MODEL UPDATE BY RADAR- AND VIDEO-BASED PERCEPTIONS OF ENVIRONMENTAL VARIATIONS

NORBERT O. STÖFFLER

TU München, Lehrstuhl für Prozeßrechner,
Prof. Dr.-Ing. G. Färber, 80290 München, Germany,
e-mail: `stoffler@lpr.e-technik.tu-muenchen.de`

THOMAS TROLL

TU München, Lehrstuhl für Hochfrequenztechnik,
Prof. Dr.-Ing. J. Detlefsen, 80290 München, Germany,
e-mail: `troll@hfs.e-technik.tu-muenchen.de`

ABSTRACT

The design of mobile robots which can cope with unexpected disturbances like obstacles or misplaced objects is an active field of research. Such an autonomous robot assesses the situation by comparing data from its sensors with an internal model of its environment. In this paper we present an object-oriented geometric model and an exemplaric model update: A new object is detected by a radar sensor and identified using a video sensor.

KEYWORDS: autonomous mobile robot, environmental modelling, microwave radar sensor

INTRODUCTION

This work is part of a research project[1] towards the development of autonomous mobile robots which can fulfil service and transport tasks in structured environments like office buildings and industrial plants. To be able to react to changes in its environment a robot has to survey its surroundings with one or more sensors. Based on the sensor data it updates an internal description (*environmental* or *world model*). This description includes the position of the robot itself, the positions of obstacles and the positions and states of other relevant objects. The model can also be considered as a set of hypotheses. These hypotheses can be tested and improved by comparison with the sensor data.

Raw sensor data is preprocessed to extract sensor-specific features. In case of video cameras these features are typically edges. In case of laser and microwave radar, which acquire 3-D range and velocity information, these features include intersection lines between object surfaces and the scanning sensor beam.

In our approach a common geometric model of the environment is used for all kinds of sensors. This avoids consistency problems between parallel sensor-specific

---

descriptions. Sensor-specific features are calculated online corresponding to the sensor type and the hypothesis to be tested.

By comparing the two sets of features - extracted and predicted ones - the hypothesis can be modified until it is approved and finally fed back into the model.

RADAR SENSOR

Using microwave radar in the field of robotics seems to be very uncommon. Such sensors are generally associated with the detection of objects like airplanes or ships over far distances. Since the propagation of microwaves is nearly independent of atmospheric conditions and coherent signal processing allows high sensitivity, it is possible to cover distances up to 100 m at reasonable transmitted power. This feature combined with the direct access to the object's velocity lets a microwave radar usefully enhance the sensing capabilities of an autonomous system, as other self-illuminating sensors like laser scanners and ultrasonic sensors hardly exceed a detection range of 15 m.

The project L4 is engaged in the development and evaluation of an experimental 94 GHz sensor. It is designed to measure distances as well as velocities of objects. The distances are determined via the time of flight of single pulses. A very short pulse width of 1.7 ns results in a radial resolution of 25 cm. The distance of resolved objects can be measured with an accuracy of 5 mm. The velocity is directly accessible via the Doppler effect. This effect requires a coherent signal source to guarantee a fixed relation between transmitted and reflected signal. In our case an IMPATT oscillator, phase-locked by a Gunn oscillator, guarantees high phase stability at sufficient peak power. The carrier frequency of 94 GHz yields corresponding Doppler frequencies of 625 Hz per 1 m/s, which permit accurate velocity measurement within a short observation time. For the angular resolution of $1.5°$ the radar beam is focused by a Fresnel type lens of 170 mm diameter. As shown in figure 1, this sharp beam is deflected by a mirror for three-dimensional imaging of the surroundings. The implementation of the sensor is described in detail in [3], the most important system parameters are shown in table 1.

μ

°

°

$\pm$
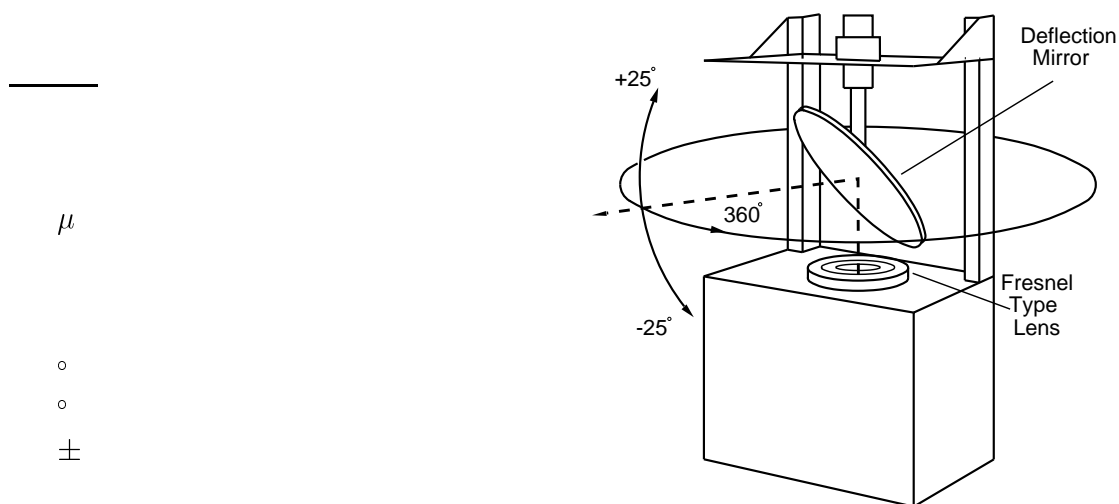
Based on this system various applications like navigation and observation, which

correspond to typical perception tasks of an autonomous robot, have been implemented in the last years [6].

VIDEO SENSOR

For the video sensor a standard industrial CCD camera is used. The projection of a 3D point in the scene into the corresponding 2D pixel of the video image is characterized by the model of a pinhole camera with radial distortions. The underlying camera parameters are determined by a specific calibration process (see [1] for a brief description). The image preprocessing extracting edges as image features is based on the image analysis system HORUS.

MODEL STRUCTURE

To permit sensor independent abstractions the model structure which is examined by the project Q5 is based on three-dimensional solid modelling techniques. Solid bodies are stored by a polyhedral boundary representation. If possible, sensor-specific features are calculated from this boundary representation using a corresponding sensor model. Intersections between e.g. a scanning radar beam and the boundaries can easily be calculated online. In the case of a video sensor an exact sensor model is difficult to obtain because the sensor data depends on various factors like surface properties and illumination. Furthermore an exact calculation would be very time-consuming. Therefore a simplified approach is used. The sensor model is reduced to perspective projection and edges of good visibility are represented separately by three dimensional line-segments, so-called *video-edges*. These video-edges are based on the same set of vertices as the boundary description but form only a subset of the boundary edges. This dualism reduces the prediction of video-features to a mere visibility test. It also allows a simplification of the boundary representation to exclusively convex polygons which additionally facilitates the visibility calculation.

To permit realtime access to the model appropriate index structures are necessary. In first experiments demonstrating localization in a static environment, a two-dimensional spatial-tree has proven to be an efficient index structure for otherwise unrelated model elements [5].

To allow more complex perception tasks and non-static environments, a hierarchic structure with additional symbolic information is currently examined [2] (figure 2a). Model elements are aggregated to *named objects*. Because it is neither possible nor necessary to describe the complete environment of the robot in terms of distinguishable, named objects, a pseudo-object called *background* is introduced. It encompasses all elements without special object assignment.

The description of a named object is built up recursively. An object can contain other objects which are termed *member-objects*. Object and member-object are connected by a joint which exhibits exactly one rotatory or translatoric degree of freedom, following the conventions used in manipulator kinematics. To deal with unknown states during a prediction the space potentially being occupied by a moving member-object is stored as an additional polyhedron, called *mask*. During the visibility test this polyhedron is used to literally mask out potentially hidden features.

Each branch in the object-tree carries its own boundary polygons and video-edges. Geometrically identical objects form an *object class*. The invariant parts of an object description are stored only once for each class; the objects (i.e. the instances of a class) differ in their individual positions and joint states.

MODEL UPDATE

In our approach the model update is divided into several widely independent perception tasks. They interact with the model on different levels of abstraction according to the parts of the model they regard as hypothetic (figure 2b).
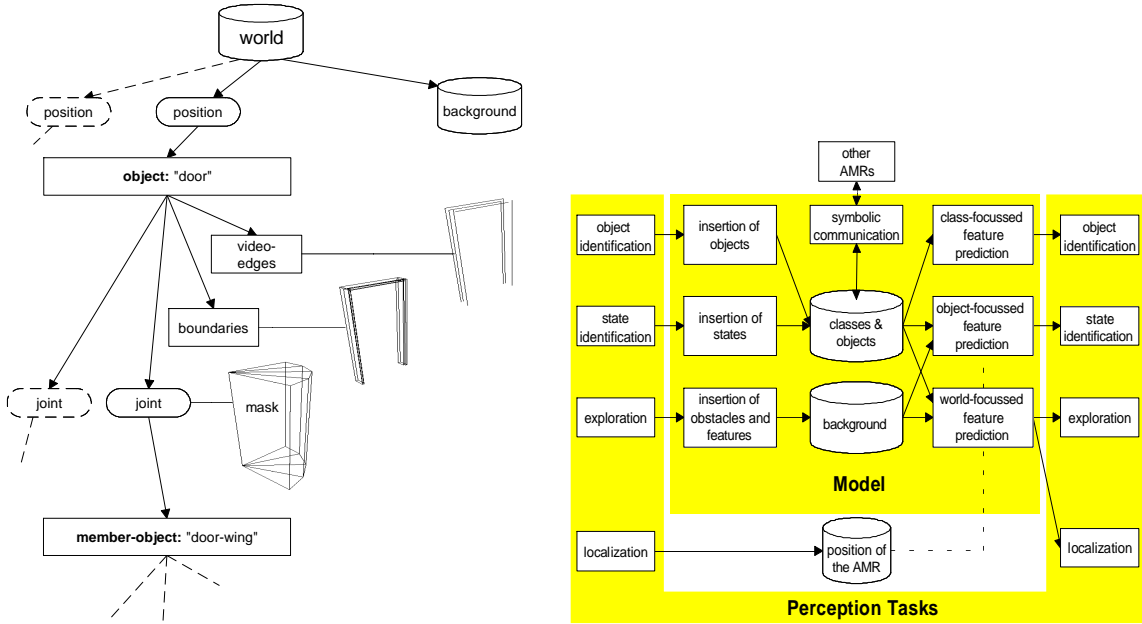

Figure 2. (a) Model structure, (b) Model update by parallel perception tasks

Each perception task is implemented by a separate client module which commands at least one sensor, extracts its own relevant features from the sensor data and simultanously requests predictions from the model server. Then it compares the two sets, interprets the difference and updates the model accordingly. Interferences between the quasi-parallel model accesses of different tasks are avoided by private communication channels containing local copies of hypothetic model subsets.

A common parameter which is assumed to be known by several perception tasks is the position of the robot itself. Localizing with a microwave radar or a video sensor is accomplished by establishing correspondences between the sensor data and model information followed by estimating the position of the sensor using least square optimization [5]. In case of the radar sensor the two-dimensional echo map is matched with the intersection lines of the solid body boundaries and the scanning plane. In case of the video-sensor extracted edges are matched with the projected *video-edges* supplied by the model.

In consequence of the radar's far detection range a localization can be carried out with a coarse or without any a-priori hypothesis about the robot's position. In contrast to that, localization with the video-sensor already needs a good position hypothesis.

Once a valid position is found and has to be successively updated, a tracking approach which uses the spatio-temporal restrictions of the robot-position can be applied. The last estimated or dynamically extrapolated position is used as a position hypothesis for the next prediction. If the cycle-time is short enough in comparision with the speed of the robot, model information can be used to reduce measurement time by defining regions of interest. Both types of localization, the initial localization with the radar sensor and the successive localization with radar or video-sensor have been demonstrated in several environments [5, 7].

Severe mismatch of the features predicted for the current robot-position and the sensor data indicates a variation of the environment. Therefore a second task evaluates these mismatches and initiates object identification. If none of the known object classes can be matched, the object is inserted into the model as part of the background. This at least prevents collisions and further mismatches.

The applied algorithms for video-based object identification and localization are described e. g. in [1, 4]. Predictions corresponding to single object classes are requested from the model and matched with the sensor data.

In addition to the model-update done by the various perception tasks information can also be updated on a symbolic level. Independently operating robots communicate about environmental variations by exchanging object names and attributes, i. e. states and positions, via a symbolic communication medium.

EXPERIMENTAL EXAMPLE

The shown experiment was carried out at the experimental industrial plant of the *Institut für Werkzeugmaschinen und Betriebswissenschaften (iwb)* using the experimental mobile platform MAC1 (figure 3a). The platform is equipped with a radar and a video sensor and several computers connected by ethernet and TCP/IP. All modules, i. e. model and perception tasks, are implemented by RPC-servers respectively clients. The platform is moving along a clearance between the machine-tools which is expected not to be obstructed (figure 4a). The radar sensor is scanning an area of 5 m x 40° in front of the vehicle. To cover the space down to the floor the radar beam is inclined by 12.5°. In figure 3b radar echos which can be matched with predictions of the model are shown in light grey. Echos depicted in dark grey can not be matched and indicate the presence of a yet unknown object. Based on this coarse position hypothesis the video-based object identification consisting of an object recognition followed by a fine localization is initiated.
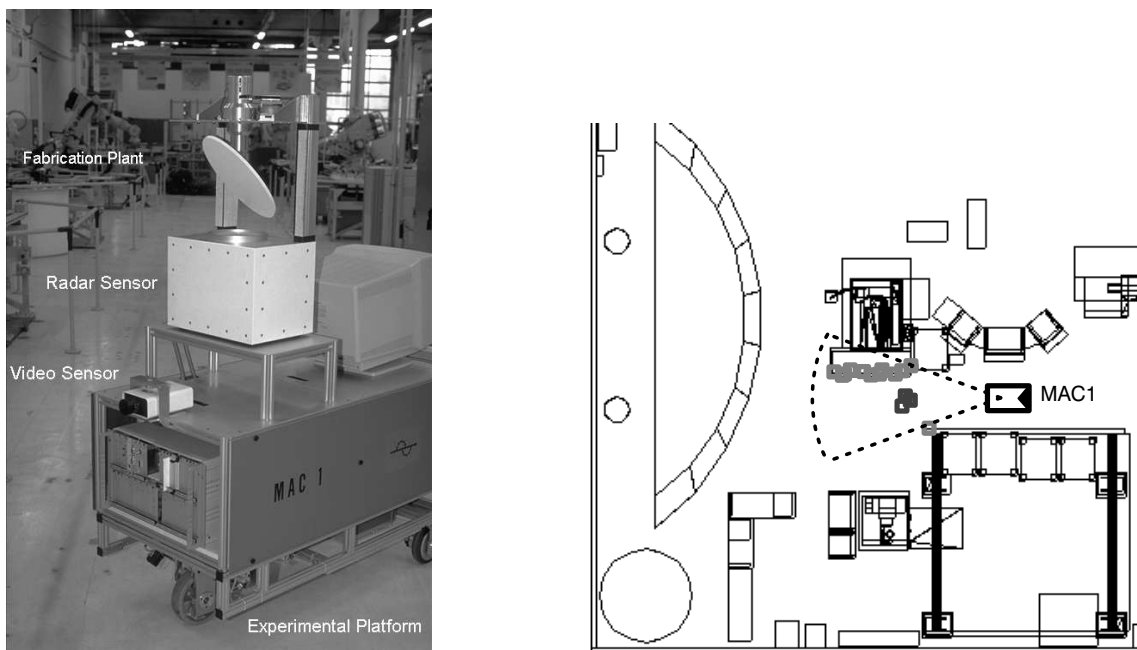

Figure 3. (a) Experimental platform MAC1, (b) correlation of radar echos and model

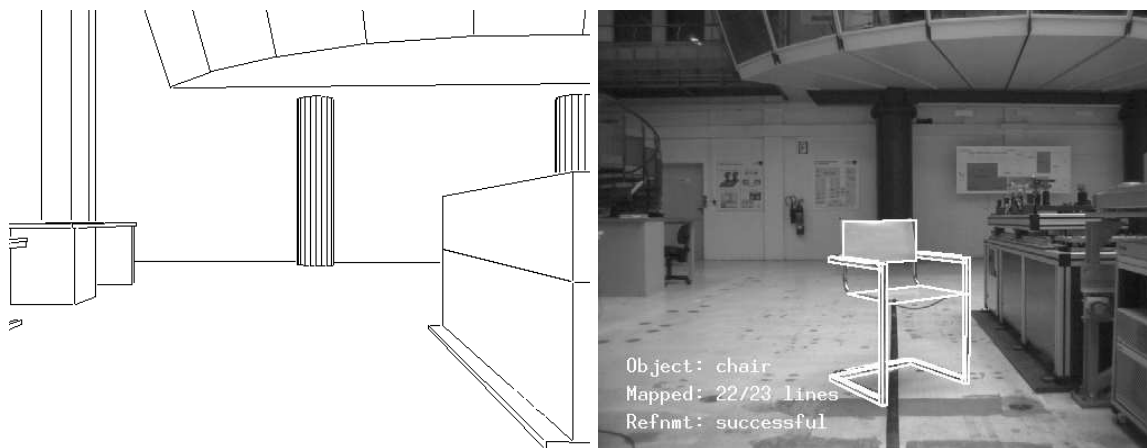The video-image and a feature-prediction corresponding to the class "chair" and

Figure 4. (a) Model before the update, (b) identified object

the final object position and are shown in figure 4b. Once the object is identified it is instantiated in the model.

CONCLUSION

The results of the experiments show the suitability of the described model structure to support concurrent perception tasks on a mobile robot. The division of model updates into several widely independent perception task facilitates combining the advantages of very different sensor types. The range data of the radar sensor is exploited in two ways to improve the performance of object identification. By defining a region of interest the computation time for the image preprocessing is typically reduced by the factor 3 and the initial position hypothesis yields additional constraints for object identification.

REFERENCES

1. S. Blessing, S. Lanser, and C. Zierl. Vision-based Handling with a Mobile Robot. In *6. ISRAM*, Montpellier, May 1996.
2. Alexa Hauck and Norbert O. Stöffler. A Hierarchic World Model supporting Video–Based Localization, Exploration and Object Identification. In *2. Asian Conference on Computer Vision*, pages 176–180, 1995.
3. M. Lange and J. Detlefsen. 94 GHz Three-Dimensional Imaging Radar Sensor for Autonomous Vehicles. *IEEE Trans. on Microwave Theory Tech.*, 39(8):819–827, May 1991.
4. Stefan Lanser, Olaf Munkelt, and Christoph Zierl. Robust video-based object recognition using cad models. In U. Rembold, R. Dillmann, L.O. Hertzberger, and T. Kanade, editors, *Proc Conf. Intelligent Autonomous Systems*, pages 529–536. IOS Press, 1995.
5. Achim Ruß, Stefan Lanser, Olaf Munkelt, and Michael Rozmann. Kontinuierliche Lokalisation mit Video- und Radarsensorik unter Nutzung eines geometrisch-topologischen Umgebungsmodells. In *9. Fachgespräch "Autonome Mobile Systeme", München*, pages 313–327, 1993.
6. M. Rožmann and J. Detlefsen. Environmental exploration based on a three-dimensional imaging radar sensor. In *Proc. IROS*, pages 422–429, Rayleigh, NC, 1992.

7. M. Rožmann and J. Detlefsen. Standortbestimmung in Innenräumen mit einem hochauflösenden 94-GHz-Radarsensor. In H. J. Eberhardt, editor, *8. Radarsymposium der DGON*, pages 43–50, Köln, 1993. Verlag TÜV Rheinland.