

Prof. zw. dr hab. n.t. Wojciech S. Mokrzycki
Wydział Matematyczno-Przyrodniczy. Szkoła Nauk Ścisłych
Instytut Informatyki, UKSW
ul. Wóycickiego 1/3, 01-938 Warszawa

Warszawa. 2017-XII-31.

RECENZJA

rozprawy doktorskiej mgr inż Bogdana Hrasymowicza-Boggio
pt. **Rozpoznawanie obiektów w chmurach punktów na potrzeby robotyki mobilnej**
wykonana na zlecenie Rady Wydziału Mechatroniki PW

1. Zakres, cel i teza Rozprawy

Rozprawa dotyczy rozumienia otoczenia przez roboty mobilne w środowisku ich pracy, szczególnie zdolności interpretacji sygnałów wizyjnych. Zdolność taka umożliwia interakcję z elementami otoczenia w wielu możliwych zastosowaniach robotów usługowych. Ograniczone rozumienie danych wizyjnych przez maszyny wynika głównie z dużej trudności samego zadania rozumienia obrazu.

W szczególnym zainteresowaniu Doktoranta są techniki przetwarzania wizyjnych danych trójwymiarowych (zwłaszcza widoków 3W).

Celem Rozprawy jest zbadanie kluczowych aspektów rozpoznawania obiektów w chmurach punktów na potrzeby robotyki mobilnej, w tym przedstawienie autorskich algorytmów rozwiązujących określone problemy.

Wiodące na tym polu algorytmy (**ogólnej klasyfikacji semantycznej**) są bardzo złożone, nie w pełni skuteczne i nie radzą sobie z rozpoznawaniem nietypowego obiektu (niepodobnego do przykładów uczących). Z tych powodów Doktorant przygotował własne bazy uczące i testowe, które umożliwiły Mu przeprowadzenie stosownych testów w planowany sposób. Jednak opisu tych baz w Rozprawie nie opisał.

Rozprawa zawiera przegląd wielu istniejących metod rozpoznawania obiektów 3W oraz szczegółowe omówienie różnych algorytmów i eksperymentów. Zawarto w niej też teoretyczne i eksperymentalne uzasadnienie wad i zalet poszczególnych podejść, a także wnioski na temat możliwego ulepszenia algorytmów. Autorskie metody przedstawione w Rozprawie charakteryzują się głównym wkładem metodologicznym, implementacyjnym i testowym.

Wiedza literaturowa n.t. praktycznego wdrożenia robotów mobilnych oraz własne wyobrażenie o pożądanej i realistycznej funkcjonalności użytecznego robota, umożliwi Doktorantowi przyjęcie następujących założeń:

1. Robot powinien przetwarzać scenę szybko, najwyżej w ciągu kilku sekund.
2. Przygotowanie robota do rozpoznawania nowego obiektu, którego nie zna (w tym nietypowego), powinno być łatwe, bez konieczności dostarczenia setek czy tysięcy jego widoków.
3. Robot może działać w środowisku mało uporządkowanym, gdzie obiekty się stykają i przesłaniają.
4. W większości wypadków robot ma ograniczony zbiór obiektów zainteresowania (związanych z wykonywaną przez niego pracą). Nie musi rozpoznawać wszystkich, ani nawet „jak najwięcej” obiektów w środowisku, tylko skutecznie rozpoznawać te, które są istotne.

5. Robot działa w realnym, konkretnym środowisku człowieka (np. w domu, szpitalu, magazynie, itp). To naturalne środowisko charakteryzuje się pewnymi prawidłowościami, które robot może wykorzystać w jego rozumieniu (np. typowe położenia pewnych obiektów w przestrzeni lub w stosunku do innych obiektów).
6. Robot może poruszać się w obserwowanym środowisku lub poruszać kamerę, uzyskując wiele widoków sceny.
7. Wysokość nad podłożem i nachylenie sensora wizyjnego robota są znane, gdyż robot ma znaną geometrię i jest najczęściej wyposażony w akcelerometry (część sensorów wizyjnych także ma wbudowany akcelerometr),

i postawienie następującej tezy:

Teza Rozprawy: Skuteczne rozpoznawanie obiektów na potrzeby robotyki mobilnej różni się od ogólnego zadania rozpoznawania obiektów i wymaga szybko uczących się i szybko działających algorytmów, a także czerpie korzyści z wykorzystania wiedzy o świecie wykraczającej poza znajomość struktury obiektów: w tym kontekście semantycznego i ograniczeń percepcyjnych.

2. Struktura Rozprawy

Rozprawa została zredagowana na 111 stronach w 7-miu rozdziałach, z Bibliografią liczącą 115 pozycji w zdecydowanej większości wywołanych w treści Rozprawy. Jest reprezentatywna dla zagadnień rozwiązywanych w Rozprawie, choć nie jest dla zagadnień ogólnego rozpoznawania i rozumienia otoczenia. W przeglądzie literatury nie ma wzmianki o ośrodkach i badaczach krajowych z osiągnięciami w przedmiocie Rozprawy. Doktorant jest współautorem 9 publikacji (w tym 8 anglojęzycznych) i nie ma publikacji samodzielnej. Całość jest poprzedzona Streszczeniem (po polsku i angielsku), Spisem treści, Listą symboli oraz zakończona Bibliografią; spisu ilustracji (28 pozycji) ani tabel (dwie pozycje) nie zamieszczono. Bibliografia jest podana alfabetycznie, co nie ułatwia jej śledzenia i oceny jej potencjalnej ważności dla zawartości Rozprawy.

2.1.

Rozprawa ma następującą strukturę:

1. Rozdział 1 omawia założenia, cel, zakres i tezę Rozprawy.
2. Rozdział 2 jest przeglądem literatury na którą Doktorant powołuje się w zakresie klasyfikacji różnych rodzajów metod rozpoznawania obiektów. Odwołania do literatury są też w odpowiednich rozdziałach.
3. Rozdziały 3÷5 zawierają opisy eksperymentów uzasadniających postawioną tezę o użyteczności poszczególnych metodologii. W szczególności:
 - Rozdział 3 przedstawia autorskie prace dotyczące kontekstu w rozpoznawaniu obiektów.
 - Rozdział 4 prezentuje autorską metodę wnioskowania w rozpoznawaniu, uwzględniającą niepewność i niewiedzę.
 - Rozdział 5 omawia prace Doktoranta dotyczące algorytmów masowo-równoległych i metod heurystycznego przyspieszania procesu rozpoznawania.
4. Rozdział 6 prezentuje autorską implementację masowo-równoległego systemu rozpoznawania obiektów 3W w systemie mobilnego robota autonomicznego.
5. Rozdział 7 zawiera wnioski i podsumowanie.

Zaprezentowane w Rozprawie prace koncentrują się na różnych, komplementarnych aspektach rozpoznawania i klasyfikowania oraz opisie obiektów 3W na potrzeby robotyki mobilnej i były prowadzone przez Doktoranta na przestrzeni kilku ostatnich lat. Uważa On, że dalsze prace są w najbliższych latach wciąż konieczne, nim omawiana dziedzina osiągnie poziom umożliwiający wdrażanie rozbudowanych systemów praktycznych.

2.2.

Przedstawiona w rozdziale 3 metoda rozpoznawania 3W obiektów bazuje na połączeniu klasyfikacji opartej o cechy odizolowanych segmentów oraz informacji kontekstowej w postaci relacji metrycznych oraz relacji semantycznych między różnymi obiektami. Wioskowanie kontekstowe Doktorant przeprowadził stosując metodę wykorzystującą autorskie uogólnienia losowych pól Markowa.

Proces klasyfikacji składa się z 5 etapów:

1. wydobycia cech w oparciu o sąsiedztwa punktów na obiekcie;
2. segmentacji metodami 2W;
3. formułowania hipotez - obliczania dla każdego segmentu histogramu cech i korelacyjne dopasowywanie do histogramów wzorcowych; gdy wynik korelacji przekracza ustalony próg oraz segment spełnia warunki przybliżonego podobieństwa, jest formułowana „hipoteza segmentu” stanowiąca o przynależności tego segmentu do określonego obiektu;
4. wnioskowania niskopoziomowego - tworzenia metahipotez (wynikowych hipotez): jest prowadzone dopasowanie metryczne chmury punktów dla hipotez obiektów złożonych w celu ustalenia powiązania segmentów sceny, należących do tych samych obiektów oraz ich ułożenia w przestrzeni 3W.
5. wnioskowania wysokopoziomowego, realizowanego w trzech krokach:
 - tworzenie połączeń między hipotezami segmentów zgodnie ze znanymi relacjami semantycznymi,
 - grupowanie segmentów polegające na znajdowaniu podziałów sceny na rozłączne zbiory możliwie powiązanych z sobą segmentów,
 - optymalizacja (właściwe wnioskowanie, przypisanie ostatecznych etykiet do segmentów), jest wykonywana na każdym ze zbiorów wynikowych za pomocą uogólnionych losowych pól Markowa oraz algorytmu genetycznego i ma na celu znalezienie kombinacji hipotez segmentów, która maksymalizuje funkcję energii przy danej strukturze topologicznej powiązanych segmentów („najlepsza teoria sceny”).

Wykonane liczne doświadczenia pokazały użyteczność zaproponowanej metody wnioskowania.

2.3.

Rozdział 4 dotyczy uwzględniania informacji o niewiedzy w systemie rozpoznawania (opartym o wnioskowanie zaproponowane przez Doktoranta), do czego wykorzystał On teorię Dempstera-Shafera (DST) oraz uogólnione losowe pola Markowa. Pierwsze cztery etapy przetwarzania (wydobycie cech, segmentacja, formułowanie hipotez i wnioskowanie niskopoziomowe) zostały rozszerzone o algorytmy mające na celu zdobycie informacji o niepewności i przesłonięciach. Piąty etap (wnioskowanie wysokopoziomowe) został zaprojektowany od nowa w taki sposób, by za pomocą DST uwzględniać te informacje.

Teoria Dempstera-Shafera jest formalizmem wnioskowania w warunkach niepewności i została opracowana w celu umożliwienia matematycznego opisu przekonania, niepewności i konfliktu. Wiedza jest kodowana przez przypisanie mas m do podzbiorów zbioru potęgowego T wszystkich możliwych

zdarzeń. Zatem, DST dostarcza prostego sposobu wyrażania niewiedzy przez masę przypisaną bezpośrednio do przestrzeni zdarzeń T , gdyż masa ta wskazuje na „jakiś” (nieznane) zdarzenie.

Użycie DST wymagało rozwiązania problemów: agregacji przekonań pochodzących z segmentów z różnych obiektów, optymalizacji funkcji przekonania dla całej sceny (wiele znanych i nieznanymi obiektów) oraz agregacji dużej liczby źródeł wiedzy (segmentów) bez doprowadzenia do saturacji przekonania.

Wnioskowanie wysokopoziomowe prowadzone w tym rozdziale korzysta z 3-stopniowego algorytmu agregacji Dempstera-Shafera: 1) obliczanie parametrów DST, 2) etap metryczny, 3) etap semantyczny i teoria sceny. Parametry DST są obliczone dla części poszczególnych obiektów teorii sceny dla obecnych, jak i nieobecnych części. Parametry te są agregowane w metrycznym etapie metody, dla każdego obiektu stosując opisaną (ważoną) modyfikację agregacji Dempsterowskiej. Wynikowe parametry dla każdej hipotezy obiektu są przekazywane do etapu semantycznego i agregowane z parametrami „a priori”, wynikającymi z wykrytych relacji semantycznych z innymi obiektami w rozpatrywanej teorii. Parametry wyjściowe obliczone dla poszczególnych obiektów także podlegają agregacji, dając pojedynczy zestaw parametrów DST dla całej teorii sceny.

Znalezienie najlepszej teorii sceny stanowi nietrywialny problem optymalizacyjny, który rozwiązano za pomocą globalno-lokalnego algorytmu ewolucyjnego. Algorytm ten dąży do ustalenia teorii o najwyższym potwierdzeniu, obliczanym jako funkcja celu. Algorytm inicjuje populację teorii klasyfikujących wszystkie segmenty sceny (lub grupę segmentów bez powiązań zewnętrznych), która podlega ewolucji.

Proces agregacji danych według DST polega na wykonaniu następujących kroków:

- Określenie poziomów przekonania poszczególnych hipotez na podstawie obserwacji stanowiących źródła wiedzy.
- Stosowanie reguły agregacji w celu łączenia przekonania i zaprzeczenia. Każde źródło wiedzy może popierać daną hipotezę w zakresie wartości $0 \div 1$, albo jej zaprzeczać, też w zakresie $0 \div 1$. Suma przekonania i zaprzeczenia może być nie większa od 1.

Istnieją różne możliwe metody agregacji parametrów różnych źródeł wiedzy. Najbardziej znaną, którą Doktorant zastosował, jest reguła agregacji Dempstera.

Przedstawiona metoda rozpoznawania obiektów została zaimplementowana i przetestowana na chmurach punktów na trzech zestawach scen:

- Zbiór treningowy z 18 scen z obiektów pełni widoczne obiekty obserwowane z niewielkiej odległości, a także ich zestawienia podobne do warunków docelowego działania systemu.
- Zbiór testowy 14 „łatwych”, realistycznych scen ze znanymi i nieznanymi obiektami w różnych ułożeniach i odległościach, z przysłonięciami, ale ze stosunkowo uporządkowanym otoczeniem.
- Zbiór testowy 12 „trudnych” scen z licznymi, głównie nieznanymi obiektami i wieloma przysłonięciami.

Wyniki eksperymentów okazały się interesujące. Wykorzystanie informacji o niepewności obniżyło (nieznacznie) liczbę prawidłowych rozpoznań, lecz przede wszystkim znacząco obniżyło liczbę rozpoznań nieprawidłowych. Zatem, niepewność dla części znalezionych w większym stopniu wpływa na system niż niepewność dla części nie znalezionych. Jednak średnio najlepsze wyniki daje połączenie obu rodzajów niepewności.

Doktorant sugeruje rozwinięcie opisanego systemu przez dokładniejsze i pełniejsze zamodelowanie źródeł niepewności, np. uwzględniają właściwości fizyczne sensora, a także badając wpływ innych znanych czynników na dokładność algorytmów segmentacji i wyznaczania cech.

Innym elementem mogącym podlegać modelowaniu niepewności jest algorytm agregacji danych wejściowych z kamery.

2.4.

W rozdziale 5 Doktorant przedstawia prace dotyczące metod masowo-równoległych. Problemem była segmentacja obszarów płaskich lub gładkich sceny, która najczęściej opiera się na wykorzystaniu algorytmu RANSAC (*random sample consensus*) mającego istotne ograniczenia, jak: czasochłonność, mało dokładne dopasowanie do sceny, losowość czasu dopasowania, nieskuteczność dla powierzchni z nawet niewielkim zakrzywieniem. W efekcie czasu dopasowywania wielu płaszczyzn do pojedynczej sceny wynoszą zwykle kilka sekund. Możliwe wykorzystanie obliczeń masowo-równoległych do przeprowadzenia segmentacji metodą rozrostu ziarna jak metoda RANSAC jest obciążone dużą losowością i zależnością od sceny. Ale przede wszystkim algorytm rozrostu ziarna jest z natury szeregowy. Doktorant zaproponował masowo-równoległą, iteracyjną metodę rozwiązania tych trudności, w której poszczególne obszary konkurencyjnie przyłączają nie przydzielone punkty i inne obszary na podstawie funkcji ciągłości i wartości swoich indeksów. Zaproponowana metoda nadaje się do masowo-równoległego przeprowadzenia rozrostu ziarna jednocześnie dla dowolnej liczby ziaren.

Jak podaje Doktorant, na testach przeprowadzonych z użyciem karty graficznej GTX Titan Black czas przetwarzania scen składających się z ponad 100000 punktów wynosił zwykle poniżej 50 ms.

W tym rozdziale Doktorant opisuje też metody rozpoznawania oparte na deskryptorach lokalnych powierzchni. Jest do tego niezbędne dokonanie ich klasyfikacji. Utrudnieniem jest wysoki koszt obliczeniowy niektórych złożonych deskryptorów. Doktorant zaproponował metodę wybierania obszarów do obliczenia deskryptorów. Podany algorytm operuje na pełnych chmurach punktów sceny redukowanych w każdym kroku procedury i polega na:

- usunięciu dużych, gładkich obszarów, w których znalezienia istotnych elementów nie jest spodziewane,
- uproszczeniu chmury z wykorzystaniem filtru wokalowego w celu otrzymania zbioru punktów kotwiczących,
- ograniczanie liczby punktów kotwiczących na podstawie informacji o konkretnym, klasyfikowanym obiekcie,
- obliczanie i klasyfikowanie deskryptorów złożonych dla punktów kotwiczących oraz przeprowadzenie dodatkowego kroku dylacji (propagacji pozytywnej klasyfikacji na sąsiednie punkty kotwiczące).

System był badany w środowisku realnym. Porównano jakość klasyfikacji deskryptorów dla pełnej sceny, dla sceny bez obszarów gładkich oraz dla sceny z dodatkowym wykorzystaniem modeli negatywnych. W szczególności, klasyfikacji deskryptorem PFHRGB dla scen pełnych, z usuwaniem obszarów gładkich (nonflat) oraz dodatkowo uczonych modelami negatywnymi lub ujemnymi (*antimodel*). Obydwa wymienione mechanizmy (usuwanie obszarów gładkich i modele negatywne) wpłynęły pozytywnie na jakość klasyfikacji (szczególnie usunięcie obszarów gładkich). Doświadczenia pokazują wysoką użyteczność łączenia różnych prostych deskryptorów lokalnych w polepszaniu dokładności rozpoznawania.

Doktorant sądzi, że by dokonać dalszej optymalizacji czasowej i jakościowej, należy testować większą liczbę deskryptorów lokalnych i użyć dużej bazy testowych scen realnych oraz, że wartościowe byłoby także zbadanie możliwości połączenia opisanych metod z podejściem wykorzystującym detektory punktów charakterystycznych. Doktorant uważa też, że naturalnym zastosowaniem omówionych algorytmów byłoby ich połączenie z dokładną techniką weryfikacji hipotez obiektów, wykorzystującą dopasowanie metryczne (np. algorytm typu RANSAC). Wymagający, lecz dokładny algorytm dopasowania mógłby używać wyników wypracowanych przez metodę deskryptorów jako heurystykę.

2.5.

W rozdziale 6 Doktorant prezentuje zaimplementowany na Wydziale Mechatroniki PW prosty system dla robotyki usługowej, integrujący nawigację metryczną, tworzenie mapy semantyczno-metrycznej oraz rozpoznawanie obiektów 3W, wykorzystujący skaner laserowy (do nawigacji metrycznej) oraz kamerę głębi Kinect, za pomocą której jest możliwe rozpoznawanie obiektów 3W. Udział Doktoranta w tych pracach to system wizyjny, ograniczony do podstawowych, szybkich metod umożliwiający płynne działanie on-line, w warunkach stosunkowo uporządkowanego środowiska laboratoryjnego. W systemie zastosowano algorytm Kinect Fusion umożliwiający znaczne podniesienie dokładności i uzupełnienie braków w chmurze punktów. Algorytm ten łączy informację pochodzącą z wielu klatek kamery głębi, korzystając z obliczeń masowo-równoległych.

Przeprowadzone testy pokazały, że (dla scen statycznych) nawet jednosekundowa obserwacja i skanowanie klatek sceny wprowadza wyraźny wzrost dokładności chmury punktów. W jedynym cyklu robot przetwarza klatki za pomocą Kinect Fusion, po czym następuje rozpoznawanie obiektów i aktualizacja mapy semantycznej. Ostatecznie jednak przeprowadzone testy pokazały, że używanie Kinect Fusion jest wręcz niezbędne do uzyskania satysfakcjonującej jakości rozpoznań, zwłaszcza niewielkich lub czarnych obiektów.

Doktorant sądzi, że eksperymentalny system działał zgodnie z oczekiwaniami – mapa semantyczna była w większości wypadków prawidłowo aktualizowana z wykorzystaniem rozpoznań obiektów. Niedoskonałością okazał się mocno ograniczony (niewiele ponad 3 m) zasięg Kinecta, a dokładniej zakres odległości, dla którego przechwycona chmura punktów jest wystarczająco dokładna do rozpoznawania obiektów. By zwiększyć wiarygodność systemu wprowadzono prostą filtrację dolnoprzepustową przy usuwaniu obiektów z mapy. Konkluzja eksperymentu była taka, że zaproponowany system jest realizowalny w praktyce i może być użyteczny w robotyce usługowej, lecz wskazane byłoby zastosowanie kamery głębi o większym zasięgu (np. Kinect 2), a także ciągłej techniki filtracji chmury punktów.

3. Wkład Doktoranta

Wkład Doktoranta jest następujący:

1. Metoda rozpoznawania obiektów z wykorzystaniem kontekstu, mianowicie, Autor w czasie klasyfikacji i kojarzenia poszczególnych fragmentów sceny stosuje swoją metodę uwzględniania kontekstu semantycznego. To, po wykorzystaniu uogólnienia losowych pól Markowa i algorytmu ewolucyjnego, umożliwia znalezienie „teorii” najlepiej wyjaśniającej scenę (na podstawie opisów semantycznych widocznych cech obiektów oraz ich konfiguracji przestrzennej).
2. Jawne wykorzystanie w metodyce wnioskowania opartej o teorię Dempstera-Shafera informacji o ograniczeniach percepcyjnych systemu rozpoznającego obiekty. Metodyka ta umożliwia modelowanie niepewności i niewiedzy systemu wynikających z różnych utrudnień percepcyjnych (typu: odległość od obiektów, przesłonięcia obiektów).
3. Masowo-równoległy algorytm rozrostu ziarna umożliwiający szybkie znajdowanie dużych jednolitych obszarów w chmurze punktów. Umożliwia on m.in. (praktycznie na bieżąco) rozpoznawanie obiektów w chmurach zintegrowanych, będących w pełni 3W zapisem sceny i jest znacznie bardziej elastyczny w przechwytywaniu sceny i łączeniu (nakładaniu) danych.
4. Metoda heurystycznego ograniczania uwagi dla rozpoznawania obiektów z użyciem deskryptorów lokalnych (ułatwiających segmentację obrazu), użytecznego dla rozpoznawania w środowisku nieuporządkowanym. Dotyczy to metod przyspieszających analizę sceny przez kilkustopniową redukcję, mogących zamieniać lub uzupełniać powszechnie stosowane podejście ogólnych detektorów punktów charakterystycznych. Użyteczność zaproponowanych technik została potwierdzona doświadczalnie.

5. Eksperymentalnego systemu robotycznego (w sprzeczcie) z rozpoznawaniem obiektów 3W w chmurach zintegrowanych w agregacji danych z wielu klatek. Implementacja systemu była związana z różnymi wyzwaniami i unaoczniała użyteczność rozpoznawania obiektów 3W w połączeniu z mapami semantyczno-metrycznymi.
6. Opracowywanie własnych zbiorów scen do różnych eksperymentów, z powodu braku ogólnodostępnych baz testowych zawierających pożądane informacje w kontekście rozważanego obszaru robotyki mobilnej.

Jest to bardzo bogaty materiał analityczny i syntetyczny, podany przekonywująco i (w miarę) przytępnie. Zatem, Doktorant wykazał postawioną w Rozprawie tezę. Pewne braki dotyczą ubogich danych n.t. stworzonych baz testowych oraz histogramów wzorcowych wykorzystywanych w klasyfikacji.

4. Niedociągnięcia

1. W świetle postawionych założeń wydaje się, że tytuł Rozprawy został sformułowany za wąsko i powinien brzmieć: **Rozpoznawanie i klasyfikowanie obiektów w chmurach punktów na potrzeby robotyki mobilnej**, gdyż w założeniach występuje nie tylko klasyczne rozpoznawanie (znanych) obiektów, ale też klasyfikowanie (opis) obiektów nieznanymi. Doktorant założył więc istnienie w systemie robotycznym mechanizmu umożliwiającego automatyczne uczenie się (co najmniej) typowych relacji w trakcie etapu rozpoznającego systemu.
2. Oczekiwanie Doktoranta, że opracuje metodę rozpoznawania radzając sobie, jeśli potrzebne będzie rozpoznawanie konkretnego, nietypowego obiektu (czyli takiego, który nie jest podobny do licznych przykładów uczących) jest bezpodstawne. Termin „rozpoznanie” oznacza bowiem operację znalezienia obiektu wcześniej znanego, opisanego. Jeśli bowiem obiekt nie jest podobny do licznych przykładów uczących, jest niezbędne włączenie algorytmu do procesów widzenia poznawczego (klasyfikacji czy budowania opisów obrazów i obiektów). Problem jednak w tym, że Doktorant nie wypowiada się na temat modeli obiektów (scen) ani ich budowy (struktury wewnętrznej). Nie wiadomo też nic, jak modele (sceny treningowe) były uzyskiwane. Utrudnia to istotnie możliwość odniesienia się do autorskich projektów na każdym z etapów rozpoznawania (kategoryzacji).
3. Niektóre podstawowe pojęcia używane w Rozprawie są zdefiniowane nieprecyzyjnie, zbyt ogólnikowo, np. pojęcie chmury punktów i jej struktury, mającej dla sposobu działania systemu rozpoznającego znaczenie podstawowe. W podanej definicji (str. 14) Doktorant nie podaje najistotniejszego na jej temat, że te punkty chmury są punktami **na powierzchni obmierzanego obiektu, rozmieszczone na tej powierzchni w regularnej siatce w takiej samej strukturze, jak obraz RGB tego obmierzenia**, zaś komórki tej siatki zawierają odległości tych punktów od punktu odniesienia układu pomiarowego. Obraz RGB to współrzędne trójchromatyczne odpowiednich punktów (piksli).
4. Sądzę też, że używane przez Doktoranta terminy „chmura zorganizowana” i „chmura niezorganizowana” (np. str. 15) można zastąpić trafniejszymi terminami „chmura pojedyncza” (albo „pojedynczą mapą głębi”) oraz „chmura zintegrowana” (albo „zintegrowana mapa głębi”). Pierwsza jest bowiem wynikiem akwizycji z jednego punktu obserwacji, druga jest wynikiem wielokrotnej akwizycji z wielu punktów obserwacji jednego, konkretnego obiektu czy sceny. W tych terminach „chmura zintegrowana” zawiera (a nie pozwala na) pełniejsze dane wizualne o obiekcie (scenie) niż „chmura pojedyncza”.
5. Doktorant sugeruje (str. 17), że system rozpoznający musi być uczony rozpoznawania, a więc jest konieczne istnienie fizycznych obiektów, których ten system musi się „nauczyć” i nie wspomina o innych możliwościach. Otóż, istnieją też bazy modeli obiektów (wykle w formie widoków -

aspektów) tworzone automatycznie na bazie formalnego opisu określonych obiektów; jest to baza o charakterze trwałym. Informacje o obiektach są zapisywane w nich w sposób trwały w odpowiednich strukturach, które mogą być wykorzystywane wielokrotnie i do różnych celów. Istnieje liczna literatura n.t. tworzenia baz tego typu, nie przytoczona przez Doktoranta.

6. Nie zgadzam się też z Doktorantem, że widzenie maszynowe 2W należy uznać za bardziej uniwersalne (str. 24). Ono jest jedynie bardziej rozpowszechnione i bardziej opanowane z powodu znacznie dłuższego czasu jego tworzenia. Bardziej uniwersalne jest widzenie maszynowe 3W.

5. Podsumowanie

Na uwagę zasługuje mnogość nowych pomysłów Doktoranta w zakresie rozwiązywanych zagadnień oraz mnogość przeprowadzonych obliczeń i eksperymentów, które prowadziły do coraz lepszych w sensie liczby i zakresu rozstrzyganych problemów, dokładności uzyskiwanych wyników i szybkości otrzymywanych rozwiązań, a więc szerszego zakresowo, dokładniejszego i szybszego systemu rozpoznającego. Jest to olbrzymi dorobek Doktoranta w pełni realizujący postawione w Celu Rozprawy zadanie badawcze.

Znalezione niedociągnięcia i krytyczne uwagi nie mają istotnego wpływu na zawartość i zrozumienie tekstu Rozprawy. Nie podważają też olbrzymiego dorobku Doktoranta zawartego w całej Rozprawie. Nie można też niedoceniać wiedzy i biegłości Doktoranta w zakresie omawianych w Rozprawie systemów. Zatem, Rozprawa stanowi istotny i niezaprzeczalny Jego dorobek.

Sądzę, że przedstawiona mi do rezensji Rozprawa odpowiada z nadmiarem wymogom stawianym rozprawom doktorskim. Wnoszę o dopuszczenie jej Autora do kolejnych etapów przewodu doktorskiego.

