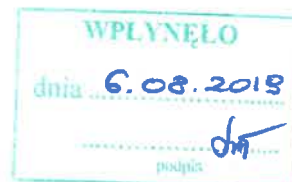


dr hab. inż. Leszek J. Chmielewski, prof. SGGW  
Wydział Zastosowań Informatyki i Matematyki  
Szkoła Główna Gospodarstwa Wiejskiego  
ul. Nowoursynowska 159, 02-776 Warszawa

29 lipca 2019 r.



## Recenzja rozprawy doktorskiej

autor rozprawy:

**mgr inż. Piotr Śmigielski**

Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej  
Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

tytuł rozprawy:

**System wizyjny dla autonomicznego robota  
poruszającego się w trzech wymiarach**

Recenzja została przygotowana w odpowiedzi na powołanie na recenzenta przez Radę Wydziału Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej Akademii Górniczo-Hutniczej im. Stanisława Staszica w Krakowie, zgodnie z uchwałą z dnia 25 kwietnia 2019 r., w związku z przewodem doktorskim Pana mgr inż. Piotra Śmigielskiego, prowadzonym w dyscyplinie informatyka. Promotorem rozprawy jest dr hab. Andrzej Bielecki, prof. AGH.

## 1 Zawartość rozprawy

### 1.1 Omówienie ogólne

Praca składa się z dwunastu rozdziałów i bibliografii.

Pierwszy rozdział to wprowadzenie do tematyki jakiej praca jest poświęcona, zawierające między innymi sformułowanie tezy pracy, oraz ogólny opis zrealizowanych rozwiązań. Tematyka pracy to zadanie nawigacji robota latającego, wyposażonego w jedną kamerę ze zmiennym położeniem obiektywu (na wprost i w dół), oraz czujniki położenia i wysokości na jakiej znajduje się robot. Nawigacja robota powinna zachodzić w możliwie dużym zakresie autonomicznie, na podstawie przede wszystkim obserwowanego otoczenia, w tym bez dostępu do mapy terenu i bez możliwości ustalenia swojego położenia niezależnie od informacji wizyjnej.

Rozdział drugi zawiera opis stanu wiedzy w dziedzinie analizy sceny na podstawie obrazów w zastosowaniach robotycznych.

W rozdziale trzecim znajduje się plan dalszych części pracy.

Rozdział czwarty zawiera omówienie procesu tworzenia wektorowej reprezentacji obiektów, których obraz robot pozyskał. Zakłada się, że obraz jest dwupoziomowy, czyli że segmentacja na plamy odpowiadające w przybliżeniu obiektom, została już przeprowadzona. Opisano kolejne etapy przetwarzania, od obrazu rastrowego, poprzez spis punktów brzegowych, wygładzanie i upraszczanie brzegów, po znajdowanie znaczących wielokątów.

W rozdziałach piątym i szóstym opisano dwie metody rozpoznawania plam opisanych jako wielokąty w formacie wektorowym, Metoda pierwsza wykorzystuje listy wierzchołków i boków wielokątów wraz z danymi o ich długościach i kątach między nimi, a druga punkty przecięcia pęku

półprostych wychodzących z centroidu z brzegiem obiektu. Każda z metod zapewnia niezależność rozpoznania od obrotu i skali.

W rozdziale siódmym przedstawiono metodę tworzenia opisu obiektu jako bryły, na podstawie obrazów, które robot uzyskuje dokonując odpowiedniego obrotu sceny. Dla obiektu wyznaczane są trzy charakterystyczne kierunki, a następnie obiekt jest fotografowany z tych kierunków. Dane z obrazów są integrowane i powstaje wektorowy opis bryły wraz z jej istotnymi szczegółami, takimi jak otwory, które mogą w dalszych etapach pozwalać na przemieszczanie się robota w ich wnętrzu.

Rozdział ósmy jest poświęcony opisowi modułu zapewniającego odnalezienie w mapie obiektów, wytworzonych w poprzednich etapach przetwarzania, tego obiektu który jest poszukiwany. Kluczem do tego jest graf bliskiego sąsiedztwa oraz poszerzony graf bliskiego sąsiedztwa. Grafy te są wykorzystywane hierarchicznie.

W rozdziale dziewiątym opisano moduł oprogramowania realizujący zadania nawigacyjne, zaś w rozdziale dziesiątym przedstawiono symulator, w którym większość zadań realizowanych przez robota (poza samą akwizycją obrazu w czasie lotu) można symulować programowo. Rozdział ten zawiera przykłady wyjaśniające działanie kolejnych modułów robota.

W rozdziale jedenastym przedstawiono konstrukcję robota latającego i towarzyszącego mu osprzętu, w którym zaimplementowano wszystkie opisane algorytmy i metody. Ten robot został poddany testom polowym, według dwóch scenariuszy, również tu opisanym.

Rozdział dwunasty zawiera konkluzje z wykonanych prac i podsumowuje całość rozprawy.

Bibliografia zawiera 92 pozycje, w znacznej części niedawne, choć znajdują się tam także pozycje klasyczne. Sześć z nich powstało ze współautorstwem Doktoranta.

## 1.2 Cel i teza pracy

Cel pracy jest postawiony następująco.

Celem, jaki przyświecał projektowi badawczemu przedstawionemu w niniejszej rozprawie, było opracowanie systemu wizyjnego umożliwiającego realizację złożonych zadań stawianych przed autonomicznymi robotami operującymi w trzech wymiarach, w szczególności latającymi. Do zadań tych należy rozpoznawanie obrazów, budowa grafowego modelu sceny oraz jego analiza jak też rozumienie oparte na zależnościach przestrzennych między obiektami znajdującymi się na niej, konstrukcja modeli trójwymiarowych tychże obiektów, lokalizowanie robota w przestrzeni oraz bezkolizyjna nawigacja w bliskim sąsiedztwie obiektów.

Autor zwraca też uwagę, że istnieje potrzeba badań nad systemami z jednym sensorem wizyjnym, aby była możliwość ich implementacji w małych robotach o ograniczonym udźwigu.

Postawiono następującą tezę:

Możliwe jest stworzenie zestawu efektywnych algorytmów strukturalnej analizy sceny w czasie rzeczywistym dla autonomicznego robota latającego wyposażonego w pojedynczą kamerę.

## 2 Omówienie treści i wyników rozprawy

### 2.1 Uwagi pozytywne

#### 2.1.1 Ważność zagadnienia

Spośród licznych zagadnień wynikających z rozwoju robotyki, kwestie działania przynajmniej częściowo autonomicznego wydają się być najbardziej istotne i jednocześnie trudne. Jest tak szczególnie wtedy, gdy autonomiczna jednostka jest niewielkich rozmiarów i ma ograniczone zasoby mocy obliczeniowej i w ogóle energii. Dlatego są ważne badania nad metodami, dzięki którym pewne wybrane działania, które w ogólności powinny być realizowane przez urządzenie wyposażone w duże zasoby inteligencji, mogłyby być wykonywane przez mniej zaawansowane urządzenia dzięki zastosowaniu przemyślanych uproszczeń i wykorzystaniu specyficznych właściwości danego zadania. Do takich zadań należy autonomiczna lub półautonomiczna nawigacja w obszarze znanym tylko częściowo, która jest przedmiotem recenzowanej pracy.

#### 2.1.2 Opis stanu wiedzy

Stan wiedzy został opisany bardzo szeroko i interesująco. Wskazuje to na dobrą orientację Autora w dziedzinie nawigacji robota latającego.

#### 2.1.3 Nowe oryginalne elementy

**Szybka segmentacja obrazu dwupoziomowego** W pracy założono, że obraz jest dany w postaci dwupoziomowego obrazu, zawierającego białe obiekty na czarnym tle lub białe krawędzie na czarnym tle. Doktorant zaznaczył, że jest to świadome, tymczasowe uproszczenie, które uniemożliwia traktowanie systemu jako więcej niż badawczego. Autor zaproponował metody wykorzystujące koncepcję wędrówki po obrazie wzdłuż brzegu między obszarem białym a czarnym lub wzdłuż białej krawędzi. Metody zapewniają wykrycie obszarów niekoniecznie wypukłym i z uwzględnieniem dziur w nich (do jednego poziomu dziur). Takie algorytmy należą do klasyki przetwarzania obrazu niskiego poziomu. Moim zdaniem ciekawa (choć spotykana w literaturze) jest propozycja autora, aby wędrówkę wokół obiektu realizować dość długimi skokami, dłuższymi niż jeden piksel, jak to było w metodach klasycznych. Opiera się to na założeniu, że obiekty są duże względem rozdzielczości obrazu, które nie musi lecz zwykle jest spełnione, i niewątpliwie przyspiesza przetwarzanie.

**Strukturalna metoda rozpoznawania obiektów – wielokąty** Metoda jest oparta na dokładnym porównywaniu liczby wierzchołków wielokątów, a następnie cyklicznym porównywaniu łańcuchów boków wielokątów, z uwzględnieniem ich długości i wzajemnych nachyleń, w poszukiwaniu najlepszego dopasowania. W tym sensie jest podobna do klasycznych metod porównywania wielokątów. Interesującym dodatkiem mającym pewne cechy oryginalności jest wykorzystanie odchylenia standardowego ilorazów długości boków; małe odchylenie wskazuje na jednolite wartości tych ilorazów, czyli na wspólną skalę przekształcenia. Różnice kątów zaś powinny być po prostu małe. Słabością tej koncepcji, z której Autor zdaje sobie sprawę, jest brak odporności na obecność małych nadmiarowych boków, jeśli segmentacja nie była dokładna, co skutkuje powstaniem różnicy w liczbie wierzchołków. Koncepcja ta została dwukrotnie opublikowana.

**Strukturalna metoda rozpoznawania obiektów – punkty brzegu** W tej metodzie korzysta się z długości serii odcinków tworzonych na pęku półprostych, przez ich wspólny początek w środku

ciężkości wielokąta i ich punkty przecięcia z konturem obiektu. Metoda porównywania jest podobna do poprzedniej, znowu szukamy cyklicznie sprawdzanego dopasowania wskazującego na jednolitą skalę jednego obiektu względem drugiego, i to jest zaleta metody. Założenie jest takie, że kąty są rozłożone na tyle gęsto, że podobieństwa się ujawnią, mimo że nie ma gwarancji, że półproste przecinają kontur w odpowiadających sobie punktach. Metoda jest ciekawa, ale ma słabości, które omówię w dalszej części recenzji.

**Konstrukcja trójwymiarowego modelu bryły** W metodzie wykorzystuje się trzy rzuty bryły, wybrane jako charakterystyczne: z góry, z przodu i z jednego boku. Kontury wyekstrahowane z rzutów, po ich uzgodnieniu, stanowią model obiektu. Koncepcja z pewnością działa dla obiektów dość dużych i mających takie cechy, jak posiadanie kształtu zbliżonego do graniastosłupa, posiadanie określonego frontu i boków, oraz podobieństwo widoków z frontu i z tyłu oraz z dwóch boków. Takie cechy zwykle posiadają budynki. Zaproponowano też uzupełnianie widoków przez cechy dodatkowe, które mają strukturę taką samą, jak podstawowy model, co tworzy załączek pewniej hierarchii (a uproszczenie polega na tym, że zatrzymujemy się na jej pierwszym stopniu). Zaletą koncepcji jest struktura na tyle prosta, że łatwo ją zalgorytmizować, i dla obiektów podobnych do budynków będzie ona dobrze działać. Trafne uproszczenia są bardzo cenne i to jest główna zaleta metody. Metoda została opublikowana. Pewne wady tej metody omawiam w dalszej części recenzji.

**Koncepcja grafu bliskiego sąsiedztwa, pierwszego i drugiego stopnia** Dużą zaletą pracy jest moim zdaniem zaproponowanie tej prostej a ekspresyjnej koncepcji do rozpoznawania obiektów (reprezentowanych jako wielokąty) wraz z ich sąsiedztwem. Graf bliskiego sąsiedztwa jest zdefiniowany z wykorzystaniem pojęcia odległości, ale bez parametrów, dzięki czemu nie ma problemu z rozmytym charakterem pojęcia „blisko”. Wartościowe jest także pojęcie rozszerzonego grafu bliskiego sąsiedztwa (przy okazji można by zaproponować bliskie sąsiedztwa dalszych rzędów). Wadą koncepcji jest wykorzystanie środka ciężkości obiektu, gdyż jest to cecha o małej odporności na błędy wyznaczenia wielokątów.

Koncepcja została opublikowana.

**Symulacja programowa części systemu** Praca na obiektach rzeczywistych wymaga nakładu pracy i środków i stwarza silne ograniczenia czasowe. Zastosowanie symulacyjnych metod do tych elementów konstrukcji robota, które nie wymagają realizacji w fizycznym sprzęcie, jest wielką zaletą pracy i wskazuje, że Doktorant zna i umie zastosować techniki efektywnego projektowania robotów.

#### 2.1.4 Badanie jakości wyników

Zadania, których rozwiązanie stanowi przedstawiony zbiór metod oraz działający robot latający, stanowią rozbudowaną hierarchię, od akwizycji obrazu, poprzez jego analizę i wykorzystanie do automatycznej budowy modeli otoczenia, użycia ich do rozpoznania obiektów otoczenia oraz do znalezienia własnego położenia robota, aż do wykorzystania tego opisu otoczenia do rozwiązania zadań lokalizacji obiektów i ich fragmentów oraz i nawigacji wśród nich. Doktorant przyjął dwie metody badawcze. Pierwsza, to testowanie niektórych modułów systemu robota w symulatorze programowym. Druga, to wykonanie konkretnych całościowych zadań nawigacyjnych przez fizycznego robota w świecie rzeczywistym.

Autor zrelacjonował pojedyncze testy wykonywania określonych zadań, zaś nie realizował serii testów aby przebadać kwestie błędów i niezawodności. Takie podejście jest uzasadnione tym



etapie badań, na jakim one się obecnie znajdują, i przy osiągniętym obecnie poziomie gotowości technologicznej.

Badania symulacyjne przeprowadzone w ramach pracy obejmowały następujące elementy: budowa modelu wektorowego na podstawie obrazu, wyodrębnianie i opis grup obiektów na podstawie modelu, rozpoznawanie modeli za pomocą algorytmu strukturalnego, wyznaczanie pozycji robota do wykonania trzech obrazów dla budowy modelu trójwymiarowego, budowa tego modelu, budowa grafowego modelu sceny z sąsiedztwami, zrozumienie sceny przez wykorzystanie reprezentacji obiektów i ich sąsiedztw, ustalenie jakie manewry nawigacyjne są możliwe na scenie. Rozumiem, że kroki polegające na fizycznym wykonaniu zdjęć z realnej sceny zostały tu zastąpione dostarczeniem zdjęć wykonanych osobno.

Opisane badania na działającym robocie w warunkach rzeczywistych obejmowały realizację dwóch określonych zadań, od startu robota i wykonania pierwszego ogólnego zdjęcia sceny z góry, poprzez segmentację tego zdjęcia (w wersji uproszczonej, jak zaznaczono powyżej, możliwej dzięki użyciu obiektów w bardzo kontrastowych kolorach), oraz wykonanie w całości zadań obejmujących te kroki, które były wykonywane w symulacji, przy czym realnie wykonywano przeloty i fizycznie pozyskiwano zdjęcia.

### 2.1.5 Publikacje

Doktorant jest współautorem sześciu publikacji w latach 2012-2017: dwóch w czasopismach z listy A i czterech prac konferencyjnych. Artykuły ukazały się w *Expert Systems with Applications*, lista A, 35 punktów, proponowana liczba punktów w nowym systemie oceny 140, oraz w *International Journal of Advanced Robotic Systems*, lista A, 20 punktów, w nowym systemie proponowane 40 punktów. Cztery publikacje ukazały się w pracach konferencyjnych, z których trzy są wydane przez wydawnictwo Springer i jedna przez IEEE – tej doktorant jest pierwszym autorem. Wszystkie są notowane w bazie Web of Science.

## 2.2 Uwagi dyskusyjne i krytyczne

### 2.2.1 Uwagi merytoryczne

**Strukturalne metody rozpoznawania** Wadę metody pierwszej polegającej na błędnym odrzuceniu wielokątów o różnych liczbach wierzchołków gdy różnica wynika z błędów w segmentacji można pokonać stosując metody niedokładnego dopasowywania grafów (ang. *inexact graph matching*). Literatura tego zagadnienia jest ogromna, podam tylko jedną pozycję [BJ00]. Tu graf jest bardzo prosty, więc zadanie nie jest bardzo trudne.

Co do metody wykorzystujące punkty na brzegu obiektu, to są dwa zastrzeżenia. Po pierwsze, punkt odniesienia będący środkiem ciężkości brzegu nie jest odporny na brakujące elementy obiektu. Oczywiście, duże i dobrze widoczne obiekty będą wolne od istotnych błędów. Może jednak warto pomyśleć o lepszej koncepcji centroidu, np. koniec najdłuższego boku, itp. Po drugie, reprezentacje wykorzystujące punkty na obwodzie były używane już w czasie wczesnych badań nad rozpoznawaniem obrazów w Polsce, w czasach Krajowych Konferencji Robotyki w latach 1988-90 [KKR90]. Niestety, trudno znaleźć konkretne publikacje z tamtego czasu, z dwóch powodów – wtedy dopiero kończył się okres żelaznej kurtyny, a w naszych okolicach Internet dopiero powstawał. Nie jest to bardzo istotny zarzut, gdyż tu Doktorant proponuje bardzo gęsty podział kątów, a takie pomysły były rzadkie w czasach, gdy komputery miały mało pamięci.

**Punkty na obwodzie** Reprezentacje wykorzystujące punkty na obwodzie były używane już w czasie wczesnych badań nad rozpoznawaniem obrazów w Polsce, w czasach Krajowych Konferencji Robotyki w latach 1988-90 [KKR90]. Niestety, trudno znaleźć konkretne publikacje z tamtego czasu, z dwóch powodów – wtedy dopiero kończył się okres żelaznej kurtyny, a w naszych okolicach Internet dopiero powstawał (17 lipca 1990).

**Wielościany, ich widoki charakterystyczne i wybrane trzy widoki** Zagadnienie znalezienia wszystkich możliwych widoków wielościanu, w tym wielościanu niewypukłego, prowadzące do pojęć sfery widokowej i widoków charakterystycznych, było szeroko badane. Pozwolę sobie wymienić jako przykłady tylko niektóre prace polskich autorów: [KM03, SM12]. W rozprawie zagadnienie widoków opisane w rozdziale 7 zostało uproszczone i można przypuszczać, że było to uproszczenie zrobione świadomie, ale byłoby lepiej, gdyby tę świadomość otwarcie pokazać.

W związku z tym trzeba stwierdzić, że wybranie tylko trzech widoków jako widoków charakteryzujących w wystarczający sposób bryłę jest oparte na wielu niejawnych założeniach. Podstawowym założeniem jest, że obiekt jest budynkiem, i dlatego prawdopodobnie można dla niego wyróżnić widok frontowy i dwa inne widoki, podstawowy z góry i tylko jeden z możliwych dwóch z boku (zawsze prawego). Dla symetrycznych budynków będących graniastosłupami to często dobre założenie, ale w ogólnym przypadku nie jest ono spełnione. Założenie to może dobrze funkcjonować, jeśli celem jest nawigacja wokół budynku w dużej odległości od niego, i ewentualnie przelecenie przez otwór w nim. Otwór wygląda tak samo z dwóch stron, o ile jest graniastosłupem. Taki model będzie jednak bezużyteczny, jeśli chcemy wlecieć do wnętrza przez okno i wylecieć tą samą drogą. Warto jawnie podawać założenia, aby mieć świadomość jakie możliwości dany model ma, a jakich nie posiada.

**Konstrukcja modelu bryły z rzutów** Moim zdaniem, opis algorytmu tworzenia opisu bryły z opisów wielokątów w trzech płaszczyznach jest błędny. Chodzi o opis podany w rozdziale 7.1 i zilustrowany rysunkiem 7.4. Coś nie zostało dobrze opisane w punktach 2 i 3 algorytmu (chodzi o błąd w opisie, nie w algorytmie). Spróbuję to wyjaśnić.

Jako rzut odniesienia bierzemy rzut z boku, z prawej strony. Bierzymy kolejne segmenty odniesienia (to wektory reprezentujące boki wielokąta reprezentującego ścianę budynku). W danym kroku rozważamy segment zielony z Rys. 7.4, czyli bierzemy (dość nietypowy) segment nie pionowy i nie poziomy. Wyznacza on płaszczyznę na której leży, prostopadłą do płaszczyzny na którą rzutujemy a równoległą do kierunku rzutowania. Na nią rzutujemy wielokąty reprezentujące dwie *pozostałe ściany*. Nie jest tak, że dwie ściany leżą na tej płaszczyźnie, jak napisano w kroku trzecim, lecz leżą na niej ich rzuty. Te rzuty nie są zdegenerowane do odcinków, bo płaszczyzna nie jest prostopadła do żadnej z dwóch ścian. Teraz, znajdujemy część wspólną rzutów reprezentacji ścian, *a nie samych ścian, jak napisano w kroku czwartym*. Zwracamy przy tym uwagę, gdzie znajduje się wewnątrz każdego wielokąta. Dobrze. Teraz, co będzie, jeśli segment odniesienia jest prostopadły do jednej z dwóch pozostałych ścian? Jeśli bryła jest prostopadłościanem (jak część budynków), to wszystkie segmenty odniesienia będą prostopadłe do jednej z pozostałych ścian. Wtedy jeden rzut będzie się redukował do odcinka, i część wspólna rzutów też będzie odcinkiem.

Na rysunkach 7.9 i 7.12 nie widać, aby reprezentacje wszystkich ścian były tylko odcinkami: są tam raczej liczne wielokąty. Zatem, opis nie pokazuje rzeczywiście zaprogramowanego algorytmu tworzenia reprezentacji brył. Zostały pominięte istotne informacje. W publikacji cytowanej w pracy jako [13] jest ten sam opis (w cytowaniu nie podano, że jest to praca konferencyjna z ICAISC 2014).

**Zniekształcenie perspektywiczne** Przedstawione metody zapisu cech obiektów i ich porównywania nie są niezależne od zniekształceń wynikających z perspektywy. W zasadzie porównań dokonuje się na obrazach wykonanych pionowo z góry, ale obiekty sąsiadujące z obiektem aktualnie nas interesującym są widziane nieco z boku, a ilościowe efekty tego faktu nie są brane pod uwagę. Można byłoby jednokrotnie oszacować wynikający stąd błąd i ewentualnie uznać go za nieistotny. Wystarczyłoby, gdyby Autor przynajmniej to skomentował. Dobrą okazję to tego stwarza Rys. 9.2 na s. 70, gdzie odchylenie kierunku padania promieni od pionu jest znaczące, a paralaksa wynikająca z różnicy wysokości na jakiej znajduje się dach budynku i jego podstawa jest wyraźna.

**Zaokrąglanie do kątów nominalnych** Na s. 34 jest napisane, że różnice kątów nominalnych będących cechami boków wielokątów są zaokrąglanie do wartości nominalnych. Wydaje się, że należałoby uniknąć dwukrotnego zaokrąglania, w reprezentacji wielokąta zapamiętywać dokładne wartości kątów z niewielką stratą pamięci, a zaokrąglać dopiero ich różnice. Błędy zaokrągleń nie dodawałyby się.

**Współrzędne punktu zastępującego dwa bliskie punkty** Na s. 32 napisano, że punkt powstający po połączeniu bliskich sobie punktów ma współrzędne uśrednione ze współrzędnych tych punktów. Czy nie byłoby lepiej, gdyby ten zastępujący punkt wyznaczyć jako punkt przecięcia prostych wyznaczonych przez odcinki mające końce w punktach zastępowanych?

## 2.2.2 Uwagi dotyczące sformułowań i nomenklatury

- S. 13<sub>9</sub>:** Nie *przeciążeniu*, lecz *przyspieszeniu*. Przeciążenie  $\equiv$  nadmierne przyspieszenie.
- S. 33<sup>1</sup>, 37<sup>1</sup> – tytuły rozdziałów:** Autor wprawdzie pisze o metodach rozpoznawania obrazów, ale w rzeczywistości chodzi o obiekty reprezentowane jako wektorowo zapisane wielokąty.
- S. 26<sub>nn</sub>:** Używane są takie terminy, jak *obiekt*, *figura*, *bryła*, *obszar*, a także *ściana* i *bok*. Przy tym, raz oznaczają one obiekty na obrazie, a innym razem w świecie rzeczywistym. Dobrze byłoby to ujednoczyć. Ustalmy najpierw, że termin *obiekt* oznacza dowolną rzecz (żeby nie mówić, że obiekt to dowolny obiekt), tak w przestrzeni fizycznej, jak i w przestrzeni obrazu. Na przykład, jednym z rozwiązań mogłoby być ustalenie, że obiekt w przestrzeni fizycznej to *bryła*, która jest *wielościannem* i ma *ściany* będące *wielobokami*, a obiekt w przestrzeni obrazu to *plama*, która może być reprezentowana w zapisie wektorowym jako *wielobok* mający *boki* reprezentowane jako wektory w tym zapisie. Te dwa wieloboki, z których jeden jest *figurą geometryczną* w przestrzeni fizycznej, a drugi w przestrzeni obrazu, są porównywane. Nie należy mylić ścian wieloboku z bokami wielokąta, jak to ma miejsce na s. 33<sub>6</sub>, gdzie napisano o kątach nachylenia ścian oraz o ich długości. Często z kontekstu jasno wynika, co mamy na myśli pisząc o fizycznym obiekcie, będącym bryłą, który na obrazie jest widoczny jako plama, a jego reprezentacją jest wielokąt zapisany wektorowo, ale nie zawsze tak jest.
- S. 29<sup>4,5</sup>:** [*kąt*] *wyznaczony przez wektor  $(B,C)$  oraz punkt  $D$* . Kąta nie wyznacza wektor i punkt, lecz dwa wektory:  $\angle BCD \equiv \angle(\vec{BC}, \vec{CD})$ .
- S. 28<sub>14</sub>:** Nie *minimalną*, lecz *maksymalną* odległość punktów od prostej. Minimalna wynosi zero.
- S. 46<sup>5</sup>:** Każda płaszczyzna jest *dwuwymiarowa*, co najwyżej może być umieszczona w przestrzeni *trójwymiarowej* i ewentualnie może znajdować się w takim położeniu, że nie jest równoległa do żadnej płaszczyzny układu odniesienia.

S. 58<sub>2</sub>, opis Rys. 8.5: Napisano: *Graf GBS<sup>A</sup> (Rys. 8.4) zawiera tylko dwa spośród trzech obiektów należących do bliskiego sąsiedztwa Obj<sub>found</sub><sup>B</sup>*. Jednak, na rysunku 8.4 pokazano wszystkie trzy obiekty.

Przy okazji warto zwrócić uwagę, że takie słowa jak *found* czy *factor* we wzorach spełniają wprawdzie rolę indeksów, ale są bliższe nazwom funkcji (jak *sin*, *min*) niż obiektów, a na pewno nie są iloczynami zmiennych *f*, *o*, *u*, *n*, *d*. Dlatego należałoby je pisać raczej tak:  $\mathit{Obj}^B_{\mathit{found}}$ .

S. 59, 60, Rysunki 8.6, 8.7: Na rysunkach nie są oznaczone obiekty o których mowa, albo w treści podpisów brak informacji, że chodzi o obiekty wyróżnione grubszą kreską. To jest w zasadzie oczywiste, ale trzeba pisać ściśle. Na rysunku 8.7 poza trzema obiektami należącymi do bliskiego sąsiedztwa obiektu *Obj<sup>A</sup>* są jeszcze trzy inne obiekty, niepotrzebne na tym rysunku.

S. 61, Rys. 8.8: Zgodność relacji przestrzennych jest duża lub bardzo duża, ale nie *pełna*. Widać wyraźnie, że obiekty sąsiednie są nałożone z pewnym błędem. Aby wiedzieć, czy jest on duży czy mały, trzeba użyć jakiejś miary i ustalić zakresy. Nie zawsze trzeba i można pisać o wszystkim, ale na pewno nie jest to zgodność pełna.

S. 70, Rys. 9.2: Widać tu pomieszanie przestrzeni fizycznej i przestrzeni obrazu, związane z niejednoznacznym przypisaniem obiektów i pojęć do tych przestrzeni, o czym napisano w uwadze dotyczącej stron 26nn. Jedne wielkości są mierzone w pikslach, inne w metrach. Jedna z odległości jest opisana w taki sposób, że wydaje się być jednocześnie wielkością daną i szukaną: 345 px, a bezpośrednio pod spodem *X* – *szukana odległość wyrażona w metrach*. Cały ten opis jest niejasny.

### 2.2.3 Drobne uwagi redakcyjne i techniczne

Uwagi te dotyczą merytorycznie nieistotnych spraw redakcyjnych, które nie wpływają na ocenę pracy, ale zamieszczam je do wiadomości Autora. Omawiam je w kolejności stron. Spis obejmuje prawie wszystkie drobne błędy i potknięcia językowe znajdujące się w pracy.

S. 5, **Abstract**: Unikajmy antropomorfizacji, stosując stronę bierną, np. *This dissertation presents the results. . .* → *In this dissertation the results. . . are presented.*

S. 9<sup>8,19</sup> i dalsze, s. 12<sup>1</sup>: Sieroty – słowa jednoliterowe pozostawione na końcu linii: „i”, „w” i podobne. Błąd typograficzny, podobnie jak wdowa, bękart i szewc. Szczególnie zwraca to uwagę w tytule rozdziału 2. Zachęcam do stosowania spacji nierozdzielającej „~” – tylda. Niestety, zbyt często zawieramy programom dokonującym składu tekstu; nawet znakomity L<sup>A</sup>T<sub>E</sub>X nie robi wszystkiego.

S. 10<sup>8</sup>: *systemów opartych o pojedynczy sensor* → *systemów opartych na pojedynczym sensorze*. Częsty błąd stylistyczny. Różnica jest mniej więcej taka, jak między zdaniami *Oprzyj się o mnie* oraz *Oprzyj się na mnie*.

S. 11, Rys. 1.1: Jaką rolę spełniają na tym rysunku kolory, bardzo wyraziste? W podpisie nie ma o tym ani słowa.

S. 14<sup>2,4</sup> i dalsze: Zamiast *ilością*, *ilości* powinno być: *liczbą*, *liczby*. W tekstach ścisłych z rzeczownikami policzalnymi łączymy *liczbę*.

S. 25<sub>2-1</sub>: Nie *bierzący* lecz *bieżący*, bo *bieżeć* i *biegać*.



- S. 26, Rys. 4.2 i 4.3: Te dwa rysunki zawierają ten sam obraz. Na drugim z nich powinno być pokazane coś innego.
- S. 29, Rys. 4.5: W opisie na rysunku słowo *kolejnych* jest powtórzone.
- S. 36, Rys. 5.3: Reprezentacja oparta o cechy na cechach. Niepotrzebne puste pole nad rysunkiem – chociaż, może to L<sup>A</sup>T<sub>E</sub>X umieścił rysunek na środku osobnej strony? Jeśli tak, to spróbowałbym zmniejszyć rysunki 5.1 i 5.2, może wtedy układ tekstu poprawiłby się.
- S. 37<sup>4</sup>: Nie *tą samą ilość boków* lecz *tę samą liczbę boków*.
- S. 37<sub>6</sub>: Anakolut: *Przyjmując wartość 5° algorytm wykona 72 pomiary*. To my przyjmujemy, a algorytm wykonuje, są to różne podmioty. Ten błąd staje się powoli smutną normą. Powinno być na przykład: *Jeśli przyjmiemy wartość 5°, algorytm wykona 72 pomiary*.
- S. 44<sub>nn</sub>: Do podkreślania fragmentów tekstu raz używana jest *kursywa*, a innym razem czcionka bezszeryfowa, do tego sprawiająca wrażenie nieco większej, niż otaczający tekst. Czym różnią się teksty wyróżnione w te dwa sposoby?
- S. 45, Rysunki 7.3 i 7.4: Na rysunku 7.3 widać silne artefakty pochodzące z procesu kompresji/dekompresji; należałoby skorzystać z wektorowej wersji rysunku. Rysunki 7.4 zawierają wiele ciekawych szczegółów, a są za małe; nie byłoby źle, gdyby każdy z nich zajmował nawet całą szerokość strony.
- S. 70<sub>1</sub>: Pojedyncze słowo zostawione w ostatniej linii akapitu to *wdowa*, zaś pojedynczy ostatni wiersz akapitu, który znajduje się na górze nowej strony, to *bękart*. Tutaj słowo *systemu* na s. 70 i słowo *sceny* na s. 89 jest jednocześnie wdową i bękartem (polecam praktyczną książkę [Cen06]).
- S. 105<sup>3-5</sup>: Tekst *Przed wszystkim ze względu na ograniczone możliwości zastosowanego preprocessingu, który... jest równoważnikiem zdania – brak orzeczenia*.
- S. 107<sub>nn</sub>, Bibliografia: Dla żadnej publikacji nie podano numeru DOI (styl BibT<sub>E</sub>X-a plain nie obsługuje DOI, ale liczne inne style robią to dobrze; można posłużyć się polem `note` i instrukcją `\href` aby uwzględnić DOI w bibliografii). Często brak informacji o tym, że dana publikacja jest pracą konferencyjną, również wtedy, gdy współautorem jest sam Doktorant.

### 3 Podsumowanie

Podsumowując, w rozprawie można wyróżnić istotne elementy oryginalne omówione w rozdziale 2.1, zaś uwagi dyskusyjne i krytyczne omówione w rozdziale 2.2 nie umniejszają w istotny sposób wartości pracy.

Cele pracy zostały osiągnięte, a teza została wykazana.

Praca jest wynikiem badań opartych na dobrej analizie stanu wiedzy. Zawiera nowe i oryginalne elementy poznawcze, implementację wynikających z nich metod w postaci robota realizującego wyznaczone zadania.

**Wniosek** Powyższy opis, uwzględniający uwagi pozytywne jak również uwagi dyskusyjne i krytyczne, uzasadnia mój ostateczny wniosek, że recenzowana rozprawa **stanowi oryginalne rozwiązanie problemu naukowego** w dyscyplinie informatyka dziedziny nauk technicznych. Zagadnienie zostało prawidłowo postawione i skutecznie rozwiązane, a rozwiązanie zostało rzetelnie zweryfikowane. Rozprawa **wykazuje ogólną wiedzę teoretyczną kandydata** w przedmiotowej dyscyplinie naukowej **oraz umiejętność samodzielnego prowadzenia pracy naukowej**.

Rozprawa spełnia wymagania obecnie obowiązującego prawa dotyczącego wymagań dla rozpraw doktorskich. Zatem, stawiam wniosek o przyjęcie rozprawy i skierowanie jej do dalszych etapów przewodu doktorskiego.



## Cytowana literatura

- [BJ00] H. Bunke, X. Jiang. Graph matching and similarity. H.-N. Teodorescu, D. Mlynek, A. Kandel, H.-J. Zimmermann, editors, *Intelligent Systems and Interfaces*, pages 281–304. Springer US, Boston, MA, 2000. doi:10.1007/978-1-4615-4401-2\_10.
- [Cen06] D. Cendrowska. *Zrób to lepiej! O sztuce komputerowego składania tekstu*. Wydawnictwo Naukowe PWN, Warszawa, 2006.
- [KKR90] *Prace III Krajowej Konferencji Robotyki*, numer 38 serii Prace Naukowe Instytutu Cybernetyki Technicznej, seria: Konferencje, Wrocław, 19-21 września, 1990. Politechnika Wrocławska.
- [KM03] M. Kowalczyk, W. S. Mokrzycki. Obtaining complete 212D view representation of polyhedra using concept of seedling single-view area. *Computer Vision and Image Understanding*, 91(3):280–301, 2003. doi:10.1016/S1077-3142(03)00098-5.
- [SM12] A. Salamonczyk, W. S. Mokrzycki. Generating view models of monotone polyhedrons for visual identification. Presentation of results of the PhD thesis. *Machine Graphics & Vision*, 21(1/4):3–17, 2012. <http://mgv.wzim.sggw.pl/MGV21.html#1-3>.