Real-Time Object Detection for Autonomous Robots

M. Pauly, H. Surmann, M. Finke and N. Liang

GMD - German National Research Center for
Information Technology, D-53754 Sankt Augustin, Germany
surmann@gmd.de

Abstract. A new kind of robots whose characteristics, objectives and operational modes drastically differ from more conventional industrial robots is gaining increased interest. This new type of robots aims to achieve a high level of flexibility, adaptability, and efficiency for acting in environments designed for humans. To meet these demands, the mobile robots need lots of information about their surroundingsat different levels of abstraction, generally in real-time. In this paper a sensor concept is described containing symbolic and sub-symbolic information delivering the required data for all tasks. Based on this information the planning, navigation, and collision avoidance of the ARIADNE robots is done. [1]

1 Introduction

Different kinds of sensors are required to control an autonomous robot effectively [6]. A mobile robot traveling in dynamic environments continuously needs e.g. the current distance to all the objects next to it. These kind of sensors should quickly register all changes in the close environment to avoid collisions. Another kind of sensor should collect informations about objects, in order to build maps or plan the missions. Therefore sensors which detect the position of an elevator/door or the position and velocity of another robot moving in the same area can be useful.

In order to equip a mobile robot with suitable sensors, there are two main problems: On the one hand the number of sensors which can be mounted on-board the robot is limited due to the restricted resources. The spatial and energy resources for the sensors are limited. Thereby, the computational capacity is restricted. Furthermore, with every new sensor the quantity of data that must be processed increases until the right information is found. This is the reason why in todays applications only a few sensors are used to collect the wide spectrum of required information from the environment.

On the other hand redundancy of information is important to correct errors. A typical type of error where sensor redundancy is needed are errors in calculating the robots current position [1]. Therefore, different kinds of sensors, mostly using different physical principles, are used, like wheel encoders and landmarks [3] for estimating the current position. A further problem is the suitable combination of data sets from different sensors, e.g. sonar sensors, laser scanner, bar-code scanner [7, 2] and their computation in real-time [5].

[1] M. Pauly and H. Surmann, M. Finke, N. Liang, Real-Time Object Detection for Autonomous Robots. Informatik Aktuell. Autonome Mobile Systeme (AMS'98), 14. Fachgesprche, Springer-Verlag, pp. 57-64, 1998.

One solution to the dilemma restricted resources versus needed redundancy is the usage of a limited number of sensors in combination with a-priori knowledge about the environment and the robot's tasks. In the ARIADNE project a multilayered sensor scheme is used. It is based on two laser scanners for each robot. In combination with a-priori knowledge from an on-board database the needed information for navigating, planning, and world-modeling is generated using the laser data. For this, a sensor scheme consisting of five levels of abstraction has been developed and validated.

The rest of this paper is organized as follows: In section 2 the ARIADNE team is presented. The sensor concept containing symbolic as well as sub-symbolic sensors is described in section 3. Finally, we conclude in section 4 with results and a summary in section 5.

2 ARIADNE

ARIADNE is a team of autonomous intelligent service robots for indoor environments. The major application areas are administrative buildings, like insurances, banks, hospitals, and offices. The variety of tasks is very large: from mail, lunch, or coffee distribution on schedule or on demand via intra-/internet during the day to patrolling at night. To support the planning and logistics of such robot service teams, the robot service group will also be integrated into the virtual reality of the GMD-Workbench, a virtual working environment [9].

The ARIADNE team consists of 3 mobile robots (Fig. 1): Odysseus, Marvin and Thor. Each of them is about 80 cm x 60 cm large and 90 cm high. The mobile platform can carry 200 kg of equipment at speeds of up to 0.8 m/s. The right and left driving wheels are mounted on a suspension in the center point of the mobile platform. Passive casters in each corner ensure stability.

The main computer on-board the robot is a Pentium PC 100 MHz with 16 MB RAM running real-time Linux. Also 2 micro-controllers are used. One of them controls the internal state of the robot and the communication with the control server. The other micro-controller manages the motors and the optical line-following-unit. Each platform is equipped with two laser scanners: one in front of the robot and one looking backwards. The mean time for a complete scan of 360 distance data is 200 ms.

Each robot is connected with a central server via a radio link. This server acts as an interface between the robot and its environment and holds a central information base. To enable the robot to reach every part of the building, the server controls doors and elevators. If a robot wants to drive to another hall or floor, it sends a request to the server to open the doors or call the elevator. For user interaction, the current jobs for the robots are administrated on the server. If a robot has free capacity, e.g. transport capacity to carry out a job, it sends a request to the server to receive a new job.

Besides autonomous navigation, which is done by using fuzzy-logic [8], the robot has the ability to drive along an optical line. This navigation is preferred in narrow passages, where the robot must drive exactly, e.g. while entering an elevator, and during docking manoeuvers.

Fig. 1. The ARIADNE team: Odysseus, Marvin, and Thor (GMD corridor in C2 basement).

3 The Sensor Concept

To build a system with an adequate number of sensors and also to get sensor data at different levels of abstraction, a five-layered sensor scheme was developed and realized (Fig. 2). In this concept different real and virtual sensors place the needed data at the robots disposal. The term virtual sensor means, that the needed sensor information is not delivered by a real physical sensor. Instead, the information is computed – in the most cases by combining data from one or more other sensors with a-priori knowledge.

Using a combination of real and virtual sensors to get the needed information, the number of sensors on-board the robot can be reduced. At the same time the quality of the information is nearly constant. A further advantage of the usage of virtual sensors is the decoupling of hardware and the needed sensor information [5].

Redundancy of data can be achieved by only a few sensors. In this case the redundant information has to be extracted, using different algorithms, out of the raw data delivered from the sensors. A disadvantage of this practice is the computation time which is needed to extract the information. Therefore in this paper a concept for processing the real sensor data at five different levels of abstraction in real-time is presented.

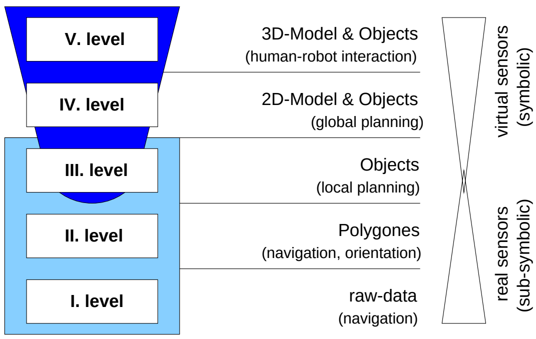

Fig. 2. Concept for a real-time spatial cognition, consisting of five different levels of abstraction.

The five levels (layers) contain all data for the different tasks of the robot (navigation, orientation, planning and human-robot interaction). The layers are divided into two parts: a sub-symbolic part, containing the original and pre-processed raw data, and the symbolic part, containing all the symbolic information about the environment.

In order to meet the real-time conditions three types of control and data exchange between the sensors, real or virtual, and the robot module are implemented: automatic delivering, delivering on demand, and sending a data-ready-signal.

For time critical tasks there is an automatic delivering of the current sensor data. If a new data set is generated by a sensor, the set is directly send to the processing modules. This kind of data exchange guarantees a short response time, like it is needed for obstacle avoidance.

To less time critical tasks the sensor only sends a message (data-ready-signal) to the modules and processing units. If the task has free resources it can get the data and process them.

Non time critical tasks, like high-level visualization tasks, look for new sensor data on their own. An advantage of this scheme is, that the current information are only computed and send on demand.

3.1 Sub-symbolic Sensors

Sub-symbolic sensors deliver all data from raw distance data up to pre-processed data, like lines or polygons. This data does not contain any information about the object itself, e.g. usability.

Level I: At the lowest level of abstraction the sub-symbolic raw-data from the real sensors are pre-processed, e.g. for navigation and collision avoidance. This data in our case consists of a set of distance values generated by a laser scanner. The data-rate at this level is 6 Hz. In the first step a reliability test is done. The data are verified using the a-priori knowledge about the environment and the data sets collected before. In the next step the information is condensed, merging similar distance points together. Thereby, the amount of data which is sent to the processes can be reduced. The information in this layer is used for velocity control and the emergency stop. If one data set is generated from a sensor at this level, it is send directly to the task handling this data.

Level II: At this layer the physical dimension of objects in the close environment are generated from the pre-processed data from level I. Therefore, the borders of the objects are approximated by lines. The real border of an object is always behind these calculated lines.

After that, lines are combined to longer lines or polygons (Fig. 3). Single lines are compared with each other. If the deviation between the compared lines is below a given limit, both lines are merged together. Thereby, complex polygons and longer line segments come to exist.

Beside this data further information about the environment is generated at this level, like the orientation of the walls in a corridor. This information is used e.g. for navigation and also for correction of the robot's orientation in the environment.

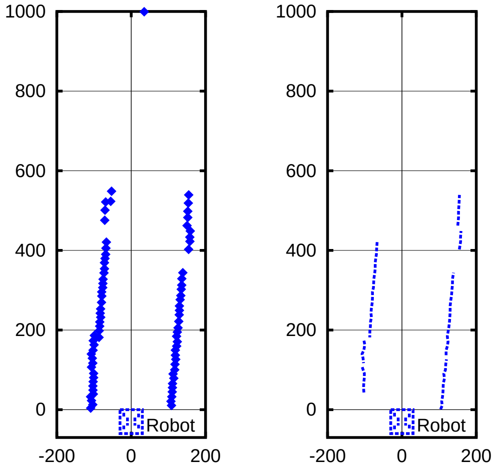

Fig. 3. Pre-processing step 1: After deleting redundant data (left side), pre-processing step 2: lines and polygons extracted from the raw-data (right side), dimensions are given in cm.

Level III: Level III is the interface between the data coming from the lowest level (generated by the real sensors) and the information located at the upper levels (virtual sensors). Hence this level contain both: symbolic and sub-symbolic information. They are combined using a fuzzy-matcher.

In a first processing step objects are generated out of the polygons from level II (Fig. 4). Therefore, the polygons are compared with objects in a database. In this database a set of representations from the objects is stored. Especially for indoor environments the number of important objects is limited. In addition to the object database the symbolic information from the fourth and fifth level is used to recognize the objects, e.g. [4].

Once an object is recognized further information from the database is attached to it. After this process the virtual sensors contain both geometric information, like the position and the outside border, and symbolic information of the object. For example, a virtual door detection sensor delivers information about the position of the door as well as information about the condition (open or closed and the room to which this door leads).

Path planning as well as position estimation is done using the information delivered from the sensors in this level. Therefore, the sensors send a signal to some processes that new data are available and deliver the current data on demand. A data-ready signal (detecting a door) is e.g. sent to a process doing the position update of the mobile robot. This kind of information should be used by the robot as soon as possible. Less time critical are the data for planning. Therefore, the sensors only send the data on demand.

3.2 Symbolic Sensors

The sensors located in the symbolic level contain not only physical information about a single object or the environment but also information about the handling, like codes to open the doors, or data about the endangerment for the robot. These data are computed from a-priori knowledge stored in a database.

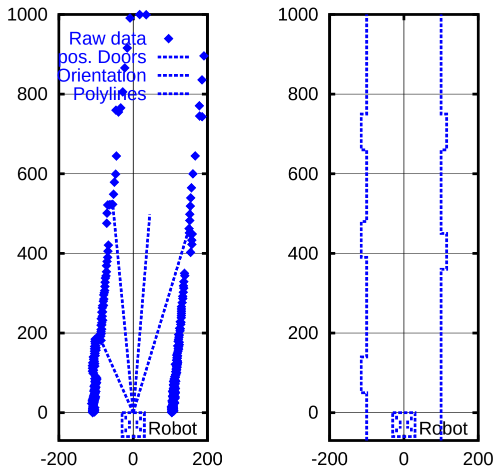

Fig. 4. Combination of polygon and real data (left side), part of the floor plan (right side).

Level IV: Symbolic information for the global planning of tasks and paths are generated from the sensors in this level. Therefore the virtual sensors combine information from a database with current information about the near environment. Typical sensors are e.g. sensors generating a two dimensional map or sensors detecting dangerous areas. Also a virtual sensor which detects dynamic objects is realized. Therefore, a-priori information about the static environment (world model) is needed.

In addition to this kind of sensor, a sensor detecting areas dangerous the robot is also important for generating plans. This sensor compounds object information, e.g. the position of doors stairs, and elevators, from level III with a-priori knowledge and data from level V, e.g. the spatial influence of specific doors or the dependence of the current traveling speed on the extension of the dangerous area.

The sensors in this level deliver the current information only on demand, in order to get a short update cycle. Since the data are mostly needed only by the planning process, the communication overhead of a continuously data sending sensor can be avoided. The consistency of the world model is checked and updated every time the robot reaches the end of a corridor or other typical landmarks.

Level V: At level V a three dimensional symbolic description of the environment is used. The 3D-model contains spatial information for the operator doing tele-monitoring and tele-maintenance. Furthermore, this representation of the environment contains auxiliary information about objects.

This object information includes besides auxiliary data for the operator also methods or scripts how the robot can interact with the object. It contains e.g. a temporal sequence of pictures taken from the object to show the history to the operator. In addition to these hints for the operator, the sensors in level V deliver data for a three dimensional spatial model as well as for temporal path- and task-planning.

The 3D-spatial sensor has the ability to give information about the current position of all robots in the environment and also their current mission. Furthermore, the current state of their mission task can be requested (waiting, running, canceled). The information of this sensor are used to plan the global path and task for the mobile robots.

4 Results

The sensor concept was implemented and tested at the robots of the ARIADNE team. The three robots Odysseus, Marvin and Thor drive in the C2 building - 3 floors connected through an elevator - at Schloss Birlinghoven, GMD. The test data were collected during a 7 day demonstration for the CeBit, the Hannover Fair and during various tests at the GMD.

In Level I the data rate was reduced, from the raw data to the calculation of the single lines, to 30% of the original data. The data rate in level I is 6 Hz. The processing of one scan took 2 ms, including the reliability test and data reduction.

In the next layer of abstraction (level II) the lines were generated with a data rate of 5 Hz. The length of single lines is, for a typical environment at GMD, 50 cm up to 2 m. The number of lines in one scan is about 7-15 (depending on the complexity of the environment), taking 1.2 ms of computational time per scan.

The objects were extracted from the lines with 1 Hz (level III), with a recognition rate for objects with, e.g. doors 97.8% and boxes 98.9%. The correctness of the object position is up to 5 cm without matching with the world model in the database. The recognition time for e.g. a door takes 1.5 ms.

The robot position and the position of new objects is updated every second. The 3D-model and information in level V is only computed on demand. It took about 250 ms for the virtual sensor to calculate the information (GMD environment with three robots).

5 Summary

In this paper a multi-layered sensor scheme was presented. Using this concept the number of sensors on-board the robot can be reduced in order to save resources. The missing real sensors are replaced by virtual sensors, which compute the needed data.

The robots get all needed data at different levels of abstraction at the different layers. Also three different levels of control as well as three different types of data exchange are realized in order to meet real-time constraints.

The sensor concept has been validated during various tests at GMD using the three robots of the ARIADNE team. Future work will improve the long term robustness of the system in dynamical environments. Furthermore, additional sensors will be integrate in this concept, like video cameras for reading door plates. With this sensor concept, the robots get the ability to autonomously collect subsymbolic and symbolic information about their environment.

References

1. J. Borenstein and L. Feng: Measurement and Correction of Systematic Odometry Errors in Mobile Robots, IEEE Transactions on Robotics and Automation, 7 (7), (1996) 869–880

2. F. Dierks: Freie Navigation autonomer Fahrzeuge, Informatik aktuell: Autonome Mobile Systeme 1994, Springer-Verlag, Berlin, (1994) 43–54
3. H.R. Everett: Sensors for Mobile Robots: Theory and Application, AK Peters, Wellesley, MA, (1995)
4. J. Kreuziger and W. Wenzel: Learning Generic Object descriptions for Autonomous Robot Systems, Proceedings of the Fifth International Symposium on Robotics and Manufacturing, ISRAM'94, Maui, (1994)
5. O. Kubitz, M. Berger and R. Stenzel: Client-Server-Based Mobile Robot Control, IEEE/ASME Transactions on Mechatronics, 3 (2), (1998) 82–90
6. M. Pauly: Ferninspektion mit mobiler Sensorik - Ein Konzept zur Unterstützung des Menschen bei der Durchführung von Überwachungsaufgaben, PhD-Thesis, Logos-Verlag, Berlin, (1998)
7. J. Porrill: Optimal Combination and Constraints for Geometrical Sensor Data, The International Journal of Robotics Research, 7 (6), (1988) 66–77
8. H. Surmann, J. Huser, L. Peters: A Fuzzy System for Indoor Mobile Robot Navigation, Fourth IEEE International Conference on Fuzzy Systems, Yokohama, Japan (1995) 83–88
9. G. Wesche, J. Wind, M. Göbel: Visualization on the Responsive Workbench, IEEE CG & A, 17 (4), (1997) 10–12