A Single−Camera Omni−Stereo Vision System for 3D Perception of Micro Aerial Vehicles (MAVs)

Carlos Jaramillo, Ling Guo, and Jizhong Xiao*
†* *Senior Member, IEEE*

***Abstract*— The limited payload and on-board computation constraints of Micro Aerial Vehicles (MAVs) make sensor configuration very challenging for autonomous navigation and 3D mapping. This paper introduces a catadioptric single-camera omni-stereo vision system that uses a pair of custom-designed mirrors (in a folded configuration) satisfying the single view point (SVP) property. The system is compact and lightweight, has a wide baseline which allows fast 3D reconstruction based on stereo calculation. The algorithm for generating range panoramas is also introduced. The simulation and experimental study demonstrate that the system provides a good solution to the perception challenge of MAVs.**

I. Introduction

Micro aerial vehicles (MAVs), such as quadrotor helicopters, keep emerging as popular platforms for unmanned aerial vehicle (UAV) research due to their structural simplicity, small form factor, their vertical take-off and landing (VTOL) capability and high maneuverability. They have many military and civilian applications, such as target localization and tracking, 3D mapping, terrain and utilities inspection, disaster monitoring, environmental surveillance, search and rescue, traffic surveillance, deployment of instrumentation, and cinematography.

The limited payload and on-board computation constraints of MAVs make sensor configuration very challenging for autonomous navigation and 3D mapping, which demand compact and lightweight sensors as well as efficient signal processing algorithms suitable for on-line implementation. The most commonly used perception sensors on quadrotors are laser scanners and cameras (i.e., monocular, stereo, fish-eye, and omni-directional cameras). Recently, red,green,blue plus depth (RGB-D) sensors like the Microsoft Kinect®, have been adopted for indoor navigation [5]. The lightweight Hokuyo® laser scanner produces accurate distance measurements on a 2D plane, but requires the quadrotor to move up and down in order to generate a 3D map. On the other hand, RGB-D sensors relieve MAVs from the burden of frequent vertical motion while generating a 3D map, but they suffer from a short range and are only suitable for indoor applications.

Vision sensor configurations for MAVs are different from those used in ground robots and high payload UAVs in the following two aspects:

1) Limitations in the sensors' size and weight are different. In MAV applications, sensors' physical dimensions and weight are always a big concern. Due to payload constraints, MAVs normally require fewer sensors that are compactly designed, while larger robots have greater freedom of sensor choice.
2) Field-of-view is different. Due to their omnidirectional motion model, MAVs require 3D observation of the surrounding environment. Conversely, most ground robots only care about frontal views.

Hence, an imaging system of lightweight compact structure, large field-of-view, and acceptable resolution is needed for MAV applications.

Omnidirectional catadioptric vision systems provide a possible solution and they have been used in autonomous flying of large UAVs [11] [10]. Researchers have proposed many different catadioptric configurations with various mirrors [7][1][3][18][12][14][9][19] and a single camera in order to produce stereo images by sacrificing spatial resolution. In [16], 9 possible folded configurations of the single-camera omni-stereo imaging system are presented and one of them is realized here. The catadioptric approach to stereo offers practical advantages for MAVs, such as reduced cost, weight, and robust pixel-disparity correspondences since a single camera does not introduce discrepancies between cameras' intrinsic parameters or synchronization issues. Omni-stereo offers the capability to recover dense omnidirectional depth maps for occupancy grids generation and MAV path planning.

In our previous work, we developed a novel catadioptric-stereo rig consisting of a perspective camera coaxially-aligned with two spherical mirrors of distinct radii (in a "folded" configuration) [14]. The spherical mirrors are available off-the-shelf but the catadioptric-stereo rig does not satisfy the single view point (SVP) constraint [15], so it can only be approximated. We extend our research to design a SVP-compliant folded, catadioptric omni-stereo system with custom designed hyperbolic mirrors. The system is compact and lightweight, satisfies the SVP property, has a wide baseline that enables fast and accurate 3D reconstruction, and has acceptable resolution for UAV applications.

This work is supported in part by U.S. Army Research Office under grant No. W911NF−09 − 1 − 0565, U.S. National Science Foundation under grants No. IIS- 0644127

Carlos Jaramillo is with the Dept. of Computer Science, The Graduate Center, The City University of New York (CUNY), 365 Fifth Avenue, New York, NY 10016 (e−mail: cjaramillo@gc.cuny.edu)

Ling Guo is with Nanjing University of Science and Technology, China (e−mail: laura0955@live.cn)

Jizhong Xiao is with the Electrical Engineering Department, The City College of New York, Convent Ave & 140th Street, New York, NY 10031 (corresponding author, jxiao@ccny.cuny.edu)

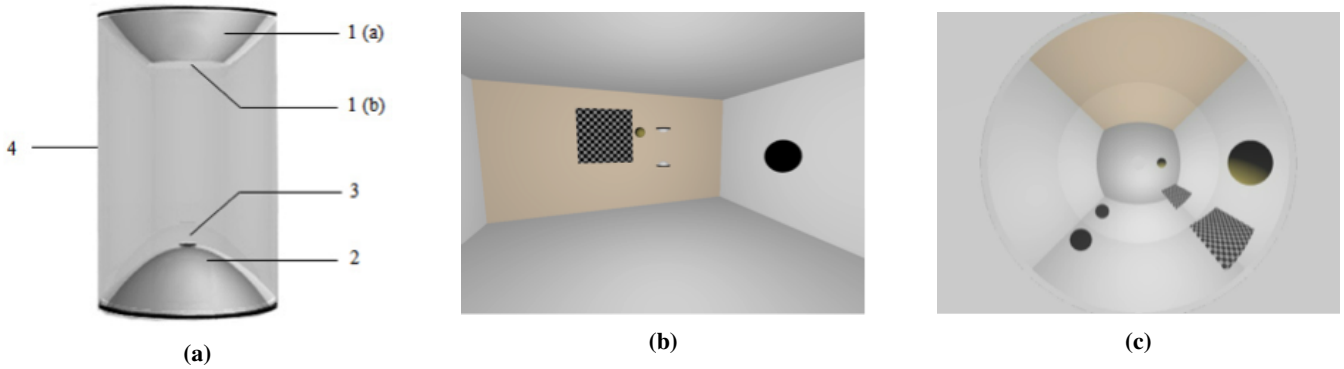


Fig. 1: Synthetic catadioptric single-camera omni-stereo system: **a)** The simulated rig, consisting of (1) hyperbolical-planar combined mirror at top, (2) hyperboloidal mirror at bottom, (3) camera (4) transparent cylinder; **b)** Simulated scenario where a chess-board, a yellow and a black ball are in the field of view of the omni-stereo vision system; and, **c)** the image captured in the scene in Fig. 1b (objects reflected on the bottom mirror appear in the inner image ring, whereas reflections from the top mirror are imaged in the outer rings). Note the radial symmetry about the image center. Depth information can be calculated from the corresponding pixel disparities.

II. Sensor Design

Fig. 1a shows the single-camera catadioptric omni-stereo vision system that has been specifically designed for our quadrotors from Ascending Technologies. It consists of 1) one hyperbolical-planar mirror on the top, 2) one hyperboloidal mirror on the bottom, and 3) a high-resolution camera inside the bottom mirror, all installed inside a 4) transparent acrylic cylinder tube. The choice of the hyperboloidal shape owes to three reasons: it is one of the four non-degenerated conic shapes satisfying the SVP constraint [15]; it allows a wider field of view than elliptical and planar mirrors; and it does not require a telescopic (orthographic) lens for imaging as with paraboloid mirrors (so our system can be downsized). In addition, the planar part of the upper mirror works as a reflex mirror, so distortion due to the dual reflection is minimal. Based on the SVP property, the system produces two radial images for visible objects, one in the inner ring and the other in the outer ring of the image plane (see Fig. 1b and Fig. 1c).

As Fig. 2 shows, the hyperboloidal part of mirror 1 forms one image point p_1 of P, and its planar part works with mirror 2 as a reflex mirror, forming the other image p_2. The two image points are on a radial line, which facilitates the stereo computation. The scale between them is determined by the position of the 3D point P and the mirror parameters.

A. SVP Constraint and Imaging Geometry

In the configuration of Fig. 3, the foci of mirror 1 are F_1 and F_1'. The foci of mirror 2 are F_2 and F_2'. We use a planar (reflex) mirror to reflect the real camera's pinhole located at F_1' as a virtual camera located at F_2'. Hence, we denote denoted as $d/2$ as the reflex mirror's distance to F_1' and to F_2', so that $d = \mid F_2' - F_1' \mid$. Since c_1 and c_2 are the distances between the two foci of mirror 1 and the two foci of mirror 2, respectively, the baseline becomes $L = c_1 + c_2 - d$. Radially symmetric mirrors conforming to SVP constraint can be described by two-parameter equations, as detailed

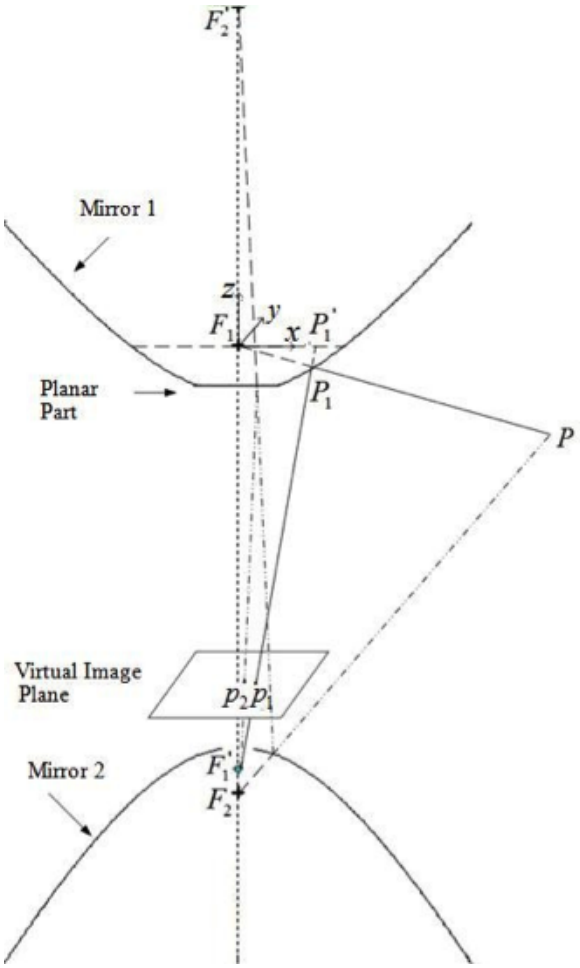


Fig. 2: Image formation of the system. A 3D point P is imaged as p_1 by reflecting on mirror 1, and as p_2 by reflecting on mirror 2 and the planar (reflex) mirror. *Courtesy of Guo [9].*

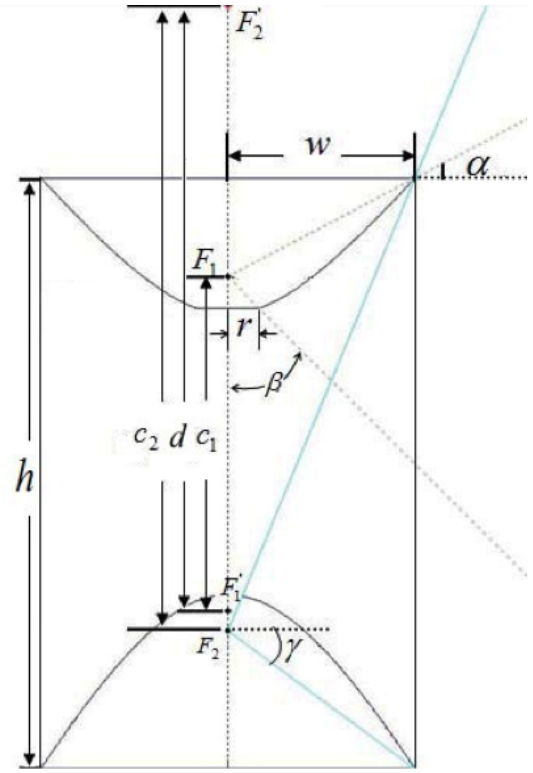


Fig. 3: Geometric model parameters and field-of-view. *Courtesy of Guo [9].*

by Baker and Nayar in [15]. Therefore, for the coordinate system with origin F_1, the respective hyperboloidal parts of mirror1 and mirror 2 can be represented by:

$$\left(Z_1 - \frac{c_1}{2}\right)^2 - (X_1^2 + Y_1^2)\left(\frac{k_1}{2} - 1\right) = \frac{c_1^2}{4}\left(\frac{k_1 - 2}{k_1}\right) \tag{1a}$$

$$\left(Z_2 - \left(d - \frac{c_2}{2}\right)\right)^2 - (X_2^2 + Y_2^2)\left(\frac{k_2}{2} - 1\right) = \frac{c_2^2}{4}\left(\frac{k_2 - 2}{k_2}\right) \tag{1b}$$

for some point $P_1 = (X_1, Y_1, Z_1)$ in mirror 1, and $P_2 = (X_2, Y_2, Z_2)$ in mirror 2. $k_1, k_2 \geq 2$ are the eccentricity-related parameters of the corresponding hyperbolic mirrors.

Different combinations of these parameters affect the system dimensions and determine the performance of the imaging system, such as field-of-view, spatial resolution and depth resolution. Taking mirror 1 as an example, its projective geometry is presented to provide a basic understanding of the imaging process. Let

$$\lambda = \frac{c_1}{R_1\sqrt{k_1 \cdot (k_1 - 2)} - k_1 Z} \tag{2}$$

so that point $P(X, Y, Z)$ in homogeneous coordinates with respect to F_1 is reflected at P_1 (transformed by λ), and then imaged as pixel $p_1 = (u_1, v_1)$ through the projection described by:

$$\begin{bmatrix} u_1 \\ v_1 \\ 1 \end{bmatrix} = M_2 M_1 \begin{bmatrix} \lambda X \\ \lambda Y \\ \lambda Z \\ 1 \end{bmatrix} \tag{3}$$

where $R_1 = \sqrt{X_1^2 + Y_1^2 + Z_1^2}$,

$$M_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & c_1 \end{bmatrix}, \text{ and } M_2 = \begin{bmatrix} f_u & s & u_0 \\ 0 & f_v & v_0 \\ 0 & 0 & \frac{1}{c_1 + \lambda Z} \end{bmatrix}$$

In this transformation chain, M_1 brings P_1 to the coordinate system of F_1' (or the camera's frame), and M_2 is the intrinsic matrix of the pinhole camera model with pixel focal lengths f_u & f_v, skew term s, and scale factor $1/(c_1 + \lambda Z)$.

B. Vertical Field-of-View (vFOV) Constraint

The horizontal FOV is 360°. However, the vertical FOVs (vFOV) of the two mirrors are different. The bottom mirror has a larger vFOV than the upper mirror as shown in Fig. 3. The overlapping area of the two mirrors' vFOVs is the area where objects can be shown in both inner and outer ring images, where the stereo calculation can be performed. In the design, the vFOVs are decided by three angles: α, β, and γ, such that:

$$vFOV = \begin{cases} \alpha + 90 - \beta & \beta < \gamma \\ \alpha + \gamma & \beta > \gamma \end{cases} \tag{4}$$

To be installed along the central axis of the quadrotor, the vision system should ensure that objects located 25cm above (or under) and at 1 meter away from the axis can be viewed. At the same time, angle β should be large enough to avoid the MAV's blades from getting imaged. The size of inner and outer ring images should be in good proportion, as well. Considering all these design factors, we select $\alpha \geq 14^o, \beta \geq 65^o, \gamma \leq 14^o$. Geometrical relations between these constraints and system parameters can be established from Fig. 3.

C. Spatial Resolution

The images acquired by our system are not resolution invariant. Indeed, an omni-camera producing resolution-invariant images has a non-analytical form of the mirror [6], and therefore, it is not suitable for fast 3D depth calculation. The spatial resolution is defined as the number of pixels per solid angle. From [15], the relationship between the resolution of upper mirror η_{m1} and that of the conventional camera η_{cam} is:

$$\eta_{m1} = g \cdot \eta_{cam} = \frac{R_1^2}{(c_1 - Z_1)^2 + X_1^2 + Y_1^2}\eta_{cam} \tag{5}$$

for a point $P_1 = (X_1, Y_1, Z_1)$ in mirror 1 with F_1 as the coordinate frame, where $R_1 = \sqrt{X_1^2 + Y_1^2 + Z_1^2}$. Since $P_1 = \lambda P$, we can substitute $R_1 = \lambda R$ in g for equation (5), to obtain:

$$\begin{aligned} g &= \frac{(\lambda R)^2}{(c_1 - \lambda Z)^2 + (\lambda X)^2 + (\lambda Y)^2} \\ &= \frac{R^2}{\left(\frac{c_1}{\lambda} - Z\right)^2 + X^2 + Y^2} \end{aligned} \tag{6}$$

which indicates how the resolution η_{cam} of a point P in the image decreases with λ^2, and (2) implies it is inversely proportional with the parameter k_1. We observe that the smaller k_1 gets, the flatter the mirror becomes, so it provides a better spatial resolution. However, a smaller k_1 requires a larger radius, w of the system in order to achieve a similar vFOV. According to our application, if $k_1 < 3$, it becomes impractical. On the other hand, the larger k_1 gets, the shorter the baseline L of the stereo system becomes. The same analysis is true for mirror 2.

D. Depth Resolution

Depth resolution shows the system's ability to distinguish two points in the 3D world. The 3D coordinates are retrievable only when the depth difference in the image plane can be detected due to pixel disparity. We study the special case of Fig. 4 for two points (w.r.t. F_1) $A_1 = (X_a, Y_a, L)$ and $A_2 = (X_a + \Delta X, Y_a, L)$ that are collinear with the bottom mirror's focus F_2, where $L = c_1 + c_2 - d$ is the baseline. A_1 and A_2 get imaged as the same pixel in the inner ring of the image (the image of mirror 2), but they appear as different pixels, $a_1 = (u_1, v_1)$ and $a_2 = (u_2, v_2)$, in the outer ring image (the image of mirror 1, top view). The depth difference between A_1 and A_2 is detectable only when $\mid u_2 - u_1 \mid > d_u$ and $\mid v_2 - v_1 \mid > d_v$, where d_u, d_v are the physical pixel sizes, respectively. A larger baseline, L, can increase the pixel disparity, and it is preferable for 3D reconstruction.

We compute a_1 and a_2 from equation (3). This leads us to conclude that depth resolution approximately decreases with the baseline L squarely. Among the three variables that constitute L, c_1 contributes more to the length of the baseline because $c_1 > c_2 - d$ according to the geometric constraint of our design. The lower limit of c_1 is given in the optimization of Section III to meet the baseline requirement.

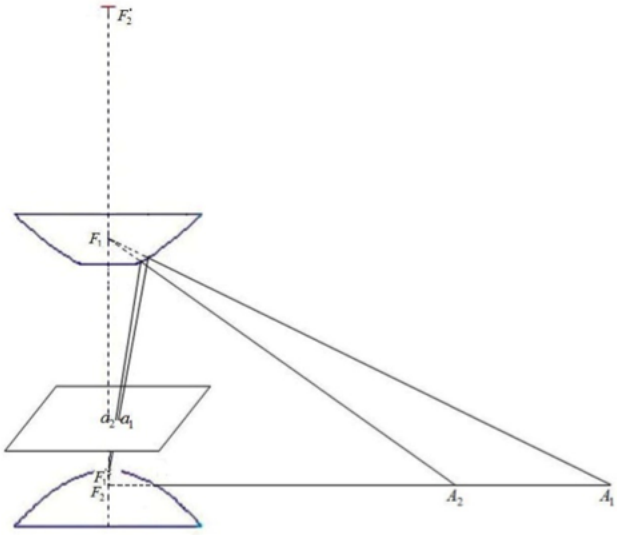


Fig. 4: Depth resolution study: two points with different depth and their projections onto the image plane.

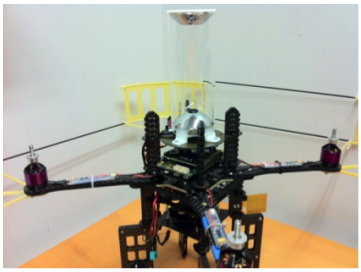
III. Optimization and Simulation

The nonlinear nature of this system makes it very difficult to balance among all performance aspects. We model this omnidirectional sensor as an optimization problem where the length of baseline, L, is the objective function, and unknown variables $xx = (c_1, c_2, k_1, k_2, d, w)$ are the parameters to be optimized under the following constraints:

a) geometrical constraints, including SVP and reflex constraint, as shown in equations (1a) and (1b);
b) physical constraints, including the rig's tube radius w, and its height h, which can be expressed by variables in xx;
c) performance constraints, where the spatial resolution and depth resolution are determined by parameters k_1, k_2, and c_1 while the vertical field-of-view (vFOV) is a function of variables in xx.

This optimization problem is solved with the `MATLAB` optimization toolbox. The found parameters are then used to build a synthetic model with POV-Ray, a open-source ray tracer that allows us to create a 3D scene, with our customized mirrors and camera models, to be rendered as an image (as the examples shown in Fig. 1). The simulation stage plays two important roles in our research: 1) to produce ground-truth images acquired by the optimized vision system, and 2) verify if the expected performance is achieved with respect to the ground-truth knowledge of the synthetic scene.

After the design optimization, we have produced two physical prototypes of the compact catadioptric omni-stereo vision system that can be installed on the Pelican quadrotor (made by Ascending Technologies). Fig. 5a shows the smaller prototype with hyperbolic mirrors of 60 mm in diameter and a MatrixVision® BlueFOX-MLC camera (752 x 480 pixel resolution with global shutter at 90fps). Fig. 5b shows the larger rig constructed with hyperbolic mirrors of 75 mm in diameter and a PointGrey® Chameleon color camera (1280x960 Y8 at 15 FPS). The rigs use acrylic glass tubes to separate the mirrors at the specified distance $h = 150$ mm from the base of mirror 2 to the top of mirror 1. The supporting caps for the mirrors were designed in 3-D CAD and printed for assembly.



(a) 60 mm diameter mirrors (installed on a AscTec Pelican quadrotor)

(b) 75 mm diameter mirrors (compared to can)

Fig. 5: Prototypes of the omni-stereo vision rig

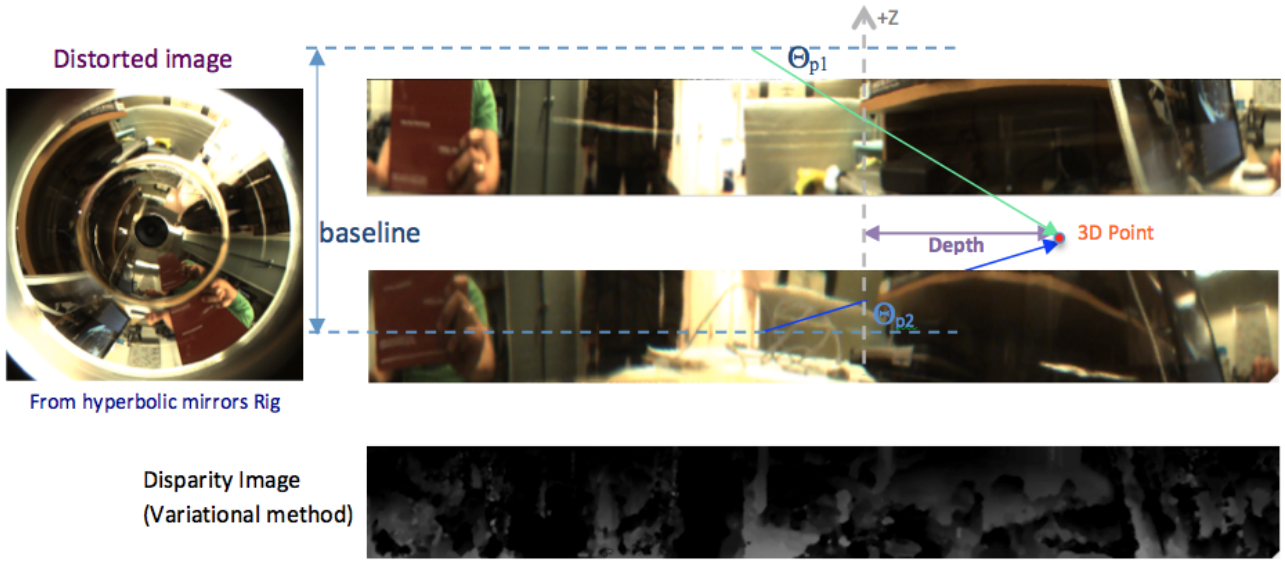


Fig. 6: After mapping pixels from the distorted image into panoramic images (with azimuth and elevation LUTs), corresponding pixels are found (encoded as a disparity map panorama), and then triangulated from their elevation angles (Θ_{p_1}, Θ_{p_2}).

IV. 3D Reconstruction from Omni-stereo Images

Stereo vision used in images is a popular method for finding pixel correspondences that can provide depth (range) information based on their disparity values. It should always be the case that the disparity of objects close to the vision system is higher than for objects that are far away. This inverse relation allows us to further triangulate these pixel correspondences (in a vertical manner) in order to compute the 3-D position of the correspondences (in the world). As shown in [8], the unwrapped panoramas contain vertical, parallel epipolar lines that facilitate the pixel correlation search.

We understand that the correspondence matching algorithm chosen is crucial for correct disparity computation. We refer the reader to [2] for a detailed description of stereo correspondence methods. After comparing various block matching algorithms, we were able to obtain denser disparity maps with the variational method introduced by [13] as it adapts the search to multi-level grids so it can find better matches in real time. However, no stereo matching algorithm (as far as we know today) is totally immune to mismatches.

First, we form the two panoramas for each mirror view (as shown in Fig. 6) by reverse-mapping 3-D points from a unit cylinder (surrounding the respective focus) to the pixel points on the distorted image. We arbitrarily fill up the panoramic images for a constant width and height (in pixels) from where we infer the corresponding angular intervals $\Delta\varphi$ and $\Delta\Theta$. We generate look-up tables (LUTs) in encode angular elevations $\Theta_{1,p_{i,j}}$, $\Theta_{2,p_{i,j}}$ and azimuth $\varphi_{p_{i,j}}$ (relative to their respective focal points F_1 and F_2) for each image point $p_{i,j} = (u_i, v_j)$.

Using stereo matching among the panoramas, we find a pair of pixels points $p_1 = (u_1, v_1)$ and $p_2 = (u_2, v_2)$ in both ring images (areas-of-interest) of the mirrors, respectively. For each azimuthal match (due to vertical correspondences), we refer to the points' back-projection angles Θ_{p_1} and Θ_{p_2} (or their complements α and β) saved in the elevation LUTs.

Given the model's baseline L, we can calculate the depth d of the match by triangulation (Fig. 7) as follows:

$$d = L\frac{sin(\alpha)sin(\beta)}{sin(\alpha+\beta)} \tag{7}$$

where $\alpha = 90^o - \Theta_{p_2}$, and $\beta = 90^o + \Theta_{p_1}$. The real-world coordinates of P become:

$$P = \begin{bmatrix} d \cdot cos(\varphi) \\ d \cdot sin(\varphi) \\ \frac{d}{tan(\alpha)} \end{bmatrix} \tag{8}$$

where φ is the azimuth angle on the XY-plane.

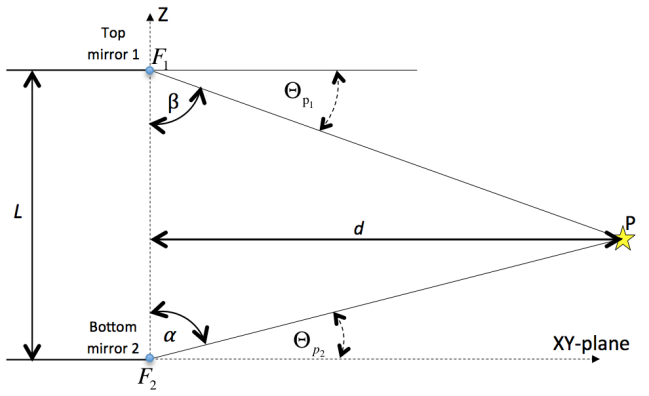


Fig. 7: Depth d for a pair of corresponding image points $\langle p_1, p_2 \rangle$ are obtained via triangulation of the elevation angles Θ_{p_1} and Θ_{p_2} obtained from back-projected rays.

Fig. 7 demonstrates the triangulation procedure that we describe here for our folded configuration of mirrors, which are coaxially aligned. It is easier to visualize this procedure as each pair of rays emanating from the origin of the viewpoints of each mirror coordinate system (the foci F_1 and F_2), which meet at point P (Fig. 6 and Fig. 7).

By implementing this process (in approximately 10 fps on a 800x600 resolution image), we give an example of a 3-D point cloud computed from a synthetic image (for clarity) as shown in Fig. 8. Observe how the depth of the ball and the yellow box differ in the scene.

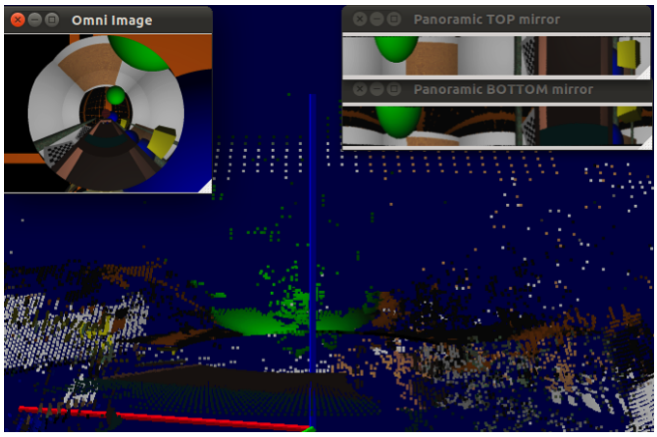


Fig. 8: A 3-D point cloud computed from a synthetic simulation of the hyperbolic mirrors rig.

V. Discussion and Future Work

The mirror shapes are optimized and custom-manufactured out of aluminum using a CNC machine, as to meet the performance specifications (e.g., FOV, resolution). However, it is not an easy job to assemble a perfect imaging system. The misalignment of top and bottom mirrors, the imprecise installation of the camera, as well as the glare from the tube are all caveats we still need to improve for better image processing and 3D reconstruction.

We have developed an adaptive supporting frame enabling us to fine tune the assembly using a spring-screw mechanism, which adds extra weight. The total weight of the smaller rig with hyperbolic mirrors of 60mm diameter is around 200 grams, which can be carried by the AscTec Pelican quadrotor whose payload is 500g. After improving the design, we are going to make a mold to produce plastic mirrors and treat the acrylic support tube with anti-glare coating. We believe this will dramatically reduce the weight and improve the image quality. The ongoing research will focus on the development of efficient software algorithms for fast 3D reconstruction (dense and sparse clouds) by taking advantage of the collinear property of corresponding points (matches). The target is to implement the algorithms onboard the AscTec Mastermind embedded computer which has a Intel® Core™ 2 Duo CPU. The 3D reconstruction software based on the omni-stereo vision system will be integrated with our motion planning and 3D mapping software [4] leading to a working system for autonomous navigation. Our software is open-source and is under continuous development. We have implemented our system with open-source tools such as the OpenCV library [2] and the Robotics Operating System (ROS) framework [17]. Our software is publicly available from our experimental Git repository at `http://robotics.ccny.cuny.edu/git/ccny-ros-pkg/ccny_experimental.git`.

References

[1] R. Bajcsy. High resolution catadioptric omni-directional stereo sensor for robot vision. *2003 IEEE International Conference on Robotics and Automation (Cat. No.03CH37422)*, pages 1694–1699, 2003.

[2] G. Bradski and A. Kaehler. *Learning OpenCV: Computer vision with the OpenCV library*. O'Reilly Media, 2008.

[3] Eduardo E.L.L. Cabral, J.C. de Souza, and M.C. Hunold. Omnidirectional stereo vision with a hyperbolic double lobed mirror. In *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, pages 0–3. Ieee, 2004.

[4] Ivan Dryanovski and William Morris. Multi-volume occupancy grids: An efficient probabilistic 3D mapping model for micro aerial vehicles. *International Conference on Intelligent Robots and Systems (IROS)*, (1):1553–1559, 2010.

[5] Ivan Dryanovski, Roberto G Valenti, and Jizhong Xiao. Fast Visual Odometry and Mapping from RGB-D Data. In *International Conference on Robotics and Automation*, volume 10031, 2013.

[6] Jose Gaspar, Claudia Deccó, Jun Jr. Okamoto, and Jose Santos-Victor. Constant resolution omnidirectional cameras. In *OMNIVIS'02 Workshop on Omni-directional Vision*, 2002.

[7] J. Gluckman and S.K. Nayar. Rectified catadioptric stereo sensors. *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2:380–387, 2000.

[8] Joshua Gluckman, Shree K. Nayar, and Keith Thoresz. Real-Time Omnidirectional and Panoramic Stereo. *Proceedings of the 1998 DARPA Image Understanding Workshop*, 1998.

[9] Ling Guo and Igor Labutov. Design and calibration of single-camera catadioptric omnistereo system for miniature aerial vehicles (MAVs). *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2010.

[10] S. Hrabar. 3D path planning and stereo-based obstacle avoidance for rotorcraft UAVs. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 807–814, September 2008.

[11] Stefan Hrabar and GS Sukhatme. Omnidirectional vision for an autonomous helicopter. In *International Conference on Robotics and Automation (ICRA)*, number April, pages 3602–3609, 2003.

[12] Gijeong Jang, Sungho Kim, and Inso Kweon. Single camera catadioptric stereo system. In *Proceedings of Workshop on Omnidirectional Vision, Camera Networks and Nonclassical cameras (OMNIVIS2005)*. Citeseer, 2005.

[13] Sergey Kosov, T. Thormählen, and HP H.P. Seidel. Accurate real-time disparity estimation with variational methods. *Advances in Visual Computing, Lecture Notes in Computer Science*, 5875:796–807, 2009.

[14] Igor Labutov, Carlos Jaramillo, and Jizhong Xiao. Generating near-spherical range panoramas by fusing optical flow and stereo from a single-camera folded catadioptric rig. *Machine Vision and Applications*, pages 1–12, September 2011.

[15] Shree K. Nayar and S. Baker. A theory of single-viewpoint catadioptric image formation. *International Journal of Computer Vision*, 35(2):175–196, 1999.

[16] S.K. Nayar and V. Peri. Folded catadioptric cameras. *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*, pages 217–223, 1999.

[17] Morgan Quigley, Brian Gerkey, Ken Conley, Josh Faust, Tully Foote, Jeremy Leibs, Eric Berger, Rob Wheeler, and Andrew Ng. ROS: an open-source Robot Operating System. In *International Conference on Robotics and Automation (ICRA)*, number Figure 1, 2009.

[18] L Su and F Zhu. Design of a novel stereo vision navigation system for mobile robots. *IEEE Robotics and Biomimetics (ROBIO)*, pages 611–614, 2005.

[19] Sooyeong Yi and N. Ahuja. An Omnidirectional Stereo Vision System Using a Single Camera. *18th International Conference on Pattern Recognition (ICPR'06)*, pages 861–865, 2006.