Creating a 3D environment map from 2D camera images in robotics

J.P. Niemantsverdriet
jelle@niemantsverdriet.nl

4th June 2003

Timorstraat 6A
9715 LE Groningen
student number: 0919462

internal advisor: prof. dr. L. Schomaker (Artificial Intelligence, R*u*G)
external advisor: dr. T. ten Kate (TNO TPD Delft)

Artificial Intelligence
Rijks*universiteit* Groningen

1 Introduction

My graduation research will take place at TNO TPD in Delft. I wanted to do a research project in robotics, and was especially interested in the field of processing sensory information. Within the project, the emphasis will be on the processing of camera images by the robot. The target is to construct a 3-dimensional map of the environment, which can be used by the robot for navigation. A framework for image analysis, developed by TNO TPD, will be used to achieve this goal. The following research questions will have to be answered:

- Which method can be used to transfer 2-dimensional camera images into a 3-dimensional representation?

- Which map representation can be used for navigation by the robot?

- How can the robot process image data in a robust way?

Besides these questions, a few more 'standard' aspects of robotics have to be implemented. Examples of these are the avoiding of obstacles and the use of sonars.

2 Theoretical background

Humans and animals rely greatly on their visial system for their navigation. Seemingly effortless we see a picture of our environment, see where we should and where we shouldn't walk and are also able to remember all this information. In robotics, the possibilities to watch the environment are far more sparse. Sonars can be used to measure distances around the robot and cameras are able to transfer images of the environment, but the possibilities of these techniques are (still?) far less than the human or animal visual system.

This project aims to develop and implement a system, which can be used by the robot to navigate in its environment by using the camera. The robot should be able to construct a 3-dimensional 'map' of its environment, just like people would do when they walk through a room and take a look at the environment. To achieve this, the robot has to be able to see certain objects and also to remember information about these objects. By combining all this information, the robot will be able to use the information from the camera images to navigate (for example to find out on what position it is, or which route it should take to another place).

However, as already noted, the possibilities to process visual information aren't anywhere near the possibilities of the human visual system. Another problem lies in the poor reliability of the odometry of most robots: they aren't very well capable of knowing their position based on motion of the wheels. So even if you are able to recognize some visual property in the environment, you can't be too sure of the robot's position.

A solution to these problems could be, to integrate all these sensory percepts. If the robot is able to combine the visual information, the odometry and the information from its sonars, it could be able to reduce the errors in

all these sensors. Some other solutions are mentioned, like using another robot which controls the position of the main robot [8], or the use of laser range finders to build the map [6].

A solution to these problems could be, to give the robot some a priori knowledge about its environment. The robot has already some model of the environment (a digital map of the floor it's on, for example), and uses the sensory information to update or refine this model.

A pitfall that has to be avoided, is trying to predict everything and build a complete representation of the environment before the robot starts to do anything. This approach, commonly referred to as "classical AI", has been replaced by an approach more based on direct feedback. An example of this approach is the subsumption architecture by Brooks [5]. This architecture breaks the processing of information into several modules, which act independently from eachother. In this project, one could imagine a module for obstacle-avoidance, one for image analysis, one for planning, et cetera. The robot drives around, first mainly guided by its obstacle-avoidance, while the image-analysis module is analyzing the environment and refining the map. Later, when the robot has achieved more knowledge about the environment, the information from the visual module can be used to navigate in a more reliable way.

3   Research question

The main research question deals with the development and implementation of a thorough method to transfer the 2-dimensional image from the camera into a 3-dimensional overview of the environment. This method also deals with the saving of the map, so the robot will be able to use older knowledge about the environment.

The research question is:

In what way can a mobile robot use its camera to refine a visual map of its environment, and use this map to navigate in a robust and reliable way?

At TNO TPD, there is a lot of knowledge and experience with processing digital images, which could be used. A few aspects of this problem are:

- measuring in images

- constructing the map from the images (so from 2-dimensions to 3-dimensions)

- data processing in a robust way (because sometimes the data can be sparse or noisy)

The modular design will give the possibility to split the implementation into several modules. If one module doesn't work in the desired way, the robot is still able to use the other modules. This design could also be useful for testing other approaches: only the relevant module or layer has to be changed; the other programmes can be reused. Especially the 'low-level' functions like obstacle avoidance are probably needed in every robotics setup.

4 Methods

To process the camera images, a framework developed by TNO TPD will be used. This framework offers a connection between *DirectShow [1]* (a *DirectX* component) with *Matlab [2]*. This framework gives the possibility to collect the images with *DirectShow* in a very fast way and to do advanced analysis on them with *Matlab*. A few filters are already available in *DirectShow*.

The robot is a Nomad Scout from Nomadic Technologies, which has had some major hardware upgrades to make the use of the above programs possible. Together, these hardware and software offer the possibility to do real-time analysis of high-quality images. The robot is equipped with a single camera system.

The research questions noted above will have to be answered using implementations in this platform. The image analysis can mostly be done by *Matlab* programming (and some aspects by the already created *DirectShow* filters). The control of the robot and the use of its sonars has to be done in C++.

To analyse the images, a lot of basic image analysis algorithms[9] have to be implemented (such as texture detection, measuring). As already mentioned, the *DirectShow* framework is able to do some basic filtering on the images. In [7] (and some other work from the same author) a probabilistic method of analysing the images is described. Such a method could be useful to create a reliable representation of the environment. There are also various examples of navigation using distinctious characteristics (also called landmarks) [10]. In [4], features from images are analysed using Principal Component Analysis.

5 Scientific relevance for AI

Navigational systems based on visual information are an important field of research within AI. A robust visual system is an important characteristic of an autonomous vehicle. The construction of such a system combines knowledge about physiology, signal analysis and programming, and is therefore a truly multi-dimensional project. The ability to refine and update a priori knowledge about the environment could also be useful on various occasions (for instance when an autonomous vehicle re-enters a formerly visited place, where some changes have occurred).

If the robot achieves its correct behaviour, this is also a very good result for the image analysis framework. Further and more advanced work could then be done afterwards.

6 Planning

The figure below denotes a global overview of the planning. The total project will reside about 20 weeks. The start of the project will be in the beginning of July 2003. The thesis will be finished in December 2003.

week	task
1-2	study literature, first experiments with the robot and the TNO framework
3-4	design of the algorithms and programs
5-15	writing the C++ and Matlab source code and testing the behaviour
15-18	evaluating the results of the robot, adjusting parameters
19-20	finishing thesis

The main part of the thesis will be written during the project. The last weeks are reserved for the finishing of the thesis and making small adjustments and corrections.

7 Resources and support

TNO will submit the infrastructure needed for this project. The robot and computer facilities are situated in the TNO TPD building in Delft, as well as facilities for writing the thesis, which can also be done at home.

The project will be implemented in C++, and will use the *Matlab/DirectShow* framework developed by TNO TPD. By using this framework, the operating system for the robot has to be Microsoft Windows (probably a "mobile" version like Windows CE). If possible, already existing software to control the motors and the sonars of the Nomadic robot will be used [3].

Support at TNO TPD will be given by dr. T. ten Kate. Prof. dr. L. Schomaker will be the advisor from AI in Groningen.

8 References

[1] *DirectShow website*. http://www.microsoft.com/Developer/PRODINFO/ directx/dxm/help/ds/default.htm.

[2] *Matlab website*. http://www.mathworks.com/products/matlab.

[3] *Nomad software*. http://nomadic.sourceforge.net.

[4] N. Vlassis Y. Motomura B. Kröse, R. Bunschoten. Appearance based robot localization. *IJCAI-99 Workshop Adaptive Representations of Dynamic Environments*, pages 53–58, 1999.

[5] R. Brooks. New approaches to robotics. *Science*, (253):1227–1232, 1991.

[6] D. Hähnel, W. Burgard, and S. Thrun. Learning compact 3d models of indoor and outdoor environments with a mobile robot. In *The fourth European workshop on advanced mobile robots (EUROBOT '01)*, 2001.

[7] D. Hähnel, R. Triebel, W. Burgard, and S. Thrun. Map building with mobile robots in dynamic environments. In *Proceedings of the IEEE International Conference on Robotics and Automation*, 2003.

[8] I. Rekleitis, R. Sim, G. Dudek, and E Milios. Collaborative exploration for the construction of visual maps. *IEEE/RSJ/International Conference on Intelligent Robots and Systems*, 3:1269–1274, 2001.

[9] L.G. Shapiro and G.C. Stockman. *Computer Vision*. Prentice-Hall, Inc., 2001.

[10] R. Sim. *Mobile Robot Localisation Using Learned Landmarks*. PhD thesis, McGill University (Montreal, Canada), http://www.cim.mcgill.ca/ simra/publications/Thesis/foe.html, 1998.