Mobile Robotics Technique for 3D Environment Reconstruction

Annalisa Milella[1], Mario Foglia[1], and Giulio Reina[2]

Abstract— **Mobile robotics can be successfully applied for 3D environment exploration and reconstruction. Here, a vehicle tracking system is described based on combining laser and camera information aiming at building maps of unstructured environments. The proposed system integrates a ground-fixed CCD camera with a phase-shifting laser to follow the vehicle moving across the environment.**

The method comprises two stages. The first stage is a pre-calibration process, for the gross estimation of camera parameters using laser data. The second stage is the tracking of the robot for map building of the environment. During the second stage, the camera calibration is refined updating its parameters through a recursive optimization using real time measurements of the vehicle positions, which are derived by the laser, automatically driven by the camera in this stage.

The novelty of the proposed method is that it performs a real-time, accurate photogrammetric calibration without employing any calibration grid, thus preserving flexibility and allowing survey of large environments.

Detailed results obtained by the simulations and preliminary experiments are described, showing the effectiveness of the system in refining the calibration parameters and accurately reconstruct the environment.

Index Terms—**Mobile robotics, laser measurements, environment reconstruction, mapping.**

I. Introduction

ONE of the main purposes of 3D Computer Vision is to estimate the pose of moving objects with respect to a known reference frame, using one or more views acquired by the imaging device. In the field of autonomous vehicles navigation, this problem is referred to as vision-based vehicle positioning and is concerned with the determination of vehicle pose using the images acquired by either an on-board or a ground-fixed camera.

Although there are techniques inferring 3D information from uncalibrated cameras, a wide range of existing algorithms for 3D reconstruction and recognition requires the knowledge of the intrinsic and extrinsic parameters of the camera [3], [6]. Camera calibration remains, therefore, a fundamental problem in Computer Vision. In the last decades, many calibration methods have been proposed. Most of them, which are referred to as photogrammetric methods, are based on the perspective model of the camera and rely on the availability of a set of image points - world points correspondences. One of the main problems to solve for the application of such methods is finding a suitable set of matching points. Typically, this is done observing a 3D object of known geometry, characterized by features, such as black squares on a white back-ground, that can be accurately detected in the observed scene [7], [8]. Although it guaranties high-precision results, this approach often requires an elaborate setup, especially for cameras surveying large rooms. Some solutions have been presented in the past, which use feature points inherent in the scene [5], or special portable patterns [1]. Other approaches, which are called self-calibration methods [4], do not employ any calibration object and rely on the correspondences between images acquired by the same device, with unvaried intrinsic factors, to recover both the internal and external parameters. Those methods are more flexible than the photogrammetric ones, but they do not always allow reliable results.

Here, an on-line semi-automatic method for accurate and flexible camera calibration is presented, aiming at developing an efficient tracking system for mobile robots and building maps of unstructured environments. The method integrates a ground-fixed CCD camera with a phase-shifting laser to calibrate the camera. Specifically, the method consists of two principal modules, the Direct Parameters Calibration module and the Collinear Points - based Vehicle Localization module. Calibration is performed in two phases: a pre-calibration phase and an automatic re-calibration phase.

The novelty of the proposed method is that it performs a real-time and precise photogrammetric calibration without employing any calibration grid, thus preserving flexibility. The method is particularly suitable for cameras surveying large rooms. At the same time, it provides an accurate system for vehicle localization and tracking, allowing map building of 3D environments. In this paper, the results of the simulation study are presented. Specifically, the rest of the paper is structured as follows. Section II introduces the basic principles of the method and gives a description of the algorithms developed. In section III, the simulation results are presented. Section IV contains experimental results.

[1] Annalisa Milella and Mario Foglia are with the Department of Mechanical Engineering, Politecnico of Bari, Viale Japigia 182, 70125 Bari, Italy (email: milella@poliba.it, mm.foglia@poliba.it).

[2] Giulio Reina is with the Department of Innovation Engineering, University of Lecce, Via per Monteroni, 73100 Lecce, Italy (corresponding author: +39 0832 297362; email: giulio.reina@unile.it).

II. Description of the Method

A vehicle can move in a large room, equipped with a target mounted vertically on the top, consisting of six red-colored spheres and a retro-reflective cylinder, disposed along the same axis.

The relative position between the spheres, the cylinder and the vehicle is supposed to be completely known. The environment of motion is surveyed by a ground-fixed camera and a phase-shifting laser mounted on a rotating platform. The camera acquires the images of the spheres, while the laser provides a measure of distance when intersecting the retro-reflective cylindrical target.

The method consists of two principal steps: pre-calibration and automatic re-calibration. In the pre-calibration phase, the vehicle is guided to reach at least four different positions across the space of motion. The camera and the laser are employed separately, respectively providing the pixel coordinates and the corresponding real world coordinates of the spheres centers. The pixel coordinates are obtained acquiring an image of the target with the camera and segmenting this image to isolate the spheres. The laser is manually guided to intersect the cylindrical retro-reflective target, thus recovering the position of the target in the real world.

Eventually, an initial number of correspondences (6 for each vehicle position) of image points-world points is available to perform a first estimate of the camera parameters. Calibration is accomplished using a Direct Parameters Calibration algorithm (DPC), mainly based on Tsai's calibration method [6].

The calibration is then refined through a successive automatic process, during which the information deriving from the camera is used to control the motion of the laser beam and contemporary reconstruct the environment geometry. In fact, once a first estimate of the camera parameters has been accomplished, the Collinear Points – based Vehicle Localization algorithm (CPVL) can be applied to localize the vehicle in its space of motion, using only the information deriving from the visual detection of the spherical targets.

In other terms, the camera is able to provide for a signal of control for the laser that is therefore automatically guided to point the vehicle and perform a new measure. That results in a new set of correspondences, which can be employed to recalibrate the system in a real time process. Once the camera parameters have been accurately determined, the system would function as a tracking system. It can be pointed out that, as it will be shown below, the CPVL algorithm is able to recover both the vehicle position and the vehicle angles of pitch and roll. Furthermore, the laser will be able to track the vehicle also in presence of variations of the vehicle quote, in a range compatible with the axial dimension of the cylindrical target.

In the following sections, the laser-based measuring system and the image segmentation algorithm used to recover the set of matching points are firstly presented. Then, the DPC and the CPVL modules are described.

A. Recovering the Set of Matching Points

For camera calibration, first of all, a set of matching points has to be recovered. To this aim, a ground-fixed camera and a rotating phase-shifting laser are used, that allow estimating respectively the image coordinates and the real world coordinates of the target described in the preceding section.

Four reference frames are fixed (see Fig. 1): the World Reference Frame (WRF) (W; X_W, Y_W, Z_W), the Laser Reference Frame (LRF) (L; X_l, Y_l, Z_l), the Camera Reference Frame (CRF) (C; X_c, Y_c, Z_c), and the Image Reference Frame (IRF) (o; x, y). The LRF has the Z_l axis and the origin L respectively coincident with the Z_W axis and the origin W of the WRF, and the X_l axis coincident with the direction of the laser beam. The CRF is the standard camera frame that is characterized by the Z_c axis coincident with the optical axis and the other two axes parallel to the image plane. The IRF is centered in the principal point o, which is defined as the intersection between the optical axis and the image plane.

Based on these reference frames, the following transformations are considered: the transformation that links the coordinates of a point P in the LRF to those of the same point in the WRF; the transformation from the coordinates of a point P in the WRF to those of the same point in the CRF and finally the relation that concerns the perspective

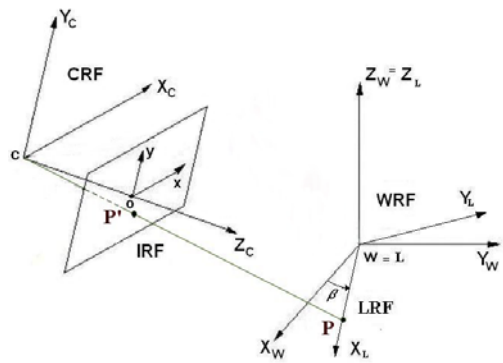

Fig. 1. Reference frames

projection of the point on the image plane.

Let us consider the first transformation. The laser performs a measure of the position of the target with respect to the LRF through a measure of distance. Knowing the orientation of the laser beam in the WRF, it is then possible to calculate the position of the target in the WRF, according to the following relations. If we denote with $[P]_L = \left[X_L^P, Y_L^P, Z_L^P\right]^T$ the coordinates of a point P in the LRF and with $[P]_W = \left[X_W^P, Y_W^P, Z_W^P\right]^T$ the coordinates of the same point in the WRF, the transformation from $[P]_L$ to $[P]_W$ can be expressed as:

$$[P]_W = R_{WL}[P]_L \tag{1}$$

where R_{WL} is the rotation matrix that represents the orientation of the LRF with respect to WRF .The coordinates of the detected retro-reflective target centroid C in the LRF

are $[C]_L = [d \quad 0 \quad 0]^T$ where d is the measure provided by the laser which takes into account the physical dimensions of the device. Knowing the angle of rotation β, from (1), the coordinates of the centroid of the retro-reflective cylinder detected by the laser can be recovered as $[C]_w = [d \cdot \cos\beta \quad d \cdot \sin\beta \quad 0]^T$. Finally, supposing that the vehicle is moving on a flat, regular surface, the coordinates of the spheres centers in the WRF, can be calculated.

The corresponding pixel coordinates are obtained from image segmentation, employing a shape-descriptor based algorithm and color information.

B. Direct Parameters Calibration (DPC)

Once the set of matching points has been obtained, a calibration algorithm can be implemented to estimate the intrinsic and extrinsic camera parameters. Here, a two-step direct parameters calibration method is used, mainly based on Tsai's algorithm [6]. Referring to the notation introduced in the preceding sections, the transformation from $[P]_W$ to $[P]_C = [X_C^P \quad Y_C^P \quad Z_C^P]^T$ can be expressed as:

$$[P]_C = T_{CW}[P]_W \tag{2}$$

where T_{CW} is the transformation matrix, which is constituted by the rotation matrix R_{CW} representing the orientation of the WRF with respect to CRF, and by the translation vector $t_{CW} = [X_C^w, Y_C^w, Z_C^w]^T$ that is a 3x1 vector whose elements are the coordinates of the origin W of the WRF in the CRF. Knowing the coordinates of a point in the CRF, we can calculate its projection on the image plane.

Assuming a pinhole model for the camera, and supposing that the CCD array is made up of a rectangular grid of photosensitive elements, we can express the coordinates of the point P with respect to a reference frame fixed on the image plane and centered in the principal point, in pixel units, as follows:

$$u = f_x \frac{X_C^P}{Z_C^P} \qquad v = f_y \frac{Y_C^P}{Z_C^P} \tag{3}$$

where fx and fy are the focal length in pixel units evaluated along x and y directions, respectively. It is useful to define here the aspect ratio s, i.e. the ratio between fx and fy. Substituting (2) in (3), the following relation is derived, which is an equation in eight unknowns:

$$\begin{aligned} &u \cdot X_W^P \cdot q_1 + u \cdot Y_W^P \cdot q_2 + u \cdot Z_W^P \cdot q_3 + \\ &u \cdot q_4 - v \cdot X_W^P \cdot q_5 - v \cdot Y_W^P \cdot q_6 - v \cdot Z_W^P \cdot q_7 - v \cdot q_8 = 0 \end{aligned} \tag{4}$$

Where

$$\left\{ \begin{array}{l} q_1 = r_{21}; q_2 = r_{22}; q_3 = r_{23}; \\ q_4 = Y_C^W; q_5 = s \cdot r_{11}; q_6 = s \cdot r_{12}; \\ q_7 = s \cdot r_{13}; q_8 = s \cdot X_C^W \end{array} \right\} \tag{5}$$

and r_{ij} are the components of the rotation matrix R_{CW}.

Writing (4) for N non coplanar points with $N \geq 7$, an homogeneous system of equations is obtained, that has a nontrivial solution up to a scale factor. This solution can be computed by the SVD decomposition of the coefficient matrix of the system. In this way, the aspect ratio, the rotation matrix and the first two components of the translation vector are determined. At this point, another system of equations can be stated to recover fx and the third component of the translation vector, as:

$$\begin{aligned} &u \cdot \left(r_{31}X_W^P + r_{32}Y_W^P + r_{33}Z_W^P + Z_C^W\right) = \\ &f_x \cdot \left(r_{11}X_W^P + r_{12}Y_W^P + r_{13}Z_W^P + X_C^W\right) \end{aligned} \tag{6}$$

For $N \geq 2$ points we obtain a system of N equations in two unknowns that can be solved in a least squares-sense with the method of the pseudo-inverse.

C. Collinear Points - Based Vehicle Localization (CPVL)

Once the perspective projection of the target is detected, the 3D position of the vehicle with respect to the CRF can be evaluated, based on the following algorithm [2]. The centers of the spheres can be seen as N collinear points, situated at known inter-point distances. Therefore, if we denote with $p = [p_1 \quad p_2 \quad p_3]^T$ the first distinguished point and with $b = [b_1 \quad b_2 \quad b_3]^T$ the unknown direction cosines in the CRF of the axis along which the spheres are aligned, we can express the position of the i-th point in the CRF as: $p_i = p + \delta_i b$, where δ_i is the known distance between the point p and the point p_i. Based on this relation and referring to the equations of the perspective projection (3), the position of the i-th point on the image plane with respect to the principal point is given by:

$$u_i = f_x \frac{p_1 + \delta_i b_1}{p_3 + \delta_i b_3} \qquad v_i = f_y \frac{p_2 + \delta_i b_2}{p_3 + \delta_i b_3} \tag{7}$$

for i=0,…, N-1, assuming $\delta_o = 0$. Writing (7) for N points, 2N equations plus the constant equation bb'=1, in six unknowns are recovered. The system can be solved in a least squares-sense using N>2 points.

Finally, knowing the transformation matrix from the CRF to the WRF, it is possible to estimate the position of the target, i.e. of the vehicle, in the WRF.

III. Simulation Results

For the simulation study a horizontal focal length of 960 px with an aspect ratio of 1 has been used. The principal point is located at the center of the image plane of 1024x768 px.

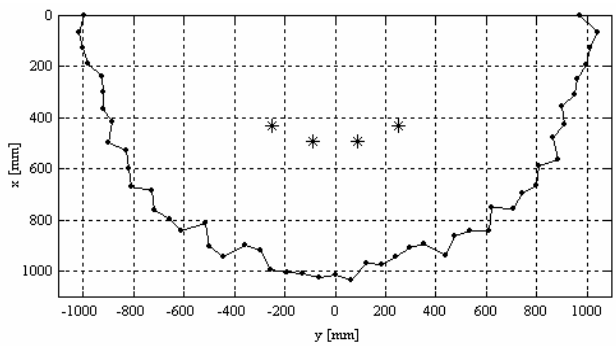

Fig. 2. Typical path generated during the calibration phase

The following extrinsic parameters have been assumed:

- CRF origin position in the WRF: [-1000; 0; 1800] mm;
- Euler angles for the rotation matrix: pan=90° tilt=125° swing=0°.

The camera surveys a working area of 2.5m x 2.0m x 1m. The target consists of six spheres of radius 15 mm at a distance of 150 mm from each other and a cylindrical part with radial and axial dimension of 15 mm and 250 mm, respectively. The dimension of the spheres is not considered in the simulation. The vehicle moves in the area in front of the laser at various radial distances; a typical generated path is shown in Fig. 2. In the same figure, the four cross-marked positions are also reported, where the vehicle was driven onto during the pre-calibration stage. An acquisition rate of 200 ms has been assumed.

Gaussian noise with 0.5 px of standard deviation has been added to image information, while uniform noise with amplitude of 2mm and of 1° has been used respectively for laser measures and laser rotation. Moreover, an angular resolution for the rotating platform of 1° has been considered. The simulation software has been run 500 times, to evaluate the variation of the calibration error during the recalibration process. At the end of each calibration simulation, the CPVL algorithm has been run for various vehicle paths and the performance of the localization method has been evaluated. In the following sections, the results of the simulation study are reported.

A. *Simulation of the Calibration System*

The following results have been obtained over 500 runs. The error for camera position is of 3.3±1.9% in correspondence of the first four positions (i.e. 24 calibration points) acquired in the pre-calibration phase and decreases at 0.37±0.31% for 118 vehicle positions (i.e. 708 calibration points). The percentage camera orientation error, instead, decreases from 0.39±0.15% to 0.06±0.03%. As to the intrinsic parameters, the focal length error passes from 3.72±2.36% to 0.34±0.28%, and the aspect ratio error from 0.29±0.22% to 0.06±0.05%. Fig. 3 - 4 report respectively the mean values of the percent-age camera position and orientation error as function of the number of acquisitions. Similar graphs were obtained for the intrinsic parameters.

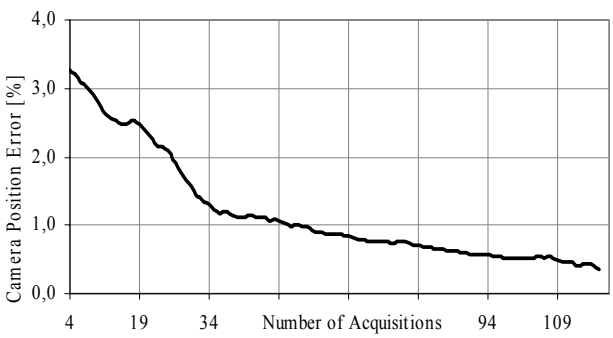

Fig. 3. Percentage camera position error.

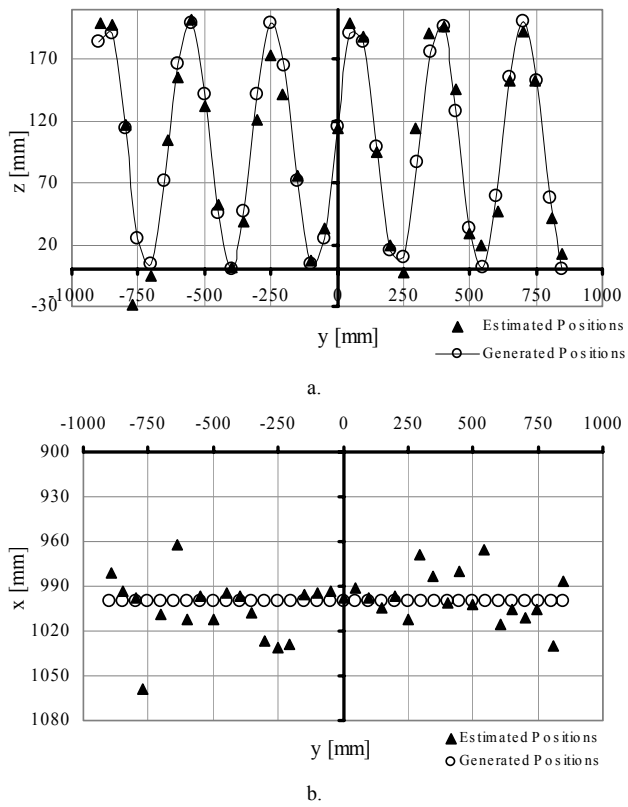

Fig. 5. Camera estimated and generated positions along a linear path in the x-y plane (a.) and a sinusoidal path in the y-z plane (b.)

B. *Simulation of Vehicle Localization and Tracking*

A vehicle path has been generated at the end of each calibration run to perform an error analysis for the localization algorithm.

The following results have been obtained as mean values over 500 runs of the CPVL algorithm, after each calibration run.

For these tests, a circular path and a square path, both at constant quote and a linear path with a quote varying along a sinusoid have been considered. An acquisition rate of 200 ms has been assumed. For the circular path of radius 1 m, the percentage position error was of 1.83±1.02%, while the laser was able to detect 33.1±6.26% of total positions. For the square path of side 1 m, the position error was of 1.74±1.40% and the laser was able to detect 24.2±3.39% of total positions.

Finally, for the linear path with a quote varying along a

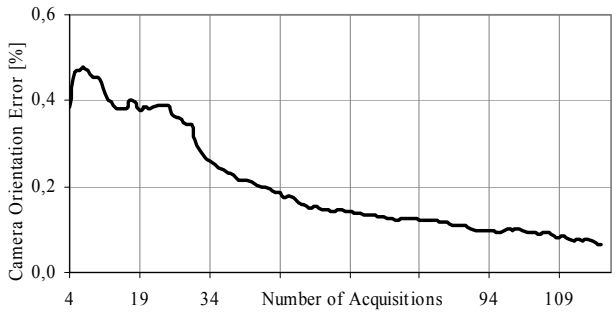

Fig. 4. Percentage camera orientation error.

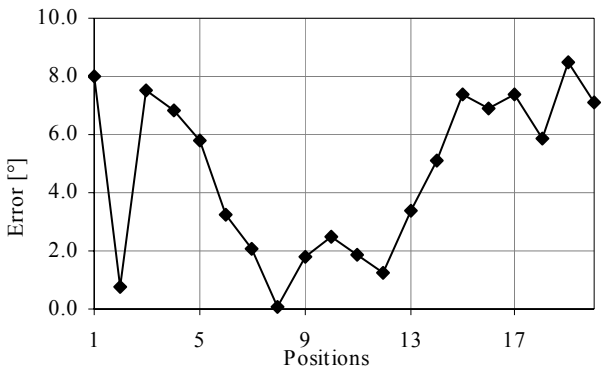

Fig. 6. Orientation error.

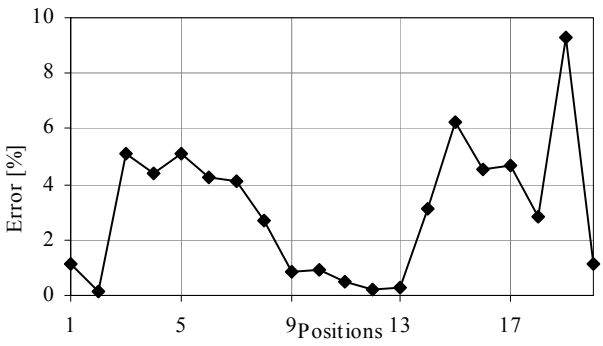

Fig. 7. Percentage distance error.

sinusoid of 200 mm of amplitude, the position error was of 1.64±1.45%, with a number of laser tracked positions that was 50.7±9.16% of total number.

Fig. 5.a-5.b report the generated positions and the positions estimated by the CPVL module in the x-y plane and in the y-z plane, respectively, for this case.

In all runs, the localization error for the measures obtained by the laser device was inferior than the one obtained by the camera of about 80%.

IV. Experimental Results

Experimental tests have been performed to prove the effectiveness of CPVL localization method.

In the experiments the vehicle was equipped with the target described in Section II. Red-colored spheres were employed which can be easily detected in the hue plane.

Once a thresholding has been performed based on the color information, a particle filtering is applied to eliminate other possible extraneous objects in the image. This filtering is based on the definition of a circularity factor, known as Hey-wood factor, which is defined as the ratio of a particle perimeter to the perimeter of the circle with the same area. The closer the shape of a particle is to a disk, the closer the Heywood circularity factor to one.

The vehicle has been controlled to drive along a semi-circumference with a radius of 50 cm, stopping at angular steps of 9°. At each position the measure estimated through the algorithm has been compared with the real one. The results of the error analysis for the various poses are shown in Fig. 6 and in Fig. 7 concerning the orientation error and the percentage distance error, respectively. A mean orientation error of 4.66° with a standard deviation of 2.78°, and a percentage distance error of 3.08±2.45 % have been obtained.

V. Conclusions

An on-line semi-automatic method for accurate and flexible camera calibration and vehicle tracking has been presented. The method integrates a ground-fixed CCD camera with a phase-shifting laser to calibrate the camera and detect a vehicle equipped with a specific target. The results of the simulation study show that the method is efficient in performing camera calibration and it can serve as a precise vehicle localization and map building system. Following the promising results of the simulation, experimental tests will be carried out on a real robotic platform in order to validate the simulation analysis and to perform a comparison with traditional calibration methods.

References

[1] J. Baltes, "Practical Camera and Color Calibration for Large Rooms," *RoboCup-99: Robot Soccer World Cup III,* Manuela Veloso, Enrico Pagello, and Hiroaki Kitano editors, pp. 148-161, New York, Springer Verlag 2000.

[2] R.M. Haralick and L.G. Shapiro, Computer and Robot Vision, Addison Wesley, 1993.

[3] D. Lowe, "Three-Dimensional Object Recognition from Two-Dimensional Images," Artificial Intelligence, Vol. 31, pp. 355-395, 1987.

[4] Q.T. Luong and O.D: Faugeras, "Self-calibration of a Moving Camera from Point Correspondences and Fundamental Matrices," International Journal of Computer Vision, Vol.22 (3), pp. 261-289, 1997.

[5] R. Matsuokaa, K. Fukuea, K. Choa, H. Shimodaa, Y. Matsumaea, K. Hongob and S. Fujiwarab, "A Study on Calibration of Digital Camera," Photogrammetric Computer Vision PCV'02, ISPRS Technical Commission, Graz (Austria), 2002.

[6] E. Trucco and A. Verri, Introductory Techniques for 3D Computer Vision, Prentice Hall, 1998.

[7] R.Y. Tsai, "A Versatile Camera Calibration Technique for High Accuracy 3D Machine Vision Metrology Using Off-the Shelf TV Cameras and Lenses,"IEEE Journal of Robotics and Automation, Vol. RA-3, No 4, pp. 323-344, 1987.

[8] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 22 No 11, pp. 1330-1334, 2000.