

Instytut Automatyki i Robotyki
Politechniki Warszawskiej
ul. Św. Andrzeja Boboli 8
02-525 Warszawa

Institute of Automatic Control and Robotics
Faculty of Mechatronics
Warsaw University of Technology
Warsaw, Poland

Tel. 22 849 96 16 Fax. 22 849 03 98 E-mail: iair@mchtr.pw.edu.pl URL: <http://iair.mchtr.pw.edu.pl>



Opiekun pracy: prof. dr hab. **Barbara Siemiątkowska**

Adam Roman Kosiorek

Praca dyplomowa inżynierska

Klasyfikacja obiektów 3D na podstawie danych RGBD

3D Object Classification Based on RGBD Images

Założenia pracy:

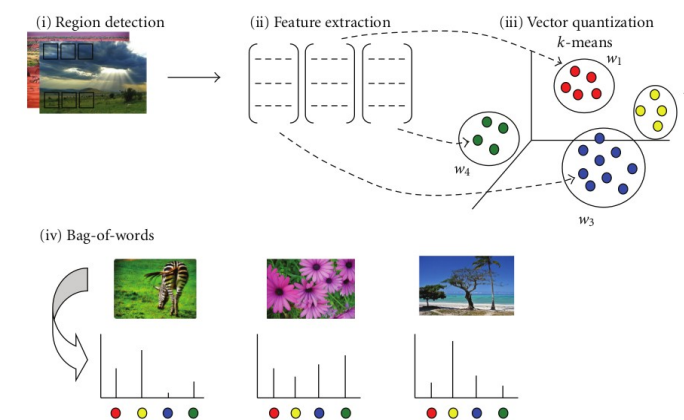
- Analizowane dane pochodzą z kamery RGBD Microsoft Kinect
- Implementacja algorytmu w języku C++ z wykorzystaniem bibliotek OpenCV i/lub PointCloud Library
- Testy aplikacji na ogólnie dostępnej, naukowej bazie danych

Zakres pracy:

- Przegląd istniejących rozwiązań dotyczących klasyfikacji obiektów na podstawie chmur punktów
- Opracowanie algorytmu klasyfikacji obiektów na podstawie chmur punktów
- Implementacja algorytmu w języku C++
- Przeprowadzenie testów szybkości oraz skuteczności klasyfikacji
- Opracowanie wniosków końcowych

Celem pracy było napisanie aplikacji do klasyfikacji obiektów 3D metodą Bag of Words na podstawie danych RGBD zarejestrowanych kamerą Microsoft Kinect. W tym celu przeprowadzono przegląd aktualnych rozwiązań i ustalono wymagania, jakie aplikacja powinna spełniać. Zaprojektowano proces przetwarzania oraz architekturę aplikacji. Następnie zidentyfikowano algorytmy, które mogły dać potencjalnie dobre rezultaty oraz przeprowadzono testy mające na celu wskazać najlepszą pod względem skuteczności oraz szybkości klasyfikacji

kombinację algorytmów. W efekcie wykonanych prac uzyskano wydajną, konfigurowalną i łatwo rozszerzalną aplikację do rozpoznawania obiektów 3D. Przewidziano możliwość i ułatwiono ponowne trenowanie klasyfikatora na innym zbiorze danych z wykorzystaniem uczenia nadzorowanego. Dzięki temu można rozpoznawać dowolne kategorie obiektów wskazane przez użytkownika aplikacji.



Rys 1: Metoda Bag of Words w przetwarzaniu obrazów. Tsai, 2012.

Do wykrywania punktów charakterystycznych wykorzystano algorytmy SIFT, ISS3D oraz Uniform Sampling. Regiony zostały opisane metodami FPFH, PFH oraz PFHRGB. Kwantyzacja została przeprowadzona algorytmem KMeans. Jako klasyfikator zastosowano SVM z jądrem RBF trenowany metodą All-vs-All.



Rys 2: Kategorie obiektów z bazy danych TOKYO, od lewej: kosz, rower, pudełko, wiadro, wózek, lodówka, laptop, drukarka, skaner, scena

W celu przetestowania aplikacji przeprowadzono przegląd dostępnych baz RGBD zawierających zdjęcia obiektów i wyodrębniono dwie z nich, jako nadające się do zastosowań w tej pracy. Skrypty w Pythonie i C++ umożliwiły ekstrakcję pojedynczych obiektów ze zdjęć i zmianę ich formatu w celu ułatwienia operacji wejścia/wyjścia w programie zasadniczym.

Uzyskana skuteczność klasyfikacji obiektów to 65.22% w przypadku zbioru 8 kategorii wyodrębnionych z Berkely 3D Object Dataset oraz 62.3% w przypadku bazy obiektów 10 kategorii udostępnionej przez Zhanga z

Descriptor	Accuracy [%]		
	ISS3D	SIFT	Uniform Sampling
FPFH	65.22	56.52	56.07
PFH	59.32	54.83	59.50
PFHRGB	63.35	52.34	53.27

University of Tokyo. Uzyskane rezultaty wskazują, że reprezentacja obiektów w formie chmur punktów jest wysoce nieefektywna. Zbiory punktów są pozbawione struktury, co zwiększa zapotrzebowanie na pamięć oraz moc obliczeniową. Najbardziej efektywne algorytmy do przetwarzania obrazów są niedostępne do zastosowań w trzech wymiarach, przez co uzyskana skuteczność jest niższa, nawet pomimo dodatkowej informacji o głębi.

Tab 1: Rezultaty uzyskane na 8 kategoriach obiektów z Berkeley 3D Object Dataset