EXPERIMENTS IN MULTISENSOR MOBILE ROBOT LOCALIZATION AND MAP BUILDING [1]

José A. Castellanos, **José M. Martínez**, **José Neira**, **Juan D. Tardós**

Dep. de Informática e Ingeniería de Sistemas, Universidad de Zaragoza, c/María de Luna 3, E-50015 Zaragoza, Spain
email: {jacaste, josemari, jneira, tardos}@posta.unizar.es

Abstract:
This article describes an experiment using a mobile robot equipped with a laser range finder and a trinocular vision system, designed to be specially well suited to test multisensor robot localization and map building strategies. A pair of theodolites were used to obtain a ground-true solution to be compared with the results obtained by processing sensor information. Experimental result are reported focusing on the mobile robot localization and map building problems.

Keywords: Mobile robots, Sensor fusion, Stereo vision, Range finders, Calibration

1. INTRODUCTION

Research in indoor mobile robotics includes a large set of problems related to answering the *where am I?* question (Leonard *et al.*, 1992): environment models, algorithms for feature extraction from sensor information, uncertainty representation and integration problems, and feature matching algorithms.

Although using experimental setups to verify the correctness, performance and robustness of systems intended to solve these problems is necessary, using them throughout the whole system development process has several drawbacks:

(1) *Exploratory experiments are time consuming.* Testing a certain feature matching strategy for debugging and tuning may be very costly. These experimental setups are quite sophisticated and complex, and such experiments may require costly hardware, software and human resources.
(2) *Non-repeatability* It is not possible to compare different strategies or alternatives for a specific problem under the same conditions. Each experiment will render different data and thus different results.
(3) *Impossibility to verify against true values.* It is not possible to determine how good your estimated robot location is, or how precise your built map is because the true location and a precise map are usually not available.

Simulation is an alternative to solve these problems only during very early stages of the system development process. It is very difficult to simulate all factors that may influence the robot motion and the obtained sensor observations.

Many researchers use datasets that provide real data and possibilitate repeatability. This article describes an experiment in robot localization and map building designed to be specially well suited to test multisensor robot localization and map building strategies.

[1] This work was partially supported by spanish CICYT projects TAP94-0390 and TAP97-0992-C02-01

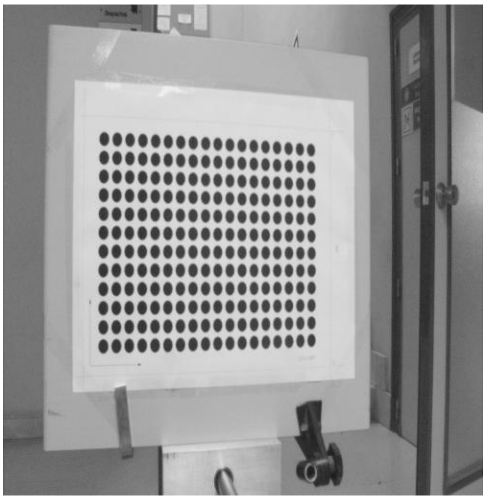

Fig. 1. Calibration pattern for the trinocular vision system.

2. THE CPS EXPERIMENT

The CPS experiment was carried out using the following equipment: a Labmate mobile robot, a laser range finder and a trinocular vision system, both mounted on the mobile robot, and a pair of theodolites, used as a precise location measurement equipment, independent from the sensors and the mobile robot.

The experiment was divided into four stages: sensor calibration, environment modelling, robot location, and environment sensing.

2.1 *Sensor Calibration*

Each of the cameras is calibrated using a pattern (fig. 1) located in three different positions with respect to the camera. These three positions are precisely measured, and allow the calibration to determine both the camera internal parameters, as well as the location of the camera with respect to the patterns (Tsai, 1987). This allows to determine the relative locations between the cameras, and between the cameras and the robot.

The laser range finder and the trinocular vision system are further calibrated using a pattern , that has a vertical edge detectable both by each of the cameras and the laser sensor (fig. 2). This allows to calculate the relative location between the laser sensor and each of the cameras.

2.2 *Environment Modelling*

The environment model obtained is a set of vertical edges, corresponding to wall corners and door frames, whose location was measured with the theodolites. The location of vertical walls was calculated using this information. The resulting 2D environment map is shown in fig. 3.

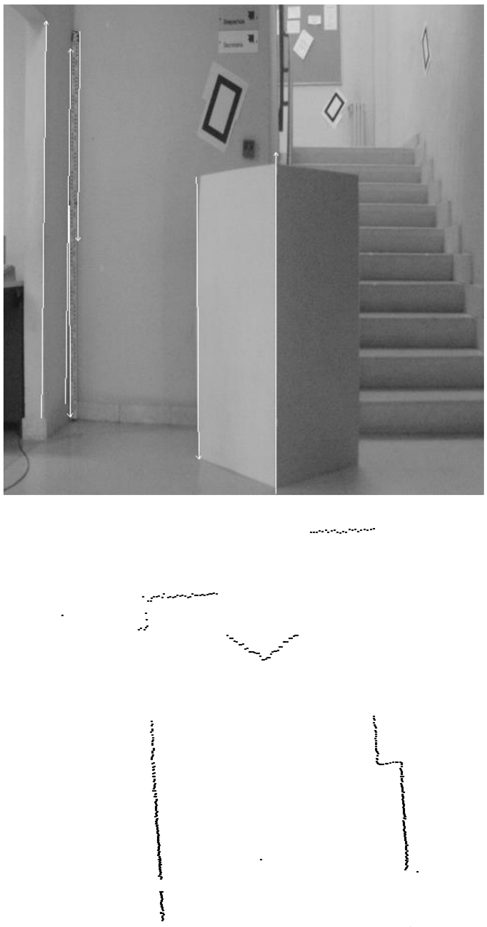

Fig. 2. Calibration pattern for laser and vision as seen by the central camera (top) and the laser range finder (bottom).

2.3 *Robot Location*

The robot was programmed to follow a trajectory in our laboratory. At each step of the path, the following data related to the robot was obtained:

(1) *Nominal robot location*, corresponding to the motion command given to the robot.
(2) *Odometric robot location*, obtained from the Labmate's odometric system.
(3) *Robot location measured with the theodolites.* Being this the most precise measurement independent from sensor observations, this is considered the "true" robot location.

In fig. 3, both the odometric and the theodolite robot location are shown, highlighting the cumulative nature of odometric errors that make the measurement useless.

2.4 *Environment Sensing*

At each step of the robot trajectory, the environment was sensed using both laser range finder and trinocular vision, as shown in figs. 4 and 5. The laser scan can be segmented to extract walls,

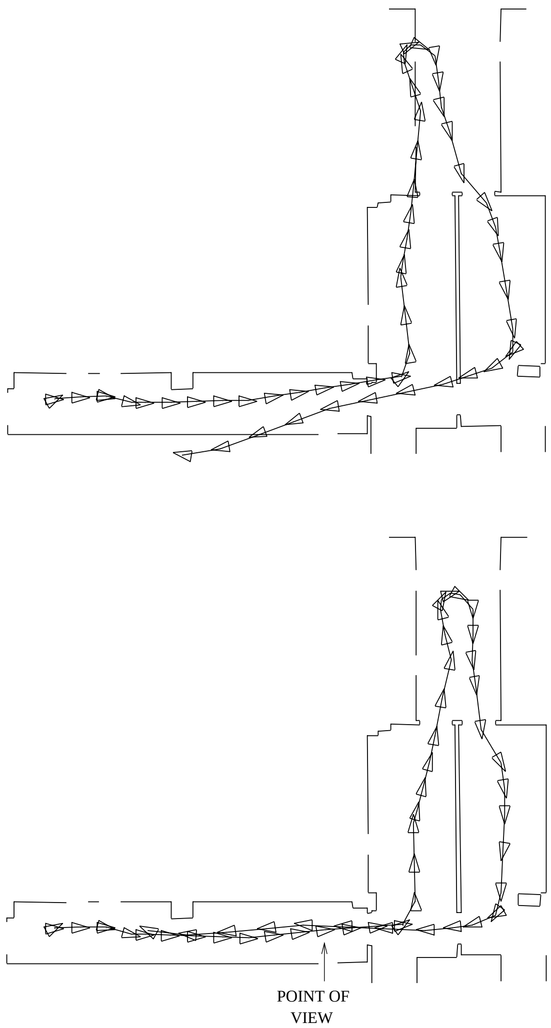

Fig. 3. Environment map and robot location according to odometry (top) and the theodolites (bottom).

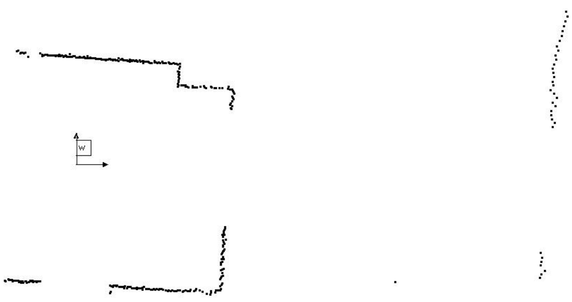

Fig. 4. Laser scan at step denoted POINT OF VIEW of the robot trajectory.

corners, door frames, etc(Castellanos *et al.*, 1996). The trinocular vision system can be used as a stereo or monocular system, to extract scene information of different types(Martínez *et al.*, 1995).

The information obtained in this way allows for a variety of localization and map building problems to work with. Both the laser scans and the trinocular system can be used to extract scene information to be contrasted with the environment model to assess the robot's position, or to be used to build an environment model that can be contrasted with the available model. Experimentation has been carried out with laser features such as points on walls and wall segments(Castellanos *et al.*, 1997), and with vertical edges obtained from monocular images(Neira *et al.*, 1997). Experiments can also be carried out using the vertical edges of the trinocular systems, and with a combination of information of both sensors(Castellanos *et al.*, 1996). In the next sections two of these experiments are presented.

3. ROBOT LOCALIZATION FROM MONOCULAR VISION USING AN *A PRIORI* MAP

Information obtained by real sensors is uncertain in nature. To represent uncertain geometric information, the 2D version of the Symmetries and Perturbations Model (Tardós, 1992; Neira *et al.*, 1996) is used. It is a probabilistic model which considers both the problem of partiality and imprecision of geometric information. In this model, a reference E is attached to every geometric element. For example, an edge is represented with a reference with the X-axis aligned with the edge, whose estimated location with respect to the robot R is given by $\mathbf{L}_{RE} = (\hat{\mathbf{x}}_{RE}, \hat{\mathbf{p}}_E, C_E, B_E)$ where $\hat{\mathbf{x}}_{RE}$ is the estimated location vector of the segment. The real location of the segment is obtained as:

$$\mathbf{x}_{RE} = \hat{\mathbf{x}}_{RE} \oplus \mathbf{d}_E = \hat{\mathbf{x}}_{RE} \oplus B_E^T \mathbf{p}_E \qquad (1)$$

where $\oplus$ represents the composition of location vectors (the inverse is represented by $\ominus$); $\mathbf{p}_E \sim N(\hat{\mathbf{p}}_E, C_E)$ is the perturbation vector of the segment, which takes into account sensor imprecision; and B_E is the self-binding matrix of the entity, which takes into account its symmetries. In the case of an edge there is symmetry of translation along the edge, therefore uncertainty along the X-axis is not considered, thus:

$$\mathbf{d}_E = \begin{bmatrix} 0 \\ d_y \\ d_\phi \end{bmatrix} \;;\; B_E = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \;;\; \mathbf{p}_E = \begin{bmatrix} d_y \\ d_\phi \end{bmatrix}$$

The problem of estimating the robot location by establishing correspondences between vertical edges extracted from an image and features (fundamentally corners and door frames) of an *a priori* map can be stated as follows: let $\mathbf{L}_{WR} = (\hat{\mathbf{x}}_{WR}, \hat{\mathbf{d}}_R, C_R, I_3)$ be the estimated location of the mobile robot; let $\mathbf{x}_{WM}$ represent the location of a vertical edge (a 2D point) according to the *a priori* map. Let $\mathbf{L}_{RE} = (\hat{\mathbf{x}}_{RE}, \hat{\mathbf{p}}_E, C_E, B_E)$ be the estimated location of the edge according to the

Fig. 5. Trinocular Image Set at step denoted POINT OF VIEW of the robot trajectory.

monocular camera. Using only the central image of the trinocular system, our 2D observation of the vertical edge is a projection line between the 2D point and the camera.

Assuming that edge E corresponds to model point M, we can use the observation to improve the estimation of the robot location. For this purpose we use an Extended Information Filter (EIF), stated as follows: the state $\mathbf{x}$ to be estimated is the perturbation vector of the robot location $\mathbf{d}_R$; the partial observation $\mathbf{y}_k$ of the state is the perturbation vector of the edge E, $\mathbf{p}_E$; each observation $\mathbf{y}_k$ is related to the state $\mathbf{x}$ by an implicit nonlinear function of the form $\mathbf{f}_k(\mathbf{x}, \mathbf{y}_k) = 0$. The implicit function states that if observation E corresponds to model feature M, their location must coincide up to symmetries. Since $\mathbf{f}_k$ is nonlinear, we use a first order approximation of $\mathbf{f}_k$:

$$\mathbf{f}_k(\mathbf{x}, \mathbf{y}_k) \simeq \mathbf{f}_k(\hat{\mathbf{x}}, \hat{\mathbf{y}}_k) + H_k(\mathbf{x} - \hat{\mathbf{x}}) + G_k(\mathbf{y}_k - \hat{\mathbf{y}}_k)$$

$$H_k = \left.\frac{\partial \mathbf{f}_k}{\partial \mathbf{x}}\right|_{(\hat{\mathbf{x}}, \hat{\mathbf{y}}_k)} \quad ; \quad G_k = \left.\frac{\partial \mathbf{f}_k}{\partial \mathbf{y}}\right|_{(\hat{\mathbf{x}}, \hat{\mathbf{y}}_k)}$$

In a first step, all observations having only one candidate feature whose location can be considered compatible are fused in the estimation of the robot location (we perform a hypothesis test on their relative location). This process limits the possibilities of accepting an incorrect match, and it is repeated until none of the remaining observations has a suitable candidate. All accepted pairings are fused using the EIF, and in this way the estimated robot location is corrected.

Figure 6 shows the projection of the observed vertical edges at a robot location, and their predicted location after re-estimating the robot location. Figure 7 shows the resulting estimation of the robot location when this process is repeated along the whole trajectory. It can be seen that the resulting estimation of the robot location is highly coherent with the *a priori* map.

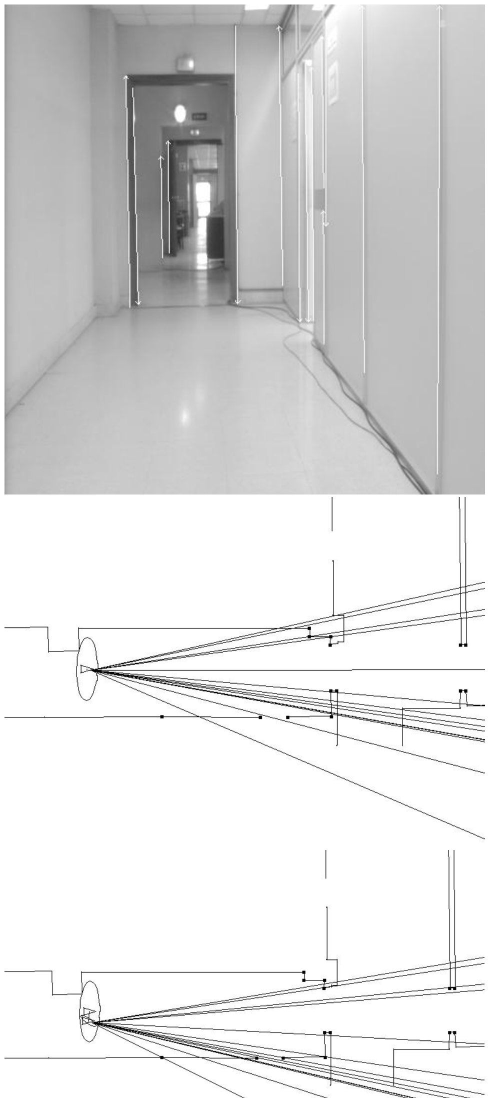

Fig. 6. Vertical edges obtained by the central camera (top) and estimated robot location before (middle) and after (bottom) fusion.

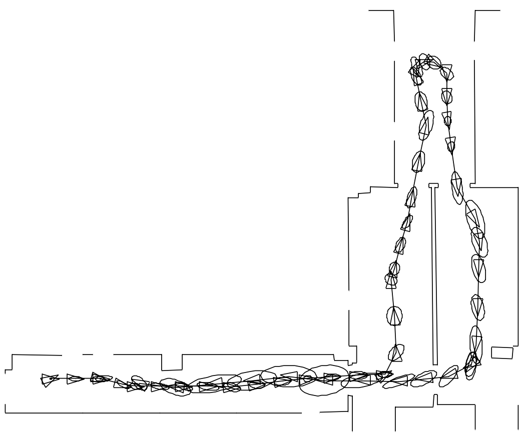

Fig. 7. Estimated robot location and *a priori* map using monocular vision (uncertainty magnified $\times$ 3).

4. ROBOT LOCALIZATION AND MAP BUILDING USING LASER

One of the most complex problems in map building stems out of the fact that the robot needs to build an environment map and simultaneously locate itself within the map. In our work, the construction of a global map of the navigation area is considered simultaneously to the relocation of the robot during navigation (Castellanos *et al.*, 1997). The problem is formulated from an Extended Kalman Filter (EKF) point of view, thus, prediction and estimation phases are considered. The plant model is obtained from the displacement of the robot whilst the measurement model derives from the pairing between observed features (i.e. laser segments) in the local map at time k with those of the global map at time $k-1$.

Laser segments are obtained by application of a segmentation technique (Castellanos and Tardós, 1996) to the set of data gathered by the laser rangefinder. A laser segment is represented by an uncertain location $\mathbf{L}_{RE} = (\hat{\mathbf{x}}_{RE}, \hat{\mathbf{p}}_E, C_E, B_E)$ with respect to the robot R. A length, calculated from its endpoints, is also associated to each segment.

The set of geometric entities involved in the problem of simultaneous map building and localization of the robot is composed of a mobile robot, with an attached reference R and a set of N laser segments, each of them represented by an attached reference $E_i,\ i \in \{1 \ldots N\}$ with respect to a global reference W. The system state vector is composed of the perturbation vectors of the robot and of each of the features in the map. Defining the subvector:

$$\mathbf{x}_M = [\mathbf{p}_{E_1} \ \ \mathbf{p}_{E_2} \ \ \cdots \ \ \mathbf{p}_{E_N}]^T$$

we have:

$$\mathbf{x} = \begin{bmatrix} \mathbf{d}_R \\ \mathbf{x}_M \end{bmatrix} \ ; \ C(\mathbf{x}) = \begin{pmatrix} C_R & C_{RM} \\ C_{RM}^T & C_M \end{pmatrix}$$

where elements in the diagonal represent the covariance of the subvectors and the off-diagonal elements represent the cross-covariance matrices between them.

From odometry measurements we obtain the relative displacement of the mobile robot between two intermediate points of the predefined trajectory. Let R_{k-1} and R_k be the references attached to these points respectively, we might calculate the location of the robot at time k from the composition:

$$\mathbf{x}_{WR_k} = \mathbf{x}_{WR_{k-1}} \oplus \mathbf{x}_{R_{k-1}R_k}$$

with:

$$\mathbf{x}_{R_{k-1}R_k} = \hat{\mathbf{x}}_{R_{k-1}R_k} \oplus \mathbf{d}_{R_{k-1}R_k}$$

where $\hat{\mathbf{x}}_{R_{k-1}R_k}$ represents the displacement of the robot as estimated by dead-reckoning and $\mathbf{d}_{R_{k-1}R_k} \sim N(\hat{\mathbf{d}}_{R_{k-1}R_k}, C_{R_{k-1}R_k})$ represents dead-reckoning errors.

The local map obtained at time k is composed of a set of features which partially might be paired with the features of the global map at time $k-1$. Compatibility between paired features is decided by calculating the Mahalanobis distance D (Bar-Shalom and Fortmann, 1988). Under the Gaussianity hypothesis, D^2 follows a χ^2 distribution. Thus, the local feature L is compatible with the global feature G only if $D^2 \leq \chi^2_{m,\alpha}$ where $\chi^2_{m,\alpha}$ is a threshold value, obtained from the χ^2 distribution with m degrees of freedom, such that the probability of rejecting a good matching is α. At each point of the robot's trajectory it is desirable to obtain as much pairings as possible because they represent the links between new observations and old stored knowledge of the navigation area.

Finally, the EKF estimates both the state vector $\mathbf{x}$ and its covariance $C(\mathbf{x})$ using the classical equations. Integration of new information produces a reduction in the location uncertainty corresponding not only to the paired map features, but also to the whole set of map features, because they are all included in the same state vector. The system covariance matrix is also updated to reestablish the relationships between map features. Non-paired local observations represent knowledge about the environment of the robot which has not yet been learned. Usually, whenever the robot changes from one well-defined navigation area to another, there are many non-paired observations which are directly added to the map. Figure 8 shows the results of this process. It can be seen that the estimated robot location resulting from the process described above is highly precise when compared to the real robot location (fig. 9).

Fig. 8. Estimated robot location and built map using laser.

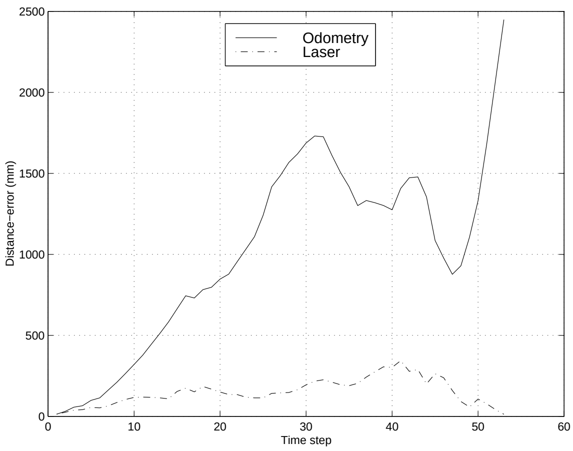

Fig. 9. Errors obtained in robot location estimation.

5. CONCLUSIONS

This paper presents an experiment in multisensor robot localization and map building that allows to use different sensors and sensor combinations to test and evaluate robot localization and map building processes. The fundamental advantage of this experiment is that it is possible to compare the obtained results with very precise measurements both of the environment and of the robot location. The use of this experiment has been exemplified by estimating the robot trajectory using vision and an a priori map, and by testing an EKF-based procedure to simultaneously estimate the robot location and build the map. We can also take advantage of a multisensor system approach (Castellanos *et al.*, 1996), which provides redundancy and assures reliability and precision of the observed features. The datasets obtained in this experiment will be made available to the scientific community at `http://www.cps.unizar.es/deps/DIIS/robot/`.

6. REFERENCES

Bar-Shalom, T. and T.E. Fortmann (1988). *Tracking and Data Association.* Academic Press In.

Castellanos, J.A. and J.D. Tardós (1996). Laser-based segmentation and localization for a mobile robot. In: *Proc. of the Second World Automation Congress.* Montpellier, France. pp. 101–108.

Castellanos, J.A., J. Neira, O. Strauss and J.D. Tardós (1996). Detecting high level features for mobile robot localization. In: *Proc. 1996 IEEE Int. Conference on Multisensor Fusion and Integration for Intelligent Systems.* Washington D.C., U.S.A. pp. 611–618.

Castellanos, J.A., J.D. Tardós and G. Schmidt (1997). Building a global map of the environment of a mobile robot: The importance of correlations. In: *Proc. 1997 IEEE Int. Conf. on Robotics and Automation.* Albuquerque, NM, U.S.A. pp. 1053–1059.

Leonard, J.J., H.F. Durrant-Whyte and I.J. Cox (1992). Dynamic map building for an autonomous mobile robot. *The International Journal of Robotics Research* **11**(4), 286–298.

Martínez, J.M, Z. Zhang and L. Montano (1995). Segment-based structure from an imprecisely located moving camera. In: *IEEE Int. Symposium on Computer Vision.* Florida. pp. 182–187.

Neira, J., J. Horn, J.D. Tardós and G. Schmidt (1996). Multisensor mobile robot localization. In: *Proc. 1996 IEEE Int. Conf. on Robotics and Automation.* Minneapolis, USA. pp. 673–679.

Neira, J., M. I. Ribeiro and J.D. Tardós (1997). Mobile robot localization and map building using monocular vision. In: *Proc. 5th Int. Symposium on Intelligent Robotic Systems.* Stockholm, Sweden. pp. 275–284.

Tardós, J.D. (1992). Representing partial and uncertain sensorial information using the theory of symmetries. In: *Proc. 1992 IEEE Int. Conf. on Robotics and Automation.* Nice, France. pp. 1799–1804.

Tsai, R.Y. (1987). A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-self TV cameras and lenses. *IEEE Journal of Robotics and Automation* **4**(3), 323–344.