An Experiment in Semantic Correction of Sensor Data

Stefan Stiene and Andreas Nüchter and Kai Lingemann and Joachim Hertzberg

Common wisdom has it that all knowledge has to go through the senses first. While this is sort of true, it is only part of the story. The other direction does also make sense: Expectation matters for perception. In semantic robot mapping, the two directions need to meet.

In prior work [1], [2], we have developed the technology for acquiring 3D geometry maps in 6DOF on a mobile robot, for interpreting data in terms of building structures (floor, walls, ceiling) and for detecting objects in the geometry data. Part of interpreting is to process the data using, e.g., matching and filtering algorithms. All these algorithms, however, were local and "syntactic" in the sense that the laser scanner data were massaged and squeezed out as good as possible, but there was no model-based feed-back from prior findings to subsequent hypotheses. There was no explicit expectation about what might be perceived.

Without a semantic model, errors in the sensor data could only be corrected locally in the sense of outlier rejection and the like. Model-based perception would allow furthermore to complete the data (I know the wall continues behind the bookcase, although I have never seen it) and to correct illusions (I can tell the image of a robot from a robot if the image is hanging high on the wall). This has been way beyond our previous approaches.

We describe here a first small step into the direction of model-based sensor data correction. It was motivated by a systematic error of our 3D laser scanner equipment, which, due to poor calibration of the pitch control servo, tends to map a ground plane to a slightly bent surface.

3D mapping of environments consists of several steps to be executed, namely 3D scan acquisition, range image registration, and global relaxation. Since every step may potentially introduce errors, we are using semantic constraints to reduce the errors in all steps. Scans are acquired by our robot in a nodding fashion of the 3D laser range finder. The controlled pitch rotation can only be performed with limited accuracy, so a horizontal plane scanned by the laser may not be perfectly horizontally adjusted in the measured data. The key idea of our horizontal scan justification is to extract scanned planes, i.e., the floor plane and the ceiling plane in the scanned data, and to readjust the 3D scan using these horizonal information, according to the following scheme:

a) Point Labeling: Using the algorithm of [3], all scan points are labeled as floor, ceiling, or object points, based on a local geometric criterion wrt. their neighbor points.

The authors are with the Knowledge Systems Research Group of the Institute of Computer Science, University of Osnabrück, Germany. {stiene, nuechter, lingemann, hertzberg}@informatik.uni-osnabrueck.de

Fig. 1. Left: Map, top view. Right, top: The unconstrained 3D mapping shows a banana-shaped form. Right, bottom: The horizontal justification and the constrained mapping lead to qualitatively correct maps.

b) Bottom Plane Extraction: For an estimate of a 3D plane, we use all points labeled as *floor*. First, an initial plane is estimated using three data points. Second, the plane is adjusted so that the mean point-to-plane distance is minimal for *all* floor points.

c) Scan Justification: Floor points are horizontal in an office building and all on one plane, unless a clear jump edge is measured. The extracted 3D floor plane is used to rotate the 3D scan, such that the estimated plane – and therewith the 3D scan – is horizontal, i.e., parallel to the ground plane of the first scan, which defines the coordinate system. Hereby, pitch and roll errors are corrected.

Preliminary experiments were carried out in an indoor office environment. 33 scans, containing 88000 3D data points each, have been acquired. Fig. 1 shows the qualitative result of the constraint mapping. For quantitative results, we compared several distance measurements both in the map and in reality, using a high precision distancemeter (Leica DISTO). The accuracy of the constrained map differs only by several centimeters from ground truth, with a mean error of 2.11%, compared to 3.19% in the non-constrained case.

In the end, data-driven interpretation and model-driven data correction would have to come together. Fusing them is obviously a hen-and-egg problem. In probabilistic robotics, an EM-type approach might be suitable. For the time being, we would opt for a more flexible way of integrating the different knowledge sources that are relevant for the overall process, favoring a classical blackboard architecture. In addition, floor classification comes in handy for object segmentation and interpretation as well as path planning.

References

[1] A. Nüchter, K. Lingemann, J. Hertzberg, and H. Surmann. 6D SLAM with Approximate Data Associoation. In *Proc. ICAR*, 2005.

[2] H. Surmann, A. Nüchter, and J. Hertzberg. An autonomous mobile robot with a 3D laser range finder for 3D exploration and digitalization of indoor en vironments. *Journal RAS*, 45(3 – 4), 2003.

[3] O. Wulf, K. O. Arras, H. I. Christensen, and B. A. Wagner. 2D Mapping of Cluttered Indoor Environments by Means of 3D Perception. In *Proc. ICRA*, 2004.