Problems and Solutions of Autonomous Exploration of Unknown Indoor Environments for Micro Aerial Vehicles with Onboard Stereo camera

Mantas Briliauskas, Virginijus Marcinkevičius

Institute of Data Science and Digital Technologies, Vilnius University

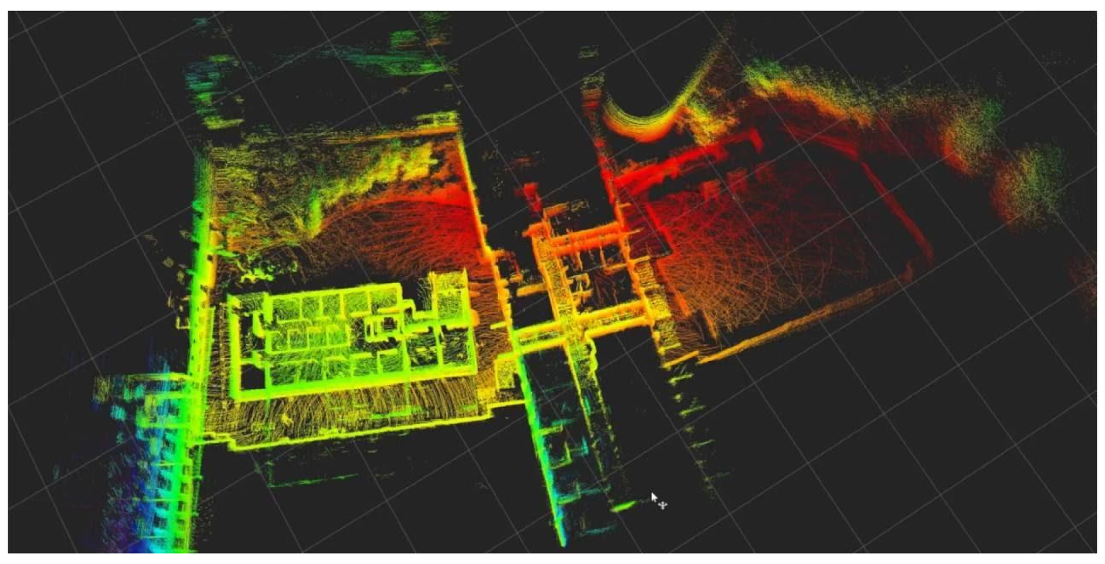
Abstract

The popularity of using micro aerial vehicles (MAV) in autonomous applications has increased significantly over the years. The usage includes agriculture, search and rescue operations, object's inspection, environmental mapping. The main goal of autonomous exploration is to produce a map of a priori unknown environment by minimizing exploration time and map uncertainty. Different SLAM algorithms make a passive localization and mapping by processing sensors' information while MAV is operating, whereas GPS signal is most likely denied or very inaccurate in the indoor environments. Navigation algorithms are employed for path planning with a goal to move MAV to required position as optimally as possible with the respect to its model physics and controls, and avoiding obstacles. Having mapping, localization and navigation, the exploration algorithms are used to find the best next view candidate to navigate MAV to in order to maximize the information gain of currently chosen action. In this survey current state-of-the-art methods and main problems of exploration algorithms for the indoor environments are reviewed.

Introduction

Autonomous exploration or active simultaneous localization and mapping (active SLAM, ASLAM) classically consists of three components: passive SLAM, next target estimation and navigation. Passive SLAM is a maximum a-posteriori (MAP) problem that is solved for belief over current MAV pose and map using KF, EKF, UDF or particle filters given sensor information, odometry or control inputs. The next exploration target is usually obtained by calculating utility for target candidates which are drawn using frontier-based methods. Since exploration-based methods suffer from position drift due to noisy sensors, loop closures are used to mitigate that problem, making active SLAM an exploration-exploitation dilemma [2].

Figure 1: Simultaneous localization and mapping.



Related works

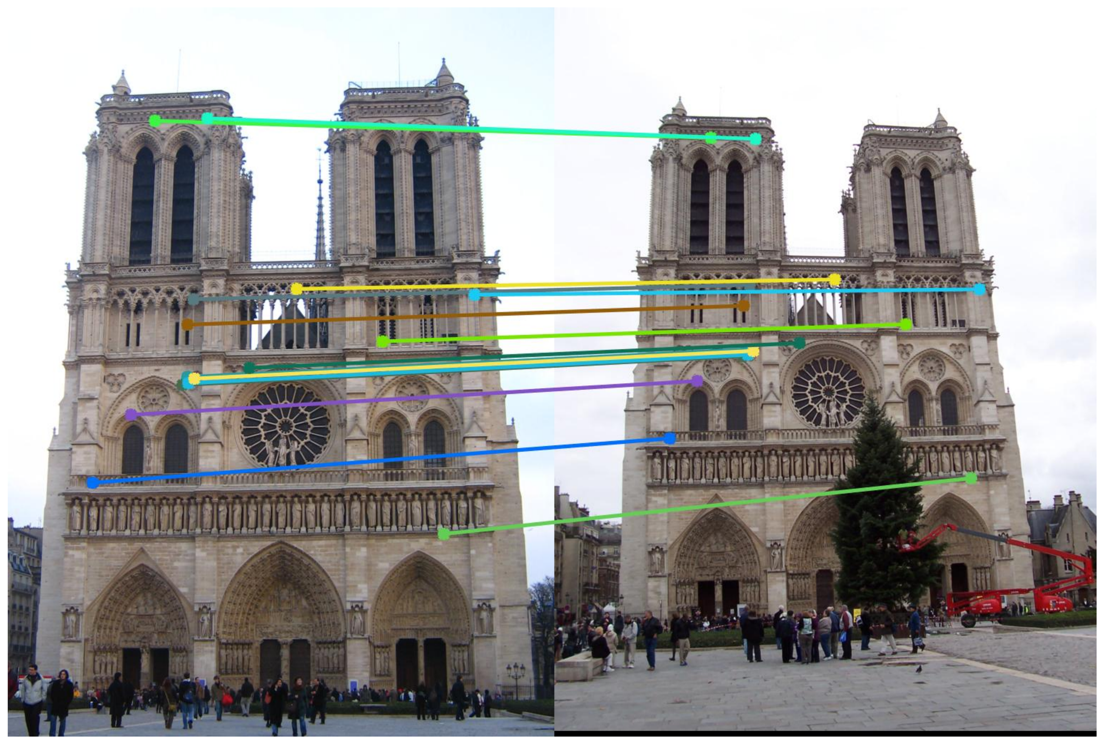
Classical visual-based active SLAM solutions use visual feature extraction from 2D images techniques (e.g. ORB-SLAM3 (2021), RTAB-Map (2019)), and by tracking these features over time (FLANN, RANSAC) try to determine their position in space along with extrinsic parameters (transformation from camera to world frame), given camera intrinsic (focal length, field of view, etc.) parameters.

Figure 2: Classical visual-based SLAM problems.

Passive SLAM problems

Cause	Issue	Solution
Sensor noise	Position drift	Loop closure
		Fusion with controls and velocity
		Adding more sensors (e.g. IMU)
Low texture environments	Lost tracking	Choose feature-rich trajectory
High velocity		Slow down rotations and translations
		Use Bundle-adjustment
		Enable tracking recovery
		Bag of words
Large 3D environment	Memory-expensive representations, slow calculations	Use octree representation (e.g. "octomap")

Figure 3: Feature matching on two images.



NN-based solutions uses DNN+RL on input image and/or RGBD information to predict next actions or directly predict map and robots pose. DNNs are also used to determine next target position [1]. Current state-of-the-art methods on some tasks outperform classic active SLAM methods trained on sufficient experience [3].

Figure 4: Comparison of Active Neural SLAM (ANS) and Occupancy Anticipator (OccAnt) on Habitat simulator [3].

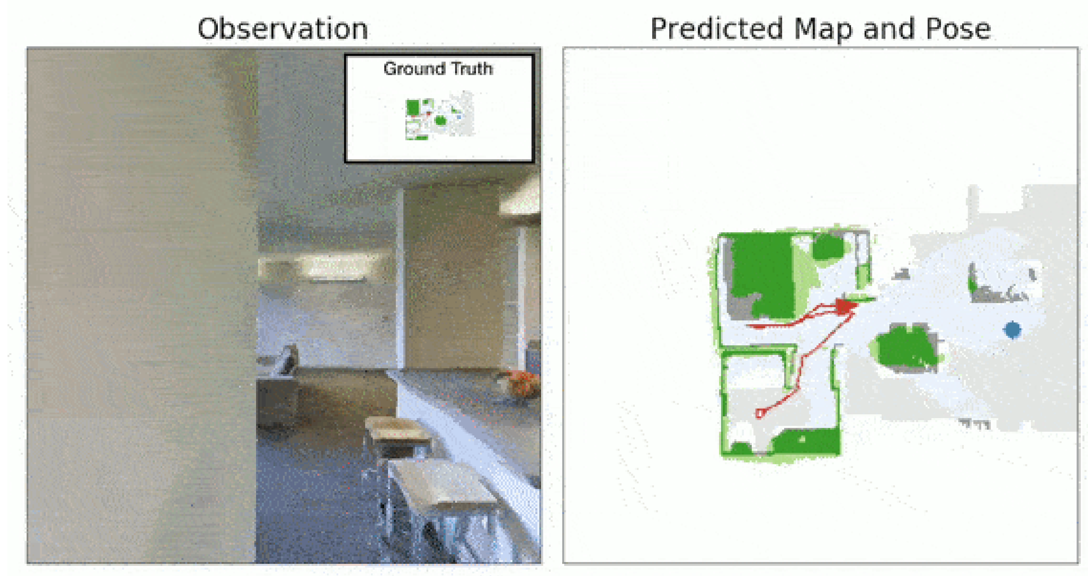
	Noisy test conditions					
	Gibson small		Gibson large		Matterport3D	
Method	Map acc.	IoU	Map acc.	IoU	Map acc.	IoU
ANS(rgb)	18.5	55	35.0	47	44.7	18
ANS(depth)	18.5	56	39.4	53	72.5	26
View-extrap.	12.0	26	28.1	27	39.4	14
OccAnt(rgb) w/o AR	21.8	66	44.2	57	65.8	23
OccAnt(depth) w/o AR	20.2	58	44.2	54	92.7	29
OccAnt(rgbd) w/o AR	16.9	45	35.6	40	76.3	23
OccAnt(rgb)	20.9	62	42.1	54	66.2	22
OccAnt(depth)	**22.7**	**71**	**50.3**	**67**	94.1	**33**
OccAnt(rgbd)	**22.7**	**71**	48.4	62	**99.9**	32
OccAnt(GT)	21.7	67	51.9	63	-	-

In recent years, navigation is added to the NN stack along odometry, making deep RL solutions end-to-end, and showing best results among competitors in navigation to point and navigation to object (search) challenges (Habitat 2021).

Conclusions

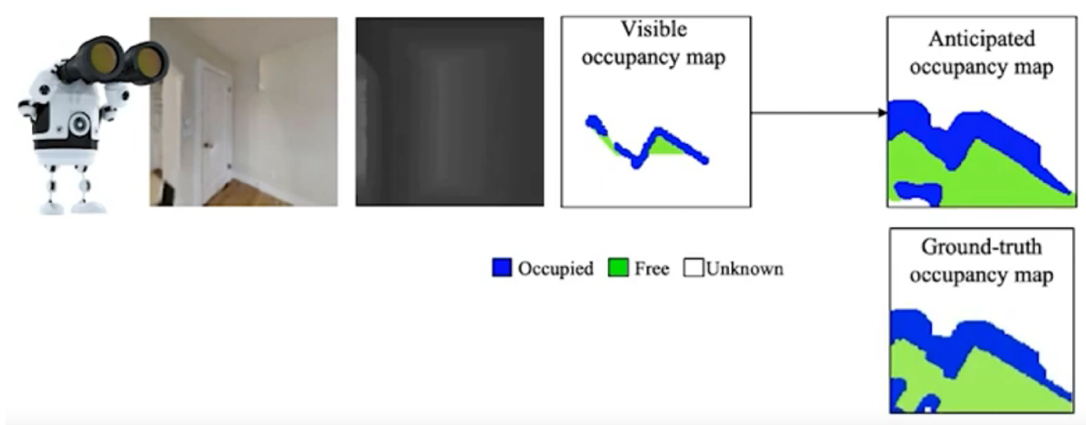
Neural networks are getting popularity in exploration and object search tasks. Navigation is also added to NN stack. Mapping involved into end-to-end makes the solution heavy [1], modular designs enables NN in active SLAM with competition-winning solutions.

Figure 5: Map and robot pose predicted using neural network.



AI-related techniques like occupancy anticipation gives great advantage of NN usage in exploration setting comparing to classical active SLAM methods. Other exploration-related fields like object-search are gaining popularity (Habitat 2019, 2020, 2021) which is not feasible without object segmentation task.

Figure 6: Anticipation of unseen regions based on visual context.



Future work

There are classical active SLAM solutions applied to MAV directly. New RL-based solutions were applied on a ground setting employing a 2.5D map for simplicity. Regardingly, RL-based active SLAM model in 3D environment for MAV could be the next challenge to research.

References

[1] D. S. Chaplot and E. Al, *Learning to explore using active neural slam*, in International Conference on Learning Representations (ICLR), 2020.

[2] J. A. Placed and E. Al, *A survey on active simultaneous localization and mapping: State of the art and new frontiers*, 2022.

[3] S. K. Ramakrishnan and E. Al, *Occupancy anticipation for efficient exploration and navigation*, 2020.