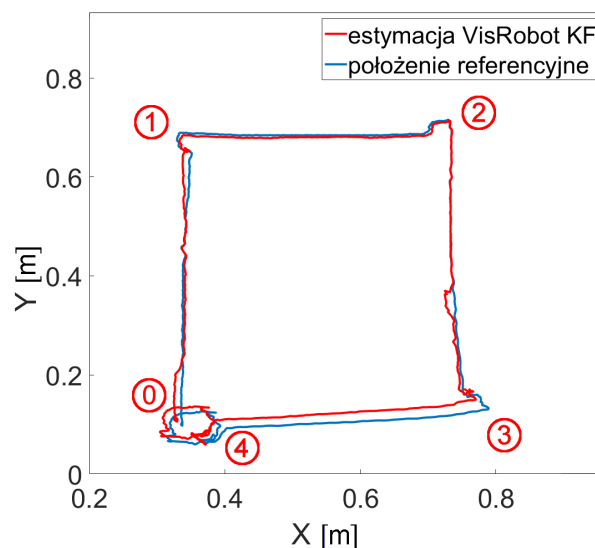
$$\mathbf{v}(k) \sim N\left(0, \begin{bmatrix} 0.1 \\ 1 \end{bmatrix}\right) \quad (5.5)$$

$$\mathbf{w}(k) \sim N(0, 1)$$

zaś estymowaną<sup>2</sup> drogę w czasie  $T_k$  próbek ( $\omega_d$  to prędkość kątowna w postaci dyskretnej) można zapisać wyrażeniem:

$$\mathbf{w}_1 = 6.521 \sum_{k=0}^{T_k} \omega_d(k) \quad (5.6)$$

Test przeprowadzono 15 razy (dla czytelności wyników zobrazowany jest tylko jeden wybrany test przedstawiony na rys. 5.21, pozostałe przypadki były analogiczne). Warto dodać, że wcześniejsze próby pomiarowe (przed kalibracją) zawierały istotne błędy w sygnale z magnetometru. Podczas ruchu po linii prostej, kiedy kierunek zmieniał się co najwyżej o kilka stopni, w danych z magnetometru widoczne było prawie liniowe narastanie kąta do wartości blisko  $20^\circ$ . Przyczyną problemów najpewniej było ulokowanie IMU zbyt blisko silników oraz ferromagnetyków lub przewodów zasilających występujących pod podłogą.



Rysunek 5.21 Estymacja położenia robota za pomocą systemu VisRobot-KF.

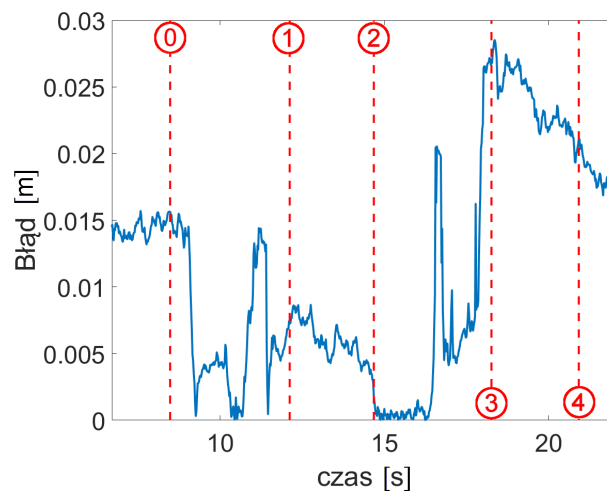
Estymowana trajektoria robota oraz trajektoria referencyjna wyznaczona za pomocą bezwzględne pomiaru położenia zostały zobrazowane na rys. 5.21. Referencyjne położenie (kolor niebieski) uzyskano na podstawie obrazu z kamery obserwującej eksperyment,

<sup>2</sup>w oparciu o postać ciągłoczasową (4.16)



natomiast estymowana przez VisRobot-KF trajektoria (kolor czerwony), jak widać, zawiera błędy położenia. Robot porusza się w miarę dokładnie po bokach prostokąta, jednak geometria napędu oraz położenie znacznika, służącego do wizualnego rozpoznawania położenia, spowodowały, że trajektoria jest nieregularna. Oznaczenia zobrazowane na rys. 5.21 w postaci punktów (0 - 4) dzielą trajektorię na sekwencje zawierające ruch do przodu oraz skręt w prawo. Po dotarciu do punktu 4, robot zatrzymuje się w miejscu przez około 5 sekund, po czym wykonuje obrót o  $360^\circ$ .

Średniokwadratowy błąd oceny położenia w analizowanym przypadku przedstawiono na rys. 5.22, gdzie także zaznaczono początki i końce sekwencji ruchu (do przodu) oraz skrętu (w prawo). Uśredniony błąd po 10 iteracjach nie przekraczał  $0.02m$ . Można zatem przyjąć, że błąd jest rzędu kilku centymetrów, a jego wartość chwilowa związana jest przede wszystkim z zakłóceniami magnetometru, których skutki można zaobserwować np. między punktami 2 i 3. Błędy te pojawiały się także w innych próbach (choć nie zawsze w tych samych momentach).

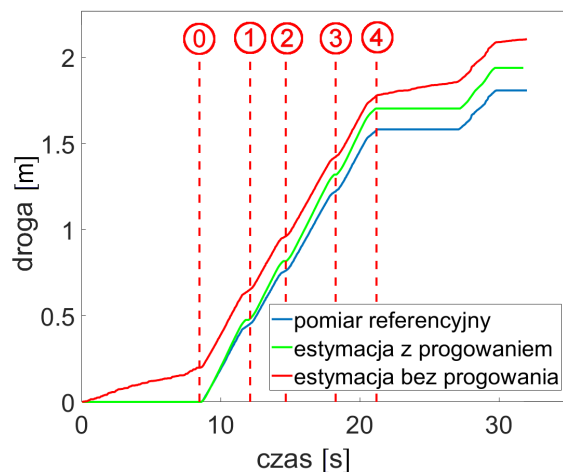


Rysunek 5.22 Chwilowy błąd estymowanego położenia robota.

Oprócz błędów związanych z chwilowym ułożeniem, na trajektorię istotnie wpływają też błędy omawianej metody określania trasy robota. Szum zawarty w danych sensorycznych wywołuje niewielki dryft położenia, czego dowodzą niewielkie zmiany położenia, kiedy robot znajduje się w spoczynku. Warto zaznaczyć, że dryft ten jest wiele rzędów mniejszy niż dryft powstały po podwójnym całkowaniu przyspieszenia bez użycia modelu robota.

Zakładając, że robot wykorzystuje prędkości robocze wyraźnie różne od zerowej (nie wykonuje powolnych ruchów postępowych), można dodatkowo zastosować progowanie estymowanej prędkości. Założono zatem, że ruch robota występuje, gdy jego prędkość przekracza  $1cm/s$ . Estymację przejechanej drogi z wykorzystaniem prezentowanej metody bez progowania oraz z progowaniem, z uwzględnieniem pomiaru referencyjnego przedstawiono na rys. 5.23.

Porównując końcowe położenia zauważyć można, że przy blisko 2 metrowej drodze metoda bez progowania prowadzi do błędu  $28cm$  (14%), natomiast z progowaniem błąd nie przekracza  $12cm$  (6%). W pozostałych przypadkach błąd metody z progowaniem nie



Rysunek 5.23 Estymacja drogi przejechanej przez robota: pomiar referencyjny, estymacja z progowaniem, estymacja bez progowania.

przekraczał 8%. Można więc przyjąć, że uzyskany błąd jest mniejszy niż 10% w warunkach laboratoryjnych, które charakteryzowały względnie płaskie i twarde podłoże oraz dynamiczny ruch robota. Podczas ruchu robota nie zaobserwowano istotnego ślizgu.

### 5.3 Analiza dokładności mapowania

Proces mapowania oparty na stereowizji zakłada, że parametry zewnętrzne kamer są znane z dużą dokładnością. W przypadku systemu VisRobot (ze zmienną bazą) założono, że precyzja położenia i ułożenia kamer związana jest z ograniczeniami metod przedstawionych odpowiednio w sekcjach 4.3, 4.2, zaś testy obu metod (wyznaczania położenia i ułożenia) zaprezentowano w poprzedniej sekcji (sekcja 5.2).

Dokładność odwzorowania otoczenia na mapie jest tu uzależniona głównie od błędów algorytmu wyznaczania dysparycji. Wrażliwość na te błędy określono na drodze analitycznej w sekcji 4.4.2. Przedstawione tam wyniki zakładają, że algorytm SGBM (oparty na dopasowaniu obszarów) poprawnie znajduje punkty będące rzutami punktu 3D na płaszczyznę obrazów. Warto zatem, na drodze odpowiedniej symulacji i eksperymentu, ocenić wrażliwość algorytmu przy określonej wielkości błędów położenia i ułożenia. Niezbędne wydaje się także przeprowadzenie ewaluacji systemu ze zmienną bazą dla zadanej sceny (mapowanego miejsca) oraz skonfrontowanie wyników z pomiarami dalmierza laserowego.

Praktyczna analiza wrażliwości stereowizji jest bardzo trudna, gdyż wymaga precyzyjnych pomiarów. Szczególnie trudne jest wyznaczenie położenia i ułożenia kamer (lewej i prawej) z dużo większą dokładnością pomiarową niż ta, uzyskiwana z wykorzystaniem dostępnych sensorów inercyjnych. Dlatego też należy rozważyć testy wrażliwości w środowisku wirtualnym.

#### 5.3.1. Wrażliwość systemu stereowizyjnego dla wirtualnej sceny

Badanie wrażliwości systemu stereowizyjnego sprowadza się do precyzyjnego rozmieszczenia lewej i prawej kamery zgodnie z szerokością bazy (przy zerowym błędzie położenia i ułożenia). Następnie dla jednej z kamer należy zmieniać ułożenie (z zadanego zakresu kąтового) i położenie (z ustalonego zakresu odległości). Obrót kamery o bardzo mały kąt (mniejszy niż  $1^\circ$ ) wymaga jednak certyfikowanego laboratorium z dedykowanymi precyzyjnymi przyrządami. Co więcej precyzyjny obrót wymaga odpowiedniego zamocowania kamery i obiektywu, aby takiemu obrotowi nie towarzyszyło przesunięcie. Problem ten pojawia się, gdyż w praktyce nie jest znane dokładne położenie osi obrotów kamery.

Wrażliwość stereowizji określono za pomocą autorskiego oprogramowania VIS służącego do analizy mapy dysparycji oraz głębi wybranego punktu, a także do zobrazowania chmury punktów. Program wyznacza też kwalifikowalność (iloraz liczby pikseli o znanej dysparycji przez liczbę wszystkich pikseli) dysparycji/głębi dla całego obrazu.

#### *Analiza problemu*

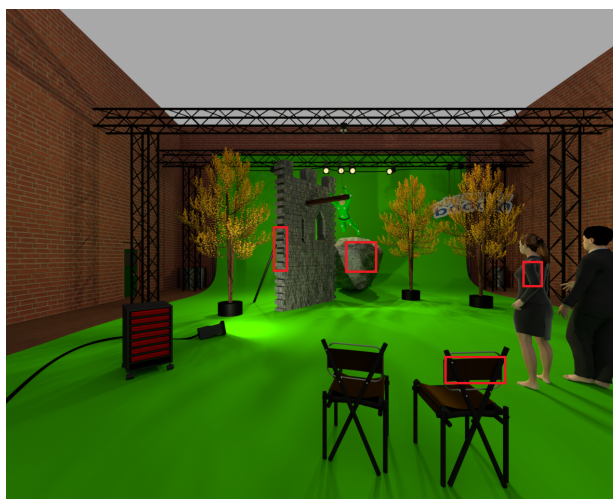
Błąd odwzorowania danego punktu 3D jest bezpośrednio związany z błędem dysparycji (pikselowego przesunięcia danego punktu na obrazie). Zgodnie z (3.39) zależność błędu dysparycji od błędu w metrach jest nieliniowa, przez co poziom błędu (w metrach) jest różny dla obiektów położonych blisko i daleko. Aby zestawić błędy odwzorowania dla obiektów skupimy się na błędzie we współrzędnych obrazu (błędzie dysparycji).

Przed właściwym eksperymentem wykonano krótki test, aby wstępnie oszacować wrażliwość stereowizji przy szerokości bazy 4cm. Wstępne pomiary wykazały, że stereowizja jest bardzo wrażliwa na błąd ułożenia. Nawet błąd ułożenia kamery rzędu  $0.5^\circ$  (i towarzyszące mu przesunięcie w trakcie pomiaru) silnie wpływa na błąd dysparycji. Aby zatem określić wrażliwość stereowizji należy ustawić kamerę z dokładnością ułożenia rzędu

0.1° bez dodatkowej translacji. Bez specjalistycznego laboratorium nietrudno o błędy pomiarowe związane z ugięciem czy przesunięciem podłoża, na którym umieszczono kamerę. Dlatego wyniki rzeczywistego pomiaru wrażliwości stereowizji należy interpretować ostrożnie zakładając, że błędy pomiaru ułożenia są stosunkowo niewielkie (rzędu 0.5°).

### *Wnioski*

Brak możliwości precyzyjnego pomiaru ułożenia i położenia wymusił zatem przeprowadzenie pomiarów w wirtualnym środowisku. Mając dowolną wirtualną scenę 3D można ustawić położenie i ułożenie wirtualnej kamery, a następnie wygenerować obraz (rendering obrazu). W takim przypadku kamerę ustawić można z bardzo dużą dokładnością. W typowych programach do tworzenia i edycji scen 3D precyzja ułożenia i położenia kamery osiąga nawet poziom 0.0001° i 0.00001m, co bezpośrednio wynika z użycia liczb zmiennoprzecinkowych pojedynczej precyzji i w zupełności wystarcza stawianym tu wymaganiom.



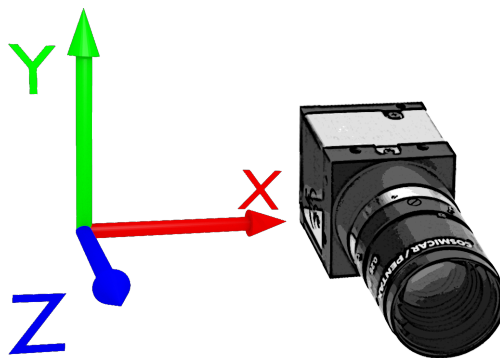
Rysunek 5.24 Render obrazu wirtualnej sceny "movie set scene" (źródło: free3d.com, licencja: Creative Commons 0).

### *Scena*

Obrazy do testowania wrażliwości stereowizji wygenerowano w programie Blender. Scena przedstawiona na rys. 5.24 posłużyła do renderowania obrazów dla wybranych ułożeń kamery. Typ kamery, rozdzielczość, ogniskową i inne parametry dobrano tak, aby jej właściwości były zbliżone do właściwości kamery używanej w testach rzeczywistych. Ponadto dobrano ustawienia renderingu, aby obraz był możliwie wysokiej jakości. Renderowanie 72 klatek zajęło kilka dni dla 2 komputerów PC z wydajnymi kartami graficznymi.

Dla odpowiednich par wyrenderowanych obrazów wyznaczono mapę dysparycji, za pomocą systemu VIS. Pomiary przeprowadzono dla bazy o szerokości 4cm, gdzie pozycja lewej kamery jest stała, natomiast prawa kamera jest dodatkowo obracana względem osi  $x$ ,  $y$ ,  $z$  oraz przesuwana w kierunku  $y$  i  $z$  (przesunięcie w kierunku  $x$  oznacza zmianę szerokości bazy). Kierunki odniesienia  $(x, y, z)$  względem kamery zobrazowano na rys. 5.25.





Rysunek 5.25 Kierunki odniesienia eksperymentu.

Tablica 5.2 Wyniki pomiaru odległości charakterystycznych punktów badanych obiektów wirtualnych ( $O_1 - O_4$ ).

	$O_1$	$O_2$	$O_3$	$O_4$
odległość [m]	1.77	2.95	7.16	9.03
rozzut [m]	0.1	0.07	0.14	0.27
dysparycja [piksele]	100	59	24	19

Ewaluację wrażliwości systemu oparto na procedurze określania kompletności i poprawności mapy oraz wartości dysparycji dla 4 wybranych punktów charakterystycznych obiektów:  $O_1$  (krzesło z prawej),  $O_2$  (osoba z lewej),  $O_3$  (mur),  $O_4$  (głaz). Analizowane punkty znajdują się w obszarach zaznaczonych na rys. 5.24, zaś wyniki pomiaru odległości tych punktów<sup>3</sup> od bazy kamer przedstawiono w tab. 5.2. Obiekty te znajdują się w różnej odległości od kamery oraz w różnych fragmentach kadru. Zatem, aby ułatwić porównanie błędu mapowania tych obiektów, wizualizacja wrażliwości stereowizji będzie przedstawiana w postaci błędu dysparycji (nie zaś głębi).

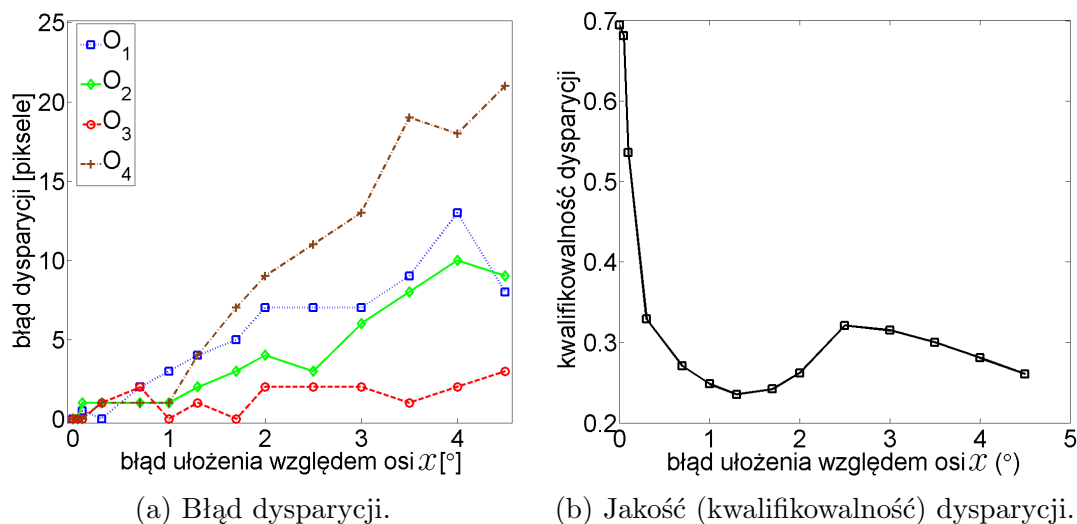
#### *Analiza wpływu błędów ułożenia*

Na rys. 5.26 przedstawiono wrażliwość stereowizji na błąd ułożenia w osi  $x$ . Błąd dysparycji, pokazany na rys. 5.26a, jest bardzo mały dla błędu ułożenia w kategorii obrotu względem osi  $x$  mniejszego od  $0.3^\circ$ . Dla większych wartości takiego obrotu, błąd dysparycji wyraźnie rośnie. Jednakże dla krzesła  $O_1$  położonego blisko kamery błąd ten rośnie nieznacznie, zaś dla głazu  $O_4$  błąd gwałtownie rośnie. Można przyjąć, że im bardziej odległy jest dany obiekt tym większa jest wrażliwość (błąd dysparycji).

<sup>3</sup>Punkt taki obejmuje obszar kilku pikseli więc jego odległość od bazy kamer zawiera się w pewnym przedziale.



Na rys. 5.26b z kolei zaobserwować można, że kwalifikowalność dysparycji gwałtownie spada już przy wartości obrotu wokół osi  $x$  około  $0.3^\circ$ , co potwierdza wysoką wrażliwość stereowizji w tym aspekcie. W zakresie błędu ułożenia  $1.5^\circ - 2.5^\circ$  występuje pewna poprawa kwalifikowalności dysparycji, co najpewniej spowodowane jest faktem, że na skutek składowej ruchu kamery w pionie, algorytm stereowizyjny SGBM znajduje duże podobieństwo między danym punktem należącym do obiektu a punktem powyżej lub poniżej.



Rysunek 5.26 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zakresie błędu ułożenia względem osi  $x$  dla 4 obiektów.

Na podstawie podanego eksperymentu, ale dotyczącego obrotu względem osi  $y$ , którego wyniki pokazano na rys. 5.27 można zauważyć, że błędy stereowizji w tym przypadku mają inny charakter. Niezależnie od odległości obiektu od kamery błąd dysparycji rośnie liniowo wraz ze wzrostem błędu ułożenia. Jednakże powyżej pewnej wartości błędu ułożenia (względem osi  $y$ ) obiekty odległe  $O_3$  (mur) i  $O_4$  (głaz) mają zerową dysparycję, która prowadzi do maksymalnego błędu odległości. Kwalifikowalność dysparycji maleje wraz ze wzrostem błędu ułożenia względem osi  $y$ , ale krzywa kwalifikowalności jest mniej stroma w porównaniu z krzywą kwalifikowalności dla błędu ułożenia względem osi  $x$ .

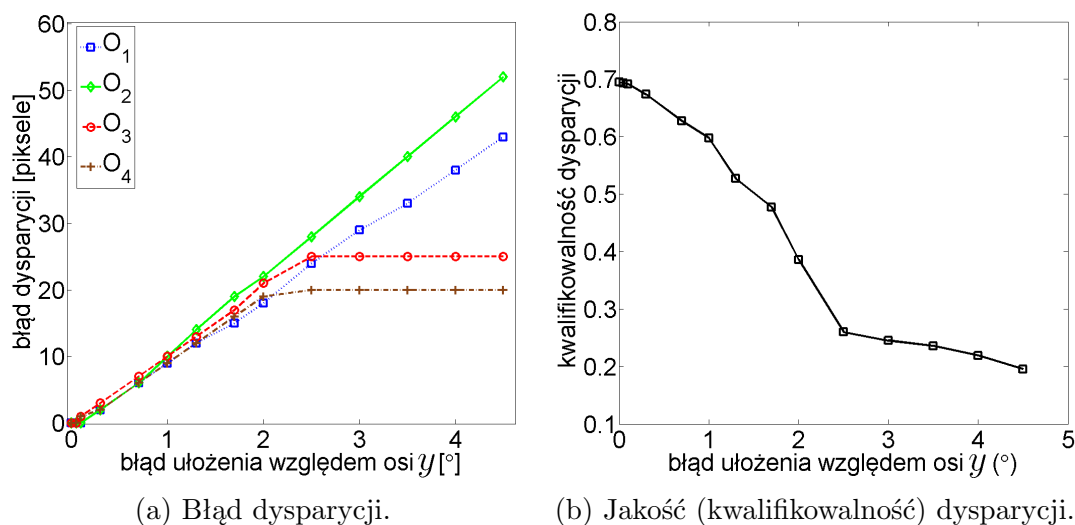
W przypadku obrotu kamery wokół osi  $z$  błąd dysparycji jest względnie najmniejszy (w porównaniu z poprzednimi przypadkami), co widać na rys. 5.28a. Błędy są większe dla obiektu  $O_1$  (krzesło) i  $O_2$  (osoba) w porównaniu z obiektami  $O_3$  (mur) i  $O_4$  (głaz), które mają praktycznie zerowy błąd dysparycji niezależnie od błędu ułożenia. Błąd ten jest silnie zależny od położenia obiektu w kadrze, przez co rośnie wraz z odległością od środka kadru (punktu głównego). Na podstawie rys. 5.28b można zauważyć, że krzywa kwalifikowalności dysparycji jest względnie duża przy błędach ułożenia mniejszych od około  $2^\circ$ .

#### Analiza wpływu błędów położenia

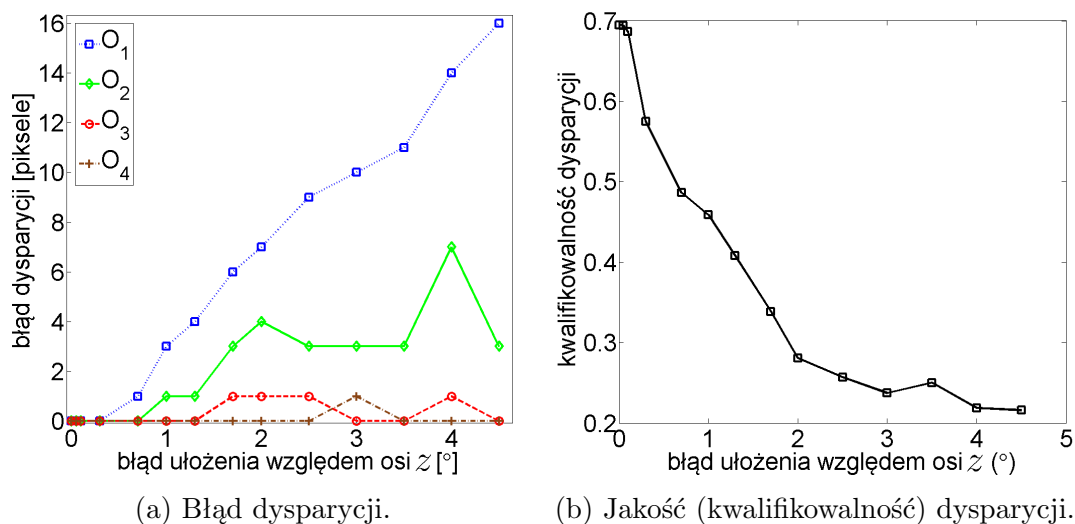
Podczas testów sprawdzono też wrażliwość stereowizji na błędy położenia dla dużego zakresu wartości w przedziale  $\langle 0.01mm; 200mm \rangle$ . Błąd dysparycji spowodowany błędem

położenia w kierunku  $y$  pokazano na rys. 5.29a. Błąd ten jest także zależny od obecności w obrazie pionowych elementów (obecność gładkich lub wzorzystych elementów o powtarzającym się wzorcu w kierunku pionowym). Duże przesunięcie w pionie powoduje, że algorytm stereo (SGBM) nie wyznacza dysparycji. Natomiast przy analogicznym przesunięciu elementu pionowego taka sytuacja nie występuje.

Ogólnie poziom błędu dysparycji jest dużo mniejszy niż w przypadku obecności błędów ułożenia. W przypadku każdego z badanych obiektów na skutek przesunięcia kamery o  $50\text{mm}$  występuje ledwie zauważalny błąd dysparycji. Jednakże przy tym poziomie błędu położenia kwalifikowalność dysparycji jest już niestety mała. Z drugiej strony należy

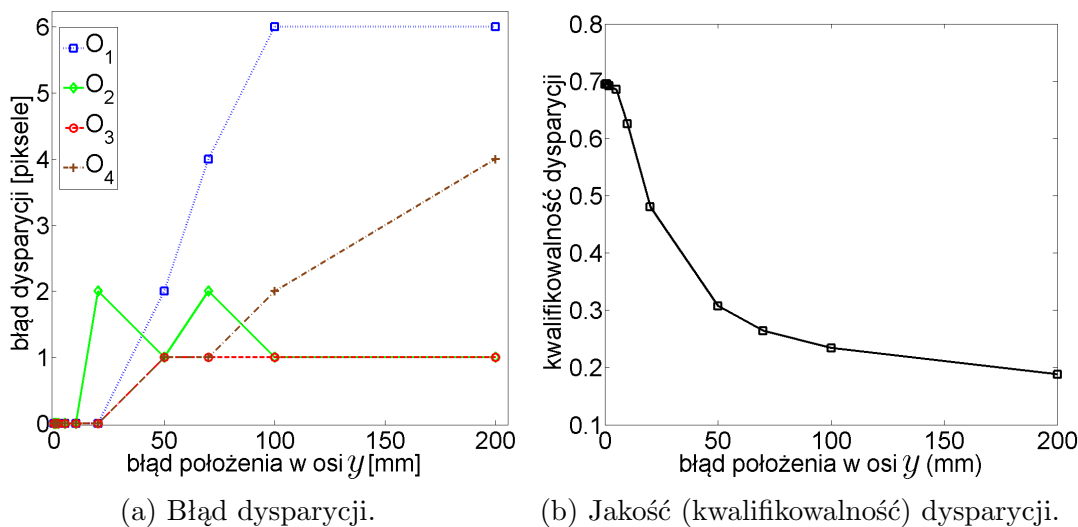


Rysunek 5.27 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zakresie błędu ułożenia względem osi  $y$  dla 4 obiektów.



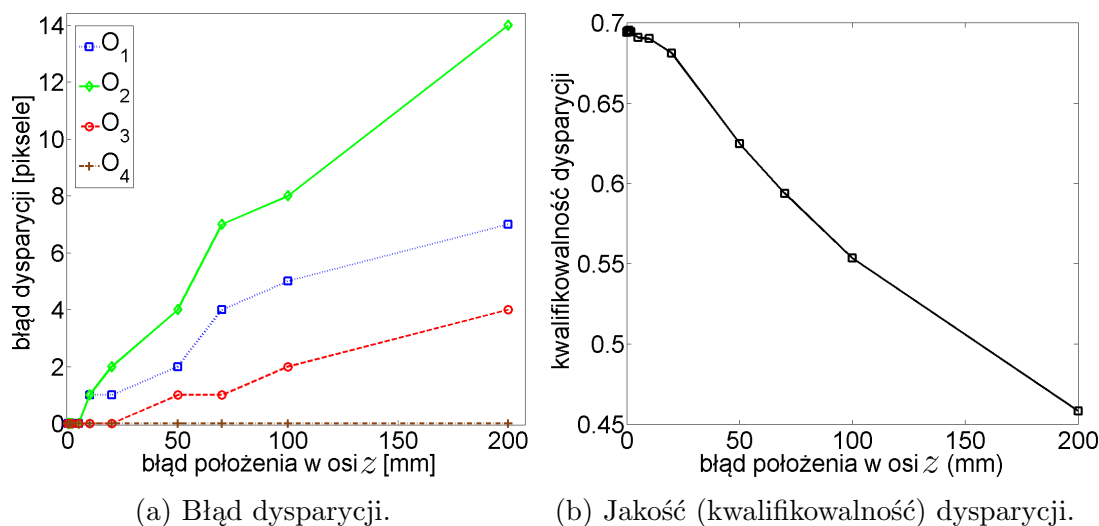
Rysunek 5.28 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zakresie błędu ułożenia względem osi  $z$  dla 4 obiektów.

uznać, że przy błędzie położenia względem osi  $y$  nieprzekraczającym  $50\text{mm}$  mapa jest poprawna i zawiera niewielki błąd.



Rysunek 5.29 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w funkcji błędu położenia w kierunku  $y$  dla 4 obiektów.

W przypadku reprezentowanym przez rys. 5.30a, przeanalizowano wpływ błędu położenia w kierunku  $z$  na błąd dysparycji. Wartości błędu są także niewielkie w porównaniu z błędami spowodowanymi rotacją kamery. Ponadto zaobserwować można podobieństwo tych krzywych dla wszystkich czterech obiektów. Poziom błęd dysparycji rośnie (odcinkami) praktycznie liniowo, a tempo wzrostu zależy od odległości od środka kadru, zaś kwalifikowalność dysparycji maleje wraz ze wzrostem błędu położenia (a nawet przy największym zakresie błędu widoczny jest wysoki poziom kwalifikowalności).



Rysunek 5.30 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu położenia w kierunku  $z$  dla 4 obiektów.

Podsumowując, wyniki przeprowadzonych testów wrażliwości stereowizji dla wirtualnej sceny są w ogólności zgodne z analizą przedstawioną w sekcji 4.4.3. Co ważne, przesunięcie w pionie (wzdłuż osi  $y$ ) oraz pochylenie każdej z kamer (obrót względem osi  $x$ ) utrudnia określenie rzutu tego samego punktu w obu obrazach i określenie dysparycji, choć testy nie wskazują na istnienie wartości progowej błędu, powyżej której algorytm SGBM działa niepoprawnie.

Analizując rys. 5.26 - 5.30 widać jednakowoż, że błąd ułożenia względem osi  $x$  już przy wartości  $0.3^\circ$  skutkuje istotnym brakiem w zbiorze pomiarów. Podobnie przy błędzie położenia przekraczającym  $50mm$ , także utrudniającym określenie dysparycji (w algorytmie SGBM), mapa zawiera mało punktów pomiarowych. Błędy ułożenia i położenia względem osi  $z$  wpływają negatywnie na dysparycję w niewielkim stopniu. Natomiast błąd ułożenia względem osi  $y$  umiarkowanie pogarsza kwalifikowalność mapy, jednak w bardzo dużym stopniu zwiększa błąd dysparycji. Oczywiście podane progowe wartości błędów zależą od ogniskowej (zgodnie ze wzorem (3.39)), która w tym przypadku ma wartość  $f = 3.5mm$ .

### *Wnioski*

Analizując powyższe wnioski można stwierdzić, że błędy ułożenia i położenia powinny być ograniczone odpowiednio do  $0.3^\circ$  oraz  $50mm$ . Natomiast niespełnienie tych warunków skutkuje mapą, która ma małą kwalifikowalność i istotny błąd dysparycji. Należy podkreślić, że na błędy ułożenia (w szczególności względem osi  $y$ ) najbardziej narażone są oceny obiektów odległych, w mniejszym stopniu obiekty oddalone od środka kadru. Najmniejszy błąd mapowania występuje bowiem dla elementów bliskich znajdujących się w środku kadru.

Testy z użyciem wirtualnej sceny pozwoliły na precyzyjny pomiar wrażliwości. Choć renderowana scena nie odzwierciedla w pełni warunków oświetleniowych, szumu kamery, zniekształceń, itp., wobec braku narzędzi umożliwiających precyzyjne pozycjonowanie kamery, w kolejnej sekcji udało się jednak sprawdzić, jaka może być wrażliwość stereowizji w rzeczywistych warunkach.

#### **5.3.2. Wrażliwość systemu stereowizyjnego dla sceny rzeczywistej**

Pomiarów wrażliwości stereowizji dokonano również w pomieszczeniu  $5 \times 5m$ , w którym ulokowano 4 obiekty w różnej odległości od układu kamer. Te obiekty to:  $O_1$  – prostopadłościan (kolorowe opakowanie),  $O_2$  – lampa na statywie,  $O_3$  – krawędź parawanu,  $O_4$  – kolumna ścienna. Dla obserwowanych obiektów wybrano punkty pomiarowe na krawędziach każdego obiektu. Odległość każdej kamery od danego punktu pomiarowego została zmierzona dalmierzem laserowym, którego dokładność wynosi  $2mm$ . Należy zaznaczyć, że pomiar w rzeczywistości ma nieco mniejszą dokładność ze względu na ruchomy punkt zamocowania dalmierza. Odległości obiektów od kamer pomierzone za pomocą dalmierza przedstawiono w tab. 5.3. Stanowiły one wartości referencyjne do badania wrażliwości stereowizji.

Dodatkowo dokonano wielokrotnego pomiaru środka obiektu, którego wartość średnia najlepiej oddaje odległość obiektu od kamery. Ogólną statystykę pomiaru (w postaci wartości średnich) przedstawiono w tab. 5.4.



W ramach testów sprawdzono wpływ rotacji platformy (prawej kamery) względem osi  $x$  oraz  $y$ , a także wpływ translacji w kierunku  $y$  oraz  $z$ . Ze względu na ograniczoną precyzję ustawiania platformy, nie zbadano wpływu małych obrotów względem osi  $z$ . Warto zaznaczyć, że z powodu braku precyzyjnych narzędzi, pomiary wykonano tylko dla kompletnej jednostki nośnej (wraz z kamerą). Zarówno obrót, jak i przesunięcie wykonywano poprzez ruch śrub (ze skokiem gwintu równym  $1mm$ ) umieszczonych w odpowiednich otworach dolnej platformy widocznej na rys. 5.18. Nie można zatem utożsamiać obrotu całej platformy z obrotem kamery, gdyż obrotem zawsze w badaniach towarzyszyła niewielka translacja kamery. Na wykresach 5.31a - 5.34 zaprezentowano wyniki dla 4 obserwowanych obiektów rzeczywistych.

### Ułożenie

Rys. 5.31 dotyczy przypadku, kiedy rotacji wokół osi  $x$  towarzyszy przesunięcie, a mianowicie przy obrocie  $0.7^\circ$  powstaje niewielka translacja w kierunku  $y$  oraz  $z$ , odpowiednio  $0.5mm$  oraz  $0.45mm$ . Natomiast większa rotacja, np.  $2.1^\circ$  skutkuje też przesunięciem  $3mm$  w kierunku  $y$  oraz  $2.7mm$  w kierunku  $z$ . Podobnie w przypadku obrotu wokół osi  $z$  zilustrowanego na rys. 5.32, wraz z obrotem  $0.5^\circ$  względem osi  $z$  występuje przesunięcie  $0.6mm$  w kierunku  $x$  oraz (około)  $0.002mm$  w kierunku  $y$ , zaś przy obrocie  $2^\circ$  powstaje dodatkowe przesunięcie o  $2.5mm$  w kierunku  $x$  oraz około  $0.04mm$  w kierunku  $y$ .

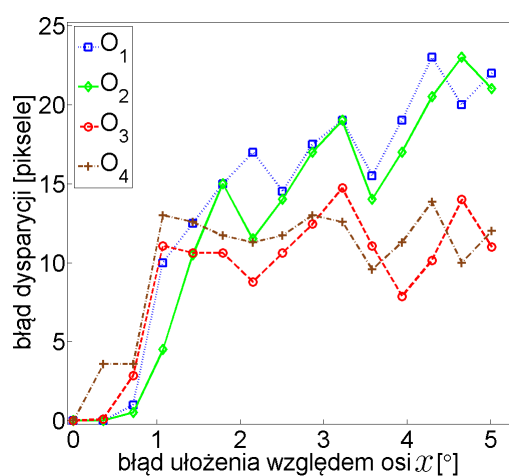
Zależność błędu dysparycji oraz jej kwalifikowalności od błędu ułożenia (prawej) kamery względem osi  $x$  jest zobrazowana na rys. 5.31. Poziom błąd dysparycji dla większych wartości błędu ułożenia jest zbliżony do poziomu odpowiedniego pomiaru dla sceny wirtualnej (rys. 5.27). Jednakże zagregowane błędy pomiarowe skutkują nagłym wzrostem błędu dysparycji dla niewielkiego błędu ułożenia. Ponadto dla większych wartości błędu ułożenia można zaobserwować niemonotoniczność błędu dysparycji, podobny dla praktycznie wszystkich badanych obiektów. Wykres kwalifikowalności dysparycji jest podobny do wyniku uzyskanego dla sceny wirtualnej. W tym przypadku niedokładność pomiaru skutkuje mniejszą kwalifikowalnością dla większych błędów ułożenia. Warto też zaznaczyć, że zawartość analizowanych scen, rzeczywistej i wirtualnej, były inne, zatem nie należy szczegółowo porównywać wyników kwalifikowalności dysparycji (w tym wartości maksymalnej i minimalnej).

Tablica 5.3 Wyniki pomiaru odległości badanych czterech rzeczywistych obiektów.

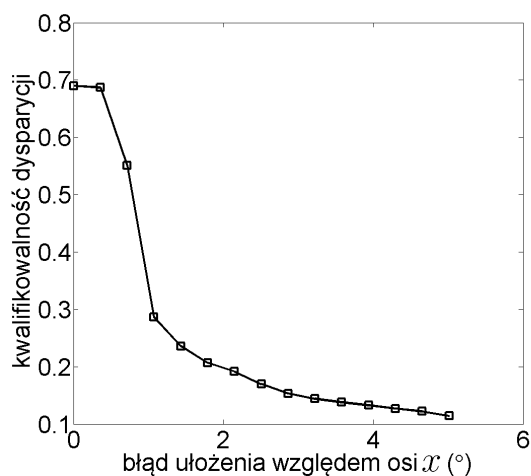
	$O_1$	$O_2$	$O_3$	$O_4$
lewy górny [m]	1.461	2.105	3.038	4.831
lewy dolny [m]	1.425	2.14	2.972	4.737
prawy górny [m]	1.446	2.09	3.018	4.853
prawy dolny [m]	1.413	2.13	2.964	4.73

Tablica 5.4 Średni pomiar odległości środka badanych obiektów od kamery.

	$O_1$	$O_2$	$O_3$	$O_4$
odległość [m]	1.4366	2.112	2.9818	4.787
rozrzut [m]	0.048	0.05	0.121	0.123
dysparycja [piksele]	62.6	42.6	30.1	18.8



(a) Błąd dysparycji.



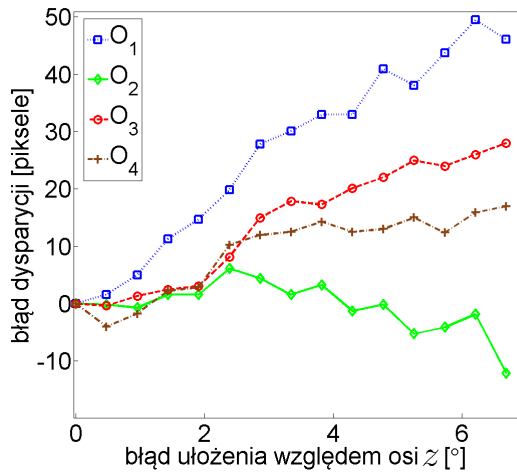
(b) Jakość (kwalifikowalność) dysparycji.

Rysunek 5.31 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w funkcji błędów ułożenia w osi  $x$  dla 4 rzeczywistych obiektów.

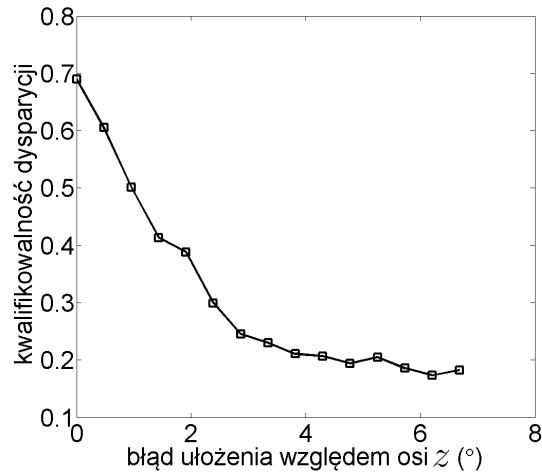
Zgodnie z prawami geometrii, testy przeprowadzone w wirtualnej przestrzeni wykazały, że błąd ułożenia (obrót) względem osi  $z$  skutkuje błędem dysparycji, który jest tym większy im dalej od środka ekranu znajduje się dany obiekt. Wyniki analogicznego pomiaru rzeczywistej sceny zobrazowano na rys. 5.32. Błąd jest mały dla obiektów w środku kadru, a dużo większy (i rośnie gwałtownie z błędem ułożenia) dla obiektów położonych blisko krawędzi kadru. Niemniej różnorodny rozkład rzeczywistych błędów pomiarowych powoduje wahania. Krzywa kwalifikowalności dysparycji prawie monotonicznie maleje wraz ze wzrostem błędów ułożenia w sposób analogiczny do innych przypadków.

### Położenie

Przy badaniu wrażliwości stereowizji względem przesunięcia, udało się wykonać pomiar z dokładnością rzędu  $0.3\text{mm}$ . Jednakże wskutek wpływu dodatkowej (obocznej) niedokładności ułożenia kamery podczas pomiaru, nie można wyizolować (wzrostu) błędów dysparycji wynikających z samego (wzrostu) błędów położenia. Analizując błędy położenia wzdłuż osi  $y$  (rys. 5.33) można wnioskować, że wpływ dodatkowego (obocznego) błędów



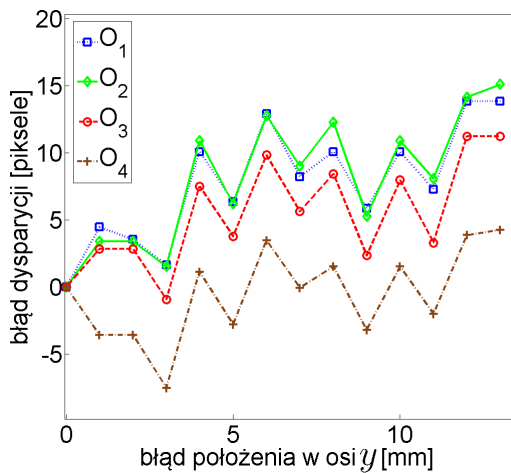
(a) Błąd dysparycji.



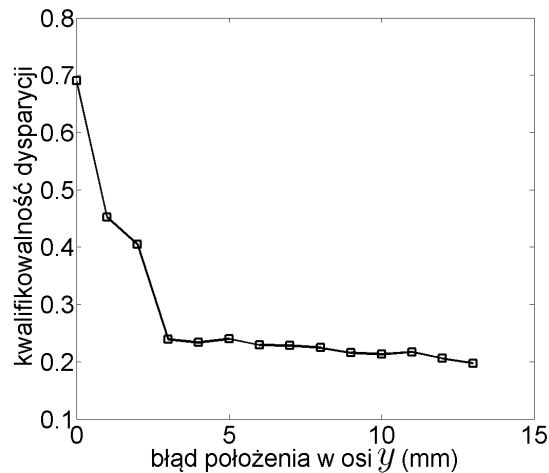
(b) Jakość (kwalifikowalność) dysparycji.

Rysunek 5.32 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu ułożenia względem osi  $z$  dla 4 rzeczywistych obiektów.

jest większy od błędu położenia, o czym świadczy duża amplituda wahań błędu dysparycji podobna dla wszystkich badanych obiektów. Natomiast kwalifikowalność dysparycji wskazuje, że błąd położenia poniżej  $3\text{mm}$  (przy ograniczonym błędzie ułożenia) skutkuje akceptowalnym poziomem kwalifikowalności. Można też wnioskować, że z obocznym błędem ułożenia związać można pewien próg błędu, powyżej którego algorytm SGBM nie znajduje już podobieństwa między punktami w lewym obrazie i ich odpowiednikami w prawym obrazie kamery.



(a) Błąd dysparycji.



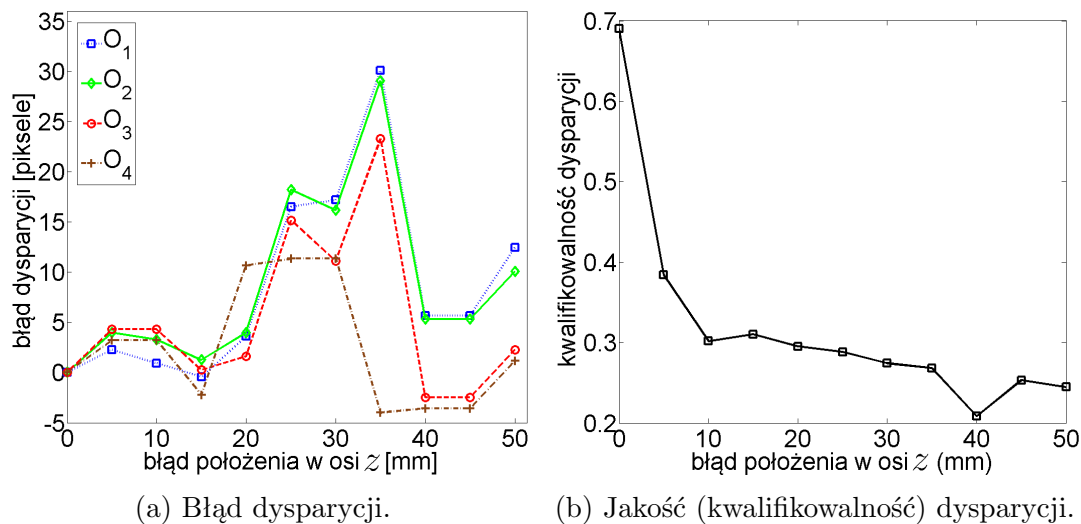
(b) Jakość (kwalifikowalność) dysparycji.

Rysunek 5.33 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu położenia w kierunku  $y$  dla 4 rzeczywistych obiektów.

Jak wykazano poprzez badania dokonane na wirtualnej scenie, błąd dysparycji rośnie wraz ze wzrostem błędu położenia w osi  $z$  w zależności od położenia obiektu w kadrze.



Wraz ze wzrostem błędu położenia kwalifikowalność dysparycji maleje bardzo łagodnie i nie osiąga wartości mniejszej niż 0.45 w badanym zakresie. W pomiarze rzeczywistej sceny (odzworowanym na rys. 5.34), wskutek wpływu obocznych błędów ułożenia, nie obserwuje się prostego wzrostu błędu dysparycji wraz ze wzrostem błędu położenia w osi  $z$ . Natomiast porównując kwalifikowalność dysparycji przy błędach położenia wzdłuż osi  $y$  i  $z$  (odpowiednio rys. 5.33b i rys. 5.34b) zauważyć można, że stereowizja jest mniej wrażliwa na błędy położenia w osi  $z$  (patrz zakresy zmienności błędu położenia), podobnie jak przy pomiarze sceny wirtualnej.



Rysunek 5.34 Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu położenia w kierunku  $z$  dla 4 rzeczywistych obiektów.

### Wnioski

Mimo występowania dużych błędów pozycjonowania kamery, badanie wrażliwości stereowizji dla sceny rzeczywistej pozwala ocenić wpływ błędów położenia oraz ułożenia (przy zjawisku obocznego niewielkiego błędu położenia) na proces mapowania. Biorąc odpowiednią poprawkę na te błędy uznać można, że wykonane pomiary dla sceny rzeczywistej i wirtualnej są zgodne. Stereowizja jest ogólnie bardzo wrażliwa na błędy ułożenia. Aby otrzymać mapę wysokiej jakości (o dużej kwalifikowalności) i błędzie dysparycji rzędu 1 piksela, konieczna jest precyzja ułożenia rzędu  $0.1^\circ$  i położenia rzędu  $10\text{mm}$ . Względnie dobrą mapę można otrzymać nawet przy dokładności ułożenia  $0.3^\circ$  i położenia  $50\text{mm}$ .

Implementacja systemu VisRobot oparta na niedrogich urządzeniach pomiarowych pozwala uzyskać jedynie dokładność ułożenia rzędu  $1^\circ$  i położenia rzędu  $8\%$  (opcja z progowaniem) w stosunku do szerokości bazy, co przedstawiono odpowiednio w sekcjach 5.2.1 oraz 5.2.2. Warunki te pozwalają na uzyskanie mapy niskiej jakości ze względnie dużym błędem dysparycji. Zatem poprawa mapy związana z użyciem zmiennej bazy wystąpi przede wszystkim dla punktów, których wartość dysparycji (dla pierwotnej bazy) jest bliska 0.

Przykładowo dla analizowanej wirtualnej sceny, przy początkowej bazie  $1\text{cm}$  obiekt  $O_4$  ma dysparycję równą 2. Błąd pomiaru związany z kalibracją to co najmniej  $\pm 0.25$

Tablica 5.5 Zakresy dysparycji związane ze zmienną bazą użyte w wirtualnych scenach.

	$B_0$	$B_1$	$B_2$	$B_3$
szerokość bazy	5cm	15cm	45cm	135cm
zakres dysparycji	[4, 34]	[12, 102]	[36, 305]	[108, 913]

piksela, zaś błąd dyskretyzacji to  $\pm 0.5$  piksela (co opisano w sekcji 3.4.3). Dlatego faktyczna dysparycja zawiera się w przedziale  $\langle 1.25, 2.75 \rangle$  piksela (czyli  $2 \pm 0.75$ ), co odpowiada przedziałowi odległości  $\langle 7.80, 17.15 \rangle m$  zgodnie z wzorem (3.39). Dla bazy 8cm obiekt  $O_4$  ma dysparycję równą 19, więc faktyczna dysparycja  $O_4$  zawiera się w przedziale  $\langle 18.25, 19.75 \rangle$ , któremu odpowiada przedział odległości  $\langle 8.69, 9.40 \rangle m$ . Zatem przy założeniu, że błąd ułożenia wynosi  $0.5^\circ$ , to według przeprowadzonej analizy wrażliwości, błąd dysparycji nie przekracza około 5 pikseli. Faktyczna dysparycja obiektu  $O_4$  zawiera się w przedziale  $\langle 24, 14 \rangle$ , co odpowiada odległości  $\langle 7.14, 12.26 \rangle m$ . Zatem przy szerszej bazie, błędzie ułożenia  $1^\circ$  i niewielkim błędzie położenia, odwzorowanie dalekiego obiektu jest dokładniejsze w porównaniu z pierwotną bazą. Oczywiście liczba punktów pomiarowych (kwalifikowalność dysparycji) będzie w takiej sytuacji mniejsza.

### 5.3.3. Dokładność mapowania wirtualnego otoczenia

Wirtualne środowisko jest bardzo przydatne, gdyż pozwala nie tylko na dokładne pozycjonowanie kamery w przestrzeni, ale także umożliwia dokładniejsze generowanie mapy głębi. Dzięki temu można też porównać dane pomiarowe z dokładniejszymi danymi odniesienia (dane *ground truth*). Alternatywą jest dostarczenie rzeczywistych danych odniesienia pochodzących, np. z zaawansowanego lidar. Jednak nawet lidar wysokiej precyzji nie zapewni dokładności mapowania choćby zbliżonej do wirtualnej mapy głębi.

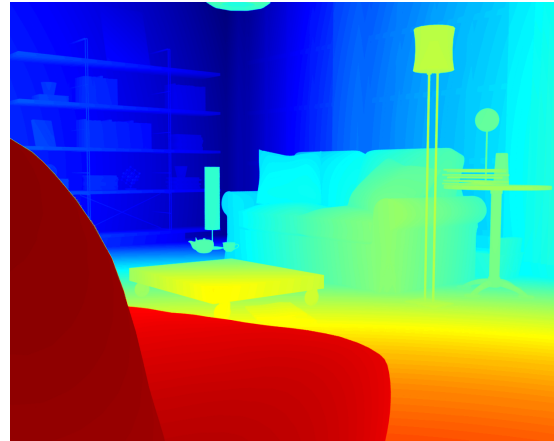
W ramach dysertacji dokonano ewaluacji stereowizji ze zmienną bazą wykorzystując środowisko wirtualne (mapy głębi). W tym celu przeprowadzono obliczenia dla wielu wirtualnych scen. Jednakże brak dostępności darmowych scen dedykowanych dla programu Blender, który wykorzystano we wcześniejszych obliczeniach, skutkowało koniecznością użycia programu 3D Studio Max z dodatkiem V-Ray, dla którego można już znaleźć liczne darmowe sceny.

Na początku procedury testowania umieszczono wirtualną kamerę, aby kadr obejmował możliwie duży obszar środowiska, a także aby umożliwić mapowanie ze zmienną bazą. Za pomocą wirtualnej linijki zmierzono odległość do najdalszego i najbliższego obiektu od kamery, aby ocenić maksymalną i minimalną użyteczną bazę. W wyznaczonym zakresie wybrano szerokości 4 baz ( $B_0 - B_3$ ), gdzie każda następna baza jest trzykrotnie szersza od poprzedniej. Przykładowe zakresy dysparycji dla jednej z analizowanych wirtualnych scen przedstawiono w tab. 5.5 (scena 7674 (Open3dmodel.com, 2015)). Oczywiście algorytm SGBM pozwala na analizę dysparycji w ograniczonym zakresie, co w praktyce (dla danej bazy) oznacza to, że wyjściowa dysparycja nie przekracza wartości 160 pikseli (co opisano w sekcji 4.4.1).





(a) Obraz wirtualnego otoczenia.



(b) Obliczona głębina odniesienia.

Rysunek 5.35 Renderowany obraz wirtualnego otoczenia i jego wyznaczona głębina odniesienia.

Jak wspomniano, wirtualna mapa głębi będąca odniesieniem w tym badaniu jest generowana w programie 3D Studio Max. Program normuje zakres głębi względem maksymalnej i minimalnej odległości sceny. Zatem wirtualna mapa głębi jest odpowiednio przeskalowana przed porównaniem jej z wynikami stereowizji. Przykładowy obraz wirtualnej kamery oraz wirtualna mapa głębi zostały zobrazowane na rys. 5.35. Dla zapewnienia czytelności głębię przedstawiono z użyciem odcieni nasyconych kolorów. Ciepłe kolory opisują małą głębnię (z minimum dla koloru czerwonego), zaś zimne kolory związane są z dużą głębnią (i maksimum dla koloru niebieskiego).

W ramach testu dokładności mapowania, porównano mapy głębi uzyskane dla baz  $B_0 - B_3$  z wirtualną mapą głębi. W rzeczywistej mapie głębi występują obszary o nieznannej głębni, które należy wykluczyć. Dokładność mapowania  $\Delta m$  obliczono jako sumę absolutnych różnic głębi

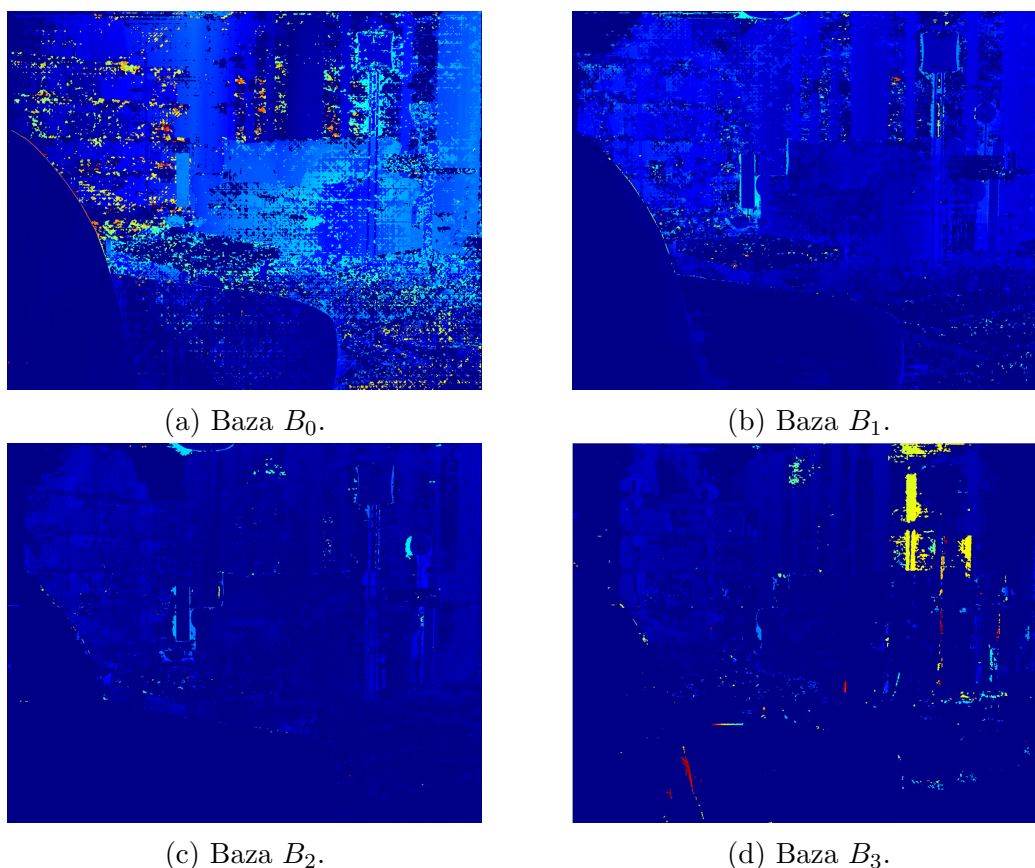
$$\Delta m = \sum_{x,y} |(d_v(x,y) - d_r(x,y))| \quad (5.7)$$

gdzie  $d_v$  i  $d_r$  to odpowiednio wirtualna i rzeczywista mapa głębi, tj. głębina punktów  $(x, y)$ , zaś  $x$  i  $y$  to poziome i pionowe współrzędne mapy głębi.

#### Przykładowa analiza stereowizyjna

Poza testem dokładności mapowania, dla każdej bazy zrealizowano wizualizację mapy różnicowej (błędu głębi), która reprezentuje różnicę między wirtualną mapą głębi a mapą głębi wyznaczoną za pomocą stereowizji. Dla czytelności użyto odcieni nasyconych kolorów. Im cieplejszy kolor tym większy błąd mapowania. Przykład map różnicowych dla sceny 7674, którą rozważano wcześniej, zaprezentowano na rys. 5.36.

Analizując rys. 5.36a można zaobserwować mały błąd mapowania dla bliskich obiektów i stosunkowo duży błąd mapowania obiektów dalekich. Dodatkowo w losowych miejscach można zauważyć małe obszary o bardzo dużych błędach mapowania (kolor żółty i czerwony). Natomiast w przypadku baz  $B_1$  i  $B_2$  występują tylko pojedyncze duże błędy,



Rysunek 5.36 Mapa dokładności mapowania dla czterech analizowanych baz

których na rys. 5.36b i 5.36c nie można wyraźnie przypisać obiektom o określonej odległości od kamery. Mapowanie przy największej bazie  $B_3$  charakteryzuje się najmniejszym błędem w obszarze obiektów dalekich (rys. 5.36d).

Warto zaznaczyć, że na rys. 5.36d żółte obszary są widoczne w miejscu, gdzie znajduje się okno. Nie oznacza to jednak wzrostu błędu mapowania dla względnie dalekich obiektów, a jedynie to, że powierzchnie przezroczyste lub odbijające światło są jednym z głównych źródeł problemów stereowizji. W tym przypadku algorytm SGBM poprawnie wyznaczył odległość wykraczającą poza pomieszczenie, przy czym wirtualna mapa odniesienia posługuje się faktycznym wymiarem pomieszczenia. Opisana sytuacja występuje jedynie przy dużej bazie, gdyż zakres mniejszych baz wyklucza możliwość mapowania obiektów znajdujących się daleko (poza pomieszczeniem). Zatem faktyczna dokładność mapowania otoczenia przy największej bazie jest nawet większa niż wskazywałaby to mapa różnicowa.

#### *Wyniki statystyczne*

Badanie dokładności mapowania zrealizowano dla 30 różnych (dostępnych i darmowych) scen wirtualnych. W każdym przypadku zapewniono jednakowe ustawienia wirtualnej kamery oraz dostosowano warunki oświetleniowe. Ze względu na różnice w rozmiarach scen pozycje wirtualnych kamer dobrano stosownie do odległości tak, aby minimalne wartości dysparycji dla każdej bazy pokrywały się z minimalnymi wartościami dysparycji

Tablica 5.6 Błąd mapowania dla czterech szerokości bazy uśredniony po 30 realizacjach.

baza	$B_0$	$B_1$	$B_2$	$B_3$
błąd mapowania	1	0.592	0.429	0.24

z tab. 5.5. Uśredniony błąd mapowania po 30 realizacjach dla baz  $B_0 - B_3$  przedstawiono w tab. 5.6. Błąd mapowania został dodatkowo unormowany względem największej wartości tego błędu. Zatem przy zmiennej bazie uśredniona dokładność mapowania przy bazie  $B_3$  wzrasta ponad czterokrotnie ( $1/0.24$ ) względem pierwotnej bazy. Jest to dowód potwierdzający tezę pracy, że zmienna baza poprawia jakość odwzorowania (co najmniej czterokrotnie), co należy uznać za istotny czynnik poprawy jakości mapowania. Poprawa ta dotyczy wszystkich punktów na mapie, których głębia jest określona.

Warto wspomnieć, że baza  $B_3$  w tym badaniu jest aż 27-krotnie większa od bazy  $B_0$ . Jednak nie uzyskuje się poprawy jakości mapowania tego rzędu, co może być spowodowane faktem, że całkowite błędy odwzorowania (występujące np. na krawędziach obiektów) pogarszają ocenę jakości mapy.

#### 5.3.4. Mapowanie wybranego otoczenia rzeczywistego

Przeprowadzono również badania, jak proponowany system działa dla przykładowego otoczenia rzeczywistego. Doświadczenie przeprowadzono w Laboratorium Integracji Systemów Automatyki wydziału ETI Politechniki Gdańskiej, gdzie znajduje się oczujnikowana makieta Pomorskiej Kolei Metropolitalnej (rys. 5.37). Ze względu na różnorodność ele-



Rysunek 5.37 Zdjęcie mapowanego pomieszczenia z makieta kolejową w Laboratorium Integracji Systemów Automatyki.

mentów oraz duże rozmiary makiety, możliwe było uzyskanie mapy bogatej w szczegóły

z uwzględnieniem dużych odległości. Aby zobrazować makietę, robota ustawiono na płaskiej powierzchni umieszczonej powyżej makiety. Efekt zmiennej bazy uzyskano poprzez przemieszczanie robota po powierzchni.

Warto zaznaczyć, że rozmiary pomieszczenia determinują dobór szerokości baz. Aby zaobserwować istotną poprawę jakości należy tak ustawić bazę  $B_0$ , aby otrzymać maksymalną dysparycję dla obiektów znajdujących się najbliżej kamer, zaś dla najdalszych obiektów - zerową (lub bardzo małą) dysparycję. Aby uzyskać opisany efekt w mapowanym pomieszczeniu konieczne byłoby ustawienie pierwotnej bazy o szerokości rzędu kilku mm. Taka konfiguracja nie jest jednak realizowalna dla podstawowego skalibrowanego układu stereo, gdzie obrazy z dwóch kamer są rejestrowane jednocześnie. Z drugiej strony realizacja małej bazy przez przesunięcie kamery i zapis pary obrazów w różnych chwilach czasu jest obciążona dużymi błędami ułożenia opisanymi w poprzedniej sekcji 5.3.2. Dlatego też doświadczenie mające na celu prezentację systemu, który łączy mapy dysparycji dla różnych baz stereo, zrealizowano dla większych szerokości baz, przy stałym błędzie pozycjonowania kamer.

### *Eksperymenty*

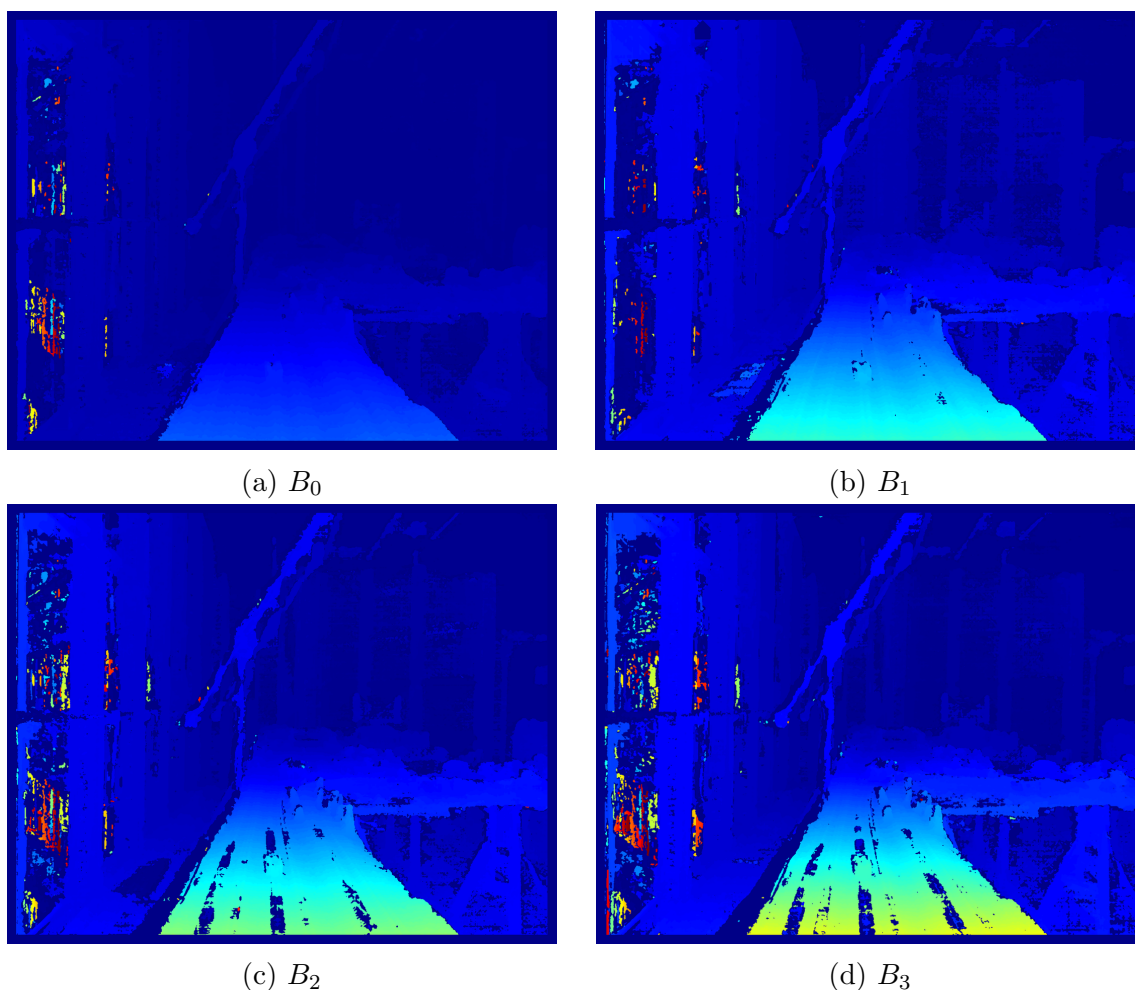
Elementy mapowanego pomieszczenia są widoczne na rys. 5.37, który jest obrazem z lewej kamery w początkowej pozycji. Eksperyment realizowano według schematu działania systemu VisRobot opisanego w sekcji 4.1, wykorzystując przy tym wyestymowane parametry modelu oraz filtrację Kalmana opisaną w sekcji 5.2.2. Zapisano obrazy z kamer dla 4 baz o szerokościach:  $B_0=1cm$ ,  $B_1=4cm$ ,  $B_2=8cm$ ,  $B_3=16cm$ . Mapy dysparycji wyznaczono dla wszystkich 4 baz. Wizualizację wyników (rys. 5.38) uzyskano w oparciu o moduł VIS. Dla zapewnienia czytelności zakres zmienności dysparycji przedstawiono w postaci odcieni nasyconych kolorów<sup>4</sup>.

Na rys. 5.38a można zaobserwować mapę dysparycji przy bazie  $B_0$ , która jest względnie gładka. W części obrazu, gdzie występuje makietka kolei nie ma istotnych obszarów o nieznannej dysparycji, jednakże zakres zmian barw jest mały, co świadczy o małej rozdzielczości wgłębnej. Torowisko znajdujące się blisko układu kamer można dostrzec porównując mapę dysparycji z oryginalnym kolorowym obrazem (rys. 5.37). Natomiast analogiczne porównanie obrazów w przypadku obiektów znajdujących się daleko od układu kamer jest w zasadzie niemożliwe. Rozdzielczość przestrzeni barw, której odpowiada rozdzielczość dysparycji, jest w tym przypadku zbyt mała.

Mapa dysparycji przy bazie  $B_1$  widoczna na rys. 5.38b jest także gładka. Porównując mapy utworzone przy bazach  $B_0$  i  $B_1$  zaobserwować można różnicę rozdzielczości przestrzeni barw, która odpowiada większej rozdzielczości dysparycji. Przy bazie  $B_1$  obiekty znajdujące się blisko układu kamer są rozróżnialne od obiektów odległych. Ponadto zwiększenie rozdzielczości wgłębnej pozwala na rozróżnienie (choć z trudnością) pojedynczych obiektów położonych względnie blisko układu kamer.

W przypadku map dysparycji przy bazach  $B_2$  i  $B_3$  (odpowiednio na rys. 5.38c i 5.38d) można zauważyć obszary o nieznannej dysparycji w torowisku (położone blisko układu kamer). Jest to najprawdopodobniej spowodowane powtarzającym się wzorcem, jaki stanowią tory. Przy większej bazie algorytm SGBM skutkuje błędną interpretacją, gdyż sąsiednie linie torów są zbyt podobne do siebie, co nie pozwala wyznaczyć dysparycji.

<sup>4</sup>Kolorem niebieskim opisujemy małą dysparycję reprezentującą również duże odległości.



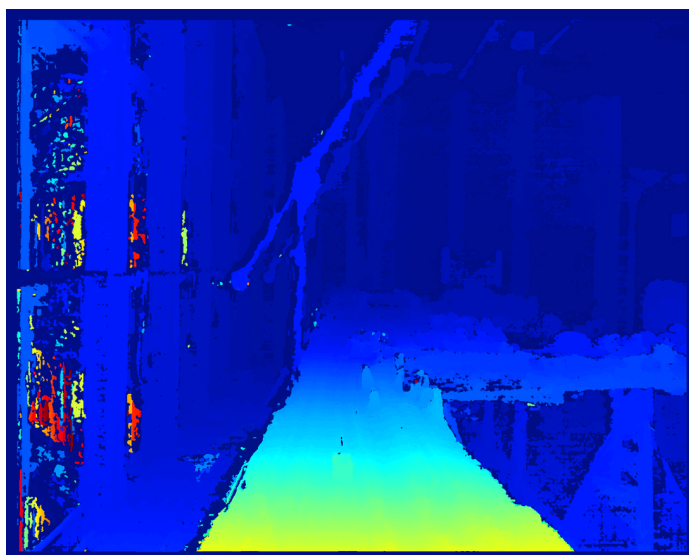
Rysunek 5.38 Mapy dysparycji przy czterech bazach stereowizyjnych.

Natomiast rozdzielczość przestrzeni barw oraz rozdzielczość dysparycji wyraźnie rosną wraz ze wzrostem szerokości bazy. Na rys. 5.38d rozdzielczość przestrzeni barw jest największa. Obserwacja obrazu pozwala na rozróżnienie obiektów znajdujących się względnie daleko od układu kamer. Podobnie można zinterpretować położenie niebieskiego wagonu (zaznaczonego w kółku na rys. 5.37) i pojedynczych drzew.

W systemie VisRobot złożenie map dysparycji przy bazach  $B_0 - B_3$  daje kompletną mapę zobrazowaną na rys. 5.39. Rozdzielczość przestrzeni barw jest taka sama jak w przypadku mapy przy bazie  $B_3$ , ale nie zawiera obszarów o nieznaney dysparycji, które są widoczne na rys. 5.38d. Przy innym doborze szerokości baz może wystąpić sytuacja, kiedy wynikowa mapa dysparycji ma większą rozdzielczość niż składowe mapy dysparycji. Jednakże w tym przypadku szerokości bazy dobrano tak, aby uniknąć negatywnego wpływu błędów związanych z niedokładnym położeniem i ułożeniem kamer. Przykład ten także potwierdza tezę, że proponowane rozwiązanie poprawia odwzorowanie otoczenia oparte na danych stereowizyjnych.

Finalną (złożoną) mapę dysparycji można wyświetlić operatorowi robota. Należy przy tym ograniczyć jej 'zakres' przez zakrycie pikseli, stosując (4.32). Takie podejście prowadzi do zakrycia bliskich elementów oraz obserwację tylko wystarczająco dalekich obiektów





Rysunek 5.39 Wynikowa mapa dysparycji.

Tablica 5.7 Ocena jakości głębi dla czterech baz.

baza	$B_0$	$B_1$	$B_2$	$B_3$
jakość głębi	0.45	0.73	0.65	0.62

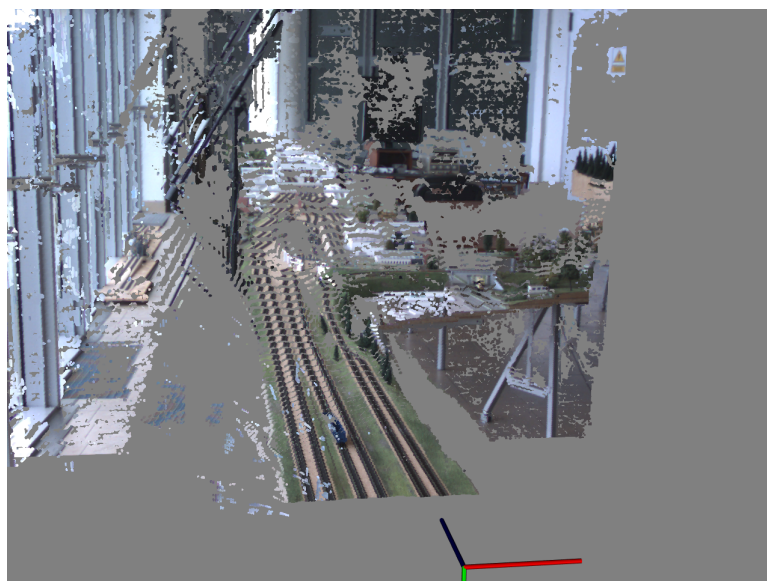
(z powiększoną rozdzielczością wgłębną). Aby wyświetlić optymalną głębię wybiera się optymalną bazę zgodnie z kryterium (4.33). W analizowanym eksperymencie optymalną jest baza  $B_1$ , co wynika z tab. 5.7.

Obserwując mapy dysparycji na rys. 5.38a - 5.38d zauważyć można, że ocena głębi przy bazie  $B_0$  jest niska ze względu na duże obszary (liczbę pikseli) o zerowej dysparycji  $d$  (nie widać konturów obiektów położonych daleko). W przypadku baz  $B_1 - B_3$  wszystkie najdalej położone obiekty mają niezerową dysparycję, co oznacza, że szerokość bazy  $B_1$  jest wystarczająca do odwzorowania głębi badanego pomieszczenia (oczywiście błąd odwzorowania obiektów dalekich przy bazie  $B_1$  jest większy w porównaniu z wynikami przy szerszych bazach). Natomiast przy szerokich bazach  $B_2$  i  $B_3$  występują już obszary o nieokreślonej dysparycji, co obniża ocenę jakości głębi. Wynik ten dokumentuje również użyteczność kryterium (4.33), które poprawnie ocenia jakość głębi.

Po złożeniu map dysparycji dla poszczególnych baz uzyskano chmurę punktów, której wizualizację przedstawiono na rys. 5.40. Na mapie widoczna jest oryginalna pozycja kamery na przecięciu osi czerwonej, zielonej i niebieskiej (w dolnej części obrazu), obiekty znajdujące się w otoczeniu, a także nieliczne błędne punkty, związane są np. z gładkimi obszarami i powtarzającym się wzorcem.

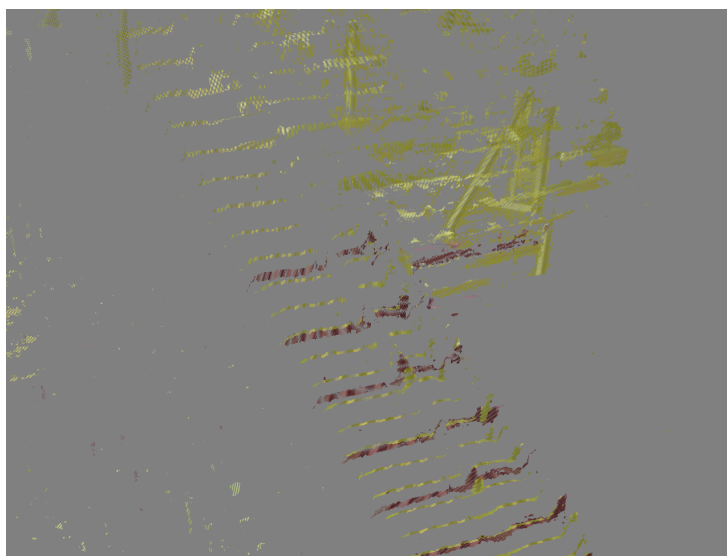
W ramach eksperymentu, odwzorowanie 3D (chmura punktów) ustawiono tak, aby wirtualna kamera dawała obraz zbliżenia na obiekty bliskie oraz widok z góry. Dzięki temu dokonano wizualizacji zmian rozdzielczości wgłębnej. Aby obrazy były czytelne zmodyfikowano kolorystykę zobrazowania tak, aby za pomocą wybranych kolorów wyizolować





Rysunek 5.40 Mapa 3D w systemie ze zmienną bazą.

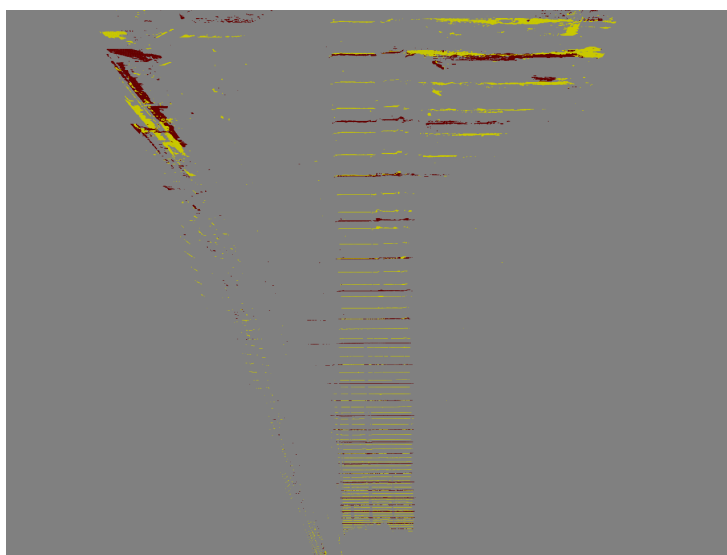
punkty związane z konkretną bazą. Dodatkowo zastosowano tu 50% poziom przezroczystości. Na rys. 5.41 i rys. 5.42 widoczne są fragmenty mapy, gdzie punkty związane z mniejszą bazą  $B_0$  zaznaczono kolorem czerwonym, zaś punkty związane z większą bazą  $B_1$  zaznaczono kolorem żółtym. Nietrudno zauważyć, że rozdzielczość pomiaru jest wyraźnie większa dla większej bazy, przez co można dokładniej określić położenie danego punktu na mapie.



Rysunek 5.41 Mapa 3D w rzucie perspektywicznym z wyróżnieniem punktów wyznaczonych dla 2 sąsiednich baz (czerwony  $B_0$ , żółty  $B_1$ ).

## Wnioski

Eksperyment rzeczywisty pozwala zaobserwować poprawność działania systemu VisRobot. Analiza poszczególnych map dysparycji i końcowej mapy daje możliwość wizualnej oceny poprawy jakości stereowizji przy użyciu zmiennej bazy. Ponadto wizualizacja rzutów mapy 3D z zaznaczonymi punktami związanymi z sąsiadującymi bazami unaocznia, jak zmienia się rozdzielczość głębi na skutek zmiany szerokości bazy. Widać, że przechodząc do większych baz uzyskujemy uwidatnienie dalekich obiektów.



Rysunek 5.42 Rzut poziomy (z góry) mapy 3D z wyróżnieniem punktów wyznaczonych dla 2 sąsiednich baz (czerwony  $B_0$ , żółty  $B_1$ ).

W ramach eksperymentu nie podjęto się kwantyfikowania stopnia polepszenia mapy ze względu na dostępną precyzję pomiarową pozycji kamer. Natomiast w środowisku wirtualnym dowiedziono, że jest ona istotna oraz w jakim stopniu wpływa na wyniki stereowizyjnej oceny głębi (dysparycji i kwalifikowalności), jak również wykazano, że w opisanym eksperymentalnym przypadku występuje czterokrotny wzrost dokładności.

## 5.4 Podsumowanie

Testy wrażliwości stereowizji w środowisku wirtualnym i rzeczywistym, a także obrazy mapowanego wybranego środowiska, pokazują że system ze zmienną bazą może być skuteczny w mapowaniu ograniczonego środowiska i daje możliwości zwiększenia zakresu mapowanego środowiska oraz rozdzielczości wgłębnej.

Przy wysokiej dokładności położenia i ułożenia kamer, system VisRobot istotnie poprawia jakość odwzorowania otoczenia. W przedstawionym teście w środowisku wirtualnym uzyskano poprawę co najmniej czterokrotną. W przypadku błędów typowych dla sprzętu o względnie niskim koszcie, stwierdzona duża wrażliwość stereowizji na błędy pozycjonowania kamer sprawia, że w systemie zwiększa się dokładność mapowania, ale tylko w ograniczonym zakresie dysparycji. Bez względu na błąd pozycjonowania kamer, system VisRobot ze zmienną bazą zwiększa zasięg mapowanego terenu w obszarach, w których dysparycja wynosi zero przy pierwotnej bazie.

Ocena jakości głębi w systemie VisRobot umożliwia określenie optymalnej bazy względem jakości mapy głębi zgodnie z kryterium (4.33). Dzięki temu ludzki operator robota może dokładniej (głębiej) obserwować analizowaną przestrzeń za pomocą zastosowanej techniki stereoskopowej. Przy szerokich bazach uzyskuje się w ten sposób efekt 'pogłębienia przestrzennego'.

W ramach prowadzonych badań przedstawiono także wyniki testów związanych z określeniem położenia i ułożenia w systemie VisRobot. Ze względu na silny negatywny wpływ szumów własnych akcelerometru oraz błędów usuwania składowej grawitacji z sygnału przyspieszenia oceniono wpływ filtracji sygnału na wynik podwójnego całkowania danych. Badania symulacyjne potwierdziły, że generalnie filtracja dolnopasmowa przyspieszenia nie wpływa na położenie (wyznaczone przez podwójne całkowania danych przyspieszenia). Przeprowadzono także testy wrażliwości na szum akcelerometru oraz błąd wynikający z usuwania składowej grawitacji, co pozwoliło na ilościowe określenie maksymalnych błędów, przy których system działa względnie poprawnie.

Wyznaczanie położenia platformy systemu VisRobot oparto na modelu dynamiki robota. Dzięki temu dane z akcelerometru stają się użyteczne przy wyznaczaniu położenia. Parametry modelu określono na drodze estymacji własną metodą laboratoryjną, która wykrywa położenie kolorowego znacznika (umieszczonego na robocie) w obrazie z kamery. Wyniki detekcji oraz estymacji parametrów zweryfikowano na drodze doświadczalnej.

Wyznaczone dane o położeniu oraz ułożeniu pochodzące z sensorów inercyjnych pozwalają na wyznaczenie trajektorii robota. Testy rzeczywiste pokazały, jaki jest typowy błąd położenia robota, wpływający bezpośrednio na położenie kamer w stereowizji, przy ograniczonej precyzji dostępnego sprzętu.

Reasumując, testy przeprowadzone i omówione w ramach niniejszej dysertacji dowodzą poprawności koncepcji oraz skuteczności działania systemu VisRobot. Analiza zebranych wyników pozwala na ocenę dokładności działania systemu, oraz determinuje, jakie warunki należy spełnić, aby uzyskać wysoką jakość trójwymiarowego odwzorowania otoczenia.



---

## ZAKOŃCZENIE

Problem podjęty w niniejszej dysertacji dotyczy trójwymiarowego odwzorowania otoczenia przez roboty mobilne, opiera się na stereowizji, z wykorzystaniem zmiennej bazy stereo. Ze względu na istnienie robotów mobilnych dostosowanych do różnych środowisk różnorodnych pod względem, budowy, napędu, sensorów itp., analizowany problem ograniczono do wybranego typu robota mobilnego. Zawężenie to wpływa oczywiście na możliwości pomiarowe oraz na tworzoną wynikową mapę 3D. Argumentacja przy wyborze platformy robota oraz metody podejścia do odwzorowania otoczenia została przedstawiona w ramach przeglądu rozwiązań opisanego w rozdziale 2. Natomiast w odniesieniu do szczegółowych metod i algorytmów kierowano się charakterystyką teoretycznych podstaw pomiarów stereowizyjnych omówioną w rozdziale 3.

### *6.1 Zrealizowane założenia*

Główny cel postawiony w pracy, czyli opracowanie metody odwzorowującej obiekty ograniczonego środowiska na trójwymiarowej mapie cyfrowej z wykorzystaniem robotów mobilnych zaopatrzonych w stereowizję, udało się osiągnąć dla wybranego rodzaju robota mobilnego. Autorskie laboratoryjne rozwiązanie stanowi złożoną procedurę opartą na algorytmie stereowizyjnym SGBM i zakłada zmienną bazę stereo, która uzyskiwana jest poprzez przemieszczanie robota. Dlatego znaczną część pracy poświęcono problemowi estymacji położenia i ułożenia robota, których dokładność jest szczególnie istotna, ponieważ silnie wpływa na precyzję odwzorowania otoczenia. Zaproponowana metoda oparta na filtracji Kalmana została sprawdzona zarówno w warunkach symulacyjnych, jak i rzeczywistych. W szczególności skupiono się na estymacji położenia w oparciu o dane akcelerometryczne oraz dynamiczny model ruchu robota.

Istotną motywacją do użycia akcelerometru było sprawdzenie użyteczności łatwo osiągalnych danych przyspieszenia do określania położenia robota, gdyż w literaturze istnieją liczne prace naukowe, z których część przeczy a część potwierdza użyteczność akcelerometrów w tym kontekście. Wyznaczanie pozycji w oparciu o dane akcelerometryczne wymaga podwójnego całkowania. Wskutek istnienia szumów pomiarowych, scałkowane dane obarczone są bardzo dużym błędem, dlatego też najpierw przeprowadzono analizę możliwości filtracji dolnopasmowej danych z akcelerometru. Ostatecznie studium wrażli-

wości oceny położenia pozwoliło na określenie maksymalnego poziomu szumu sygnałów akcelerometrycznych, przy którym całkowanie jest jeszcze użyteczne.

Estymacja położenia robota została w pracy zrealizowana za pomocą odpowiednio zaadaptowanej metody obserwatorowej opartej na filtracji Kalmana. Wykorzystując zidentyfikowany model dynamiki robota (na drodze estymacji parametrów), model akcelerometru, dane akcelerometryczne a także sygnał sterujący robotem (wymuszenie), odpowiednio zaprojektowany filtr umożliwił użyteczną estymację położenia robota mimo istnienia złożonych i silnych szumów pomiarowych. Metodę zweryfikowano w warunkach rzeczywistych z wykorzystaniem laboratoryjnej platformy sprzętowej. Dodatkowo parametry modelu robota użytego podczas testów wyznaczono autorską metodą estymacji opartą na przetwarzaniu danych z kamery CMOS (z wykrywaniem kolorowego znacznika).

Ewaluacja opracowanej koncepcji stereowizyjnej obejmowała zarówno rozważania analityczne, jak i symulacje w środowisku wirtualnym i rzeczywistym. Na drodze analitycznej wyznaczono wrażliwość stereowizji na błędy położenia i ułożenia robota, co pozwoliło określić, w jaki sposób błędy te propagują się przy tworzeniu mapy dysparycji, przy czym ostateczny skutek zależy od współrzędnych obrazu. Ewaluację wrażliwości stereowizji przeprowadzono także w środowisku wirtualnym i rzeczywistym. Tutaj też, analizując uzyskane wyniki, można zauważyć wyraźną zależność błędu odwzorowania od współrzędnych ekranu. Błędy algorytmu SGBM i inne problemy towarzyszące porównywaniu obrazów skutkowały jednak niemonotonicznym przebiegiem (wahaniem) błędu dysparycji w funkcji błędu położenia lub ułożenia. Dodatkowo, analiza wrażliwości pokazała, jak określoność dysparycji, z którą związana jest jakość mapy, zmienia się wraz ze wzrostem błędu położenia i ułożenia. Analogiczne testy przeprowadzono w środowisku rzeczywistym, jednak w tym przypadku nieliniowości i ograniczenia sprzętowe powodują bardzo duże nieregularności widoczne w ocenie błędu dysparycji. Jednak przeprowadzona analiza ogólnie pozwoliła określić wymagania nakładane na dokładność pozycji robota dla potrzeb realizowanej stereowizji.

Testy dokładności procedury mapowania ze zmienną bazą stereo zrealizowano dla otoczenia wirtualnego i rzeczywistego. Dla wybranych 30 wirtualnych scen, w każdym przypadku wyznaczono wirtualną mapę głębi (jako mapę odniesienia) oraz dla czterech wybranych baz określono odrębne mapy głębi. Dzięki porównaniu tych map z mapą odniesienia można było określić błąd odwzorowania. Na tej podstawie zweryfikowano ilościowo, w jakim stopniu wykorzystanie zmiennej bazy poprawia docelowe odwzorowanie otoczenia. Analogiczne testy w środowisku rzeczywistym nie pozwoliły na dokładną ocenę jakości odwzorowania otoczenia. Niemniej, testy te pokazały użyteczność opracowanego systemu stereowizyjnego ze zmienną bazą.

## **6.2 Odniesienie do tej pracy**

Cele pracy zrealizowano w oparciu o zdefiniowane główne tezy badawcze oraz dwie tezy pomocnicze. Treść tezy głównej oznacza, że odpowiednie przetwarzanie informacji z sensorów inercyjnych oraz kamer stereoskopowych umieszczonych na robotach mobilnych pozwala na adekwatne odwzorowanie otoczenia na mapie 3D. W pracy wykazano, że opracowany system pozwala odwzorować otoczenie na mapie trójwymiarowej. Dokładność odwzorowania została oceniona i zweryfikowana w warunkach symulacyjnych i rze-

czywistych. Można dzięki temu stwierdzić, że otoczenie jest odwzorowywane adekwatnie do precyzyjnego pozycjonowania kamer (położenia i ułożenia) opartego na danych inercyjnych oraz ograniczeń stereowizji, co potwierdza tezę główną.

Pierwsza teza pomocnicza jest związana z użytecznością koncepcji zmiennej bazy przy odwzorowaniu otoczenia. Teza ta zakłada, że istotnym czynnikiem poprawy jakości w systemie stereowizyjnym jest zastosowanie obrazów z kamer o zmiennej bazie stereoskopowej. W systemie VisRobot użyto zmiennej bazy przez kontrolowane przemieszczenie robota z kamerami, przez co akwizycja lewego i prawego obrazu (pary stereo) jest realizowana w różnych lokalizacjach robota. Na podstawie symulacji i eksperymentów rzeczywistych stwierdzono, że zmienna baza pozwala zwiększyć użyteczny zakres pomiarowy. Eksperymenty potwierdziły co najmniej czterokrotny wzrost dokładności dla badanych 30 środowisk wirtualnych. Skutkiem zaproponowanej metody jest także odwzorowanie dalszego otoczenia, które zwykle jest poza zakresem pomiarowym pojedynczej (mniejszej) bazy stereo.

Metoda zmiennej bazy zwiększa rozdzielczość wglębną, co pozwala na dokładniejsze odwzorowanie otoczenia i poprawę jakości systemu stereowizyjnego. Poprawa ta wystąpi, jeśli spełnione będą dwa warunki. Po pierwsze algorytm stereowizyjny (SGBM) musi wykryć odpowiednie punkty w lewym i prawym obrazie, które są odpowiednimi rzutami punktu 3D. Po drugie błąd położenia i ułożenia kamer musi być wystarczająco mały.

W opracowanym systemie VisRobot pierwsza (najmniejsza) baza stereo jest realizowana dla dobrze skalibrowanej pary obrazów (typowe zastosowanie stereowizji). Natomiast kolejne większe bazy są tworzone z większym błędem położenia i ułożenia. Zatem taka konfiguracja daje możliwość bezpośredniego porównania standardowej stereowizji i systemu ze zmienną bazą. Na drodze eksperymentów wykazano, że błędy ułożenia i położenia nie przekraczające odpowiednio  $0.3^\circ$  i  $50mm$ , przy ogniskowej  $3.5mm$ , dają istotną poprawę jakości systemu stereowizyjnego. Dla większych błędów, nie przekraczających jednak progów  $1.0^\circ$  i  $200mm$ , odpowiednio dla ułożenia i położenia, występuje nieznaczna poprawa jakości odwzorowania. Dla większych błędów w ogólności należy przyjąć, że dochodzi do poszerzenia zakresu pomiarowego, ale błąd odwzorowania jest duży. Zatem teza ta jest potwierdzona warunkowo (z ograniczeniami).

Warto podkreślić, że błąd odwzorowania nie jest stały dla obiektów widocznych na obrazie. W zależności od tego, który błąd (ułożenia czy położenia) jest dominujący (względem osi  $x$ ,  $y$ ,  $z$ ) oraz jakie jest położenie obiektu w kadrze, wynikowy błąd odwzorowania może istotnie się różnić. Odpowiednie przypadki zostały wyznaczone analitycznie oraz sprawdzone na drodze symulacji i rzeczywistych eksperymentów.

Druga teza pomocnicza sugeruje, że istnieje optymalne ustawienie bazy obrazów stereoskopowych pod względem możliwości poprawnej oceny głębi obrazu. W opracowanym systemie VisRobot zdefiniowano kryterium oceny jakości głębi (4.33), co pozwoliło porównać jakościowo mapy głębi przy różnych bazach oraz znaleźć najlepszą mapę głębi, którą można zobrazować operatorowi robota. W tym celu zaprojektowano moduł wizualizujący użytkownikowi obraz 3D, co zawarto w sekcji 4.5. Moduł ten wyświetla tylko te fragmenty par obrazów, które - zgodnie z regułą  $1/30$  - statystyczny człowiek może zobaczyć przy użyciu dostępnych technik stereoskopowych. Przy zbyt małej bazie użytkownik nie odczuje przestrzenności obrazu 3D, zaś przy zbyt dużej bazie pojawiają się obszary nieprawidłowe. Zatem w kontekście drugiej tezy pomocniczej istnieje optymalny dobór bazy, przy którym suma zakrytych pikseli (o zbyt dużej dysparycji) i liczba pikseli o

zerowej dysparycji jest najmniejsza. W praktyce nawet dla dużej bazy wciąż może występować wiele punktów o zerowej dysparycji (kiedy odległość punktu od bazy przekracza przewidywany zakres pomiarowy). Optymalne ustawienie można wyznaczyć dopiero po wyznaczeniu map dysparycji dla całego dostępnego zakresu baz. W systemie VisRobot optymalną mapę głębi określa się na podstawie dostępnych kilku (np. 8) baz. W praktyce im większa jest liczba pomiarów (baz), tym większa szansa na znalezienie najlepszej mapy głębi.

### **6.3 Główne osiągnięcia pracy**

W ramach niniejszej pracy rozwiązano wiele problemów naukowych i technicznych. Do najważniejszych osiągnięć należy zaliczyć:

- odwzorowanie otoczenia na trójwymiarowej mapie cyfrowej w oparciu o sterowizję ze zmienną bazą
- ewaluację systemu VisRobot w warunkach symulacyjnych i rzeczywistych
- wykazanie, że zmienna baza stereowizyjna pozwala na istotną poprawę dokładności odwzorowania otoczenia
- zbadanie wrażliwości stereowizji i określenie na tej podstawie wymagań precyzji pozycjonowania kamer oraz robota
- estymację ułożenia robota na podstawie fuzji danych inercyjnych
- estymację położenia robota na podstawie danych ułożenia, danych akcelerometrycznych oraz sygnału wymuszenia
- potwierdzenie użyteczności akcelerometru przy wyznaczaniu położenia – na podstawie badań symulacyjnych i rzeczywistych
- zoptymalizowaną dla ludzkiego operatora wizualizację obrazu stereoskopowego w systemie stereowizyjnym ze zmienną bazą.

### **6.4 Krytyczna analiza pracy**

Badania przeprowadzone w ramach niniejszej pracy pokazały wielką wrażliwość stereowizji na błędy ułożenia. Wykorzystanie fuzji danych z sensorów inercyjnych i uzyskanie ułożenia z małym błędem, który nie przekracza  $1.0^\circ$ , nie wystarczyły do otrzymania zadowalającej poprawy jakości odwzorowania. Warto też podkreślić, że brak precyzyjnych narzędzi pomiarowych wymagał wprowadzenia wielu kompromisowych rozwiązań procedury badawczej. Również jakość dostępnego sprzętu w dużym stopniu decydowała o przyjęciu określonych założeń przy projektowaniu opracowanego systemu VisRobot i jego praktycznej implementacji.

Planując dalszą realizację projektu w oparciu o wiedzę na temat wrażliwości stereowizji i ograniczeń sprzętowych, należałoby rozważyć i wprowadzić technikę SfM, co mogłoby wpłynąć pozytywnie na dokładność położenia i ułożenia. Oczywiście dopiero po przeprowadzeniu odpowiednich testów można będzie stwierdzić, w jakim stopniu technika ta poprawi dokładność pozycji robota. Zwiększenie przewidywanych kosztów może wiązać się z potrzebą włączenia większej mocy obliczeniowej oraz zastosowania nowej, cięższej platformy robota mobilnego z rozbudowanymi systemami.

W celu uniknięcia doświadczonych ograniczeń metodologicznych i sprzętowych, należałoby zwiększyć precyzję określania ułożenia, na przykład przez zastosowanie droższych i dokładniejszych sensorów inercyjnych. Aby uzyskać znaczącą poprawę (o kilka rzędów



wielkości), pozycjonowanie można oprzeć na multilateracji z użyciem większej liczby robotów mobilnych. Każdy robot mógłby być zaopatrzony w znacznik i dedykowany system optyczny, który za pomocą lasera precyzyjnie określa ułożenie i położenie (odległości między robotami). Przewidywany koszt wytworzenia takiego robota może jednak być wielokrotnie większy i wiązać się z gruntownymi zmianami konstrukcji robotycznej.

Na zakończenie warto zauważyć, że problem odwzorowania otoczenia można także rozwiązać korzystając z metod aktywnych, jak Kinect lub droższy lidar. Obecnie doświadczamy obecności wielu aktywnych, radiowych systemów, które zakłócają inne systemy, oraz generalnie działają negatywnie na organizmy żywe. Wybór rozwiązania pasywnego, takiego jak stereowizja był konsekwencją takiego poglądu. Przez setki tysięcy lat natura wypracowała pasywny narząd wzroku dla wszystkich organizmów poruszających się podczas dnia w środowisku lądowym. W metody aktywne zaopatrzone są organizmy żyjące przede wszystkim bez dostępu światła (jak np. nietoperze). Zatem odpowiednia stereowizja z właściwymi sensorami również powinna być co najmniej wystarczająca przy realizacji codziennych zadań bez niepotrzebnego zaśmiecania środowiska.



---

## SPIS RYSUNKÓW

2.1	Bezzałogowy samolot Ikhana ( <i>źródło: NASA licencja: Public Domain</i> ). . . . .	9
2.2	Bezzałogowy samolot RQ-1 Predator ( <i>źródło: Departament Obrony USA licencja: Public Domain</i> ). . . . .	10
2.3	Mały dron Phantom 4 ( <i>źródło: DJI licencja: Creative Commons 0</i> ). . . . .	11
2.4	Autonomiczny robot podwodny Echo Voyager ( <i>źródło: DARPA licencja: Public Domain</i> ). . . . .	13
2.5	Robot gaśniennicowy Talon ( <i>źródło: Departament Obrony USA licencja: Public Domain</i> ). . . . .	14
2.6	Robot humanoidalny FLASH ( <i>źródło: lirec.eu, prawa autorskie: Lirec Group</i> ). . . . .	15
2.7	Robot humanoidalny Nao ( <i>źródło: Ubahnverleih wikipedia licencja: Creative Commons 0</i> ). . . . .	16
2.8	Robot czworonożny BigDog ( <i>źródło: DARPA licencja: Public Domain</i> ). . . . .	16
2.9	Robot SnakeBot ( <i>źródło: NASA licencja: Public Domain</i> ). . . . .	17
3.1	Schemat działania SLAM. . . . .	45
3.2	Model otworkowy kamery. . . . .	48
3.3	Równoważny model otworkowy kamery. . . . .	48
3.4	Zniekształcenia radialne. . . . .	50
3.5	Geometria epipolarna. . . . .	51
3.6	Rzutowanie odcinka na płaszczyzny nieskalibrowanych kamer. . . . .	52
3.7	Rektyfikacja obrazów z kamer. . . . .	53
3.8	Dysparycja obrazów po rektyfikacji obrazów z kamer. . . . .	53
4.1	Schemat działania systemu VisRobot ze zmienną bazą. . . . .	59
4.2	Kalibracja magnetometru. . . . .	62
4.3	Schemat algorytmu wyznaczania ułożenia oparty na danych inercyjnych. . . . .	63
4.4	Przyspieszenia podczas obrotu sensora względem osi $x$ , $y$ , $z$ . . . . .	65
4.5	Błąd związany z obrotem kamery względem osi $x$ oraz $y$ . . . . .	76
4.6	Błąd związany z obrotem kamery względem osi $z$ . . . . .	76
4.7	Błąd związany z przesunięciem kamery w kierunku $z$ . . . . .	77
4.8	Błąd głębi dla różnych baz stereo. . . . .	79
4.9	Teoretyczny przebieg funkcji wiarygodności w zależności od odległości. . . . .	79
4.10	Funkcja wiarygodności z uwzględnieniem stałego błędu położenia i ułożenia. . . . .	80

4.11	Funkcja wiarygodności uwzględniająca stały błąd ułożenia oraz zmienny błąd położenia baz $B_1$ - $B_3$ . . . . .	81
4.12	Obszar uśredniania głębi na granicach zakresu sąsiadujących baz. . . . .	81
4.13	Postrzeganie głębi w obrazie stereoskopowym. . . . .	84
5.1	Przyspieszenie robota względem osi $y$ . . . . .	89
5.2	Ułożenie robota względem osi $x$ (niebieski), $y$ (zielony), $z$ (czerwony). . . . .	90
5.3	Przyspieszenie robota wzdłuż osi $x$ spowodowane zmianą ułożenia. . . . .	90
5.4	Przyspieszenie robota. . . . .	91
5.5	Położenie robota. . . . .	91
5.6	Generowane szумы użyte w symulacji. . . . .	92
5.7	Szum typu błądzenie losowe. . . . .	92
5.8	Dwukrotnie całkowany szum typu błądzenie losowe. . . . .	93
5.9	Wpływ szumu akcelerometru na ocenę położenia robota. . . . .	94
5.10	Wpływ szumu ułożenia na ocenę położenia robota. . . . .	94
5.11	Symulowany sygnał wymuszenia robota. . . . .	96
5.12	Symulowane położenie bez obecności szumu pomiarowego. . . . .	96
5.13	Symulowany szum akcelerometru. . . . .	97
5.14	Estymacja stanu oparta na obserwatorze Leuenbergera dla typowego szumu akcelerometru. . . . .	97
5.15	Estymacja stanu oparta na filtracji Kalmana dla typowego szumu z akcelerometru. . . . .	98
5.16	Studium czułości estymacji położenia dla obserwatora Leuenbergera (niebieski) i filtru Kalmana (zielony) w obecności szumu pomiarowego w postaci. . . . .	99
5.17	Trajektoria robota w systemie VisRobot. . . . .	99
5.18	Platforma testowa robota mobilnego. . . . .	100
5.19	Mierzone położenie oraz prędkość. . . . .	102
5.20	Dopasowanie odpowiedzi modelu do danych pomiarowych (prędkości). . . . .	102
5.21	Estymacja położenia robota za pomocą systemu VisRobot-KF. . . . .	104
5.22	Chwilowy błąd estymowanego położenia robota. . . . .	105
5.23	Estymacja drogi przejechanej przez robota: pomiar referencyjny, estymacja z progowaniem, estymacja bez progowania. . . . .	106
5.24	Render obrazu wirtualnej sceny "movie set scene" (źródło: free3d.com, licencja: Creative Commons 0). . . . .	108
5.25	Kierunki odniesienia eksperymentu. . . . .	109
5.26	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zakresie błędu ułożenia względem osi $x$ dla 4 obiektów. . . . .	110
5.27	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zakresie błędu ułożenia względem osi $y$ dla 4 obiektów. . . . .	111
5.28	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zakresie błędu ułożenia względem osi $z$ dla 4 obiektów. . . . .	111
5.29	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w funkcji błędu położenia w kierunku $y$ dla 4 obiektów. . . . .	112
5.30	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu położenia w kierunku $z$ dla 4 obiektów. . . . .	112



5.31	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w funkcji błędu ułożenia w osi $x$ dla 4 rzeczywistych obiektów. . . . .	115
5.32	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu ułożenia względem osi $z$ dla 4 rzeczywistych obiektów. . . . .	116
5.33	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu położenia w kierunku $y$ dla 4 rzeczywistych obiektów. . . . .	116
5.34	Bezwzględny błąd dysparycji oraz analiza jakościowa mapy dysparycji w zależności od błędu położenia w kierunku $z$ dla 4 rzeczywistych obiektów. . . . .	117
5.35	Renderowany obraz wirtualnego otoczenia i jego wyznaczona głębokość odniesienia. . . . .	119
5.36	Mapa dokładności mapowania dla czterech analizowanych baz . . . . .	120
5.37	Zdjęcie mapowanego pomieszczenia z makietą kolejową w Laboratorium Integracji Systemów Automatyki. . . . .	121
5.38	Mapy dysparycji przy czterech bazach stereowizyjnych. . . . .	123
5.39	Wynikowa mapa dysparycji. . . . .	124
5.40	Mapa 3D w systemie ze zmienną bazą. . . . .	125
5.41	Mapa 3D w rzucie perspektywicznym z wyróżnieniem punktów wyznaczonych dla 2 sąsiednich baz (czerwony $B_0$ , żółty $B_1$ ). . . . .	125
5.42	Rzut poziomy (z góry) mapy 3D z wyróżnieniem punktów wyznaczonych dla 2 sąsiednich baz (czerwony $B_0$ , żółty $B_1$ ). . . . .	126



---

## SPIS TABLIC

2.1	Porównanie algorytmów SLAM opartych na danych wizyjnych. . . . .	27
4.1	Ogólne parametry kinematyczne opisujące rozważanego robota mobilnego.	67
4.2	Skrócony zapis rozkazów sterowania w systemie VisRobot przy 4 bazach.	72
5.1	Wartości parametrów modelu robota opisanego wzorem (4.15). . . . .	103
5.2	Wyniki pomiaru odległości charakterystycznych punktów badanych obiektów wirtualnych ( $O_1 - O_4$ ). . . . .	109
5.3	Wyniki pomiaru odległości badanych czterech rzeczywistych obiektów. . .	114
5.4	Średni pomiar odległości środka badanych obiektów od kamery. . . . .	115
5.5	Zakresy dysparycji związane ze zmienną bazą użyte w wirtualnych scenach.	118
5.6	Błąd mapowania dla czterech szerokości bazy uśredniony po 30 realizacjach.	121
5.7	Ocena jakości głębi dla czterech baz. . . . .	124





---

## WYKAZ WAŻNIEJSZYCH SKRÓTÓW I OZNACZEŃ

Skróty	
AHRS	Attitude and Heading Reference System
AIS	Automatic Identification System
AOV	Angle of View
BLC	Backlight Compensation
BM	Block Matching
CCD	Charge-Coupled Device
CMOS	Complementary Metal-Oxide Semiconductor
CPU	Central Processing Unit
DARPA	Defense Advanced Research Projects Agency
DC	Direct Current
DCM	Direct Cosine Matrix
EKF	Extended Kalman Filter
FPS	Frames Per Second
GPU	Graphics Processing Unit
HLC	Highlight Compensation
HSV	Hue Saturation Value
IMU	Inertial Measurement Unit
MEMS	Micro Electro Mechanical Systems
NASA	National Aeronautics and Space Administration
PIAP	Przemysłowy Instytut Automatyki i Pomiarów
PWM	Pulse Width Modulation
ROV	Remotely Operated Vehicle
SAR	Synthetic Aperture Radar
SfM	Structure from Motion
SGBM	Semi-Global Block Matching
SLAM	Simultaneous Localization and Mapping
SUAV	Small Unmanned Aircraft Vehicle
UAV	Unmanned Aircraft Vehicle
UKF	Unscented Kalman Filter
UUV	Underwater Unmanned Vehicle
WDR	Wide Dynamic Range

Oznaczenia	
$\mathbb{R}$	zbiór liczb rzeczywistych
$\mathcal{T}$	transformacja układu współrzędnych
$T$	transpozycja macierzy
<i>rank</i>	rzęd macierzy
$\  \cdot \ $	norma
*	sprzężenie
$\bullet$	iloczyn skalarny
$\times$	iloczyn wektorowy
$\sigma$	odchylenie standardowe
$R$	macierz obrotu
$T$	wektor translacji
$q$	kwaternion
$G(s)$	transmitacja układu ciągłego (dziedzina Laplace'a)
$A, B, C, D$	macierze stanu
$\dot{\mathbf{x}}(t)$	pochodna wektora stanu po czasie
$\hat{\mathbf{x}}$	estymata stanu
$\mathbf{v}$	szum procesowy
$\mathbf{w}$	szum pomiarowy
$K$	wzmocnienie Kalmana
$f_s$	częstotliwość próbkowania
$f_{stop}$	częstotliwość odcięcia
$\mathbf{g}$	wektor grawitacji
$g$	przyspieszenie grawitacyjne ( $9.81m/s^2$ )
$a_l, a_s, a_t$	składowe przyspieszenia: wzdłużna, boczna, poprzeczna
$\phi, \theta, \psi$	obrót względem osi $x, y, z$
$\eta$	szum sensora
$\omega$	prędkość obrotowa
$\epsilon$	przyspieszenie obrotowe
$k_m$	stała mechaniczna
$k_e$	stała elektryczna
$J$	moment bezwładności stojana
$B_v$	współczynnik tarcia
$R_r$	rezystancja silnika
$L_r$	indukcyjność silnika
$T_m$	mechaniczną stałą czasową silnika
$T_e$	elektryczną stałą czasową silnika
<i>ratio</i>	przełożenie przekładni silnika
$f$	ogniskowa
$c_x, c_y$	współrzędne punktu głównego kamery
$M$	macierz parametrów wewnętrznych
$\pi^p, \pi^l, \pi^r$	plaszczyny: epipolarna, kamer (lewej i prawej)
$o_l, o_r$	środki rzutowania kamery (lewej i prawej)
$d$	dysparycja

$\mathcal{P}$	paralaksa
$z$	głębina obrazu
$B$	baza stereowizyjna
$Q$	macierz odwzorowania
$I_L, I_R$	luminancja obrazu lewego i prawego



---

## BIBLIOGRAFIA

- Abdolahi, A. i Aghaebrahimi Samani, H. (2012). An efficacious method to assemble a modern multimodal robotic team: dilemmas, challenges, possibilities and solutions. In *Introduction to Modern Robotics*, pages 151–172. CreateSpace Independent Publishing Platform.
- Agrawal, D. K., Woodhouse, J., i Seshia, A. (2013). Modeling nonlinearities in MEMS oscillators. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 60(8):1646–1659.
- Ahrary, A. i Ishikawa, M. (2006). Self-Localization of autonomous sewer robots by using a stereo camera and a laser scanner. In *IEEE International Conference on Networking, Sensing and Control*, pages 78–83, Ft. Lauderdale, USA.
- Albarbar, A., Badri, A., Sinha, J. K., i Starr, A. (2009). Performance evaluation of MEMS accelerometers. *Measurement: Journal of the International Measurement Confederation*, 42(5):790–795.
- Almabrouk, H., Kharroubi, M., Tounsi, F., Mezghani, B., i Bernard, Y. (2016). Macro model analysis of a single mass 6-DOF Inertial Measurement Unit system. In *11th International Design & Test Symposium (IDT)*, pages 290–295, Hammamet, Tunezja.
- Ang, W. i Khoo, S. (2004). Physical model of a MEMS accelerometer for low-g motion tracking applications. In *International Conference on Robotics and Automation*, volume 2, pages 1345–1351, New Orleans, USA.
- Angeli, A., Filliat, D., Doncieux, S., i Meyer, J. A. (2008). Fast and incremental method for loop-closure detection using bags of visual words. *IEEE Transactions on Robotics*, 24(5):1027–1037.
- Angrisani, L., D’Alessandro, G., D’Arco, M., Paciello, V., i Pietrosanto, A. (2015). Autonomous recharge of drones through an induction based power transfer system. In *IEEE International Workshop on Measurements & Networking (M&N)*, pages 1–6, Coimbra, Portugalia.
- Austin, R. (2011). *Unmanned Aircraft Systems: UAVS Design, Development and Deployment*. John Wiley & Sons, Chichester, UK.

- Bai, M., Zhuang, Y., i Wang, W. (2010). Stereovision based obstacle detection approach for mobile robot navigation. In *2010 International Conference on Intelligent Control and Information Processing*, pages 328–333, Dalian, Chiny.
- Bailey, T. i Durrant-Whyte, H. (2006a). Simultaneous localization and mapping (SLAM): Part I. *IEEE Robotics and Automation Magazine*, 13(2):99–110.
- Bailey, T. i Durrant-Whyte, H. (2006b). Simultaneous localization and mapping (SLAM): Part II. *IEEE Robotics and Automation Magazine*, 13(3):108–117.
- Beard, R. W. i McLain, T. W. (2012). *Small Unmanned Aircraft: Theory and Practice*. Princeton University Press, Princeton, USA.
- Beeby, S., Ensell, G., Kraft, M., i White, N. (2004). *MEMS Mechanical Sensors*. Artech House, Boston, USA.
- Bekir, E. (2007). *Introduction to Modern Navigation Systems*. Recherche. World Scientific Publishing, Londyn, UK.
- Belter, D., Łabecki, P., i Skrzypczyński, P. (2012). Estimating terrain elevation maps from sparse and uncertain multi-sensor data. In *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 715–722, Guangzhou, Chiny.
- Belter, D., Łabecki, P., Fankhauser, P., i Siegwart, R. (2016). RGB–D terrain perception and dense mapping for legged robots. *International Journal of Applied Mathematics and Computer Science*, 26(1):81–97.
- Ben-Tzvi, P. i Xu, X. (2010). An embedded feature-based stereo vision system for autonomous mobile robots. In *IEEE International Workshop on Robotic and Sensors Environments (ROSE)*, pages 176–181, Phoenix, USA.
- Besl, P. i McKay, N. D. (1992). A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256.
- Bier, A. i Luchowski, L. (2009). Error analysis of stereo calibration and reconstruction. In *4th International Conference on Computer Vision/Computer Graphics Collaboration Techniques*, pages 230–241, Rocquencourt, Francja.
- Bignoli, A., Romanoni, A., i Matteucci, M. (2018). Multi-view stereo 3D edge reconstruction. In *IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 867–875, Milan, Włochy.
- Bono, R., Caccia, M., Spirandelli, E., i Veruggie, G. (1999). ROV exploration of the keel of the Campbell Ice Tongue in Antarctica. In *Oceans '99 MTS/IEEE. Riding the Crest into the 21st Century. Conference and Exhibition*, volume 2, pages 563–566, Seattle, USA.
- Borenstein, J., Everett, H. R., Feng, L., i Wehe, D. (1997). Mobile robot positioning: Sensors and techniques. *Journal of Robotic Systems*, 14(4):231–249.

- Bradski, G. i Kaehler, A. (2013). *Learning OpenCV: Computer Vision in C++ with the OpenCV Library*. O'Reilly Media, Inc., Sebastopol, USA, 2nd edition.
- Cairns, E. J. i Albertus, P. (2010). Batteries for electric and hybrid-electric vehicles. In *Annual Review of Chemical and Biomolecular Engineering*, volume 1, pages 299–320, Berkeley, USA.
- Campion, G., Bastin, G., i D'Andrea-Novell, B. (2006). Structural properties and classification of kinematic and dynamic models of wheeled mobile robots. In *IEEE International Conference on Robotics and Automation*, pages 462–469, Ft. Lauderdale, USA.
- Chandra, K. P. B., Gu, D.-W., i Postlethwaite, I. (2010). SLAM using EKF,  $EH_\infty$  and mixed  $EH_2/H_\infty$  filter. In *IEEE International Symposium on Intelligent Control*, pages 818–823, Yokohama, Japonia.
- Choi, I. i Ha, J. (2010). Easy calibration method for omni-stereo camera system. *International Conference on Control Automation and Systems (ICCAS)*, pages 2315–2317.
- Chrysostomou, D., Kyriakoulis, N., i Gasteratos, A. (2010). Multi-camera 3D scene reconstruction from vanishing points. In *2010 IEEE International Conference on Imaging Systems and Techniques*, pages 343–348, Saloniki, Grecja.
- Czubenko, M., Kowalczyk, Z. Z., i Ordys, A. (2015). Autonomous driver based on an intelligent system of decision-making. *Cognitive Computation*, 7(5):569–581.
- Das, B., Chakrabarti, A., Kasari, P. R., i Sharma, N. (2013). Sensorless speed control of DC machine. In *International Conference on Power, Energy and Control (ICPEC)*, pages 572–577, Sri Rangalatchum Dindigul, Indie.
- Dauwalter, C. i Ha, J. (2005). Magnetically suspended MEMS spinning wheel gyro. *IEEE Aerospace and Electronic Systems Magazine*, 20(2):21–26.
- Davison, A. J., Reid, I. D., Molton, N. D., i Stasse, O. (2007). MonoSLAM: Real-time single camera SLAM. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(6):1052–1067.
- Dennis, R. i Edwards, J. (2013). Considering endoscopic design: a snakebot prototype. *IEEE Pulse*, 4(4):30–35.
- Diebel, J. i Reutersward, K. (2004). Simultaneous localization and mapping with active stereo vision. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 4, pages 3436–3443, Sendai, Japonia.
- Dubrovic, F. S. i Scherbatyuk, A. P. (2016). About accuracy estimation of AUV single-beacon mobile navigation using ASV, equipped with DGPS. In *OCEANS 2016 - Shanghai*, pages 1–4, Shanghai, Chiny.
- EASA (2015). Wprowadzenie ram regulacyjnych dotyczących wykonywania operacji z użyciem dronów. Europejska Agencja Bezpieczeństwa Lotniczego.

- Edwan, E., Zhang, J., i Loffeld, O. (2012). Angular motion and attitude estimation using fixed and rotating accelerometers configuration. In *IEEE/ION Position, Location and Navigation Symposium*, pages 8–14, Myrtle Beach, USA.
- Esrafilian, O. i Taghirad, H. D. (2016). Autonomous flight and obstacle avoidance of a quadrotor by monocular SLAM. In *4th International Conference on Robotics and Mechatronics (ICROM)*, pages 240–245, Tehran, Iran.
- Estrada, C., Neira, J., i Tardos, J. D. (2005). Hierarchical SLAM: real-time accurate mapping of large environments. *IEEE Transactions on Robotics*, 21(4):588–596.
- Everett, H. (1995). *Sensors for Mobile Robots*. A.K. Peters, Natick, USA.
- Furukawa, Y. i Ponce, J. (2007). Accurate, dense, and robust multi-view stereopsis. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Minneapolis, USA.
- Futami, S., Nagai, T., i Kurematsu, A. (2002). 3D modeling of real environments by autonomous mobile robot with stereo vision. In *International Conference on Image Processing*, volume 1, pages 856–859, Rochester, USA.
- Gallup, D., Frahm, J.-M., Mordohai, P., i Pollefeys, M. (2008). Variable baseline/resolution stereo. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1–8, Anchorage, USA.
- Gelfand, N., Ikemoto, L., Rusinkiewicz, S., i Levoy, M. (2003). Geometrically stable sampling for the ICP algorithm. In *Proceedings of International Conference on 3-D Digital Imaging and Modeling*, pages 260–267, Banff, Kanada.
- Georgiades, C., German, A., Hogue, A., Hui Liu, i Prahacs, C. (2004). AQUA: an aquatic walking robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3525–3531, Sendai, Japonia.
- Gilbert, H. B., Celik, O., i O'Malley, M. K. (2010). Long-term double integration of acceleration for position sensing and frequency domain system identification. In *IEEE/A-SME International Conference on Advanced Intelligent Mechatronics*, pages 453–458, Montreal, Kanada.
- Godfrey, A., Hourigan, T., i OLaighin, G. M. (2007). Pendulum analysis of an integrated accelerometer to assess its suitability to measure dynamic acceleration for gait applications. In *29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4891–4894, Lyon, Francja.
- Goel, P., Roumeliotis, S., i Sukhatme, G. (1999). Robot localization using relative and absolute position estimates. In *IEEE/RSJ International Conference on Intelligent Robots and Systems. Human and Environment Friendly Robots with High Intelligence and Emotional Quotients.*, pages 1134–1140, Kyongju, Korea Południowa.
- Grygiel, R. i Bieda, R. (2014). Metody wyznaczania kątów z żyroskopów dla filtru komplementarnego na potrzeby określania orientacji IMU. *Przegląd Elektrotechniczny*, 90(9):217–224.



- Gui, P., Tang, L., i Mukhopadhyay, S. (2015). MEMS based IMU for tilting measurement: Comparison of complementary and kalman filter based data fusion. In *10th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, pages 2004–2009, Auckland, Nowa Zelandia.
- Han, L., Qian, H., Xing, K., i Xu, Y. (2016). Heavy-payload omnidirectional robot. In *IEEE International Conference on Real-time Computing and Robotics (RCAR)*, pages 579–584, Angkor Wat, Kambodża.
- Hansen, P., Alismail, H., Rander, P., i Browning, B. (2012). Online continuous stereo extrinsic parameter estimation. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 1059–1066, Providence, USA.
- Hardt von der, H.-J., Wolf, D., i Husson, R. (1996). The dead reckoning localization system of the wheeled mobile robot ROMANE. In *IEEE/SICE/RSJ International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 603–610, Washington, USA.
- Harms, J. (2015). Terrestrial gravity fluctuations. *Living Reviews in Relativity*, 18(1):3.
- Harris, C. G. i Pike, J. M. (1988). 3D positional integration from image sequences. *Image and Vision Computing*, 6(2):87–90.
- Hartley, R. i Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*. Cambridge University Press, Nowy York, USA.
- Hirschmuller, H. (2008). Stereo processing by semiglobal matching and mutual information. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 30(2):328–341.
- Holdaway, C. (2000). Open-loop step motor control system. US Patent 6,016,044.
- Hong, S. (2012). A comparison of SLAM algorithms based on a graph of relations. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2089–2095, St. Louis, USA.
- Hu, M. i You, Z. (2010). A resonant magnetometer based on MEMS. In *International Conference on Electrical and Control Engineering*, pages 4944–4947, Wuhan, Chiny.
- Huang, H., Zhu, D., i Yuan, F. (2012). Dynamic task assignment and path planning for multi-AUV system in 2D variable ocean current environment. In *24th Chinese Control and Decision Conference (CCDC)*, pages 3660–3664, Taiyuan, Chiny.
- Isidori, A. (1995). Elementary theory on nonlinear feedback for single-input single-output systems. In *Nonlinear Control Systems*, pages 137–218. Springer-Verlag, London, UK.
- Jeon, J., Kim, K., Kim, C., i Ho, Y.-S. (2001). A robust stereo-matching algorithm using multiple-baseline cameras. In *IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, volume 1, pages 263–266, Victoria, Kanada.

- Jia, Y., Li, M., An, L., i Zhang, X. (2003). Autonomous navigation of a miniature mobile robot using real-time trinocular stereo machine. In *IEEE International Conference on Robotics, Intelligent Systems and Signal Processing*, volume 1, pages 417–421, Changsha, Chiny.
- Johansen, T. A. i Perez, T. (2016). Unmanned aerial surveillance system for hazard collision avoidance in autonomous shipping. In *International Conference on Unmanned Aircraft Systems (ICUAS)*, pages 1056–1065, Arlington, USA.
- Jones, G. R., Lee, D., Holliman, N. S., i Ezra, D. (2001). Controlling perceived depth in stereoscopic images. In *Proceedings of 8th Stereoscopic Displays and Virtual Reality Systems*, pages 42 – 53, San Jose, USA.
- Joubert, J. W., de Beer, D., i de Koker, N. (2016). Combining accelerometer data and contextual variables to evaluate the risk of driver behaviour. *Transportation Research Part F: Traffic Psychology and Behaviour*, 41(1):80–96.
- Julier, S. i Uhlmann, J. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*, 92(3):401–422.
- Junchuan, L., Yuru, Z., i Zhen, L. (2006). Selection of cameras setup geometry parameters in binocular stereovision. In *IEEE Conference on Robotics, Automation and Mechatronics*, pages 1–6, Bangkok, Tajlandia.
- Kalman, R. E. (1959). On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3):481–492.
- Kędzierski, J., Kaczmarek, P., Dziergwa, M., i Tchoń, K. (2015). Design for a robotic companion. *International Journal of Humanoid Robotics*, 12(01):1–25.
- Kitajima, K., Masuzawa, H., Miura, J., i Satake, J. (2010). A hierarchical SLAM for uncertain range data. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 144–149, Salt Lake City, USA.
- Kowalczyk, Z. (1992). Competitive identification for self-tuning control: Robust estimation design and simulation experiments. *Automatica*, 28(1):193–201.
- Kowalczyk, Z. i Czubenko, M. (2015). Przegląd robotów humanoidalnych. *Pomiary, Automatyka, Robotyka*, 19(4):33–41.
- Kowalczyk, Z. i Domżański, M. (2013). Asynchronous networked estimation system for continuous time stochastic processes. In *Conference on Control and Fault-Tolerant Systems (SysTol)*, pages 300–305, Nicea, Francja.
- Kowalczyk, Z. i Merta, T. (2013). Stereo image visualization for VISROBOT system. In *18th International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 794–799, Międzyzdroje, Polska.
- Kowalczyk, Z. i Merta, T. (2014). Wizualizacja obrazu stereowizyjnego w systemie VISROBOT. *Pomiary, Automatyka, Kontrola*, 10(1):803–808.

- Kowalczyk, Z. i Merta, T. (2015). Evaluation of position estimation based on accelerometer data. In *10th International Workshop on Robot Motion and Control (RoMoCo)*, pages 246–251, Poznań, Polska.
- Kowalczyk, Z. i Merta, T. (2016a). Evaluating the position of a mobile robot using accelerometer data. In *Advanced and Intelligent Computations in Diagnosis and Control*, volume 386, pages 131–143. Springer International Publishing, Londyn, UK.
- Kowalczyk, Z. i Merta, T. (2016b). Three-dimensional mapping for data collected using variable stereo baseline. In *21st International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 1082–1087, Międzyzdroje, Polska.
- Kowalczyk, Z. i Merta, T. (2017). Estimation of DC motor parameters using a simple CMOS camera. In *22nd International Conference on Methods and Models in Automation and Robotics (MMAR)*, pages 872–877, Międzyzdroje, Polska.
- Kozłowski, K., Dutkiewicz, P., i Wróblewski, W. (2012). *Modelowanie i Sterowanie Robotów*. Wydawnictwo Naukowe PWN, Warszawa, Polska.
- Kreindler, E. i Sarachik, P. (1964). On the concepts of controllability and observability of linear systems. *IEEE Transactions on Automatic Control*, 9(2):129 – 136.
- Krishnamurthy, P. i Khorrami, F. (2003). Robust adaptive voltage-fed permanent magnet step motor control without current measurements. *IEEE Transactions on Control Systems Technology*, 11(3):415–425.
- Kumar, S. (2009). Binocular stereo vision based obstacle avoidance algorithm for autonomous mobile robots. In *IEEE International Advance Computing Conference*, pages 254–259, Patiala, Indie.
- Kushner, H. J. (1962). Optimal stochastic control. *IRE Transactions on Automatic Control*, 7(5):120–122.
- Kuzmin, M. (2018). Review. Classification and comparison of the existing SLAM methods for groups of robots. In *22nd Conference of Open Innovations Association (FRUCT)*, pages 115–120, Jyväskylä, Finlandia.
- Labbé, M. i Michaud, F. (2014). Online global loop closure detection for large-scale multi-session graph-based SLAM. In *IEEE International Conference on Intelligent Robots and Systems*, pages 2661–2666, Chicago, USA.
- Lategahn, H., Geiger, A., i Kitt, B. (2011). Visual SLAM for autonomous ground vehicles. In *IEEE International Conference on Robotics and Automation*, pages 1732–1737, Shanghai, Chiny.
- Lee, J.-H., Uk-Jin, J., i Hong, Y.-S. (2016). Indoor navigation for an automatic guided vehicle with beacon based relative distance estimation. In *International Conference on Ubiquitous and Future Networks (ICUFN)*, pages 424–429, Wiedeń, Austria.
- Levander, O. (2017). Autonomous ships on the high seas. *North American IEEE Spectrum*, 54(2):54–59.

- Li, Y., Dempster, A., Li, B., Wang, J., i Rizos, C. (2006). A low-cost attitude heading reference system by combination of GPS and magnetometers and MEMS inertial sensors for mobile applications. *Journal of Global Positioning system*, 5(1):88–95.
- Liu, T., Yuan, D., Zhao, H., i Yin, J. (2017). Multi-view stereo via geometric expansion and depth refinement. In *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, pages 555–560, Macau, Chiny.
- Lu-Fang, G., Yu-Xian, G., i Sheng, F. (2007). Simultaneous localization and mapping for autonomous mobile robots using binocular stereo vision system. In *International Conference on Mechatronics and Automation (ICMA)*, pages 326–330, Harbin, Chiny.
- Luenberger, D. G. (1964). Observing the state of a linear system. *IEEE Transactions on Military Electronics*, 8(2):74–80.
- Ma, Y., Kosecka, J., Soatto, S., i Sastry, S. (2006). *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer Verlag, Nowy York, USA.
- Madgwick, S. O. H., Harrison, A. J. L., i Vaidyanathan, A. (2011). Estimation of IMU and MARG orientation using a gradient descent algorithm. In *IEEE International Conference on Rehabilitation Robotics*, pages 1–7, Zurich, Szwajcaria.
- Mahon, I., Williams, S. B., Pizarro, O., i Johnson-Roberson, M. (2008). Efficient view-based SLAM using visual loop closures. *IEEE Transactions on Robotics*, 24(5):1002–1014.
- Martinelli, A., Tomatis, N., Tapus, A., i Siegwart, R. (2003). Simultaneous localization and odometry calibration for mobile robot. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 2, pages 1499–1504, Las Vegas, USA.
- Martinez-Gomez, J., Fernandez-Caballero, A., Garcia-Varea, I., Rodriguez, L., i Romero-Gonzalez, C. (2014). A taxonomy of vision systems for ground mobile robots. *International Journal of Advanced Robotic Systems*, 11(7):1–26.
- McAree, O., Aitken, J. M., i Veres, S. M. (2016). A model based design framework for safety verification of a semi-autonomous inspection drone. In *UKACC 11th International Conference on Control*, pages 1–6, Belfast, UK.
- Meng, X., Wang, S., Cao, Z., i Zhang, L. (2016). A review of quadruped robots and environment perception. In *35th Chinese Control Conference (CCC)*, pages 6350–6356, Chengdu, Chiny.
- Merlin, P. W. (2009). *Ikhana: Unmanned aircraft system western states fire missions. Monographs in aerospace history*. CreateSpace Independent Publishing Platform, Washington, USA.
- Min, S., Rixin, H., i Daojun, W. (2007). Precision analysis to 3D reconstruction from image sequences. In *ISPRS Workshop on Updating Geo-spatial Databases with Imagery*, pages 141–146, Beijing, Chiny.

- Mohd-Yasin, F., Nagel, D. J., i Korman, C. E. (2009). Noise in MEMS. *Measurement Science and Technology*, 21(1):1–22.
- Mokri, Y. i Yasamin, M. (2005). Omni-stereo vision system for an autonomous robot using neural networks. In *Canadian Conference on Electrical and Computer Engineering*, pages 1590–1593, Saskatoon, Kanada.
- Montemerlo, M. i Thrun, S. (2007). *FastSLAM: A Scalable Method for the Simultaneous Localization and Mapping Problem in Robotics*. Springer-Verlag, Berlin Heidelberg.
- Nakabo, Y., Mukai, T., Hattori, Y., Takeuchi, Y., i Ohnishi, N. (2005). Variable baseline stereo tracking vision system using high-speed linear slider. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 1567–1572, Barcelona, Hiszpania.
- Nakajoh, H., Osawa, H., Miyazaki, T., Hirata, K., Sawa, T., i Utsugi, H. (2012). Development of work class ROV applied for submarine resource exploration in JAMSTEC. In *2012 Oceans - Yeosu*, pages 1–5, Yeosu, Korea Południowa.
- Nedevschi, S., Marita, T., Danescu, R., Oniga, F., Frentiu, D., i Pocol, C. (2002). Camera calibration error analysis in stereo measurements. *Journal of Control Engineering and Applied Informatics (CEAI)*, 4(2):21–28.
- Nez, A., Fradet, L., Laguillaumie, P., Monnet, T., i Lacouture, P. (2016). Comparison of calibration methods for accelerometers used in human motion analysis. *Medical Engineering and Physics*, 38(11):1289–1299.
- Nise, N. S. (2015). *Control Systems Engineering*. Wiley, Jefferson City, USA.
- Oleari, F., Lodi Rizzini, D., Kallasi, F., Aleotti, J., i Caselli, S. (2016). Issues in high performance vision systems design for underwater interventions. In *42nd Annual Conference of the IEEE Industrial Electronics Society*, pages 945–950, Florence, Włochy.
- Open3dmodel.com (2015). Romantic Living Room Interior 3d Max Model Free - 7674. data dostępu: 2018-03-26.
- Pang, G. i Liu, H. (2001). Evaluation of a low-cost MEMS accelerometer for distance measurement. *Journal of Intelligent and Robotic Systems: Theory and Applications*, 30(3):249–265.
- Pawitan, G. A. H., Mutijarsa, K., i Adiprawita, W. (2016). Three-wheeled omnidirectional robot controller implementation. In *International Conference on Information Technology Systems and Innovation (ICITSI)*, pages 1–4, Bandung, Indonezja.
- Peng, S., Lam, C., i Cole, G. (2003). A biologically inspired four legged walking robot. In *IEEE International Conference on Robotics and Automation*, volume 2, pages 2024–2030, Taipei, Tajwan.
- Pfeifer, R., Lungarella, M., i Iida, F. (2007). Self-organization, embodiment, and biologically inspired robotics. *Science*, 318(5853):1088–1093.

- Qian, N. i Lo, C.-Y. (2015). Optimizing camera positions for multi-view 3D reconstruction. In *International Conference on 3D Imaging (IC3D)*, pages 1–8, Liege, Belgia.
- Rodriguez, N. E. N. (2011). *Advanced Mechanics in Robotic Systems*. Mobile Robots. Springer, Londyn, UK.
- Rusinkiewicz, S. i Levoy, M. (2001). Efficient variants of the ICP algorithm. In *International Conference on 3-D Digital Imaging and Modeling (3DIM)*, pages 145–152, Quebec City, Kanada.
- Sagitov, A. i Gerasimov, Y. (2017). Towards DJI Phantom 4 realistic simulation with gimbal and RC controller in ROS/Gazebo environment. In *10th International Conference on Developments in eSystems Engineering (DeSE)*, pages 262–266, Paryż, Francja.
- Samsuri, S., Zamzuri, H., Abdul Rahman, M., Mazlan, S., i Abd Rahman, A. H. (2014). Computational cost analysis of extended Kalman filter in simultaneous localization and mapping (EKF-SLAM) problem for autonomous vehicle. *Applied Mechanics and Materials*, 10(17):1–8.
- Santos, J. M., Portugal, D., i Rocha, R. P. (2013). An evaluation of 2D SLAM techniques available in Robot Operating System. In *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 1–6, Linkoping, Szwecja.
- Scharstein, D., Hirschmüller, H., Kitajima, Y., Krathwohl, G., Neši, N., Wang, X., i Westling, P. (2014). High-resolution stereo datasets with subpixel-accurate ground truth. In *German Conference on Pattern Recognition, Lecture Notes in Computer Science*, pages 31–42, Münster, Niemcy.
- Scharstein, D. i Szeliski, R. (2002). A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *International Journal of Computer Vision*, 47(1):7–42.
- Sheard, K., Scaysbrook, I., i Cox, D. (2008). MEMS sensor and integrated navigation technology for precision guidance. In *IEEE Position Location and Navigation Symposium*, pages 1145–1151, Monterey, USA.
- Shimadu, T., Kasakawa, K., i Suzuki, S. (2012). Evaluation of sensibility to robots approach using stereovision. In *Proceedings of SICE Annual Conference*, pages 2212–2217, Akita, Japonia.
- Shortelle, K., Graham, W., i Rabourn, C. (1998). F-16 flight tests of a rapid transfer alignment procedure. In *IEEE Position Location and Navigation Symposium*, pages 379–386, Palm Springs, USA.
- Siciliano, B. i Khatib, O. (2008). Wheeled Robots. In *Springer Handbook of Robotics*, pages 391–410. Springer Berlin Heidelberg, Berlin, Niemcy.
- Siegel, M. i Nagata, S. (2000). Just enough reality: comfortable 3-D viewing via microstereopsis. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(3):387–396.

- Siegwart, R. i Nourbakhsh, I. R. (2004). *Introduction to Autonomous Mobile Robots*. MIT Press, Londyn, UK.
- Sim, R., Elinas, P., Griffin, M., Shyr, A., i Little, J. J. (2006). Design and analysis of a framework for real-time vision-based SLAM using Rao-Blackwellised particle filters. In *Canadian Conference on Computer and Robot Vision*, Guangzhou, Chiny.
- Sinopoli, B., Schenato, L., Franceschetti, M., Poolla, K., Jordan, M., i Sastry, S. (2004). Kalman filtering with intermittent observations. *IEEE Transactions on Automatic Control*, 49(9):1453–1464.
- Sipos, M., Paces, P., Rohac, J., i Novacek, P. (2012). Analyses of triaxial accelerometer calibration algorithms. *IEEE Sensors Journal*, 12(5):1157–1165.
- Soberi, M. S. B. M. i Zakaria, M. Z. B. (2016). Autonomous ship hull inspection by omnidirectional path and view. In *IEEE/OES Autonomous Underwater Vehicles (AUV)*, pages 38–43, Tokyo, Japonia.
- Stasse, O., Davison, A. J., Sellaouti, R., Yokoi, K., i Aist, C. (2006). Real-time 3D SLAM for humanoid robot considering pattern generator information. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 348–355, San Diego, USA.
- Stefano, L. D., Marchionni, M., i Mattoccia, S. (2004). A fast area-based stereo matching algorithm. *Image and Vision Computing*, 22(12):983–1005.
- Szeliski, R. i Scharstein, D. (2004). Sampling the disparity space image. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3):419–425.
- Tabibi, S., Geremia-Nievinski, F., i T. van Dam (2017). Statistical comparison and combination of GPS, GLONASS, and multi-GNSS multipath reflectometry applied to snow depth retrieval. *IEEE Transactions on Geoscience and Remote Sensing*, 55(7):3773–3785.
- Titterton, D. H. i Weston, J. L. (2005). *Strapdown Inertial Navigation Technology*. The Institution of Engineering and Technology, Stevenage, UK.
- Todescato, M., Carron, A., Carli, R., Franchi, A., i Schenato, L. (2016). Multi-robot localization via GPS and relative measurements in the presence of asynchronous and lossy communication. In *European Control Conference (ECC)*, pages 2527–2532, Aalborg, Dania.
- Tomono, M. (2009). Robust 3D SLAM with a stereo camera based on an edge-point ICP algorithm. In *IEEE International Conference on Robotics and Automation*, pages 4306–4311, Kobe, Japonia.
- Trojnacki, M., Dąbek, P., Kacprzyk, J., i Hendzel, Z. (2015). Motion stabilization system of a four-wheeled mobile robot for teleoperation mode: Experimental investigations in indoor environment. *Advances in Intelligent Systems and Computing*, 323(1):117–128.

- Tzafestas, S. G. (2014). *Introduction to Mobile Robot Control*. Elsevier, Londyn, UK.
- Vikranth, T., Raghavendra, C. S., i Kumari, T. S. (2016). Autonomous air & underwater vehicle. In *IEEE/OES Autonomous Underwater Vehicles (AUV)*, pages 69–75, Tokyo, Japonia.
- Walas, K., Nowicki, M., Ferstl, D., i Skrzypczyński, P. (2016). Depth data fusion for simultaneous localization and mapping - RGB-DD SLAM. In *IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pages 9–14, Baden, Niemcy.
- Wang, Y., Adams, S., Thorp, J., MacDonald, N., Hartwell, P., i Bertsch, F. (1998). Chaos in MEMS, parameter estimation and its potential application. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 45(10):1013 – 1020.
- Weatherington, D. i Deputy, U. (2005). Unmanned aircraft systems roadmap 2005-2030. Deputy, UAV Planning Task Force, OUSD (AT&L).
- Wei, J., Resch, B., i Lensch, H. P. A. (2016). Dense and occlusion-robust multi-view stereo for unstructured videos. In *13th Conference on Computer and Robot Vision (CRV)*, pages 69–76, Victoria, Kanada.
- Wolm, U. P., Chen, X. Q., Chase, J. G., Pettigrew, W., i Hann, C. E. (2008). Analysis of a PM DC motor model for application in feedback design for electric powered mobility vehicles. In *15th International Conference on Mechatronics and Machine Vision in Practice*, pages 640–645, Auckland, Nowa Zelandia.
- Woodman, O. (2007). An introduction to inertial navigation. Report 696. University of Cambridge.
- Wu, Y. i Pei, L. (2017). Gyroscope calibration via magnetometer. *IEEE Sensors Journal*, 17(16):5269–5275.
- Xu, G., Du, S., i Xue, J. (2016). Precise 2D point set registration using iterative closest algorithm and correntropy. In *Proceedings of the International Joint Conference on Neural Networks*, volume 10, pages 4627–4631, Vancouver, Kanada.
- Zhao, Y., Hou, X., Jia, L., i Ma, S. (2010). The obstacle avoidance system for mobile robot based on binocular stereo vision. In *8th World Congress on Intelligent Control and Automation*, pages 6461–6465, Jinan, Chiny.
- Zilly, B. F., Kluger, J., i Kauff, P. (2011). Production rules for stereo acquisition. *Proceedings of the IEEE*, 99(4):590–606.
- Zivkovic, Z. i der Heijden, F. (2006). Efficient adaptive density estimation per image pixel for the task of background subtraction. *Pattern Recognition Letters*, 27(7):773–780.
- Zollhöfer, M., Stotko, P., Görlitz, A., Theobalt, C., Nießner, M., Klein, R., i Kolb, A. (2018). State of the art on 3D Reconstruction with RGB-D Cameras. *Computer Graphics Forum (Eurographics State of the Art Reports 2018)*, 37(2).