

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Evidenčné číslo: FEI-100863-111119

**POROVNANIE SPRACOVANIA 3D DÁT ANALYTICKÝM
METÓDAMI A NEURÓNOVÝMI SIEŤAMI**

BAKALÁRSKA PRÁCA

2023

Maroš Kocúr

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Evidenčné číslo: FEI-100863-111119

**POROVNANIE SPRACOVANIA 3D DÁT ANALYTICKÝM
METÓDAMI A NEURÓNOVÝMI SIEŤAMI**
BAKALÁRSKA PRÁCA

Študijný program:	Robotika a kybernetika
Názov študijného odboru:	kybernetika
Školiace pracovisko:	Ústav robotiky a kybernetiky
Vedúci záverečnej práce:	prof. Ing. Jarmila Pavlovičová, PhD.
Konzultant:	Ing. Miroslav Kohút

Bratislava 2023

Maroš Kocúr



ZADANIE BAKALÁRSKEJ PRÁCE

Autor práce:	Maroš Kocúr
Študijný program:	robotika a kybernetika
Študijný odbor:	kybernetika
Evidenčné číslo:	FEI-100863-111119
ID študenta:	111119
Vedúci práce:	Ing. Miroslav Kohút
Vedúci pracoviska:	prof. Ing. Jarmila Pavlovičová, PhD.
Miesto vypracovania:	Ústav robotiky a kybernetiky

Názov práce: **Porovnanie spracovania 3D dát analytickými metódami a neurónovými sieťami**

Jazyk, v ktorom sa práca vypracuje: slovenský jazyk

Špecifikácia zadania: Na základe rastúcej potreby práce s 3D dátami dochádza k neustálemu testovaniu a výskumu nových možností algoritmov. Súčasná práca sa venuje algoritmu RANSAC pomocou ktorého je možné matematicky opísať rôzne zoskupenia bodov. Modely je následne možné použiť pri ďalšom spracovaní 3D dát. Cieľom práce je porovnať analytickú metódu implementácie RANSAC modelu s jeho rozšírením, ktoré implementuje aj modely neurónových sietí na dosiahnutie lepších výsledkov.

Úlohy:

1. Popíšte a analyzujte základnú problematiku práce s 3D dátami a neurónovými sieťami.
2. Analyzujte algoritmus RANSAC a možnosti jeho rozšírenia pomocou NN modelu.
3. Implementujte vami vybrané algoritmy.
4. Otestujte implementované algoritmy na zvolených dátach.
5. Zdokumentujte a vyhodnoťte výsledky.

Termín odovzdania práce: 02. 06. 2023

Dátum schválenia zadania práce: 12/08/2022

Zadanie práce schválil: **doc. Ing. Eva Miklovičová, PhD.**
garantka študijného programu

SÚHRN

SLOVENSKÁ TECHNICKÁ UNIVERZITA V BRATISLAVE
FAKULTA ELEKTROTECHNIKY A INFORMATIKY

Študijný program:	Robotika a kybernetika
Autor:	Maroš Kocúr
Bakalárska práca:	Porovnanie spracovania 3D dát analytickým metódami a neurónovými sieťami
Vedúci záverečnej práce:	prof. Ing. Jarmila Pavlovičová, PhD.
Konzultant:	Ing. Miroslav Kohút
Miesto a rok predloženia práce:	Bratislava 2023

Na základe rastúcej potreby práce s 3D dátami dochádza k neustálemu testovaniu a výskumu nových možností algoritmov. Súčasná práca sa venuje algoritmu RANSAC pomocou, ktorého je možné matematicky opísať rôzne zoskupenia bodov. Modely je následne možné použiť pri ďalšom spracovaní 3D dát. Cieľom práce je porovnať analytickú metódu implementácie RANSAC modelu s jeho rozšírením, ktoré implementuje aj modely neurónových sietí na dosiahnutie lepších výsledkov.

Kľúčové slová: RANSAC, PCL, neurónová sieť

ABSTRACT

SLOVAK UNIVERSITY OF TECHNOLOGY IN BRATISLAVA

FACULTY OF ELECTRICAL ENGINEERING AND INFORMATION TECHNOLOGY

Study Programme:	Robotics and cybernetics
Author:	Maroš Kocúr
Bachelor's thesis:	Comparison of 3D Data Processing by Analytical Methods and Neural Networks
Supervisor:	prof. Ing. Jarmila Pavlovičová, PhD.
Consultant:	Ing. Miroslav Kohút
Place and year of submission:	Bratislava 2023

Based on the growing need to work with 3D data, there is constant testing and research of new algorithm possibilities. The current work is devoted to the RANSAC algorithm, with the help of which it is possible to mathematically describe various groupings of points. The models can then be used for further processing of 3D data. The aim of the work is to compare the implementation of the analytical method of the RANSAC model with its extension, which also implements neural network models to achieve better results.

Keywords: RANSAC, PCL, neural network

Pod'akovanie

I would like to express a gratitude to my thesis supervisor.

Obsah

Úvod	1
1 PCL	2
1.1 Použitie	2
1.2 Befor installation	2
1.3 Inštalácia PCL	2
1.4 Používanie PCL	3
1.5 Kompilácia projektu	3
2 Point Cloud	4
2.1 General	4
2.2 Použitie	4
2.3 Konverzia do 3D povrchu	4
3 Neural Network	5
3.1 History	5
3.2 What is it	5
3.3 Vývoj	5
4 RANSAC	6
4.1 Overview	6
4.2 Algoritmus	6
4.3 RANSAC 3D	7
4.4 Neural guided RANSAC	7
5 KINECT	8
5.1 Ako funguje	8
5.2 Výstup z KINECTU	8
5.2.1 List potrebných Knižnic pre Ubuntu 22.04 LTS	8
5.3 Python script	9
6 Implementovanie RANSAC algoritmu	10
7 Vystupy	11
8 Implementovanie RANSAC alfgoritmu s použitím NN	12

9 Recitácia	13
10 Kod	14
Záver	15
Zoznam použitej literatúry	16

Zoznam obrázkov a tabuliek

Obrázok 3.1	Priklad neuronovej siete	6
Obrázok 4.1	Priklad algoritmu v 2D rovine	7
Obrázok 5.1	Výstupný Point Cloud z Kinectu	9

Zoznam algoritmov

Zoznam výpisov

1	Vytvorenie PCD pomocou Kinectu	14
---	--	----

Úvod

V tejto práci sa zameriame na vyhľadavanie rôznych tvarov v Point Cloude, ktorý si sami naskenujeme pomocou Kinect v2. Porovname algoritmus RANSAC do ktorého implementujeme neuronové siete a porovname efektivitu a presnosť tohto algoritmu. Taktiež sa naučíme pracovať s Point Cloud Library a hardwarom Kinect v2 na operačnom systéme Ubuntu 22.04 LTS.

1 PCL

PCL je samostný, vysoko škaloateľný, otvorený projekt pre 2D a 3D obrázky a spracovávanie point cloud. Knižnica je cross platform, napisana v jazyku C++ a Pythone. Najčastejšie sa používa na operačnom systéme Linux. Existujú package aj pre macOS a Windows vytvorené tretími stranami. My v tomto projekte budeme používať Ubuntu 22.04. LTS a verzia PCL je 1.12.1. Knižnica je vydaná pod BSD licenciami čo znamená, že je voľne použiteľná na úroveň komerčné účely a za účelom výskumu.

1.1 Použitie

PCL je open-source knižnica s algoritmami pre point cloud spracovanie úloh a spracovanie 3D geometrie, aké sa vyskytuje v trojdimenzionálnom strojovom videní. Knižnica má algoritmy na filtrovanie, odhady funkcie, rekonštrukcie povrchov, 3D registrácie, prispôsobenie modelu, vyhľadávanie objektov a segmentácie. PCL má vlastný formát na ukladanie point cloud dát - PCD, ale podporuje načítanie a ukladanie dát do rôznych iných formátov. Algoritmy sa používajú na perception v robotike na filtrovanie zašumených dát, spájanie 3D dát, segmentovanie dôležitých častí v priestore, extrahovanie kľúčových bodov a výpočet deskriptorov na rozpoznávanie objektov vo svete na základe ich geometrického vzhľadu a vytváranie povrchov z point cloudu a vykresľovanie ich.

1.2 Befor installation

PCL potrebuje niekoľko knižníc tretích strán na fungovanie, ktoré musia byť nainštalované. Väčšina matematických operácií je implementovaných v Eigen knižnici. Vizualizačný model pre 3D point cloudy je na základe VTK. Knižnica Boost je použitá na zdieľanie pointerov a FLANN knižnica pre rýchle hľadanie v okolí na základe algoritmu k-nearest. Ďalej sme nainštalovali OpenNI 2 knižnicu, ktorá nám zabezpečuje komunikáciu s Kinectom2

1.3 Inštalácia PCL

PCL je dostupný na mnoho distribúcií Linuxu ako Ubuntu, Debian, Fedora, Gentoo a Arch Linux. PCL na distribúciách Ubuntu a Debian môžeme nainštalovať pomocou.

```
sudo apt install libpcl-dev
```

Na Windowse sa PCL inštaluje pomocou vcpkg package manažéra vytvoreného Microsoftom.

```
PS> ./vcpkg install pcl
```

MacOS má Homebrew package manažéra ktorý podporuje inštaláciu packagov, ktoré Apple

alebo Linux nedokáže nainštalovať. brew install pcl Toto su odporúčane inštalacie pre PCL na daných operačných systémoch.

1.4 Používanie PCL

Na používanie PCL si potrebujeme v našom kóde vložiť potrebné knižnice. Po nainštalovaní sa v našom systéme nastaví premenné, takže stačí nám použiť príkaz v C++, `include <pcl/“názovknižnice“.h>`.

1.5 Kompilácia projektu

Vytvoríme si súbor CMakeList.txt v ktorom zadefinujeme potrebné premenné aby make vedel najst' cestu ku knižnici a vedel aké súbory ma skompilovať. Ďalej vytvoríme priečinok s názvom build, v terminály vojdeme do neho a pomocou príkazu cmake “cesta k CMakeList.txt“, si vytvoríme makefile. Ďalej používame len príkaz make ktorý nam vytvorí súbor pomocov ktorého môžeme spustiť projekt.

2 Point Cloud

2.1 General

Je to súbor bodov v priestore. Body môžu reprezentovať 3D tvary alebo objekty. Každý bod má karteziánske súradnice (X,Y,Z). Point cloud je generovaný pomocou 3D skenera alebo pomocou softwaru na fotogrametriu, ktorý meria veľ'a bodov na externom povrchu objektov okolo. My v tomto projekte použijeme Kinect 2 (odsek 6). Point cloud sa používa v 3D modelovaní, metrológii, meranie kvality výrobkov a rôzne vizualizácie. Point cloud sa často zarovnáva s 3D modelmi alebo inými point cloudmi ako registrácia množín bodov.

2.2 Použitie

Pre priemyselnú metrológiu alebo inšpekciu pomocou priemyselnej počítačovej tomografie možno mračno bodov vyrobeného dielu zosúladiť s existujúcim modelom a porovnať, aby sa skontrolovali rozdiely. Geometrické rozmery a tolerancie možno získať aj priamo z point cloudu.

2.3 Konverzia do 3D povrchu

V geografických informačných systémoch sú point cloudi jedným zo zdrojov využívaných na tvorbu digitálneho výškového medelu terénu. Používajú sa aj na generovanie 3D modelov mestského prostredia. Drony sa často využívajú na nazbieranie RGB obrázkov ktoré sa neskôr pomocou algoritmu strojového videnia ako je AgiSoft Photoscan, PixčD alebo DroneDeploy používajú na vytvorenie RGB point cloudu, kde sa môže požiť vzdialenostná a objemová aproximácia.

3 Neural Network

3.1 History

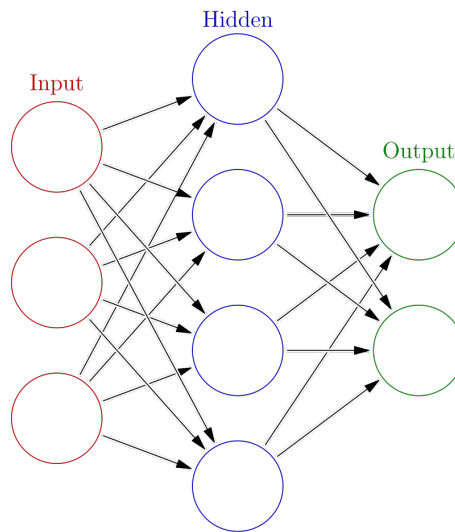
V posledných rokoch, najlepšie systémy s aplikáciou umelej inteligencie - ako sú rozpoznávače reči v mobilných zariadeniach alebo automatické prekladače majú dobré výsledky v technike nazývajúcej sa "deep learning". Deep learning je v podstate iné meno pre aplikáciu umelej inteligencie s názvom Neural network, ktorý sa vyvíja takmer 80 rokov. NN boli prvý krát navrhnuté v roku 1944, dvoma výskumníkmi z Chicagskej univerzity, ktorí v roku 1952 prešli na MIT a založili oddelenie kognitívnych vied.

3.2 What is it

NN sú myslené na robenie strojového učenia, v ktorom sa počítač učí vykonávať úlohy na základe analyzovania trenovacích príkladov, ktoré sú zvyčajne ručne označené. Systém na rozpoznávanie objektov by mohol mať prístup k tisíciam obrázkov označených ako auto, dom, guľa a podobne. NN sieť hľadá vizuálne patery v obrázku ktoré majú vzajomný vzťah s konkrétnym označením. Voľne modelovaná NN na ľudskom mozgu má tisíce až milióny jednoduchých node, ktoré sú husto prepojené. Väčšina siete sa dnes organizuje do vrstiev node a tie posielajú údaje vpred, čo znamená že dáta nimi prechádzajú iba v jednom smere. Jedna noda môže byť pripojená k viacerým nodam vo vrstvách pod ňou z ktorej prijíma dáta a viacej node vo vrstve nad ňou kde spracované dáta posielajú. Ku každému vstupnému prepoju sa pridá číslo známe ako "váha". Ak je sieť aktívna, noda prijíma rôzne dáta z každého vstupu a vynásobí ich pridelenou váhou, potom všetky výsledky sčíta ak je výsledne číslo pod threshold hranicou, noda nepošle žiadne dáta do ďalšej vrstvy. Ak číslo prekračuje threshold hodnotu, tak noda "výstrelí", čo znamená že pošle súčet vstupov na všetky výstupy. Keď sa NN trénuje, všetky váhy a thresholdy sú nastavené na náhodnú hodnotu. Trénovacie dáta sú posielané do spodnej vstupnej vrstvy a prechádzajú cez nasledujúce vrstvy, násobia sa a sčítavajú sa v zložitými cestami, pokiaľ radikálne transformované nedojdu na vrchnú výstupnú vrstvu. Počas tréningu sa váhy a prahové hodnoty neustále upravujú kým tréningové údaje s podobnými štítkami neprinášajú podobne výstupy.

3.3 Vývoj

Neuronové siete opísané v roku 1944 mali váhy a threshold hodnoty, ale neboli usporiadané do vrstiev a výskumníci nešpecifikovali trénovací mechanizmus. Výskumníci dokázali ukázať, že NN dokáže v princípe vypočítať hocikakú funkciu ako digitálny počítač. Výsledkom bola viac neurová veda ako počítačová a cieľom bolo poukázať, že ľudský mozog môžeme považovať za výpočtové zariadenie.



Obr. 3.1: Příklad neuronovej siete

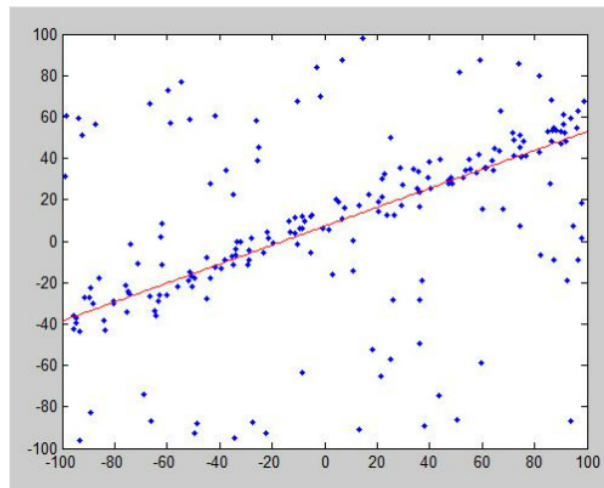
4 RANSAC

4.1 Overview

RANSAC je algoritmus vytvorený Fischlerom a Bollesom, určuje všeobecný prístup k odhadu parametrov, s veľkým podielom outliers v stupnom datasete. Na rozdiel od iných výkonných algoritmov odhadu, ako napríklad M-estimators a meródov najmenších štvorcov s prepojením na strojové učenie. RANSAC bol vytváraný komunitou ľudí používajúci strojové učenie. RANSAC je vzorkovacia technika ktorá generuje kandidáta na minimalny počet pozorovani potrebných na zistenie odhadu parametrov ležiacich pod modelom. Na rozdiel od ostatných vzorkovacích algoritmov, ktoré používajú čo najviac bodov ako môžu, RANSAC používa najmenší počet bodov ako môže.

4.2 Algoritmus

1. Vybranie minimum náhodných bodov potrebných na určenie parametrov modelu
2. Vyriešenie parametrov pre model
3. Určenie koľko bodov z množiny všetkých bodov leží s preddefinovanou ϵ .
4. Ak zlomok bodov ležiacich v preddefinovanej ϵ , presahuje preddefinovaný prah τ , prehodnotí parametre modelu použitím všetkých identifikovaných inliers a ukončí.
5. Inak, opakuj kroky 1 až 4, s maximálnym N opakovaniami.



Obr. 4.1: Příklad algoritmu v 2D rovine

Modre sú body z datasetu, pomocou RANSAC algoritmu sa určili 2 body, ktoré majú vo svojom subsete najmenej bodov ležiacich mimo alfy.

4.3 RANSAC 3D

Podobne ako pri opise vyššie RANSAC 3D funguje s výberom náhodných troch bodov na ktorých zostaví rovinu a spočíta body ležiace v rovine a body ležiace mimo roviny. Algoritmus sa opakuje n-opakovaní a výstupom je najlepší model.

4.4 Neural guided RANSAC

5 KINECT

Kinect je vstupné zariadenie snímajúce pohyb, vyrobené Microsoftom. Zariadenie obsahuje RGB kamery a infračervený projektor a detektor ktorý monitoruje hĺbku priestoru na základe štrukturovateľného svetla alebo na základe času trvajúceho svetlu dopadnúť na objekt, vďaka ktorému vie kinect poskytnúť rekognizáciu miest v reálnom čase. Kinect sa používa hlavne v hernom priemysle, ale používa sa taktiež na komerčné a akademické účely pretože poskytuje mapovanie priestoru a je lacnejší než profesionálne zariadenia.

5.1 Ako funguje

Infračervený projektor na kinecte posiela modulované infračervené svetlo ktoré je zachytené sensormi. Infračervené svetlo ktoré sa odrazí od bližších objektov má kratší čas letu ako svetlo ktoré sa odrazí od vzdialenejších objektov, takže sensor sníma ako vymodulovaný vzor bol deformovaný z času letu svetla, pixel po pixeli. Čas priletu meranej hĺbky touto metódou môže byť presnejšie vypočítaný v kratšom čase, čo zabezpečí viac snímkov za sekundu. Hneď ako kinect naskenuje hĺbkovú fotografiu, použije metódu zisťovania hrán k vytýčeniu bližších objektov z pozadia fotky.

5.2 Výstup z KINECTU

Na získanie výstupu z Kinectu je potrebné si nainštalovať Open-source knižnicu LibFreenect2, ktorá je určená na získavanie videí z Kinectu verzie 2. Knižnica disponuje vzorovým kódom, ktorý po spustení zapne Viewer a okno rozdelí na štyri časti a v nich uvidíme výstupy Infra Red kamery, farebnej kamery a depth kamery. Nato aby sme knižnicu mohli používať potrebovali sme doinštalovať potrebné knižnice a súčasti. Na získanie snímky z Kinectu sme použili Python script, ktorý nám vytvorí point cloud súbor, v ktorom budeme hľadať požadované tvary.

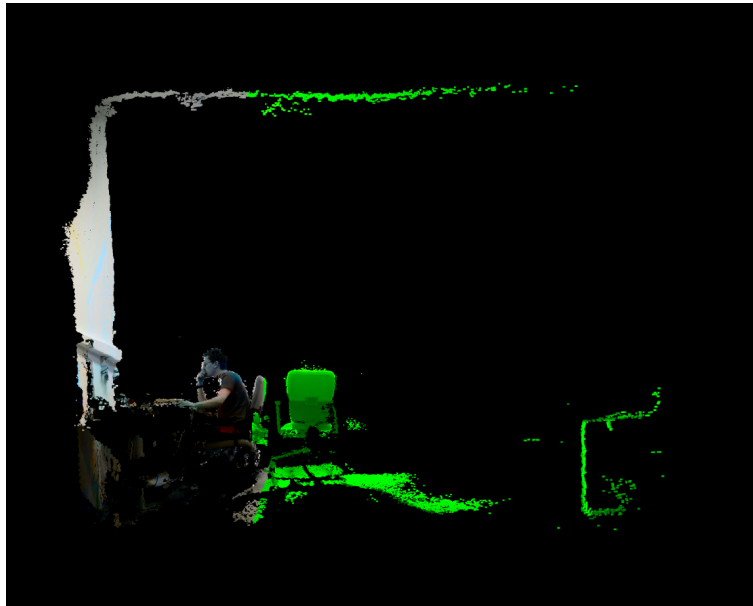
5.2.1 List potrebných Knižníc pre Ubuntu 22.04 LTS

- libusb
- TurboJPEG
- OpenGL
- OpenCL (optional)
- CUDA (optional, NVIDIA only)
- VAAPI (optional)

- OpenNI2

5.3 Python script

Z githubu si cloneme `freenect2-python`, ktorý obsahuje scripty pomocov ktorých vieme získať snímky z Kinectu. Script `dump-pcd.py` nám vytvorí 2 point cloudy. Po menšej úprave v kóde výstupom je jeden Point cloud a snímka priestoru.



Obr. 5.1: Výstupný Point Cloud z Kinectu

6 Implementavanie RANSAC algoritmu

7 Vystupy

8 Implementovanie RANSAC algoritmu s použitím NN

9 Recitácia

Citujem všetky zdroje v **bibliography.bib**, [1, 2, 3].
Good luck.

10 Kod

```
from freenect2 import Device, FrameType
import numpy as np

# Open the default device and capture a color and depth frame.
device = Device()
frames = {}
with device.running():
    for type_, frame in device:
        frames[type_] = frame
        if FrameType.Color in frames and FrameType.Depth in frames:
            break

# Use the factory calibration to undistort the depth frame and register the RGB
# frame onto it.
rgb, depth = frames[FrameType.Color], frames[FrameType.Depth]
undistorted, registered, big_depth = device.registration.apply(
    rgb, depth, with_big_depth=True)

# Combine the depth and RGB data together into a single point cloud.
with open('output.pcd', 'wb') as fobj:
    device.registration.write_pcd(fobj, undistorted, registered)

with open('output_big.pcd', 'wb') as fobj:
    device.registration.write_big_pcd(fobj, big_depth, rgb)
```

Výpis 1: Vytvorenie PCD pomocou Kinectu

Záver

Conclusion is going to be where?

Here.

Zoznam použitej literatúry

1. HARDESTY, Larry (zost.). *Explained: Neural networks: metodický materiál pro autory vysokoškolských kvalifikačních prací* [online]. MIT: MIT News Office, 2014-04-14 [cit. 2014-04-14]. Dostupné z : <https://news.mit.edu/2017/explained-neural-networks-deep-learning-0414>.
2. RADU BOGDAN RUSU, Steve Cousins (zost.). *Point cloud library (pcl): 3D is here*. 2011 IEEE international conference on robotics and automation: IEEE. Dostupné tiež z: https://scholar.google.com/citations?view_op=view_citation&hl=en&user=U_NEZHQAAAAJ&citation_for_view=U_NEZHQAAAAJ:u-x6o8ySG0sC.
3. *Point cloud library*. Dostupné tiež z: <https://pointclouds.org/>.