3-D Mapping of Natural Environments With Trees by Means of Mobile Perception

Pekka Forsman and Aarne Halme

Abstract—In this paper, a method of generating a three-dimensional (3-D) geometric model for large-scale natural environments with trees is presented. The environment mapping method, which uses range images as measurement data, consists of three main phases. First, geometric feature objects are extracted from each of the range images. Second, the relative coordinate transformations (i.e., registrations) between the sensor viewpoint locations, where the range data are measured, are computed. Third, an integrated map is formed by transforming the submap data into a common frame of reference. Tree trunks visible in the range images are modeled with cylinder segments and utilized as reference features for registration computation. The final integrated 3-D model consists of the cylinder segments representing the visible sections of the tree trunks, as well as of the ground elevation data. The constructed environment map can be utilized as, for example, a virtual task environment for outdoor robotic machines such as new-generation forest working machines or service robots.

Index Terms—Geometric feature extraction, range image processing, viewpoint registration, 3-D environment mapping.

I. INTRODUCTION

In natural outdoor environments, several characteristics make the three-dimensional (3-D) mapping task particularly challenging. Changes in foreground and background illumination complicate the extraction of an object's contour from camera images. Occlusion in cluttered environments may change the visible scene contents drastically when the sensor is displaced to even the smallest extent. Acquiring a good initial estimate for the distance traveled between perception positions may be difficult due to rough outdoor terrain. This paper addresses a method for constructing a 3-D model for forest environments by means of range image data acquired from a number of viewpoints within the environment. The method utilizes feature object models extracted from the range images to estimate spatial relationships between the viewpoint locations. In the case of forest environments, the feature objects are cylinder models corresponding to the visible sections of the tree trunks.

A majority of the probabilistic methods for implementing a simultaneous localization and mapping (SLAM) system can be classified into two main categories. The first includes the statistical state space approaches, where location estimates of the map objects and the mobile robot are maintained and updated within a Kalman filter mechanism. A single location hypothesis along with a Gaussian uncertainty approximation is stored for each map object [1]–[4]. The second, which comprises the expectation-maximization (EM) family of algorithms, provides a relaxation of the single hypothesis—Gaussian noise framework of the Kalman filter-based methods. The EM algorithm is a hill

climbing routine in likelihood space, which alternates two steps: the expectation step (E-step) and the maximization step (M-step) [5]–[7]. In the context of SLAM, the E-step corresponds to the estimation of the viewpoint locations by means of the best available map. Thereafter, the M-step, which recomputes the map by using the viewpoint locations computed during the E-step, is executed.

Lu and Milios proposed an extension to the Kalman filter-based mapping methods [8]. In their method, global locations of the map objects are estimated through observations of the relative relationships between the map objects. Characteristic of the published implementations of this approach is that the map objects in the stochastic environment model correspond to the past locations of the perception sensor augmented with the corresponding observation data [6], [8], [9].

The method proposed in this paper consists of three main phases. First, for all range images, feature object models are extracted from the dense range image data. The feature objects (here cylinder models) are mapped to the horizontally leveled reference frame of the corresponding scanning location to form a submap. As the second phase of the mapping method, relative alignment coordinates among the partly overlapping submaps are computed with respect to the x-, y-, and z-translation degrees of freedom, as well as with respect to the z-rotation axis. As the third, the global locations of the map objects (i.e., submaps) are estimated for building an integrated large-scale map. For computing the large-scale map, the approach proposed by Lu and Milios [8] is applied. The mapping method is demonstrated by constructing an integrated 3-D model for a small (40 meters square) forest area. In the experiment, six high-resolution panoramic range images served as measurement data.

In what follows, Section II proposes the 3-D mapping method. The description of the method is divided into three subsections. The first considers the extraction of cylinder segments from range image data, while the second considers the alignment of object maps; the third introduces an estimation method of building up an integrated map. In Section III, experimental results are described. Finally, concluding remarks are presented in Section IV.

II. GENERATION OF A 3-D HYBRID MAP FOR A FOREST ENVIRONMENT

The hybrid map to be generated consists of both the geometric object models and the ground elevation data. Tree trunks visible in the range images are modeled with cylinder segments and utilized as reference features for registration computation. By augmenting the feature extraction phase for modeling other types of geometric objects, the method can be applied to mapping other types of environments as well. For example, by extracting wall-ceiling corner line segments to augment the feature set the method can be applied to map indoor office environments [10].

A. Extraction of Cylinder Models From Range Image Data

The range images used as measurement data in this study have a size of 3000×720 measurements in horizontal and vertical directions, respectively, yielding, in total, 2.16 million data points. The corresponding field of view of the images is $330 \times 79°$. The sensor is moved from one measurement position to another by simply carrying it or moving it on a mobile platform. In each case, the a priori estimates of the relative displacements of the sensor between neighboring measurement positions are supposed to be relatively inaccurate. Because of this, registration of measurement positions with respect to each other would be very difficult based on the original measurement data points. Huber *et al.* proposed a method of extracting local signatures

from range image data for large-scale map building [11]. The signatures describe the local shape of the object surface at a particular location. The method requires visibility to the same part of the object surface within the two range images to be registered. In forest environments, the visible contents of the scene changes substantially when the sensor is displaced, for example, at the other side of a group of trees. In such a case, it may not be possible to extract enough local signatures within the two range images to enable the computation of a relative alignment estimate for the two perception positions. We propose an alternative approach in which cylinder segments are first extracted from the range images. Then, the centerline segments of the cylinder models are utilized to produce the data points for registration computation (compare Fig. 3). There are two advantages to using this procedure. First, the amount of data for registration computation is reduced dramatically. Second, registration is based on features of a particular geometry, which can be reliably distinguished from the range image data. As a key difference from the method presented in [11], visibility of the same sections of object surfaces is not required from the two scanning locations to be registered. Instead, visibility of any part of the surface of the same symmetric feature objects (here cylinder segments) is required for viewpoint registration, which will be discussed in more detail in Section II-B.

The base line approach for modeling the profile of a cylindrical object uses the width of the range segment for estimating the width of the object. The technique can yield good results if a dense range scan is cast across the object surface. However, this requirement limits the allowable distance between the sensor and target objects. Högström and Wernersson proposed a method for modeling forest scenes, which uses the width of the range segment for approximating the width of the tree trunk [12]. In the method, the range measurements were transformed into a 2-D histogram by projecting them to the xy-plane. Major peaks in the histogram were assumed to correspond to the tree trunks visible in the range image. The measurements within each of the significant peaks in the xy-histogram were labeled as belonging to the particular tree trunk. Then, the data points labeled for a tree object hypothesis were fitted to a smoothed cubic spline for estimating the centerline of the tree trunk. The radius of the tree trunk was determined by studying the range discontinuities around the centerline approximation. The method has two major drawbacks. First, it relies on the distinctive peaks of accumulated data points in the xy-histogram for identifying the tree trunk objects. This has the consequence that a major part of the tree trunk must be visible to the sensor to generate a clear peak in the histogram. Second, the estimate of the width of the cylindrical object determined from the width of the range segment may vary significantly as a function of the measurement distance. The method, to be proposed below, is capable of generating cylinder models for small sections of the tree trunks also. Moreover, the accuracy of the estimated width of the cylindrical objects is rather insensitive to the measurement distance.

As the first step of the method, the range image is segmented into the separate areas of the image that correspond to smooth surfaces of desired size. In this study, smooth areas sized between 6 cm and 1 m were selected for further processing. The given size span corresponds to the width of the tree trunks in the environment. For the segmentation, the approximate orientation of the target objects, i.e., tree trunks, in the range image must be known. In our experiments, all the target objects were known to be vertical (within about $\pm 30°$ bounds). Cylindrical objects oriented nonvertically can be segmented in the same way. In each case, scan lines, or, more generally, consecutive sequences of range measurements within the range image that are approximately perpendicular to the expected object orientation, should be selected for segmentation.

As the second step, circle models are estimated from the segmented range image data to determine the cross-section of the visible parts of
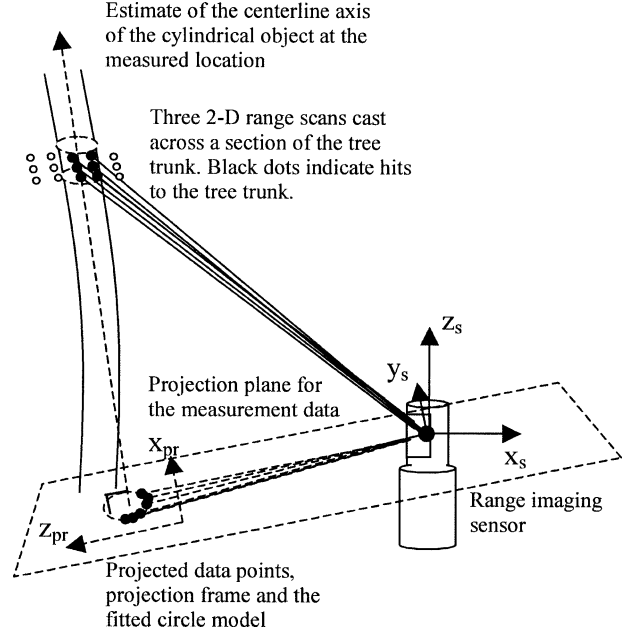

Fig. 1. Illustration of the computation of a circle model approximation for the cross section of a location on the tree trunk. As indicated in the figure, also measurement points contained within (a couple of) neighboring segments both above and below the segment under consideration are used for estimating the circle model.

the tree trunks. For estimating the circle models, we project the data points to a plane, perpendicular to the center axis of the approximated cylinder model of the tree trunk at that location. Another constraint, selected for the projection plane, is that it passes through the origin of the sensor frame. The principle of data projection and circle model approximation for the cross-section of a cylindrical target object (tree trunk), based on the projected data points, is illustrated in Fig. 1.

A circle model for approximating the cross section of a cylindrical object is given by

$$
\begin{aligned}
(z_{pr} - z_0)^2 + (x_{pr} - x_0)^2 &= r^2 \\
\Leftrightarrow z_{pr}^2 + x_{pr}^2 - 2z_0 z_{pr} - 2x_0 x_{pr} &= r^2 - z_0^2 - x_0^2 \\
\Leftrightarrow \left(\frac{1.0}{v}\right)\left(z_{pr}^2 + x_{pr}^2\right) - \frac{2z_0}{v} z_{pr} - \frac{2x_0}{v} x_{pr} &= 1.0 \quad (1)
\end{aligned}
$$

where $v = r^2 - z_0^2 - x_0^2$ and x_{pr}, as well as z_{pr}, correspond to the coordinates of the projected measurement points, x_0 and z_0 are the unknown coordinates of the circle center with respect to the projection frame and r is the unknown circle radius to be estimated. A linear equation with respect to the unknown parameters can be formed with the following substitutions

$$
a = \frac{1.0}{v}, \quad b = -\frac{2z_0}{v}, \quad c = -\frac{2x_0}{v} \quad (2)
$$

yielding

$$
a\left(z_{pr}^2 + x_{pr}^2\right) + b z_{pr} + c x_{pr} = 1.0. \quad (3)
$$

The estimates of the circle model parameters, i.e., the coordinates of the circle centre point, x_0 and z_0, and the radius of the circle, r, can be computed from the least squares estimates of the parameters as follows:

$$
\hat{x}_0 = -\frac{\hat{c}}{2\hat{a}}, \quad \hat{z}_0 = -\frac{\hat{b}}{2\hat{a}}
$$

$$
\hat{r} = \sqrt{\frac{1.0}{\hat{a}} + \hat{x}_0^2 + \hat{z}_0^2} = \sqrt{\frac{4\hat{a} + \hat{c}^2 + \hat{b}^2}{4\hat{a}^2}}. \quad (4)
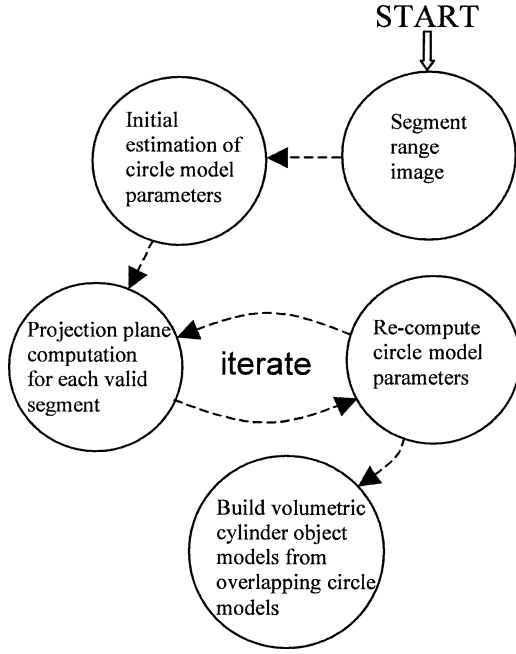$$

Fig. 2. Schematic diagram of the iterative process for modeling the narrow, elongated objects with circular cross sections visible in the range image.

Fig. 3. Two sequences of xyz-data points sampled from the cylinder segment models corresponding to the visible sections of the same tree trunk observed from the origins of the submap frames **A** and **B**.

After computing a circle model, the circularity of the object contour will be tested by studying the symmetry of the residual error of the data points with respect to the estimated circle contour. A symmetrical enough distribution of the error indicates a (nearly) circular cross section of the object. However, if the average residual error exceeds the experimentally predefined limit the estimated circle model is rejected. A similar procedure, called the region test, has been proposed to determine the proper order of a polynomial surface model to be fitted to the segmented range data [13].

After the estimation of the circle model parameters within the 2-D data projection frame, the 3-D coordinates of the circle model center location are computed. The computation of the x-, y-, and z-coordinates of the circle model center location with respect to the reference frame of the corresponding submap is carried out by utilizing the average projection distance of the data points. The 3-D coordinates of the circle center point are computed by back projecting the estimated 2-D location coordinates in the direction of the projection plane normal. The projection distance corresponds to the average of the projection distances of the data points involved in circle model estimation.

The iterative process for extracting cylinder segment models from range images is illustrated in Fig. 2. At the beginning of the process, the initial orientation estimate of the projection plane normal can be obtained from the overlapping segment data or, as used in the context of tree trunk modeling, from the assumption that the objects are roughly vertical. During the following phases of the iterative method, the orientation of the cylindrical object at the measured location is estimated as the direction of a straight line fitted to the circle center point data of the overlapping, nearby circle models. After the initial phase, the computation of projection plane parameters and the subsequent estimation of circle models are repeated until the model parameters stabilize. Usually, two iterations (in addition to the initial estimation phase) are required. Finally, volumetric cylinder object (or truncated cone segment object) models are assembled from the overlapping circle models. There are two primary motivations to form the cylinder segments from the overlapping circle models. First, a more compact representation of the feature based environment model can be generated; second, separate circle models originating from very small visible sections of the
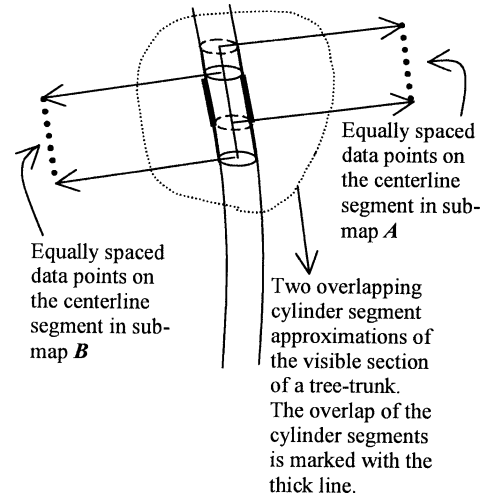
tree trunks or, more importantly, from the measurements hitting the leaves of the trees or the ground, can be filtered with this procedure.

B. Finding an Optimal Alignment Between Two Submaps

In this work, a large-scale environment model is built from local submaps by importing the data contained within the local maps into the common frame of reference of the integrated map. The first step in this process is the determination of the most likely match of the local submaps with respect to the other (overlapping) submaps. The most likely match is defined as the coordinate transformation that maximizes the likelihood score. It is computed as a function of the shortest distances between each of the data points in the first map with respect to all the data points in the second map [14].

Due to the fact that the number of "raw" data points within each of the submaps (i.e., within each of the range images) is very large, and that the a priori relative alignment estimate between the two maps is poor, the cylinder feature objects are utilized to generate data points for registration computation. The data points are generated by sampling from the centerline of each of the cylinder segment models within a submap. This procedure facilitates an implicit solution to the data association problem, which is solved implicitly by finding the maximum likelihood alignment within the registration search space. Note that the EM-algorithm-based mapping methods [5]–[7] solve the data association problem implicitly within the iterative process for finding the maximum likelihood map.

The situation where the physical object corresponds to a section of a tree trunk is illustrated in Fig. 3. In the figure, the point sets for the overlapping part will be sampled from slightly different physical spots of the tree trunk. The uncertainty, due to the mismatch of the closest points in data sets **A** and **B**, can be evaluated as the function of the spacing Δ_d between the two consecutive point samples

$$\varepsilon_d = \frac{\Delta_d}{\sqrt{12}} \tag{5}$$

where ε_d is the standard deviation of the uncertainty of the sampled data points and Δ_d the spacing between two consecutive sampled data points. Equation (5) corresponds to the standard deviation of a rectangular distribution within the interval $\pm(1/2)\Delta_d$. For example, if the sampling interval is 50 mm, the uncertainty, which is related to the assumption that the nearest data point in set **B** with respect to a data point in set **A** actually corresponds to the same physical location on the centerline axis of the tree trunk, becomes 14.4 mm. The analysis
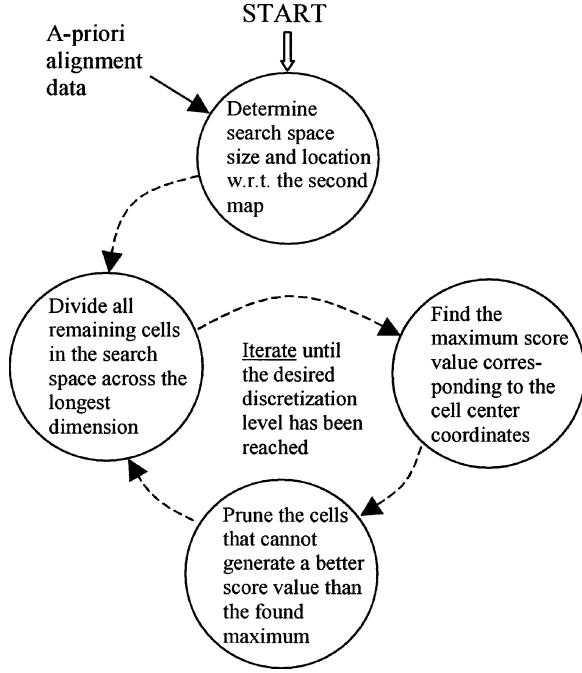
Fig. 4. Schematic diagram of the divide-and-conquer alignment search method. The method is able to find the location in the multidimensional (here four-dimensional) discrete search space, which corresponds to the maximum score value.
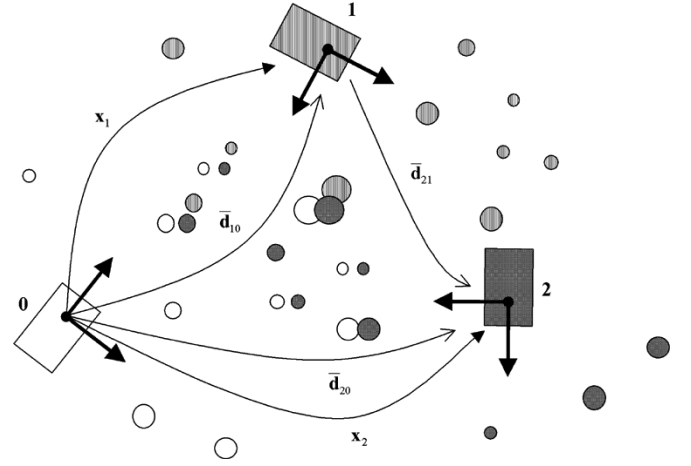

Fig. 5. Illustration of the map construction problem. The circles of a particular texture describe the feature objects extracted from the measurement data acquired from the corresponding sensor location, which is symbolized with the box of the same texture.

is relevant to the approximation of the uncertainty of the data points extracted from the overlapping parts of the geometric feature models. The uncertainty of the data points in set A that have no counterpart in set B has to be considered separately [14].

To find the globally optimal alignment between two submaps, the likelihood score (see [14] for details) is first computed for the predicted position of the first submap with respect to the second submap. Then the transformation space is divided into rectilinear cells (at a coarse resolution) and each cell is tested as to whether it could contain a position that is better than the best position found so far. The cells that cannot contain a position that could contribute to a higher score value are removed from the search space, while the other cells (i.e., stretches in the relative alignment space) are divided into two parts along the longest search direction and tested in the same way. The search for the location in the discrete search space that yields the maximum "match" score value between the two 3-D point sets is illustrated in Fig. 4.

After having reached the desired level of accuracy within the discretized search space, a subgrid cell size solution for the registration problem can be computed. Olson proposed that the log-likelihood values at the close neighborhood of the maximum value could be fit to the parabolic equation to yield the parameters of a Gaussian distribution [14]. In the experiments, we found this kind of uncertainty approximation unsatisfactory when the original distribution had a rough, relatively nonsymmetric form. Instead, we computed the center of mass of the connected grid cell cloud around the cell that had the maximum score value to serve as the registration parameter estimate. The uncertainty of the estimate could then be computed as the weighted sample standard deviation of the connected grid cell cloud. Computation of the alignment estimate can be described with the following equations:

$$\overline{T} = \frac{\sum_{k=1}^{n} l(T_k)T_k}{\sum_{k=1}^{n} l(T_k)}, \qquad \sigma \approx \sqrt{\frac{\sum_{k=1}^{n} l(T_k)(T_k - \overline{T})^2}{\sum_{k=1}^{n} l(T_k)}} \qquad (6)$$

where T_k corresponds to a Cartesian coordinate transformation and $l(T_k)$ is the corresponding weight (or likelihood) of the transformation under consideration. The symbol n corresponds to the number of cells in the connected grid cell cloud. The estimate and related uncertainty are computed separately for each degree-of-freedom of the registration search space.

C. Construction of an Integrated Global Map

In the first mapping phase, cylinder feature object data were extracted from the range images, while local submaps were formed from the feature object data. Then, relative coordinate transformations were computed for each of the overlapping submap pairs. In the final mapping phase, global location coordinates will be estimated for the submaps. The problem is illustrated in Fig. 5. In Fig. 5, $\overline{\mathbf{d}}_{10}$, $\overline{\mathbf{d}}_{20}$, and $\overline{\mathbf{d}}_{21}$ are observations of relative alignment coordinates between a pair of submaps. The vectors $\mathbf{x}_1$ and $\mathbf{x}_2$ correspond to the unknown global locations of the submaps number one and two. The origin of the global frame of reference has been fixed to coincide with the reference frame of the submap having the index zero. The observations in the example of Fig. 5 can be embedded into the measurement equation yielding

$$\overline{\mathbf{d}} = \begin{bmatrix} \overline{\mathbf{d}}_{10} \\ \overline{\mathbf{d}}_{20} \\ \overline{\mathbf{d}}_{21} \end{bmatrix} = \mathbf{Hx} + \mathbf{w} \qquad (7)$$

which can be further expanded

$$\begin{bmatrix} \overline{d}x_{10} \\ \overline{d}y_{10} \\ \overline{d}z_{10} \\ \overline{d}\phi_{10} \\ \overline{d}x_{20} \\ \overline{d}y_{20} \\ \overline{d}z_{20} \\ \overline{d}\phi_{20} \\ \overline{d}x_{21} \\ \overline{d}y_{21} \\ \overline{d}z_{21} \\ \overline{d}\phi_{21} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \\ \phi_1 \\ x_2 \\ y_2 \\ z_2 \\ \phi_2 \end{bmatrix} + \mathbf{w}. \qquad (8)$$

In (8), for example, $\overline{d}\phi_{10}$ is an observation of the heading angle of the first submap frame with respect to the zeroth submap frame (i.e., with respect to the global frame of reference). Now, the minimum variance
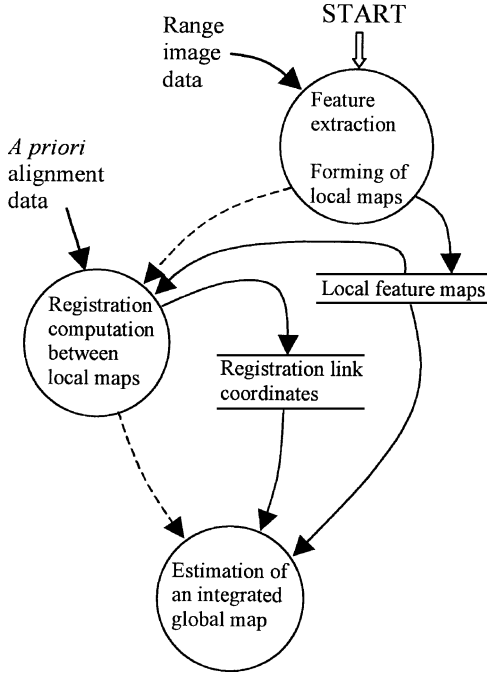
Fig. 6.   Schematic diagram of the environment mapping method. In the diagram, solid and dashed line arrows indicate data access and flow of control, respectively.

estimate of the global coordinates of the perception positions can be computed by [8]

$$\hat{\mathbf{x}} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{C}^{-1}\overline{\mathbf{d}} \qquad (9)$$

and the associated covariance as

$$\mathbf{C_x} = \left(\mathbf{H}^T\mathbf{C}^{-1}\mathbf{H}\right)^{-1}. \qquad (10)$$

The uncertainty of the observation vector $\overline{\mathbf{d}}$ is approximated with a Gaussian distribution with covariance matrix $\mathbf{C}$.

The minimum variance estimation method presented above is a batch process, solving all the location variables at the same time. Every time the perception sensor is moved to a new location, a new submap is created, which is then aligned with the existing neighboring submaps. The observations of relative alignments, which are judged as having a high probability of being correct, are augmented to the $\overline{\mathbf{d}}$ vector and $\mathbf{C}$ and $\mathbf{H}$ matrices while increasing their dimensions correspondingly. The equations above are then solved again to produce new estimates of the global submap locations. A schematic diagram of the overall mapping method is shown in Fig. 6.

A check as to whether all the previously computed registration values are really correct can now be made by studying the residual values of the optimal map. The residual values are computed as the difference between the registered relative alignment coordinates and the relative alignment values derived from the estimated global coordinates of the corresponding submaps in the submap pair under consideration (compare Fig. 5). The vector of residual values for the submap pair (i, j) is given as

$$\mathbf{e}_{(i,j)} = \overline{\mathbf{d}}_{ij} - (\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j) \qquad (11)$$

where $\overline{\mathbf{d}}_{ij}$ corresponds to the observation of the ith submap location with respect to the jth submap location and $\hat{\mathbf{x}}_i$ and $\hat{\mathbf{x}}_j$ are the estimates of the global submap locations. Now, if the absolute values of the components of $\mathbf{e}_{(i,j)}$ are about the same order of magnitude as the uncertainty of the corresponding observations, the integrated map can be presumed to be topologically correct.
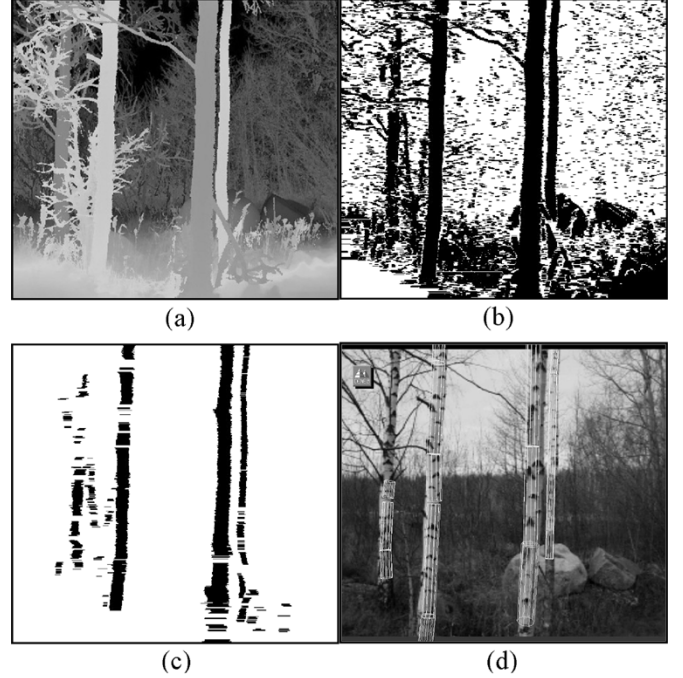

Fig. 7.   Sequence of images corresponding to different stages of the feature modeling method. In (a), the range image from a group of trees is presented. Ranges between 3.5 and 10 m are shown with varying shades of gray. In (b), horizontal segments extracted from the range image are presented. In (c), the valid circle models computed from the measurement points stored within the range segments shown in (b) are depicted. In (d), a camera view of the modeled environment overlaid with the 3-D cylinder model representations, computed from the range image data of (a), is presented.

III. EXPERIMENTAL RESULTS

The experimental system consists of the Riegl LMS-Z210 range-imaging sensor mounted either on a manually maneuvered chart or at the top of a tripod. The chart was used as the sensor platform for the cylinder feature extraction experiments presented in Section III-A and the tripod for the forest mapping experiments presented in Sections III-B and III-C. During the forest mapping experiments, the sensor was leveled in an upright position so that the tilt angles were zero. The vertical sensor orientation results from two facts. First, the tripod did not have tilt sensors to measure the exact inclination angles with respect to the horizontal level; instead, it had a bubble-spirit glass to determine the exact vertical orientation. Second, the vertical sensor orientation offers an ideal scanning sector to acquire a panoramic coverage of the surrounding environment. From the mapping method point of view, the sensor could be placed in any orientation. The only restriction of the mapping method in its present form is that the exact inclination angles of the sensor are determined explicitly and accurately. Accurate measurement of the inclination angles can be easily performed in practice because the sensor platform is expected to be stationary during the acquisition of a range image. When the sensor was mounted on the chart, the inclination angles were measured and correspondingly considered in the kinematics equations of the system.

A. Modeling of Cylindrical Objects From Range Image Data

To demonstrate the first step of the feature object modeling method, the range image, presented in Fig. 7(a), was segmented. During the segmentation, horizontal scan lines were processed in order to find connected sets of data points with a nearly equal distance from the range sensor. Another criterion was set for the segment width, which had to
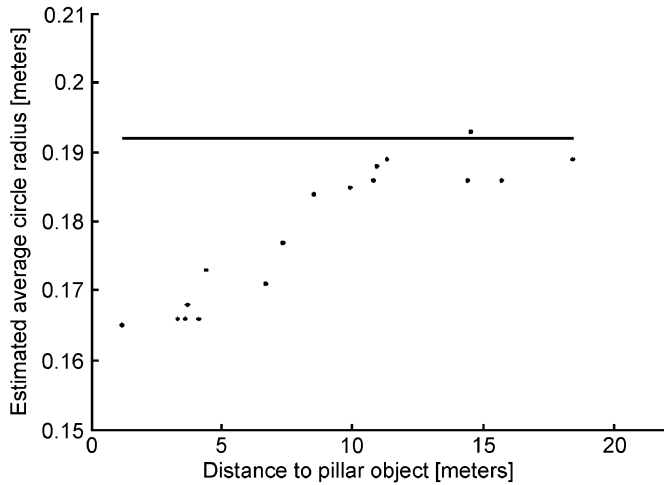
Fig. 8. In the figure, the correlation between the measurement distance and the computed radius of a pillar object is shown. The radius of the cylinder segment model for the pillar object was computed as the average of the radii of the individual circle models belonging to the particular object. The manually measured radius (0.192 m) is indicated with a horizontal line.

be between 0.06 and 1 m. The results of applying the segmentation method can be seen in Fig. 7(b). The range segments in Fig. 7(b), for which a valid circle model could be computed, are shown in Fig. 7(c).

In Fig. 7(c), the width of the estimated circle models is clearly smaller than the width of the corresponding range segments, shown in Fig. 7(b). This is because the Riegl LMS-Z210 range scanner used in this study is designed for measurement distances up to 350 m. Therefore, even a minor share of the emitted laser pulse hitting the surface of a close target yields a valid range measurement. As neighboring laser pulses considerably overlapped each other in the experiments, range segments wider than the underlying object were acquired. Interestingly, in another forest mapping research paper, the width of the range segments corresponding to the cross section of tree trunks was found to be systematically smaller than the true tree width [12]. That could have been the consequence of a less powerful laser range measuring system, which requires that more than half of the laser beam hits the object surface for a valid range measurement.

In Fig. 7(d), a reality augmented virtual camera view of the generated cylinder segment models estimated from the range image data of Fig. 7(a), is presented. The (uncalibrated) camera image was taken with a camera placed at the top of the scanner housing.

The precision of the circle models, computed for estimating the cross section of cylindrical objects, was studied in an indoor mapping experiment. In the experiment, pillar objects of the same size were modeled altogether 17 times across their full length, from the floor level up to the ceiling at a height of about 2.3 m. The measuring distance between the perception sensor and a pillar object varied from 1.2 to 18.4 m. In Fig. 8, the correlation between the computed cylinder radii against the distance from the perception sensor is depicted. As can be seen in Fig. 8, a relatively accurate estimate for the pillar radius was computed when the pillar object was located far away from the sensor. With close measurement distances, a radius value, which is too small, was obtained. A possible cause for the phenomenon is the wide laser beam pattern (beam width 4.2 cm + 3 cm per 10 m of measured distance) and consequent folding of the beam along the object surface. The folding may produce a nonsymmetric error component to the measured beam hit point coordinates, which distorts the circle model parameters. A thorough investigation of the phenomenon would require detailed information of the range measurement method proprietary to the manufacturer of the range-imaging sensor. Nevertheless, it can be argued

that, with a considerably narrower laser beam, the problem would be largely eliminated.

B. Alignment of Local Object Maps

In the mapping method, the local object maps, each formed from a single range image, are aligned with respect to four degrees of freedom. The two remaining degrees of freedom, i.e., the tilt angles of the sensor platform, are measured explicitly. The tilt angles are utilized to transform the range image data to a horizontally leveled reference frame for constructing a local object map. If the tilt angles cannot be measured accurately, for example, they could, in principle, be estimated during the registration phase. However, this would not be feasible in many cases because the computation time for finding the maximum likelihood alignment within a six-dimensional search space would become excessively long.

To demonstrate the registration method, two feature object maps were formed each from a single range image acquired from different locations within the forest test environment. In the map building experiment, the range-imaging sensor was manually maneuvered between the scanning locations. During the experiment, the operator of the sensor was asked to assess the distance traveled between the measurement positions. This assessment served as the a priori estimate of the relative alignment between the corresponding feature maps (or equivalently between the corresponding measurement positions).

In Fig. 9, the evolution of score distribution as a function of search space cell size is depicted. The size of each of the six grid maps is 10 meters square. In the experiment, the a priori alignment estimate was 4.0, -2.0, and 1.0 m along the x-, y-, and z-directions, respectively, and $180.0°$ around the z-rotation axis. The computed relative alignment coordinates at the end of the search (corresponding to the image on the right bottom of Fig. 9) were 4.08, -3.69, and 0.14 m and $161.30°$. The related sample standard deviation was 0.04, 0.03, and 0.15 m and $0.24°$. In Fig. 9, only the grid cells belonging to the *connected* cell cloud around the highest score value have been visualized. Usually, there are also some other cells within the registration search space that are able to survive the cell pruning phase. Eventually, they could be utilized to form a multiple hypotheses registration result. Within the score distribution images, the x- and y-coordinates of the highest score value and the center of mass of the connected cell cloud are marked with a simple and an ended cross hair, respectively. The lengths of the vertical and horizontal lines of the ended cross hair correspond to the computed sample standard deviation of the grid cell cloud.

The calculation cost of the method depends on the number of grid cells to be tested during the divide-and-conquer search process. The calculation cost increases exponentially as a function of the number of dimensions in the search space. The cost is also dependent on the size of the search space along each of the search directions. The size of the search space typically corresponds to the size of the uncertainty of the a priori alignment estimate. The distribution of the feature objects within the maps to be registered has also an impact on the calculation cost. The more evenly the feature objects are distributed within the two maps, the larger the number of cells at each discretization level of the search process that are able to survive the cell pruning phase. The computation of the alignment estimate in Fig. 9 took several minutes on a 1.3 GHz PC computer.

C. Construction of an Integrated Map for a Small Forest Area

In the forest mapping experiment, the range-imaging sensor was manually maneuvered between measurement locations. Consequently, the subjectively assessed a priori alignment estimates for pairs of neighboring submap locations were quite inaccurate. In order to enable the detection of the correct relative alignment coordinates, the dimensions of the registration search space had to be given quite large values:
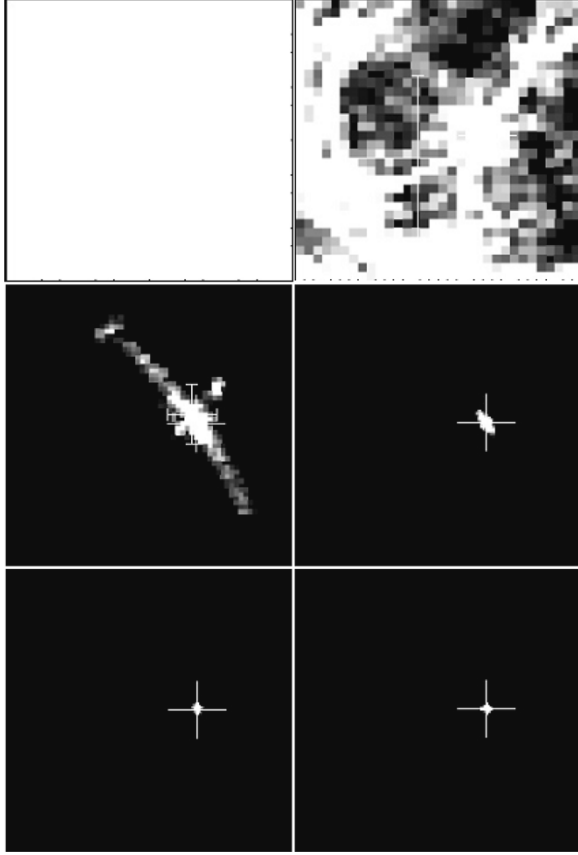
Fig. 9.   Score distribution for search space cell sizes of 0.64, 0.32, 0.16, 0.08, 0.04, and 0.02 m is depicted for the registration pair "RS" in pictures from the upper left image to the image on the right side of the bottom row. In the images, the score values of the cells around the peak score value of the four-dimensional connected score cloud have been accumulated onto the x- and y-coordinate plane of the search space. Note that in the uppermost left image corresponding to the coarsest search level, the score distribution is almost even across the search space. This is due to the cluttered, outlier rich data sets corresponding to the submaps "R" and "S." The directions of the x- and y-coordinate axis are upwards and to the left, respectively.

± 5.0 m in x- and y-directions, ± 1.25 m along the z-direction and $\pm 60.0\,^\circ$ around the a priori heading direction. Six range images and ten subjectively assessed relative displacement coordinates between the six measurement locations served as measurement data for constructing an environment map for the forest area. The range images and corresponding measurement locations were labeled T, S, R, O, Q, and N. The neighboring submap pairs for which a relative alignment estimate was computed were ST, RT, QT, RS, OS, QS, OR, QR, OQ, and NO. A camera-view of the test environment is shown in Fig. 10.

The computation of the global map of the forest area was carried out gradually by starting with the submaps "T" and "S," for which the first integrated map was computed. Then the submap "R" was added to the submap data set and the second map-building phase was started. The relative observation coordinates linking submap "R" to submap "S" were projected to the global reference frame by using the estimated global heading angle of submap "S" computed during the first map-building phase.

In the third phase, submap "Q" was added to the submap data set. The observation coordinates linking it to the submaps "R" and "S" as well as the previously computed relative observation coordinates (i.e., "RS") were projected onto the global reference frame. For the projection the global heading angle estimates of submaps "S" and "R," computed during the second phase of the map building process, were used.

Fig. 10.   Camera view of the forest test environment. The image was taken from above the first scanning position, labeled T, toward the scanning position R (compare Fig. 11).

TABLE I
ESTIMATED SUBMAP LOCATIONS WITH RESPECT TO GLOBAL
FRAME OF REFERENCE

Sub-map label	S
Global sub-map coordinates corresponding to the sub-map set "T, S"	4.43, -1.46, 0.68, -38.71
"T, S, R"	4.42, -1.46, 0.64, -38.68
"T, S, R, Q"	4.42, -1.46, 0.71, -38.66
"T, S, R, Q, O"	4.42, -1.46, 0.50, -38.66
"T, S, R, Q, O, N"	4.42, -1.46, 0.50, -38.66

R	Q
X	X
5.29, -6.90, 0.78, 122.62	X
5.30, -6.90, 0.65, 122.61	10.52, -3.98, 1.27, 127.96
5.30, -6.90, 0.74, 122.61	10.52, -3.98, 1.28, 128.01
5.30, -6.90, 0.74, 122.61	10.52, -3.98, 1.28, 128.01

O	N
X	X
X	X
X	X
15.33, -4.49, 1.55, 264.76	X
15.33, -4.49, 1.55, 264.76	16.61, -9.60, 2.30, 302.94

Units are meters for the translation degrees of freedom (x, y, z) and degrees for the z-rotation angle (ϕ).

The integration of a new submap was repeated altogether five times, until all the six submaps were included in the computation of the global map. The global locations of the submaps, computed during each of the five environment mapping phases have been given in Table I. The origin of the global frame of reference coincides with the reference frame of the submap labeled "T." As can be seen in the table, the coordinates of the global submap locations change very little as new submaps are incorporated into the integrated map. It is a consequence of both accurate alignment of neighboring submaps and a topologically correct map. In Fig. 11, an elevation map, corresponding to the estimated global submap locations, is shown. The evolution of the combined map of the forest environment, consisting of the ground elevation information and the cylinder approximations of the visible section of the tree trunks, is depicted in Fig. 12. In Fig. 13, a bird's-eye view of the 3-D realization of the combined elevation/feature map is shown.
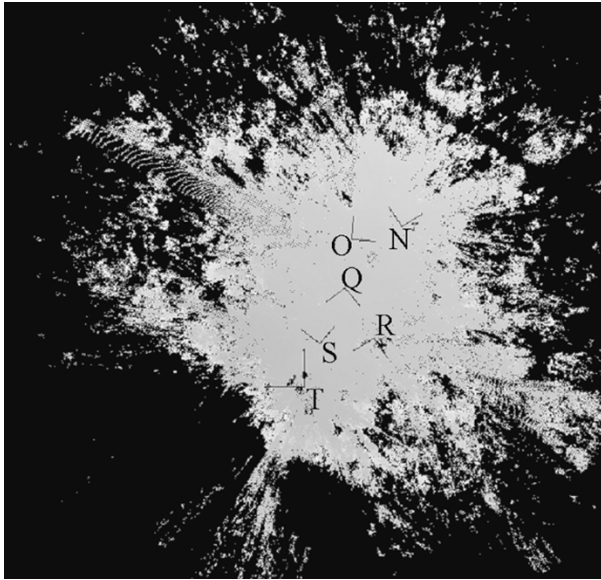
Fig. 11. Size of the grid map is 600×600 pixels corresponding to an area of 60 meters square. In the image, black indicates no registered elevation value. The first measurement position "T," corresponding to the global frame of reference of the integrated map, is indicated by the larger xy-coordinate axis symbol. The other five scanning positions are marked with smaller symbols.

Fig. 13. Bird's-eye view of the combined 3-D elevation/feature model of the small forest area. The size of a single grid element in the elevation map is 0.6 m. A model of a mobile robot (not used in the experiments) is placed at the origin of the global coordinate system. The field of view of the image sequence shown in Fig. 12 has also been sketched in the figure.
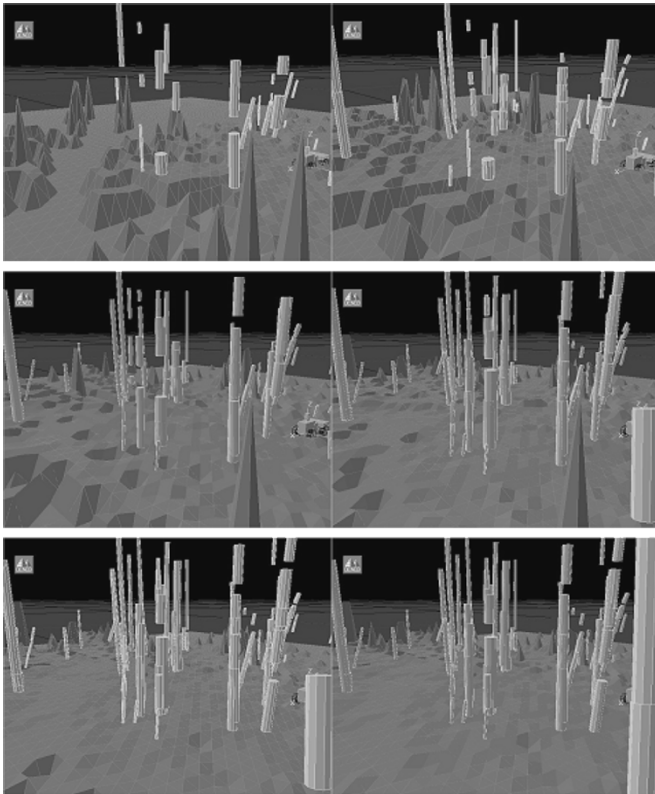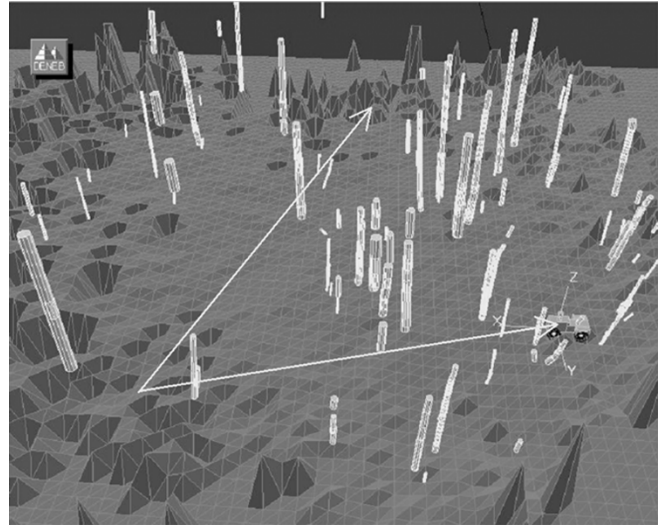

Fig. 12. View of the 3-D representation of the environment model for the forest test site. The images starting from the upper left corner correspond to the integrated map constructed from the submap sets "T," "T,S," "T,S,R," "T,S,R,Q," "T,S,R,Q,O," and "T,S,R,Q,O,N," respectively. The peaks in the elevation map, due to multiple hits to the branches of the trees, can also be seen in some of the figures.

At the end of the experiment, it was discovered that the a priori estimates of the relative translation displacements between the scanning positions deviated from the computed values from a few centimeters up to 4.5 m, while the difference between the a priori and computed heading angles deviated from a few degrees up to $32\,^\circ$. The capability to find the correct alignment between a pair of submaps, even when the initial alignment estimate has a large error, is due to the globally optimal registration method. The prerequisite of the registration method to converge with the correct solution is that the solution resides within the branch and bound search space. Further experimental results of the presented 3-D environment mapping method can be found in [10].

IV. CONCLUSION

In this work, a method for 3-D localization and mapping of large-scale (outdoor) environments is presented. The measurement data was collected with a new type of two-axis range-imaging sensor. The achieved results are partly specific to this category of perception systems, especially with respect to the feature extraction phase of the method. On the other hand, computation of the relative alignment (i.e., registration) between a pair of submaps is a generic technique applicable to models constructed by means of other mapping approaches and perception systems. Currently, the geometric object data, which is transformed to the common frame of reference, usually constitutes only a partial, fragmented model of the underlying physical objects. Merging the partial models (cylinder segments) into a more complete, smooth model representation remains as future work.

The presented work is a part of the project for the development of Workpartner, the prototype of a mobile service robot designed to work interactively with people in outdoor environments [15]. There are two main applications planned for the utilization of the constructed maps. First, the large-scale integrated map can serve as the basic navigational reference of an autonomous mobile robot. The second application is related to the development of a high-level communication interface for future service robots. Cognitive communication between people and robots requires a common understanding of the operational environment. The models, constructed using the presented method, can offer the basic media for the human-machine dialog. Furthermore, in this application, the automated extraction of tree trunks and similar objects helps to visualize the environment for building its object content further.

Beyond the robotics field of science, the developed 3-D mapping method could be applied elsewhere, such as in forest surveying. However, at the present stage of the development of the method and considering the characteristics of the applied range sensor, the precision requirements of the tree trunk models might be difficult to meet.

REFERENCES

[1] P. Moutarlier and R. Chatila, "Stochastic multisensory data fusion for mobile robot location and environment modeling," in *Proc. 5th Int. Symp. Robotic Research*, Aug. 1989, pp. 85–94.

[2] R. Smith *et al.*, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*, I. J. Cox and G. T. Wilfong, Eds, New York: Springer-Verlag, 1990, pp. 167–193.

[3] J. Guivant and E. Nebot, "Optimization of the simultaneous localization and map-building algorithm for real-time implementation," *IEEE Trans. Robot. Autom.*, vol. 17, no. 3, pp. 242–257, Jun. 2001.

[4] H. J. Feder, J. J. Leonard, and C. M. Smith, "Adaptive concurrent mapping and localization using sonar," in *Proc. IEEE/RSJ Int. Conf. Robotic System (IROS)*, Victoria, B. C., Canada, Oct. 1998, pp. 892–898.

[5] S. Thrun *et al.*, "A probabilistic approach to concurrent mapping and localization for mobile robots," *Auton. Robots*, vol. 5, pp. 253–271, 1998.

[6] S. Thrun *et al.*, "Integrating topological and metric maps for mobile robot navigation: a statistical approach," in *Proc. AAAI-98*, 1998, pp. 998–995.

[7] W. Burgard *et al.*, "Sonar-based mapping with mobile robots using EM," presented at the *16th Int. Conf. Machine Learning (ICML'99)*, Bled, Slovenia, 1999.

[8] F. Lu and E. E. Milios, "Globally consistent range scan alignment for environment mapping," *Autonomous Robots*, vol. 4, no. 4, pp. 333–349, 1997.

[9] J.-S. Gutmann and K. Konolige, "Incremental mapping of large cyclic environment," in *Proc. IEEE Int. Symp. Computational Intelligence in Robotics and Automation*, 1999, pp. 318–325.

[10] P. Forsman, "Three-dimensional localization and mapping of static environments by means of mobile perception," Ph.D. dissertation, Dept. Automations Syst. Technol., Helsinki Univ. Technol., Finland, 2001.

[11] D. Huber *et al.*, "3-D map reconstruction from range data," in *Proc. IEEE Int. Conf. Robotic Automation*, San Francisco, CA, Apr. 2000, pp. 891–897.

[12] T. Högström and Å. Wernersson, "On segmentation, shape estimation and navigation using 3-D laser range measurements of forest scenes," in *Proc. Preprints 3rd IFAC Symp. Intelligent Autonomous Vehicles*, Madrid, Spain, Mar. 1998, pp. 547–552.

[13] P. J. Besl and R. C. Jain, "Segmentation through variable-order surface fitting," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 2, pp. 167–192, Mar. 1988.

[14] C. F. Olson, "Probabilistic self-localization for mobile robots," *IEEE Trans. Robot. Autom.*, vol. 16, no. 1, pp. 55–66, Feb. 2000.

[15] A. Halme *et al.*, "Workpartner: interactive human-like service robot for outdoor applications," *Int. J. Robot. Res.*, vol. 22, no. 7–8, pp. 627–640, Jul.–Aug. 2003.

Hybrid Kinematic and Dynamic Simulation of Running Machines

Wei Hu, Duane W. Marhefka, and David E. Orin

Abstract—Dynamic simulation requires the computationally expensive calculation of joint accelerations, while in kinematic simulation these accelerations are known based on a given trajectory. This paper describes a hybrid kinematic and dynamic simulation method that can be applied to the simulation of running machines to speed up the computations over that of a dynamic simulation. This is possible because much of the time the legs of a running machine are in the air and their trajectories are directly specified and tightly controlled. The method is more flexible than dynamic simulation alone because it allows joints to be either motion-controlled or force-controlled. It is general to all robotic systems with tree structures, and fully motion-controlled or force-controlled kinematic loops. It should work best for machines with appendages that are motion-controlled, such as those encountered in underwater and space manipulation.

Index Terms—Closed loop, hybrid simulation, legged locomotion, robot dynamics.

I. INTRODUCTION

Dynamic simulation has been an important tool for controller and mechanical structure design for the next generation of robotic systems. It has been especially valuable in the intelligent controller design for a quadruped running machine [1], [2]. Computational efficiency continues to be an important performance criterion for dynamic simulation, especially with the increased complexity of the systems and applications being considered. A new high-performance articulated leg design for a quadruped running machine contains kinematic loops [3], and this structure places greater demands on computational efficiency. The use of simulation in the evolutionary search for dynamic running gaits in [2] has also shown the need for improved computational speed.

Basically, there are two important methods for dynamic simulation of open tree-structure systems [4]. The composite rigid body (CRB) algorithm by Walker and Orin is based on the efficient recursive Newton–Euler computations of inverse dynamics [5]. This algorithm is of complexity $O(n^3)$ and works best for systems with a small number of links. Featherstone developed the articulated-body (AB) algorithm with $O(n)$ complexity. It uses the AB inertia to relate the acceleration of a link to the force applied to the link [6]. The hybrid simulation algorithm developed in this paper is based on the AB method, but it uses aspects of the CRB method to compute the inertia for motion-controlled appendages.

When the leg of a running machine is in the air, the trajectory of the leg is planned and a tight proportional derivative (PD) controller is set to move the leg along this trajectory [1], [2]. To make use of this property to improve the efficiency of the simulation, our paper presents a new hybrid dynamic and kinematic simulation method. In hybrid simulation (forward dynamics), some of the joints are motion-controlled in addition to those that are force-controlled. Motion-controlled joints are

W. Hu and D. E. Orin are with the Department of Electrical and Computer Engineering, The Ohio State University, Columbus, OH 43210 USA (e-mail: hu.130@osu.edu; orin.1@osu.edu).

D. W. Marhefka is with Kent Displays, Kent, OH 44240 USA.