We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

A 3D Omnidirectional Sensor
For Mobile Robot Applications

Rémi Boutteau, Xavier Savatier, Jean-Yves Ertaud and Bélahcène Mazari
Institut de Recherche en Systèmes Electroniques Embarqués (IRSEEM)
France

1. Introduction

In most of the missions a mobile robot has to achieve – intervention in hostile environments, preparation of military intervention, mapping, etc – two main tasks have to be completed: navigation and 3D environment perception. Therefore, vision based solutions have been widely used in autonomous robotics because they provide a large amount of information useful for detection, tracking, pattern recognition and scene understanding. Nevertheless, the main limitations of this kind of system are the limited field of view and the loss of the depth perception.

A 360-degree field of view offers many advantages for navigation such as easiest motion estimation using specific properties on optical flow (Mouaddib, 2005) and more robust feature extraction and tracking. The interest for omnidirectional vision has therefore been growing up significantly over the past few years and several methods are being explored to obtain a panoramic image: rotating cameras (Benosman & Devars, 1998), muti-camera systems and catadioptric sensors (Baker & Nayar, 1999). Catadioptric sensors, i.e. the combination of a camera and a mirror with revolution shape, are nevertheless the only system that can provide a panoramic image instantaneously without moving parts, and are thus well-adapted for mobile robot applications.

The depth perception can be retrieved using a set of images taken from at least two different viewpoints either by moving the camera or by using several cameras at different positions.

The use of the camera motion to recover the geometrical structure of the scene and the camera's positions is known as Structure From Motion (SFM). Excellent results have been obtained during the last years with SFM approaches (Pollefeys et al., 2004; Nister, 2001), but with off-line algorithms that need to process all the images simultaneous. SFM is consequently not well-adapted to the exploration of an unknown environment because the robot needs to build the map and to localize itself in this map during its world exploration.

The in-line approach, known as SLAM (Simultaneous Localization and Mapping), is one of the most active research areas in robotics since it can provide a real autonomy to a mobile robot. Some interesting results have been obtained in the last few years but principally to build 2D maps of indoor environments using laser range-finders. A survey of these algorithms can be found in the tutorials of Durrant-Whyte and Bailey (Durrant-Whyte & Bailey, 2006; Bailey & Durrant-Whyte, 2006).

Vision-based SLAM algorithms are generally dedicated to monocular systems which are cheaper, less bulky, and easier to implement than stereoscopic ones. Stereoscopic systems have, however, the advantage to work in dynamic environments since they can grab simultaneously two images. Calibration of the stereoscopic sensor enables, moreover, to recover the Euclidean structure of the scene which is not always possible with only one camera.

In this chapter, we propose the design of an omnidirectional stereoscopic system dedicated to mobile robot applications, and a complete scheme for localization and 3D reconstruction. This chapter is organized as follows. Section 2 describes our 3D omnidirectional sensor. Section 3 is dedicated to the modelling and the calibration of the sensor. Our main contribution, a Simultaneous Localization and Mapping algorithm for an omnidirectional stereoscopic sensor, is then presented in section 4. The results of the experimental evaluation of each step, from calibration to SLAM, are then exposed in section 5. Finally, conclusions and future works are presented section 6.

2. System overview

2.1 Sensor description

Among all possible configurations of central catadioptric sensors described by Nayar and Baker (Baker & Nayar, 1999), the combination of a hyperbolic mirror and a camera is preferable for the sake of compactness since a parabolic mirror needs a bulky telecentric lens.

Although it is possible to reconstruct the environment with only one camera, a stereoscopic sensor can produce a 3D reconstruction instantaneously (without displacement) and will give better results in dynamic scenes. For these reasons, we developed a stereoscopic system dedicated to mobile robot applications using two catadioptric sensors as shown in Figure 1.
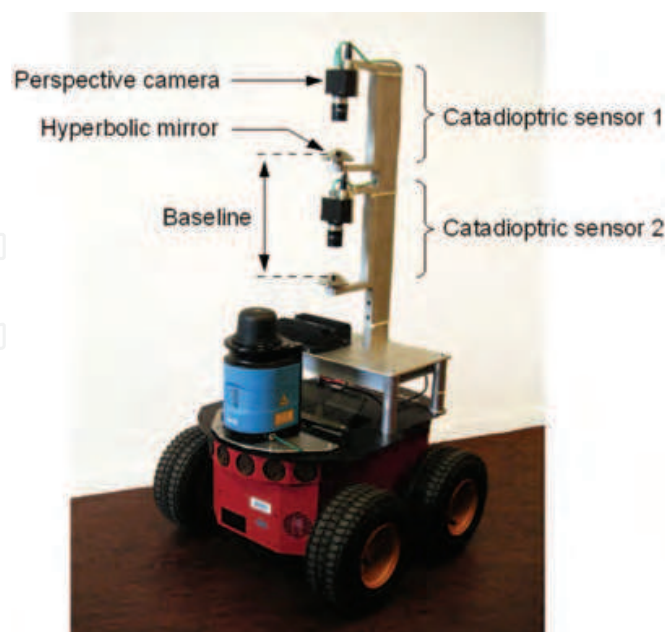

Fig. 1. View of our catadioptric stereovision sensor mounted on a Pioneer robot. Baseline is around 20cm for indoor environments and can be extended for outdoor environments. The overall height of the sensor is 40cm.

2.2 Imposing the Single-Viewpoint (SVP) Constraint

The formation of images with catadioptric sensors is based on the Single-Viewpoint theory (Baker & Nayer, 1999). The respect of the SVP constraint permits the generation of geometrically correct perspective images. In the case of a hyperbolic mirror, the optical center of the camera has to coincide with the second focus **F′** of the hyperbola located at a distance of *2e* from the mirror focus as illustrated in Figure 2. The eccentricity *e* is a parameter of the mirror given by the manufacturer.
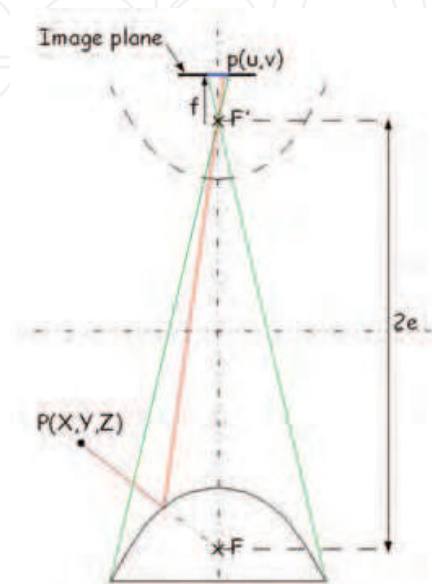

Fig. 2. Image formation with a hyperbolic mirror. The camera center has to be located at *2e* from the mirror focus to respect the SVP constraint.

A key step in designing a catadioptric sensor is to respect this constraint as much as possible. To achieve this, we first calibrate our camera with a standard calibration tool to determine the central point and the focal length. Knowing the parameters of both the mirror and the camera, the image of the mirror on the image plane can be easily predicted if the SVP constraint is respected as illustrated in Figure 2. The expected mirror boundaries are superposed on the image and the mirror has then to be moved manually to fit this estimation as shown in Figure 3.
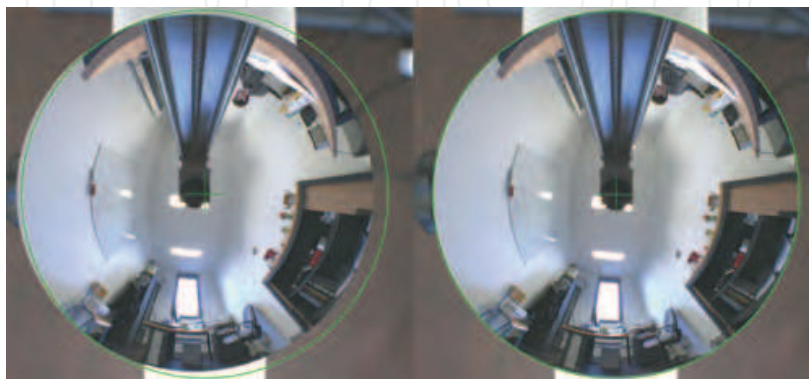

Fig. 3. Adjustment of the mirror position to respect the SVP constraint. The mirror border has to fit the estimation (green circle).

3. Modelling of the sensor

The modelling of the sensor is a necessary step to obtain metric information about the scene from the camera. It establishes the relationship between the 3D points of the scene and their projections into the image (pixel coordinates). Although there are many calibration methods, they can be classified into two main categories: parametric and non-parametric. The first family consists in finding an appropriate model for the projection of a 3D point onto the image plane. Non-parametric methods associate one projection ray to each pixel (Ramalingram et al., 2005) and provide a "black box model" of the sensor. They are well adapted for general purposes but they make more difficult the minimization algorithms commonly used in computer vision (gradient descent, Gauss-Newton, Levenberg-Marquardt, *etc*).

3.1 Projection model

Using a parametric method requires the choice of a model, which is very important because it has an effect on the complexity and the precision of the calibration process. Several models are available for catadioptric sensors: complete model, polynomial approximation of the projection function and generic model.

The complete model relies on the mirror equation, the camera parameters and the rigid transformation between them to calculate the projection function (Gonzalez-Barbosa & Lacroix, 2005). The large number of parameters to be estimated leads to an error function which is difficult to minimize because of numerous local minima (Mei & Rives, 2007). The polynomial approximation of the projection function was introduced by Scaramuzza (Scaramuzza et al., 2006), who proposed a calibration toolbox for his model. The generic model, also known as the unified model, was introduced by Geyer (Geyer & Daniilidis, 2000) and Barreto (Barreto, 2006), who proved its validity for all central catadioptric systems. This model was then modified by Mei (Mei & Rives, 2007), who generalized the projection matrix and also took into account the distortions. We chose to work with the unified model introduced by Mei because any catadioptric system can be used and the number of parameters to be estimated is quite reasonable.
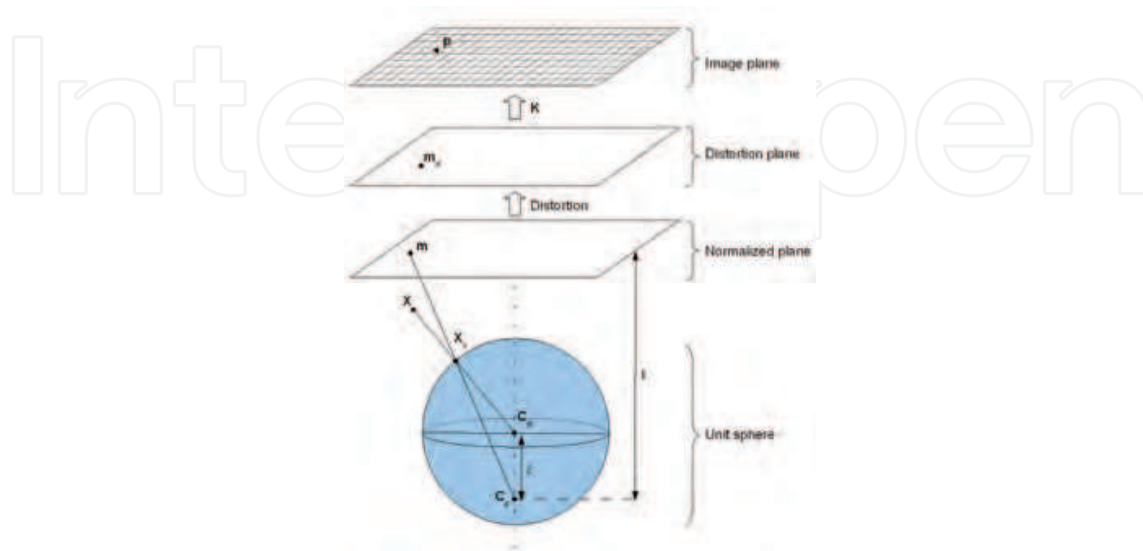

Fig. 4. Unified projection model.

As shown in Figure 4, the projection $\mathbf{p} = \begin{bmatrix} u & v \end{bmatrix}^{\mathrm{T}}$ of a 3D point $\mathbf{X}$ with coordinates $\begin{bmatrix} X_w & Y_w & Z_w \end{bmatrix}^{\mathrm{T}}$ in the world frame can be computed using the following steps:

- The coordinates of the point $\mathbf{X}$ are first expressed in the sensor frame by the rigid transformation W which depends on the seven parameters of the vector $\mathbf{V_1} = \begin{bmatrix} q_w & q_x & q_y & q_z & t_x & t_y & t_z \end{bmatrix}^{\mathrm{T}}$. The first four parameters are the rotation $\mathbf{R}$ parameterised by a quaternion and the three others are those of the translation $\mathbf{T}$. The coordinates of $\mathbf{X}$ in the mirror frame are thus given by:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \mathbf{R} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \mathbf{T} \tag{1}$$

- The point $\mathbf{X} = \begin{bmatrix} X & Y & Z \end{bmatrix}^{\mathrm{T}}$ in the mirror frame is projected from the center onto the unit sphere giving $\mathbf{X_S} = \begin{bmatrix} X_S & Y_S & Z_S \end{bmatrix}^{\mathrm{T}}$. This point is then projected onto the normalized plane from a point located at a distance ξ from the center of the sphere. These transformations are combined in the function H which depends on only one parameter: $\mathbf{V_2} = \begin{bmatrix} \xi \end{bmatrix}$. The projection onto the normalized plane, written $\mathbf{m} = \begin{bmatrix} x & y \end{bmatrix}^{\mathrm{T}}$ is consequently obtained by:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \dfrac{X_S}{Z_S + \xi} \\ \dfrac{Y_S}{Z_S + \xi} \end{bmatrix}, \text{ with } \begin{bmatrix} X_S \\ Y_S \\ Z_S \end{bmatrix} = \begin{bmatrix} \dfrac{X}{\sqrt{X^2 + Y^2 + Z^2}} \\ \dfrac{Y}{\sqrt{X^2 + Y^2 + Z^2}} \\ \dfrac{Z}{\sqrt{X^2 + Y^2 + Z^2}} \end{bmatrix} \tag{2}$$

- Distortions are then added to the point $\mathbf{m}$ using the distortion function D, which depends on 5 coefficients: 3 for radial distortions and 2 for tangential distortions. Let $\mathbf{V_3} = \begin{bmatrix} k_1 & k_2 & k_3 & k_4 & k_5 \end{bmatrix}^{\mathrm{T}}$ be the parameter vector containing these coefficients, $\rho = \sqrt{x^2 + y^2}$, and $\mathbf{m_d} = \begin{bmatrix} x_d & y_d \end{bmatrix}^{\mathrm{T}}$ the resulting point. Its coordinates are obtained by the following equation:

$$\mathbf{m_d} = \begin{bmatrix} x(1 + k_1\rho^2 + k_2\rho^4 + k_5\rho^6) + 2k_3xy + k_4(\rho^2 + 2x^2) \\ y(1 + k_1\rho^2 + k_2\rho^4 + k_5\rho^6) + 2k_4xy + k_3(\rho^2 + 2y^2) \end{bmatrix} \tag{3}$$

- Final projection is a perspective projection involving the projection matrix $\mathbf{K}$. This matrix contains 5 parameters: the generalized focal lengths γ_u and γ_v, the

coordinates of the principal point u_0 and v_0, and the skew α. Let K be this projection function, and $\mathbf{V_4} = \begin{bmatrix} \alpha & \gamma_u & \gamma_v & u_0 & v_0 \end{bmatrix}^{\mathrm{T}}$ be the parameter vector. The projection $\mathbf{p} = \begin{bmatrix} u & v \end{bmatrix}^{\mathrm{T}}$ of the 3D point $\mathbf{X}$ is given by equation (4).

$$\mathbf{p} = \mathbf{K} \cdot \mathbf{m_d} = \begin{bmatrix} \gamma_u & \gamma_u\alpha & u_0 \\ 0 & \gamma_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \mathbf{m_d} \tag{4}$$

Let $\mathbf{V}$ be the global parameter vector, i.e. $\mathbf{V} = \begin{bmatrix} \mathbf{V_1}^{\mathrm{T}} & \mathbf{V_2}^{\mathrm{T}} & \mathbf{V_3}^{\mathrm{T}} & \mathbf{V_4}^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$. The global projection function of a 3D point $\mathbf{X}$, written $P(\mathbf{V}, \mathbf{X})$, is obtained by chain composition of the different functions presented before:

$$P(\mathbf{V}, \mathbf{X}) = K \circ D \circ H \circ W(\mathbf{V}, \mathbf{X}) \tag{5}$$

These steps allow the computation of the projection onto the image plane of a 3D point knowing its coordinates in the 3D space. In a 3D reconstruction framework, it is necessary to do the inverse operation, i.e. to compute the direction of the luminous ray corresponding to a pixel. This step consists in computing the coordinates of the point $\mathbf{X_S}$ belonging to the sphere given the coordinates $\begin{bmatrix} x & y \end{bmatrix}^{\mathrm{T}}$ of a 2D point on the normalized plane. This step of retro projection, also known as lifting, is achieved using formula (6).

$$\mathbf{X_S} = \begin{bmatrix} \dfrac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} x \\[3mm] \dfrac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} y \\[3mm] \dfrac{\xi + \sqrt{1 + (1 - \xi^2)(x^2 + y^2)}}{x^2 + y^2 + 1} - \xi \end{bmatrix} \tag{6}$$

3.2 Calibration

Calibration consists in the estimation of the parameters of the model which will be used for 3D reconstruction algorithms. Calibration is commonly achieved by observing a planar pattern at different positions. With the tool we have designed, the pattern can be freely moved (the motion does not need to be known) and the user only needs to select the four corners of the pattern. Our calibration process is similar to that of Mei (Mei & Rives, 2007). It consists of a minimization over all the model parameters of an error function between the estimated projection of the pattern corners and the measured projection using the Levenberg-Marquardt algorithm (Levenberg, 1944; Marquardt, 1963).

If n is the number of 3D points $\mathbf{X_i}$, $\mathbf{x_i}$ their projections in the images, we are looking for the parameter vector $\mathbf{V}$ which minimizes the cost function $E(\mathbf{V})$:

$$E(\mathbf{V}) = \frac{1}{2} \sum_{i=1}^{n} \left[P(\mathbf{V}, \mathbf{X_i}) - \mathbf{x_i} \right]^2 \qquad (7)$$

Our calibration tool was developed in C++ using the computer vision library OpenCV and can be freely downloaded from our website (Boutteau, 2009). It does not require any commercial software and a particular attention has been given to the optimization of the computation time since a calibration with 10 images does not exceed 2 minutes.

3.3 Relative pose estimation

The estimation of the intrinsic parameters presented in the previous section allows to establish the relationship between 3D points and theirs projections for each sensor of the stereoscopic system. To obtain metric information from the scene, for example by triangulation, the relative pose of the two sensors has to be known.

This step is generally performed by a pixel matching between both images followed by the estimation of the essential matrix. This matrix, originally introduced by Longuet-Higgins (Longuet-Higgins, 1981), has the property to contain information on the epipolar geometry of the sensor. It is then possible to decompose this matrix into a rotation matrix and a translation vector, but the last one can only be determined up to a scale factor (Bunschoten & Kröse, 2003). The geometrical structure of the scene can consequently be recovered only up to this scale factor.

Although in some applications, especially for 3D visualization, the scale factor is not needed, it is required for preparation of intervention or for navigation. To accomplish these tasks, the size of the objects and their distance from the robot must be determined. The 3D reconstruction has therefore to be Euclidean.

Thus, we suggest in this section a method to estimate the relative pose of the two sensors, with a particular attention to the estimation of the scale factor. The estimation of the relative pose of two vision sensors requires a partial knowledge of the environment to determine the scale factor. For this reason, we propose a method based on the use of a calibration pattern whose dimensions are known and which must be visible simultaneously by both sensors.

Let $(\mathbf{C_1}, \mathbf{x_1}, \mathbf{y_1}, \mathbf{z_1})$ and $(\mathbf{C_2}, \mathbf{x_2}, \mathbf{y_2}, \mathbf{z_2})$ be the frames associated with the two sensors of the stereoscopic system, and $\mathbf{M}$ be a 3D point, as shown in Figure 5.
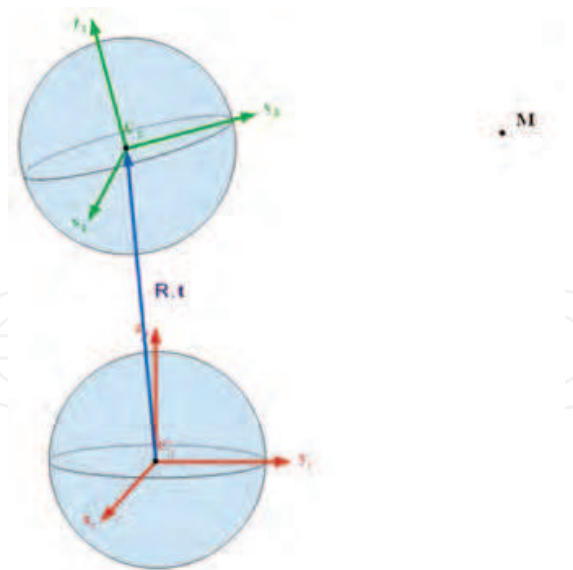
Fig. 5. Relative pose estimation principle.

The point $\mathbf{M}$ with coordinates $\begin{bmatrix} x_2 & y_2 & z_2 & 1 \end{bmatrix}^{\mathrm{T}}$ in the frame associated with the second sensor has for coordinates in the first sensor frame:

$$
\begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_2 \\ y_2 \\ z_2 \\ 1 \end{bmatrix}
\tag{8}
$$

where $\mathbf{t} = \begin{bmatrix} t_x & t_y & t_z \end{bmatrix}^{\mathrm{T}}$ and $\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$ correspond to the pose of the second sensor with respect to the first one. With n control points, equation (8) yields to the following system:

$$
\begin{bmatrix}
x_1^1 & y_1^1 & z_1^1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & x_1^1 & y_1^1 & z_1^1 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & x_1^1 & y_1^1 & z_1^1 & 0 & 0 & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
x_1^n & y_1^n & z_1^n & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & x_1^n & y_1^n & z_1^n & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & x_1^n & y_1^n & z_1^n & 0 & 0 & 1
\end{bmatrix}
\cdot
\begin{bmatrix}
r_{11} \\ r_{12} \\ r_{13} \\ r_{21} \\ r_{22} \\ r_{23} \\ r_{31} \\ r_{32} \\ r_{33} \\ t_x \\ t_y \\ t_z
\end{bmatrix}
=
\begin{bmatrix}
x_2^1 \\ y_2^1 \\ z_2^1 \\ \vdots \\ x_2^n \\ y_2^n \\ z_2^n
\end{bmatrix}
\tag{9}
$$

The principle of the relative pose estimation consists therefore in computing the rigid transformation between the two sensors, knowing the coordinates of a set of points in the two frames. The coordinates of the points are however not known directly since the location of the pattern is not known. The pose of the pattern has to be determined in the two frames using the Levenberg-Marquardt algorithm to minimize the error between estimated projection and projections extracted from the images. The resolution of equation (9) then gives the relative pose $(\mathbf{R}, \mathbf{t})$ of the two sensors including the scale factor.

4. Simultaneous Localization and Mapping

3D reconstruction of a scene using the motion of a sensor addresses two problems: localization and mapping. Localization consists in estimating the trajectory of a robot in a known map (Thrun et al., 2001). The aim of mapping is the creation of a map of the environment using measurements from the sensors embedded on a robot knowing its trajectory (Thrun, 2003). When neither the trajectory of the robot, nor the map of the environment are known, localization and mapping problems have to be considered simultaneously: it is the issue of SLAM (Simultaneous Localization And Mapping).

The first approach to solve the SLAM problem is to assume the motion known (by odometry or by the command law) even if this one is corrupted by noise. The position of visual landmarks can consequently be predicted. Sensors embedded on the mobile robot, for example laser range-finders, provide measurements of its environment. These observations are then used to update the model containing the coordinates of the visual landmarks and the positions of the robot. These steps (prediction/observation/update) are implemented using the Kalman filter or one of its derivatives.

The second approach is to optimize the geometrical structure of the scene and the positions of the sensor using the bundle adjustment method. A synthesis on bundle adjustment algorithms was published by Triggs (Triggs et al., 1999). Bundle adjustment provides more accurate results than Kalman filters (Mouragnon et al., 2009) but needs more computing time. In most of the applications, this algorithm is used off-line to obtain a very accurate model, but it is also possible to apply it iteratively. Although bundle adjustment is commonly used with conventional cameras, there are very few works on its adaptation to omnidirectional sensors. The main works in this field are those of Lhuillier (Lhuillier, 2005) and Mouragnon (Mouragnon et al., 2009) who suggest to find the structure of the scene and the motion of a catadioptric sensor by a local bundle adjustment followed by a global one to obtain more accurate results. Their works highlight the difficulty of estimating the scale factor, although it is theoretically possible with a non-central sensor.

The contribution of this section deals with the application of a bundle adjustment algorithm to an omnidirectional stereoscopic sensor previously calibrated to solve the ambiguity on the scale factor. Bundle adjustment relies on the non-linear minimization of a criterion, so a first estimation of the parameters as to be found to ensure the convergence. Before the presentation of the bundle adjustment algorithm, we thus expose our initialization step.

4.1 Initialization

Estimating the camera motion, also called ego-motion, requires to relate the images of the sequence grabbed during the motion. Relating images consists in localizing in the images the projections of a same 3D point of the scene. This step is decisive because the precision of

the motion relies on the quality of this matching. The visual landmarks are detected in a majority of works by the Harris corner detector since it has good performances on luminosity change. This detector is however very sensitive to scale change and it can thus fail with large motion. To avoid this, Lowe (Lowe, 2004) has presented a very interesting approach for the detection and the matching of regions of interest: the Scale Invariant Feature Transform (SIFT). The SIFT principle is to detect features from images which are invariant to scale change, rotation, and small point of view change. A descriptor, which corresponds to the orientation histogram, is then associated to the features and the matching can be achieved by comparison of their Euclidean distances.

Once the images are related, the epipolar geometry can be estimated between the two times $k-1$ and k. The epipolar geometry is interesting since it gives information on the relative pose of two vision sensors. Several works are dedicated to the estimation of the epipolar geometry for catadioptric sensors. Some of them (Pajdla et al., 2001; Gonzalez-Barbosa & Lacroix, 2005; Mariottini & Prattichizzo, 2005) give analytical solution to the estimation of the epipolar curves. Their methods need nevertheless to introduce the mirror equations and the proposed solutions are thus specific to the kind of sensor used.

Other works (Bunschoten & Kröse, 2003; Negishi, 2004) rely on the use of panoramic images, i.e. unwrapped images, and consider the epipolar curves as the intersection of the epipolar plane with a cylinder representing the image plane. This approach is interesting because the idea is to generalize the notion of epipolar geometry to panoramic sensors.

We suggest to generalize the epipolar geometry for all central sensors using the model of the equivalence sphere. With this model, the coplanarity constraint initially defined for perspective cameras (Longuet-Higgins, 1981) can be transposed to all central sensors. As shown in Figure 6, if the points $\mathbf{X_{S1}}$ and $\mathbf{X_{S2}}$ correspond to the projection of the same 3D point $\mathbf{X}$ onto the two spheres, then $\mathbf{C_1}$, $\mathbf{C_2}$, $\mathbf{X_{S1}}$, $\mathbf{X_{S2}}$ and $\mathbf{X}$ lie in the same plane.
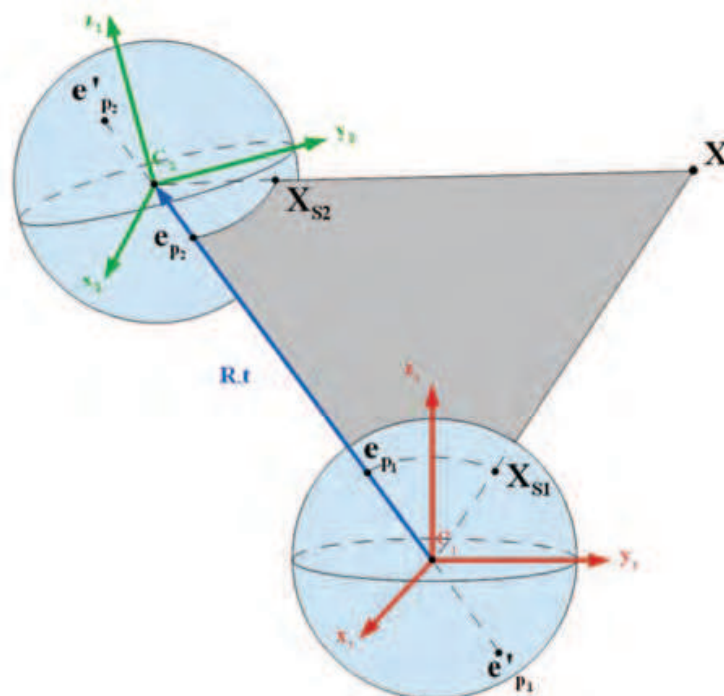

Fig. 6. Epipolar geometry for central sensors.

Let $\mathbf{t}$ be the translation vector of the sensor and $\mathbf{R}$ be the rotation matrix, the coplanarity constraint can be expressed in the coordinate system $(\mathbf{C_2}, \mathbf{x_2}, \mathbf{y_2}, \mathbf{z_2})$ as:

$$\mathbf{X_{S2}^T R(t \times X_{S1})} = 0 \tag{10}$$

The coplanarity constraint (10) can be rewritten in the matrix form as:

$$\mathbf{X_{S2}^T E X_{S1}} = 0 \tag{11}$$

where $\mathbf{E} = \mathbf{RS}$ is the essential matrix first introduced by Longuet-Higgins (Longuet-Higgins, 1981) and $\mathbf{S}$ is an antisymmetric matrix characterizing the translation:

$$\mathbf{S} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \tag{12}$$

The essential matrix $\mathbf{E}$ can be estimated from a set of matched points using the eight-point algorithm (Longuet-Higgins, 1981). Given two lifted points $\mathbf{X_{S1}} = \begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix}^T$ and $\mathbf{X_{S2}} = \begin{bmatrix} x_2 & y_2 & z_2 \end{bmatrix}^T$ corresponding to the same 3D point $\mathbf{X}$, (11) becomes for each pair of matched points:

$$x_2 x_1 e_{11} + x_2 y_1 e_{12} + x_2 z_1 e_{31} + \cdots + z_2 z_1 e_{33} = 0 \tag{13}$$

where $\mathbf{E} = \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}$.

Introducing the vector $e = \begin{bmatrix} e_{11} & e_{12} & e_{13} & e_{21} & e_{22} & e_{23} & e_{31} & e_{32} & e_{33} \end{bmatrix}^T$ and using n pairs of matched points, the set of equations (13) can be expressed in the matrix form as:

$$\mathbf{Ae} = 0 \tag{14}$$

With more than eight points, a least squares solution can be found by singular value decomposition (SVD) of $\mathbf{A}$. Because of the least squares estimation, the estimated matrix may not respect the two constraints of an essential matrix: two of its singular values have to be equal, and the third has to be zero (Hartley & Zisserman, 2004). A constraint enforcement step is therefore necessary and is achieved by the method described by Hartley (Hartley & Zisserman, 2004). Given a 3x3 matrix $\mathbf{E} = \mathbf{UDV}^T$, where $\mathbf{D} = diag(a, b, c)$ with $a \geq b \geq c$. The closest essential matrix to $\mathbf{E}$ in Frobenius norm is given by:

$$\hat{\mathbf{E}} = \mathbf{U\hat{D}V}^T \tag{15}$$

where $\hat{\mathbf{D}} = diag((a+b)/2, (a+b)/2, 0)$.

The estimation of the essential matrix may fail when there is one outlier among the set of points used in equation (14). The solution is therefore found by using the RANSAC algorithm. The main difference between perspective cameras and omnidirectional sensors comes from the computation of the error to determine if a point belongs to the consensus set. In the perspective case, the error is defined as the distance d between the point $\mathbf{X_2}$ and the epipolar line $\mathbf{l}$ corresponding to the point $\mathbf{X_1}$ as shown in Figure 7. In the omnidirectional case, the computation of the error is more difficult since the epipolar line becomes an epipolar curve (see Figure 8). We therefore suggest to work on the equivalence sphere using the coplanarity constraint defined by equation (11) to compute this error. Given a point $\mathbf{X_{S1}}$ on the first sphere, the normal of the epipolar plane in the second sensor frame is given by $\mathbf{N_2} = \mathbf{E}\mathbf{X_{S1}}$. The vectors $\mathbf{X_{S2}}$ and $\mathbf{N_2}$ are orthogonal when the coplanarity constraint is respected. The error e can consequently be computed as the angular error given by:

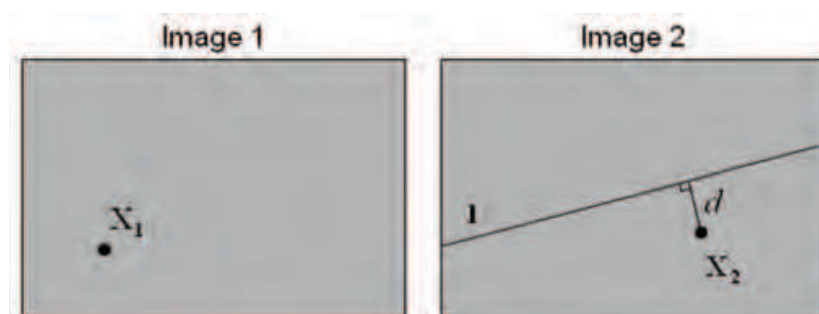$$e = \mathbf{X_{S2}^{T}}\mathbf{N_2} \tag{16}$$

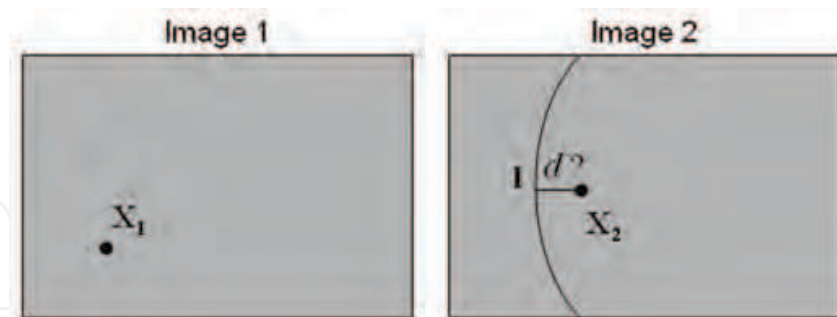Fig. 7. Computation of the error in the perspective case.

Fig. 8. Computation of the error in the omnidirectional case.

The essential matrix does not give directly the displacement of the sensor since it is a combination of the rotation matrix $\mathbf{R}$ and the antisymmetric matrix $\mathbf{S}$ (see equation 12). The essential matrix has therefore to be decomposed to retrieve the rotation $\mathbf{R}$ and the translation $\mathbf{t}$. After the constraint enforcement step, the essential matrix can be written as $\mathbf{E} = \mathbf{U}\text{diag}(1,1,0)\mathbf{V}^{\mathrm{T}}$. Let $\mathbf{u_i}$ be the ith column of $\mathbf{U}$ and $\mathbf{v_i}$ the ith column of $\mathbf{V}$, there are four possible solutions for the rotation $\mathbf{R}$ and two solutions for the translation vector $\mathbf{t}$ (Wang & Tsui, 2000):

$$
\begin{cases}
\mathbf{R_1} = (-\mathbf{u_2}, \mathbf{u_1}, \mathbf{u_3})\mathbf{V}^{\mathrm{T}} \\
\mathbf{R_2} = (-\mathbf{u_2}, \mathbf{u_1}, -\mathbf{u_3})\mathbf{V}^{\mathrm{T}} \\
\mathbf{R_3} = (\mathbf{u_2}, -\mathbf{u_1}, \mathbf{u_3})\mathbf{V}^{\mathrm{T}} \\
\mathbf{R_4} = (\mathbf{u_2}, -\mathbf{u_1}, -\mathbf{u_3})\mathbf{V}^{\mathrm{T}} \\
\mathbf{t_1} = \mathbf{v_3} \\
\mathbf{t_2} = -\mathbf{v_3}
\end{cases}
\tag{17}
$$

Two solutions for the rotation matrix can be easily eliminated because they do not respect the property of a rotation matrix $\det(\mathbf{R}) = 1$. It remains consequently four possible combinations of rotation and translation. In the perspective case, the right solution is trivial since it can be obtained after the triangulation of a point by checking the reconstructed point is in front of both cameras. This notion of "in front of the cameras" does not exist anymore in the omnidirectional case since points from the entire 3D space can have a projection onto the image plane. The right solution could be recovered using other sensors, for example odometers, but it is preferable to use only visual data to be totally independent from the mobile robot. For each possible combination, the right solution is found by triangulating the points, and then reprojecting them onto the image plane. The computation of the reprojection error allows the determination of the correct solution since the others give totally aberrant projections. The right solution is the one which gives the minimal reprojection error since the others are totally aberrant as shown in Figure 9.
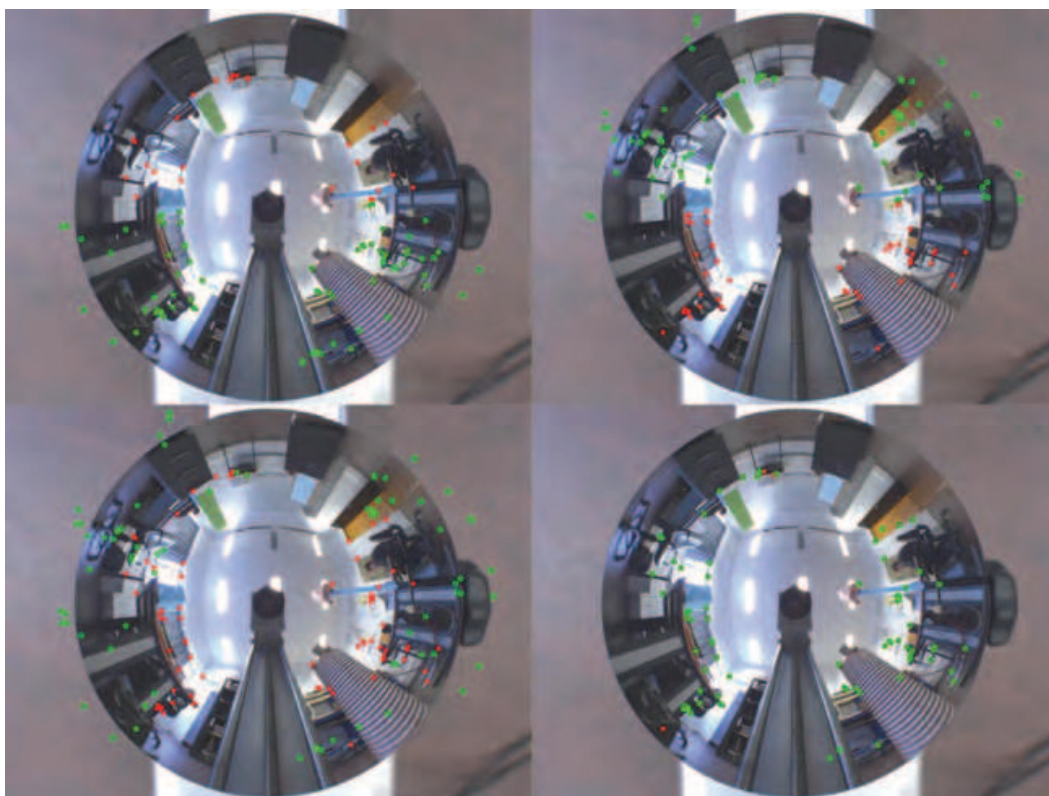

Fig. 9. Test of the four possible solutions. Green points are the reprojected points and red points are the real 2D projections.

The estimation of the essential matrix, followed by its decomposition, enables to retrieve the motion of the system, but only up to a scale factor. In a large majority of works, this ambiguity is removed by the use of odometry data (Bunschoten & Kröse, 2003). This approach is interesting since it gives directly the result, but it is preferable to use only visual data to avoid any communication with the mobile robot. The stereoscopic structure of our system can be used to retrieve the scale factor since the baseline is known. The 3D coordinates $\mathbf{X_{k-1}}$ and $\mathbf{X_k}$ of a set of points can actually be computed at times $k-1$ and k and the scale factor can be retrieved by the computation of the norm of $\mathbf{t}$, given by:

$$\mathbf{t} = \mathbf{X_k} - \mathbf{R}\mathbf{X_{k-1}} \tag{18}$$

4.2 Bundle Adjustment

The algorithms presented in the previous section provide initial values for the motion of the sensor and for the 3D coordinates of points. These values are however not sufficiently accurate to be used directly. An excessive error will provide, because of its accumulation, a wrong 3D model. An extra optimisation step, the bundle adjustment, is therefore necessary. This method is well known for perspective cameras (Triggs et al., 2000) but only few works are dedicated to omnidirectional sensors. In this section, we propose a bundle adjustment algorithm which takes into account the specificity of our sensor: its omnidirectional field of view and its stereoscopic structure.

Bundle adjustment consists in minimizing the cost function corresponding to the error between estimated projections of 3D points and their measured projections. Let m be the number of positions of the stereoscopic system and n be the number of 3D points, the cost function can be written as:

$$F(\mathbf{V}, \mathbf{X}) = \frac{1}{2} \sum_{j=1}^{m} \sum_{i=1}^{n} \left[P(\mathbf{V}_j, \mathbf{X}_i) - \mathbf{x}_{ij} \right]^2 \tag{19}$$

where $\mathbf{V}_j$ is the parameter vector of the jth camera, $\mathbf{X}_i$ is the coordinates of the ith 3D point in the world frame, $\mathbf{x}_{ij}$ is the projection of the ith 3D point in the image of the jth camera and $P(\mathbf{V}_j, \mathbf{X}_i)$ is the predicted projection of point i in image j.

Using a stereoscopic sensor allows to recover the scale factor, which is impossible with a monocular sensor. The projection function defined by equation (5) has therefore to be modified to take into account the stereoscopic structure of the sensor. Thus, the projection function provides now four elements: the coordinates u_{low} and v_{low} of the projection of $\mathbf{X}$ onto the image plane of the first sensor, and the coordinates u_{high} and v_{high} of its projection onto the image plane of the second sensor. The system is no longer considered as two separate sensors but as a global device. The relative pose of the two sensors $(\mathbf{R}, \mathbf{t})$ is consequently added to the model parameters as shown in Figure 10.
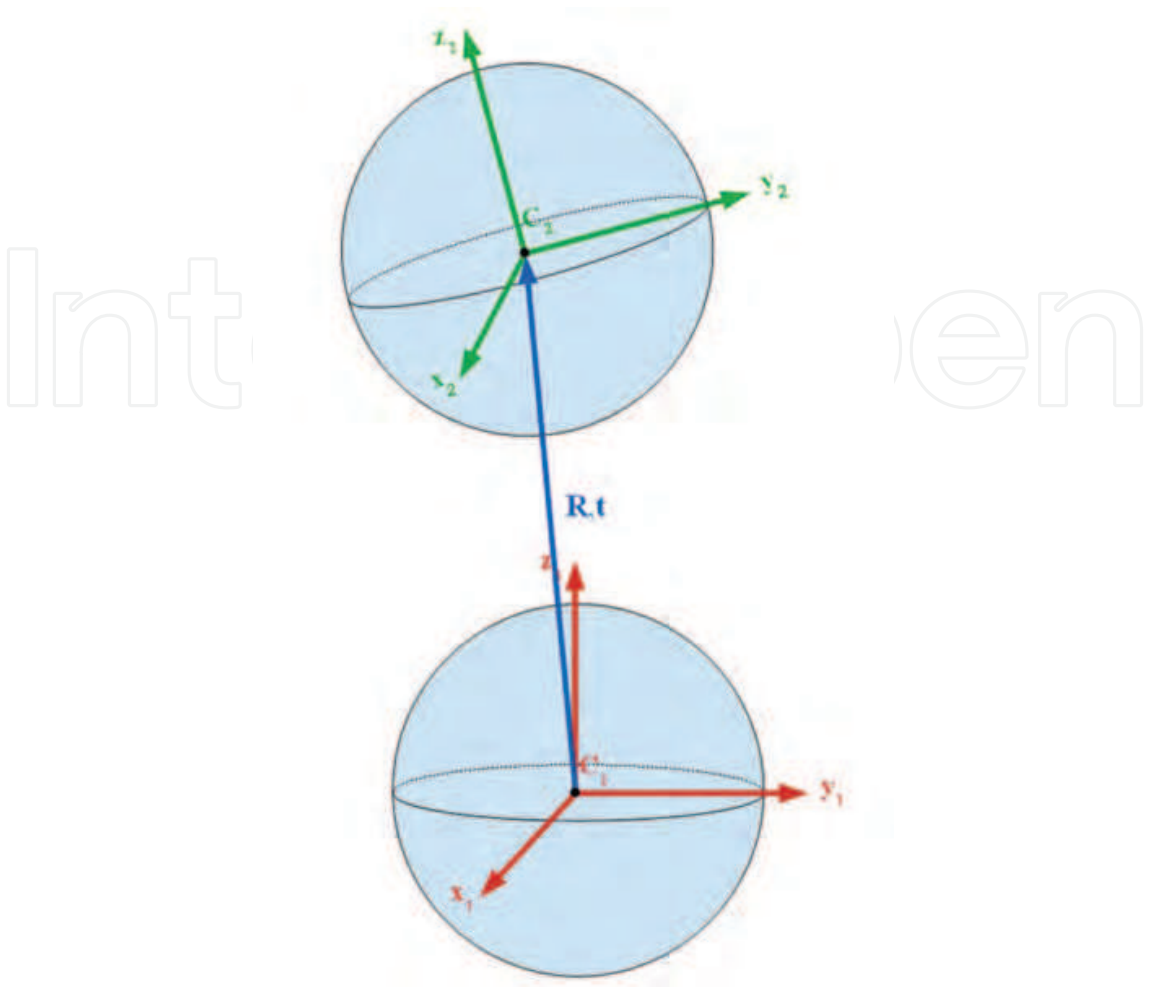
Fig. 10. Rigid transformation between the two sensors of the stereoscopic system.

Let $\mathbf{V_5} = \begin{bmatrix} q_{w12} & q_{x12} & q_{y12} & q_{z12} & t_{x12} & t_{y12} & t_{z12} \end{bmatrix}^{\mathrm{T}}$ be the new parameter vector, the rotation being parameterised by a quaternion. An extra rigid transformation, written $C(\mathbf{V_5})$, is added to the initial projection function to express coordinates either in the low sensor frame or in the high sensor frame:

$$P(\mathbf{V},\mathbf{X}) = K \circ D \circ H \circ C \circ W(\mathbf{V},\mathbf{X}) \tag{20}$$

where $\mathbf{V} = \begin{bmatrix} \mathbf{V}_1^{\mathrm{T}} & \mathbf{V}_2^{\mathrm{T}} & \mathbf{V}_3^{\mathrm{T}} & \mathbf{V}_4^{\mathrm{T}} & \mathbf{V}_5^{\mathrm{T}} \end{bmatrix}^{\mathrm{T}}$

The minimization of the reprojection error defined by equation (19) is carried out by the Levenberg-Marquardt algorithm. The key step of this algorithm lies in the resolution of the augmented normal equation:

$$\left(\mathbf{J}^{\mathrm{T}}\mathbf{J} + \lambda\mathbf{I}\right)\boldsymbol{\Delta} = -\mathbf{J}^{\mathrm{T}}\mathbf{e} \tag{21}$$

where λ is a real number varying from iteration to iteration and $\mathbf{I}$ is the identity matrix.

The resolution of equation (21) requires the computation of the Jacobian $\mathbf{J}$ of the projection function. The sensor is calibrated so the parameters that have to be estimated are the poses $\mathbf{V}_1^j$ of the cameras and the coordinates $\mathbf{X}_i$ of the 3D points. Thus, the partial derivatives which have to be estimated are the derivative of the projection function whith respect to the pose of the system:

$$\frac{\partial \mathbf{P}}{\partial \mathbf{V}_1^j} \tag{22}$$

as well as the derivative of the projection function with respect to the coordinates of the points:

$$\frac{\partial \mathbf{P}}{\partial \mathbf{X}_i} \tag{23}$$

The form of the Jacobian matrix used in equation (21) is illustrated in Figure 11 for three poses of the sensors and four points.
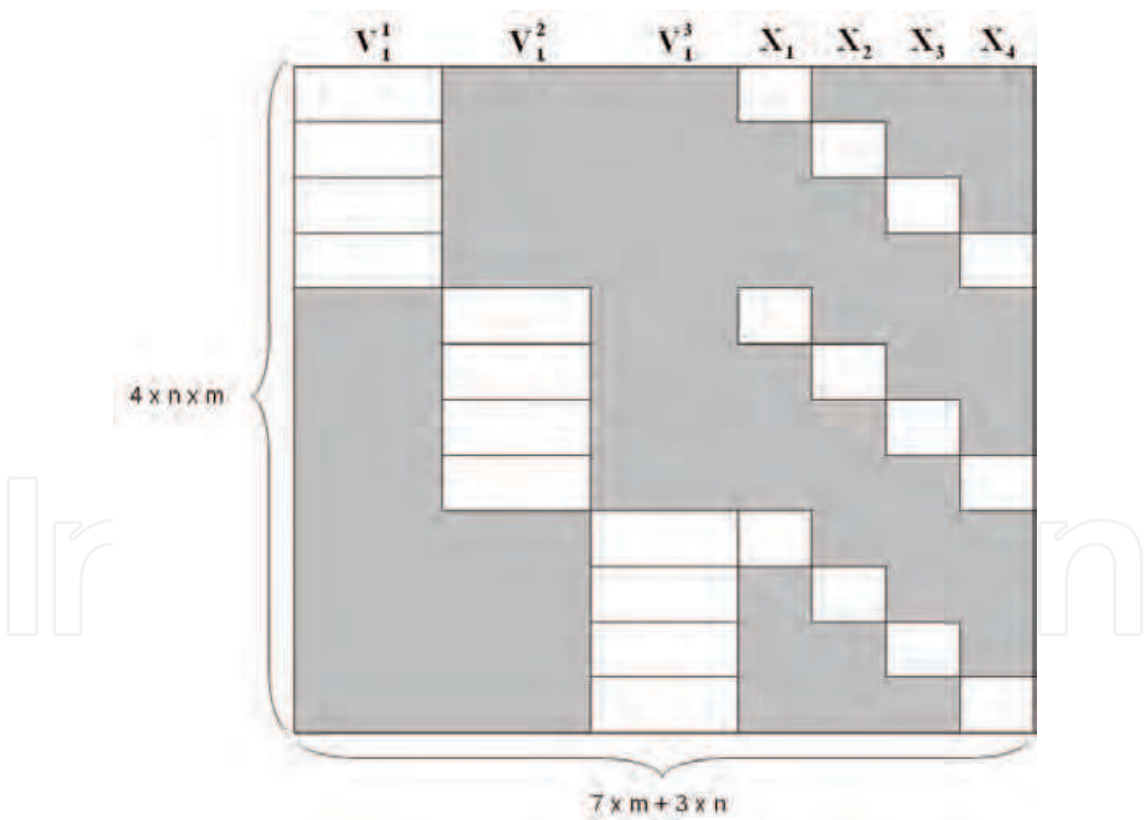
Fig. 11. Form of the Jacobian matrix for 3 poses and 4 points.

The iterative resolution of the augmented normal equation as described in (Levenberg, 1944; Marquardt, 1963) minimizes the reprojection error and provides a good estimate of the poses of the system and of the coordinates of the 3D points.

5. Experimental results

The stereoscopic omnidirectional sensor was mounted on a Pioneer robot with a Sick LD-PDS laser range-finder used to provide the ground truth. Each step, from calibration to 3D reconstruction and motion estimation, was evaluated on real images and without prior knowledge to evaluate the system in real conditions. Some results obtained with synthetic images are also presented to validate results that need a perfect knowledge of the 3D points coordinates in the sensor frame.

5.1 Calibration of the sensor

The calibration step was evaluated by computing the Root of Mean Squares (RMS) distances between estimated and theoretical projections of a set of 3D points. As it is very difficult to know precisely the coordinates of a 3D point in the sensor frame, we used synthetic images. The sensor was simulated in POV-Ray, a ray-tracing software, to generate omnidirectional images containing a calibration pattern as shown in Figure 12 and these images were used to calibrate the sensor.

Fig. 12. Synthetic omnidirectional images used for the calibration

As the projection model and the 3D coordinates of reference points are perfectly known in this case, their theoretical projection can easily be computed and compared to the projection obtained by the estimated model. Table 1 shows the mean error and the standard deviation obtained on a set of 150 points.

Mean error (pixels)	0.24
Standard deviation (pixels)	0.11

Table 1. Calibration results on synthetic images

The calibration was then evaluated on real images. Two sets of omnidirectional images containing a calibration pattern were taken. We used the first set to calibrate the sensor as described in section. The second set was used to compute the error between estimated projections of the grids points computed using estimated model and their measured projections extracted from the images. Table 2 summarizes the results obtained on this set of real images.

Mean error (pixels)	0.44
Standard deviation (pixels)	0.26

Table 2. Calibration results on real images

The error on real images is greater than the error on synthetic images, but both are very low since they are less than half a pixel. These results highlight the good estimation of the model

parameters even with noisy data since we used a real system which is not perfect. This can be explained by the introduction of the distortions into the model which allows some misalignments between the mirror and the camera.

5.2 Relative pose estimation

The estimation of the relative pose was evaluated on synthetic images for a vertical setup of the two sensors, i.e. the same setup as the real system. Several pairs of images containing a calibration were taken as shown in Figure 13 and the relative pose was evaluated with the method presented in section 3.3. The distance between the two sensors was defined with 10cm and they are supposed perfectly aligned. The result of the relative pose estimation is presented in Table 3.
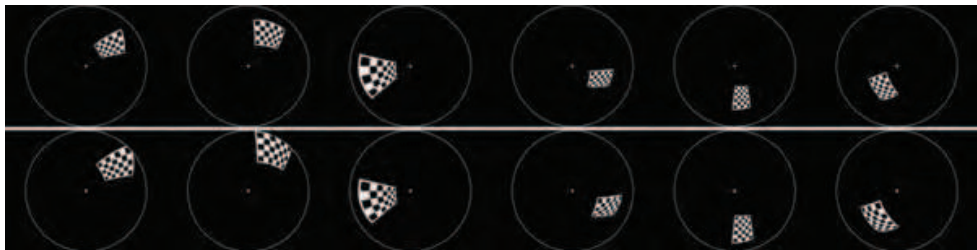
Fig. 13. Pairs of images used for the estimation of the relative pose

Rotation	$\mathbf{R} = \begin{bmatrix} 1.0000 & -0.0001 & 0.0006 \\ 0.0001 & 1.0000 & 0.0005 \\ -0.0006 & -0.0005 & 1.0000 \end{bmatrix}$
Translation	$\mathbf{T} = \begin{bmatrix} -0.413 \\ 0.142 \\ 99.948 \end{bmatrix}$

Table 3. Results of the estimation of the relative pose on synthetic images

The rotation matrix is almost an identity matrix since the sensors are perfectly aligned. The norm of the translation vector is 99.95mm instead of 100mm, so an error of 0.05%. These results are particularly good especially since they involve the calibration results needed for the estimation of the pose of the calibration pattern.

The estimation of the relative pose was then evaluated on real images. The sensor was mounted on a graduated rail and was moved by steps of 10cm. At each position, an omnidirectional image was grabbed with the aim of computing the displacement of the sensor according to the first position using five calibration patterns placed in the room as shown in Figure 14. Table 4 summarizes the results.

Displacement (mm)	100	200	300	400	500	600
Estimation (mm)	100.13	200.77	296.48	393.66	494.70	594.60

Table 4. Calibration results on real images

The average error is greater than the one obtained with synthetic images but it is less than 0.9%. This increase can be explained by the noise in real images but also by the accumulation of the errors because the estimation of the relative pose involves the calibration values.

Once the relative pose is estimated, the essential matrix can be computed and used to check the epipolar geometry properties. In Figure 14, for each pixel on the left image (red crosses), the corresponding epipolar curve (green curves) is drawn on the right image and vice versa.
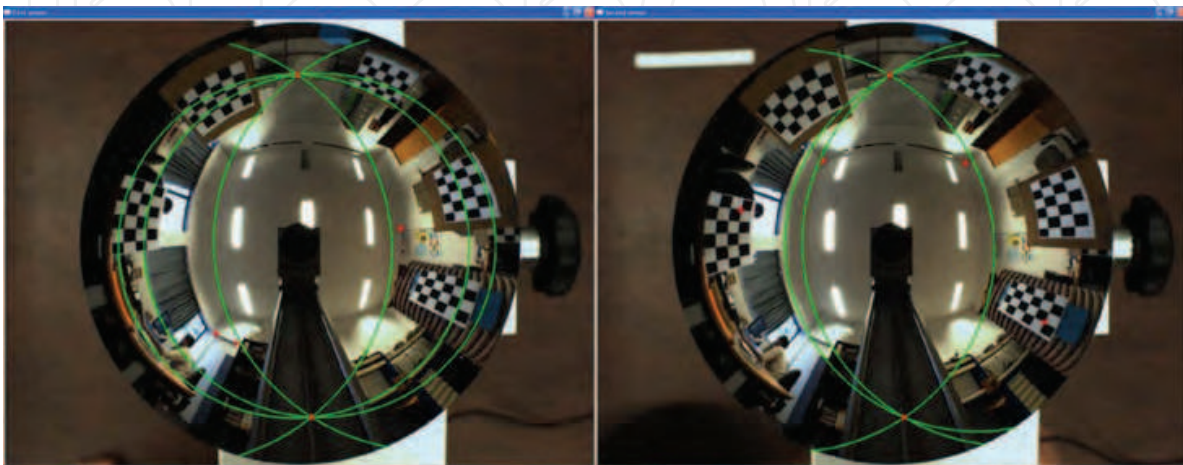
Fig. 14.Epipolar curves (green) corresponding to selected pixels (red crosses)

5.3 Triangulation

The combination of the calibration of the two sensors and of their relative pose estimation was evaluated by triangulation. The coordinates of the 3D points of five grids disposed on three walls of a room were estimated. Images used for this step are presented in Figure 15.
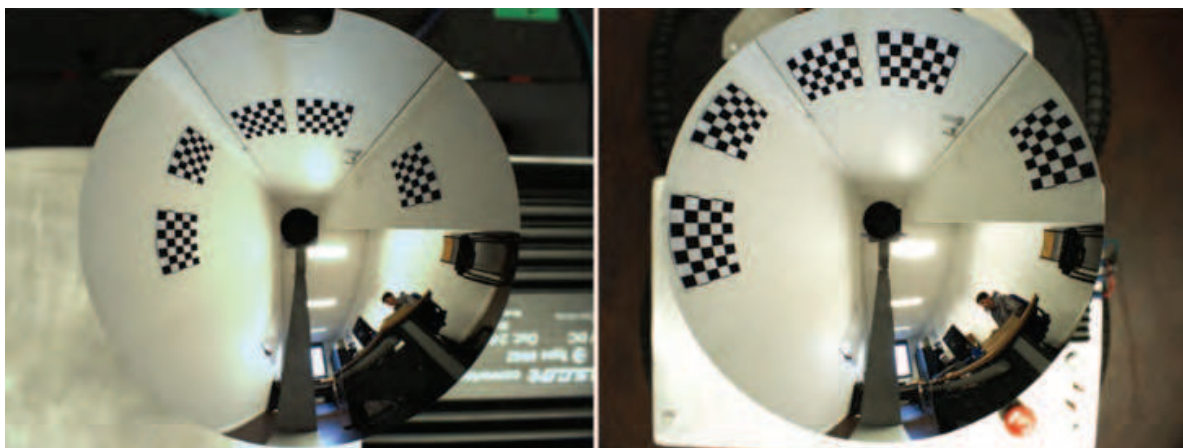
Fig. 15. Stereo pair used for the 3D reconstruction experimentation.

For each grid, the four corners were selected manually on the two images and an automatic corners detector was applied to extract all grid points. The 3D position of these points was evaluated by triangulation and is displayed in Figure 16. The position of the points is compared to the ground truth obtained by the laser range-finder. A linear regression was performed on raw laser data (in blue) to obtain the position of the walls (red lines).
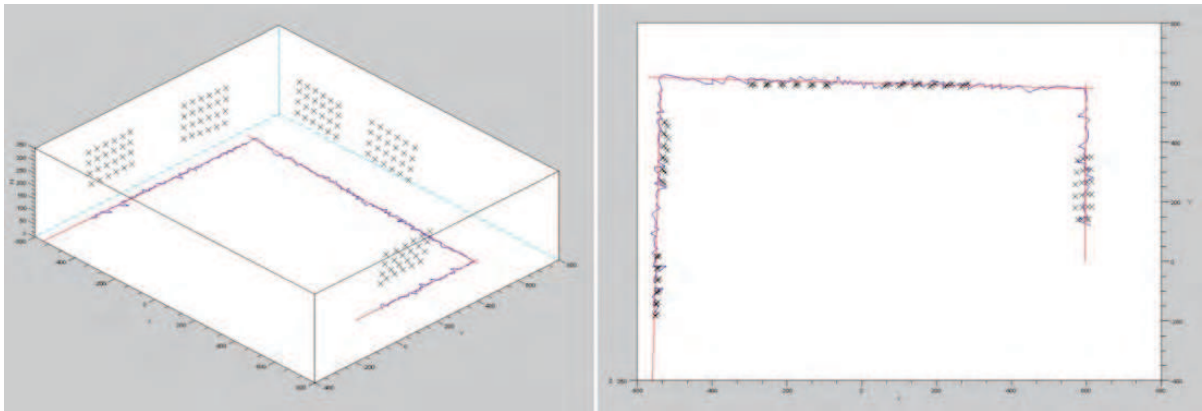
Fig. 16. Position of the five grids (crosses) compared to ground truth (blue and red lines).

For each grid, the mean error and the standard deviation of the position of its points were computed and summarized in Table 5.

Grid number	1	2	3	4	5
Mean error (mm)	5.843	15.529	13.182	3.794	12.872
Standard deviation (mm)	2.616	4.500	2.420	2.504	7.485

Table 5. Triangulation results.

The 3D error is very low, around 1% at 1 meter. The good results obtained in triangulation imply a reliable estimation of the sensor parameters and of the relative pose. Due to the resolution of catadioptric sensors, this error will nevertheless increase with the distance and will be around 10% at 10 meters.

5.4 Simultaneous Localization And Mapping

Our SLAM algorithm was evaluated on a real video sequence. The mobile robot was moved along the front of a building on a distance of 30 meters. The environment used is particularly rich since there are both manmade and natural items as shown in Figure 17.

At each position of the robot, a pair of omnidirectional images was grabbed and a map of the environment was acquired by the laser range-finder. These maps are then used to compute the real trajectory (ground truth) of the robot since this kind of sensor is very accurate.
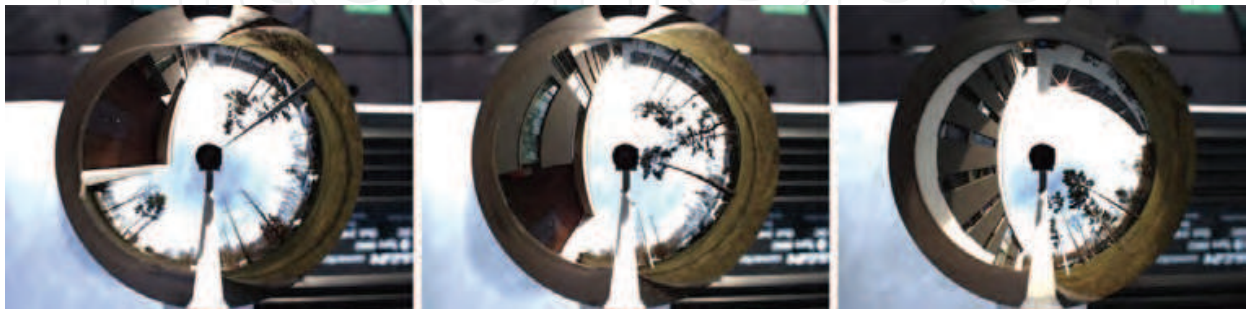

Fig. 17. Some of the pictures of the video sequence used for the Simultaneous Localization and Mapping.

The knowledge of the poses of the robot allows merging the local maps to obtain the global map of the environment, as shown in Figure 18. Thus, a comparison between the estimations provided by the laser range-finder, by the SLAM algorithm and by odometry can be achieved by observation of the global map, as shown in Figure 18.
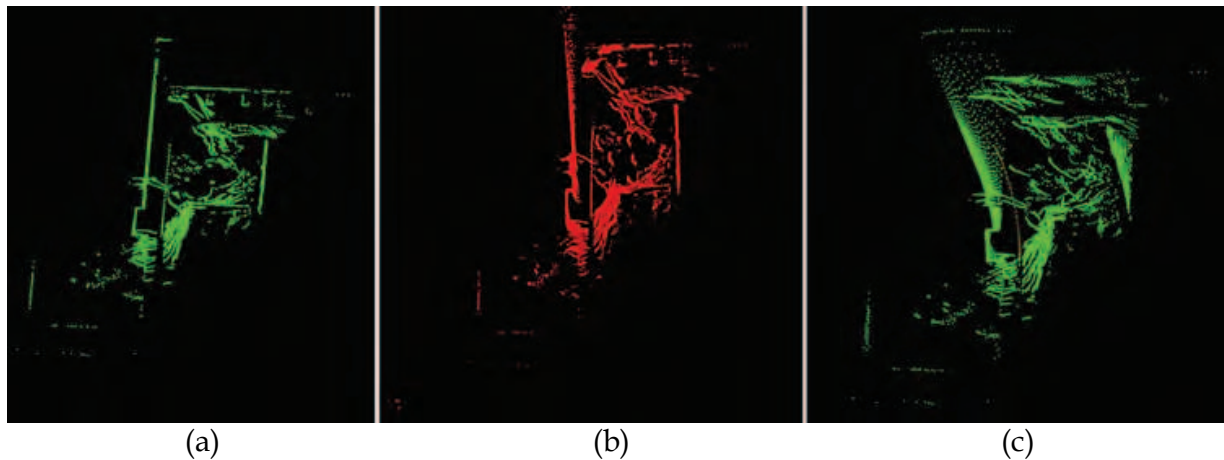

| (a) | (b) | (c) |

Fig. 18. Global map of the environment using: (a) laser data, (b) SLAM algorithm, (c) odometry data

As shown in Figure 18, the global map of the environment obtained by the laser range-finder is the best result since local maps are perfectly coincident. Our SLAM algorithm provides a slight deviation but the global map is far better than the map that can be obtained by odometry. The mean error on this sequence is around 12%.

6. Conclusion and future work

This paper has been devoted to the design of an innovative vision sensor dedicated to mobile robot application. Combining omnidirectional and stereoscopic vision offers many advantages for 3D reconstruction and navigation that are the two main tasks a robot has to achieve. In this article we have highlighted that the best stereoscopic configuration is the vertical one as it simplifies the pixel matching between images.
A special care has been put on the sensor calibration to make it flexible since it only requires the use of a planar calibration pattern. Experimental results highlight a high accuracy, which foreshadows good results for the following algorithms.
The fundamental issue of Simultaneous Localization And Mapping (SLAM) was then addressed. Our solution to this problem relies on a bundle adjustment between two displacements of the robot. The initialization of the motion and the coordinates of 3D points is a prerequisite since bundle adjustment is based on a non-linear minimization. This step is a tricky problem to which we answered by the generalization of the epipolar geometry for central sensors using the unified model. Our experimental results on SLAM are promising but the error is higher than the one expected further to the calibration results.
Our future work will focus on the improvement of our SLAM method by adding a global bundle adjustment to avoid error accumulation. A lot of challenges in SLAM are moreover always open, for instance SLAM based only on vision systems, SLAM taking into account six degrees of freedom, or SLAM for large-scale mapping.

7. References

Bailey, T. & Durrant-Whyte, H (2006). Simultaneous Localization and Mapping (SLAM): Part II. *Robotics and Automation Magazine,* Vol. 13, No. 3, (September 2006) 108-117.

Baker, S. & Nayar, S.K. (1999). A Theory of Single-Viewpoint Catadioptric Image Formation. *International Journal of Computer Vision (IJCV),* Vol. 35, No. 2, (November 1999) 175-196, ISSN 0920-5691

Barreto, J.P. (2006). A Unifying Geometric Representation for Central Projection Systems. *Computer Vision and Image Understanding,* Vol. 103, No. 3, (September 2006) 208-217, ISSN 1077-3142

Benosman, R. & Devars, J. (1998). Panoramic Stereovision Sensor, *Proceedings of the International Conference on Pattern Recognition (ICPR)*, pp. 767-769, Brisbane, Australia, August 1998, IEEE Computer Society, Los Alamitos

Boutteau, R. (2009). 3D Reconstruction for Autonomous Navigation, *http://omni3d.esigelec.fr/doku.php/thesis/r3d/start*

Bunschoten, R. & Kröse, B. (2003). Robust Scene Reconstruction from an Omnidirectional Vision System. *IEEE Transactions on Robotics and Automation,* Vol. 19, No. 2, (April 2003) 351-357, ISSN 1042-296X

Durrant-Whyte, H & Bailey, T. (2006). Simultaneous Localization and mapping: Part I. *Robotics and Automation Magazine,* Vol. 13, No. 2, (June 2006) 99-110, ISSN 1070-9932

Geyer, C. & Daniilidis, K. (2000). A Unifying Theory for Central Panoramic Systems and Practical Implications, *Proceedings of the European Conference on Computer Vision*, pp. 445-461, ISBN 3-540-67685-6, Dublin, Ireland, June 2000, Springer, Berlin

Gonzalez-Barbosa, J.J. & Lacroix, S. (2005). Fast Dense Panoramic Stereovision, *Proceedings of the International Conference on Robotics and Automation (ICRA)*, pp. 1210-1215, ISBN 0-7803-8914-X, Barcelona, Spain, April 2005, IEEE

Harley, R. & Zisserman, A. (2004). *Multiple View Geometry in Computer Vision,* Cambridge University Press, ISBN 0521540518

Levenberg, K. (1944). A method for the solution of certain problems in least squares. *Quarterly of Applied Mathematics,* Vol. 2, 164-168

Lhuillier, M. (2005). Automatic Structure and Motion using a Catadioptric Camera, *Proceedings of the 6th Workshop on Omnidirectional Vision, Camera Networks and Non-classical* Cameras, pp. 1-8, ISBN, Beijing, China, October 2005

Longuet-Higgins, H.C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature,* Vol. 293, (September 1981) 133-135

Lowe, D.G. (2004). A Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision (IJCV),* Vol. 60, No. 2, (November 2004) 91-110, ISSN 0920-5691

Mariottini, G.L. & Prattichizzo, D. (2005). The Epipolar Geometry Toolbox : multiple view geometry and visual servoing for MATLAB. *International IEEE Robotics and Automation Magazine,* Vol. 12, No. 4, (December 2005) 26-39, ISSN 1070-9932

Marquardt, D.W. (1963). An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied Mathematics,* Vol. 11, No. 2, 431-441

Mei, C. & Rives, P. (2007). Single View Point Omnidirectional Camera Calibration from Planar Grids, *Proceedings of the International Conference on Robotics and Automation (ICRA)*, pp. 3945-3950, ISBN 1-4244-0601-3, Roma, Italy, April 2007, IEEE

Mouaddib, E.M. (2005). Introduction à la vision panoramique catadioptrique. *Traitement du Signal,* Vol. 22, No. 5, (September 2005) 409-417, ISSN 0765-0019

Mouragnon, E.; Lhuillier, M.; Dhome, M.; Dekeyser, F.; Sayd, P. (2009). Generic and Real-Time Structure from Motion using Local Bundle Adjustment. *Image and Vision Computing Journal,* Vol. 27, No. 8, (July 2009) 1178-1193, ISSN 0262-8856

Negishi, Y.; Miura, J.; Shirai, Y. (2004). Calibration of Omnidirectional Stereo for Mobile Robots, *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS),* pp. 2600-2605, ISBN 0-7803-8463-6, Sendai, Japan, September 2004

Nister, D. (2001). Automatic Dense Reconstruction from Uncalibrated Video Sequence. *PhD Thesis, Royal Institute of Technology KTH, Stockholm, Sweden*

Pajdla, T.; Svoboda, T.; Hlavac, V. (2001). Epipolar Geometry of Central Panoramic Catadioptrics Cameras, In: *Panoramic Vision: Sensors, Theory and Applications,* (Gries, D. & Schneider, F.B.), (73-102), Springer, ISBN 978-0-387-95111-9, Berlin, Germany

Pollefeys, M; Gool, L.; Vergauwen, M; Verbiest, F.; Cornelis, K. ; Tops, J. (2004). Visual Modeling with a Hand-Held Camera. *International Journal of Computer Vision (IJCV),* Vol. 59, No. 3, (September-October 2004) 207-232, ISSN 0920-5691

Ramalingram, R.; Sturm, P.; Lodha, S.K. (2005). Towards Complete Generic Camera Calibration, *Proceedings of the International Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 767-769, ISBN 0-7695-2372-2, San Diego, USA, June 2005, IEEE Computer Society, Los Alamitos

Scaramuzza, D.; Martinelli, A.; Siegwart, R. (2006). A Flexible Technique for Accurate Omnidirectional Camera Calibration and Structure From Motion, *Proceedings of the International Conference on Computer Vision Systems (ICVS),* pp. 45-52, ISBN 0-7695-2506-7, New York, USA, January 2006, IEEE Computer Society, Los Alamitos

Thrun, S.; Fox, D.; Burgard, W.; Dellaert, F. (2001). Robust Monte Carlo Localization for Mobile Robots. *Artificial Intelligence,* Vol. 128, No. 1-2, (May 2001) 99-141

Thrun, S. (2003). Robotic mapping: A survey, In: *Exploring Artificial Intelligence in the New Millenium,* (Kakemeyer, G. & Nebel, B.), (1-29), Morgan Kaufmann, ISBN 1-55860-811-7, San Francisco, USA

Triggs, B.; McLauchlan, P.; Hartley, R.; Fitzgibbon, A. (1999). Bundle Adjustment – A modern Synthesis. *Vision Algorithms: Theory and Practice,* Vol. 1883, (September 1999) 298-372, ISSN 0302-9743

Wang, W. & Tsui, H.T. (2000). A SVD decomposition of essential matrix with eight solutions for the relative positions of two perspective cameras, *Proceedings of the International Conference on Pattern Recognition (ICPR),* pp. 362-365, ISBN, Barcelona, Spain, September 2000, IEEE Computer Society, Los Alamitos

Mobile Robots Navigation

Edited by Alejandra Barrera

ISBN 978-953-307-076-6

Hard cover, 666 pages

Publisher InTech

Published online 01, March, 2010

Published in print edition March, 2010

Mobile robots navigation includes different interrelated activities: (i) perception, as obtaining and interpreting sensory information; (ii) exploration, as the strategy that guides the robot to select the next direction to go; (iii) mapping, involving the construction of a spatial representation by using the sensory information perceived; (iv) localization, as the strategy to estimate the robot position within the spatial map; (v) path planning, as the strategy to find a path towards a goal location being optimal or not; and (vi) path execution, where motor actions are determined and adapted to environmental changes. The book addresses those activities by integrating results from the research work of several authors all over the world. Research cases are documented in 32 chapters organized within 7 categories next described.

INTECH
open science | open minds