HTBLuVA Wiener Neustadt
Höhere Lehranstalt für Informatik

DIPLOMARBEIT

Localisation via ML Methods

Ausgeführt im Schuljahr 2019/20 von:

Rolle x Ida Hönigmann	5AHIF
Rolle y Peter Kain	5AHIF

Betreuer / Betreuerin:

MMag. Dr. Michael Stifter

Wiener Neustadt, am September 16, 2019/20

Abgabevermerk: Übernommen von:

Eidestattliche Erklärung

Hiermit erkläre ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst und keine anderen als die im Literaturverzeichnis angegeben Quellen und Hilfsmittel verwendet habe. Insbesondere versichere ich, dass ich alle wörtlichen und sinngemäßen Übernahmen aus anderen Werken als solche kenntlich gemacht habe.

Wiener Neustadt am September 16, 2019/20

Verfasser / Verfasserinnen:

Ida HÖNIGMANN Peter KAIN

Contents

Acknowledgement

We would like to thank ...

Kurzfassung

asdf

Abstract

asdf

Chapter 1

Introduction

Author: Ida Hönigmann

Robots are getting more and more mobile. While a few years ago their usage was mostly limited to aid factory automation, robots have found widespread adoption in a multitude of industries, such as self driving cars and autonomous delivery drones. A challenge frequently encountered is navigating in unknown environments, which either requires the robot to sense specific characteristics of its surroundings or to communicate with some external system.
The problem of navigation has been looked at from many different angles. One popular approach in mobile robotics is to use the GPS, an external positioning system. In order to determine the position of a robot using the GPS, it has to establish communication with at least four satellites. The exact position of each satellite as well as the current time is broadcast by the satellites. By measuring the time needed for the signal to reach the robot, the position can be calculated up to three meters accurately.
However, in some cases positioning a robot using external positioning methods is not possible. In the case of the GPS this can be due to obstacles interfering with the radio signals send by the satellites, for example occurring inside a building. In comparison, we focus on a system that can navigate in outdoor as well as in indoor environments.

1.1 Goal

The goal of this diploma thesis is to implement a system which can localize a robot using no other sensors than a camera. This limitation was purposely chosen as our system will be used by future robotic students at our school and many robot systems used in the field of education are only poorly equipped with sensors that are able to detect its environment. One sensor used in the field of educational robotics is the either already equipped, or easily mountable camera.
As part of our work we not only want to implement an easy to use API for future

robotic students, but to also show the possibilities and advantages of machine learning in localisation.
In order to accomplish precise localisation in various different surroundings, we plan on implementing a neural network. The neural network should take images, taken by the camera, as an input, and outputs the relative distance to any object shown in the images. By using machine learning we hope to be less dependent on a specific situation or setup in comparison to different camera based localisation methods. For example the localisation should work on objects varying in size and shape, as well as in different situations of lighting.

1.2 Motivation

In July 2019 the two authors of this work participated in the aerial tournament at the Global Conference on Educational Robotics held in Norman, Oklahoma. One of the two main challenges encountered at this tournament was landing a drone next to some randomly placed object, which colour, shape and size was known in advance.
The second challenge we faced was flying from one side of randomly placed cardboard boxes to the other. The cardboard boxes, representing a mountain, are placed in one of various configurations, one of which can be seen in figure 1.1.

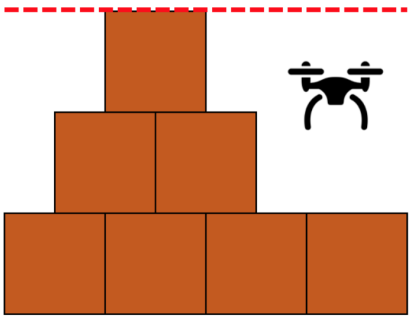

Figure 1.1: Seven cardboard boxes, representing a mountain, are placed in a random configuration. The team scores points if the drone passes the mountain while staying under the height limit indicated by the red dotted line.

The drones used at this aerial tournament are equipped with a camera, while lacking any other sensor that can be used to detect the obstacles and game items. Therefore we needed to be able to detect the distance to the object and the cardboard boxes using only the camera. At the Global Conference on Educational Robotics we decided to detected the object based on its colour, but had to invest quite some time to tweak the values to get the localisation working correctly. Therefore we want to research and implement a method that is more robust than the colour based one.

1.3 Outlook / Perspective

The objective of this work is to create a system which uses machine learning methods in localising objects. After having trained the system, it should reliably return the x, y and z distances to an object, shown in two pictures taken from different angles.
It is planned that the distance will be measured from the second camera position to the centre of object.
[TODO: Foto von Drohne mit Objekt]

1.4 Equipment

[TODO: Welche Drohnen verwenden wir?, Kamera auf der Drohne]

Chapter 2

Study of Literature

Author:

2.1 Different Approaches to the Problem

[Input: Paper (gleiches Problem ohne NN) finden - Peter]
[Input: Video (eine Kamera, Entfernung zu Punkt (größe bekannt)) finden (ohne NN) - Peter]

2.1.1 LIDAR

2.1.2 Structure from Motion

2.1.3 Feature Tracking

2.2 Depth perception

[TODO: humans, two eyes - gleich wie bei unserem Aufbau]

2.2.1 Depth sensation

[TODO: Pigeons, deer, children (visual cliff)]

Chapter 3

Methodology

Author:

3.1 Stereo Camera

[warum verwenden wir keine stereo kamera?, weil keine vorhanden]

3.2 Image preprocessing

[schwarz-weiß? auflösung? filter? object detection?]

3.3 Neural Network

[TODO: Funktionsweise]

3.4 Generating data

[blender]

Chapter 4

ROS2

Author:

4.1 What is ROS?

[Ubuntu?]
[Core Concepts]

4.1.1 Nodes

[Wie verwenden wir das? / Wie macht es unsere Arbeit leicher?]

4.1.2 ...

4.2 Why use ROS?

4.3 Comparing ROS and ROS2

[python2 < python3]

Chapter 5

Implementation

Author:

5.1 Generating test data

Good test data is one of the most important things in machine learning. The system knows only what is depicted in the training data, which is why it is important to include as many aspects of the problem as possible in this data.
Since machine learning needs a lot of data in order to solve the given task it can be tiresome to generate and label all this data by hand. Therefore we decided to simulate some objects using a computer graphics software called Blender.
Blender allows for relatively easy generation of training data by providing a Python API.
[TODO: Image camera setup, lightning, objects in Blender] [TODO: Renders and labels for example objects]

5.2 OpenCV

[TODO: was ist das? wofür wird es allgemein verwendet? wofür verwenden wir es?] [TODO: Codebeispiele?]

5.3 Neural Network

[TODO: Erklärung - zuerst Tensorflow, dann selbst implementiert in C++]

5.3.1 Tensorflow

[TODO: was ist Tenserflow, welche Sprachen werden unterstützt, wie wird es (sonst noch) verwendet?]

5.3.2 Alternatives to Tensorflow

[+ warum verwenden wir ausgerechnet Tenserflow]

5.3.3 Structure of our Neural Network

[wie lesen wir Daten ein, wie viele layer, was ist der output (maximale entfernung? z.B. 10m)]

5.3.4 C++ Implementation

5.3.5 Technical difficulties

Chapter 6

Experiment 1

Author:

6.1 Environment

[TODO: viele Fotos] [TODO: Hintergrundfarbe, Untergrundfarbe, Struktur (Hintergrund und Untergrund), Beleuchtung (Art, Helligkeit, Richtung, mehrere Lichtquellen, welche Lichtquellen, ...)]

6.2 Setup

[TODO: viele Fotos] [TODO: welche Objekte (Größe, Farbe, wie viele (mindestens 3)?, ...), Kameras, Entfernung zu Objekten]

6.3 Sequence of Events

6.4 Results

Chapter 7

Lessons learned

Author:

Chapter 8

Experiment 2

Author:

8.1 Environment

[TODO: viele Fotos]

8.2 Setup

[TODO: viele Fotos]

8.3 Materials

[TODO: viele Fotos]

8.4 Sequence of Events

8.5 Results

Chapter 9

Conclusion

Author: