A Unified Framework for Uncertainty, Compatibility Analysis, and Data Fusion for Multi-Stereo 3-D Shape Estimation

Mariolino De Cecco, Marco Pertile, Luca Baglivo, Massimo Lunardelli, Francesco Setti, and Mattia Tavernini

***Abstract*—This paper describes the uncertainty analysis performed for the reconstruction of a 3-D shape. Multiple stereo systems are employed to measure a 3-D surface with superimposed colored markers. The procedure comprised a detailed uncertainty analysis of all measurement phases, and the uncertainties evaluated were employed to perform a compatibility analysis of points acquired by different stereo pairs. The compatible acquired markers were statistically merged in order to obtain the measurement of a 3-D shape and an evaluation of the associated uncertainty. Both the compatibility analysis and the measurement merging are based on the evaluated uncertainty.**

***Index Terms*—Multiple stereo systems, uncertainty evaluation, 3-D shape reconstruction.**

I. Introduction

THREE-DIMENSIONAL shape reconstruction using vision systems is a technique widely used to reconstruct spatial objects. Its flexibility has fostered the development of a significant number of algorithms and methods that are available in the literature. However, the increasing resolution of CCD and CMOS is bringing actual measurement performance toward limits that were not available until a few years ago. In 3-D vision metrology, calibration, matching, and uncertainty estimation need to be developed at even more reliable levels in order to follow increased resolution with a comparable level of accuracy. Conversely, increasing computation power is enabling these applications in real-time computing (RTC).

As regards the hardware setup, one main difference is that between multi-camera and multi-stereo. Multiple stereo systems reconstruct shapes by associating camera couples. The use of multiple pairs of cameras allows the reconstruction of different portions visible to each pair and partially overlapping. Compared with the multi-camera procedure, the multi-stereo approach allows a better match between the two views, which are commonly very close to each other. However, the short baseline is prone to high depth uncertainty. In order to increase shape accuracy, the different parts can be matched using Iterative Closest Point (ICP) methods [11], [12]. Then, for each point, a compatibility analysis can be performed with their neighbors in order to fuse each estimate coming from the different couples.

This paper describes a method for uncertainty estimation, compatibility analysis, and position fusion of 3-D points reconstructed from multi-stereo systems. The algorithm can be used in RTC.

Several methods can be used to match the information on different cameras: shape detection, edge detection, correlation analysis, marker matching in the two views, etc. For instance, [1] describes a method for surface reconstruction that employs a Lagrangian polynomial for surface initialization and a quadratic variation method to improve the results. [2] recovers the first approximation of the shape through the object silhouettes seen by the multiple cameras. Then, the shape is improved by a carving approach, employing local correlation analyses between the images taken by different cameras. This approach is based on the hypothesis that if a 3-D point belongs to the object surface, its projection into the different cameras that really see it will be closely correlated. [3] presents a method for the spatial grouping of 3-D points viewed by multiple stereo systems. The grouping algorithm comprises a 3-D space-compressing step in order to map the 3-D points into a space of even density, which allows easier grouping by a neighborhood approach; a subsequent decompressing step preserves the adjacencies of the compressed space and helps the fusion of the grouped points seen by the different cameras.

One of the most important aspects of the multi-stereo approach is the fusion of data coming from different stereo pairs. The process of merging images requires techniques that decide whether points should be merged or not; if they are to be merged, the results must account for each estimation.

One drawback of the previously mentioned approaches is that they do not evaluate the uncertainty of the reconstructed object. If a multiple stereo system is used to perform measurements, a region of confidence of the measured 3-D points or objects should definitely be evaluated to a desired level of confidence. The method described in this paper develops a symbolic uncertainty estimation that merges the measurements performed with different stereo pairs and yields the uncertainty associated with the measured quantities.

Manuscript received July 30, 2009; revised September 25, 2009; accepted October 7, 2009. Date of current version October 13, 2010. The Associate Editor coordinating the review process for this paper was Dr. Alessandro Ferrero.

M. De Cecco, L. Baglivo, M. Lunardelli, F. Setti, and M. Tavernini are with the Department of Mechanical and Structural Engineering, University of Trento, 38100 Trento, Italy (e-mail: mariolino.dececco@ing.unitn.it).

M. Pertile is with the Department of Mechanical Engineering, University of Padova, 35131 Padova, Italy (e-mail: marco.pertile@unipd.it).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIM.2010.2060930

Each step of the measurement process is affected by uncertainty, which propagates to the final 3-D estimates. Uncertainty has different sources, such as noise in image acquisition, defocusing, intrinsic and extrinsic calibration of the parameters of each camera, triangulation, choice of the points to be merged, and the merging method for the final fusion of 3-D parts.

In [4], an uncertainty analysis is presented for a binocular stereo reconstruction, but it does not describe a method to compare and fuse the measurements of the different stereo pairs. In addition, in our method, the covariance of the parameters estimated during the calibration phase is obtained by means of a Monte Carlo simulation, thereby avoiding linearization. Moreover, the correlation between the different parameters is analyzed in depth, giving rise to a covariance that can be considered sparse but not simply diagonal, as in [4].

A method that takes uncertainty into consideration in order to choose the best combination of camera pairs for stereo triangulation is described in [5]. In this case, however, the uncertainties associated with the intrinsic and extrinsic camera calibration parameters are not taken into account, and a simplified geometrical uncertainty estimation and a propagation algorithm that makes use of scalar instead of vector quantities are employed. In this way, cross-correlations between the different sources of uncertainty are neglected. The resulting criteria for camera pair association considers the relative position between the cameras and each point, thus leading to association zones within the field of view and neglecting the uncertainty associated with the matched point lying on the CCD of each stereo pair. This depends on the relative view of the object seen by each camera, i.e., from the orientation of the tangent surface to the object with respect to the camera. In our method, the contribution of each camera pair is used for the final result.

This paper presents a detailed uncertainty analysis performed by applying the general method described in the *Guide to the Expression of Uncertainty in Measurement* (GUM) [6] and its supplement 1 [7]. The described procedure also includes a statistical compatibility analysis performed before the fusion of the different stereo pairs.

The experimental verification used to assess achievable results is based on the acquisition of markers superimposed on the shape to be reconstructed by means of pairs of cameras. The centroid of each marker is detected on each camera, and matching is performed by means of epipolar geometry. Depth evaluation is performed for each pair. Then, the compatibility of the points measured by different stereo pairs is verified. Lastly, the fusion of compatible points is performed on a common reference frame for all cameras. Uncertainty evaluation associates a covariance matrix with each 3-D point reconstructed by each stereo pair. The information contained in the uncertainty ellipsoid is the basis for verifying compatibility, merging compatible points, and estimating their uncertainty. In such a process, each point reconstructed in the 3-D space is not only identified by its coordinates, but is also associated with an uncertainty ellipsoid deriving from the whole reconstruction process.

This information is also available for other purposes, e.g., surface interpolation, which examines not only the position of a point, but also its uncertainty. This can significantly improve interpolation and therefore shape reconstruction. Uncertainty can improve ICP synchronization of point clouds by giving confidence to each point according to its accuracy [11]. Bundle adjustment would also give more confidence to accurate reprojections, thus leading to a more accurate structure and relative pose (or motion depending on the application) reconstruction [15].

In the following sections, the study method is first described (Sections II–V) with a detailed uncertainty analysis, and is then applied to the measurement of a 3-D shape with superimposed markers. Experimental results are presented in Section VI. These results are an important extension of those presented in [16] as they include a bundle adjustment and an ICP method in order to verify the overall proposed method.

II. Stereo Camera Model

As described in [8], the stereo system comprises two cameras, 1 and 2. Each camera has a corresponding frame of reference with its z-axis aligned with the optical axis. Considering the model of each camera, the generic position of a point feature within the field of view of both cameras may be written as

$$ {}^{i}\mathbf{X} = {}^{i}\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda_i \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \lambda_i \mathbf{x}_i \tag{1} $$

where i is 1 or 2 depending on which camera is considered; ${}^{1}\mathbf{X}$ (or ${}^{2}\mathbf{X}$) is the point position expressed in frame 1 (or 2) associated with camera 1 (or 2); $\mathbf{x}_1$ (or $\mathbf{x}_2$) is the projection of the point ${}^{1}\mathbf{X}$ (or ${}^{2}\mathbf{X}$) with an ideal camera aligned like camera 1 (or 2) with a focal length equal to 1 (in length units); and $\lambda_i \in \mathbb{R}^+$ is a scalar parameter associated with the depth of the point.

Each camera is characterized by a set of intrinsic parameters that are evaluated during camera calibration, as described below, and defines the functional relationship between projection $\mathbf{x}_i$, expressed in length units, and projection $\mathbf{x}'_i$, expressed in pixels (x'_i and y'_i are the column number and row number, respectively, from the upper left corner of the sensor). An ideal pinhole camera reveals the following direct model.

$$ \begin{cases} x' = f_m \cdot s \cdot y_i + x'_0 \\ y' = -f_m \cdot x_i + y'_0 \end{cases} \Leftrightarrow $$

$$ \mathbf{x}'_i = \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & f_m \cdot s & x'_{0,i} \\ -f_m & 0 & y'_{0,i} \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \mathbf{K} \cdot \mathbf{x}_i. \tag{2} $$

The inverse model becomes

$$ \begin{cases} x = \frac{y'_i}{-f_m} + \frac{y'_{0,i}}{f_m} \\ y = \frac{x'_i}{f_m \cdot s} - \frac{x'_{0,i}}{f_m \cdot s} \end{cases} \Leftrightarrow $$

$$ \mathbf{x}_i = \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & -\frac{1}{f_m} & \frac{y'_{0,i}}{f_m} \\ \frac{1}{f_m \cdot s} & 0 & -\frac{x'_{0,i}}{f_m \cdot s} \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \mathbf{K}^{-1} \cdot \mathbf{x}'_i \tag{3} $$

where $f_m = f \cdot Sx$; $s = Sy/Sx$; and $Sx = pixels/length\ unit$ along the x axis (not x'); $Sy = pixels/length\ unit$ along

the y axis (not y'); f is the focal length in length units; $x'_{0,i}$ and $y'_{0,i}$ are the distances (in pixel columns and rows, respectively) between the upper left corner and the principal point (intersection of the optical axis with the sensor).

III. Triangulation

When both cameras of the stereo system are calibrated, the 3-D position of a feature point in space may be measured using a triangulation algorithm. In this paper, the algorithm of the middle point is used for triangulation. In theory, when a point feature in space $\boldsymbol{X}$ is acquired by both cameras, the preimage lines that project point $\boldsymbol{X}$ into the sensors should intersect in point $\boldsymbol{X}$ itself. In practice, due to measurement uncertainty, the lines do not intersect. Thus, the algorithm, starting from projected points $\mathbf{x}'_i$, finds 3-D points $\mathbf{X}_{1,s}$ and $\mathbf{X}_{2,s}$ with the minimum distance belonging to the preimage lines of cameras 1 and 2, respectively. Points $\mathbf{X}_{1,s}$ and $\mathbf{X}_{2,s}$ define a segment orthogonal to the two skew preimage lines. Middle point $\mathbf{X}_m$ of this segment is selected as the measured 3-D point of the feature.

Equation (1) can be used to find two generic points, $\mathbf{X}_1$ and $\mathbf{X}_2$. Each of these points belong to the corresponding preimage line and are expressed in the reference frame associated with the corresponding camera

$$ {}^1\mathbf{X}_1 = \lambda_1 \cdot \mathbf{x}_1, \quad {}^2\mathbf{X}_2 = \lambda_2 \cdot \mathbf{x}_2. \tag{4} $$

Both points in the world reference frame yield

$$ {}^w\mathbf{X}_1 = \lambda_1 \cdot {}^w_1\mathbf{R} \cdot \mathbf{x}_1 + {}^w\mathbf{P}_{01}, \quad {}^w\mathbf{X}_2 = \lambda_2 \cdot {}^w_2\mathbf{R} \cdot \mathbf{x}_2 + {}^w\mathbf{P}_{02} \tag{5} $$

where ${}^w_i\mathbf{R}$ is the rotation matrix from frame i to the world frame and ${}^w\mathbf{P}_{0i}$ is the origin of frame i expressed in the world frame.

In order to find points $\mathbf{X}_{1,s}$ and $\mathbf{X}_{2,s}$ with minimum distance, the following cost function g is defined and then minimized by imposing gradient $[(\partial g/\partial\lambda_1) \quad (\partial g/\partial\lambda_2)]$ equal to zero

$$ \begin{aligned} g &= \left\| {}^w\mathbf{X}_1 - {}^w\mathbf{X}_2 \right\|^2 = \left({}^w\mathbf{X}_1 - {}^w\mathbf{X}_2 \right)^T \left({}^w\mathbf{X}_1 - {}^w\mathbf{X}_2 \right) \\ &= \left(\lambda_1 \cdot {}^w_1\mathbf{R} \cdot \mathbf{x}_1 + {}^w\mathbf{P}_{01} - \lambda_2 \cdot {}^w_2\mathbf{R} \cdot \mathbf{x}_2 - {}^w\mathbf{P}_{02} \right)^T \\ &\quad \cdot \left(\lambda_1 \cdot {}^w_1\mathbf{R} \cdot \mathbf{x}_1 + {}^w\mathbf{P}_{01} - \lambda_2 \cdot {}^w_2\mathbf{R} \cdot \mathbf{x}_2 - {}^w\mathbf{P}_{02} \right) \end{aligned} $$

where superscript T indicates matrix transposition. Defining ${}^w\mathbf{P}_{12} = {}^w\mathbf{P}_{01} - {}^w\mathbf{P}_{02}$, which is the origin of frame 1 with reference to the origin of frame 2 and expressed in the world frame, and using the combination of rotation matrices yield

$$ \begin{aligned} g = {} & \lambda_1^2 \cdot \mathbf{x}_1^T \cdot \mathbf{x}_1 - 2 \cdot \lambda_1 \cdot \lambda_2 \cdot \mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 - 2 \cdot \lambda_1 \cdot \mathbf{x}_1^T \cdot {}^1\mathbf{P}_{21} \\ & - 2 \cdot \lambda_2 \cdot \mathbf{x}_2^T \cdot {}^2\mathbf{P}_{12} + \lambda_2^2 \cdot \mathbf{x}_2^T \cdot \mathbf{x}_2 + {}^w\mathbf{P}_{12}^T \cdot {}^w\mathbf{P}_{12} \end{aligned} \tag{6} $$

where ${}^1\mathbf{P}_{12} = {}^1_w\mathbf{R}\,{}^w\mathbf{P}_{12} = -{}^1\mathbf{P}_{21}$; ${}^2\mathbf{P}_{12} = {}^2_w\mathbf{R}\,{}^w\mathbf{P}_{12}$; ${}^1\mathbf{P}_{21}$ is the origin of frame 2 with reference to the origin of frame 1 and expressed in frame 1; ${}^2\mathbf{P}_{12}$ is the origin of frame 1 with reference to the origin of frame 2 and expressed in frame 2; ${}^1_w\mathbf{R}$ (or ${}^2_w\mathbf{R}$) is the rotation matrix from the world frame to frame 1 (or 2); and ${}^1_2\mathbf{R} = {}^2_1\mathbf{R}^T$ is the rotation matrix from frame 2 to frame 1.

Taking partial derivatives and assigning a value of zero to the gradient yield the following equation system:

$$ \begin{aligned} \frac{\partial g}{\partial \lambda_1} &= \lambda_1 \cdot \left(2 \cdot \mathbf{x}_1^T \cdot \mathbf{x}_1 \right) + \lambda_2 \cdot \left(-2 \cdot \mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 \right) \\ &\quad - \left(2 \cdot \mathbf{x}_1^T \cdot {}^1\mathbf{P}_{21} \right) = \mathbf{0} \\ \frac{\partial g}{\partial \lambda_2} &= \lambda_1 \cdot \left(-2 \cdot \mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 \right) + 2 \cdot \lambda_2 \cdot \left(\mathbf{x}_2^T \cdot \mathbf{x}_2 \right) \\ &\quad - \left(2 \cdot \mathbf{x}_2^T \cdot {}^2\mathbf{P}_{12} \right) = \mathbf{0}. \end{aligned} $$

The solutions of this system are the $\lambda_{1,s}$ and $\lambda_{2,s}$ values that define the minimum distance segment between the two skew preimage lines

$$ \lambda_{1,s} = \frac{\left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 \right) \cdot \left(\mathbf{x}_2^T \cdot {}^2\mathbf{P}_{12} \right) + \left(\mathbf{x}_2^T \cdot \mathbf{x}_2 \right) \cdot \left(\mathbf{x}_1^T \cdot {}^1\mathbf{P}_{21} \right)}{\left(\mathbf{x}_1^T \cdot \mathbf{x}_1 \right) \cdot \left(\mathbf{x}_2^T \cdot \mathbf{x}_2 \right) - \left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 \right)^2} \tag{7} $$

$$ \lambda_{2,s} = \frac{\left(\mathbf{x}_1^T \cdot \mathbf{x}_1 \right) \cdot \left(\mathbf{x}_2^T \cdot {}^2\mathbf{P}_{12} \right) + \left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 \right) \cdot \left(\mathbf{x}_1^T \cdot {}^1\mathbf{P}_{21} \right)}{\left(\mathbf{x}_1^T \cdot \mathbf{x}_1 \right) \cdot \left(\mathbf{x}_2^T \cdot \mathbf{x}_2 \right) - \left(\mathbf{x}_1^T \cdot {}^1_2\mathbf{R} \cdot \mathbf{x}_2 \right)^2}. \tag{8} $$

Thus, the extreme points $\mathbf{X}_{1,s}$ and $\mathbf{X}_{2,s}$ of the minimum distance segment are

$$ \begin{aligned} {}^w\mathbf{X}_{1,s} &= \lambda_{1,s} \cdot {}^w_1\mathbf{R} \cdot \mathbf{x}_1 + {}^w\mathbf{P}_{01}, \\ {}^w\mathbf{X}_{2,s} &= \lambda_{2,s} \cdot {}^w_2\mathbf{R} \cdot \mathbf{x}_2 + {}^w\mathbf{P}_{02} \end{aligned} $$

and the middle point associated with the point feature is

$$ \mathbf{X}_m = \frac{\mathbf{X}_{1,s} + \mathbf{X}_{2,s}}{2}. \tag{9} $$

IV. Uncertainty Analysis

In the triangulation algorithm previously described, triangulated point $\mathbf{X}_m$ is computed from the values of the following quantities: 1) $\mathbf{x}'_i$ (i = 1, 2), which are the projections of 3-D point $\mathbf{X}$ in cameras 1 and 2, and are presumed to be known from the measurement (the evaluation of the uncertainty of $\mathbf{x}'_i$ is described in Section IV-B); 2) $f_{m,i}$, s_i, $x'_{0,i}$, and $y'_{0,i}$, which are the intrinsic calibration parameters of the cameras; 3) ${}^w\mathbf{P}_{0i}$, which are the origin positions of the camera frames; 4) ${}^w_i\mathbf{R}$, which are the rotation matrices from camera frame i to the world frame. ${}^w\mathbf{P}_{0i}$ and ${}^w_i\mathbf{R}$ are the extrinsic calibration parameters of the cameras. Both intrinsic and extrinsic parameters with their uncertainties are evaluated by the calibration, as described in Section IV-A. Each rotation matrix can be conveniently expressed by a set of three Euler angles defining rotations around three different axes: α_i, β_i, and γ_i. Thus, for each camera, the extrinsic parameters are the three values of ${}^w\mathbf{P}_{0i}$ and the three Euler angles α_i, β_i, and γ_i.

A. Calibration Uncertainty

The parameters previously defined in the camera model are estimated through camera calibration. The procedure used is similar to that proposed by Tsai [9], with a planar target that

translates the orthogonal to itself, generating a 3-D grid of calibration points. At the first step, the parameters are obtained by a pseudo-inverse solution of a least-squares problem employing points on the calibration volume and image points. After this first estimation of intrinsic and extrinsic parameters, an iterative optimization is performed in order to minimize the errors between the acquired image points and the projections of the 3-D calibration points on the image plane with the estimated parameters.

Before the algorithm of the calibration can be applied, optical radial distortions are estimated and adjusted by rectifying distorted images. Radial distortion coefficients are estimated by the compensation of the curvature induced by radial distortion on the calibration grid [10].

Camera parameter uncertainties are evaluated propagating by the uncertainties of the 3-D calibration points and those of image points [4], [9]. A Monte Carlo simulation is used.

There are various reasons for the deviation between the measured image points and the projection of the 3-D calibration points on the image plane: 1) simplification of the camera model; 2) camera resolution; 3) dimensional accuracy of the calibration grid; and 4) geometrical and dimensional accuracy of grid translation. As the motion of the grid to generate a calibration volume is not perfectly orthogonal to the optical axis of the camera, a bias is induced in the uncertainty distribution of the grid points so that the uncertainty becomes nonsymmetric. Two more parameters are therefore introduced to characterize the horizontal and vertical deviations from orthogonality. The first parameter, α_R, considers the angle between the translation direction and grid rows. In the ideal motion of the grid, the value of α_R is 90°. The second parameter, β_C, considers the angle between the translation direction and the grid columns; also, the ideal value here is 90°.

Summarizing, the calibration routine consists of four steps.

1) Estimation and adjustment of optical radial distortions.
2) First estimation of parameters, which defines the imperfection of the 3-D calibration target (α_R, β_C). These values are achieved by minimizing iteratively the deviation between the measured image coordinates of the calibration points and those reconstructed projecting volume points. The principal point is assumed to lie in the middle of the image. With this assumption, once the systematic deviations from orthogonality have been compensated by α_R and β_C, extrinsic and intrinsic parameters can be derived from a pseudo-inverse solution [9].
3) Final iterative optimization of all camera parameters (including principal points) is performed, iteratively minimizing the deviation between the measured image points and those reconstructed projecting the 3-D calibration points. This step supplies the final estimation of intrinsic and extrinsic parameters. Standard deviation σ of the residuals after the projection is combined with resolution uncertainty, and is used to evaluate the uncertainty associated with the image points used in step 4.
4) Lastly, through a Monte Carlo simulation, the uncertainties of the image points (as evaluated in step 3) and the 3-D calibration points are propagated in order to evaluate the uncertainty of the calibration parameters.

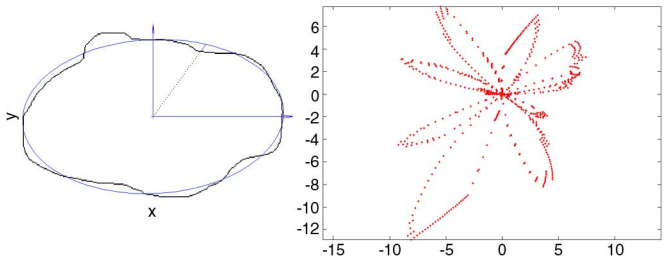

Fig. 1. Example of a badly segmented elliptical marker and the fitted one (on the left), and distribution Δ_b (on the right).

Steps 2 and 3 usually require 5–6 iterations each, and always less than 10 iterations, while the Monte Carlo simulation in step 4 used 10^5 iterations.

B. Matching Uncertainty

Point $\mathbf{X}$ in the 3-D space is defined as the centroid of a circular marker; for this reason, the determination of projection $\mathbf{x}_i'$ on the CCD of point $\mathbf{X}$ is always affected by uncertainty. First, digitalization and successive binarization of the image deforms the circular shape into a polygonal shape, and the centroid of these two shapes is not the same. Second, the marker, which was originally a circle, is deformed in order to adhere to the surface of the target; as a first approximation, the deformed marker is expressed as an ellipse. Third, due to perspective effects, an ellipse, which is not perpendicular to the optical axis of the camera, is projected on the CCD as an ovoid.

A simplified model of the perspective geometry identifies each marker projection as an ellipse. This ellipse can be fitted by the covariance matrix of the distribution of the pixels recognized as markers. The projected marker can then be compared with the corresponding covariance ellipse, and the 2-D distance of the two boundaries (Δ_b) can be computed as a function of angle α^*. This distribution, $\Delta_b(\alpha^*) = b_{real}(\alpha) - b_{fit}(\alpha)$, is a map $\mathbb{R} \rightarrow \mathbb{R}^2$. Fig. 1 shows an example of a badly segmented elliptical marker and the fitted one, and distribution Δ_b.

Two parameters D_C and C_b, which express the "difference" between the projected ovoid and the estimated ellipse can be computed; $D_C = [x_{D_C} \quad y_{D_C}]^T \in \mathbb{R}^2$ is the displacement between the centroids of the segmented marker and the fitted ellipse, and C_b is the covariance matrix of distribution Δ_b.

The uncertainty of the centroid of the segmented marker is represented by a covariance matrix, which is a function of these two parameters

$$\mathbf{U}_{meas} = f(D_C, C_b) = a \begin{bmatrix} x_{\Delta_C} & 0 \\ 0 & y_{\Delta_C} \end{bmatrix} + bC_b.$$

The larger the difference between the projected ovoid and the estimated ellipse, the larger the uncertainty associated with the computed centroid. In this function, parameters (a) and (b) are evaluated by a calibration procedure that uses a grid of circular photolithographic markers. This grid is moved in a set of known positions and orientations, and the computed centroid of the segmented marker is compared with the projected reference on the CCD. In order to have a large set of views, two kinds of grid

are used: the first is a planar surface and the second is the lateral surface of a cylinder.

C. Uncertainty Propagation

The uncertainty evaluation for the triangulated point $\mathbf{X}_m$ of each stereo pair becomes an uncertainty propagation problem that employs the functional model between input and output quantities

$$\mathbf{X}_m = f\left(\mathbf{x}'_i, f_{m,i}, s_i, x'_{0,i}, y'_{0,i}, {}^w\mathbf{P}_{0i}, \alpha_i, \beta_i, \gamma_i\right) \quad \mathrm{i} = 1, 2. \tag{10}$$

Several uncertainty propagation methods are known. Each of them is based on a theory (i.e., probability, possibility, or evidence theory), which can express uncertainty by corresponding suitable means (i.e., probability density functions, fuzzy variables, or random-fuzzy variables). In this paper, according to the GUM [6], the uncertainty is analyzed according to the probability theory and is expressed by probability density functions (PDFs). In order to calculate the propagated uncertainty of triangulated position $\mathbf{X}_m$, and taking into account the contributions of all uncertainty sources, the method based on the formula expressed in the GUM [6] is used. This method is selected instead, for example, of the Monte Carlo propagation approach, to increase computing speed and to allow RTC implementation. The propagation formula uses the sensitivity coefficients obtained from the linearization of the mathematical model; this method is based on the hypothesis that a probability distribution, assumed or experimentally determined, can be associated with every uncertainty source considered, and that a corresponding standard uncertainty can be obtained from the probability distribution.

The GUM proposes a formula for the calculation of the uncertainty to be associated with output quantities $\mathbf{X}_m$, obtainable as an indirect measurement of all input quantities $\mathbf{U}_{out} = \mathbf{c} \cdot \mathbf{U}_{in} \cdot \mathbf{c}^T$, where $\mathbf{U}_{in} \in \mathbb{R}^{24\times24}$ is the covariance matrix associated with the input quantities, which are 24 in this application; $\mathbf{U}_{out} \in \mathbb{R}^{3\times3}$ is the covariance matrix associated with the output quantities, which are the three components of $\mathbf{X}_m$; and $\mathbf{c} \in \mathbb{R}^{3\times24}$ is the matrix of the sensitivity coefficients achievable from partial derivatives of $f()$ with respect to input variables

$$c_{i,j} = \frac{\partial f_i}{\partial input_j}.$$

In this application, the following assumptions are made.

a) The 24 scalar input quantities are considered in this order:

$$\mathbf{x}_1'^T, \mathbf{x}_2'^T, f_{m,1}, s_1, x'_{0,1}, y'_{0,1}, f_{m,2}, s_2, x'_{0,2}, y'_{0,2}, \alpha_1, \beta_1, \gamma_1, {}^w\mathbf{P}_{01}^T, \alpha_2, \beta_2, \gamma_2, {}^w\mathbf{P}_{02}^T$$

b) The two components of $\mathbf{x}'_i$ of each camera are assumed to be cross-correlated and not correlated with any other input quantity.

c) The intrinsic calibration parameters of each camera are assumed to be cross-correlated and not correlated with the corresponding parameters of the other camera or any other input quantity.

d) The extrinsic calibration parameters of each camera are assumed cross-correlated among themselves and not correlated with the corresponding parameters of the other camera and all other input quantities.

These assumptions allow us to build the 24×24 covariance matrix of scalar input quantities, putting six reduced dimension covariance matrices along the diagonal of $\mathbf{U}_{in}$ and assigning zero values to all other elements of $\mathbf{U}_{in}$. The six matrices are the following: 1)$\mathbf{U}_{\mathrm{meas},1} \in \mathbb{R}^{2\times2}$, associated with measurement $\mathbf{x}'_1$ of camera 1; 2) $\mathbf{U}_{\mathrm{meas},2} \in \mathbb{R}^{2\times2}$, associated with measurement $\mathbf{x}'_2$ of camera 2; 3) $\mathbf{U}_{\mathrm{int},1} \in \mathbb{R}^{4\times4}$, associated with intrinsic parameters $f_{m,1}$, s_1, $x'_{0,1}$, and $y'_{0,1}$ of camera 1; 4) $\mathbf{U}_{\mathrm{int},2} \in \mathbb{R}^{4\times4}$, associated with intrinsic parameters $f_{m,2}$, s_2, $x'_{0,2}$, and $y'_{0,2}$ of camera 2; 5) $\mathbf{U}_{\mathrm{ext},1} \in \mathbb{R}^{6\times6}$, associated with extrinsic parameters α_1, β_1, γ_1, and ${}^w\mathbf{P}_{01}^T$ of camera 1; and 6) $\mathbf{U}_{\mathrm{ext},2} \in \mathbb{R}^{6\times6}$, associated with extrinsic parameters α_2, β_2, γ_2, and ${}^w\mathbf{P}_{02}^T$ of camera 2.

The propagation model between the input and output quantities described in Section III, although not very simple, does have the advantage of being explicit. Thus, it is possible to compute explicitly the sensitive coefficients as symbolic expressions, and it is not necessary to evaluate them numerically, which often happens with complex applications.

V. Compatibility Analysis

In non-ideal conditions, the stereo systems at different positions provide different measurements of the same feature (like the center of mass of a colored spot on a surface). Each measurement comes with its uncertainty, and a fusion process can combine them in a single best-estimated one with the associated fused uncertainty. Before points measured from different stereo systems can be fused, it is necessary to state whether they are associated with the same feature or, statistically speaking, whether they belong to the same distribution. A compatibility analysis of the measured points is therefore performed. A compatibility test on two points, $\mathbf{X}_1$ and $\mathbf{X}_2$ with covariances $\mathbf{C}_1$ and $\mathbf{C}_2$, is based on the consideration that the difference $\mathbf{X}_1 - \mathbf{X}_2$ is distributed with zero mean and covariance $\mathbf{C}_1 + \mathbf{C}_2$. On the Gaussian assumption, the Mahalanobis Distance (MD) $D^2 = (\mathbf{X}_1 - \mathbf{X}_2)^T(\mathbf{C}_1 + \mathbf{C}_2)^{-1}(\mathbf{X}_1 - \mathbf{X}_2)$ has a χ^2 distribution with a degree of freedom ν equal to the dimension of vectors $\mathbf{X}$. Once a confidence level α' has been chosen, it is stated that the two points are compatible if $D^2 \leq \chi^2(\nu, \alpha')$. Let $\mathbf{X}_{i,m}$ be the ith 3-D point measured by stereo system m with covariance $\mathbf{C}_{i,m}$. The analysis is made up of the following steps: from measured points sets $\sum_m$ and $\sum_n$ of stereo system m and n, respectively, each $\mathbf{X}_{i,m}$ is associated with the point of $\sum_n$ having the minimum MD to $\mathbf{X}_{i,m}$; if the compatibility test is passed, the association is accepted and the associated couple is fused, yielding the best estimate

$$\mathbf{X}^*_{k,mn} = \mathbf{C}_{i,n}(\mathbf{C}_{i,m} + \mathbf{C}_{j,n})^{-1}\mathbf{X}_{i,m} + \mathbf{C}_{i,m}(\mathbf{C}_{i,m} + \mathbf{C}_{j,n})^{-1}\mathbf{X}_{j,n}$$

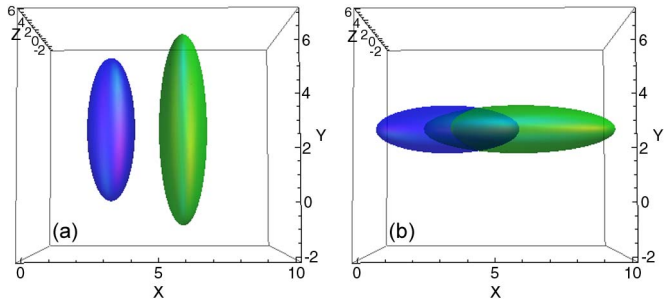

Fig. 2. Examples of different compatibility measures between the same couple of points, with the covariance ellipsoids of equal axis length but different orientation. In a) the MD is 4.5 and in b) the MD is 0.36.

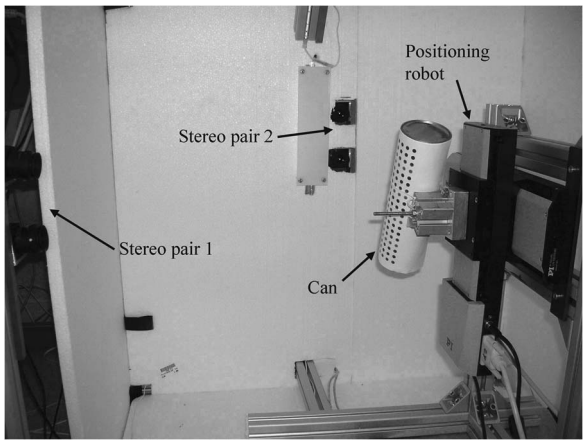

Fig. 3. Experimental setup.

and its covariance matrix

$$\mathbf{C}^{*}_{k,mn} = \mathbf{C}_{i,n}(\mathbf{C}_{i,m} + \mathbf{C}_{j,n})^{-1}\mathbf{C}_{i,m}$$

otherwise, $\mathbf{X}_{i,m}$ and its covariance matrix $\mathbf{C}_{i,m}$ are kept as the best estimate of the feature; the process between all these best estimates is iterated (including points not associated between the two sets), and a new set $\sum_p$ is obtained. Two examples of the compatibility test are shown in Fig. 2.

Ambiguous cases may occur when one point of set $\sum_n$ is compatible with two or more points of set $\sum_m$. In this case, the point of $\sum_n$ is eliminated. For this reason, threshold α' must be tuned in order for both to keep the cases of ambiguity low and not to lose useful information. In this paper, the best compromise of the confidence level α' was identified as equal to 68.3%.

VI. Experimental Results

As described in the experiment in [16], an inclined can provided with colored markers on its lateral surface and positioned by a Cartesian robot is acquired by two stereo pairs, spaced nearly 90° apart, as shown in Fig. 3. Starting from an initial position, the can is translated along a straight trajectory to a final position, and the markers on its surface are acquired in both the initial and final positions.

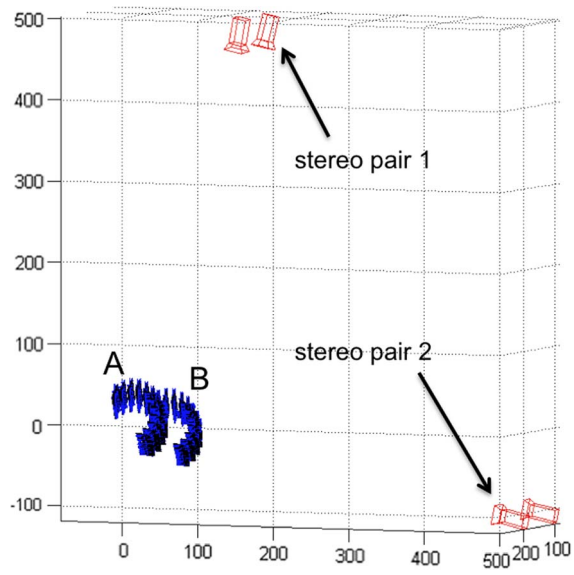

Fig. 4. Points acquired by two stereo pairs.

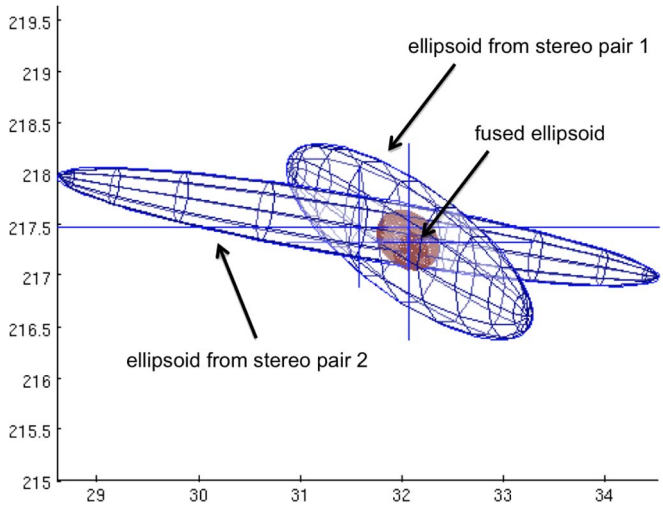

Fig. 5. Covariance ellipsoids of two corresponding points (obtained from the same marker) acquired by the two stereo pairs.

A. Multi-Stereo

The acquired colored markers yield two sets of points, $\sum_1$ and $\sum_2$, for each stereo pair, and these two sets are reconstructed in both positions A and B. The reconstructed 3-D sets of $\sum_1$ and $\sum_2$ in positions A and B are shown in Fig. 4 with their corresponding covariance ellipsoids. For each position, a compatibility test between $\sum_1$ and $\sum_2$ is made. The points that passed the compatibility tests have overlapping ellipsoids (shown in Fig. 5) for two corresponding points (1 marker), one belonging to $\sum_1$ and the other to $\sum_2$. For all corresponding points, their covariance ellipsoids can be fused, as explained in Section V, in order to obtain a fused point with its covariance ellipsoid as shown in Fig. 5 for one marker. Clearly, the uncertainty associated with the fused point is significantly reduced with reference to uncertainties obtained from a single stereo pair. This useful result is particularly true for the two stereo pairs in question since they are spaced 90° apart.

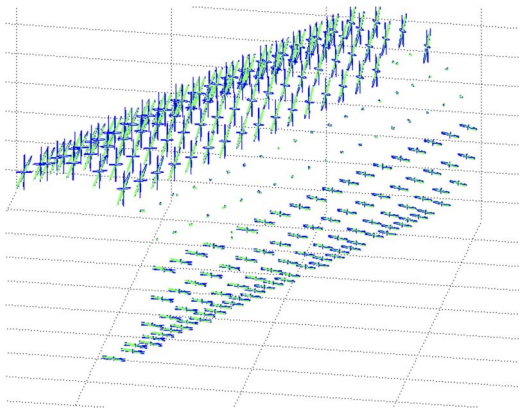

Fig. 6. Covariance elliposids of all the markers in position A.

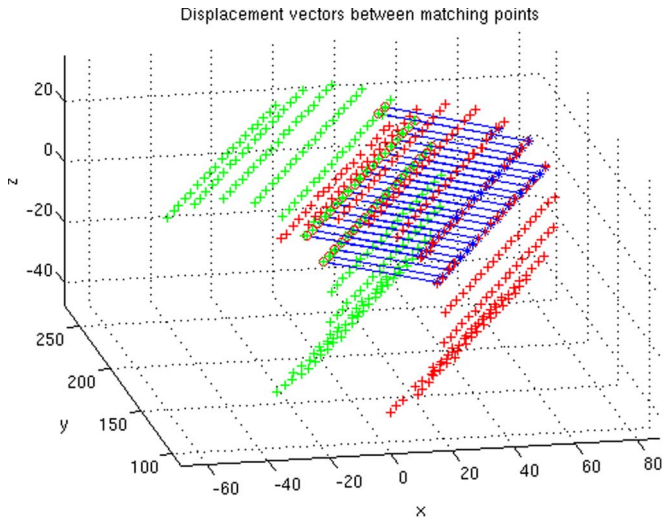

Fig. 7. Blue vectors show the estimated displacements of the matching points of A (green markers) and B (red markers).

The resulting uncertainty for all the markers of position A is shown in Fig. 6, in which the smaller ellipsoids in the center belong to the fused points, which are those compatible between $\sum_1$ and $\sum_2$. It is important to note that in this paper, the ellipsoids shown are smaller than those in [16] because we applied a bundle adjustment algorithm [13], [14] to the cameras. As a consequence, the uncertainty relative to the extrinsic parameters (angles and positions) of the stereo pairs is greatly reduced, thus leading to a minor propagation effect up to the uncertainty of the position of the 3-D markers.

After computing the covariance of each marker, stating compatibility, and fusing points in both positions A and B, we measured the translation of the can from A to B by computing the mean displacement between the corresponding points in the two positions. Point matching is computed with an ICP algorithm that also employs color information for a more robust result. The reference displacement superimposed by the Cartesian robot is 38 mm, with 1 micrometer of spherical overall accuracy. The measured mean displacement is 37.95 mm. Its uncertainty is computed with the squared root of the maximum eigenvalue of the covariance of displacements (the blue arrows in Fig. 7) multiplied by the coverage factor for a confidence level of 95.5%. The resulting value is 0.14 mm. It is fully compatible with the entity of the uncertainty of the points estimated by the propagation and fusion method, showing that the initial parameter calibration was good.

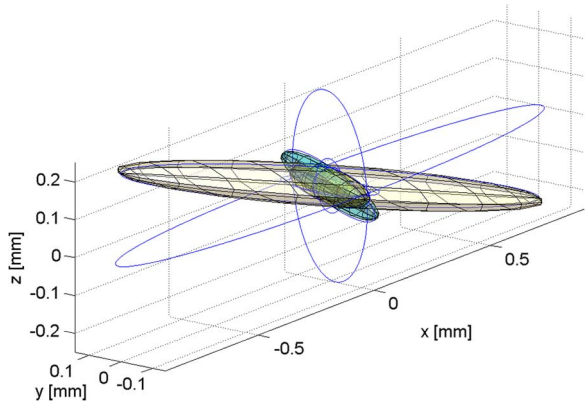

Fig. 8. Ellipsoids of the uncertainties obtained by the multi-stereo (the smaller one), multi-camera (the blue one), and single stereo aproaches.

B. Comparison With Multi-Camera and Single Stereo

The results of the proposed multi-stereo procedure are compared with those obtained by only one stereo system, i.e., without information fusion, and with those obtained by a multi-camera procedure. The latter method is described in [8] and treats the whole system as an array of cameras, with no preferred pairing. In general, considering n feature points and m single cameras grouped in an array, the λ_j, associated with point j viewed by the first camera of the array, can be computed with the following formula:

$$\lambda_j = -\left(\frac{\sum_{i=2}^{m} \left(\widehat{\mathbf{x}_i^j}\, {}^i\mathbf{P}_{1i} \right)^T \widehat{\mathbf{x}_{i\,1}^{j\,i}} \mathbf{R} \mathbf{x}_1^j}{\sum_{i=2}^{m} \left\| \widehat{\mathbf{x}_i^j}\, {}^i\mathbf{P}_{1i} \right\|^2} \right)^{-1}$$

where $\mathrm{j} = 1, \ldots, \mathrm{n}$; $\mathrm{i} = 1, \ldots, \mathrm{m}$; ${}^i_1\mathbf{R}$ is the rotation matrix from camera frame 1 to camera frame i; ${}^i\mathbf{P}_{1i}$ is the origin of frame 1 with reference to the origin of frame i and expressed in frame i; and $\widehat{\mathbf{x}_i^j}$ is the cross product, i.e., $\widehat{\mathbf{x}_i^j}\, {}^i\mathbf{P}_{1i} = \mathbf{x}_i^j \times {}^i\mathbf{P}_{1i}$. Lastly, each 3-D point can be computed by the following expression:

$${}^w\mathbf{X}_j = \lambda_j \cdot {}^w_1\mathbf{R} \cdot \mathbf{x}_1^j + {}^w\mathbf{P}_{01}, \quad \text{with } \mathrm{j} = 1, \ldots, \mathrm{n}.$$

Also, in the multi-camera and single stereo approach, the measurement of the can displacement is performed by computing the displacement mean value of feature points. The multi-camera approach yields a displacement equal to 37.94 mm, with an uncertainty of 0.2 mm (the squared root of the maximum eigenvalue of the covariance of displacements multiplied by the coverage factor for a confidence level of 95.5%), while the single stereo method yields a displacement of 37.98 mm with a corresponding uncertainty of 1.22 mm. A comparison among the covariance matrixes obtained by the three considered approaches is shown in Fig. 8.

Clearly, both multi-stereo and multi-camera approaches yield superior uncertainty results than the single stereo method. The best uncertainty results are obtained by the described multi-stereo approach. This behavior can be explained with the uncertainty sensitive fusion that is carried out in the multi-stereo approach. The measurement fusion of more than one stereo pair gives more weight to the stereo pairs with smaller uncertainty and reduces the effects of low-accuracy measurements.

VII. Conclusions

This paper presents a method for the reconstruction of a 3-D shape by means of multiple stereo systems, estimating and taking into account uncertainty. By means of uncertainty, the compatibility of each of the 3-D points acquired by the different stereo pairs can be estimated, thus allowing us to merge the measurements of the different stereo pairs statistically. This yields a more robust measurement of a 3-D shape while reducing its associated uncertainty at the same time. Thanks to its explicit nature, the algorithm can be applied in RTC.

Experimental results show that the combination of the estimates of each stereo pair noticeably improves the evaluated uncertainties. The measurements performed by the multi-stereo, multi-camera, and single stereo approaches are associated with uncertainties equal to 0.14 mm, 0.2 mm, and 1.22 mm, respectively. These values prove the superior accuracy of the multi-stereo approach, which explicitly weighs each stereo pair with its evaluated uncertainty.

Although the method may be considered quite exhaustive for the purpose of shape-estimation-based point clouds, the problem of outliers is not covered. When the surface is estimated by means of polynomial descriptors or mesh, the presence of outliers causes significant problems. Outliers caused by mismatching can be filtered by various methods such as disparity constraints; nevertheless, methods able to estimate the uncertainty of the matching phase would represent an important step forward.

Acknowledgment

The authors would like to thank Andrea Rao for his valuable collaboration.

References

[1] M. Celenk and R. A. Bachnak, "Multiple stereo vision system for 3-D object reconstruction," in *Proc. IEEE Int. Conf. Syst. Eng.*, Aug. 1990, pp. 555–558.

[2] C. H. Esteban and F. Schmitt, "Multi-stereo 3D object reconstruction," *Comput. Vis. Image Understanding*, vol. 96, no. 3 Spec. Iss., pp. 367–392, Dec. 2004.

[3] S. Nedevschi, R. Danescu, D. Frentiu, T. Marita, F. Oniga, and C. Pocol, "Spatial grouping of 3D points from multiple stereovision sensors," in *Proc. IEEE Int. Conf. Netw. Sens. Control*, Mar. 2004, vol. 2, pp. 874–879.

[4] J. Chen, Z. Ding, and F. Yuan, "Theoretical uncertainty evaluation of stereo reconstruction," in *Proc. II ICBBE*, May 2008, pp. 2378–2381.

[5] J. Amat, M. Frigola, and A. Casals, "Selection of the best stereo pair in a multi-camera configuration," in *Proc. IEEE Int. Conf. Robot. Autom.*, Washington, DC, May 2002, pp. 3342–3346.

[6] BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, Guide to the Expression of Uncertainty in Measurement, Geneva, Switzerland: Int. Org. Standardization 1995.

[7] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, Evaluation of Measurement Data—Supplement 1 to the Guide to the Expression of Uncertainty in Measurement—Propagation of Distributions Using a Monte Carlo Method.

[8] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision*. New York: Springer-Verlag, 2004.

[9] B. K. P. Horn, private communication, 2000.

[10] F. Devernay and O. Faugeras, "Straight lines have to be straight," *Mach. Vis. Appl.*, vol. 13, no. 1, pp. 14–24, Aug. 2001.

[11] E. Trucco, A. Fusiello, and V. Roberto, "Robust motion and correspondence of noisy 3-D point sets with missing data," *Pattern Recognit. Lett.*, vol. 20, no. 9, pp. 889–898, Sep. 20, 1999.

[12] D. W. Eggert, A. Lorusso, and R. B. Fisher, "Estimating 3-D rigid body transformations: A comparison of four major algorithms," *Mach. Vis. Appl.*, vol. 9, no. 5/6, pp. 272–290, Mar. 1997.

[13] J. Zhang, M. Boutin, and D. G. Aliaga, "Robust bundle adjustment for structure from motion," in *Proc. ICIP*, Atlanta, GA, Oct. 8–11, 2006, pp. 2185–2188.

[14] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment—A modern synthesis," in *Proc. Int. Workshop Vis. Algorithms, ICCV*, 1999, pp. 298–372.

[15] M. Irani and P. Anandan, "Factorization with uncertainty," in *Proc. 6th Eur. Conf. Comput. Vis.*, vol. 1842, *LNCS*, Dublin, Ireland, 2000, pp. 539–553.

[16] M. De Cecco, M. Pertile, L. Baglivo, G. Parzianello, M. Lunardelli, and F. Setti, "Uncertainty analysis for multi-stereo 3D shape estimation," in *Proc. AMUEM*, Bucharest, Romania, Jul. 6–7, 2009, pp. 22–27.

Mariolino De Cecco was born in Pescara, Italy, in 1969. He received the electronic engineering degree (first-class honors) from the University of Ancona, Ancona, Italy, in 1995 and the Ph.D. degree in mechanical measurement from the University of Padova, Padova, Italy, in 1998.

He is currently an Associate Professor of Mechanical Measurements and Robotics and Sensor Data Fusion at the University of Trento, Trento, Italy, and is responsible for the university's EU projects, namely VERITAS and AGILE. His interests include mechanical measurements, sensor fusion, vision systems, signal processing, mobile robotics, adaptive control, and space mechanisms and qualification.

Marco Pertile received the M.Sc. degree (with honors) in mechanical engineering and the Ph.D. degree in space science and technology from Padova University, Padova, Italy, in 1998 and 2002, respectively.

In 2005 and 2007, he was a temporary Professor in Mechanical and Thermal Measurement with Padova University. Since 2005, he has given several lectures in Mechanical and Thermal Measurement and Space Robotics. His research interests include dynamic and kinematic analysis of space mechanisms, calibration and performance optimization of measurement systems, measurement uncertainty expression and evaluation, and development of control systems (hardware and software).

Dr. Pertile received the Award Santini from the Italian Association for Aeronautics and Astronautics for best 2007 paper (space section).

Luca Baglivo was born in 1978. He received the degree in mechanical engineering (automation) and the Ph.D. degree in space science, technology, and measurement from the Mechanical Engineering Department, University of Padova, Padova, Italy, in 2003 and 2007, respectively.

Currently, he has a Research Grant in sensor fusion techniques for mechanical systems at the University of Trento, Trento, Italy. His interests include mobile robots control, sensor fusion techniques for AGV localization, and signal processing of mechanical systems.

Massimo Lunardelli received the engineering degree from the University of Padua, Padua, Italy, in 2005.

Since June 2005, he has been working with the University of Padua, the University of Trento, Trento, Italy, and some companies in machine vision applications. His research interests include camera-based precise 3-D measurements, camera calibration, and laser camera systems.

Francesco Setti received the B.S. degree in industrial engineering and the M.S. in mechatronical engineering from the University of Trento, Trento, Italy, in 2005 and 2006, respectively. He is currently working toward the Ph.D. degree in mechanical measurement in the CISAS, University of Padua, Padua, Italy.

He was recently a Visiting Student at the ISR at Instituto Superior Tecnico of Lisboa, Lisboa, Portugal. His research interests include image segmentation, multiple-view vision measurement, image processing for industrial and robotics applications, structure from motion, and motion segmentation.

Mattia Tavernini was born in 1983. He received the B.S. degree in industrial engineering and the M.S. degree in mechatronical engineering from the University of Trento, Trento, Italy, in 2006 and 2009, respectively. He is currently working toward the Ph.D. degree in the University of Trento.

His interests include mobile robots localization, pattern recognition, and data clustering.