Fusion Engineering and Design 88 (2013) 1969–1972

Contents lists available at ScienceDirect

Fusion Engineering and Design

journal homepage: www.elsevier.com/locate/fusengdes

Introducing artificial depth cues to improve task performance in ITER maintenance actions

C.J.M. Heemskerk[a,*], P.T. Eendebak[b], G.Y.R. Schropp[a], H.V. Hermes[c], B.S.Q. Elzendoorn[d], A.J. Magielsen[e]

[a] Heemskerk Innovative Technology, Sassenheim, The Netherlands
[b] TNO Netherlands Organisation for Applied Scientific Research, The Netherlands
[c] TU Eindhoven, The Netherlands
[d] FOM Institute DIFFER (Dutch Institute for Fundamental Energy Research), Association EURATOM-FOM, Partner in the Trilateral Euregio Cluster, P.O. Box 1207, 3430 BE Nieuwegein, The Netherlands[1]
[e] NRG, P.O. Box 25, 1755 ZG Petten, The Netherlands

HIGHLIGHTS

- Operators regard lack of 3D perception as primary factor hindering remote maintenance.
- Promising techniques to improve depth perception are depth gauges and stereo vision.
- First experiment shows artificial depth gauges need to be designed with care.
- Second experiment shows that stereo vision improves task performance significantly.

ARTICLE INFO

Article history:
Available online 16 March 2013

Keywords:
Maintenance
ITER
Remote handling
Task performance
Depth perception
Experimental procedure

ABSTRACT

Maintenance operations on ITER tokamak components will be largely performed by remote handling. In previous work it was shown that representative maintenance tasks could be performed significantly faster with direct visual feedback than with camera feedback. In post-test interviews, operators indicated that they regarded the lack of 3D perception as the primary factor hindering their performance.

This paper discusses various techniques to improve depth perception in teleoperation, including stereo vision, head tracking, virtual camera views and depth gauges. The most promising techniques were tested. Performance metrics included time-to-complete, path analysis and operator work-load.

In a first experiment, artificial depth gauges views were tested in a 1:1 scale hardware testbed with mechanical master-slave manipulators handled by experienced operators. Robust real-time image processing was achieved with marker-based objects. The simple depth gauge and graphical overlay did not significantly improve task performance. Operators commented on their view of the task being "obstructed" by the graphical overlay, and the depth gauge was judged not very informative.

In a second experiment, real time tracking was combined with VR display including stereo and head tracking. While stereo was found to improve the task performance significantly over the 1 camera (mono) condition, head tracking unexpectedly did not.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

During the operational lifetime of ITER, in-vessel components will be exposed to high neutron fluxes and become contaminated with toxic materials such as beryllium and tritium. Therefore, maintenance of exposed components has to be performed remotely. Due to the experimental nature of the ITER reactor, the exact condition of the components at the time of maintenance is hard to predict. As the maintenance tasks will be complex and largely non-repetitive, human interaction with and supervision of the maintenance operations is required. Master-slave manipulator systems are the preferred solution to perform this type of remote handling (RH) maintenance.

During the execution of ITER RH maintenance procedures, limited visual and haptic feedback is available. In previous work [1,2], we established that simple, ITER representative maintenance tasks could be performed significantly faster with direct visual feedback than with camera feedback. In post-test interviews, operators indicated that they regarded the lack of 3D perception as the primary factor hindering their performance.

* Corresponding author.
E-mail address: c.heemskerk@heemskerk-innovative.nl (C.J.M. Heemskerk).
[1] www.differ.nl.

0920-3796/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.fusengdes.2013.02.034

This paper first discusses various techniques that could be applied to improve depth perception in teleoperation, including stereo vision, head tracking, virtual camera views and depth gauges.

2. Artificial depth cues

In everyday life, humans use many different cues to perceive depth. An overview is presented in [3]. Monocular cues provide depth information when viewing a scene with one eye. Some of these cues only work in direct observation, while others also work when viewing the scene by video. An example of the first is accommodation: when we try to focus on an object far away, some muscles in the eye stretch the lens to make it thinner and change the focal length. The muscle action is processed in the brain and interpreted as distance/depth. Accommodation only works in direct observation and is only effective at distances less than 2 m. An example of the second is motion parallax: when an observer (e.g. a camera on a slave robot) moves, the apparent motion of objects against the background provides information on their relative distance.

A well-known binocular depth cue is stereopsis, which combines the visual images of two eyes, slightly different due to the distance between the eyes. There are many other visual cues that humans use in depth perception, such as occlusion, known size, shadows, texture, hazing and retinal disparity, but those are not further considered here.

A third category of depth cues considered here is artificially generated depth cues, intentionally designed to improve depth observation in remote control situations. Using video overlay techniques, artificial depth indicators can help an operator to identify and track critical task elements in a scene. Virtual reality techniques can generate synthetic camera viewpoints, and help an operator to "look around the corner" or create an artificial close-up of a tight situation. Combining real time tracking techniques with video overlays and VR, we can create Augmented Reality views, in which task data, real and virtual images are combined to improve situational awareness.

To evaluate the influence of artificial depth cues on depth perception and task performance, we test promising techniques on small scale ITER representative tasks in a VR set-up and in a 1:1 HW mock-up. The tasks are based on objects and tools from the ITER RH Code of Practice [6].

Task performance in RH maintenance depends on the complexity of the task, on the environmental conditions and on the skill and experience of the operator. We designed a simple benchmark product in which four ITER compatible pop-up bolts have to be addressed with a bolt runner (Fig. 1).

The key question to be answered in these experiments was whether the introduction of artificial depth cues helps the operator to improve the task performance.

3. Depth cue overlays on mono video

3.1. Setup

The first test was performed using the RH test facility at NRG (Nuclear Research Group, Petten). The facility has two standard mechanical through-the-wall master-slave arms, in a configuration similar to a classical hot cell laboratory setup. The manipulator arms provide direct, mechanical, one-to-one scale haptic feedback.

A single camera showed a frontal close-up view of the test setup on the workbench. A second camera was added at a roughly 45° viewing angle and used to generate reliable real time depth information for the artificial depth cue (Fig. 2). This second image was not shown to the operator.

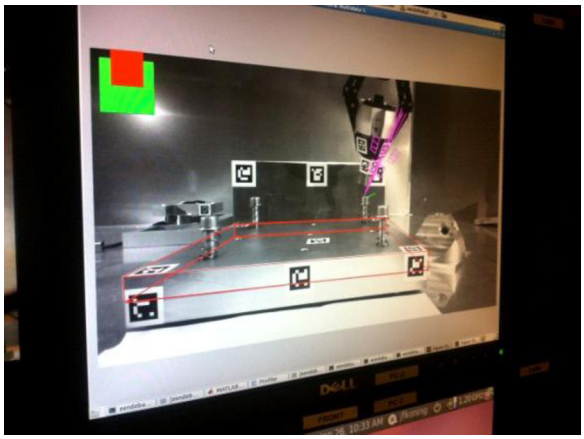

Fig. 1. Operator view of the 4-bolt benchmark, with artificial depth cue – the red bar on the top left and task cue – a magenta circle indicating (but also obscuring) the target location at the bolt head. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.2. Experimental conditions

Two visual conditions were tested: no depth cues (N) and depth cues (D). In both conditions, the operator has to rely on the mono (front) camera view; there is no direct view on the scene provided. The experiments were done with alternating conditions in order to prevent outliers and beginner's luck. Counterbalancing of conditions was done to eliminate learning and fatigue effects. The tests were performed with four experienced ('expert') operators, who work with similar manipulator arms on a daily basis. Only right-handed operators participated.

3.3. Task

The task was deliberately kept simple in order to test the impact of the 3D depth cue on visual feedback impact rather than the skill of the operator using the arms. A summary of the task sequence:

1. Lift the bolt runner from its starting place.
2. Move to the goal (a specific bolt).
3. Place the bolt runner in this hole.
4. Repeat according to sequence.
5. Replace the bolt runner in starting place.

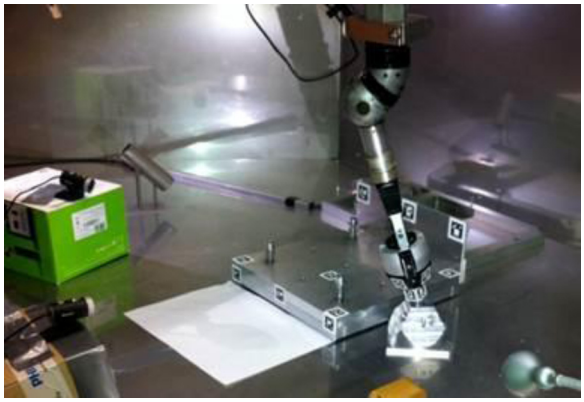

Fig. 2. Test setup of the first experiment: slave manipulator holding bolt runner. Front and side view cameras observe the scene. Tracking markers are placed on tool and task board.

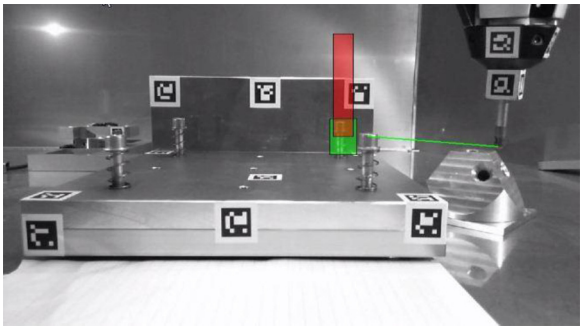

Fig. 3. New depth cues based on operators' comments after the first test: error bar (red) now displayed next to the bolt head, new task cue (green line without end circle). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

3.4. Experimental procedure

The operators were instructed to work with care and precision. The time needed to complete the task was measured, but precision and safety were deemed "more important". The gripper is already attached to the bolt runner (thus avoiding the grasp operation). After completing the task it is not necessary to release the bolt runner.

Four operators each performed eight runs under two conditions (a total of 16 runs), with a variation in the bolting pattern. Within one subject's test, there were no repetitions in the pattern. After the task, participants discussed their test experience by filling in the NASA TLX questionnaire. Despite 16 repetitions, the whole experiment was designed to take no longer than 60 min per participant.

3.5. Results

Robust real-time image processing (see Section 4.2) was achieved with marker-based objects. In the first set of tests, no significant difference could be found between the 'no depth cues (N)' and 'depth dues (D)' conditions. The simple depth gauge and graphical overlay of the 'depth cues' condition did not improve task performance. Operators commented on the placement of the depth indicator and their view of the bolt head "obstructed" by the task cue overlay (Fig. 3).

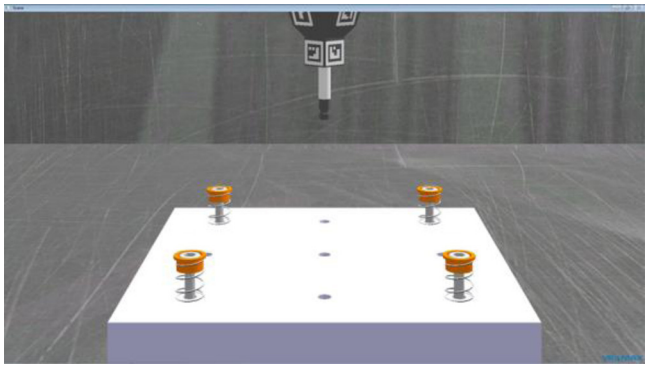

Fig. 4. VR version of the four bolt test.

Fig. 5. Experiment 2 setup: operator wearing stereo glasses looks at VR scene on the monitor, while manipulating the bolt runner via HW arm.

4. Stereo and active head tracking

4.1. Setup and experimental conditions

A VR version of the four-bolt test product was created (Fig. 4). The VR setup was linked to the real time tracking software on a simplified HW testbed with a passive manipulator (Fig. 5).

Five conditions were tested: direct view (DV), Mono VR (MONO), stereo (AS), head tracking (HT) and combination of stereo and head tracking (ASHT). Counterbalancing was done and the participants were asked to perform a practice trial before the actual test.

4.2. Tracking system

Video was recorded using Microsoft Lifecam Cinema cameras. The video streams were processed with OpenCV and video4rda. For tracking the objects we used AR markers as a tool for quickly recognizing and locating objects in a scene. For marker detection we used the framework as described in [4,5]. This framework is able to detect square markers with a 5×5 bit pattern on a white background. The head tracking software was based on an active marker (IR) and visual processing (Fig. 6).

For all relevant objects in the scene there is a CAD model available, including the positions of the AR markers. The pose of the objects with respect to the cameras is estimated initially using a 2D–3D pose estimation algorithm from OpenCV. Next, the initial pose estimations from all cameras are combined in a multi-camera multi-object pose estimation algorithm. The code was custom developed for this purpose and uses sparse Levenberg–Marquardt optimization. Visualization of the scene is performed with VR4Robots.

4.3. Data analysis

Key performance measures were time-to-complete, path deviation and safety (errors/retries). The time taken by a participant to perform the task was measured using the computer clock using manual indication of start and finish times and double checked with a stopwatch. Path deviation was measured by post processing the recorded motion tracks.

Using the tracking system described in Section 4.2 we could record for each experiment the exact position of the tool used (for the VR experiments the data can be extracted from the software). In Fig. 7 the *XYZ*-positions for a single track are plotted. From these tracks other quantities are calculated such as: total distance travelled, total horizontal distance travelled, total curve energy.

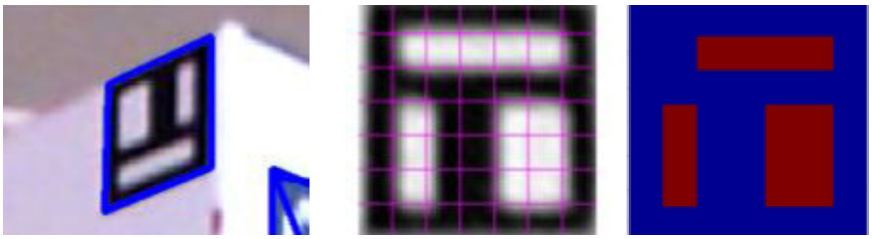

Fig. 6. Left: detected marker, centre: warp to normal form, right: extracted bit pattern.

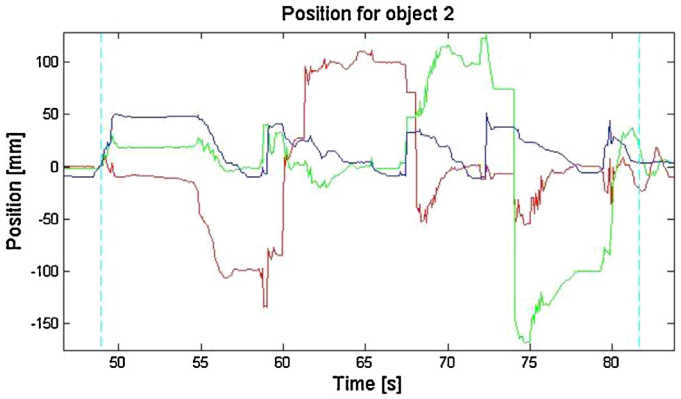

Fig. 7. *XYZ* track for single operation.

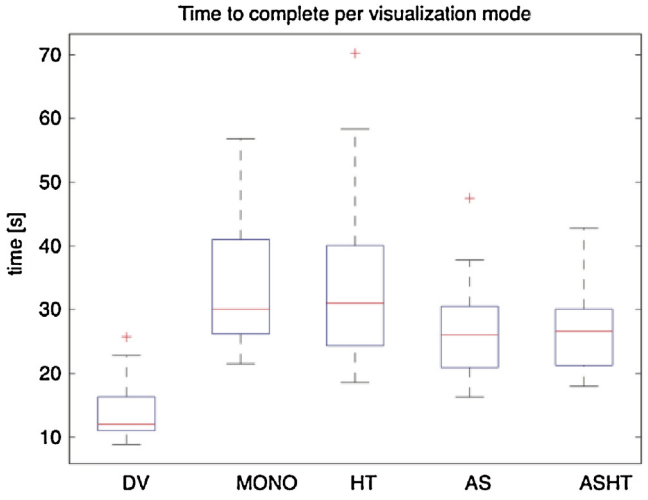

Fig. 8. Box plot of the time to complete under stereo (AS and ASHT) shows a significant improvement over MONO: 25 vs. 30 s average.

The resulting indicators can be analyzed using an interactive tool that is able to make scatter plots, histograms and density plots for all indicators.

4.4. Results

While the task execution time is significantly longer with all artificial views than with direct view (DV), stereo (AS) shows a significant improvement compared to MONO (Fig. 8). Head tracking (HT) was expected to show a similar improvement, but this could not be established. Limitations in tracking performance (delay, resolution) and setup may have had a negative influence on this result.

5. Discussion

The cameras used in the tests were relatively simple webcams. Resolution and frame rates are expected to be roughly equivalent to real cameras to be used in ITER RH, where the choice and quality of cameras is limited by radiation compatibility. In general, low quality camera images make it more difficult to estimate distances and depth in comparison with direct vision. Also, humans typically use binocular vision, which cannot be obtained from a single camera. Retinal disparity and stereopsis, two of the most important cues to depth, can thus not be used. This makes interpreting depth in low resolution single camera images extra hard.

6. Future research

In the ITER hot cell maintenance situation it will be possible to derive accurate 3D pose information on robots and robot-held tools from slave arm encoders and CAD models. Other objects will have to be actively tracked. This information is expected to lead to better quality Augmented Reality, which in turn will contribute to improved task performance.

Further development is needed towards robust 3D object tracking without markers, automatic object recognition, visual detection of deformation, contamination and damage and possibly real (camera based) stereo.

Acknowledgments

This work, supported by NWO, ITER-NL and the European Communities under the contract of the Association EURATOM/FOM, was carried out within the framework of the European Fusion Programme. The views and opinions expressed herein do not necessarily reflect those of the European Commission.

References

[1] C.J.M. Heemskerk, B.S.Q. Elzendoorn, A.J. Magielsen, G.Y.R. Schropp, Verifying elementary ITER maintenance Actions with the MS2 Benchmark product, Fusion Engineering and Design 86 (2011) 2064–2066.
[2] G.Y.R. Schropp, C.J.M. Heemskerk, A.M.L. Kappers, W.M. Bergmann Tiest, B.S.Q. Elzendoorn, D. Bult, Influence of visual feedback on human task performance in ITER Remote Handling, Fusion Engineering and Design 87 (2012) 808–812.
[3] http://en.wikipedia.org/wiki/Depth_perception (accessed 11.09.12).
[4] M. Hirzer, Marker detection for augmented reality applications, in: Institute for Computer Graphics and Vision, Graz University of Technology, Austria, 2008 http://studierstube.icg.tugraz.at/thesis/marker_detection.pdf
[5] ArUco: A Minimal Library for Augmented Reality pplications based on OpenCV, http://www.uco.es/investiga/grupos/ava/node/26
[6] ITER RH Code of Practice, ITER_D_2E7BC5 v1.2.