

Information-theoretic compression of pose graphs for laser-based SLAM

Henrik Kretzschmar and Cyrill Stachniss

Abstract

In graph-based simultaneous localization and mapping (SLAM), the pose graph grows over time as the robot gathers information about the environment. An ever growing pose graph, however, prevents long-term mapping with mobile robots. In this paper, we address the problem of efficient information-theoretic compression of pose graphs. Our approach estimates the mutual information between the laser measurements and the map to discard the measurements that are expected to provide only a small amount of information. Our method subsequently marginalizes out the nodes from the pose graph that correspond to the discarded laser measurements. To maintain a sparse pose graph that allows for efficient map optimization, our approach applies an approximate marginalization technique that is based on Chow–Liu trees. Our contributions allow the robot to effectively restrict the size of the pose graph. Alternatively, the robot is able to maintain a pose graph that does not grow unless the robot explores previously unobserved parts of the environment. Real-world experiments demonstrate that our approach to pose graph compression is well suited for long-term mobile robot mapping.

Keywords

SLAM, long-term, pose graph, compression, mutual information

1. Introduction

Maps of the environment are needed for a wide range of robotic applications including transportation and delivery tasks, search and rescue, or efficient automated vacuum cleaning. The capability of building an appropriate model of the environment allows for designing robots that can efficiently operate in complex environments only based on their on-board sensors and without relying on external reference systems. In the past, several effective approaches to robot mapping have been developed. A popular approach to address the simultaneous localization and mapping (SLAM) problem models the poses of the robot as nodes in a graph. Spatial constraints between poses that result from observations or odometry are encoded as edges. Often, graph-based approaches marginalize out features or local grid maps and reduce the problem to trajectory estimation without prior map knowledge, followed by mapping with known poses.

Most of the SLAM approaches assume that map learning is carried out as a preprocessing step and that the robot then uses the acquired model for tasks such as localization and path planning. A robot that has to extend the map of its environment during long-term operation cannot apply

most of the existing graph-based mapping approaches since their complexity grows with the length of the robot's trajectory. The reason for this is that standard graph-based approaches constantly add new nodes to the graph. As a result, memory and computational requirements grow over time, preventing long-term mapping applications. A continuously growing graph slows down graph optimization and makes it more and more costly to find constraints between the current pose and former poses, i.e. to identify loop closures. Note that there exist also incremental methods for online corrections that perform partial optimizations. These methods are mostly orthogonal to our contribution.

The contribution of this paper is an information-theoretic approach to lossy pose graph compression to allow graph-based SLAM systems to operate over extended periods of time. Figure 1 depicts a motivating example. The top image shows the pose graph and the resulting map obtained by a standard graph-based approach to SLAM. The bottom

Department of Computer Science, University of Freiburg, Germany

Corresponding author:

Cyrill Stachniss, Department of Computer Science, University of Freiburg, Geroges Köhler-Allee 79, Freiburg 79110, Germany.
Email: stachnis@informatik.uni-freiburg.de

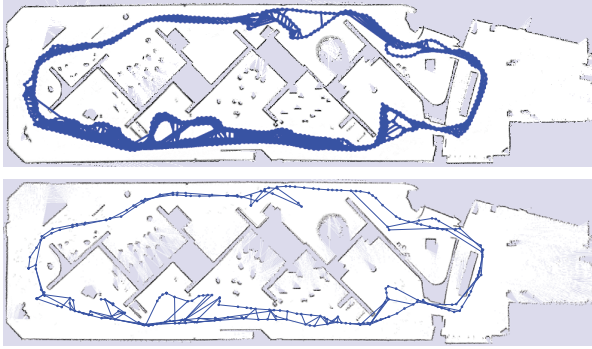


Fig. 1. The goal of our work is to compress the SLAM pose graph (top: 2,049 nodes, 7,190 edges) to maintain a sparse pose graph (bottom: 250 nodes, 324 edges), minimizing the loss of information in the graph and the resulting map. Although the sparsified pose graph exhibits substantially fewer nodes and edges, the resulting grid maps are highly similar. In this example, only 1.2% of the grid cells changed their most likely state (free, occupied, unknown) due to our compression technique. Most of the cells that changed their most likely state were seen as free before applying our compression method and are seen as unknown after applying our technique. This applies, for example, to some of the cells in the partially visited room at the right-hand side of the map. The standard SLAM approach required 32 min to process the dataset, whereas our approach needed 17 min.

image displays the corresponding pose graph along with the map generated by our information-theoretic compression approach. The images illustrate that significantly fewer laser scans are required to provide a comparable mapping result. We present an approach to select laser scans for removal such that the expected loss of information with respect to the map is minimized. Our unbiased selection applies the information-theoretic concept of mutual information to determine the laser scans that should be removed. In order to keep the pose graph compact, the pose nodes corresponding to the discarded laser scans need to be eliminated from the pose graph. We achieve this by applying an approximate marginalization scheme that maintains the natural sparsity pattern that is observed in the context of SLAM. Our approach is highly relevant to long-term mapping, particularly when the robot frequently re-traverses already visited areas. Our method allows the robot to maintain a pose graph that does not grow unless the robot explores previously unobserved parts of the environment, while minimizing the expected loss of information.

This paper extends the work of Kretschmar et al. (2011) and Stachniss and Kretschmar (2011) as it provides a more detailed description of the presented approach and a more comprehensive experimental evaluation. In addition to that, this paper relaxes the assumption of our prior work that the a priori probability of range readings is governed by an exponential distribution. Instead, the distribution of range readings is learned.

2. Related work

In the context of the SLAM problem, Lu and Milios (1997) were the first to perform global map optimization using a pose graph. Since then, a large variety of approaches to minimize the error in pose graphs have been proposed, for example, Estrada et al. (2005), Folkesson et al. (2005), Grisetti et al. (2010), Konolige and Agrawal (2008), and Ni et al. (2007). Most of these approaches to SLAM do not provide a means to effectively prune the graph. Instead, they add more and more nodes to the graph over time. Some approaches, for example that of Folkesson et al. (2005), group nodes into rigid local sub-maps. Others, such as the ATLAS framework by Bosse et al. (2003), subdivide the map into connected frames that contain maps that capture the local environment. Typically, these methods do not discard nodes that store information about the environment and therefore do not prevent the graph from growing.

One way to reduce the number of nodes in the graph is to sample the trajectory of the robot at an appropriate spatial decimation as done by Eustice et al. (2006). Similarly, the method of Konolige and Agrawal (2008) only adds a new node to the graph if it is not spatially close to any existing node. Konolige and Bowman (2009) presented an approach to lifelong mapping that uses a single stereo camera and that is able to update the map when the environment changes. Their method discards views based on a least-recently-used algorithm. The above-mentioned techniques do not rely on information-theoretic concepts to determine which measurements to discard. Similar to that, hierarchical techniques such as the techniques by Grisetti et al. (2010), Estrada et al. (2005), or Ni et al. (2007) have been employed to bound the computational requirements by optimizing only higher levels of the hierarchy.

In contrast to that, Davison (2005) analyzed mutual information, particularly in the case of Gaussian probability distributions, to guide image processing. In the vision community, Snavely et al. (2008) aim to find a skeletal subgraph with the minimum number of interior nodes that spans the full graph while achieving a bound on the full covariance. Their technique is used for reconstructing scenes based on large, redundant photo collections.

Kaess and Dellaert (2009) consider the information gain of measurements in the state estimate within the iSAM framework. In contrast to that, our approach estimates the mutual information of laser scans and the occupancy grid map, thus considering the effect on the resulting grid map explicitly. Ila et al. (2010) propose to only incorporate non-redundant poses and informative constraints based on the relative distance between poses in information space and the expected information gain of candidate loop closures. As opposed to our maximum-likelihood approach to SLAM based on pose graphs, their method applies an information filter and does not marginalize out poses that were already added. Eade et al. (2010) presented a view-based monocular SLAM system that reduces the complexity of the graph by

marginalization and subsequent suppression of edges incident to nodes of high degrees. Their heuristic discards the constraints that most agree with the current state estimate. This, however, introduces a bias into the system.

When discarding laser range scans, our approach marginalizes out the corresponding pose node from the pose graph. Exact marginalization, however, results in a dense pose graph and thus we apply an approximate marginalization scheme that is based on Chow–Liu trees. Other researchers in robotics use Chow–Liu trees to approximate multi-dimensional probability distributions. For instance, Cummins and Newman (2008) extend the bag-of-words approach by learning a model for the sensory data in the form of a Chow–Liu tree. Furthermore, Chli and Davison (2009) use a Chow–Liu tree approximation to divide large visual maps into a fully hierarchical correlation and clustering structure.

The idea of using mutual information or expected information gain in the context of mapping has also been investigated for other means. For example, Kim and Eustice (2011) propose a system for finding loop closures by combining the information gain with saliency. Bachrach et al. (2011), Kollar and Roy (2008), and Stachniss et al. (2005) use this concept to evaluate target locations during exploration to estimate the amount of novel information that measurements are expected to provide. Furthermore, He et al. (2008) estimate the increase of uncertainty within the planning system to select appropriate actions for a quadcopter.

3. Brief introduction to graph-based SLAM

Graph-based approaches to SLAM model the poses of the robot as nodes in a graph. The edges of the graph model spatial constraints between the nodes. These constraints arise from odometry measurements and from feature observations or scan matching. The so-called SLAM front-end interprets the sensor data to extract the constraints. The so-called SLAM back-end typically applies optimization techniques to estimate the configuration of the nodes that best matches the spatial constraints.

Our laser-based front-end applies correlative scan matching as proposed by Olson (2009) to estimate a constraint between the current node and the previous node. Our method also generates loop closure hypotheses by matching the current laser scan against a set of scans that is determined by the relative positional uncertainties and then rejects false hypotheses using the spectral clustering approach described by Olson (2008). Our approach incrementally optimizes the pose graph while adding the poses and the constraints to it. Once the poses are estimated, the laser scans are used to render an occupancy grid map of the environment. The robot therefore stores the laser scans that correspond to the pose nodes in the pose graph.

The back-end aims at finding the spatial configuration \mathbf{x}^* of the nodes that maximizes the log likelihood of the

observations. Let $\mathbf{x} = (x_1^\top, \dots, x_n^\top)^\top$ be a vector where x_i describes the pose of node i , and let z_{ij} and Ω_{ij} be the mean and the information matrix of an observation of node j seen from node i assuming Gaussian noise. Furthermore, let $\mathbf{e}_{ij}(\mathbf{x})$ be an error vector which expresses the difference between an observation and the current configuration of the nodes and let \mathcal{C} be the set of pairs of nodes for which a constraint exists. Assuming the constraints to be independent, we have

$$\mathbf{x}^* = \underset{\mathbf{x}}{\operatorname{argmax}} \sum_{(i,j) \in \mathcal{C}} \mathbf{e}_{ij}(\mathbf{x})^\top \Omega_{ij} \mathbf{e}_{ij}(\mathbf{x}). \quad (1)$$

Our approach applies the technique proposed by Grisetti et al. (2010), which uses sparse Cholesky factorization to efficiently solve the system of linearized equations that is obtained from Equation (1).

4. Discarding uninformative laser scans

The first contribution of this paper is an approach to select the laser scans that are most informative with respect to the map estimate. Our technique aims at minimizing the expected loss of information in the resulting map without introducing a bias during the selection of the laser scans. Such a technique is important to allow for long-term robot mapping since a robot that keeps all scans will run out of resources at some point. In addition to that, our method can be used to directly implement an any-space SLAM system. Whenever the memory limit is reached, our algorithm discards the laser scans that are expected to be least informative about the map and marginalizes out the corresponding pose nodes.

4.1. Finding the most informative subset of laser scans

We define the map M as a random variable describing the state of the world. It is highly correlated to the random variables $Z_{1:t}$ describing the laser scans $z_{1:t}$ recorded at the poses $x_{1:t}$. We use Z_i^j to refer to the random variable of an individual beam of laser scan Z_i . To estimate the state of the world m , we consider the posterior probability distribution of the map M given the laser measurements $z_{1:t}$. In this section, we are interested in finding the subset $Z^* \subseteq Z_{1:t}$ of at most n laser measurements that is expected to result in the smallest uncertainty about the map M .

Following the notation of MacKay (2003), the average reduction in the uncertainty of the map M due to a set Z of laser measurements is given by the mutual information

$$I(M; Z) = H(M) - H(M | Z), \quad (2)$$

where H is the entropy. Hence, we want to find the subset $Z^* \subseteq Z_{1:t}$ of at most n laser measurements such that the mutual information of the map M and the subset Z^* is maximized, i.e.

$$Z^* = \underset{Z \subseteq Z_{1:t}, |Z| \leq n}{\operatorname{argmax}} H(M) - H(M | Z). \quad (3)$$

The conditional entropy $H(M | Z)$ of the map M given the set Z of measurements is the expected value, over the space of all possible measurements, of the conditional entropy of the map given the individual measurements z , leading to

$$H(M | Z) = \int_z p(z) H(M | Z = z) dz. \quad (4)$$

4.2. Efficiently approximating mutual information

Computing the map entropy $H(M | Z)$ given in Equation (4) is infeasible in general since integrating over the space of all possible combinations of up to n laser measurements is practically impossible. Furthermore, computing the map entropy $H(M | Z = z)$ given a set of measurements z requires model assumptions about the world.

We therefore make the following assumptions to efficiently compute the entropy $H(M | Z)$. We ignore the distribution over \mathbf{x} and operate on the most likely estimate, which is given in Equation (1). Like most approaches in robotics, we furthermore assume that the laser measurements and the individual laser beams are independent and we model the map M as a standard occupancy grid map. An occupancy grid map is a grid of independent discrete binary random variables C that take the values $\text{Val}(C) = \{\text{'free'}, \text{'occupied'}\}$. The entropy of an occupancy grid map M given a set of measurements z is then given by

$$H(M | Z = z) = \sum_{C \in M} H(C | Z = z) \quad (5)$$

$$= - \sum_{C \in M} \sum_{c \in \text{Val}(C)} P(C = c | z) \log P(C = c | z). \quad (6)$$

Furthermore, we need to estimate the probability of sensing an object at a certain range given no further information about the environment. One option is to assume that the probability of sensing an object decreases with range, similar to most works on robot localization. The prior probability of the range of the j th beam of measurement z_i , denoted as z_i^j , without any knowledge of the map M can then be described by the exponential distribution

$$p(z_i^j) = \begin{cases} \eta \lambda e^{-\lambda z_i^j} & z_i^j \leq z_{\max}, \\ 0 & z_i^j > z_{\max}, \end{cases} \quad (7)$$

where z_{\max} denotes the maximum range of the scanner, λ is a parameter of the measurement model, and η is a normalizing constant.

Alternatively, the distribution $p(z_i^j)$ can be inferred from data. To do so, we used our own datasets as well as several publicly available datasets from the Radish dataset repository (see Howard and Roy 2003) and computed statistics about laser range measurements averaged over all beams, scans, and datasets. The resulting distribution is shown in Figure 2. The plots suggest that for range readings larger

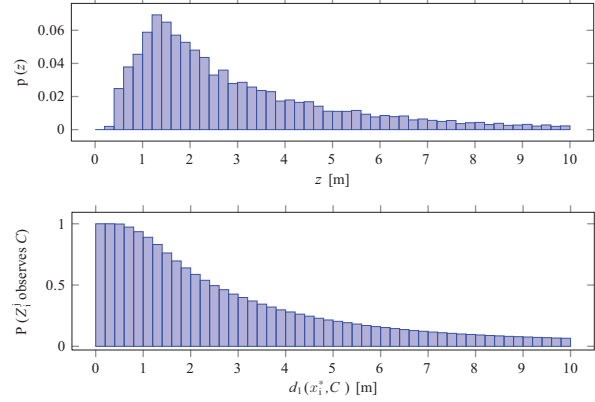


Fig. 2. Statistics of laser range measurements extracted from a series of robotic datasets. Top: Learned probability distribution $p(z)$ of the laser range readings. Bottom: Probability of measuring a cell C that is $d_1(x_i^*, C)$ meters away from the sensor location x_i^* .

than 1.5 m, the learned distribution is fairly similar to the exponential distribution given in Equation (7). In the experimental evaluation, we consider both models.

The aforementioned assumptions allow us to efficiently compute the conditional entropy $H(M | Z = z)$ of the occupancy grid map M given a measurement z . To facilitate the computation of the entropy $H(M | Z)$ defined in Equation (4), we propose to consider the measurement outcomes of a laser beam with respect to a particular grid cell, instead of integrating over the space of all possible range measurements. There are three possible measurement outcomes of a laser beam with respect to a particular grid cell that is located along the ray of the beam. The laser beam either traverses the cell and thus observes the cell as free, the laser beam ends in the cell and thus observes the cell as occupied, or the laser beam does not observe the cell. Given no prior map information, the probability distribution of the outcome can be computed by integrating over the density $p(z_i^j)$, leading to

$$P(Z_i^j \text{ does not observe } C) = \int_0^{d_1(x_i^*, C)} p(z_i^j) dz_i^j \quad (8)$$

$$P(Z_i^j \text{ observes } C \text{ as occupied}) = \int_{d_1(x_i^*, C)}^{d_2(x_i^*, C)} p(z_i^j) dz_i^j \quad (9)$$

$$P(Z_i^j \text{ observes } C \text{ as free}) = \int_{d_2(x_i^*, C)}^{z_{\max}} p(z_i^j) dz_i^j, \quad (10)$$

where $d_1(x_i^*, C)$ is the distance between the pose x_i^* from which the laser scan Z_i is taken and the closest border of the grid cell C in the direction of the j th beam, i.e. the border where the beam enters the cell. Similarly, $d_2(x_i^*, C)$ is the distance to the border of the grid cell C in the direction of the j th beam that is furthest away from x_i^* , i.e. the border where the beam leaves the cell.

Exploiting the probability distribution of the measurement outcome expressed in Equations (8)–(10) allows us

to avoid integrating over the space of all measurements z . Instead, we consider the effects of the measurements on the estimates of the grid cells, yielding the mutual information

$$I(C; Z) = H(C) - \sum_{z' \in \mathcal{A}_Z} P(z') H(C | z') \quad (11)$$

of the grid cell C and the set Z of laser measurements. \mathcal{A}_Z is the set of all possible measurement outcomes z' with respect to the grid cell C of all laser scans that are recorded close enough to potentially measure the cell.

In general, the number of possible combinations of grid cell measurement outcomes is exponential in the number k of laser scans, which is illustrated in Figure 3. It is therefore infeasible to enumerate all of the combinations in a tree. However, our approach relies on a standard inverse measurement model, $p(c | z_i^j)$, for laser range scanners that updates each cell using the values l_{free} , l_{occ} , and l_0 . Most importantly, these values do not depend on the sensed range. Therefore, the effect of a set of observations on a particular cell exclusively depends on the number of free and occupied observations, i.e. the histogram of measurement outcomes. Furthermore, the effect does not depend on the order in which the measurements were obtained. This allows us to combine all nodes in the tree that represent the same histogram of measurement outcomes. See Figure 3 for an illustration. The number of measurement outcomes is given by the number of ways to sample k elements from the set of three laser measurement outcomes with replacement and disregarding different orderings, leading to

$$\# \text{outcomes}(k) = \sum_{i=0}^k \left(\sum_{j=0}^{k-i} 1 \right) = \frac{k^2}{2} + \frac{3}{2}k + 1 \in \mathcal{O}(k^2), \quad (12)$$

which is quadratic in k .

Note that the probability of each of these outcomes would be given by a multinomial distribution if the free and occupied measurements were equally likely for all laser scans. However, Figure 2 suggests that the probability of observing a cell depends on the distance from the sensor to the cell. Algorithm 1 computes the probabilities of the outcomes by propagating the probability mass through the graph, using a hash table $P(\langle \cdot, \cdot, \cdot \rangle)$ that is indexed by the measurement histograms. The number of probabilities that the algorithm needs to consider is given by

$$3 \sum_{i=1}^k \# \text{outcomes}(i) = \frac{3}{2}k^3 + \frac{9}{2}k^2 + 3k \in \mathcal{O}(k^3), \quad (13)$$

leading to cubic complexity in the number k of measurements that potentially observe the grid cell C .

The number k of scans that the algorithm needs to consider is typically small because our technique discards laser scans online while building the graph (see Section 4.3). Furthermore, the maximum measurement range of laser scanners restricts the set of scans that can measure a cell. We can

Algorithm 1 Compute the probabilities of the measurement outcomes with respect to a grid cell.

Input: set of k laser measurements Z , cell C

Output: probabilities $P(\langle \cdot, \cdot, \cdot \rangle)$ of the outcomes (free, occupied, not observed)

$P(\langle 0, 0, 0 \rangle) = 1$

for $r = 1 \dots k$ **do**

for all $\langle f, o, u \rangle$ with $f + o + u == r$ **do**

$P(\langle f + 1, o, u \rangle) \leftarrow$

$P(\langle f, o, u \rangle) P(Z_i^f \text{ observes } C \text{ as free});$

$P(\langle f, o + 1, u \rangle) \leftarrow$

$P(\langle f, o, u \rangle) P(Z_i^f \text{ observes } C \text{ as occupied});$

$P(\langle f, o, u + 1 \rangle) \leftarrow$

$P(\langle f, o, u \rangle) P(Z_i^f \text{ does not observe } C);$

end for

end for

return $P(\langle \cdot, \cdot, \cdot \rangle)$

reduce the computational burden by only considering the l laser scans that most affect the mutual information between the measurements and the grid cell, i.e. the scans with the highest likelihood of measuring C . These are the l scans that are recorded closest to C . The approximation of using only l instead of k scans is appealing because it yields a complexity that is linear in k since choosing these l scans is linear in k . However, the approximation neglects a subset of the scans when computing the mutual information between the measurements and the grid cell.

Finally, the mutual information $I(M; Z)$ of the map M and the set Z of laser scans is given by

$$I(M; Z) = \sum_{C \in M} I(C; Z). \quad (14)$$

All terms needed to compute the subset Z^* of laser scans that maximizes the mutual information of the map M and the subset Z^* given in Equation (3) are specified and can be approximated efficiently.

4.3. Greedily discarding laser scans online

Computing the optimal subset Z^* of n measurements that most reduces the uncertainty of an estimate of the state of the world has been shown by Krause and Guestrin (2005) to be at least NP-hard. Fortunately, the problem is sub-modular. Hence, greedily selecting measurements results in obtaining a set of measurements that is at most a constant factor $(1 - e^{-1} \approx 0.63)$ worse than the optimal set. Motivated by this insight, our approach estimates the subset Z^* by successively discarding laser scans. In each iteration, it identifies and discards the laser scan that is expected to be least informative.

Our approach is able to restrict the size of the pose graph, resulting in an any-space SLAM system. Alternatively, our approach is capable of discarding uninformative laser scans by setting a threshold on the expected information gain of

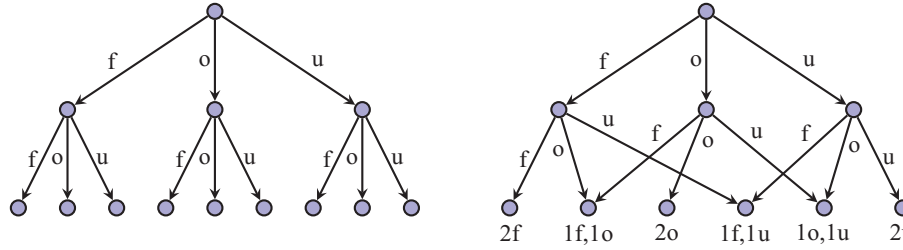


Fig. 3. Space of measurement outcomes of a grid cell given $k = 2$ laser scans, where free is denoted by f, occupied by o, and unknown by u. Left: In general, there are 3^k measurement outcomes, which is exponential in k . Right: The number of measurement outcomes is quadratic in k when taking into account the measurement model.

Algorithm 2 Compress the pose graph such that it has at most n nodes and laser scans.

```

while number of laser scans > n do
  for all remaining laser scans  $Z_i \in Z$  do
     $IG_{Z_i} \leftarrow I(M; Z) - I(M; Z \setminus \{Z_i\})$  // see Equation (14)
    // and Algorithm 1
  end for
   $\tilde{Z} \leftarrow \operatorname{argmin}_{Z_i} IG_{Z_i}$ 
   $Z \leftarrow Z \setminus \{\tilde{Z}\}$  // discard laser scan  $\tilde{Z}$ 
  marginalize out node corresponding to  $\tilde{Z}$  // see Section 5
end while

```

Algorithm 3 Discard laser scans less informative than ε and marginalize out the corresponding pose nodes.

```

repeat
  for all remaining laser scans  $Z_i \in Z$  do
     $IG_{Z_i} \leftarrow I(M; Z) - I(M; Z \setminus \{Z_i\})$  // see Equation (14)
    // and Algorithm 1
  end for
   $\tilde{Z} \leftarrow \operatorname{argmin}_{Z_i} IG_{Z_i}$ 
  if  $IG_{\tilde{Z}} < \varepsilon$  then
     $Z \leftarrow Z \setminus \{\tilde{Z}\}$  // discard laser scan  $\tilde{Z}$ 
    marginalize out node corresponding to  $\tilde{Z}$  // see
    // Section 5
  end if
until  $IG_{\tilde{Z}} \geq \varepsilon$ 

```

laser scans. Pseudo-code of our greedy approach to online pose graph compression, as laser scans are acquired by the robot, is sketched in Algorithm 2 and 3. In practice, most of the mutual information values $I(C, Z)$ do not change when incorporating new laser scans. Therefore, these values can be cached and do not need to be recomputed in each step. The algorithm only needs to recompute $I(C; Z)$ for cells that are in the measurement range of laser scans that are incorporated into or discarded from the belief.

5. Maintaining a sparse pose graph

The previous section presented our method to estimate the mutual information between the laser measurements and the

grid map to identify laser scans that provide little information about the map. Our approach subsequently marginalizes out the nodes from the pose graph that correspond to the discarded laser scans. This section presents an approximate marginalization technique that maintains a sparse pose graph that allows for efficient map optimization.

A pose graph can be seen as a Gaussian Markov random field (GMRF) that models the belief of the robot. In this view, each pose is a random variable that is represented as one node in the GMRF and each constraint between two poses in the pose graph is a binary potential between the nodes in the GMRF. Marginalizing out a pose node from the graph implies summarizing the information stored in the edges that connect that node and storing that information in the edges between nodes that are kept. The main problem of exact marginalization is that it introduces new edges between all pairs of variables that are directly related to the eliminated variables, adding a so-called elimination clique to the graph. This destroys the natural sparsity pattern that typically occurs in SLAM. Our approach aims at preserving sparsity by means of approximate marginalization while, at the same time, seeking to minimize the loss of information. Here, we assume measurements that yield full rank constraints in the pose graph, such as the constraints obtained from laser range data, and refer the reader to the work by Cheeseman and Smith (1986) for a discussion on how to compound and merge geometric constraints with associated errors.

5.1. Marginalization ceases sparsity

The GMRF framework helps us understand the effects of marginalization on the belief of the robot. Suppose the belief is given by the joint Gaussian probability distribution $p(\alpha, \beta)$ over two stacked vectors α and β which together comprise the poses of the robot. In information form, the belief $p(\alpha, \beta)$ can be expressed as

$$p(\alpha, \beta) = \mathcal{N}^{-1} \left(\begin{bmatrix} \xi_\alpha \\ \xi_\beta \end{bmatrix}, \begin{bmatrix} \Omega_{\alpha\alpha} & \Omega_{\alpha\beta} \\ \Omega_{\beta\alpha} & \Omega_{\beta\beta} \end{bmatrix} \right), \quad (15)$$

where ξ are information vectors and Ω information matrices. To remove β from the belief, we compute the marginal density $p(\alpha) = \int p(\alpha, \beta) d\beta$. In information

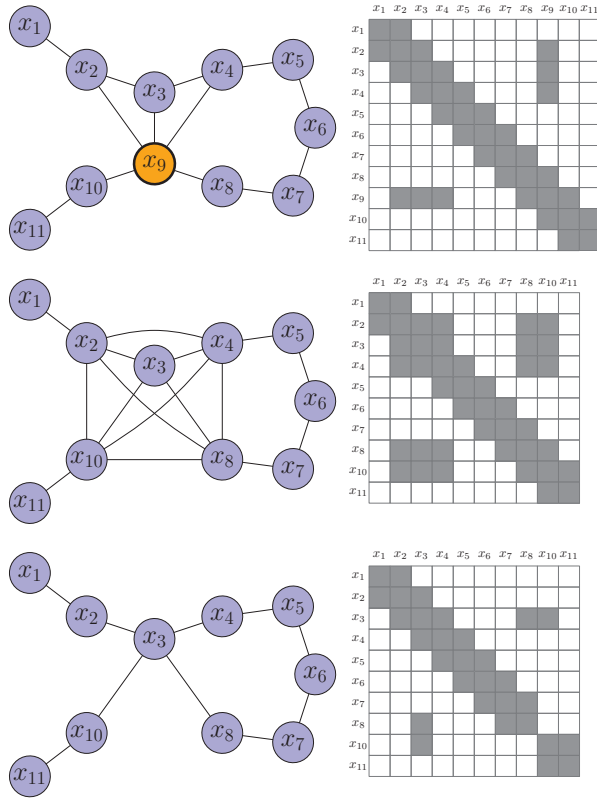


Fig. 4. Eliminating a pose node from the belief of the robot. The left column depicts the GMRFs and the right column shows the corresponding information matrices, where gray shades indicate non-zero values. Top: x_9 is slated for removal. Middle: Belief after having marginalized out node x_9 . The former neighbors of x_9 form an elimination clique making the graph dense. Bottom: Belief resulting from Chow–Liu tree approximation of the elimination clique.

form, marginalization requires computing the Schur complement over the variables which should be kept. Hence, the parameters of $p(\alpha)$ are

$$\xi = \xi_\alpha - \Omega_{\alpha\beta} \Omega_{\beta\beta}^{-1} \xi_\beta, \quad (16)$$

$$\Omega = \Omega_{\alpha\alpha} - \Omega_{\alpha\beta} \Omega_{\beta\beta}^{-1} \Omega_{\beta\alpha}. \quad (17)$$

Unfortunately, the Schur complement in Equation (17) introduces new constraints between all pairs of variables that are directly related to the eliminated variables, adding a so-called elimination clique to the graph. This fill-in destroys the natural sparsity pattern that is typical to SLAM problems. See Eustice et al. (2006) for more details and Figure 4 for an illustration.

Representing a dense matrix typically requires significantly more memory resources than representing a sparse matrix. Furthermore, the density adversely affects the computational costs of subsequent marginalizations and of the underlying pose graph optimization. In the worst case, the number of dependencies increases quadratically with the number of variables. Hence, exact marginalization ceases

the sparsity of the belief and therefore introduces a complexity that is not suited for long-term map learning.

5.2. Approximate marginalization

We propose to reduce the number of edges in the elimination cliques, that emerge when marginalizing out pose nodes. We motivate this technique by the fact that this part of the graph has just become dense due to the fill-in resulting from marginalization.

To measure the effect of the approximation, we make use of the relative entropy or Kullback–Leibler divergence, which is a standard measure of the difference between two probability distributions p and q . It is given by

$$D_{\text{KL}}(p \parallel q) = \int_x p(x) \log \frac{p(x)}{q(x)} dx. \quad (18)$$

Let $\tilde{\mathbf{x}} = (x_1^\top, \dots, x_k^\top)^\top$ be the stacked vector of the set of variables in the elimination clique. Furthermore, let $p(\tilde{\mathbf{x}})$ be the density which arises from the constraints within the clique. The key idea of our approach is to reduce the computational burden by approximating the density $p(\tilde{\mathbf{x}})$ with a distribution that has fewer conditional dependencies. This implies treating some pairs of variables as conditionally independent.

5.3. Chow–Liu tree approximation

We propose to reduce the number of constraints in the elimination cliques by locally approximating the density $p(\tilde{\mathbf{x}})$ with a probability distribution $q(\tilde{\mathbf{x}})$ such that each variable is only conditioned upon one of the other variables:

$$p(\tilde{\mathbf{x}}) = p(x_k) \prod_{i=1}^{k-1} p(x_i \mid x_{i+1}, \dots, x_k) \quad (19)$$

$$\approx p(x_k) \prod_{i=1}^{k-1} p(x_i \mid x_{i+1}) = q(\tilde{\mathbf{x}}). \quad (20)$$

Consequently, the graphical model of $q(\tilde{\mathbf{x}})$ is tree-shaped. Note that we do not propose to transform the entire graph into a tree, but only the resulting elimination clique. Importantly, our approximation preserves the global graph structure and, in particular, global loop closures.

Chow and Liu (1968) propose to approximate a probability distribution $p(\tilde{\mathbf{x}})$ by a distribution $q_{\text{opt}}(\tilde{\mathbf{x}})$ such that each variable is conditioned upon at most one other variable and such that the Kullback–Leibler divergence between $p(\tilde{\mathbf{x}})$ and $q_{\text{opt}}(\tilde{\mathbf{x}})$ is minimized. Let the mutual information graph of a probability distribution be a graph such that each edge between two nodes x_i and x_j has weight equal to the mutual information $I(x_i; x_j)$ of these variables. Chow and Liu proved that the optimal approximation $q_{\text{opt}}(\tilde{\mathbf{x}})$ with first-order tree dependence to the probability distribution $p(\tilde{\mathbf{x}})$ has the same structure as the maximum-weight spanning tree of the mutual information graph of $p(\tilde{\mathbf{x}})$.

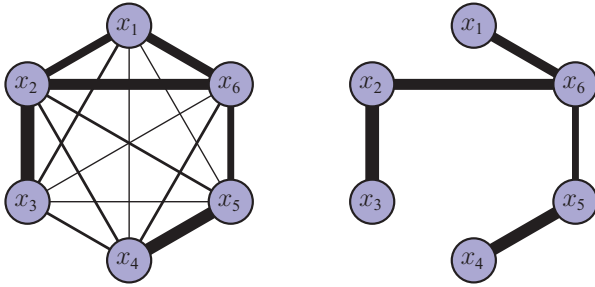


Fig. 5. Chow-Liu tree approximation. Left: Graphical model of a probability distribution, where the mutual information between the pairs of random variables is indicated by the thickness of the edges. Right: Chow-Liu tree approximation.

The maximum-weight spanning tree of the mutual information graph can be computed efficiently using Kruskal's algorithm (see Figure 5 for an illustration). As shown by Davison (2005), the mutual information $I(x_i; x_j)$ of two normally distributed random variables x_i and x_j is given by

$$I(x_i; x_j) = \frac{1}{2} \log_2 \left(\frac{|\tilde{\Sigma}_{ii}|}{|\tilde{\Sigma}_{ii} - \tilde{\Sigma}_{ij} \tilde{\Sigma}_{jj}^{-1} \tilde{\Sigma}_{ji}|} \right), \quad (21)$$

where $\tilde{\Sigma}_{ij}$ refers to the entry in the covariance matrix $\tilde{\Sigma}$ that relates x_i and x_j .

The covariance matrix $\tilde{\Sigma}$ of the elimination clique is obtained by marginalizing the inverse of the information matrix Ω of the linearized system. However, to avoid the costly matrix inversion, we approximate $\tilde{\Sigma}$ by marginalizing Ω and then inverting the resulting sub-matrix. This approximation of the clique covariance matrix neglects the constraints in the pose graph that are in the remainder of the pose graph. Consequently, the tree-shaped approximate marginalization is local to the elimination clique. Inverting the sub-matrix is efficient since it is typically small due to the sparsity of the graph. The tree-shaped approximation allows us to use the elimination clique distribution, which is in information form, to compute the non-linear constraints that are kept in our pose graph. We compute these non-linear constraints using the corresponding marginals of the clique distribution. Summing up, our technique allows for efficient marginalization and preserves sparsity in the pose graph while seeking to minimize the loss of information.

6. Experimental evaluation

To evaluate the presented approach, we carried out several experiments using a real ActivMedia Pioneer-2 robot equipped with a SICK laser range finder. In addition to that, we applied our method to popular SLAM datasets. We compare our approach with the performance of the same technique when no scans are discarded, which is referred to as the 'standard approach'. Our experiments demonstrate that our approach to pose graph compression is well suited for

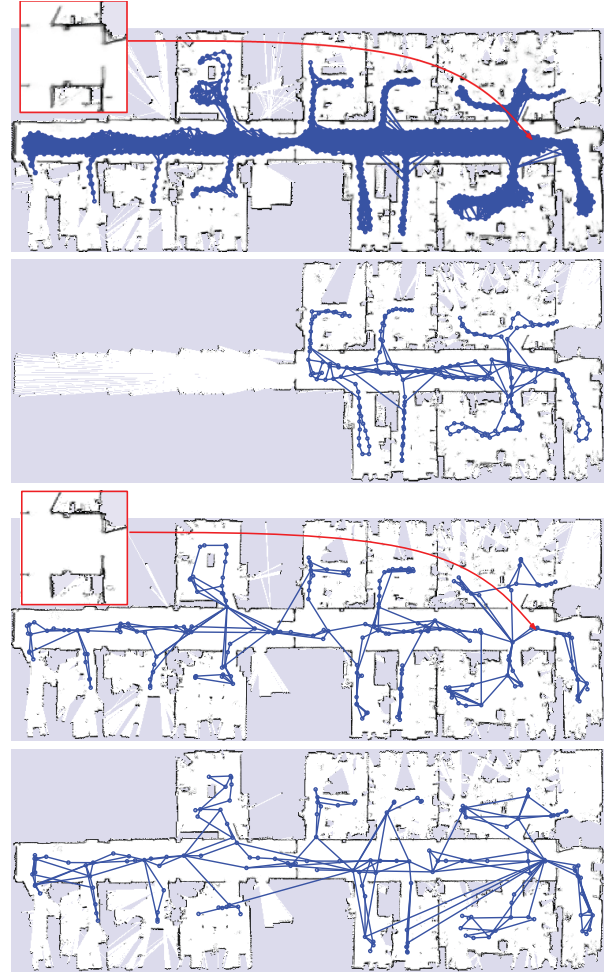


Fig. 6. The robot moved around in an office environment for an extended period of time, visiting the rooms and the corridor many times. First: Standard approach, 2,597 laser scans, 15,695 edges. Second: Our approach at an intermediate time step, 200 laser scans, 264 edges. Third: Our approach, 200 laser scans, 315 edges. Fourth: Our approach when setting a threshold for the mutual information, 148 laser scans, 250 edges instead of a fixed value for n .

long-term mobile robot mapping and for standard SLAM problems.

We evaluated our approach using four datasets. We used a self-recorded dataset in which the robot traveled in our lab environment for an extended period of time (see Figure 6) and a previously recorded dataset from a computer science building in Freiburg, which was also used by Grisetti et al. (2007) (see Figure 7). Furthermore, we used the FHW dataset (see Figure 1) and the Intel Research Lab dataset (see Figure 8) by Dirk Hähnel.

Figure 6 shows four maps resulting from our experiments. The first image depicts the map and the pose graph obtained by the standard approach. The second and the third image show the maps and pose graphs resulting from our approach when restricting the size of the pose graph to 200 nodes. More precisely, the second image depicts the state

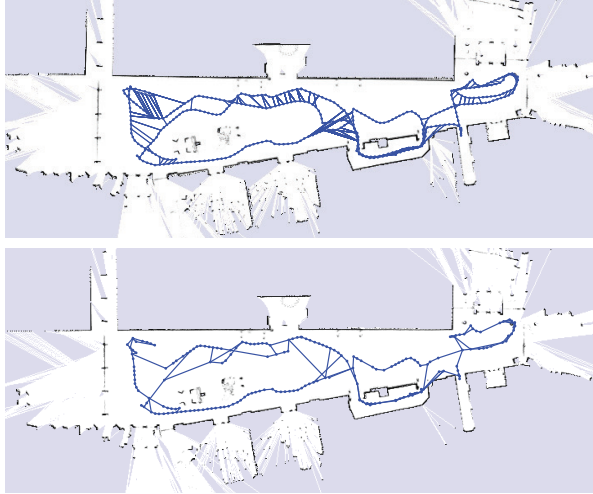


Fig. 7. Map and pose graph obtained when mapping the FR101 dataset. Top: Standard approach, 408 nodes, 723 edges. Bottom: Our approach, 200 nodes, 246 edges. Since the robot does not frequently re-traverse previously visited terrain, few scans were discarded.



Fig. 8. Map and pose graph obtained when mapping the Intel Research Lab. Left: Standard approach, 1,802 laser scans, 3,916 edges. Right: Our approach preserves the sparsity of the pose graph, 250 laser scans, 349 edges. Arrows indicate locally blurred areas or small alignment errors in the map obtained by the standard approach. In contrast to our approximate marginalization technique, exact marginalization would result in 13,052 edges.

of our approach before the robot has entered the left side of the corridor. The 200 nodes are used to model the right part only. The third image shows the pose graph modeling the entire environment. Note that our approach has redistributed the nodes in the environment, still complying with the limit of 200 nodes. Finally, the fourth image shows the map when setting a threshold on the mutual information.

6.1. Approximate marginalization

We quantitatively evaluated the estimate of the pose graphs generated by our approach to graph compression. We analyzed the mean and uncertainty estimates of the robot's poses and compared them to the corresponding estimates generated by the standard approach. Figure 9 depicts the

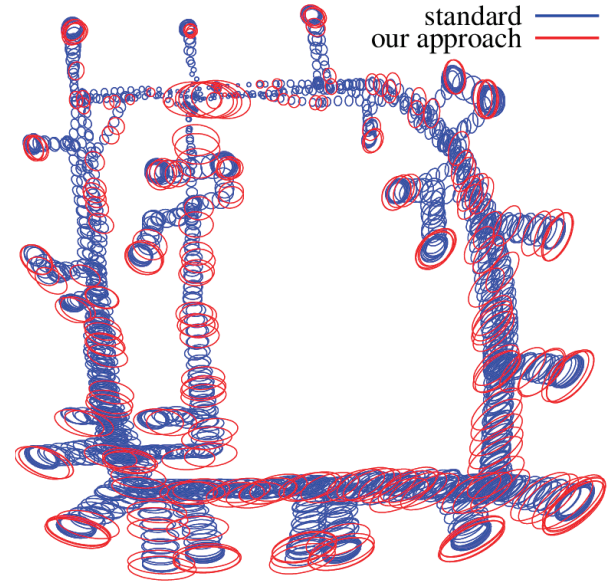


Fig. 9. Intel Research Lab: 3σ covariance ellipses of the poses. Blue: Standard approach. Red: Our approach. The estimates of our approach are typically larger since less observations are used.

3σ covariance ellipses of the pose estimates. In this experiment, our approach kept less than 9% (349 of 3,916) of the edges of the original pose graph. As a result, 2.8% of the probability mass of the original pose graph is not covered by our approximation, and the covariance estimates are 41% more conservative.

6.2. Measurement probability distribution

We briefly analyzed the effects on the compressed pose graph when using different measurement probability distributions $p(z)$. We performed the same mapping experiment using the exponential probability distribution given in Equation (7) as well as using the probability distribution that was learned from data (see Figure 2). Since the exponential function and the learned function are similar except of the first 1.5 m, we expected that the effect on the resulting pose graph and the grid map is small. As can be seen in Figure 10, different nodes have been selected, but the overall structure of both pose graphs is similar in terms of topology and local density of the nodes. The same applies to the resulting occupancy grid maps. Given our experiments, we cannot argue that one approach outperforms the other.

6.3. Memory requirements and runtime

Our approach to pose graph compression causes a computation time overhead that only pays off when the robot frequently re-traverses previously visited parts of the environment. We evaluated the memory requirements of our approach in terms of the size of the resulting pose graphs and measured the runtime of our SLAM system on a

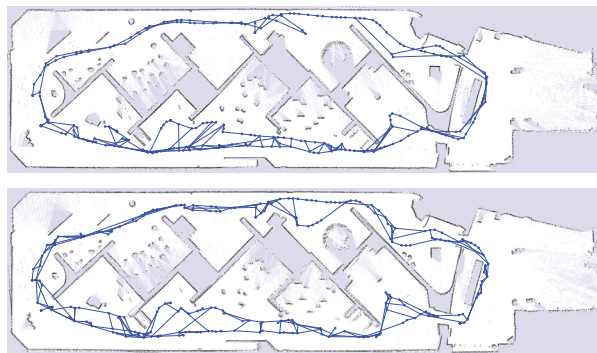


Fig. 10. Maps and pose graphs obtained using different measurement probability distributions $p(z)$ when mapping the FHW dataset. Top: The exponential probability distribution given in Equation (7). Bottom: The probability distribution learned from robotic datasets plotted in Figure 2. The resulting node densities are fairly similar in both settings.

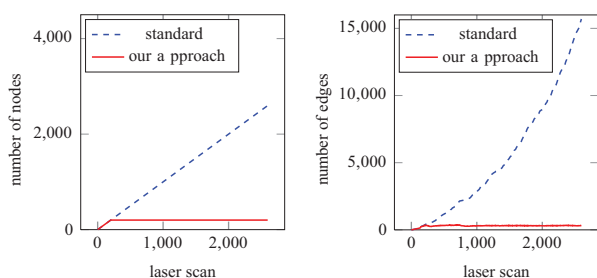


Fig. 11. Size of the pose graph when moving around in an office environment for an extended period of time. In this experiment, our approach restricted the total number of nodes to 200.

2.6 GHz Intel CPU system. Note that our current implementation is not optimized for speed. A significant speedup can be expected by caching intermediate results (see Section 4.3).

When the robot stays in the same environment and rarely explores new terrain, the loop closing component of the SLAM front-end, which performs scan matching to find constraints between the current scan and all former scans that were recorded in the vicinity of the robot, needs to consider an ever-increasing number of nodes in each step. Furthermore, the pose graph optimization scales with the number of edges, which in this case increases roughly quadratically. In contrast to the standard approach, our approach compresses the pose graph such that the number of nodes in the graph remains low.

In the experiment depicted in Figure 6, the robot moved around in our lab environment for an extended period of time. The corresponding plots shown in Figure 11 and 12 clearly suggest that the experiment leads to an explosion in terms of memory requirements when using the standard approach. In the experiment depicted in Figure 11, the size of the graph was restricted to 200 nodes. In the experiment corresponding to Figure 12, a threshold on the mutual information was set and, as a result, the complexity does

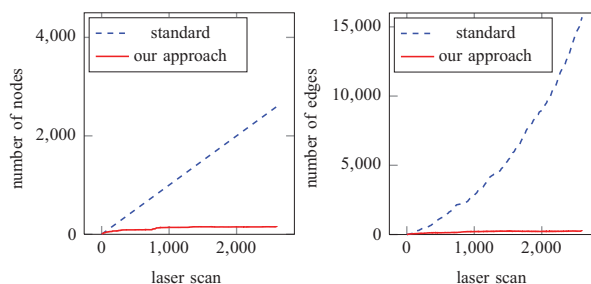


Fig. 12. Size of the pose graph when moving around in an office environment for an extended period of time. In this experiment, our algorithm discarded all laser scans whose expected information gain was below a given threshold.

Table 1. Percentage of cells that changed their most likely state when applying our compression approach.

Dataset	Number of scans	Number of scans (our approach)	Changed grid cells (%)
Intel Research Lab	1,802	349	0.9%
FR079 long term	2,597	148	1.6%
FHW museum	2,049	250	1.2%

not grow as long as the robot does not explore new terrain. Our compression method kept 148 of 2,597 laser scans. The graph generated by our approach has 250 edges, whereas the graph obtained by the standard approach has 15,695 edges. The standard approach required a total of 129 min to process the dataset. Despite the computation time overhead, the SLAM system only needed 87 min when applying our graph compression technique.

When mapping the FHW dataset (see Figure 1), our approach retained 250 of 2,049 laser scans and 324 of 7,190 edges. The standard approach required 32 min, whereas our approach needed 17 min. When mapping the Intel Research Lab dataset (see Figure 8), our approach kept 250 of 1802 laser scans and 349 of 3916 edges. The standard approach required 9 min, whereas our approach needed 26 min. When mapping the Intel Research Lab dataset, the computation time overhead of our method does not pay off because the robot mostly explores new terrain and only revisits the comparably small corridor.

6.4. Effects on the most likely grid map

We furthermore analyzed the effects of our pose graph compression approach on the resulting occupancy grid maps. We therefore compared the maps at a resolution of 10 cm and counted the number of cells that changed their most likely state (free, occupied, and unknown) when applying our compression technique.

We evaluated the grid maps using three different datasets. Table 1 summarizes the results. When mapping the Intel Research Lab, our pruning approach retained 349 of 1,802

laser scans, but only 0.9% of the cells changed. In the long-term experiment, our method kept 148 of 2,597 laser scans, and 1.6% of the cells changed. When mapping the FHW, our approach maintained 250 of 2,049 scans, and 1.2% of the cells changed. These results suggest that the changes in the most likely maps are small. Most of the cells that changed their most likely state were seen as free before applying our compression method and are seen as unknown after applying our technique because no beam covers the cells anymore. This effect can, for example, be seen in Figure 1 where some of the cells in the partially visited room at the right-hand side of the map are unobserved.

6.5. Scan matching and grid map quality

We furthermore evaluated the effects of our online pose graph compression technique on the quality of the resulting grid maps. In theory, the more observations are available, the better the estimate. Ignoring measurements will lead to a belief with higher uncertainty. However, today's occupancy grid-based mapping approaches typically rely on some kind of scan alignment techniques such as scan matching, which aligns the scans to extract spatial constraints. These systems have disadvantages when it comes to long-term mapping. Whenever the robot obtains a laser measurement, the scan matcher aims at aligning the new scan with existing scans in order to solve the data association problem. The probability that the scan matcher thereby makes a small alignment error is non-zero. A scan that is incorporated at a slightly wrong position blurs the map. As a result, the probability that the scan alignment procedure poorly aligns subsequent scans increases since matching is performed with misaligned scans. Hence, the probability of making alignment errors increases with the number of incorporated scans. In the long run, the map tends to become increasingly blurred and the mapping approach can diverge.

Figures 6 and 8 depict the maps and graphs obtained from the long-term experiments conducted in our office environment and the Intel Research Lab dataset. The grid maps generated by the standard approach exhibit visibly more blur in several parts of the maps, which can be attributed to the scan alignment procedure. In general, the more often the robot re-traverses already visited terrain, the more blur is added to the maps. In contrast to the standard approach, our method produces maps with less blurred obstacle boundaries, particularly when the robot frequently re-traverses already visited places.

Although we do not claim that such maps are better estimates of the environment from an information-theoretic point of view, our compression technique facilitates scan matching and therefore mitigates the risk of divergence in the mapping process. All in all, the maps resulting from our graph compression technique are well suited for navigation tasks such as localization and path planning.

7. Conclusion

In this paper, we presented a method for information-theoretic compression of pose graphs in graph-based SLAM. This is relevant for long-term mapping applications. Our approach seeks to select the most informative set of laser scans and allows for restricting the size of the pose graph either based on a memory limit, resulting in an any-space mapping system, or based on a threshold on the minimum amount of information that a laser scan is expected to provide. Our approach maximizes the mutual information between the laser measurements and the resulting occupancy grid map. When marginalizing out pose nodes from the graph, exact marginalization leads to a densely connected graph. To preserve the sparsity of the pose graph, we proposed to approximate the elimination cliques by local Chow–Liu trees, seeking to minimize the loss of information. Real world experiments illustrate the effectiveness of our method for computing compressed pose graphs in the context of graph-based SLAM.

Funding

This work has partly been supported by the DFG (grant number SFB/TR-8), the EC (grant number FP7-260026-TAPAS), and Microsoft Research, Redmond, WA.

Acknowledgements

We would like to thank Giorgio Grisetti and Maximilian Beinhofer for fruitful discussions. Furthermore, we thank Dirk Hähnel for providing the FHW and the Intel Research Lab datasets.

References

- Bachrach A, Prentice S, He R and Roy N (2011) RANGE - robust autonomous navigation in GPS-denied environments. *Journal of Field Robotics* 28: 644–666.
- Bosse M, Newman PM, Leonard JJ and Teller S (2003) An ATLAS framework for scalable mapping. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 1899–1906.
- Cheeseman P and Smith P (1986) On the representation and estimation of spatial uncertainty. *The International Journal of Robotics Research* 5: 56–68.
- Chli M and Davison AJ (2009) Automatically and efficiently inferring the hierarchical structure of visual maps. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Kobe, Japan, pp. 387–394.
- Chow C and Liu C (1968) Approximating discrete probability distributions with dependence trees. *IEEE Transactions on Information Theory* 14: 462–467.
- Cummins M and Newman P (2008) FAB-MAP: Probabilistic localization and mapping in the space of appearance. *The International Journal of Robotics Research* 27: 647–665.
- Davison A (2005) Active search for real-time vision. In *Proceedings of the International Conference on Computer Vision (ICCV)*, vol. 1, Beijing, China.
- Eade E, Fong P and Munich M (2010) Monocular graph SLAM with complexity reduction. In *Proceedings of the IEEE/RSJ*

- International Conference on Intelligent Robots and Systems (IROS)*, Taipei, Taiwan, pp. 3017–3024.
- Estrada C, Neira J and Tardós J (2005) Hierarchical SLAM: Real-time accurate mapping of large environments. *IEEE Transactions on Robotics* 21: 588–596.
- Eustice R, Singh H and Leonard J (2006) Exactly sparse delayed-state filters for view-based SLAM. *IEEE Transactions on Robotics* 22: 1100–1114.
- Folkesson J, Jensfelt P and Christensen H (2005) Vision SLAM in the measurement subspace. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Barcelona, Spain, pp. 325–330.
- Grisetti G, Kümmerle R, Stachniss C, Frese U and Hertzberg C (2010). Hierarchical optimization on manifolds for online 2D and 3D mapping. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Anchorage, AK, pp. 273–278.
- Grisetti G, Stachniss C and Burgard W (2007) Improved techniques for grid mapping with Rao–Blackwellized particle filters. *IEEE Transactions on Robotics* 23: 34–46.
- He R, Prentice S and Roy N (2008) Planning in information space for a quadrotor helicopter in a gps-denied environments. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Los Angeles, CA, pp. 1814–1820.
- Howard A and Roy N (2003) The robotics data set repository (Radish). [Http://radish.sourceforge.net/](http://radish.sourceforge.net/).
- Ila V, Porta J and Andrade-Cetto J (2010) Information-based compact pose SLAM. *IEEE Transactions on Robotics* 26: 78–93.
- Kaess M and Dellaert F (2009) Covariance recovery from a square root information matrix for data association. *Journal of Robotics and Autonomous Systems* 57: 1198–1210.
- Kim A and Eustice R (2011) Combined visually and geometrically informative link hypothesis for pose-graph visual SLAM using bag-of-words. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, San Francisco, CA, pp. 1647–1654.
- Kollar T and Roy N (2008) Efficient optimization of information-theoretic exploration in slam. In *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, Chicago, IL, pp. 1369–1375.
- Konolige K and Agrawal M (2008) FrameSLAM: From bundle adjustment to realtime visual mapping. *IEEE Transactions on Robotics* 24: 1066–1077.
- Konolige K and Bowman J (2009) Towards lifelong visual maps. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, St. Louis, MO, USA, pp. 1156–1163.
- Krause A and Guestrin C (2005) Near-optimal nonmyopic value of information in graphical models. In *Proceedings of Uncertainty in Artificial Intelligence (UAI)*.
- Kretschmar H, Stachniss C and Grisetti G (2011) Efficient information-theoretic graph pruning for graph-based SLAM with laser range finders. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, San Francisco, CA, USA, pp. 865–871.
- Lu F and Milios E (1997) Globally consistent range scan alignment for environment mapping. *Autonomous Robots* 4: 333–349.
- MacKay D (2003) *Information Theory, Inference, and Learning Algorithms*. Cambridge: Cambridge University Press.
- Ni K, Steedly D and Dellaert F (2007) Tectonic SAM: Exact, out-of-core, submap-based SLAM. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Rome, Italy, pp. 1678–1685.
- Olson E (2008) *Robust and Efficient Robotic Mapping*. Ph.D. thesis, MIT, Cambridge, MA, USA.
- Olson E (2009) Real-time correlative scan matching. In *Proceedings of the IEEE International Conference on Robotics & Automation (ICRA)*, Kobe, Japan, pp. 4387–4393.
- Snaveley N, Seitz S and Szeliski R (2008) Skeletal graphs for efficient structure from motion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Anchorage, AK, pp. 1–8.
- Stachniss C, Grisetti G and Burgard W (2005) Information gain-based exploration using Rao–Blackwellized particle filters. In *Proceedings of Robotics: Science and Systems (RSS)*, Cambridge, MA, USA, pp. 65–72.
- Stachniss C and Kretschmar H (2011). Pose graph compression for laser-based SLAM. In *Proceedings of the International Symposium of Robotics Research (ISRR)*, Flagstaff, AZ, USA.