To appear in Proceedings of the IEEE/RSJ Internationals Conference on Intelligent Robots and Systems
Edmonton, Alberta, Aug. 2-6, 2005

Calibrating an Active Omnidirectional Vision System

Nicholas D. Jankovic and Michael D. Naish
Dept. of Mechanical and Materials Engineering
The University of Western Ontario
London, Ontario, Canada
ndjankov@uwo.ca, naish@eng.uwo.ca

***Abstract* – This paper describes a straightforward process for calibrating an active vision system containing both pinhole perspective and omnidirectional cameras. The perspective cameras can be easily calibrated using standard methods. Unfortunately, these methods are not suitable for omnidirectional cameras. Methods that rely on iterative least squares optimization, using a set of known image-world correspondences, are adopted for omnidirectional cameras. To ensure unbiased estimation of camera parameters, an omnidirectional calibration rig is employed so that nearly the entire field of view contains known calibration points. Measurement uncertainties collected from each stage of calibration are then combined to estimate the overall system uncertainty. This calibration process is evaluated experimentally by estimating the location of known points using triangulation, where the results achieved are comparable with the estimated system uncertainties.**

Index Terms – Calibration, triangulation, omnidirectional cameras and active vision.

I. Introduction

Conventional cameras have been applied to many diverse areas over the past several decades, some of which include machine vision, robot navigation and surveillance. In recent years, omnidirectional cameras are finding increasing application in many visual tasks. These cameras are typically able to view an entire hemisphere, cylinder or more. Having such a large field of view (FOV) provides advantages in some situations, improving tasks such as motion estimation or target detection. A modular vision system composed of both standard and omnidirectional cameras has been developed so that the benefits of each type may be utilized [1]. The goal of this system is to detect and track targets anywhere within its FOV and estimate their positions in space. Such a system can be used in a variety of surveillance situations for tracking people. It is also applicable to mobile robot vision for providing more robust localization and obstacle avoidance.

Any vision system that intends to extract information from its images must undergo calibration to some degree. Most often, a standard camera can be approximated well by the pinhole camera model and there have been many calibration methods developed to extract the camera's parameters [2–6]. Approaches for calibrating standard cameras are common and many methods have freely available implementations [7–11]. Unfortunately, these methods are not suitable for omnidirectional cameras due to their geometry and large distortions. Existing omnidirectional calibration methods have not yet evolved to the same level of support.

This paper describes the approach taken to calibrate a system composed of standard and omnidirectional cameras. The goal is to illustrate how an omnidirectional system can be calibrated using conceptually simple and straightforward means, while maintaining good accuracy. The following section describes the vision system in more detail. Section III discusses the applied calibration approach. Calibration results and analysis are presented in Section IV, followed by conclusions and recommendations in Section V.

II. System Description

The vision system under study is composed of three different camera modules aligned vertically to form a camera tower, as shown in Fig. 1. The bottom module is a catadioptric camera, which combines a standard camera pointed up at a convex mirror, providing a FOV greater than a hemisphere [12]. At the top of the stack is a fisheye module, which provides an unobstructed FOV slightly greater than a hemisphere and is composed of a standard camera looking through a fisheye lens converter [13]. A nearly spherical peripheral view can be obtained by arranging these two omnidirectional cameras in opposing directions. The third module has a standard perspective camera fixed to an active pan/tilt platform, which allows it to pan continuously, enabling a high level of detail to be resolved in any direction.

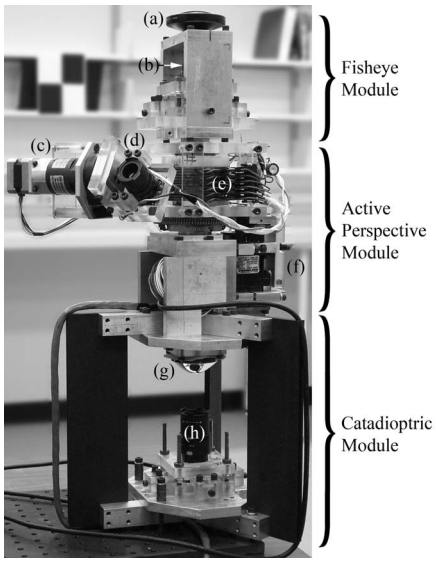

Fig. 1 Initial prototype. a) Fisheye lens, b) 4mm camera, c) tilt motor system, d) 12mm camera, e) slip ring, f) pan motor system, g) hyperbolic mirror, h) 3.6mm camera.

Determining the position of a point using vision requires that two cameras (with a known baseline) see the same point or that one camera sees a point from two different (known) positions. Finding the point in each of the two images that relates to the real world point is commonly referred to as the correspondence problem [2]. Using a vertical system layout helps to simplify this problem since the active camera can be aligned relative to the target as seen in an omnidirectional image. Once the two cameras see the same point, they can then estimate its position using triangulation. The system can potentially be extended to include more than one of each camera type, allowing three basic triangulation modalities: 1) between two omnidirectional cameras, 2) between two active cameras and 3) between an omnidirectional camera and an active camera. The focus herein will be on the latter approach.

III. Calibration Process

The calibration process discussed herein is intended to be conceptually simple and easy to implement. This process is applicable to general omnidirectional systems containing any combination of fisheye, catadioptric and perspective cameras. First, each individual camera is calibrated with respect to the pinhole model. Then the omnidirectional cameras are aligned with their accompanying optics. Following the collection of image data points using a calibration rig, the internal parameters are determined for the fisheye and catadioptric cameras using their respective methods. External parameters are then determined and, finally, the motion system can be calibrated.

A. Pinhole Calibration

The first step is to determine each camera's internal parameters with respect to the pinhole model [2]. After positioning and focusing an omnidirectional camera, the perspective camera portion is separated from its accompanying optics (i.e., fisheye lens or convex mirror). The focusing and zooming rings should be locked to prevent the internal parameters from changing. A number of methods are available for calibrating a pinhole camera, but the method developed by Zhang [6] and implemented by Jean-Yves Bouguet as part of the Intel OpenCV Library [9] or as a Matlab Toolbox [7] is adopted herein. This implementation provides good results while being quick and simple to use. All that is needed are several images of a planar checkerboard pattern.

B. Camera Alignment

The internal pinhole camera parameters allow the camera to be precisely aligned with a mirror or fisheye lens. Aligning the camera with a fisheye lens converter is straightforward. Covering the lens with a thin tissue results in an image with a large white circle and a black border. Fitting an ellipse to the circular fisheye image allows the lens to be adjusted so that it coincides with the calibrated image centre. Ellipse fitting also allows the camera orientation to be determined and adjusted. When the ellipse becomes a circle, then the principal axes of the camera and fisheye lens will be collinear.

Aligning a catadioptric camera becomes slightly more involved if a single effective viewpoint is to be preserved. This requires that the camera centre must coincide with the second focal point of the hyperbolic mirror. To capture a circular image, the mirror is placed in front of a black background and surrounded with thin white paper so that the background remains visible while the reflection at the mirror edge remains white. Ellipse fitting is then used to align the camera; however, unlike the fisheye case, the size of the ellipse must also be matched to determine the appropriate distance from the mirror. Fig. 2 shows how the camera parameters and mirror diameter are used to determine the distance using simple geometry. The radius r_i of the mirror rim in ideal pixel coordinates is found using similar triangles (Eq. 1). It is assumed that the mirror's radius r_m, semiaxis a and semiaxis b are known. The mirror rim is located at a distance z (Eq. 4) above the mirror centre and the rim must be located at a distance h (Eq. 2) from the camera centre [12]. Here, e (Eq. 3) represents the eccentricity of the mirror and f is the calibrated focal length in pixel coordinates.

$$r_i = (r_m / h) f \tag{1}$$

$$h = 2e + z \tag{2}$$

$$e = \sqrt{a^2 + b^2} \tag{3}$$

$$z = \sqrt{a^2 \left(r_m^2 / b^2 + 1 \right)} - e \tag{4}$$

The resulting ellipse must be distorted from the ideal pinhole coordinates to that of the captured image to ensure correct distance representation. These distortion parameters are given in the calibration results from Step A. By continually fitting ellipses to a live video stream, the camera can be adjusted until the desired ellipse has been achieved.

C. Calibration Rig

In order to calibrate an omnidirectional camera, structured data must be extracted from the image. In the context of this paper omnidirectional cameras are assumed to be stationary, hence, calibration methods based on motion are not viable.

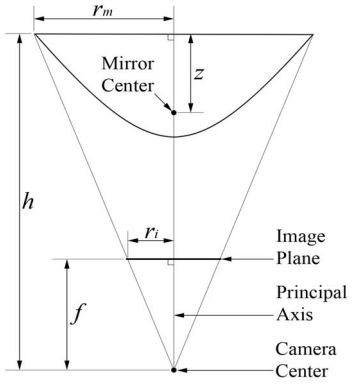

Fig. 2 Catadioptric mirror alignment geometry.

The remaining methods involve computations based on known points, similar to many classical calibration methods [2–5] or based on lines or circles [14, 15]. The former approach is simpler conceptually and is easily implemented using a general optimization framework.

Generating precise point correspondences requires an accurate calibration rig. Ideally the rig provides known world points that can be sampled evenly across the entire omnidirectional FOV, ensuring uniformity. A very large planar rig is not suitable in this case because the FOV is greater than 180° and no points would be observed near the image circumference. An alternative is to use many smaller targets distributed precisely throughout the environment, but this approach involves a great deal of effort and time.

An example of a suitable calibration rig is shown in Figure 3. It is composed of two thin perpendicular planes fixed to a precision rotation stage, allowing a column of points to be swept across nearly the entire FOV for both omnidirectional camera types. This approach is similar to the one used in [13], where only a single plane is swept across the FOV. The required accuracy of the rig largely depends on the resolution of the cameras being calibrated, as well as its distance from the camera. Extracting points from the images may have to be done manually since distortions can prevent conventional automated methods from locating the grid intersections. If the cameras provide sufficient resolution, then a semi-automated approach can be applied.

D. Fisheye Calibration Method

Fisheye lenses inherently demonstrate significant amounts of distortion. It may not always be possible to obtain lens equations from the manufacturer, thus numerical estimation of the parameters is required. Since the pinhole camera model cannot be applied to cameras with a FOV greater than 180°, another camera model must be used. Bakstein and Pajdla [13] have developed a hybrid model that combines both the stereographic and equisolid camera models (Eq. 5).

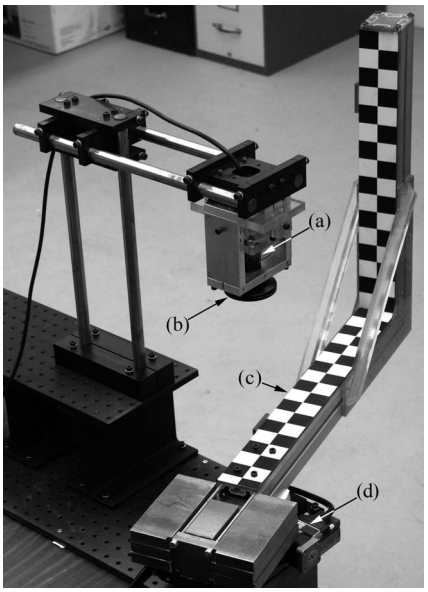

Fig. 3 Fisheye camera calibration set-up. The perspective camera (a) and fisheye lens converter (b) are pointed downward so that the calibration planes (c) are clearly visible. The precision rotation stage (d) sequentially moves the rig to known positions. A similar approach is also used for the catadioptric camera.

$$r = a\left(\tan\frac{\theta}{b}\right) + c\left(\sin\frac{\theta}{d}\right). \tag{5}$$

The radial distance r of a point on the image plane is a function of the angle θ that is made between the incident ray and the principal axis. The four parameters, in addition to six external parameters (for both rotation and translation), are determined by minimizing the distance between the imaged point and its corresponding mapped world point using the Levenberg-Marquardt (LM) non-linear least-squares algorithm [16]. Implementations of the LM algorithm can be found for MATLAB [17] and C/C++ [18], among others.

E. Catadioptric Calibration Method

A catadioptric camera is calibrated in much the same manner as a fisheye camera, except that the equations differ. Instead of mapping incoming light rays directly to the image plane, the world points are mapped to the hyperbolic mirror surface (Eq. 6) using a non-linear function F^+ (Eq. 7) [12].

$$\mathbf{X}_h = F^+(\mathbf{X}_m). \tag{6}$$

Let $\mathbf{v}$ represent any point vector $\mathbf{X}_m$, so that:

$$F^+(\mathbf{v}) = \frac{b^2\left(e\mathrm{v}_3 + a\|\mathbf{v}\|\right)}{b^2\mathrm{v}_3^2 - a^2\mathrm{v}_1^2 - a^2\mathrm{v}_2^2}, \tag{7}$$

where a and b are the semiaxis mirror parameters and e is the eccentricity defined in Eq. 3. Here, $\mathbf{X}_m$ represents the target point located on the calibration rig with respect to the mirror's focal point and $\mathbf{X}_h$ is the corresponding point on the hyperbolic surface. This point is then mapped to the image plane using the pinhole camera model, calibrated in Step A. The catadioptric parameters are then refined through minimization using the LM algorithm mentioned in Step D.

F. Relative Camera Positions

Next, the external parameters must be determined for a completed assembly; these relate each camera's position (or baseline) to the other cameras in the system. This is accomplished by having each camera observe and acquire calibration points relative to a global coordinate system. The external parameters are then solved using the calibration methods in Steps D and E, while the internal parameters remain fixed. A large planar calibration pattern can be suspended such that all of the cameras can extract points from it. An example of such a pattern is shown in Figure 6. Experiments have shown that good estimates for camera orientation are obtained this way, but the relative positions can only be determined within centimetre range accuracy. Therefore, it is better to measure the distances between the

camera centres if the cameras are within close proximity; in this case accuracies in the millimetre range can be attained.

G. Motion System Calibration

Calibration of the drive system is also necessary if an active camera is to provide reliable orientation readings. Angular accuracy relies mostly on the optical encoder or stepper motor resolution. If precise specifications are not known or depending on the type of motor system that is used, then it may be necessary to calibrate the pan and tilt axes of the active cameras. Any additional gearing may increase the backlash, which will also have to be accounted for. There are two measures of particular interest for each axis: the resolution and repeatability. There are a number of ways to calibrate the active camera motors. One approach is to use a laser pointer mounted to the camera platform, which is then jogged by known amounts. The performance can then be quantified by casting the dot onto a ruler at a known distance,.

These simple steps for calibration provide a useful starting point and guide for those who are new to the area of omnidirectional vision and wish to develop systems incorporating these camera technologies. An advantage of this process is that there are no major technical or theoretical difficulties, making the software easy to implement.

IV. Results and Analysis

This section discusses the results of the calibration approach described in Section III when applied to the omnidirectional vision system described in Section II. Simple Firewire board cameras are used to extract 640×480 pixel images. The fisheye camera is composed of a 4mm monofocal lens looking through a Nikon FC-E8 fisheye lens converter, where the catadioptric camera observes the reflection from a Neovision H3G hyperbolic mirror using a varifocal lens set to 3.6mm. The active perspective camera simply uses a 12mm monofocal lens. The pan and tilt axes are both equipped with 8000 pulse per revolution incremental encoders for positional feedback.

During the calibration process, the internal parameters for each camera are estimated; however, it is also desirable to evaluate the accuracy of these parameters. When conducting triangulation, usually the target's depth is the quantity of most interest; therefore, accuracy is evaluated in relation to the estimated depth. The next section describes the geometric model used to compute the theoretical position errors, followed by the parameter uncertainties for each camera, which are derived from the calibration results. Finally, the triangulation experiment is described and comparisons are made between theoretical and experimental errors.

A. Theoretical Error

By knowing the relative geometry of each camera and their internal parameters, it is possible to project a ray from the camera centre in the direction the target. However, measurement errors prevent the projected rays from intersecting at the target's exact location in space. Figure 4 illustrates the basic triangulation geometry, where the active and fisheye camera pair is used as an example. The positive error is defined as the difference between the maximum and actual depth estimations, which is a function of the ray angles β and θ for each active and omnidirectional camera, respectively. Similarly, the negative error relates to the minimum depth estimation. Ray angle variations are expressed as the angular uncertainties, $\delta\beta$ and $\delta\theta$. Uncertainties also apply to the relative camera positions (i.e., δZ and δR). However, these values are typically in the order of millimetres and have a negligible effect on the depth error estimation when the target distance is in the order of metres; therefore, the position uncertainty values can be excluded.

B. Calibration Uncertainty

1) Active Camera: This camera is calibrated as described in Section III-A. Taking the calibration points and projecting them onto the image plane allows the standard deviation to be determined for the pixel error, being the difference between a measured and projected image point. For the active camera, the standard deviation is found to be 0.22 pixels. A reasonable uncertainty, for projecting a ray though an image point, is three times the standard deviation, $\pm 3\sigma$. This corresponds to about $\pm 0.017°$ when converted to an angular uncertainty. Since the camera is rigidly attached to the tilt encoder axis, which itself has an uncertainty of $\pm 0.045°$, the total angular uncertainty, $\delta\beta$, for the ray is about $\pm 0.062°$.

2) Omnidirectional Cameras: In a similar manner to that of the active camera, the pixel uncertainties are computed for the omnidirectional cameras using the calibration results obtained in Sections III-D and III-E. Again, three standard deviations are used to determine the uncertainties, which are summarized in Table I. Two sets of values are given because the angular resolution for each camera changes non-linearly over its respective field of view. Therefore, the angular uncertainty near the image centre will differ from the uncertainty near the image circumference. Additional angular uncertainty is introduced by misalignment that may remain

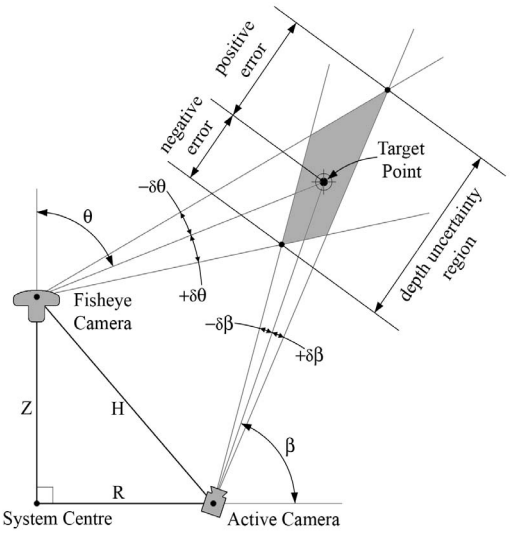

Fig. 4 Triangulation geometry of projected image rays.

TABLE I
RESOLUTION AND UNCERTAINTY FOR OMNIDIRECTIONAL CAMERAS

	Fisheye Camera		Catadioptric Camera	
	At 0°	At 91°	At 27°	At 104°
Standard Deviation	0.33 pix	0.33 pix	0.44 pix	0.44 pix
Angular Resolution	0.38°/pix	0.45°/pix	0.63°/pix	0.27°/pix
Angular Uncertainty	±0.38°	±0.45°	±0.83°	±0.36°
Orientation Uncertainty	±0.06°	±0.06°	±0.06°	±0.06°
Total Uncertainty	±0.44°	±0.51°	±0.89°	±0.42°

following the calibration discussed in Section III-F. This uncertainty corresponds to approximately ±0.06° for both the fisheye and catadioptric camera.

3) Uncertainty Region: Referring to Figure 4, the depth uncertainty region is the absolute sum of the positive and negative errors. This region can be computed for the working range of each active and omnidirectional camera pair using the angular uncertainty values given above. Figure 5 shows how the size of this region varies for an active and fisheye camera pair as a function of the target's distance from the system centre and its angle with respect to the system's vertical axis. In other words, Figure 5 demonstrates that the estimation error varies with the target's position. The uncertainty increases with increasing distance, as well as with decreasing angle between the two projected rays. The angle between the two rays approaches zero as they become collinear with the hypotenuse, H, at which point depth can no longer be estimated. The non-linear omnidirectional resolution also influences the surface shape.

C. Triangulation Experiment

A large calibration pattern is used to obtain images of corresponding target points. In fact, this is the same pattern used to determine the omnidirectional camera orientation in Section III-F. Each corner of the pattern is known accurately within about 1 mm relative to the board and the position of the board is known accurately within 1 cm relative to the system. The pattern is placed at three different distances from the system in order to obtain points over a large range. Figure 6 illustrates the experimental setup used. Note that, a transit level is used to ensure proper orientation and alignment of the target board after repositioning.

Triangulation of the intersection point for two rays is typically computed as the midpoint of the shortest line joining each ray. This approach produces the best results when using two identical cameras. Since the angular precision of the active camera is far greater than either omnidirectional camera, then the direction of the projected active camera ray is also known very accurately. Assuming that the active camera ray intersects the target, then the target position is estimated as a point lying on the active camera ray that is nearest to the omnidirectional camera ray; therefore, the omnidirectional ray only contributes to the depth component of the estimation.

D. Error Analysis

Calibration accuracy can be assessed by comparing experimental triangulation data with the theoretical error limits over the operating range of each camera pair. Figure 7

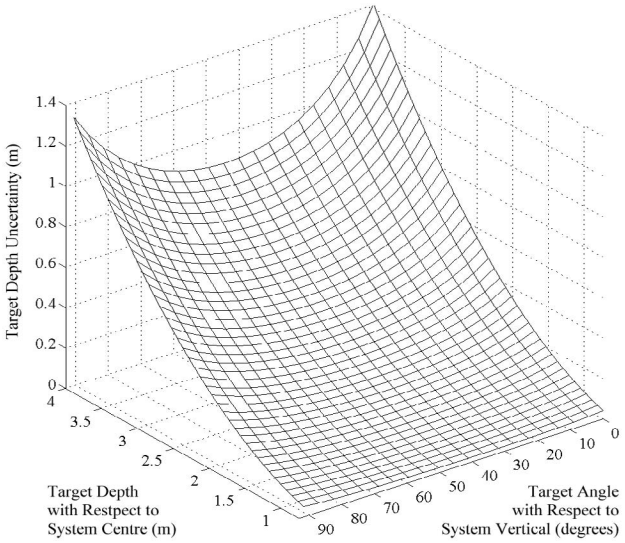

Fig. 5 Depth uncertainty region. This plot illustrates how the triangulation accuracy changes over the working range of an active and fisheye camera pair.

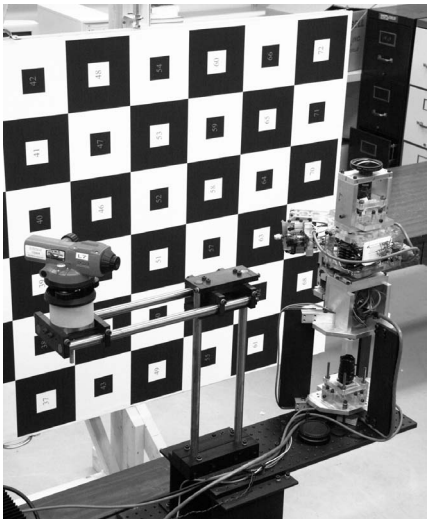

Fig. 6 Experimental setup for triangulation.

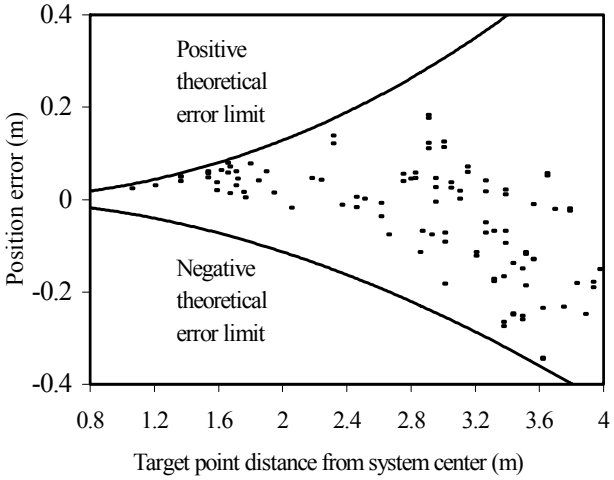

Fig. 7 Comparison of theoretical error limits and experimental depth error points for the active-fisheye camera pair.

shows sample data taken from the active camera and fisheye camera pair. The data points shown correspond to actual errors for target points ranging between 70° and 80° from the vertical axis. This range simplifies the presentation, while displaying a sizable number of data points. The two curves correspond to the positive and negative error limits for a vertical angle for 70°. From the figure, it is seen that the experimental data is within the theoretical error limits, where even points approaching 80° are still within the error limits of 70°. These results are typical over the operating range of the system except when the points are at distances less than 1.2m. In this case, a few points exceed the limits, likely due to the active camera's inability to clearly focus on near objects. The placement of the target points is assumed to be exact because the positional uncertainty of the target is an order of magnitude less than the depth estimation uncertainty.

Even though the greatest error occurs in depth (i.e., extending radially from the system centre), the accuracy in the orthogonal (i.e., tangential) directions is also considered. According to Section IV-C, the target point coincides with the projected active camera ray; however, observed angular errors reach almost 5 times that of the total angular uncertainty predicted in Section IV-B. The active camera's ability to resolve fine detail allows it to exceed the accuracy of the target placement. More precise positioning of the target would greatly reduce this tangential error, while having minimal effects on the depth error.

Nevertheless, the worst-case test data shows a maximum tangential error of 0.02 m, corresponding to a distance of 3.9 m from the system. Even though this error is larger than anticipated, it is an order of magnitude smaller than any related depth estimations; thus, showing that depth is still the limiting factor in estimating a target's 3D position. Similar results are also observed for the catadioptric and active camera pair, but are not shown to simplify presentation.

V. Conclusions

This paper describes a straightforward process for calibrating a vision system containing both the perspective and omnidirectional cameras. The perspective cameras are calibrated using a standard and readily available method, whereas the omnidirectional cameras are calibrated through optimization of points taken from a calibration rig. This rig can be used interchangeably between the fisheye and catadioptric modules. It also provides image points that cover the entire omnidirectional FOV, especially near the circumference where large distortions occur, allowing accurate estimation of calibration parameters. These parameters are then solved using a general LM non-linear least-squares optimization framework.

The above triangulation results indicate that some residual systematic errors remain after calibration, showing that some further improvements can still be made to the calibration process. However, these results show that the position estimations are within the specified uncertainties of the system hardware and test equipment. If greater accuracy is required, then higher resolution cameras and encoders should be used.

The consequence of using a conceptually simple calibration approach is that the task becomes significantly more labour intensive, largely due to the reliance on classical calibration methods that require known point data. It may be the case that this point data must be entered manually when insufficient camera resolution prevents automatic extraction. The calibration process described in this paper also relies heavily on precisely constructed calibration targets, which may increase costs. This calibration process is specifically intended for use within a laboratory setting for evaluating the potential performance of an omnidirectional vision system so that subsequent development and analysis can be carried out.

References

[1] N.D. Jankovic and M.D. Naish, "Developing a modular active spherical vision system," in *IEEE International Conference on Robotics and Automation (ICRA'05)*, Barcelona, Spain, pp. 1246–1251, April 2005.

[2] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge, UK: University Press, 2004.

[3] J. Heikkilä, "Geometric camera calibration using circular control points," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1066–1077, 2000.

[4] R.K. Lenz and R.Y. Tsai, "Techniques for calibration of the scale factor and image centre for high accuracy 3-D machine vision metrology," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 10, no. 5, pp. 713–720, 1988.

[5] J. Salvi, X. Armangue and J. Batlle, "A comparative review of camera calibrating methods with accuracy evaluation," *Pattern Recognition*, vol. 35, pp. 1617–1635, 2002.

[6] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 22, no. 11, pp. 1330–1334, 2000.

[7] J. Bouguet, *Camera Calibration Toolbox for Matlab*, Oct. 15, 2004, [on-line] Available: http://www.vision.caltech.edu/bouguetj/calib_doc/

[8] J. Heikkilä, *Camera Calibration Toolbox for Matlab*, Oct 17. 2000, [on-line] Available: http://www.ee.oulu.fi/~jth/calibr/

[9] Intel Research, *Open Computer Vision Library*, Aug. 18 2004, [on-line] Available: http://sourceforge.net/projects/opencvlibrary/

[10] Reg Willson, *Tsai Camera Calibration Software*, Oct. 28, 1995, [on-line] Available: http://www-2.cs.cmu.edu/afs/cs.cmu.edu/user/rgw/www/TsaiCode.html

[11] Z. Zhang, *A Flexible New Technique for Camera Calibration*, Aug. 10, 2002, [on-line] Available: http://research.microsoft.com/~zhang/calib/

[12] T. Svoboda, T. Pajdla and V. Hlavac, "Epipolar geometry for panoramic cameras," in *Proc. of the Fifth European Conference on Computer Vision*, Freiburg, Germany, pp. 218–232, June 1998.

[13] H. Bakstein and T. Pajdla, "Panoramic mosaicing with a 180º field of view lens," in *Proc. of Third Workshop on Omnidirectional Vision*, pp. 60–67, June 2, 2002.

[14] R. Swaminathan and S.K. Nayar, "Nonmetric calibration of wide-angle lenses and polycameras," *IEEE Transactions of Pattern Analysis and Machine Intelligence*, vol. 22, no. 10, pp. 1172–1178, 2000.

[15] X. Ying and Z. Hu, "Catadioptric camera calibration using geometric invariants," in *Proc. of Ninth IEEE Intl. Conf. on Comp. Vision*, vol. 2, pp. 1351–1358, Oct. 13–16, 2003.

[16] K. Madsen, H.B. Nielsen and O. Tingleff, *Methods for Non-Linear Least Squares Problems*, 2nd ed., Lyngby, Denmark, Technical University of Denmark, 2004.

[17] MathWorks, *MATLAB 7*, Jan 28, 2005, [on-line] Available: http://www.mathworks.com/

[18] M. Lourakis, *Levenberg-Marquardt non-linear least squares algorithms in C/C++*, Nov. 23, 2004, [on-line] Available: http://www.ics.forth.gr/~lourakis/levmar/