Multisensoric Active Spatial Environment Exploration and Modeling

Axel Walthelm, Amir Madany Mamlouk
Medizinische Universität zu Lübeck
Institut für Technische Informatik
Ratzeburger Allee 160, 23538 Lübeck, Germany
{walthelm,madany}@iti.mu-luebeck.de

Abstract. *This paper presents new results for position tracking and map building in a special sensor oriented 3D world model for a mobile robot. The world model itself and the tracking algorithm are briefly explained and the performance of position tracking is discussed. For active online map building, the map building module has to decide where to explore next. To allow the action selection mechanism to mediate between different competing modules, a method to compute an attention distribution for gaze directions is presented.*

1. Introduction

Service–robots have to operate within a complex and partially unknown environment. Therefore they need to build up a model of their environment (a world model), using a limited set of multipurpose sensors. This model has to perform the task of mediating between sensors, goals and actions. When designing such a world model, it is essential to have a good understanding of strong points and weaknesses of the sensors, the essential questions the world model has to answer to fulfill goals and the available actions to acquire more information.

One of the most important tasks a world model has to perform is to maintain orientation and location of the robot in relation to its environment. Several approaches to solve this problem have been presented [1][2]. These approaches divide into two classes: landmark–based and grid–based. Landmark–based methods use some sort of high level features like artificial markers or walls. It is basically possible for them to use three dimensional information [4]. But these methods strongly depend on having high quality landmarks. They are not robust against detecting false or not detecting existing landmarks. Grid–based methods use low level features, typically occupancy of areas. Grid–based methods are restricted to two dimensional maps, due to the drastic increase of cells for three dimensions at a given resolution. Relatively few work has been done on active localization methods [3].

We have presented a world model which robustly uses low level features similar to the grid–based methods, but which is capable of spatial, three dimensional modeling [5]. It uses a 180^o SICK laser range finder mounted on a pan–tilt unit in combination with ultrasonic sensors. Fig. 1 shows the experimental hardware platform. While using a given environmental model to perform position tracking works quite well for 2D and 3D, building up a 3D model is much more difficult. This is because for building the map, position tracking already has to work sufficiently well. Due to the two dimensional scanning characteristic of the laser range finder the decision where to look at becomes vitally important for the case of 3D map building.

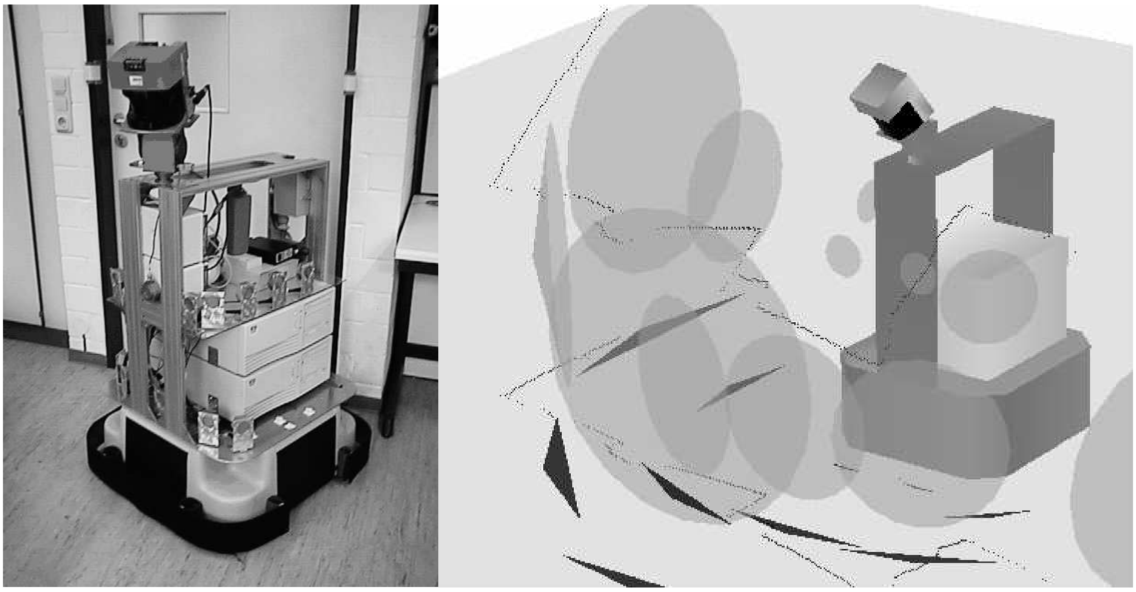

Figure 1. Left: Experimental Mobile Service-Robot MAVERIC (*M*obile *A*utonomous *V*ehicle to *E*xperiment upon *R*obotic *I*ndoor *C*hores) with SICK laser range finder and camera mounted on a pan-tilt unit. 16 ultrasonic sensors are arranged at different heights. **Right:** Visualization of sensor readings. Laser scanner measurements are displayed as dark dots. Half transparent circular areas indicate possible positions of ultrasonic echos. Black dots and triangles on the floor are downward projected "shadows" of these measurements.

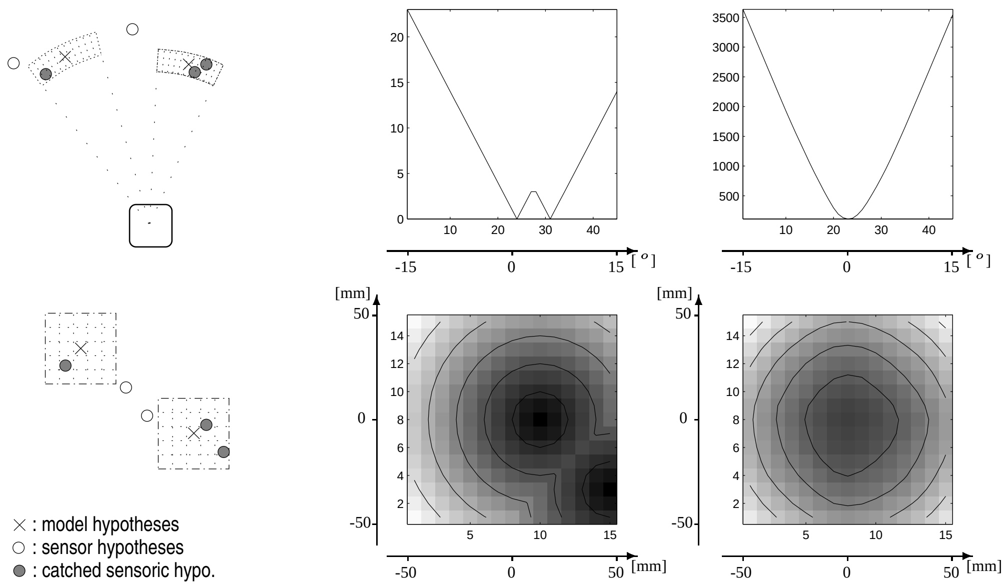

Figure 2. Left: Local areas around model hypotheses are searched for matching sensoric hypotheses. **Center:** Error distribution of a single model hypothesis, quantized for parameters $\Delta\theta$ and $(\Delta x, \Delta y)$. **Right:** Local errors are summed into a global error distribution. **Top:** Computation of vote vectors for $\Delta\theta$. **Bottom:** Vote matrices for $(\Delta x, \Delta y)$.

2. Position Tracking

The world model consists of a set of high precision occupancy hypotheses samples derived from laser range finder measurements. Rapid access to this set is done via an octree data structure. Efficient operations on octrees include insertion and deletion of elements and searching arbitrary volumes for elements. Thus by using a guess of how accurate the actual position estimation of θ resp. (x, y) is, for every actual sensoric hypothesis (measurement) a corresponding volume can be searched for potentially matching model hypotheses (fig. 2). All model hypotheses with potential matches generate a vote of the local resulting error of possible changes for θ resp. (x, y). This is done using a quantization of the ranges, i.e. a vote is a vector resp. a matrix. By propagating these imprecise local votes along the octree data structure, a robust overall vote is computed.

For an analysis of the stability of our tracking algorithm, we created a world model consisting of 13000 model hypotheses. The hypotheses have been build up from five different positions around the rooms center. The area of this world is ca. 10×10 m. For the analysis we took a previously recorded run of MAVERIC. On this run, position is tracked for each of the 460 horizontal laser scans and a detailed stability analysis at seven chosen points is performed. Results can be seen at the top of fig. 3. On the left we tested the stability regarding angle errors in a range of $\pm20^o$ with an increment of 4^o. The number of necessary iterations to correct the error are color coded: black if one, white if more than ten iterations are needed for correction. On the right side, the same has been performed for x/y-direction. An 11×11 error matrix with a range of ±1000 mm and an increment of 200 mm was chosen. The estimated maximal position error, i.e. the search range, was set constantly to ±200 mm and $\pm20^o$.

At many positions an error of 20^o and 1 m (!) can still be corrected successfully and in the average case errors of about 6^o or 600 mm can be fixed fast and accurate.

At the bottom of fig. 3 results for a run with a tilting laser scanner are plotted. Due to gaps and inaccuracies in world model[1] and the smaller average intersection of scan points and model hypotheses the results are slightly inferior but can still be used in most situations.

3. Map Building and Gaze Decisions

The world model, i.e. the map is built by taking over measured occupancy hypotheses which have no matching model hypotheses yet into the model. A single observation of a sensoric hypothesis is not always sufficient to qualify as a good model hypothesis. Therefore multiple observations are necessary to model a hypothesis which has the right to vote for position tracking.

Ultrasonic sensors are used to remove model hypotheses that apparently are no longer valid. This removes moving objects from the model. The same is done for the laser range finder, but due to the small volume covered, this sensor is of much less value for this task.

Obviously it is important to have a good position estimation when extending or changing the world model. Additionally it is crucial not to loose track of the current position and orientation, because MAVERIC has no other means of determining its global position yet.

Therefore the activities of MAVERIC have to be influenced or even determined by the needs of the map building and position tracking modules. This concerns $(pan, tilt)$, i.e. the direction MAVERIC looks at and the "point of view", i.e. (x, y, θ). We suggest the following set of rules and criterias to modulate these parameters:

[1] Note that the robot seems to drive through an object at the right side of the room; a table present at the time of map building was removed.

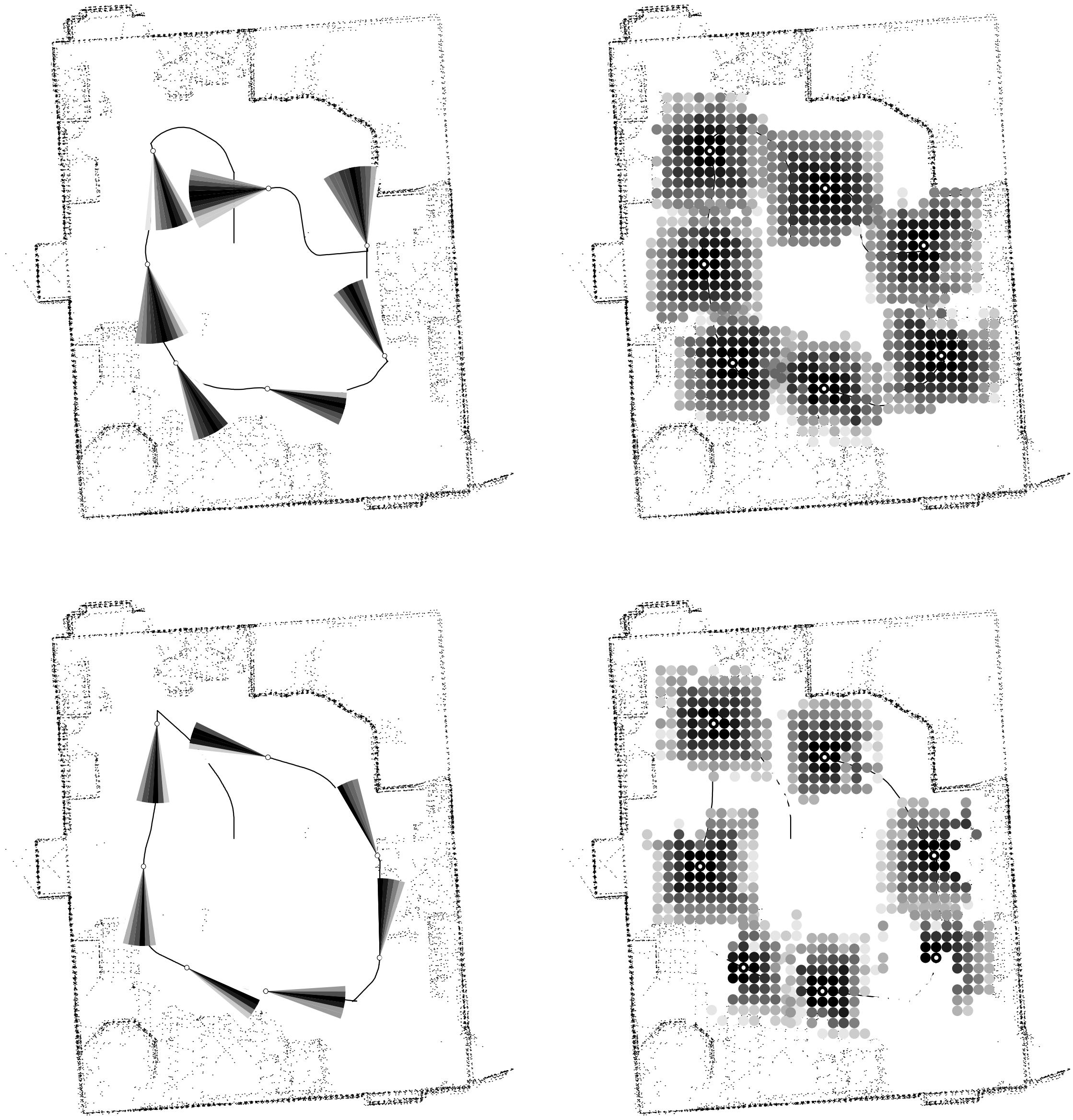

Figure 3. Stability analysis of the region sampling method. Tested positions are marked with small white dots. The solid line displays the trajectory of MAVERIC. Gray tones represent the number of iterations needed to correct the corresponding error (black:1, white:$\geq$ 10) to less then 100 mm and 3^o. **Top:** Horizontal laser scanner. **Bottom:** Tilted laser scanner. **Left:** Angle error in a range of $\pm 20^o$, increment of 4^o. The cones are aligned in driving direction. **Right:** Displacement error in a range of $\pm$1000 mm and an increment of 200 mm.

1. If position tracking is of poor quality then limit the speed $(\dot{x}, \dot{y}, \dot{\theta})$. In the worst case the robot must not move any more.
2. If a region of $(pan, tilt)$ contains hypotheses especially suited for position tracking, or
3. if a region of $(pan, tilt)$ does not contain hypotheses, or
4. if a region of $(pan, tilt)$ contains hypotheses of low certainty then look to this region.

While rules 2 and 4 support position tracking, rule 3 ensures an interest to explore unknown regions. Rule 4 is necessary to give new hypotheses a good chance to qualify for position tracking voting.

An indecisive vote is characterized by a low number of matched sensor hypotheses, a low total number of sensor hypotheses, multiple (local) minima for the overall vote matrix (especially for x,y), and in that the resulting correction vector is larger than the expected maximum error.

To find a set of points in the world model which potentially can be seen from the current point of view of the robot, occlusion should be taken into consideration. E.g. it is futile to try to look at points on the other side of a wall. As an efficient approximation of this goal, the field of view is quantized along pan and tilt and the octree is searched for the closest element within every quantization area. Searching all points within every quantization area that are near the known closest point is possible with similar computational costs.

To compute potentially interesting gaze directions based on a set of points, we have to take into consideration that the laser range finder of MAVERIC is a 180^o scanner. At a given pan-tilt-head position $(pan_0, tilt_0)$ a set of positions can be measured. So our aim would be to find the point $(pan, tilt)$ where the sum of utility of all potentially measurable points of the world model is maximal. Usually there are other reasons for the robot to look somewhere. Therefore it is reasonable not only to compute one best point, but to determine a set of potentially useful gaze directions and their expected utility.

Let's consider an arbitrary point $\vec{x}$ in robot coordinates. This point can be expressed in terms of pan p, tilt t, scan angle α and distance d. Equally $\vec{x}$ can be expressed in terms of pan_0 (p_0), $tilt_0$ (t_0) and distance, thus we get

$$\vec{x} \quad = \quad \mathrm{R}_z(p)\mathrm{R}_x(t)\left(\mathrm{R}_z(\alpha)\,\vec{d}+\vec{h}\right) = \mathrm{R}_z(p_0)\mathrm{R}_x(t_0)\left(\vec{d}+\vec{h}\right)$$

where $\vec{d} = (0 \quad d \quad 0)^T$, $\vec{h} = (0 \quad 0 \quad h)^T$, h is the distance of the laser scanner center from the t axis and $R_x(\phi)$, $R_z(\phi)$ are three dimensional rotation matrices around x resp. z axis with angle ϕ. As a simplification we assume h to be sufficiently small to be considered zero, which is valid if $d \gtrsim 6h$ (ca. 1 m for MAVERIC). This results in:

$$tilt \quad = \quad \arctan\left(\,\sin(t_0),\, \cos(t_0)\,(\cos(p)\;\cos(p_0) + \sin(p_0)\;\sin(p)\,)\right)$$

Fig. 4 visualizes this function for varying parameter $tilt_0$. Changes in pan_0 do only shift this function along the pan axis. To the right, fig. 4 shows the implemented result for three single hypotheses, the situation after adding the first scan to the world model and the situation in a well filled world model.

4. Conclusion

Active perception needs action selection. To explore an unstructured environment, many components like collision avoidance, position tracking and map building have to cooperate. A sensor

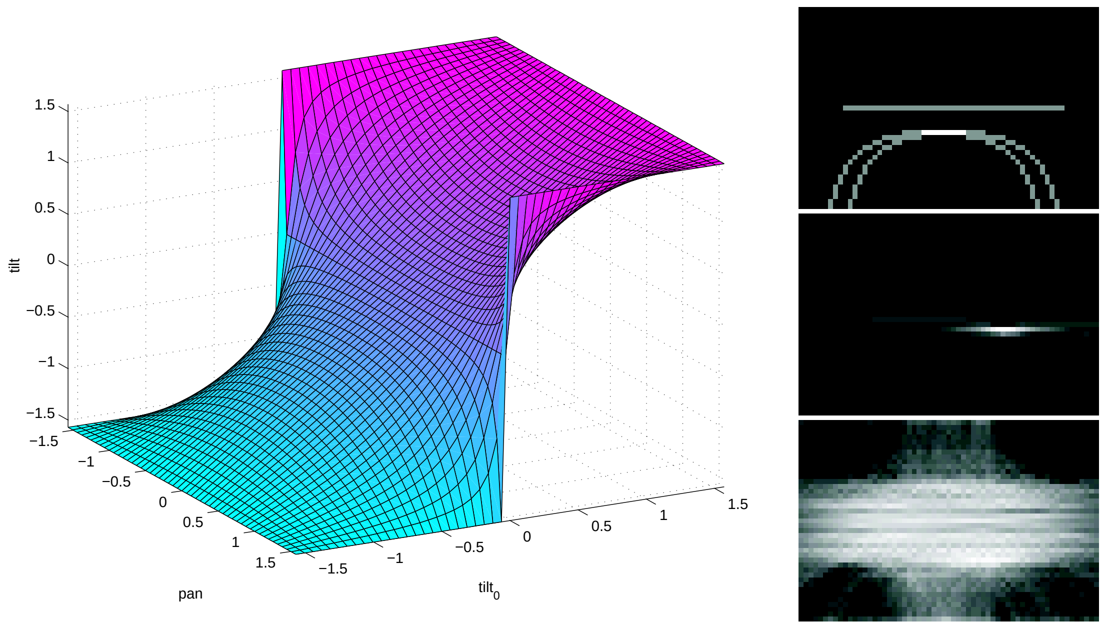

Figure 4. Left: Function $f_{pan_0,tilt_0} : pan \rightarrow tilt$ for $pan_0 = 0$. Changes in pan_0 shift the function along pan. **Right:** This function is implemented to compute an attention distribution for $(pan, tilt) \in [-120^o, 120^o] \times [-60^o, 60^o]$ with a resolution of 61×41. **Top:** Three visible points. **Center:** Situation after a first scan in the world model. **Bottom:** Situation in the complete world model.

oriented world model suited for position tracking was presented. The architecture for active autonomous map building was outlined and a way to compute an attention distribution for the map building module was presented.

References

[1] J. Borenstein, H.R. Everett, L. Feng: "Navigating Mobile Robots: Systems and Techniques", 1996 A K Peters Ltd.

[2] J.-S. Gutmann, W. Burgard, D. Fox, K. Konolige: "An Experimental Comparison of Localization Methods", International Conference on Intelligent Robots and Systems (IROS 98), Victoria, Canada, October 1998

[3] W. Burgard, D. Fox, S. Thrun: "Active Mobile Robot Localization", Fifteenth International Joint Conference on Artificial Intelligence (IJCAI'97), Nagoya, Japan 1997

[4] U. Frese, M. Hörmann, B. Bäuml, G. Hirzinger: "Global konsistente visuelle Lokalisation ohne vorgegebene Karte", 15. Fachgespräch Autonome Mobile Systeme (AMS '99), München, November 26-27, 1999

[5] A. Walthelm, R. Kluthe, A. Madany Mamlouk: "Ein 3D Weltmodell zur teilaktiven Positionsverfolgung in komplexen dynamischen Umgebungen.", 15. Fachgespräch Autonome Mobile Systeme (AMS '99), München, November 26-27, 1999