1 Abstract

With this project, we collect and estimate ground traversability for a no-wheel crocodile-like robot. We move the robot in a simulated environment recording its pose. Later, we crop for each height map used to generate the ground a patch such as it includes booth the robot in the center and its footprint in the case of maximum advancement.

Our approach is based on an already existing methodology that we further expanded with a smaller deep convolutional neural network based on residual connection and the squeeze and excitation operator. The network is unaware of the robot's locomotion and physical characteristics. Then, we evaluate the results by visualizing different datasets and custom patches using GRAD-CAM to highlight and discover the strength and weakness of the model.

2 Introduction

Effective identification of traversable terrain is essential to operate mobile robots in the environment. Today, there two main different approaches used today in the industry to correctly move a robot in a new environment: online and offline. The first one uses local sensors to map new surroundings while the second equip the mobile robot with an already labeled map of the terrain.

In most indoor scenarios, specific hardware such as infrared or LIDAR sensors is used to perform online mapping while the robot is exploring for the first time. This is the case of most vacuum cleaner.

Indoor scenarios share similar features across different places shifting the problem from which ground can be traversable to which obstacle must be avoided. For example, the floor is always flat without any bumps or holes due to is artificial design. Usually, traversability must be estimated on the fly due to the high number of possible obstacles and to the layout of the objects in each room may not be persistent in time.

On the other hand, outdoor scenarios may have less artificial obstacle but they have a homogeneous ground making challenging to estimate where the robot can properly travel. Moreover, a given patch may not be traversable by all direction due to the not uniform features of the terrain. But, a map of the ground can be obtained easily by using third-party services such as google maps or mapped with a flying drone.

These two different scenarios have different challenges. Usually, in control environments such us indoor it is easier to move the robot on the ground but harder to perform obstacle avoidance while in outdoors scenario it is hard to first estimate where the robot is able to move. Furthermore, data gathering may not be straight forward. In indoors terrain, data is collected most of the times by driving the robot in the environment by a human or an artificial controller. While in the outdoors scenario the data is usually gather using simulations since it is faster.

Our approach aims to estimate traversability on uneven ground, mostly outdoors scenarios. Our frameworks have two main phases, we first run different simulations by spawning the robot on custom created maps with different groud configurations, walls, bumps, and slopes. We record the robot pose at each time, then we crop a patch from the height map used in each run such that it included the robot in the center and its footprint for a selected minimum advancement. This value is calculated by observing the advancement of the given robot on flat ground and taking its half. So, each patch contains the current robot position on the map and the future position in the future if the robot will move at the select value. To create the dataset for the classifier, we label each patch using the selected minimum advancement.

The report is organized as follow, the next chapter introduces the related work, Chapter 2 describes our approach, Chapter 3 talk in deep about the implementation details, Chapter 4 shows the results and Chapter 5 discuss clusion and future work The learning and perception of traversability is a fundamental competence for both organisms and autonomous mobile robots since most of their actions depend on their mobility [8]. Visual perception is known to be used in most all animals to correctly estimate if an environment can be traversed or not. Similar, a wide array of autonomous robots adopt local sensors to mimic the visual properties of animals to extract geometric information of the surrounding and plan a safe path through it. Different methodologies have been proposed to collect the data and then learn to correctly navigate the environment. Most of the methodologies rely on supervised learning, where first the data is gathered and then a machine learning algorithm is trained sample to correctly predict the traversability of those samples.

Among the huge numbers of methods proposed, there are two categories based on the input data: geometric and appearance based methods.

Geometric methods aim to detect traversability using geometric properties of surfaces such as distances in space and shapes. Those properties are usually slopes, bumps, and ramps. Since nearly the entire world has been surveyed at 1 m accuracy [6], outdoor robot navigation can benefit from the availability of overhead imagery. For this reason, elevation data has also been used to extract geometric information. Recently, [2], the work we base on, proposed a full pipeline to estimate traversability using only elevation data in the form of height maps.

Elevation data can also be estimated by flying drones. [4] proposed a collaborative search and rescue system in which a flying robot that explores the map and creates an elevation map to guide the ground robot to the goal. They utilize an 'on-the-spot training' using a convolutional neural network to segment the terrain in different traversable classes.

Whereas appearance methods, to a greater extent related to camera images processing and cognitive analyses, have the objective of recognizing colors and patterns not related to the common appearance of terrains, such as grass, rocks or vegetation. Those images can be used to directly estimate the traversability cost.

Historically, the collected data is first preprocessed to extract texture features that are used to fit a classic machine learning classified such us an SVM [8] or Gaussian models [6]. Those techniques rely on texture descriptors, for example, Local Binary Pattern [7], to extract features from the raw data images obtained from local sensors such as cameras. With the rise of deep learning methods in computer vision, deep convolution neural network has been trained directly on the raw RGB images bypassing the need to define characteristic features.

One recent example is [1] where a deep neural network was training on real-world hiking data collected using head-mounted cameras to teach a flying drone to follow a trail in the forest. Geometric and appearance methods can also be used together to train a traversability classifier. [3] proposes the first on-the-spot training method that uses a flying drone to gather the data and train an estimator in less than 60 seconds.

Data can also extract in simulations, where an agent interacts in an artificial environment. Usually, no-wheel legged robot that is able to traverse harder ground, can benefits from data gathering in simulations due to the high availability. [5] proposed a locomotion planning that is learned in a simulation environment.

References

[1] Jérôme Guzzi Alessandro Giusti et al. "A Machine Learning Approach to Visual Perception of Forest Trails". In: (2016).

[2] R. Omar Chavez-Garcia et al. "Learning Ground Traversability from Simulations". In: (2017). DOI: `10.1109/LRA.2018.2801794`. eprint: `arXiv:1709.05368`.

[3] Jeffrey Delmerico et al. "On-the-spot Training" for Terrain Classification in Autonomous Air-Ground Collaborative Teams". In: (2017).

[4] Julia Nitsch Jeffrey Delmerico Elias Mueggler and Davide Scaramuzza. "Active Autonomous Aerial Exploration for Ground Robot Path Planning". In: (2016).

[5] Tobias Klamt and Sven Behnke. "Anytime Hybrid Driving-Stepping Locomotion Planning". In: (2017).

[6] Boris Sofman et al. "Improving Robot Navigation Through Self-Supervised Online Learning". In: (2006).

[7] Matti Pietikainen Timo Ojala and Topi Maenpaa. "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns." In: (2002).

[8] E. Ugur and E. Sahin. "Traversability: A case study for learning and perceiving affordances in robots". In: (2010).