Technische Universität Berlin

Fakultät IV - Elektrotechnik und Informatik

Fachgebiet AOT
Prof. Dr. Sahin Albayrak

Master's Thesis Proposal

A Vision-Based Navigation Approach for MAVs with Powers of Recall

Friedrich Clemens Valentin Mangelsdorf
Matrikel–Nummer 356686

Supervisor Dr. rer. nat. Yuan Xu

Table of Contents

1 Motivation

Im Bereich ... (Warum muss es eine neue L"osung/ einen neuen Ansatz geben)

The advances in drone technology over the past years have been just as fantastic as associated application possibilities in the public, private and voluntary sector of today's and future economies. Especially the branches infrastructure, transport, insurance, media and entertainment, telecommunication, agriculture, security and mining offer great potential for the commercial use of drones. [?] Drones exhibit substantial advantages over ground vehicles as they are unaffected by many obstacles and largely independent of infrastructure. Destinations can be reached by the shortest route, waiting times can be avoided and the effort and risk of getting to places that are difficult to access can be reduced. In addition, from the privileged perspective of a bird, onboard sensors are able to record extensive data of high quality as well as increased speed due to the fast aerial maneuverability of drones. On the downside, drones as smaller aircraft vehicles can move much less weight which, among other things, limits mission payload and battery capacity, which in turn restricts flight range and duration. In order for the use of drones to pay off economically, efficiency must therefore be maximized. One key to increase efficiency is to shift functions from a human operator to the drone itself, i.e., increase the drone's degree of autonomy. This becomes particularly clear in the example of delivery applications. Because drones can load significantly fewer packages and have to recharge or refuel more often than a conventional delivery truck, drone delivery systems would only really pay off if most decisions are made autonomously by the individual delivery drone or even by the collective of the drone swarm. Consequently, expensive human labor could be reduced and room for mathematical optimization could be created.

While in the commercial context, already today drones are proving their economic and safety-related value in aerial inspection services of more controlled and undisturbed environments of "large industrial sites" such as agriculture, construction, infrastructure, utilities, and mining, they have not yet have not yet really been able to assert themselves for commercial applications in the open-world environment of our daily lives. This is particularly true for drones with autonomous functions.

drone-in-a-box -In heavy industry, automation and drone use is now seen as the standard in inspection practices -Our 2021 solution, combining powerful autonomous robots with AI-powered visual data management is the nuanced and advanced solution industrial sites today can rely on to reduce risk, costs and environmental impact.

Apart from legal restrictions and lack of acceptance in the population [Quelle], decisive reasons for this are of a technical nature. For example, autonomous navigation methods are not yet robust enough for reliable deployment in densely populated urban areas.[**?**] The master's thesis is intended to make a contribution here.

————BIS HIER

In common parlance, the term "drone" is often used to refer to the aircraft class of unmanned aerial vehicles (UAVs), which can be further devided into the sub-classes of

This proposal deals with the sub-class of micro aerial vehicles (MAVs) which represent ... . Classification... Current research revolves around... batteries, lalala and more autonomy, especially in navigation.

As a basic functionality of MAVs, navigation comprises the main task of achieving a desired pose or position while performing necessary sub-tasks at the same time. The necessity of implementing individual sub-tasks depend either on the environment, e.g., obstacle avoidance and coordination with other agents, or on the application, e.g., shooting plant seeds []. Consequently, state of the art allows the degree of autonomy of MAVs in some environments and applications and not in others.

When again taking drone delivery as an example, projects are already realized in rural areas where the airspace is mainly undisturbed []. Here, navigating through waypoints only relying on GNSS without any implementation of obstacle avoidance and agent-coordination may be sufficient.

In contrast, urban areas are full of unstructered obstacles and other agents which result in a high uncertainty that cannot be planned in advance.

Only a high level of autonomy in navigation which has not yet been achieved can robustly cope with the challenges of this environment.

Recent research on autonomous MAV navigation is mainly based on deep learning which allows to perceive and reason the immediate environment. State-of-the-art navigation methods achieve a high spatial understanding of the environment by feeding convolutional neural networks (CNN) with vision or depth data.

This research aims to develop a simple navigation method that extend this spatial perception onto temporal extension by serially connecting a CNN with a long-short-term-memory (LSTM) network.

Based on the assumption that powers of recall are crucial for humans when navigating, I am convinced that future autonomous navigation systems will also encompass this ability. The navigation method will be tested in simulation and real world in a simplified test scenario, which, however, requires the MAV to remember the expansion and relative motion of obstacles while considering its own elapsed acceleration.

2 Objectives

Im Rahmen ... (Was will ich ueberhaupt mit meiner Abreit erreichen? Etwas verbessern, entwickeln, vergleichen...)

The following objectives should be achieved within the framework of the master's thesis:

Many advanced methods for autonomous navigation of MAVs already exist. However, they are not sophisticated enough to conquer open-world environments of high uncertainty. In current research, deep learning techniques that empower MAVs with perception and reasoning abilities constitute the best approach to face the uncertainty of these environments.

State-of-the-art, vision- or depth-based navigation methods integrate feedforward, deep convolutional neural networks (CNNs) that map the current color or depth image to action. [Sources]

By deploying CNNs in this way a high, spatial perception and reasoning of the immediate surrounding of the MAV can be achieved. Yet, the mere comprehension of space, however good it is, may not be sufficient to robustly deploy autonomously navigating MAVs in open-world environments. Therefore, I aim to develop a navigation method that besides spatial also includes temporal comprehension.

To my knowledge, researchers from ETH Zurich have come up with very impressing work with respect to autonomous MAVs.

Current paper: ...

Want to take a step back to the work of Loquercio, Kaufmann et. al. [**?**] from 20??: They developed a vision-based method that navigates a MAV through a drone racing track with possibly dynamically moving race gates. Thereby, they achieved a high reliability and agility at high speeds.

drone racing is a good test environment... - reactive, inherent problem of high-level goal formulation in deep learning policies

Taking their work as a basis, by adding temporal comprehension, I expect the MAV to entfalten das folgende Potential:

- As mentioned above, the navigation method theoretically does not require a high-level goal formulation since the reactive targeting the next gate already results in completing the race track. Practically, in the event that at any time the next gate is not located in the frame of view (FOV) of the MAV's onboard camera, the output of the deep neural network (DNN) is not defined and the lack of a high-level goal becomes evident. This is also the case, even if the MAV has seen the next gate in the past (see figure ??).

- In addition, humans cannot estimate the velocity and direction of motion of themselves, obstacles or other agents based on a short blink with their eyes but they need to observe over at least a short period of time. - Not only localization but also situational reasoning is strengthened by memory, e.g., a car driver observes a child that runs from the sidewalk through the parking cars onto the road and has enough reaction time or even anticipation to brake. Without any power of recall, the driver may start braking in anticipation when the child is still on the sidewalk, but may stop braking after the child disappears behind a parking car.

- dynamic trajs

In the navigation method of this research, I want to meet the lack of high-level goal planning by introducing powers of recall to navigation. To my knowledge, Kelchtermans and Tuytelaars [**?**] are the only ones who used a recurrent neural network for memory abilities in autonomous, vision-based UAV navigation. [**?**] - reason why their work is not what i want.. However, they only tested their method in simulation and did not comprehensively evaluate their results. If the method, after passing the first gate, could remember that it has seen the second gate before, the method could plan to navigate through the second gate based on elapsed images.

In my method, I plan to use a LSTM-CNN which is a deep convolutional neural network serially connected to a long-short-term-memory (LSTM) neural network.

While the CNN has the ability to perceive and reason spatial structures of the immediate environment, the LSTM is able to establish connections through time.

In other words, the CNN is responsible to predict waypoints or generate trajectories based on single images, whereas the LSTM empower the method to recall and remember by evaluating the temporal structure between the predictions of the CNN.

Besides the above steep curve scenario, this memory would show great benefit in situations when the quadcopter lost track of the goals and it can recall elapsed images. In case of the application in urban areas, for example, quadcopters could remember obstacles that were visible but have become occluded and thus, could better anticipate. Or after an evasive maneuver, the quadcopter could return to the actual path much faster because he memorizes its maneuver. In that sense, the memory of the quadcopter is another form of localization in the environment, which is not global but namely local. In addition, memory could enable better optimization, e.g., the imitation learning of optimal trajectories which are not only spatial but also temporal objects. However, this research, in a simplified scenario, should only prove if powers of recall are applicable and generally useful for the autonomous navigation of MAVs.

3 Task Packages

Based on the objectives, the following task packages are defined:

- Familiarization with the subject area. This includes evaluating relevant papers and taking into account the components and implementations that are already available. The work result is a summary of relevant work.
- Developing an Approach ...
- Analysis, design and drafting of an architecture
- Definition of suitable test scenarios
- Implementation of the architecture, implementation
- Documentation of the architecture and the program code
- ———————————————————
- Shift simulation from Gazebo to the more photo-realistic Flightmare
- Generate training data in simulation
- Shift data input pipeline from tensor flow to pytorch
- Design and drafting of the RCNN in pytorch
- Train the RCNN
- Definition of suitable test scenarios in simulation
- Implementation
- Documentation of the architecture and the program code

4 Time Schedule

Die Bearbeitung dauert maximal 6 Monate. (Am besten eignet sich ein Gannt Diagramm)

5 Organizational Matters

- Language of the thesis: English
- Text processing system: LaTeX
- Programming languages: C++, Python
- Supervisor: Dr. rer. nat. Yuan Xu
- Reviewers: Prof. Dr. Sahin Albayrak, Dr.- Ing. Stefan Fricke

6 Annex

6.1 Steep curve scenario

For example, a section of the racetrack that consist of two successive gates, in between a steep curve could not successfully be navigated through by the method, even if both gates have already appeared on images. Before the MAV navigate through the first gate, both gates are in the FOV of the camera. After it has flown through the first gate, because the curve is too steep, the second gate is out the FOV and the navigation method has no goal to be achieved.

Bibliography

[1] Maria Musterfrau. Ja, auch dieser Bibtex-Eintrag muss gelöscht werden... *ACM Trans. Inf. Syst.*, 12(1):92–115, 1994.

[2] Max Mustermann. Bitte lösche diesen Eintrag und schreibe deinen eigenen Bibtex-einträge hier rein. In *ABC '99: Proceedings of the eighth international conference on ABC*, pages 330–334, New York, NY, USA, 1999. ACM Press.