Photogrammetrie Fernerkundung Geoinformation, volume 5, pp. 451-464, 2014.

PFG 2014 / X, 0323 – 0332 **Article**
Stuttgart, XXX 2014

Omnidirectional Perception for Lightweight MAVs using a Continuously Rotating 3D Laser Scanner

DAVID DROESCHEL, DIRK HOLZ, AND SVEN BEHNKE

Keywords: MAVs, 3D Laser Scanner, Scan Registration, Obstacle Detection

Summary: Micro aerial vehicles (MAV) are restricted in their size and weight, making the design of sensory systems for these vehicles challenging. We designed a small and lightweight continuously rotating 3D laser scanner—allowing for environment perception in a range of 30 m in almost all directions. This sensor is well suited for applications such as 3D obstacle detection, 6D motion estimation, localisation, and mapping.

Reliably perceiving obstacles in the surroundings of the MAV is a prerequisite for fully autonomous flight in complex environments. Due to varying shape and reflectance properties of objects, not all obstacles are perceived in every 3D laser scan (one half rotation of the scanner). Especially farther away from the MAV, multiple scans may be necessary in order to adequately detect an obstacle. In order to increase the probability of detecting obstacles, we aggregate acquired scans over short periods of time in an egocentric grid-based map. We register acquired scans against this local map to estimate the motion of our MAV and to consistently update the map.

In experiments, we show that our approaches to pose estimation and laser scan matching allow for reliable aggregation of 3D scans over short periods of time, sufficiently accurate to improve detection probability without causing inaccuracies in the estimation of the position of detected obstacles. Furthermore, we assess the probability of detecting different types of obstacles in varying distances from the MAV.

Zusammenfassung: *Omnidirektionale Wahrnehmung für leichte MAVs mittels eines kontinuierlich rotierenden 3D-Laserscanners*

Dieser Artikel beschreibt einen kleinen und leichten kontinuierlich rotierenden 3D-Laserscanner, der für die dreidimensionale Wahrnehmung von Hindernissen auf einem Micro Aerial Vehicle (MAV) entwickelt wurde. Der Sensor ermöglicht eine nahezu omnidirektionale Wahrnehmung der Umgebung in einer Entfernung von bis zu 30 Metern.

Eine Voraussetzung für die vollständige Autonomie von MAVs in komplexen Umgebungen ist die zuverlässige Wahrnehmung von Hindernissen. Durch unterschiedliche Formen und optische Eigenschaften, wie Reflektanz, werden nicht alle Hindernisse in jedem Laserscan wahrgenommen. Speziell bei weiter entfernten Objekten sind mehrere Scans notwendig um ein Hindernis zu detektieren.

Um die Detektionswahrscheinlichkeit zu erhöhen, werden Entfernungsmessungen des Sensors über kurze Zeitfenster in einer egozentrischen 3D Gitterkarte aggregiert. Neue 3D-Scans werden gegen diese Karte registriert um die Eigenbewegung des Fluggeräts zu schätzen.

Experimente zeigen, dass der Ansatz zur Positionsbestimmung eine robuste Aggregation von 3D-Scans über kurze Zeitfenster ermöglicht. Dadurch wird die Detektionswahrscheinlichkeit erhöht, ohne Ungenauigkeiten in der Position erkannter Hindernissen zu verursachen.

1 Introduction

In recent years, lightweight micro aerial vehicles (MAV) such as quadrotors attracted much attention in the field of aerial robotics because of their relatively low cost, ease of control, and compatibility with everyday indoor and outdoor environments. The size and weight limitations of such platforms, however, pose a problem for the design of their sensory systems. Most of today's lightweight MAVs are equipped with ultrasound distance sensors and air pressure sensors for estimating height above the ground, inertial sensors for estimating attitude, magnetometers for estimating the heading direction, GPS for absolute position estimates (outdoors), and cameras, e.g. for estimating visual odometry. While these small

 www.schweizerbart.de
DOI: XXX XXX/XX/XXXX $ XXX0

and lightweight sensors provide valuable information, they do not suffice for obstacle detection and safe navigation. Only few systems are equipped with 2D laser range finders (LRF) that measure distances in a plane around the MAV (TOMIĆ et al., 2012; GRZONKA et al., 2009; BACHRACH et al., 2009; SHEN et al., 2011).

2D laser range sensors are widely used for mobile robots navigating on flat ground—due to their accurate distance measurements even in bad lighting conditions and their large field-of-view (FoV). For robots acting in 3D environments or driving on more complex terrain, three-dimensional laser scanning sensors are popular. For instance, many autonomous cars perceive obstacles by means of a rotating laser scanner with a 360° horizontal FoV, allowing for detection of obstacles in all directions (URMSON et al., 2008; MONTEMERLO et al., 2008). Up to now, such 3D laser scanners are rarely used on lightweight MAVs due to their payload limitations.

In order to enable navigation in complex 3D environments for lightweight MAVs, we designed a continuously rotating 3D laser scanner that is minimalistic in terms of size and weight and measures distances of up to 30 m in almost all directions. Figure 1 shows the sensor mounted on our MAV.

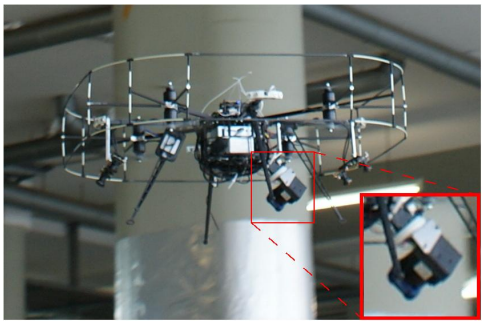


Fig. 1: The 3D laser scanner mounted on our MAV.

We use the laser scanner to perceive obstacles around the MAV by aggregating the distance measurements over a short period of time in an MAV-centric local multiresolution 3D map. The map models occupancy in all directions around the MAV and, thus, can cope with dynamic obstacles and changing environments as the regions where changes take place are updated with high frequency. We estimate the 6D motion of the MAV—relative to the egocentric map—by registering 3D laser scans with this local map. Figure 3 illustrates the architecture of our approach.

2 Related Work

The use of MAVs in recent remote sensing and robotics research varies largely in the level of autonomy—ranging from basic hovering and position holding (BOUABDALLAH et al., 2004) over trajectory tracking and waypoint navigation (PULS et al., 2009) to fully autonomous navigation (GRZONKA et al., 2012). Similarly, the complexity of environments where MAVs operate ranges from flight arenas instrumented with motion capture systems and external computing, over outdoor flights in open spaces where GPS is available, to indoor flights in restricted spaces. Limiting factors for increasing the level of autonomy and/or the complexity of environments for lightweight MAVs are onboard sensing and onboard processing power.

Particularly important for fully autonomous operation of MAVs is the ability to perceive obstacles and avoid collisions. Most autonomous MAVs, however, cannot adequately perceive their surroundings and, hence, cannot avoid all collisions. Instead, collision detection is often restricted to the two-dimensional measurement plane of laser range finders (GRZONKA et al., 2012) or to the limited FoV of forward-facing cameras (MORI & SCHERER, 2013; ROSS et al., 2013). Most often, collision avoidance is neglected altogether, e.g., by flying in a certain height when autonomously flying between waypoints.

One way to extend the FoV for obstacle detection is to combine multiple sensors. TOMIĆ et al. (2012), for example, present an autonomous MAV that perceives the environment using a stereo camera pair mounted in forward direction and a 2D laser range scanner mounted horizontally (TOMIĆ et al., 2012). Still, their perceptual field does not include the space below, above, and behind the MAV. Most similar to our work is the work of SCHERER et al. (2012) and COVER et al. (2013). The authors describe a system that is used to au-

tonomously explore rivers using visual localisation and laser-based 3D obstacle perception. Similar to their approach, we aim at perceiving as much of the surroundings as possible in order to obtain almost omnidirectional obstacle detection.

In contrast to related approaches, we are able to aggregate distance measurements from consecutive 3D scans in an egocentric 3D multiresolution map and efficiently align new 3D scans with it to estimate the 6D motion of the MAV. By aggregating the scans, we obtain a higher measurement density in the map and a higher probability of detecting obstacles.

3 System Setup

Our platform is based on the open source *MikroKopter* octocopter kit, with a co-axial arrangement of rotors. The onboard computer (Intel Core i7-3820QM 2.7 GHz, 8 GB RAM) has ample computing power for sensor data processing and navigation planning. As middleware, we employ the Robot Operating System ROS (QUIGLEY et al., 2009).

Besides the 3D laser scanner, our MAV is equipped with two stereo camera pairs (see Figure 2a) to estimate the motion of the MAV during scan acquisition.

3.1 *3D Laser Scanner*

Our continuously rotating 3D laser scanner consists of a Hokuyo UTM-30LX-EW 2D laser range finder (LRF) which is rotated by a Robotis Dynamixel MX-28 servo actuator to gain a three-dimensional FoV. As shown in Figure 2, the scanning plane is parallel to the axis of rotation, but the heading direction of the scanner is twisted slightly away from the direction of the axis—in order to enlarge its FoV.

The 2D LRF is electrically connected by a slip ring, allowing for continuous rotation of the sensor. The axis of rotation is pitched downward by 45° in forward direction, which places the core of the MAV upwards behind the sensor as depicted in Figure 2. Hence, the sensor can measure in all directions, except for a conical blind spot pointing upwards behind the robot.

The 2D laser scanner has a size of $62\,\text{mm} \times 62\,\text{mm} \times 87.5\,\text{mm}$ and a weight of 210 g. Together with the actuator (72 g) and the slip ring, the total weight of the 3D scanner is approximately 400 g.

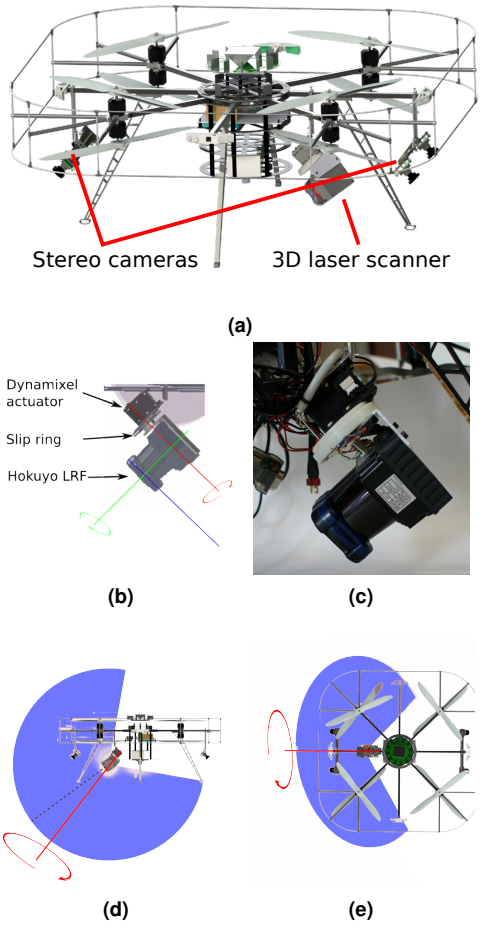


Fig. 2: CAD drawings of (a) our MAV and (b) the continuously rotating laser scanner. The Hokuyo 2D LRF is mounted on a bearing and rotated around the red axis. Its mirror is rotated around the green axis, resulting in a 2D measurement plane (blue). (c) Photo of the sensor. (d + e) CAD drawings illustrating the FoV of individual scans of the laser scanner (blue) from side and top view. The black dashed line illustrates the centre of the measurement plane. The 2D LRF is rotated around the red axis.

The Hokuyo LRF has an apex angle of 270° and an angular resolution of 0.25°, resulting

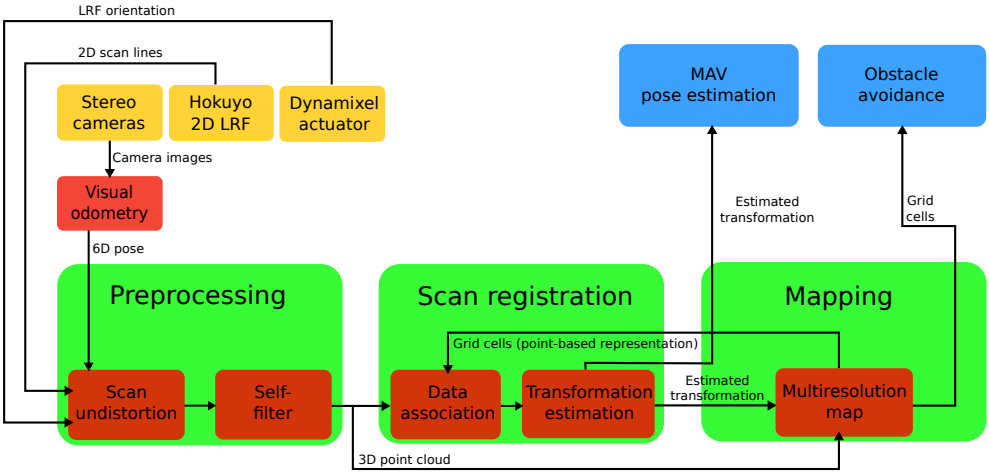


Fig. 3: An architectural overview of our system. The LRF measurements are processed in preprocessing steps described in Section 4. The resulting 3D point cloud is used to estimate the transformation between the current scan and the map as described in Section 5. Registered scans are stored in a local multiresolution map.

in 1080 distance measurements per 2D scan, called a *scan line*. The measurement accuracy is specified by the manufacturer with $\pm 30\,mm$ at $0.1 - 10\,m$ distance ($\pm 50\,mm$ at $10 - 30\,m$). The Dynamixel actuator rotates the 2D LRF at one rotation per second, producing 40 scan lines and 43,200 distance measurements per full rotation. Slower rotation is possible if a higher angular resolution is desired. For our setup, a half rotation leads to a full 3D scan of most of the environment. Hence, we can acquire 3D scans with up to 21,600 points with 2 Hz.

3.2 *Camera System*

The MAV is equipped with two stereo camera pairs pointing forward and backward with a pitch angle of 45° (see Figure 2a). Each stereo pair consists of two UEye 1221LE-M cameras with Lensagon BF2M15520 fisheye lenses. The lenses allow for a wide field-of-view up to 185° and the baseline between the cameras in a stereo pair is 20 cm. Image acquisition of all four cameras is initiated by a hardware trigger with 18 Hz, which allows for synchronised data acquisition. The hardware trigger is released by the onboard computer by using a general-purpose input/output. Thus, we gain time synchronisation between the clock on the onboard computer and the camera images. In addition, we also synchronise the internal clock of the Hokuyo laser scanner with the PC board by actively resetting the clock and then using time differences.

The mutual orientations of the cameras in a stereo pair are determined in advance with the method of SCHNEIDER & FÖRSTNER (2013). We calibrate the camera system to the laser scanner by manually labelling points in a laser scan and its corresponding pixel in the camera images. We select 50 scan points and their corresponding camera pixel, preferring spatially distributed points in corners and on edges in the scene. The 3D coordinates for the camera points are calculated as described in (SCHNEIDER & FÖRSTNER, 2013). The resulting transformation is calculated using singular value decomposition (SVD).

4 Preprocessing

In order to calculate a 3D point cloud from the scan lines originated by the LRF, a 3D point P_l in the LRF's coordinate frame is transformed to P_b in the base coordinate frame of the MAV by

$$P_b = T_l T_r T_c P_l. \tag{1}$$

Here, T_l is a static transformation between

the base frame of the MAV and the link where the 3D laser scanner is mounted and T_c is the static transformation between the 2D LRF and the bearing. T_r is a continuously changing transformation that takes into account the bearing's orientation. Its rotational parts are measured using the encoder positions reported by the Dynamixel actuator. Figure 4 shows a resulting scan of an outdoor environment.

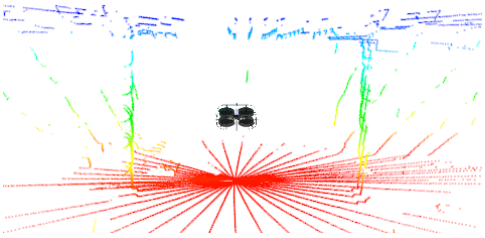


Fig. 4: A 3D scan of an outdoor environment acquired with our continuously rotating laser scanner. The colour of the points encodes the distance from the ground.

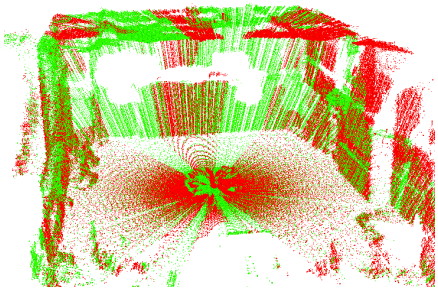


Fig. 5: Accumulated 3D scans of an indoor environment. The colour encodes the different half rotations of the scanner. Moving the optical centre of the 2D laser range finder away from the rotation axis of the actuator results in different self-occlusions of the scans from the first (green) and the second (red) half rotation.

The offset between the scanned plane section and the rotation axis results in a different FoV for two different half rotations which is shown in Figure 5. In this way, occlusion from small parts of the MAV is reduced significantly.

4.1 *Multi-Echo Detection*

The Hokuyo UTM-30LX-EW is able to measure up to three echoes of a single emitted light pulse. The number of echoes for a light pulse depends on the surface of the measured objects, i.e., shape and reflectivity. For example, transparent materials, vegetation or edges of buildings often yield more than one echo. Often, the second echo comes from a structure in the original pulse direction, behind a partial occlusion, which means that it can be treated as an additional distance measurement. Measurements from the first and the second echo are shown in Figure 6.

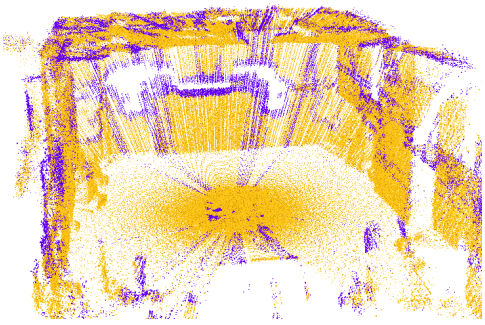


Fig. 6: Accumulated 3D scans of an indoor environment. Measurements from first echo (yellow) and the second echo (purple). In case of partial occlusions, e.g., by the MAV itself, multi-echo detection leads to an increase of distance measurements.

4.2 *Scan Aggregation*

Preprocessing raw laser scans to form 3D scans (see Section 4) considers only the rotation of the scanner w.r.t. to the MAV. The motion of the MAV during acquisition is not taken into account (so far). As a consequence, the environmental structures measured in 3D scans are not consistent. In particular, first and last scan line appear considerably disconnected, structures in consecutive scans drift. We account for this effect by undistorting 3D scans in two steps.

First, measurements of individual scan lines are undistorted with regards to the rotation of the 2D LRF around the servo rotation axis (red axis in Figure 2). Here, the rotation between

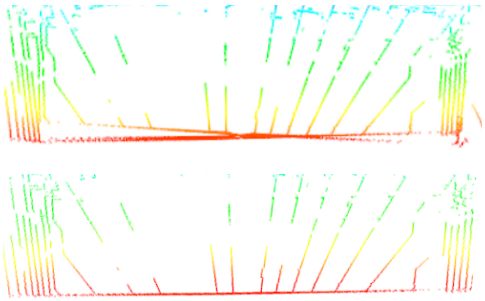


Fig. 7: Example of deskewing a 3D laser scan acquired in an indoor environment with flat ground from a side view. Sensor movement during scan acquisition yields distorted 3D scans (top). We deskew the scan based on the motion estimate (bottom).

the acquisition of two scan lines is distributed over the measurements by using spherical linear interpolation.

Second, we compensate for the motion of the MAV during acquisition of a full 3D scan. To this end, we incorporate a visual odometry estimate from the two stereo cameras. Here, a keyframe-based bundle adjustment is performed (SCHNEIDER et al., 2013) on the synchronised images with 18 Hz. Since the update rate of the 2D LRF is 40 Hz, we linearly interpolate between the estimates of the visual odometry.

The 6D motion estimate is used to assemble the individual 2D scan lines of each a half rotation to a 3D scan. Figure 7 illustrates the effect of scan undistorting.

4.3 *Self-Filter*

The laser range finder measures also points on the MAV. These points are excluded from further processing by checking the egocentric point coordinates (in the base coordinate frame of the MAV) against a simplified CAD model of the MAV. Furthermore, distance measurements that are most likely caused by the veiling effect when scanning the edge of an object are removed.

5 Scan Registration

We register 3D laser range scans with a local multiresolution map to estimate the motion of the MAV. After aggregating scans to a full 3D scan over one half rotation and transforming them to compensate for the sensor motion, the generated 3D scan is aligned to the so far built map by means of the Iterative Closest Point (ICP) algorithm. The map is incrementally updated by inserting every registered 3D scan. It is initialised using the first 3D scan acquired.

Correspondences are assigned using the point-based representation in the cells of our local grid map (Section 5.1) and the ICP algorithm estimates a transformation between the scan and the map, describing the displacement between them. We benefit from the multiresolution property of our map, which allows aligning a 3D scan in a coarse-to-fine approach. Hence, we start assigning correspondences and estimating the transformation at the coarsest level. The resulting transformation is used as initialisation for the registration on the next finer level and so forth.

5.1 *Local Multiresolution Map*

Distance measurements from the sensor are accumulated in a robot-centric 3D multiresolution map with increasing cell sizes from the MAV centre. The representation consists of multiple MAV-centred 3D grid-maps with different resolutions. On the finest resolution, we use a cell length of 0.25 m. Each grid-map is embedded in the next level with coarser resolution and doubled cell length. An example of a built local multiresolution map is shown in Figure 9.

We use a hybrid representation, storing 3D point measurements along with occupancy information in each cell. Point measurements of consecutive 3D scans are stored in fixed-sized circular buffers, allowing for point-based data processing and facilitating efficient nearest-neighbour queries.

Figure 8 shows a one-dimensional schematic illustration of the map organisation. We aim for efficient map management for translation and rotation. To this end, individual grid cells are stored in a circular buffer to allow shifting of elements in constant time. We interlace multiple circular buffers to obtain a map with three dimensions. The length of the circular buffers depends on the resolution and the size of the map. In case of a translation of the MAV, the circular buffers are shifted whenever necessary to

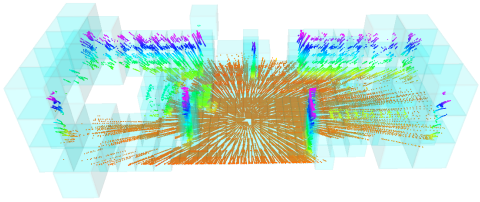


Fig. 9: A grid-based local multiresolution map with a higher resolution in the close proximity to the sensor and a lower resolution with increasing distance. The point colour encodes the distance from the ground.

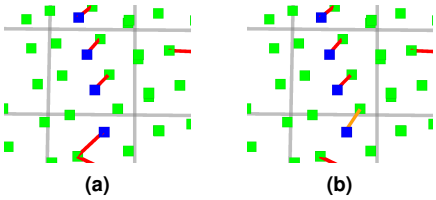


Fig. 10: Assigning point correspondences. (a) For every point of a 3D scan (blue), a corresponding map point (green) is initially assigned from the cell's point list (red line). (b) If the distance to neighbouring cells is smaller than the distance to the initial assignment, closer points might be found in the neighbouring cell (orange line).

maintain the egocentric property of the map. In case of a translation equal to or larger than the cell size, the circular buffers for respective dimensions are shifted. For sub-cell-length translations, the translational parts are accumulated and shifted if they exceed the length of a cell.

Since we store 3D points for every cell for point-based processing, individual points are transformed into the local coordinate frame of a cell when adding points, and back to the map's coordinate frame when accessing point coordinates. Every cell in the map stores a list of 3D points from the current and previous 3D scans. This list is also implemented by a fixed-sized circular buffer. If the capacity of the circular buffer is exceeded, old measurements are discarded and replaced by new measurements.

Since rotating the map would require moving all cells, our map is oriented independently of the MAV orientation. We maintain the orientation between the map and the MAV and use it to rotate measurements when accessing the map.

Besides the scan registration approach for pose estimation and map updates, the map is used for obstacle avoidance and local navigation planning (NIEUWENHUISEN & BEHNKE, 2014).

5.2 *Data Association*

When using the ICP algorithm for scan registration, corresponding points between the model and the current point cloud are assigned, usually by building a space-partitioned data structure from the model point cloud. In contrast, we continuously maintain our data structure for efficient nearest-neighbour queries to assign correspondences. Every point from a newly acquired 3D scan is directly assigned to a cell in the map in constant time. The closest point in terms of the Euclidean distance from the point list of this cell is initially assigned as corresponding point.

As illustrated in Figure 10, points in neighbouring cells might be closer to the measured point than the initially assigned point. Consequently, we extend the search to neighbouring cells, if the distance to the initial assignment is larger than the distance to the border of a neighbouring cell.

Since acquired 3D scans of the scene and the aggregated local map differ in terms of structure and point density, especially when parts of the scene have previously been occluded, individual assigned correspondences can be incorrect. These incorrect correspondences distort the transformation estimation and need to be filtered. Thus, we reject correspondences by the following criteria:

- *Asymmetric correspondences:* We check for symmetry in the assignments: for a correspondence from a scan point d_i to a map point m_i, we check if d_i is the closest point to m_i in the scan point cloud. Otherwise, the correspondence is rejected.

- *One-to-many correspondences:* In case multiple points in the 3D scan correspond to the same point in the map, we keep only the one correspondence with the smallest point-to-point distance and reject all others.

- *Correspondence trimming:* Correspondences are rejected by only considering the best θ_t percent of the assigned cor-

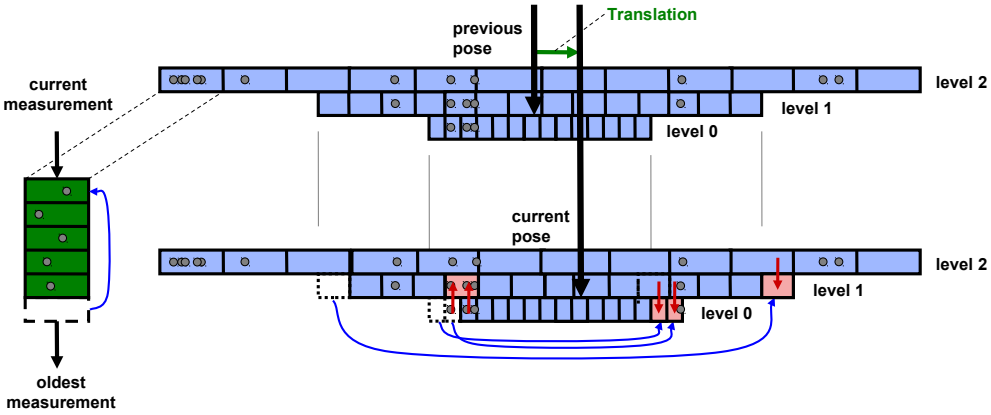


Fig. 8: One-dimensional schematic illustration of the hybrid local multiresolution map. Along with the occupancy information, every grid-cell (blue) maintains a circular buffer (green) with its associated measurement points (grey). The map is centred around the MAV and in case of a MAV motion, ring buffers are shifted according to the translational parts of the movement, maintaining the egocentric property of the map. Cells at coarser levels are used to retain points from vanishing cells at finer levels and to initialise newly added cells (red arrows). Due to the implementation using ring buffers, cells vanishing on one side of the map become new cells at the other side (blue arrows).

respondences (ranked by point-to-point distance).

- *Distance rejection:* Correspondences are rejected if the point-to-point distance exceeds a threshold θ_d.

5.3 *Transformation Estimation*

With N assigned corresponding point pairs (m_i, d_i), we determine the displacement between the points of a scan d_i and the map points m_i by finding a rigid transformation T that minimises

$$E(\mathbf{T}) = \sum_{i=1}^{N} ||\mathbf{m}_i - \mathbf{T}d_i||^2 \qquad (2)$$

using a closed-form solution using singular value decomposition (BESL & MCKAY, 1992).

In each ICP iteration, correspondences are re-assigned, the transformation best aligning the corresponding points is applied to the scan, and the following termination criteria are checked:

- $E(\mathbf{T})$ is smaller than a given threshold θ_r,
- the difference between T_k and T_{k-1} is smaller than θ_ϵ, or
- the number of iterations exceeds θ_i,

where T_k and T_{k-1} are the estimated transformations from the current and the previous iteration, respectively.

6 Experiments and Results

Due to varying shape and reflectance properties of objects, not all obstacles are perceived in every 3D laser scan (one half rotation of the scanner). Especially farther away from the robot, multiple scans may be necessary in order to adequately detect an obstacle.

In experiments, we assess the probability of detecting different types of obstacles in varying distances from the MAV. Furthermore, we assess the quality of the produced local map. The intuitions behind these experiments are the following: if a certain object can only be perceived (at least once) in n 3D laser scans, it is sufficient for reliable collision avoidance if our local mapping approach can reliably aggregate n 3D laser scans without inducing inconsistencies in the egocentric obstacle map (e.g. blurring effects due to drifts in the pose estimates). Obviously, whether or not an obstacle can be avoided also depends on the distance to the obstacle and movement direction and speed of the

Tab. 1: Absolute trajectory error (ATE) and run-time of our registration method, in comparison to visual odometry (VO), and Generalized-ICP (GICP). Referring to the root mean square error (RMSE), our approach achieves better results while being significantly faster.

	ATE (m)				run-time (ms)	
	RMSE	mean	median	max	mean	max
VO	0.151	0.134 (± 0.059)	0.129	0.324		
GICP	0.033	0.030 (± 0.013)	0.030	**0.079**	1432 (± 865)	5673
ours	**0.030**	**0.028** (± 0.015)	**0.026**	0.093	**311** (± 90)	**376**

MAV. Both can be neglected if the sensor is able to detect all types of obstacles in the immediate vicinity of the MAV (and the MAV is not flying too fast).

6.1 *Obstacle Detection Probabilities*

For assessing the probability of detecting objects in the vicinity of the robot, we have chosen seven test obstacles differing, amongst other characteristics, in size (diameter), colour and material (reflectivity), and transparency. Referring to the experiment setup in Figure 11a, the objects are mounted on a tripod holder. The MAV is positioned with distances to the holder of 1 m to 10 m. For each distance, a total of 30 3D scans are captured. We visually inspect the acquired data and count the 3D scans in which at least a part of the object is visible in the distance measurements and estimate the average detection probability. For the estimation of the detection probabilities, we considered roughly the same lengths for all obstacles (1 m, which is also the minimum safety distance during navigation). That is, the probabilities primarily depend on diameter and reflection properties rather than object length.

As can be seen in the plots (Figure 11b) and the detailed results (Figure 11c), we are able to detect all types of obstacles at 1 m distance, and with an aggregation period of 10 s up to 3 m (for the transparent plexiglass tube). Obstacles with better visibility like the metal rod, the cardboard and plastic tubes, are reliably detected up to 10 m away from the robot.

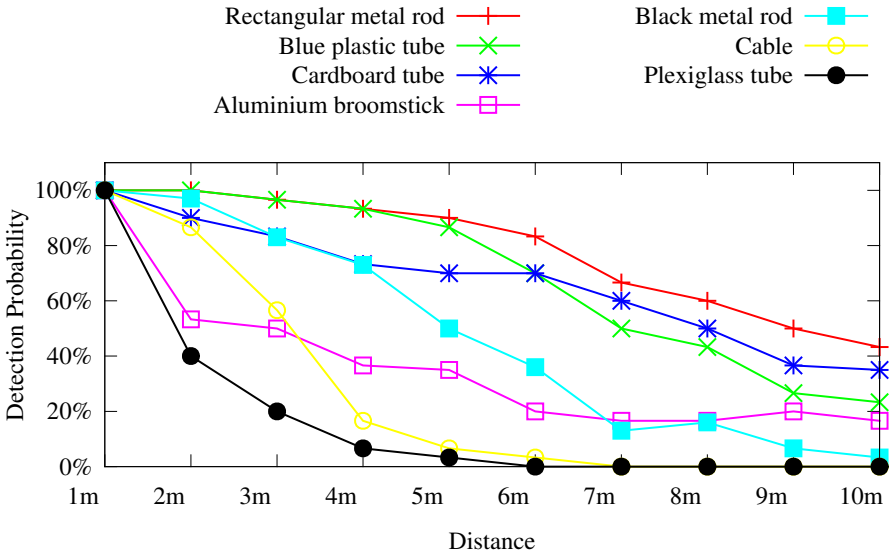
Based on the achievable results, we distinguish, respectively, different obstacle types (and distances) and detection probabilities: objects that can be reliably detected when aggregating over 2.5 s (5 scans) are considered safe and easy to detect (green in table Table 11c), obstacles that cannot be detected at least once in 10 s (20 scans) are considered especially dangerous and very hard to detect (red in Table 11c), and obstacles of moderate detection probability (yellow in table Table 11c) can be reliably handled by scan aggregation when not flying too fast.

6.2 *Scan Matching and Aggregation*

In a second experiment, we evaluate the accuracy of the scan registration in an indoor motion capture (MoCap) system. It provides accurate pose information of the MAV at high rates (100 Hz) but is restricted to a small capture volume of approximately $2\,\text{m} \times 2\,\text{m} \times 3\,\text{m}$. As error metric, the absolute trajectory error (ATE) is computed, based on the estimated and the ground-truth trajectory from the MoCap system. The reference implementation provided by STURM et al. (2012) was used to compute the error. Throughout the experiments, we used five levels for the multiresolution map and a cell length of 0.125 m at the finest level, yielding a cell length of 2 m at the coarsest level. The parameters θ_r, θ_i and θ_ϵ are manually determined. In this experiment, θ_r=1 cm, θ_ϵ=0.001 cm, θ_i=15, θ_d=1 m, and θ_t=80% showed best results.

The data set for evaluation is a 50 seconds flight sequence containing 100 3D scans, where the MAV was controlled by a human operator, taking off and landing at two locations in the

(a) Left: Different object types, from left to right: cable (∅0.75 cm), aluminium broomstick (∅2.7 cm), plexiglass tube (∅5 cm), cardboard tube (∅7.5 cm), rectangular metal rod (∅4 cm), blue plastic tube (∅11 cm). Middle and right: experiment setup with object holder, measuring tape and flying multicopter.



(b) Change of detection probabilities over increasing distance for the different objects (standing).

Object	Distance									
	1m	2m	3m	4m	5m	6m	7m	8m	9m	10m
Rectangular metal rod	100.0	100.0	96.6	93.3	90.0	83.3	66.6	60.0	50.0	43.3
Blue plastic tube	100.0	100.0	96.6	93.3	86.6	70.0	50.0	43.3	26.6	23.3
Cardboard tube	100.0	90.0	83.3	73.3	70.0	70.0	60.0	50.0	36.6	35.0
Aluminium broomstick	100.0	53.3	50.0	36.6	35.0	20.0	16.6	16.6	20.0	16.6
Black metal rod	100.0	97.0	83.0	73.0	50.0	36.0	13.0	16.0	6.6	3.3
Cable	100.0	86.6	56.6	16.6	06.6	03.3	00.0	00.0	00.0	00.0
Plexiglass tube	100.0	40.0	20.0	06.6	03.3	00.0	00.0	00.0	00.0	00.0

(c) Detailed detection probabilities in percent (standing MAV).

Fig. 11: Assessing the probabilities of detecting different types of objects in different distances, measured over 30 3D scans for each obstacle and distance. We count the 3D scans in which at least a part of the object is visible in the measurements and estimate the average detection probability.

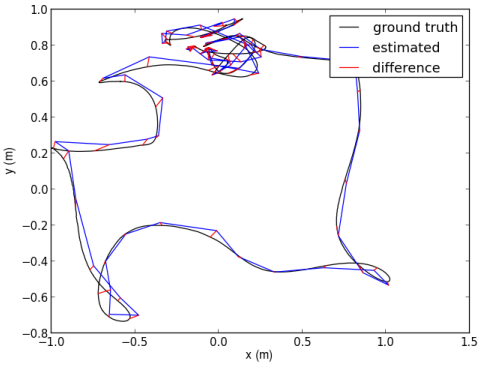


Fig. 12: Absolute trajectory error of the scan registration using the multiresolution map (blue) compared to ground-truth data from the MoCap system (black). Points of the trajectory are projected on the xy-plane.

MoCap volume. The speed during the experiment was varying from $0.1\,m/s$ to $1.1\,m/s$. Figure 12 shows the ATE of our multiresolution scan registration method, compared to the MoCap trajectory.

In quantitative experiments, we compare our method to a state-of-the-art registration method, the Generalized-ICP (SEGAL et al., 2009). In addition, we evaluate the accuracy of the visual odometry that is used to undistort acquired 3D scans. The mean, standard deviation and maximum ATE of all three methods are summarized in Table 1. The results indicate that both scan registration methods improve the motion estimate from the visual odometry and that the trajectory generated by our method has a slightly lower ATE compared to the Generalized-ICP. The run-times of both methods for this experiment on a single core of an Intel Core i7-3820QM (2.7 GHz) processor, are also summarized in Table 1, showing that Generalized-ICP is computationally much more expensive than our method.

In a third experiment, we assess the quality of our grid-based map. Consecutive 3D scans are aligned with the map and 3D points are added to respective cells. Figure 13 shows the point-based representation of the map at different time steps, accounting for an increasing density of the map after adding consecutive 3D scans. As obstacle, the blue plastic tube shown in Figure 11a was used. The MAV was hov-

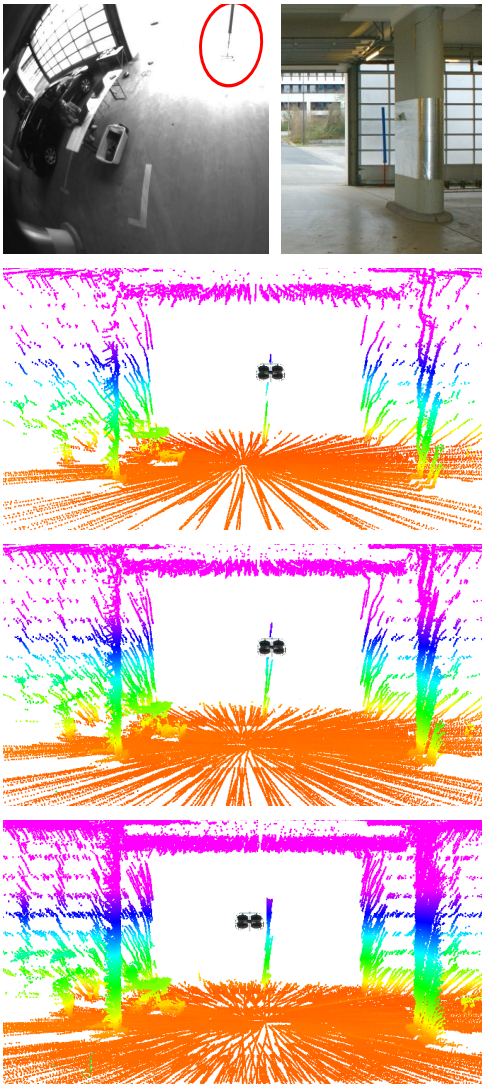


Fig. 13: Aggregating 3D scans in a scene of the third experiment. Photos show the sensed pole (red ellipse) and the surroundings from an on-board camera with fish-eye lens (top left) and an external camera (top right). Scan aggregation increases the detection probability and the point density in the map (from top to bottom: map after aggregating 5, 10 and 20 scans.

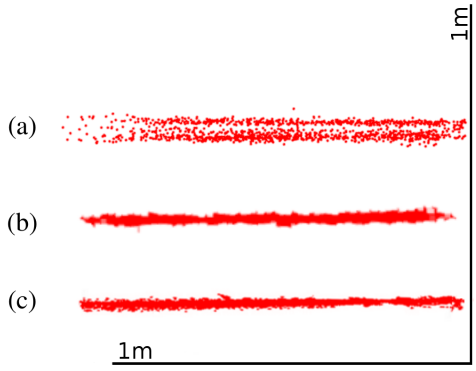


Fig. 14: A cut-out part of the floor from a side view after scan aggregation (length of the segments: 1 m). (a) using only visual odometry; (b) visual odometry combined with scan registration using Generalized-ICP; (c) visual odometry with our multiresolution scan registration.

ering at a height of 2 m with a distance to the obstacle of 8 m. The experiment shows that we can reliably and accurately track the MAV's movement over 10 s and aggregate the acquired 20 3D scans to a dense and sharp map. Since ground-truth data, e.g., from a MoCap system was not available in this experiment, we evaluate the different methods by inspecting the variation of points in a planar area. Figure 14 shows a part of the floor in the resulting point-based representation from a side view. It can be seen that using scan registration decreases the thickness of the floor significantly, indicating an improved motion estimate. Similar to the results of the second experiment, the resulting floor patch generated by aggregating scans using our method is slightly thinner. Note that in normal operation, scans are only added if they properly align with the map, i.e., $E(\mathbf{T})$ is smaller than θ_r in Eq. (1). For this experiment, we added every scan to the map to have a fair comparison to the motion estimate from visual odometry.

7 Conclusion

We designed a small and lightweight continuously rotating 3D laser scanner that is particularly well suited for the use in MAVs. The sensor allows for measuring distances of up to 30 m in almost all directions with a minimal blind spot. For each light pulse, up to three echoes are reported which is advantageous in case of transparent material, vegetation, or edges of buildings.

We use the sensor to perceive obstacles in the vehicle's vicinity by building a grid-based obstacle map. We estimate the motion of our MAV by registering 3D laser scans with the map.

In experiments, we showed that our approaches to pose estimation and laser scan matching allow for reliably aggregating 3D scans over short periods of time, accurately enough to improve detection probability and without causing inaccuracies in the estimation of the position of detected obstacles.

Overall, we can build dense and sharp 3D obstacle maps and estimate the vehicle's trajectory by 3D scan registration.

As an outlook on ongoing and future work, we integrated the sensor in our MAV and conducted experiments with dynamical obstacles, showing that the MAV is able to omnidirectional perceive obstacles and to react on them (NIEUWENHUISEN et al., 2013).

Acknowledgements

This work has been supported partially by grant BE 2556/7-1 of German Research Foundation (DFG).

References

BACHRACH, A., HE, R. & ROY, N., 2009: Autonomous flight in unstructured and unknown indoor environments. – European Micro Aerial Vehicle Conf (EMAV), 1–8.

BESL, P. J. & MCKAY, N. D., 1992: A method for registration of 3-D shapes. – IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) **14** (2): 239–256.

BOUABDALLAH, S., MURRIERI, P. & SIEGWART, R., 2004: Design and control of an indoor micro quadrotor. – Robotics and Automation (ICRA), IEEE International Conference on.

COVER, H., CHOUDHURY, S., SCHERER, S. & SINGH, S., 2013: Sparse tangential network (SPARTAN): Motion planning for micro aerial vehicles. – Robotics and Automation (ICRA), IEEE International Conference on.

GRZONKA, S., GRISETTI, G. & BURGARD, W., 2009: Towards a navigation system for autonomous indoor flying. – Robotics and Automation (ICRA), IEEE International Conference on.

GRZONKA, S., GRISETTI, G. & BURGARD, W., 2012: A fully autonomous indoor quadrotor. – IEEE Trans. on Robotics **28** (1): 90–100.

MONTEMERLO, M., BECKER, J., BHAT, S., DAHLKAMP, H., DOLGOV, D., ETTINGER, S., HAEHNEL, D., HILDEN, T., HOFFMANN, G., HUHNKE, B., JOHNSTON, D., KLUMPP, S., LANGER, D., LEVANDOWSKI, A., LEVINSON, J., MARCIL, J., ORENSTEIN, D., PAEFGEN, J., PENNY, I., PETROVSKAYA, A., PFLUEGER, M., STANEK, G., STAVENS, D., VOGT, A. & THRUN, S., 2008: Junior: The stanford entry in the urban challenge. – Journal of Field Robotics **25** (9): 569–597.

MORI, T. & SCHERER, S., 2013: First results in detecting and avoiding frontal obstacles from a monocular camera for micro unmanned aerial vehicles. – Robotics and Automation (ICRA), IEEE International Conference on.

NIEUWENHUISEN, M. & BEHNKE, S., 2014: Hierarchical planning with 3D local multiresolution obstacle avoidance for micro aerial vehicles. – Proceedings of the Joint Int. Symposium on Robotics (ISR) and the German Conference on Robotics (ROBOTIK).

NIEUWENHUISEN, M., DROESCHEL, D., SCHNEIDER, J., HOLZ, D., LÄBE, T. & BEHNKE, S., 2013: Multimodal obstacle detection and collision avoidance for micro aerial vehicles. – Proc. of European Conference on Mobile Robots (ECMR).

PULS, T., KEMPER, M., KUKE, R. & HEIN, A., 2009: GPS-based position control and waypoint navigation system for quadrocopters. – Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on.

QUIGLEY, M., CONLEY, K., GERKEY, B. P., FAUST, J., FOOTE, T., LEIBS, J., WHEELER, R. & NG, A. Y., 2009: Ros: an open-source robot operating system. – ICRA Workshop on Open Source Software.

ROSS, S., MELIK-BARKHUDAROV, N., SHANKAR, K. S., WENDEL, A., DEY, D., BAGNELL, J. A. & HEBERT, M., 2013: Learning monocular reactive uav control in cluttered natural environments. – Robotics and Automation (ICRA), IEEE International Conference on.

SCHERER, S., REHDER, J., ACHAR, S., COVER, H., CHAMBERS, A. D., NUSKE, S. T. & SINGH, S., 2012: River mapping from a flying robot: state estimation, river detection, and obstacle mapping. – Autonomous Robots **32** (5): 1–26.

SCHNEIDER, J. & FÖRSTNER, W., 2013: Bundle Adjustment for Omnidirectional Camera Systems with Points at Infinity Including System Calibration. – Z. f. Photogrammetrie, Fernerkundung und Geoinformation: to appear.

SCHNEIDER, J., LÄBE, T. & FÖRSTNER, W., 2013: Incremental real-time bundle adjustment for multi-camera systems with points at infinity. – ISPRS Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W2.

SEGAL, A., HAEHNEL, D. & THRUN, S., 2009: Generalized-ICP. – Proc. of Robotics: Science and Systems (RSS).

SHEN, S., MICHAEL, N. & KUMAR, V., 2011: Autonomous multi-floor indoor navigation with a computationally constrained micro aerial vehicle. – Robotics and Automation (ICRA), IEEE International Conference on, 2968 –2969.

STURM, J., ENGELHARD, N., ENDRES, F., BURGARD, W. & CREMERS, D., 2012: A benchmark for the evaluation of rgb-d slam systems. – Intelligent Robots and Systems (IROS), IEEE/RSJ International Conference on.

TOMIĆ, T., SCHMID, K., LUTZ, P., DOMEL, A., KASSECKER, M., MAIR, E., GRIXA, I., RUESS, F., SUPPA, M. & BURSCHKA, D., 2012: Toward a fully autonomous UAV: Research platform for indoor and outdoor urban search and rescue. – Robotics Automation Magazine, IEEE **19** (3): 46–56.

URMSON, C., ANHALT, J., BAE, H., BAGNELL, J. A. D., BAKER , C. R., BITTNER, R. E., BROWN, T., CLARK, M. N., DARMS, M., DEMITRISH, D., DOLAN, J. M., DUGGINS, D., FERGUSON, D., GALATALI, T., GEYER, C. M., GITTLEMAN, M., HARBAUGH, S., HEBERT, M., HOWARD, T., KOLSKI, S., LIKHACHEV, M., LITKOUHI, B., KELLY, A., MCNAUGHTON, M., MILLER, N., NICKOLAOU, J., PETERSON, K., PILNICK, B., RAJKUMAR, R., RYBSKI, P., SADEKAR, V., SALESKY, B., SEO, Y.-W., SINGH, S., SNIDER, J. M., STRUBLE, J. C., STENTZ, A. T., TAYLOR , M., WHITTAKER, W. R. L., WOLKOWICKI, Z., ZHANG, W. & ZIGLAR, J., 2008: Autonomous driving in urban environments: Boss and the urban challenge. – Journal of Field Robotics Special Issue on the 2007 DARPA Urban Challenge, Part I **25** (8): 425–466.

Address of the Author:

DAVID DROESCHEL, Autonomous Intelligent Systems Group, Computer Science Institute VI, University of Bonn, Friedrich-Ebert-Allee 144, 53113 Bonn, Germany
http://www.ais.uni-bonn.de
droeschel@ais.uni-bonn.de