

Autonomous Exploration with a Low-Cost Quadcopter using Semi-Dense Monocular SLAM

Lukas von Stumberg¹, Vladyslav Usenko¹, Jakob Engel¹, Jörg Stückler², and Daniel Cremers¹

Abstract—Micro aerial vehicles (MAVs) are strongly limited in their payload and power capacity. In order to implement autonomous navigation, algorithms are therefore desirable that use sensory equipment that is as small, low-weight, and low-power consuming as possible. In this paper, we propose a method for autonomous MAV navigation and exploration using a low-cost consumer-grade quadcopter equipped with a monocular camera. Our vision-based navigation system builds on LSD-SLAM which estimates the MAV trajectory and a semi-dense reconstruction of the environment in real-time. Since LSD-SLAM only determines depth at high gradient pixels, texture-less areas are not directly observed. We propose an obstacle mapping and exploration approach that takes this property into account. In experiments, we demonstrate our vision-based autonomous navigation and exploration system with a commercially available Parrot Bebop MAV.

I. INTRODUCTION

Most autonomous micro aerial vehicles (MAVs) to-date rely on depth sensing through e.g. laser scanners, RGB-D or stereo cameras. Payload and power capacity are, however, limiting factors for MAVs, such that sensing principles are desirable that require as little size, weight, and power-consumption as possible.

In recent work, we propose large-scale direct simultaneous localization and mapping (LSD-SLAM [1]) with handheld monocular cameras in real-time. This method tracks the motion of the camera towards reference keyframes and at the same time estimates semi-dense depth at high gradient pixels in the keyframe. By this, it avoids strong regularity assumptions such as planarity in textureless areas. In this paper, we demonstrate how this method can be used for obstacle-avoiding autonomous navigation and exploration for a consumer-grade MAV. We integrate our approach on the recently introduced Parrot Bebop MAV, which is available for only approx. 500\$ and comes with a 30 fps high-resolution fisheye video camera and integrated attitude sensing and control.

Our proposed two-step exploration strategy is specifically and directly suited for semi-dense reconstructions as obtained with LSD-SLAM. A simple but effective local exploration strategy, coined *star discovery*, safely discovers free and occupied space in a local surrounding of a specific position in the environment. It specifically takes the depth measurement principle of LSD-SLAM based on motion parallax

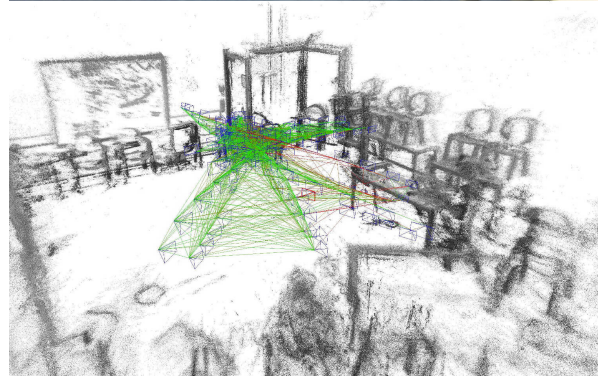


Fig. 1: We propose vision-based autonomous navigation and exploration for small low-cost micro aerial vehicles equipped with monocular cameras. Our approach is based on large-scale direct SLAM which determines a semi-dense reconstruction of the environment. We integrate our approach on the commercially available Parrot Bebop MAV.

into account. A global exploration strategy then determines interesting volume for further local explorations in order to sequentially discover novel parts of the environment. We demonstrate the properties of our exploration strategy in several experiments with the Parrot Bebop.

II. RELATED WORK

Autonomous exploration by mobile robots has been investigated over many years, mainly relying on laser scanner sensors. Yamauchi [2] proposed in his seminal work the so-called frontier-based exploration strategy that favors exploring the frontiers of the unexplored space in the map. Some methods define a utility function [3], [4], e.g., on paths or view poses that, for instance, trade-off discovered area with travel costs. The approaches in [5] combines the probabilistic measure of information gain with travel cost in a measure of utility. For exploration in 3D, Joho et al. [6] proposed an approach that measures information gain in

¹The authors are with the Computer Vision Group, Computer Science Institute 9, Technische Universität München, 85748 Garching, Germany {stumberg, usenko, engelj, cremers}@in.tum.de

²The author is with the Computer Vision Group, Visual Computing Institute, RWTH Aachen University, 52074 Aachen, Germany stueckler@vision.rwth-aachen.de

multi-level surface maps. In [7], such an approach is made in a Rao-Blackwellized particle filter framework for SLAM with 2D laser range finders. The proposed method takes measurements and travel costs along exploration paths into account. Rekleitis et al. [8] optimize a utility function that favors the reduction of uncertainty in the map, and at the same time tries to achieve a fast exploration of the map.

Very recently, autonomous exploration of the environment with flying robots has attracted attention. Nuske et al. [9] explore rivers using an MAV equipped with a continuously rotating 3D laser scanner. They propose a multi-criteria exploration strategy to select goal points and traversal paths. Heng et al. [10] propose a two-step approach to visual exploration with MAVs using depth cameras. Efficient exploration is achieved through maximizing information gain in a 3D occupancy map. At the same time, high coverage of the viewed surfaces is determined along the path to the exploration pose. In order to avoid building up a dense 3D map of the environment and applying standard exploration methods, Shen et al. [11] propose a particle-based frontier method that represents known and unknown space through samples. This approach also relies on depth sensing through a 2D laser scanner and a depth camera. Yoder and Scherer [12] explore the frontiers of surfaces measured with a 3D laser scanner. Desaraju et al. [13] use a monocular camera and a dense motion stereo approach to find suitable landing sites.

We propose an exploration method which is suitable for light-weight, low-cost monocular cameras. Our visual navigation method is based on large-scale direct SLAM which recovers semi-dense reconstructions. We take special care of the semi-dense information and its measurement process for obstacle mapping and exploration.

III. AUTONOMOUS QUADROCOPTER NAVIGATION USING MONOCULAR LSD-SLAM

We built on the TUM ARDrone package developed by Engel et al. [14] which has been originally developed for the Parrot ARDrone 2.0. We transferred the software to the new Parrot Bebop platform which comes with similar sensory equipment and onboard control.

A. Hardware Platform

The Parrot Bebop is equipped with an IMU built from 3-axis magnetometer, gyroscope, and accelerometer. It measures height using an ultrasonic sensor, an air pressure sensor and a vertical camera, similar to the Parrot ARDrone 2.0. The MAV is equipped with a fisheye camera with wide 186° field-of-view. The camera provides images at 30 frames per second. A horizontally stabilized region-of-interest is automatically extracted in software on the main processing unit of the MAV, and can be transmitted via wireless communication together with attitude measurements.

B. State Estimation and Control

The visual navigation system proposed in [14] integrates visual motion estimates from a monocular SLAM system with the attitude measurements from the MAV. It filters both

kinds of messages using a loosely-coupled Extended Kalman filtering (EKF) approach. Since the attitude measurements and control commands are transmitted via wireless communication, they are affected by a time delay that needs to be compensated using the EKF framework. Waypoint control of the MAV is achieved using PID control based on the EKF state estimate.

In monocular SLAM, the metric scale of motion and reconstruction estimates are not observable, so the scale must be estimated from different sensor modalities. On the MAV, this is achieved by probabilistically fusing ultrasonic and air pressure measurements and adapting the scale of the SLAM motion estimate to the observed metric scale.

C. Vision-Based Navigation Using Monocular LSD-SLAM

LSD-SLAM [1] is a keyframe based SLAM approach. It maintains and optimizes the view poses of a subset of images, i.e. keyframes, extracted along the camera trajectory. In order to estimate the camera trajectory, it tracks camera motion towards a reference keyframe. If the camera moved too far from its reference keyframe, a new keyframe is generated which becomes a new reference keyframe. For motion tracking, the current image is aligned with the reference keyframe using direct image alignment. This requires depth in either of the images, which we estimate from stereo correspondences between the two images within the reference keyframe. We estimate depth in the current reference keyframe from temporal stereo correspondences based on the tracked motion. The poses of the keyframes are made globally consistent by mutual direct image alignment and pose graph optimization.

A key feature of LSD-SLAM is the ability to close trajectory loops within the keyframe graph. In such an event, the view poses of the keyframes are readjusted to compensate for the drift that is accumulated through tracking along the loop. This especially changes the pose of the current reference keyframe that is used for tracking, also inducing a change in the tracked motion estimate.

Note that the tracked motion estimate is used to update the EKF that estimates the MAV state which is fed into the control loop. At a loop closure, this visual motion estimate would update the filter with large erroneous velocities which would induce significant errors in the state estimate. In turn this could cause severe failures in flight control. We therefore compensate for the changes induced by loop-closures with an additional pose offset T_{offset} on the visual motion estimate $T_{\text{cam}}^{\text{world}}$ before feeding it into the EKF, i.e. we use the modified estimate $T_{\text{offset}} T_{\text{cam}}^{\text{world}}$. The visual motion estimate is determined from the tracked motion of the camera towards the reference keyframe $T_{\text{cam}}^{\text{ref}}$ and the optimized pose $T_{\text{ref}}^{\text{world}}$ of the keyframe, i.e. $T_{\text{cam}}^{\text{world}} = T_{\text{ref}}^{\text{world}} T_{\text{cam}}^{\text{ref}}$. Before and after the loop closure, the visual motion estimate should remain the same, inducing an update of the offset to

$$\tilde{T}_{\text{offset}} = T_{\text{offset}} T_{\text{ref}}^{\text{world}} \left(\tilde{T}_{\text{ref}}^{\text{world}} \right)^{-1}, \quad (1)$$

where $T_{\text{ref}}^{\text{world}}$ and $\tilde{T}_{\text{ref}}^{\text{world}}$ are the pose estimates for the keyframe before and after the loop closure, respectively.

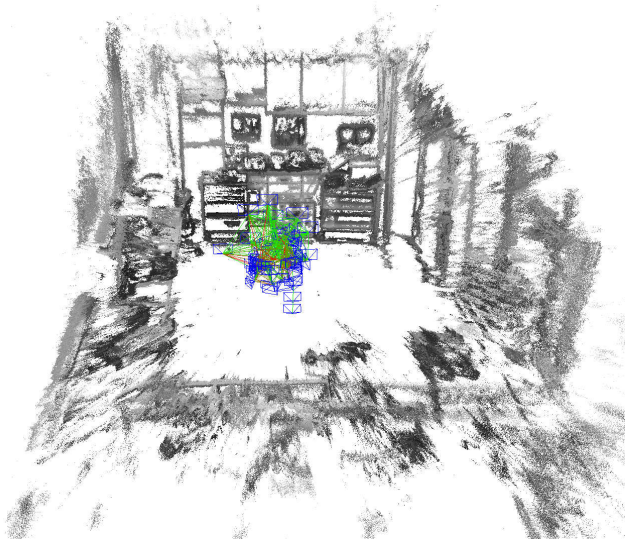


Fig. 2: Example of the semi-dense reconstruction by LSD-SLAM from an initial look-around maneuver.

In order to initialize the system, the MAV performs a simple look-around maneuver in the beginning by flying a 360° turn on the spot while hovering up and down by several centimeters. In this way, the MAV already obtains an initial keyframe map with a closed trajectory loop (s. Fig. 2).

IV. AUTONOMOUS OBSTACLE-FREE EXPLORATION WITH SEMI-DENSE DEPTH MAPS

Autonomous exploration has been a research topic for many years targeting exploration of both 2D and 3D environments. In most 3D scenarios an exploration strategy works with a volumetric representation of the environment, such as a voxel grid or an octree, and uses laser-scanners or RGB-D cameras as sensor to build such a representation.

In this paper we devise an exploration strategy that builds on a fundamentally different type of sensor data – semi-dense depth maps estimated with a single moving monocular camera. The difference to previously mentioned sensors lies in the fact that only for the image areas with strong gradients the depth can be estimated. This means that especially initially during exploration, large portions of the map will remain unknown. The exploration strategy has to account for the motion parallax measurement principle of LSD-SLAM.

This section first discusses the results of volumetric fusion of semi-dense depth maps and discusses its properties. We propose a local exploration strategy that reduces the number of unknown voxels while navigating in the parts of the environments that are guaranteed to be free. Finally, we use this local strategy within a global exploration strategy that allows for exploring larger environments with a sensor that provides semi-dense depth maps.

Exploration is started using the initializing scheme for LSD-SLAM as described in Sec. III. We assume that the MAV is initialized with sufficiently large obstacle-free space around it. After the completed look-around maneuver, we in-

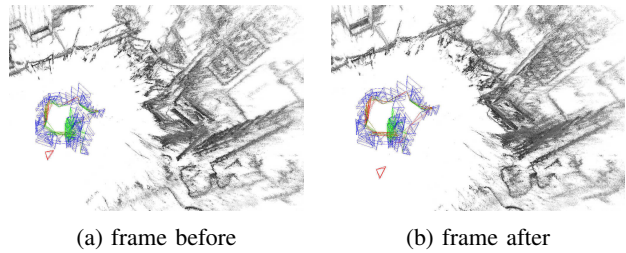


Fig. 3: Example of a LSD-SLAM loop closure.

tegrate all obtained semi-dense depth maps in the keyframes from their optimized poses into an obstacle map.

A. Occupancy Mapping with Semi-Dense Depth Maps

In this work we use OctoMap [15] that provides an efficient implementation of hierarchical 3D occupancy mapping in octrees. We directly use the semi-dense depth maps reconstructed with LSD-SLAM to create the 3D occupancy map. All keyframes are traversed and the measured depths are integrated via ray-casting using the camera projection model.

Since LSD-SLAM performs loop closures, keyframe poses will change from their estimate of when they have been integrated into the occupancy map. Hence, the map will become outdated. Thus, we periodically regenerate the map from scratch using the updated keyframe poses, since there is no efficient way to correct for the changes directly in the map. This operation may last for several seconds, but the MAV can hover on the spot and wait until proceeding with the exploration.

An alternative approach would be to represent the obstacle map in multiple local occupancy maps that move locally with one or several keyframes. This approach, however, would require a strategy to decide which keyframes should be subsumed in a local map, and the run-time of the obstacle look-up would scale with the number of the local maps. In this work, for simplicity, we decide for using a single map with acceptably short exploration breaks.

Each voxel in the occupancy map stores the probability of being occupied in log-odds form. In order to determine if a voxel is free or occupied, a threshold is applied on the occupancy probability (0.86 in our experiments). During the integration of a depth measurement, all voxels along the ray in front of the measurement are updated with a probability value for missing the voxel. The voxel at the depth measurement in turn is updated with a hit probability value. Note that LSD-SLAM also outputs the variance of each depth estimation. Although measurements with a high variance can be very noisy, they still contain information about the vicinity of the sensor. Therefore we insert free space at a reduced distance for these pixels and do not perform a hit update. Fig. 9 shows an occupancy map obtained using LSD-SLAM.

By using semi-dense reconstructions, we do not make strong assumptions such as planarity about the properties of the environment in textureless areas. On the other hand, the

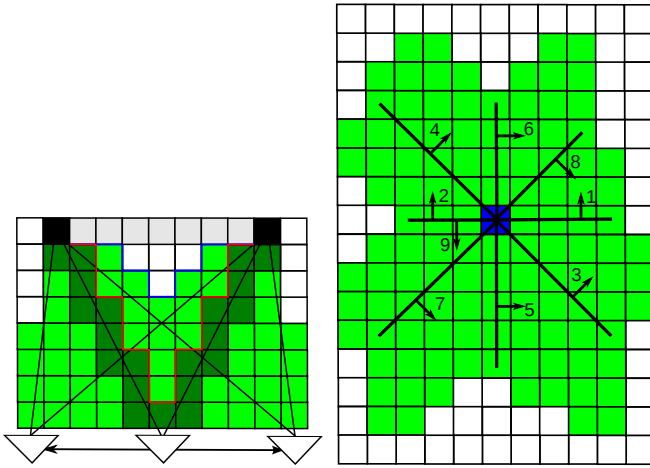


Fig. 4: Left: With LSD-SLAM, only semi-dense depth measurements at edges and high-texture areas can be obtained. It provides no depth in textureless areas such as walls. This creates indentations of unknown space in the occupancy map (red line). Lateral motion towards the measurable points, however, allows for reducing the indentations. Right: Local star discovery exploration strategy.

use of semi-dense reconstruction in visual navigation leads to indentations of unknown volume (s. Fig. 4, left). Importantly, these indentations can be removed through lateral motion towards the measurable structures – an important property that we will exploit in our exploration strategy.

B. Obstacle-Free Local Exploration through Star Discoveries

One possible approach for an exploration strategy would be a variant of next best view planning which measures the discovered unknown voxels on the path. This procedure would be computationally very expensive, since for each potential view poses many ray-casting operations would need to be performed. We propose a simpler but effective local exploration strategy that we call *star discovery*, which discovers the indentations in the unknown volume around a specific position in the map.

In star discovery, the MAV flies a star-shape pattern (s. Fig. 4, right). In order to generate motion parallax for LSD-SLAM and to discover indentations in the unknown volume, the MAV flies with a 90° heading towards the motion direction. Clearly, the MAV can only fly as far as the occupancy map already contains explored free-space.

The star-shape pattern is generated as follows: We cast m rays from a specific position in the map at a predefined angular interval in order to determine the farthest possible flight position along the ray. The traversability of a voxel is determined by inflation of the occupied voxels by the size of the MAV. In order to increase the success of the discovery, we perform this computation at l different heights and choose the result of maximum size. With n being the maximum edge length of the bounding box in voxels, the runtime needed for calculating the star discovery is in $O(n \cdot m \cdot l + n^3)$. First the

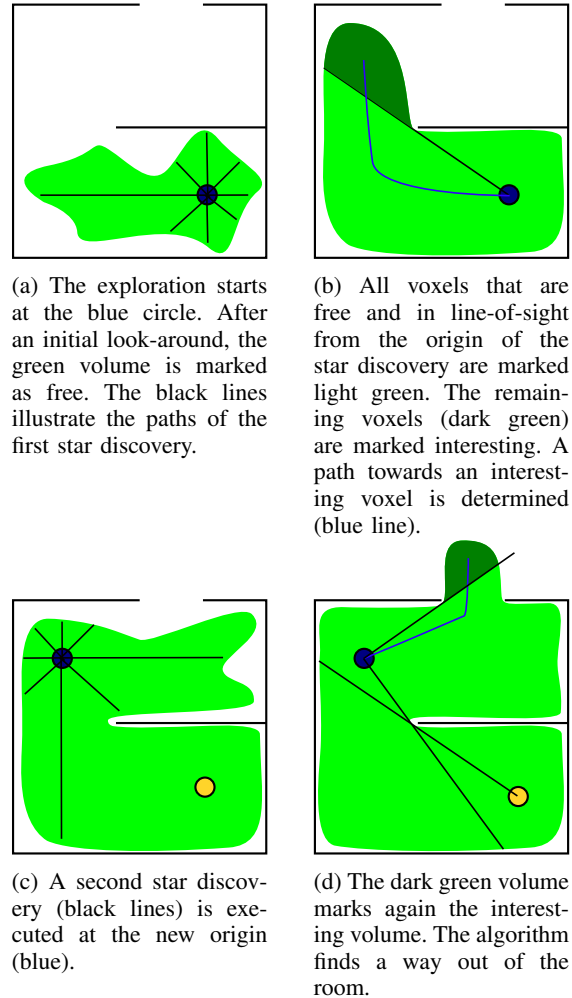


Fig. 5: Four steps of the proposed exploration strategy.

inflated map is created which takes $O(9n^3) = O(n^3)$. The process requires $m \cdot l$ raycasts, each taking $O(n)$ time.

One important property of this local exploration strategy is that it allows for precomputing several exploration moves in advance. Only if the star discovery is fully executed, we redetermine the occupancy map from the updated LSD-SLAM keyframe map. This also enables to postpone loop-closure updates towards the end of the exploration process, and provides a full 360° view from the center position of the star discovery.

Another simple strategy one might think of is flying an outward facing ellipse of maximum size. This however will result in LSD-SLAM losing tracking in most cases because the drone will only see few or no gradients when it flies close to an obstacle while facing it. When performing a star discovery the drone is in most cases able to see a much bigger part of its surroundings and is therefore less likely to lose tracking.

C. Global Exploration

Obviously, a single star discovery from one spot is not sufficient to explore arbitrarily shaped environments, as only

positions on the direct line-of-sight from the origin can be reached (s. Fig. 5). This induces a natural definition of interesting origins for subsequent star discoveries. We denote a voxel *interesting* if it is known to be free, but it is not in line-of-sight of any origin previously used for a star discovery.

We determine the interesting voxels for starting a new star discovery as follows: For every previously visited origin of a star discovery, we mark all free voxels in direct line-of-sight as visited. Then all free voxels in the inflated map are traversed and the ones that have not been marked are set to interesting. With m being the number of star discovery origins, the whole algorithm runs in $O(n^3 \cdot (m + hor^2 \cdot ver))$, where hor and ver are the number of voxels inflated in the horizontal and vertical directions. We bound n by the number of voxels along the longest direction of the bounding box of the occupancy map.

Afterwards, we use Dijkstra’s algorithm in the occupancy map to reach one of the interesting voxels. We look at several random voxels within the largest connected component of interesting voxels and choose the one from which we can execute the largest star discovery afterwards. Running Dijkstra’s algorithm takes $O(n^3 \log(n^3) + 27n^3)$ as there are n^3 nodes and less than $27n^3$ edges. This is equal to $O(n^3 \cdot 3 \log(n)) = O(n^3 \log(n))$.

V. RESULTS

We evaluate our approach on a Parrot Bebop MAV in two differently sized and furnished rooms (a lab and a seminar room). We transmit the live image stream of the horizontally stabilized images of the Bebop to a control station PC via wireless communication. The images are then processed on the control station PC to implement vision-based navigation and exploration based on LSD-SLAM.

All experiments were executed completely autonomously.

We recommend viewing the accompanying video of the experiments at <https://youtu.be/fWBsDwBJD-g>.

A. Qualitative Evaluation of the Exploration in Real-World Environments

1) *Star Discovery*: In a first experiment, we demonstrate autonomous star discovery in our lab room. There was no manual interaction except triggering the discovery and the landing at the end. At first, the MAV performs a look-around maneuver. In Fig. 6a one can see the semi-dense reconstruction of the room obtained with LSD-SLAM at this point. Fig. 7a shows the corresponding 3D occupancy map. Although the look-around already discovered a lot of free voxels, from Figs. 8a and 8c indentations in the unknown volume can be observed which are typical to the occupancy map after just a look-around. Based on this occupancy map, a star discovery is planned (s. Fig. 9). Figure 10 depicts the free voxels in the map inflated by the size of the Bebop. In this case, we used three voxels in horizontal direction and one voxel in vertical direction to inflate the map.

Fig. 11 shows the planned waypoints of the star discovery overlaid with the actual trajectory estimate obtained with

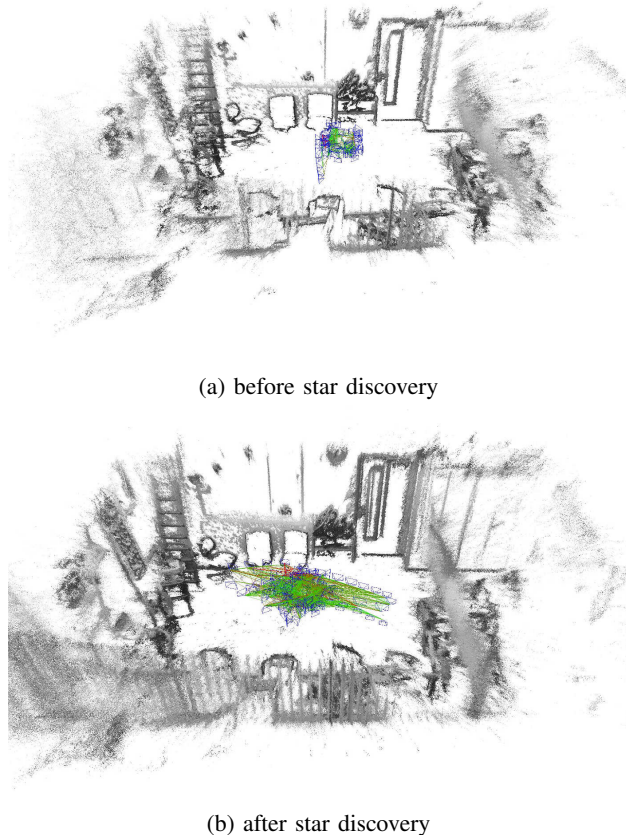


Fig. 6: Semi-dense reconstruction before and after the star discovery in the first experiment.

LSD-SLAM. As can be seen from Fig. 6b the reconstruction becomes much denser through the star discovery.

2) *Full Exploration Strategy*: In a second experiment, we demonstrate a star discovery with subsequent repositioning at an interesting voxel in a larger seminar room.

First, the MAV took off, initialized the scale, and performed a look-around maneuver. The resulting occupancy maps with free and traversable voxels are shown in Fig. 12. Afterwards, the MAV executed a star discovery. Fig. 13 shows the planned discovery motion and the flown trajectory estimated with LSD-SLAM. We explain the differences by LSD-SLAM pose graph updates.

After the star discovery, we obtain the maps and interesting voxels in Fig. 14. The largest connected component found by our algorithm is the one outside the room. The MAV planned a path towards it and autonomously executed it. In Fig. 15 we depicted the planned path and the actually flown trajectory estimated with LSD-SLAM.

After reaching the interesting point the battery of the drone was empty and it landed automatically. The step that our algorithm would have performed next is the star discovery depicted in Fig. 16.

B. Quantitative Evaluation

Table I gives results on the run-time of various parts of our approach and properties of the LSD-SLAM and occupancy

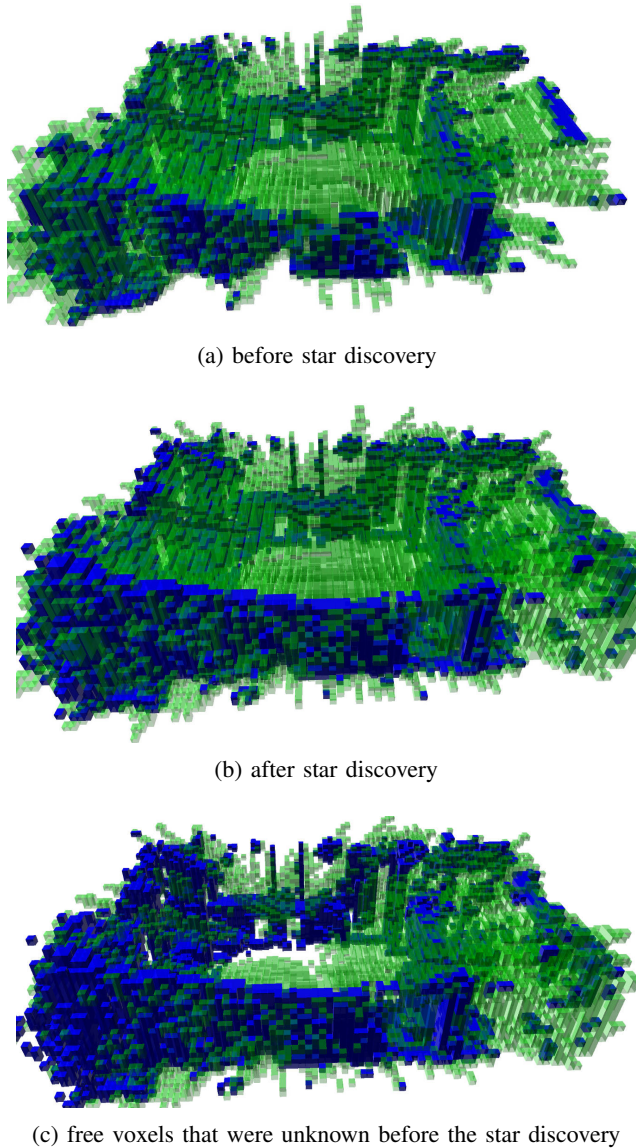


Fig. 7: 3D occupancy map before and after the star discovery in the first experiment as well as the difference between them. Occupied voxels are shown blue, free voxels green.

mapping processes for the two experiments. The creation of the occupancy map is visibly the most time-consuming part of our method, especially at later time steps when the semi-dense depth reconstruction becomes large. In the second experiment modified parameters were used for the creation of the occupancy map and for marking interesting points. While they proved to perform better for finding ways to a new center point, they also further increased the time consumption. The remaining parts are comparatively time efficient and can be performed in a couple of seconds. Our evaluation also shows that star discoveries significantly increase the number of free voxels in the map.

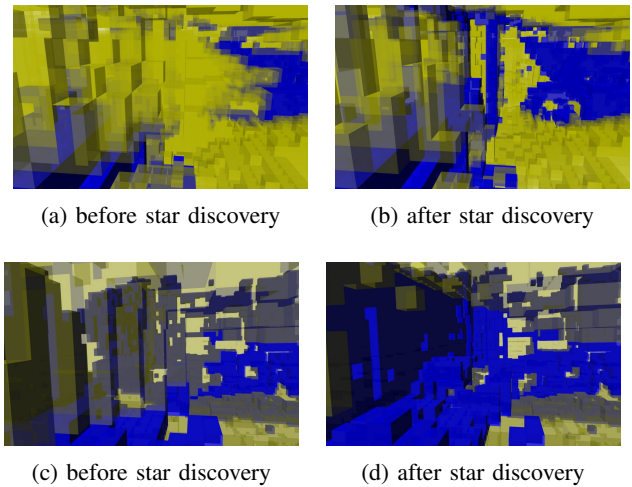


Fig. 8: Indentations in unknown volume before and after the star discovery in the first experiment. Occupied voxels are blue and unknown voxels are yellow.

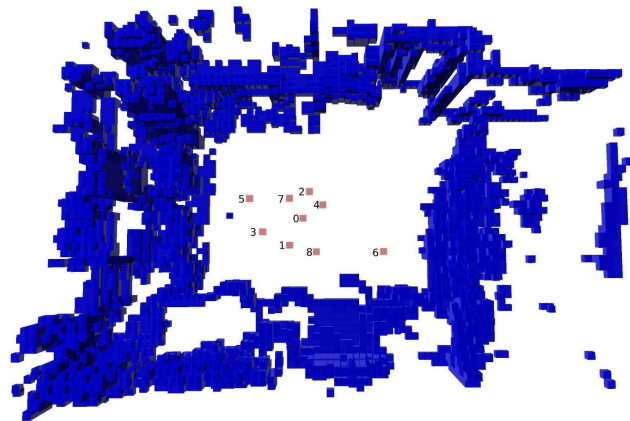


Fig. 9: Exploration plan of the star discovery in the first experiment. Occupied voxels are blue, approached voxels are red and numbered according to their approach order (0: origin).

VI. CONCLUSIONS

In this paper, we proposed a novel approach to vision-based navigation and exploration with MAVs. Our method only requires a monocular camera, which enables low-cost, light-weight, and low-power consuming hardware solutions. We track the motion of the camera and obtain a semi-dense reconstruction of the environment in real-time using LSD-SLAM. Based on these estimates, we build 3D occupancy maps which we use for planning obstacle-free exploration maneuvers.

Our exploration strategy is a two-step process. On a local scale, star discoveries find free-space in the local surrounding of a specific position in the map. A global exploration strategy determines interesting voxels in the reachable free-space that is not in direct line-of-sight from previous star discovery origins. In experiments, we demonstrate the performance of LSD-SLAM for vision-based navigation of an

TABLE I: Quantitative results on run-time and occupancy map statistics for the two experiments.

experiment	1		2		
	look-around	star discovery	look-around	star discovery	new origin
occupancy map	5.65	13.06	4.25	38.43	41.09
inflating map	0.69	0.76	1.40	2.15	2.86
mark voxels in sight	0.19	0.20	1.42	4.93	4.52/6.63
way to new origin	0.0065	0.0087	0.016	0.024	0.065
#voxels in bounding box	195048	211968	449565	728416	1312492
#free voxels	36071	46021	75113	106944	159294
#num occupied voxels	8259	11477	6102	9673	10816
#free ÷ #known	0.81	0.80	0.92	0.92	0.94
#free ÷ #bounding box	0.18	0.22	0.17	0.15	0.12
#keyframes (approx.)	66	162	54	236	257
total #points	15411528	37828296	3152358	13776972	15002889

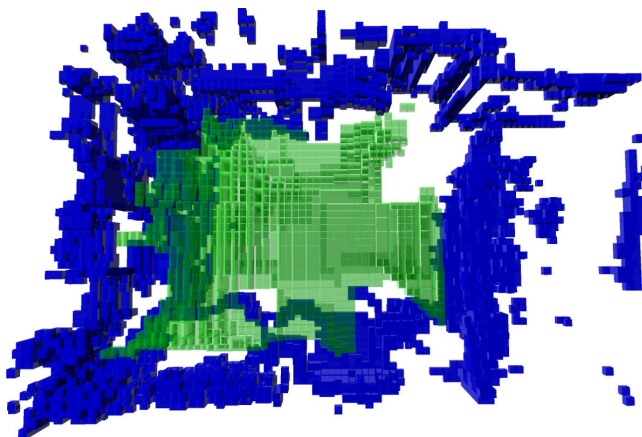


Fig. 10: Inflated free-space map before the star discovery in the first experiment. Free voxels are depicted in green.

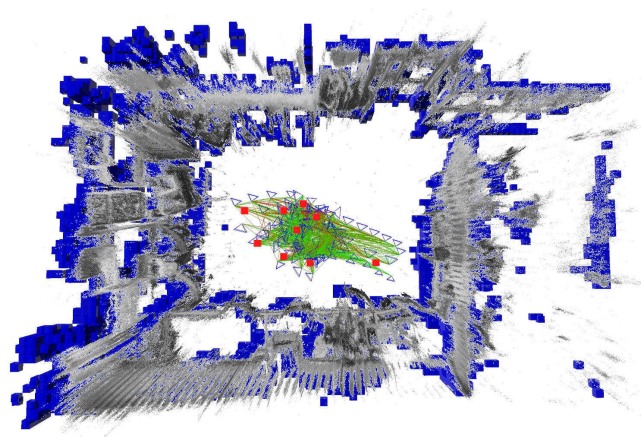


Fig. 11: 3D occupancy map manually overlaid with semi-dense reconstruction, planned waypoints of the star discovery and executed trajectory in the first experiment. Planned waypoints in red, occupied voxels in blue, actual trajectory as estimated by LSD-SLAM in green.

MAV. We give qualitative insights and quantitative results on the effectiveness of our exploration strategy.

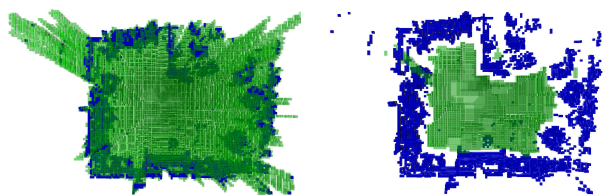
The success of our vision-based navigation and exploration method clearly depends on the robustness of the visual tracking. If the MAV moves very fast into regions where it observes mostly textureless regions, tracking can become difficult. A tight integration with IMU information could benefit tracking, however, such a method is not possible with the current wireless transmission protocol for visual and IMU data on the Bebop.

Also a more general path planning algorithm based on the next best view approach is desirable. This however requires a more efficient way to refresh the occupancy map when pose graph updates happen.

In future work we will extend our method to Stereo LSD-SLAM [16] and tight integration with IMUs. We may also use the method for autonomous exploration on a larger MAV with onboard processing.

REFERENCES

[1] J. Engel, T. Schöps, and D. Cremers, “LSD-SLAM: Large-scale direct monocular SLAM,” in *European Conference on Computer Vision (ECCV)*, 2014.



(a) Occupancy map with free voxels (green). (b) Occupancy map with traversable voxels (green).

Fig. 12: 3D occupancy maps after the look-around maneuver in the second experiment. Occupied voxels in blue.

[2] B. Yamauchi, “A frontier-based approach for autonomous exploration,” in *IEEE Int. Symp. on Computational Intelligence in Robotics and Automation*, 1997.

[3] H. H. Gonzalez-Banos and J.-C. Latombe, “Navigation strategies for exploring indoor environments,” *International Journal of Robotics Research*, vol. 21, no. 10-11, pp. 829–848, 2002.

[4] N. Basilico and F. Amigoni, “Exploration strategies based on multi-criteria decision making for searching environments in rescue operations,” *Autonomous Robots*, vol. 31, no. 4, pp. 401–417, 2011.

[5] W. Burgard, M. Moors, C. Stachniss, and F. Schneider, “Coordinated

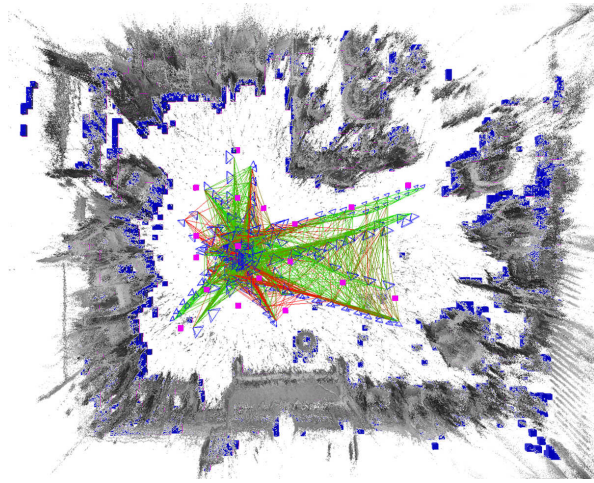


Fig. 13: Planned waypoints of the star discovery (pink voxels) and the flown trajectory estimated with LSD-SLAM (green) in the second experiment.

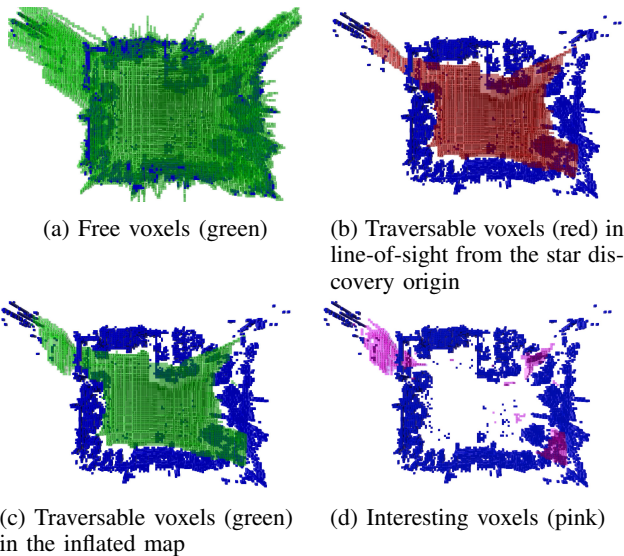
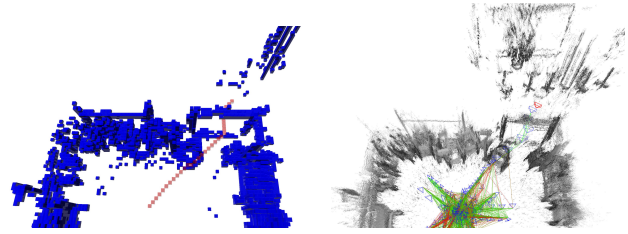


Fig. 14: Occupied (blue), free, traversable, and interesting voxels in the second experiment.



(a) Planned path (red) to the found next star discovery origin. (b) Flown trajectory (blue) to the interesting point estimated with LSD-SLAM.

Fig. 15: Planned and executed path in the second experiment.



Fig. 16: Plan (red voxels) for the second star discovery in the second experiment.

multi-robot exploration.” *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 376–386, 2005.

[6] D. Joho, C. Stachniss, P. Pfaff, and W. Burgard, “Autonomous exploration for 3D map learning,” in *Autonome Mobile Systeme (AMS)*, 2007.

[7] C. Stachniss, G. Grisetti, and W. Burgard, “Information gain-based exploration using rao-blackwellized particle filters,” in *Proc. of Robotics: Science and Systems (RSS)*, Cambridge, MA, USA, 2005.

[8] I. M. Rekleitis, “Single robot exploration: Simultaneous localization and uncertainty reduction on maps (slurm),” in *CRV*, 2012.

[9] S. Nuske, S. Choudhury, S. Jain, A. Chambers, L. Yoder, S. Scherer, L. Chamberlain, H. Cover, and S. Singh, “Autonomous exploration and motion planning for a uav navigating rivers,” *Journal of Field Robotics*, 2015.

[10] L. Heng, A. Gotovos, A. Krause, and M. Pollefeys, “Efficient visual exploration and coverage with a micro aerial vehicle in unknown environments,” in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 1071–1078.

[11] S. Shen, N. Michael, and V. Kumar, “Autonomous indoor 3D exploration with a micro-aerial vehicle,” in *IEEE International Conference*

on Robotics and Automation (ICRA), 2012, pp. 9–15.

[12] L. Yoder and S. Scherer, “Autonomous exploration for infrastructure modeling with a micro aerial vehicle,” in *Field and Service Robotics*, 2015.

[13] V. R. Desaraju, N. Michael, M. Humenberger, R. Brockers, S. Weiss, J. Nash, and L. Matthies, “Vision-based landing site evaluation and informed optimal trajectory generation toward autonomous rooftop landing,” *Autonomous Robots*, vol. 39, no. 3, pp. 445–463, 2015. [Online]. Available: <http://dx.doi.org/10.1007/s10514-015-9456-x>

[14] J. Engel, J. Sturm, and D. Cremers, “Scale-aware navigation of a low-cost quadcopter with a monocular camera,” *Robotics and Autonomous Systems (RAS)*, vol. 62, no. 11, pp. 1646–1656, 2014.

[15] A. Hornung, K. Wurm, M. Bennewitz, C. Stachniss, and W. Burgard, “OctoMap: an efficient probabilistic 3d mapping framework based on octrees,” *Autonomous Robots*, vol. 34, no. 3, pp. 189–206, 2013.

[16] J. Engel, J. Stückler, and D. Cremers, “Large-scale direct SLAM with stereo cameras,” in *Int. Conf. on Intell. Robots and Systems (IROS)*, 2015.