Simultanous Mapping and Localization of Rescue Environments

Gleichzeitiges Kartieren und Lokalisieren von Rescue-Umgebungen

Hartmut Surmann, Kai Pervölz, Fraunhofer AIS, Sankt Augustin,
Andreas Nüchter, Kai Lingemann, Joachim Hertzberg, University of Osnabrück,
Matthias Hennig, Dresden University of Technology

Summary Deploying rescue workers in an urban setting is often a perilous, time-, power-, and force-consuming job, and robot systems to assist in this effort are needed. A fundamental task for rescuer is to localize and salvage injured persons. To this end, robotic systems are used for mapping a site and for remote inspection of suspicious objects. The mobile robot Kurt3D is the first rescue robot that is capable of mapping its environment in 3D and self localize in all six degrees of freedom, i. e., considering its *x*, *y* and *z* positions and the roll, yaw and pitch angles. This capability allowed the small robot to win the Vice World Championship at the RoboCup Rescue competition 2004 at Lisbon. ►►► **Zusammenfassung** Der Einsatz von Rettungskräften in Ballungsgebieten ist eine gefährliche, Zeit-, Energie- und kraftaufwändige Arbeit, bei der Robotersysteme zur Unterstützung benötigt werden. Eine fundamentale Aufgabe der Retter ist die Lokalisierung und Bergung verletzter Personen. Robotersystem werden für die Kartierung der Umgebung und der Inspektion von verdächtigen Objekten verwendet. Der mobile Roboter Kurt3D ist der erste Rescue-Roboter, der in der Lage ist, seine Umgebung dreidimensional zu kartieren und sich in der Karte selbst zu lokalisieren, und zwar in allen 6 Freiheitsgeraden, d. h. unter Berücksichtigung der drei Raumpositionen *x*, *y* und *z* und den Dreh-, Schwenk- und Neigewinkeln. Diese Fähigkeit ermöglichte es dem kleinen Roboter, die Vize-Weltmeisterschaft auf dem RoboCup Rescue Wettbewerben 2004 in Lissabon zu gewinnen.

KEYWORDS I.2.9 [Robotics] Mobile Robots, 6D-SLAM, RoboCup Rescue, Mapping, Localization

1 Introduction

For protecting humans, it is nowadays important to build robots that are able to operate in earthquake, fire, explosive and chemical disaster areas. The community of Urban Search and Rescue Robotics (USAR) grows very fast. Many robots are manufactured, both from research institutes and from industry. However, until now, there have been no systems that can reliably map their environment in 3D under real-time conditions. The mobile robot Kurt3D was presented at RoboCup Rescue 2004 in Lisbon (Fig. 1). The robot is capable of mapping its environment in 3D and self localize in six degrees of freedom, i. e., considering its *x*, *y* and *z* positions and the roll, yaw and pitch angles (6D SLAM) in real-time.

Kurt3D's driving and mapping system consists of four major parts. First is a non linear feed forward PI controller. Second is the precise planar pose tracking algorithm HAYAI. Using 2D laser scans, features that correspond to natural landmarks are extracted and paired with features of previous scans. Third, a fast and reliable 3D scan matching procedure, employing a point reduction and approximate nearest neighbor search, is used to generate 3D maps and relocalize the robot. The basis of the 3D scan matching is the well know iterative closest points (ICP) algorithm. The fourth module is an interactive, semi-automatic control program that enables the operator to interfere with the mapping process for corrections. This paper focuses on Kurt3D's driving, mapping and localization modules and their interaction in the rescue context, devising the similarities between HAYAI and 6D SLAM.

1.1 Rescue Robotics Systems

Current rescue robots are mainly designed for searching for victims and paths through rubble

The paper is an extended version of the SSRR 2005 Best paper awarded paper "Mapping of Rescue Environments with Kurt3D" [18]

Figure 1 The mobile robot Kurt3D equipped with the 3D laser range finder as presented at RoboCup 2004. The scanners technical basis is a SICK 2D laser range finder (LMS-200).

that would be quicker to excavate, for structural inspection and for detection of hazardous material [7]. These robots are designed to go a bit deeper than traditional search equipment, i. e., cameras mounted on poles [7]. The actual operating range of current rescue robots is 5–20 m. The robots, e. g., the microtracs "micro-VGTV" and "Solem" [8], are small tanks that are connected with the operator by wire for transmitting a video signal. In fact, cameras are the only sensors, thus mapping the environment is basically impossible. Other rescue robot systems are based on tank-like chassis, too. Some prototype rescue robots are presented at the different RoboCup competitions where a tendency to enhance the sensor equipment can be observed, e. g., using 2D laser scanners, CO_2 sensors or infrared cameras. Another strategy is to improve the driving mobility with flexible tracked vehicles.

While mapping environments is a large research field in mobile robotics, only a little work has been done in the *automatic* mapping of rescue environments, as presented in [4].

1.2 3D Mapping

Instead of using 3D scanners which yield consistent 3D scans in the first place, some groups have attempted to build 3D volumetric representations of environments with 2D laser range finders. For instance, Thrun et al. [16] use two 2D laser range finders for acquiring 3D data. One laser scanner is mounted horizontally, the other vertically. The latter one grabs a vertical scan line which is transformed into 3D points based on the current robot pose. The horizontal scanner is used to compute the robot pose. The precision of 3D data points crucially depends on that pose and on the precision of the scanner. The same argument applies to the work of Wulf et al. who let the scanner rotate around the vertical axis while moving the robot [17].

A few other groups use high accurate, expensive 3D laser scanners [5; 12]. The RESOLV project aimed at modeling interiors for virtual reality and tele-presence [12]. They used a RIEGL laser range finder on robots and the ICP algorithm for scan matching [3]. The AVENUE project develops a robot for modeling urban environments [5], using a CYRAX laser scanner. Nevertheless, in their recent work they do not use data of the laser scanner in the robot control architecture for localization [5]. The research group of M. Hebert has reconstructed environments using the Zoller+Fröhlich laser scanner and aims to build 3D models without initial position estimates, i. e., without odometry information [6].

1.3 RoboCup Rescue

The goal of the urban search and rescue (USAR) robot competitions is to increase awareness of the challenges involved in search and rescue applications, provide objective evaluation of robotic implementations in representative environments, and promote collaboration between researchers. It requires robots to demonstrate their capabilities in mobility, sensory perception, planning, mapping, and practical operator interfaces while searching for simulated victims in unstructured environments. As robot teams begin demonstrate repeated successes against the obstacles posed in the arenas, the level of difficulty is increased accordingly so that the arenas provide a stepping-stone from the laboratory to the real world [9]. During the competition each discovered victim is scored with maximal 50 points (depending on the map quality and used sensors) whereas the three different type of arenas influence the score according to their difficulty with a scaling factor 0.5 (yellow), 0.75 (orange) or 1.0 (red). Each operator reduce the score by $1/(n+1)^2$, i. e., 1/4 for one operator and 1/9 for two operators.

2 The Exploration Robot Kurt3D

Kurt3D[1] (Fig. 1) is a mobile robot platform with a size of 45 cm (length) × 33 cm (width) × 26 cm (height) and a weight of 15.6 kg, both indoor as well as outdoor models exist. Equipped with the 3D laser range finder, the height increases to 47 cm and the weight increases to 22.6 kg. Two 90 W motors (short-term 200 W) are used to power the 6 wheels. Compared to the original Kurt3D robot platform, the outdoor version has larger wheels, where the middle ones are shifted outwards. Front and rear wheels have no tread pattern to enhance rotating. Kurt3D operates for about 4 hours with one battery charge (28 NiMH cells, capacity: 4500 mA h). The core of the robot is an Intel-Centrino-1400 MHz with 768 MB RAM and a Linux operating system. An embedded 16-bit CMOS

[1] Videos of explorations with Kurt3D can be found at: http://www.ais.fraunhofer.de/ARC/kurt3D/index.html

microcontroller is used to process commands to the motor. A CAN interface connects the laptop with the microcontroller.

The 3D laser range finder (Fig. 1) is built on basis of a 2D range finder by extension of a mount and a standard servo motor [13]. The 2D laser range finder is attached to the mount in the center of rotation for achieving a controlled pitch motion. The servo is connected on the left side (Fig. 1). The 3D laser scanner operates for up to 5 h (Scanner: 17 W, 20 NiMH cells with a capacity of 4500 mA h, Servo: 0.85 W, 4.5 V with batteries of 4500 mA h) on one battery pack. The area of 180°(h) × 90°(max. 120°)(v) is scanned with different horizontal (181, 361, 721) and vertical (128, 176, 256, 400, 500) resolutions. A plane with 181 data points is scanned in 13 ms by the 2D laser range finder (rotating mirror device). Planes with more data points, e. g., 361, 721, duplicate or quadruplicate this time. Thus a scan with 361 × 176 data points needs 4.5 seconds. In addition to the distance measurement the 3D laser range finder is capable of quantifying the amount of light returning to the scanner, resulting in a black/white reflectance image of the scene. Scanning the environment with a mobile robot is done in a stop-scan-go fashion.

Kurt3D is equipped with 2 × 4 super bright LEDs and two fluorescent tubes to illuminate the surroundings. The LEDs are attached to the two pan-and-tilt Logitech QuickCam 4000 cameras. Additional 8 NiMH cells are used to power the light.

3 The feed forward PI Controller

The controller of Kurt3D was originally developed for the high speed version of Kurt3D. Localizing a robot that drives with high speed (e. g., 4 m/s) does not only call for a very fast tracking algorithm to keep being localized. It also requires a vehicle that is able to drive safely with such velocities in an indoor environment, including the ability to react in real time to sudden changes of the surrounding, like opening doors or people walking around. The robot control implemented on Kurt3D consists of two parts, namely a fast motor controller and behaviors for set value computation.

Designing a control program for autonomous mobile robots consists of implementing a controller or set of behaviors including a set value computation for the controller. Hereby, the goal is to adjust the input signal to external disturbances, e. g., fluctuations of battery voltage and friction. Appropriate set values, e. g., the robot speed or turning velocity, are mapped by a controller or set of behaviors to the actors, i. e., its motors. The control cycle of Kurt3D's motors runs with 100 Hz.

In order to model the physical characteristics, different step responses and the speed of the motors with no-load and on carpet have been measured. Fig. 3 shows the step response to maximum velocity, following the characteristics of a PT_1 element. Fig. 4 presents the speed of the two motors when operating with no-load and on carpet moving straight forward. Every discrete PWM value from 0 up to 1023 is given to the system for one second, and the reaction, i. e., the speed at the end of the second, is measured. The system shows non-linearities in the upper speed range. The no-load measurement and the one on carpet differ only by a constant offset. The

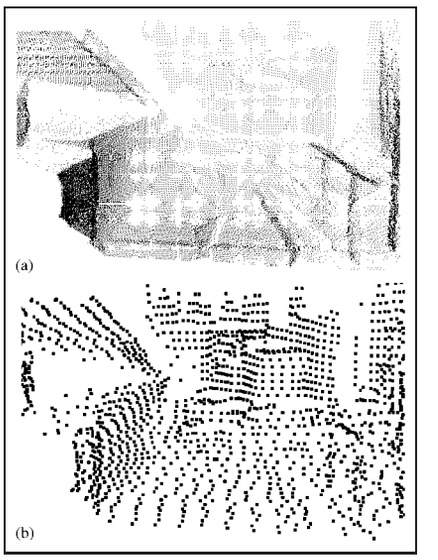

Figure 2 (a) A view of a 3D scene (66 785 3D data points). (b) Subsampled version (points have been enlarged), 6700 data points.

controller works under the assumption that this offset stays constant throughout the complete velocity interval, since it was not possible to measure the whole speed range on carpet due to physical limitations of the testing environment (the longest corridor was too short). The inverse of the lookup table (Fig. 4) together with the offset build the base for the open loop controller. A given speed and steering angle is scaled by two feed forward terms F_v and F_ω, (see Fig. 5) and mapped to a PWM value according to the lookup table and increased by the offset. This open loop controller is able to drive the robot but has a steady state error because of different loads and battery states and has also problems with small speed values.

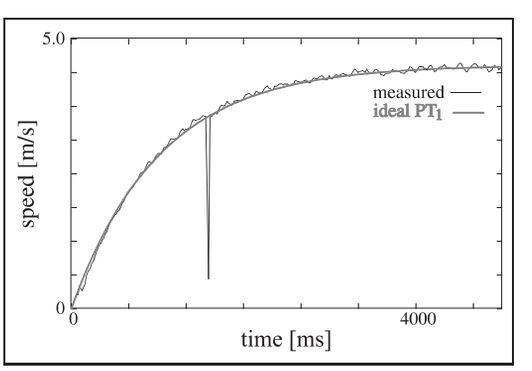

Figure 3 Step response to a velocity of 3.5 m/s (measured on carpet), approximating $PT_1 : b\left(a(1 - e^{-ax})\right)^{-1}$. The measurement error at $t = 1200$ is due to the odometry: We measure the ticks per second at the motor, not at the wheels; errors occur if the transmission belt slips.

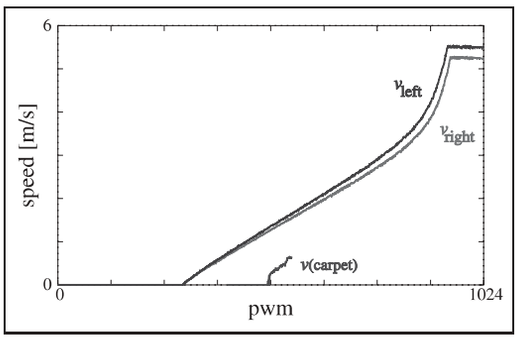

Figure 4 Speed of left and right wheels as a function of incoming PWM signals, measured with no-load. In comparison, the robot's speed when driving on carpet differs approximately by a constant offset.

The fast reaction time of the overall system and the reliable motor control is formed by closing the loop and adding a PI-term. By design, PI algorithms generate a correction term if and only if a control deviation exists. The P-term decreases the reaction time to changing set values, e. g., different speed values. The I-term avoids steady state errors but leads to a small overshot. The combination with the feed forward terms means that the part of the output values which is generated from the PI part is relative small, nearly zero for a constant input value. A D-term in the controller part is not necessary since changing speed values leads directly to changing output values according to the feed forward terms. Furthermore, the influence of noise and measurement errors of the current speed to the control output is reduced. In the case of an D-term this errors leads to oscillation, whereas in case of the feedforward controller the PI-terms have only a partly influence to the control output (compare Fig. 3 [11]). F_v is set to 1 according to the design of the lookup table and offset. An additional feed forward term is needed for the angular velocity ($F_\omega = \pi^{-1}$ cm) according to the large friction during turning and rotations of the robot. The value can be regarded as an offset from straight moving to rotating the robot. Finally, the motors receive the bandpass filtered PWM signals as input. Fig. 6 shows the overall system, containing the motor controllers for the left and right wheels as well as the state transformation formulas.

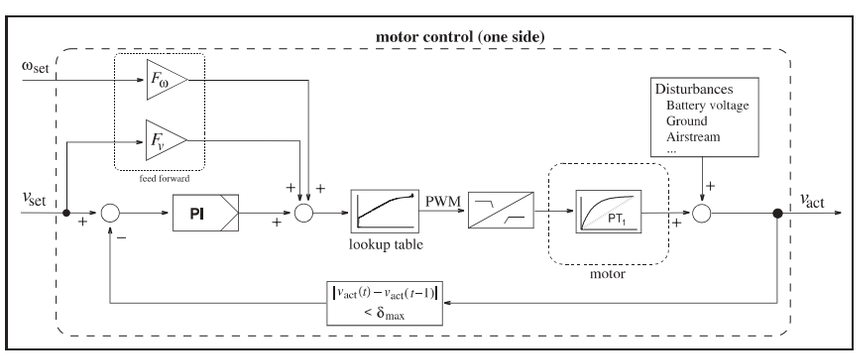

Figure 5 Control unit for one motor, based on a feed forward PI controller.

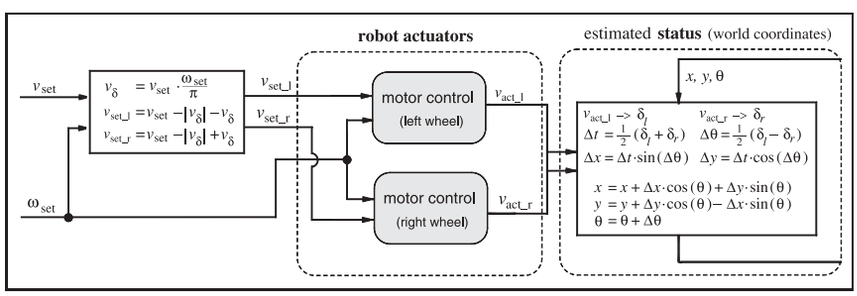

Figure 6 Schematic overview of the robot's control system, containing two motor controllers for the left and right wheel.

We have implemented three alternatives for the computation of the set values for Kurt3D. Depending on the application, tele-operated control by a human operator, fuzzy control for autonomous high-speed driving or globally stable robot control for driving precisely to specified coordinates is used. For the rescue application particularly the tele-operation mode with a joystick is important. The joystick signals are directly mapped to the reference velocity v_{set} and the reference turning velocity ω_{set}. In our experience, it is difficult for human operators to control Kurt3D manually beyond a speed of 1 m/s.

4 Pose Tracking with HAYAI

This section briefly describes the newly developed algorithm HAYAI (*Highspeed And Yet Accurate Indoor/outdoor-tracking*). The matching algorithm is based on the following scheme:

(1) Detect features within scan $\mathcal{R}$, yielding feature set M (*model set*). Likewise compute set D (*data set*) from a previous scan $\mathcal{S}$.
(2) Search for pairwise corresponding features from both sets, resulting in two subsets $\check{M} \subseteq M$ and $\check{D} \subseteq D$.
(3) Compute the pose shift $\Delta\boldsymbol{p} = (\Delta x, \Delta y, \Delta\theta)^T$ as the optimal transformation for mapping $\check{D}$ onto $\check{M}$.
(4) Update the robot's pose $\boldsymbol{p}_n \xrightarrow{\Delta\boldsymbol{p}} \boldsymbol{p}_{n+1}$
(5) Save the current scan as new reference scan $\mathcal{R} \leftarrow \mathcal{S}$.

4.1 Data Filtering

Scanning is noisy and small errors may occur, namely Gaussian noise and salt and pepper noise. The latter one arises for example

at edges where the laser beam of the scanner hits two surfaces, resulting in a mean and erroneous data value. Furthermore reflections, e. g., at glass surfaces, lead to suspicious data. We propose two fast filtering methods to modify the data in order to enhance the quality of each scan, typically containing 181 data points. The data reduction, used for reducing Gaussian noise, works as follows: The scanner emits the laser beams in a spherical way, such that the data points close to the source are more dense. Multiple data points located close together are joined into one point. The number of these so-called *reduced points* is one order of magnitude smaller than the original one. For eliminating salt and pepper noise, a median filter removes the outliers by replacing a data point with the median value of the n surrounding points (here: $n = 7$). The neighbor points are determined according to their index within the scan, since the laser scanner provides the data sorted in a counter-clockwise direction. The median value is calculated with regard to the Euclidian distance of the data points to the point of origin. In order to remove noisy data but leave the remaining scan points untouched, the filtering algorithm replaces a data point with the corresponding median value if and only if the Euclidian distance between both is larger than a fixed threshold (e. g., 200 cm).

4.2 Extraction and Matching of Features

As described above, the scan matching algorithm computes a transformation $\Delta \boldsymbol{p}$ such that a *set of features*, extracted from the first scan, is mapped optimally to a feature set of the second scan. In order to be usable for a pose tracking algorithm, these features have to fulfill two requirements: First, they have to be *invariant* with respect to rotation and translation. Second, they have to be *efficiently* computable in order to satisfy real time constraints.

Using the inherent order of the scan data allows the application of linear filters for a fast and reliable feature detection. HAYAI chooses *extrema* in the polar representation of a scan as natural landmarks. These extrema correlate to corners and jump edges in Cartesian space. The usage of polar coordinates implicates a reduction by one dimension, since all operations deployed for feature extraction are fast linear filters, operating on the sequence $(r_i)_{i\in\mathbb{N}}$ of range values of a scan $S = \big((\varphi_i, r_i)\big)_{i=1,\ldots,N}$.

Given a one dimensional filter $\Psi = [\psi_{-1}, \psi_0, \psi_{+1}]$, the filtered value r_i^{Ψ} of a scan point r_i $(i = 2, \ldots, N-1)$ is defined as $r_i^{\Psi} = \sum_{k=-1}^{1} \psi_k r_{i+k}$. For feature detection, the scan signal is filtered as follows:

(1) Sharpen the data in order to emphasize the significant parts of the scan, i. e., the extrema, without modifying the residual scan, by applying a sharpen filter of the form $\Psi_1 = [-1, 4, -1]$.
(2) Compute the derivation signal by using a gradient filter $\Psi_2 = [-\frac{1}{2}, 0, \frac{1}{2}]$.
(3) Smooth the gradient signal to simplify the detection of zero crossings with a soften filter $\Psi_3 = [1, 1, 1]$.

The diagram on the left of Fig. 7 illustrates the effects of the used filters.

After generating the sets of features M, D from both scans, a matching between both sets has to be calculated. Instead of solving the hard optimization problem of searching for an optimal match, we use a heuristic approach, utilizing inherent knowledge about the problem of matching features, e. g., the fact that the features' topology cannot change fundamentally from one scan to the following. The basic aim is to build a matrix of possible matching pairs, based on an error function defining the distance between two points $\boldsymbol{m}_i, \boldsymbol{d}_j$, with $\boldsymbol{m}_i = (m_i^x, m_i^y)^T$ in Cartesian, or $(m_i^{\varphi}, m_i^r)^T$ in polar coordinates, respectively. ($\boldsymbol{d}_j$ analogously). The resulting matrix $w_{i,j}$ denoting feature correspondences is simplified until the match is non-ambiguous. Fig. 7 shows the match of two scans.

4.3 Pose Calculation

Given two sets of features

$$\check{M} = \{\, \boldsymbol{m}_i \mid \boldsymbol{m}_i \in \mathbb{R}^2,\ i = 1, \ldots, N_m \}$$

and

$$\check{D} = \left\{\, \boldsymbol{d}_i \;\middle|\; \boldsymbol{d}_i \in \mathbb{R}^2,\ i = 1, \ldots, N_d \right\},$$

the calculation of the optimal transformation for mapping D onto M is an optimization problem of the error function $E(\boldsymbol{R}, \boldsymbol{t})$ which is given by

$$\sum_{i=1}^{N_m} \sum_{j=1}^{N_d} w_{i,j} ||\boldsymbol{m}_i - (\boldsymbol{R}\boldsymbol{d}_j + \boldsymbol{t})||^2 \qquad (1)$$

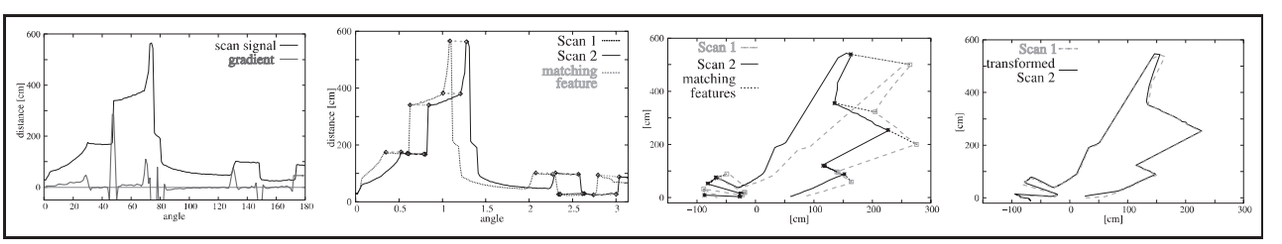

Figure 7 From left to right: (1) Application of the feature detection filters. (2) and (3) Pairing of corresponding features. (φ, r) representation (2) vs. the euclidian one (3). (4) Transformed scan.

$$\propto \frac{1}{N}\sum_{i=1}^{N} ||\boldsymbol{m}_i - (\boldsymbol{R}\boldsymbol{d}_i + \boldsymbol{t})||^2, \tag{2}$$

since the matching is non-ambiguous. The first step of the computation is to decouple the calculation of the rotation $\boldsymbol{R}$ from the translation $\boldsymbol{t}$ using the centroids of the points belonging to the matching:

$$\boldsymbol{c}_m = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{m}_i, \boldsymbol{c}_d = \frac{1}{N}\sum_{i=1}^{N} \boldsymbol{d}_i$$

and

$$M' = \{\boldsymbol{m}'_i = \boldsymbol{m}_i - \boldsymbol{c}_m\}_{1,\ldots,N},$$
$$D' = \{\boldsymbol{d}'_i - \boldsymbol{d}_i - \boldsymbol{c}_d\}_{1,\ldots,N}.$$

In order to minimize the sum $E(\boldsymbol{R}, \boldsymbol{t})$, all terms have to be minimized. By solving the equation

$$\frac{\partial}{\partial \Delta\theta} E(\boldsymbol{R}_{\Delta\theta}) = 0$$

for a 2D rotation $\boldsymbol{R}_{\Delta\theta} = \boldsymbol{R}$, the optimal rotation $\Delta\theta$ is calculated as

$$\arctan\left(\frac{\sum_{i=1}^{N}\left(m'^x_i d'^x_i + m'^y_i d'^y_i\right)}{\sum_{i=1}^{N}\left(m'^y_i d'^x_i - m'^x_i d'^y_i\right)}\right).$$

With given rotation, the translation $\Delta\boldsymbol{t} = (\Delta x, \Delta y)^T$ is calculated as

$$\boldsymbol{c}_m - \underbrace{\begin{pmatrix} \cos\Delta\theta & \sin\Delta\theta \\ -\sin\Delta\theta & \cos\Delta\theta \end{pmatrix}}_{=\boldsymbol{R}_{\Delta\theta}} \cdot \boldsymbol{c}_d.$$

5 3D Mapping and 6D Robot Relocalization

Multiple 3D scans are necessary to digitalize environments without occlusions. To create a correct and consistent model, the scans have to be merged into one coordinate system. This process is called registration. If the localization of the robot with the scanner were precise, the registration could be done directly based on the robot pose. However, relative self localization is erroneous, even with HAYAI, so the geometric structure of overlapping 3D scans has to be considered for registration.

5.1 6D Registration of 3D Scans

The following method for registration of point sets is part of many publications, so only a brief summary is given here [14]. The complete algorithm was invented in 1992 and can be found, e. g., in [3]. The method is called *Iterative Closest Points (ICP) algorithm*. The procedure considers all six degrees of freedom, i. e., the roll, yaw and pitch orientation and the x, y, and z position of the robot

Given two independently acquired sets of 3D points, M ($|M| = N_m$) and D ($|D| = N_d$), which correspond to a single shape, we aim to find the transformation consisting of a rotation $\boldsymbol{R}$ and a translation $\boldsymbol{t}$ which minimizes the cost function shown in Eq. (1). Note that this time the vectors are in 3D space and $\boldsymbol{R}$ has to be an orthonormal 3×3 matrix. Now, $w_{i,j}$ is assigned 1 if the i-th point of M describes the same point in space as the j-th point of D. Otherwise $w_{i,j}$ is 0. Two things have to be calculated: First, the corresponding points, and second, the transformation $(\boldsymbol{R}, \boldsymbol{t})$ that minimizes $E(\boldsymbol{R}, \boldsymbol{t})$ on the base of the corresponding points.

The ICP algorithm calculates iteratively the point correspondences. In each iteration step, the algorithm selects the closest points as correspondences and calculates the transformation $(\boldsymbol{R}, \boldsymbol{t})$ for minimizing the function given in Eq. (1). The assumption is that in the last iteration step the point correspondences are correct. In every iteration the optimal transformation $(\boldsymbol{R}, \boldsymbol{t})$ has to be computed. Like before, Eq. (1) can be reduced to Eq. (2). The difficulty of the minimization problem is to enforce the orthonormality of matrix $\boldsymbol{R}$. The following method, first published by Arun et al., is based on singular value decomposition (SVD) [1]. It is robust and easy to implement, thus we give a brief overview here:

The conversion of Eq. (1) to Eq. (2) holds in 3D space, too. The algorithm computes the optimal rotation by $\boldsymbol{R} = \boldsymbol{V}\boldsymbol{U}^T$. Hereby the matrices $\boldsymbol{V}$ and $\boldsymbol{U}$ are derived by the singular value decomposition $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\ \boldsymbol{V}^T$ of a correlation matrix $\boldsymbol{H}$. This (3×3) matrix $\boldsymbol{H}$ is given by

$$\boldsymbol{H} = \sum_{i=1}^{N} {m'_i}^T d'_i = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix} \tag{3}$$

with $S_{xx} = \sum_{i=1}^{N} m'_{ix} d'_{ix}$, $S_{xy} = \sum_{i=1}^{N} m'_{ix} d'_{iy}, \ldots$.

Since rotations are length preserving, i. e., $||\boldsymbol{R}\boldsymbol{d}'_i||^2 = ||\boldsymbol{d}'_i||^2$, the error function $E(\boldsymbol{R}, \boldsymbol{t})$ is expanded to

$$\sum_{i=1}^{N} \left\|\boldsymbol{m}'_i\right\|^2 - 2\sum_{i=1}^{N} \boldsymbol{m}'_i \cdot \boldsymbol{R}\boldsymbol{d}'_i + \sum_{i=1}^{N} \left\|\boldsymbol{d}'_i\right\|^2.$$

The rotation affects only the middle term, thus it is sufficient to maximize

$$\sum_{i=1}^{N} \boldsymbol{m}'_i \cdot \boldsymbol{R}\boldsymbol{d}'_i = \sum_{i=1}^{N} {\boldsymbol{m}'_i}^T \boldsymbol{R}\boldsymbol{d}'_i. \tag{4}$$

Using the trace of a matrix, Eq. (4) can be rewritten to obtain

$$\mathrm{tr}\left(\sum_{i=1}^{N} \boldsymbol{R}\boldsymbol{d}'_i {\boldsymbol{m}'_i}^T\right) = \mathrm{tr}\left(\boldsymbol{R}\boldsymbol{H}\right).$$

Hereby, matrix $\boldsymbol{H}$ has to be defined as in Eq. (3). Now we have to find the matrix $\boldsymbol{R}$ that maximizes $\mathrm{tr}\,(\boldsymbol{R}\boldsymbol{H})$. Assume that the singular value decomposition of $\boldsymbol{H}$ is $\boldsymbol{H} = \boldsymbol{U}\boldsymbol{\Lambda}\ \boldsymbol{V}^T$, with $\boldsymbol{U}$ and $\boldsymbol{V}$ orthonormal 3×3 matrices and $\boldsymbol{\Lambda}$ a 3×3 diagonal matrix without negative elements. Suppose $\boldsymbol{R} = \boldsymbol{V}\boldsymbol{U}^T$. Then, $\boldsymbol{R}$ is orthonormal and

$$\boldsymbol{R}\boldsymbol{H} = \boldsymbol{V}\boldsymbol{U}^T\boldsymbol{U}\boldsymbol{\Lambda}\ \boldsymbol{V}^T = \boldsymbol{V}\boldsymbol{\Lambda}\ \boldsymbol{V}^T$$

is a symmetric, positive definite matrix. Arun, Huang and Blostein provide a lemma to show that

$$\mathrm{tr}\left(\boldsymbol{R}\boldsymbol{H}\right) \geq \mathrm{tr}\left(\boldsymbol{B}\boldsymbol{R}\boldsymbol{H}\right),$$

for any orthonormal matrix $\boldsymbol{B}$. Therefore the matrix $\boldsymbol{R}$ is optimal. Proving the lemma is straightforward, using the inequation of

Chauchy-Schwarz [1]. The optimal translation is calculated as: $\boldsymbol{t} = \boldsymbol{c}_m - \boldsymbol{R}\boldsymbol{c}_d$.

5.2 ICP-based 6D SLAM

To digitalize environments, multiple 3D scans have to be registered. After registration, the scene has to be globally consistent. A straightforward method for aligning several 3D scans is *pairwise matching*, i. e., the new scan is registered against the scan with the largest overlapping areas. The latter one is determined in a preprocessing step. Alternatively, *incremental matching* could be used, i. e., the new scan is registered against a so-called *metascan*, which is the union of the previously acquired and registered scans. Each scan matching has a limited precision. Both methods accumulate the registration errors such that the registration of a large number of 3D scans leads to inconsistent scenes and to problems with the robot localization.

After matching multiple 3D scans, errors have accumulated and a closed loop will be inconsistent. Our 6D SLAM algorithm detects a closing loop by registering the last acquired 3D scan with earlier acquired scans. If a registration is possible, the computed error is distributed over all 3D scans. A second step minimizes the global error. The registration of one scan is followed by registration of all neighboring scans, such that the error is minimized. In an iterative fashion a consistent model is produced. Details of the full algorithm can be found in [15].

5.3 ICP Speedups

The computational requirements are reduced by two methods: First we reduce the 3D data, i. e., we compute point clouds that approximate the scanned 3D surface and contain only a small fraction of the 3D point cloud. Second is the fast approximation of the closest point with *k*d-trees for the ICP algorithm.

Data reduction for the ICP algorithm is done using the proposed filters of Section 4.1. Without filtering, a few outliers may lead to multiple wrong point pairs during the 3D matching phase and results in an incorrect 3D scan alignment. Reduction and filtering are done in every single 2D scan slice while scanning, they are implemented as online algorithms and run in parallel to the 3D scan acquisition. In the end, the data for the scan matching are collected from every third scan slice. This fast vertical reduction yields a good surface description.

*k*D-trees are a generalization of binary search trees. Every node represents a partition of a point set to the two successor nodes. The root represents the whole point cloud and the leaves form a disjunct partition of the set. These leaves are called buckets. Furthermore, every node contains the limits of the represented point set. Searching in *k*d-trees is done recursively. For a given 3D point $\boldsymbol{p}_q$, a comparison with the separating plane has to be performed in order to decide on which side the search must continue. This procedure is executed until the leaves are reached. There, the algorithm has to evaluate all bucket points. However, the closest point may be in a different bucket, if and only if the distance to the limits is smaller than the one to the closest point in the bucket. In this case backtracking has to be performed (Fig. 8).

Arya et al. introduce the following notion for approximating the nearest neighbor [2]: Given an $\varepsilon > 0$, then the point $\boldsymbol{p} \in D$ is the $(1+\varepsilon)$-approximate nearest neighbor of the point $\boldsymbol{p}_q$ if and only if $\|\boldsymbol{p} - \boldsymbol{q}\| \leq (1+\varepsilon)\|\boldsymbol{p}^* - \boldsymbol{q}\|$, whereas $\boldsymbol{p}^*$ denotes the true nearest neighbor, i. e., p is within a relative error of ε of the true nearest neighbor. In every step the algorithm records the closest point $\boldsymbol{p}$; the search finishes if the distance to the unanalyzed leaves is larger than $\|\boldsymbol{p}_q - \boldsymbol{p}\|/(1+\varepsilon)$. Fig. 9 shows an example where the gray cell doesn't have to be analyzed, since the

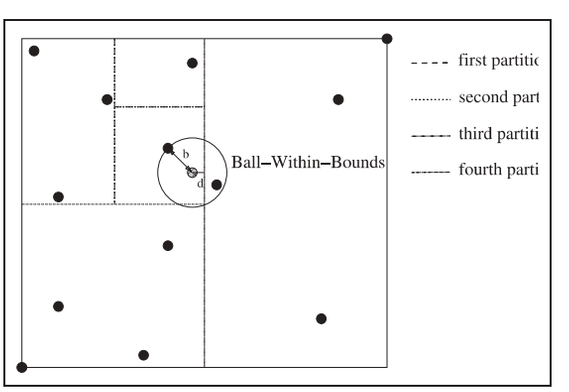

Figure 8 Construction of a *k*d-tree.

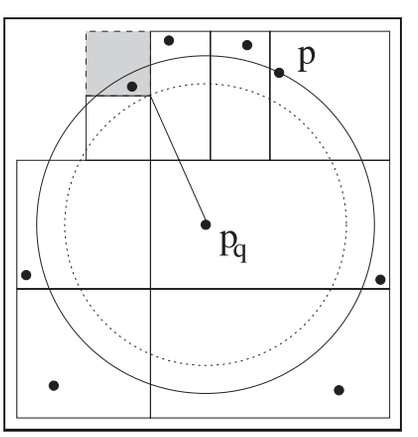

Figure 9 The $(1+\varepsilon)$-approximate nearest neighbor. The search algorithm doesn't have to analyze the gray cell, since the point $\boldsymbol{p}$ satisfies the approximation criterion.

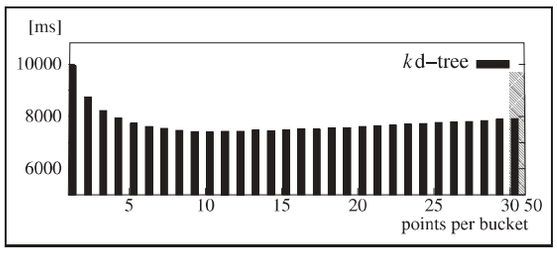

Figure 10 Run time of the ICP algorithm using *k*d-trees with different bucket sizes. The minimal time is reached for 10 points per bucket.

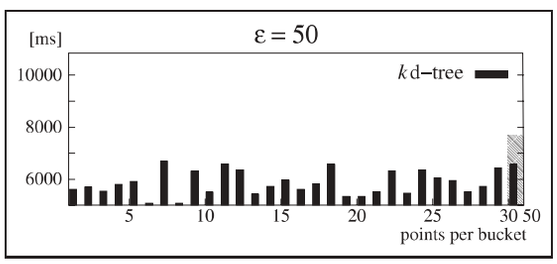

Figure 11 Computing time for Approximate *k*d-tree search.

point $\boldsymbol{p}$ satisfies the approximation criterion. Fig. 10 and Fig. 11 show the computation time for matching two 3D scans using *k*d-trees and approximate *k*d-trees with $\varepsilon = 50$, respectively, for different bucket sizes. In [10] we show that the quality of the scan matching is not affected by the approximation, due to the large number of points and the iterative fashion of the 3D scan matching.

6 The Software Architecture

Fig. 12 sketches the software architecture. It is a client-server architecture where client and server are connected by wireless LAN. On the robot's side a 100 Hz control loop is running, adjusting the motor velocities. In this loop odometry and HAYAI are processed and merged with a Kalman filter. The processing time of HAYAI is around 3% of the CPU load. Thus there is enough time to compress the camera images to JPEQ and to transmit the data.

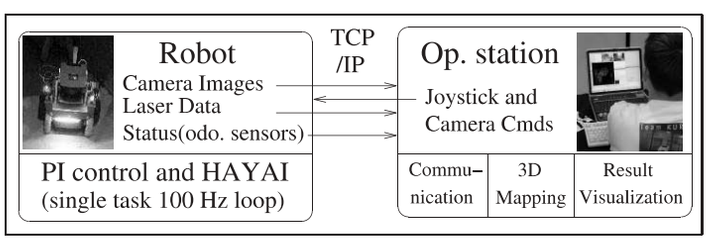

Figure 12 System Overview: The robot runs a 100 Hz loop for motor control, scan processing, image encoding and transmission. The Kurt3D server executes three threads that are responsible for communicating, 3D mapping and result visualization.

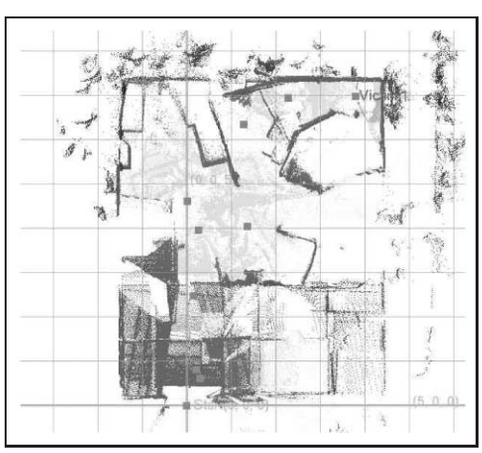

Figure 13 A 3D map of the rescue arena during RoboCup 2004 as a point cloud (top view). The points on the ground have been colored in light grey. The 3D scan positions and a found victim are marked.

At the operator station the robot is teleoperated. Three processes are executed in parallel: The communication between remote joystick control and robot, the 3D mapping, and the interactive, OpenGL-based map viewer. The latter one enables the operator to intervene in the map generation process, e. g., manually correcting the initial 6D pose of an acquired 3D scan and restarting the matching algorithm. During competition, the operator tried to minimize the number of acquired 3D scans to save time for victim detection. The virtual camera pose in the 3D mapping window is freely adjustable, though a view from top is used by default, since it provides the best situation awareness.

7 Results/Conclusions

The proposed algorithms have been evaluated at RoboCup Rescue 2004 in Lisbon. Fig. 13 shows an online generated 3D map (top view). Two 3D views are given in Fig. 14. An offline-rendered animation of the acquired data can be found at http://www.ais.fhg.de/ARC/kurt3D/rr.html.

Based on the principles of 3D scan matching, we have designed a SLAM algorithm in six dimensions using loop closing and global error minimization [10]. This global error minimization is currently too slow to be deployed on a rescue system. The computing time of this relaxation algorithm is in the order of several minutes for a typical rescue arena, thus in the designated scenario we use pure 3D scan matching for mapping. Since 3D scans potentially provide a lot of information and the area is small, i. e., 6 m by 6 m, scan matching is sufficient for mapping.

This paper has presented a mobile robotic system for teleoperated 3D mapping of environments. The mapping is done by means

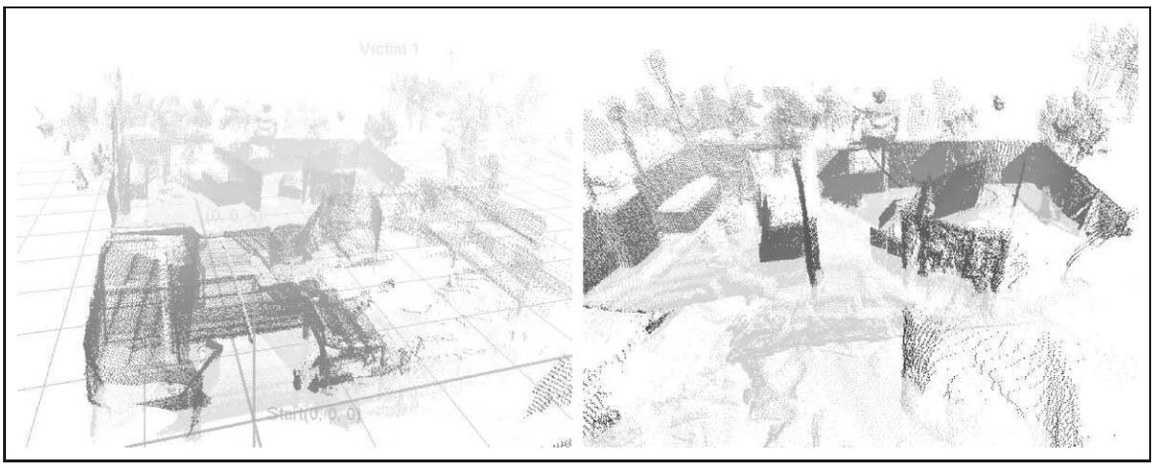

Figure 14 A 3D view of the map of Fig. 13 rendered from a pose slightly above the arena. A 1 m^2 grid is superimposed in the first map.

of a fast pose tracking algorithm and 3D scan matching. We have demonstrated our 3D mapping capabilities at the RoboCup Rescue 2004 in Lisbon, were our team won the 2nd prize. The aim of future work is combining the mapping algorithms with mechatronic robotic systems, i. e., building a robot system that can actually go into the third dimension and can cope with the red arena in RoboCup Rescue. Furthermore, we concentrate on enhancing the system's autonomy: In addition to automatic mapping, autonomous driving and exploration are planned.

Acknowledgements

Special thanks to Rainer Worst, Kiran Raj Tiruchinapalli and Thomas Christaller for supporting our work.

References

[1] K.S. Arun, T.S. Huang, and S.D. Blostein. Least square fitting of two 3-d point sets. *IEEE rRans. on Pattern Analysis and Machine Intelligence*, 9(5):698–700, 1987.

[2] S. Arya and D.M. Mount. Approximate nearest neigbor queries in fixed dimensions. In *Proc. of the 4th ACM-SIAM Symp. on Discrete Algorithms*, pages 271–280, 1993.

[3] P. Besl and N. McKay. A method for Registration of 3-D Shapes. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb 1992.

[4] St. Carpin, H. Kenn, and A. Birk. Autonomous Mapping in the Real Robots Rescue League. In *Proc. of RoboCup 2003: Robot Soccer World Cup VII*, 2004.

[5] A. Georgiev and P.K. Allen. Localization methods for a mobile robot in urban environments. *IEEE Trans. on Robotics and Automation (TRO)*, 20(5):851–864, Oct 2004.

[6] M. Hebert, M. Deans, D. Huber, B. Nabbe, and N. Vandapel. Progress in 3-D Mapping and Localization. In *Proc. of the 9th Int'l Symp. on Intelligent Robotic Systems, (SIRS '01)*, Toulouse, France, July 2001.

[7] R.R. Murphy. Activities of the rescue robots at the world trade center from 11–21 sep 2001. *IEEE Robotics & Automation Magazine*, 11(3):851–864, Sep 2004.

[8] R.R. Murphy. Rescue robotics for homeland security. *CACM, Special Issue on Homeland Security*, 27(3):66–69, Mar 2004.

[9] NIST. Urban search and rescue robot competions. *http://www.isd.mel.nist.gov/projects/USAR/competitions.htm*, May 2005.

[10] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann. 6D SLAM with Approximate Data Associoation. In *Proc. of the 12th Int'l Conf. on Advanced Robotics (ICAR '05)*, (submitted).

[11] K. Ogata. *Modern Control Engineering (4th Edition)*. Prentice Hall, Nov 2001.

[12] V. Sequeira, K. Ng, E. Wolfart, J. Goncalves, and D. Hogg. Automated 3D reconstruction of interiors with multiple scan–views. In *Proc. of SPIE, Electronic Imaging '99, The Society for Imaging Science and Technology/SPIE's 11th Annual Symp.*, San Jose, CA, USA, Jan 1999.

[13] H. Surmann, K. Lingemann, A. Nüchter, and J. Hertzberg. A 3D laser range finder for autonomous mobile robots. In *Proc. of the 32nd Int'l Symp. on Robotics (ISR '01)*, pages 153–158, Seoul, Korea, Apr 2001.

[14] H. Surmann, A. Nüchter, and J. Hertzberg. An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor en vironments. *Robotics and Autonomous Systems*, 45(3–4):181–198, Dec 2003.

[15] H. Surmann, A. Nüchter, K. Lingemann, and J. Hertzberg. 6D SLAM A Preliminary Report on Closing the Loop in Six Dimensions. In *Proc. of the 5th IFAC Symp. on Intelligent Autonomous Vehicles (IAV '04)*, Lisabon, Portugal, July 2004.

[16] S. Thrun, D. Fox, and W. Burgard. A real-time algorithm for mobile robot mapping with application to multi robot and 3D mapping. In *Proc. of the IEEE Int'l Conf. on Robotics and Automation (ICRA '00)*, San Francisco, CA, USA, Apr 2000.

[17] O. Wulf, K.O. Arras, H.I. Christensen, and B.A. Wagner. 2D Mapping of Cluttered Indoor Environments by Means of 3D Perception. In *Proc. of the IEEE Int'l Conf. on Robotics and Automation (ICRA '04)*, pages 4204–4209, New Orleans, USA, Apr 2004.

[18] A. Nüchter, K. Lingemann, J. Hertzberg, H. Surmann, K. Pervölz, M. Hennig, K.R. Tiruchinapalli, R. Worst, and T. Christaller, Mapping of Rescue Environments with Kurt3D, in *Proceedings of the International Workshop on Safety, Security and Rescue Robotics (SSRR '05)*, pages 158–163, Kobe, Japan, June 2005, (best paper award).


1


2


3


4


5


6

1 Dr.-Ing. Dipl.-Inf. Hartmut Surmann heads a research group within the Fraunhofer Institute for Autonomous Intelligent Systems (AIS). He received his diploma in computer science from the Department of Computer Science, University of Dortmund and his PhD in electrical engineering from the Faculty of Electrical Engineering, University of Dortmund, Germany, in 1989 and 1994, respectively. His primary research interests are autonomous mobile robotics and computational intelligence. He has published several papers in the area of design, architecture, hardware, sensors and application of mobile robotics and fuzzy logic controllers. He is a member of the GI and VDE and received several awards, e. g., the FUZZ-IEEE/IFES'95 robot intelligence award and the PhD award for his thesis from the German AI institutes in 1996. He gives lectures in control theory and robotics at FH Bonn-Rhein-Sieg.
Address: Fraunhofer AIS, Schloss Birlinghoven, 53754 Sankt Augustin, Germany, Tel.: +49-2241-142518, Fax: +49-2241-1442518, E-Mail: hartmut.surmann@ais.fraunhofer.de

2 Dipl.-Ing. Kai Pervölz is doing a master by research course in autonomous systems at the University of Applied Sciences Bonn-Rhein-Sieg (Germany) in cooperation with the Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Sankt Augustin, Germany). He received the diploma degree in mechatronics from the Aachen University of Applied Sciences in 2003. His primary research interests are autonomous robotics, 3D data analysis, and image processing.
Address: Fraunhofer AIS, Schloss Birlinghoven, 53754 Sankt Augustin, Germany, Tel.: +49-2241-141011, Fax: +49-2241-142342, E-Mail: kai@pervoelz.de

3 Dipl.-Inf. Andreas Nüchter is a research associate at University of Osnabrück and a PhD student at University of Bonn. Past affiliations were with the Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Sankt Augustin) and University of Bonn. He received the diploma degree in computer science from University of Bonn in 2002 and was awarded with the best paper award by the German Society of Informatics (GI) for his thesis. His research interests include reliable robot control, 3D vision, and laser scanning technologies. He is a member of the GI and the IEEE.
Address: University of Osnabrück, Institute for Computer Science, Albrechtstraße 28, 49069 Osnabrück, Germany, E-Mail: nuechter@informatik.uni-osnabrueck.de

4 Dipl.-Inf. Kai Lingemann is a research associate and PhD student at University of Osnabrück. Past affiliations were with the Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Sankt Augustin) (Germany), University of Bonn (Germany) and Kyoto University (Japan). He received the diploma degree in computer science from University of Bonn in January 2004. His research interests include mobile robot localization, reliable robot control, 3D vision, and laser scanning technologies.
Address: University of Osnabrück, Institute for Computer Science, Albrechtstraße 28, 49069 Osnabrück, Germany, E-Mail: lingemann@informatik.uni-osnabrueck.de

5 Prof. Dr. rer. nat. Joachim Hertzberg is a full professor for computer science at the University of Osnabrück, where he is heading the Knowledge-Based Systems research group. He holds a diploma and a doctorate degree in Computer Science, both from the University of Bonn. Past affiliations and stays were with the Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Sankt Augustin), with the Universities of Bonn (Germany), Dortmund (Germany), and Auckland (New Zealand), and with the International Computer Science Institute (ICSI), Berkeley, CA. His area of research is artificial intelligence, where he has contributed to action planning, robot localization and mapping, plan-based robot control, robot control architectures, temporal reasoning, logical reasoning about action and change, constraint-based reasoning, and applications of these techniques. He is currently a member of the Executive Council of the International Conference on Automated Planning and Scheduling (ICAPS).
Address: University of Osnabrück, Institute for Computer Science, Albrechtstraße 28, 49069 Osnabrück, Germany, E-Mail: hertzberg@informatik.uni-osnabrueck.de

6 Dipl.-Wirtsch.-Ing. Matthias Hennig is a research associate and PhD student at the Dresden University of Technology. Past affiliations were with the Artificial Intelligence Institute (Department of Computer Science, TU-Dresden) and Fraunhofer Institute for Autonomous Intelligent Systems (AIS, Sankt Augustin). He received the diploma degree in business and engineering from the Dresden University of Technology in 2005. His primary research interests are autonomous robotics, reliable robot control and image processing.
Address: Institute of Automation, Mobile Robotics Group, Dresden University of Technology, 01062 Dresden, Germany, Tel.: +49-351-463-35391, E-Mail: mhennig@ifa.et.tu-dresden.de