



Politechnika Łódzka

Instytut Informatyki

PRACA DYPLOMOWA MAGISTERSKA

Algorytmy rozpoznawania miejsc w modelowaniu nawigacji u mrówki pustynnej

Wydział Fizyki Technicznej, Informatyki i Matematyki Stosowanej

Promotor: dr inż. Przemysław Nowak

Dyplomant: Maciej Skorek

Nr albumu: 221655

Kierunek: Informatyka

Specjalność: Systemy Inteligentne i Inżynieria Oprogramowania

Łódź 20.02.2020



Instytut Informatyki

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

tel. 042 631 27 97, 042 632 97 57, fax 042 630 34 14 email: office@ics.p.lodz.pl

Streszczenie

Nawigacja stanowi istotny aspekt wielu codziennych zadań wykonywanych przez ludzi, umożliwiając dotarcie do wyznaczonego celu, które często wymaga pokonania skomplikowanej drogi w niełatwym terenie. Istnieje zatem duże zapotrzebowanie na stworzenie skutecznych algorytmów nawigacyjnych, które będą efektywnie wspierać różne dziedziny ludzkiego życia — jednym z przykładów są choćby autonomiczne pojazdy, obecnie najprężniej rozwijająca się dziedzina związana z nawigacją. Rozważając możliwe rozwiązania problemów spotykanych w tej tematyce, warto czerpać inspiracje z otaczającej nas natury, która przez lata wypracowała własne metody nawigacyjne. Jednym z takich przykładów są umiejętności mrówek pustynnych, które ze względu na trudne warunki środowiskowe opracowały efektywne techniki nawigacyjne. Techniki te rozwiązują problem precyzyjnej nawigacji, wykorzystując proste metody rozpoznawania obrazów związanych z widokami określonych lokalizacji oraz zebraną wiedzę o otoczeniu.

Niniejsza praca skupia się na zbadaniu metod rozpoznawania obrazów opracowanych na podstawie badań mrówek pustynnych. W szczególności w pracy poddano analizie wpływ doboru reprezentacji obrazów oraz metod ich porównywania na wielkość obszarów poprawnie rozpoznanych lokalizacji.

W ramach pracy wykonano środowisko wirtualne w programie V-REP, którego zadaniem było jak najdokładniejsze odzwierciedlenie środowiska naturalnego mrówek pustynnych. W oparciu o nie dokonano analizy zaproponowanych w literaturze algorytmów rozpoznawania lokalizacji. Zbadano trzy reprezentacje obrazów: wykorzystujące albo jasność wszystkich pikseli obrazu, albo wyznaczoną linię horyzontu, albo średnią jasność pikseli w kolejnych kolumnach obrazu. Zbadano także cztery metody porównywania obrazów: korelację, błąd średniokwadratowy, iloczyn skalarny oraz podobieństwo kosinusowe. Dla wszystkich możliwych zestawień reprezentacji i metod porównywania obrazów zanalizowano następnie obszary poprawnie rozpoznanych lokalizacji w okolicy mrowiska.

Otrzymane w niniejszej pracy wyniki potwierdziły słuszność wykorzystanych algorytmów, wskazując na istnienie obszarów poprawnego rozpoznania lokalizacji. W warunkach przygotowanego środowiska wirtualnego największe obszary rozpoznania uzyskano dla reprezentacji linii horyzontu. Reprezentacja ta w połączeniu z metodą porównywania wykorzystującą błąd średniokwadratowy wykazała najlepszą separację obrazów. Wyniki wskazują, że to właśnie to

połączenie jest najbardziej obiecujące w kontekście potencjalnie bardziej złożonych algorytmów nawigacyjnych.

Słowa kluczowe: modelowanie, symulacja, nawigacja, lokalizacja, rozpoznawanie miejsc, przetwarzanie obrazu, mrówki pustynne, V-REP.

Abstract

Navigation is an important aspect of many everyday tasks carried out by people. It allows them to reach their destinations, which often requires travelling long distances over difficult terrain. Therefore, there is a need to develop efficient navigational algorithms that could support various areas of human life. One example of a possible application is autonomous cars, currently the fastest growing field related to navigation. When searching for possible solutions, however, it is worth taking inspiration from the nature, as it has already developed many outstanding navigational tools. A good example are the skills exhibited by desert ants. These insects live in tough environmental conditions and thus require excellent navigational skills to survive. The skills that they have developed are an answer to the problem of precise navigation using simple methods based on view comparison and on the knowledge about surroundings of their nest.

The goal of this thesis is to examine several methods for image recognition that have been proposed based on observations of desert ants. Specifically, the thesis focuses on the effect of image representations and of methods for image comparison on the size of correctly recognised locations.

The work involved creating a virtual environment using the V-REP simulator, which would imitate as closely as possible a natural environment of desert ants. This environment was subsequently exploited for analysing various localisation algorithms. Specifically, three types of image representations were investigated: based on the intensities of all pixels in the image, based on the average intensities in consecutive columns in the image, and based on the extracted skyline. Also, four image comparison methods were examined: correlation, root-mean-square error, scalar product, and cosine similarity. For each combination of image representation and comparison method, the areas of correctly recognised locations in the vicinity of the nest were assessed.

The results obtained here have confirmed efficiency of the chosen algorithms by proving that areas of correctly recognised locations could be identified. In the virtual environment used, the largest areas were found for the skyline representation. This representation combined with the root-mean-square error as the image comparison method turned out to yield the best image separation. These results indicate that this combination seems to be the most promising in terms of more complex navigational algorithms.

Keywords: modelling, simulation, navigation, localisation, place recognition, image processing, desert ants, V-REP.

Spis treści

1	Wstęp	1
1.1	Wprowadzenie	1
1.2	Cele pracy	2
1.3	Układ pracy	3
2	Nawigacja u mrówek pustynnych	5
2.1	Główne strategie nawigacyjne	5
2.2	Reprezentacje i porównywanie obrazów	9
3	Materiały i metody	17
3.1	Symulacja komputerowa	17
3.1.1	Środowisko naturalne mrówki	17
3.1.2	Reprezentacja mrówki	21
3.2	Algorytmy rozpoznawania obrazów	22
3.2.1	Wybrane cechy	22
3.2.2	Wybrane metody porównywania obrazów	24
3.3	Przebieg badania	26
3.3.1	Dane	27
3.3.2	Badania <i>catchment areas</i> i <i>absolute areas</i>	28
4	Wyniki	31
5	Dyskusja	39
6	Wnioski	43

Rozdział 1

Wstęp

1.1 Wprowadzenie

W dzisiejszych czasach wykorzystuje się rozmaite narzędzia nawigacyjne, które stały się codziennością, wspierając ludzi w niemal każdym zadaniu. Narzędzia takie pomagają ludziom nie tylko podczas podróży, ale również wykorzystywane są w przemyśle, a także w dziedzinie astronomii. Obecnie, wraz z koncepcją autonomicznych pojazdów, bardzo prężnie rozwija się dziedzina związana z automatyzacją procesów nawigacyjnych. Opracowane rozwiązania mają w przyszłości całkowicie zastąpić człowieka, zwiększając komfort podróży oraz bezpieczeństwo pasażerów. Pojazdy autonomiczne mogą znaleźć również zastosowanie w magazynach lub niebezpiecznych akcjach ratowniczych. Algorytmy nawigacyjne są jednak wciąż problematyką, w której należy szukać lepszych i bardziej efektywnych rozwiązań.

Choć na przestrzeni lat wypracowano różne techniki nawigacji, to każda z nich posiada pewne ograniczenia. Wiele z nich bazuje na niezawodności sprzętu oraz siły sygnału. Inne zakładają wcześniejsze poznanie terenu. Obecnie najpopularniejszą metodą nawigacji jest nawigacja satelitarna wykorzystywana m.in. przez system GPS (ang. *Global Positioning System*). Nawigacja w oparciu o ten system może okazać się jednak zawodna, zwłaszcza kiedy stanowi on jedyne źródło informacji o pozycji. Jednym z często wstępujących zjawisk towarzyszących temu systemowi jest zanik sygnału. Ograniczenie stanowią również wnętrza budynków, tunele czy parkingi podziemne, gdzie sygnał nie dociera, a także lasy, w których sygnał się odbija. Problemy te stały się przyczyną poszukiwań nowych metod nawigacji, które pozwoliłyby na usprawnienie bądź zastąpienie już istniejących.

Wiele problemów człowiek stara się rozwiązać poprzez obserwację natury, która na liczne

trudności już dawno znalazła rozwiązanie. Umiejętność nawigacji posiada każde stworzenie, choć niektóre opanowały ją w szczególny sposób. Mrówki pustynne są przykładem stworzeń o wyjątkowych zdolnościach odnajdywania drogi powrotnej, doskonale sprawdzających się na trudnym terenie pustynnym. Dodatkowo ich zachowania oraz sposoby nawigacji zdają się być dużo łatwiejsze do naśladowania niż te, którymi kieruje się człowiek. Różnica między liczbą neuronów w mózgu człowieka w porównaniu do mrówki przekracza wartość 80 miliardów. Mimo to owady te bardzo sprawnie poruszają się po terenie, wykorzystując minimalny zbiór informacji o otoczeniu. Niezwykła miniaturyzacja przy jednoczesnej bardzo wysokiej wydajności czyni mrówki obiektem zainteresowania podczas poszukiwania algorytmów nawigacyjnych o niskiej złożoności obliczeniowej. Jest to bardzo istotne zwłaszcza z punktu widzenia zasilania robota lub pojazdu, bowiem niska złożoność obliczeniowa pozwoli na dłuższe i szybsze działanie urządzenia. Adaptacja techniki nawigacyjnej mrówek może znaleźć zastosowanie w trudno dostępnych miejscach bądź stanowić mechanizm awaryjny dla klasycznych metod.

1.2 Cele pracy

Celem niniejszej pracy było zbadanie wpływu doboru algorytmu nawigacyjnego na wielkość obszarów poprawnie rozpoznanych lokalizacji. Praca skupia się jedynie na aspekcie rozpoznania lokalizacji na podstawie zapamiętanych obrazów. Niniejsze studium stanowi wstęp do dalszych badań dotyczących nawigacji z wykorzystaniem mechanizmów zaobserwowanych u mrówek pustynnych.

Osiągnięcie powyższego celu uzyskano, realizując następujące cele szczegółowe:

1. **Utworzenie wirtualnego środowiska imitującego naturalne środowisko mrówek pustynnych.** Aby zapewnić możliwość badania algorytmów nawigacyjnych, potrzebne było środowisko, które symulowałoby obrazy widziane przez mrówkę. W zakresie pracy należało wykonać analizę środowiska naturalnego mrówek pustynnych oraz przygotować zestaw parametrów definiujących otoczenie mrówki. W kolejnym kroku należało wykonać program służący do tworzenia sceny w wirtualnym środowisku na podstawie zdefiniowanych parametrów. Na tym etapie należało również wygenerować scenę, która stanowiła podstawę dalszych badań.
2. **Stworzenie narzędzia umożliwiającego pobranie obrazów z wirtualnego środowiska.** Ten etap prac zakładał utworzenie narzędzia pozwalającego na ekstrakcję

zestawu obrazów widzianych przez mrówkę w scenie wirtualnego środowiska. W ramach prac należało następnie użyć tego narzędzia do przygotowania zbioru obrazów na potrzeby dalszych analiz.

3. **Implementacja algorytmów rozpoznawania obrazów.** W tej części pracy należało zaimplementować szereg algorytmów porównywania obrazów opisanych w literaturze dotyczącej problematyki nawigacyjnej mrówek pustynnych. W ramach algorytmów trzeba było zaprogramować metody ekstrakcji cech z obrazu oraz metody porównywania obrazów.
4. **Przeprowadzenie analizy algorytmów.** W końcowej fazie na podstawie zaimplementowanych algorytmów trzeba było dokonać analizy wielkości obszarów poprawnego rozpoznania lokalizacji. Otrzymane wyniki w kolejnym kroku stanowiły podstawę do weryfikacji i porównania algorytmów.

1.3 Układ pracy

Praca została podzielona na 6 rozdziałów.

Rozdział 1. stanowi wstęp do problematyki niniejszej pracy. Zawiera on informacje określające tematykę pracy, definicję celów oraz jej układ.

W rozdziale 2. opisane zostały techniki nawigacyjne wykorzystywane przez mrówki pustynne. W tej części pracy przedstawione zostały definicje związane z omawianą dziedziną, jak i ogólne strategie nawigacyjne tych mrówek. Zaprezentowane tu informacje odnoszą się do dotychczasowych wyników badań i obserwacji związanych z nawigacją opartą o obrazy.

Rozdział 3. przedstawia wirtualne środowisko mrówek pustynnych, szczegółową implementację algorytmów oraz przebieg samego badania. W tej części opisany został sposób wykonania sceny wirtualnego środowiska, doboru jej parametrów oraz trójwymiarowych modeli imitujących trawy i drzewa. Dodatkowo zdefiniowane zostały reprezentacje cech obrazów oraz metody ich porównywania. Na koniec opisano sposób pobrania obrazów ze sceny, a także metodologię wykonanych eksperymentów.

W rozdziale 4. zaprezentowane zostały wyniki uzyskane w ramach przeprowadzonych badań.

Rozdział 5. stanowi dyskusję dotyczącą otrzymanych wyników.

W rozdziale 6. zawarte zostały najistotniejsze wnioski oraz możliwe dalsze ścieżki rozwoju badań nawigacji opartej na rozpoznawaniu obrazów.

Rozdział 2

Nawigacja u mrówek pustynnych

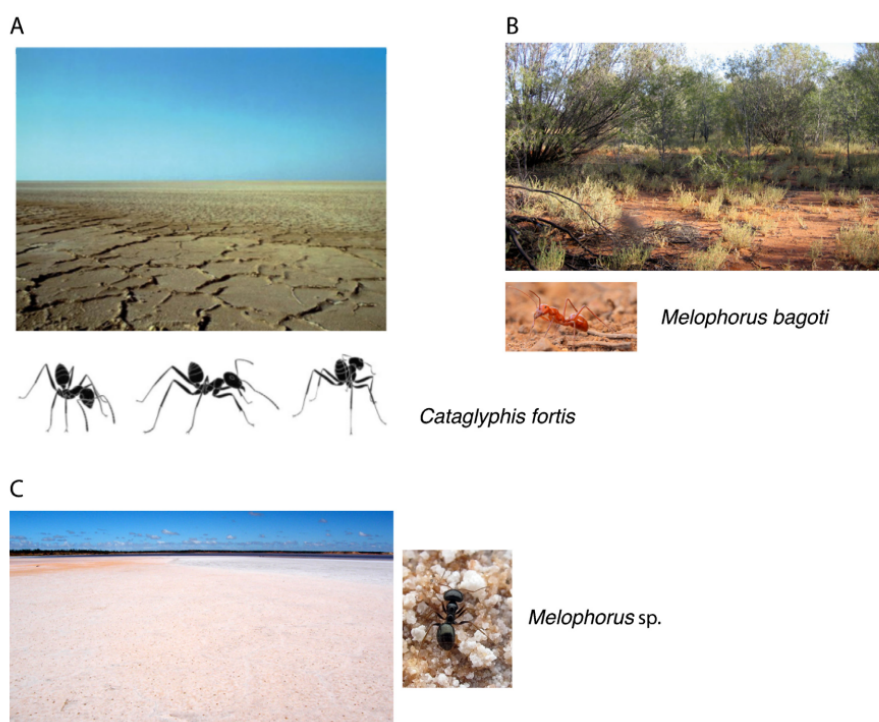
2.1 Główne strategie nawigacyjne

Ze względu na wymagające warunki środowiskowe mrówki pustynne stały się swoistymi ekspertami w nawigacji w znacznym stopniu opartej na rozpoznawaniu obrazów. Do tej pory przeprowadzono wiele badań testujących możliwości tych stworzeń, odkrywając schematy ich działania oraz wyciągając rozmaite wnioski. Wyniki tych badań pozwoliły zrekonstruować ogólny zestaw mechanizmów wykorzystywanych przez mrówki pustynne w kontekście nawigacyjnym.

Ograniczenia środowiska pustynnego

Wśród wielu spostrzeżeń dotyczących tych insektów przede wszystkim warto zwrócić uwagę na szereg podstawowych ograniczeń związanych z nawigacją na terenach pustynnych bądź okołopustynnych. Ze względu na wysokie temperatury, sięgające powyżej 40° Celsjusza, rozwiązania typowe dla mrówek leśnych oparte na oznaczaniu drogi, na przykład poprzez ślady feromonowe, nie znajdują tutaj zastosowania. Takie warunki znacznie ograniczają komunikację między mrówkami, a co za tym idzie ich współpracę. W efekcie, w przeciwieństwie do mrówek leśnych, gatunki pustynne znacznie częściej żerują samotnie.

Dodatkowo, warunki pustynne, w zależności od położenia geograficznego, mogą znacznie różnić się pod względem szaty roślinnej. Rysunek 2.1 przedstawia jak bardzo odmienne wizualnie może być otoczenie mrówki w zależności od jej geograficznego występowania. Mrówka *Melophorus bagoti* zamieszkuje tereny pustynne środkowej Australii, które porośnięte są tra-



Rysunek 2.1: Porównanie cech wizualnych różnych środowisk naturalnych charakterystycznych dla danego gatunku mrówki pustynnej. Źródło: [1]

wami i niskimi krzewami. Mrówki *Cataglyphis fortis* oraz znaczna część rodziny *Melophorus* zamieszkują natomiast solniska, występujące w Afryce oraz Australii, które całkowicie pozbawione są roślinności. Brak wyraźnych punktów odniesienia znacząco utrudnia możliwość rozróżnienia pozycji w środowisku. Mimo to mrówki rozwiązały problem nawigacji w tym surowym otoczeniu.

Dotychczasowe obserwacje

Do tej pory dokonano wielu obserwacji, badając różne aspekty nawigacji mrówki pustynnej.

Po pierwsze zauważono, że mrówki podążają swoimi wyuczonymi trasami — korytarzami, które pozostają unikatowe względem innych mrówek w mrowisku [2]. W ramach korytarzy może znajdować się wiele ścieżek [2]. Co więcej, pojedyncza mrówka może wykorzystywać wiele korytarzy prowadzących do tego samego punktu, a dokładne ścieżki w ramach korytarzy nie są zawsze takie same [2, 3].

Po drugie, wyniki badań wskazują na to, że mrówki nie posiadają umiejętności reprezentacji ścieżek w formie mapy. Wiedza mrówki na temat ścieżki wydaje się ściśle powiązana

z informacją kontekstową na temat kierunku ruchu: od (ang. *outbound*) lub do (ang. *inbound*) mrowiska. Znajomość ścieżki prowadzącej od mrowiska do źródła pożywienia nie oznacza umiejętności wykorzystania tej ścieżki jako powrotnej do mrowiska [2, 4].

Po trzecie, mrówki wykorzystują informacje wzrokowe podczas nawigacji między mrowiskiem a punktem, w którym znajduje się pożywienie. Co więcej, mechanizm oparty na informacjach wzrokowych nie jest jedynym mechanizmem nawigacyjnym stosowanym przez mrówki i może być od innych mechanizmów do pewnego stopnia niezależny [2].

Po czwarte, mrówki, które nie posiadają wiedzy na temat otoczenia mrowiska, wykonują spacer w formie krótkich tras o charakterze spiralnym wokół mrowiska, zwane „*learning walks*” [5]. Takie zachowania zauważono u nowych mrówek w mrowisku jak i u doświadczonych mrówek, które zostały przeniesione do innego otoczenia.

Po piąte, mrówki podczas *learning walks* systematycznie wykonują obroty, ustawiając się w kierunku mrowiska [5, 6, 7]. Zastosowanie takiego mechanizmu pozwala mrówkom zebrać znaczną ilość informacji o otoczeniu mrowiska. Mając możliwość obserwacji mrowiska pod różnym kątem i z różnych pozycji, mrówka może zapamiętać cechy charakterystyczne dla różnych lokalizacji, z których przeprowadzone zostały obserwacje.

Zdobyte do tej pory informacje pozwoliły określić domniemane elementy składowe algorytmu zawiadującego nawigacją mrówki pustynnej. Zaproponowane zostały trzy podstawowe mechanizmy:

1. Całkowanie ścieżki (ang. *path integration*)

Mechanizm całkowania ścieżki polega na sumowaniu wykonywanych kroków i zapamiętaniu sumy jako globalnego wektora powrotnego do mrowiska. Wniosek o wykorzystaniu przez mrówki tego mechanizmu znalazł swoje korzenie w hipotezie o biologicznej możliwości sumowania wykonywanych przez mrówkę przemieszczeń. Zachowanie to miałoby pozwolić mrówce powrócić z dowolnie oddalonego punktu jedynie dzięki znajomości wektora globalnego. Umiejętność ta wymaga jednak sumowania wykonanych kroków z uwzględnieniem ich relatywnego kierunku względem mrowiska. Dalsze eksperymenty wskazały, że mrówka szacuje swoje ustawienie kątowe względem mrowiska, wykorzystując między innymi percepcję polaryzacji promieni słonecznych [8, 9]. Niestety, wraz ze wzrostem liczby wykonanych kroków, rozwiązanie oparte na całowaniu ścieżki podatne jest na akumulację błędów [10]. Skutkiem tego mrówka wykorzystująca jedynie ten

mechanizm mogłaby ominąć mrowisko w drodze powrotnej. Co więcej, w zależności od warunków środowiska, mrówka podatna jest także na wszelkie siły zewnętrznie zmieniające jej pozycję, jak na przykład podmuchy wiatru, które mogą przenieść mrówkę na względnie duże odległości [11] — takie przeniesienie sprawia, że wskazania wektora globalnego przestają być aktualne.

Pomimo wszelkich wad tego rozwiązania, warto jednak zaznaczyć, że mechanizm ten dobrze sprawdza się na niewielkich dystansach, gdzie skumulowany błąd jest znikomy.

2. Nawigacja oparta o widoki (ang. *view-based navigation*)

Jak łatwo można zauważyć, mechanizm całkowania ścieżki nie jest wystarczający do precyzyjnej nawigacji mrówki. Umiejętność rozpoznawania widoków przez mrówkę rozwiązuje problem błędu we wspomnianym mechanizmie. Wiele badań wykonanych na mrówkach wykazuje, że posiadają one umiejętność tworzenia skojarzeń widzianych obrazów z informacją sterującą powrotem do mrowiska. W ramach nawigacji opartej na rozpoznawaniu obrazów zaproponowano dwie grupy metod.

Pierwsza grupa metod, zwana naprowadzaniem względem punktów orientacyjnych (ang. *landmark guidance*), wymaga porównania aktualnie widzianego obrazu z obrazem zapamiętanym w określonym punkcie środowiska, najczęściej w pobliżu mrowiska, i posłużeniu się różnicą pomiędzy tymi dwoma obrazami do oszacowania kierunku ruchu w stronę tego punktu.

Druga grupa metod, zwana nawigacją w oparciu o wektory lokalne (ang. *local vector navigation*), bazuje na hipotezie, że mrówki, wykonując *learning walks*, zapamiętują podczas obrotów obraz widziany z danej pozycji w kierunku mrowiska oraz w jakiś sposób kojarzą go z wektorem globalnym odpowiadającym tej pozycji [12]. Obrazy takie, dalej nazywane obrazami referencyjnymi, stanowią punkty odniesienia, dzięki którym mrówka może później skorygować swoją drogę. Rozwiązanie polega tym, że jeżeli mrówka dotrze do miejsca, w którym zapamiętała pewien obraz referencyjny, może wykorzystać skojarzoną z tym miejscem informację w celu określenia dalszego kierunku ruchu. Podczas pojedynczej drogi mrówki może wystąpić wiele takich miejsc, których rozpoznanie wpłynie na finalną ścieżkę do celu [13]. Działanie to uzasadnia różnorodność ścieżek, ponieważ proces nauki jest procesem ciągłym, wobec czego zapamiętane drogi mogą podlegać nieznacznym fluktuacjom [14].

Warto jednak podkreślić, że oba mechanizmy — całkowanie ścieżki i nawigacja oparta o widoki — działają do pewnego stopnia niezależnie i w zależności od różnorodności otoczenia mrowiska jeden z nich może przyjmować rolę dominującą [1, 15].

3. Systematyczne przeszukiwanie (ang. *systematic search*)

Ostatni mechanizm stanowi działanie awaryjne w sytuacji, gdy mrówka uzna, że się zgubiła. W tym przypadku mrówka rozpoczyna poszukiwanie obszarów, które rozpoznaje. Przeszukiwanie, w uproszczeniu, ma charakter zwiększających się pętli wokół punktu początkowego z powrotami do tego punktu. Co ważne, specyfika działania różni się nieznacznie w sytuacji, gdy mrówka zgubi się w okolicy mrowiska czy dobrze znanego punktu z pożywieniem, a w dużej odległości od wcześniej znanych miejsc [1].

Badania, podczas których przenoszono mrówki w zupełnie nowe, nieznanne dla nich miejsca, wskazują kluczowe aspekty warunkujące rozpoczęcie procesu poszukiwań. Środowisko naturalne w znaczącym stopniu wpływa na możliwość rozpoznania otoczenia jako nieznanego, a co za tym idzie na możliwość zastosowania mechanizmu systematycznego przeszukiwania. Zauważono, że mrówki, które zasiedlają tereny pustynne, nieposiadające żadnych punktów charakterystycznych, po przeniesieniu znacznie dłużej podążają za wektorem powrotnym zanim uznają, że się zgubiły — w przypadku środowiska o znikomej ilości odróżnialnych punktów charakterystycznych mrówki przemieszczają się aż do całkowitego zredukowania wektora powrotnego. Z kolei gatunki zamieszkujące środowiska o bogatszej szacie roślinnej rozpoczynały systematyczne poszukiwania już po osiągnięciu 50–75% długości wektora powrotnego, czyli nawet dwukrotnie wcześniej w porównaniu z mrówkami zamieszkującymi środowiska pozbawione roślinności. Wyniki te potwierdzają, że ilość wizualnie różnych punktów charakterystycznych ma krytyczne znaczenie dla nawigacji mrówki pustynnej, oraz pozwalają dostrzec zależność pomiędzy mechanizmami całkowania ścieżki a nawigacją opartą o widoki [1].

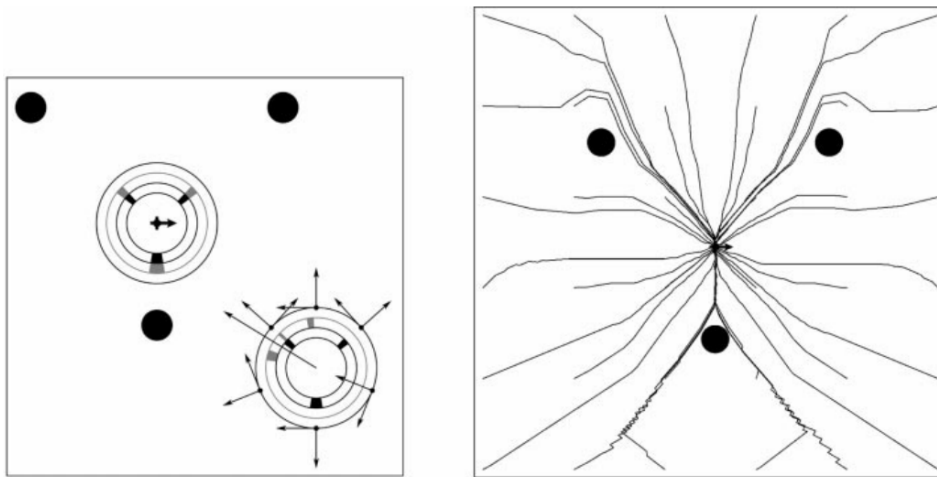
2.2 Reprezentacje i porównywanie obrazów

Mechanizm nawigacji oparty na rozpoznawaniu obrazów stanowi najbardziej intrygujący aspekt całego algorytmu nawigacyjnego stosowanego przez mrówki pustynne. W procesie tym można wyodrębnić dwa zagadnienia. Pierwsze z nich to sposób reprezentacji obrazu, czyli sposób, w

jaki obraz widziany przez mrówkę zostaje przetworzony w efektywny zbiór informacji zwany cechami obrazu. Drugie zaś dotyczy bezpośrednio sposobu porównania informacji pochodzących z dwóch obrazów i określenia miary ich podobieństwa.

Zainteresowanie sposobami, w jakie oba te zagadnienia realizowane są przez mrówki pustynne, przekształciło się w szereg badań naukowych. Celem wykonywanych analiz było uzyskanie informacji o tym, w jaki sposób otoczenie wpływa na decyzje podejmowane przez mrówkę. W badaniach wykorzystywano możliwości przysyłania części panoramy widzianej przez mrówkę oraz możliwość zmiany lokalizacji mrówki. W wyniku wielu obserwacji dotyczących zachowań mrówki w zmodyfikowanym otoczeniu zidentyfikowano wiele możliwych źródeł kluczowych informacji wizualnych.

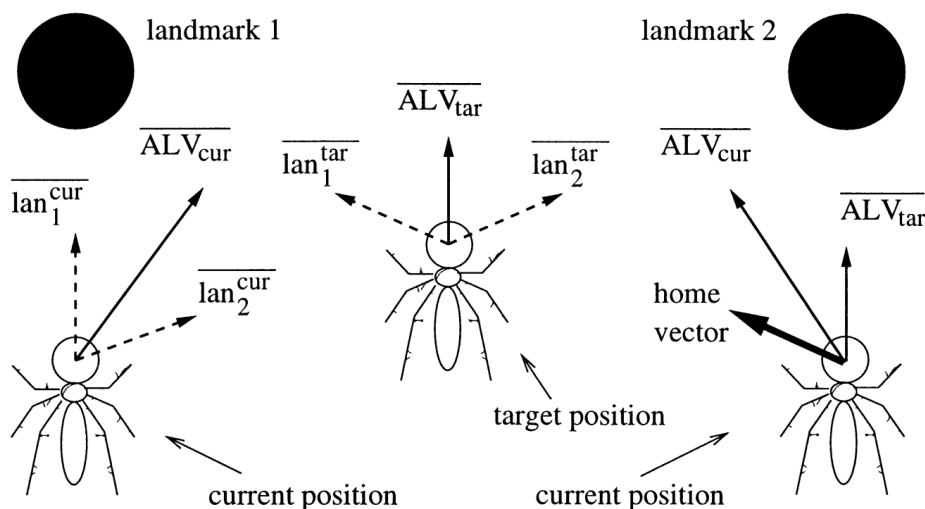
Pierwszym kluczowym wnioskiem było spostrzeżenie, że mrówki prawdopodobnie zapamiętują obrazy w formie relatywnie nieprzetworzonych widoków — tzw. migawek (ang. *snapshots*). Ponieważ w ramach mechanizmu nawigacji opartej o widoki mrówka porównuje aktualnie widziany obraz z zapamiętanymi wcześniej widokami, manipulacje najbardziej znaczącymi obiektami w otoczeniu mrówki pozwoliły określić, jakie cechy obrazów wykorzystywane są do porównań. W efekcie powstała koncepcja modelu nawigacyjnego opartego na obliczaniu sumy wektorów wynikających z przesunięć obiektów między obrazami. Model ten nazwano modelem migawkowym (ang. *snapshot model*) ze względu na charakter porównywania aktualnego widoku z zapamiętaną migawką. Model został zaklasyfikowany do metod nawigacyjnych względem punktów orientacyjnych. Rysunek 2.2 przedstawia graficzną interpretację modelu. Model zakłada, że obraz pochodzący z receptorów wzrokowych mrówki składa się z dwóch rodzajów sekcji: sekcje pełne odpowiadają miejscom, gdzie znajdują się obiekty, zaś sekcje puste to miejsca, gdzie na obrazie nie widać żadnego punktu charakterystycznego. Algorytm opiera się o porównywanie sekcji w obrazie i obliczenie dla każdej z sekcji wektora odpowiedniego przesunięcia, a następnie zsumowaniu wyznaczonych w ten sposób wektorów przesunięć w wynikowy wektor wskazujący kierunek do celu. Przesunięcia obliczane dla każdej z sekcji tworzone są na podstawie dwóch informacji: przesunięcia kąтового między porównywanymi sekcjami oraz stosunku wielkości tych sekcji. Stosunek wielkości sekcji przekłada się na jeden z trzech możliwych wektorów odśrodkowych: w sytuacji, gdy aktualnie widziana sekcja jest większa niż sekcja referencyjna, wektor wskazuje do środka, odpowiadając ruchowi w tył, z kolei w przypadku, gdy aktualna sekcja jest mniejsza od referencyjnej, wskazuje on na zewnątrz, co odpowiada ruchowi w przód, natomiast gdy sekcje są dopasowane — wektor jest



Rysunek 2.2: Graficzna reprezentacja modelu migawkowego, wykorzystanego w konfiguracji z trzema punktami charakterystycznymi. Lewy panel przedstawia sposób obliczania wektora skierowanego w stronę mrowiska. Czarne koła obrazują pozycję dużych obiektów — punktów charakterystycznych w scenie. Schemat porównywania widoków został zaprezentowany dla dwóch pozycji: pierwszej, w której pobrano obraz referencyjny (centrum, pomiędzy czarnymi kołami) oraz drugiej — nieco oddalonej (prawy dolny róg). Wewnętrzny okrąg przedstawia cechy widoku zapamiętane w punkcie referencyjnym (obrazowane przez czarne kwadraty), natomiast środkowy okrąg przedstawia cechy aktualnego widoku (szare kwadraty). W punkcie referencyjnym widoki są całkowicie dopasowane. Wraz ze wzrostem odległości od punktu referencyjnego, jak widać w przypadku drugiej pozycji, wektory przesunięć pomiędzy odpowiadającymi sobie sekcjami tworzą sumarycznie wektor powrotny, który został zaprezentowany jako strzałka rozpoczynająca się w środku okręgów. Prawy panel przedstawia trajektorie powrotne do mrowiska otrzymane poprzez zastosowanie modelu migawkowego. Na obrazie widoczne są otrzymane ścieżki, które zaczynają się od różnych punktów początkowych w środowisku i kierują do mrowiska. Źródło: [10]

zerowy. Warto zauważyć, iż model zakłada, że brak idealnego dopasowania sekcji, niezależnie od skali różnicy, zawsze skutkuje wektorem jednostkowym. Wektor do celu określony jest zatem jako znormalizowana średnia ważona wektorów przesunięcia kątownego, gdzie wagą jest relacja wielkości sekcji [10]. Zastosowanie tego algorytmu posiada jednak pewne istotne wady. Przede wszystkim konieczna jest tutaj wiedza dotycząca pozycji azymutalnej, aby określić relacje między sekcjami. Kolejną wadą jest jednostkowe podejście do określania wektorów różnic wielkości między sekcjami [10].

Jako rozwinięcie tego podejścia zaproponowano rozwiązanie wprowadzające proporcję do długości wektorów odśrodkowych. W rozwinięciu tym długości wektorów obliczane są na



Rysunek 2.3: Graficzna reprezentacja metody *Average Landmark Vector*, wykorzystanej w konfiguracji z dwoma punktami charakterystycznymi. Czarne okręgi obrazują pozycję dużych obiektów — punktów charakterystycznych w scenie. Po lewej stronie przedstawiona została metoda obliczania uśrednionego wektora z wektorów wskazujących na poszczególne punkty charakterystyczne. Środkowa część obrazu pokazuje uśredniony wektor w punkcie referencyjnym. Po prawej stronie obrazu przedstawiono sposób porównywania uśrednionych wektorów w celu oszacowania kierunku obrotu mrówki. Źródło: [10]

podstawie stosunku wielkości obserwowanych sekcji. Podejście to pozwoliło rozwiązać problem ominięcia punktu docelowego w sytuacji, gdy porównywane obrazy nie są dokładnie takie same. Algorytm ten otrzymał nazwę *Proportional Vector Model* [10].

Pomimo wielu zalet założenia o możliwości zapamiętywania relatywnie nieprzetworzonych obrazów przez mrówkę techniki te są nie tylko bardziej złożone pamięciowo, ale ze względu na większą ilość danych do porównania — także bardziej złożone obliczeniowo. Odpowiedzią na te problemy było zaproponowanie algorytmu o nazwie *Average Landmark Vector*. Algorytm ten zakłada jedynie potrzebę zapamiętania uśrednionego wektora obliczonego na podstawie wektorów skierowanych w stronę charakterystycznych obiektów widzianych w danej lokalizacji. W kolejnym kroku poprzez porównanie owego zapamiętanego uśrednionego wektora z uśrednionym wektorem wyliczonym w oparciu o aktualnie widziany obraz można oszacować kierunek ruchu w stronę celu (rys. 2.3) [10].

Kolejnym podejściem do problemu nawigacji mrówek w terenie jest rozwiązanie oparte na zapamiętanej puli obrazów zarejestrowanych wzdłuż pokonywanej trasy [2]. Porównywaniu

podlega aktualnie widziany obraz oraz jego przekształcenia wynikające z różnych przesunięć kątowych (uzyskiwane w wyniku obracania się mrówki w miejscu) ze wszystkimi obrazami zapamiętanymi w puli. Najlepsze dopasowanie wyznacza tym samym kąt, o jaki powinna obrócić się mrówka, by dalej kontynuować wędrówkę. Proces ten jest cyklicznie powtarzany aż do momentu dojścia do celu. Dzięki takiemu podejściu wynikiem algorytmu jest trajektoria zbudowana z krótkich odcinków, która — jak wskazują badania — z sukcesem może doprowadzić do celu, nawet w środowiskach zupełnie odmiennych od tych naturalnych dla życia mrówki pustynnej [16].

Pozytywne wyniki zastosowania przedstawionych technik nawigacyjnych w symulacjach skutkują zatem pytaniem, w jaki sposób mrówka faktycznie reprezentuje obrazy oraz jaki algorytm ich porównywania najlepiej odzwierciedli mechanikę nawigacji tych stworzeń?

W literaturze zaproponowano różne możliwe cechy obrazów, z czego można wyróżnić cztery wymieniane najczęściej.

Pierwszą, najbardziej ogólną cechą były wszystkie informacje oryginalnie zawarte w obrazie, a zatem stanowiące całe widoki w nieprzetworzonej formie. Jak zauważono już na początku, stanowi to duże wyzwanie obliczeniowe.

Druga cecha wynika z wiedzy o wysokiej wrażliwości receptorów wzrokowych mrówki na fale o długości odpowiadające kolorom: niebieskiemu i zielonemu. Dzięki temu możliwość dokładnego rozpoznawania linii horyzontu jest bardzo prawdopodobna. Zatem bezpośrednią cechą obrazu są jednowymiarowe dane na temat kształtu linii horyzontu. Seria badań testująca zachowania mrówek w przeniesionej lokalizacji z odtworzeniem sztucznej linii horyzontu wskazuje, że owady te w dużym stopniu polegają raczej na holistycznej informacji o linii horyzontu niż na informacji o wysokościach punktów szczególnych [13].

Trzecią z zaproponowanych cech jest średnia jasność pikseli w kolejnych kolumnach obrazu. Cecha ta, podobnie jak linia horyzontu, uwzględnia holistyczny charakter informacji. Co więcej, ukierunkowana jest na utworzenie swoistego odzwierciedlenia linii horyzontu, opierając się jednak na stosunku elementów nieba względem ziemi [17].

Czwarta cecha to środek masy elementów widzianych przez mrówkę. Specyfika ta została pierwotnie zauważona podczas badań pszczoł, jednak kolejne badania wykazały, że mechanizm ten występuje powszechnie u wielu owadów. Ich wyniki wskazują, że mrówka posiada zdolność rozpoznania kierunku do celu poprzez abstrakcję środka masy widzianych obiektów na scenie. Przeprowadzono badania, które polegały na wytrenowaniu mrówki w otoczeniu zawierającym

jeden obiekt — punkt charakterystyczny, gdzie środek obiektu znakował miejsce docelowe mrówki [18]. Kolejne próby modyfikujące obiekt charakterystyczny poprzez zmianę jego szerokości nie wskazywały spadku poprawności nawigacji insekta. Warto zauważyć, że obiekt modyfikowano symetrycznie, a zatem zmianie nie ulegał środek masy. Wyniki te potwierdzają zdolność mrówki do nawigacji w oparciu o ekstrakcję środka masy, jednak nie przedstawiają użycia algorytmu w bardziej rozbudowanym środowisku.

Tak zdefiniowane cechy stanowią w kolejnym kroku dane, na podstawie których można oszacować kierunek ruchu. Rozwiązania wykorzystywane do porównywania obrazów mają za zadanie dostarczyć dwie kluczowe informacje: pierwsza z nich to informacja o oczekiwanym przemieszczeniu, druga zaś to informacja o docelowym obrocie mrówki.

Wiele publikacji z sukcesem ukazuje potencjał informacji zawartej w obrazie, wykorzystując różne techniki porównywania danych. W literaturze można odnaleźć dwie funkcje, które najczęściej stosowane są w symulacjach komputerowych z wybranymi z przedstawionych wcześniej cech i dają wysoce zadowalające wyniki.

1. Korelacja

Pierwszą propozycją porównywania obrazów jest wykorzystanie korelacji niecentrowanej określonej wzorem:

$$\Phi(S, T) = \frac{\sum_{i=1}^n S_i T_i}{\sqrt{\sum_{i=1}^n S_i^2} \sqrt{\sum_{i=1}^n T_i^2}} \quad (2.1)$$

gdzie:

S, T – reprezentacja obrazu w postaci jednowymiarowej tablicy danych.

W ogólnym przypadku wynikiem korelacji niecentrowanej jest wartość należąca do przedziału $\langle -1.0; 1.0 \rangle$, jednak w przypadku obrazów przedział ten zawęża się do $\langle 0.0; 1.0 \rangle$, jako że wszystkie elementy S i T są nieujemne. Wartość korelacji równa 1.0 odpowiada całkowitemu dopasowaniu obrazów i spada wraz ze spadkiem podobieństwa obrazów, co odpowiada sytuacji oddalania się od miejsca docelowego. Funkcja ta została wykorzystana między innymi w badaniu opisanym w artykule [17], gdzie została użyta w celu porównania obrazów reprezentowanych za pomocą linii horyzontu oraz średniej jasności kolumn obrazu. W ramach badania utworzono bazę obrazów, które zostały sztucznie wygenerowane w oparciu o przygotowane środowisko wirtualne. Obrazy zostały zebrane

w lokalizacjach odpowiadających punktom na hipotetycznej siatce pokrywającej scenę, które były rozłożone równomiernie w odległościach 0.1 m, oraz przy stałej orientacji kamery. Jak pokazały wyniki, dla obu reprezentacji uzyskano zbliżone rezultaty, jeśli chodzi o wielkości obszarów, w których obrazy były poprawnie rozpoznawane.

2. Root Mean Square Error

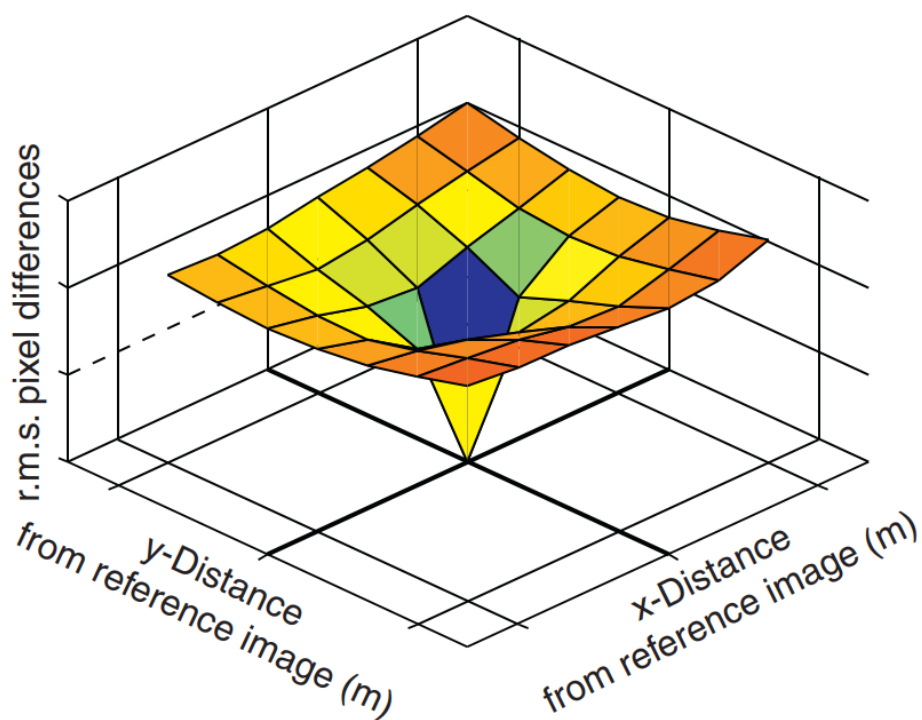
Drugą możliwością jest wykorzystanie funkcji pierwiastka błędu średniokwadratowego określonego wzorem:

$$RMSE = \sqrt{\frac{\sum_i^n (S_i - T_i)^2}{N}} \quad (2.2)$$

gdzie:

S, T – reprezentacja obrazu w postaci jednowymiarowej tablicy danych.

Wartość błędu równa 0.0 oznacza całkowite dopasowanie porównywanych obrazów. Badania przeprowadzone na zbiorze obrazów pobranych ze środowiska naturalnego mrówek w nieprzetworzonej formie oraz z wyekstrahowanymi z nich liniami horyzontu wskazują, że dane te są wystarczające do odtworzenia kierunku trasy mrówki [19]. Jak zostało to zaprezentowane na rysunku 2.4, w okolicy punktu referencyjnego błąd znacznie spada [20]. Obszar wokół punktu referencyjnego, w którym obraz został poprawnie rozpoznany, tworzy tzw. *catchment area*.



Rysunek 2.4: Graficzna reprezentacja wartości błędu w okolicy punktu referencyjnego. Wykres przedstawia wyniki porównywania widoków zarejestrowanych w pobliżu punktu referencyjnego z widokiem zarejestrowanym w tym punkcie przy użyciu funkcji pierwiastka błędu średniokwadratowego. Mniejsze wartości wskazują na lepsze dopasowanie obrazów. W punkcie referencyjnym wartość jest równa 0. Źródło: [20]

Rozdział 3

Materiały i metody

3.1 Symulacja komputerowa

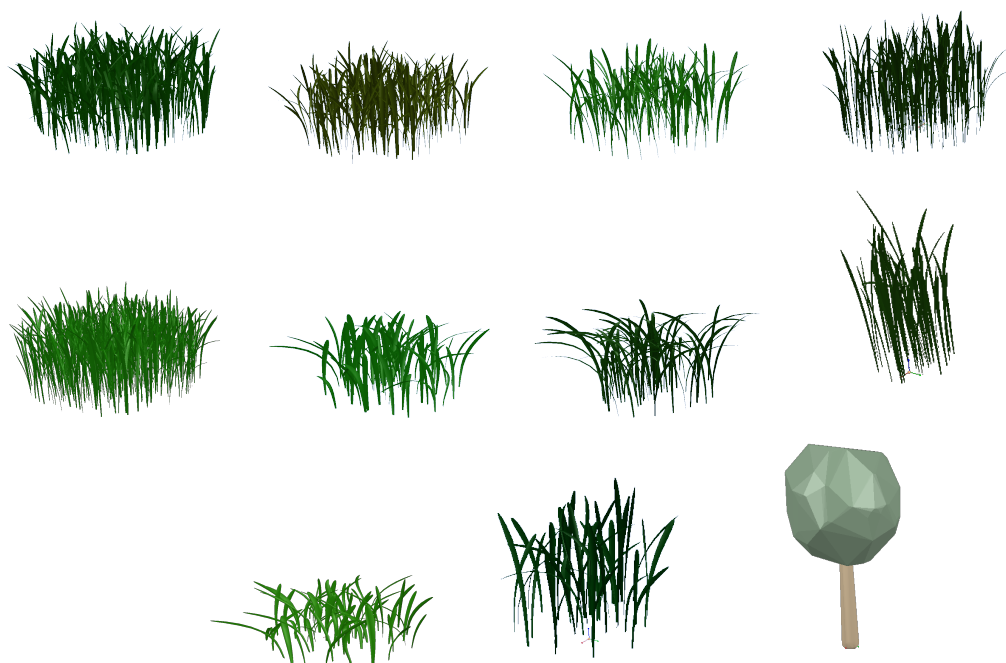
Pierwszym krokiem umożliwiającym wykonanie analiz domniemanych algorytmów nawigacyjnych wykorzystywanych przez mrówki pustynne było przygotowanie środowiska wirtualnego symulacji.

W celu przygotowania symulacji naturalnych warunków życia mrówki pustynnej wykorzystano narzędzie V-REP firmy Coppelia Robotics [21]. Narzędzie to przeznaczone jest do przeprowadzania dokładnych symulacji układów robotycznych. Głównymi zaletami oprogramowania V-REP jest możliwość zastosowania różnych silników fizyki, które charakteryzują się wysoką dokładnością obliczeniową. Co więcej, narzędzie wspiera wiele języków programowania i pozwala na zdalne sterowanie symulacją. Istotnym atutem narzędzia V-REP jest jego dostępność na wielu platformach systemowych.

3.1.1 Środowisko naturalne mrówki

Jak zostało powiedziane w rozdziale 2, środowisko życia mrówek pustynnych może znacznie się różnić ze względu na miejsce geograficzne ich występowania. Środowiska można podzielić ze względu na gęstość szaty roślinnej znajdującej się w okolicy mrowiska: w najbardziej surowych warunkach ilość roślinności jest znikoma — są to tzw. środowiska rzadkie, natomiast zupełnym przeciwieństwem tych warunków są tzw. środowiska gęste, które posiadają znaczną ilość różnorodnej roślinności.

Uwzględniając badania wskazujące, że mrówki w środowisku gęstym znacznie bardziej

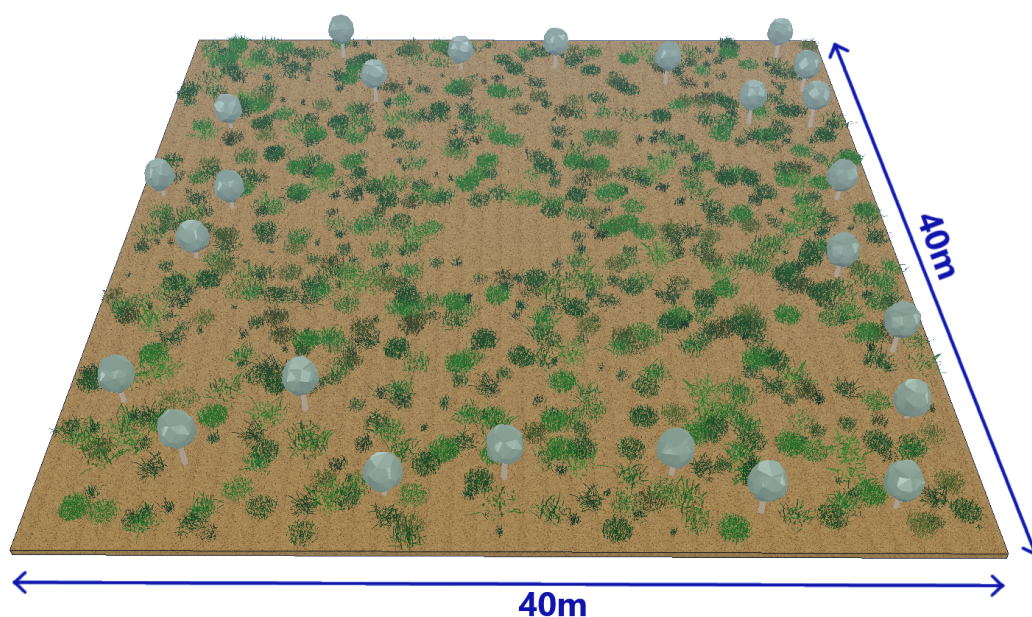


Rysunek 3.1: Trójwymiarowe modele roślinności wykorzystane w symulowanej scenie. Model drzewa dostępny jest bezpośrednio w narzędziu V-REP. Modele traw pochodzą ze źródła: <https://www.ronenbekerman.com/free-grass-model-texture>, Dostęp: 10.09.2018.

polegają na zapamiętanych widokach [1], podstawą niniejszej pracy było stworzenie symulacji takiego środowiska. W wyniku przeprowadzonego przeglądu literatury zawierającej opisy tego rodzaju środowisk zamieszkiwanych przez mrówki wyodrębniono zestaw najistotniejszych parametrów sceny otaczającej mrowisko. W kolejnym kroku utworzony został dedykowany program generujący scenę wirtualną w oparciu o podane parametry. Istotny element stanowi tu scena bazowa, która podlega zmianom aż do uzyskania finalnej postaci. Program, wykorzystując wspomnianą scenę, dokonuje odpowiedniej duplikacji i przemieszczeń obiektów, które się w niej znajdują. Wspominane obiekty są trójwymiarowymi modelami traw i drzew przygotowanymi do celów niniejszej pracy. Rysunek 3.1 przedstawia wszystkie wspomniane modele roślinności wykorzystane w badaniu.

Szczegółowa analiza informacji na temat rozkładu obiektów w otoczeniu mrowiska dostępnych w literaturze pozwoliła na właściwe dobranie parametrów sceny, dzięki czemu uzyskano bardzo dokładne odzwierciedlenie środowiska zamieszkiwanego przez mrówki pustynne. Parametry te przedstawione są poniżej w formie ogólnego opisu scenerii.

Scena została utworzona w formie kwadratu o boku długości 40 m (rys. 3.2). Na płasz-

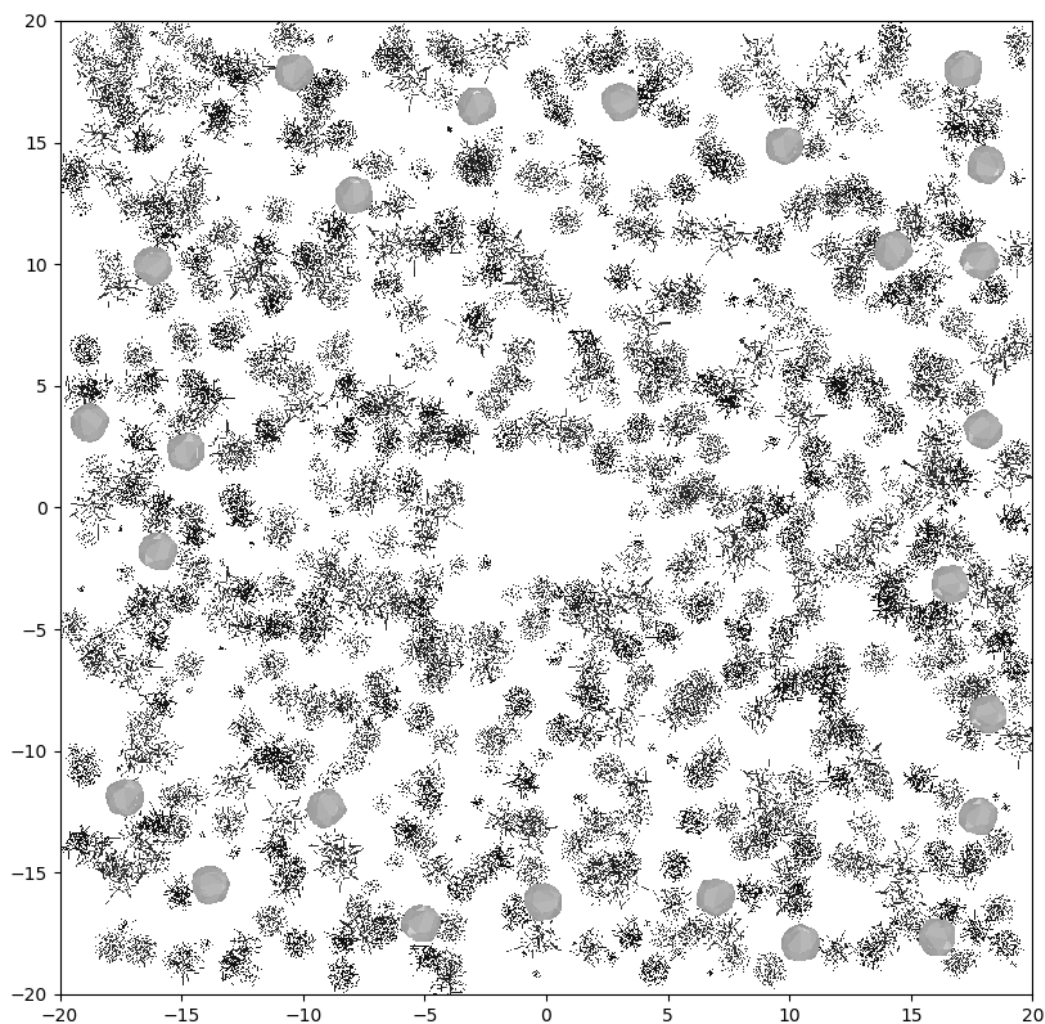


Rysunek 3.2: Wygenerowana scena symulująca naturalne warunki życia mrówki pustynnej w środowisku gęstym.

czyzną ziemi nałożona została tekstura imitująca piasek. Punktem centralnym sceny jest mrowisko, które jednakże nie posiada reprezentacji graficznej.

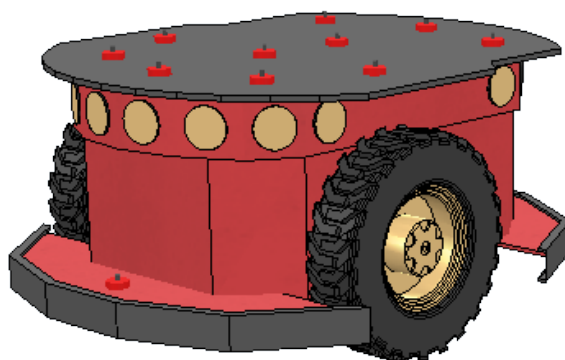
Najbliższa roślinność składająca się z niskich krzewów i traw znajduje się w odległości większej niż 3 m od mrowiska [22]. Najbliższe drzewa natomiast znajdują się w odległości większej niż 15 m, co stanowi wartość pośrednią obliczoną na podstawie dostępnych w literaturze opisów środowiska naturalnego mrówki [22, 2].

W środowiskach gęstych średnia liczba traw na jednostkę powierzchni wynosi 2 trawy/m². Oczekiwana całkowita liczba traw na scenie równa jest zatem 3200. Uwzględniając jednak problem złożoności obliczeniowej w przetwarzaniu każdego kroku symulacji oraz wygląd modeli traw, liczba ta została zredukowana do 800. Wartość ta została dobrana na podstawie szacunkowych obliczeń, które wskazują, że zaprezentowane modele roślinności niskiej są dwukrotnie większe pod względem szerokości i głębokości niż rzeczywista roślinność pustynna. Liczba drzew wysokich została ograniczona do 25. Ta wartość z kolei była inspirowana symulacjami przeprowadzonymi w ramach innych badań, w których zakładano występowanie łącznie 50 drzew wysokich i niskich krzewów [2]. Na potrzeby niniejszych badań niskie krzewy zostały zastąpione trawami, w wyniku czego przedstawiona wartość została ograniczona o połowę.



Rysunek 3.3: Widok sceny z góry prezentujący rozkład roślinności w otoczeniu mrowiska.

Całość roślinności została rozlokowana w sposób losowy, wykorzystując do tego celu rozkład równomierny, zgodnie z wynikami analizy dotyczącej rozkładu traw w środowisku życia mrówki *Melophorus bagoti* [14]. Podczas obliczania pozycji modeli traw i drzew na scenie przyjęto, że środki modeli traw muszą być odległe od siebie o co najmniej 0,5 m, środki drzew — o 3,5 m, a odległość między środkiem trawy a drzewem nie może być mniejsza niż 1 m. Dodatkowo algorytm zakłada, że drzewa nie mogą wystąpić bliżej niż 1 m od granicy sceny, a trawy — nie bliżej niż 0,5 m. Dane te zostały dobrane eksperymentalnie, uwzględniając wielkości modeli poszczególnych obiektów. Rysunek 3.2 prezentuje finalną scenerię symulacji, natomiast na rysunku 3.3 przedstawiono widok sceny z góry, który umieszczony został w układzie kartezjańskim, gdzie mrowisko jest punktem centralnym układu.



Rysunek 3.4: Trójwymiarowy model robota *Pioneer 3-DX* symulujący mrówkę w scenie.

3.1.2 Reprezentacja mrówki

W wirtualnym środowisku dodano również model imitujący mrówkę. W tym celu wykorzystano trójkołowy model robota mobilnego *Pioneer 3-DX* dostępny w narzędziu V-REP (rys. 3.4). Robot ten wyposażony jest w napęd typu różnicowego, który pozwala nim sterować. Co więcej, robot został odpowiednio dostosowany, aby jak najwierniej imitować mrówkę — w celu zachowania odpowiedniej relacji wielkości pomiędzy nim a sceną został on pomniejszony do wymiarów zbliżonych do faktycznych rozmiarów mrówki: jego wysokość została ustalona na 1 cm [23].

Do modelu został dodany sferyczny czujnik wizyjny o kącie widzenia równym 360° , pozwalający na pobranie panoramicznego widoku z pozycji robota. Czujnik został umieszczony na wysokości 0,74 cm i ustawiony równolegle do płaszczyzny robota. W kolejnym kroku dostosowane zostały jego parametry, aby były zgodne z aktualną wiedzą na temat właściwości układu wzrokowego mrówki.

Jak wskazują badania wykonane na mrówkach, rozdzielczość widzenia mrówki równa jest $3,7^\circ$, całkowity kąt widzenia w poziomie to 300° , zaś kąt widzenia w pionie — 76° [24]. Wiele wykonanych do tej pory symulacji komputerowych w celu uproszczenia obliczeń przybliżyła te wartości: w badaniach najczęściej przyjmowana jest rozdzielczość widzenia mrówki równa 4° , kąt widzenia w poziomie równy 360° bądź 300° oraz kąt widzenia w pionie — równy 76° [23, 25].

Uwzględniając powyższe źródła informacji, przyjęto następujące parametry jako referencyjne: rozdzielczość widzenia — 4° , całkowity kąt widzenia w poziomie — 360° oraz całkowity kąt widzenia w pionie — 76° . Wybrane parametry zostały bezpośrednio przełożone na cha-



Rysunek 3.5: Nieprzetworzony widok pobrany z czujnika wizyjnego symulowanej mrówki.

rakterystykę czujnika. Szerokość widzenia w pikselach równa jest zatem 90 px, a wysokość — 19 px. Ze względu na różne metody ekstrakcji cech z obrazu widoki w momencie pobrania nie ulegają przetworzeniu i pobierane są ze środowiska w postaci kolorowej (rys. 3.5).

3.2 Algorytmy rozpoznawania obrazów

Na podstawie dostępnej literatury i analizy przedstawionej w rozdziale 2 zdefiniowane zostały poniższe metody ekstrakcji cech oraz funkcje porównywania obrazów.

3.2.1 Wybrane cechy

W pracy postanowiono zbadać trzy metody ekstrakcji cech przedstawione poniżej. W przyjętych wzorach symbol A oznacza reprezentację macierzową obrazu w postaci kolorowej (RGB), symbol I odnosi się do jasności pikseli obrazu monochromatycznego, zaś jego indeksy dolne opisują wymiary macierzowej reprezentacji danych: oznaczenie x odnosi się do kolumn, a y — do wierszy macierzy. Symbole X oraz Y oznaczają natomiast całkowitą wielkość macierzy w danym wymiarze.

INTENS

INTENS (ang. *intensity*) jest najprostszą metodą ekstrakcji cech z obrazu opartą na możliwościach wzrokowych mrówki, polegającą na uzyskaniu jasności poszczególnych pikseli. Przyjmując pewne uproszczenie, bazujące na wiedzy o zdolności separacji nieba względem innych obiektów, obraz analizowano jedynie w oparciu o kanał niebieski. Uproszczenie to jest wynikiem analizy rozkładu jasności pikseli (w kanale niebieskim i zielonym) obiektów sceny (rys. 3.8). Metoda ekstrakcji cech *INTENS* została zdefiniowana wzorem:

$$INTENS = I_{x,y} = \phi(A) \quad (3.1)$$

gdzie:



Rysunek 3.6: Obraz widziany przez mrówkę pustynną uzyskany na podstawie cechy INTENS.



Rysunek 3.7: Wynik przetworzenia obrazu widzianego przez mrówkę na podstawie cechy AVGINTENS.

$$\phi : A_{RGB} \rightarrow A_B$$

Rysunek 3.6 przedstawia wynik przetworzenia obrazu zgodnie z powyższym wzorem.

AVGINTENS

Ekstrakcja cech obrazu na podstawie metody *AVGINTENS* (ang. *average intensity*) polega na obliczeniu średniej jasności pikseli w kolejnych kolumnach obrazu. Sposób ekstrakcji definiuje poniższy wzór:

$$AVGINTENS = \bar{I}_x = \frac{\sum_y I_{x,y}}{Y} \quad (3.2)$$

gdzie:

$$I_{x,y} = INTENS(A)$$

Y – liczba wierszy w pojedynczej kolumnie obrazu

Wynik ekstrakcji został przedstawiony na rysunku 3.7.

SKYLINE

SKYLINE (ang. *skyline*) jest metodą ekstrakcji cech opartą na obliczeniu pozycji linii horyzontu na obrazie. Działanie przekształcenia polega na obliczeniu pozycji punktu, który graniczy z niebem, dla każdej kolumny obrazu. Decyzja o klasyfikacji punktu jako przynależącego do linii horyzontu została oparta na wartości progowej. W przypadku zaklasyfikowania wielu takich punktów w jednej kolumnie linię horyzontu tworzy ten, który znajduje się najwyżej. Sposób ekstrakcji metodą *SKYLINE* został określony wzorem:

$$SKYLINE = \tau(I_x) \quad (3.3)$$

gdzie:

$$\tau : I_x \rightarrow \operatorname{argmax}(C(I_x))$$

$$C : I_x \rightarrow \begin{cases} 1, & \text{dla } I_{x,y} > c \\ 0 & \end{cases}$$

$$I_{x,y} = \operatorname{INTENS}(A)$$

c – wartość progowa

Wartość progę wynosi 220 i została dobrana arbitralnie na podstawie jasności obiektów sceny (rys. 3.8).

Wynik operacji *SKYLINE* został przedstawiony na rysunku 3.9.

3.2.2 Wybrane metody porównywania obrazów

W pracy postanowiono zbadać cztery metody porównywania obrazów przedstawione poniżej.

CORR

Pierwszą metodę porównywania obrazów stanowi korelacja (CORR, ang. *correlation*), która jest określona następującym wzorem:

$$CORR = \frac{\sum_{x,y} (I_{x,y} - \bar{I}) \cdot (J_{x,y} - \bar{J})}{\sqrt{\sum_{x,y} (I_{x,y} - \bar{I})^2} \cdot \sqrt{\sum_{x,y} (J_{x,y} - \bar{J})^2}} \quad (3.4)$$

gdzie:

I, J – reprezentacje cech porównywanych obrazów

X – liczba kolumn w pojedynczym wierszu obrazu

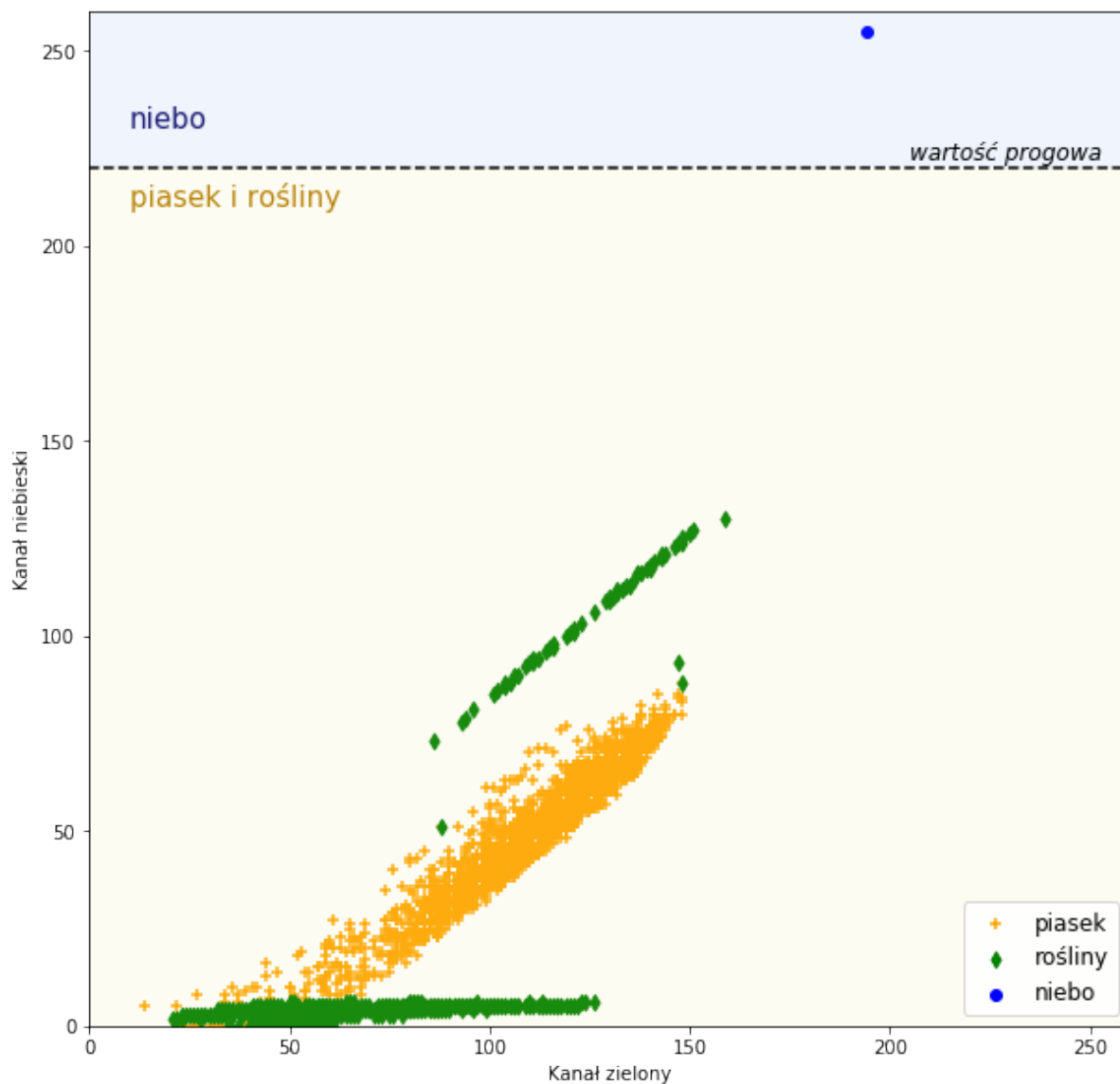
Y – liczba wierszy w pojedynczej kolumnie obrazu

RMSE

Drugą metodą jest błąd średniokwadratowy (RMSE, ang. *root-mean-square error*) określony poniższym wzorem:

$$RMSE = \sqrt{\frac{\sum_{x,y} (I_{x,y} - J_{x,y})^2}{X \cdot Y}} \quad (3.5)$$

gdzie poszczególne oznaczenia są takie same jak w przypadku korelacji.



Rysunek 3.8: Rozkład jasności pikseli obiektów sceny wirtualnej. Dane przedstawione na wykresie zostały obliczone na podstawie wybranych 8 obrazów pobranych w okolicy mrowiska. Przerywana linia przedstawia arbitralnie dobraną wartość progową separującą piasek i rośliny względem nieba.



Rysunek 3.9: Wynik przetworzenia obrazu widzianego przez mrówkę na podstawie cechy SKYLINE. Górna część rysunku przedstawia położenie linii horyzontu na obrazie. Dolna część rysunku przedstawia obliczony wektor cech, gdzie jasność piksela odwzorowuje wysokość położenia linii horyzontu w danej kolumnie obrazu, przy czym jaśniejsze punkty oznaczają wyższe położenia.

DOT

Trzecia metoda to iloczyn skalarny (DOT, ang. *dot product*), określony następującym wzorem:

$$DOT = \sum_{x,y} I_{x,y} \cdot J_{x,y} \quad (3.6)$$

gdzie poszczególne oznaczenia są takie same jak w przypadku korelacji.

COS

Ostatnią metodę stanowi podobieństwo kosinusowe (COS, ang. *cosine similarity*), które jest określone poniższym wzorem:

$$COS = \frac{DOT(I_{x,y}, J_{x,y})}{\sqrt{\sum_{x,y} I_{x,y}^2} \cdot \sqrt{\sum_{x,y} J_{x,y}^2}} \quad (3.7)$$

gdzie poszczególne oznaczenia są takie same jak w przypadku korelacji.

Zarówno iloczyn skalarny jak i podobieństwo kosinusowe to metody, które stanowią uproszczenie korelacji.

3.3 Przebieg badania

W celu oceny działania przedstawionych technik nawigacji przyjęto, że mrówka ma możliwość zapamiętania całych obrazów i nie posiada żadnych ograniczeń pojemności pamięciowej. Uproszczenie zakładające, że zapamiętane przez mrówkę obrazy nie podlegają żadnemu zaszumieniu bądź zniekształceniu, ograniczyło oddziaływanie czynników zewnętrznych na uzyskane

wyniki. Zabieg ten pozwolił skoncentrować się na zbadaniu wpływu wybranych algorytmów na poprawność rozpoznawania pozycji.

Do celów badań przygotowano dwa zestawy obrazów pobranych ze środowiska wirtualnego. Pierwszy zestaw stanowiły obrazy widziane przez mrówkę w równomiernie rozłożonych punktach znajdujących się na hipotetycznej siatce pokrywającej całe środowisko. Punkty takie określane będą jako „punkty siatki”, a zarejestrowane tam obrazy — jako *grid views*. Drugi zestaw to zapamiętane obrazy referencyjne zebrane w odpowiednio wybranych punktach środowiska zwanych „punktami referencyjnymi”. Obrazy te będą z kolei określane mianem *reference views*. Badania miały za zadanie sprawdzić, w jakiej odległości od punktu referencyjnego mrówka jest w stanie poprawnie rozpoznać swoją pozycję, bazując jedynie na zapamiętanych obrazach.

3.3.1 Dane

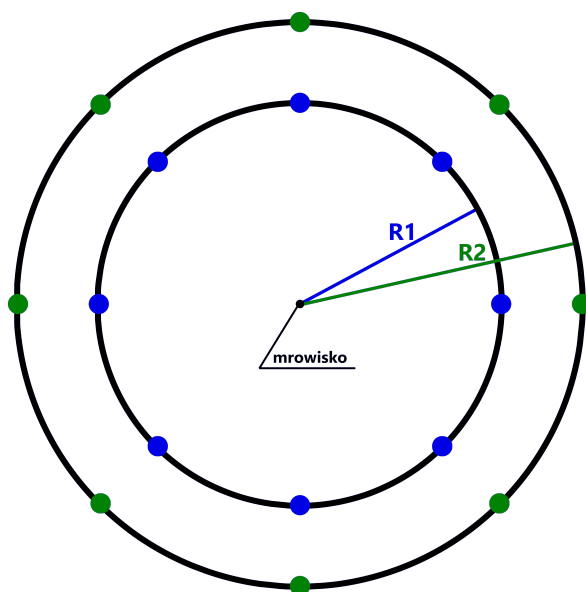
Podstawą badań wykonanych w ramach niniejszej pracy były dane zebrane z przygotowanego wcześniej środowiska wirtualnego. W ramach prac związanych z przedstawioną analizą przygotowany został program pobierający dane w formie obrazów widzianych przez mrówkę. Obrazy były pobierane za pomocą czujnika wizyjnego umieszczonego na robocie symulującym mrówkę. Wspomniane dane można podzielić na dwie zasadnicze części:

1. *Grid views*

Pierwsza część obejmuje obrazy zebrane w punktach rozłożonych równomiernie w formie kwadratowej siatki, której punktem centralnym było mrowisko. Wielkość siatki ograniczała przestrzeń o powierzchni 100 m², a obrazy pobierane były z krokiem 0,1 m w obu wymiarach. Należy podkreślić, iż wszystkie obrazy pobierane były w taki sposób, że mrówka skierowana była w stronę mrowiska. W wyniku działania programu wygenerowanych zostało 10201 obrazów.

2. *Reference views*

Drugą częścią danych były obrazy zabrane w pobliżu mrowiska, które reprezentowały obrazy zapamiętane podczas procesu nauki mrówki. Program umożliwiał pobieranie tych obrazów w różnych odległościach od mrowiska, jednak za każdym razem odpowiednie pozycje były rozłożone na dwóch koncentrycznych okręgach, w środku których znajdowało się mrowisko. Na każdym z okręgów pobierano obrazy z 8 równomiernie rozłożonych



Rysunek 3.10: Graficzna reprezentacja punktów referencyjnych względem mrowiska. Punkty wyróżnione na okręgach reprezentują pozycje, z których zostały pobrane obrazy referencyjne. Kolorem niebieskim oznaczone zostały punkty znajdujące się na bliższym okręgu, a zielonym — na okręgu dalszym.

punktów, co zostało zilustrowane na rysunku 3.10. Podobnie jak w przypadku *grid views* mrówka za każdym razem skierowana była w stronę mrowiska.

Do celów niniejszej pracy obrazy referencyjne zostały pobrane na okręgach o promieniu $R_1 = 0,5$ m oraz $R_2 = 1$ m. Przyjęte wartości inspirowane są badaniami dotyczącymi zakresami wędrówek mrówek w ramach *learning walks* [5]. Dodatkowo, przyjęte wartości zostały dobrane w taki sposób, aby nie nachodziły na linię pobliskiej roślinności. Całkowita liczba wygenerowanych obrazów to 16. Obrazy referencyjne na obu okręgach pobrane pod tym samym kątem względem mrowiska ułożone były współliniowo. Pobrane obrazy referencyjne zostały zaprezentowane na rysunku 3.11.

3.3.2 Badania *catchment areas* i *absolute areas*

Badania polegały na określeniu wielkości obszarów, na których zostały rozpoznane poszczególne obrazy referencyjne: pierwsze dotyczyło obszarów, na których każdy z obrazów referencyjnych został rozpoznany poprawnie, a drugie — obszarów, na których każdy z obszarów referencyjnych został rozpoznany bez względu na to, czy jego rozpoznanie w danym miejscu



Rysunek 3.11: Obrazy referencyjne pobrane ze środowiska wirtualnego. W lewej kolumnie umieszczone są obrazy pobrane na bliższym okręgu, a w prawej — na dalszym. W każdym wierszu znajdują się pary obrazów pobranych pod tym samym kątem względem mrowiska — obrazy współliniowe. Nazwy widoków referencyjnych zostały przedstawione w formacie zawierającym kolejno: identyfikator obrazu referencyjnego, pozycję x i y (względem mrowiska) oraz kąt obrotu robota.

było właściwe. W celu eliminacji danych, które mogłyby zostać zakwalifikowane jako fałszywie dodatnie (ang. *false positive*), obszar obu badań zawężono do odległości 2,5 m od mrowiska w każdym wymiarze. Odległość ta została dobrana w taki sposób, aby odpowiednie punkty siatki nie wychodziły na linię traw. Badania zostały wykonane niezależnie dla każdego zestawienia reprezentacji obrazu i metody porównywania obrazów.

W ramach pierwszego badania dla każdego *grid view* wyznaczono najpierw najlepiej dopasowany obraz referencyjny. Porównywano zatem aktualnie widziany obraz w danym punkcie siatki ze wszystkimi zapamiętanymi obrazami referencyjnymi. Wynikiem takiego porównania był numer najlepiej dopasowanego obrazu referencyjnego. W rezultacie utworzona została reprezentacja siatki, w której każdemu punktowi przypisany został numer pojedynczego, najlepiej dopasowanego obrazu referencyjnego. Następnie, wykorzystując tak otrzymaną reprezentację siatki, obliczono wielkości rejonów punktów referencyjnych (ang. *catchment areas*). W tym celu zastosowano algorytm *flood fill*, którego zadaniem było obliczenie liczby połączonych ze sobą elementów reprezentacji siatki o tej samej przypisanej wartości. Punktem początkowym poszukiwań był ten element reprezentacji siatki, który odpowiadał danemu punktowi referencyjnemu.

W ramach drugiego badania zbadano całkowitą wielkość obszarów, w których poszczególne obrazy referencyjne zostały rozpoznane, niezależnie od poprawności wyniku rozpoznania. Obliczenia wykonano dla każdego obrazu referencyjnego poprzez zsumowanie liczby wszystkich elementów reprezentacji siatki z przypisanymi wartościami odpowiadających temu obrazowi. Obszary takie w niniejszej pracy zostały określone jako „całkowite obszary dopasowania obrazów referencyjnych” (ang. *absolute areas*).

Wyniki uzyskane w ramach obu badań stanowiły zatem sumę odpowiedniej liczby elementów reprezentacji siatki dla każdego punktu referencyjnego. Zostały one na koniec zamienione na powierzchnię obszarów, wykorzystując informację o powierzchni odpowiadającej pojedynczemu elementowi.

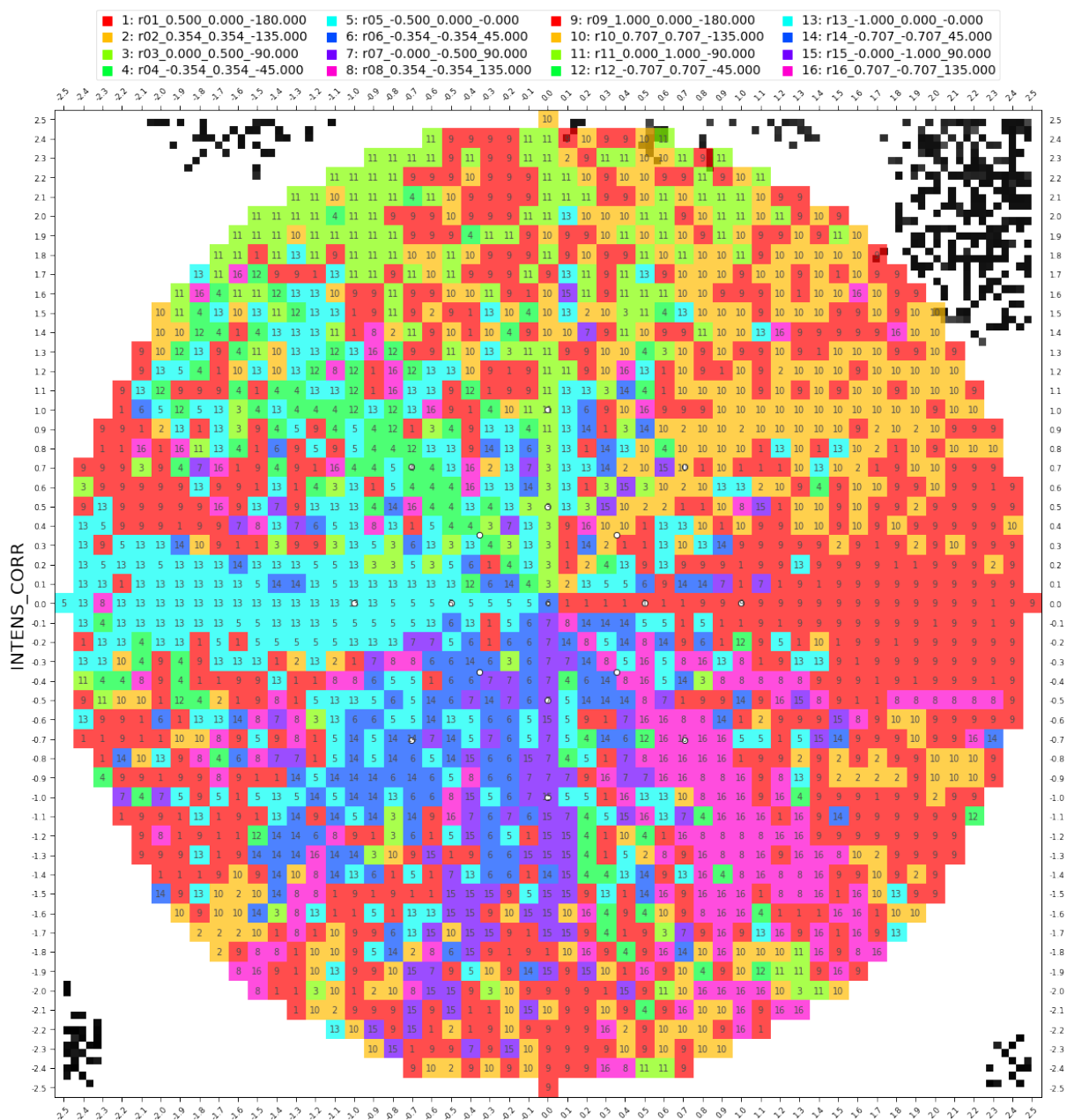
Rozdział 4

Wyniki

Wiele prac wskazuje na skuteczność metody korelacyjnej (CORR) zastosowanej w parze z całymi obrazami (INTENS). Zatem w pierwszym etapie przeprowadzono badania, które pozwoliły na ocenę tego rozwiązania.

Wykres 4.1 przedstawia obszar wokół mrowiska o wymiarach 5×5 m, gdzie mrowisko stanowi punkt centralny symbolizowany przez białą kropkę, natomiast punkty referencyjne zostały wyróżnione jasnymi punktami z czarną otoczką. Nazwy widoków referencyjnych zamieszczone na górze wykresu są zgodne z nomenklaturą przedstawioną na rysunku 3.11. Na wykresie zostały również uwzględnione trawy, które znalazły się w badanym obszarze — ich widok z góry został przedstawiony w postaci zbioru czarnych pikseli. Istotą wykresu jest pokazanie danych o najlepiej dopasowanych obrazach referencyjnych w sąsiedztwie mrowiska. Obrazy porównywane względem referencyjnych stanowiły tę część *grid views*, która pochodziła z części siatki ograniczonej okręgiem o promieniu $R = 2,5$ m względem mrowiska. Każdy punkt reprezentacji siatki zawiera numer najlepiej dopasowanego obrazu. Rozkład kolorów przypisanych do poszczególnych widoków referencyjnych został odpowiednio dobrany, aby odzwierciedlał pozycję kątową względem mrowiska. Obrazy współliniowe — pobrane z obu okręgów pod tym samym kątem względem mrowiska — reprezentowane są tym samym kolorem na wykresie.

Graficzna interpretacja statystyk dotyczących informacji przedstawionych na wykresie 4.1 została ukazana na rysunku 4.2. Rysunek ten przedstawia wielkości powierzchni, na których dopasowane zostały poszczególne obrazy referencyjne: panel A dla każdego obrazu referencyjnego przedstawia obszar, na którym dopasowania zostały uznane za poprawne, zaś panel B — całkowity obszar dopasowania danego obrazu referencyjnego. Innymi słowy, na panelu A



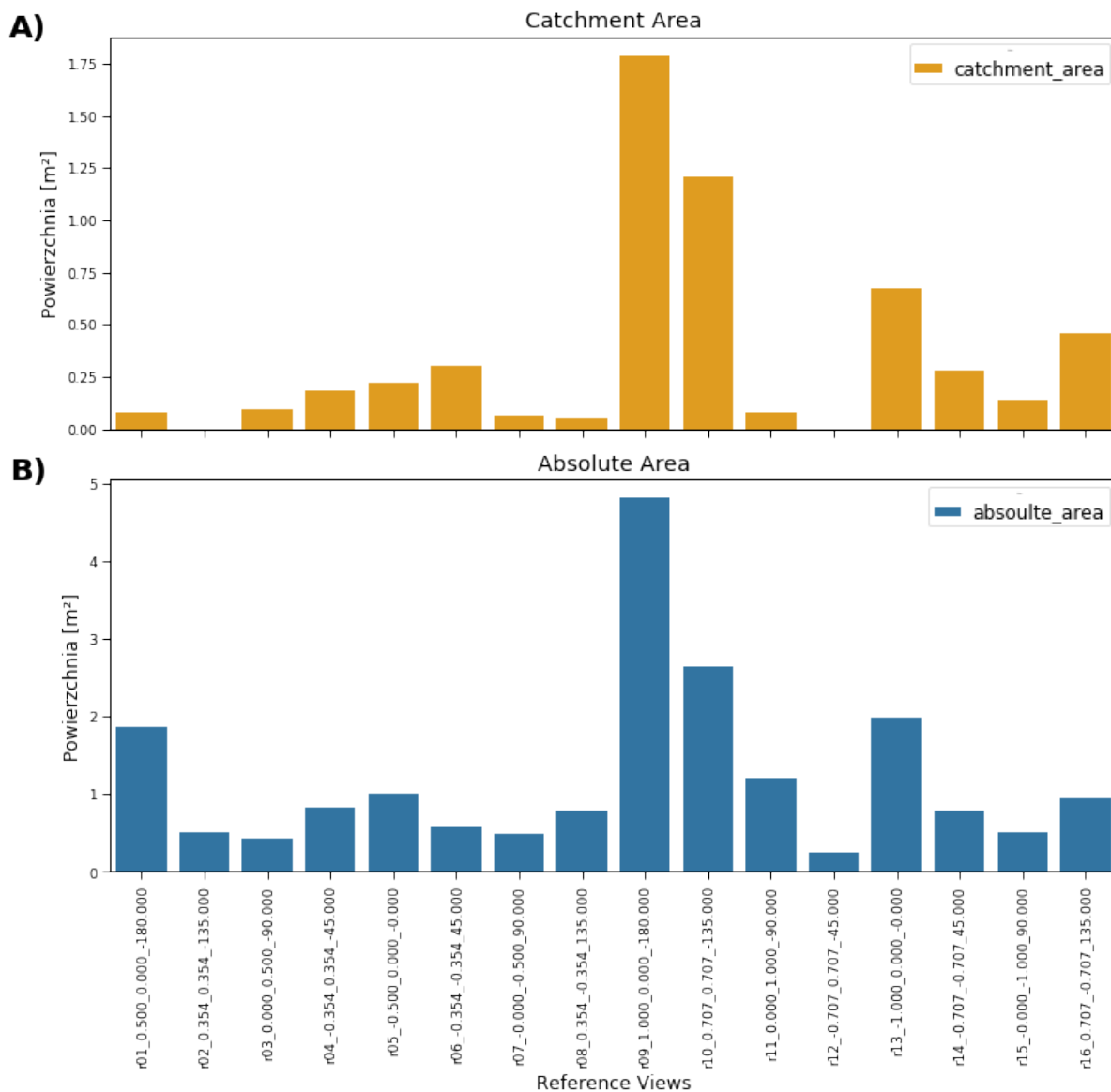
Rysunek 4.1: Graficzna reprezentacja rozkładu dopasowanych obrazów referencyjnych w punktach siatki wyznaczona dla reprezentacji INTENS i metody CORR. Numery przypisane poszczególnym elementom odpowiadają numerom najlepiej dopasowanych obrazów referencyjnych. Numerom odnoszącym się do współliniowych obrazów referencyjnych odpowiada ten sam kolor.

widać rozkład *catchment areas* dla każdego obrazu referencyjnego, natomiast na panelu B — rozkład *absolute areas*.

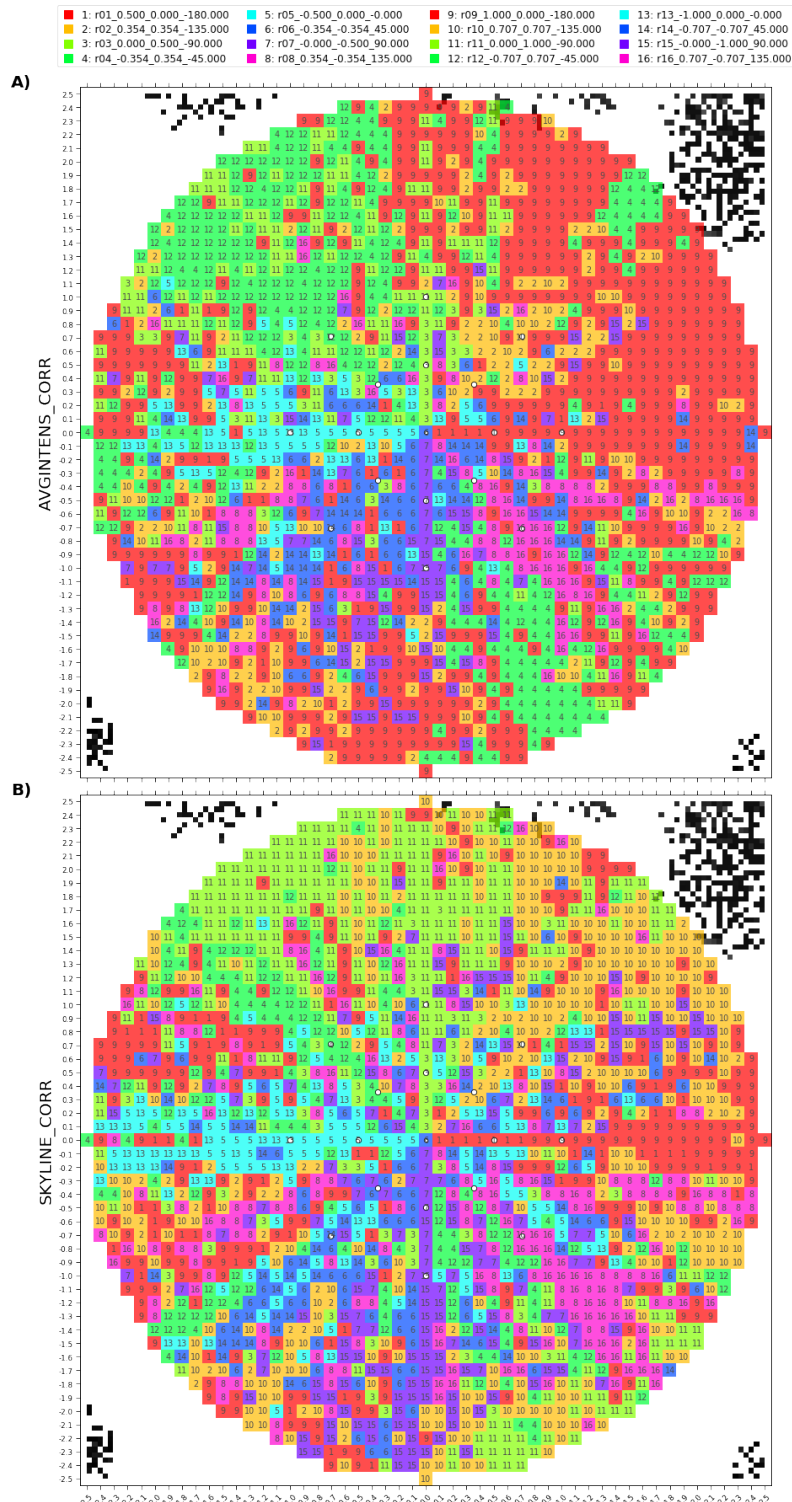
Dla widoku 'r09' wartość *catchment area* jest największa i sięga aż $1,75 \text{ m}^2$. Kolejna z wartości przypadła obrazowi 'r10' i jest mniejsza aż o $0,5 \text{ m}^2$ względem 'r09'. Można zauważyć, że całkowity obszar rozpoznania dla tych dwóch widoków jest również największy: dla obrazu 'r09' jest to aż 5 m^2 , co stanowi około 25% całego obszaru badań, zaś dla obrazu 'r10' wartość ta równa jest 3 m^2 (około 15% całego obszaru). Z kolei widoki, które były najgorzej rozpoznane, to 'r02' i 'r12' — ich *catchment areas* były równe zeru. Obrazy te były rozpoznane jedynie w nielicznych punktach siatki, które jednak nie tworzyły większej, spójnej całości (rys. 4.1).

W kolejnym kroku zbadano, jak zmienia się rozpoznanie widoków wokół mrowiska przy zastosowaniu metody korelacji z użyciem innych reprezentacji obrazów. Rysunek 4.3 przedstawia rozpoznane obrazy referencyjne w okolicy mrowiska przy wykorzystaniu reprezentacji AVGINTENS (panel A) i SKYLINE (panel B). Porównując oba wykresy na tym rysunku oraz wykres na rysunku 4.1, można zauważyć różnice w rozkładzie dopasowanych widoków: dla AVGINTENS dominują kolory czerwony i ciemno-zielony, dla INTENS — kolory czerwony i żółty, a dla SKYLINE — kolory czerwony, żółty i jasno-zielony, choć rozkład barw jest tu bardziej równomierny. Obszary poprawnego rozpoznania dla reprezentacji SKYLINE i INTENS są uformowane w kształcie lejków w większości sięgających granicy okręgu badań, natomiast w przypadku reprezentacji AVGINTENS można zauważyć, że obszary poprawnego rozpoznania występują jedynie w bliskiej okolicy obrazów referencyjnych, zaś granicy badań sięgają jedynie obrazy dominujące. Niezależnie od zastosowanej metody reprezentacji obrazu istnieją jednak pewne spójne obszary poprawnego rozpoznania obrazów referencyjnych.

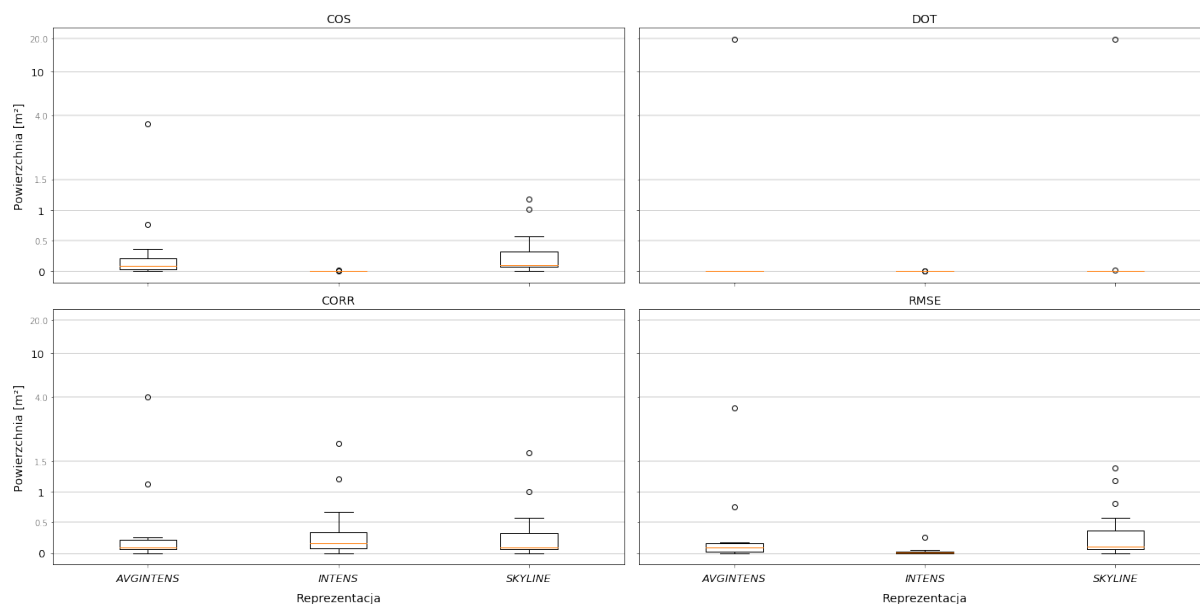
W celu wybrania najlepszej kombinacji reprezentacji obrazu i metody porównywania wykonano analizę rozkładu wielkości *catchment areas* dla każdego z widoków referencyjnych względem zarówno metody porównywania obrazów jak i sposobu ich reprezentacji. Wyniki dotyczące wszystkich możliwych zestawień reprezentacji obrazu i metody porównywania pokazano na czterech wykresach pudełkowych zamieszczonych na rysunku 4.4. Każdy z zamieszczonych tam wykresów przedstawia odpowiednie zakresy wielkości *catchment areas*. Wszystkie pudełka skonstruowane są w ten sposób, że właściwe pudełko ograniczone jest przez pierwszy i trzeci kwartył wielkości obszarów, mediana przedstawiona jest w postaci poziomej kreski w kolorze pomarańczowym, zaś wąsy sięgają ostatniej wartości, która mieści się w zakresie granicy właściwego pudełka powiększonej o 1,5-krotność rozstępu ćwiartkowego; elementy odstające



Rysunek 4.2: Wyznaczone wielkości obszarów występowania dopasowań obrazów referencyjnych dla reprezentacji INTENS i metody CORR. Panel A przedstawia wielkości *catchment areas*, zaś panel B ukazuje całkowite powierzchnie dopasowań obrazów referencyjnych na badanym obszarze — *absolute areas*.



Rysunek 4.3: Graficzna reprezentacja rozkładu dopasowanych obrazów referencyjnych w punktach siatki wyznaczona dla reprezentacji AVGINTENS (A) i SKYLINE (B) oraz metody CORR.



Rysunek 4.4: Porównanie wielkości *catchment areas* dla wszystkich rozważanych w niniejszej pracy reprezentacji obrazów i metod ich porównywania.

zostały z kolei oznaczone w postaci kropek.

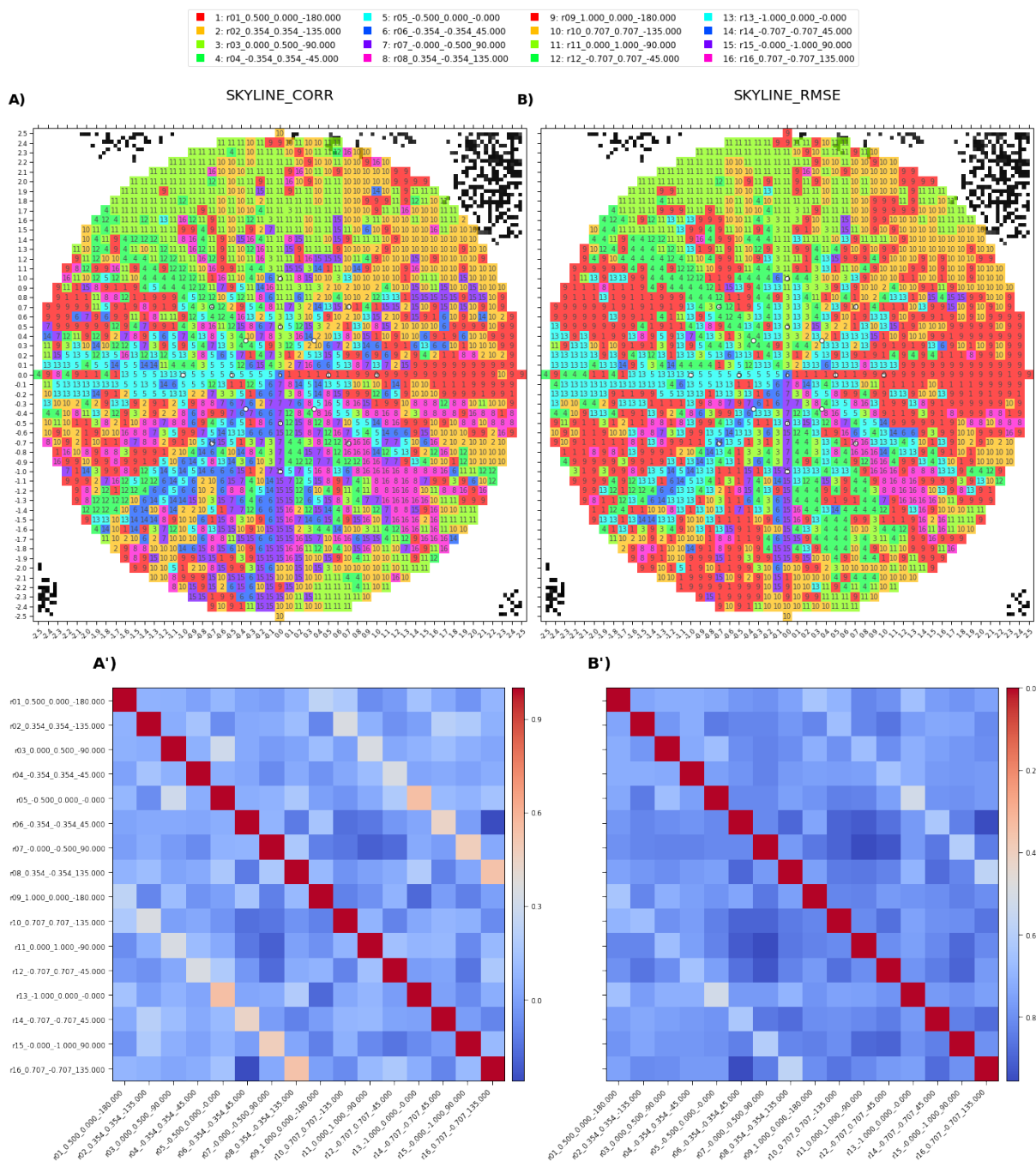
Na rysunku 4.4 widać, że metoda DOT osiąga najgorsze wyniki niezależnie od wybranego sposobu reprezentacji obrazu: wielkości *catchment areas* są zerowe bądź zawierają jedno dominujące dopasowanie, które stanowi przez to element odstający. W przypadku reprezentacji INTENS można zauważyć, że *catchment areas* osiągają największe wartości dla metody CORR, zaś dla innych metod uzyskują wartości bliskie zeru. Jeśli chodzi o reprezentacje AVGINTENS i SKYLINE, to reprezentacja SKYLINE osiąga większe bądź równe wielkości *catchment area* w porównaniu z reprezentacją AVGINTENS: są one większe dla metod RMSE i CORR, natomiast dla metody COS są one porównywalne. Co więcej, reprezentacja SKYLINE jest najbardziej stabilna: uzyskiwane przez nią wyniki są zbliżone bez względu na metodę porównywania obrazów. Niezależnie od wybranej reprezentacji oraz metody porównywania wartości mediany nie przekraczają powierzchni $0,5 \text{ m}^2$.

Dzięki stabilnej charakterystyce reprezentacji SKYLINE można było porównać metody RMSE i CORR. Metoda COS została pominięta w tym porównaniu ze względu na duże podobieństwo do metody CORR, natomiast metoda DOT została pominięta ze względu na bardzo słabe wyniki.

Na rysunku 4.5 przedstawione zostały obszary rozpoznania widoków w okolicy mrowiska dla obu wspomnianych metod. Panele A i A' dotyczą metody CORR, zaś panele B i B'

— metody RMSE. Na wykresach A' i B' zaprezentowano podobieństwa obrazów referencyjnych względem siebie przy użyciu danej metody porównywania. Wykresy na osiach posiadają nazwy poszczególnych obrazów referencyjnych. Każdy kwadrat przedstawia podobieństwo między dwoma takimi obrazami. Skala wartości podobieństwa (CORR) lub błędu (RMSE) została przedstawiona po prawej stronie wykresu. Kolorem czerwonym oznaczono największe dopasowanie, a kolorem niebieskim — najmniejsze.

Porównując wykresy A i B, można zauważyć, że obszary dopasowań dla obu metod są podobne. Analizując wykresy podobieństwa obrazów referencyjnych (A' i B'), można zauważyć, że istnieją pary różnych obrazów, które wykazują wysoki poziom podobieństwa. W metodzie CORR widać, że obrazy współliniowe pobrane pod tym samym kątem względem mrowiska wykazują większe relatywne podobieństwo niż w metodzie RMSE. Co więcej, oba wyniki cechuje podobny relatywny rozkład wartości, widoczny w postaci charakterystycznych wzorów na wykresie.



Rysunek 4.5: Porównanie wyników metod CORR oraz RMSE dla reprezentacji SKYLINE. Na panelach A i B przedstawiony został rozkład dopasowań obrazów referencyjnych w punktach siatki. Panele A' i B' przedstawiają natomiast obliczone podobieństwa obrazów referencyjnych względem siebie dla każdej z powyższych metod.

Rozdział 5

Dyskusja

Opisane w niniejszej pracy badania pozwoliły określić obszary dopasowań obrazów referencyjnych dla różnych reprezentacji obrazów oraz metod ich porównywania.

Wyniki uzyskane dla korelacji w powiązaniu z reprezentacją INTENS potwierdzają rezultaty opisane w innych artykułach. Zaobserwowane obszary dopasowań okazały się ciągle i skupione w pobliżu odpowiednich punktów referencyjnych. Oczekiwany wynikiem byłoby uzyskanie równego podziału całej przestrzeni na wycinki koła, gdzie każdemu punktowi siatki zostaje przypisany numer najbliższego względem niego punktu referencyjnego. Można zauważyć, na podstawie rysunku 4.1, że rozkład dopasowań dąży do tego oczekiwanego układu. Co więcej, poprawne dopasowania występują w punktach pobrania obrazu referencyjnego bądź z nim graniczą. Osiągnięty wynik jest zatem zgodny z postawioną w artykule [17] hipotezą dotyczącą spadku współczynnika korelacji wraz ze wzrostem przesunięcia.

Analiza wykresu 4.1 wraz z wykresem 4.2 wskazuje na wysoką dominację obrazów 'r09' i 'r10'. Jak można zauważyć, obrazy, które zostały błędnie rozpoznane, nie zostały pobrane w pobliżu punktów referencyjnych. Błędne rozpoznania zazwyczaj związane były z punktami siatki zlokalizowanymi pomiędzy parami współliniowych punktów referencyjnych — czyli na granicy oczekiwanych wycinków koła. Wynikać to może z powtarzalności struktury tekstury ziemi, która może być powielona w dwóch zupełnie odległych miejscach wirtualnego środowiska.

Analizując wykres 4.2, można zauważyć, że wielkości powierzchni *catchment areas* i *absolute areas* są mniejsze dla punktów referencyjnych znajdujących się bliżej mrowiska. Wynik ten jest uzasadniony rozkładem punktów referencyjnych: punkty leżące na bliższym okręgu

rozlokowane są na mniejszej przestrzeni i charakteryzują się mniejszymi odległościami między sobą, natomiast punkty leżące na dalszym okręgu cechują większe odległości między nimi, a to z kolei skutkuje większymi obszarami poprawnego dopasowania.

Na wykresie 4.3 przedstawiono obszary dopasowania obrazów referencyjnych dla różnych reprezentacji danych. Różnice w rozkładzie obszarów dopasowania wskazują na duże znaczenie wyboru reprezentacji danych. Jak można zauważyć, reprezentacja SKYLINE (wykres 4.3-B) cechuje się węższymi i bardziej jednolitymi *catchment areas* w porównaniu do reprezentacji AVGINTENS (wykres 4.3-A) i INTENS (wykres 4.1). Taka charakterystyka rozkładu może być w dużym stopniu związana z niezależnością reprezentacji SKYLINE względem jasności obrazu. Podobne zjawisko pokazano w artykule [17], gdzie stwierdzono, że w przypadku reprezentacji AVGINTENS znacznie trudniej jest rozróżnić obrazy — w artykule tym zauważono mianowicie, że współczynniki korelacji dla tej reprezentacji osiągały wysokie wartości niezależnie od porównywanych obrazów.

Na wykresie 4.3 widać, że w przypadku błędnego dopasowania obrazem dopasowywanym najczęściej był obraz 'r09'. Sugerować to może, że obrazy referencyjne są trudno rozróżnialne między sobą, co może być spowodowane niską wariancją jasności pikseli na niektórych fragmentach obrazu, takich jak niebo czy ziemia.

Podsumowanie wszystkich reprezentacji obrazów i metod ich porównywania zaprezentowane na wykresie 4.4 pokazuje różnice w wielkości *catchment areas*. Wyniki zaprezentowane dla metody DOT zawierają głównie zerowe wartości niezależnie od wybranej reprezentacji. Wyniki osiągnięte dla metody COS odbiegały za to w znacznym stopniu od wyników dla metody DOT, osiągając wartości mediany na poziomie $0,1 \text{ m}^2$. Metoda COS różni się od metody DOT występowaniem dodatkowych czynników powodujących normalizację danych. Powyższe wyniki wskazują zatem na duży wpływ normalizacji na poprawność rozpoznania. Co więcej, metoda CORR, różniąc się od metody COS dodatkowym centrowaniem danych, osiąga dla reprezentacji INTENS dużo lepsze wyniki niż metoda COS. Warto zauważyć, że reprezentacja INTENS uzyskuje niezerowe wyniki jedynie dla tej metody.

Szerokości pudełek w większości zestawień przedstawionych na wykresie 4.4 wskazują na dużą rozpiętość wyników. Jest to skutkiem dominacji jednych obrazów względem innych podczas rozpoznawania (tak jak było to wcześniej wspomniane w przypadku obrazu 'r09'). Reprezentacja SKYLINE okazała się tu najbardziej stabilna, osiągając zbliżone wyniki niezależnie od wybranej metody porównania. Zaprezentowane wyniki wskazują, że reprezentacja SKYLINE

jest najlepszym wyborem podczas porównywania metod rozpoznawania.

Wybór reprezentacji SKYLINE pozwolił na porównanie metod CORR i RMSE. Obie metody cechuje podobny charakter rozkładu obszarów rozpoznania (wykres 4.5-A,B). Na obu wykresach widoczne są charakterystyczne wąskie wycinki koła odpowiadające obszarom, gdzie obrazy zostały rozpoznane poprawnie. Wykresy 4.5-A',B' przedstawiają wysoką separację obrazów referencyjnych, co potwierdzają przedstawione powyżej rozkłady. Obie metody mają podobne możliwości rozróżnienia obrazów. Na wykresie 4.5-A' widoczne są dodatkowe, równoległe do przekątnej, linie o wyższym współczynniku podobieństwa między obrazami. Wynika to ze współliniowości widoków. W przypadku metody RMSE (wykres 4.5-B') linie te wskazują na mniejsze podobieństwo między obrazami współliniowymi w porównaniu do metody CORR. Z punktu widzenia problemu nawigacji opartego o wektory lokalne cecha ta jest bardziej pożądana ze względu na potrzebę rozpoznania konkretnego obrazu: ponieważ w tym podejściu obraz jest skojarzony z dokładną informacją o dalszym działaniu, niepoprawne rozpoznanie może znacząco wpłynąć na jakość nawigacji mrówki.

Rozdział 6

Wnioski

Mechanizmy nawigacyjne mrówek pustynnych stanowią inspirację w poszukiwaniu nowych technologii nawigacyjnych. Dzięki obserwacjom zachowań tych mrówek podczas powrotów do mrowiska sformułować można różne hipotezy dotyczące wykorzystywanych przez nie mechanizmów. Hipotezy te można następnie testować, badając działanie tych mechanizmów w symulowanym środowisku. Takie badania były właśnie przedmiotem niniejszej pracy.

Podsumowując, z pracy można wyciągnąć następujące wnioski:

- Wykorzystanie symulacji komputerowych pozwala w wysokim stopniu odwzorować środowisko naturalne mrówek pustynnych, jednocześnie pozwalając badać jedynie wybrane aspekty mechanizmów nawigacyjnych. Wykorzystanie wirtualnego środowiska jest znane z innych prac opisanych w literaturze. Przedstawione w niniejszej pracy wyniki zgodne są z rezultatami zaprezentowanymi w owych innych pracach. Potwierdza to słuszność wykorzystania symulacji komputerowych jako narzędzia badań, a zarazem ukazuje uniwersalność badanych mechanizmów.
- Wyniki przedstawione w niniejszej pracy można uznać za obiecujące. Wskazują one na trafność doboru reprezentacji obrazów i metod ich porównywania, co potwierdzają z kolei wyznaczone obszary poprawnego rozpoznania obrazów referencyjnych. Na wykresach przedstawiających dopasowanie obrazów widać, że pojawiają się jednolite obszary tworzące drogi skierowane w stronę mrowiska.
- Analiza wielkości rozpoznanych obszarów dla różnych reprezentacji obrazów i metod ich porównywania pokazuje, że ich odpowiedni wybór ma duże znaczenie. Z przeprowa-

dzonych badań widać, że reprezentacja SKYLINE osiąga wysokie wyniki niezależnie do metody porównywania obrazów. Co więcej, reprezentacja ta jest niezależna od jasności obrazu, co jest pożądanym aspektem ze względu na możliwe niewielkie zmiany jasności środowiska wraz z upływem czasu. Dodatkowo, reprezentacja SKYLINE w połączeniu z metodą RMSE pozwala na lepsze rozróżnienie obrazów niż w połączeniu z metodą CORR.

- Badania pokazują, że metoda DOT nie sprawdza się niezależnie od użytej reprezentacji obrazu. Wskazuje to na znaczący wpływ normalizacji danych na wyniki. Podobną wadę można zaobserwować w przypadku reprezentacji INTENS, która uzyskuje poprawne wyniki jedynie w zestawieniu z metodą CORR, czyli metodą charakteryzującą się centrowaniem danych względem wartości średniej.
- Niezależnie od wskazanych wad, większość kombinacji reprezentacji obrazów i metod ich porównywania osiąga satysfakcjonujące wyniki. Wartość mediany powierzchni *catchment areas* dla wielu metod jest bliska $0,3 \text{ m}^2$. Przyjmując, że obszar rozpoznania miałby kształt koła wokół punktu referencyjnego, jego promień przy takiej powierzchni byłby równy ok. $0,3 \text{ m}$. Wartość ta stanowi w przybliżeniu 30-krotność długości mrówki, co wydaje się zadowalającym rezultatem.
- Pomimo, że wyniki wyglądają obiecująco, na obrazach przedstawiających rozkład dopasowań wciąż można zauważyć błędne rozpoznania. Kolejnym krokiem, który pozwoliłby zwiększyć obszar poprawnych dopasowań obrazów, mogłoby być dodatkowe uwzględnienie pozycji kątowej mrówki. Takie podejście potencjalnie umożliwiłoby podczas porównań wczesne odrzucenie obrazów referencyjnych, które pobrane były pod kątem znacznie różniącym się od aktualnego kąta obrotu mrówki. Jako że stworzenia te posiadają receptory dostarczające informacje o pozycji kątowej, podejście to można byłoby uznać za uzasadnione.
- Kolejnym ciekawym aspektem, który należałoby w przyszłości rozważyć w badaniach, jest wpływ rozkładu obrazów referencyjnych. Warto przeprowadzić analizę, jak zmieniłyby się dopasowania, gdyby obrazy referencyjne nie były współliniowe lub gdyby było ich więcej.
- Dodatkowo, uwagę należałoby poświęcić również analizie, jak duży wpływ na osiągnięte

wyniki miało wygenerowane środowisko wirtualne. Szczególnie interesującym tematem jest wpływ rozkładu obiektów na scenie oraz zastosowanej tekstury ziemi na wyniki.

Spis rysunków

2.1	Porównanie cech wizualnych różnych środowisk naturalnych charakterystycznych dla danego gatunku mrówki pustynnej	6
2.2	Graficzna reprezentacja modelu migawkowego	11
2.3	Graficzna reprezentacja metody <i>Average Landmark Vector</i>	12
2.4	Graficzna reprezentacja wartości błędu w okolicy punktu referencyjnego	16
3.1	Trójwymiarowe modele roślinności wykorzystane w symulowanej scenie	18
3.2	Scena symulująca naturalne warunki życia mrówki pustynnej	19
3.3	Widok sceny prezentujący rozkład roślinności w otoczeniu mrowiska	20
3.4	Model robota symulujący mrówkę w scenie	21
3.5	Widok pobrany z czujnika wizyjnego symulowanej mrówki	22
3.6	Obraz widziany przez mrówkę po przetworzeniu na podstawie cechy INTENS	23
3.7	Obraz widziany przez mrówkę po przetworzeniu na podstawie cechy AVGINTENS	23
3.8	Rozkład jasności pikseli obiektów sceny wirtualnej	25
3.9	Obraz widziany przez mrówkę po przetworzeniu na podstawie cechy SKYLINE	26
3.10	Graficzna reprezentacja punktów referencyjnych względem mrowiska	28
3.11	Obrazy referencyjne wykorzystane w badaniach	29
4.1	Graficzna reprezentacja rozkładu dopasowanych obrazów referencyjnych w punktach siatki wyznaczona dla reprezentacji INTENS i metody CORR	32
4.2	Wyznaczone wielkości obszarów dopasowań obrazów referencyjnych dla reprezentacji INTENS i metody CORR	34
4.3	Graficzna reprezentacja rozkładu dopasowanych obrazów referencyjnych w punktach siatki wyznaczona dla reprezentacji AVGINTENS i SKYLINE oraz metody CORR	35

- 4.4 Porównanie wielkości *catchment areas* dla wszystkich reprezentacji i metod . . . 36
- 4.5 Porównanie wyników metod CORR oraz RMSE dla reprezentacji SKYLINE . . 38

Bibliografia

- [1] K. Cheng, P. Schultheiss, S. Schwarz, A. Wystrach, and R. Wehner, “Beginnings of a synthetic approach to desert ant navigation,” *Behav. Processes*, vol. 102, pp. 51–61, Feb. 2014.
- [2] B. Baddeley, P. Graham, P. Husbands, and A. Philippides, “A Model of Ant Route Navigation Driven by Scene Familiarity,” *PLoS Comput Biol*, vol. 8, Jan. 2012.
- [3] S. Sommer, C. von Beeren, and R. Wehner, “Multiroute Memories in Desert Ants,” *Proceedings of the National Academy of Sciences*, vol. 105, pp. 317–322, Jan. 2008.
- [4] R. Wehner, M. Boyer, F. Loertscher, S. Sommer, and U. Menzi, “Ant navigation: one-way routes rather than maps,” *Curr. Biol.*, vol. 16, pp. 75–79, Jan. 2006.
- [5] P. N. Fleischmann, M. Christian, V. L. Müller, W. Rössler, and R. Wehner, “Ontogeny of learning walks and the acquisition of landmark information in desert ants, *Cataglyphis fortis*,” *J. Exp. Biol.*, vol. 219, no. Pt 19, pp. 3137–3145, 2016.
- [6] M. Müller and R. Wehner, “Path integration provides a scaffold for landmark learning in desert ants,” *Curr. Biol.*, vol. 20, pp. 1368–1371, Aug. 2010.
- [7] P. N. Fleischmann, R. Grob, R. Wehner, and W. Rössler, “Species-specific differences in the fine structure of learning walk elements in *Cataglyphis* ants,” *J. Exp. Biol.*, vol. 220, no. Pt 13, pp. 2426–2435, 2017.
- [8] Wehner, Michel, and Antonsen, “Visual navigation in insects: coupling of egocentric and geocentric information,” *The Journal of experimental biology*, vol. 199, no. Pt, pp. 129–140, 1996.

- [9] M. Wittlinger, R. Wehner, and H. Wolf, “The ant odometer: stepping on stilts and stumps,” *Science*, vol. 312, pp. 1965–1967, June 2006.
- [10] D. Lambrinos, R. Möller, T. Labhart, R. Pfeifer, and R. Wehner, “A mobile robot employing insect strategies for navigation,” *Robotics and Autonomous Systems*, vol. 30, pp. 39–64, Jan. 2000.
- [11] A. Wystrach and P. Graham, “What can we learn from studies of insect navigation?,” *Animal Behaviour*, vol. 84, Apr. 2012.
- [12] M. Collett, T. Collett, S. Bisch-Knaden, and R. Wehner, “Local and Global Vectors in Desert Ant Navigation,” *Nature*, vol. 394, July 1998.
- [13] P. Graham and K. Cheng, “Ants use the panoramic skyline as a visual cue during navigation,” *Current Biology*, vol. 19, pp. R935–R937, Nov. 2009.
- [14] M. Kohler and R. Wehner, “Idiosyncratic Route-Based Memories in Desert Ants, *Melophorus bagoti*: How Do they Interact with Path-Integration Vectors?,” *Neurobiology of learning and memory*, vol. 83, pp. 1–12, Feb. 2005.
- [15] E. L G Legge, A. Wystrach, M. Spetch, and K. Cheng, “Combining sky and Earth: Desert ants (*Melophorus bagoti*) show weighted integration of celestial and terrestrial cues.,” *The Journal of experimental biology*, vol. 217, Oct. 2014.
- [16] B. Baddeley, P. Graham, A. Philippides, and P. Husbands, “Holistic visual encoding of ant-like routes: Navigation without waypoints,” *Adaptive Behaviour*, vol. 19, pp. 3–15, 2011.
- [17] K. Basten and H. Mallot, “Simulated visual homing in desert ant natural environments: Efficiency of skyline cues,” *Biological cybernetics*, vol. 102, pp. 413–25, Mar. 2010.
- [18] J. L. Woodgate, C. Buehlmann, and T. S. Collett, “When navigating wood ants use the centre of mass of a shape to extract directional information from a panoramic skyline,” *J. Exp. Biol.*, vol. 219, no. Pt 11, pp. 1689–1696, 2016.
- [19] A. Philippides, B. Baddeley, K. Cheng, and P. Graham, “How might ants use panoramic views for route navigation?,” *Journal of Experimental Biology*, vol. 214, pp. 445–451, Feb. 2011.

-
- [20] J. Zeil, “Visual homing: an insect perspective,” *Current Opinion in Neurobiology*, vol. 22, pp. 285–293, Apr. 2012.
- [21] “Robot simulator V-REP: create, compose, simulate, any robot..” url: <http://www.coppeliarobotics.com/>, dostep: 2018-07-10.
- [22] A. D. Dewar, A. Philippides, and P. Graham, “What is the relationship between visual environment and the form of ant learning-walks? An in silico investigation of insect navigation,” *Adaptive Behavior*, vol. 22, pp. 163–179, June 2014.
- [23] P. Ardin, F. Peng, M. Mangan, K. Lagogiannis, and B. Webb, “Using an Insect Mushroom Body Circuit to Encode Route Memory in Complex Natural Environments,” *PLOS Computational Biology*, vol. 12, p. e1004683, Feb. 2016.
- [24] S. Schwarz, A. Narendra, and J. Zeil, “The properties of the visual system in the Australian desert ant *Melophorus bagoti*,” *Arthropod structure & development*, vol. 40, pp. 128–34, Oct. 2010.
- [25] A. Wystrach, A. Dewar, A. Philippides, and P. Graham, “How do field of view and resolution affect the information content of panoramic scenes for visual navigation? A computational investigation,” *J. Comp. Physiol. A Neuroethol. Sens. Neural. Behav. Physiol.*, vol. 202, pp. 87–95, Feb. 2016.