The 2009 IEEE/RSJ International Conference on
Intelligent Robots and Systems
October 11-15, 2009 St. Louis, USA

Fast 3D Mapping by Matching Planes Extracted from Range Sensor Point-Clouds

Kaustubh Pathak, Narunas Vaskevicius, Jann Poppinga, Max Pfingsthorn, Sören Schwertfeger, Andreas Birk
Dept. of Computer Science, Jacobs University Bremen, 28759 Bremen, Germany
k.pathak@jacobs-university.de, a.birk@jacobs-university.de

***Abstract*—This article addresses fast 3D mapping by a mobile robot in a predominantly planar environment. It is based on a novel pose registration algorithm based entirely on matching features composed of plane-segments extracted from point-clouds sampled from a 3D sensor. The approach has advantages in terms of robustness, speed and storage as compared to the voxel based approaches. Unlike previous approaches, the uncertainty in plane parameters is utilized to compute the uncertainty in the pose computed by scan-registration. The algorithm is illustrated by creating a full 3D model of a multi-level robot testing arena.**

I. Introduction

Encouraged by the success of 2D simultaneous localization and mapping (SLAM) [1], there have been many recent attempts (ref. [2], [3], [4], [5], [6], [7]) to extend the methodology to 3D. Since onboard 3D odometry is usually lacking or inadequate, an essential part of the mapping procedure is finding the relative pose offset of the robot between two successive range sensor samples using scan-matching. Typical 3D sensors can be categorized as: 1) large field-of-view (FOV) laser range finders (LRF) mounted on a rotating platforms [3] having a large scanning time of around a minute, and 2) time-of-flight (TOF) sensors like the Swiss-ranger [8] and PMD [9] with a much restricted FOV, but being able to provide several scans per second. These sensors provide a 3D range scan as a point-cloud of noisy spatial coordinates, numbering typically between 10^4 and 10^6.

The scan-matching algorithms can be broadly classified as 1) point-based, e.g. iterative closest point (ICP) [10], [7], or voxel-based 3D-NDT [6], and 2) feature-based— where the most popular geometric features are surfaces represented as planar-patches [11], [12], [5], [13], [14]. Although point/voxel based methods do not assume any structure in the environments, they are computationally expensive and converge to local minima when the scene overlap is not high, and sensitivity to outliers [15]. Maps created by these methods are neither easily visualized nor are particularly suitable for 3D path-planning. Planar patches offer good visualization and an impressive compression of data (typically 95%) as compared to point-clouds. As shown in this paper, the feature-correspondence problem can also be solved faster and more robustly due to this compression. As in [5], the approach in our article is also based on plane-features. However, our approach obviates the ICP step necessary for the pose change prediction in their work. The basic steps of the algorithms are compared in Fig. 1. Combining plane-correspondence determination and pose registration in one step and removal of ICP leads to savings in computation time and an increase in robustness. In fact, most of the previous 3D mapping approaches have been off-line due to the excessive time required by ICP. With the new approach presented in this article, the pose registration can be computed "fast." This claim can be made more precise as follows:

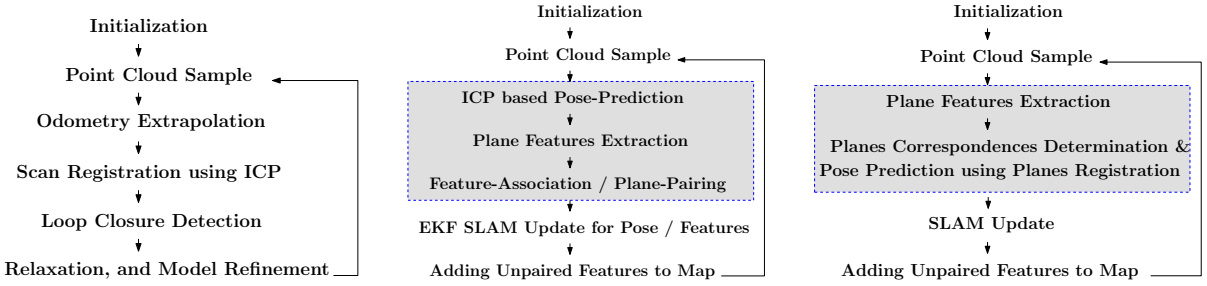
1) For large FOV ($\sim 270^\circ \times 180^\circ$) actuated LRFs, the pose-registration time of $4-10$ sec. with planes on an AMD Turion 1.6 GHz computer is much less than the typical time of 30 sec. needed to collect one scan.
2) For small FOV ($\sim 50^\circ \times 50^\circ$) TOF sensors, the pose-registration time is typically less than 0.1 sec.

In this article, we focus only on the steps marked with a box in Fig. 1. The well known EKF-SLAM step can be the same as in [5] and is outside the scope of this article. A plane extraction step is also common in both approaches and it will not be discussed in detail in this article. The reader is referred to the authors' previous work [16], [17] for details.

The work by Kohlhepp et al [2], [19], [13] is also closely related to ours. In [19], surfaces are extracted from range images obtained by a rotating laser range finder (LRF) and registered together. A local module determines the correspondences and computes transforms, and a global module detects loop-closure and distributes uncertainty using an elastic graph. For this article, only the local module is relevant which is discussed again in [13]. Similar to Sec. II-D1 in this article, their approach uses feature-directions to compute rotation between successive views. However, no mention is made in their work of using the uncertainties in the extracted feature-parameters. For estimating translation, they resort back to point features, which is essentially the same as ICP. By contrast, this article uses only plane-features and does not return to the domain of point-features – even translation is obtained using planes. This allows for detection of dominant directions of uncertainties, as detailed in Sec. II-D2. In [13], many heuristic measures for estimating correspondences between features, in particular the ground, across views are discussed. By contrast, the plane correspondence algorithm presented by us does not give any special status to the ground.

The feature-correspondence finding problem has been addressed in the past using graph-theoretic techniques [20], [21]. However, the latter work offered scant experimental data in form of only one pair of scans matched and no statistical measures for the variance of registration estimation. Another such approach is found in [2], [19], [13]. where two sets of planar or quadratic patches are matched using attribute-graphs . In that work, similarity metrics were formulated based on several attributes like shape-factor, area ratio, curvature-histogram, inter-surface relations etc., and a bounded tree search was performed to give a set of correspondences which maximized the metric. The result is refined using an evolutionary algorithm, which is computation-time intensive. An ICP-like algorithm working with planar-patch normals instead of points was presented in [15]. Planes were matched based on proximity in terms of inter-normal angle and Euclidean distance but the overall consistency of the scene was not considered. Again, no statistical measures of the variance of the solution were provided.

In contrast to the aforementioned approaches, the algorithm presented in this paper use only planar patches and maximizes the overall geometric consistency within a search-space to determine correspondences between planes. The search-space is pruned using



(a) The approach of Nüchter et al [7]. (b) The approach of Weingarten et al [18]. (c) The approach presented in this article.

Fig. 1. Comparison of algorithm structures. The ICP step in Weingarten et al's approach has been removed in our approach.

criteria such as overlap, size-similarity, and agreement with odometry, if available. For all these tests, only the plane parameter covariance matrix is employed, without the need to refer back to the original point-cloud. Our approach is fast and its reliability as well as computation-time increases with the number of planes. There is the option of using additional attributes like intensity, color, etc. for making the correspondence-finding step more reliable. Finally, the covariance matrix of the solution is computed which identifies the principal uncertainty directions. This information is indispensable for subsequent refinement processing like loop-closing or EKF-SLAM, although these are outside the scope of this paper.

This paper is organized as follows: Sec. II begins by presenting the basic notation and problem statement. Sec. II-C presents a new algorithm for finding plane correspondences. Sec. II-D solves the least-squares determination of pose-registration using the found correspondences. Sec. III shows a real-life example of mapping. The paper is concluded in Sec. IV.

II. Plane-Segment Extraction and Matching

The scan-matching based on plane-segments consists of the following three steps:

1) **Planes extraction from raw point-clouds**: This procedure is based on region-growing in a range-image scan followed by a least-squares estimation of the parameters of planes. The covariances of the plane-parameters are computed as well. The details may be found in the previously published work of the authors [16].
2) **Pose-registration by plane-matching**: This step consists of two substeps:
   a) Finding the correspondences between plane-segments in the two scans to be matched. These two scans may be successive samples for normal registration or may be non-successive, if a loop is being closed.
   b) After the correspondences have been decided on, finding the optimal rotation and translation which aligns the corresponding set of planes. This gives the pose change of the robot between the scans.
3) **Polygonization**: This step consists of polygonizing each plane-segment by finding the boundary of each surface-patch so that the surface can be compactly described. This step is crucial for visualization of the result, however, if only pose registration is desired, it may be omitted. It is also described in [16].

We now provide an overview of plane-matching, i.e. the second step.

A. Notation

A plane $P(\hat{\mathbf{m}}, \rho)$ is given by the equation $\hat{\mathbf{m}} \cdot \mathbf{r} = \rho$, where ρ is the signed distance from the origin in the direction of the unit plane normal $\hat{\mathbf{m}}$. We see that $P(\hat{\mathbf{m}}, \rho) \triangleq P(-\hat{\mathbf{m}}, -\rho)$. To achieve a consistent sign convention, we define planes as $P(\hat{\mathbf{n}}, d)$, where, $d \triangleq |\rho| \geq 0$, and $\hat{\mathbf{n}} \triangleq \sigma(\rho)\,\hat{\mathbf{m}}$, where, $\sigma(\rho) = -1$ if $\rho < 0$ and $+1$ otherwise. If $\rho = 0$, then we choose the maximum component of $\hat{\mathbf{n}}$ to be positive. The latter case is unlikely to occur in practice in the sensor-frame, because such a plane, which is parallel to the line of sight of the range sensor, is unlikely to be detected by it.

For denoting frames and relative transforms, we use the notation of [22]. The j-th sample is associated with a frame $\mathcal{F}_j$, in which the set of extracted planes is denoted as ${}^{j}\mathcal{P}$. Suppose we are given two samples ${}^{j}\mathcal{P}$ and ${}^{k}\mathcal{P}$. Usually they are successive, but they may be non-successive, for example during loop-closing. Using the procedure described in [17], one can compute an isotropic uncertainty measure σ^2 which is derived (usually, using the trace operation) from the 4×4 covariance matrix $\mathbf{C}$ of the plane parameters $\hat{\mathbf{n}}$ and d associated with it. Thus a plane-set ${}^{k}\mathcal{P}$ is an ordered set of triplets given by

$$ {}^{k}\mathcal{P} \triangleq \{\, {}^{k}\mathcal{P}_i \langle\, {}^{k}\hat{\mathbf{n}}_i,\ {}^{k}d_i,\ {}^{k}\sigma_i^2 \rangle,\ i = 1 \ldots \mathsf{N}_k \}, \tag{1} $$

B. Decoupling of Rotation and Translation

If the robot (more precisely, the sensor mounted on the robot) moves from $\mathcal{F}_j$ to $\mathcal{F}_k$, i.e. rotates by ${}^{j}_{k}\mathbf{R}$ and translates by ${}^{j}_{k}\mathbf{t}$ between samples j and k (resolved in $\mathcal{F}_j$), then the Cartesian coordinates ${}^{j}\mathbf{p}$ and ${}^{k}\mathbf{p}$ of the same physical point observed from the two frames are related by

$$ {}^{j}\mathbf{p} = {}^{j}_{k}\mathbf{R}\ {}^{k}\mathbf{p} + {}^{j}_{k}\mathbf{t}. \tag{2} $$

Substituting the above in the plane equation, one can derive that if ${}^{j}\mathcal{P}_i$ and ${}^{k}\mathcal{P}_i$ represent the same physical plane, the plane parameters observed in the two planes are related by

$$ {}^{j}\hat{\mathbf{n}}_i = {}^{j}_{k}\mathbf{R}\ {}^{k}\hat{\mathbf{n}}_i \tag{3} $$

$$ {}^{j}\hat{\mathbf{n}}_i^{\mathsf{T}}\ {}^{j}_{k}\mathbf{t} = {}^{j}d_i - {}^{k}d_i \tag{4} $$

The above equations show that the rotation and the translation components have been nicely decoupled.

C. Determining Correspondences

Given two plane-sets ${}^{j}\mathcal{P}$ of size N_j and ${}^{k}\mathcal{P}$ of size N_k, the problem here is to find which plane indices correspond to each other, i.e. which represent the same physical plane. The overlap is not given, and so some planes in one set may not have any corresponding plane in the other set. We present a new matching algorithm called Plane Registration based on Rotation of a Unit Sphere (PRRUS). The algorithm scales as $O(\mathsf{N}^4)$, where N is the average of N_j and N_k. Unlike RANSAC [14], this algorithm has no random component and relies on two observations 1) the number of high-evidence (detected using their parameter covariance matrix)

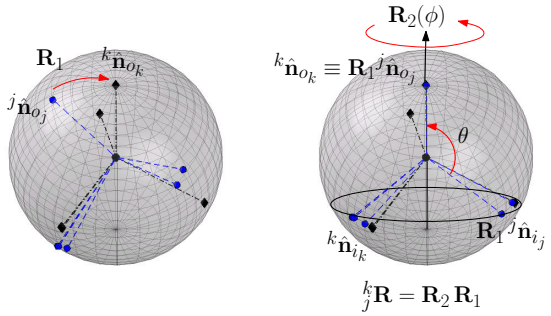


Fig. 2. Finding correspondences: steps 1 and 2.

planes in a scene, are much less than the number of points, and 2) When one pair of corresponding planes is found, only one degree of freedom is left for finding the rest of the correspondences.

PRRUS finds the normal correspondences which give the most geometric consistency, i.e. maximizes the satisfaction of (3). As illustrated in Fig. 2, it divides the problem in two steps:

a) Step 1: Find the pair of planes (o_j^*, o_k^*) which maximizes the size of the set Ω_1 of plane-pairs with the same relative angles to them. In other words,

$$(o_j^*, o_k^*) \triangleq \max_{(o_j, o_k)} \#\Omega_0(o_j, o_k), \tag{5}$$

$$\Omega_0(o_j, o_k) \triangleq \{(i_j, i_k) \mid {}^j\hat{\mathbf{n}}_{o_j} \cdot {}^j\hat{\mathbf{n}}_{i_j} \approx {}^k\hat{\mathbf{n}}_{o_k} \cdot {}^k\hat{\mathbf{n}}_{i_k}\} \tag{6}$$

The approximate equality-test uses the χ^2 test since the covariances of the normals is known. In (6), we also remove planes, where the dot-product is close to ± 1, i.e. planes parallel or anti-parallel to ${}^j\hat{\mathbf{n}}_{o_j}$ and ${}^k\hat{\mathbf{n}}_{o_k}$. Furthermore, we apply additional checks within the maximization step to test for overlap and size-similarity, and prune out pairs in $\Omega_0(o_j, o_k)$ which fail these checks. This first step takes the worst case time of $O(\mathsf{N}_j^2 \mathsf{N}_k^2)$.

After this step, we compute the rotation matrix $\mathbf{R}_1$, such that

$$\mathbf{R}_1 {}^j\hat{\mathbf{n}}_{o_j^*} = {}^k\hat{\mathbf{n}}_{o_k^*}. \tag{7}$$

We then compute the consensus set of triplets Ω_1,

$$\Omega_1 \triangleq \{(i_j, i_k, \phi_i) \mid (i_j, i_k) \in \Omega_0(o_j^*, o_k^*)\} \tag{8}$$

$$\phi_i \triangleq \measuredangle(\mathbf{R}_1 {}^j\hat{\mathbf{n}}_{i_j},\, {}^k\hat{\mathbf{n}}_{i_k}) \text{ about } {}^k\hat{\mathbf{n}}_{o_k^*}. \tag{9}$$

b) Step 2: Now it remains to find the most consistent ϕ. Find the biggest cluster in the distribution of ϕ_i within Ω_1; call the mean of the cluster ϕ^*. This gives us the final consensus set

$$\Omega_2 \triangleq \{(i_j, i_k) \mid (i_j, i_k, \phi^*) \in \Omega_1\} \tag{10}$$

The set Ω_2 is our set of resolved correspondences, to which a least-squares rotation and translation is fitted next.

D. Pose-registration by plane-matching

Statement of the Problem Given noisy plane-sets ${}^j\mathcal{P}$ and ${}^k\mathcal{P}$, with correspondences between planes known and given by the common index i, find the optimum rotation ${}^j_k\mathbf{R}$ and the optimum translation ${}^j_k\mathbf{t}$. We provide a solution below.

1) Optimum Rotation: To find the optimum rotation between ${}^j\mathcal{P}$ and ${}^k\mathcal{P}$, we maximize the following value-function

$$\max_{{}^j_k\mathbf{R}} \zeta_r \triangleq \frac{1}{2} \sum_{i=1}^{\mathsf{N}} ({}^j\sigma_i^2 + {}^k\sigma_i^2)^{-1}\, {}^j\hat{\mathbf{n}}_i^\mathsf{T}\, {}^j_k\mathbf{R}\, {}^k\hat{\mathbf{n}}_i \tag{11}$$

This is the well-known Wahba's problem [23]. To solve this we parameterize the rotation ${}^j_k\mathbf{R}$ with quaternions ${}^j_k\check{\mathbf{q}}$ and proceed as in [24]. Using the above, one can reformulate (11) as

$$\max_{{}^j_k\check{\mathbf{q}}} \zeta_r \triangleq \frac{1}{2}\, {}^j_k\check{\mathbf{q}}^\mathsf{T} \mathbf{K}\, {}^j_k\check{\mathbf{q}}, \tag{12}$$

where the matrix $\mathbf{K}$ is depends on ${}^j\hat{\mathbf{n}}_i$ and ${}^k\hat{\mathbf{n}}_i$, $i = 1 \ldots \mathsf{N}$. The optimum quaternion ${}^j_k\check{\mathbf{q}}$ is then the eigenvector of the matrix $\mathbf{K}$ corresponding to its maximum eigenvalue $\bar{\mu}(\mathbf{K})$. The 4×4 covariance of this optimum ${}^j_k\check{\mathbf{q}}$ can be computed as

$${}^j_k\mathbf{C}_{\check{\mathbf{q}}\check{\mathbf{q}}} = -\left(\mathbf{K} - \bar{\mu}(\mathbf{K})\, \mathbf{I}_4\right)^+, \tag{13}$$

where, $\mathbf{X}^+$ represents the Moore-Penrose inverse of the matrix $\mathbf{X}$ [25]. This covariance matrix can then be transformed to roll-pitch-yaw (RPY) space by using the Jacobian of the transform between the quaternion and the RPY space. Finally, we note that at least two non-parallel pairs of planes are required to fully determine rotation.

2) Optimum Translation: Equations (4) can be stacked-up to give

$$\mathbf{M}\, {}^j_k\mathbf{t} = \mathbf{d}, \tag{14a}$$

where,

$$\mathbf{M}_{\mathsf{N}\times3} \triangleq \begin{bmatrix} {}^j\hat{\mathbf{n}}_1^\mathsf{T} \\ \vdots \\ {}^j\hat{\mathbf{n}}_\mathsf{N}^\mathsf{T} \end{bmatrix}, \qquad \mathbf{d}_{\mathsf{N}\times1} \triangleq \begin{bmatrix} {}^jd_1 - {}^kd_1 \\ \vdots \\ {}^jd_\mathsf{N} - {}^kd_\mathsf{N} \end{bmatrix} \tag{14b}$$

Due to its intuitive nature and fast closed-form solution, we will solve Eq. (14a) with ordinary least squares (LS). A diagonal weighting matrix $\mathbf{W}$ is defined as

$$\mathbf{\Sigma} \triangleq \begin{pmatrix} {}^j\sigma_1^2 + {}^k\sigma_1^2 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & {}^j\sigma_\mathsf{N}^2 + {}^k\sigma_\mathsf{N}^2 \end{pmatrix}, \tag{15}$$

$$\mathbf{W} \triangleq \left(\mathbf{\Sigma}^{-1}\right)^{1/2}. \tag{16}$$

Then the LS solution minimizes $\|\mathbf{W}(\mathbf{M}\, {}^j_k\mathbf{t} - \mathbf{d})\|$. If $\mathbf{M}$ is full rank, the least-squares optimum translation is ${}^j_k\mathbf{t} = \left(\mathbf{M}^\mathsf{T}\mathbf{W}^2\mathbf{M}\right)^{-1}\mathbf{M}^\mathsf{T}\mathbf{W}^2\mathbf{d}$.

Unlike rotation, in general we need $\mathsf{N} \geq 3$ mutually non-parallel planes to find ${}^j_k\mathbf{t}$. The above formula is not a good way to compute the solution because $\mathbf{M}$ may be ill-conditioned, may be rank-deficient, or $\mathsf{N} < 3$. A more general way to solve the equation is presented next. We define

$$\hat{\mathbf{M}} \triangleq \mathbf{W}\mathbf{M}, \qquad \hat{\mathbf{d}} \triangleq \mathbf{W}\mathbf{d}. \tag{17}$$

Let the singular-value decomposition of $\hat{\mathbf{M}}$ be given by $\mathbf{U}_{\mathsf{N}\times\mathsf{N}}\mathbf{\Lambda}_{\mathsf{N}\times3}\mathbf{V}_{3\times3}^\mathsf{T}$. $\mathbf{\Lambda}$ has non-negative singular values λ_i^2 arranged in descending order. The column unit vectors of $\mathbf{U}$ are denoted $\mathbf{u}_i$, $i = 1 \ldots \mathsf{N}$ and the column unit vectors of $\mathbf{V}$ are denoted $\mathbf{v}_i$, $i = 1 \ldots 3$.

Let $\mathsf{N}_{\hat{\mathbf{M}}} \leq 3$ be the *effective* rank of $\hat{\mathbf{M}}$. If the largest singular value $\lambda_1^2 < \epsilon_1$, then the effective rank is 0. The parameter ϵ_1 is dependent on machine accuracy. In our computations, we have taken it as 10^{-7}. If $\lambda_1^2 \geq \epsilon_1$, then the effective rank is found by finding the count of all singular values $\lambda_i^2 > \lambda_1^2/\bar{c}$ in the diagonal matrix $\mathbf{\Lambda}$, where $\bar{c}$ is the maximum allowable condition number of the matrix. We have set $\bar{c} \approx 200$. In practice ϵ_1 and $\bar{c}$ are quite important parameters for obtaining good translation estimates and also for identifying the directions in which the translation estimate is the most uncertain.

Then the best rank $\mathsf{N}_{\hat{\mathbf{M}}}$ approximation of $\hat{\mathbf{M}}$ is

$$\tilde{\mathbf{M}} = \sum_{i=1}^{\mathsf{N}_{\hat{\mathbf{M}}}} \lambda_i^2 \hat{\mathbf{u}}_i \hat{\mathbf{v}}_i^{\mathsf{T}}, \qquad \mathsf{N}_{\hat{\mathbf{M}}} \leq 3. \tag{18}$$

The span of the orthogonal unit vectors $\hat{\mathbf{u}}_i,\ i = 1 \ldots \mathsf{N}_{\hat{\mathbf{M}}}$ gives the best approximation for the range-space of $\hat{\mathbf{M}}$. Therefore, the closest we can get to $\hat{\mathbf{d}}$ is $\tilde{\mathbf{d}} = \sum_{i=1}^{\mathsf{N}_{\hat{\mathbf{M}}}} (\hat{\mathbf{u}}_i \cdot \hat{\mathbf{d}}) \hat{\mathbf{u}}_i$, which gives the corresponding translation estimate

$$\begin{aligned} {}^{j}_{k}\mathbf{t} &= \sum_{i=1}^{\mathsf{N}_{\hat{\mathbf{M}}}} \lambda_i^{-2} (\hat{\mathbf{u}}_i \cdot \hat{\mathbf{d}}) \hat{\mathbf{v}}_i & (19) \\ &\triangleq \hat{\mathbf{M}}^{+} \hat{\mathbf{d}}, \qquad \hat{\mathbf{M}}^{+} \triangleq \sum_{i=1}^{\mathsf{N}_{\hat{\mathbf{M}}}} \lambda_i^{-2} \hat{\mathbf{v}}_i \hat{\mathbf{u}}_i^{\mathsf{T}}. & (20) \end{aligned}$$

This is also the minimum 2-norm solution of the LS problem regardless of the rank of $\mathbf{M}$ mentioned in [26].

Note that for directions $\hat{\mathbf{v}}_i,\ i = (\mathsf{N}_{\hat{\mathbf{M}}} + 1) \ldots 3$, we have no information about the translation. One option is to keep these components 0 and inject large uncertainty along those directions in the covariance matrix. However, if an odometry estimate ${}^{j}_{k}\mathbf{t}_{\mathrm{O}}$, along with its covariance matrix ${}^{j}_{k}\mathbf{C}_{\mathbf{tt},\mathrm{O}}$ is available, we can use it *only* for these missing components. In this case we have

$$\begin{aligned} {}^{j}_{k}\mathbf{t} &= \hat{\mathbf{M}}^{+}\mathbf{d} + \sum_{i=\mathsf{N}_{\hat{\mathbf{M}}}+1}^{3} ({}^{j}_{k}\mathbf{t} \cdot \hat{\mathbf{v}}_i) \hat{\mathbf{v}}_i, \\ &\triangleq \hat{\mathbf{M}}^{+}\mathbf{d} + \mathbf{M}_{\mathrm{O}}\, {}^{j}_{k}\mathbf{t}_{\mathrm{O}}, \quad \mathbf{M}_{\mathrm{O}} \triangleq \sum_{i=\mathsf{N}_{\hat{\mathbf{M}}}+1}^{3} \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^{\mathsf{T}}. \end{aligned} \tag{21}$$

The covariance of odometry is only very roughly known because it depends on vehicle model, unknown slippage etc., and a usual fusion of odometry translation and plane-matching translation using this covariance should be avoided. The nice thing about this solution is that it automatically detects the directions for which the translation is uncertain, and resorts to odometry only for these directions. Needless to say, if such a case occurs, i.e. the translation is not fully determined, the uncertainty in translation increases, which needs to be then mitigated by loop-closing. Finally, we can write the estimate of the covariance matrix for translation as follows

$$\begin{aligned} {}^{j}_{k}\mathbf{C}_{\mathbf{tt}} &= \hat{\mathbf{M}}^{+} \mathbf{W} \boldsymbol{\Sigma} \mathbf{W}^{\mathsf{T}} (\hat{\mathbf{M}}^{+})^{\mathsf{T}} + \mathbf{M}_{\mathrm{O}}\, {}^{j}_{k}\mathbf{C}_{\mathbf{tt},\mathrm{O}}\, \mathbf{M}_{\mathrm{O}}^{\mathsf{T}}, \\ &= \mathbf{M}^{+} \boldsymbol{\Sigma} (\mathbf{M}^{+})^{\mathsf{T}} + \mathbf{M}_{\mathrm{O}}\, {}^{j}_{k}\mathbf{C}_{\mathbf{tt},\mathrm{O}}\, \mathbf{M}_{\mathrm{O}}^{\mathsf{T}} \end{aligned} \tag{22}$$

where the last equation comes from simplification using (16) and (17).

III. Multi-Level Lab Arena Example

(a) (b) (c)

Fig. 3. The robot collecting data in the locmotion test arena in form of a high bay rack

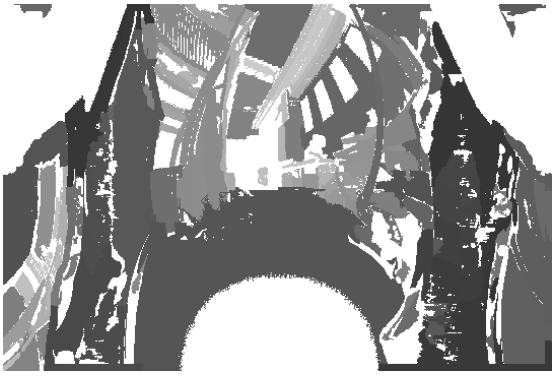


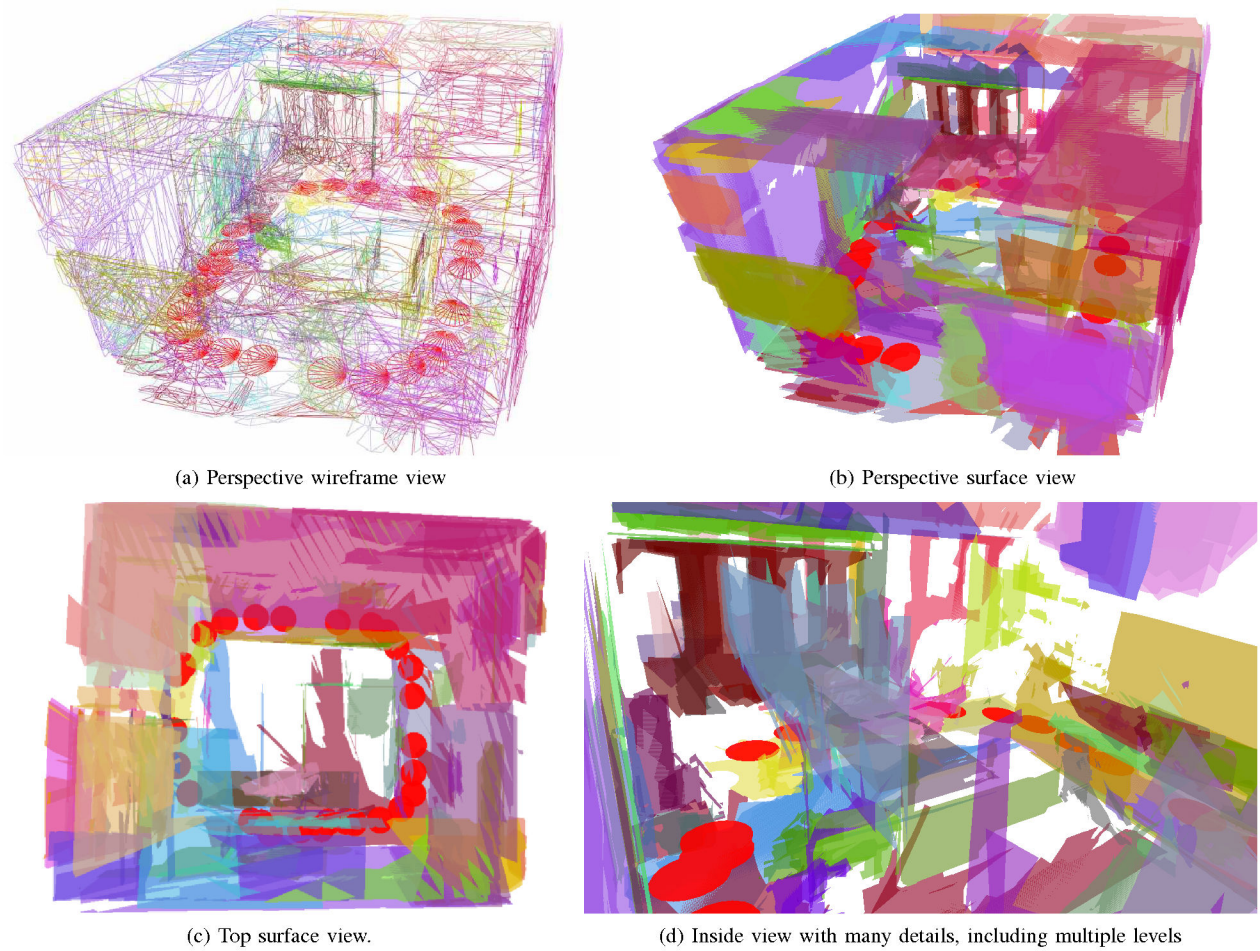
Fig. 4. The 541×361 range-image corresponding to the first scan, with the planar regions identified. This image actually lies on the surface of a sphere, and displaying it as a planar rectangle leads to an obvious distortion.

The Jacobs University rescue robot is equipped with an actuated laser range-finder (ALRF). The ALRF has a horizontal field of view of 270° of 541 beams. The sensor is rotated (pitched) by a servo from -90° to $+90^{\circ}$ at a spacing of 0.5°. This gives a 3D point-cloud of a total size of $541 \times 361 = 195301$ per sample. The maximum range of the sensor is about 20 meters. The mobile robot was teleoperated and stopped occasionally to take scans. The time to take one full scan is about $T_{\text{scan}} \approx 32$ seconds. Every scan corresponds to a 541×361 range-image as shown in Fig. 4. Planar patches were extracted from these range images using a region-growing algorithm described in [16]. For each planar surface, a covariance matrix of plane parameters was computed, as described in [17]. A total of 29 usable scans were taken as the robot was driven around the multi-level arena shown in Fig. 3. The front camera images of some of these locations are shown in Fig. 5.

The result of the plane-matching is shown in Fig. 6 and a view of this map from inside the arena is shown in Fig. 6(d). This map is clearly well aligned and is easy to understand by a human. The rotation and translation trajectory of the robot computed by plane-matching is shown in Fig. 7. This figure also shows the $n\sigma$ bounds of the uncertainties in the pose variables, where $n = 100, 200$ respectively. These bounds, like in any other scan-matcher, assume that all the found plane correspondence are correct, and therefore tend to be optimistic.

Fig. 5. The front camera view of the scene at the locations where the robot took a scan. Only every fourth scan is shown to save space. The full video may be accessed at [27].

The run-times for the various steps described in Sec. II are shown in Table I. These are for an AMD Turion 2×64 machine 1.6 GHz with 960 MB RAM running OpenSUSE 10.3 O/S. Compared to the previously mentioned total time for taking one scan, the registration method is indeed "fast": mean time for extraction and plane-matching together $\approx 0.23\, T_{\text{scan}}$. We reiterate that the polygonalization step is needed mainly for visualization. Fig. 8 shows the dependence of the



(a) Perspective wireframe view

(b) Perspective surface view

(c) Top surface view.

(d) Inside view with many details, including multiple levels

Fig. 6. The lab model. The robot location is shown by a red circle at the height of the sensor. Corresponding planes are drawn with the same color. Planar surface patches are drawn semi-transparent to show internal detail. The left wall in Fig. 6(c), shows some translation offset where the loop was closed, but almost no rotational misalignment. The translation misalignment can be corrected by loop-relaxation, which is, however, outside the scope of this paper. An X3D model and video files for this dataset can be found at [27]

TABLE I
PLANE-MATCHING RESULTS FOR 29 SCANS: MEAN ± STD

Nr. of planes extracted	110 ± 30	
Nr. of planes after filtering	16 ± 5	
Nr. of corresponding planes found	8 ± 3	
Time for plane extraction	2.67 ± 0.3	sec.
Time for plane-matching	4.867 ± 4.771	sec.
Time for polygonalization	3.701 ± 0.885	sec.

plane-matching computation time on the average number of high-evidence planes in the scene.

Results for a Small FOV TOF Sensor: The results for a scene recreated using a Swiss-Ranger is presented in [28] and could not be included in this paper due to size restrictions. This data-set also provides a comparsion with ICP and highlights its behavior of local minimum convergence.

IV. CONCLUSIONS

A fast method for 3D mapping using planar patches was presented and experimentally validated. The method decouples rotation and translation determination, and is able to compute the uncertainty in pose-registration using the uncertainties in the plane parameters. It provides an efficient alternative to point/voxel-based approaches in predominantly planar environments, and furthermore represents the map in a compressed and easily understandable form.

V. ACKNOWLEDGMENTS

This work was supported by the German Research Foundation (DFG).

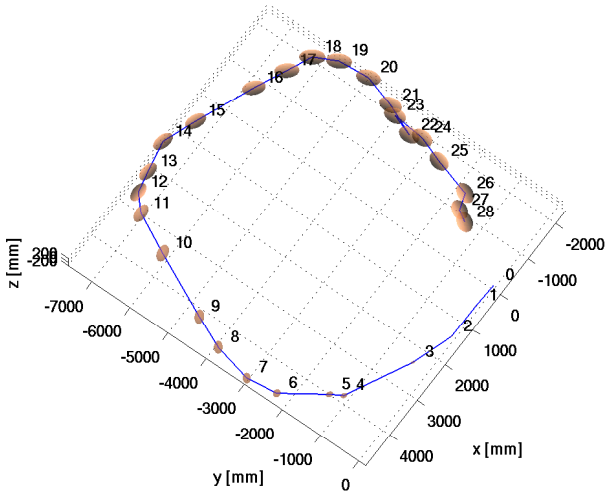
REFERENCES

[1] H. Durrant-Whyte and T. Bailey, "Simultaneous localization and mapping: Part I," *IEEE Robotics and Automation Magazine*, pp. 99–108, June 2006.

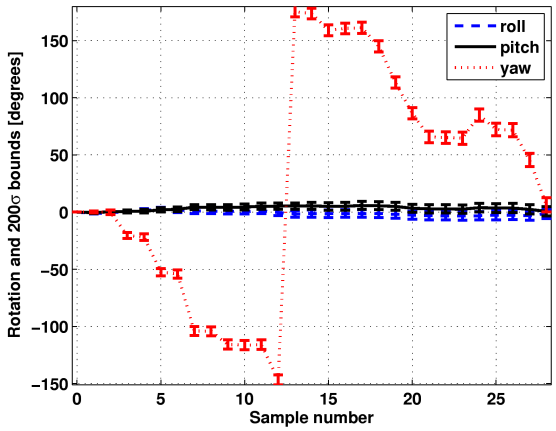
[2] D. Fischer and P. Kohlhepp, "3D geometry reconstruction from multiple segmented surface descriptions using neuro-fuzzy similarity measures," *Journal of Intelligent and Robotic Systems*, vol. 29, pp. 389–431, 2000.

[3] H. Surmann, A. Nuechter, and J. Hertzberg, "An autonomous mobile robot with a 3d laser range finder for 3d exploration and digitalization of indoor environments," *Robotics and Autonomous Systems*, vol. 45, no. 3-4, pp. 181–198, 2003.

[4] S. Thrun, D. F. D. Haehnel, M. Montemerlo, R. Triebel, W. Burgard, C. Baker, Z. Omohundro, S. Thayer, and W. Whittaker, "A system for volumetric robotic mapping of abandoned mines," in *Proc. IEEE International Conference on Robotics and Automation (ICRA)*, Taipei, Taiwan, 2003.



(a) XYZ plot with 100σ ellipsoids. Note how the major axes of ellisoids gradually become parallel to corridor directions.



(b) Roll, pitch, yaw and their 200σ bounds. Note the wrapping of yaw at ± 180.

Fig. 7. Computed robot trajectory and uncertainty.

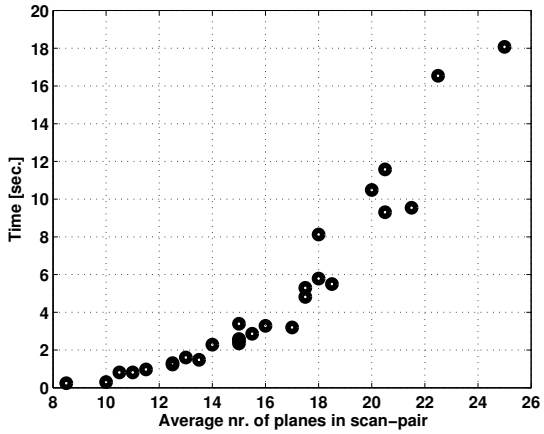


Fig. 8. The dependence on the computation time of the plane-matching algorithm on the average number of planes in the scan-pair after filtering out low evidence planes.

[5] J. Weingarten and R. Siegwart, "3D SLAM using planar segments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Beijing, 2006.

[6] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *Journal of Field Robotics*, vol. 24, no. 10, pp. 803–827, 2007.

[7] A. Nüchter, K. Lingemann, and J. Hertzberg, "6D SLAM– 3D mapping outdoor environments," *Journal of Field Robotics*, vol. 24, no. 8/9, pp. 699–722, 2007.

[8] CSEM, *The SwissRanger, Manual V1.02*, CSEM SA, 8048 Zurich, Switzerland, 2006. [Online]. Available: http://www.swissranger.ch

[9] *PhotonIC (R) PMD 3k-S*, PMD Technologies, 2008. [Online]. Available: http://www.pmdtec.com

[10] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 239–256, Feb 1992.

[11] O. D. Faugeras and F. Lustman, "Motion and structure from motion in piecewise planar environment," *Int. Journal of Pattern Recognition and Artificial Intelligence*, vol. 2, no. 3, pp. 485–508, 1988.

[12] R. Fisher, "Geometric constraints from planar surface patch matching," *Image and Vision Computing*, vol. 8, no. 2, pp. 148–154, 1990.

[13] P. Kohlhepp, G. Bretthauer, M. Walther, and R. Dillmann, "Using orthogonal surface directions for autonomous 3d-exploration of indoor environments," in *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 2006, pp. 3086–3092.

[14] D. Fontanneli, L. Ricciato, and S. Soatto, "A fast RANSAC-based registration algorithm for accurate localization in unknown environments using LIDAR measurements," in *IEEE Int. Conf. on Automation Science and Engineering (CASE)*, 2007, pp. 597–602.

[15] D. Viejo and M. Cazorla, "3D plane-based egomotion for SLAM on semi-structured environment," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS)*, 2007, pp. 2761–2766.

[16] J. Poppinga, N. Vaskevicius, A. Birk, and K. Pathak, "Fast plane detection and polygonalization in noisy 3D range images," in *IEEE Int. Conf. on Intelligent Robots and Systems (IROS)*, Nice, France, 2008.

[17] K. Pathak, N. Vaskevicius, and A. Birk, "Revisiting uncertainty analysis for optimum planes extracted from 3D range sensor point-clouds," in *IEEE Int. Conf. on Robotics and Automation*, Kobe, Japan, 2009.

[18] J. Weingarten, "Feature-based 3D SLAM," Ph.D. dissertation, EPFL, Lausanne, Switzerland, 2006. [Online]. Available: http://library.epfl.ch/theses/?nr=3601

[19] P. Kohlhepp, P. Pozzo, M. Walther, and R. Dillmann, "Sequential 3D-SLAM for mobile action planning," *Intelligent Robots and Systems, 2004. (IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 1, pp. 722–729 vol.1, Sept.-2 Oct. 2004.

[20] W. E. L. Grimson, *Object recognition by computer: the role of geometric constraints*. Cambridge: MIT Press, 1990.

[21] W. He, W. Ma, and H. Zha, "Automatic registration of range images based on correspondence of complete plane patches," in *Fifth Int. Conf. on 3-D Digital Imaging and Modeling (3DIM'05)*. IEEE, 2005, pp. 470–475.

[22] J. J. Craig, *Introduction to robotics – Mechanics and control*. Prentice Hall, 2005.

[23] M. D. Shuster, "The generalized Wahba problem," *The Journal of the Astronautical Sciences*, vol. 54, no. 2, pp. 245–259, April-June 2006.

[24] B. K. P. Horn, "Closed-form solution of absolute orientation using unit quaternions," *Journal of the Optical Society of America*, vol. 4, no. 4, pp. 629–642, 1987.

[25] K. Kanatani, *Statistical Optimization for Geometric Computation*. New York: Dover Publications, Inc., 2005.

[26] G. H. Golub and C. F. V. Loan, "An analysis of the total least squares problem," *SIAM Journal of Numerical Analysis*, vol. 17, no. 6, pp. 883–893, 1980.

[27] K. Pathak and A. Birk, "3D map of Jacobs University locmotion test arena in form of a high bay rack," 2009. [Online]. Available: http://robotics.jacobs-university.de/datasets/Lab3D-2009/

[28] J. Poppinga, N. Vaskevicius, K. Pathak, and A. Birk, "Plane matching using Swiss-Ranger and comparison to ICP," 2009. [Online]. Available: http://robotics.jacobs-university.de/datasets/SRLab