Obtaining the Inverse Distance Map from a Non-SVP Hyperbolic Catadioptric Robotic Vision System

Bernardo Cunha, José Azevedo, Nuno Lau, and Luis Almeida

LSE-IEETA/DETI, Universidade de Aveiro, Portugal
{mbc,jla,lau,lda}@det.ua.pt

Abstract. The use of single viewpoint catadioptric vision systems is a common approach in mobile robotics, despite the constraints imposed by those systems. A general solution to calculate the robot centered distances map on non-SVP catadioptric setups, exploring a back-propagation ray-tracing approach and the mathematical properties of the mirror surface is discussed in this paper. Results from this technique applied in the robots of the CAMBADA team (Cooperative Autonomous Mobile Robots with Advanced Distributed Architecture) are presented, showing the effectiveness of the solution.

Keywords: Omnidirectional vision, robot vision, visualization.

1 Introduction and Related Work

The use of a catadioptric omni-directional vision system based on a regular video camera pointed at a hyperbolic mirror is a common solution for the main sensorial element found in a significant number of autonomous mobile robot applications. This is the case of the Middle Size Robocup Competition, where most of the teams adopt this approach for their robots vision sub-system [1-5]. This ensures an integrated perception of all major target objects in the robots surrounding area, allowing a higher degree of maneuverability at the cost of higher resolution degradation with growing distances away from the robot [6] when compared to non-holonomic setups. For most practical applications, this setup requires the translation of the planar field of view, at the camera sensor plane, into real world coordinates at the ground plane, using the robot as the center of this system. To simplify this non-linear transformation, most practical approaches choose to create a mechanical geometric setup that ensures a symmetrical solution by means of single viewpoint (SVP) approach [1][2][5]. This calls for a precise alignment of the four major points comprising the vision setup: the mirror focus, the mirror apex, the lens focus and the center of the image sensor. It also demands the sensor plane to be both parallel to the ground field and normal to the mirror axis of revolution, and the mirror foci to be coincident with the effective viewpoint and the camera pinhole respectively [7]. This approach generally precludes the use of low cost video cameras, due to the commonly found problem of translational and angular misalignment between the CCD sensor and the lens plane and focus. In this paper we describe a general solution to calculate the robot centered distances map on non-SVP catadioptric setups, exploring a back-propagation ray-tracing approach, also known as "*bird's eye view*", and the mathematical properties of

U. Visser et al. (Eds.): RoboCup 2007, LNAI 5001, pp. 417–424, 2008.

the mirror surface [8][9]. This solution effectively compensates for the misalignments that may result either from a simple mechanical setup or from the use of low cost video cameras. Results from this technique, applied to the robots of the CAMBADA team (Cooperative Autonomous Mobile Robots with Advanced Distributed Architecture), are presented.

2 The Framework

In the following discussion we will assume a specific setup comprising a catadioptric vision module mounted on top of a mechanical structure (figure 1a)). It includes a low cost Fire-I BCL 1.2 Unibrain camera with a 3.6mm focal distance inexpensive lens. The main characteristics of this sensor can be depicted in figure 1b).

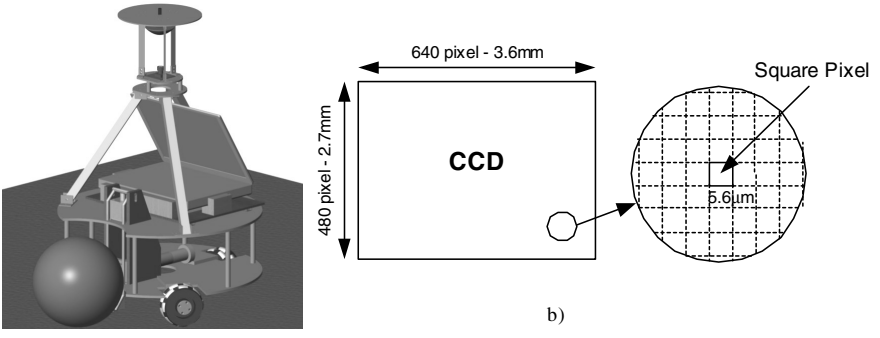

Fig. 1. a) The robot setup with the top catadioptric vision system. b) The Unibrain camera CCD main characteristics.

The used mirror has a hyperbolic surface, described by the following equation:

$$\frac{y^2}{1000} - \frac{\left(x^2 + z^2\right)}{1000} = 1 \quad (\text{mm}) \,. \tag{1}$$

where ***y*** is the mirror axis of revolution and ***z*** is the axis parallel to a line that connects the robot center to its front. Height from the mirror apex to the ground plane is roughly 650mm. Some simplifications will also be used in regard with the diffraction part of the setup. The lens has a narrow field of view and must be able to be focused at a short distance. This, together with the depth of the mirror, implies a reduced depth of field and therefore an associated defocus blur problem [6]. Fortunately, since spatial resolution of the acquired mirror image is significantly reduced with distance, this problem has a low impact in the solution when compared with the low-resolution problem itself. A narrow field of view, on the other hand, also reduces achromaticity aberration and radial distortion introduced by the lens. Camera/lenses calibration procedures are a well-known problem and are widely described in the literature [10][11] – e.g Zhang's method. We will also assume that the pinhole model can provide an accurate enough approach for our practical setup, therefore disregarding any radial distortion of the lens.

3 Discussion

3.1 Initial Approach

Lets assume a restricted setup as in fig. 2. Assumptions of this setup are as follows:

- The origin of the coordinate system is coincident with the camera pinhole through which all light rays will pass;
- **i**, **j** and **k** are unit vectors along axis ***X***, ***Y*** and ***Z***, respectively;
- The ***Y*** axis is parallel to the mirror axis of revolution;
- CCD major axis is parallel to the ***X*** system axis;
- CCD plane is parallel to the **XZ** plane;
- Mirror foci do not necessarily lie on the ***Y*** system axis;
- The vector that connects the robot center to its front is parallel and have the same direction as the positive system ***Z*** axis;
- Distances from the lens focus to the CCD plane and from the mirror apex to the **XZ** plane are ***htf*** and ***mtf*** respectively and can be readily available from the setup and from manufacturer data.
- Point ***Pm***(m_{cx}, ***0***, m_{cz}) is the intersection point of the mirror axis of revolution with the **XZ** plane;
- Distance unit used throughout this discussion will be the millimeter.

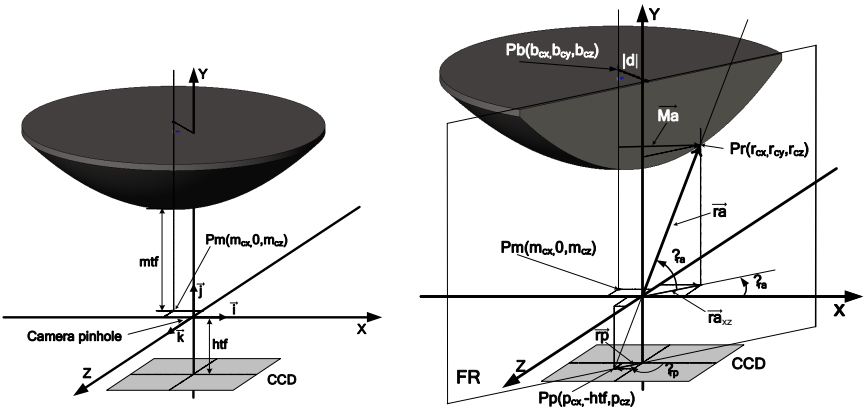

Fig. 2. left) The restricted setup with its coordinate system axis (*X, Y, Z*), (*mirror*) and (*CCD*). The axis origin is coincident with the camera pinhole (figure objects are not drawn to scale). right) A random pixel in the CCD sensor plane is the start point for the back propagation ray.

Mapping equation (1) it into the defined coordinate system, we get

$$y = \sqrt{1000 + (x - m_{cx})^2 + (z - m_{cz})^2} + K_{off} \quad \text{where} \quad k_{off} = mtf - \sqrt{1000}\ . \qquad (2,3)$$

Assuming a randomly selected CCD pixel (X_x,X_z), at point ***Pp***(p_{cx}, ***-htf***, p_{cz}) (fig. 2 b)), the back propagation ray that crosses the origin, may or may not intersect the mirror surface. This can be evaluated from the ray vector equation, solving for ***y=mtf+md***, where ***md*** is the mirror depth, and obtaining the distance module from the mirror

center. If this module is greater than the mirror maximum radius then the ray will not intersect the mirror and the selected pixel will not contribute to the distance map.

Assuming now that this particular ray will intersect the mirror surface, we can then define a plane **FR**, normal to **XZ** and containing this line, equated by

$$z = x\tan(\alpha_{ra}). \tag{4}$$

The line containing position vector **ra,** can then be expressed as a function of ***X*** as

$$y = x\tan(\beta_{ra})/\cos(\alpha_{ra}). \tag{5}$$

Substituting (4) and (5) into (2) we get the equation of the line of intersection between the mirror surface and plane **FR**. ***Pr***, can then be determined from the equality

$$\frac{x\tan(\beta_{ra})}{\cos(\alpha_{ra})} = \sqrt{1000 + (x - M_{cx})^2 + (x\tan(\alpha_{ra}) - M_{cz})^2} + K_{off}. \tag{6}$$

which can be transformed into a quadratic equation of the form

$$ax^2 + bx + c = 0. \tag{7}$$

where

$$a = \left(1 + k_{tn}{}^2 - k_{tc}{}^2\right). \quad b = 2\left(k_{tc}k_{off} - k_{tn}M_{cz} - M_{cx}\right). \tag{8,9}$$

$$c = 1000 + M_{cz}{}^2 + M_{cx}{}^2 - K_{off}{}^2. \tag{10}$$

and

$$k_{tc} = \frac{\tan(\beta_{ra})}{\cos(\alpha_{ra})} \qquad k_{tn} = \tan(\alpha_{ra}). \tag{11}$$

Having found ***Pr***, we can now consider the plane **FN** (fig. 3 a)) defined by ***Pr*** and by the mirror axis of revolution. The angle of the normal to the mirror surface at point ***Pr*** can be equated from the derivative of the hyperbolic function at that point, as a function of **|Ma|,**

$$\frac{\partial h}{d|M_a|} = \frac{|M_a|}{\sqrt{1000 + |M_a|^2}} \qquad \beta_{tm} = \tan^{-1}\left(\frac{|M_a|}{\sqrt{1000 + |M_a|^2}}\right) - \pi/2. \tag{12}$$

This normal line intercepts the **XZ** plane at point ***Pn.*** The angle between the incident ray and the normal at the incidence point can be obtained from the dot product between the two vectors, -**ra** and **rn**. Solving for ϕ_{rm}:

$$\phi_{rm} = \cos^{-1}\left(\frac{r_{cx}(r_{cx} - n_{cx}) + r_{cy}(r_{cy} - n_{cy}) + r_{cz}(r_{cz} - n_{cz})}{|ra||rn|}\right). \tag{13}$$

The reflection ray vector, **rt**, starts at point ***Pr*** and lies on a line going through point ***Pt.*** Its line equation will therefore be

$$P = (r_{cx}i + r_{cy}j + r_{cz}k) + u((t_{cx} - r_{cx})i + (t_{cy} - r_{cy})j + (t_{cz} - r_{cz})k). \tag{14}$$

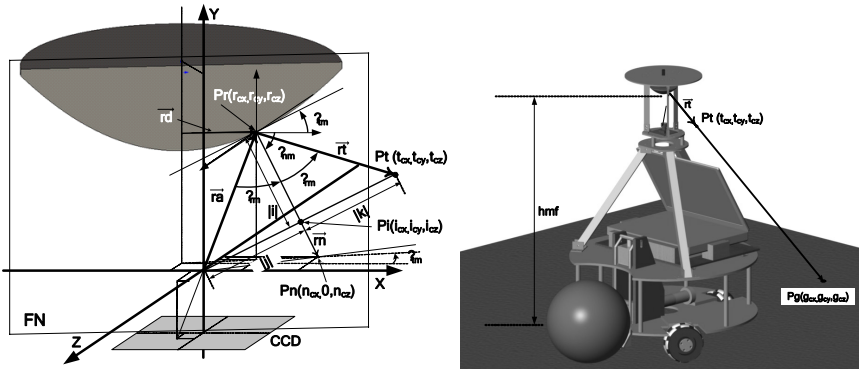

Fig. 3. left) Determining the normal to the mirror surface at point ***Pr*** and the equation for the reflected ray. right) (***Pg***) will be the point on the ground plane for the back-propagation ray.

The point ***Pg*** can then be obtained from the mirror to ground height ***hmf***, and from the ground plane and **rt** line equations (fig. 3 b)), which, evaluating for ***u***, gives

$$u = \frac{(mtf - hmf) - r_{cy}}{t_{cy} - r_{cy}}. \tag{15}$$

3.2 Generalization

To generalize this approach we must now consider the following misalignment factors: 1) The CCD plane may not be parallel to the **XZ** plane; 2) The CCD minor axis may not be correctly aligned with the vector that connects the robot center to its front; 3) The mirror axis of rotation may not be normal to the ground plane.

The first of these factors may result from two different sources: the CCD plane not being parallel to the lens plane; and the mirror axis of rotation being not normal to the CCD plane. Since both effects result in geometrical transformations of the setup, we will integrate these two contributions in the CCD plane, therefore providing a simpler solution. The second of the misalignment factors, on the other hand, can also be integrated as a rotation angle around the ***Y*** axis. To generalize the solution for these two correction factors, we will assume a CCD center point translation offset given by (***-dx, 0, -dy***), and three rotation angles applied to the sensor: γ, ρ and θ, around the ***Y'***, ***X'*** and ***Z'*** axis respectively. These four geometrical transformations upon the original ***Pp*** pixel point can be obtained from the composition of the four homogeneous transformation matrices,

$$R_x(\rho) \bullet R_y(\gamma) \bullet R_z(\theta) \bullet T = \begin{bmatrix} t1_{\rho\gamma\theta} & t2_{\rho\gamma\theta} & t3_{\rho\gamma} & d_x \\ t1_{\rho\theta} & t2_{\rho\theta} & t3_{\rho} & 0 \\ t1_{\rho\gamma\theta} & t2_{\rho\gamma\theta} & t3_{\rho\gamma} & d_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{16}$$

The new start point ***Pp'(p'cx, p'cy, p'cz)***, already translated to the original coordinate system, can therefore be obtained from the following three equations:

$$\begin{aligned} p'_{cx} &= p_{cx}(\cos(\gamma)\cos(\theta) + \sin(\rho)\sin(\gamma)\sin(\theta)) + p_{cz}(\sin(\gamma)\cos(\rho)) + d_x \\ p'_{cy} &= p_{cx}(\cos(\rho)\sin(\theta) + p_{cz}\sin(\rho) - htf \\ p'_{cz} &= p_{cx}(-\sin(\gamma)\cos(\theta) + \sin(\rho)\cos(\gamma)\sin(\theta)) + p_{cz}(\cos(\gamma)\cos(\rho)) + d_z \end{aligned} \tag{17}$$

Analysis of the remaining problem can now follow from (5) substituting ***Pp'*** for ***Pp***.

Finally we can also deal with the third misalignment pretty much in the same way. We just have to temporary shift the coordinate system origin, assume the original floor plane equation defined by its normal vector **j**, and perform a similar geometrical transformation to this vector. This time, however, only rotation angles **ρ** and **θ** need to be applied. The new unit vector **g**, will result as

$$\begin{aligned} g_{cx} &= -\sin(\theta) \\ g_{cy} &= \cos(\rho)\cos(\theta) - mtf + hmf \\ g_{cz} &= \sin(\rho)\cos(\theta) \end{aligned} \tag{18}$$

The rotated ground plane can therefore be expressed in Cartesian form as

$$g_{cx}X + g_{cy}Y + g_{cz}Z = g_{cy}(mtf - hmf) \tag{19}$$

Replacing the **rt** line equation (14) for the ***X***, ***Y*** and ***Z*** variables into (19), the intersection point can be found as a function of ***u***. Note that we still have to check if **rt** is parallel to the ground plane – which can be done by means of the **rt** and **g** dot product. This cartesian product can also be used to check if the angle between **rt** and **g** is obtuse, in which case the reflected ray will be above the horizon line.

3.3 Obtaining the Model Parameters

Some of the parameters needed to obtain the distance map can be measured from the setup itself, e.g., the ground plane rotation relative to the mirror base. A half degree and 0.5mm precision has been proven enough for practical results. Other parameters can be extracted from algorithmic analysis of the image or from a mixed approach. Consider, for instance, thin lens law

$$f = \frac{g}{1 + G/B} \ . \tag{20}$$

G/B is readily available from the diameter of the mirror outer rim in the sensor image; ***g*** can be easily obtained from the practical setup while ***f*** and the actual pixel size are defined by the sensor and lens manufacturers. Since ***G/B*** is also the ratio of distances between the lens focus and both the focus plane and the sensor plane, the ***g*** value can also be easily obtained. The main image features used in this automatic extraction are the mirror outer rim diameter and eccentricity, the center of the mirror image, the center of the robot image, and both the radius, distance and eccentricity of the game field lines – mainly the mid-field circle, lateral and area lines.

4 Support Visual Tools and Results

Although misalignment parameters can actually be obtained from a set of features in the acquired image, the resulting map can still present minor distortions. This is due to the fact that spatial resolution on the mirror image greatly degrades with distance. Since parameter extraction depends on feature recognition on the image, degradation of resolution actually places a bound on feature extraction fidelity. To allow further trimming of these parameters, two simple image feedback tools have been developed.

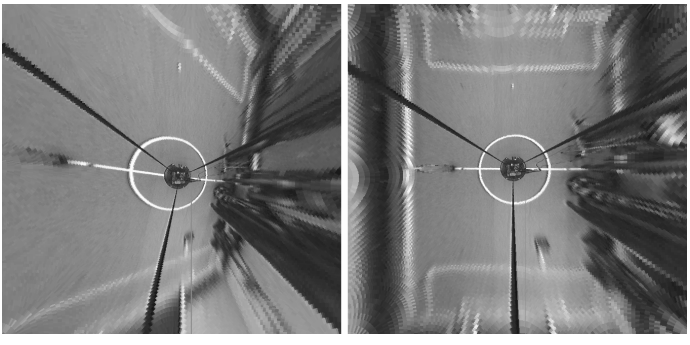

Fig. 4. Acquired image after reverse-mapping into the distance map. On the left, the map was obtained with all misalignment parameters set to zero. On the right, after automatic correction.

The first one creates a reverse mapping of the acquired image into the real world distance map. A fill-in algorithm is used to integrate image data in areas outside pixel mapping on the ground plane. (fig. 4).

The second generates a visual grid with 0.5m distances between both lines and columns, which is superimposed on the original image. This provides an immediate visual clue for the need of possible further distance correction (fig. 5). Since the mid-field circle used in this setup has exactly an outer diameter of 1m, incorrect distance map generation will be emphasized by grid and circle misalignment.

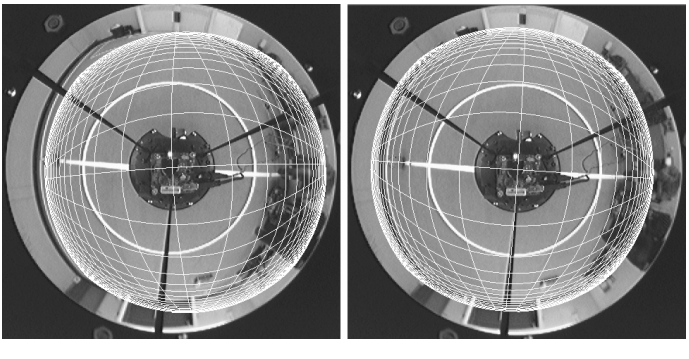

Fig. 5. A 0.5m grid, superimposed on the original image. On the left, with all correction parameters set to zero. On the right, the same grid after geometrical parameter extraction.

Comparison between real distance values measured at more than 20 different field locations and the values taken from the generated map, have shown errors always bellow twice the image spatial resolution. These results are perfectly within the required bounds for the robot major tasks, namely object localization and self-localization on the field.

5 Conclusions

Use of low cost cameras in a general-purpose omni-directional catadioptric vision system, without the aid of any precision adjustment mechanism, will normally preclude the use of a SVP approach. To overcome this limitation, this article explores a "*birds eye view*" algorithm to obtain the ground plane distance map in the CAMBADA football robotic team. Taking into account the intrinsic combined spatial resolution of mirror and image sensor, the method provides viable and useful results that can actually be used in practical robotic applications. This method is supported by a set of image analysis algorithms that can effectively extract the parameters needed to obtain a distance map with an error within the resolution bounds. Further trimming of these parameters can be manually and interactively performed, in case of need, with the support of a set of visual feedback tools that provide the user with an intuitive solution for analysis of the obtained results.

References

1. Zivkovic, Z., Booij, O.: How did we built our hyperbolic mirror omni-directional camera - practical issues and basic geometry. Intelligent Systems Laboratory Amsterdam, University of Amsterdam, IAS technical report IAS-UVA-05-04 (2006)
2. Wolf, J.: Omnidirectional vision system for mobile robot localization in the Robocup environment. Master's thesis, Graz, University of Technology (2003)
3. Menegatti, E., Nori, F., Pagello, E., Pellizzari, C., Spagnoli, D.: Designing an omnidirectional vision system for a goalkeeper robot. In: Birk, A., Coradeschi, S., Tadokoro, S. (eds.) RoboCup 2001. LNCS (LNAI), vol. 2377, pp. 78–87. Springer, Heidelberg (2002)
4. Menegatti, E., Pretto, A., Pagello, E.: Testing omnidirectional vision-based Monte Carlo localization under occlusion. In: Intelligent Robots and Systems (IROS 2004). IEEE/RSJ, vol. 3, pp. 2487–2493 (2004)
5. Lima, P., Bonarini, A., Machado, C., Marchese, F., Marques, C., Ribeiro, F., Sorrenti, D.: Omni-directional catadioptric vision for soccer robots. Robotics and Autonomous Systems 36(2-3), 87–102 (2001)
6. Baker, S., Nayar, S.K.: A theory of single-viewpoint catadioptric image formation. International Journal of Computer Vision 35(2), 175–196 (1999)
7. Benosman, R., Kang, S.B.: Panoramic Vision. Springer, Heidelberg (2001)
8. Blinn, J.F.: A Homogeneous Formulation for Lines in 3D Space. In: SIGGRAPH 1977, pp. 237–241 (1977)
9. Foley, J.D., van Dam, A., Feiner, S.K., Hughes, J.F.: Computer Graphics: Principles and Practice in C, 2nd edn. Addison-Wesley Professional (1995)
10. Zhang, Z.: A flexible new technique for camera calibration. IEEE Transactions on Pattern Analysis and Machine Intelligence 22(11), 1330–1334 (2000)
11. Hartley, R., Zisserman, A.R.: Multiple View Geometry in Computer Vision, Cambridge University Press (2003)