

Omnidirectional Depth Computation from a Single Image*

Radu Orghidan[†] and El Mustapha Mouaddib[‡]

[‡]*Centre de Robotique*

Electrotechnique et d'Automatique

Université de Picardie Jules Verne, Amiens, France.

mouaddib@u-picardie.fr

Joaquim Salvi[†]

[†]*Computer Vision and Robotics Group*

Institute of Informatics and Applications

University of Girona, Girona, Catalonia, Spain.

{radu,qsalvi}@eia.udg.es

Abstract—Omnidirectional cameras offer a much wider field of view than the perspective ones and alleviate the problems due to occlusions. However, both types of cameras suffer from the lack of depth perception. A practical method for obtaining depth in computer vision is to project a known structured light pattern on the scene avoiding the problems and costs involved by stereo vision. This paper is focused on the idea of combining omnidirectional vision and structured light with the aim to provide 3D information about the scene. The resulting sensor is formed by a single catadioptric camera and an omnidirectional light projector. It is also discussed how this sensor can be used in robot navigation applications.

Index Terms—catadioptrics, omnidirectional vision, calibration, structured light, 3D reconstruction

I. INTRODUCTION

The omnidirectional vision sensors enhance the field of view of traditional cameras by means of special optics, structures of still or gyrotory cameras or combinations of lenses and mirrors. Yagi [21] surveyed the existing techniques for building cameras with a wide field of view and Svoboda [19] proposed several classifications of the existing omnidirectional cameras according to their most important features.

The catadioptric sensors use at least one mirror coupled to a conventional camera. The catadioptric cameras can be classified depending on the way they gather the light rays. When all the observed light rays cross into a point, called focus, the sensors are known as Single View Point (SVP). The class of catadioptric sensors that have a SVP was derived by Baker and Nayar [1]. The SVP is a desirable property that enables distortion-free reconstruction of panoramic images in a familiar form for the human users. The catadioptric sensors that do not possess a single focal point (non-SVP) are less used but proved to be helpful for applications with specific requirements such as prescribed distortions [10] or with linear projection constraints [4]. Still, neither the standard cameras nor the catadioptric ones can provide depth information of the scene when used independently.

Stereoscopic vision combines separate images taken from distinct points of view and permits to visually per-

ceive depth. Stereo catadioptric sensors are special structures of mirrors and lenses designed for obtaining depth from images with a wide field of view. In order to obtain distinct points of view of the scene the camera is pointed towards a structure of convex [13] [9] [2] or planar [8] mirrors. The 3D information is obtained by triangulation. This method leans on the assumption that the correspondences of the points between the observed images can be accurately found. However, correspondence matching is deteriorated in the case of catadioptric sensors because the resolution of the omnidirectional images is lower than the resolution of the conventional ones since the number of the scene points is significantly different while both images are represented using the same number of pixels. A solution to this problem is the use of a structured light pattern projected onto the scene [18] [17]. Using this technique is similar to placing visible landmarks in the scene so that image points can be identified and matched faster.

It is noticeable that the use of 360 degrees images and of scene-depth information is ideal for robot navigation tasks. An efficient approach to the navigation problem is to choose a reasonable tradeoff between localization accuracy and travelled distance. Similarly, in real life a high precision of movements is required in small spaces such as narrow halls or offices while a low accuracy is needed in wide areas where the possible obstacles are more spaced. Consequently, robot navigation can be divided in two branches as shown by Gaspar [20]: topological navigation and visual path following. Topological navigation gives a qualitative characterization of the robot's global position using omnidirectional images. This technique is suitable for navigation tasks that implies travelling large distances with low precision of localization. Visual path following is mainly used for short-distance segments that require high accuracy navigation.

The goal of this paper is to present an omnidirectional sensor that provides 3D information using a single camera. From this point of view, a robot navigation application is also discussed. The sensor is formed by a single-camera catadioptric configuration with an embedded omnidirectional structured light projector. By mounting the omnidirectional sensor on a mobile robot applications such as 3D map building, robot navigation and localization, active surveillance with real-time object detection or 3D

*This work is partially supported by the Spanish project CICYT TIC 2003-08106-C02-02 and by the AIRE mobility grant provided by the Generalitat of Catalunya that allowed a four month stay in the CREA lab from Amiens, France

reconstruction can be performed within a horizontal field of view of 360 degrees. The sensor design and the calibration of the whole system are detailed in section II. The experimental results are shown in section III. The article ends with conclusions.

II. SENSOR DESIGN AND CALIBRATION

As explained in the introductory section, the main goal of the proposed sensor is to retrieve 3D information from an omnidirectional image overcoming the matching problem risen in stereo vision. The solution, presented in Figure 1, is to couple to the omnidirectional camera a structured light projector with a field of view of 360 degrees. A more compact sensor can be build by placing the light projector within the blind zone of the omnidirectional camera as shown in [14] where a similar sensor was described and analyzed by simulation. However, for the realization of the first prototype of the physical sensor the two parts have been separated for more maneuverability.

The third dimension of the scene is perceived by crossing the light rays emitted by the laser with the ones observed by the camera, in other words, performing a triangulation. The points of the pattern projected by the laser are reflected by the conical mirror and become light-spots on the scene. The parabolic mirror reflects the whole scene into the camera and all the spots enlightened by the laser can be immediately identified (more details in section III). With the models for both components of the sensor a precise triangulation can be carried out. In the following, we describe the modelling process and parameters estimation (calibration).

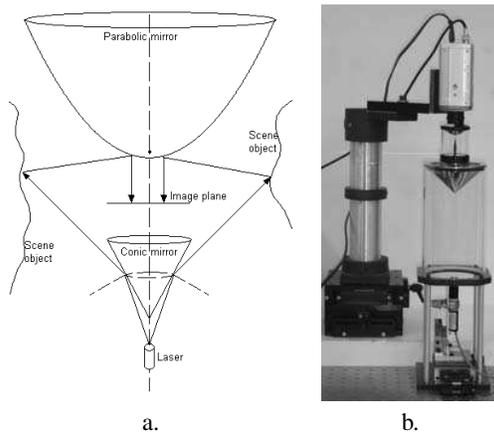


Fig. 1. a. Catadioptric omnidirectional camera with embedded structured light projector. b. Laboratory prototype.

The traditional approach for calibrating a structured light system takes two steps. The camera is calibrated at first and the light projector is subsequently calibrated based on information provided by the camera. In the case of structured light systems using perspective cameras Huynh [11] and Forest [3] proposed calibration methods based on the cross ratio invariance under perspective projection. Their

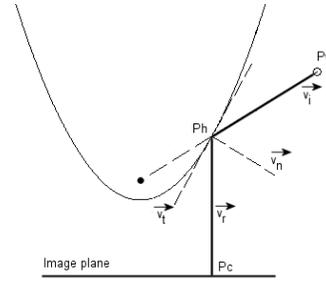


Fig. 2. Image formation in a catadioptric camera composed by a paraboloidal mirror with an camera provided with telecentric lenses.

methods provide a direct image to world transformation matrix for the points that belong to the light stripes of the laser projector. Since our intention was to model the light projector and the camera independently the two steps calibration method was preferred.

A. Omnidirectional Camera Calibration

Two different models have been tested for the omnidirectional camera used for building the sensor. The first model takes into account the fact that the reflecting surface is known (parabolic mirror) and the camera is provided with orthographic lenses. The second model is based on the assumption that the pair camera-mirror possesses a SVP and consequently can be represented by the equivalent projection on the sphere defined by Geyer and Daniilidis [7] [6]. Both models were used for remapping the image points onto the scene and the outcomes of the models were compared as shown in section III.

The first model which considers the image formation based on the reflection of light into a paraboloidal surface assumes that the mirror is perfectly specular i.e. the angle of incidence is equal to the angle of reflection. Note that, by using this method, the SVP property becomes a consequence and not a constraint which means that the model can be applied for any kind of mirror shapes no matter if they possess a SVP or not.

Consider P_w a point on the scene, P_c the point in the image plane and the function describing the mirror surface as shown in Figure 2. The rotational symmetry of the paraboloid permits, without loss of generality, to reduce the study of the optical behavior of such a surface from three-dimensional Cartesian coordinates system (x, y, z) to a two dimensional one (r, z) , where $r = \sqrt{x^2 + y^2}$. Equation (1) models the parabolic shape in the (r, z) coordinate system.

$$z = \frac{r^2}{4 \cdot a} \quad (1)$$

Let P_h be the point on the mirror where the point P_w is reflected i.e. P_h belongs to the tangent \bar{v}_t at the mirror surface whose normal \bar{v}_n is the bisecting line of the angle $\widehat{P_w P_h P_c}$. Also said, the incident vector $\bar{v}_i = \overline{P_h P_w}$ is the symmetric of the reflected vector $\bar{v}_r = \overline{P_c P_h}$ with respect to \bar{v}_n . This relation is described by equation (2).

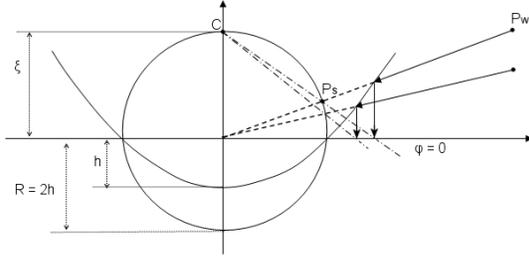


Fig. 3. Image formation using the projective equivalence of a SVP catadioptric projection with the projection on the sphere.

$$\bar{v}_i = 2 \cdot (\bar{v}_r \cdot \bar{v}_n) - \bar{v}_r \quad (2)$$

Consequently, in the case of a parabolic mirror with telecentric lenses, for finding the image of a point P_w with known coordinates the problem reduces to finding the point P_h on the mirror surface so that the vectors \bar{v}_i and \bar{v}_r satisfy equation (2). The same method can be adapted for the case of a hyperbolic mirror with a perspective camera and it was described by Gaspar in [5]. Considering the world points with known coordinates, placed onto a calibration pattern, their projection onto the image plane can be calculated using an initial guess for the parameters of the model. The difference between the positions of the calculated image points and the positions of the real image points is the calibration error of the model. Minimizing the above error by means of an iterative algorithm such as Levenberg-Marquardt the omnidirectional camera is calibrated. The parameters of the model are: a , the coefficient of the parabola; $\alpha_u, \alpha_v, u_0, v_0$, the intrinsic camera parameters; $r_X(\phi), r_Y(\theta), r_z(\varphi)$, and t_x, t_y, t_z , the six extrinsic parameters that model respectively the orientations and the translations between the world coordinate system and the mirror coordinate system. The orientation vectors are functions of the three angles (ϕ, θ, φ) which define the rotation on each axis and are expressed in radians while the translations are measured in millimeters, as detailed in [16].

The previously described method provides a simple way to model a catadioptric camera but has two drawbacks: it assumes that the mirror is perfectly parabolic and it requires a high computation time.

The next model for the omnidirectional camera stands on the projective equivalence between any central projection and the projection onto the sphere followed by the projection to a plane from a point [6]. Remark that this model does not assume that the mirror shape is known but considers that the omni-system possesses a SVP.

If $P_w = [x_w, y_w, z_w]$ is a scene point and $P_s = [x_s, y_s, z_s]$ is the intersection of the light ray emitted by the point P_w with the sphere of radius $R = 2h$ (see Figure 3) then we can write equation (3) where all points are represented with respect to the sphere coordinate system.

$$\begin{cases} x_s = \lambda \cdot x_w \\ y_s = \lambda \cdot y_w \\ z_s = \lambda \cdot z_w \end{cases} \quad (3)$$

Since the points belong to the sphere, equation (4) also holds.

$$x_s^2 + y_s^2 + z_s^2 = R^2 \quad (4)$$

The perspective projection of P_s on the image plane from a point $C = [0, \xi]$ produces a point $P_i = [x, y]$ as expressed in equation (5)

$$\begin{cases} \frac{x_s}{\xi - z_s} = \frac{x}{\xi + \varphi} \\ \frac{y_s}{\xi - z_s} = \frac{y}{\xi + \varphi} \end{cases} \quad (5)$$

Adding the intrinsic camera parameters $\alpha_u, \alpha_v, u_0, v_0$, the pixel coordinates of the image points are shown in equation (6)

$$\begin{cases} u = \frac{\alpha_u(\xi + \varphi)x_w}{\xi\sqrt{x_w^2 + y_w^2 + z_w^2} - z_w} + u_0 \\ v = \frac{\alpha_v(\xi + \varphi)y_w}{\xi\sqrt{x_w^2 + y_w^2 + z_w^2} - z_w} + v_0 \end{cases} \quad (6)$$

The parameters of the model are ξ , which depends on the eccentricity, φ which is a function of both the eccentricity and the scale. As in the first model, $\alpha_u, \alpha_v, u_0, v_0$ are the intrinsic camera parameters and $r_X(\phi), r_Y(\theta), r_z(\varphi)$, and t_x, t_y, t_z are the extrinsic parameters used for changing the reference system from the world to the sphere. The calibration is done by means of a minimization algorithm for reducing the error on the image plane.

B. Omnidirectional laser projector calibration

The circular laser pattern projected into the conical mirror is reflected on the walls of the calibration box along an imaginary conical shape, called "light cone". Unlike the camera, the light projector does not provide "image points" therefore no correspondences can be established. The bright spots on the scene are observed by the calibrated omnidirectional camera which possesses an unique center of projection. This property allows calculating the direction of the light source for each image point. Besides, the size and location of the calibration box is known so the 3D absolute position of the bright spots can be determined by intersecting the known planes with the lines along which the light came into the mirrors' focus. A set of such spots can be used for calibrating the pair laser-mirror which is equivalent to estimating the light cone.

If the laser is perfectly aligned with the conical mirror, meaning that the rotation axis of both the conical mirror and the light cone overlap, then the light cone is circular with the same rotation axis as the mirror. However, such a perfect alignment is difficult to guarantee and a more general shape should be considered. Since the perspective projection of a circle placed on a plane π onto a plane that is not parallel with π is an ellipse it is straight forward that the shape we are looking for must be a revolution surface whose intersection with the plane perpendicular

on the omnidirectional camera optical axis is and ellipse. This shape is the elliptical cone that is described by the parametric equation (7):

$$\begin{cases} x_L = (a_L \cdot \frac{h+z_w}{h}) \cdot \cos(\alpha) \\ y_L = (b_L \cdot \frac{h+z_w}{h}) \cdot \sin(\alpha) \\ z_L = z_w \end{cases} \quad (7)$$

The parameters of the model are: a_L, b_L , the coefficients of the ellipse and h , the height of the cone. A point on the cone $P_L(x_L, y_L, z_L)$ is the calibrated model of the world point $P_w(x_w, y_w, z_w)$. The light cone was obtained by fitting the model to the previously calculated laser spots.

III. EXPERIMENTAL RESULTS

Due to their optical properties, catadioptric sensors having a SVP were chosen for building the omnidirectional camera of the proposed sensor. Commonly, the central projection constraint is fulfilled by the use of a parabolic mirror in conjunction with an orthographic camera or a hyperbolic mirror with a perspective camera. The first configuration that preserves the SVP independently of the translation between the mirror and the camera, was finally preferred.

The omnidirectional light projector is formed by a laser which emits a circular pattern and is pointed to a conical mirror so that the projected light covers the entire field of view of the catadioptric camera. The proposed projector can be seen as a reversed omni-camera where the light flows in the opposite sense. So, the projector benefits of the attributes revealed by previous studies of catadioptric cameras based on the conical mirror shape. Lin and Bajcsy [12] pointed out that the conical mirror can be used for building true SVP configurations with the advantage that it preserves image points brightness better than other mirrors since it does not distort the image in longitudinal directions. Yagi [21] highlighted the fact that the conical mirror on vertical section behaves like a planar mirror and consequently provides a much better resolution than any other omni-mirror shape. Baker and Nayar [1] proved that the curved mirrors (such as parabolic, hyperbolic, etc.) increase defocus blur because of their bend. Consequently, the cone bears out to be the ideal shape of mirror to be used for building the structured light projector.

A. Omnidirectional Camera Calibration

The mirror used for the omnidirectional camera was provided by Remote Reality [15]. The reflecting part is made, for compactness, of two mirrors: a paraboloidal one and a spherical one, as shown in Figure 4. Due to the existence of the spherical mirror, a third coordinate system centered in the focal point of the sphere had to be considered for building the first calibration model. With respect to the absolute position of the paraboloid reference frame the new coordinate system suffers a translation along the Z axis and a rotation of π radians around the Y axis which is reflected in a sign change of the u coordinate.

Since our attention was focused on calibration and reconstruction no efforts have been dedicated for implementing

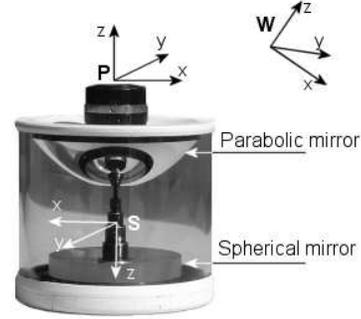


Fig. 4. The reflecting part of the catadioptric camera used for the sensor contains a paraboloidal and a spherical mirror. The three coordinate systems (c.s) are: **W**, the world c.s.; **P**, the parabolic mirror c.s. and **S**, the spherical mirror c.s.

a robust image segmentation method. Currently, a semi-automatic extraction of points is performed being based on a correlation algorithm which finds out the position of an interesting point searching through the approximative area indicated by the user. The results are given in pixel coordinates. Future improvements in the image processing will be aimed for the automatic detection of the calibration points with sub-pixel accuracy.

The calibrated parameters and the average calibration error for the two models are shown in Table I and Table II, respectively. The parameters obtained for the second model indicate that the mirror is not paraboloid since $\xi > 1$ and $\varphi > 0$ (see [6]). The calibration was performed using a set of 56 points equally distributed on the four walls of a cube placed around the sensor. The side of the cube has 50cm and the distance between a point and any of its neighbors on the same plane is 6.5cm. The calibration error was estimated using a new set of 28 test points located on the same cube and is the average of the distances between the predicted projections and the corresponding points in the real image.

The bright contour in the image corresponding to the shape produced by the laser pattern on the calibration planes was sampled at equal distances and 56 points were remapped on the scene using both models. The distances between the resulting 3D points for each model were measured on each calibration plane and are shown in Figure 5. The four bumps correspond to the planes of the calibration box and the larger discrepancy is produced in the central area of each plane.

B. Omnidirectional laser projector calibration

The light cone obtained by the projection of the circular laser pattern into the conical mirror was fitted to the back projections on the scene of the bright spots imaged by the camera. A set of 100 points identified in the omni-image were used for the calibration. The parameters for the calibrated model are presented in Table III and the resulting cone is shown in Figure 6. Note that there is only a small difference between the parameters a_L and b_L which means that the cone basis is near to a circle, in other words, the

TABLE I
THE CALIBRATED PARAMETERS FOR THE REFLECTION MODEL.

a	α_u	α_v	u_0	v_0	t_x	t_y	t_z	T_x	T_y	T_z	Average Error
11.09	20.40	20.49	399.14	281.39	-0.0008	3.14	-3.08	242.06	249.75	-120.89	3.58px

TABLE II
THE CALIBRATED PARAMETERS FOR THE SPHERICAL MODEL.

ξ	φ	α_u	α_v	u_0	v_0	t_x	t_y	t_z	T_x	T_y	T_z	Average Error
1.10	1.35	193.33	194.25	400.67	279.21	-0.06	0.003	-3.14	-228.05	261.86	-126.88	3.46px

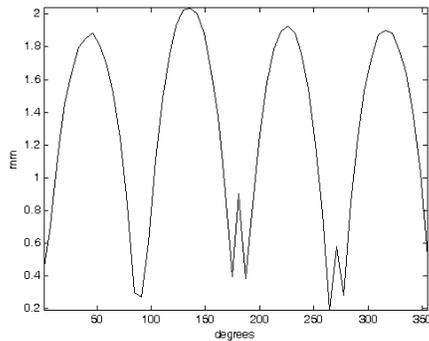


Fig. 5. The distances between a set of points extracted from the omnidirectional image and back-projected onto the scene using the two calibration models described in section II. The average distance is 1.089mm within a range of 1m.

TABLE III
THE LIGHT CONE PARAMETERS.

a_L	b_L	h	Average Error
571.27	568.89	-400.01	7.69%

lasers' optical axis was almost aligned with the conical mirror axis. The average distance error between the real points and the corresponding ones on the surface of the modelled cone represents 7.69% of the distance from the sensor to the 3D points.

C. Omnidirectional depth computation

Depth is calculated by triangulation. The laser projection on the scene is detected all over the image taken by the omnidirectional camera. Since the bright spots belong both to the light cone and to the rays crossing in the mirror focal point their 3D position can be calculated. The original image is presented in Figure 7.a. The third dimension was calculated within a field of view of 360 degrees and the 3D points formed the contour depicted in Figures 7.b.

IV. CONCLUSIONS

This work includes the detailed presentation of a novel stereo catadioptric sensor that provides 3D information for omnidirectional images. Mathematical formulation and

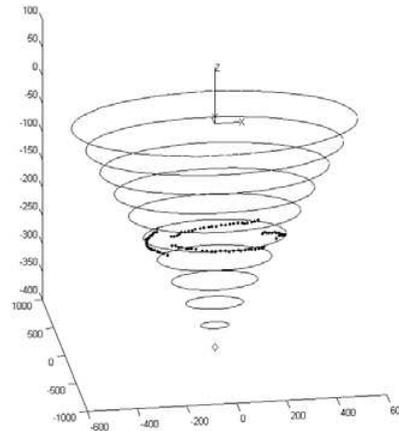


Fig. 6. Light cone obtained by calibration. The dots represent the points calculated using the calibrated omnidirectional camera.

specifications of the intrinsic and extrinsic parameters of the sensor as well as a calibration method were discussed.

The main advantage of such sensors is that depth is calculated simultaneously around a field of view of 360 degrees which makes them suitable for real-time applications. The sensor uses only one camera which reduces computation and manufacturing costs and provides a more robust system. The correspondence problem that is present in most stereo vision systems is solved much faster by introducing structured light projection since the matching is directly obtained. Moreover, the sensor is suitable for both topological and path following navigation models. At large distances the laser spots becomes difficult to identify so the localization has to be done using the information retrieved from the omnidirectional image. The structured light pattern can be precisely recognized for obstacles in the proximity of the robot and depth information becomes available.

ACKNOWLEDGMENT

Many thanks to professor Yasushi Yagi for the advices that inspired the work and the results presented in this article. Special thanks to José António Gaspar for the useful discussions we had about several subjects of this paper.

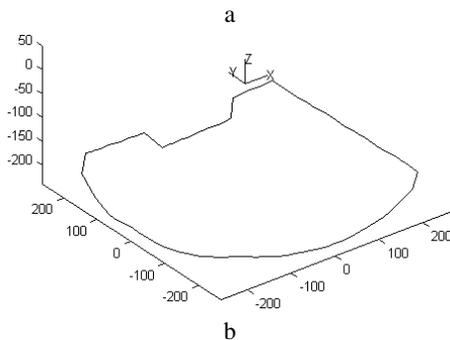
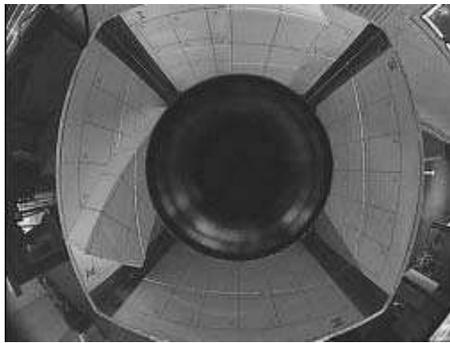


Fig. 7. Omnidirectional depth computation from a single image. a. Image of the scene containing the laser pattern. b. Lateral view of the 3D contour obtained from the lightened scene points.

REFERENCES

- [1] S. Baker and S.K. Nayar. A theory of catadioptric image formation. *Sixth International Conference on Computer Vision*, pages 35–42, 1998.
- [2] M. Fiala and A. Basu. Feature extraction and calibration for stereo reconstruction using non-svp optics in a panoramic stereo-vision sensor. In *Proceedings of the Third Workshop on Omnidirectional Vision*, pages 79–86, 2002.
- [3] Josep Forest. *New methods for triangulation-based shape acquisition using laser scanners*. PhD thesis, Universitat de Girona, 2004.
- [4] J. Gaspar, C. Decco, J. Okamoto, and J. Santos-Victor. Constant resolution omnidirectional cameras. In *In Proceedings on Third Workshop of Omnidirectional Vision*, pages 27–34, 2 June 2002.
- [5] Jose Antonio Gaspar. *Omnidirectional Vision for Mobile Robot Navigation*. PhD thesis, Technical Superior Institute, Technical University of Lisbon, 12 2002.
- [6] C. Geyer and K. Daniilidis. Equivalence of catadioptric projections and mappings of the sphere. In *OMNIVIS00*, pages 91–96, 2000.
- [7] C. Geyer and K. Daniilidis. A unifying theory for central panoramic systems and practical applications. *ECCV 2000 (Dublin, Ireland)*, pages 445–461, June 2000.
- [8] J. Gluckman and S.K. Nayar. Planar catadioptric stereo: Geometry and calibration. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1(1):I: 22–28, 23–25 June 1999.
- [9] J. Gluckman and S.K. Nayar. Rectified catadioptric stereo sensors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2):224–236, February 2002.
- [10] R.A. Hicks and R.K. Perline. Geometric distributions for catadioptric sensor design. In *CVPR01*, pages 1:584–589, 2001.
- [11] Du Q. Huynh. Calibration of a structured light system: a projective approach. pages 225 – 230, 17–19 June 1997.
- [12] S. S. Lin and R. Bajcsy. The true single view point (svp) configuration for omni-directional view catadioptric system using cone mirror. Technical report ms-cis-00-24, Computer and Information Science Department, University of Pennsylvania., Philadelphia, PA, USA, 11 2001.
- [13] S.A. Nene and S.K. Nayar. Stereo with mirrors. *Sixth International Conference on Computer Vision*, pages 1087–1094, 4–7 Jan. 1998.
- [14] Radu Orghidan, Joaquim Salvi, and El Mustapha Mouaddib. Calibration of a structured light-based stereo catadioptric sensor. *OMNIVIS 2003*.
- [15] Remote Reality. <http://www.remotereality.com/>.
- [16] J. Salvi, X. Armangué, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *PR*, 35(7):1617–1635, July 2002.
- [17] J. Salvi, J. Batlle, and E. Mouaddib. A robust-coded pattern projection for dynamic 3d scene measurement. *Pattern Recognition Letters*, (19):1055–1065, September 1998.
- [18] J. Salvi, J. Pagès, and J. Batlle. Pattern codification strategies in structured light systems. *Pattern Recognition*. In press.
- [19] Tomas Svoboda and Tomas Pajdla. Panoramic cameras for 3d computation. *Proceedings of the Czech Pattern Recognition Workshop*, pages 63–70, February 2000.
- [20] Nial Winters, Jose Antonio Gaspar, Gerard Lacey, and Jose Santos-Victor. Omnidirectional vision for robot navigation. pages 21–28, 12 June 2000.
- [21] Y. Yagi. Omnidirectional sensing and its applications. *IEICE Trans. Inf. and SYST.*, E82-D(3):568–578, 1999.