Collaborative Object Picking and Delivery with a Team of Micro Aerial Vehicles at MBZIRC

Matthias Nieuwenhuisen, Marius Beul, Radu Alexandru Rosu, Jan Quenzel,
Dmytro Pavlichenko, Sebastian Houben, and Sven Behnke

Abstract— Picking and transporting objects in an outdoor environment with multiple lightweight MAVs is a demanding task. The main challenges are sudden changes of flight dynamics due to altered center of mass and weight, varying lighting conditions for visual perception, and coordination of the MAVs over unreliable wireless connections.

At the Mohamed Bin Zayed International Robotics Challenge (MBZIRC) teams competed in a Treasure Hunt where three MAVs had to collaboratively pick colored disks and drop them into a designated box. Only little preparation and test time on-site required robust algorithms and easily maintainable systems to successfully achieve the challenge objectives.

We describe our multi-robot system employed at MBZIRC, including a lightweight gripper, a vision system robust against illumination and color changes, and a control architecture allowing to operate multiple robots safely. With our system, we—as part of the larger team NimbRo of ground and flying robots—won the Grand Challenge and achieved a third place in the Treasure Hunt.

I. INTRODUCTION

Aerial manipulation—especially picking, transporting, and delivering objects—became an area of much interest in recent years. Micro aerial vehicles (MAV) are well suited to quickly deliver small, but valuable objects, e.g., spare parts or medical substances. A particular advantage of employing aerial vehicles to detect and pick objects is that—in contrast to ground vehicles—they can reach otherwise hard to access or even dangerous areas.

To facilitate the development in this field, one of the tasks at the Mohamed Bin Zayed International Robotics Challenge 2017 (MBZIRC) was collaborative picking with MAVs. The Treasure Hunt task was to find and pick colored, ferromagnetic discs from the ground of an outdoor arena and to deliver them to a designated box in a predefined drop zone. Fig. 1 shows one of our MAVs while picking a moving object.

Teams were provided with rough specifications of the objects, i.e., diameter, height above ground, maximum weight of $500\,\mathrm{g}$, and the possible colors, in advance. The drop box was specified by its approximate dimensions. Nevertheless, the exact arena setup—including colored markings on the ground making color-based perception challenging—was not known in advance and teams had to develop robust and flexible systems.
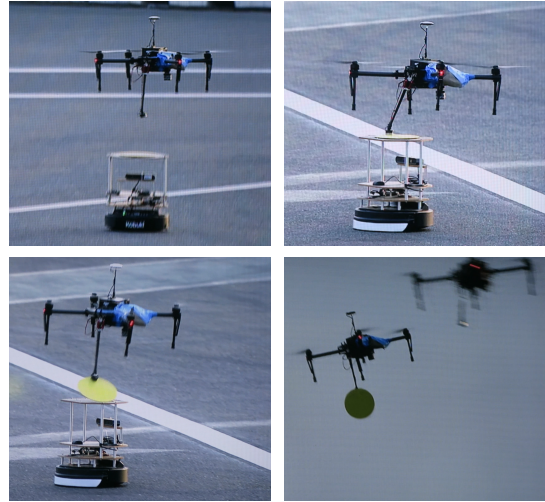
Fig. 1. Picking a dynamic object. Our MAV follows the yellow disc with visual servoing. The telescopic rod and the ball joint of our electromagnetic gripper allow compliant picking without disturbing attitude control of the MAV. The picked objects were delivered to a drop box up to $75\,\mathrm{m}$ away.

In contrast to lab experiments or controlled field tests, the particular challenge of the competition was the much reduced testing time. The arena was only accessible for teams at assigned time slots. In total, individual teams had two rehearsal slots of 35 minutes and four competition trials—two for the Treasure Hunt and two for the Grand Challenge. The systems had to be set up in the arena in only five minutes in each run. Consequently, complex algorithms that need extensive fine-tuning or are prone to failing in some cases are not an option for a competition system. Thus, we focused on simple but robust approaches and tried to identify and cover as many issues in advance as possible. Experiences gained during trials had to be incorporated into the system without additional testing before the next trial. Hence, the system complexity needed to be as low as possible to eliminate error sources. Furthermore, quickly changing overcast and light sandstorms changed illumination and navigation conditions significantly from trial to trial.

Due to these challenging conditions, only very few teams managed to score in this task autonomously. Our main contributions are

- robust detection of objects with only roughly specified color under varying lighting conditions,
- relative navigation while picking and dropping,
- lightweight and flexible picking hardware,
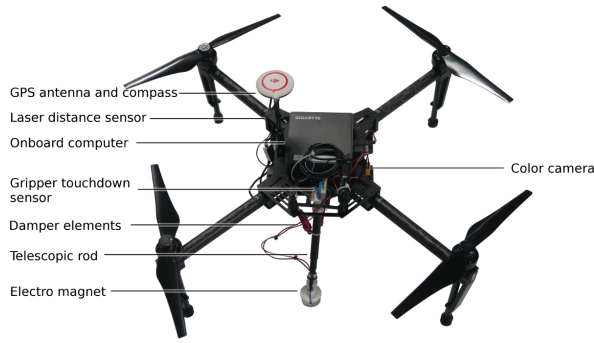- coordination of multiple MAV with and without WiFi connections, and

Fig. 2. Closeup of our MAV. The Matrice 100 is equipped with a down-pointing color camera for object and drop box detection. Objects are picked with an electromagnetic gripper on a telescopic rod. A small lidar sensor measures the distance to the ground. All calculations are performed on a powerful onboard PC.

- evaluation of our integrated multi MAV system in challenging conditions at MBZIRC.

II. RELATED WORK

Aerial manipulation has been investigated by multiple research groups.

Morton et al. [1], for example, developed an MAV with manipulation capabilities for outdoor use. The MAV is equipped with a 3-DoF arm which is operated in hover mode without object perception or autonomous flight. An MAV with a 2-DoF robot arm that can lift relatively heavy weights is presented by the authors [2]. The controller explicitly models the changes in the vehicle dynamics by attaching a heavy object. The object positions are known beforehand. We employ a trajectory generator that uses a very simple dynamics model with frequent replanning on top of a model-free attitude controller to achieve robustness against changes in flight dynamics.

Ghadiok et al. [3] built a lightweight quadrotor for grasping objects in indoor environment. Similar to our work, they have a lightweight and compliant gripper to cope with uncertainties during grasping. Target objects are equipped with infrared beacons; we detect objects based on coarse color specifications. The ETH Zürich MBZIRC team Electronic Treasure Hunters describe a preliminary state of their approach to solving the challenge in [4]. They employ an electro-permanent magnetic gripper and color blob detection for visual servoing.

Some work exists on cooperative transport of objects [5], [6]. However, this work assumes a permanently connected object, i.e., no picking and placing.

III. SYSTEM SETUP

Our picking MAV, depicted in Fig. 2, is based on the DJI Matrice 100 quadcopter platform. This platform is designed for research and development—and consequently offers an easy ROS integration. We equipped the basic platform with a small, but powerful Gigabyte GB-BSi7T-6500 onboard PC with an Intel Core i7-6500U CPU running at $2.5/3.1\,\mathrm{GHz}$ and $16\,\mathrm{GB}$ of RAM. For object and drop box detection, we employ a downward facing Point Grey BFLY-U3-23S6C-C color camera with a wide-angle lens.

For allocentric localization, we use the filter result from the MAV flight control unit (FCU) fusing GNSS, barometer, and IMU data. A small Garmin Lidar Lite measures the distance to the ground to allow for exact, drift-free navigation close to the ground. To avoid electromagnetic interference between components—in particular USB3 and GPS—the core of our MAV is wrapped in electromagnetic shielding material. This increases the system stability significantly.

Our gripper is an electromagnet with a telescopic rod. The rod is passively extended to its full length by gravity and can be shortened up to $31\,\mathrm{cm}$ when in contact with an object. A switch detects shortening the rod. Two dampers avoid fast oscillations while still allowing the rod to align with the gravity vector. The gripper weighs $220\,\mathrm{g}$ including mounting and electronics.

All three MAVs are similar in hard- and software—also most of the configuration is derived from a single copter ID—to simplify the handling of multiple robots in stressful competition situations.

To make all components easily transferable between the test area at our lab and also different arenas on site, we defined all coordinates (x, y, z, θ_{yaw}) in a field-centric coordinate system. The center and orientation of the current field were broadcasted by a base station PC to all active MAVs—and the ground robot in the Grand Challenge. Furthermore, the base station PC, an Intel NUC equipped with a DJI N3 module, constantly measures its GNSS position and broadcasts position correction offsets to eliminate larger position deviations caused by atmospheric effects.

The communication among the MAVs and the base station is conducted over a stationary WiFi infrastructure. For robustness, we employ a UDP protocol that we developed for connections with low-bandwith and high-latency [7].

IV. VISUAL PERCEPTION

The mission plan demands for robust perception in two phases: when sweeping the field at a very high speed, the copter has to detect and track the pickable objects with low latency; after arrival in the drop zone the box has to be reliably detected and tracked during approach.

A. Object Detection

The challenge rules define two kinds of pickable objects: thin ferromagnetic disks with a diameter of $20\,\mathrm{cm}$ and thin rectangular objects with a size of $200 \times 20\,\mathrm{cm}$. The former were colored in red, green, blue, and yellow while the latter were exclusively orange. Since little detail was given beforehand about the competition arena and, thus, possible distracting objects, the detection algorithm was based on both color and shape information where the specific color was supposed to be trained quickly on-site when the actual objects were available. The learned color distribution is also able to model the effect of different lighting conditions and reflective object surfaces.

Briefly, the camera image is scaled down depending on the copter altitude and a bird's-eye perspective transform is applied in order to account for its attitude (see Fig. 3
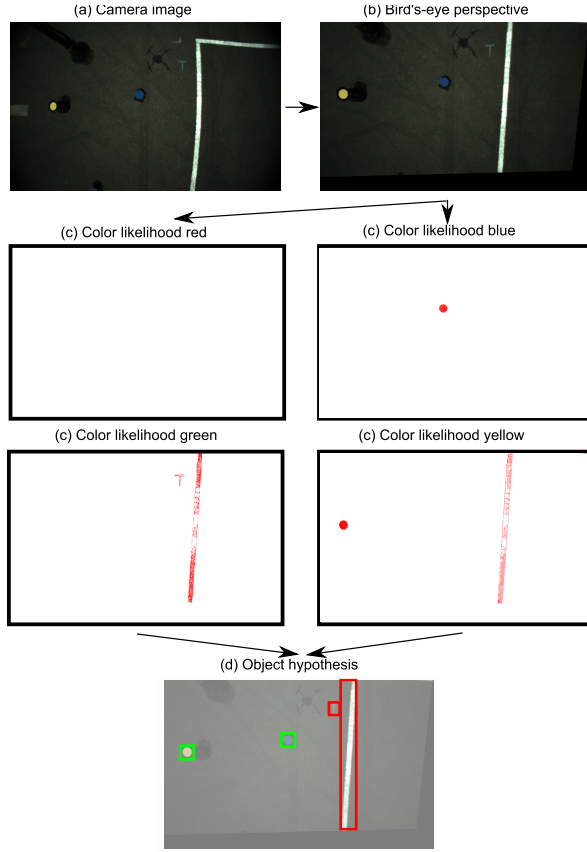
Fig. 3. Overview of the object detection pipeline: (a) original camera image, (b) undistorted bird's-eye representation, (c) color likelihood images, (d) detection hypothesis in green (accepted) and red (discarded).

(b)). Please note, that during maneuvering the camera can be significantly tilted. A pixel-wise transform is computed assigning the likelihood of belonging to one of the relevant colors which results in one likelihood image per color (see Fig. 3 (c)). A blob detection method identifies the connected regions which are then filtered by several shape (aspect ratio, convexity, size) and color (average likelihood, contrast to background) criteria (see Fig. 3 (d)).

The initial image transform serves to simplify the detection problem but also to limit the computational burden. Let therefor be r the magnitude of the shorter side of a detectable object in meters and h the MAV altitude obtained by relative barometric and laser range measurements. The image is scaled by

$$s := max \left(min \left(30 \cdot \frac{h}{rf}, 1. \right), 0.1 \right),$$

where f is the camera focal length. For later convenience, s is rounded to the first position after the decimal point, hence, s may only obtain one of ten possible values.

For defining the bird's-eye transformation, let $\hat{r}_z$ be the IMU gravitational vector in camera coordinates. The rotation matrix

$$\hat{R} := (\hat{r}_x, \hat{r}_y, \hat{r}_z)$$
$$\text{with } \hat{r}_x := \begin{pmatrix} 0 & 1 & 0 \end{pmatrix}^T \times \hat{r}_z,$$
$$\hat{r}_y := \hat{r}_z \times \hat{r}_x,$$

describes the rotation from the camera frame into a camera frame where the image plane is aligned with the ground plane, i.e., the matrix

$$M := sK_c\hat{R}K_c^{-1}$$

yields a pixel coordinate transform into the bird's-eye representation via homogeneous coordinates. Finally, taking the lens distortion into account, we arrive at

$$(u, v) \mapsto M \begin{pmatrix} d(u,v) \\ 1 \end{pmatrix} \tag{1}$$

with an invertible radial-tangential lens undistortion function d operating on the image coordinates (u, v). K_c is given by the camera intrinsics. In order to efficiently execute this transform, please note that d is static such that a pixel-wise lookup table for each of the ten possible rounded scale factors s can be precomputed. The second part of the mapping in equation (1) is linear-projective and can be computed very efficiently.

For detection processing, the color image is transformed into HSV space. As pixel-wise color likelihood we use a max-mixture of Gaussians model:

$$\max_i \left(exp \left\{ - (x - c_i)^T \, diag \left(\sigma_h^2, \sigma_s^2, \sigma_v^2 \right) (x - c_i) \right\} \right) \tag{2}$$

where $x = (x_h, x_s, x_v)^T$ denotes the three channel pixel value, $c_i = (c_{i,h}, c_{i,s}, c_{i,v})^T$ are a number of trained prototype pixel values, and $\sigma_h, \sigma_s, \sigma_v$ three hyperparameters. During training, the channel-wise mean of all pixels from a manually labeled object detection is computed and stored in a single prototype pixel value c_i. In order to efficiently calculate (2), a lookup-table is set up where the HSV color space is sampled with a grid of $20 \times 20 \times 20$ points.

The resulting images contain a point-wise likelihood for each of the detectable colors. As blob detection we use the implementation by Nistér and Stewénius [8] of the maximally-stable extremal regions (MSER) algorithm which yields a number of initial hypotheses. In order to select the final detections, we regard

- the size: the number of pixels of the region,
- the aspect ratio of its oriented bounding box,
- the convexity: the ratio of the number of pixels over the area of their convex hull,
- the color: the average likelihood in the region, and
- the background: the average discrepancy between the likelihood of pixels inside the region and pixels sampled from a surrounding circle.

A classifier can be trained on these quantities, but for the scope of this venture, the selection criteria were manually and individually tuned which allowed us to better follow the behavior of the detector and quickly adapt it in case of failure.

Upon a significantly confident detection, the algorithm switches to a tracking mode where only a window around the last known object position is searched for only the identified color. This enables a much faster detection rate, in particular during the picking maneuver. In close proximity

to the ground when the object is expected to be only partly visible in the image, the shape criteria are ignored when filtering the hypothesis.

It is further possible to use this algorithm to detect whether an object is attached to the gripper by filtering for very large detections in the specific color.

B. Drop Box Detection

In contrast to the pickable objects, the drop box was not specified by the challenge rules. Hence, we deployed a very general approach, only assuming that the box was rectangular and would provide some contrast to the surrounding ground. It is explicitly not assumed that the box would be uniformly colored[1]. Nevertheless, the dimensions of the box are parametrized. As for the object detection, the camera image is transformed to a bird's-eye perspective to account for the copter attitude. A Hough transform of the resulting gradient image yields line segments that are combined in a RANSAC-like procedure. In order to combine only promising pairs of line segments, a hash table with the line orientation as key is set up and only approximately perpendicular line segments are sampled. Testing the rectangularity, aspect ratio, and size of all RANSAC hypotheses provides the detection.

V. STATE ESTIMATION

At the onboard PC, we use state estimation filters for maintaining a height offset between the measured height over ground and the barometer, and for estimating the position and velocity of (faster) moving objects during picking. Our generic filter design does not make any model assumptions and all dimensions are treated independently. Thus, we can employ the same filter with different dimensionality for both use cases. We modified our state estimation filter from [9] by replacing the acceleration-based prediction step with a constant velocity assumption.

A. Laser Height Correction

The operation close to the ground during picking makes a good height estimate over ground obligatory. The Matrice 100 provides absolute GNSS altitude and a barometric height to a starting position. While the first is usually not very accurate, especially close to the ground, the second is prone to drift over time. Hence, we employ a laser distance sensor in order to correct the drift. The laser measurements close to the ground are very noisy, at greater heights they are assumed to be not reliable due to bright sunlight, but if available their measurement is correct. In contrast, the barometer is very reliable and locally consistent. Thus, we maintain an offset between laser height and barometric measurements and use this offset to correct the barometer drift. To acquire the correct heights, we first transform the laser measurements into an attitude-corrected frame. If the resulting measurement is between $0.1\,\mathrm{m}$ to $6\,\mathrm{m}$, we use this value to correct the height offset. The advantage of this approach is that even without laser measurements over longer periods of time, the MAV can safely navigate in higher altitudes, e.g., to explore
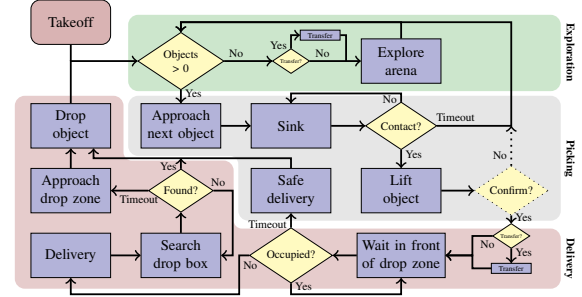

Fig. 4. Overview of our state machine. To avoid any false negatives, the dotted part was shortcut during competition.

or deliver objects, but the filter still converges quickly to the correct height over ground when picking.

B. Object Tracking

To track dynamic objects, we filter their positions and velocities in an allocentric frame to omit the orientation dimension. Furthermore, in contrast to our MAV state estimation filter, we only incorporate position measurements, letting the filter predict velocities without explicit correction. Outputs of the filter are allocentric 2D position and velocity estimates, used to intercept the target objects. For the very slow moving objects used in the actual MBZIRC challenges, estimating the object velocities was not necessary such that we omitted their estimation in favor of filter stability.

VI. NAVIGATION AND CONTROL

Whereas the competition arena is of rectangular shape without larger obstacles and with good GNSS coverage, picking small objects from the ground and the coordination of a team of multiple collaborating robots pose challenges for navigation and control.

The top-level coordination is achieved by is a state machine running at $50\,\mathrm{Hz}$, depicted in Fig. 4. The state machine selects navigation targets and configures the perception and navigation modules and the hardware. After takeoff, and when the list of detected objects is empty, the system starts to explore the arena in a spiral pattern at a height of $4\,\mathrm{m}$. The maximum exploration speed is $6\,\mathrm{m/s}$. Immediately after object detection, we approach the closest object for picking[2]. When reaching a position above the detected object, we confirm its color and reconfigure the object perception to use its fast tracking mode with only one color. With visual servoing, the MAV descents within a cone around the object center until a) contact of the gripper with the object is detected, b) the measured distance to the ground from the laser falls below a safety threshold, or c) the object is not perceivable any more. If the object is no longer perceived, the MAV enters the exploration mode, in the other cases it ascents and starts visual confirmation that an object is attached. Please note that during the MBZIRC we disabled the visual object confirmation and assumed that every picking attempt was successful as false positives were far less problematic than false negatives regarding the scoring scheme. Furthermore, the gripper was much more reliable than expected.

---

[1]As a matter of fact, it was uniformly colored.

[2]Due to the high exploration velocity, it is possible to observe multiple objects before switching to the approach state.
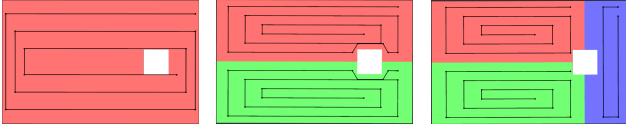
Fig. 5. Sectors for safe operation of multiple MAVs. We divide the arena into one to three sectors, depending on the number of active MAVs, all with access to the drop zone (white rectangle). The black lines depict the exploration patterns.

To drop objects, the MAV enters the drop zone at a height of $8\,\mathrm{m}$ and starts a local exploration flight to detect the drop box. If the drop box is detected, the MAV descents to $1\,\mathrm{m}$ height and drops the object. After a timeout, the MAV drops the object at the predefined center of the drop zone. As long as the drop zone is occupied, the MAV waits outside. It enters the drop zone close to its border after a timeout to drop the object safely for partial points.

The allocentric navigation is based on GNSS positions—bias corrected with help of the base station—in a field-centric coordinate system. Positions in the arena, e.g., endpoints of exploration trajectories, starting points of picking attempts, and the drop zone, are directly approached by means of our time-optimal trajectory generator [10] with maximum velocity. Except for exploration this is $8.33\,\mathrm{m/s}$. We use a generic motion model in our trajectory generator combined with frequent replanning. Our controller generates attitude setpoints which are executed by the Matrice 100 onboard controller. This approach is independent from the accurate weight and other parameters that change when picking or dropping objects and, thus, robust but still efficient.

To ensure safe operation of multiple MAVs flying at high speed, we divide the arena into sectors (see Fig. 5). Sectors are derived from the number of active robots and their IDs. Within their sector, the MAVs are allowed to freely navigate below a maximum altitude. Outside their sectors, the MAVs transfer at assigned higher altitudes on straight lines.

We assume that wireless connections between agents are unreliable. Consequently, we designed our system in a way to stay operative without communication, but to employ knowledge about the other agents to operate more efficiently, if available. Our system has no central control instance or explicit negotiation between agents. The MAVs broadcast selected parts of their knowledge, namely a) allocentric 3D position, b) current navigation target, c) detected objects outside of own operation sector, d) if the MAV is flying or landed. The received information is integrated into the individual world models. If the team communication is reliable, the agents can enter the drop zone immediately if unoccupied and follow dynamic objects into neighboring sectors while picking. In case of disturbed communication, fallback strategies are in place, e.g., more conservative navigation in other sectors and time slots for entering the drop zone.

VII. EVALUATION

The competition objective was to pick metal discs with a radius of $10\,\mathrm{cm}$ on a $20\,\mathrm{cm}$ stand in the colors red, green, blue, and moving discs in yellow, and deliver them to a
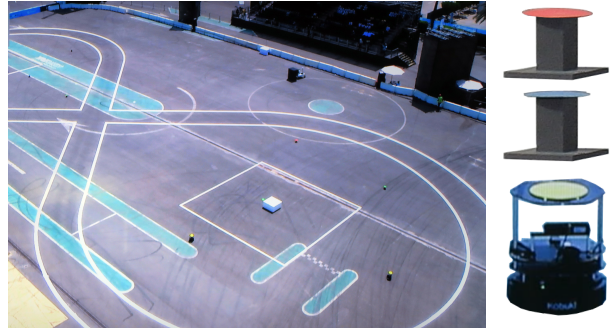

Fig. 6. Arena at MBZIRC. The colored discs randomly distributed over the arena had to be detected, picked, and dropped into the white box or surrounding drop zone. Many white lines and colored markings on the ground posed a challenge for object detection. Right: Closeup of some objects.

designated drop zone. The elevated position of the disks made it necessary to pick the objects during flight without the possibility to land nearby. A white box with ground plane of one square meter—the preferred place to drop objects—was placed in the drop zone of size $10 \times 10\,\mathrm{m}$. Dropping objects next to the box but into the drop zone was awarded with half of the points. Overall, 16 discs were randomly distributed over the $100 \times 60\,\mathrm{m}$ arena. Furthermore, three larger elongated objects of orange color were placed in the arena, which we were able to detect, but did not attempt to pick. The maximum challenge duration was 20 minutes. Fig. 6 shows the arena setup in the Grand Challenge. Videos of our evaluation can be found on our website[3].

In the first attempt of the first trial, we began to explore the arena with three MAVs simultaneously. The trial was canceled because of very strong winds with a speed of up to $9\,\mathrm{m/s}$. Qualitatively, all MAVs followed their assigned exploration trajectories until then.

In the second attempt of the first trial, we explored the arena with three MAVs and successfully picked two discs on moving bases. One of the discs could be delivered into the drop box. Before the second disc could be delivered, the referees called a reset and the MAV landed with the disc still attached. Due to conservative safety distances to the ground, we could not pick that disc after the reset. Furthermore, two MAVs arrived at the drop zone at the same time and were kept in a deadlock situation. Modifications to the system during the competition were not allowed, so we could only address these issues between trials. This was the fourth-best result of all 36 trials—18 teams with two trials per team where the better trial counted for the final score—in the Treasure Hunt and worth a third place.

The second trial took place with very strong wind. Objects were detected reliably and the descent of the MAVs was stable despite the wind, but the MAVs always had an offset of a few centimeters into the wind direction when picking.

In the first trial of the Grand Challenge, we started with three MAVs, one failed directly at takeoff due to a hardware defect. The other two explored the arena and started picking and delivering objects. As the field was not covered in full

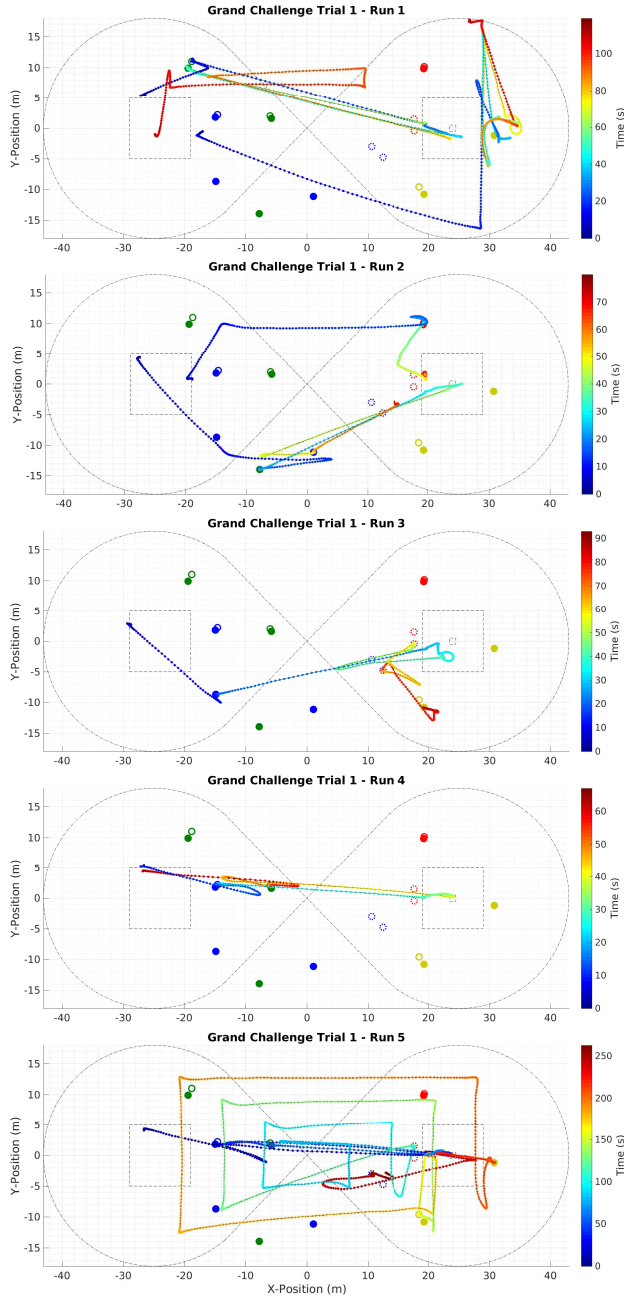[3]www.ais.uni-bonn.de/videos/ECMR_2017_Picking

Fig. 7. Treasure Hunt in the Grand Challenge. Each image shows the trajectories of the active MAVs during the Grand Challenge (separated by 4 resets). Solid disks represent successful and rings show missed picks. The dotted disks indicate disks lying on the ground. The left rectangle is the starting zone, the right one the drop zone including the drop box. In Run 1 and Run 2, two MAVs were active. In Runs 3-5, only one MAV was active since the other one worked erroneous. It flew way to high so we had to call a reset. The following colored disks were picked (p) and missed (m) during the Grand Challenge: Run 1: m-p; Run 2: p-p-p-p (the blue and the red disk had to be put on the ground, because we called a reset. Each disk is attempted to be picked twice later); Run 3: p-m-m-p (the yellow disk was picked during a reset and had to be put back on the cart); Run 4: m-m; Run 5: p-p-m-m-p-m. Total time airborne is 624 s.

due to one missing MAV, we reconfigured the system to use only two MAVs in a reset. After a second reset due to another hardware problem, the remaining MAV operated on the whole arena. We successfully picked nine discs and were able to deliver seven of them—six into the drop zone and one into

the drop box. Two discs were still attached to MAVs during a reset and, thus, were lying on the ground after resuming the trial. Overall, we scored 10.5 points and reached a second place in this Grand Challenge subtask. Fig. 7 details our trial. We canceled the second Grand Challenge trial due to severe hardware issues without a score.

VIII. CONCLUSION

Operating complex robotic systems without manual adaptation to the current situation and with virtually no testing time is very challenging. Whereas many highly sophisticated state-of-the-art algorithms to all subproblems of the challenge exist, simpler and failsafe solutions are often key to success. The complexity of the task is represented in the final results: From 18 teams participating in the Treasure Hunt only four were able to autonomously achieve partial task fulfillment. Five more teams were able to deliver at least one object with manual control. We came in third in the Treasure Hunt after the second competition day, and, while winning the Grand Challenge overall—in collaboration with an MAV landing on a moving target and a ground robot operating a valve—we achieved the second highest score in the Treasure Hunt sub-challenge out of 14 participating teams.

We addressed many possible issues in advance; still, unforeseen challenges occur during actual competitions, e.g., the unexpected strong wind and deadlock situations. The system could be robustified by adding more elements of randomness to the algorithms to prevent repetitive failing.

REFERENCES

[1] K. Morton and L. F. Gonzalez, "Development of a robust framework for an outdoor mobile manipulation UAV," in *Proc. of IEEE Aerospace Conference (AEROCONF)*, 2016.

[2] S. Kim, S. Choi, and H. J. Kim, "Aerial manipulation using a quadrotor with a two dof robotic arm," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2013.

[3] V. Ghadiok, J. Goldin, and W. Ren, "Autonomous indoor aerial gripping using a quadrotor," in *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2011.

[4] A. Gawel, M. Kamel, T. Novkovic, J. Widauer, D. Schindler, B. Pfyffer von Altishofen, R. Siegwart, and J. Nieto, "Aerial picking and delivery of magnetic objects with MAVs," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2017.

[5] N. Michael, J. Fink, and V. Kumar, "Cooperative manipulation and transportation with aerial robots," *Autonomous Robots*, vol. 30, no. 1, pp. 73–86, 2011.

[6] A. Tagliabue, M. Kamel, S. Verling, R. Siegwart, and J. Nieto, "Collaborative transportation using MAVs via passive force control," in *Proc. of IEEE Int. Conf. on Robotics and Automation (ICRA)*, 2017.

[7] M. Schwarz, T. Rodehutskors, D. Droeschel, M. Beul, M. Schreiber, N. Araslanov, I. Ivanov, C. Lenz, J. Razlaw, S. Schüller, D. Schwarz, A. Topalidou-Kyniazopoulou, and S. Behnke, "NimbRo Rescue: Solving disaster-response tasks through mobile manipulation robot Momaro," *Journal of Field Robotics*, vol. 34, no. 2, pp. 400–425, 2017.

[8] D. Nistér and H. Stewénius, "Linear time maximally stable extremal regions," in *Proc. of European Conf. on Computer Vision (ECCV)*, 2008.

[9] M. Nieuwenhuisen, J. Quenzel, M. Beul, D. Droeschel, S. Houben, and S. Behnke, "ChimneySpector: Autonomous MAV-based indoor chimney inspection employing 3D laser localization and textured surface reconstruction," in *Proc. of Int. Conf. on Unmanned Aircraft Systems (ICUAS)*, 2017.

[10] M. Beul and S. Behnke, "Fast full state trajectory generation for multirotors," in *Proc. of Int. Conf. on Unmanned Aircraft Systems (ICUAS)*, 2017.