Bachelor Project

Czech Technical University in Prague

F3 Faculty of Electrical Engineering
Department of Cybernetics

Drone detection using neural networks from combined RGB camera and LiDAR data

Adam Škuta

Supervisor: Matouš Vrba
Supervisor–specialist: Martin Saska
Field of study: Mathematical Engineering
Subfield: Mathematical Modelling
February 2017

Acknowledgements

Děkuji ČVUT, že mi je tak dobrou *alma mater*.

Declaration

Prohlašuji, že jsem předloženou práci vypracoval samostatně, a že jsem uvedl veškerou použitou literaturu.

V Praze, 10. February 2017

Abstract

Keywords: word, key

Supervisor: Matouš Vrba
Ústav X,
Uliční 5,
Praha 99

Abstrakt

Klíčová slova: slovo, klíč

Překlad názvu: Moje bakalářka se
strašně, ale hrozně dlouhým předlouhým
názvem — Cesta do tajů kdovíčeho

Contents

Figures

Tables

Chapter **1**

Introduction

The goal of this thesis is to prove whether a usage of LiDAR data coupled with images from RGB is useful for the localization of UAVs in contrast to the usage of image data alone. The LiDAR and RGB camera will be mounted on top of the scanner UAV. All the measurements will be taken inside a virtual environment, with a realistic UAV and sensor simulation. The dataset will then be preprocessed and used as the input to a Convolutional neural network for the object detection. The preprocessing will be as following:

- Coordinate transformation for the non-matching coordinate systems

- Projection of 3d points into a 2d image

- Using a Convolutional neural network to generate more points in a sparse LiDAR pointcloud

Infographic of the process

Chapter 2

Convolutional neural networks

A convolutional neural network will be used in two problems in the thesis. First one is used on a sparse LiDAR pointcloud generating more points and therefore making it more dense. Second one is used for object, int this case drone, detection, using RGB and RGBD data as the input.

2.1 Sparse to dense

Sparse to dense is a Convolutional neural network written in PyTorch. It predicts the depth measurements from sparse depth dataset.

Chapter 3

Sensors

3.1 Coordinate systems

In order to correctly label the data for training, a position of the second drone in relation to the camera mounted on the first one is required. The API call in AirSim returns a position in relation to its starting point. Therefore a transformation from the starting point of the second drone to the camera mounted on the first one is required. We can write this transformation as follows:

$$\mathbf{T} = \mathbf{T}_{d1}^{c} \mathbf{T}_{s1}^{d1} \mathbf{T}_{s2}^{s1} \tag{3.1}$$

where:

- $\mathbf{T}_{s2}^{s1}$ is transformation from the starting point of the second drone to the starting point of the first drone

- $\mathbf{T}_{s1}^{d1}$ is transformation from the starting point of the first drone to the body of the first drone

- $\mathbf{T}_{d1}^{c}$ is transformation from the body of the first drone to the cameras coordinate system

insert picture showing the positions of the drones

Transformation matrix $\mathbf{T}$ can generally be described as follows:

$$\mathbf{T} = \begin{bmatrix} \mathbf{R} & \mathbf{p} \\ \mathbf{0}^{T} & 1 \end{bmatrix} \tag{3.2}$$

where:

- **R** is a 3x3 rotation matrix

- **p** is a 3x1 translation column vector

- $\mathbf{0}^T$ is a 1x3 row vector of zeros

◼ 3.2 Camera Model

For the creation of the bounding boxes used for training a transformation from coordinate system of the camera to the pixel values of the image needs to be defined. For this task a pinhole camera model is used. The transformation is then defined as follows:

$$
\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} f\frac{x}{z} + c_x \\ f\frac{y}{z} + c_y \end{bmatrix}
\tag{3.3}
$$

where:

- x' and y' are pixel coordinate values on the image

- x,y,z are coordinate values of a point to be transformed

- f is focal length of the camera

- c_x,c_y are offsets on the image plane

Picture of pin-hole camera

Chapter 4

Dataset

The dataset for this work can be generated in two ways. The first is real-life drone shots mixed with point clouds from LiDAR mounted on top of a drone. The second is generating a dataset using a realistic virtual environment where a drone, camera and LiDAR are being emulated very close to their real-life counterparts. An advantage to this approach is that a great variety of environments can be chosen a lot of them often inaccessible otherwise (power plant, airport, snowy mountains out of season etc.). Therefore this approach will be chosen for the task.

4.1 Unreal Engine

Unreal Engine is a software tool used for creating realistic 3d environments, most often used as a video game engine. It is written in C++ and open-source supporting a variety of pre-built environments and assets. For this work three diferent environments will be used for the creation of the dataset:

> citation https://www.unrealengine.com/en-US/features

-
  > exact name. Snow

-
  > exact name. Park

▪

> exact name. City centre

Pictures of the environments

Together

> exact number of pictures

pictures and labels were generated using two drones. One drone was equipped with RGB camera and LiDAR sensor and was responsible for taking the pictures and pointclouds from LiDAR. The second one was used as a model for drone detection.

4.2 AirSim

Open-source plugin for Unreal Engine called AirSim was used for the generation of the dataset. It simulates realistic flight motions of drones as well as

airsim zdroj

seven types of sensors , including RGB camera and LiDAR used for this task. AirSim supports both a C++ API as well as Python API, latter which was used for controlling the motion and capturing the dataset. Location of the second drone was generated through API call, which produces a location of the drone in global coordinate system of the map, which is later transformed to the local coordinates of the first drone carrying the LiDAR and RGB

Transformacna matica?

sensors.. The capturing drone traveled on each map on a 3d cube grid:

> Grafika kocky po ktorej lietal dron

4.3 Camera Model

Part I

Your Party

Chapter **5**

Heading on Level 0 (chapter)

Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. $\sin^2(\alpha) + \cos^2(\beta) = 1$. If you read this text, you will get no information $E = mc^2$. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. $\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$. This text should contain all letters of the alphabet and it should be written in of the original language. $\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$. There is no need for special content, but the length of words should match the language. $a\sqrt[n]{b} = \sqrt[n]{a^n b}$.

5.1 Heading on Level 1 (section)

Hello, here is some text without a meaning. $d\Omega = \sin\vartheta d\vartheta d\varphi$. This text should show what a printed text will look like at this place. If you read this text, you will get no information. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. $\sin^2(\alpha) + \cos^2(\beta) = 1$. This text should contain all letters of the alphabet and it should be written in of the original language $E = mc^2$. There is no need for special content, but the length of words should match the language. $\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$.

▪ 5.1.1 Heading on Level 2 (subsection)

Hello, here is some text without a meaning. $\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$. This text should show what a printed text will look like at this place. $a\sqrt[n]{b} = \sqrt[n]{a^n b}$. If you read this text, you will get no information. $d\Omega = \sin\vartheta d\vartheta d\varphi$. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. This text should contain all letters of the alphabet and it should be written in of the original language. There is no need for special content, but the length of words should match the language. $\sin^2(\alpha) + \cos^2(\beta) = 1$.

▪ Heading on Level 3 (subsubsection)

Hello, here is some text without a meaning $E = mc^2$. This text should show what a printed text will look like at this place. $\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$. If you read this text, you will get no information. $\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. $a\sqrt[n]{b} = \sqrt[n]{a^n b}$. This text should contain all letters of the alphabet and it should be written in of the original language. $d\Omega = \sin\vartheta d\vartheta d\varphi$. There is no need for special content, but the length of words should match the language.

Heading on Level 4 (paragraph). Hello, here is some text without a meaning. This text should show what a printed text will look like at this place. $\sin^2(\alpha) + \cos^2(\beta) = 1$. If you read this text, you will get no information $E = mc^2$. Really? Is there no information? Is there a difference between this text and some nonsense like "Huardest gefburn"? Kjift – not at all! A blind text like this gives you information about the selected font, how the letters are written and an impression of the look. $\sqrt[n]{a} \cdot \sqrt[n]{b} = \sqrt[n]{ab}$. This text should contain all letters of the alphabet and it should be written in of the original language. $\frac{\sqrt[n]{a}}{\sqrt[n]{b}} = \sqrt[n]{\frac{a}{b}}$. There is no need for special content, but the length of words should match the language. $a\sqrt[n]{b} = \sqrt[n]{a^n b}$.

5.2 Lists

5.2.1 Example for list (itemize)

- First item in a list
- Second item in a list
- Third item in a list
- Fourth item in a list
- Fifth item in a list

Example for list (4*itemize)

- First item in a list
  - First item in a list
    - First item in a list
      - First item in a list
      - Second item in a list
    - Second item in a list
  - Second item in a list
- Second item in a list

5.2.2 Example for list (enumerate)

1. First item in a list
2. Second item in a list
3. Third item in a list
4. Fourth item in a list
5. Fifth item in a list

■ Example for list (4*enumerate)

1. First item in a list
    a. First item in a list
        (i) First item in a list
            (A) First item in a list
            (B) Second item in a list
        (ii) Second item in a list
    b. Second item in a list
2. Second item in a list

■ 5.2.3 Example for list (description)

First item in a list

Second item in a list

Third item in a list

Fourth item in a list

Fifth item in a list

■ Example for list (4*description)

First item in a list

    First item in a list

        First item in a list

            First item in a list

            Second item in a list

        Second item in a list

    Second item in a list

Second item in a list

Foo	Bar
foo1	bar1
foo2	bar2

Table 5.1: Foobar.
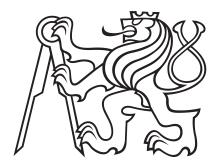
Figure 5.1: Black logo of the CTU in Pragueueue.
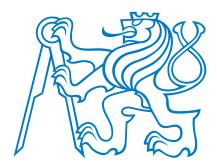
Figure 5.2: Blue logo of the CTU in Pragueueue.

Chapter **6**

Conclusions

6.1 Test — this is just a little test of something in the table of contents

6.1.1 Yes, table of contents

Theorem 6.1. 1. *Bla*

2. *Blo*

Proof. 8 Bla

1. Blo

□

Appendices

Appendix **A**

Index

A

affine, 7

C

Cardano, 9

E

extrinsic, 9

F

field, 18

free, 18

function, 6

G

Gaussian, 23

Germain, 17

H

holomorphic, 13

I

ideal, 6

isomorphism, 11

L

Leibniz–Poisson, 22

M

matrix, 6

modulus, 20

monoid, 10, 20

N

natural, 21

null, 16

O

open, 11

P

point, 17

positive, 16

prime, 16

S

subset, 13

▍T

triangle, 23

▍U

universal, 7

▍V

von Neumann, 10

Appendix B

Bibliography

[Lei97]   H. Leibniz, *Problems in abstract calculus*, Journal of Real Representation Theory **66** (1997), 302–339.

[SW05]   X. Suzuki and A. Wu, *Maximality methods in rational graph theory*, Archives of the Indian Mathematical Society **20** (2005), 204–223.

[TLJ92] Z. Thomas, H. Lindemann, and A. Jones, *Subgroups for a random variable*, Journal of Axiomatic Galois Theory **15** (1992), 520–528.

Katedra: matematiky

Akademický rok: 2008/2009

ZADÁNÍ BAKALÁŘSKÉ PRÁCE

Pro: Tomáš Hejda

Obor: Matematické inženýrství

Zaměření: Matematické modelování

Název práce: Spřátelené morfismy na sturmovských slovech / Amicable Morphisms on Sturmian Words

Osnova:

1. Seznamte se se základními pojmy a větami z teorie symbolických dynamických systémů.

2. Udělejte rešerši poznatků o sturmovských slovech: přehled ekvivalentních definic sturmovských slov, popis morfismů zachovávajících sturmovská slova, popis standardních párů slov.

3. Zkoumejte vlastnosti párů spřátelených sturmovských morfismů, pokuste se popsat jejich generování a počty v závislosti na tvaru jejich matice.

Doporučená literatura:

1. M. Lothair, Algebraic Combinatorics on Words, Encyclopedia of Math. and its Applic., Cambridge University Press, 1990

2. J. Berstel, Sturmian and episturmian words (a survey of some recent result results), in: S. Bozapalidis, G. Rahonis (eds), Conference on Algebraic Informatics, Thessaloniki, Lecture Notes Comput. Sci. 4728 (2007), 23-47.

3. P. Ambrož, Z. Masáková, E. Pelantová, Morphisms fixing a 3iet words, preprint DI (2008)

Vedoucí bakalářské práce: Prof. Ing. Edita Pelantová, CSc.

Adresa pracoviště: Fakulta Jaderná a fyzikálně inženýrská
                               Trojanova 13 / 106
                               Praha 2

Konzultant:

Datum zadání bakalářské práce: 15.10.2008

Termín odevzdání bakalářské práce: **7.7.2009**

V Praze dne 17.3.2009

....

        Vedoucí katedry   Děkan