OMNIDIRECTIONAL STEREOVISION SYSTEM WITH TWO-LOBE HYPERBOLIC MIRROR FOR ROBOT NAVIGATION

Fabiano Rogério Corrêa
Departamento de Engenharia Mecatrônica e Sistemas Mecânicos
Escola Politécnica of the University of São Paulo
Av. Prof. Mello Moraes, 2231
05508-9000 São Paulo, SP, Brazil
fabiano.correa@poli.usp.br

Vitor Campanholo Guizilini
Departamento de Engenharia Mecatrônica e Sistemas Mecânicos
Escola Politécnica of the University of São Paulo
Av. Prof. Mello Moraes, 2231
05508-9000 São Paulo, SP, Brazil
vitor.guizilini@poli.usp.br

Jun Okamoto Junior
Departamento de Engenharia Mecatrônica e Sistemas Mecânicos
Escola Politécnica of the University of São Paulo
Av. Prof. Mello Moraes, 2231
05508-9000 São Paulo, SP, Brazil
jokamoto@usp.br

Abstract. *Vision systems are widely used in autonomous robots. Images are a rich source of information about the surroundings of robots and can be used to calculate distances of the objects in the scene to avoid collision, for tracking of objects and localization. In particular, omnidirectional vision systems have gained popularity and several different implementations for mobile robots. With those systems, a robot can have information of the environment in all directions with only one image. In a previous implementation an omnidirectional system with a hyperbolic mirror was used as a range sensor for a mobile robot navigation task. In that case, the robot needed to move to acquire a pair of images for the stereovision processing. To eliminate this movement, we present here a new mirror design, a two-lobe hyperbolic mirror to acquire one image with information suited for stereo processing as the same scene appears distinctly in each lobe of the mirror. This new system is presented with all the mathematical formulation required to the design of the two-lobe mirror. All these are explained in detail and the results of the stereovision processing algorithm that extracts range information from the omnidirectional image is shown for simulated images.*

Keywords: *omnidirectional vision, stereovision, mobile robot, mirror design*

1. Introduction

As computers increased their processing velocity and algorithms of image processing were improved, images have been used as a rich source of information to solve many problems in mobile robotics. Currently, vision systems are widely used in navigation of mobile robots (DeSouza *et al.,* 2002). Autonomous mobile robots rely on their sensors to cope with dynamic environments.

Different vision systems were proposed for mobile robots, such as omnidirectional vision systems that have a field of view of 360 degrees. It can be designed in many ways, but catadioptric systems that use a camera and a convex mirror have gained enormous popularity. It is compact and has proved to be useful in robotics applications (Delahoche *et al.* 1997; Winters *et al.*, 2000; Yagi *et al.*, 1995; Patel *et al.*, 2002; Matsumoto *et al.*, 1999), as it does not require any movement of the camera to acquire information in all directions. One application of this system in robotics is as a range sensor.

To calculate distances using images, it is necessary that the same scene in the real world be represented in at least two different perspectives in the images. In our previous implementation of omnidirectional stereovision (Correa, Deccó and Okamoto, 2003), a hyperbolic mirror was used with a conventional camera to acquire range information for a mobile robot navigation task. As the vision system had just one camera, the robot should move about 300 mm to acquire a pair of images for the stereovision processing. Although errors in the algorithm due to the displacement were minor, it had a dependency of the odometry.

To eliminate the movement of the robot, we present here a new mirror design, a two-lobe hyperbolic mirror to acquire one image with information suited for stereo processing as the same scene appears distinctly in each lobe of the mirror. This design was first proposed by Southwell (1996). It consists of two coaxial hyperboles. Each hyperbolic mirror has two foci: one virtual inside the mirror and one real outside where camera's focus must be positioned. In the case of our two-lobe mirror the real foci of each lobe are coincident.

All stereo algorithms have three basic steps: feature extraction, matching and triangulation. Matching is the step that is more prone to errors. It consists in the correspondence of pixels of different images (or in this case, different lobes). Some restrictions could be applied in the search for the best match. A restriction that is normally used is the epipolar geometry. In our mirror, the epipolar lines are radial in the image thus facilitating the search.

Also, matching in stereo algorithms determines the type of the results. In this implementation we use dynamic programming to increase the density of the resultant map although the computational cost is increased. With dynamic programming it is possible to include many restrictions as continuity and uniqueness in the pixel correspondence.

This new system is presented with all the mathematical formulation required to the design of the two-lobe mirror. Several issues had to be solved, such as amount of resolution in each image, epipolar restriction due to differences in the images, limit in the field of view, among others.

The paper follows this specific structure: in Section 2, we present the design of the mirror and all properties considered in the design. Section 3 contains the stereo algorithm and the matching implementation using dynamic programming. Results are presented in Section 4.

2. Design of a two-lobe hyperbolic mirror

Hyperbolic mirrors when used with conventional cameras have the property of having a single-viewpoint. With this property the projection of points in the image is easily calculated. It means that the incidence of rays of light necessarily passes through the foci of the mirror.

When the mirror is composed by two-lobes, the design must consider that the real foci of those two lobes must be coincident so that they have the single-viewpoint property.

Figure 1 shows the projection properties of this type of mirror.

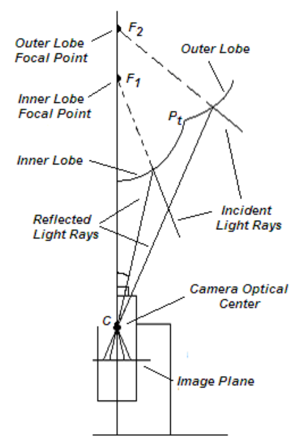

Figure 1. Projections on the two-lobe hyperbolic mirror

So we must determine the intersection of two hyperboles with the restriction that they possess a common real focus. Equation (1) shows the generic equation of a bi-dimensional hyperbole.

$$z = \frac{a}{b}\sqrt{x^2 + b^2} \tag{1}$$

where a, b and $c = \sqrt{a^2 + b^2}$ are parameters of the hyperbole.

We must add the difference between the foci of both hyperboles to the equation of the superior hyperbole so that the real foci to be coincident. Equation (2) shows the equation that determines the intersection between the hyperboles and eq. (3) the radius where it occurs.

$$\left[\left(\frac{a_1}{b_1}\right)^2 - \left(\frac{a_2}{b_2}\right)^2\right]x^2 + \left[\left(a_1{}^2 - a_2{}^2\right) - \left(c_2 - c_1\right)^2\right] = 2\left(\frac{a_2}{b_2}\right)\left(c_2 - c_1\right)\sqrt{x^2 + b_2{}^2} \tag{2}$$

$$r = \sqrt{\frac{\left(C^2 - 2AB\right) \pm \sqrt{\left(2AB - C^2\right)^2 - 4A^2\left(B^2 - C^2 b_2{}^2\right)}}{2A^2}} \tag{3}$$

where:

$$\begin{aligned} A &= \left[\left(\frac{a_1}{b_1}\right)^2 - \left(\frac{a_2}{b_2}\right)^2\right] \\ B &= \left[\left(a_1{}^2 - a_2{}^2\right) - \left(c_2 - c_1\right)^2\right] \\ C &= 2\left(\frac{a_2}{b_2}\right)\left(c_2 - c_1\right)\sqrt{x^2 + b_2{}^2} \end{aligned} \tag{4}$$

To choose appropriate parameters for the hyperboles, desired geometric properties must be defined. It is necessary to determine the maximum and minimum distance that the mirror can cover. Another property is the amount of pixels related to the two lobes that reproduces the same part of the ambient, that is used by the stereo algorithm. Furthermore, the size of the mirror is a important property as the vision system is mounted on a mobile robot.

The area of the image that could be used to produce stereo information is determined by the intersection between rays of light that passes in focus and the point that separate one lobe from the other. This also limits the maximum and minimum range that could be covered by the mirror.

It is important that both the mirrors' foci are precisely aligned with the camera's focal point. Here we don't consider some deviations in the projection's geometry that would be matter of future work.

P_t represents the point where both hyperboles intersect, representing the transition where one image ends and the other starts. Knowing this, vector $\boldsymbol{t}$ is determined by the union of points P_t and C (which is the optical center of the camera). The intersection of $\boldsymbol{t}$ and the image plane λ determines I_t, the image point of this transition.

P_t also determines the minimal and maximum range of this catadioptric system, which is related with the parameters of the hyperboles. This range is obtained through vectors $\boldsymbol{t_{min}}$ and $\boldsymbol{t_{max}}$, which are the vectors created respectively by the union of points (F_2, P_t) and (F_1, P_t). And the intersection of $\boldsymbol{t_{min}}$ and $\boldsymbol{t_{max}}$ with the floor plane ξ determines points R_{min} and R_{max}, the minimum and maximum range of this catadioptric system. With these two points it is possible to find their correspondents in the other hyperbole, creating the first correspondence between the two views in the image and setting the range where the stereo algorithm will work. $\boldsymbol{i_{min}}$ and $\boldsymbol{i_{max}}$ represent the union between points (R_{min}, F_1) and (R_{max}, F_2), and with them the points P_{min} and P_{max}, their intersection with hyperbole H_1 and H_2, can be determined.

We limit the maximum radius of the mirror to 35 mm. The desirable field of view is determined around 0.5 m and 6 m.

3. Stereovision algorithm

3.1. Basics

Research in stereo algorithms is mostly done with conventional vision systems (Brown *et al.*, 2003). In the literature, when an omnidirectional vision system is used to acquire stereo information, the image is unwrapped based in the mirror profile so that traditional solutions could be used. Basically, a stereo algorithm can be divided in three parts: feature extraction, matching and triangulation.

The choice of the feature to be extracted is somewhat related to the quality of the final result and type of the matching process. A feature could be a point, a line or curve, or an area or object in the image.

In the matching process, the features extracted in one image of the stereo pair must be corresponded to points in the second one based in a correlation measure. In this step, constraints are applied to reduce computational cost in the search for correlation. Restrictions normally used are epipolar geometry, uniqueness of correspondence and continuity of vicinity. A correlation function must be chosen to compare candidate points. The matching could be area-based or feature-based. This part of the stereo algorithm is subject to a lot of errors producing false matching.

Given the matched features, the triangulation is calculated based on geometrical properties of the vision system to determine the distance between the camera and the point in the real world that corresponds to the features.

The specific implemented stereo algorithm is described bellow. Here, we use directly the omnidirectional image.

3.2. Restrictions applied to the matching

Some restrictions must be applied in the search for correspondence. Epipolar constraints reduce the search of two to one dimension. Epipolar geometry determines a geometric place where the matching point must be. It is calculated by a given point in one image with respect to some point in the space and characteristics of the vision system. The epipolar geometry for catadioptric system was described by Svoboda (1998).

For each pixel in the interior lobe of the mirror, the search for the correlated is done under a radial line that passes in it. This line is the epipolar line.

Figure 2 shows three epipolar lines traced in the omnidirectional image. The feature extracted in the image is based by the pixels in the inner lobe.

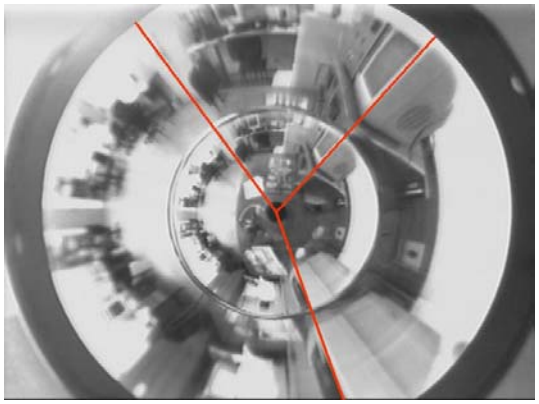

Figure 2. Epipolar lines in the omnidirectional image

3.3. Triangulation

Once we have the correspondent points (i_1,j_1) and (i_2,j_2) determined by the matching, it is possible to calculate the triangulation. The pixel coordinates in image plane is translated to the center of the omnidirectional image. It is necessary to transform the pixel coordinates to mm with the extrinsic camera parameters. This coordinates are used in the equations below.

To find the vector that point to a real object in space, given a pixel in image plane, the problem to be solved is a system equation of the intersection between a line (ray of light) and a hyperbole (mirror's surface).

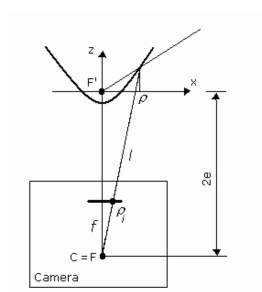

Figure 3. Catadioptric system

Line l represents the ray of light that reflects in the mirror's surface and goes to the camera's focus marking a pixel in the image plane:

$$l : z = \frac{f}{\rho_i}\rho - 2c \tag{5}$$

where ρ_i is the root square of the sum of squared coordinates of the pixel in the image plane transformed in mm and f is the focal length of the camera in mm.

Hyperbole h is the mirror's surface represented by eq. (1). The intersection of line l with hyperbole h produces a point that with the mirror's focus represent a vector that points to a point in the real environment. Intersection is obtained through substitution of z of eq (1) in eq (5). Solving for ρ**,** we obtain:

$$\rho = \frac{\frac{f}{\rho_i}.e \pm a\sqrt{\left(\frac{f}{\rho_i}\right)^2 + 1}}{\left(\frac{f}{\rho_i}\right)^2 - \left(\frac{a}{b}\right)^2} \tag{6}$$

The height z could be obtained substituting ρ in eq. (1). Applying this equation to the two pixels correspondent one in each lobe, we obtain vectors whose origin are the foci of the mirror and that intersects in a specific point in space. Then, it is possible to calculate this intersection to determine the distance d of point X projected in the floor. The baseline is the distance between the foci of the two lobes and is used in all calculus of distance.

3.4. Matching using Dynamic Programming

In our previous implementation, we applied an area-based matching that used two different images. Restrictions such as uniqueness and continuity were not considered then. To add those restrictions in the matching process, we decided to use Dynamic Programming (DP) to solve the correspondence problem. This approach is commonly used to find an optimal path through a surface. It divides an otherwise time-consuming task in smaller parts, storing the results to use in the next step. Although results may not be locally optimal, it has the property of minimizing globally a determined cost function that is created to fill mathematically the surface. Here, an analogy of the surface is made when each epipolar line is organized as a matrix n_1 x n_2, where n_x is the number of pixels in the view of each lobe.

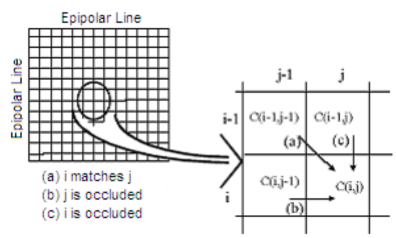

Figure 4. Dynamic Programming Matrix

Once the DP matrix is created, it has to be filled with values created by the cost function, representing how close that cell of the matrix is to a correct match between its line and column. There are several ways of creating a cost function, and this paper relies on three constrains to determine the correct match. These constrains are organized by their computational cost, and if one cell of the matrix fails in any of them it does not go to the next one, saving time and thus allowing a faster method to achieve the desired result. The constraints are as follows:

Intersection Constraint: It only indicates if the match is possible or not between two pixels. It is based on the fact that two incident rays, each one touching a different lobe, will only intersect each other if the angle of incidence of the outer lobe is smaller than the inner lobe.

Single Pixel Constraint: Even with shading and other light variables, same points in space are supposed to have similar gray scale in both images. So, the similarity between two pixels in different images can determine how close they are to a match.

$$I_{(i,j)} = \sqrt{(I_i - I_j)^2} \tag{5}$$

where I_i is the intensity of gray scale of pixel in lobe i. A threshold determines if the intensity of a pixel could be correspondent to another.

Area Pixel Constraint: The region around the pixel is used in order to create correspondences. It is possible to find the corresponding range of each pixel in one image based on floor plane. This range varies as the angle of the reflected light ray increases. This procedure makes it possible to find the pixels that cover the same range in both images, and this interval of pixels (different in each image) is the one used in this constraint.

$$I_{A(i,j)} = \sum_{k=i-N}^{i+N} \left[\frac{(I_k + I_{k+1})}{2} - \frac{\sum_{u=j_k}^{j_{k+1}} I_u}{w_k - v_k} \right]^2 \tag{6}$$

where w_k and v_k are the size of windows of search. The size of the window is proportional to the resolution of each lobe that reproduces the geometry of the mirror.

When the function cost is completely defined, the DP matrix can be implemented, and then the optimal path can be found. In order to do this, another matrix has to be created, when each cells depends directly on the value of the others around it, according to the following equation:

$$F(i,j) = \min\left(F_{(i,j-1)} + C_{(i,j)}, F_{(i-1,j)} + C_{(i,j)}, F_{(i-1,j-1)} + C_{(i,j)}\right) \tag{7}$$

where $C_{(i,j)}$ is the basic cost of the current cell and $F_{(i-1,j-1)}$ the influence of neighbor cells.

Now that the second DP matrix is ready, the optimal can be found by moving through the matrix starting in the bottom right and going left until the value of the cell starts rising, which defines a match. This correspondence pair is stored for later use, and the algorithm goes up one line, starting the process again.

		G	A	A	T	T	C	A	G	T	T	A
	0	0	0	0	0	0	0	0	0	0	0	0
G	0	1	1	1	1	1	1	1	1	1	1	1
G	0	1	1	1	1	1	1	1	2	2	2	2
A	0	1	2	2	2	2	2	2	2	2	2	3
T	0	1	2	2	3	3	3	3	3	3	3	3
C	0	1	2	2	3	3	3	4	4	4	4	4
G	0	1	2	2	3	3	3	4	4	5	5	5
A	0	1	2	3	3	3	3	4	5	5	5	6

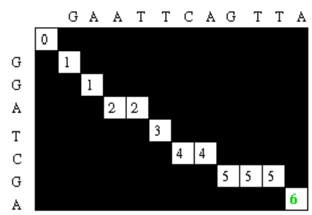

Figure 5. Example of DP Matrix

The matching points stored in during the Dynamic Programming step allow the reconstruction of the environment, using information from both images to find the real cartesian coordinates of each pixel. This is a simple step, because the real point is the intersection of the incident light rays that create each one of the two matched pixels. The process used to find these light rays has been described before, and through it the real point can be easily determined.

4. Results

In Section 2 was presented a mirror design. That mirror was recently lathed in the Ultra-Precision Lathing Laboratory of São Carlos - USP. Using an ultra-precision lathe machine does not require polishing the aluminum mirror in the end of the process.

Figure 6 shows the mirror and fig. 7 represents the environment as seen with the monochromatic camera pointing at the mirror.

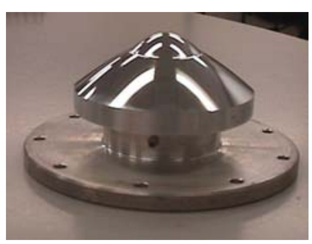

Figure 6. Hyperbolic mirror with two lobes

Figure 7. Omnidirectional image

This vision system was mounted in a mobile robot such that the mirror is placed in a height of approximately 57,8 cm above the floor. Table 1 describes some properties of the mirror.

Table 1. Parameters of the mirror

	Radius (mm)
Internal Lobe Radius	11.4
External Lobe Radius	30.0
Maximum Field of View	~1000
Minimum Field of View	~6000

The stereo algorithm implemented was validated using virtual images. The initial results presented here were obtained using POV-Ray software to model the vision system and the environment. The environment created is composed by two boxes as is seen in fig. 8.

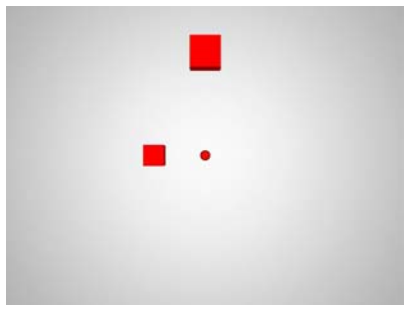

Figure 8. Environment with two boxes

The results presented in fig. 9 show the contours of the objects in the environment and over it the map determined by the stereo algorithm. The result has a precision that is adequate to be used in an autonomous navigation of a mobile robot.

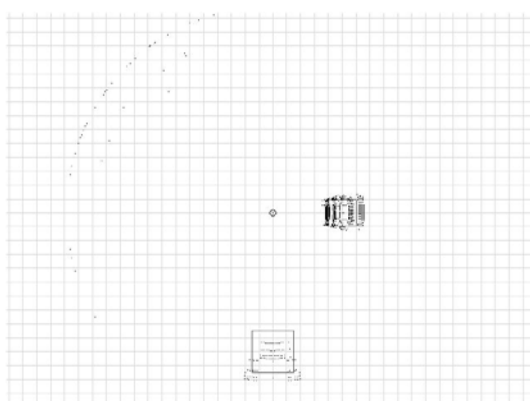

Figure 9. Results from stereovision

5. Conclusions

Omnidirectional vision systems are very useful in autonomous mobile robots. We use this system to produce range information for navigational tasks.

An stereo algorithm was developed to work with a two-lobe hyperbolic mirror. With that mirror design we could produce stereo information in one image. To obtain a dense map of the environment as the result from stereo, we employ dynamic programming in the matching process.

We designed and lathed a two-lobe hyperbolic mirror. Results from stereo algorithm were presented based in virtual images made in POV-Ray software. The results are adequate to be used in autonomous navigation.

6. Acknowledgements

We gratefully acknowledge support from agencies Capes and CNPq in this research.

7. References

Birchfield, S. and Tomasi, C. Depth Discontinuities by Pixel-to-Pixel Stereo. **International Journal of Computer Vision,** 35(3): 269-293, December 1999.

Brown, M. Z., Burschka, D., Hager, G. Advances in computational stereo. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, vol. 25, no. 8, pg. 993-1007, August 2003.

Correa, F.R., Deccó, C.C.G., Okamoto, J. Obtaining range information with an omnidirectional vision system. In: CONGRESSO BRASILEIRO DE ENGENHARIA MECÂNICA (COBEM), XVII, São Paulo, 2003.

Cox, I., Hingorani, S., Rao, S., and Maggs, B. A maximum likelihood stereo algorithm. Computer Vision, Graphics and Image Processing, 63(3): 542--567, 1996.

Delahoche, L., Pégard, C., Marhic, B., Vasseur, P.; A navigation system based on an omnidirectional vision sensor. **IEEE Proceedings IROS 97.**

DeSouza, G. N., Kak, A. C.; Vision for mobile robot navigation: a survey. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, v.24, n.2, p.237-267, February 2002.

Matsumoto, Y., Ikeda, K., Inaba, M., Inoue, H.; Visual navigation using omnidirectional view sequence.

Patel, S., Jung, S., Ostrowski, J. P., Rao, R., Taylor, C. J.; Sensor based door navigation for a nonholonomic vehicle. **Proceedings of the 2002 IEEE, International Conference on Robotics & Automation**, Washington, DC., May 2002.

Southwell, M. F. D,. Basu, A. and Reyda, J. "Panoramic stereo", In ICPR, pp. 378-382, Vienna, 1996.

Svoboda, T., Pajdla, T., Hlavác, V. Epipolar geometry for panoramic cameras. In: EUROPEAN CONFERENCE ON COMPUTER VISION, 5th, Freiburg, Germany, 1998. **Proceedings**, p. 218-232.

Yagi, Y., Nishizawa, Y., Yachida, M.; Map-based navigation for a mobile robot with omnidirectional image sensor COPIS. **IEEE Transactions on Robotics and Automation,** vol. 11, no. 5, October 1995.

Winters, N., Gaspar, J., Lacey, G., Santos-Victor, J.; Omni-directional vision for robot navigation, 2000.

Zhu, Z. Omnidirectional stereo vision, Workshop on Omnidirectional Vision Applied to Robotic Orientation and Nondestructive Testing (NDT), The 10th IEEE International Conference on Advanced Robotics, August 22-25, 2001, Budapest, Hungary (invited talk).

8. Responsibility notice

The authors are the only responsible for the printed material included in this paper.