Vision-based Robotics Control Experiment on ETS VII

Maarten Vergauwen, Reinhard Koch, Tinne Tuytelaars and Luc Van Gool
ESAT–PSI, K.U.Leuven,
Kard. Mercierlaan 94, B-3001 Heverlee, Belgium
Phone: +32-16-32.10.64, Fax: +32-16-32.17.23
vergauwe | koch | tuytelaa | vangool@esat.kuleuven.ac.be

Abstract

This contribution describes the vision-based robotic control (VBRC) experiments executed on the Japanese research satellite ETS-VII. The VBRC experiments were designed to enhance image quality, refine calibration of different system components, facilitate robot-operation by automatically refining the robot-pose and provide data for robot-calibration.

1 Introduction

The vision-based robotic control (VBRC) experiments were executed on the Japanese research satellite ETS-VII in conjunction with the Visual Interactive Autonomy Bi-Lateral Experiments (VIABLE) between ESA and NASDA.

The VIABLE[1] project is the first collaboration between ESA and NASDA with the aim to test the Interactive Autonomy (IA) concept for space robotics and to investigate advanced vision-based techniques for robot-control and calibration.

The ETS-VII satellite is equipped with a 6-DOF robot manipulator and two sets of cameras. The VIABLE experiments had access to a taskboard that allows several tasks to be executed by the manipulator. The taskboard contains a set of 3-point calibration markers with known 3D positions in the taskboard reference frame.

The ETS-VII onboard vision system consists of two sets of cameras. The arm monitor camera (AMC) is mounted on the first joint and the arm hand camera (AHC) is mounted on the end effector of the robot arm. Each set contains two cameras, one primary and one redundant unit. Both can be utilized as stereo head with 60 mm baseline. Each camera records a grey level image with 668x480 pixel resolution with fixed focal length. The images are compressed with JPEG by a factor of 8.6 to yield a frame rate of 4 images per second on the video downlink. Two NTSC video channels allow access of two camera images simultaneously.

[1]The VIABLE consortium consisted of the Belgian companies TRASYS-SPACE and SAS and the K.U.Leuven departments PMA and ESAT-PSI.

2 VBRC Experiments

The VBRC experiments were designed to

- enhance the image quality to allow better visual control,
- refine calibration of the system components (intrinsic camera parameters, eye-hand calibration) based on the calibration markers,
- perform on-line pose estimation procedures and guide the robot by automatically refining the robot pose,
- aid the operator during the experiments with visual clues using augmented reality techniques,
- provide material for post-mission robot calibration and testing of advanced methods for uncalibrated vision experiments.

3 Viable station setup

An important part of the VBRC experiments is the capability of the Vision Tools to allow operator intervention during execution of a vision

Proc. Fifth International Symposium on Artificial Intelligence, Robotics and Automation in Space, 1-3 June 1999 (ESA SP-440)

task. Image processing and computer vision is a process with possibly many sources for errors that can not all be modeled beforehand. Therefore a user-friendly interface was developed to assist the VBRC tasks. The interface allowed the operator to interact with the vision system to guide and help the automatic processing. While the human operator is very good at interpreting the scene and recognizing qualitative information, the vision system is good at precise quantitative measurements when given the appropriate input data.

To simulate and verify the VIABLE experiments a photo-realistic 3D model of the taskboard and the robot was constructed. This model served as reference for the IA path planning (in ROBCAD) and the VBRC visual simulator (in OpenInventor[1]). The model allowed the realistic visual simulation of all aspects of the experiments.

Verification of this simulation was performed with a mockup taskboard of scale 1:1. It contained all visually significant parts and served as a realistic testbed for the VBRC experiments.

4 Enhancing the image quality

A first set of experiments evaluated the impact of the imaging conditions in space (degradation of the images due to noise, image compression, direct sunlight, etc.) and derived parameters for image preprocessing. Analysis of the images that were taken for this purpose yielded a set of parameters for image-enhancement filters. Evaluation led to the following filter sequence:

1. **a median filter.** This non-linear filter effectively removes spikes and noise in the image but preserves the edges. It was chosen for its capacity to remove the ringing that typically occurs around the edges of an image when JPEG compression is used. Because the JPEG-ringing was quite severe, a window-size of 5 was used for most images.

2. **a binomial filter.** This low-pass filter smoothes the image to remove noise. It has the advantage over standard mean filtering that its frequency response has no ripples.

3. **a sharpening filter.** This unsharp-masking filter cancels the smoothing of the edges caused by the previous filter.

4. **radial distortion compensation.** This filter undoes the quite severe radial distortion of the images.

5. **aspect ratio compensation.** This procedure restores the original aspect ratio of the image which was changed due to the conversion to NTSC.

These preprocessing filters were applied to all incoming images before further processing.

5 Calibration

Online calibration is one of the crucial needs in the VIABLE project because no a priory calibration of the intrinsic camera parameters, the eye-hand, or robot pose is available. Only approximate calibration parameters could be obtained from the specification documents and from a limited set of images taken while the system was still on ground. We therefore designed a set of calibration experiments that verified and refined the approximate calibration from images during the flight segment. These experiments are explained in paragraph 5.2 and 5.3 but first a procedure for invariant-based feature extraction is discussed. This procedure is used in other experiments (like pose estimation or robot calibration) as well.

5.1 Invariant-based feature extraction

The online calibration relies on the 3-point markers and their given 3D position on the taskboard. One of the novel strategies in this task is the reliable extraction of image features and finding the corresponding 3D features. Correspondences are found robustly and completely automatically by exploiting viewpoint invariant relations. Two strategies were used.

1. If enough markers are visible, marker points are extracted as ellipse centers. Collinear points are found by computing the cross-ratio of all sets of 4 points. Lines are formed exploiting the viewpoint-invariant properties of the cross-ratio of 4 collinear points (see[2]). Figure 1 shows the marker points and lines found back by this approach.

2. If the camera is closer to the taskboard and only one 3-point marker is visible, a different approach is used. Because the ellipses

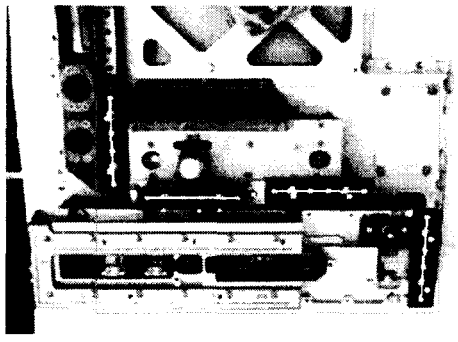

Figure 1: Marker points (found as ellipse centers) are grouped into lines using viewpoint invariant relations. The lines for each marker block are superimposed (in white) over the image for visual confirmation. Correspondences between the lines and marker blocks are also computed which yields 2D-3D relations.

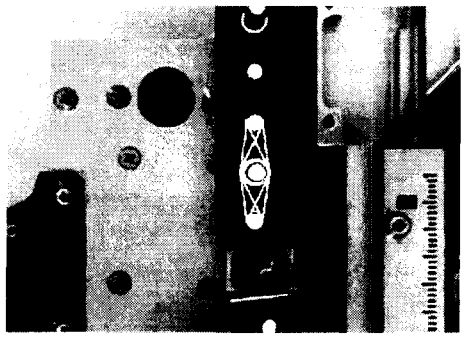

Figure 2: Common tangent lines of ellipses yield tangent points that are invariant under projective transformations. The ellipses and their common tangent lines are superimposed (in white) over the image for visual confirmation. This yields 2D-3D relations.

can be extracted more reliably in this case, we can use them (and not only their center points) to find enough 2D-3D correspondences. The fact is exploited that tangent points of two ellipses with a common tangent are invariant under projective transformations. In figure 2 the common tangent point and lines are superimposed over one of the 3-point markers.

5.2 Camera intrinsic calibration

Calibrating the intrinsic parameters of the camera is an important task in every application where measurements in the image are used to compute 3D spatial information like camera poses or 3D reconstructions.

Based on two images of a calibration pattern that were taken by the cameras before the satellite was launched the intrinsic parameters of the cameras were computed.

During the flight segment images of the 3-point markers were taken by the AHC. These markers served as a calibration pattern. The result of the processing of these images was consistent with the precomputed values of both intrinsic parameters and radial distortion.

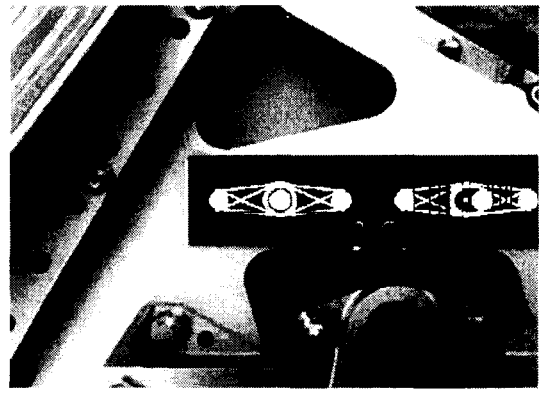

Figure 3: Setup for the eye-hand calibration experiment. The robot is touching the GPF and the AHC is above a 3-point marker. The common tangent points are found and based on these 2D-3D relations the camera pose is computed. This yields the eye-hand calibration of the camera.

5.3 Eye-hand calibration

For robot guidance from images the relative transformation between the cameras and the robot tip frame – the *eye-hand calibration* – has to be known. A procedure was developed especially targeted towards the ETS-VII robot. When the robot executes the procedure to grasp the grapple-fixture (GPF), it comes into contact with the taskboard in a predefined position and orientation. In this specific pose, the cameras are approximately aligned with 3-point markers. These markers are exploited to compute the camera poses with the second technique explained in paragraph 5.1. Based on these computed camera poses and the fixed robot pose, the eye-hand calibration can be calculated. Figure 3 shows the setup of this calibration experiment.

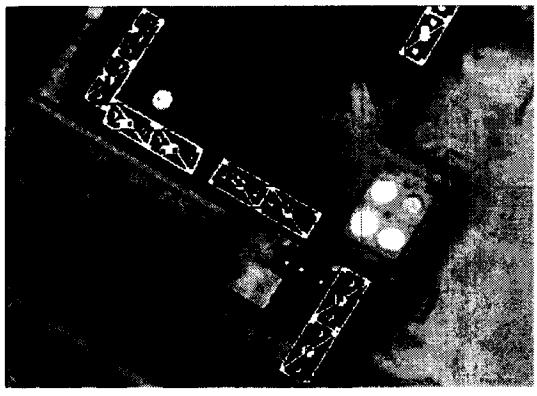

Figure 4: From automatically found 2D-3D relations the camera pose is computed. Parts of the model are superimposed over the real image and give a very good and intuitive verification of the calibration accuracy.

6 Pose estimation and on-line robot guidance

Several experiments concerning pose estimation and on-line robot guidance were performed during the flight segments.

6.1 Calculating pose from known markers

A first experiment consisted of calculating the robot pose from the known 3-point markers. The robot moved to a position where different markers were visible. Using the invariant relations described in paragraph 5.1, 2D-3D relations were found. These relations were the input for a robust camera pose estimation algorithm. An immediate verification of the current calibration status and the accuracy of the computed position could be supplied to the operator by superimposing parts of the given CAD-model with the actual images, using the calculated position. An example of this superimposition can be seen in figure 4.

A second step in this experiment consisted in moving the camera to a position much closer to one of the 3-point markers. The robot was intentionally positioned in a pose not perfectly above the marker. The second invariant method of paragraph 5.1 was used to calculate the camera- (and using the eye-hand transformation also the robot-) pose. Parts of the model were reprojected into the actual image to verify the calculation (figure 5). Using the computed pose, a relative translation- and orientation-change was computed to position the robot perfectly above the 3-point marker (figure 6).

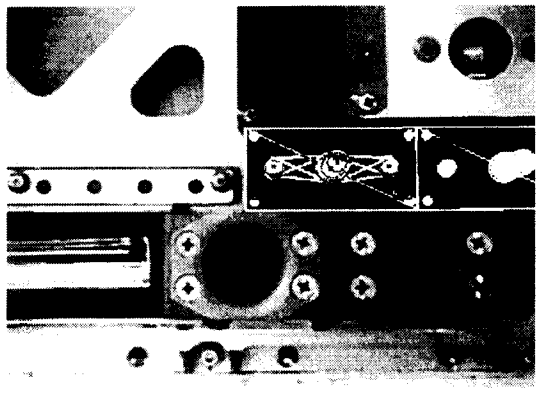

Figure 5: From automatically found 2D-3D relations the camera pose is computed. Verification of the result is possible by reprojection of parts of the CAD-model in the image.

6.2 Insertion of GPF into a hole

The ETS-VII robot has the possibility to attach a grapple-fixture (GPF) to its end effector and insert it into different holes and a slider on the taskboard. Usually, positioning of the robot is done manually by the operator who uses the artificial markers as a visual clue. During the VI-ABLE experiments we showed that positioning could be done automatically using the image of the hole or slider only. This is especially important for the case of the slider because its exact position is unknown due to possible previous motions. Using an ellipse-fitting algorithm the hole or slider was extracted and the center point

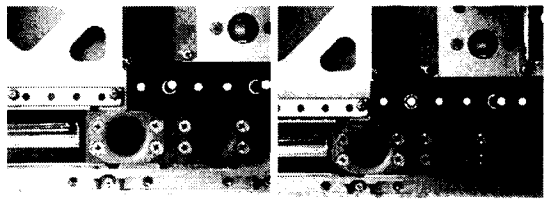

Figure 6: After a relative motion from the current position (left image), automatically computed by the vision-tools, the AHC ends up perfectly above the 3-point marker (right image). This is verified visually by the fact that the central marker tip is centered perfectly with the outer marker ring.

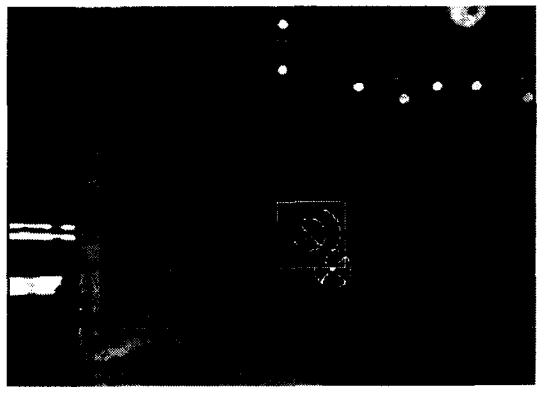

Figure 7: The vision system computes the current impact point of the GPF. The center point of the slider is extracted automatically and the relative movement is computed to position the GPF above the slider. The predicted impact point is shown to fit into the hole.

was found. This allowed the algorithm to compute a relative update of the current pose to position the GPF perfectly above the hole or slider. The image was augmented with the current impact point of the GPF (the point where the GPF would hit the taskboard if it were lowered from its current position) and the estimated impact point after relative motion. During operations the robot was deliberately mispositioned above both hole and slider. The algorithm managed to automatically update the pose to allow insertion. Figure 7 shows both current (misplaced) and estimated impact point.

7 Taskboard calibration and reconstruction

7.1 Calibration of 3-point markers

The 3-point markers on the taskboard are important for vision-based algorithms. The calculation of the camera pose from 2D-3D relations, found by the algorithm, needs the exact 3D coordinates of these markers. Experiments were designed which could retrieve this information.

Because a good estimate of the 3D coordinates of the markers was supplied to us by NASDA, a quick and easy check on the consistency of this data could be performed. We moved the robot over the taskboard to different positions for which different markers were visible in the images. We computed the camera pose based on the markers and reprojected the given 3-point markers in the original image. The estimated mean reprojection error was below a pixel which confirmed the consistency of the marker positions.

The coordinates of the 3-point markers can also be explicitly retrieved from images. This is what was done in another experiment. Three different images, taken from three different poses, showed the same 3-point markers. Based on the given pose of the robot and the eye-hand calibration the camera poses were computed. Based on the identification of the markers given by the invariants, multiple-view matches were found. The markers could then be reconstructed in 3D by triangulation. The resulting data was consistent with the given 3D information (up to the accuracy of the reconstruction of 2.23 mm in x, 1.45 mm in y and 0.84 mm in z).

7.2 Taskboard reconstruction

In an advanced experiment we investigated novel techniques for calibration based on image data alone, without the need to know precise 3D calibration markers. Based on a sequence of images taken from different view points, one can obtain a metric calibration (up to a constant scale factor) of the cameras and the scene (see [3]). This technique allows the handling of a priory unknown objects with little calibration information. For these experiments we recorded predefined image sequences during the flight segment and evaluated these techniques in the post processing phase.

Figure 8 shows some results of the reconstruction of the slider-area of the taskboard. The figure shows views of the reconstruction without any manual refinement. In a post processing step it is easy to obtain reconstructions of parts of the taskboard by human interaction in the image only, using the computed depth data.

8 Robot calibration

8.1 Robot calibration

Robot calibration is a procedure which aims at improvement of the robot accuracy by modifying the robot positioning software, rather than changing or altering the design of the robot or its control system[4]. The procedure that is followed to obtain this goal is

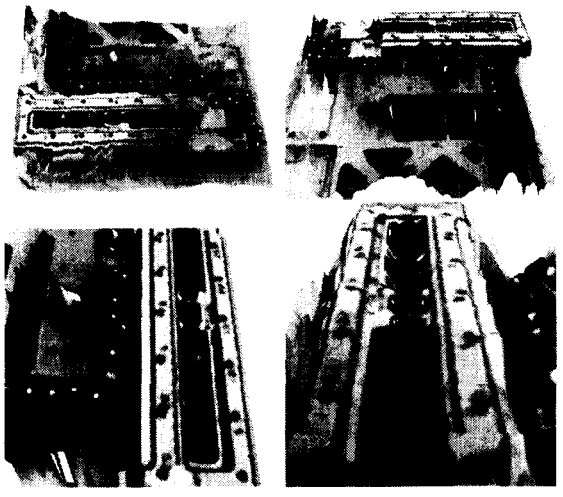

Figure 8: Different artificial views from the reconstruction of the slider-area of the taskboard.

- position the robot in different poses, trying to excite all possible modes,
- measure these poses with a measurement system,
- compute the difference between these measured poses and the pose computed from the joints telemetry by the forward kinematics model of the robot.

If all modes are excited sufficiently this allows to identify updates to be made to the current model.

Standard robot calibration procedures obtain pose measurements from external measuring systems. In the case of the ETS-VII robot no such system is available. Instead we computed the robot poses from the 3-point markers as explained in paragraph 6.1.

Since the taskboard on which all VBRC experiments were conducted is placed in one corner of the ETS-VII satellite, we could excite only a limited range of values in joint space. Ongoing evaluation will show if all joint offsets and link lengths can be identified or if computing a subset will yield better results.

8.2 Repeatability study

During the last temporal window data was gathered to perform a brief repeatability study of the robot. The arm was moved to a certain position where three 3-point markers were visible. Then the robot was commanded to move a certain distance in a direction and return to the same position. This procedure was executed for three orthogonal directions and each time an image was taken by the AHC. Using the second procedure from paragraph 5.1 the camera- (and using the eye-hand transformation the robot-) pose was computed.

The result of the computations is displayed in table 1. The first pose is the reference pose. Poses 2, 3 and 4 are poses obtained after a motion in x, y and z direction respectively. For each pose the computed position of the camera is shown.

This data clearly shows quite some difference in the x and y direction. It indicates a repeatability-error of a few mm in the robot movements but because of the very limited amount of motions and repetitions no statistical conclusions can be drawn from this data.

pos	x	y	z
pos 1	415.496	152.332	435.001
pos 2	419.604	152.325	435.071
pos 3	417.232	151.359	434.834
pos 4	417.920	154.069	435.442

Table 1: Result of the repeatability study

Acknowledgments

We acknowledge support from the Belgian IUAP4/24 'IMechS' project.

References

[1] Open Inventor Architecture Group, Open Inventor C++ Reference Manual, ISBN 0-201-62491-5, 1994.

[2] J. Mundy and A. Zisserman, "Machine Vision", In J.L. Mundy, A. Zisserman, and D. Forsyth (eds.), *Applications of Invariance in Computer Vision*, Lecture Notes in Computer Science, Vol. 825, Springer-Verlag, 1994.

[3] M. Pollefeys, R. Koch, M. Vergauwen en L. Van Gool, Metric 3D Surface Reconstruction from Uncalibrated Image Sequences, In *Proceedings SMILE Workshop (post-ECCV'98)*, LNCS 1506, pp.138-153, Springer-Verlag, 1998.

[4] K. Schröer, Handbook on Robot Performance Testing and Calibration (IRIS-project), ISBN 3-8167-5200-4, Fraunhofer IRB Verlag, 1998.