

A Computer Vision Sensor for Panoramic Depth Perception*

Radu Orghidan¹, El Mustapha Mouaddib², and Joaquim Salvi¹

¹ Institute of Informatics and Applications, Computer Vision and Robotics Group
University of Girona, Girona, Spain
{radu, qsalvi}@eia.udg.es

² Centre de Robotique, Électrotechnique et Automatique
Université de Picardie Jules Verne, Amiens, France
mouaddib@u-picardie.fr

Abstract. A practical way for obtaining depth in computer vision is the use of structured light systems. For panoramic depth reconstruction several images are needed which most likely implies the construction of a sensor with mobile elements. Moreover, misalignments can appear for non-static scenes. Omnidirectional cameras offer a much wider field of view than the perspective ones, capture a panoramic image at every moment and alleviate the problems due to occlusions. This paper is focused on the idea of combining omnidirectional vision and structured light with the aim to obtain panoramic depth information. The resulting sensor is formed by a single catadioptric camera and an omnidirectional light projector.

1 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 [14] surveyed the existing techniques for building cameras with a wide field of view and Svoboda [13] proposed several classifications of the existing omnidirectional cameras according to their most important features.

The catadioptric sensors use reflecting surfaces (convex or planar mirrors) coupled to a conventional camera and are usually classified depending on the way they gather the light rays. When all the observed light rays converge into a point, called focus, the sensors are known as Single View Point (SVP) [1]. The SVP enables distortion-free reconstruction of panoramic images in a familiar form for the human users.

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 [3] or planar [5] mirrors. The results obtained by stereoscopic vision depend on the accuracy of matching the points between the observed images. Structured light based techniques

* 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.

are a particular case of stereo vision where one of the cameras is replaced by a pattern projector [12]. Using this technique is similar to placing visible landmarks in the scene so that image points can be identified and matched faster.

This paper presents an omnidirectional sensor that provides 3D information using structured light. 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 reconstruction can be performed within a horizontal field of view of 360 degrees. The sensor design and the calibration of the whole system is detailed in section 2. The experimental results are shown in section 3. The article ends with conclusions, presented in section 4.

2 Sensor Geometry

In the proposed solution, see Figure 1, the omnidirectional camera is coupled with a structured light projector that has 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 [8] 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.

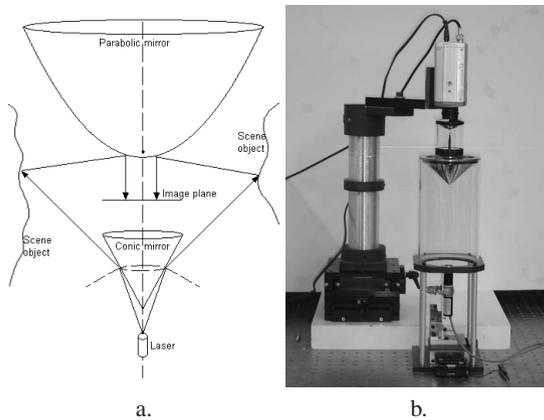


Fig. 1. a. Catadioptric omnidirectional camera with embedded structured light projector. b. Laboratory prototype.

The circular pattern projected by the laser is reflected by the conical mirror and becomes a light-stripe on the scene. The parabolic mirror reflects the scene into the camera and the laser-stripe can be immediately identified. With the models for both components of the sensor a precise triangulation can be carried out.

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. A method based on the cross ratio invariance under perspective projection providing a direct image to world transformation was proposed by Huynh [6]. Since our intention was to model the light projector and the camera independently the two steps calibration method was preferred.

2.1 Omnidirectional Camera Model

Assuming that the pair camera-mirror possesses a SVP, the omnidirectional camera can be modelled as the projection onto the sphere followed by the projection to a plane, as stated by Geyer and Daniilidis in [2]. Another way of approaching camera calibration is by considering the mirror surface as a known revolution shape and modelling it explicitly, for instance considering that the reflecting surface is a paraboloid and the camera is orthographic. Both models were tested and the comparative results were reported in [9]. The omni camera used for this work has a SVP but contains two reflecting surfaces so the first mentioned method was preferred.

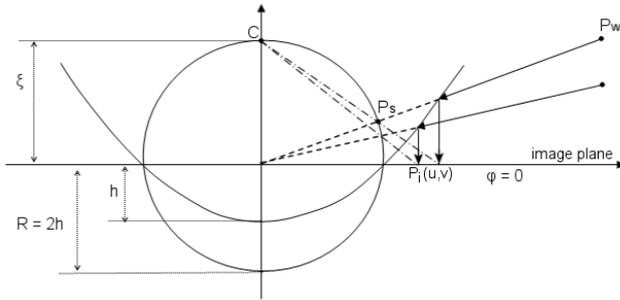


Fig. 2. Image formation using the projective equivalence of a SVP catadioptric projection with the projection on the sphere.

The calibration is performed using a set of known 3D points distributed on the four walls of a cube placed around the sensor. Consider a scene point $P_w = [x_w, y_w, z_w]$, and $P_s = [x_s, y_s, z_s]$ the intersection of the light ray emitted by the point P_w with the sphere of radius $R = 2h$ (see Figure 2). We can write equation (1) where all points are represented with respect to the camera coordinate system.

$$\begin{cases} x_s = \lambda \cdot x_w \\ y_s = \lambda \cdot y_w \\ z_s = \lambda \cdot z_w \end{cases} \tag{1}$$

Since the points belong to the sphere: $x_s^2 + y_s^2 + z_s^2 = R^2$.

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 (2)

$$\begin{cases} \frac{x_s}{\xi - z_s} = \frac{x}{\xi + \varphi} \\ \frac{y_s}{\xi - z_s} = \frac{y}{\xi + \varphi} \end{cases} \tag{2}$$

Adding the intrinsic camera parameters $\alpha_u, \alpha_v, u_0, v_0$, the pixel coordinates of the image points are shown in eq. (3)

$$\begin{cases} u = \frac{\alpha_u(\xi+\varphi)x_w}{\xi\sqrt{x_w^2+y_w^2+z_w^2}-z_w} + u_0 \\ v = \frac{\alpha_v(\xi+\varphi)y_w}{\xi\sqrt{x_w^2+y_w^2+z_w^2}-z_w} + v_0 \end{cases} \quad (3)$$

The parameters of the model are ξ , which depends on the eccentricity; φ which is a function of both the eccentricity and the scale; $\alpha_u, \alpha_v, u_0, v_0$, 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 orientation and the translation between the world coordinate system placed in the upper corner of the first calibration plane and the camera 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 [11].

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 model of the omnidirectional camera is calibrated.

2.2 Omnidirectional Laser Projector Model

The omnidirectional light projector is formed by a laser which emits a circle 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 [7] 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 [14] 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.

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. Since the locations of the calibration planes are known, the 3D coordinates of the laser-stripe lying on those planes can be determined. A set of such points can be used for calibrating the pair laser-mirror.

Ideally, when the laser is perfectly aligned with the conical mirror, the 3D shape formed by the reflected laser pattern can be imagined as a circular cone, called “laser-cone”. Unfortunately, the precision of obtaining the coordinates of the bright spots is bounded by the catadioptric camera calibration accuracy and by its resolution. Moreover, a perfect alignment of the laser and the conical mirror is difficult to guarantee so a

more general shape than the circular cone 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 can be deduced that a suitable shape for modelling the laser-cone is a revolution surface whose intersection with the plane perpendicular on the omnidirectional camera optical axis is and ellipse. This shape, the elliptic cone, was used in [9] and proves to be more accurate than the circular cone. Still, for a large amount of noise, the elliptical cone can not be uniquely determined.

Therefore, the general quadratic surface was chosen for modelling the laser projection. Consider $P_{wi}(x, y, z)$ the bright spots on the calibration walls with known coordinates. The quadratic surface that passes through all the points is represented in eq. 4. Let H be the matrix that contains the coordinates of the points, A the matrix of the parameters and F the free term matrix. Writing $H \cdot A = F$, the matrix A can be obtained by $A = (H' \cdot H)^{-1} \cdot H' \cdot F$. This is a simple method for calibrating the omni projector. Since no iterations are needed it is much faster than the iterative minimization methods. However, its main drawback is that the matrix H can not be controlled and, for noisy data, it is likely to be singular.

$$\begin{bmatrix} x_1^2 & y_1^2 & z_1^2 & 2x_1y_1 & 2x_1z_1 & 2y_1z_1 & 2x_1 & 2y_1 & 2z_1 \\ \vdots & \vdots \\ x_i^2 & y_i^2 & z_i^2 & 2x_iy_i & 2x_iz_i & 2y_iz_i & 2x_i & 2y_i & 2z_i \\ \vdots & \vdots \\ x_n^2 & y_n^2 & z_n^2 & 2x_ny_n & 2x_nz_n & 2y_nz_n & 2x_n & 2y_n & 2z_n \end{bmatrix} \cdot \begin{bmatrix} a_{11} \\ a_{22} \\ a_{33} \\ a_{12} \\ a_{13} \\ a_{23} \\ \beta_{31} \\ \beta_{32} \\ \beta_{33} \end{bmatrix} = \begin{bmatrix} \vdots \\ -1 \\ -1 \\ -1 \\ \vdots \end{bmatrix} \quad (4)$$

Therefore, a more robust method for finding the parameters of the general quadratic surface must be considered. Lets assume, without loss of generality, that the world reference system is placed such that the calibration planes are perpendicular on the X and Y axis. The intersections of the quadratic with the calibration planes are arcs described by a subinterval of the parameter domain: the arcs contained in the planes perpendicular on the X and Y axis provide information on the parameters of the quadratic with $x = ct$ and $y = ct$, respectively. Writing the quadratic as in eq. 5, its intersection with the planes X and Y are shown in eq. 6 and eq. 7, respectively. The parameters of the arcs for each plane are obtained by fitting the corresponding points into the subsequent equations. Taking into account that the 3×3 matrix is symmetric, the full set of parameters of the quadratic surface can be retrieved from equations 6 and 7.

$$[x \ y \ z] \cdot \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ y \\ z \end{bmatrix} + [x \ y \ z] \cdot \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + f = 0 \quad (5)$$

$$[y \ z] \cdot \begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} y \\ z \end{bmatrix} + [y \ z] \cdot \begin{bmatrix} P_x \\ Q_x \end{bmatrix} + R_x = 0 \quad (6)$$

$$[x \ z] \cdot \begin{bmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{bmatrix} \cdot \begin{bmatrix} x \\ z \end{bmatrix} + [x \ z] \cdot \begin{bmatrix} P_y \\ Q_y \end{bmatrix} + R_y = 0 \quad (7)$$

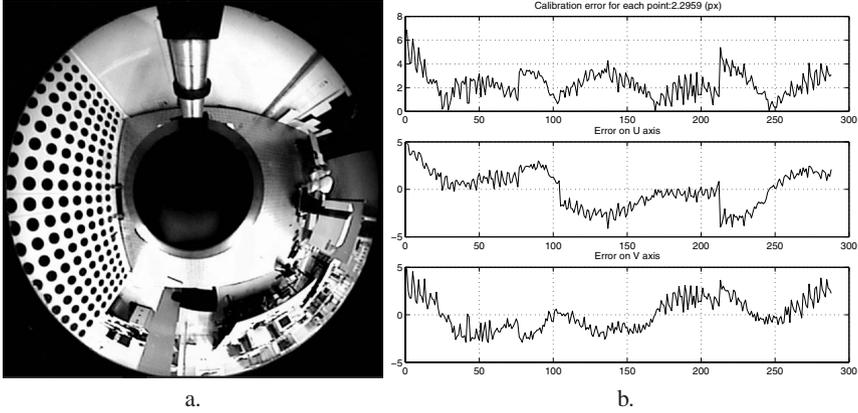


Fig. 3. a. One of the calibration planes. b. Error for the 285 calibration points, measured in pixels.

Table 1. The calibrated parameters for the omni camera.

ξ	φ	α_u	α_v	u_0	v_0	r_x	r_y	r_z	t_x	t_y	t_z
1.06	-9.64	-32.53	33.24	429.51	292.72	0.02	0.01	-0.009	-26.45	-0.82	-754.1

Dividing the calibration in two parts the number parameters to be simultaneously minimized decreases which leads to a robust calibration method.

3 Experimental Results

The system was build using off the shelf components. The optics and the mirror used for the omnidirectional camera were provided by Remote Reality [10]. The camera is a Sony SSC-DC198P with the ccd of 1/3". The laser and its optics are produced by Lasiris, the diode power is 3mW and produces red light with a wavelength of 635nm.

The camera calibration is performed using a set of 285 dots distributed on the four planes placed around the sensor. The distance between the centers of any two adjacent dots on the same plane is 6cm and the height of the calibration plane is 80cm. A semi-automatic point extraction method is performed. For each plane, several dots are selected by the user and their centers are determined with sub-pixel accuracy. The centers of the remaining dots are automatically found with the same precision. The calibrated parameters of the camera-model are listed in the Table 1. The average calibration error is $\mu = 2.3\text{px}$ and the sample standard deviation $\sigma = 2.542$.

The conical mirror used for building the laboratory prototype has a height $h = 4.4$ cm and the cone aperture angle is $\beta = 52$ degrees. The laser projects a circular cone with a fan angle $\alpha = 11.4$ degrees. Given that the relation between the two angles is $\beta \approx 0.5(\alpha + \pi/2)$, the laser is reflected along a very flat surface which can be approximated to a plane: $ax + by + cz + d = 0$, see Figure 4.b. The center of the laser stripe is determined with sub-pixel accuracy using the peak detection method described by Forest [4] and the discrete points are used for calibrating the parameters of the plane: $a = -0.13$, $b = -0.001$, $c = 1$ and $d = 78.99$.

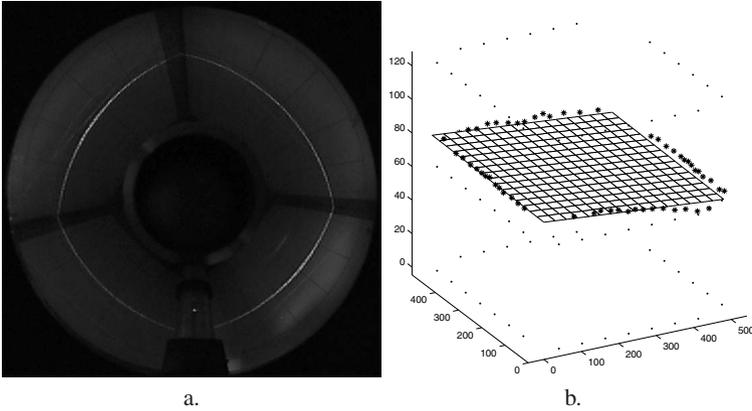


Fig. 4. a. Projection of the laser pattern. b. Flat surface fitted to a set of discrete points from the laser stripe. The three dotted rectangles are the points used for calibrating the camera.

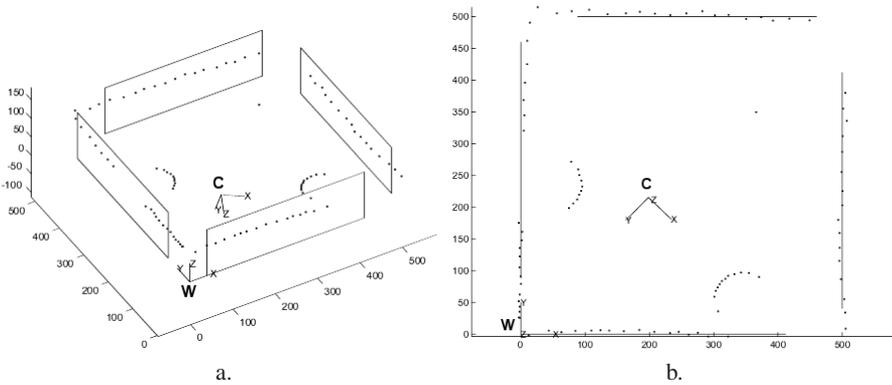


Fig. 5. Omnidirectional 3D profile obtained along the laser stripe. The dots stand for the reconstructed 3D points. a. Lateral view b. Upper view.

With the sensor surrounded by four planes depth was calculated using a set of discrete points of the laser pattern. For a scene containing two cylinders the result is presented in Figures 5 with the two cylindrical shapes correctly identified. It is also noticeable that the points on the walls fall on the corresponding planes. In terms of accuracy, the radius of the cylinder was measured and has 93cm while the range finder returned a result of 95cm.

4 Conclusions

It is noticeable that the use of 360 degrees images and of scene-depth information is ideal for robot navigation tasks. Starting from this observation we combine the advantages of omnidirectional vision and structured light. We presented here the geometry

and the calibration for a prototype of a panoramic range finder. The two omnidirectional systems that compose the sensor are calibrated and the resulting model is used for measuring depth in a real scene. The accuracy of the sensor is enhanced by the use of sub-pixel accuracy techniques at calibration and reconstruction stages. The results obtained are encouraging and prove that this sensor can be used in real robot navigation and depth perception applications.

References

1. S. Baker and S.K. Nayar. A theory of catadioptric image formation. *IEEE Int. Conf. on Computer Vision, ICCV*, pages 35–42, 1998.
2. K. Daniilidis C. Geyer. A unifying theory for central panoramic systems and practical applications. *Sixth European Conference on Computer Vision*, pages 445–461, June 2000.
3. M. Fiala and A. Basu. Feature extraction and calibration for stereo reconstruction using non-svp optics in a panoramic stereo-vision sensor. In *Workshop on Omnidirectional Vision*, pages 79–86, 2002.
4. J. Forest, J. Salvi, E. Cabruja, and C. Pous. Laser stripe peak detector for 3d scanners. a fir filter approach. In *International Conference on Pattern Recognition*, volume 3, pages 646–649, Cambridge, United Kingdom, August 2004.
5. J. Gluckman and S.K. Nayar. Planar catadioptric stereo: Geometry and calibration. *IEEE Computer Vision and Pattern Recognition*, 1(1):I: 22–28, 23-25 June 1999.
6. Du Q. Huynh. Calibration of a structured light system: a projective approach. In *IEEE Computer Vision and Pattern Recognition*, pages 225–230, 17-19 June 1997.
7. 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.
8. Radu Orghidan, Joaquim Salvi, and El Mustapha Mouaddib. Calibration of a structured light-based stereo catadioptric sensor. *Workshop on Omnidirectional Vision*, IEEE Conf. on Computer Vision and Pattern Recognition - Volume 7, 2003.
9. Radu Orghidan, Joaquim Salvi, and El Mustapha Mouaddib. Omnidirectional depth computation from a single image. *IEEE International Conference on Robotics and Automation, ICRA*, April 18-22, 2005.
10. Remote Reality. <http://www.remotereality.com/>.
11. J. Salvi, X. Armangue, and J. Batlle. A comparative review of camera calibrating methods with accuracy evaluation. *Pattern Recognition*, 35(7):1617–1635, July 2002.
12. 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.
13. Tomas Pajdla Tomas Svoboda. Panoramic cameras for 3d computation. *Proceedings of the Czech Pattern Recognition Workshop*, pages 63–70, February 2000.
14. Y. Yagi. Omnidirectional sensing and its applications. *IEICE Trans on Information and Systems*, E82-D(3):568–578, 1999.