CAD Based Vision for Guiding a Mobile Robot*

C. Beltran[2], M. Ayromlou[1], W. Ponweiser[1], O. Madsen[3], M. Vincze[1], A. Gasteratos[2]
[1]Institute of Flexible Automation, Vienna University of Technology
{ma,wp,vm}@infa.tuwien.ac.at, robvision.infa.tuwien.ac.at
[2]Laboratory for Integrated Advanced Robotics, University of Genoa
{cbeltran,antonis}@lira.dist.unige.it
[3]Department of Production, Aalborg University, i9om@iprod.auc.dk

Abstract We introduce a system developed during the Esprit project RobVision (**ROB**ust **VI**sion for **S**ensing in **I**ndustrial **O**perations and **N**eeds), to guide a walking robot through a big steel construction like a ship section. The basic idea is to continuously generate a robot position and orientation (pose) signal by combining visual sensing information from the environment with predetermined CAD-information. These pose signals are used by the trajectory planner in the robot controller for further movements. Robust feature detection and real-time capability are key issues to reach an acceptable performance of the overall system. Our pose update cycle time achieved was 120ms but could be enhanced. Due to appearing jerks of the pneumatic walking robot, accelerometers are used for stabilization. Experiments show that our approach is feasible and reaches the positioning accuracies demanded.

1 Introduction

Quality assurance and intelligent products are key roads to success in global competition. Supervising and automatically measuring the quality of parts in the production of large structures can reduce work costs by up to 20% (expected by Odense Steel Shipyard Ltd. (OSS), the end user of this project). The inspection of these structures needs automated systems to position the tools necessary in a large environment (e.g. mobile robots). For both, quality measurement and positioning of inspection tools, the key is to generate the 3D pose of objects. To find this 3D pose a vision sensor system is developed. Since Industries are using CAD systems to design parts or working areas this model knowledge is applied to initialize the vision process. The CAD system provides features to the vision module that tries to find these features in the images. Two alternative image-processing techniques are implemented, a monocular and a binocular, to take advantage of the arising redundancy. The emphasis of using two different vision methodes and CAD model data is to enhance the robustness of the vision process to make the overall system reliable. To ensure correct feature detection, model and image cues (e.g. geometrical order of features or feature intensity) are integrated [6] [7]. 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. Portsmouth Technology Consultants Ltd. (Portech), one of the industrial project partner constructs the 8-legged walking robot (see Figure 1).

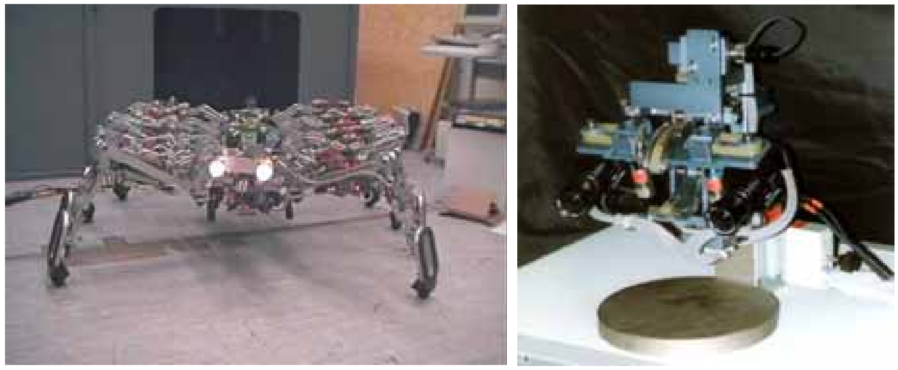
2 Sytem Overview

The entire system is shown in Figure 2. It consists of several modules colored in different greyscales. Each subsystem is provided by one of the project partners and fulfills all the functions drawn inside the according module. The next subsections give an overview of the modules and their respective functions.

2.1 C2V

C2V (Cad to Vision), developed by the Department of Production, Aalborg University, provides the CAD system. The main task is to generate geometrical features for the vision system. These features

* This work has been mainly supported by the RobVision project Esprit 28867 and is partly supported by the Austrian Science Foundation (FWF) under grant P13167-MAT and Project GRD1-1999-10693 FlexPaint.



(a) (b)

Figure 1: (a) the 8-legged pneumatic walking and climbing robot; (b) the Stereohead

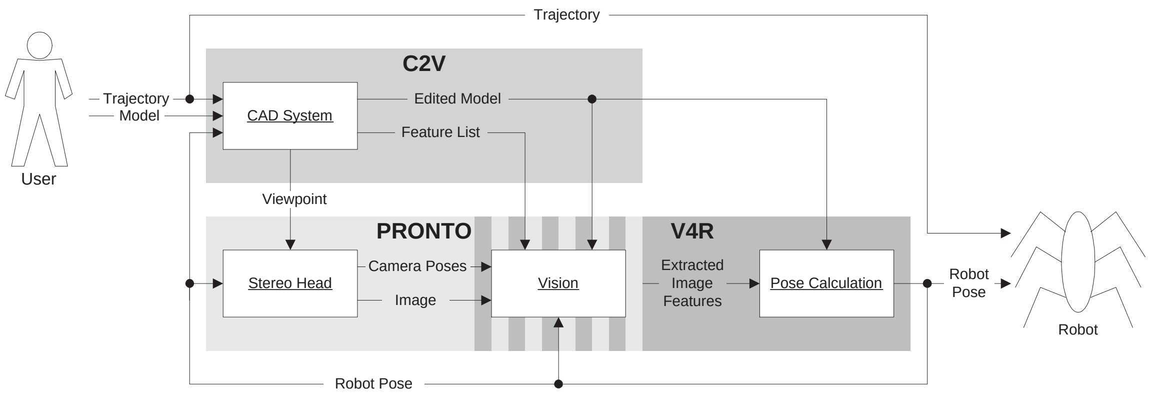


Figure 2: System Overview of the RobVision Subsystems

specify the shape and location of geometrical entities such as lines, junctions and regions that the cameras can expect to see while the robot is moving inside the structure.

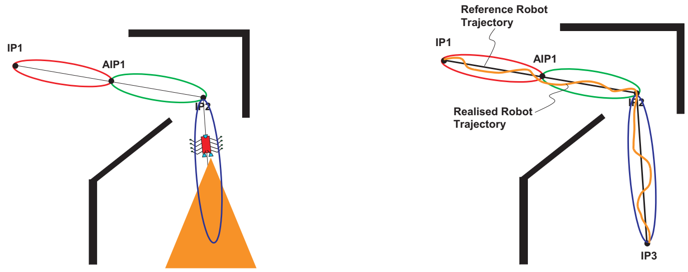
As it appears from Figure 2 the first input of the CAD system is a *CAD model* that is the geometric model of the structure inside which the robot has to move. As a second input, a *Reference Robot Trajectory* is delivered containing a specification of the task or trajectory that the robot has to perform. The user generates this trajectory off-line by using a trajectory planner developed by OSS. The last input are *Robot Poses* computed by the pose calculation component using the features found by the vision system. The main output from the CAD system are model and view data. The *Model data* contain a specification of relevant information from the underlying geometrical model, such as the feature topology. The *Camera Viewpoint* specifies a 3D point (x,y,z) in world co-ordinates towards which the stereo head ought to point. The *Feature List* contains a set of robust features that the vision systems can expect to find when looking at the specified viewpoint.

The CAD-model and the reference robot trajectory are computed prior to the operation of the Rob-Vision system. In this off-line phase of C2V (C2VoffLine), good viewpoints along the referece robot trajectory are determined and the features associated with these views are derived and generated. This approach divides the trajectory into areas where sets of features can be used to determinate the robot pose (Figure 3 a).

In the online phase of C2V (C2VonLine) the estimated robot pose generated by V4R is used to identify in which of the areas generated by the off-line system the robot is presently located (Figure 3 b). The according viewpoint and the features associated with this area are sent to the vision systems. This approach requires that the actual deviations of the robot trajectory are less than the areas generated by C2VoffLine. Simulations show that deviation of 500 mm or more, depending on the situation, are acceptable.

2.2 PRONTO

PRONTO developed by Laboratory for Integrated Advanced Robotics, University of Genoa, is the software/hardware module that is responsible for the stereo head task which in brief consists of *Head control, Head stabilization, Head calibration* and the *acquisition of 3D feature data*. For each of the



(a) (b)

Figure 3: The (a) Reference- and the (b) Realised- Robot Trajectory. The ellipses denote areas along the trajectory where a constant set of features is visible.

3D edge junctions PRONTO applies a stereovision algorithm to measure the feature depths. A Hough technique is implemented to extract the lines belonging to the junction. A weighted LMS (Least Mean Square) method is used to relate them to the features provided by the CAD system [3]. Then a closed loop method is followed, so that by moving simultaneously the three degrees of freedom of the head 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.

After this high precision depth measurements, the vision algorithm based on tracking (V4R) is used and PRONTO switches to its second operation mode, where it concentrates its computational efforts to stabilize the head and to keep the gaze of the cameras fixated on the viewpoint recommended by C2V. The robot pose generated by V4R is used every 240 ms to calculate the required movements of the head. The stabilization is performed using angular accelerometers that react to the movements of the robot. This consents PRONTO 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 producing a compensation movement. Both, stabilization and gaze orientation have the purpose to guarantee a good image quality to increase the robustness of the tracking algorithm described in the next section.

2.3 V4R

V4R (Vision for Robotics), developed by Institute of Flexible Automation, Vienna University of Technology consists of two functions. The first task of V4R is the monocular 2D feature search and tracking. The second unit of V4R is the pose calculation component to calculate the pose of the robot relative to a reference system.

The emphasis of the vision method is basically to provide robust features in real-time, i.e., keeping the processing time lower than the frame rate of the camera (40ms). To handle the real-time constraint a windowing method is applied to limit the processing time [4]. Robustness can be achieved by including image and model information - so-called cues - into the vision process. The method used is a combination of cue integration strategies (with cues like intensity values, color, texture) and the RANSAC-method for Edge finding [1] [7]. V4R is presently able to track edge features like lines, junctions, ellipses and arcs. In near future regions will be included.

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 like, e.g., connectedness of some features can assure that the right feature is found. Furthermore some other model cues like color information of 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 of 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 for a tracking sequence is given in Figure 4.

The second unit of V4R is the pose calculation which computes the position and orientation of the robot related to the environment object, i.e. the ship. This is done by fitting the 2D and 3D image features extracted to the according model received from the CAD system. The algorithm employed

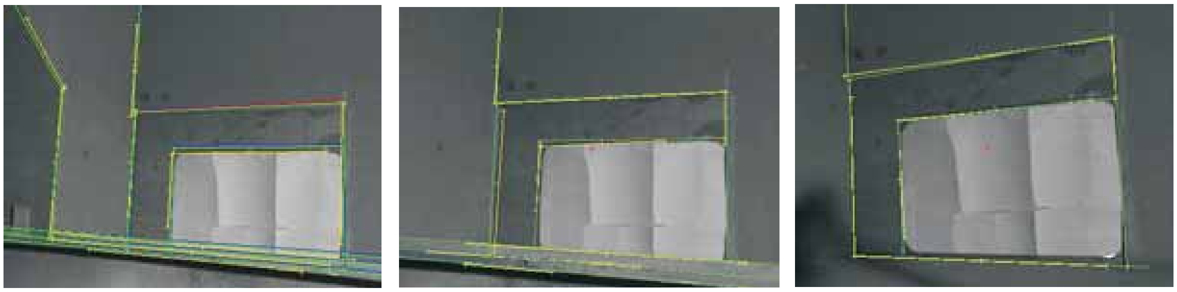
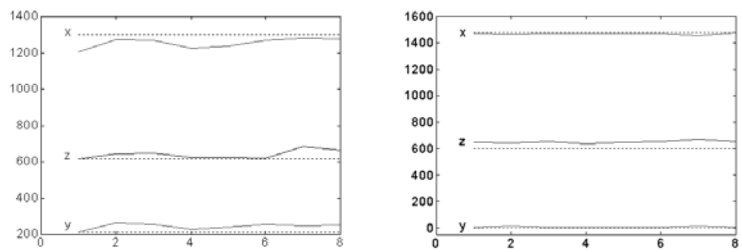


Figure 4: A tracking sequence along the robot trajectory.

is based on a method proposed by Wunsch [8] and was modified for providing accuracy estimation of the calculated pose as well as outlier detection [5].

3 Results

This section gives an overview of the tests for evaluating the whole system. A mock-up was built with structures that can be found in a big ship environment. A big challenge is to find a large number of "good" views to enable a successful task completion, that is, views containing a large number of robust features. The importance hereby is not only to generate good views along the path but also good alternative views for one robot pose to be on the secure side 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's why feature redundancy is a basic issue. The CAD system generates the features seen in the image to deliver it to the vision. The two vision methods are able to extract the 3D coordinates of the junctions [2] and the 2D positions of edges and junctions in the image. From all the features extracted a pose is calculated [8].



(a) (b)

Figure 5: Robot pose measurements in a fix pose (a) diagonal, (b) parallel

Two kind of tests were performed. The first group had the goal to *mesure the accuracy of the system at different fix poses* within the mock-up. With a ruler a reference position was measured. Figure 5 shows the results of different trials to calculate the position in two poses. The first one (Figure 5 a) corresponds to the robot standing diagonal in the mock-up and the second one (Figure 5 b) corresponds to the robot standing parallel. Table 1 and Table 2 summarise the measurements and give the standard deviations and the mean deviation from the reference measurements. The mean values are relatively high due to consistent off-set, which might be caused by the difficulty to access the origin of the measurement coordiante system. A further uncertainty is the actual geometry of the mock-up. The bending of the plates causes a few centimetres deviation over the model. As can be seen, the three dimensional standard deviation is in one case 35.72 mm and in the other case 4.64 mm, measurements that fulfil the specification for a ship building aplication.

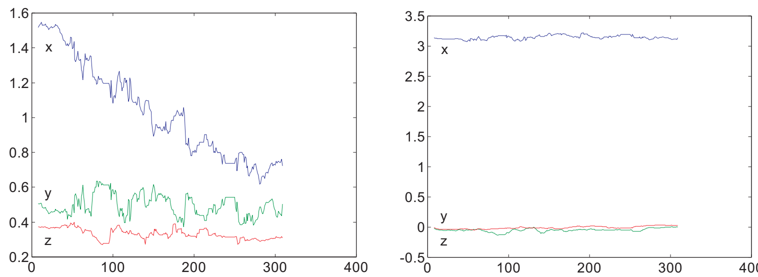
	X[mm]	Y[mm]	Z[mm]	roll[deg]	pitch[deg]	yaw[deg]	3D pos[mm]
std	28.17	16.91	23.56	0.35	0	0	35.72
mean	45.87	33.87	25	2.13	0	0	71.91

Table 1: The standard deviation (std) of the measurements and the mean distance from the reference.

	X[mm]	Y[mm]	Z[mm]	roll[deg]	pitch[deg]	yaw[deg]	3D pos[mm]
std	5.49	3.29	8.48	0	0.92	0.35	4.64
mean	4.88	8.62	51.25	0	0.62	0.13	52.49

Table 2: The standard deviation (std) of the measurements and the mean distance from the reference.

The second group of tests had the goal to test the *performance and reliability* of the pose calculation with the robot walking along a trajectory. Figure 6 gives an overview of the pose calculation recorded along a path in the x-direction. As can be seen the pneumatic walking robot produces many jerks, in particular for the translational degrees of freedom. It can be observed that the body of the robot sometimes moves up 15 cm when releasing a leg. These 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, used in one of the vision methods, such features are then lost. This fact reduces the reliability of the system. The camera head stabilization implemented lower the optical flow and increases the robustness of feature finding.



(a) (b)

Figure 6: Robot pose output along a trajectory, (a) position, (b) orientation

The tests executed 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 stabilization 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.

References

[1] Fischler M.A., Bolles R.C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography; Communications of the ACM Vol.24(6), pp.381-395 (1981).

[2] Gasteratos A., Martinotti R., Metta G., Sandini G., Precise 3D Measurements with a High Resolution Stereo Head. IWISPA 2000, Pula, Croatia, 2000, 171-176.

[3] Gasteratos A. , Sandini G., On the Accuracy of the Eurohead, LIRA - TR 2/00, July 2000.

[4] Hager G. and Toyama K. (1998): The XVision-System: A Portable Substrate for Real-Time Vision Applications, Computer Vision and Image Understanding, 69 (1), 23-37.

[5] Kraus K.: Photogrammetry Volume 2, Advanced Methods and Applications, Dümmler/Bonn, pp.205-215, 1997.

[6] M. Vincze: Robust Tracking of Ellipses at Frame Rate; Pattern Recognition 34(2), 487-498, 2001.

[7] M. Vincze, M. Ayromlou, W. Kubinger: An Integrating Framework for Robust Real-Time 3D Object Tracking; Int. Conf. on Vision Systems, Gran Canaria, S. 135-150, January 13-15, 1999.

[8] Wunsch P., Modellbasierte 3-D Objektlageschätzung für visuell geregelte Greifvorgänge in der Robotik, PhD Thesis, Munich University of Technology, pp.37-47, 145-149, 1997.