Exploration of Unknown or Partially Known Environments*

Darius Burschka, Christof Eberst

Institute of Process Control Computers
Prof. Dr. -Ing. G. Färber
Technische Universität München
80280 Munich, Germany
e-mail: {*burschka*|*eberst*}*@lpr.e-technik.tu-muenchen.de*

Abstract This paper describes our present results in the generation of a local 3D environmental model primarily based on the line-based stereo-vision and its planned extension to other sensor systems. It describes the way the explored data is processed to reconstruct the environmental description and to support the sensor data processing. We present the way the sensor data is obtained, stored and modified in our *dynamic local map* (DLM) of the environment and how it is processed by the *predictive spatial completion* (PSC) module to prefer specific matches in ambiguous situations and to control the exploration of a local area. The described system is developed to stabilize and filter the uncertain sensor information. Missing features due to physical or procedural limitations are predicted and verified or rejected by the applied sensors.

1 Introduction

1.1 Motivation

The ability to operate in partially known as well as dynamic or unknown structured indoor environments is essential for an *autonomous mobile robot* (AMR) to fulfill its tasks. A geometrical model of the environment helps the AMR to interpret the sensor information. This model can be a-priori known or be explored by the sensor system. A particular task requires a specific level of abstraction of the used information.

A localization task requires only the spatial properties of particular features to compare their modeled with the current pose.

The reconstruction of the object faces requires a knowledge about the relationships between the extracted features.

The sensor information underlies physical and procedural limitations reducing the precision and the reliability of the detected features. The kind of ambiguities differ depending on the applied sensor system. In case of a line-based stereo system the geometrical and relational constraints between the detected features do not solve the correspondence problems definitely (section 2.1). The detected information must be verified with additional information uncorrelated to the one used in the current step. It can be derived from a third camera or an environmental model representing the accumulated information from previous sensor readings from different positions (section 2.4) and from an a-priori known model.

The explored information consists of a feature-based description of the environment. It must be further processed to recover the relationships between the detected features.

1.2 Approach

The interaction between an environmental model and an active module completing the explored information is a new approach compensating some sensor limitations. The presented approach is based on a division of the environmental model into a stable, global and a dynamic, local description.

The global part containing stable geometric as well as symbolic elements is described in [2]. In the following it is referred to as GSM (*geometric symbolic model*).

The local, feature oriented part contains the uncertain information and is called *dynamic local map* (DLM).

The information flow between these two parts is controlled by the *predictive spatial completion* module (PSC) (fig. 1).

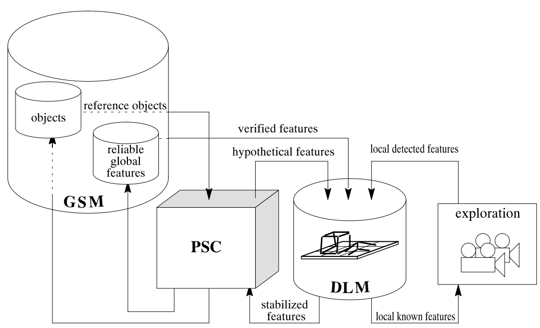

Figure 1: Information flow.

The DLM stores the current sensor information and stabilizes it by evaluating succeeding sensor readings.

*This work was partially supported bei the *Deutschen Forschungsgemeinschaft* as part of an interdisciplinary research project on "Information Processing in Autonomous Mobile Robots" *Sonderforschungsbereichs 331*

It stores only the description of a local area instead of a global model in case of the GSM. The stored information is used to accelerate the sensor data processing by a fast verification of the known features in the local sensor view before the time-consuming stereo-matching algorithms are applied. The stored sensor information during an operation in a local area results in a feature-based sensor-specific model of the local environment.

The PSC uses the information stored in the DLM to generate additional features based on reference objects stored in the GSM (fig. 1). These features are subsequently verified or rejected by the applied sensors. The hypothetical features generated by the PSC are used in the sensor data processing to prefer particular matches. Often these features cannot be detected without the PSC because of physical and procedural limitations. The interaction between DLM and PSC improves the sensor data processing significantly.

The GSM stores the verified sensor information and the a-priori environmental model [2]. This information is used to initialize the grid elements of the DLM, when a new area is approached. This structure helps the sensor to involve the a-priori and already explored information for the first orientation in a new area. All sensor systems are connected directly to the DLM, which exchanges its information with other modules like GSM and PSC (section 2.3).

This approach can be extended to fuse information from different sensor systems. The applied sensors store their sensor-specific features in the local map. The PSC merges the different environmental descriptions into a unified world model and it completes the missing features in the particular feature descriptions.

2 Acquisition and handling of the environmental information

2.1 Feature extraction

Currently, we apply a line-based binocular stereo-system for the exploration. Our system allows a 360 degree scan of the environment from a single position, which is helpful for initial path planning. The cameras are attached to a camera mount, capable of turns in azimuth and elevation. The vergence angle between the cameras can also be controlled.

There are different kinds of image processing applied to the camera images depending on the aspired task. The edge detection must be applied to the whole image in case of an *exploration* task, where new features are expected, whereas it can be reduced to windows around the predicted features in case of the *localization*. In our system the edge detection is applied to whole image. The found edges in both pictures are matched in consideration of geometrical (epipolar lines) [1] and relational (neighborhood) constraints. These restrictions are not sufficient to solve all correspondence problems between the two images (fig. 2). In some cases an additional information is required to determine the correct matches. This information is derived from the three-dimensional features of the local environment stored in the DLM.

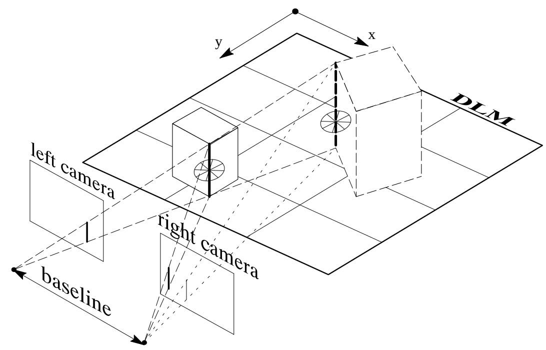

Figure 2: Two unresolved matching results inserted into the DLM with attached visibility maps.

The information stored in the DLM helps to establish known correspondences between the detected features in the images. It accelerates the sensor data processing, because only the new and unknown correspondences must be established in a time-consuming matching algorithm. The stored information gives important clues about correct matches in ambiguous situations, which are verified from different positions in consecutive sensor readings. Each feature description includes a confidence value accumulated with each successful matching within the DLM. The sensor data processing can determine the amount of the reliance on the known feature depending on its *confidence*, *accuracy* and *age* stored in the DLM.

2.2 Feature representation

The detected features are stored in the DLM as three-dimensional lines with additional attributes describing their confidence, accuracy, age, orientation in space, visibility in form of a visibility map and their source.

The *confidence* describes the accumulated information from the sensor data processing about the correctness of the feature. Features resulting from definite correspondences between only two image lines have a higher confidence than features obtained from multiple matching candidates.

The *accuracy* describes the estimated error in position of each endpoint of the line. This value is also estimated during the sensor data processing based on the location of the feature within the sensor range.

The *age* of the feature describes the last successful matching of the map content to a sensor view. Features aged beyond a limit are removed from the DLM. The "age" of an feature is only modified, when the feature is within the sensor view.

The *visibility map* allows a prediction of free space between the features required for navigation. This map is actualized by the DLM itself, when a new feature is matched to an already stored one. It marks the range of angles perpendicular to the orientation of the feature from which it was seen.

The *source* of a feature is stored as a bitmap, where each bit represents a specific source. In this way it is possible to determine the source and the matching history of the modified feature. This can be a combination of a verified hypothesis from the PSC, a certain feature from the GSM and the exploration result.

2.3 Possible sources of information

The features stored in the DLM originate from different sources. A new grid element in the DLM is initialized with the certain information from the GSM (fig. 1). This information is put together from a-priori knowledge and already explored, stabilized information stored by the PSC module during an earlier operation in this area (section 3.4).

The most important source of information are the applied sensor systems, which verify the stored information with the present situation to incorporate the recent changes in the environment to the model.

The PSC stores hypothetical features in the DLM resulting from assumptions about the scene and from the reconstruction of the known objects, which references are stored in the GSM (section 3.2).

2.4 Handling of the input data

The local detected features are continuously compared with the old map content to improve the accuracy of the stored information and to reject wrong features caused by false correspondences. An aging mechanism removes the false features if a feature is not found in an adjustable amount of succeeding sensor views. The way the age of a feature is modified (section 2.2) allows a correction of the position of a feature extracted in the past, if the vehicle is still operating in the local area. This allows an exploration during a normal mission. In this case there is often not enough time for an accurate exploration of the entire area, if it is not relevant for the current path planning. A repeated detection of a feature in several passes during different missions can also stabilize its description in the DLM.

The features are stored according to their geometrical position. Each query in the DLM returns only the features in a local sensor view.

2.5 Structure of the DLM

The function of the DLM is similar to that of a computer cache. The "real world" is mapped on the DLM allowing a fast access to the local information. The DLM's structure is optimized for continuously changing data from applied sensor systems. It allows an adaptation to the obstacles' distribution in the world and simplifies the modifications of the content, when a new area is approached. The local detected features are matched within the DLM with the already explored features to improve their accuracy and confidence. The following requirements were considered in the structure of the DLM: limited storage space, fast access time and simple modification of the content. The resulting structure of the DLM consists of an coarse grid dividing the local area into single grid elements, which size is adapted to the sensor range. Each grid element contains an octree (fig. 3). This structure combines the advantages of both structures. The octree adapts to the obstacles' distribution and the grid structure allows simple modifications of new elements without time-consuming transfer-operations within the tree. The features are stored at different depths of the resulting octree depending on their size. They are stored as deep as possible to increase the selectivity of the query functions. It is desirable to reduce the possible matching candidates in each access. In a query all octree branches in the DLM are intersected with the sensor cone representing the sensor view to extract only the relevant features. Therefore, features stored in upper levels of the octree appear almost always, when the grid element is entered, whereas the features stored in the lower branches appear only, if they are within the sensor view.

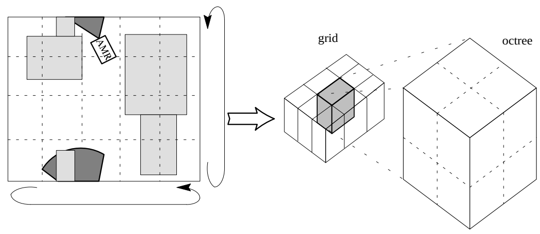

Figure 3: Each grid element can be independently mapped to a local section of the environment (cache structure).

2.6 Resulting environmental model

The information stored in the DLM consists of a pure sensor-specific feature description of the local environment. The stored features originate from different sources: GSM (a-priori and certain features), PSC (hypothetical features) and sensor (explored). These features have different accuracy and confidence depending on the structure of the environment, physical and procedural limitations. These values are modified in consecutive sensor frames.

At present we apply a DLM capable of storing three-dimensional lines resulting from a line-based stereo-matching. These line attributes and their geometrical appearance determine the location where they are stored in the DLM. The DLM will be extended to store other kinds of features, like planes for a laser-range camera, etc. A combination of different sensor systems will help to reduce the physical limitation of the particular sensor. The function of the PCS will be extended not only to complete single descriptions, but to combine them to an unified model.

3 Reconstruction of spatial structures

The capabilities of the DLM resemble the ones of the applied sensor system. Our sensor based feature extraction is merely capable of establishing memberships of features to planes. Therefore, the DLM does not store any information about structures and objects. This information must be recovered by the PSC to reconstruct the spatial structures. It is useful for e.g. a hidden-line removal or a detection of objects relevant for a mission.

The hidden-line removal simplifies the localization in the environment by removing invisible and occluded feature parts in the local sensor view. In this way the wrong correspondences between the model and the sensor readings are strongly reduced. This operation requires a higher level of abstraction from the stored information. The sensor-specific features must be assembled to planes, clusters and objects.

3.1 Clustering of explored features

Newly explored features are tested for their interrelationships to others. Adjacent features are tested whether they are related with others and assembled to structures as polygons, corners, and clusters. If one closed substructure is found, the equivalent feature on a higher level is derived. This means, that e.g. a face is derived from a polygon of CCD-edges. If a closed description of a structure is found, faces, structures and objects are generated (section 3.4).

3.2 Generation of hypothetical completions

A spatial structure, that can be assembled incompletely from the explored data, triggers a hypothesis about its completion. The missing features are predicted. These features show a hypothetical status. Even if they are plausible, their can be wrong or, since they are based on explored data, inaccurate. Predicted features have to be verified with the sensor readings. This is done in the DLM. Features can be generated by three approaches:

- The features are compared with common substructures as concave or convex edges and parallel structures. Simple completions can be derived.
- Features that give a clue to the presence of an object of known type initiate an object completion and recognition process. Each possible interpretation of the feature as a part of an object leads to the generation of a hypothesis about the type and pose of the underlying object. Explored features in the vicinity of the expected object are used to reduce uncertainty and ambiguities about the object and its position. For the remaining plausible hypotheses, missing features of the object description are predicted. The validated features are exploited for the object identification.
- Multiple sensors can give a clue to the underlying structure. Structures, expected due to the data of one or more sensors can be used to derive hypothetical features for other sensors. These features are predicted to be validated by the according sensor. In the PSC features of multiple sensors can be exploited to predict new features. The match of these features with the sensor data of the congruent sensor usually leads to an adaptation of the feature description. This can be considered as a fusion of sensor data, even if one sensor can provide its data just with delay.

3.3 Verification and deduction

Verification of the predicted features is done in the DLM. The predicted features are compared with the sensor data. A feature is supposed to be verified if it can be matched with detected ones or if a feature is matched at the predicted place in further frames. Verified and falsified features are used to draw conclusions about the hypothesis and the underlying structure. Verified features can easily be clustered with others since the structure, or object, it is based on, is known.

3.4 Transfer of found features and structures into the GSM

Features as well as verified clusters, structures and objects are transfered into the GSM. Extracted faces and structures are exploited for hidden line calculation to enhance the localization. This information is also important for correct initialization of the DLM, when the surrounding area is reentered. The verified information has not to be explored repeatedly, but must be merely verified in a less time-consuming algorithm.

4 Application of the gained information for vehicle and sensor control

4.1 Support of reliable feature extraction

The tight coupling between the DLM and the sensor data processing is motivated by the poor accuracy and the ambiguities of the features detected by the sensor data processing in a single frame. To achieve the accuracy sufficient for model building and mission planning, the DLM stabilizes the extracted features by averaging their description over a time period. The hypotheses generated by the PSC reduce some physical sensor limitations by involving knowledge about the environment in the matching process (section 2).

4.2 Free space estimation for navigation and obstacle avoidance

Navigation in structured indoor environments requires a knowledge about the free space around the vehicle in consideration of the newest sensor information. The newly explored features are often very inaccurate or can be artefacts, hindering the exact path planning. However, they indicate the possible presence of an obstacle and have to be considered for the motion control of the robot. The key of the path planning is the extraction of possible pathways. Pathways are defined by a gap between features that is wide enough to pass through. This scene can be misleading since these features can represent a gap or one solid plane. A promising approach is to combine geometrical indicators applied in the PSC with visibility tests in the DLM. If a feature that lies behind an assumed plane is not hidden, the plane assumption can be rejected - a gap is more plausible. Extracted concave or convex structures combined with the visibility range and the vehicle's position helps to distinguish between a pathway and a solid surface. However, transparent material is a problem. A laser radar can be applied to resolve some of the remaining ambiguities. Poor illuminated objects can often be completed faster due to the predicted features. The knowledge of their shape can be exploited for path planning.

4.3 Determine regions of interest

The vicinity of the hypothetical features represents an area that includes features with a high possibility. Checking this area promises a detection rate far higher than average. The verification of a predicted feature reflects a high amount of relevant information, since its relationship with other features is already known. Another way to get a high increase in information with a minimum of costs is to improve the description of uncertain features. This let it seem desirable to focus the interest on these regions. The robot can drive to or focus its cameras on the regions where features appear in a high density to explore potentially poor known environments.

4.4 Map construction

The environment is perceived by the sensor systems in form of single features. This representation is adopted by the DLM to be stabilized in succeeding sensor readings. The requirements for a fast access and elementary modifications imply an unattached structure. The GSM stores the environmental description at multiple levels of abstraction to serve the heterogeneous demands of the object recognition, hidden-line removal and localization. In unknown environments these abstract levels must be gained from the low-level information contributed by the sensors. This is achieved by the PSC based on assumptions verified by the DLM.

5 Conclusion

The presented structure forms a core of our exploration system. It is capable of handling uncertain information in structured indoor environments. The cooperation between the *dynamic local map* (DLM) and the *predictive spatial completion* (PSC) module results in a robust reconstruction of the environment and an acceleration of the sensor data processing. A *DLM-cycle* (query and update of the stored information) resulting in approximately 100 features out of several thousands averages 20 ms an a DEC's Alpha-workstation with a 21064 processor (SPECint92 114, SPECfp92 162). The amount of features that have not been detected in prior frames lies below 20 % - 40 % depending on the distance between the frames and the accuracy of the position estimation. The presented approach leads to a hierarchical representation of the environment consisting of features, faces, structures and objects. This has been found useful for different kinds of sensors as well as for locomotion and exploration tasks. With each sensor reading the exploration provides fragmentary knowledge about the environment, which is completed in successive readings. As soon as there is appropriate information in a particular level it is abstracted to the next higher level.

References

[1] O. Faugeras. *Three-Dimensional Computer Vision*. Massachusetts Institute of Technology, The MIT Press, Cambridge, Massachusetts London, England, 1993.

[2] A. Hauck and N. O. Stöffler. A hierarchic world model supporting video-based localisation, exploration and object identification. In *2. Asian Conference on Computer Vision, Singapore, 5. – 8. Dec.*, 1995.

[3] X. Lebegue and J. Aggarwal. Extraction and interpretation of semantically significant line segments for a mobile robot. *Proc. IEEE Int. Conf. on Robot. and Autom*, pages 1778 – 1785, May 1992.

[4] G. Magin and A. Ruß. Supporting real-time update of an environment representation for autonomous mobile robots. In *6. EuroMicro Workshop on Real-Time Systems*, 1994.

[5] G. Magin, A. Ruß, D. Burschka, and G. Färber. A dynamic 3d environmental model with real-time access functions for use in autonomous mobile robots. *Robotics and Autonomous Systems*, 14:119–131, 1995.