

CALIBRATION OF A STRUCTURED LIGHT-BASED STEREO CATADIOPTRIC SENSOR *

Radu Orghidan[†], Joaquim Salvi[†], El Mustapha Mouaddib[‡]

[†] Computer Vision and Robotics Group, Institute of Informatics and Applications, Universitat de Girona, Avda. Lluís Santalo, s/n. 17071 Girona, Catalonia (Spain)

[‡] Centre de Robotique, Electrotechnique et d'Automatique, Université de Picardie Jules Verne, Rue du Moulin Neuf 80000 Amiens, France

ABSTRACT

Catadioptric sensors are combinations of mirrors and lenses made in order to obtain a wide field of view. In this paper we propose a new sensor that has omnidirectional viewing ability and it also provides depth information about the nearby surrounding. The sensor is based on a conventional camera coupled with a laser emitter and two hyperbolic mirrors. Mathematical formulation and precise specifications of the intrinsic and extrinsic parameters of the sensor are discussed. Our approach overcomes limitations of the existing omnidirectional sensors and eventually leads to reduced costs of production.

Keywords: omnidirectional view, hyperbolic mirror, laser, structured light, stereo vision

1. INTRODUCTION

A large field of view is an attractive goal in computer vision. Potential applications are in the domains of surveillance, teleconferencing, unknown environment exploration, tourism, advertising etc. Moreover, egomotion estimation and object tracking can largely benefit from enhancing the field of view of imaging systems. The ideal omnidirectional device is a system having the capability to see simultaneously 360° in all directions. In this way a panoramic image of the environment is continuously processed. There are many proposals for building omnidirectional sensors: a) special lenses such as fish-eye lenses; b) multiple image acquisition by means of rotating cameras or using structures of many cameras with complementary fields of view; and c) cameras with especial mirrors (planar, hyperbolic, parabolic, spherical, dual convex and conic) [3].

A common constraint upon the omnidirectional sensors modelling requires that all the imaged rays pass through a unique point called single view-point (SVP). This is an important feature because it permits the distortion-free

reconstruction of the image. Consequently, the omnidirectional sensors can be classified depending on whether they have a SVP or not. Nayar and Baker [4] explored the class of SVP catadioptric sensors that produce SVP.

Omnidirectional view can be achieved by means of fish-eye lenses with the disadvantage that such lenses introduce radial distortion that complicates the mathematical model. The obtained images have good resolution on the central region but have poor resolution on the marginal region. Furthermore, fish-eye lenses do not meet the SVP constraint. High resolution panoramic views can be obtained using multiple images provided by a rotating camera. The main drawback of this method is that the panoramic image is not formed with a single capture. Moreover, omnidirectional view is formed from several partial images, which can lead to discontinuities in the resulting image if the camera is not precisely controlled involving lower robustness and constraining the use of these sensors in robot applications. Cameras with complementary fields of view overcome some problems of the previous method because there are not moving parts but the sensor becomes more expensive and difficult to calibrate since it involves several cameras. The combinations of mirrors (catoptrics) and lenses (dioptrics) is called catadioptrics. A 360° horizontal field of view is achieved at every capture using a single camera looking into a convex mirror. For instance, Pegard and Mouaddib [2] built such a sensor with a conic mirror while Yamazawa et al. [4] used a hyperbolic mirror. Both obtained good results for robot navigation and localization.

Stereo catadioptric sensors are special structures of catadioptrics used to obtain depth from two omnidirectional images. Nene and Nayar [10] studied several stereo configurations using the mirrors which have a SVP. Also, Gluckman et al. [9] designed a catadioptric stereo sensor based on two parabolic mirrors with two cameras. Southwell et al. [7] studied the idea of triangulation by means of a catadioptric sensor with a bi-lobbed mirror. Later, Ollis [6] simulated various configurations using two hyperbolic mirrors with one and

* Work funded by the Spanish project CICYT TAP99-0443-C05-01

two cameras. Depth measurements were obtained by triangulation. The correspondence problem was solved by window matching between both images by means of a correlation algorithm considering the curvature of the mirrors. Fiala and Basu [5] built such a sensor using embedded spherical mirror lobes. Depth measurements were also obtained by means of feature extraction and triangulation.

Besides, structured light is a technique widely used to measure 3D surfaces by the projection of a sequence of light patterns [11]. The measuring surface is emphasised by the apparent deformations of the imaged pattern from the projected one. Such light patterns are projected by means of slide projectors, digital projectors or laser light emitters which forms different shape of patterns. Slides are difficult to use with calibrated systems because of the deformations caused by the functioning temperature of the projector lamp. Besides, digital projectors have focusing problems which limits the depth range of the measuring area. Finally, laser light offers attractive features such as monochromatic light which permits the use of optical filters for a straight-forward image segmentation. Laser projectors are thin and can be placed inside compact devices. The experiments conducted by Baldwin and Basu [8] demonstrate that bright lights improves matching, therefore, laser light seems a good choice. However, laser emitters can only project simple types of patterns such as points, lines and grids.

This paper presents a new catadioptric panoramic sensor with an embedded structured light projector. In section 2, several possible configurations are compared and the mathematical model of the most desirable one is analyzed. Section 3 deals with the design of the optimal size of the mirrors. The calibration method is described in section 4. Then, section 5 presents the simulated results of the proposed sensor. The article ends with conclusions.

2. SENSOR DESIGN

Catadioptric devices can be classified related to the SVP constraint. For instance, spherical shaped mirrors do not meet such constraint and the lack of a perspective point makes more complicated the image reconstruction process. Nevertheless, spherical mirrors are preferred for manufacturability reasons and for specific properties such as the compact size or the equal characteristics over azimuth [5][7][8]. The second category contains the mirrors that meet the SVP constraint: the parabola, the ellipse and the hyperbola. The parabolic mirror works in the same way as the parabolic antenna. The incoming rays pass through the focal point and are then reflected parallel to the rotating axis of the parabola. Therefore, in order to use the SVP property, the image must be formed by a camera with orthographic lenses. This is an expensive type of camera and increases the price of the catadioptric device. Besides, the ellipsoid mirror has a small practical use since the reflecting part and the

camera must be on the inner side of the surface providing an ineffective field of view. Finally, the hyperbolic mirror can be observed using a common perspective camera, therefore is considered an attractive choice.

In addition, stereo omnidirectional using structured light can be reached by means of combinations of two hyperbolic mirrors, a camera and a laser projector (figure 1). The mirrors are symmetrical about the z axis so the bi-dimensional graphical representation is allowed. The first configuration (fig. 1.a) is based on two identical mirrors located along a baseline which can be fixed depending on the accuracy desired. The light pattern is projected onto

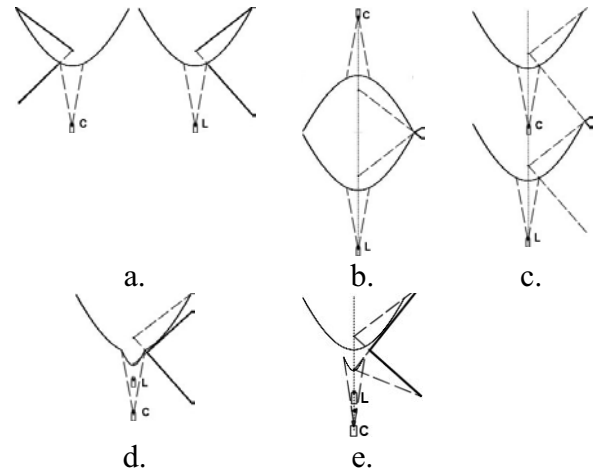


Figure 1. Stereo omnidirectional systems (the dotted lines bound the camera and the laser self-occlusion regions).

one of the mirrors while the other mirror grabs the scene. The mirrors are called “laser mirror” and “camera mirror”, respectively. Such a setting, makes invaluable a part of the image because of the scene occlusion caused by the laser projector. This problem can be solved by placing both mirrors back to back along the rotation axis (see fig 1.b). Such a system removes the occlusion problem but the laser has to be projected on the peripheral size of the mirror (in order to be seen by the camera mirror) where a major lens distortion is present leading to a poorer resolution of the system. Placing the two mirrors by a single translation along the vertical axis (see fig 1.c), permits the laser to be projected nearby the centre of the mirror where a better resolution is given but the death area of the camera mirror is wasted leading to a larger system. A natural solution is to put the laser mirror inside the death cone given by the camera mirror (Fig 1.d) obtaining a compact sensor, yet, with the inconvenient that the laser mirror occludes part of the vision field of the camera mirror. In order to avoid this problem and to enhance the baseline, the two mirrors can be separated (Fig 1.e). Taking into account the previous observations, a new catadioptric sensor is proposed (see fig.2). The sensor is composed of two hyperbolic mirrors, a laser projector and a camera. The camera and the laser emitter are both represented using the pinhole model since the

projector can be seen as an inverted camera. The laser emits a circle onto the laser mirror which is projected forward 360° all around the measuring scene and backwards to the camera mirror into the image plane. Depth information is retrieved by analyzing the 2D projection on the camera of the points of the scene enlightened by the laser pattern. These points are isolated by an optical filter and further image segmentation techniques. The camera mirror complies with the SVP criteria, therefore the line on which lays an object point P_w can be determined from its projective image point P_i . The laser mirror benefits of the same properties, that is: the line along which a point of the pattern is projected can also be defined. Depth information is obtained by crossing both lines. Hereafter, a complete description of the sensor is detailed.

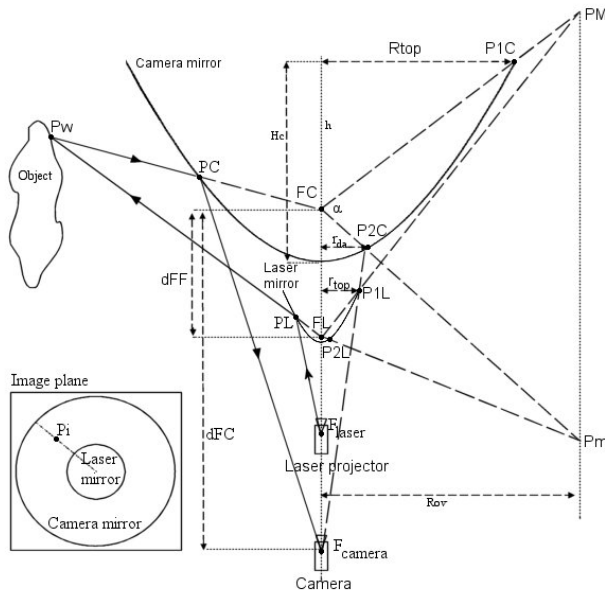


Figure 2. Stereo catadioptric sensor with structured light projector.

The angle of view (α) of the sensor is the angle between $\overline{FC, P1C}$ and $\overline{FC, P2C}$, where $FC(x_0, y_0, z_0)$ is the focal point of the camera mirror, $P1C$ is the highest reflecting point of the mirror and $P2C$ is the projection of point P_m on the camera mirror. The angle α and the range of view (Rov) determine the position of the highest (PM) and lowest (Pm) points of scene imaged by the sensor. These points are reflected on the laser mirror on points $P1L$ and $P2L$, respectively. The focal points of the laser mirror (FL) and the camera mirror (FC) are aligned along the rotation axis and translated a distance dFF .

The hyperbolic shapes of the camera and the laser mirrors are modeled by equations (1) and (2), respectively.

$$\frac{(z + \sqrt{a^2 + b^2})^2}{a^2} - \frac{x^2 + y^2}{b^2} = 1 \quad (1)$$

$$\frac{((z - dFF) + \sqrt{a_L^2 + b_L^2})^2}{a_L^2} - \frac{x^2 + y^2}{b_L^2} = 1 \quad (2)$$

Solving equation (1) for z two solutions are found, which are the equations of the upper (3) and lower (4) hyperbolas.

$$z_1 = \frac{-e \cdot b + \sqrt{a^2 b^2 + a^2 y^2 + a^2 x^2}}{b} \quad (3)$$

$$z_2 = \frac{-e \cdot b - \sqrt{a^2 b^2 + a^2 y^2 + a^2 x^2}}{b} \quad (4)$$

where, $e = \sqrt{a^2 + b^2}$.

The parameters of each mirror are supposed to be linked by $k = a/b$, where k is a constant value. The focal point of the camera (F_{camera}) is placed in the focus of the lower hyperbola, at distance $dFC = 2e$ from FC .

The light beam emitted by an infinitely small point of the pattern hits the laser mirror in a point (PL) that can be calculated by solving the system formed by equations (2) and the equation of the line defined by the light beam. According to the SVP constraint, PL is projected along the line $\overline{FL, PL}$ at an object point P_w which is projected onto the camera mirror at $PC(x_{PC}, y_{PC}, z_{PC})$, obtaining equation (5) and (6). $P_w(x_w, y_w, z_w)$ is determined by triangulation since it belongs to both lines $\overline{FL, PL}$ and $\overline{FC, PC}$.

$$\frac{(z_{PC} + \sqrt{a^2 + b^2})^2}{a^2} - \frac{x_{PC}^2 + y_{PC}^2}{b^2} = 1 \quad (5)$$

$$\begin{cases} x_{PC} - x_0 = \frac{y_{PC} - y_0}{y_{PC} - y_w} = \frac{z_{PC} - z_0}{z_{PC} - z_w} \\ x_{PC} - x_w = \frac{y_{PC} - y_w}{y_{PC} - y_w} = \frac{z_{PC} - z_w}{z_{PC} - z_w} \end{cases} \quad (6)$$

There are two solutions of the system but only the point closer to the emitting light source must be considered. The coordinates of PC as a function of P_w are given by equation (7).

$$\begin{cases} x_{PC_{1,2}} = \frac{bx_w(e \cdot z_w + a \cdot \|P_w\|)}{a^2 \cdot \|P\|^2 - z_w^2 \cdot e^2} \\ y_{PC_{1,2}} = \frac{y_w \cdot x_{1,2}}{x_w} \\ z_{PC_{1,2}} = \frac{z_w \cdot x_{1,2}}{x_w} \end{cases} \quad (7)$$

where, $\|P_w\| = \sqrt{x_w^2 + y_w^2 + z_w^2}$.

Besides, the points FL , PL and P_w are on the same line. Since the coordinates of the points FC , PC , FL and PL are known, the point P_w is uniquely determined.

3. DESIGN OF THE OPTIMAL MIRROR SIZE

An important concern when designing the sensor is its size. The sensor is designed to be used in indoor

environments so it has to be small. Besides, image processing requires a good resolution which is improved by using large mirrors. A compromise must be set up in order to obtain optimal performances.

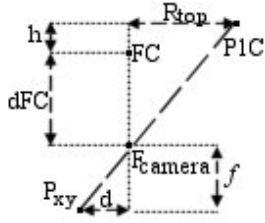


Figure 3. The relation between R_{top} , the size of the sensor and the intrinsic parameters of the camera

In figure 3, P_{xy} is the image point of the camera mirror rim on the CCD, d is the distance from the centre of the CCD to P_{xy} , f is the focal distance of the camera and h is the vertical distance from the focal point of the mirror to its border. Ideally, the camera mirror is imaged by the camera as a disc tangent to the CCD border, therefore, several constraints have to be taken into account during the design process of the sensor:

Constraint 1. The point PIC must be on the line $\overline{FC, PM}$ (fig 1.) so that the camera mirror is able to see the highest points of the required field of view.

Constraint 2. The points of the camera mirror rim must be close to the line $\overline{P_{xy}, PIC}$ for an optimally sized image in the camera.

Constraint 3. From several suitable hyperbolic shapes for the camera mirror the one with the bigger “death area” is chosen since it will provide a larger cone for placing the structured light projector.

A numeric algorithm produces all hyperbolic shapes within a given range for specific values of α , R_{top} and R_{ov} . Every shape is compared with the ideal one and the shapes with an error below a threshold are kept. According to third constraint, the desired one is selected. The relation between r_{da} and FF is shown in figure 3. Values for several shapes have been calculated and are presented in Table 1. The most suitable one depends on the final application of the system.

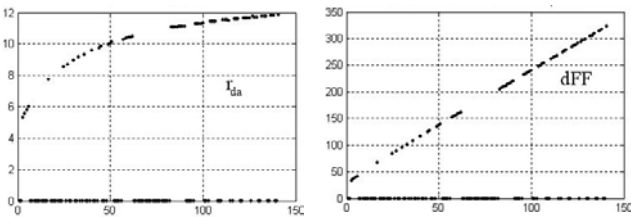


Figure 4. Values of r_{da} and dFF for mirror shapes with the a parameter varying from 0 up to 150.

Table 1. The parameters for suitable sensors

α	a	b	R_{top}	H	r_{top}	dFC
70°	143	45.6	30	28.16	7.46	300.24
	132	50	40	37.24	9.61	282.30
	141	57.3	50	46.29	11.78	304.40
90°	150	41.2	30	35.57	4.61	311.11
	148	46.5	40	47.14	5.94	310.29
	143	50.3	50	58.68	7.1652	303.21
100°	146	37.7	30	40.56	3.4983	301.59
	144	42.4	40	53.81	4.4779	300.26
	146	47.0	50	67.06	5.4105	306.81

4. CALIBRATION

In the following, the calibration of the sensor is done in two steps considering first the calibration of the pair camera and camera mirror (see section 3.1) which is then used for the calibration of the pair laser and laser mirror (see section 3.2).

3.1. The pair camera and camera mirror. The pair formed by the camera and its coupled mirror is modelled by equation (8).

$$P_i = {}^{2D}K_C \cdot {}^C K_M \cdot F({}^M K_W \cdot {}^W P) \quad (8)$$

where:

${}^W P$ is a known object point respect to the world coordinate system,

P_i is the image point corresponding to the projection of ${}^W P$,

${}^M K_W$ is the transformation matrix that relates the world coordinate system with respect to the mirror coordinate system,

${}^C K_M$ is the transformation matrix that relates the mirror coordinate system with respect to the camera coordinate system,

${}^{2D}K_C$ is the projection matrix that models the internal geometry of the camera by a set of intrinsic parameters:

$${}^{2D}K_C = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (9)$$

where: $\alpha_u = -f \cdot k_u$, $\alpha_v = -f \cdot k_v$

So, the projection of an object point ${}^W P$ onto the mirror (obtaining the point PC) is defined by a non-linear function detailed in equation (10).

$$F(P_w) = \begin{bmatrix} 1 & y_w & z_w \\ x_w & x_w & x_w \end{bmatrix}^T \cdot \frac{bx_w(e \cdot z_w + a \cdot \|P_w\|)}{a^2 \cdot \|P\|^2 - z_w^2 \cdot e^2} \quad (10)$$

Besides, the world coordinate system might coincide with the camera coordinate system: ${}^M K_W = {}^M K_C$ and ${}^C K_M = {}^W K_M = \text{inv}({}^M K_W)$.

Therefore, equation (8) can be written as:

$$P_i = {}^{2D}K_C \cdot {}^C K_M \cdot F(\text{inv}({}^C K_M) \cdot {}^W P) \quad (11)$$

and, expanding the matrices we obtain equation (12):

$$\begin{bmatrix} s \cdot x_i \\ s \cdot y_i \\ s \end{bmatrix} = \begin{bmatrix} \alpha_u & 0 & u_0 & 0 \\ 0 & \alpha_v & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} F_x \\ F_y \\ F_z \\ 1 \end{bmatrix} \quad (12)$$

Summarizing equation (12), it is obvious that there are seven unknowns: a) the four intrinsic parameters of the camera; b) the translation vector corresponding to the transformation matrix between the mirror and the camera coordinate system; and c) the a and b parameters that define the hyperbolic shape of the camera mirror.

The calibration method obtains the values of these seven unknowns by means of a non-linear minimization of equation (12) based on Levenberg-Marquard and a set of known object points (calibrating pattern).

3.2. The pair laser and laser mirror. The laser source emits an optical circle forming a cone which is determined by the laser pose and the aperture angle of the lens. That cone is then projected onto the laser mirror towards the 3D scene spreading the light in a set of optical rays. Every optical ray intersects with known calibrating planes obtaining a set of 3D points ${}^W P$ which are imaged by the camera system. Besides the aperture angle γ is known and given by manufacturers. Then, the relationship between the laser parameters and the set of 3D points is determined by equation (13).

$$\tan(\gamma) = \frac{\sqrt{b_L^4 (e \cdot z_W + a_L \sqrt{x_W^2 + y_W^2 + z_W^2})^2 \cdot (x_W^2 + y_W^2)}}{((b_L z_W)^2 - (x_W a_L)^2 - (a_L y_W)^2)^2} \quad (13)$$

$$|dFL| - |dFF| = \frac{z_W b_L^2 (e \cdot z_W + a_L \sqrt{x_W^2 + y_W^2 + z_W^2})}{((b_L z_W)^2 - (x_W a_L)^2 - (a_L y_W)^2)^2}$$

The coordinates of the points ${}^W P$ are obtained from their 2D projection on the image plane and the parameters of the camera system given by the calibrating algorithm explained in section 3.1. Equation (13) is composed by only four unknowns parameters: a) the mirror shape parameters a_L , b_L ; b) dFF , the translation between the world coordinate system and the focal point of the laser mirror and c) dFL , the distance between the laser source point and the focal point of the and the corresponding mirror. The solution of such non-linear system is also obtained by a minimization algorithm based on Levenberg-Marquard.

5. EXPERIMENTAL RESULTS

The whole model presented in section 3 has been simulated in Matlab with the aim of analysing the accuracy of the calibrating algorithm in the presence of Gaussian noise. The complete virtual sensor used is shown in figure 3. The reflecting shapes of the two mirrors can be observed. The dotted lines represent the

incoming and outgoing laser rays. The sensor is placed in a virtual room whose walls plane equations are known. The room is 2x3 meters and 2.5 meters height. At calibration time, the sensor was placed at 1 meter height from the floor.

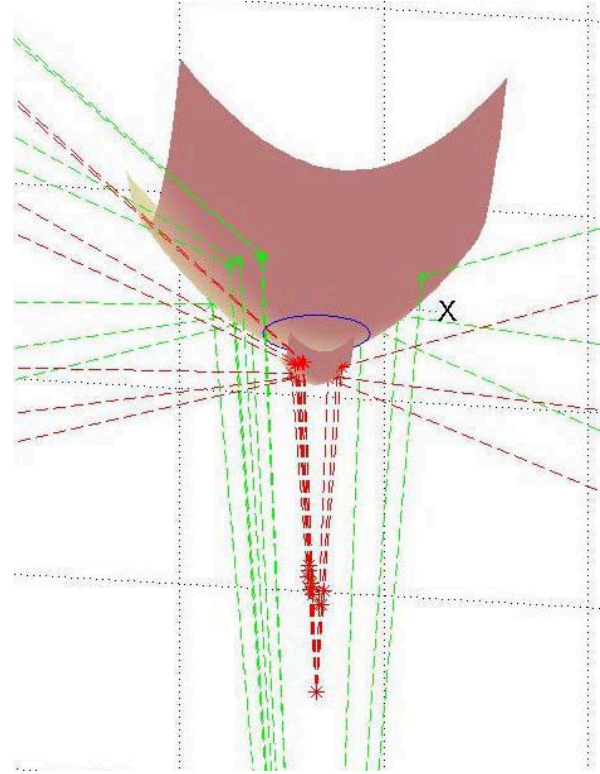


Figure 4. Virtual sensor used for calibration

A set of 200 iterations has been provided to the calibrating algorithm considering a Gaussian noise from $\sigma = 0$ up to $\sigma = 2$ in increments of 0.1.

In Figure 5 is presented the discrepancy between the 3D object points (expressed in millimetres) given by the proper model (with zero noise) with respect to the model obtained by calibration in the presence of noise. Besides, Figure 6 shows the same discrepancy but with respect to the image points, represented in pixels.

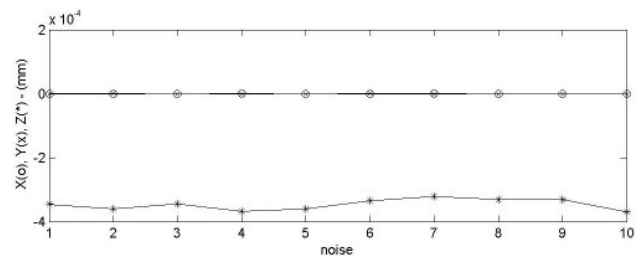


Figure 5. Distance between the pattern 3D points and the reconstructed 3D points by means of the model obtained by calibration in the presence of noise.

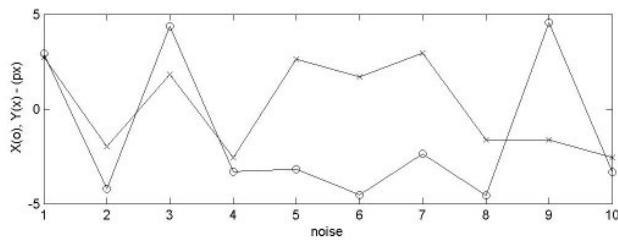


Figure 6. Distance between the noiseless image points and the reconstructed 2D points using the model obtained by calibration in the presence of noise.

6. CONCLUSIONS

Passive stereo catadioptric sensors are structures built in order to provide 3D information using panoramic images. Depth is calculated by triangulation observing the same features from different points of view. The main advantage of such sensors is that depth is calculated simultaneously around a field of view of 360° which makes them suitable for real-time applications. However, stereo techniques have a major limitation related to the high computational cost of the image segmentation stage and the further matching solving process. Introducing structured light projection, the observed points are identified faster. Moreover, by using a monochromatic laser projector, the pattern can be isolated by means of optical filters. This paper proposes a new device that permits the acquisition of three dimensional information from a unique image avoiding the correspondence problem present in most stereo vision systems. Moreover, the use of monochromatic light implies a rapid and accurate image segmentation reducing the computational requirements. A set of suitable lenses and mirror shapes and several possible panoramic catadioptrics sensors are presented and, finally, a new stereo catadioptric sensor based on laser structured light projection is introduced and studied. This work provides extensive information on the modelling and using of such sensors and represents a compulsory step in the designing process.

5. REFERENCES

- [1] S. Baker and S.K. Nayar, "A theory of catadioptric image formation", *Sixth International Conference on Computer Vision*, Page(s): 35 -42, 1998.
- [2] K. Yamazawa, Y. Yagi, M. Yachida, "Omnidirectional imaging with hyperboloidal projection", *Proceedings of the 1993 IEEE/RSJ International Conference on Intelligent Robots and Systems '93*, Volume: 2 , Page(s): 1029 -1034, 1993
- [3] Y. Yagi, M.Yachida, "Omnidirectional Sensing and Its Applications", *IEICE Trans. Inf. & SYST.*, Vol.E82-D, No.3, March 1999.
- [4] C. Pegard, E.M. Mouaddib, "A mobile robot using a panoramic view", *Proceedings. of IEEE International*

- Conference on Robotics and Automation*, page(s): 89 -94 vol.1, 22-28, Apr. 1996.
- [5] M. Fiala, A. Basu, "Feature extraction and calibration for stereo reconstruction using non-SVP optics in panoramic stereo-vision sensor", *Proceedings of the IEEE Workshop on Omni directional Vision*, Page(s): 79 -86, 2002.
- [6] M. Ollis, H. Herman and S. Singh, "Analysis and Design of Panoramic Stereo Vision Using Equi-Angular Pixel Cameras", *tech. report CMU-RI-TR-99-04*, Robotics Institute, Carnegie Mellon University, January, 1999.
- [7] D. Southwell, A. Basu, M. Fiala, M., J. Reyda, "Panoramic stereo", *Proceedings of the 13th International Conference on Pattern Recognition*, Volume: 1, Page(s): 378 -382 1996.
- [8] J. Baldwin, A. Basu, "3D Estimation Using Panoramic Stereo", *Proceedings of 15th International Conference on Pattern Recognition*, Page(s): 97 -100 vol.1, 2000.
- [9] J. Gluckman, S.K. Nayar and Thoresz K.J., "Real-Time Omnidirectional and Panoramic Stereo", *Proceedings of the 1998 DARPA Image Understanding Workshop*, Monterey, California, November 1998.
- [10] S.A. Nene, S.K. Nayar, "Stereo with mirrors", *Sixth International Conference on Computer Vision*, Page(s): 1087 -1094, 1998.
- [11] J. Battle, E. Mouaddib and J. Salvi. A Survey: Recent Progress in Coded Structured Light as a Technique to Solve the Correspondence Problem. *Pattern Recognition* 31(7), pp 963-982, July 1998.