The 2010 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 18-22, 2010, Taipei, Taiwan

Design and calibration of single-camera catadioptric omnistereo system for Miniature Aerial Vehicles (MAVs)

Ling Guo, Igor Labutov, Jizhong Xiao*, *Senior Member, IEEE*

***Abstract*—Stereo system plays an important role in the navigation of MAVs. In this paper, we design a single-camera catadioptric omnistereo system for MAV, which consists of one hyperboloidal mirror, one hyperboloidal-planar combined mirror, and one conventional camera. System parameters are optimized based on the analysis of constraints and each parameter's influence on performance. Projective model of this system is derived, which provides a foundation for sphere-based calibration algorithm. It calibrates not only the conventional camera parameters, but also the mirror parameters. We also prove that a minimum of two spheres are needed to calibrate the seven parameters.**

I. INTRODUCTION

Miniature unmanned aerial vehicle (UAV), such as the quadrotor helicopter, provides a flexible platform for surveillance, remote sensing and other information acquisition applications. But its limited payload makes sensor configuration very challenging, especially when it's used in an environment where GPS-signal is not available. In this circumstance, visual sensors play an important role in UAV localization and navigation.

Catadioptric camera, an implementation of omnidirectional vision with aligned mirrors and lens, is widely used in both ground robot [1] and UAV [2]. Its category with SVP (single view point) property is preferred, since only rays that go towards mirror foci, and then are reflected by mirror, are captured by camera and thus geometrically correct projection is guaranteed. Omnistereo cameras with two mirrors and two cameras [3] and then with one camera [4-6] are researched for depth computation in a large field of view, although the latter is at the cost of spatial resolution. Considering the system size and performance, the configuration in [5,6] is practical for our MAV application because of its compact design and reasonable baseline. More possible configurations of the single-camera omnistereo imaging system are listed in [7]. The requirements of visual sensor configuration for MAVs are different from that for ground robots and high payload UAVs in the following two aspects: (1) Limitations in the sensors' size and weight; (2) 3D observation of surrounding environment required, not only the front view. So an imaging system with compact structure, large field of view and acceptable performance is needed in MAV applications.

Calibration is another important topic for catadioptric system. The calibration method of catadioptric camera can be point-based, line-based and sphere-based, which study projective models for different elements on the calibration target. Point-based method uses corners with known coordinates on chessboard [8]. Line-based method establishes the relationship between lines on chessboard or in 3D space and their conic images mapped on the image plane [9, 10]. The equivalent model for catadioptric projection [11] provides the basis for sphere-based method that has spheres as calibration target. Line-based and sphere-based methods are more flexible, since they neither require the prior knowledge like corner coordinates, nor deal with the external parameters of the calibration target. The latter outperforms the former because of the higher conic fitting accuracy[9].

In this paper, a catadioptric single-camera omnistereo imaging system designed for our AscendTech quadrotor is introduced, with the goal of being lightweight and compact, having acceptable resolution and allowing fast 3D reconstruction based on stereo calculation. System parameters are presented and their influence on the weight limitation (i.e., 300gram) and performance is analyzed. The projective model of this system is derived and a calibration algorithm based on this model is presented. We also prove that a minimum of two spheres are required. POV-Ray is used to simulate the design and produce images for sphere-based calibration.

The paper is organized as follows. Section II describes system configuration, its parameters and the optimization procedure. Section III presents the equivalent model of the imaging geometry and the calibration method. Section IV shows the experimental results using the POV-Ray simulated images. Conclusion is drawn in section V.

II. SYSTEM DESIGN

A. System Overview

The single-camera omnistereo imaging system, as shown in Fig. 1, consists of a conventional camera and two mirrors: one hyperboloidal mirror on the bottom and one hyperboloidal-planar combined mirror on the top.

This work was supported in part by the U.S. Army Research Office under grant No.W911NF-09-1-0565, and U.S. National Science Foundation under grants No.CNS-0619577 and No. IIS-0644127.

Ling Guo was with Nanjing University of Science and Technology, China, and is now with the City College of New York as a visiting post-doc research associate (e-mail: laura0955@ live.cn).

Igor Labutov is with the Electrical Engineering Department, City College of New York, Convent Ave & 140th Street, New York, NY 1003 (e-mail: igor.labutov@gmail.com).

*Jizhong Xiao is with the Electrical Engineering Department, City College of New York, Convent Ave & 140th Street, New York, NY 10031(corresponding author, phone: 212-650-7268; fax: 212-650-8249; e-mail: jxiao@ccny.cuny.edu).

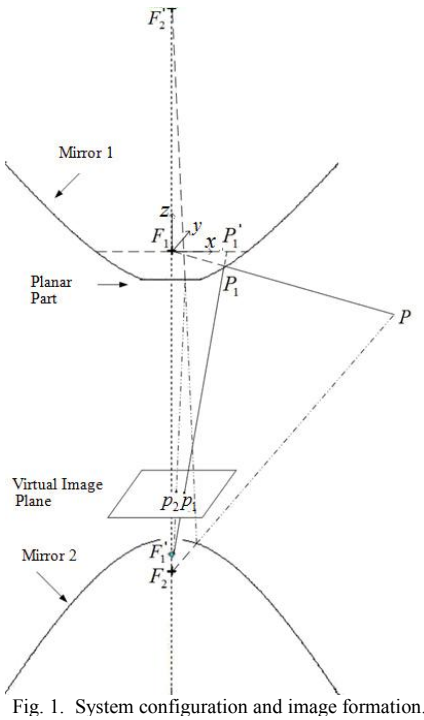

Fig. 1. System configuration and image formation.

Based on SVP property, the system produces two radial images for any visible object, one in the inner ring and the other in the outer ring of the image plane (see Fig. 6). As shown in Fig. 1, the two images of P are on the radial line, which facilitates stereo computation, and the scale between them is determined by object position and mirror parameters.

B. Parameters and Their Impact on Performance

1) SVP Constraint and Imaging Geometry: Mirrors conforming to SVP constraint can be described with two-parameter equations, as detailed in [12]. In the configuration illustrated by Fig. 2, the foci of mirror 1 are F_1 and F_1'. The foci of mirror 2 are F_2 and F_2'. F_1' is also the optical center of the conventional camera. To have the planar part as a reflex mirror, its distance to F_1' and F_2' should be equal, noted as $d/2$. So under the coordinate system with origin F_1, the hyperboloidal parts of mirror1 and mirror 2 can be represented by the following equations:

$$(Z-\frac{c_1}{2})^2-(X^2+Y^2)(\frac{k_1}{2}-1)=\frac{c_1^2}{4}(\frac{k_1-2}{k_1}) \quad (1)$$

$$(Z-(d-c_2/2))^2-(X^2+Y^2)(\frac{k_2}{2}-1)=\frac{c_2^2}{4}(\frac{k_2-2}{k_2}) \quad (2)$$

where c_1 and c_2 are the distance between two foci of mirror 1 and mirror 2 respectively, which determine the baseline together with d. k_1 and k_2 are eccentricity-related parameters and $k_1, k_2 > 2$.

Different combinations of the five parameters affect both system dimensions and its performance. Before further analysis, projective geometry, taking mirror 1 as an example, is presented to provide a basic understanding of the imaging process. The point $P(X,Y,Z)$ in Fig. 1 under F_1 coordinate system is imaged as pixel $p_1(u_1,v_1)$ through the projection described by the following equation:

$$\begin{bmatrix}u_1\\v_1\end{bmatrix}=M_3\cdot M_2\cdot M_1\cdot\begin{bmatrix}\lambda X\\\lambda Y\\\lambda Z\\1\end{bmatrix} \quad (3)$$

where $\lambda=\frac{c_1}{\sqrt{k_1\cdot(k_1-2)\cdot R-k_1\cdot Z}}$, $R=\sqrt{X^2+Y^2+Z^2}$,

$M_1=\begin{bmatrix}1&0&0\\0&1&0\\0&0&c_1\end{bmatrix}$, $M_2=\begin{bmatrix}f_0&0&0&0\\0&f_0&0&0\\0&0&1&0\end{bmatrix}$, $M_3=\begin{bmatrix}r&s&u_0\\0&1&v_0\\0&0&1\end{bmatrix}$.

The parameters in M_2 and M_3 are the same as in pin-hole camera model. The equation for p_2 imaged by mirror 2 can be derived in the same way.

2) Field of View (FOV) Constraint: The horizontal FOV is 360°. The vertical FOVs (VFOVs) of the two mirrors are usually different. The intersection of the two mirrors' VFOVs is the area where objects can be shown on both inner & outer ring images, and the stereo calculation can be performed. The VFOVs are decided by three angles: α, β and γ, as Fig. 2 shows, where :

$$VFOV=\begin{cases}\alpha+90-\beta & \beta<\gamma\\ \alpha+\gamma & \beta>\gamma\end{cases}. \quad (4)$$

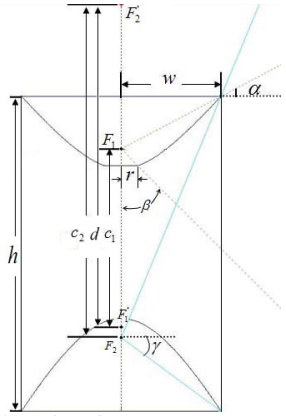

Fig. 2. Geometric illustration of system parameters and field of view.

To be installed along the central axis of our MAV, the system should ensure that at the distance of 1m to the central axis, objects 25cm above or under the MAV can be seen in the image. At the same time, angle β should be large enough to avoid MAV blades being imaged and keep the size of inner and outer ring images in good proportion. Considering all these design factors, we select $\alpha \ge 14^o, \beta \ge 65^o, \gamma \le 14^o$. Geometrical relations between these constraints and system parameters can be established from Fig. 2.

3) Spatial Resolution: the image acquired by our system is not resolution invariant. However, the omnicamera producing resolution-invariant images has non-analytical form of the mirror [13], and thus is not suitable for fast 3D depth calculation.

Spatial resolution is defined as the number of pixels per solid angle. From [12], the relationship between the resolution of upper mirror R_c and that of the conventional camera R_o is:

$$R_c = g * R_o = \frac{R_1^2}{(c_1 - Z_1)^2 + (X_1^2 + Y_1^2)} R_o, \quad (5)$$

where (X_1, Y_1, Z_1) is the coordinate of P_1 under F_1 coordinate frame. From Fig. 1 and equation (3), we get:

$$g = \frac{\lambda^2 \cdot R^2}{(c_1 - \lambda Z)^2 + \lambda^2 (X^2 + Y^2)} = \frac{R^2}{(c_1/\lambda - Z)^2 + (X^2 + Y^2)}, \quad (6)$$

which means approximately the resolution decreases with the square of k_1. The smaller k_1, the flatter the mirror and hence the better spatial resolution. But it might increase the radius of the system. According to our application, k_1 smaller than 3 is impractical. On the other hand, the larger k_1, the shorter baseline of the stereo system. It's the same case for mirror 2.

4) Disparity and Depth Resolution: Disparity is known as the difference between the two projected pixels of a point. Taking point P in Figure 1 as an example, the disparity is $p_1 - p_2$. According to the projective model in section 1), the coordinates (X, Y, Z) of P can't be computed from the coordinates $p_1(x_1, y_1)$ and $p_2(x_2, y_2)$ in image plane when $|p_1 - p_2|_\infty < d_{xy}$, where d_{xy} is the pixel size. From equation (3), a larger baseline $L = c_1 - d + c_2$ can increase disparity and thus is preferable for 3D reconstruction.

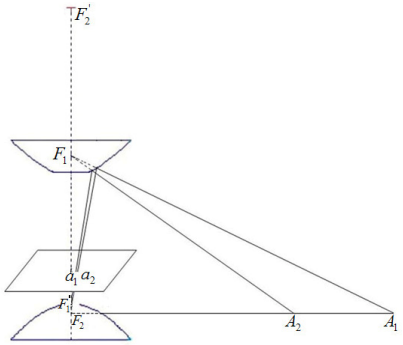

Fig. 3. Two points with different depth and their projections.

Depth resolution shows the system's ability to distinguish two points in 3D world, e.g. A_1 and A_2 in Figure 3. Let's study a special case when $A_1(X_a, Y_a, c_1 - d + c_2)$ and $A_2(X_a + \Delta X, Y_a, c_1 - d + c_2)$ are collinear with the focal point of mirror 2, F_2. They are imaged as the same pixel in the inner ring of the image, but as different pixels in the outer ring image, $a_1(x_1, y_1)$ and $a_2(x_2, y_2)$, respectively. The depth difference between them is undetectable when $|a_1 - a_2|_\infty < d_{xy}$, so the larger baseline L also helps to increase the depth resolution (Fig. 7). Among the three variables that constitute L, c_1 contributes more to the length of baseline than $c_2 - d$ according to the geometric constraint (see Fig. 2). So the lower limit of c_1 should be given in optimization to meet the baseline requirement.

C. Optimization and Simulation

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

(1) geometrical constraints, including SVP and reflex constraint, as shown in equation (1) and (2);
(2) physical constraints, including height h and tube radius w, where h can be expressed by variables in xx;
(3) performance constraints, where the spatial resolution and depth resolution are determined by k_1, k_2 and c_1, while field of view is a function of variables in xx.

This optimization problem is solved by Matlab toolbox and the result is simulated with POV-Ray.

III. Calibration

A. Equivalent Projective Model

The imaging model in section Ⅱ explains how points are projected into pixels physically, while the equivalent projective model proposed in [14] provides foundation for a category of methods for calibration, that's line-based and sphere-based calibration. As described in [11] and shown in Fig. 4, the projection of catadioptric camera is equivalent to a projection of a sphere (B) onto a unit sphere first and then onto a plane positioned at $l + m$ from the projective center at l on the main axes. For example, the mapping from $P(X, Y, Z)$ to $P_1'(x_1', y_1')$ by mirror 1 in Fig. 1 can be regarded as projections described with the following equation:

$$\begin{bmatrix} x_1' \\ y_1' \\ 1 \end{bmatrix} = \begin{bmatrix} -(l+m) & 0 & 0 \\ 0 & -(l+m) & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -l \end{bmatrix} \cdot \begin{bmatrix} X/R \\ Y/R \\ Z/R \\ 1 \end{bmatrix} \quad (7)$$

where l and $l + m$ are decided by mirror parameters: $l = \sqrt{k_1(k_1 - 2)}/(k_1 - 1)$ and $(l + m) = c_1/(k_1 - 1)$.

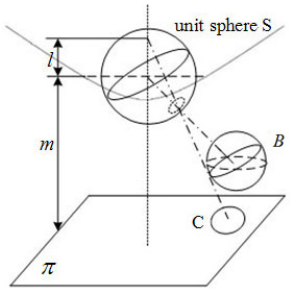

Fig. 4. Equivalent model for sphere projection.

In Fig. 4, a sphere B in space is mapped into a circle on unit sphere, then projected into a conic C on plane π, and finally imaged into another conic by perspective camera. So here we introduce the equivalent projective model for our

single-camera omnistereo system, which consists of three steps: 1) expressing circles mapped on the two unit spheres centering at foci of the upper and lower mirrors; 2) computing the conic matrices according to equivalent model, and 3) deriving the final conic image from the matrices.

We adopt this notation convention for simplicity: use a character with different subscript to distinguish the variables for mirror1 and mirror2 (e.g., c_1 and c_2 are the focal length for mirror1 and mirror2), respectively; and the character without subscript (e.g., c) for both when we don't have to distinguish them.

1) Circles Mapped on Unit Spheres: assume a sphere with radius rr for calibration is placed at (X_0, Y_0, Z_0) , the circle mapped on a unit sphere can be regarded as intersection of the sphere and the plane: $n_x \cdot X_s + n_y \cdot Y_s + n_z \cdot Z_s + d_0 = 0,$ where (n_x, n_y, n_z) is the unit normal vector. So the parameters for plane intersecting with sphere S_1 are:

$$n_{x1} = \frac{-X_0}{\sqrt{X_0^2 + Y_0^2 + (c_1 - Z_0)^2}},\ n_{y1} = \frac{-Y_0}{\sqrt{X_0^2 + Y_0^2 + (c_1 - Z_0)^2}},$$
$$n_{z1} = \frac{c_1 - Z_0}{\sqrt{X_0^2 + Y_0^2 + (c_1 - Z_0)^2}},\ d_{01} = \frac{\sqrt{X_0^2 + Y_0^2 + (c_1 - Z_0)^2 - rr^2}}{\sqrt{X_0^2 + Y_0^2 + (c_1 - Z_0)^2}}. \quad (8)$$

The parameters for sphere S_2 can also be derived.

2) Conic Equations: assume (X_{s1}, Y_{s1}, Z_{s1}) and (X_{s2}, Y_{s2}, Z_{s2}) are points on the two circles, taking the one on S_1 as an example, three equations are satisfied:

$$X_{s1}^2 + Y_{s1}^2 + Z_{s1}^2 = 1, \quad (9)$$
$$n_{x1} \cdot X_{s1} + n_{y1} \cdot Y_{s1} + n_{z1} \cdot Z_{s1} + d_{01} = 0, \quad (10)$$
$$\lambda_1 \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = M_2 \cdot M_1 \cdot M_0 \cdot \begin{bmatrix} X_{s1} \\ Y_{s1} \\ Z_{s1} \\ 1 \end{bmatrix},\ M_0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -l_1 \end{bmatrix},$$
$$M_1 = \begin{bmatrix} -(l_1 + m_1) & 0 & 0 \\ 0 & -(l_1 + m_1) & 0 \\ 0 & 0 & 1 \end{bmatrix}, M_2 = \begin{bmatrix} f_0 & 0 & 0 \\ 0 & f_0 & 0 \\ 0 & 0 & c_1 \end{bmatrix}. \quad (11)$$

Eliminate λ_1, substitute X_{s1}, Y_{s1}, Z_{s1} expressed with other variables into (9), and we get the conic equation after perspective projection by the camera:

$$(x, y, 1) \cdot CM_1 \cdot (x, y, 1)^T = 0, \quad (12)$$

where the conic matrix

$$CM_1 = \begin{bmatrix} A_1 & B_1 & D_1 \\ B_1 & C_1 & E_1 \\ D_1 & E_1 & F_1 \end{bmatrix} \quad (13)$$

$$A_1 = (l_1^2 - 1) \cdot n_{x1}^2 + (d_{01} + l_1 \cdot n_{z1})^2;$$
$$B_1 = (l_1^2 - 1) \cdot n_{x1} \cdot n_{y1};$$
$$C_1 = (l_1^2 - 1) \cdot n_{y1}^2 + (d_{01} + l_1 \cdot n_{z1})^2;$$
$$D_1 = -f_1 \cdot n_{x1} \cdot (d_{01} \cdot l_1 + n_{z1});$$
$$E_1 = -f_1 \cdot n_{y1} \cdot (d_{01} \cdot l_1 + n_{z1});$$
$$F_1 = f_1^2 \cdot (d_{01}^2 - n_{z1}^2) \cdot$$

Here l_1 and $l_1 + m_1$ have the same meaning as those in equation (7), and $f_1 = -f_0 \cdot (l_1 + m_1) / c_1 = -f_0 / (k_1 - 1) \cdot$

The conic matrix CM_2 projected from unit sphere S_2 has the similar form except the signs of some terms.

3) Final Image: The conics with matrix CM_1 and CM_2 are finally transformed from camera coordinate system into image coordinate system, that is:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = M_3 \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} r & s & u_0 \\ 0 & 1 & v_0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (14)$$

From (12) and (14) we get:

$$(u, v, 1) \cdot M_3^{-1} \cdot CM \cdot M_3 \cdot (u, v, 1)^T = 0 \cdot \quad (15)$$

Equation (15) describes conics in the final image, where $M_3^{-1} \cdot CM \cdot M_3$, noted as CM', is the conic matrix.

The projective model can be summarized like this: the sphere with a radius of rr centered at (X_0, Y_0, Z_0) is firstly mapped into two circles on imaginary unit spheres, expressed with equation (9) and (10) for unit sphere S_1, then the circles are projected into conics on the image plane but under camera coordinate system, which are described with matrix CM_1 and CM_2, and finally they are transformed into conics under pixel coordinate system, as matrix CM' shows.

B. Calibration Method

From the above imaging model we get eight parameters to calibrate: f_0, r, s, u_0, v_0 that belong to the camera, and k_1, k_2, $(c_1 - d + c_2)$ that belong to the two mirrors. Calibration is the inverse procedure of the derivation of projective model. So here we present the three-step calibration method as: 1) fitting conics from the image to get the conic matrices; 2) recovering camera parameter from the equations established from conic matrices; 3) solving mirror parameters. The calibration procedure is detailed as follows.

1) Conic Fitting: A sphere in 3D space can be imaged as either ellipse, or hyperbola or even as parabola, depending on its location relative to the vision system. Basically it's difficult to fit a conic without knowing its shape. We narrow down conic to ellipse by placing the sphere object at certain locations where the following inequations are satisfied:

$$(l_1^2 - 1) \cdot (1 - n_{z1}^2) + (d_{01} + l_1 \cdot n_{z1})^2 > 0,$$
$$(l_2^2 - 1) \cdot (1 - n_{z2}^2) + (d_{02} - l_2 \cdot n_{z2})^2 > 0,$$

which are derived from the sufficient and necessary condition for a conic to be an ellipse: $B^2 - AC < 0$. The inequations are satisfied easily for our system for the following reasons:

(1) The first terms in the inequations are always negative, but $(l^2 - 1) = -1/(k-1)^2, k > 2$, is always greater than -1, and decreases rapidly with the increase of k. According to our design, k is no less than 3, so the absolute value of this term is much smaller than 1;

(2) The second terms of the inequations are always positive, and d has the same sign with n_z from the discussion in

section A. If we put a sphere with 25cm radius at the point where $\sqrt{X_0^2+Y_0^2} > 50cm,$, d is close to 1 and thus the second term is close to or even greater than 1.

Ellipse fitting can be solved using the algorithm in [15] based on the edge pixels on ellipses detected by Canny edge detector, and five-parameter result is obtained: $(cx, cy, ax, ay, theta)$, where (cx, cy) is the ellipse center, (ax, ay) the long and short axis respectively, and $theta$ the rotation angle. Obviously ellipse matrices CM' can be retrieved from the five parameters.

2) Camera Parameter Recovery: From retrieved CM' and equation (15), the matrix CM can be expressed with CM' and camera parameters. So the entries of matrix CM_1 and CM_2 can be expressed with the entries of CM'.

A conic matrix like CM with six elements has only five free elements, since CM multiplied by a constant is also the matrix for the same conic. So we examine ratio of the elements rather than a single element. CM_1 and CM_2 are not in exactly the same form, but they conform to the following equations:

$$\frac{D}{E}-\frac{E}{D}=\frac{A-C}{B}, \tag{16}$$

$$(\frac{B}{E}-\frac{A}{D})\cdot f^2=(\frac{F}{D}-\frac{E}{B})\cdot(l^2-1)\cdot \tag{17}$$

From CM' and equation (16), we get an equation with four unknown parameters, which obviously need at least four ellipses to get solved. Since one calibration sphere produces two ellipses in the image, a minimum of two spheres are required. LM algorithm [16] is used to solve the nonlinear equations. The initial solution is given by the method in [9].

From equation (17) and the parameters we solved from (16), we can get the ratio between f^2 and l^2-1, which is f_0^2 from the discussion in section A.

3) Mirror Parameter Recovery: In addition to the parameters of c, k or l, m, the omnistereo system has one more parameter c_1-d+c_2 to calibrate, which defines the relative distance between the two mirrors. With retrieved camera parameters and CM', CM can be computed. We observe that each ellipse provides five equations with four unknown parameters n_x, n_y, k, d_0, because $n_x^2+n_y^2+n_z^2=1$, and both f and l are a function of k. But the initial estimation of the unknown parameters in nonlinear optimization algorithms (e.g., LM algorithm) is not as easy as what we did to solve camera parameters. If we have a prior knowledge about k, all these parameters can be solved easily. After n_x, n_y, n_z, d_0 are solved, substitute them into expressions of d_{01} and d_{02} in equation (8), then the four unknown variables X_0, Y_0, Z_0, c_1-d+c_2 are solved. Notice that, in this step, only two spheres are needed to get the parameter c_1-d+c_2 calibrated, but more spheres will benefit calibration accuracy.

IV. EXPERIMENTS

Experiments are conducted to test the optimized system design and the performance of the calibration algorithm. System parameters are firstly optimized as detailed in section II, and then exported into POV-Ray for simulation. The system performance is tested to verify that all design requirements are satisfied. In second phase, the simulated images with calibration spheres are produced in POV-Ray to test each key part of the calibration algorithm, including conic fitting, retrieval of parameters of conventional camera and computation of mirror parameters.

A. Design

The final design of the catadioptric single-camera omnistereo imaging system is shown in Fig. 5, which has a height of 15cm and radius of 3.6cm. The acylic tubing weighs about 73g. Plus plated mirrors on plastic mould and camera, the weight of the whole system is able to be below 300g. The focal length of the camera is 1.8cm and the CCD resolution is 1600*1200, which are changeable in POV-Ray. To test the field of view, a box-like scene is created and a chessboard of 50cm long and wide is placed 1m away from the vertical central axis of the imaging system, with its upper edge at c_1+25 cm. Fig. 6 shows the complete chessboard and thus proves the expected angle is achieved. The other two angles for field of view are tested in the same way.

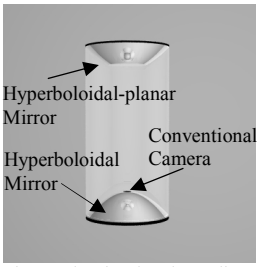

Fig. 5. The simulated catadioptric single-camera omnistereo system.

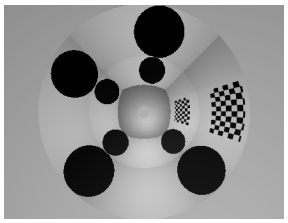

Fig. 6. The chessboard image to test field of view.

Depth resolution for the special case illustrated in Fig. 3 is studied and shown in Fig. 7. The x-axis of Fig. 7 represents the distance of A_2 to the central axis of the imaging system. The y-axis represents the pixel distance between the projections of A_1 and A_2 (i.e., a_1, a_2), which reflects the depth resolution. The dots on line 1 shows the case when the length of the line segment ($A_1 A_2$) equals to 75 cm. Line2, 3 represent the cases when length $A_1 A_2$ equals to 50cm and 25cm, respectively. In Figure 7, the first dot on line 1 shows that the pixel distance is close to 17 pixels when A_1 is positioned at 1m away from the central axis and the length of $A_1 A_2$ is 75cm, (i.e., A_2 is positioned at 175cm away from the central axis). The last dot on line 3 is worth noticing, which shows the pixel distance is less than one pixel, indicating the depth resolution is less than 25cm when A_1 is 4m away from the vision system.

In addition, points on solid lines are computed using the

optimized baseline. The points on dashed lines are computed using smaller baseline value. It demonstrated that the depth resolution increases with the increase of the baseline L.

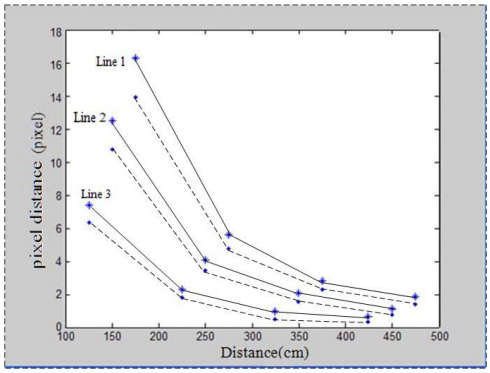

Fig. 7. Depth resolution and its relation with L.

B. Calibration

Images are created with POV-Ray for calibration, as illustrated in Fig. 6 (without chessboard), which have a number of ellipses imaged by four 25cm radius spheres. Using the calibration method in section III-B, after canny edge detection and ellipse fitting, we can calculate both camera and mirror parameters. Ten images with a resolution of 1600*1200 are used for test and the result is as shown in table I, which is very close to the ground truth value.

TABLE I
CALIBRATION RESULTS

Parameters	Results	Ground Truth
f_0	1.8132 ± 0.0109 cm	1.8 cm
u_0	794.9554 ± 3.2532 pixels	800 pixels
v_0	595.8406 ± 4.0873 pixels	600 pixels
r	0.9996 ± 0.0008	1.0
s	0.0001 ± 0.0003	0
$c_1 - d + c_2$	13.26 ± 0.04 cm	13.15 cm

From experiment we notice that increasing image resolution or making the imaged ellipses as large as possible is helpful to improve calibration accuracy, because the ratio between the long and short axis of the ellipses is close to 1 from equation (13). For our upper mirror, $l_1 = 0.9772$, the maximum ratio is 1.0401, which means if an ellipse has a long axis of only 50 pixels, its short axis is of the same length due to pixel quantilization. This small error will be passed to ellipse fitting, then enlarged by the nonlinear equations in calibration and finally result in a big error in calibration results.

V. CONCLUSION

In this paper, we present a single-camera omnistereo imaging system designed for MAV with limited payload. The vision system provides a light-weight and compact sensing tool for stereo image acquisition and fast 3D calculation. It consists of a conventional camera, one hyperboloidal mirror and one hyperboloidal-planar combined mirror. System parameters are optimized taking into consideration of the constraints in SVP, field of view, physical dimensions and resolution. A three-step sphere-based calibration algorithm is proposed based on the derivation of projective model. We prove that a minimum of two spheres are needed to do the calibration. The factors that affect the accuracy of calibration are analyzed. Experiments with POV-Ray images validate the expected performance of the design and also show that sphere-based calibration method is practical and effective.

REFERENCES

[1] Z. Taha, J. Y. Chew and H. J. Yap, "Omnidirectional vision for mobile robot navigation," Journal of Advanced Computational Intelligence and Intelligent Informatics, vol.14, no. 1, 2010, pp. 55–62.

[2] S. Hrabar and G. S. Sukhatme, Omnidirectional vision for an autonomous helicopter. in *Proc. IEEE Int. Conf. Robotics and Automation*, Taiwan, 2004, pp. 558–563.

[3] H. Koyasu, J. Miura, and Y. Shirai, "Recognizing moving obstacles for robot navigation using real-time omnidirectional stereo vision," *Robotics and Mechatronics*, vol. 14, no. 2, pp. 147–156, 2002.

[4] S. Lin and R. Bajcsy, "High resolution catadioptric omni-directional stereo sensor for robot vision", in *Proc. IEEE Int. Conf. Robotics and Automation, Taipei*, 2003, pp. 1694-1699.

[5] E. L. L. Cabral, J. C. de Souza Junior and M. C. Hunold, "Omnidirectional stereo vision with a hyperbolic double lobed mirror," in *Proc. 17th Int. Conf. Pattern Recognition,* Cambridge, 2004, pp 1-4.

[6] L. Su and F. Zhu, "Design of a novel stereo vision navigation system for mobile robots," in *Proc. IEEE Int. Conf. Robotics and Biomimetics*, Hong Kong and Macao, 2005, pp. 611–614.

[7] S. Yi and N. Ahuja, "An omnidirectional stereo vision system using a single camera," in *Proc. 17th Int. Conf. Pattern Recognition*, Hong Kong, 2006, vol. 4, pp. 861–865.

[8] C. Luo, L. Su, F. Zhu and Z. Shi, "A versatile method for omnidirectional stereo camera calibration based on BP algorithm," *Lecture Notes In Computer Science*, vol. 3972, pp. 383–389, 2006.

[9] X. Ying and Z. Hu, "Catadioptric Camera Calibration Using Geometric Invariants," *IEEE Trans. PAMI* ,vol. 26, no. 10, pp. 1260–1271, 2004.

[10] L. Zhang, X. Du, Y. Zhu and J. Liu, "Central catadioptric camera calibration with single image," in *Proc. 34th IEEE int. Conf. Acoustics, Speech and Signal*, Taiwan, 2009, pp. 1253–1256.

[11] C. Geyer and K. Daniilidis, "A unifying theory for central panoramic systems and practical implications," in *Proc. 6th European Conf. on Computer Vision*, Dublin, 2000, pp. 445–461.

[12] S. Baker and S. K. Nayar, "A theory of single-viewpoint catadioptric image formation," *Int. Journal of Computer Vision*, vol. 35, no. 2, pp. 1–22, 1999.

[13] J. Gaspar, C. Decco, J. Okamoto Jr. and J. Santos-victor, "Constant resolution omnidirectional cameras", in *Proc. 3rd workshop on omnidirectional vision,* Copenhagen and Denmark, 2002, pp. 27–34.

[14] H. Nagahara, K. Yoshida and M. Yachida, "An omnidirectional vision sensor with single view and constant resolution", in *Proc. IEEE 11th Int. Conf. Computer Vision,* rio de janeiro, Brazil, pp. 1–8, 2007.

[15] A. Fitzgibbon, M. Pilu and R. B. Fisher, "Direct least squares fitting of ellipses," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 5, pp. 476–480, 1999.

[16] K. Madsen, H. B. Nielsen and O. Tingleff, "Methods for non-linear least squares problems," (2nd ed.) [Online]. Available: http://www2.imm.dtu.dk/pubdb/views/edoc_download.php/3215/pdf/imm3215.pdf.