A Stereoscopic Eye-in-Hand Vision System for Remote Handling in ITER

Longchuan Niu[a], Liisa Aha[a], Jouni Mattila[a], Atanas Gotchev[b], Emilo Ruiz[c]

[a]*Laboratory of Automation and Hydraulic Engineering, Tampere University of Technology, Tampere, Finland*
[b]*Laboratory of Signal Processing, Tampere University of Technology, Tampere, Finland*
[c]*Fusion for Energy, Barcelona, Spain*

Abstract

The International Thermonuclear Experimental Reactor (ITER) maintenance is performed by means of remote handling (RH) systems and with aid of user interfaces such as haptic and joystick devices, virtual reality (VR) systems, and camera views. Many RH operations involving RH equipment, such as robotic manipulator arms, require millimeter accuracy, but camera views are often occluded or of poor quality, and might be unavailable during sensitive steps that require accurate, close-up views. Moreover, the VR system may not reflect the current scene accurately, as physical conditions may have changed under the harsh environment. The purpose of this research was to prototype and evaluate a novel software system, called 3D Node, that locates and detects the position and orientation of a piece of RH equipment or reactor element with respect to a stereo camera pair. The detection information is utilized to adjust the motion trajectories of a robotic manipulator arm. The 3D Node features stereo-camera calibration, target depth mapping, target position and orientation detection, and online target tracking. This paper reports on the 3D Node demonstration on the ITER Divertor RH use case and discusses the system applicability to other ITER RH systems.

Keywords:
Stereoscopic, Eye-in-Hand, Remote Handling, Computer Aided Teleoperation

1. Introduction

Performing accurate ITER RH maintenance operations inside dark and highly radioactive chambers, where human access is impossible, is extremely demanding. The RH operator can utilize a number of user interfaces for commanding and controlling the RH equipment [1], e.g. a robotic manipulator arm. Other auxiliary interfaces involve live images of the RH equipment and its environment, Computer-Aided Teleoperation (CAT) used in master-slave teleoperation [2], and virtual reality (VR) representing the movements of the RH equipment and its environment [3].

VR displays visual information based on the measured pose of the manipulators and pre-constructed virtual models. Due to the harsh environment, the VR representation may not exactly reflect the actual scene, as physical conditions may have changed, e.g. through material deformation due to extreme heat, or small drifts in the poses of the components to be manipulated. Thus, the pre-defined motion trajectories of the RH equipment have to be adjusted by other interfaces, e.g. a robot perception unit.

The purpose of the study herein is to prototype and demonstrate new means to assist RH operators to successfully perform ITER RH operations. A robot perception unit, namely 3D Node, was designed and developed to introduce new operator assisting features. The new features are based on detection of a target, i.e. a piece of RH equipment or reactor element, and recognition of its position and orientation in a relation to the environment using stereo camera images.

During ITER RH operations, a number of RH operations are identified in which 3D Node information could be helpful. This information could be valuable for updating VR models and implementing augmented reality and synthetic viewing functionalities. However, in this paper we consider 3D Node's usage merely in adjustment of the motion trajectories of RH equipment. A subset of operations related to the Divertor Cassette Locking System (CLS) operations is considered, and the use of 3D Node therein is demonstrated.

2. System Architecture

As stated, the RH operator utilizes multiple software systems and user interfaces during an RH operation. As seen in the Figure 1, the novel software system, 3D Node, requires an interface to a manipulator control system and a stereo-camera pair that is attached to, for example, the manipulator arm. The 3D Node receives images from the stereoscopic cameras through GigE Vision protocol and the pose of the manipulator robot's tool center point (TCP) from the manipulator control system. Additionally, it can receive operator input and provide visual feedback to the operator through its graphical user interface (GUI). 3D Node provides the target pose, i.e. position and orientation of the viewed target, to the manipulator control system.

The 3D Node is designed to comply with ITER remote handling control system (RHCS) requirements and is integrated at a later stage of its development into the RHCS as one of its components. Moreover, interfaces to the other systems, such as providing calibration information to VR, can be developed.

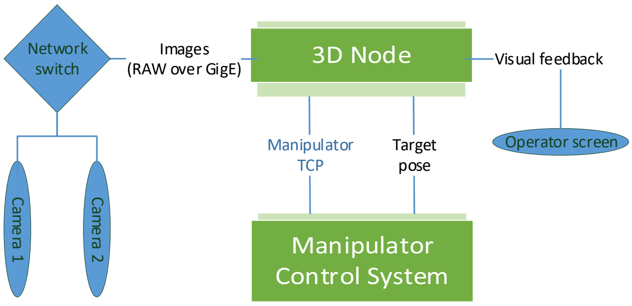


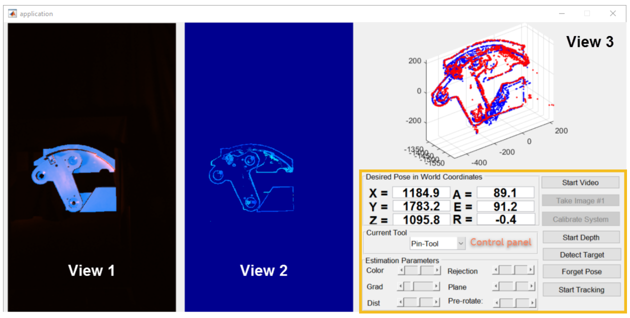
Figure 1: *Top-level architecture*

3. 3D Node Software

3.1. Operation Modes

3D Node features five modes: calibration mode, video mode, depth mode, detection mode, and tracking mode. The 3D Node GUI has four main parts: the operator control panel and three views. Items displayed in the views depend on the selected mode (two examples in Figure 2). The purpose and functionality of each mode is elaborated in the following.

(a) *Calibration mode*



(b) *Detection mode*

Figure 2: *Sample GUI views: Calibration and Detection modes*

Calibration mode is required to calibrate the stereo cameras and the relative position of the cameras with respect to the robot manipulator TCP, i.e. hand-eye calibration, prior to the actual RH operations. In ITER, this would be performed in the Hot Cell facility. In calibration mode, 3D Node captures stereo images of the scene while recording a current manipulator TCP. Images should contain a calibration pattern from diverse locations and angles. 3D Node performs the camera calibration with the aid of stereo images. The hand-eye calibration is also performed based on the stereo images and poses of the manipulator TCP.

Video mode is utilized for inspecting the camera views to confirm that the target object is not occluded by other objects, that lighting is sufficient, and that nothing prevents target detection during the RH operations. In video mode, 3D Node shows the images from both cameras.

Depth mode is used for checking the geometry of the scene or validating the correctness of stereo camera calibration. In depth mode, 3D Node visualizes a depth map of the scene that it has generated.

Detection mode is for detecting the target object and estimating its real pose. In detection mode, 3D Node determines the pose of the target object and aligns a rendered image of the target in the estimated position with the real camera view of the target. If the alignment is correct, the real image and rendered image should correspond to each other as seen in View 1 of Figure 2(b). In addition, 3D Node displays the desired pose of the manipulator TCP in the selected RH operation. The pose values are Cartesian positions in millimeters for X, Y and Z, and in degrees for orientation in Euler angles A, E and R. The pose is updated on RH operator command through the 3D Node control panel (Figure 2). These values can be utilized for adjusting the motion trajectory of a manipulator arm.

Tracking mode is utilized when the manipulator is moved around to inspect the environment. It differs from video mode, as in tracking mode 3D Node illustrates the rendered image of the target on the camera view.

3.2. Method

3.2.1. Depth from Stereo

Estimation of 3D scene geometry from parallel calibrated cameras is known as depth from stereo. In hazardous ITER environments, robust estimation of depth values is crucial. Instead of conventional rectification based methods, we employ the *plane-sweeping depth estimation* method, which uses calibrated camera parameters [4]. The method controls well the trade-off between precision and computational speed.

3.2.2. Advanced Sampling

As illustrated in [5], an important step for pose estimation is the fine alignment between the sensed target point cloud and the reference point cloud.. This is done by utilizing a state-of-the-art edge point iterative closest point (ICP) algorithm.

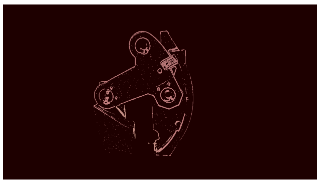
Conventional edge-point ICP samples its model point cloud only once. In order to improve the robustness and accuracy of the fine alignment, we used a left-to-right correspondence check and dynamic CAD model resampling as a mechanism to reduce outliers in the model point clouds [4].

An example of a target object is given in Figure 3(a). We render the image to find strong edges, then prepare a point cloud using the left-to-right correspondence check. The sensed point cloud after sampling is shown in Figure 3(b).

3.2.3. Pose Estimation

The 3D Node estimates the target pose based on the stereo camera images and camera poses in the manipulator base frame. The camera pose is calculated by rigid body transformation between the cameras and the manipulator TCP, which is known as

(a) given image



(b) depth after sampling

Figure 3: *Sampling of CAD model*

hand-eye calibration. The details are presented in [4], [5]. We adopt Tsai's method [6] for the hand-eye calibration.

4. Proof-of-Concept Demonstration

4.1. Demonstration Equipment and Setup

As indicated in Figure 4, in the demonstration setup we used the Comau Smart NM45-2.0 robot as the manipulator with two cameras attached to its wrist. The target object utilized in the demonstration was a test mock-up, which is a 1:1 replica of the Divertor Cassette Locking System (CLS).

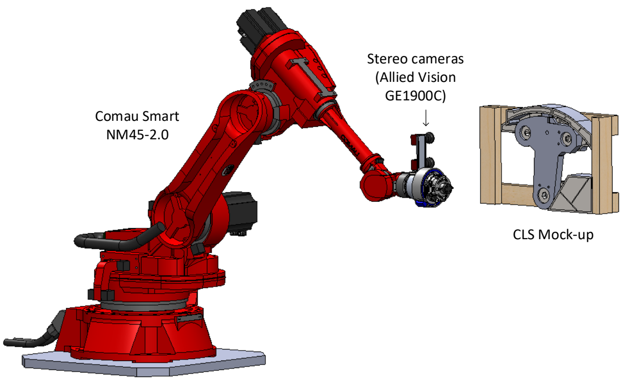


Figure 4: *Comau Smart NM45-2.0 with stereoscopic camera and CLS Mockup*

The stereo cameras are mounted on an adjustable mounting plate, allowing reconfiguration for particular environments. The accuracy of camera positioning is not an issue as long as the camera field of view is clear. The camera calibration process recovers the underlying stereo camera position every time the camera configuration changes or, for example, when a collision occurs.

The stereo cameras are arranged vertically. This is due to the dimensions between the tool exchanger and the Comau robot wrist. At ITER, the cameras could also be positioned horizontally depending on the manipulator. This is not an issue as the developed 3D Node system works similarly regardless of the camera arrangement.

The current setting is optimized for depth sensing at a range of 400–1500 mm. This is mostly defined by adjusting the camera lenses to deliver optimal sharpness at these distances. For the distance between the cameras, a stereoscopic baseline of 100 mm was chosen as a practical compromise.

A pair of industrial machine vision digital cameras (Allied Vision GE1900C) was used for the demonstration. The native resolution of the camera is 1920 x 1080, and the sensor size is 1" with an effective pixel size of 7.4 μm. However, such cameras are not usable inside the actual ITER environment due to high levels of radiation; radiation tolerant cameras typically have a lower spatial resolution. We use the "Pixel Binning" feature of the cameras in order to decimate the original resolution as well to automatically convert images to a grayscale format. The resulting effective resolution of 960 x 540 is close to that of the standard radiation tolerant camera.

The Comau Smart NM45-2.0 robot payload capacity is 45kg. In the 3D Node system demonstration, it operates the Divertor RH equipment tool prototypes, i.e. pin tool and jack tool (Figure 5). Tool weights are 16 kg for the pin tool and 33.5 kg for the jack tool.

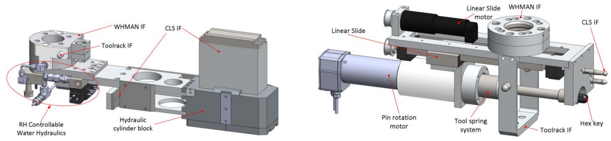
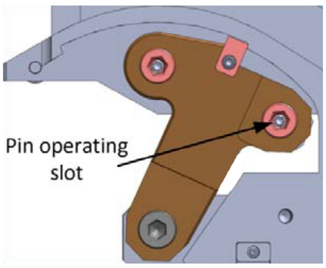


Figure 5: *Cassette Locking System tools: jack tool and pin tool*

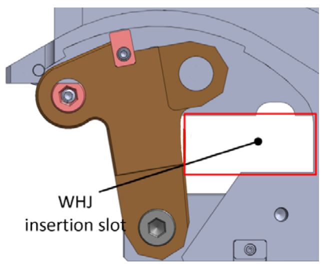
The Comau control system communicates with the 3D Node through an Ethernet connection at 1 Hz. The Comau control system sends the pose values of the manipulator TCP to the 3D Node. Received pose values from the 3D Node will be used to guide the manipulator arm in RH operations. At this moment, the target pose is only displayed within the 3D Node GUI and not sent directly to the Comau control system. Later, the 3D Node communication interfaces will be implemented to comply with ITER RH network communication protocols.

4.2. Demonstration Cases

The pin tool and the jack tool are utilized in the CLS operations for unlocking and locking the cassette, and cassette compression, respectively. Figure 6 indicates the location for these operations. Inserting tools into their corresponding slots requires millimeter accuracy in order to guarantee that the operations are performed properly. For example, the horizontal clearance between the jack tool and the slot in the cassette knuckle shown in Figure 6(b) is approximately 7 mm. Therefore, we selected these two use cases to validate the functionality of the 3D Node and to give a proof-of-concept demonstration.



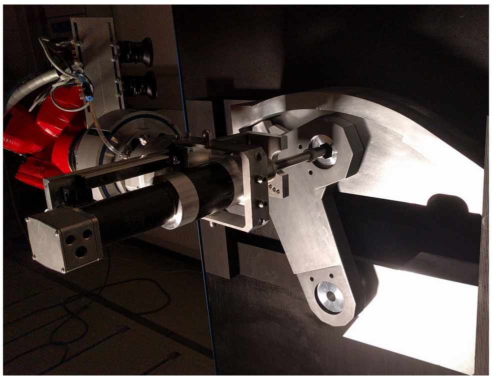
(a) Location of pin operating slot



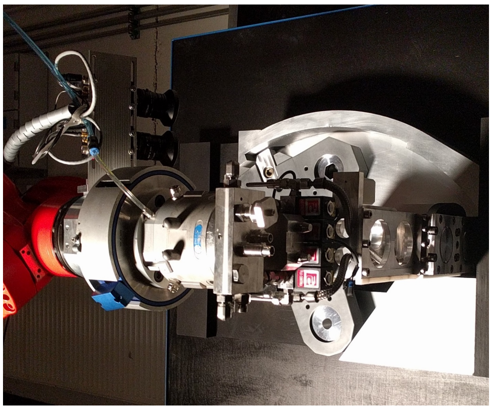
(b) Jack tool insertion slot

Figure 6: *Locking mechanism, tool operation location*

The purpose of the demonstration was to determine whether the 3D Node can help the RH operator to execute the RH operations in a physically unknown environment. One use case involves inserting the pin tool into the operating slot shown in



(a) Pin tool insertion



(b) Jack tool insertion

Figure 7: *Tool operations in the CLS mock-up scene*

Figure 6(a) and the other is inserting the jack into the insertion slot shown in Figure 6(b). Just before the demonstration, the target, i.e. the CLS mock-up, was randomly placed, which emulates unpredicted target movement during ITER maintenance operation. The operations were performed in a dark laboratory room with only a single adjustable light source pointed to the target. Prior to any operation, camera and hand-eye calibration were performed.

In both cases the operation sequence is the same when using the 3D Node. At first, the video mode can be utilized for inspecting the scene and the depth mode for validating the correct camera and hand-eye calibration. The detection mode is then utilized to find the actual pose of the target. According to the detected target pose values and calculated pose of the operated tool, the RH operator drives the Comau manipulator and the operated tool towards the calculated pose. As the tool tip reaches the desired pose, the operator can finalize the tool insertion in a peg-in-hole manner by simply driving along the Z-axis(depth) in the manipulator tool frame. The demonstration results are shown in Figure 7. The successful operation from the insertion of both tools validates that the pose values given by the 3D Node are accurate.

5. Discussion

In order to assess the accuracy of the pose estimation algorithm, we performed a series of experiments [4]. For the observation range between 600 and 1200 mm, the relative accuracy from the repeatability test, i.e. deviation from re-measurement of the same position, is approxmately 0.5-1 mm with respect to the position and 0.2-0.4 degrees with respect to the angle, providing that the target object has a planar surface.

Other sources of errors on the accuracy come from camera calibration, hand-eye calibration, and robot calibration. The stereoscopic camera calibration shows excellent stability, and the pixel reprojection error is about 0.15 pixels. Should there be higher resolution radiation tolerant cameras in the future, this would naturally improve the results of the camera calibration. The major proportion of hand-eye calibration error comes from the absolute accuracy of the robot, payload, and path of movement, i.e. possible backlash. Therefore, the selection and calibration of the manipulator are very important to ensure precise end-to-end movement.

6. Conclusions

3D Node is designed to fulfil generic ITER vision system requirements and can be easily integrated to any RHCS. The state-of-the-art pose estimation algorithm is developed to ensure good accuracy and robustness that can be achieved under dark and harsh conditions and with fairly low resolution cameras. The demanding test cases demonstrated its applicability in RH operations. The overall accuracy of the current system is highly dependent on the precision of the manipulator. It can be improved by robot calibration and fine tuning of the hand-eye calibration.

Acknowledgement

The work leading to this publication has been funded in part by Fusion for Energy and TEKES under Grant F4E-GRT-0689. This publication reflects the views only of the authors, and Fusion for Energy or TEKES cannot be held responsible for any use which may be made of the information contained herein.

References

[1] J. Tuominen, A. Muhammad, J. Mattila, L. Aha, H. Saarinen, M. Siuko, D. Hamilton, L. Semeraro, Command and control application framework for interoperable heterogeneous iter remote handling devices, Fusion Engineering and Design 86 (9-11) (2011) 2067–2070.

[2] M. Viinikainen, J. Tuominen, P. Alho, J. Mattila, Improving the performance of dtp2 bilateral teleoperation control system with haptic augmentation, Fusion Engineering and Design 89 (9-10) (2014) 2278–2282.

[3] S. Esque, J. Mattila, M. Siuko, M. Vilenius, J. Järvenpää, L. Semeraro, M. Irving, C. Damiani, The use of digital mock-ups on the development of the divertor test platform 2, Fusion Engineering and Design 84 (2-6) (2009) 752–756.

[4] L. Niu, S. Smirnov, J. Mattila, A. Gotchev, E. Ruiz, Robust pose estimation with a stereoscopic camera in harsh environments, Electronic Imaging 2018 (9) (2018) 1–6.

[5] L. Niu, O. Suominen, M. M. Aref, J. Mattila, E. Ruiz, S. Esque, Eye-in-hand manipulation for remote handling: Experimental setup, in: IOP Conference Series: Materials Science and Engineering, Vol. 320, IOP Publishing, 2018, p. 012007.

[6] R. Y. Tsai, R. K. Lenz, Real time versatile robotics hand/eye calibration using 3d machine vision, in: Robotics and Automation, 1988. Proceedings., 1988 IEEE International Conference on, IEEE, 1988, pp. 554–561.