Model and Vision Based Pose Estimation for Mobile Robots

O. Madsen[3], M. Ayromlou[1], C. Beltran[2], A. Gasteratos[2], W. Ponweiser[1], M. Vincze[1]

[1] Institute of Flexible Automation, Vienna University of Technology, Gusshausstr. 27-29/361, A-1040 Vienna, Austria

[2] DIST Laboratory for Integrated Advanced Robotics, University of Genoa, Viale Causa 13, I-16145 Genoa, Italy

[3] Department of Production, Aalborg University, Fibigerstraede 16, 9210 Aalborg, Denmark

Abstract. A system is introduced which can perform automatic, real-time estimation of the pose of a mobile robot moving in a known environment – in this case a ship section. In the system a pose estimate is generated continuously while the robot is moving based on the combination of visual sensing information and CAD-information. To reach an acceptable performance of the overall system a special emphasis is placed on robust feature detection and real-time capabilities. The update cycle time is 120ms but this could be enhanced. Experiments show that the approach is feasible and reaches the demanded positioning accuracy's.

Keywords: Mobile robots, pose estimation, model based vision.

1 Introduction

This paper presents a system for automatic, real-time estimation of the pose (position and orientation) of a mobile robot moving in a steel structure as found in ships. This pose estimation system is needed to perform closed loop control of the path of the robot. A number of different technologies have been used to generate pose estimations, e.g. vision, acoustic, laser, and radar (see [1] for an analysis of the state-of-the-art of the technologies used in this area). However, presently no technology exists which fulfils general demands to speed, accuracy and prize. The system presented in this paper (in the following referred to as the RobVision[1] system) is designed for controlling a mobile robot moving in a ship section, where the environment is relatively structured, and reliable CAD-models of the ship section exist. The system presented is based on the use of vision. Two alternative image-processing techniques are used, a monocular and a binocular, to take advantage of the arising redundancy. To ensure correct feature detection, visual sensing information from the environment is integrated with predetermined CAD-information [2],[3]. After feature extraction, the pose estimation algorithm can use both 3D and 2D-feature information found in the images to calculate the current pose and send it to the robot.

2. The RobVision System

Figure 1 shows the main outline of the RobVision system. As it appears from this figure1, the RobVision system consists of 4 main components:

CAD-System:	Generating references to the other components.
Stereo head	Controlling the cameras such that they will point at areas in the environment with good features and stabilising the camera movements while the robot is moving.

[1] This work has been mainly supported by the RobVision project Esprit 28867.

Vision	identifying the location of different features in the environment
Pose calculation	computing the pose of the robot by combining CAD-information with the identified features.

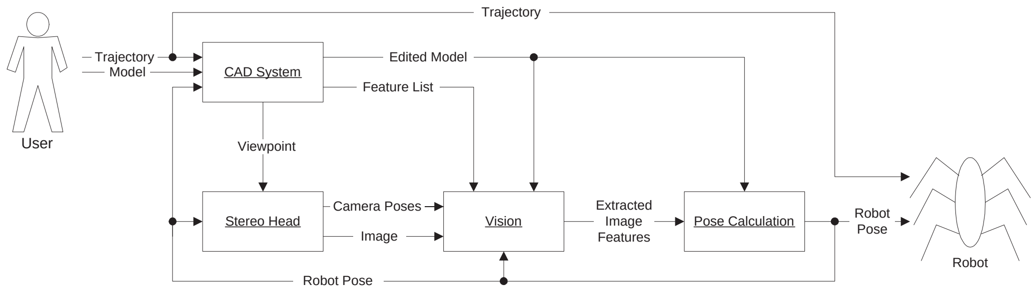


Figure 1. RobVision system components.

The basic mode of operation of the RobVision system is as follows: First, the user plans a desired trajectory of the robot. The robot is then put in a well known location and the movements are started. Information on the pose of the robot in its starting point, together with the desired robot trajectory and a geometrical model of the environment is used by the **CAD-System** to compute a viewpoint (specifying where the robot should look) and a feature list (specifying what the cameras are expected to see). The **Stereo Head** is then moved such that the cameras are looking in the desired direction, and the **Vision** component attempts to identify the features specified by the **CAD-system**. This information is sent to the **Pose Calculation** that computes the pose of the robot by combining the CAD-information with the identified features.

In the next sections the components are described in more details.

2.1 CAD System

The task of the CAD system is to provide the vision component with geometrical features. These features specify the shape and location of geometrical entities such as lines, junctions and regions that the cameras are expected to see while the robot is moving inside a steel structure. As it appears from Figure 1 the CAD system has three types of input:

- A *CAD model*, which is a geometric model of the steel structure inside which the robot is going to move.
- A *Reference Robot Trajectory* containing a specification of the task or trajectory that the robot has to follow.
- *Robot Poses* computed by the pose calculation component.

The CAD-model and the reference robot trajectory are computed prior to the operation of the RobVision system, whereas the robot pose is computed while the robot is moving. An example of a CAD-model and a Robot Reference Trajectory is shown in Figure 2.

The main output from the CAD system is model and view data containing:

- A *Camera Viewpoint*, specifying a 3D point (x,y,z) in world coordinates towards which the stereo head ought to point.
- A *Feature List* containing a set of robust features which the vision systems are expected to find when looking at the specified viewpoint.
- *Model data* containing a specification of relevant information from the underlying geometrical model.

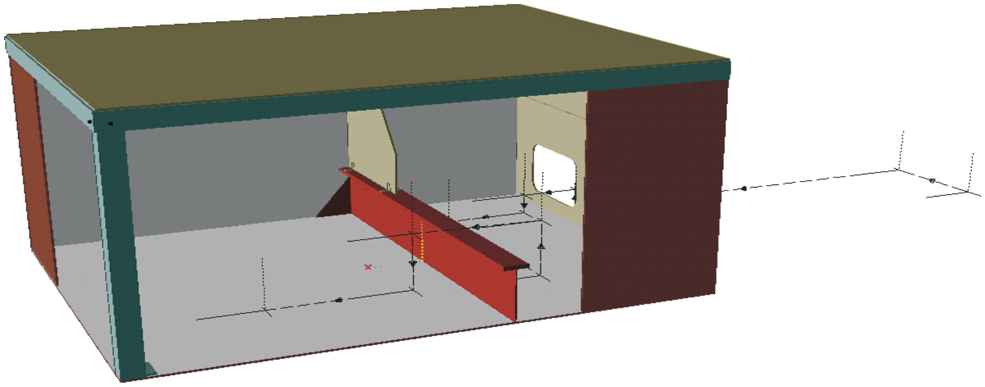


Figure. 2. A CAD-model and the Reference Robot Trajectory

2.3 Stereo Head

The stereo head is a 3-degree of freedom head mounted on the mobile robot. The head is equipped with two cameras. The task of the stereo head is to realise the correct camera orientation and image stabilisation [4]. The necessary input is the viewpoint selected from the CAD system and the actual robot pose. From this input the corresponding axis positions are extracted, and the joints are moved accordingly.

The RobVision system uses the Eurohead of LIRA-Lab (see Figure 3). For the control of the pan, tilt and vergence, four harmonic drive actuators are used. These actuators were chosen according to their mechanical characteristics which, due to their harmonic drive gearing, provide high reduction ratios in a single stage, zero backlash and high precision.

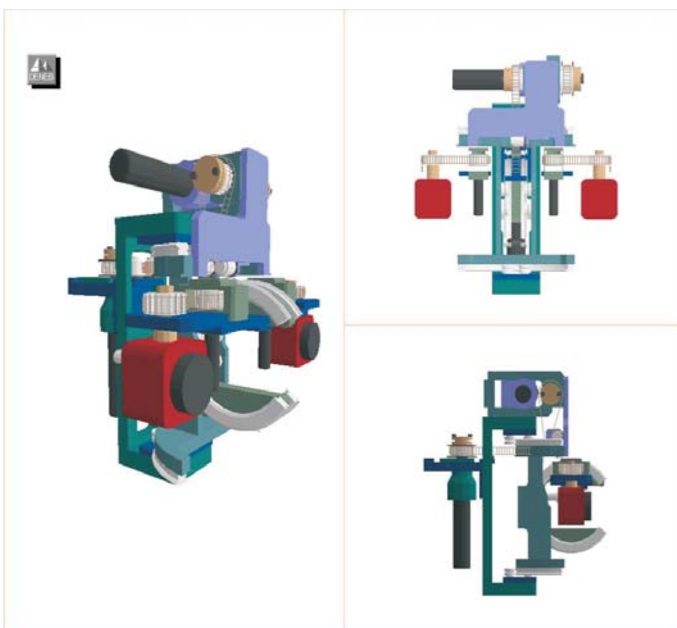


Figure 3. The Eurohead: Mechanical design

During the movement of the robot, control algorithms are used to stabilise the head and to keep the gaze of the cameras fixated on the viewpoint recommended by the CAD system. The stabilisation is performed using angular accelerometers that react to the movements of the robot. In this way it is possible to calculate the angular velocity by integrating the acceleration in a 40 ms cycle. This velocity is then multiplied by a predefined factor and directly introduced in the head motors which produce a compensation movement.

The gaze of the head is maintained in the direction of the viewpoint in a 240 ms cycle using the robot pose feedback. During this operation the inertial stabilisation is deactivated in order not to confuse the inertial sensors. Stabilisation and gaze orientation both have the purpose of guaranteeing a good image quality to increase the robustness of the tracking algorithm described in the next section.

2.3 **Vision**

The basic goal of the vision component is to extract the positions of geometrical features from the camera image. To do this, information about what features should be identified is needed together with feature models specifying the expected shape and pose of these features. Furthermore, input as the image itself, the current camera viewpoint, and an estimated robot pose, is given to enable the projection of the features into the image. This projection provides a first guess for feature extraction.

To fulfil this vision task, two different approaches are applied: a monocular and a binocular approach.

2.3.1. The monocular approach

The first vision approach is a monocular search and tracking strategy which is based on a local windowing technique. Local techniques have the advantage of saving computing time and therefore to keep the search and tracking cycle-time low [5]. In the RobVision system it is kept lower than the frame rate of the camera (40 ms).

Robustness is achieved by including image and model information - so-called cues - into the vision process. The RobVision system is presently able to track edge features like lines, junctions, ellipses and arcs.

For the first search of features the projection of the model into the image indicates an approximate position of the features. Since the model contains the model of the features and their topological relations to each other, checking of the actual geometric relations can assure that the right feature is found. Furthermore, other model cues like colour information about the object can be used to eliminate wrong feature candidates. Once a feature is found for the first time, this feature is then tracked in the next cycle. Additional information about the features found in the image (image cues like intensity) can be stored to facilitate the search for the same feature in the next tracking cycle. An example of a tracking sequence is given in Figure 4.

2.3.1. The binocular approach

The second method uses stereovision. The main idea of this approach is to measure the depths of edge junctions. A Hough technique is implemented to extract the lines on the image planes of the stereo pair. A weighted LMS (Least Mean Square) method is used to relate them to the features provided by the CAD system [6]. Then a closed loop method is followed, where the three degrees of freedom of the head is moved such that the junction is put at the principal point of the image in both images. When this is the case, the two cameras are verging on the certain junction and the direct kinematics of the head are applied in order to determine the 3D position of the junction relative to the head.

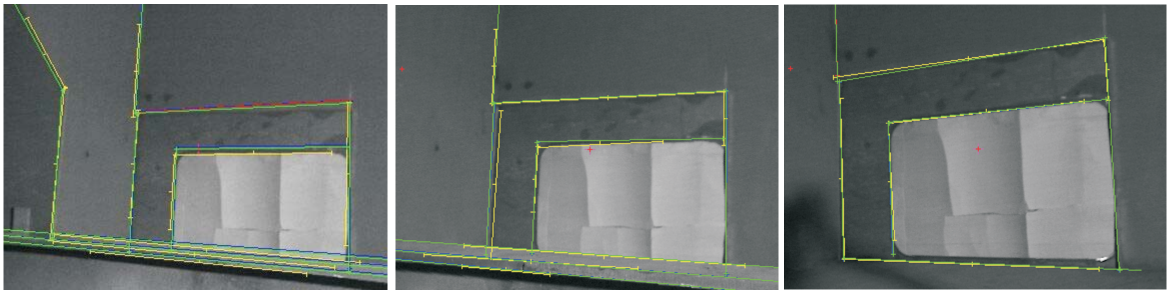


Figure 4. A tracking sequence along the robot trajectory: Projection of the model and result of feature finding

2.4 Pose Calculation

The pose calculation function of the RobVision system computes the position and orientation of the robot relative to the environment object, i.e. the ship. This is done by fitting the extracted image features to the model received from the CAD system. The algorithm employed is based on a method proposed by Wunsch [7] and was modified to provide accuracy estimation of the calculated pose as well as outlier detection [8]. The Wunsch-algorithm was chosen because it is fast and therefore well suited for object tracking.

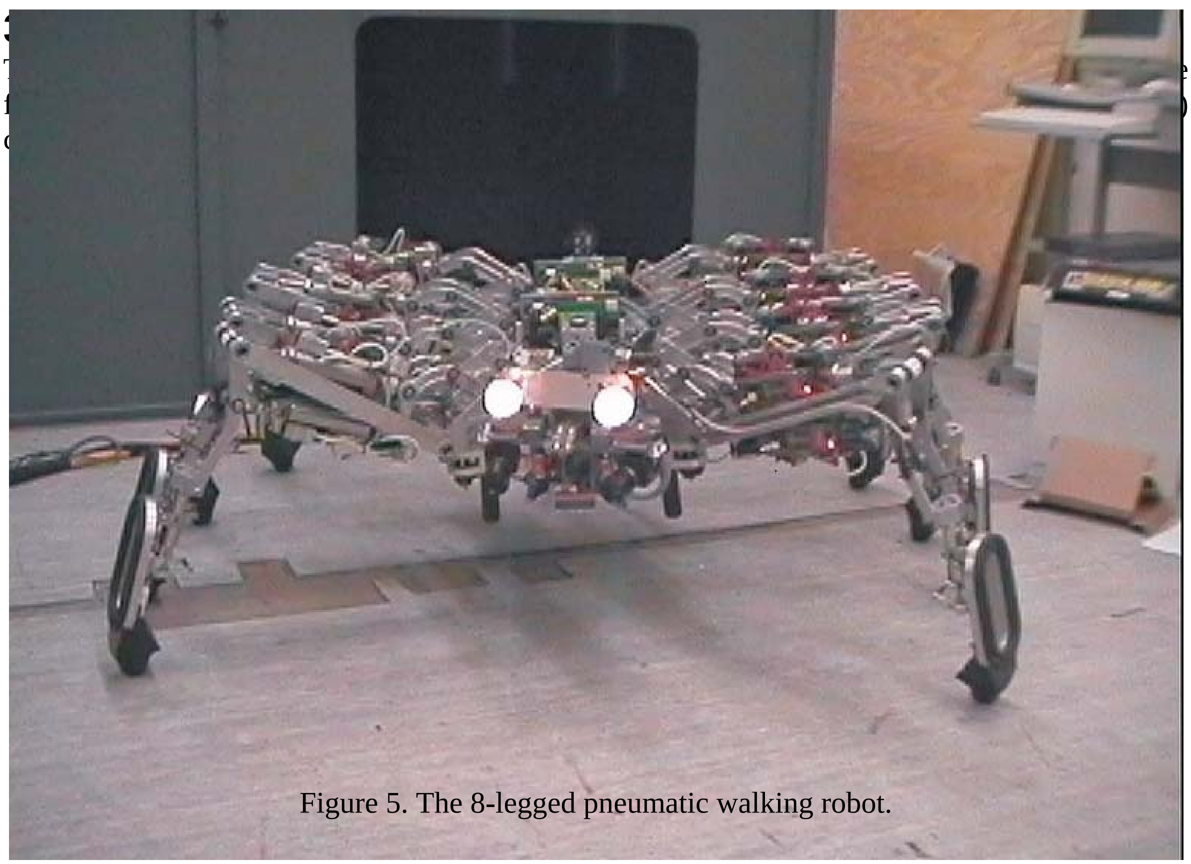


Figure 5. The 8-legged pneumatic walking robot.

Several trajectories were examined to test the robustness of the system and its accuracy. The average position accuracies reached are summarised in Table 1.

X [mm]	Y [mm]	Z [mm]	3D position [mm]
5.49	3.29	8.48	10.62

Table 1. Average position accuracy results

It is a big challenge to find a large number of "good" views to enable a successful task completion. Generally spoken, a good view is a view containing a large number of robust features. The importance is not only to generate good views along the path but also to have good alternative views for one robot pose if pose determination with the first view fails. Malfunction can happen since the mock-up is made of big metal plates welded together, which is a big defiance for the vision. Some welding causes irregular edge features, contrast is sometimes poor and additional features on the plate surfaces can affect pose calculation results negatively. That is why feature redundancy is a basic issue. Figure 6 shows an example of a view. The **CAD system** generates the features seen in the image and delivers it to the vision (Figure 6a). The **binocular part of the Vision** component is able to extract the 3D coordinates of the junctions (Figure 6b), and the **monocular part** the 2D position of edges and junctions in the image (Figure 6c). From all the extracted features a pose is calculated.

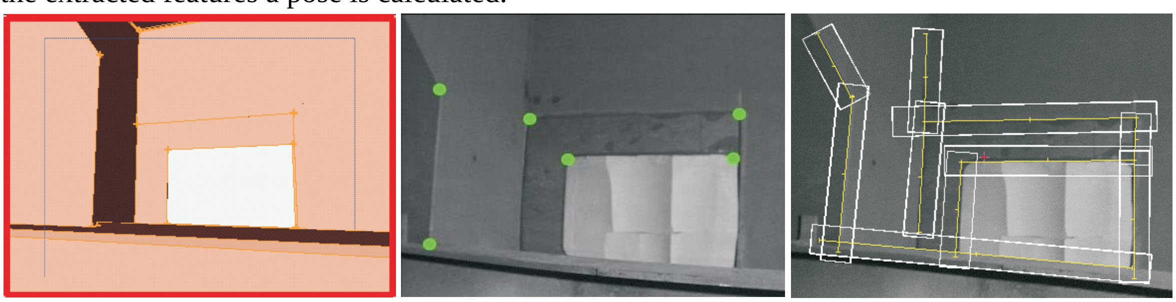
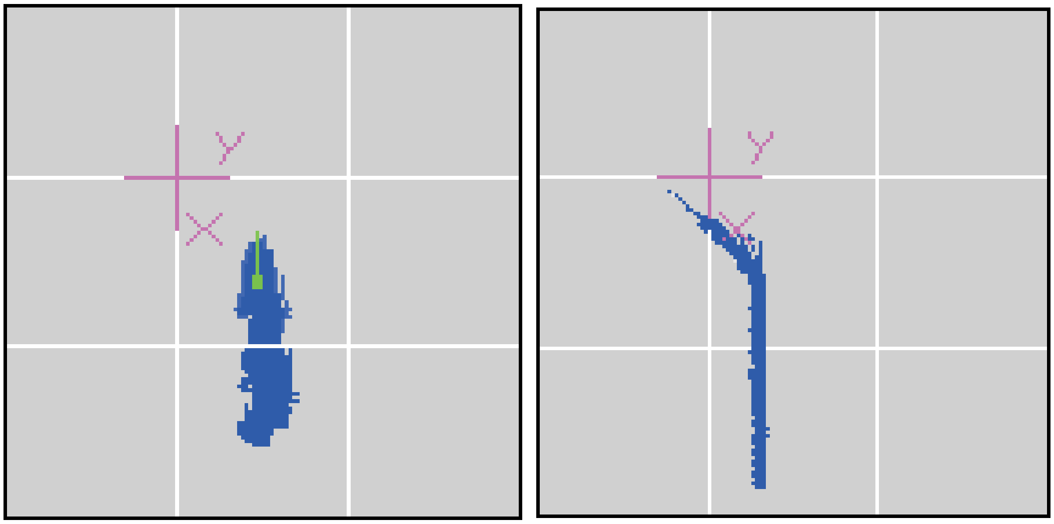


Fig. 6. (a) View with features generated by the CAD-system, (b) 3D junctions detected by the binocular part of the Vision component, (c) lines in their searching windows generated by the monocular part

Figure 7a gives an overview of the pose calculation recorded along a path. As can be seen, the pneumatic walking robot makes many jerks. Especially fast changes in the orientation of the robot cause big deviations of the feature positions in the image. Because of the fast but spatially restricted windowing technique such features are then lost. This fact reduces the reliability of the system. The implemented head stabilisation lowers the optical flow and increases the robustness of feature finding. In Figure 7b the motion of a trolley is compared to the motion of the robot. The smoother movement causes much less deviations of the poses along the trajectory since features can be tracked more easily. The resulting robot poses are therefore continuous.

The executed tests show the system benefits from the redundancies implemented. As long as enough features can be tracked, the system is able to re-find lost features. Also wrong detected image features are filtered out from the pose calculation. The stabilisation of the pose increases by the use of 3D features, and the continuous update of new image features from the CAD database makes the overall system robust and reliable.



(a) (b)

Fig. 7. Projection of (a) robot poses (b) trolley poses along a trajectory into the x/y plane

6 Summary

An automatic system which can determine the pose of a mobile robot in real-time has been described. It has been shown that applying CAD-model information in combination with visual sensing input enables the extraction of continuous robot poses. Two separate vision systems utilising different approaches guarantee enhancement through redundancies. Furthermore, synchronisation between all the subsystems is managed to achieve an pose update cycle time of 120 ms. The jerks of the pneumatic walking robot are a big challenge for the vision, but is partly compensated by including accelerometer data.

As a conclusion it can be stated that CAD based vision is a potential position-sensor in normal industrial environments. To reach the reliability demanded multiple views, redundancies in image processing strategies and feature validation techniques are necessary. Still, enhancements are necessary to achieve an overall reliable system behaviour.

Future work includes the integration of linear accelerometers and inclinometers to the existing gyros, in order to perform a more efficient and accurate image stabilisation. Moreover, since the robustness of the pose estimation method relies on the redundancy of the measured features, more 3D features extracted by stereoscopic vision, such as lines in 3D [6] and the normal of surfaces, are still under investigation. This should enhance the pose estimation in areas with poor features. Extensions of feature validation techniques to reach more robustness of the vision is another important issue. Together with the improvements of the vision capabilities, future work is also necessary to enhance error recovery and fault tolerance behaviours. Due to the high modularity of the system, the integration with extra sensors can be considered feasible and easy to implement.

References

1. “Current State Analysis”, deliverable from Esprit-project RobVision (deliverable 1.1).
2. M. Vincze: Robust Tracking of Ellipses at Frame Rate; Pattern Recognition 34(2), 487-498, 2001.

3. M. Vincze, M. Ayromlou, W. Kubinger: An Integrating Framework for Robust Real-Time 3D Object Tracking;
4. Panerai F., Metta G., Sandini G., Visuo-inertial Stabilization in Space-variant Binocular Systems, Robotics and Autonomous Systems, 30 (2000) 195-214.
5. Hager G. and Toyama K. (1998): The XVision-System: A Portable Substrate for Real-Time Vision Applications,
6. Gasteratos A., Sandini G., On the Accuracy of the Eurohead, LIRA - TR 2/00, July 2000.
7. Wunsch P., Modellbasierte 3-D Objektlageschätzong für visuell geregelte Greifvorgänge in der Robotik, PhD Thesis, Munich University of Technology, pp.37-47, 145-149, 1997.
8. Kraus K.: Photogrammetry Volume 2, Advanced Methods and Applications, Dümmler/Bonn, pp.205-215, 1997.