# Cooperative Mutual 3D Laser Mapping and Localization 

Julian Ryde and Huosheng Hu<br>Department of Computer Science, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England $\{$ jryde, hhu\}@essex.ac.uk


#### Abstract

A 3D laser scanner is built by adding a rotating mirror to a conventional 2D scanner. The scanners are deployed on four robots to build full 3D representations of an indoor environment. An original representation mechanism referred to as occupancy lists, rather than standard 2D free space grids, is used to maintain the 3D map. Localization is done by extracting horizontal sub-ceiling cross-sections. Taking cross-sections from near the ceiling in this way results in more reliable and time invariant maps. Experimental results show inter-robot sightings and the sharing of map data aid mapping by improving the reliability of localization. Mapping with four robots reduced the average position error from 0.35 m for single robot operation to 0.1 m when cooperating.

Index Terms-Mutual, Cooperative, Localization, Mapping, 3D laser scanner


## I. Introduction

Effective cooperation and 3D sensing will help release robots from the research laboratory into the world at large. Usually localization and mapping has taken place in 2D, mainly due to limitations of the sensors in the case of laser range finders, or processor speed and algorithms for that of stereoscopic vision [1]. To overcome the 2D laser scanner dimensional limitation researchers typically nod or rotate the entire laser scanner sweeping out an area over which the 3D range data is gathered [3].

Recently in a drive to test the benefits of 3D sensors researchers have mounted 2D laser scanners on nodding or rotating mechanisms to obtain 3D scans [9], [10]. Alternatively, two laser scanners mounted with their scan planes orthogonal [4] are also popular. The main problems with nodding or rotating approaches are difficulties in hardware implementation and high power consumption as the 2D scanners are heavy. Consequently a rotating mirror prototype has been built which produces 3D scans with a field of view of 100 by $180^{\circ}$, is light, has low power consumption and is easily deployed on conventional robotics platforms.

A number of groups are undertaking research into 3D laser mapping however no group is performing cooperative 3D laser mapping, the closest is [11]. The benefits of full 3D mapping are abundant and so the rapid expansion of this field is inevitable. The detection of negative and over-hanging obstacles greatly enhances avoidance behavior. Once 3D maps of environments have been built they can be customized for different robots. For instance various 2D maps may be built for robots of different sizes or with 2D sensors at different heights. Severely cluttered non-manifold environments such as search and rescue situations may be reliably mapped. Maps based upon the higher volumes of rooms nearer the
ceiling remain accurate for longer and an unoccluded view of the ceiling is usually readily accessible to a robot even in crowded environments, [2]. The disadvantages of 3D sensing technologies are slower acquisition time and the geometric increase in data that needs to be processed.

In this paper, the mutual localization approach discussed in Section III-A coupled with the 3D laser range finder prototype pushes this research into the new area of 3D cooperative localization and mapping. Combining mutual localization with the data from multiple 3D laser scanners enables full 3D mapping of indoor environments. This would prove vital for a number of industries such as nuclear decommissioning, search and rescue scenarios, surveying as built structures and maps for mobile robots.

The rest of this paper is organized as follows. Section II presents the system framework of the research carried out in this paper. Section III details the system components, mainly focused on mutual localization and scan alignment. Section IV describes the design of our 3D laser scanner. Section V includes the experimental results and is followed by Section VI discussing their implications along with future research directions.

## II. System Framework

An overview of the multi-robot system is delineated in Fig. 1. In this multi-robot system each robot can exchange two types of information, map updates and sightings. The map updates are filtered 3D scans and the inter-robot sightings are a list of other robots observed by that robot. For our research these inter-robot sightings are made possible by mounting cylinders covered with retroreflective tape above the laser scanners. Retroreflective materials are relatively rare in most environments and combining this with the expected height above the floor makes this a reliable method for robot detection. Mutual detection events (where a pair of robots observe one another) increase this reliability still further as both robots must measure the same separation from each other. These sightings are used to both reduce the search space and enhance the reliability of the map updates. In Fig. 1 there are two roles for the robots as either observers or mapping robots. The observers are the robots whose coordinate system the scans are merged into and for improved reliability and accuracy they can be stationary whilst the other robots map around them. This, however this is not a necessary constraint as maps have been produced with all robots moving. Depending on which robot made the sighting various pose constraints are generated in the the coordinate system of the observer.


Fig. 1. Framework and robot communication of the global map $M_{a b c}, 3 \mathrm{D}$ sensor scans $S_{k}$, inter-robot sightings and movement commands, $U_{k}$.

A standard observation, although not uniquely identifying a single pose, vastly reduces the number of poses that need to be considered to a loci of points in pose space. Consider two robots, the stationary observer robot, S and the mapping roaming robot, R . There are three observational situations to consider in descending order of typical occurrence frequency in mapping missions. The first is if $S$ observes R at a position of $\left(R_{x}, R_{y}\right)$ then the loci of poses is given by $\left(R_{x}, R_{y}, t\right)$, where $t$ is a parametric variable used to describe the loci that varies from 0 to $2 \pi$. This in effect says that the position of R is fixed and its orientation is unknown. Secondly if R observers S at a $(s, \beta)$ then the loci of poses of R in the coordinate frame of $S$ gives rise to a loci of poses the positions of which are along the circumference of a circle and with each pose there is associated a single orientation and this is expressed in parametric form as

$$
\left(\begin{array}{c}
S_{x}+s \cos t  \tag{1}\\
S_{y}+s \sin t \\
\pi-\beta-t
\end{array}\right)
$$

again with the parameter $t$. Finally the rarest and most informative observation is when both S and R observe each other and in this situation the relative pose is determined with both high accuracy and certainty. In the situations where observations give rise to a range of poses then localization may be performed by map matching along the loci of poses by simply testing values of the parameter $t$. This not only dramatically reduces the search space for possible poses but also decreases the likelihood of high correspondence false map matches. This approach gives localization in four degrees of freedom $x, y, z$ and $\theta$ the horizontal orientation. The two degrees of freedom missing for a fully unconstrained 3D pose may be acquired from either the assumption that the robot is on a flat floor or that the robot is capable of determining the direction of the acceleration due to gravity.

## III. System Components

## A. Mutual Localization

The premise for mutual localization is that rather then merely observing robots as beacons each robot observes and is itself observed simultaneously, [5]. The simultaneous nature delivers enhanced performance for relative pose determination. The usual benefits to cooperative localization are applicable such as bounded errors and higher certainty. Additionally, ensuring that robots may observe team-mates and be observed themselves means that simultaneous mutual localization events can occur. These events allow superior relative pose determination. Firstly, the mutual localization is robust to spurious readings because simple checks on the validity of the mutual pose are available; for instance the separation of the robots should be similar as measured by both observing robots. Secondly, the accuracy in the pose does not deteriorate with separation, a very useful property. Increasing separation merely decreases the frequency of mutual localization events.

Robot detection is accomplished by mounting retroreflective cylinders above the origin of the laser scanner as shown in Fig. 2. The laser range finder can only distinguish the retroreflective tape for ranges exceeding 0.4 m . At ranges less than this the laser spot is comparably bright even on diffusely reflecting surfaces. Thus it is necessary to mount the beacons further than 0.4 m from the origin of the laser scanner to guarantee detection even when robots are very close. Another advantage to mounting the beacons high is increased visibility during experiments in cluttered rooms. It was found that most of the moveable objects and clutter in a room is quite low for example chairs and tables so even when this clutter interrupts the line of sight between the robots; the beacons, being some way from the ground, may still be detected by the laser scanners.

Mutual localization is accomplished by first projecting the observed beacon's 3D position onto the $x y$-plane which is parallel to the floor of the room. Once projected onto the horizontal plane the 2 D mutual localization algorithm may be used. This method assumes that the beacons are always the same height above the floor, reasonable in the case for many indoor environments with flat horizontal floors. Given the relative pose of the mapping robot with respect to the observing robot multiple 3D laser scans may be combined to produce a local map.

Conventional 2D occupancy grids however do not extend well into 3D in terms of their computational requirements. For instance a typical room of 10 by 10 by 4 m at a resolution of 0.05 m requiring only a 40 thousand cell representation in 2D needs 3.2 million cells in 3D. This orders of magnitude jump has ramifications for both memory requirements and performance. A number of observations of the nature of the 3D data gained from 3D mapping experiments enable some important reductions in the size of the representation without significant loss of information. One of these observations is that difference between the number of occupied and
unoccupied cells in 3D for navigable environments is large with many more cells free than occupied. This relative rarity of occupied cells makes their information content higher, or in others words observations of occupied cells are more significant than those of free cells and so consequently should be stored and manipulated in preference. The large majority of the unoccupied cells means they have low significance especially when it comes to matching sensors scans with prior maps. Thus their low significance and abundance makes them ideal candidates for removal without the loss of too much information. Unoccupied cells have been historically maintained in occupancy grids because the grids were 2D and the sensors, often sonar, were very unreliable so every bit of information had to be gleaned from the sensor data and maintained. On the other hand, 3D laser scanners are much more reliable and are almost guilty of producing too much data, thus we can afford to discard some of it for the sake of a compact representation.

Instead of holding the cell probabilities in a 3D array all the cells with a high probability of occupancy are maintained as a list. Not only is there considerable saving in storage and manipulation but there does not have to be strict bounds on the environment as is the case with a 3D array. Experimentally it has been found that the typical room considered above can be represented in this manner by 20 thousand cells at resolution of 0.1 m . The representation does depend on the clutter of the room and grows with the number of observable surfaces but most importantly on the cell size. Cells that are not listed are assumed to be empty, in an occupancy list representation the information lost is the ability to discern between freespace and unobserved space. This distinction dwindles as the map grows and the coverage of the environment rises. Further compression of the 3D data sets is possible by extracting features that are common in indoor environments such as edges and planes, however this is at the expense of a loss of generality and is not done in this work.

In this instance the representation is implemented by using a hash table with the index key being the cell's integer position and the value, the probability of occupancy. This allows the probability at a given position to be looked up quickly which is typically the most common operation on the map. In this way the environment is represented as a dense feature set rather than an occupancy grid. Each feature is a voxel that usually produces a laser return and so there are very many of these features typically tens of thousands in a room at a resolution of 0.1 m . This representation scales much better with environment size. Data points are recorded for voxels containing surfaces in the environment and as these surfaces are 2 D then the number of voxels required to express them scales as the square of the environmental size.

Due to the fidelity and narrow beam width of the laser scanner a straightforward raytrace model was employed. When updating the occupancy list all occupied cells along the path of the laser ray have their probabilities decreased whilst the cell containing the end point of the ray has its probability increased.


Fig. 2. Beacon positioning and calculating 3D coordinates from range data.


Fig. 3. 3D rotating mechanism deployed on a Pioneer 2 and 4 robots mapping.

## B. Scan alignment

Much effort has been expended in the area of scan matching and global localization [7], [8]. Mostly to improve performance in generating the position probability density function through a variety of sample techniques. Usually this localization or scan matching has to be performed on a fairly regular basis (every meter or so) because the accuracy of odometric data deteriorates so rapidly. In our case because inter-robot sightings are often available relative pose determination is much faster. In this way reliable local maps can be generated quickly and they have a better chance of unambiguously localizing in the global map. This makes exhaustive search global localization a potentially viable approach. Scan matching depends on a good overlap between subsequent scans which comes from good sensor coverage and wide visibility of features both of which are delivered by the higher regions of a 3D scan.

Floor-level 2D maps are unreliable and quickly become out of date when furniture is moved however, in this work map data is extracted as a horizontal cross-section just below the ceiling of the room. This volume of a room changes infrequently and is more easily observed especially in cluttered environments. The ceiling is 3.5 m high and the horizontal cross-section chosen was the volume from 2.8 to 3.3 m in height and the resulting 2 D occupancy grid is produced by iterating through the occupied voxels of the 3D map and inserting a corresponding occupied cell in the 2D crosssection if one is not already present.

A plethora of global localization algorithms given a prior


Fig. 4. Alignment of two submaps with angular resolution of $0.5^{\circ}$ and cell size of 0.1 m . The scans are 3D however the cross-sections are shown containing voxels from $0-2.8$. The surface plot shows the voxel overlap count for the best orientation as a function of position.
map exist however in this research an exhaustive search approach was implemented. This has a number of advantages, the entire pose probability density function may be recorded and inspected, it is very robust and suffers only from its slow execution speed. The impact of this on real-time operation may be vastly reduced especially if a position estimate is available which is almost always the case. It was found that execution speed of an unoptimized algorithm meant that global localization could take place on maps of size 20 m by 20 m with pose resolution of 0.1 m and $1^{\circ}$ in the order of minutes on a 2 GHz processor. For the implementation of the exhaustive search algorithm a subset of the data is matched by extracting the cross-section near the ceiling. A match metric distribution, Fig. 4, is built up by stepping through various positions and rotations with respect to the map. The match metric is a count of the coincident cells for that particular rotation and translation.
Typically, pose resolutions of 0.1 m and $0.5^{\circ}$ are used. The pose with the best correlation metric is then chosen with an example in Fig. 4 illustrating the alignment of two consecutive scans. Also shown in Fig. 4 is the resulting match and the match distribution function where the best match ratio out of all orientations is plotted at each position.

## IV. 3D Scanner Design

The 3D laser range scanner consists of a SICK LMS 200, rotating reflecting surface prototype and software to convert the range scans into 3D data. After gearing down
the mirror rotation period is 40 seconds because the laser beam is reflected from the mirror the angular velocity of the laser scans is double that of the mirror and so a full scan is completed every 20 seconds. The slow rotation speed contributes to a very low power consumption of 0.01 W . Feedback of the angular velocity of the mirror to the robot is accomplished by using the rotation sensor in a cannibalized PS2 ball mouse. A particular advantage of this approach is ease of deployment, in that the rotating mirror mechanism can simply be placed in front of the laser scanner, as in Fig. 3 , and the PS2 mouse plugged in with no further connections required, even the power supply may be self-contained.

The scan frequency is limited by the data rate and at the moment is 0.05 Hz . This data rate is $34.8 \mathrm{~kb} / \mathrm{s}$, the maximum for the serial interface, resulting in 13 horizontal scans per second. The SICK LMS 200 can support a $500 \mathrm{~kb} / \mathrm{s}$ data rate using a USB interface. At this faster rate, full $3 \mathrm{D} 1^{\circ}$ by $1^{\circ}$ scans should be possible at 0.5 Hz .

In Fig 2 the effect of the mirror is to bend part of the $x y$-coordinate plane to a new elevation illustrated in gray. The following equations, which reference the values indicated in Fig. 2, indicate the conversion between the $r, \theta$ and $\phi$ coordinates, measured by the laser scanner, and 3D Cartesian coordinates.

$$
\begin{align*}
& x=(r \cos \theta-d) \cos \phi+d  \tag{2}\\
& y=r \sin \theta  \tag{3}\\
& z=(r \cos \theta-d) \sin \theta \tag{4}
\end{align*}
$$



Fig. 5. Photograph of mapped room.


Fig. 6. Typical 3D scan rendered with further points a darker gray.

Where the value $d$ is the separation between the origin of the laser scanner and the axis of the rotating mirror. The range and bearing as measured by the laser scanner are $r$ and $\theta$. The angle of the plane (gray region in Fig. 2) to the horizontal introduced by reflecting the laser ray from the rotating mirror in front of the scanner is indicated by $\phi$.

## V. Experimental Results and Analysis

To visualize the map data a 3 D scene is generated consisting of cubes or spheres representing occupancy and the results rendered using a ray tracer (Fig 6). Depth information can also be added by coloring the voxels by their distance from the observer, as in Fig. 8. However, despite these efforts it still remains difficult to effectively display 3D maps on two dimensional static media. An example of such a 3D image is displayed in Fig. 6 and a photograph from the corresponding view point is shown in Fig. 5. Each sphere in Fig. 6 represents a laser range return and is colored by its distance from the view point, with those closer being lighter. One of the main problems encountered in obtaining 3D scans was pixel mixing. This is alleviated to a certain extent by multiple observations from different vantage points and probabilistic updates to the map. Pixel mixing may also be helped by the


Fig. 8. Perspective view depthmap.
median filter approach used in [9]. Other attempts to improve the quality of laser range data for 3D imaging [6] show that substantial improvements are possible. One of the main advantages of these maps is that they enable the avoidance of obstacles not visible to 2D scanners, for example negative obstacles (downward stairs) and tables. Because volumes near the ceilings of indoor environments tend to be more time invariant then these ceiling based maps retain their accuracy for longer. Previous success with this idea employing an upward pointing camera has been obtained, [2], however our work uses the sub-ceiling volume as represented by an occupancy list. This is amply illustrated in Fig. 7 which contains overhead depthmap views of the 3D map. The grid lines are 1 m apart and the occupied voxels closer to the floor are a lighter gray. The second image view in Fig. 7 contains all map data between the floor and the ceiling ( 3.5 m high). The floor and ceiling data are filtered out to aid visualization. The third image in Fig. 7 shows only occupied voxels above 2.8 m . The latter has much better agreement with the architectural schematic shown in Fig. 7 which indicates that it is depicting the stable aspects of the environment such as walls and pillars. These will not change over time and so this map remains an accurate representation for longer. As the data in this map is over 2.8 m high then it will not be affected by people and the visibility is better, namely that in typical indoor environments features that are higher can be observed from more positions than those nearer to the floor, this is due to the increase in clutter at lower heights. Both of the overhead view depthmaps show close agreement with the schematic in Fig. 7 thus confirming the quality of the mapping process. The difference in mapping performance between the robots acting alone and cooperatively is highlighted in Fig. 9. The true position was found by manually aligning the laser scans and using prior knowledge of the environment. An experiment involving 4 robots taking 20 scans each was performed. Fig 9 shows the distributions of position errors for this experiment. When the robots are acting cooperatively, sharing map information and sightings, the error distribution is highly concentrated around 0.1 m however when the robot data is considered in isolation the number of errors greater than than 0.2 increases as may be seen by the area under the curve in Fig. 9. To summarize the use of multiple cooperating robots improves the accuracy of


Fig. 7. Architectural schematic of mapped area. Overhead view depthmaps from various height cross-sections of the 3D occupancy list at resolution 0.1 m . The first of all data and the second of voxels above 2.8 m but below the ceiling. Gridlines are separated by 1 m .


Fig. 9. Distribution of robot position errors for single and multiple robots
the map building process slightly but increases the reliability significantly.

## VI. Conclusion and Future Work

For decades mobile robots have been shackled by the 2D assumption, which has been necessary due to the absence of suitable 3D sensors. Even today 3D sensing technology is problematic and expensive. The two main approaches to 3D sensing are vision or laser based. The visual approach suffers from having to solve a difficult correspondence problem for each data point. The laser based approach has numerous advantages including relative insensitivity to light conditions, low realtime processing demands and the production of dense uniform 3D data.

In this paper a rotating mirror mechanism has been added to a standard SICK LMS 200 laser scanner which produces 3D scans. These 3D scans coupled with inter-robot sightings and exhaustive search scan alignment have produced full 3D environmental representations with resolutions of 0.1 m . It has been found that four robots operating cooperatively significantly improved mapping by increasing the reliability of localization. Mapping with four robots reduced the average position error from 0.35 m to 0.1 m .

Further work would include mounting the laser scanner so that it is facing up into the mirror to allow fuller scans. Employing the higher data rate interface to the scanner will
increase the possible scan frequency from 0.05 Hz to 0.5 Hz . The 3D occupancy grids produced by mutual localization, although interesting for humans to inspect, will mainly be used by other robots. The global localization accuracy is a good indicator of their fidelity and suitability for further use by other robots with global poses accurate to 0.1 m and $1^{\circ}$.

In this way a 3D mapping system has been devised with robots able to enter a completely unknown indoor environment, produce a 3D internal representation and then globally localize themselves using the relatively static geometric ceiling maps.

## REFERENCES

[1] A. Davison and Y. Cid and N. Kita. Real-Time 3D SLAM with WideAngle Vision. In IAV, 2004.
[2] W. Burgard, A. B. Cremers, D. Fox, D. Hahnel, G. Lakemeyer, D. Schulz, W. Steiner, and S. Thrun. Experiences with an interactive museum tour-guide robot. Artificial Intelligence, 114(1-2):3-55, 1999.
[3] A. Howard, A. Parker, E. Lynne, and G. S. Sukhatme. The SDR experience: Experiments with a large-scale heterogenous mobile robot team (extended abstract). Int. Symposium on Experimental Robotics, June 2004.
[4] A. Howard, D. F. Wolf, and G. S. Sukhatme. Towards 3D mapping in large urban environments. In IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pages 419-424, Sendai, Japan, Sept. 2004.
[5] J. Ryde and H. Hu. Laser based simultaneous mutual localisation for multiple mobile robots. In Proc. of Int. Conf. Mechatronics and Automation, pages 404-409, Niagara Falls, Canada, July 2005.
[6] J. Tuley and N. Vandapel and M. Hebert. Analysis and Removal of Artifacts in 3-D LADAR Data. In IEEE International Conference on Robotics and Automation, 2005.
[7] P. Jensfelt and S. Kristensen. Active global localization for a mobile robot using multiple hypothesis tracking. IEEE Transactions on Robotics and Automation, 17(5):748-760, 2001.
[8] C. Kwok, D. Fox, and M. Melia. Real-Time Particle Filters. Proc. of the IEEE, 92(2), 2004.
[9] K. Lingemann, H. Surmann, A. Nüchter, and J. Hertzberg. Indoor and outdoor localization for fast mobile robots. In Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems, pages 2185-2190, Sendai, Japan, Sept. 2004.
[10] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann. Heuristicbased laser scan matching for outdoor 6D SLAM. In Advances in artificial intelligence. 28th annual German Conf. on AI, Sept. 2005.
[11] R. Rocah and J. Dias and A. Carvalho. Cooperative Multi-Robot Systems: a study of Vision-based 3-D Mapping using Information Theory. Robotics and Autonomous Systems, 53(3-4):282-311, 31 Dec. 2005.

