Towards advanced robotic manipulation for nuclear decommissioning: a pilot study on tele-operation and autonomy

Naresh Marturi[1,2]   Alireza Rastegarpanah[2]   Chie Takahashi[2]   Rustam Stolkin[2]
Jeffrey A. Kuo[3]   Yasemin Bekiroglu[2]

Abstract—**We present early pilot-studies of a new international project, developing advanced robotics to handle nuclear waste. Despite enormous remote handling requirements, there has been remarkably little use of robots by the nuclear industry. The few robots deployed have been directly teleoperated in rudimentary ways, with no advanced control methods or autonomy. Most remote handling is still done by an aging workforce of highly skilled experts, using 1960s style mechanical Master-Slave devices. In contrast, this paper explores how novice human operators can rapidly learn to control modern robots to perform basic manipulation tasks; also how autonomous robotics techniques can be used for operator assistance, to increase throughput rates, decrease errors, and enhance safety. We compare humans directly teleoperating a robot arm, against human-supervised semi-autonomous control exploiting computer vision, visual servoing and autonomous grasping algorithms. We show how novice operators rapidly improve their performance with training; suggest how training needs might scale with task complexity; and demonstrate how advanced autonomous robotics techniques can help human operators improve their overall task performance. An additional contribution of this paper is to show how rigorous experimental and analytical methods from human factors research, can be applied to perform principled scientific evaluations of human test-subjects controlling robots to perform practical manipulative tasks.**
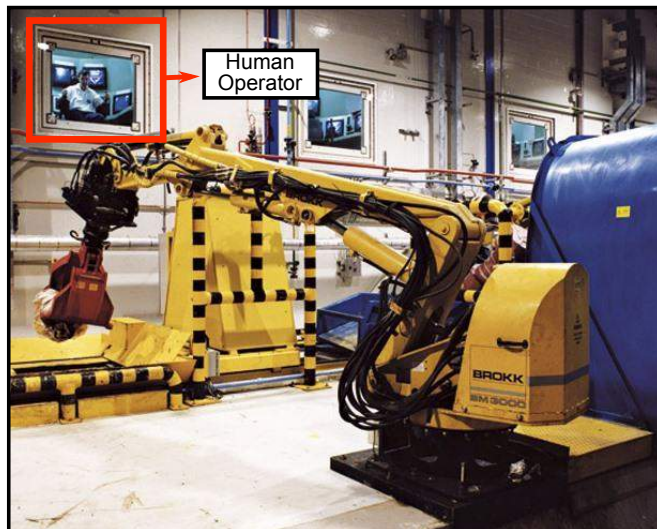
Figure 1. Human operator controlling a Brokk robot for pick and place operations inside a radioligically active environment. The operator is viewing the robot through a 1.2m thick lead-glass window, and controlling each joint with a separate lever, while guessing the inverse kinematics from experience.

I. INTRODUCTION

This paper discusses the potential of applying advanced robotics technologies to the many manipulation tasks involved in nuclear decommissioning. The human factors issues of tele-operation are explored; learning effects of novice robot operators are observed; and direct tele-operation is compared against the potential advantages of exploiting advanced autonomous robotics and computer vision methods. This paper reports early findings, of preliminary experiments, performed as part of a major new international project called RoMaNS (Robotic Manipulation for Nuclear Sort and Segregation) [1]. RoMaNS is a project funded by the European Commission's H2020 programme, with collaboration by five labs in three countries, including the nuclear agencies of UK and France.

Nuclear decommissioning, and the safe disposal of nuclear waste, is a global problem of enormous societal impor-

tance. Decommissioning the legacy waste inventory of the UK alone, represents the largest environmental remediation project in the whole of Europe, and is expected to take at least 100 years to complete, with estimated clean-up costs as high as £220billion (around $300billion) [2]. The quantities of legacy nuclear waste are vast, with intermediate level waste (ILW) alone amounting to 1.4 million cubic metres in the UK. Nuclear cleanup is a worldwide humanitarian issue (saving the environment for future generations) that must be faced by any country that has engaged in nuclear activities.

Much of the legacy nuclear waste is very old. Nuclear operations in USA and UK began in the 1940s, and greatly accelerated in both countries following the first USSR atomic bomb test in 1949. UK pioneered peaceful use of atomic energy, with the world's first industrial scale civil nuclear power plant coming online at the UK Sellafield site in 1956. Thus, in several countries, legacy nuclear waste materials and facilities can be more than two-thirds of a century old. At the UK Sellafield site, 69,600 cubic metres of legacy ILW waste must be placed into 179,000 storage containers. To avoid wastefully filling expensive high-level containers with low-level waste, many old legacy containers must be cut open, and their contents "sorted and segregated" [3]. This engenders an enormous requirement for complex remote manipulations,

since all of this waste is too hazardous to be approached by humans. In most nuclear plants worldwide, the vast majority of remote manipulation is performed by an aging (mean age >50 in UK) workforce of very highly skilled human operators, using mechanical Master-Slave Manipulator (MSM) devices. Such devices date back to at least 1949 [4] and at many nuclear sites have changed little in design since the 1960s. They comprise a "master" arm, grasped and moved by a human in a safe zone, which is mechanically linked to a system of gearing and cables which passes through a concrete shielding wall to a "slave" arm and gripper in the active contaminated zone. Perhaps surprisingly to a lay-person, remarkably few real robots have been deployed in the nuclear industry, and these have predominantly been directly tele-operated in rudimentary ways [5]. One such example can be seen in Fig. 1, where a human operator is tele-operating a hydraulic Brokk robot arm. Such robots, widely trusted in the industry due to their ruggedness and reliability, do not actually have joint-encoders, and no inverse-kinematics solving is possible for enabling Cartesian work-space control via a joystick.

Remote operation of dual-arm mobile manipulators for nuclear decommissioning tasks have been proposed [6], [7] and attempts have been made to deploy tele-operated robots at nuclear disaster sites [8], with robots controlled by viewing through a window, or via CCTV cameras mounted on or around the robot. A major difficulty is situational awareness, especially the lack of depth perception which affects accuracy and repeatability [9]. Direct tele-operation by humans is also slow, while the legacy waste inventory that needs processing is astronomical. We suggest that modern computer vision techniques are now robust enough to significantly enhance throughput rates, accuracy and reliability, by enabling partial or even full automation of many nuclear waste manipulation tasks.

Machine vision systems are already being used for a wide variety of industrial processes [10], where they provide information about scenes and objects (size, shape, colour, pose) which can be used to control a robot's trajectory in the task space [11], [12]. However, we are not aware of such visual-servoing technques being applied in the (highly conservative) nuclear domain. Some machine vision applications have been proposed for nuclear facilities. In [3], a vision-based solution is presented for identification and classification of nuclear waste items. [13] proposed a vision-based method for controlling the trajectory of inspection robots for nuclear reactors. Techniques for estimating radiation levels and nondestructive testing of piping systems can be found in [14] and [15], respectively.

However, comparatively little attention has been given to automating nuclear sort and segregation tasks, e.g., integration of tele-operation and automation using a dual-arm robot based on force feedback only [16]. As an initial step towards this, in this paper we realise a semi-autonomous framework for grasping and stacking objects. We use a model-based object tracker [17], [18] to track and estimate object poses, which are then fed to a robot trajectory planner to execute autonomous grasping. This is then combined with human-supervised motion for placement of grasped objects, making the overall task semi-autonomous.

We believe that a greater understanding of the underlying processes is necessary, before nuclear manipulation tasks can be safely automated. Therefore, an additional objective of this paper is to identify the problems that emerge during directly tele-operated tasks, and to then investigate how much might be gained by human-supervised autonomy technologies, over conventional direct tele-operation. We present a preliminary study which compares the performance of vision-guided semi-autonomy versus direct tele-operation, during basic grasping and manipulation tasks. We perform two different tasks using a 7 DoF articulated robot equipped with a two finger gripper. We first investigate a point-to-point dexterity task, which requires human subjects to control the position and orientation of the robot end-effector to reach a set of predefined goal positions as quickly as possible. We analyse the subjects' performance in executing the task repetitively and discuss the factors affecting the performance. We next investigate a more complex object grasping and stacking task, intended to represent Box Encapsulation (manipulating waste items into safe storage containers) which is a major challenge for the industry. We measure and compare the performance of humans on the grasping and stacking task, when using conventional direct tele-operation, and when employing a semi-autonomous system, where the human supervises at a high level, while the vision-guided robot automatically performs complex tasks such as grasping and trajectory execution.

Rest of this paper is structured as follows. We introduce our experimental set-up along with detailed task descriptions in Section II. In Section III we present our methodology, including procedures for data collection and analysis, and we also explain the technical details of our visual servoing method for semi-autonomous manipulation. Section IV presents the experimental evaluations, followed by a discussion and conclusions in Section V.

II. EXPERIMENTAL SET-UP

Two different tasks are designed to study various factors effecting performance of fully supervised tele-operated handling and vision-guided semi-autonomous manipulations. The developed experimental set-up along with the designed test-rig to perform this study are detailed first in this section. Later we present the two tasks in detail.

A. Robotic Test Platform

The experimental set-up used to perform both tele-operated as well as semi-automatic tests is shown in Fig. 2. It consists of a collaborative 7 DoF lightweight robot arm: KUKA LBR iiwa 14 R820, a Schunk PG plus-70 2-finger parallel jaw gripper with a stroke size of 68 mm and a commercial Logitech c920 USB camera. The gripper is attached to the
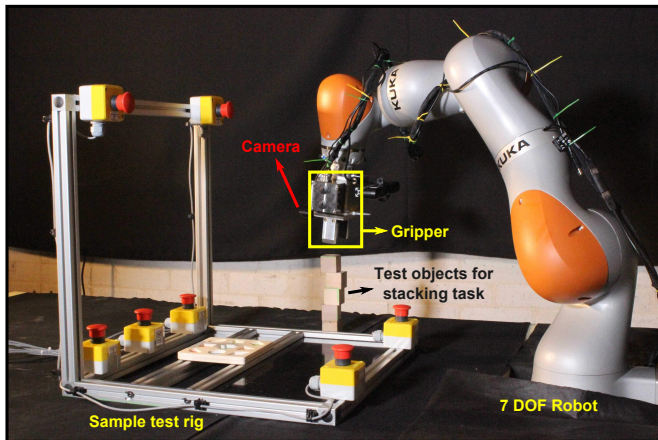
Figure 2. Experimental set-up showing various components used in this work. The test rig with seven buttons, used for the point-to-point dexterity task, can be seen on the left.
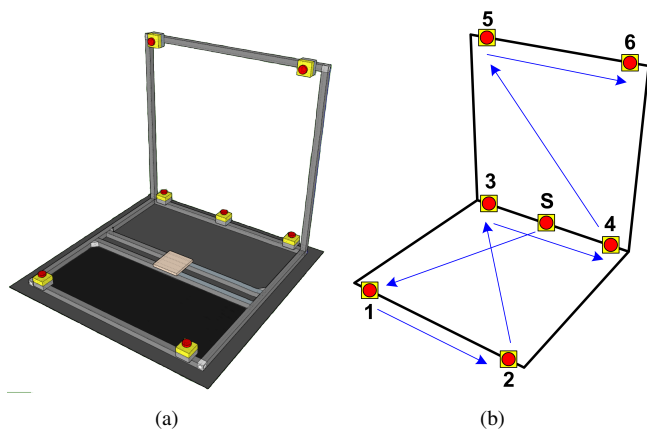

Figure 3. (a) CAD model of the test-rig used for point-to-point tele-operation task. (b) Button-press order for navigation with arrows illustrating navigation directions required for the point-to-point task.

tool center point of the robot and the camera is mounted on top of the gripper whose optical axis coincides with the z-axis of the gripper frame. The vision-guided control architecture for semi-automatic tests was implemented in C++ and is executed from a remote computer running Windows (8 GB RAM, 2.3 GHz Intel core i7 CPU). The communication between the PC and the robot controller has been realised using UDP and the gripper control commands are transmitted over serial port. Real-time matrix computations are performed using ViSP [17].

B. Test-bed for Tele-operation

The tele-operation test-bed is an aluminium frame consisting of seven industrial emergency stop push buttons marked as $S \cdots 6$ as shown in Fig 3. The outer dimensions of the frame are $50 \times 50 \times 50$ cm^3 and the diameter of each button head is 40 mm. Each button on the frame is electrically connected to the work computer via U-HID (USB) interface board in order to record the corresponding button press event (with a beep sound) as well as the overall task time. Time recording automatically starts when the start button S is pressed and the total time is saved with a task label when the last button *i.e.,* button-6 is pressed. The directions to navigate from point-to-point are pre-selected and are shown with the arrows in Fig. 3(b).

A software user interface (UI) as shown in Fig. 4 has been implemented in order to provide various functionalities to tele-operate the robot as well as to control other system components. Firstly, it allows the operator to jog the robot in both Cartesian and joint spaces with different speeds. However, for the sake of task simplicity only Cartesian jogging in world coordinate frame is used in this work. Next, it allows to configure and jog the robot using a standard USB keyboard that acts like a low level haptic device in moving the robot. Finally it allows to open and close the gripper whenever required and provides views from all the scene cameras.
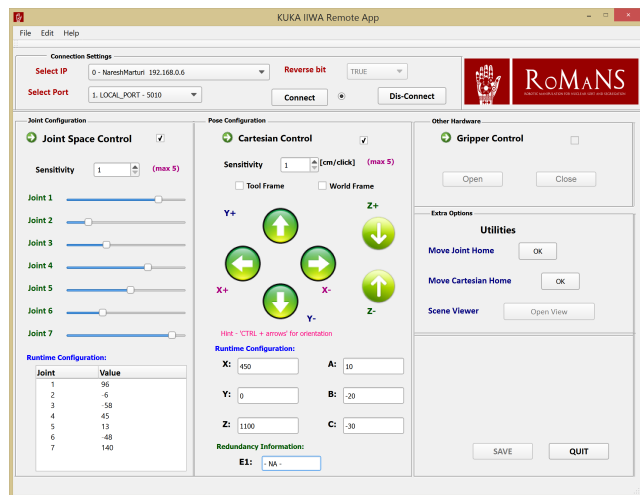

Figure 4. Our user interface for controlling various components in the system.

C. Task Description

The two tasks designed to analyse the differences between fully supervised tele-operated handling and vision-guided semi-autonomous manipulations are detailed below.

1) Point-to-point Dexterity Task: The initial task of point-to-point robot navigation has been designed to evaluate the performance of human subjects who are required to execute the task by passive vision *i.e.,* by looking at the camera-provided views of the workspace (see Fig. 5). This is usually the case in nuclear environments where tele-operated robots are used with several camera views of the application area. This task requires the participants to move the robot's end effector from button-to-button on the test-rig in a specific order described by the arrows in Fig. 3. To successfully complete this task, participants need to change both the position and orientation of the end effector to be able to reach the buttons.
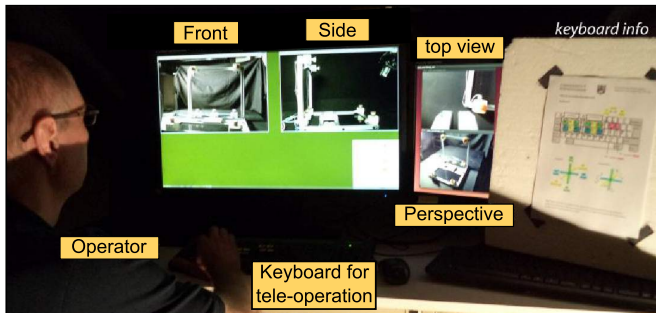
Figure 5. Human operator performing the point-to-point tele-operation task by looking at multiple camera views of the work space.

2) Block Stacking: The second task is stacking of wooden blocks seen in Fig. 2. This task has been designed to replicate packaging scenario in nuclear facilities, where the operators are required to load various objects in the containers sequentially. In this work, we perform this task to study and analyse the major differences between tele-operation and vision-guided semi-autonomy. To carry out human subject performance evaluation, we recruited multiple participants who manoeuvred the robot from a pre-fixed distance using the developed software UI to stack various wooden blocks presented in the workspace. For comparison, we semi-automated the task by using vision information as feedback in localising the arm to a pre-grasp location. To this extent, we implement a CAD model-based tracking to track the object pose in robot frame. In both cases, we evaluate the performance in terms of task repeatability, handling and placement accuracy, and overall stacking time.

III. METHODOLOGY

In this section, we first describe the data acquisition procedure for tele-operation experiments, then detail the implementation of vision-guided semi-autonomous block stacking.

A. Tele-operation: Data Acquisition and Analysis

1) Procedure for the Point-to-point Dexterity Task: Six participants from the students and staff at the University of Birmingham voluntarily participated in this experiment . All had normal or corrected to normal vision, and had no known neurological motor deficits (self-reported). Participants took a 10 minutes training session followed by the main experiments. In the training session, they were provided with an instruction set of the task and asked to familiarize with the robot control. In the main experiment, they were asked to complete the task three times without any certain time limitation. They were allowed to have a break between the tree trials.

After the end of final session, participants were asked to answer the computerized version of the NASA Task Load Index (NASA-TLX) assessment as well as system usability scale (SUS) questionnaires. The NASA-TLX survey is widely used in human factor research [19]. The TLX consists of questionnaires where a participant subjectively rates the workload on 6 different sub-scales: Mental Demand: how mentally demanding was the task, Physical Demand: how physically demanding the task was, Temporal Demand: how hurried or rushed the pace of the task was, Effort: how hard the participant has to work to accomplish the level of performance, Frustration Level: how insecure, disorganised, irritated, stressed, or annoyed the participant was, and Overall Performance: how successful in accomplishing the task the participant was, after conducting the specific task. The raw scores are within a 100 points range (very low: 0 to very high: 100). In this study, we employed the computer versions of the TLX and only raw scores were analysed [20], [21]. In general, the TLX scores indicate the task difficulties. At each trial, the task performance was measured based on several factors, the time taken to complete the task, the number of collisions with the environment, the number of missed buttons, the number of unacceptable critical incidents such as user failure to cause the robot to stop with wrong input.

2) Procedure for the Block Stacking Task: This experiment involves tele-operating the robot to remotely pick-up and stack 5 cube-shaped wooden blocks. The dimensions of each block are $4 \times 4 \times 4$ cm^3. A group of 5 new participants were recruited for this task. Each participant had 10 minutes training to get themselves familiarise with the robot jogging. Similar to the previous experiment, participants tele-operated the robot using keyboard functions. However, in this case the participants were asked to complete the task by directly viewing the robot standing two meters far from the robot. The blocks were situated in the workspace ensuring that they were within reach of the robot and also they were sufficiently apart from each other such that the gripper fingers do not collide with neighbouring blocks. Apart from that, to have a fair comparison, blocks were placed at the same positions for all trials and for all participants.

For the main experiment, each participant was asked to perform the task three times and during each trial, similarly to the previous task, the time taken to complete the task, the number of collisions with the environment, the number of dropped objects, and the number of unacceptable critical incidents were recorded. The task has been considered successful only when the stacked objects remain still after placing the final block.

B. Visual-guided Semi-automatic Stacking of Blocks

The overall task is divided into two phases: 1) automatic navigation of robot and grasping of objects, and 2) stacking blocks at a pre-defined location using the prior knowledge of block dimensions. At the first hand, for automatic navigation, we use the images provided by the camera mounted on top of the parallel jaw gripper for estimating the objects' pose in the robot frame. In turn this information will be used as a feedback to navigate the robot end-effector to a desired location *i.e.*, to the pre-grasp location. This intermediate location has been selected to avoid the black-spots for tracking *i.e.*, since the camera is mounted on top of the gripper, the object will be moved out of the camera field of view when

the robot end-effector is positioned immediately above the object and can induce errors in the estimated poses. Once the robot is at pre-grasp location with correct pose (provided by tracker), it will be automatically re-directed to the grasp location that provides a stable grasp. The stable grasps are identified by the last estimated pose from the tracker. Once the object is grasped successfully, robot navigates to a pre-defined stacking location. Objects are stacked based on the number of the object that is being handled ($\kappa = 1 \cdots 5$) and the object height ($\Omega = 40$ mm) *i.e.*, the z location to place the object is $\kappa * \Omega$ and, x and y locations are pre-supplied and remain same for each block.

1) 3D Model-based Tracking: As mentioned before, we use a model-based object tracker [17] to track each block's pose. It uses 3D CAD models to track and estimate the object's pose in real-time. The underlying idea behind the tracker is to obtain a camera pose for which the projected CAD model fits best with the 2D image contours of the object. This process involves estimating a rigid transformation between the virtual camera frame and the tracked object frame. Previously, these types of methods were studied for tracking objects in complex environments and are proved to be efficient in tackling various computer vision issues such as intensity variations, partial occlusions etc. [22], [23].

Tracking a 3D model of an object in an image counts back to the classical pose estimation problem. The basic idea is to project the model using an initial pose, propagate it to the next frame and perform a 1D search along the model edges to update the pose. In general, the pose matrix $^c\mathbf{M}_o$ links the 3D object features $\mathbf{P}$ in the world frame to their corresponding projections $\mathbf{p}$ in the image by

$$\mathbf{p} = Pr(^c\mathbf{M}_o, \mathbf{P}) \tag{1}$$

where, $Pr(\cdot)$ is the projection operator. Assuming, the camera intrinsic parameters $\mathbf{K}$ are known, (1) becomes

$$\mathbf{p} = \mathbf{K}^c\mathbf{M}_o\mathbf{P} \tag{2}$$

If we assign a virtual camera with the object model, the objective is to estimate the pose of the camera $^c\mathbf{M}_o$ to fit the model. In order to estimate $\mathbf{r}$ (6 degree of freedom pose vector computed from $^c\mathbf{M}_o$), in this work, we solely rely on the object edge information and use the *virtual visual servoing*(VVS) framework proposed by Marchand and Chaumette [24]. The idea is to estimate the transformation parameters by minimising the error $\mathbf{\Delta}$ between the current values $\mathbf{s}(\mathbf{r})$ obtained by forward projection of model using pose $\mathbf{r}$ and the ground truth values $\mathbf{s}^*$ measured in the image and is given by

$$\mathbf{\Delta} = \sum_{i=1}^{n} (s_i(\mathbf{r}) - s_i^*)^2 \tag{3}$$

The regulation of $\mathbf{\Delta}$ requires linking temporal variations of $s_i(\mathbf{r})$ with the velocity screw of the virtual camera defined by pose $\mathbf{r}$. This is achieved by using an image Jacobian matrix $\mathbf{J_s}$. This process is similar to the classical visual

Figure 6. Initialisation of trackers for all blocks in the very first image. The wire frames illustrate the object CAD models. These are projected onto the image using initial poses supplied by the human operator.

servoing and hence, we refer the reader to [24] for further details. Alternatively, regulation of (3) can be optimised by using iterative optimisation algorithms such as Quasi-Newton, Levenberg-Marquardt etc.

In this work, we track five different cube-shaped blocks present in the workspace. Initially, the robot moved to a pre-defined home position, which is fixed throughout the process. Trackers are initialised simultaneously for each object in the very first image using the user supplied initial poses as shown in Fig.6. In the case of highly cluttered scenes due to presence of redundancy edge information, the tracked poses are highly unstable. In order to minimise these phenomena, we use a Kalman Filter to predict and update the final pose after VVS, thus smoothing the changes.

2) Pose-based Visual Controller: If we decompose the overall task into 5 different subtasks of pick (autonomous) and place (pre-defined) of objects, we need a reliable visual controller (using pose features obtained from tracking) to navigate the arm to the grasping locations of each object. In order to accomplish this, we implement a pose-based visual controller in this work. Simplifying the details, we below show how this can be done for one object.

Let $\mathcal{R}_c$, $\mathcal{R}_o$ and $\mathcal{R}_{o^*}$ be the frame attached to the camera and the current and desired object frames, respectively. The tracker provides the homogeneous transformations $^c\mathbf{M}_o$ and $^c\mathbf{M}_{o^*}$ between the camera and the current and desired object frames, respectively. In order to move the robot to the desired object pose $\mathbf{r}^*$ (computed from $^c\mathbf{M}_{o^*}$) the error given by (4) has to be regulated.

$$^o\mathbf{M}_{o^*} = {^c\mathbf{M}_o}^{-1}{^c\mathbf{M}_{o^*}} \tag{4}$$

Let $\mathcal{R}_R$ be the frame associated with the robot base. Then the homogeneous transformation between the object and the base frame is

$$^R\mathbf{M}_o = {^R\mathbf{M}_c}\,{^c\mathbf{M}_o} \tag{5}$$

where, $^R\mathbf{M}_c$ denotes position and orientation of the camera in the robot base frame. Using (5), the error to be regulated is

$$\mathbf{e} = \left(^R\mathbf{t}_o - {}^R\mathbf{t}_{o*}, \theta\mathbf{u}\right) \tag{6}$$

where, $^R\mathbf{t}_o$ and $^R\mathbf{t}_{o*}$ are translation vectors derived from their corresponding homogeneous transformations and $\theta\mathbf{u}$ is the angle-axis representation of rotation.

Now by linking the time variation of current features $\mathbf{s} = {}^R\dot{\mathbf{M}}_o = \left(^R\dot{\mathbf{t}}_o, \theta\dot{\mathbf{u}}\right)^\top$ to the robot velocity, we get

$$\underbrace{\begin{bmatrix} ^R\dot{\mathbf{t}}_o \\ \theta\dot{\mathbf{u}} \end{bmatrix}}_{\dot{\mathbf{s}}} = \underbrace{\begin{bmatrix} \mathbf{J_v} & \mathbf{0} \\ \mathbf{0} & \mathbf{J_w} \end{bmatrix}}_{\mathbf{J}} \underbrace{\begin{bmatrix} \boldsymbol{v} \\ \boldsymbol{w} \end{bmatrix}}_{\boldsymbol{v}} \tag{7}$$

where, $\boldsymbol{v}$ and $\boldsymbol{w}$ are linear and angular velocities, respectively and $\mathbf{J_v}$ and $\mathbf{J_w}$ are linear and angular Jacobians. In this work we chose $\mathbf{J_v} = \mathbf{I}$ and $\mathbf{J_w}$ as provided in [25]. Using (7), he final control law is given by:

$$\mathbf{v} = -\lambda\mathbf{J}^\dagger \left(^R\mathbf{M}_o - {}^R\mathbf{M}_{o*}\right) \tag{8}$$

where, $\mathbf{J}^\dagger$ the pseudo inverse of the Jacobian matrix and λ is a positive gain to ensure the exponential convergence. Equation (8) navigates the robot to a pre-grasp location at which the visual control stops. The final pose of the object during the visual servoing process is retained for grasping purposes from whereon the robot follows a pre-defined motion until placing the object at a desired location.

IV. Experimental Evaluation

In this section we present the experimental results obtained for both tele-operation and vision-guided semi-autonomous tasks along with a detailed analysis investigating the task performance.

A. Results for the Point-to-point Dexterity Task

The performance of all 6 participants during point-to-point dexterity task has been evaluated based on the criteria listed in Section III. Although some participants failed to press the buttons a number of times during their training session, no such instances were observed during the main trials. As shown in Fig. 7(a), the averaged time-to-completion across 6 participants were 272.8 seconds at the first trial and then reduced to 199.7 seconds at the third trial. The repeated measure ANOVA (Analysis of Variance) was used on the time-to-completion data in order to evaluate whether the performance significantly changed or not across three trials [26]. Mauchly's test indicated that the assumption of sphericity for ANOVA is satisfied: $\chi^2(2) = 2.17$, $p = .338$. The ANOVA test shows that there was a significant learning effect across the three trials: $F(2, 10) = 8.739$, $p = .006 < .05$. These results suggest that the task performance can be constantly improved through the repetition.

Fig. 8 shows the NASA-TLX scores. The Mental Demands are significantly high compared with other sub-scales. This seems to reflect that the point-to-point dexterity task with button pressings required participants to construct the 3D

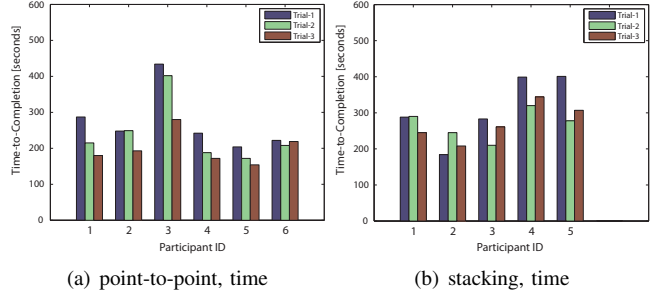(a) point-to-point, time   (b) stacking, time

Figure 7. Time spent to complete a) the point-to-point dexterity task and b) the block stacking task for all participants. For each participant task completion time in each trial is shown.
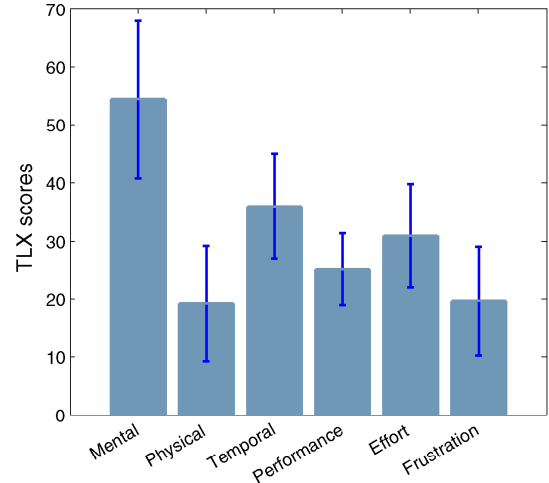

Figure 8. NASA Task Load Index scores for the point-to-point dexterity task, averaged across 6 participants.

perception of the remote workspace through the 2D images of the live camera-feeds. At the same time, participants also needed to control the robot arm by tele-operation, which is the 2D space configuration and intuitively difficult to operate corresponding to the 3D space. These operations require high cognitive load and functioning. On the contrary, the physical demands and frustration were relatively low compared with other sub-scales, suggesting that the tele-manipulation could reduce the physical tiredness for such a repetitive task. This trend might depend on the experimental design, *i.e.*, no-time limit for the completion. Participants could focus more on their performance rather than the temporal demand.

Overall results suggest that designed experimental set-up is sufficiently convenient for the participants and they improved their performance in terms of completion time over the trials.

B. Results for the Block Stacking Task

Fig. 7(b) shows the time-to-completion of all 5 participants for the stacking task. The averaged time-to-completions across 5 participants were 311.0 seconds at the first trial and then slightly reduced to 273.0 seconds at the third trial. Similarly to the previous task, the repeated measure ANOVA was

Participant ID	Π		α		β		γ	
	Av.	Std.	Av.	Std.	Av.	Std.	Av.	Std.
1	274.3	25.4	3.6	3.2	0.6	0.5	0.6	0.5
2	212.3	30.7	0.3	0.5	0.3	0.5	0.0	0.0
3	251.3	37.4	2.6	3.0	0.6	0.5	0.3	0.5
4	354.3	40.5	2.0	3.4	1.6	0.5	0.0	0.0
5	328.6	64.2	1.6	2.8	0.3	0.5	0.0	0.0

Av. = Average over 3 trials, Std. = Standard deviation.

used on the time-to-completion data. Mauchly's test indicated that the sphericity assumption is satisfied: $\chi^2(2) = 2.17$, $p = .257$. In contrast to the point-to-point dexterity task, the results show that there was no significant effect across the three trials: $F(2, 8) = 1.797$, $p = .227$. This can be explained by the fact that in the block stacking task, there were unpredictable, or accidental, factors due to the interactions between the end effector and the blocks; e.g., sometimes as more blocks were stacked not sufficiently well aligned, the tower could become unbalanced and easily collapse. Overall, the number of critical incidents (collisions and drops) were relatively high compared with the point-to-point dexterity task. Table I shows the task completion time (Π), the number of collisions (α), the number of dropped objects (β), the number of critical incidents (γ) for each trial. The number of accidents was relatively high compared with the point-to-point dexterity task. Although the task completion time reflected the effect of such unpredictable factors, the impact on the task performance was huge especially within small number of trials. We plan to extend the evaluations to deal with such variations in performance by e.g., designing a task score considering not only time-completion but also weight, or penalty, for unwanted behaviour.

Even though the participants directly viewed the workspace (which we expect would provide good depth perceptions compared with the point-to-point dexterity task), there were difficulties to construct straight and stable block towers. This seemed to have a large impact on the task performance, which possibly was one of the main reasons that the learning effect was disappeared across three trials. Overall, this task was more difficult to achieve compared to the previous one and require more trials for participants to excel. To avoid such accidents, providing the participants with more information, e.g. force feedback, can be helpful.

C. Results for Vision-guided Stacking of Blocks

We have conducted experiments in order to evaluate the semi-autonomous robotic performance in performing the block stacking task. As mentioned before, this task is decomposed into two phases: automatic navigation and grasping of blocks using vision feedback and stacking blocks at a predefined location. In order to have a fair evaluation, the blocks were placed in similar locations as for the tele-operated task. Trackers are automatically initialised from the user defined
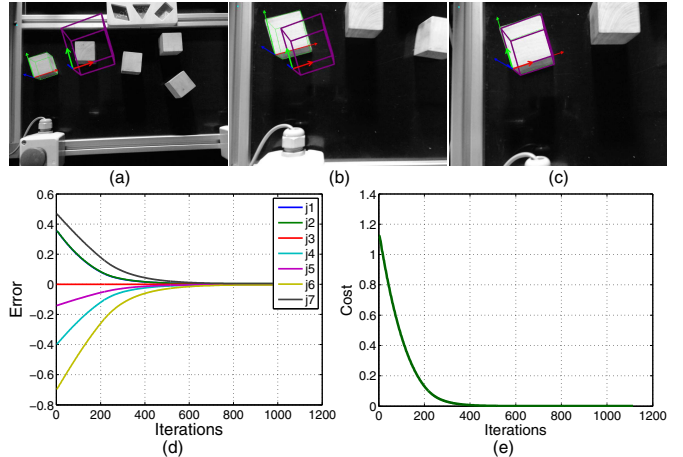

Figure 9. (a)-(c) Series of images obtained during the automatic arm navigation to pre-grasp location for block 1. Purple wireframe represents the desired pose and the green one represents current pose. Pose errors at (a) initial and (b) final iterations during the visual servoing process can be seen. (d) and (e) illustrate the error for all seven joints and cost variations respectively, during visual servoing for block 1. Cost is considered to be sum of squared error and is used to verify the task convergence.

initial poses. Then the robot has been automatically navigated to the pre-grasp pose, which is accomplished by regulating the positional error. Fig. 9(a)-(c) illustrate a series of images obtained during the robot navigation with visual feedback. Error and cost variations during the task for the first object are shown in Fig. 9(d) and (e), respectively. Here cost is considered to be the sum of squares of error $i.e.$, $\mathbf{e}^2$ and has been used to check the process convergence. Fig. 10 shows a series of images during the stacking of blocks. In order to evaluate the performance, this task has been executed 11 times and the corresponding time for each trial has been shown in Fig. 11. On average it requires 88 seconds (in robot T1 mode) to stack 5 different blocks, over three times faster than the tele-operation experiments. Apart from that the task has also been monitored for collisions and failures. The overall performance directly depends on the success of tracking system. Due to constant lighting and cluster-less environment, the trackers succeeded in all 11 trials and hence, no task failures or collisions were observed. Also, unlike with human tests, there were no shortcomings in the depth perception, which we think is the main reason behind reliable performance. However, in either of the cases $i.e.$, both semi-autonomous and human tests, integrating tactile information with grasping can improve the overall performance [27].

V. CONCLUSION

This paper has investigated human performance in executing remote manipulation tasks by teleoperating a robot. We have also presented and demonstrated a system for semi-autonomous manipulation using visual servoing. Performance of human test-subjects using both systems was evaluated and compared. Our results show how performance improves with training, and suggest how training requirements scale
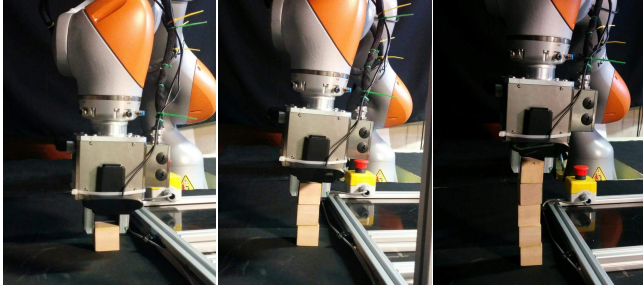
Figure 10. Series of images illustrating the stacking of various blocks during the semi-automated task. Rightmost image illustrates the task success.
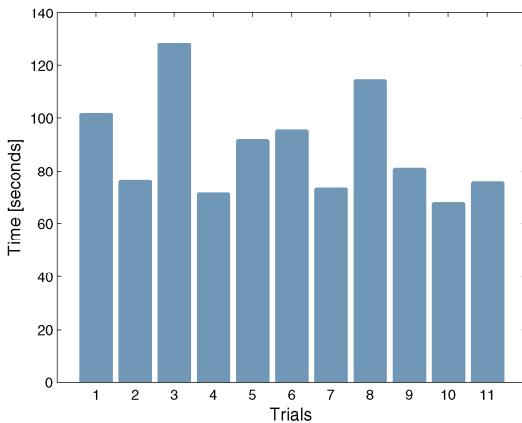

Figure 11. Overall time taken for semi-automated block stacking during 11 runs. On average the stacking of 5 blocks took 88 seconds, more than three times faster than teleoperation mode.

with task complexity. Our results also demonstrate how the incorporation of autonomous robot control methods can reduce workload for human operators, while improving task completion time, repeatability and precision. We plan to extend this initial pilot-study with additional experiments, including greater numbers of participants and additional manipulation example tasks. We also plan to extend our study to include advanced haptic feedback control devices.

ACKNOWLEDGEMENTS

The authors would like to thank Mr. Joshua Price of NNL Ltd. for designing the test-bed.

REFERENCES

[1] . European Commission, *RoMaNS - Robotic Manipulation for Nuclear Sort and Segregation*. http://www.h2020romans.eu/, 2015.

[2] Nuclear Decommissioning Authority, "Nuclear provision: explaining the cost of cleaning up BritainâĂŹs nuclear legacy," *Corporate Report*, 2015.

[3] A. Shaukat, Y. Gao, J. A. Kuo, B. A. Bowen, and P. E. Mort, "Visual classification of waste material for nuclear decommissioning," *Robotics and Autonomous Systems*, vol. 75, pp. 365–378, 2016.

[4] R. Goertz, "Master-slave manipulator," *Technical Report, Argonne National Laboratory, USA*, March 7th, 1949.

[5] J. Vertut, *Teleoperation and robotics: applications and technology*. Springer Science & Business Media, 2013, vol. 3.

[6] C. J. Taylor, A. Chotai, and D. Robertson, "State dependent control of a robotic manipulator used for nuclear decommissioning activities," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*, Oct 2010, pp. 2413–2418.

[7] M. Bakari, K. Zied, and D. Seward, "Development of a multi-arm mobile robot for nuclear decommissioning tasks," *International Journal of Advanced Robotic Systems*, vol. 4, no. 4, pp. 387–406, 2007.

[8] *http://spectrum.ieee.org/automaton/robotics/industrial-robots/fukushima-robot-operator-diaries*. IEEE, 2011.

[9] J. Iqbal, A. M. Tahir *et al.*, "Robotics for nuclear power plants?challenges and future perspectives," in *Applied Robotics for the Power Industry (CARPI), 2012 2nd International Conference on*. IEEE, 2012, pp. 151–156.

[10] J. L. Sanz, *Advances in machine vision*. Springer Science & Business Media, 2012.

[11] N. Marturi, V. Ortenzi, J. Xiao, M. Adjigble, R. Stolkin, and A. Leonardis, "A real-time tracking and optimised gaze control for a redundant humanoid robot head," in *IEEE International Conference on Humanoid Robots*, 2015, pp. 467–474.

[12] N. Marturi, B. Tamadazte, S. Dembélé, and N. Piat, "Visual servoing schemes for automatic nanopositioning under scanning electron microscope," in *IEEE International Conference on Robotics and Automation*, 2014, pp. 981–986.

[13] J.-Y. Park, B.-H. Cho, and J.-K. Lee, "Trajectory-tracking control of underwater inspection robot for nuclear reactor internals using time delay control," *Nuclear engineering and design*, vol. 239, no. 11, pp. 2543–2550, 2009.

[14] C.-H. Baek, S. J. An, H.-i. Kim, S.-W. Kwak, and Y. H. Chung, "Development of a pinhole gamma camera for environmental monitoring," *Radiation Measurements*, vol. 59, pp. 114–118, 2013.

[15] A. Ettemeyer, "New techniques for nondestructive testing of piping systems in nuclear power plants," *Nuclear engineering and design*, vol. 139, no. 2, pp. 183–185, 1993.

[16] C. Kapoor and D. Tesar, "Integrated teleoperation and automation for nuclear facility cleanup," *Industrial Robot: An International Journal*, vol. 33, no. 6, pp. 469–484, 2006.

[17] E. Marchand, F. Spindler, and F. Chaumette, "Visp for visual servoing: a generic software platform with a wide class of robot control skills," *IEEE Robot. Autom. Mag.*, vol. 12, no. 4, pp. 40–52, 2005.

[18] V. Ortenzi, N. Marturi, R. Stolkin, J. A. Kuo, and M. Mistry, "Vision-guided state estimation and control of robotic manipulators which lack proprioceptive sensors," in-press.

[19] A. Steinfeld, T. Fong, D. Kaber, M. Lewis, J. Scholtz, A. Schultz, and M. Goodrich, "Common metrics for human-robot interaction," in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*. ACM, 2006, pp. 33–40.

[20] S. G. Hart, "Nasa-task load index (nasa-tlx); 20 years later," in *Proceedings of the human factors and ergonomics society annual meeting*, vol. 50, no. 9. Sage Publications, 2006, pp. 904–908.

[21] N. H. P. R. Group *et al.*, "Task load index (nasa-tlx) v1. 0 computerised version," *NASA Ames Research Centre*, 1987.

[22] S. Dembélé, N. Piat, N. Marturi, and B. Tamadatze, "Gluing free assembly of an advanced 3d structure using visual servoing," in *Micromechanics and Microsystems Europe 2012*, 2011.

[23] A. Comport, E. Marchand, M. Pressigout, F. Chaumette *et al.*, "Real-time markerless tracking for augmented reality: the virtual visual servoing framework," *Visualization and Computer Graphics, IEEE Transactions on*, vol. 12, no. 4, pp. 615–628, 2006.

[24] É. Marchand and F. Chaumette, "Virtual visual servoing: a framework for real-time augmented reality," in *Computer Graphics Forum*, vol. 21, no. 3. Wiley Online Library, 2002, pp. 289–297.

[25] E. Malis, F. Chaumette, and S. Boudet, "21/2d visual servoing," *Robotics and Automation, IEEE Transactions on*, vol. 15, no. 2, pp. 238–250, 1999.

[26] C. N. von Ende, "Repeated-measures analysis," *Design and analysis of ecological experiments. Oxford University Press, Oxford*, pp. 134–157, 2001.

[27] Y. Bekiroglu, J. Laaksonen, J. A. Jorgensen, V. Kyrki, and D. Kragic, "Assessing grasp stability based on learning and haptic data," *IEEE Transactions on Robotics*, vol. 27, no. 3, pp. 616–629, 2011.