Evaluation of a factorized ICP based method for 3D mapping and localization

Ondřej Jež

Abstract
This article is presenting a method for simultaneous localization and mapping (SLAM) of mobile robots in six degrees of freedom (DOF). The localization and mapping task is equal to the registration of multiple 3D images into a common frame of reference. For this purpose, a method based on the Iterative Closest Point (ICP) algorithm was developed. The SLAM method originally implemented in this research was using solely 6DOF ICP based registration. The computing effort and the registration quality issues of such solution were examined and in order to accelerate and improve the quality of the time-demanding 6DOF image registration, an extended method was developed. The major extension is the introduction of a factorized registration, extracting 2D representations of vertical objects called leveled maps from the 3D point sets, ensuring these representations are 3DOF invariant. The leveled maps are registered in 3DOF using ICP algorithm, allowing pre-alignment of the 3D data for the subsequent robust 6DOF ICP based registration. The extended method is presented in this article, followed by the evaluation using real 3D data acquired in different indoor environments, examining the benefits of the factorization and other extensions as well as the performance of the original ICP method. The factorization gives promising results compared to a single phase 6DOF registration in regularly structured environments. Also, the disadvantages of the method are discussed, proposing possible solutions.

Keywords: navigation, mobile robots, simultaneous localization and mapping, Iterative Closest Point algorithm

Introduction

The concept of navigation is as old as the human civilization; it is connected to human migration and the development of the means of transportation. In order to navigate themselves through miles of open ocean, the Polynesian sailors had to memorize extensive facts such as the motion of stars, weather influences and directions of swells. European explorers were initially using dead reckoning and slowly developed celestial navigation, using passive landmarks (planets and stars) to find ship's current position. Today's navigation methods used in transportation are mainly dependent on artificial landmarks (such as beacons) or even active systems of moving landmarks (satellites).

The development of navigation methods for mobile robots started in a similar way as the development of human navigation: first robots used dead reckoning for pose estimation in known environment; further development of the robotic navigation went hand in hand with the developments in sensing technologies. Today, mobile robots are equipped with a variety of sensing devices such as cameras, range sensors (ultrasound, optical, microwave), tactile sensors, gyros etc. which allow the solution of the task of localization and exploration. The state of technology makes it possible for us to advance further with the theory and strategy to navigate the robot in the environment. Current sensing devices are able to explore the surroundings of robots in such a precision and detail that the possibilities of using this information to navigate the robot are so far unused. In the current scientific discourse in the field of mobile robot's autonomy and navigation, the use of three dimensional environment perception is slowly breaking through. This is the area on which the research focuses, trying to bring a new strategy and make use of new technological potential.

The research documented in this article intended to cover some of the major challenges that are yet to be approached or further investigated in the navigation methods for mobile robots. The focus is on one of the most challenging ways of navigating mobile robots in unknown environments: using spatial 3D range measurements. The existing methods for localization and map building are very computing-power demanding especially due to the high data flow and difficult optimization tasks. Also quality issues are often to be resolved when working with real data. The work summarized in this article was aiming to innovate the state of the art in both aspects.

In this article, the navigation method will be first discussed, focusing mainly on the details of the innovation. Also, since the method is based on the use of the Iterative Closest Point algorithm, this algorithm will be also presented before proceeding to the evaluation results. The main objective of this article is to introduce the reader to the developed method and to present the experiments which were carried out to evaluate the method.

State of the art

The task of navigation in unknown environments implies three basic subtasks: *localization, mapping and path planning* [5]. These three subtasks correspond to the fact that we have to obtain information about the environment, find our position related to some frame of reference in the environment, and decide where we should and where we are able to move and how to do it. This is a very complicated process consisting of data acquisition and preprocessing and various optimization tasks such as data registration, occlusion detection, error diffusion and optimal path planning. In our case the environment is apart from global characteristics (such as gravity and other physical laws, possible presence of objects of known characteristics etc.) unknown. When designing a navigation system which operates in unknown environments, the first two subtasks – localization and mapping – can be treated as simultaneous processes. Since we have no prior map in which we are determining our position, in every moment we get some information about the environment, we update our maps and also our current position on this updated map. Thus the localization and map building form one procedure called simultaneous localization and mapping (SLAM). In case of three dimensional environments this task will be solved in six degrees of freedom (three translations and three rotations). Although the position of a mobile robot is expressed in six degrees of freedom, its possible trajectory is often limited by the terrain and ground constrains, thus allowing us only certain movements and achievable locations. This is what makes the task of path planning of wheeled vehicles more difficult than e.g. path planning of an aircraft; on the other hand the proximity of an observable environment allows the self-localization with respect to the environment structures [3].

When proceeding with SLAM in unknown environments, we have to measure new range information of the environment once in a while during robot's movement in order to keep track of the current position and to expand the current map. The transformation of this new set of information (image, range data) into the original coordinate system (in which the built map is referenced) is called image registration and is serving for both purposes of SLAM (we compare the new data to the reference map and consequently we integrate it). Data registration is the key step in the SLAM algorithm operation and in fact in this research, such registration method was developed and experimentally verified.

Methods for registering two 3D images

In order to explore the environment and to create a map which is representing the environment without occlusions, there have to be multiple scans taken from different perspectives. After each image is taken and the semantic analysis is performed, the data set has to be processed to expand the current map of the environment. This step is called image (scan) registration and the principle is that the current image has to be transformed into the original frame of reference. This step is based on finding the optimal rotation and translation for which the matching of the new scan into the global frame of reference with previous scans is most consistent. This means that the point clouds in the new scan referring to environment areas which appear also in previous scans are as close as possible to the corresponding point clouds in the map. This rotation and translation in fact corresponds to the change in robot's pose, thus the robot is localized in the global map of the environment after each image registration. The operation of scan registration (matching) is starting with initial values of the rotation and translation taken from the odometry of the vehicle, while the matching (registering) algorithm is from the localization point of view correcting the errors which have integrated from the vehicle's odometers measurements. Here we can actually observe the meaning of the simultaneous localization and mapping since image registration serves to both purposes. The registration procedure is in general a minimization algorithm, which minimizes the error caused by incorrect placement of new scans into the global map based on global frame of reference. The error function E can be in general expressed as in equation (1), where R and t are the rotation and translation variables, M is a model set of previously scanned points and D is a data set of new points from the current scan.

$$
\begin{aligned}
&E = f(R,t,M,D) \\
&M = \{m_1,...,m_i\}, D = \{d_1,...,d_j\} \qquad (1) \\
&i \in N, j \in N
\end{aligned}
$$

The data registration can be done in an automated way using correspondence based methods. The mainstream in the 3D data matching is the use if Iterative Closest Point algorithm (ICP) proposed in 1992 by P. Besl and N.D. McKay [2]. In general, the ICP method is minimizing the least-square point correspondence sum (see equation 0) proposed by Arun et al. in [1]. This is done though finding corresponding points in data sets and building a special matrix from which the optimum rotation is calculated. The determined rotation is applied on the data and the whole process is done iteratively until the best (optimum) match is reached when the calculated rotation matrix is close to ones matrix. The method decouples the calculation of translation from the estimation of the rotation matrix. The navigation method proposed in this paper is using the Iterative Closest Point algorithm and therefore this method will be presented in a greater detail in next section.

There are also other methods for minimizing the function expressed in equation 0. One interesting method is parallel evolutionary registration method proposed by Robertson and Fisher [12]. This method is using an evolutionary search while the chromosomes used in this search consist of six parameters: three angles representing the rotation of the system and three compounds of the translation. First the chromosomes are initialized using a random function. This is followed by the evaluation of the fit function, the results of which are used in the following phase: adapted crossover and mutation. In the algorithm, there are in fact parallel processes for the evaluation of the fit function accelerating the computation for the whole population, since this could be seen as a bottleneck of the algorithm. The criteria – fit function is in fact very similar to the one used in the ICP algorithm and comes from Arun's publication on Least-squares fitting of two 3-d point sets [1]. Hence the nearest neighbors have to be determined in order to calculate the fit function: in this matter the evolutionary search algorithm is limited by the time demanding nearest neighbor search which also appears to be the bottleneck of the ICP based methods.

Iterative Closest Point Algorithm

As already mentioned, the image registration is from the computational point of view the most complicated task of all processes in three dimensional mapping and model building. In order to align multiple datasets, the algorithm naturally requires processing of all data points in an organized way; since the sets in case of 3D images contain thousands of points, this process could be very computing-effort demanding.

The Iterative Closest Point (ICP) algorithm is an iterative aligning algorithm, as already mentioned it was first proposed in 1992 by P. Besl and N.D. McKay [2]. The application of this algorithm in 6DOF SLAM by Fraunhofer AIS institute is very comprehensive and it inspired the early development of the SLAM method presented in this article to great extent [9]. The algorithm serves to merge new scans into one – reference – coordinate system. Let's assume that we have a data set *D* as it is defined in the equation (1), which is partly overlapping with the existing model *M* .

As stated in previous section, the criteria function which is to be minimized is a function *E*. In the case if ICP, this function is expressed in equation (2), where $w_{i,j}$ is 1 if the point d_j in *D* describes the same point as m_i in *M* (the points are corresponding), otherwise it is set to 0.

$$E(R,t) = \sum_{i=1}^{m}\sum_{j=1}^{n} w_{i,j}\left\| m_i - (Rd_j + t)\right\|^2 \tag{2}$$

In each iterative step, the algorithm selects closest points as corresponding and calculates the transformation (*R,t*) according to the minimization of *E*(*R,t)*.

There are four known methods for the minimization of this function [9]. Easy to implement are quaternion based method and singular value decomposition method (SVD). Both are applied by Fraunhofer AIS, in later work SVD is preferred due to its robustness and an easy implementation [9]. The algorithm supposes decoupling of rotation *R* from translation *t* which is done by using centroid points given in equation (3).

$$c_m = \frac{1}{N}\sum_{i=1}^{N} m_i \qquad c_d = \frac{1}{N}\sum_{j=1}^{N} d_j$$
$$M' = \{m'_i = m_i - c_m\}, D' = \{d'_j = d_j - c_d\} \tag{3}$$
$$i \in N, j \in N$$

In order to minimize the rotation matrix *R,* a matrix *H* has to be created. This matrix is defined as in equations (4) and (5). Its elements $S_{xx},...,S_{zz}$ have to be calculated by multiplying the corresponding model and data points (these are the closest points). The H is calculated as in equation (6).

$$S_{xx} = \sum_{i=1}^{N} m'_{ix} d'_{ix}, S_{xy} = \sum_{i=1}^{N} m'_{ix} d'_{iy} \tag{4}$$

$$H = \sum_{i=1}^{N} m'_i \cdot d'_i = \begin{pmatrix} S_{xx} & S_{xy} & S_{xz} \\ S_{yx} & S_{yy} & S_{yz} \\ S_{zx} & S_{zy} & S_{zz} \end{pmatrix} \tag{5}$$

After obtaining the matrix H, it has to be decomposed e.g. using singular value decomposition method:

$$H = U\Lambda V^T \tag{6}$$

The final *R* is calculated in equation (7), where the matrices *U* and *V* are obtained by a singular value decomposition of matrix *H.* Final translation *t* is derived from equation (8).

$$R = VU^T \tag{7}$$

$$t = c_m - Rc_d \tag{8}$$

The most time and computation demanding task of the ICP method is the determination of the closest points necessary for calculating the elements of matrix H. Using brute force, this is a task of searching for the closest points results in a complexity O(n) single steps given the number of points n; the total processing time is in the order of tens of hours when using current computers. Though there are methods to speed up the search of corresponding points, commonly used in 3D graphics programming. Majority of them is based on the structuring of point clouds using planes, thus creating a tree structure. Nuchter et al. are using octrees, box decomposition trees and kd-trees for accelerating the search [10]. The kd-trees are a generalization of binary trees, where every nod represents a partition of a point set to the two successor nodes. Nuchter et al. have optimized kd-trees, so that the buckets are being chosen in such a way that the separating planes create two subgroups of equal amounts of points and the planes are also perpendicular. The algorithm of searching for the closest point is iterative, the point is compared to the separating plane and the decision about the direction of searching is determined from this. The implemented kd-tree acceleration reduces the search space and this can be expressed in the big O notation as O(nc) where 0 < c < 1.

After estimating the optimum rotation of the data set, the centered data points are rotated using this matrix and the iterative closest point algorithm is run again. The optimum state is reached when the rotation matrix output from the SVD or quaternion based evaluation is close to ones matrix, therefore the data set is not to be transformed any further. Therefore the iterative character of the ICP is in the iterative running of the process to find the rotation matrix which is used for rotating of the centered data. The method finishes when the output rotation is not transforming the data any further and this state is expected to be the optimum.

Performance considerations

In the registration of two range images using an ICP based method, data points are iteratively used to find a corresponding point in the model set. This is obviously the most challenging part of the registration algorithm from the computing effort point of view. The search can be accelerated by limiting the search space using tree structures or by projections of data points to model surfaces. Accelerations of other parts of the ICP algorithm are less likely to benefit the overall computing requirements of the method. The only other iteratively performed task is the estimation of the optimum rotation matrix R either using a quaternion based method or using the singular value decomposition method. This task though requires computing time of approximately one to two lower orders compared to the building of matrix H though finding the correspondences. The inevitable problem in computing requirements of the ICP based 6DOF SLAM is therefore the search for the correspondences. The difficulty of such task is proportional to the number of points for which the closest point search has to be performed and the number of iterations which have to be done to find the optimum. In 6DOF SLAM, the size of the 3D data sets is a major factor. Also, the time required for the optimization is very close to be directly proportional to the number of iterations of the ICP algorithm. One can therefore distinguish three major ways of improvement of the 6DOF SLAM ICP based methods: first, which is partly indirect, is to reduce the overall number of iterations of the ICP algorithm. This can be done if we ensure that the two point sets are somehow pre-aligned so that these are close to the optimum match. The iterative closest point algorithm has to perform less iterations and this could save the processing time. Second major way to improve the time required for matching the sets is the acceleration of the closest point search. This approach can be seen e.g. in the Fraunhofer team's research and the improvements are quite significant [9]. The last way of reducing the computing effort requirement would be a complex and robust data reduction. Data reduction could though introduce system error since the registration would be performed on the simplified data set, neglecting details in the scans.

Platform development

The most common method of measuring range data in 3D is the use of inclined 2D time of flight based laser range-finders. Time of flight (TOF) ranging systems determine the range by measuring the time required for a pulse of emitted energy to travel to and from an object which is expected to reflect the emitted energy. Due to the precision and availability of laser based TOF sensors, these are now dominating the use in mobile 3D perception. SICK LMS 200 sensor was selected in this research as it is capable of data acquisition at scanning frequency of 75 plane scans per second, transmitting 180 values for each scan with angular range of 180 degrees and resolution 0.5° (even scans start at 0° and odd at 0.5°, with 1° step between values). The maximum distance range of the scanner is 80 meters and the systematic error at the distance of eight meters is ±15 mm.

In order to measure distances in three dimensions, the sensor has to be moved or inclined while scanning the planar ranges. In this research, the objective was to develop a robot-independent platform, which could be installed on different robot platforms available at our department. The orientation of the scanner in such platform was also an issue: the analysis performed by Wulf et al. concluded that the orientation of the scanner is very influential on where the highest density of scanned points would be [14]. In each application, this highest point density area should be directed to the area of interest, which in case of this research would be in front of the robot. This would imply the rolling inclination of the scanner, though there were other criteria which were influencing the selection of the inclination method. A very important factor was the field of view. Another aspect was the possibility of 3DOF SLAM use during the movement of the robot which requires horizontal alignment of the scanning plane with the ground. Taking all these factors into account, the "pitching" inclining method was selected: the principle of the pitching inclined mechanism can be seen on the platform equipped robot in Fig. 1. The actuating device powering the inclining mechanism is a DC motor equipped with an incremental sensor. The main operation regime is precise angular velocity regulation, meaning that the scanner is continuously inclined. The resulting 3D image in case of continuous inclination is not in the regular matrix form since the measured points are positioned as if the scanned lines were tilted.

The 3D range scanning system was accustomed mainly for one experimental robot: Universal Telepresence and Autonomous Robot (UTAR). This robot was developed especially for experimental objectives. Most of its components were upgraded during this research. The robot can operate in both interior and exterior environment, allowing movement in light terrain (the maximum height of an obstacle is ten centimeters). Its maximum velocity is two km/h; the drivers are digitally controlled, allowing readings from incremental encoders connected to both motors. It is equipped with an industrial PC and the communication is performed by a WiFi MIMO module. Another robot which is being used for the 3D scanning is the HERMES omni directional platform based robot. The mechanism of this scanning platform is using a servo motor for inclining the sensor. This platform is also suitable for experiments since it allows movement of the robot in three degrees of freedom. UTAR robot is shown in Fig. 1.

Fig. 1: UTAR robot equipped with the inclining platform

Two stage leveled map accelerated 6DOF SLAM

A block diagram of the developed extended navigation method is shown in Fig. 2. In this diagram one can differentiate three main parts of the navigation system. The first part is naturally the sensor subsystem, using the above described scanner and inclination module. Then in the second part (not highlighted in the diagram), the 3D data are preprocessed and reduced. The SLAM core is the next step of the navigation method and it is highlighted in the Figure. This is the part where the major innovation takes place; its details will be explained in the following sections. The output from the SLAM core is the localization of the robot in six degrees of freedom (roll, pitch, yaw angles and x, y, z translations) and the created model of the environment could also be perceived as an output.

As previously stated, the major issue from the computing effort point of view is the number of closest point correspondences which are to be found when performing the 6DOF registration. In this research the main concern was to implement such means that the robustness of the 6DOF ICP would be preserved while the number of iterations in this time-demanding step would be reduced. In this method the model of the system composes of two sub-models: the leveled map model and the 3D environment model. This corresponds to the fact that the SLAM core composes of two phases: in the first phase, 2D leveled maps are extracted (factorized) from the 3D data and these are registered in a 3DOF ICP based matching algorithm. The registration is therefore initially done in fewer dimensions and thus fewer degrees of freedom, decoupling the registration in certain degrees of freedom from the overall robust but slow 6DOF SLAM. The 3D data are then transformed according to the results in 3DOF and passed to the robust 6DOF ICP based matching algorithm.

Factorizing the task

Factorization as a term relates to multiple areas of mathematics. Simple factorization is in fact a decomposition of an object into a product of other objects or factors. Similarly, in statistical analysis, factor analysis is a method which is widely used for determining or reducing the number of dimensions - variables - which are influencing the measured characteristic. The method presented in this article - decomposition of the 6DOF registration problem into two separate tasks, reducing the number of dimensions in the first step - follows both patterns in which the term factorization has been constituted. The factorization itself in this research is based on detecting such characteristics in the environment that the task of the localization can be decomposed into two separate phases. If such decomposition is expected to bring positive performance change, the dimension reduction must significantly reduce the effort necessary for the competition of the registration task.

In this solution, one aspect of the physical environment is being used to decouple the 3DOF registration from the 6DOF SLAM: the gravity. It is quite easy to measure the robot's pose in two degrees of freedom: the pitch and yaw inclination. These are easy to measure when the robot is still, which is the case when acquiring 3D images. Since we can measure these angles, the 3D data can be pre-aligned so that the two dimensional data which are to be extracted for the 3DOF SLAM can be independent on the pitch and roll inclination of the robot. This 2D extracted data set used in this research was called leveled map since it is aligned with the horizontal plain. In other words the leveled map created from the 3D data is invariant to the pitch and yaw inclination of the robot since we can measure these angles and align the 3D data in these two degrees of freedom before the extraction. Then we have to ensure that the 2D leveled map is also invariant to the remaining degree of freedom – the z ("upward") translation. This implies the type of objects extracted into the leveled map - vertical objects. It is assumed that the environment has such characteristics that in most times the robot will be able to see vertical objects in all heights (that is z translations), therefore the extracted leveled map will be independent on the robot's pose in terms of z translation. This actually holds true for most single-floor indoor environments and also for some outdoor environments.

The decoupled three degrees of freedom are therefore the remaining DOFs: horizontal translations x and y and the yaw angle ψ. In the first phase of the registration, these three degrees of freedom are registered. The leveled map extraction is described in more detail in the previous publication of EUROS robotics symposium and can be found in [6]. Briefly, the vertical structures are extracted by first splitting the data into vertical columns, which is followed by a search for equidistant obstacles within the columns. This process is powered by Combsort sorting algorithm.

When the leveled 2D map of vertical objects is constructed, it is passed to the 3DOF ICP algorithm. First, the leveled maps have to be centered using the computed centroids. In order to accelerate the search of the closest points in the ICP algorithm, a kd-tree is built for the leveled model set and then closest points of the leveled data points are queried. The kd-tree implementation uses splitting rules proposed and implemented by D. Mount and S. Arya in [8]. After the closest points are determined, the matrix H is calculated and finally the rotation R is obtained using the SVD algorithm. This is iteratively repeated until the rotation matrix is close to the ones matrix (the criteria set for this state is when each non-diagonal element is smaller than a preset limit ε). After this is fulfilled, the final rotation (which is the product of the particular rotation matrixes from all iterations) is determined and also the translation is computed. All steps are identical to the Iterative Closest Point algorithm described earlier. The result from this ICP registration is the rotation and translation in 3DOF (yaw angle – rotation around z axis, x and y translation). The reduction of the dimension is to significantly reduce the number of points for which the closest points are to be found when compared to registering of 3D data sets. That should naturally reduce the time required for the calculation of the H matrix.

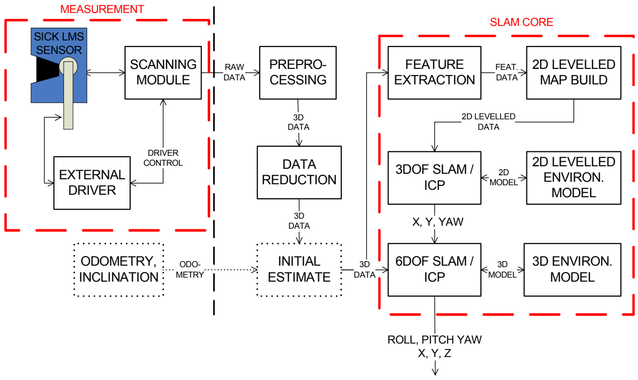

Fig. 2: Block diagram of the navigation system

Registering the pre-aligned 3D data

After the matching of the leveled map, the obtained estimate of the yaw rotation is applied to the 3D data set. Since the ICP works with centered data sets, the application of the translation is not necessary at this point. The input 3D data should also be aligned in the two rotation angles: the roll angle and the pitch angle. These were also the source of information for the leveled map building and are necessary for this method to accelerate the classical ICP. The two angles correspond to the inclination of the robot and could be easily measured in any environment where gravity is present. The pre-alignment is applied on the data set by applying the rotation on each of the data points. This transformation is expected to significantly reduce the number of iterations in which the following registration method is to determine the optimum rotation and translation to correctly match the data set to the model. The ICP algorithm is applied on the 3D data. Initially, data are centered and a kd-tree is built for all model points, using the sliding midpoint splitting rule proposed as in the leveled map matching, proposed and implemented in [8] by D. Mount and S. Arya. The optimal rotation R is iteratively calculated, finally resulting in the full 6DOF match of the data. Then the data can be used to expand the overall model of the environment.

Unique matching of model points in the closest point search

The originally implemented one-phase ICP algorithm had to be modified according to quality issue which was encountered when evaluating the "simple" ICP based registration. The problem which had to be dealt with is in the fact that the theoretical iterative closest point algorithm supposes that there are corresponding points in the data pairs and also the number of points corresponding to the same physical objects in the environment is identical for both point sets which are being registered. Point density correction algorithm was not implemented in this research due to loss of information reasons. Therefore the point density is changing with respect to the distance of objects from the scanner; due to the spherical character of the beam spread this density reduction is very significant. One cannot assume that the scans will be acquired from reasonably close locations as this would undermine the whole task of localization in unknown environments: the robot's movement in this environment would be limited by the poses from which the scans are to be acquired. This situation is also problematic in areas where in two point sets we are looking for closest points of points which are close to the border of the overlap of the sets. In this case, it could happen that for the data points which do not have real correspondences in the model (meaning that the objects represented by the data are not present in the model set), correspondences of closest points will still be found but will introduce significant errors as the real correspondences do not exist. Such situation could be avoided if the overlaps would present 100% of the sets, meaning that the overlap would be known and the transformation of the sets could hence be determined, therefore the registration would be no longer necessary. This is though not a realistic solution.

A possible solution to this problem is the establishing of a mechanism which enforces the uniqueness of the use of model points in the closest point queries. Such mechanism was implemented in the navigation method and it has become part of the closest point query. It works according to the following principle: when a closest point is found, it is first checked whether it has already been allocated for another - previous query. If this is true, the distance between the queried points and the closest point are compared and the queried point which is closer to the found closest point is registered in the records of found closest points. The matrix H is therefore updated at the end of the overall process of finding closest points for the data. Otherwise the matrix would already reflect the previously added points which would sometimes (in case a closer data point was queried) no longer be valid and it would have to be corrected to drawback the previous update. This modification is not without an impact on the time-wise performance of the ICP algorithm.

Experiments

Multiple experiments were carried out to verify the performance of the method from two main perspectives: one is the computing effort which is required to match model and data range images, second perspective is the quality performance of the method. The experiments will be structured in the following way: first, an analysis under close-to-artificial conditions will be discussed, evaluating the performance of the method under predefined close-to-optimum conditions. This analysis is though based on data acquired in real environment. The following three evaluations will represent three different environment model conditions. These correspond to the changing difficulty of the image registration task. The three evaluation environments were selected with regards to the expected weaknesses of the developed method: the main changing attribute of the environments is the structuring factor - the presence of vertically consistent structures in the environment. Another aspect is the presence of randomly scattered objects and the last and most challenging aspect is the horizontal structuring of the environment meaning the presence of multiple floor zones.

Real data with arbitrary transformation

The experiments usually start by testing of the implemented methods on artificially created objects. The reasoning behind this strategy is that for such artificial objects, one knows the correct solution (the ground truth) and therefore method's performance compared to the ground truth solution can be assessed. One way to do this would be creating an artificial point cloud with known characteristics, applying a known homogenous transformation on this point cloud and registering the original and the transformed one. The disadvantage of such step would be that the object would most likely not contain characteristics of real environments in which the method would be expected to function when in real operation. For this reason, a real 3D image was used instead of an artificial one. A known homogenous transformation was applied to this point cloud and the created and original point clouds were then registered. Since such point clouds contain equal number of points from the same measurement and there are ideal corresponding pairs for all points of both sets, the mentioned ideal ICP application conditions are fulfilled and the method should perform as ideally. The performance is expected to be the same as for the artificial object – the point clouds should be registered using the same but opposite homogenous transformation to the one applied on the point cloud. For the testing real point cloud, an office environment was selected since it contains both vertical structures (walls, furniture) and also scattered object of various types. The data set for this experiment contained approximately 33 thousands of points after removing inclining construction from the 3D data. This data set contains an office room with open door and a small part of a hallway.

The first experiment was to confirm the correct functioning of the ICP algorithm. This test was passed successfully as the calculated translation or rotation was an inverse of the initially applied transformation. The next step was the testing of the computing time and number of ICP iterations dependency on the scale of transformation between the model and data sets. Since the translation is decoupled from the iterative algorithm, the manipulated transformation was rotation around z axis. The dependency of the registration duration on the scale of the rotation between model and data point sets for different algorithm configurations is shown in Fig. 3 (the evaluations were performed on a Pentium M 1.8 GHz machine). One can observe that the number of iterations of the ICP registration and the overall computing time when factorization using leveled maps is not introduced in the algorithm can be approximated by a logarithmic function. When leveled maps are introduced, the number of iterations as well as the computing time is constant for all scales of rotation. The scale of the acceleration when the factorization is introduced gradually grows as the registration duration using 6DOF ICP only rises, resulting in a significant performance improvement for transformations over approx. 4°. When uniqueness of model points use in the closest point search is introduced, the duration of the registration almost doubles, although since in this experiment the conditions are ideal, quality improvement is simply not possible (all matches are perfect).

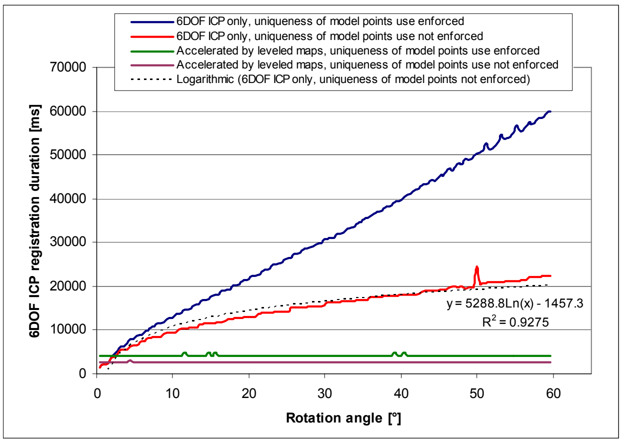

Fig. 3: Duration of the 6DOF ICP registration as a function of the applied transformation's rotation angle

Experiments in real operation

After verifying the method on real data using arbitrary transformation to create two registered sets, the method was verified using again real data (measurements of real environments) though this time the transformation took place physically in the environment – moving and rotating the scanning platform. The experiments were carried out in four different physical environments: a hallway with simple vertical structuring and absence of reflective materials, an office environment with transparent and reflective materials and scattered objects, a laboratory environment with many scattered objects creating occlusions and finally in a multi-level environment with reflective materials and scattered objects enclosed by iron bar.

Hallway environment

The hallway environment structuring contains regular vertical planar patterns. The size of the vertical planes present in the environment is much higher than the resolution of the scanner. Ten scans were taken in this environment, using different transformations between acquiring model and data sets (both translations and rotations). For most pairs, "simple" 6DOF ICP based SLAM converged to and optimum which was relatively close to the ground truth (measured at the time of data acquisition), although the different angle of view and varying point densities between the data and model sets have introduced an error in the registration. Such error was eliminated when the uniqueness of the model points use in the closest point search was enforced. The impact on the computing time of the registration was significant; the difference in the quality of the resulting match can be seen in Fig. 4. The difference was especially visible in the rotation determination, where the uniqueness of model point's use introduced a significant improvement. For all measured point set pairs, the computing effort requirements were compared between the original 6DOF ICP based SLAM and the implemented acceleration based on the factorization of the algorithm using leveled maps. An example of leveled maps of the hallway environment (registered and unregistered) is shown in Figures 5 and 6. The improvement of the computing time requirement when leveled maps were introduced varied for five of the range image pairs between 14% and 46% (in average 28%). There was one case where the introduction of leveled maps decelerated by -8%. This outlier was in fact a pair where the model and data sets were transformed only applying translation to the scanning platform. This actually appeared to be the weakness of the ICP based method since the centroids calculated to determine the translation were not representing identical part of environment given the varying density, point of views etc. This error was partly overcome in leveled maps by introducing a two pass extension where the data were filtered based on the first registration, partly eliminating the objects present only in one of the sets and correcting the centroids. Such solution is though increasing the computing time especially if implemented in the 6DOF registration.

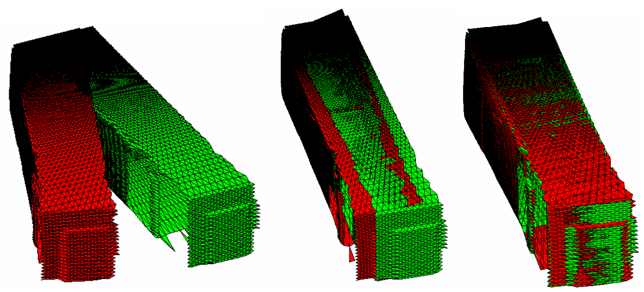

Fig. 4: Hallway experiment: unregistered point pairs (left); registered solution without model point use uniqueness (middle) and registered solution with uniqueness introduced (right)

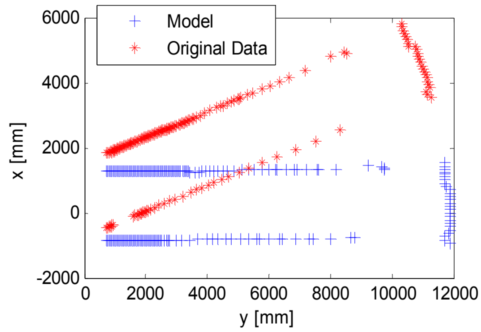

Fig. 5: Unregistered leveled maps of the hallway

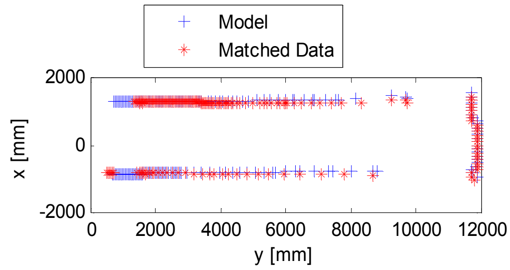

Fig. 6: Registered leveled maps of the hallway

Laboratory and office environments

The following two experiments have shared some characteristics of the environment: both the laboratory and office environments contained plenty of structured objects of irregular shapes. In the laboratory environment, these objects actually blocked a significant part of the scanner's viewing angle. To verify the behavior of the method, two pairs of data were acquired and registered for both the office room and the laboratory. The difference between the two pairs was in the scale of the transformation applied on the scanning platform. For the first pair, a rotation of approx 20° (resp. 15° for the laboratory) was applied and for the second pair a rotation of 45° (resp. 35° for the lab) was applied. The translation between sets within a pair was within the estimated accuracy of odometry – maximum 0,5 m. The results were ambiguous: for the first pair in both environments, the registration was partly successful. The quality of the match for the office was perfect rotation-wise but error was introduced by erroneous determination of the translation. The results of the registration are presented in Fig. 7. The duration of the registration was 57 s using factorization and 73 s using pure 6DOF ICP SLAM (both with uniqueness of model points' use in the closest point search). The result for the laboratory environment was partly acceptable: the rotation around z axis was determined correctly but the solution was also rotated around y axis – this introduced errors. The registration of second pairs with larger transformation resulted in only partial registration for the office room and again partly erroneous registration in the lab. In the office, the transformation of 45° in the second pair was beyond the capabilities of the implemented method. Leveled maps have not significantly accelerated the registration of the two scan pairs acquired in the laboratory, this is most likely due to the occlusions and structuring of the environment.

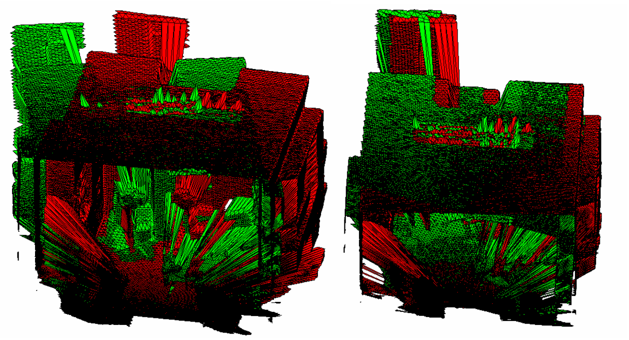

Fig. 7: Registration of data sets acquired in office environment: scale of rotation approximately 20° (left: unregistered; right: registered)

Multi-floor environment with difficult structuring

In the last experiment presented in this article, 3D range data were acquired in an environment which was expected to be extremely difficult for the registration using both the original and the extended method: a multi-storey area with stairway and various objects of irregular shapes in the view of the scanning platform. Also, the stairway was lined with a glass border, possibly introducing multiple reflections depending on the angle of incidence of the laser beam. Another feature of the environment was a presence of bars which were enclosing the scattered objects in the bottom part of the scans, resulting in regular structuring of the 3D data sets in this part of the image. In the first experimental data acquisition, the transformation applied on the platform between acquiring the model and data sets was a rotation around z-axis direction of approximately 15°. The registration when using leveled maps took approx. 203 ms, performing ten ICP iterations in the 3DOF SLAM. The rotation determined by this 3DOF SLAM was of approximately 18°, the data set was therefore slightly "over-rotated". This is shown in Fig. 8. The subsequent registration in 6DOF after pre-aligning the sets took approximately 46 s, completing 33 ICP iterations and determining rotation of 21°, again over-rotating the solution. When using the simple 6DOF ICP based SLAM without factorization, the registration took approx. 100 s, completing 72 ICP iterations and reaching same solution. The acceleration in this case was quite significant; the quality of the match was though limited. The second pair of point clouds was acquired by applying rotation around z-axis of approx. 45°, for this pair the registration encountered a local minimum before reaching the correct solution. The leveled maps have diverged from the correct solution. Such transformation is therefore beyond the capabilities of the implemented method especially due to the character of the environment.

Fig. 8: Multi-storey environment data and model sets: transformation approx. 15° rotation around z-axis direction; (left: unregistered; right: registered)

It has to be noted that for all experiments where the use of the feature enforcing uniqueness of model points use in the closest point search is not explicitly mentioned, the feature was enabled during the registration. This is because it significantly improves the quality of the registrations.

Conclusion

The implemented navigation method is in fact a hybrid two stage solution which uses factorization to accelerate the method for simultaneous localization and mapping in 6DOF. This is done by extracting 3DOF invariant vertical structures (landmarks) from the 3D data, creating a 2D map of vertical objects and registering these in 3DOF prior to the robust and computing time demanding registration in 6DOF. The aim of such modification was to pre-align the 3D data in 3DOF and to reduce the time required for the 6DOF SLAM algorithm in this way while preserving the robustness of the original 6DOF ICP based method. The practical implementation of such a method requires use of accurate and inexpensive technologies of gravity vector measurement, initially aligning the data with a horizontal plane by hardware means. Selecting vertical structures for extraction ensures maximum height invariance of the created 2D data for the 3DOF registration. Features focusing on the quality of the final registration were also implemented within the method, such as the enforcement of uniqueness of model points use in the closest point search of the ICP algorithm. This partly eliminated the errors introduced by occlusions and different fields of view of acquired images. The closest point search was also accelerated using kd-trees.

Overall, one can conclude that the implemented method gives very promising results in terms of the expected ICP behavior. This means that when factorization described in this article is introduced in the SLAM method, the quality and robustness of the 6DOF ICP based SLAM is preserved and the overall registration process is often significantly accelerated. In ideal conditions and in vertically structured environments, the implemented method is of a great benefit in terms of the computation power requirements. Also, the leveled maps can be used as a separate mapping output since these maps depict the vertical structuring of the environment and may be viewed as reduced 2D maps. The difficulties of the implemented 6DOF SLAM are tightly connected to the general issues of the ICP use in the SLAM algorithms. On one hand the navigation method preserves the robustness of a general 6DOF ICP based SLAM but on the other hand it also shares the disadvantages of the ICP. The issues seem to be appearing when registering images of randomly and finely structured environments and also in registrations where the scale of transformation between the data and the model sets is too large. In both mentioned situations occlusions are present in the point clouds and there are also objects appearing only in one of the sets which are being registered, sometimes introducing error in the rotation determination but much more often in the translation calculation. The calculation of translation between the point clouds seems to be the most important issue of the ICP based method as resolving it within the ICP framework appears to be difficult.

In the continuation of this research, this issue could be resolved by introducing other than ICP based registration method into the navigation system solely to help to determine the correct translation. Other option would be the development of a data reduction method which would initially determine common objects in data and model set and also correct the areas with uneven point density.

The implemented navigation method is though not expected to be operating as a standalone module on which the robot would be completely reliant. The research in the Laboratory of Telepresence and Robotics builds on the development of a modular system applicable on different robots allowing multiple configurations and levels of autonomy and teleoperation support. This is mainly due to the fact that teleoperation is still one of the most practical ways in which the robot can fulfill complicated tasks yet unsuitable for autonomous regimes. The current state of this research now allows the implementation of the developed and experimentally verified navigation method in the existing robot prototypes, supporting teleoperation and mapping in unknown environments. The module is expected to provide 3D images for the HUD displays of the operator's console. The localization information will be combined with other pose calculations e.g. coming from the inertial navigation subsystem and GPS module correcting the errors such as translation offsets and will inform the operator about the robot's position. The reliance on multiple systems is in fact the major advantage of the interconnected modular approach to navigation.

Acknowledgements

I would like to thank to Assoc. Prof. Luděk Žalud Ph.D. and Prof. František Šolc CSc. for their support during this research. This project was also supported by Intelligent Systems in Automation Grant MSM0021630529 and also by the Ministry of Education of the Czech Republic under Project 1M0567.

References

[1] ARUN, K. S.; HUANG, T.; BLOSTEIN, S. D. Least-squares fitting of two 3-d point sets. In *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.9, no.5, ISSN 0162-8828. pp 698-700

[2] BESL, P. J.; MCKAY, N. D. A method for registration of 3-d shapes. In *Proceedings of IEEE Transactions on Pattern Analalysis and Machine Intelligence* 14(2), Feb 1992. Pp 239-256.

[3] BORENSTEIN, J. EVERETT, H. R. FENG, L. *"Where am I?" Sensors and Methods for Mobile Robot Positioning.* University of Michigan, Oak Ridge National Lab D&D Program, United States Department of Energy, USA, 1996

[4] CORMEN, T.; LEISERSON, C.; RIVEST, R.; STEIN, C. *Introduction to Algorithms.* Mcgraw-Hill College 1994, ISBN 0070131430

[5] EVERETT, H. R. *Sensors for Mobile Robots.* A.K. Peters Ltd., Natick, USA 1995.

[6] JEŽ, O. 3D Mapping and Localization Using Leveled Map Accelerated ICP. In European Robotics Symposium 2008. *Springer Tracts in Advanced Robotics*. Springer, 2008. Pp. 343-353. ISBN: 978-3-540-78315-2.

[7] MARTIN, C. M.; MORAVEC, H. P. *Robot Evidence Grids.* The Robotics Institute Carnegie Melon University, Pittsburg, 1996

[8] MOUNT, D.; ARYA, S.: *ANN: Approximate Nearest Neighbour Searching. Programming Manual.* online http://www.cs.umd.edu/~mount/ANN/ (1/10/2007)

[9] NUCHTER, A.; LINGEMANN, K.; HERTZBERG, J.; SURMANN, H. 6D SLAM with Approximate Data Association, in *Proceedings of the 12th International Conference on Advanced Robotics (ICAR '05)*, ISBN 0-7803-9178-0, Seattle, USA, July 2005a. Pp 242 – 249.

[10] NUCHTER, A. LINGEMAN, K. SURMANN, H.; HERTZBERG, J. 6D SLAM – Preliminary Report on Closing the Loop in Six Dimensions. In *Proceedings of the IFAC/EURON Symposium, Portugal, 5 - 7 July 2004*. Elsevier, Amsterdam 2005. Pp. 203- 208

[11] PULLI, K. Multiview Registration for Large Data Sets. In *Proceedings of the 2nd International Conference on 3D Digital Imaging and Modeling (3DIM '99)*, Ottawa, Canada, October 1999. Pp 160 – 168.

[12] ROBERTSON, C.; FISHER, B. *Parallel Evolutionary Registration of Range Data*. Computer Vision and Image Understanding, Vol. 87, Elsevier Science Inc., New York, USA, 2002. Pp. 39 – 50

[13] WULF, O.; BRENNEKE, C.; WAGNER, B.. : Colored 2D Maps for Robot Navigation with 3D Sensor Data. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, September 28-October 2, 2004, Sendai, Japan.

[14] WULF, O.; WAGNER, B.. Fast 3D-Scanning Methods for Laser Measurement Systems. In *Proceedings of the International Conference on Control Systems and Computer Science (CSCS14)*, Bucharest, Romania, July 2-5, 2003.

Ing. Ondřej Jež

Dept. of Control and Instrumentation
Faculty of Electrical Engineering and Communication
Brno University of Technology
Kolejní 2906/4
612 00 Brno
Czech Republic
Tel.: +420 603 208 380
E-mail: ondrej.jez@phd.feec.vutbr.cz