Stereo Vision-based Obstacle Avoidance for Micro Air Vehicles using an Egocylindrical Image Space Representation

R. Brockers[a], A. Fragoso[b], and L. Matthies[a]

[a]Jet Propulsion Laboratory, Caltech, Pasadena, CA
[b]Graduate Aerospace Laboratories, Caltech, Pasadena, CA

ABSTRACT

Micro air vehicles which operate autonomously at low altitude in cluttered environments require a method for on-board obstacle avoidance for safe operation. Prior approaches can be divided between purely reactive approaches, mapping low-level visual features directly to headings to maneuver the vehicle around the obstacle, and deliberative methods that use on-board 3-D sensors to create a 3-D, voxel-based world model, which is then used to generate collision free 3-D trajectories. In this paper, we use forward-looking stereo vision with a large horizontal and vertical field of view and project range from stereo into a novel robot-centered, cylindrical, inverse range map we call an *egocylinder*. With this implementation we reduce the complexity of our world representation from a 3D map to a 2.5D image space representation, which supports very efficient motion planning and collision-checking. Configuration space expansion is done very efficiently on the egocylinder as an image processing function. Deploying a fast reactive motion planner directly on the configuration space expanded egocylinder image, we demonstrate the effectiveness of this new approach experimentally in an indoor environment.

Keywords: Micro air vehicles, obstacle avoidance, vision, egocylinder

1. INTRODUCTION

Micro air vehicles (MAVs) require on-board obstacle detection and avoidance systems with minimal size, weight, power, complexity, and cost, using sensors with a large field of view for maneuvering in cluttered spaces. Vision-based approaches have excellent potential to address these needs for many applications. In prior work,[1] we used stereo vision for forward-looking depth perception and showed that stereo disparity maps can be used for MAV motion planning, with configuration space (C-space) obstacle expansion done in image space. In this approach, we generated dynamically feasible trajectories in 3-D Cartesian space, which were evaluated by projecting candidate trajectories into image space to determine whether they intersect obstacles. This is a very efficient approach to geometric representation and collision checking, and the overall approach is quite effective where the goal is obstacle avoidance rather than mapping.

In this paper, we exchange the planar representation we used before with a cylindrical inverse range image we call an *egocylinder*, and perform C-space expansion directly on the egocylinder. This novel representation has the advantage of a constant angular resolution around the vehicle body frame supporting directional motion planning in a natural way. It also allows for quick integration of additional range sensors with a peripheral field of view (e.g. side looking stereo cameras), and temporal propagation to augment the field of regard with previous observations.

To reduce the computational cost of motion planning for low-speed flight, we implement a simple reactive method to select directions toward more distant goals that stay within free-space shown by the egocylinder.

Further author information: (Send correspondence to R. Brockers)
R. Brockers: E-mail: brockers@jpl.nasa.gov
A. Fragoso: E-mail: afragoso@caltech.edu
L. Matthies: E-mail: lhm@jpl.nasa.gov

2. RELATED WORK

There are a number of papers that discuss advantages and disadvantages of various passive and active sensor options for MAV obstacle avoidance,[1,2] with recent examples of MAV systems that use multiple types of sensors.[3,4]

Vision-based approaches can be characterized based on how perception, scene representation, and motion planning and control are implemented. Common approaches use either optical flow, monocular or stereo depth perception, or a combination of these. Optical flow based methods typically incorporate reactive control algorithms that maneuver the vehicle based on the optical flow input. Control algorithms for provably stable wall-following and corridor-following behavior have been developed this way;[5] however, navigation that requires a discrete choice among alternate directions requires higher-level perception or reasoning. Machine learning methods have been deployed to map optical flow and other monocular visual features into reactive obstacle avoidance behaviors,[6] but it is difficult to generalize this approach to work in a wide variety of conditions. The majority of MAV obstacle detection use depth perception. Methods using monocular depth perception from a structure from motion (SfM) approach have been used for MAVs,[7,8] but require aircraft motion to measure depth and have poor depth perception near the focus of expansion. Stereo vision overcomes these limitations, works well in many outdoor settings, and small, fast stereo implementations are progressing.[9–11]

The predominant implementation for scene representation has been 2-D or 3-D Cartesian probabilistic grid maps, which can be used with motion planning algorithms that vary from reactive to deliberative and from local to global.[3,10,12,13] These methods are particularly useful for mapping, exploration, or obstacle avoidance in areas that require memory of previously examined areas. However, grid maps in general require a lot of storage and computation, and are hence undesirable for small platforms. If obstacle avoidance is the main focus for the motion planner, less expensive representation and planning algorithms are possible. Representations that are polar in nature, which match the polar angular resolution of the depth sensors[14,15] can be used to reflect the emphasize on objects in the proximity of the vehicle. There has been work[16] that used the advantage of such a polar perspective implementation by fusing depth data from two on-board stereo pairs in a cylindrical inverse range map centered on the vehicle. This work introduced C-space obstacle expansion of an image space depth map, though in a limited fashion based on an assumed ground plane. In our previous work,[1] we generalized the C-space expansion to be based on the actual depth at each pixel and developed the first combination of an image space depth representation with a dynamics-aware motion planner; feasible trajectories were generated in 3-D and projected into image space to do collision checking testing directly in image space with the C-space expanded depth map. This approach to obstacle representation and collision checking is fast and showed good potential in experiments; however, the image space representation was limited to a fronto-parallel planar representation and the CL-RRT motion planning algorithm was computationally expensive.

3. TECHNICAL APPROACH

With future extensions in mind that include the extension of the field of regard by adding side ways looking cameras, we introduce a cylindrical image representation that allows fusing range inputs from multiple range sources onto one common image-based, polar-perspective, cylindrical world representation (*egocylinder*) (Fig. 1).

3.1 Egocylinder world representation

After a set of new images has been acquired, a block matching stereo algorithm[17] is applied to the images from the stereo cameras, and the resulting disparity image is projected into the egocylinder world representation.

The egocylinder is defined as an image-based, inverse-range representation that is centered on the MAV. It can be visualized as a disparity map that is wrapped around the vehicle (cylinder surface image) (Fig. 1).

To project world points into the cylinder surface image, we model the egocylinder with an origin C within the vehicle body frame, a forward direction vector A, and a horizontal and vertical vector H and V:

$$H = [f_{hp}, 0, h_c] \quad (1)$$

$$V = [0, f_{vp}, v_c] \quad (2)$$

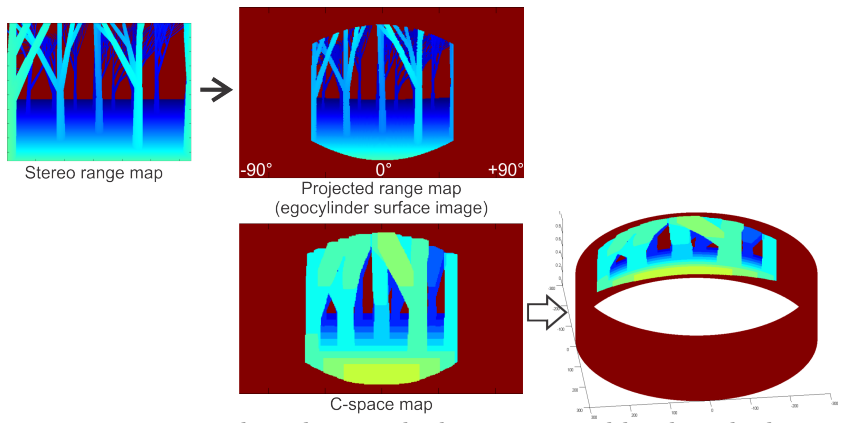

Figure 1: Stereo disparity map projected on the *egocylinder*, represented by the cylinder surface image. C-space expansion takes place as an image processing function.

h_c and v_c are the pixel coordinates of the forward direction within the cylinder surface image, whereas f_{hp} and f_{vp} are the horizontal and vertical focal length in pixel, needed for the projection of a world point P_b in the vehicle body frame onto the egocylinder ($P_b \mapsto p_{cyl}$):

$$P_{cyl} = R_A(P_b - C) \tag{3}$$

$$\hat{P}_{cyl} = P_{cyl}/\sqrt{P_{cyl_x}^2 + P_{cyl_z}^2} \tag{4}$$

P_{cyl} is P_b expressed in cylinder world coordinates with the origin at the center of the cylinder, where the z-axis is pointing in the viewing direction and the y-axis is pointing in the direction of the cylinder. $\hat{P}_{cyl}$ is the projection of P_{cyl} onto a unit cylinder.

Assuming that the cylinder surface image fully wraps completely around the cylinder, the horizontal and vertical focal length and the center pixel coordinates h_c, v_c that correspond to the viewing direction are defined by the image dimensions:

$$f_{vp} = f_{hp} = 1/tan(2 * \pi/cols) \tag{5}$$

$$h_c = cols/2; v_c = rows/2 \tag{6}$$

$$p_{cyl} = \begin{pmatrix} atan(\hat{P}_{cyl_x}/\hat{P}_{cyl_z}) * f_{hp} + h_c \\ \hat{P}_{cyl_y} * f_{vp} + v_c \end{pmatrix} \tag{7}$$

Which finally allows to determine p_{cyl}, the image coordinates of the projection of P_b into the cylinder surface image.

For calculating a corresponding disparity value within the cylinder surface image, we retain a conversion factor fxb within the cylinder model that is equivalent to the product of focal length and baseline of the stereo cameras. Note, that we only consider the radial distance within the x/z plane of the cylinder world coordinate frame for conversion into the egocylinder disparity d:

$$d = fxb/\sqrt{P_{cyl_x}^2 + P_{cyl_z}^2} \tag{8}$$

3.2 C-space expansion

C-space expansion is implemented as an image processing function applied directly to the cylinder surface image. This follows the approach in[1] with the difference that disparity measures the radial distance of a world point as described above.

3.3 Motion planning

The C-space expanded egocylinder allows the aircraft to be treated as a single point in space for collision checking. Currently, we are evaluating the innovations in the perception system with a simple planning algorithm that will be safe if there are no major perceptual errors. At low, constant velocity, we make the simplifying assumption that the vehicle can turn with negligible distance traveled, so the velocity can be transformed into a planning horizon, for example based on stopping distance estimates. This horizon can be checked against the inverse range at any pixel in the egocylinder. Given a goal direction, a simple egocylinder search produces a flight direction that heads closest to the goal while avoiding collisions within the planning horizon (Figure 2). This approach is very fast, safe, and allows us to focus on evaluating perception; ongoing work is developing a more sophisticated image space motion planning algorithm.

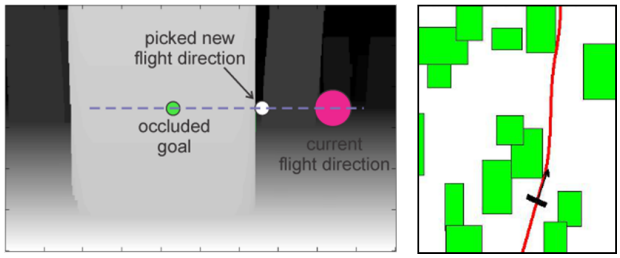

Figure 2: Reactive motion planner: Left: Selected flight direction to avoid obstacle; Right: Simulated flight through cluttered environment

4. RESULTS

We conducted experimental trials in an indoor environment to test our approach. This allowed us to place an obstacle within the flight path of the vehicle and gather ground truth data for obstacle location and vehicle trajectory from a VICON motion tracking system. For vehicle control, we used a state estimation filter that fused VICON pose information with high-rate IMU data.

Figure 3: Asctec Hummingbird with stereo camera head and Odroid XU4 flight computer.

We implemented our approach on an Asctec Hummingbird that was equipped with an Odroid XU4 flight computer (Figure 3), which hosts an 8 core Exynos5422 cell phone based system on a chip (SoC). Our stereo camera setup uses two Matrix Vision BlueFox-MLC200wG cameras (752x480, gray scale) that are hardware synchronized. All software was running on-board the Odroid XU4 using ROS for message passing. The vision pipeline including stereo processing, projection onto the egocylinder, and C-space expansion was executed at 5Hz with an image resolution of 384x240.

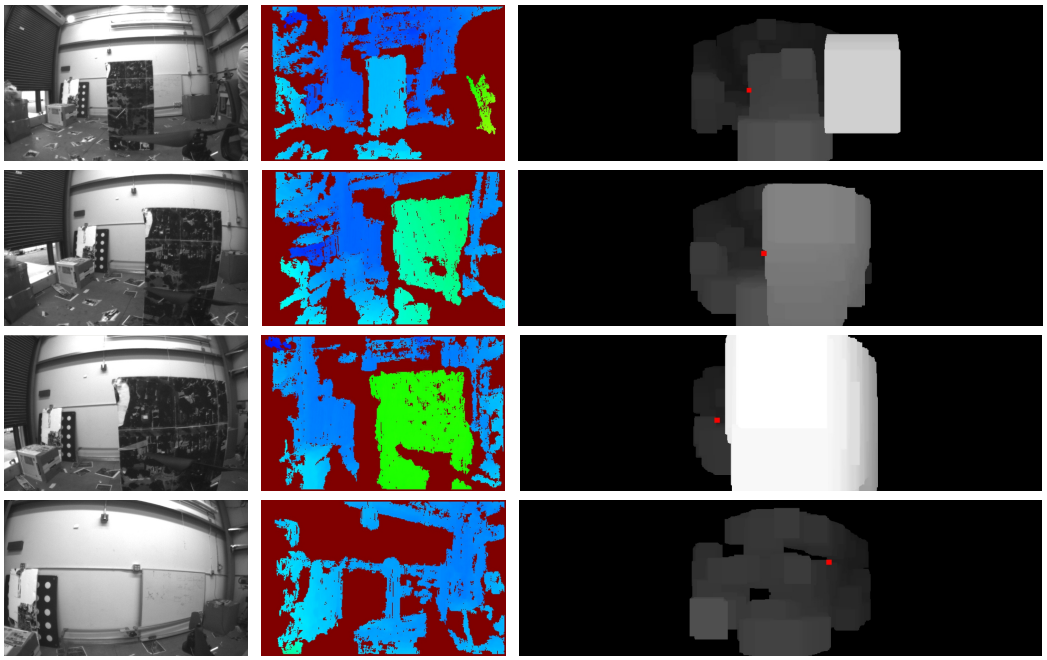

Figure 4: Visual Processing pipeline during obstacle flyby. The vehicle was commanded to fly to a goal behind the rectangular obstacle in the middle of the top left image. Left: Raw images from left stereo camera (reference camera); Middle: false-color stereo disparity maps (closer objects have warmer colors); Right: Egocylinder images, which spans from 0-360 degrees around the vehicle (closer objects are labeled with brighter colors). Selected flight direction is labeled in red.

Figure 4 illustrates the visual processing pipeline during an avoidance maneuver. Initially the vehicle is commanded to fly towards a goal behind a large rectangular obstacle. While it approaches the obstacle, it selects a flight path to the left of the obstacle for avoidance, and once the avoidance maneuver was completed and the obstacle was no longer in view, the vehicle turned back to the goal location. Figure 5 illustrates the executed

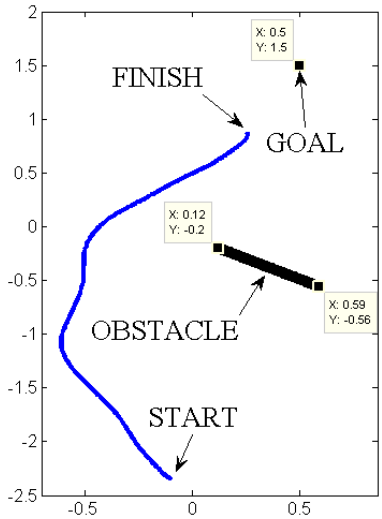

Figure 5: Indoor obstacle avoidance experiment in a controlled VICON environment for ground truth. Blue: Trajectory the Hummingbird flew, Black: obstacle within flight path.

trajectory around the obstacle. Note, that the goal is declared as reached when the vehicle is in within a certain distance to the given goal waypoint.

To illustrate the computational efficiency of our approach, we summarized the processor and memory usage in Table 1 and Table 2. Overall processor usage for all user processes is 28.3 %. More than half of the processing resource is still available. The most resource consuming process is in fact our block matching stereo algorithm, whereas the collision validation part of the motion planner uses only minimal resources.

Table 1: Total processor usage and idle percentage.

	Total usage
Processor user space processes	28.3%
Processor idle	54.8%

Table 2: Processor usage percentage per process of the total 800% available. Note, that the percentage does not state on which core the process is running.

Process	**Processor Usage**	**Memory Usage**
JPLV stereo	136%	0.5%
Reactive Planner: Image Search	58%	48%
Egocylinder representation	43.2%	0.2%
State estimation filter	34.7%	13.3 %
Asctec Autopilot communication	27.8%	0.1%
Camera drivers	31.4%	0.4%
Reactive Planner: Flight direction validation	4.3%	0.1%
total	335.4%	62.6%

5. CONCLUSION

The egocylinder representation is a computational efficient implementation of a world representation for local motion planning. It replaces conventional 3D world representations with a compact 2.5D image-based representation, which only requires - in combination with an inverse range coding of object distance - a fixed, relatively small memory footprint at any time. It can accommodate very effectively future integration of additional range sensors to extend the field of regard (e.g. stereo cameras), while still allowing to execute C-space expansion and collision checking as a fast image processing function.

The disparity-space reactive planner extends the advantages of the C-space expansion method and egocylinder to the planning regime. Overall, this choice of representation demonstrates decreased planning latency and complexity compared to world-coordinate methods.

We implemented our obstacle avoidance framework on a small micro air vehicle using a credit card size SoC flight computer, demonstrating that our approach is suitable to run on very small embedded hardware with significant SWaP constraints.

Acknowledgments

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] Matthies, L., Brockers, R., Kuwata, Y., and Weiss, S., "Stereo vision-based obstacle avoidance for micro air vehicles using disparity space," in [*2014 IEEE International Conference on Robotics and Automation (ICRA)*], 3242–3249 (May 2014).

[2] Kendoul, F., "Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems," *Journal of Field Robotics* **29**(2), 315 – 378 (2012).

[3] Droeschel, D., Nieuwenhuisen, M., Beul, M., Holz, D., Stuckler, J., and Behnke, S., "Multi-layered mapping and navigation for autonomous micro air vehicles," *Journal of Field Robotics* (2015).

[4] Nuske, S., Choudhury, S., Jain, S., Chambers, A., Yoder, L., Scherer, S., Chambelain, L., Cover, H., and Singh, S., "Autonomous exploration and motion planning for an unmanned aerial vehicle navigating rivers," *Journal of Field Robotics* **32**(8), 1141 – 1162 (2015).

[5] Keshavan, J., Gremillion, G., Alvarez-Escobar, H., and Humbert, J. S., "Autonomous vision-based navigation of a quadrature in corridor-like environments," *International Journal of Micro Air Vehicles* **7**(2), 111 – 123 (2015).

[6] Del, D. and et al, "Vision and learning for deliberative monocular cluttered flight," in [*10th Conference on Field and Service Robotics*], (2015).

[7] Alvarez, H., Paz, L. M., Sturm, J., and Cremers, D., "Collision avoidance for quadratures with a monocular camera," in [*Experimental Robotics: the 14th International Symposium on Experimental Robotics*], (2015).

[8] Daftry, D., Dey, D., Sandhawalia, H., Zeng, S., Bagnell, J. A., and Hebert, M., "Semi-dense visual odometry for monocular navigation in cluttered environments," in [*IEEE Intl. Conf. on Robotics and Automation, Workshop on Recent Advances in Sensing and Actuation for Bioinspired Agile Flight*], (May 2015).

[9] Goldberg, S. B. and Matthies, L., "Stereo and IMU assisted visual odometry on an OMAP3530 for small robots," in [*IEEE Conf. on Computer Vision and Pattern Recognition, Workshop on Embedded Computer Vision*], (2011).

[10] Schmid, K., Lutz, P., Tomic, T., Mair, E., and Hirschmuller, H., "Autonomous vision-based micro air vehicle for indoor and outdoor navigation," *Journal of Field Robotics* **31**(4), 537 – 570 (2014).

[11] Kuhn, M., Moser, S., Isler, O., Gurkaynak, F. K., Burg, A., Felber, N., Kaelin, H., and Fichtner, W., "Efficient ASIC implementation of a real-time depth mapping stereo vision system," in [*IEEE 46th Midwest Symposium on Circuits and Systems*], (2003).

[12] Shen, S., Michael, N., and Kumar, V., "3d indoor exploration with a computationally constrained mav," in [*Robotics: Science and Systems*], (2011).

[13] Fraundorfer, F., Heng, L., Honegger, D., Lee, G. H., Meier, L., Tanskanen, P., and Pollefeys, M., "Vision-based autonomous mapping and exploration using a quadrotor MAV," in [*IROS*], (2012).

[14] Bajracharya, M., Howard, A., Matthies, L., Tang, B., and Turmon, M., "Autonomous off-road navigation with end-to-end learning for the LAGR program," *Journal of Field Robotics* **26**(1), 3 – 25 (2009).

[15] Oleynikova, H., Honegger, D., and Pollefeys, M., "Reactive avoidance using embedded stereo vision for mav flight," in [*IEEE Intl. Conf. on Robotics and Automation*], (May 2015).

[16] Otte, M. W., Richardson, S. G., Mulligan, J., and Grudic, G., "Path planning in image space for autonomous robot navigation in unstructured outdoor environments," *Journal of Field Robotics* **26**(2), 212 – 240 (2009).

[17] Hirschmüller, H., Innocent, P. R., and Garibaldi, J., "Real-time correlation-based stereo vision with reduced border errors," *Int. J. Comput. Vision* **47**, 229–246 (Apr. 2002).