

Dynamically Feasible Motion Planning for Micro Air Vehicles Using an Egocylinder

Anthony T. Fragoso, Cevahir Cigla, Roland Brockers
and Larry H. Matthies

Abstract Onboard obstacle avoidance is a challenging, yet indispensable component of micro air vehicle (MAV) autonomy. Prior approaches for deliberative motion planning over vehicle dynamics typically rely on 3-D voxel-based world models, which require complex access schemes or extensive memory to manage resolution and maintain an acceptable motion-planning horizon. In this paper, we present a novel, lightweight motion planning method, for micro air vehicles with full configuration flat dynamics, based on perception with stereo vision and a 2.5-D egocylinder obstacle representation. We equip the egocylinder with temporal fusion to enhance obstacle detection and provide a rich, 360° representation of the environment well beyond the visible field-of-regard of a stereo camera pair. The natural pixel parameterization of the egocylinder is used to quickly identify dynamically feasible maneuvers onto radial paths, expressed directly in egocylinder coordinates, that enable finely detailed planning at extreme ranges within milliseconds. We have implemented our obstacle avoidance pipeline with an Asctec Pelican quadcopter, and demonstrate the efficiency of our approach experimentally with a set of challenging field scenarios.

A. T. Fragoso (✉)
California Institute of Technology, Pasadena, CA, USA
e-mail: afragoso@caltech.edu

C. Cigla · R. Brockers · L. H. Matthies
NASA Jet Propulsion Laboratory, California Institute of Technology,
Pasadena, CA, USA
e-mail: cevahircigla@gmail.com

R. Brockers
e-mail: brockers@jpl.nasa.gov

L. H. Matthies
e-mail: lhm@jpl.nasa.gov

1 Introduction

Micro air vehicles possess size, weight, and power (SWaP) constraints on computing resources that require efficient obstacle detection, representation, and avoidance algorithms in order to successfully operate in unknown environments. Accordingly, vision-based techniques using one or more cameras, particularly binocular stereo, have emerged as a popular and successful method for obstacle detection due to their light weight, compactness, and ability to generate dense depth maps of a scene. These advantages have been extended into the internal onboard representation of obstacle environments through the use of egospace data structures [7], such as the disparity images and egocylinders, that are based on the geometry of camera projection, possess a favorable resolution pattern tailored to visual detection, and offer efficient collision-checking and temporal fusion of depth data.

Egospace obstacle representations also simplify reactive obstacle avoidance because their underlying geometry reduces any radially-aligned path to a single pixel, which in turn can be evaluated and selected against depth data using a lightweight pixel scan. Egospace also admits full configuration flat dynamics, of which the quadcopter and commonly-used car and airplane models are examples, and allow all control inputs and vehicle states to be determined uniquely in terms of the trajectories as they appear, in egospace coordinates, in relation to obstacles.

In this paper, we extend reactive egospace trajectory selection for micro air vehicles with negligible dynamics [1] to incorporate full configuration flat dynamics. We present experimental verification on a quadcopter platform in *a priori* unknown environments, using stereo obstacle detection, temporal filtering of depth data, and an egocylindrical obstacle and motion planning representation.

2 Related Work

Traditional approaches for MAV motion planning rely on flatness-based trajectory generation (as in [14]) over a three-dimensional voxel model of the environment. [12] populates a uniform resolution occupancy representation of an unstructured environment, which is used to identify a minimum-jerk trajectory using layered short-range and long-range planners based on convex segmentation followed by a graph search. For a known indoor environment, [17] achieves high-speed flight within a volumetric world model by seeding a polynomial motion planner with the output of a preliminary, dynamics-free RRT* search. Although almost any standard motion-planning method can be used within a voxel representation, uniform grids scale very poorly in size in large outdoor environments and complex access is required to distribute resolution in a more favorable way [9].

Egospace representations, on the other hand, can encode collision-free radial paths using a single pixel and offer an efficient trajectory search space for dynamics-free, reactive motion planning in unstructured environments [1]. Matthies et al.[13]

uses a disparity image representation to efficiently collision-check the segments of a closed loop RRT (CL-RRT) motion planner after a projection, but perform trajectory selection and generation in world coordinates using forward integration of a non-linear plant model at severe computational expense. Although reactive methods in egospace [1] are extremely lightweight, they scale poorly to high speeds because full trajectories are neither known nor collision-checked in advance—even if an instantaneous collision-free action is identified, there is no way to practically verify if it will remain dynamically feasible and trackable in an unknown environment. Deliberative motion planning performed entirely within egospace representations, rather than simply projected into egospace for collision-checking, is introduced and characterized in [7] for robots with configuration flat dynamics—completeness properties are established for the egospace equivalents of motion planners in world coordinates, and motion primitives expressed directly in egospace coordinates extend the advantages of world trajectory libraries to also encompass fast collision-checking and trajectory selection.

Depth data can be obtained from active depth sensors or passive stereo matching. Stereo matching approaches work well both indoors and outdoors and have an adjustable depth range that provides computational flexibility for the design of the perception system. Depth estimation is typically performed on each frame independently, which introduces errors and holes due to environmental factors as well as stereo matching failures. As a result, some degree of temporal fusion is required to propagate reliable data over time and generate a complete and reliable scene representation for later motion planning steps—unfused egospace approaches, such as [1], require additional side-looking cameras to achieve a large field-of-regard, and suffer from flicker and missed-obstacle incidents that result in collisions. Fusion can be performed during the matching step [5, 10, 16], in image space [2, 3, 18] or on 3D voxel representations [4, 8].

In [16] and [10], temporal fusion incorporates temporal data in cost functions during estimation as an extension to spatial aggregation. SLAM techniques [5] also constrain temporal consistency in the estimation framework, where online updates are applied in key frames. In [5], a simple Gaussian model represents depth measurements and limits the search range according to the standard deviation of prior hypotheses. Multi-view filtering is another approach that improves depth maps by removing outliers and compensating holes. In [18], the frames within a specified time window (a fixed number of the most recent frames, for example) are mapped to the most recent frame according to the related depth maps. These hypotheses are merged probabilistically via probability density estimation. Cigla [3] and Cigla et al. [2] use Gaussian Mixture Models (GMM) to represent depth observations in a compact manner, with an online update and propagation approach that decreases computational complexity and memory requirements. These methods tend to produce more reliable depth maps than can be obtained using filter-based fusion approaches.

Three-dimensional models [4, 8] are also commonly used to merge multiple depth map observations and provide temporal fusion. In this regime, depth data is mapped to voxels [4] and accumulated in a surface representation [8] that produces highly accurate maps at the expense of memory usage and computational efficiency.

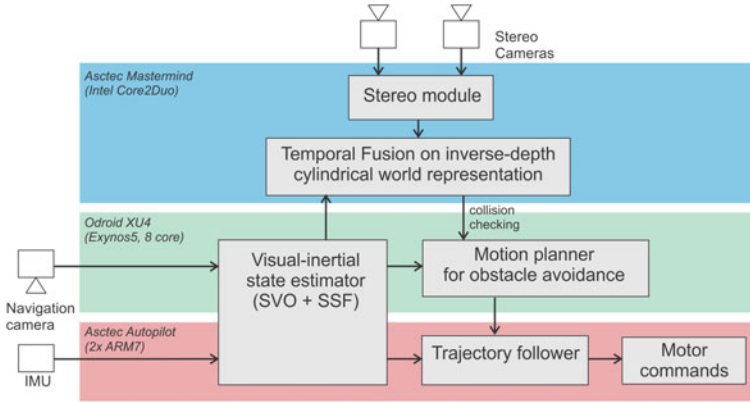


Fig. 1 System architecture for the implementation on an Asctec Pelican quadrotor that is equipped with an Asctec Mastermind and an Odroid XU4 computing board

3 System Overview

The perception, representation, and motion planning modules are implemented within a pipeline architecture onboard an Asctec Pelican quadcopter, with vision and perception tasks performed on an Asctec Mastermind embedded computer and motion planning and state estimation performed on an Odroid XU4 (Fig. 1). After acquiring stereo imagery of a scene using a set of forward-looking cameras, depth data is calculated, temporally filtered, and used to populate an egocylinder structure in which obstacles are artificially expanded in order to abstract the vehicle to a point mass for planning. The expanded egocylinder is then passed to a two-stage motion planner, which first performs a pixel scan to identify a candidate flight path, and then attempts to identify a dynamically feasible, collision-free maneuver onto the candidate path using minimum-jerk paths directly in image space. Once a motion plan is successfully found, a low-level flatness-based controller calculates feedforward inputs and tracks the reference trajectory until a new motion plan is required. Vehicle pose is provided to both the temporal filtering and motion planning modules using visual odometry from a downward-looking camera (SVO [6]) fused with IMU data using the SSF framework [19].

3.1 Perception and Representation

The egocylinder representation of the scene [1] is used to represent the environment for collision-checking, which avoids much of the memory and computation expense used by voxel maps to represent free space. A full 360° field-of-regard is achieved by integrating depth maps forward as the vehicle moves within the environment.

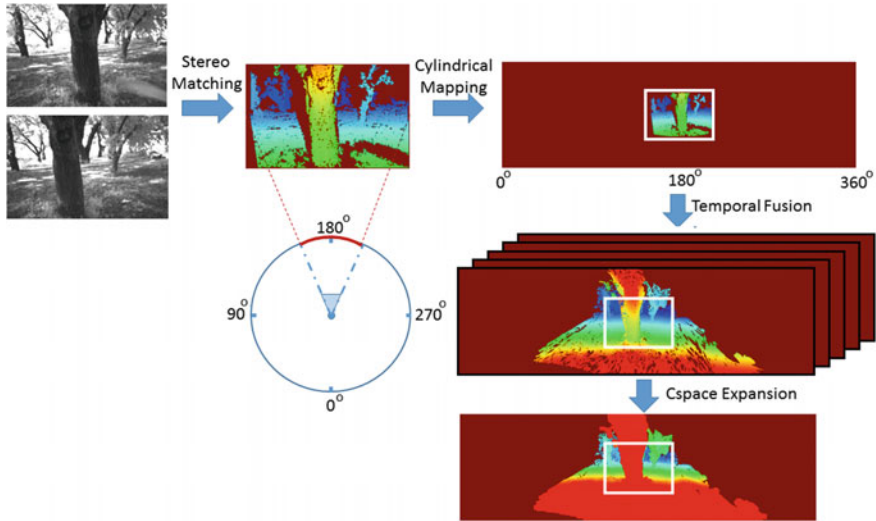


Fig. 2 The visual perception pipeline uses stereo matching to acquire depth data, which is then mapped onto an egocylinder for temporal fusion and C-space expansion directly in egocylinder coordinates. Areas of the egocylinder visible to the stereo cameras are enclosed within a white rectangle for illustration

Visual perception for collision avoidance is based on the approach shown in Fig. 2. The most recent stereo vision-based depth map estimate is first mapped onto an egocylinder representation fixed to the frame of the left camera—the optical center of the left camera maps to the center of the structure, with the forward direction aligned with its viewing axis. The internal representation of the egocylinder is a disparity image which wraps around the cylinder structure, which is called the egocylinder image.

Obstacles are then expanded into configuration space (C-space) [13] by a characteristic radius of the vehicle, directly on the coordinates of the fused egocylinder map, in order to abstract the vehicle to a point mass and simplify path planning. As motion planning proceeds and the vehicle continues to move through the environment, new depth data is observed in the central area of the egocylinder corresponding to the field-of-view of the stereo cameras.

We use GMM-based fusion [2, 3] to efficiently merge the current depth map with past observations directly in image space. In this approach, each pixel is modeled by a mixture of Gaussian distributions, in pixel coordinates (u,v) and disparity d , and propagated to the recent depth map using vehicle pose estimates. The final depth map is obtained by constraining both the standard deviation of the disparity models and an observation count to remove flicker and rely on frequently observed depth values. Cigla et al. [2] demonstrates that GMM-based fusion improves depth quality by removing large depth errors such as false obstacles, which are typically due to errors in stereo matching, and by assigning reliable depth values for empty

pixels in the recent depth map. Temporal fusion in this region of the egocylinder mitigates sensing failures, such as incorrect depth estimates or empty pixels, by constraining the temporally accumulated data. Regions of the egocylinder not visible to the stereo sensors do not receive new depth data, and are instead updated by transferring temporally accumulated models subject to a forgetting factor at each update. As the vehicle moves around and encounters obstacles at a variety of viewing angles, a more complete representation of the scene emerges as temporal data is fused into the structure.

3.2 Motion Planning

The first stage of the motion planning module replicates the approach of [1] in which collision-free, yet dynamically infeasible radial paths are represented by their pixel locations and identified using a scan over the expanded egocylinder structure. During this scan, the distance to obstacles at each pixel is compared to a predetermined planning horizon distance h , which is chosen to maintain a safe time-to-contact towards obstacles at the desired speed and can take values up to the maximum reliable sensor range depending on the expected complexity of the environment. The collision-free pixel that is closest to that of the destination is selected as a candidate path (Fig. 3), and if no safe candidate path is found the process is repeated with a lower desired speed and shorter planning horizon.

In the second stage, a trajectory generation module attempts to identify a feasible maneuver that can smoothly merge the vehicle onto the infeasible candidate path at the desired final velocity V , which greatly narrows the set of admissible trajectories that must be calculated and collision-checked. The straight-line segment that follows is known to be collision-free from the first stage and need not be checked again, and

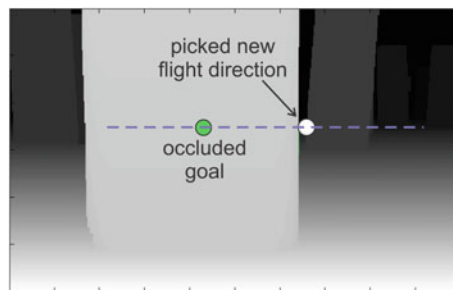


Fig. 3 The first stage of the planner identifies collision-free flight paths by checking the pixels of the egocylinder against a planning horizon chosen to guarantee a minimum time-to-contact. A goal is projected into the egocylinder (green, occluded) and the nearest collision-free pixel (white) is selected as a candidate path and passed to the trajectory generator. Here, brighter pixels are closer to the camera

the required acceleration, velocity, and position are automatically specified at the beginning and end of the turning maneuver (for position, up to an undetermined range):

$$\begin{aligned} \ddot{\mathbf{x}}(0) &= \mathbf{a}_0, & \dot{\mathbf{x}}(0) &= \mathbf{v}_0 & \mathbf{x}(0) &= \mathbf{x}_0, \\ \ddot{\mathbf{x}}(t_f) &= 0, & \dot{\mathbf{x}}(t_f) &= V\hat{\lambda} & \mathbf{x}(t_f) &= (\lambda h)\hat{\lambda}. \end{aligned}$$

Here, the final flight direction $\hat{\lambda}$ is chosen in the first planning stage, t_f is the total maneuver time, and $\lambda \in (0, 1]$ is a free parameter that modulates the distance to the end of the maneuver.

The set of possible trajectories is further narrowed by insisting on a minimum jerk maneuver, which can be shown to consist of fifth degree polynomials $(x(t), y(t), z(t)) = \sum_{i=0}^5 \mathbf{a}_i t^i$ using a straightforward application of the Pontryagin minimum principle. Once the endpoints are prescribed, the maneuver has only two remaining free parameters, t_f and λ , that can either be modulated independently or constrained to each other to produce a one-parameter family of admissible trajectories—for example, the total time can be chosen to maintain an average vector velocity associated with the candidate path (fixed speed, fixed direction) such that $t_f = (\lambda h)/V$. Minimum-jerk polynomials can be shown by direct substitution to satisfy the differentially flat quadcopter plant model of [11],

$$\begin{aligned} \ddot{x} &= (\sin \phi \cos \psi - \cos \phi \sin \theta \sin \psi) \frac{u_{\text{collective}}}{m}, \\ \ddot{y} &= (-\sin \phi \sin \psi - \cos \phi \sin \theta \cos \psi) \frac{u_{\text{collective}}}{m}, \\ \ddot{z} &= -g + (\cos \phi \cos \theta) \frac{u_{\text{collective}}}{m}, \\ \ddot{\phi} &= u_{\text{roll}}, \\ \ddot{\theta} &= u_{\text{pitch}}, \\ \ddot{\psi} &= u_{\text{yaw}}, \end{aligned}$$

where (ϕ, θ, ψ) are roll-pitch-yaw Euler angles and $u_{\text{collective}}$ is the collective motor thrust. The trajectories $(x(t), y(t), z(t))$ correspond to unique control inputs, subject to actuator saturation and a yaw angle trajectory that can be chosen independently of the translation of the vehicle, that are determined algebraically using the associated flatness relations.

The coefficients of the minimum-jerk polynomials are determined symbolically, in advance, as a solution of a small linear system of equations in terms of the initial and final vehicle states and the search parameter λ . This solution step dramatically reduces the amount of computation that needs to be performed onboard and allows for a large number of potential trajectories to be considered during a planning cycle (see Sect. 4 for runtime statistics). Evaluation for control saturation and collisions is enhanced by a simultaneous dual representation of the trajectories (Fig. 4) in world coordinates and egocylinder coordinates, which both have a simple closed form and are determined by the single search parameter—the world coordinate trajectories

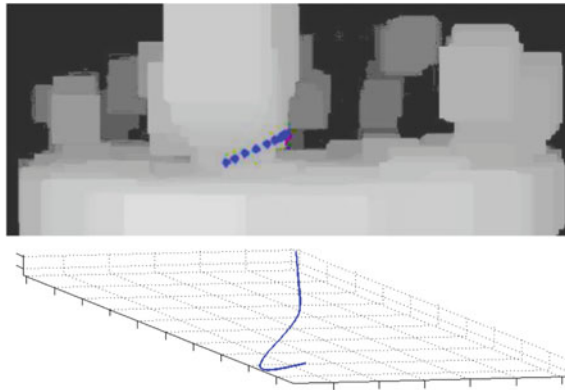


Fig. 4 Closed-form trajectories are simultaneously maintained in egocylinder coordinates for collision-checking and world coordinates for control. Here, a trajectory (top, blue dots proceeding towards the right) has been selected to avoid a tree (brighter pixels are closer) using a pixel scan and collision-check in egocylinder coordinates. The trajectory is tracked in world coordinates (blue line, bottom) by the flatness-based controller

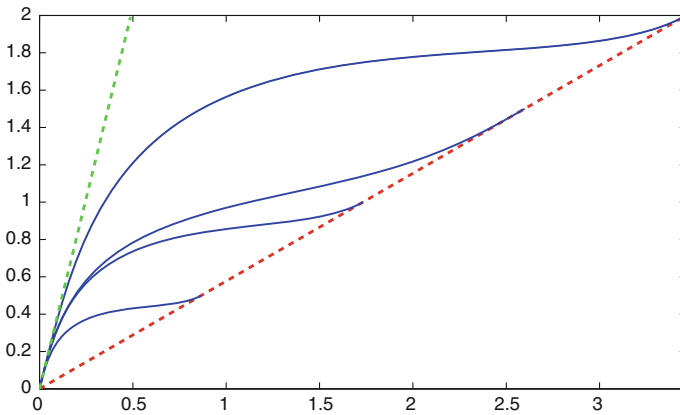


Fig. 5 The trajectory generator, shown in 2-D for clarity, attempts to merge the vehicle smoothly from the green heading (left, dashed) to the red heading (right, dashed) using minimum-jerk polynomials (blue) that are parametrized by the distance covered before merging. By decreasing the length of the maneuver, the trajectory generator can provide a more aggressive turn if a collision is detected

can be easily checked for saturation and converted into control inputs, while the egocylinder coordinates can be immediately collision-checked using a pixel-by-pixel comparison against the egocylinder data structure.

The trajectory search itself begins by attempting to merge the vehicle onto the candidate radial path at a distance corresponding to the planning horizon (Fig. 5). The proposed maneuver is collision-checked, evaluated against the control constraints, and returned if successful—otherwise, the distance to the end of the maneuver is

decreased along with the total time until either a valid trajectory is identified or the controls saturate, in which case either the total time is increased independently or a different candidate radial path is attempted. Once the trajectory is available, a low level controller tracks the reference using a standard two degree-of-freedom feed-forward design [15], with nested position/velocity and attitude loops corresponding to the slow and fast dynamics of the vehicle model.

Our receding-horizon approach does differ from an RRT-type formulation, like that of [17], in that it lacks a completeness guarantee and can become stuck in dead ends. For obstacle avoidance in unstructured environments where frequent replanning is required, however, RRT-type methods can bottleneck the pipeline [13] and limit flight speeds. Receding-horizon formulations such as ours, however, do share some degree of deliberative action with complete approaches and offer forward collision-checking unavailable to reactive approaches such as [1].

4 Results

The visual perception pipeline is implemented on the Asctec Mastermind embedded computer equipped with a 1.86 GHz Intel Core2Duo processor. The forward-looking stereo cameras are installed with a baseline of 25 cm, and frame-wise stereo disparity maps (376×240) are calculated by block matching with a search range of 100 pixels. The resolution of the egocylinder image is 660×200 . The full pipeline maintains a 10Hz update rate using both cores of the processor, with computation times for each step in the visual perception pipeline given in Table 1.

Typical results for temporal fusion on the egocylinder are illustrated in Fig. 6. The left stereo image, unfused disparity map, and the corresponding egocylinder images are shown sequentially as the vehicle approaches an obstacle. Temporal fusion yields a more complete scene representation compared to the unfused disparity maps—known regions that would be otherwise discarded at each frame are instead retained in the egocylinder. The propagation process is particularly important for obstacles at close range, which eventually become invisible to stereo matching due to a limited search range and would otherwise present a missed-detection hazard (Fig. 7).

Table 1 Onboard computation times for each step in the visual perception pipeline

Step	Time (ms)
Stereo matching	100
Cylindrical mapping	14.4
Temporal fusion	52.6
C-space expansion	4.5

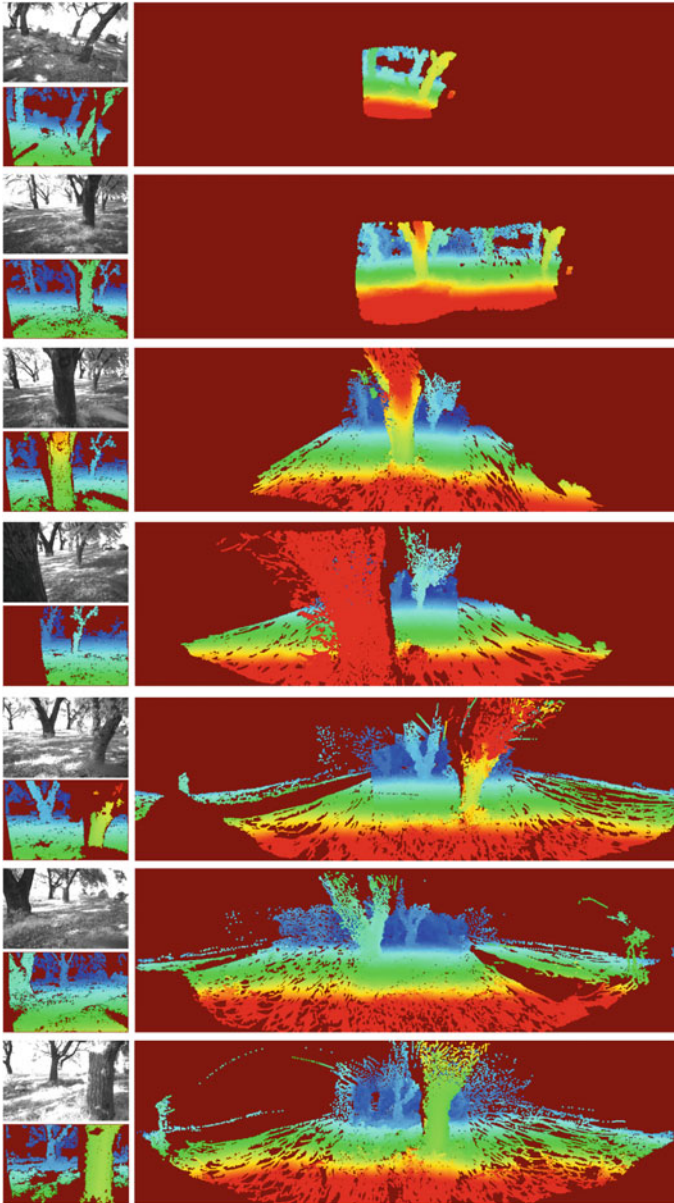


Fig. 6 The perception pipeline was tested in a forested environment. In the first column are raw images (grayscale) from the left stereo camera and the unfused disparity map, and in the second are full egocylinder maps with temporal fusion. Here, warmer colors are closer, cooler colors are farther, and the maroon background represents no data. In the fourth row, temporal fusion prevents a collision at close range (first column) by propagating obstacle data forward from an earlier detection (second column)

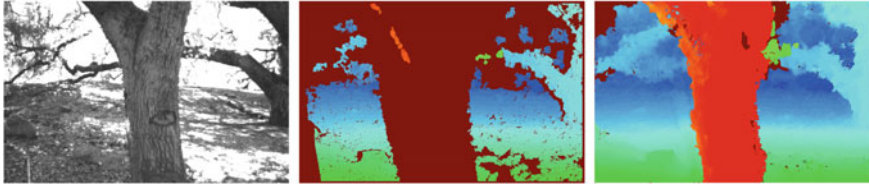


Fig. 7 While the maximum disparity calculated by the stereo algorithm is too low to resolve the nearby obstacle, temporal fusion propagates the previously observed obstacle, which is then used in the motion planner. Left: left rectified image, Middle: stereo disparity map, Right: temporally fused disparity map

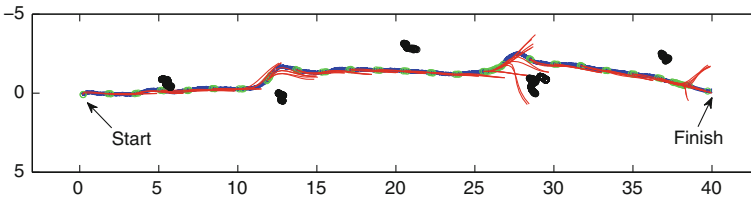


Fig. 8 While navigating through a forested environment (tree trunks at 2m height shown as black point clouds), the vehicle took evasive maneuvers around three trees and successfully reached a predetermined destination 40m away. The path (blue, vehicle travels from left to right) consists of segments of dynamically feasible trajectory segments (red arcs) calculated using the two-step procedure of Sect. 3 and followed in a receding-horizon fashion. Axis scale is in meters

Motion planning experiments were conducted at a number of challenging sites, including a 40 m course through a forested area, a small dead-end alcove in a built-up area, and an open, obstacle-free wash with distant buildings.

While maintaining a speed of 1 m/s, the vehicle encountered and successfully avoided three trees in the forest using the minimum-jerk trajectories of Sect. 3 executed in a receding-horizon fashion (Fig. 8). This environment required frequent replanning as new obstacles became visible, which was performed efficiently with average runtimes on the Odroid XU4 listed in Table 2.

The egocylinder also proved highly useful in preventing the vehicle from becoming stuck in confined areas when paired with temporal propagation of depth data. After exploring a small alcove and discovering a dead end (Fig. 9), the motion planner was able to extricate the vehicle towards an open area to its left that was invisible to the stereo pair, but already stored within the egocylinder structure and available to the motion planner.

Table 2 Onboard computation times for each step in the motion planning pipeline

Step	Time (ms)
Egocylinder scan	4.0 ^a
Trajectory generation	0.02
Collision-checking	0.2

^aFor a complete replan (motion planner can reuse old targets)

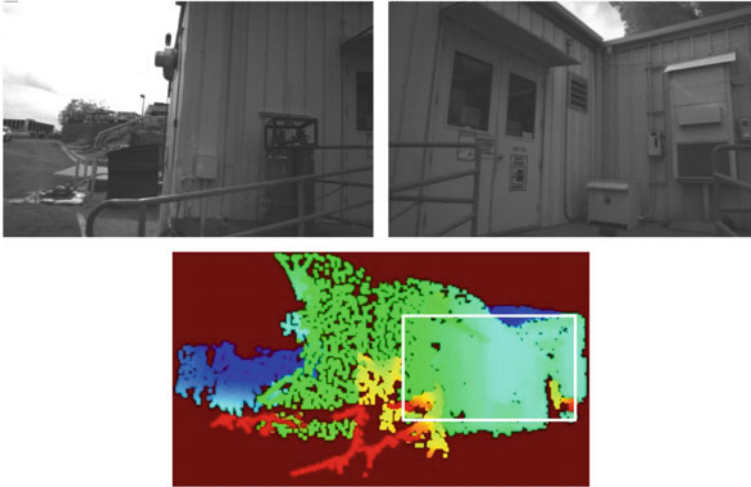


Fig. 9 When combined with temporal fusion, the egocylinder facilitates exploration of confined areas by propagating knowledge of open areas when they become invisible to the stereo camera. Here, the vehicle has reached a dead end within an alcove (top, raw images shown), but continues to represent an exit to its left within the egocylinder (blue area, bottom left) beyond the visible field-of-regard of unfused stereo (white rectangle). An escape route is also available by climbing over the building, but was excluded by a mission-level altitude ceiling

By identifying a restricted trajectory search space and exploiting the natural resolution pattern of the egocylinder, the motion planning module was also able to efficiently generate detailed, dynamically feasible trajectories at ranges near the limits of reliable stereo detection. When presented with a destination waypoint inside a parking garage over 40 m away, within 5 ms the motion planner identified a 5 m × 2 m opening into the garage as the best path towards the target and calculated a collision-free and dynamically feasible path into the building (Fig. 10).

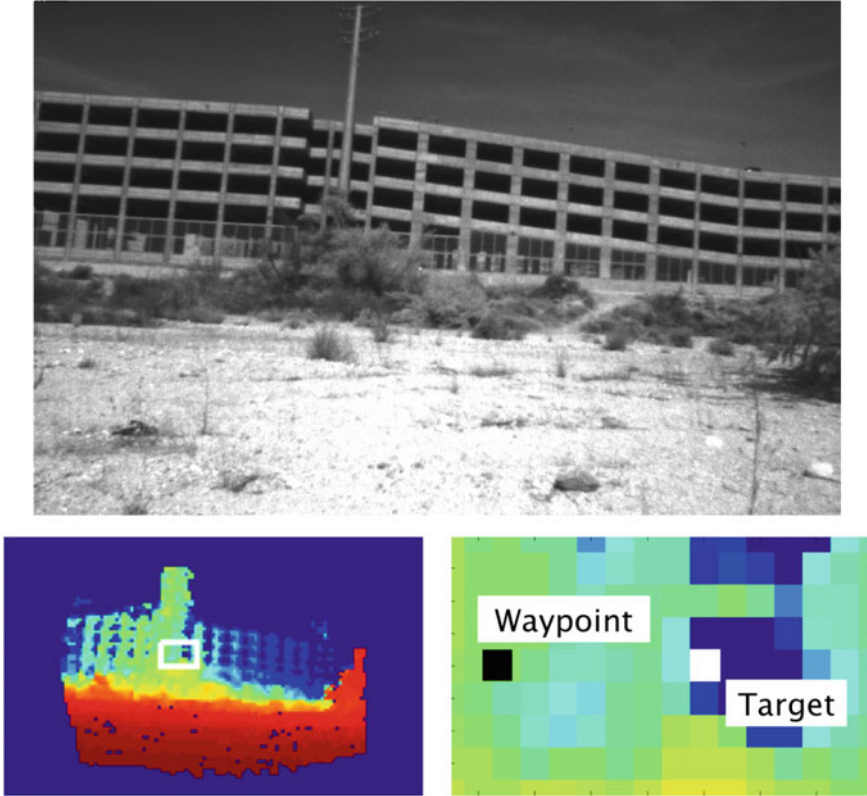


Fig. 10 Egospace representations are particularly well-suited for producing finely detailed trajectories at extreme ranges. Here, the motion planner has been assigned a waypoint behind a pillar inside a parking garage 45 m away (raw image, top). After generating and expanding an egocylinder representation (original, bottom left, zoomed for clarity, right), the motion planner projects the waypoint (black square, bottom right) into the image and identifies the closest collision-free target (white square, bottom right) as a candidate flight path for the trajectory generator

5 Conclusion

We have demonstrated an obstacle avoidance pipeline over the full dynamics of a micro air vehicle using a temporally fused egocylinder representation. In addition to being simple to access and compact to store, the egocylinder representation admits an extremely lightweight trajectory selection procedure, based on a parametrization of flight paths by pixels and differential flatness, that can quickly generate dynamically feasible turn maneuvers towards targets at the range limit of stereo sensing. This advantage is made more reliable and flexible with temporal fusion of obstacle data within the egocylinder geometry, which improves the accuracy of sensed data and

accumulates a 360° representation with a single set of stereo cameras that is useful for navigation in confined and cluttered environments.

Our perception framework also addresses a number of weaknesses of stereo vision, primarily due to the limited and fixed baseline of MAV stereo applications, and significantly increases its power as an obstacle detection tool. The resolution pattern of the egocylinder, which decreases favorably with distance without additional overhead, allows stereo data to inform motion planning and remain useful even at long ranges where it is less accurate. Detection errors at close ranges, which eventually become inevitable for raw stereo due to a limited field of view and disparity search, are prevented by propagation and storage of sensed data with real-time temporal fusion.

Acknowledgements This work was funded by the Army Research Laboratory under the Micro Autonomous Systems & Technology Collaborative Technology Alliance program (MAST-CTA). JPL contributions were carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

References

1. Brockers, R., et al.: Vision-based obstacle avoidance for micro air vehicles using an egocylindrical depth map. In: *ISER* (2016)
2. Cigla, C., Brockers, R., Matthies, L.: Image-based visual perception and representation for collision avoidance. In: *IEEE International Conference on Computer Vision and Pattern Recognition, Embedded Vision Workshop* (2017)
3. Cigla, C., et al.: Gaussian mixture models for temporal depth fusion. In: *IEEE Winter Conference on Applications of Computer, Vision* (2017)
4. Dryanovski, W., et al.: Multi-volume occupancy grids: an efficient probabilistic 3d mapping model for micro aerial vehicles. In: *IROS* (2010)
5. Engel, J., et al.: Semi-dense visual odometry for a monocular camera. In: *2013 IEEE International Conference on Computer Vision*, pp. 1449–1456. IEEE (2013)
6. Forster, C., Pizzoli, M., Svo, S.D.: Fast semi-direct monocular visual odometry. In: *IEEE International Conference on Robotics and Automation (ICRA)* (2014)
7. Fragoso, A., Matthies, L., Murray, R.: A fast motion planning representation for configuration flat robots with applications to micro air vehicles. In: *ACC* (2017)
8. Hane, C., et al.: Stereo depth map fusion for robot navigation. In: *IROS* (2011)
9. Hornung, A., et al.: Octomap: an efficient probabilistic 3D mapping framework based on octrees. *Auton. Robot.* **34**(3) (2013)
10. Hosni, A., et al.: Temporally consistent disparity and optical flow via efficient spatio-temporal filtering. In *Pacific Rim Symposium on Image and Video Technology* (2011)
11. How, J., et al.: Real-time indoor autonomous vehicle test environment. *IEEE Control Syst. Mag.* **28**(2) (2008)
12. Liu, S., et al.: High speed navigation for quadrotors with limited onboard sensing. In: *ICRA* (2016)
13. Matthies, L., et al.: Stereo vision-based obstacle avoidance for micro air vehicles using disparity space. In *International Conference on Robotics and Automation (ICRA)* (2014)
14. Mellinger, D., et al.: Minimum snap trajectory generation and control for quadrotors. In: *ICRA* (2011)
15. Powers, C., et al.: Quadrotor kinematics and dynamics. In: *Handbook of Unmanned Aerial Vehicles*. Springer, Netherlands (2015)

16. Richardt, C., et al.: Real-time spatiotemporal stereo matching using the dual-cross-bilateral grid. In: *European Conference on Computer Vision* (2010)
17. Richter, C., et al.: Polynomial trajectory planning for aggressive quadrotor flight in dense indoor environments. In: *ISRR* (2013)
18. Unger, C., et al.: Probabilistic disparity fusion for real-time motion stereo. *Mach. Vis. Appl.* **25** (2011)
19. Weiss, S., et al.: Monocular vision for long-term micro aerial vehicle state estimation: a compendium. *J. Field Robot.* **30**(5) (2013)