We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900
Open access books available

185,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Chapter 3

Towards Advanced Robotic Manipulations for Nuclear Decommissioning

Naresh Marturi, Alireza Rastegarpanah,
Vijaykumar Rajasekaran, Valerio Ortenzi,
Yasemin Bekiroglu, Jeffrey Kuo and Rustam Stolkin

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.69739

Abstract

Despite enormous remote handling requirements, remarkably very few robots are being used by the nuclear industry. Most of the remote handling tasks are still performed manually, using conventional mechanical master-slave devices. The few robotic manipulators deployed are directly tele-operated in rudimentary ways, with almost no autonomy or even a pre-programmed motion. In addition, majority of these robots are under-sensored (i.e. with no proprioception), which prevents them to use for automatic tasks. In this context, primarily this chapter discusses the human operator performance in accomplishing heavy-duty remote handling tasks in hazardous environments such as nuclear decommissioning. Multiple factors are evaluated to analyse the human operators' performance and workload. Also, direct human tele-operation is compared against human-supervised semi-autonomous control exploiting computer vision. Secondarily, a vision-guided solution towards enabling advanced control and automating the under-sensored robots is presented. Maintaining the coherence with real nuclear scenario, the experiments are conducted in the lab environment and results are discussed.

Keywords: nuclear decommissioning, robot tele-operation, robot vision, visual servoing

1. Introduction

Nuclear decommissioning, and the safe disposal of nuclear waste, is a global problem of enormous societal importance. From the world nuclear statistics, there are over 450 nuclear plants operating in the world and out of which, 186 are currently being operated within Europe [1].

INTECH
open science | open minds

© 2017 The Author(s). Licensee InTech. This chapter is distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. CC BY

At present, nuclear industry forms the main basis for approximately one quarter of the EU's total power generation, which is forecasted to be increase by at least 15% by 2030. 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. Due to the raising concerns over fossil-fuelled power generation especially with the alarming levels of greenhouse gasses and the difficulty in managing nuclear waste, many nuclear plants worldwide are undergoing some revival. While many countries plan to rejuvenate their nuclear plants, countries like UK are presently decommissioning their old nuclear plants. Nevertheless, nuclear clean-up is a worldwide humanitarian issue (saving the environment for future generations) that must be faced by any country that has engaged in nuclear activities.

Despite the fact that nuclear activities around world are increased, it is estimated that many nuclear facilities world-wide will reach their maximum operating time and require decommissioning in the coming two or three decades. Thousands of tons of contaminated material (e.g. metal rods, concrete, etc.) need to be handled and safely disposed until they no longer possess a threat. This process involves not only the cleaning costs and human hours, but also the risk of humans being exposed to radiation. 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 £220 billion (around $300 billion) [2]. Worldwide decommissioning costs are of order $trillion. Record keeping in the early days was not rigorous by modern standards, and there are now many waste storage containers with unknown contents or contents of mixed contamination levels. At the UK Sellafield site, 69,600 m^3 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.

The vast majority of these remote manipulation tasks in most of the nuclear sites around the world are still performed manually (by ageing workforce), where eminently skilled human operators control bulky mechanical Master-Slave Manipulator (MSM) devices. Usage of MSMs at nuclear plants dates back to at least 1940s and has changed slightly in design since then. Notably, few heavy-duty industrial robot manipulators have been deployed in the nuclear industry during the last decade (replacing MSMs) to be used for decommissioning tasks. However, most of these have predominantly been directly tele-operated in rudimentary ways [4]. An example can be seen in **Figure 1**, where an operator is looking through a 1.2 m thick lead-glass window (with very limited situational awareness or depth perception) and tele-operating the hydraulic BROKK robot arm by pushing buttons to control its various joints. Such robots, widely trusted in the industry due to their ruggedness and reliability, do not actually have proprioceptive joint encoders, and no inverse kinematics solving is possible for enabling Cartesian workspace control via a joystick. Instead, the robot's inverse kinematics has to be guessed from the operator's experience, which directly affects the task performance. It is not considered feasible to

http://dx.doi.org/10.5772/intechopen.69739



Figure 1. A BROKK robot, equipped with a gripper, being used for a pick and place task at the Sellafield nuclear site in UK. Human operator can be seen controlling the robot from behind a 1.6m thick lead-glass window which shields him from radiation, but significantly limits his situational awareness. For more examples, refer to www.sellafieldsites.com/solution/decommissioning.

retrofit proprioceptive sensors to the robots used in such environments: firstly, electronics are vulnerable to different types of radiation; secondly, the installation of new sensors on trusted machinery would compromise long-standing certification; thirdly, such robots are predominantly deployed on a mobile base platform (e.g. a rugged tracked vehicle) and have tasks often involving high-force interactions with surrounding objects and surfaces. Even if the robot had proprioceptive sensors, such high-force tools cause large and frequent perturbations to the robot's base frame, so that proprioceptive sensors would still be unable to obtain the robot's pose with respect to a task frame set in the robot's surroundings.

Recently, many efforts have been made to deploy tele-operated robots at nuclear disaster sites [5, 6]; with robots controlled by viewing through cameras mounted on or around the robot. Albeit the significant efforts made, overall throughput rates are deficient in tackling the real-world problems. Nevertheless, in the context of this chapter, the major difficulty is situational awareness while tele-operation, especially the lack of depth perception and effect of external disturbances on the operator (e.g. surrounding noise levels, temperature, etc.), which primarily questions accuracy and repeatability of the task being performed [7]. Also, since the legacy waste inventory that needs processing is astronomical, direct tele-operation by humans is time consuming and tedious. Up on noticing all these difficulties associated with direct tele-operation of robots in hazardous environments, it can be seen that many of the nuclear decommissioning tasks can be (semi-)automated up to an extent to improve the task completion time as well as performance.

A major helping block is the computer vision or vision sensing. 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. Moreover, by adopting external sensing (e.g. vision) not only provides quantitative feedback to control the robot manipulator but also enables to estimate its joint configuration effectively when the proprioception is absent [8]. Machine vision systems are already being used for a

wide variety of industrial processes [9], 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 [10, 11]. In the case of nuclear applications, previous studies have used vision information to classify nuclear waste [3] and to estimate radiation levels [12]. However, it is not known that any visual servoing techniques (using the tracked image information) been applied in the (highly conservative) nuclear domain. Nevertheless, we believe that a greater understanding of the underlying processes is necessary, before nuclear manipulation tasks can be safely automated.

This chapter mainly discusses 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. In this context, two (common decommissioning) tasks are investigated: (1) point-to-point dexterity task, where human subjects control the position and orientation of the robot end-effector to reach a set of predefined goal positions in a specific order, and (2) box encapsulation task (manipulating waste items into safe storage containers), where human tele-operation of the robot arm is compared to that of a human-supervised semi-autonomous control exploiting computer vision. Human subjects' performance in executing both these tasks is analysed and factors affecting the performance are discussed. In addition, a vision-guided framework to estimate the robot's full joint state configuration by tracking several links of the robot in monocular camera images is presented. The main goal of this framework is to resolve the problem of estimating robot kinematics (operating in nuclear environments) and to enable automatic control of the robot in Cartesian space.

2. Analysis of tele-operation to semi-autonomy

The role played by robots in accomplishing various tasks in hazardous environments has been greatly appreciated; mainly for preventing the humans from extreme radioactive dosage [13, 14]. As previously stated, for many years they have been used to manipulate vast amount of complex radioactive loads and contaminated waste. Despite this fact, with the growing needs and technological developments, more new and advanced robotic systems continue to be deployed. This not only signifies the task importance but also questions the ability of human workers in operating them. Most of these robots used for decommissioning tasks are majorly tele-operated with almost no autonomy or even a pre-programmed motion as in other industries (e.g. automotive). Invariably there is regular human intervention for ensuring the environment is safely secured from any unsupervised or unplanned interactions. Most of this process is not going to change; however, some tasks in this process can be semi-automated to reduce the burden on human operators as well as the task completion time. In this section, we focus on analysing various factors affecting the performance of fully supervised tele-operated handling and vision-guided semi-autonomous manipulations.

2.1. Tele-operation systems

Tele-operation systems have been in existence as an ideal master-slave system or a client-server system. Many tele-operation based tasks have been used and their importance

http://dx.doi.org/10.5772/intechopen.69739

specifically in cases of human collaborative tasks has been widely studied [14]. These studies help in providing the importance to the human inputs and provide them with distinctive role to lead in more supervised tasks. In most of the cases the human acts as a master controlling or coordinating the movement to be handled by pressing buttons or by varying controller keys (e.g. joystick) and the robot acts as a slave (model) executing the commanded trajectories [15]. A typical example of such a tele-operation system using the joystick can be seen in **Figure 2**, where an operator uses the joystick as a guiding tool to move the robot arm or performing the orientation correction of the robot by viewing it from the live camera feed (multiple views of the scene). The robot then follows the instructions as advised and reaches the commanded position.

In an advanced tele-operation set-up, the operator feels the amount of force applied or the distance in which the gripper is opening and closing to grasp an object. Further, there are possibilities to include a haptic interface controller with specific force field induced while nearing an environmental constraint. These types of systems mainly assist operators in having full control of the tele-operators, when controlling them in congested environments. However, these systems are not fully exploited for performing nuclear manipulations and still most of the nuclear decommissioning tasks are executed using joysticks. In a more traditional model, the MSM device needs the human to apply forces directly. The major challenge associated is that no error handling is included with such systems (either MSM or joystick) and instead it is the task of human operator to correct any positional errors by perceiving the motion in a camera-display, which often induces task delays.

2.2. Tele-operated tasks for nuclear decommissioning

In the context of nuclear decommissioning, two commonly tele-operated (core-robotic) tasks are analysed: positioning and stacking. Maintaining the coherence with real nuclear scenario, these two tasks are simulated in our lab environment and are detailed below.

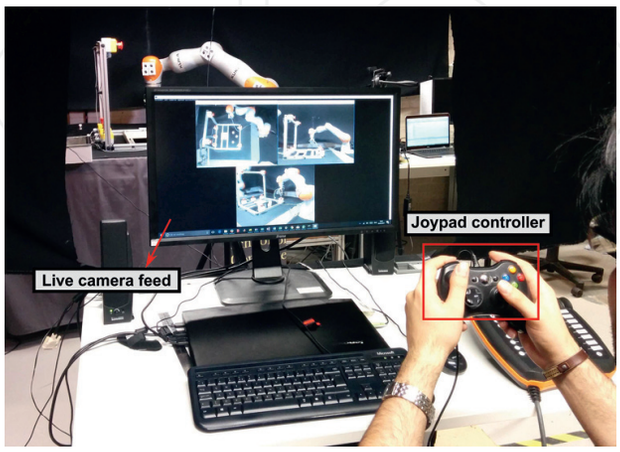


Figure 2. A human operator controlling the motion of an articulated industrial manipulator using joystick. The robot motion has been corrected by viewing it in the live camera feed.

2.2.1. Task 1: sequential positioning

This is one of the initial and majorly performed tasks, where an operator is required to manoeuvre the robot arm (end-effector) from a point to another in a specific order. While performing this task, the operators are required to control the positioning errors only from passive vision, i.e. by viewing at multiple camera displays. A special purpose tele-operation testbed containing multiple buttons has been designed to study this task (can be seen in **Figure 2**). For the sake of analysing human performance, multiple participants with almost none or limited robotic knowledge are recruited following specific criteria (explained in Section2.2.3). Each participant has been asked to move the robot end-effector from point-to-point in a designed order and while doing so, multiple parameters have been recorded (explained below) in order to analyse the task performance in terms of various effecting factors. Three specific points of action are chosen (buttons on the testbed) based on the kinematic configuration and to challenge the manipulability of the operator. Furthermore, a 'beep' sound is introduced to these three points such as to indicate the operator upon successful point-to-point positioning and completion of the task. The same three points are chosen for all the trials in order to evaluate the operator's performance over the course of repetitions. Each participant has been asked to repeat the task four times, where two trials are made in the presence of loud industrial noise such as moving machines, vibrations, etc. This has been done to analyse the operator performance in case of external environmental disturbances.

2.2.2. Task 2: object stacking

Stacking classified objects in an order into the containers is one of the vital tasks performed in the frame of decommissioning. Here it is assumed that objects (contaminated waste) have been classified beforehand and hence, waste classification process is not explained in this chapter. In general, stacking concurrently includes positioning and grasping. Underlying goal is to get hold of the (classified) objects positioned in the arbitrary locations. Operator while tele-operating has to identify the stable grasping location including collisions with the environment and has to stack the grasped object in a specific location or inside a bin. Similar to the previous task, this task has been designed to use passive vision and operator is able to control both robot and gripper movements from the joypad. However, since positioning and grasping collectively can be automated up to an extent, the task performance by direct human tele-operation is compared with a semi-autonomous vision-guided system (explained in next section). For analysing, three wooden cubes of size $4 \times 4 \times 4$ cm^3 are used as sample objects. In order to allow fair comparison, the objects to be stacked are positioned in the same locations for all trials and similar experimental conditions are maintained.

2.2.3. Data acquisition for performance analysis

Multiple factors are evaluated to analyse the human operators' performance and workload in accomplishing above mentioned tasks. A total of 10 participants (eight male and two female) were recruited with no prior hands-on experience or knowledge about the experimental set-up. All the participants had a normal to corrected vision. Previously developed software [16] has been used to interface the robot motion control with a gaming joypad, which allows the

operator to switch between and jog the robot in different frames (joint, base and tool) as well as to control the attached tool, i.e. a two finger parallel jaw gripper. An initial training was provided at the beginning of each task for each participant, focussing on detailing the safety measures as well as to get accustomed with the experimental scenario. Since passive vision has been used (emulating the real nuclear decommissioning environment), it was also necessary to ensure that the participants understand different camera views. Finally, the analysis has been performed by evaluating the following measures:

- *Observed measures*: These are intended to evaluate the operators' performance and are purely based on the data recorded during each task. The following factors are identified to estimate individual performance: success rate per task, task completion time and errors per task. These are detailed in **Table 1**.
- *Measures obtained from self-assessment*: These are intended to evaluate the operators' workload. To this purpose, *NASA Task Load Index* (NASA-TLX) forms are provided to each participant upon task completion, which involves questioning the user to rate their effort in terms of workload demand. The following measures are obtained from the completed forms: mental demand, physical demand, temporal demand, performance, effort, and frustration. The participant evaluates his/her individual performance in each task based on the influence or impact of the task and their individual comfort in pursuing them using the robot.

2.3. Semi-autonomous systems

Semi-autonomous systems are another prominent robot based approaches used for manipulation tasks in nuclear decommissioning [14]. The concept of semi-autonomy is quite similar to the tele-operation but with even more less effort or input from the human. The role of human in a semi-autonomous system is still a Master but handling only the supervisory part, i.e. initialising and monitoring. The operator gives the orders or decides the course of action to be performed by the robot, which are then executed by the system in a seemingly effortless response. For instance, human can identify the path to be followed by the robot and define that by means of an interface, which will be then executed by the robot. In some cases, the human operator can even define actions like grasping, cutting, or cleaning, etc. The system only needs the input in order to execute from the human instead of moving the entire robot like as in the previous case and in addition, since human being master can take over the control

Measure	Description
Total trials	Total number of repetitions in each task
Success rate per task	$\frac{\text{Total trials}-\text{Collisions}-\text{Perceptual misses}}{\text{Total trials}}$
Task completion time	$\frac{\sum \text{Elapsed time (completed trials)}}{\text{Completed trials}}$
Errors per task	$\frac{\text{Collisions}+\text{Perceptual misses}}{\text{Total trials}}$

Table 1. Observed measures to analyse operators' performance.

at any point of time. Most of the semi-autonomous systems rely on the external sensory information (e.g. vision, force, etc.) of the environment. The use of vision based input to manipulate the tasks and to progress through the environment has been proven effective in many cases [17]. Using visual information as a feedback to control robotic devices is commonly termed as *visual servoing* and is classified based on the type of visual features used [18]. For the sake of analysing the performance of a semi-autonomous system as well to compare the human performance in case of stacking objects, a simple position-based visual servoing scheme has been developed as in Ref. [16] to automatically manoeuvre the robot to a desired grasping location and to stack objects. A trivial visual control law has been used in combination with a model-based 3D pose matching [19]. It is always possible to use a different tracking methodology and to optimise the visual controller in many aspects. Readers can find more details about this optimisation process in Ref. [20].

2.3.1. Stacking objects by visual servoing

Overall task is decomposed into two different modules: grasping and stacking, where the former involves automatic navigation of the robot to a stable grasp location, and the latter involves placing the grasped objects at a pre-defined location. It is assumed that the object dimensions are always available from the knowledge database and the vision system is pre-calibrated. The task of automatic navigation starts by operator selection of the object, i.e. by providing an initial pose (e.g. with mouse clicks) and can be accomplished by: tracking full (six DoF) pose of the objects and by commanding the robot to a pre-grasp location using this information. The optional pre-grasp location is required only when the camera is mounted on top of the robot end-effector in order to avoid any blackspots for the vision while the robot approaches the object. This location has to be selected such that the robot can always maintain a stable grasp by moving vertically downwards without colliding with any other objects in its task space. It is worth noticing that the operator possess full control of this process by visualising the task as well as robot trajectory. **Figure 3** shows different tracked poses of an object during this process of automatic navigation to pre-grasp location.

Once the object is stably grasped, the task of stacking will be initiated automatically. During this phase the system uses its knowledge of the location to stack, i.e. the location to release the

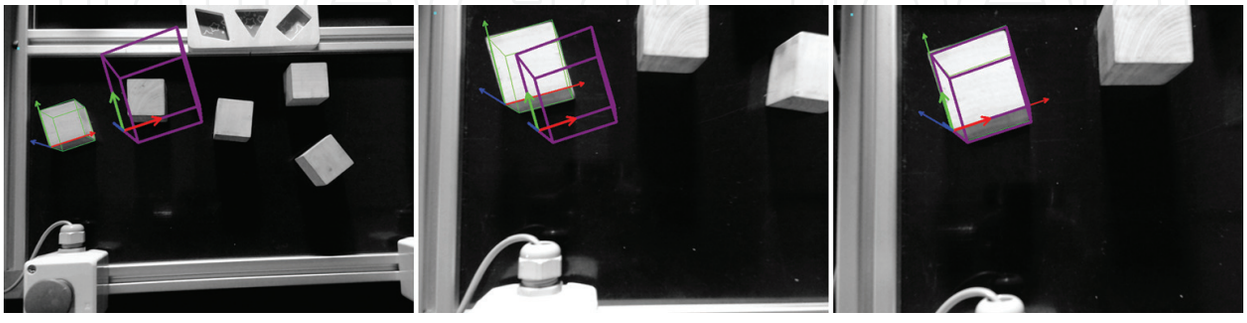


Figure 3. Series of images obtained during vision-guided navigation to grasp object 1 (first wooden cube). The wireframe in space represents the desired or initialised pose provided by the human and the wireframe on object represents the pose tracked during various iterations. Visual servoing goal is to minimise the error between those two poses such that they match with each other.

object, the number of objects already stacked and the dimensions of object being handled. Once the object is released or stacked, the robot will return to a defined home position and waits for the next human input.

2.3.2. Task analysis

Analogous to tele-operation, the task performance has been analysed by monitoring various factors such as collisions, success rate and task completion time. The robot has been commanded to stack all three blocks 10 different times and during which it only uses the images from on-board camera. In order to achieve good performances especially when using artificial vision, environmental lighting has been made seemingly stable throughout the task, which is also the case for tele-operation.

3. Vision-guided articulated robot state estimation

The overall goal of this section is to increase the operating functionality of under-sensored robots that are used in hazardous environments. As mentioned, most of the heavy-duty industrial manipulators that operate in hazardous environments do not possess joint encoders, which make it hard to automate various tasks. An inability to automate tasks by computer control of robot motions, not only means that task performances are sub-optimal, but also that humans are being exposed to high risks in hazardous environments. Moreover, during the execution of tasks, robots must interact with contact surfaces, which are typically unknown a-priori, so that some directions of motion are kinematically constrained. Our premise is that adopting external sensing, which is remote from the robot (e.g. vision using remote camera views of the robot) offers an effective means of quantitative feedback of the robot's joint configuration and pose with respect to the scene. Note that cameras can be rad-hardened in various well-known ways, and even simple distance of a remote sensor, away from the radiation source, greatly reduces impact on electronics via inverse-square law. Vision-based proprioceptive feedback can enable advanced trajectory control and increased autonomy in such applications. This can help remove humans from harm, improve operational safety, improve task performance, and reduce maintenance costs [16].

3.1. Related work

Vision information is used as the backbone of this concept and solely relying on which, the entire robot joint configuration is derived. Usually, this can be misinterpreted with classical visual servoing methods, where the robots are visually controlled using the information obtained from proprioceptive sensors. Visual servoing literature predominantly relies on accurately knowing robot states derived from joint encoders. However, Marchand et al. [21]. demonstrated an eye-to-hand visual servoing scheme to control a robot with no proprioceptive sensors. In order to compute the Jacobian of the manipulator, they estimate the robot configuration. Thus, they feed the end-effector position to an inverse kinematics algorithm for the

non-redundant manipulator. In Ref. [22], a model-based tracker was presented to track and estimate the configuration as well as the pose of an articulated object.

Alongside visual servoing, pose estimation is also related to this section. Pose estimation is classically defined for single-body rigid objects, with six DoF. On the other hand, articulated objects are composed of multiple rigid bodies and possess higher DoF (often redundant). There are also a number of kinematic (and potentially dynamic) constraints that bind together the bodies belonging to kinematic chains. Further, these constraints can also be used to locate and track the chain of robot parts. A variety of ways to track articulated bodies can be found in Refs. [23, 24]. These authors mainly focused on localising parts of the articulated bodies in each image frame, and not on the estimation of joint angles between the connected parts. Additionally, much of this work focussed on tracking parts of robots, but made use of information from the robot's joint encoders to do so, in contrast to the problem posed. A real-time system to track multiple articulated objects using RGB-D and joint encoder information is presented in Ref. [25]. A similar approach was used in Ref. [26] to track and estimate the pose of a robot manipulator. Some other notable examples can be found in Ref. [27], where the authors propose to use depth information for better tracking of objects. Recently, an approach based on regression forests has been proposed to directly estimate joint angles using single depth images in Ref. [28]. However, most of these methods require either posterior information (e.g. post-processing of entire image sequences offline to best-fit a set of object poses), or require depth images, or must be implemented on a GPU to achieve online tracking.

In summary, the use of depth information alongside standard images can improve the tracking performances. However, it also increases the computational burden and decreases robustness in many real-world applications. Our choice of using simple, monocular 2D cameras is motivated by cost, robustness to real-world conditions, and also in an attempt to be as computationally fast as possible.

3.2. Chained method to estimate robot configuration

Similar to the semi-autonomous task from Section2.3.1, a CAD model-based tracker based on *virtual visual servoing* is used to track and identify the poses of various links of the robot. We also assume that:

- the robot is always in a defined home position before initialising the task, i.e. its initial configuration is known; and
- the robot's kinematic model is available.

In turn the tracked poses are related with various transformations to estimate the entire robot configuration. There are two ways to relate camera to each tracked part, a direct path whose relationship is given by the tracking algorithm, and another path using the kinematic model of the robot. These two paths kinematically coincide, thus we enforce the following equalities to estimate the state of the robot:

$$^{C}\mathbf{M}_{obj_i} = {}^{C}\boldsymbol{T}_0 \, {}^{0}\boldsymbol{T}_{obj_i}(\boldsymbol{q}) \tag{1}$$

http://dx.doi.org/10.5772/intechopen.69739

Where, ${}^{C}\mathbf{M}_{obj_i}$ is the homogenous transformation from camera to object frame, ${}^{C}\boldsymbol{T}_0$ is the transformation from camera to world frame and ${}^{0}\boldsymbol{T}_{obj_i}(\boldsymbol{q})$ represents the transformation from world to object i frame parametrised over the joint values $\boldsymbol{q}$, i.e. ${}^{0}\boldsymbol{T}_{obj_i}(\boldsymbol{q})$ embeds the kinematic model of the robot. We track four different links of the robot as shown in **Figure 4(b)**. Therefore, for each tracked robot part, we get:

$$ {}^{C}\mathbf{M}_{obj_1} = {}^{C}\boldsymbol{T}_0\ {}^{0}\boldsymbol{T}_1(\boldsymbol{q}_1)\ {}^{1}\boldsymbol{T}_{obj_1} \tag{2} $$

$$ {}^{C}\mathbf{M}_{obj_2} = {}^{C}\boldsymbol{T}_0\ {}^{0}\boldsymbol{T}_1(\boldsymbol{q}_1)\ {}^{1}\boldsymbol{T}_2(\boldsymbol{q}_2)\ {}^{2}\boldsymbol{T}_{obj_2} \tag{3} $$

$$ {}^{C}\mathbf{M}_{obj_3} = {}^{C}\boldsymbol{T}_0\ {}^{0}\boldsymbol{T}_1(\boldsymbol{q}_1)\ {}^{1}\boldsymbol{T}_2(\boldsymbol{q}_2)\ {}^{2}\boldsymbol{T}_3(\boldsymbol{q}_3)\ {}^{3}\boldsymbol{T}_4(\boldsymbol{q}_4)\ {}^{4}\boldsymbol{T}_{obj_3} \tag{4} $$

$$ {}^{C}\mathbf{M}_{obj_4} = {}^{C}\boldsymbol{T}_0\ {}^{0}\boldsymbol{T}_1(\boldsymbol{q}_1)\ {}^{1}\boldsymbol{T}_2(\boldsymbol{q}_2)\ {}^{2}\boldsymbol{T}_3(\boldsymbol{q}_3)\ {}^{3}\boldsymbol{T}_4(\boldsymbol{q}_4)\ {}^{4}\boldsymbol{T}_5(\boldsymbol{q}_5)\ {}^{5}\boldsymbol{T}_6(\boldsymbol{q}_6)\ {}^{6}\boldsymbol{T}_{obj_4} \tag{5} $$

The state of the robot is estimated by imposing the equality given in the previous equations, and casting them as an optimisation problem. Since the robot's initial configuration is known, it is used as a seed for the first iteration of the optimisation problem and the robot's kinematic model is used to compute ${}^{0}\boldsymbol{T}_{obj_i}(\boldsymbol{q})$. The optimisation problem is then stated as:

$$ \underset{\boldsymbol{q}}{\text{minimise}} \sum\nolimits_i \boldsymbol{e}_i(\boldsymbol{q}) \text{ Subject to } |\boldsymbol{q}_i| < \boldsymbol{q}_{max} \tag{6} $$

Where,

$$ \boldsymbol{e}_i(\boldsymbol{q}) = vec\Big({}^{C}\mathbf{M}_{obj_i} - {}^{C}\boldsymbol{T}_0\ {}^{0}\boldsymbol{T}_{obj_i}(\boldsymbol{q}) \Big) \tag{7} $$

represents an error in the difference of the two paths shown in **Figure 4(a)** to define a transformation matrix from the camera frame to the tracked objects frames, and $\boldsymbol{q}_{max}$ is the joint limit. **Figure 4(a)** also depicts the overall estimation schema. The trackers return a set of matrices, i.e. one for each tracked part. The sets of equations coming from each of the four ${}^{C}\boldsymbol{M}_{obj_i}$ can be used in series to solve for subsets of joint variables, which can be called the 'chained' method. From **Figure 4(a)**, the following dependencies can be observed for each tracked object:

- first object's position obj_1RF depends only on q_1;
- second object's position obj_2RF on q_1 and q_2 ;
- third object's position obj_3RF on q_1, q_2, q_3 and q_4 ;
- finally, fourth object's position obj_4RF depends on all the six joints.

As pointed in **Figure 4(c)**, two cylindrical and two cuboid-shaped parts of a KUKA KR5 sixx robot are tracked for proof of principle. However, this choice is not a limitation, and a variety of different parts could be chosen. Nevertheless, the parts must be selected such that they provide sufficient information about all joints of the robot. Even though the concerned robot possesses proprioceptive sensors, they are not used in the estimation schema.

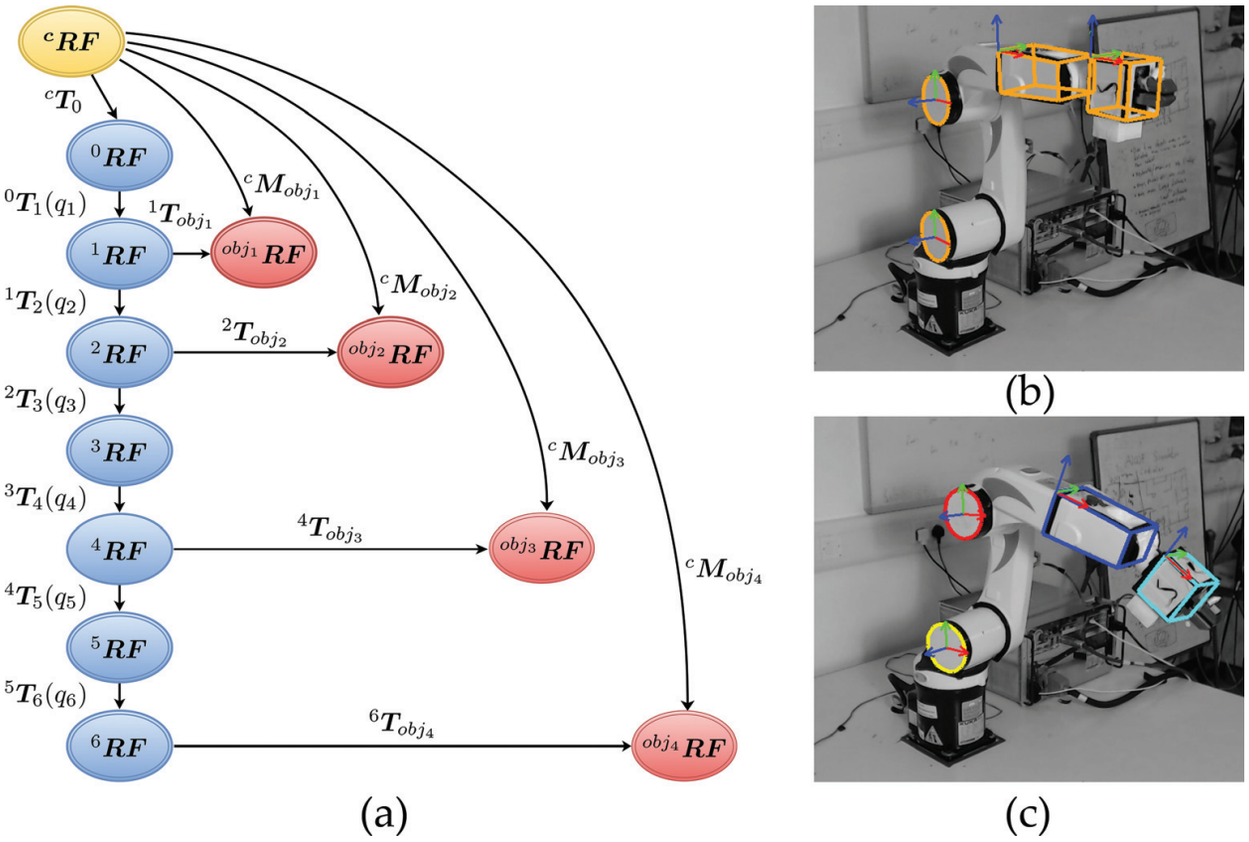


Figure 4. Illustration of the estimation work and tracking. (a) Shows the proposed state estimation model. Nodes represent reference frames and are classified top node represents camera frame, left aligned nodes represent robot frames and the right distributed nodes represent tracked object frames. The two paths leading to each tracked object frame from the camera reference frame can be seen. (b) Sequence of robot links to track and (c) tracked links in a later frame.

The chained method uses each object to estimate only a subset of joint values. These, in turn, are used as known parameters in the successive estimation problems. For example, q_3 and q_4 can be retrieved using obj_3, as in:

$$\underset{q_3\,q_4}{\text{minimise}}\ \boldsymbol{e}_3(q_1\,q_2\,q_3\,q_4)\ \text{Subject to}\ |\boldsymbol{q}_j| < \boldsymbol{q}_{max},\ j = 3.4 \tag{8}$$

In the similar fashion, other joints, i.e. q_1, q_2, q_5 and q_6 can be estimated using Eq. (6). Using only one object at a time, the quality of configuration estimation becomes highly dependent on the tracking performance for each individual part. Although it induces the advantage of being robust to single part tracking failure (producing outliers that influence the estimation of only the relative subset of angles), it adds the disadvantage of propagating the possible error of already estimated angles in subsequent estimations.

4. Experimental studies

Two different sets of experiments are conducted to validate human factor performance (explained in Section2) and to evaluate the vision-guided state estimation scheme (explained

in Section3). It is worth noting that all the experiments reported are conducted in the frame of hazardous environments. For the first set of experiments, an industrial collaborative robot KUKA lbr iiwa 14 r820 with seven DoF is used, to which a Schunk PG-70 parallel jaw gripper is connected as a tool. On the other hand, the second set of experiments is conducted using an industrial low-payload robot KUKA KR5 Sixx with six DoF. The commercial Logitech c920 cameras have been used for both experiments: for the first set of tests it is mounted on the tool and is used only for semi-autonomous tasks whereas for the second set, the camera is placed inside the workspace such that all the robot links are visible to accomplish the task. In either case, same work computer has been used and the communication between: robot and PC is realised over Ethernet, gripper and PC is realised over serial port, and camera and PC is realised over USB. ViSP library [29] has been used for the purpose of fast math computation and scene visualisation.

4.1. Analysing human factor performance

Recalling, various measures are identified to evaluate the human performance in performing different tasks. Later, operator performance is evaluated with a semi-autonomous system. In this context, first the experimental results evaluating 10 novice participants (eight male and two female) of age 30 ± 5.5 (mean $\pm \sigma$) performing the two tasks are reported. The performance of each participant is analysed and evaluated based on both the observed and self-reported measures as explained in the Section2.2.3.

4.1.1. Observed measures analysis

The observed measures of each participant have been categorized based on the time taken by the participant to fulfil the task, success rate in achieving it and the performance over the number of trials.

4.1.1.1. Sequential positioning: point-to-point dexterity task

The task was to push and release the buttons (upon hearing a beep sound) in a sequential order by jogging the robot. If the robot's tool collided with any object in the environment or if there is a perceptual miss in the target, the trial was considered as a failure. In total there are four trials for a participant. In order to increase the challenges in the task as well as to replicate a real industrial scenario, the final two trials are conducted with an audio track of industrial environment. The noise in the audio comprises of multiple tracks with continuous machinery sounds and intermittent sounds like welding, clamping, etc. **Figure 5(a)** and **(b)**, illustrates the average time taken by all the participants in reaching the desired points, i.e. in pushing the three buttons. The minimum and maximum values of all the participants are also indicated in the **Figure 2**. The influence of noise is also observed and the averaged time taken by each participant in pushing the three buttons is shown in **Figure 5(b)**. The normalised success rate computed among all participants in case of the first two trials (without noise) is 0.95 and for the latter two trials (with noise) is 0.87. Upon observing the results, it can be seen that there is a considerable amount of effect by the environmental noise on human operator in accomplishing the task. Mainly, from operators' experience, it has been found that the intermittent noise distracted their attention from the task and consequently led to reduced performance.

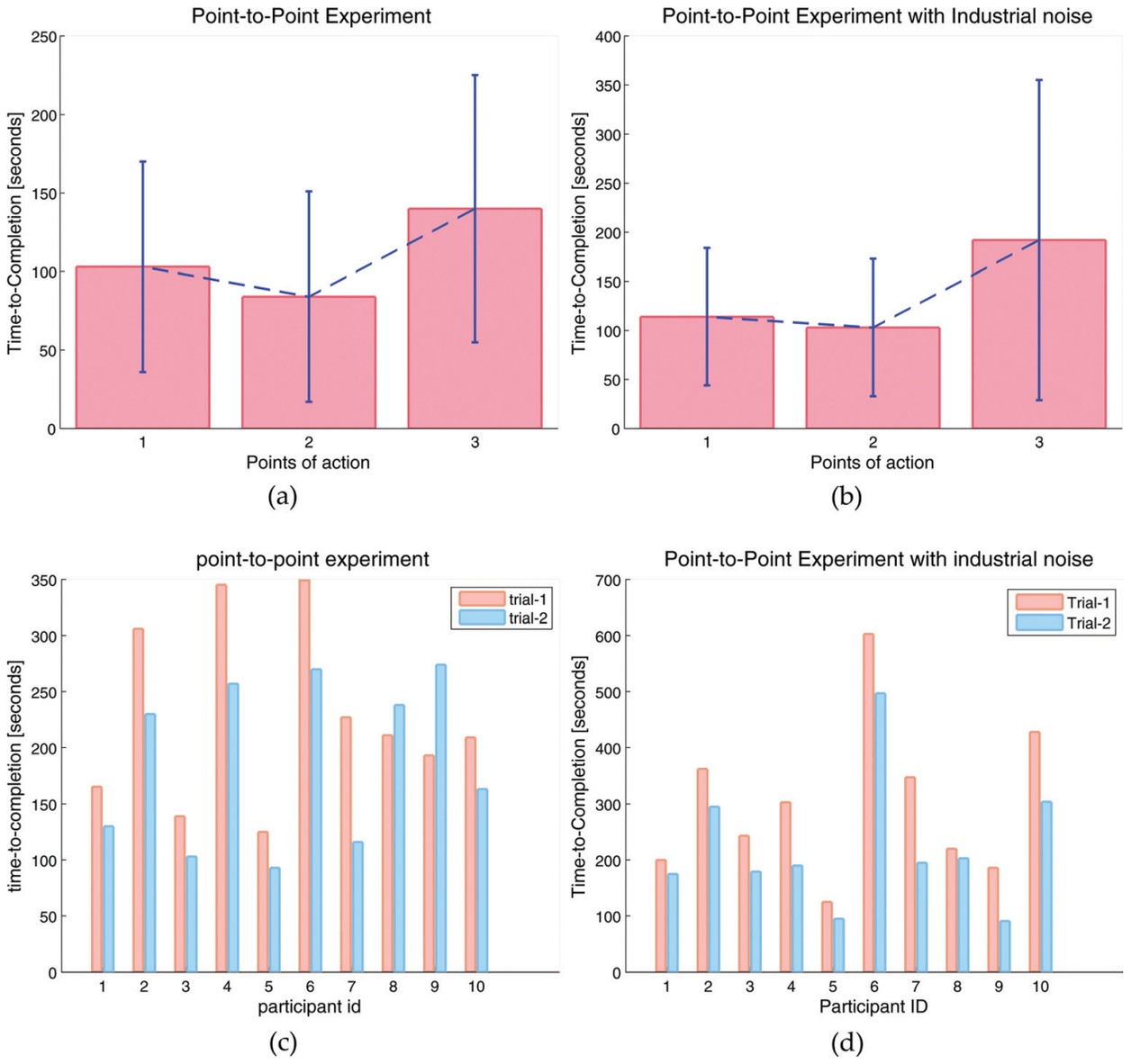


Figure 5. Illustration of the human performance in accomplishing point-to-point task over different trials at multiple conditions, i.e. in the absence and presence of environmental noise. (a) and (b) shows the average time spent in reaching multiple points of action on the test rig without (first two trials) and with (latter two trials) industrial noise, respectively. Since the location of button-3 is quite challenging to reach, participants spent more time with this point. (c) and (d) shows participants' individual time taken among all trials without and with noise, respectively. Effect of noise on task performance is clearly evident.

On the other hand, operators' learning over the tasks is also analysed. **Figure 5(c)** and **(d)** shows the time to completion for each participant over multiple trials. The repeated measure analysis of variance (ANOVA) was used on the time-to-completion data in order to evaluate whether the performance significantly changed or not across the trials. ANOVA revealed that there was a significant learning effect across all four trials. From manual observation, for example, consider participants eight and nine whose performance reduced from trial-1 to trial-2. However, as a proof of learning, the same participants' performance improved over the next two trials (even in the presence of noise). Interestingly, individual performance is

http://dx.doi.org/10.5772/intechopen.69739

affected by the task complexity, which is noticeable by analysing the time-to-completion in **Figure 5(a)** and **(b)**. All the participants find it significantly easy to reach the points of action 1 and 2, where the robot end-effector is required to be pointing down, i.e. normal to the ground plane. Task learning is clearly visible among these two trials. However, the performance reduced (even after learning) while approaching the third point of action, where the robot end-effector needs to be positioned parallel to the ground plane, which requires operator's intelligence in solving the robot's inverse kinematics so as to move appropriate joints in accomplishing the task.

4.1.1.2. Object stacking task

This task is to stack multiple objects at a defined location by controlling the robot motion. Similar to the previous, if the robot's tool collided with any object in the environment or if there is a perceptual miss in the target or if the object is not stably grasped or if the object is not successfully stacked; the trial was considered as a failure. In total, there are three trials for each participant. Since this task has been compared with semi-autonomy, all the trials are performed at noisy conditions. **Figure 6(a)** shows the average time taken among the trials by the participants in accomplishing the task and **Figure 6(b)** illustrates the individual performance. The normalised success rate computed among all participants is 0.74, which is comparatively less than the previous task. Also due to task complexity, the rate of perceptual misses observed were higher (mainly due to passive vision), specifically while positioning initial block at the specified location. Similar to the previous, ANOVA has been used to identify the learning among trials. Even though it returned significant learning behaviour, visually it can be seen that only 6 out of 10 participants (one partial) improved over trials. Also, it has been observed that many participants struggled matching or registering the camera views, which led to environmental collisions and thus leading to task failures or

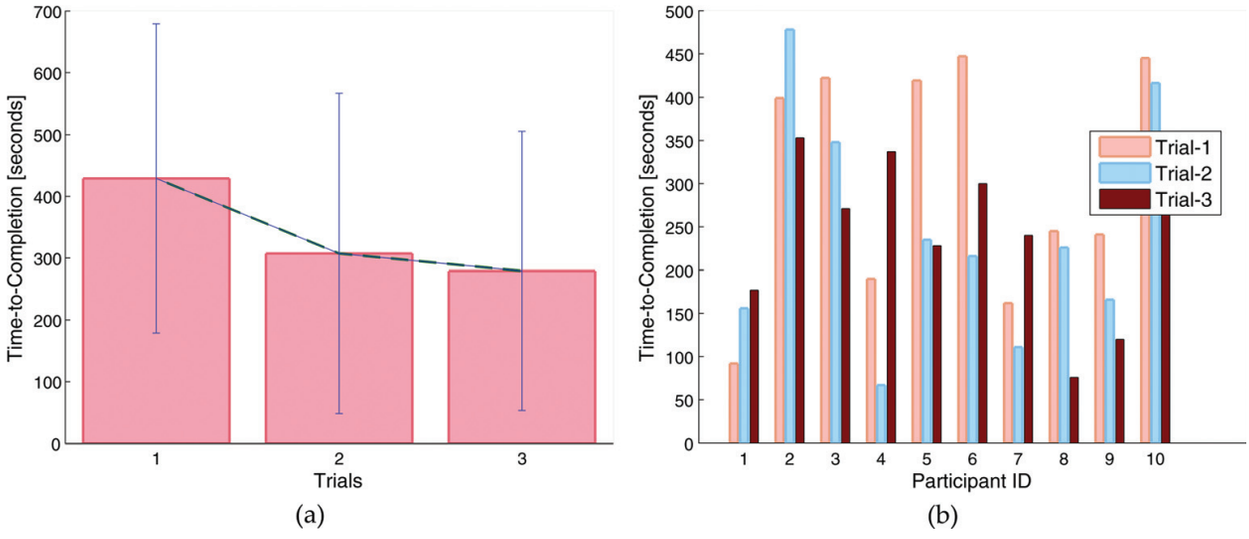


Figure 6. Illustration of the human performance in accomplishing stacking task over multiple trials. (a) Shows the average time taken among three different trials. Significant learning behaviour can be seen. (b) Illustrates the participants' individual task performance.

delays. These results clearly suggest the difficulties a human operator face in accomplishing a systemised task and therefore, the need for automation.

4.1.2. Self-assessed measures analysis

NASA TLX model was used as a base analysis for the self-reported measures. Each participant evaluated the task based on the following criterions: Mental demand, Physical demand, Temporal demand, Performance, Effort, Frustration. Two new parameters were also considered for task 1, i.e. the audio and video stress. **Table 2** and **Figure 7** report the results for both the tasks.

It is evident from the results, that all the participants found the task two to be more demanding. The Mental demands are significantly high for both the tasks, when compared with other

Self-reported measures	**Task 1**	**Task 2**
Mental demand	57.5 ± 21.24	71.5 ± 21.22
Physical demand	53.5 ± 32.14	60 ± 19.86
Temporal demand	57.5 ± 20.85	58 ± 20.30
Performance	40.5 ± 25.10	45.5 ± 18.63
Effort	50.5 ± 24.43	71.5 ± 16.17
Frustration	6.9 ± 4.33	10.6 ± 4.32
Total work load	**294 ± 29.08**	**359.5 ± 19.63**
Influence of audio	57 ± 23.47	-
Influence of video	55 ± 20	-

Table 2. Self-assessed measures for both the tasks.

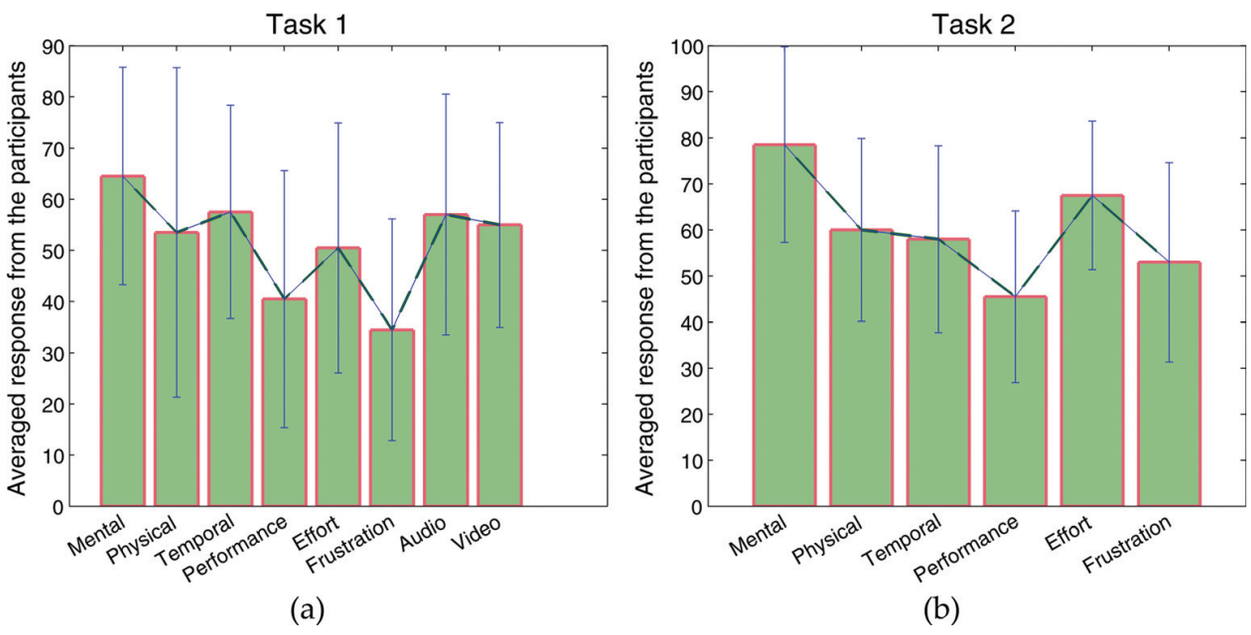


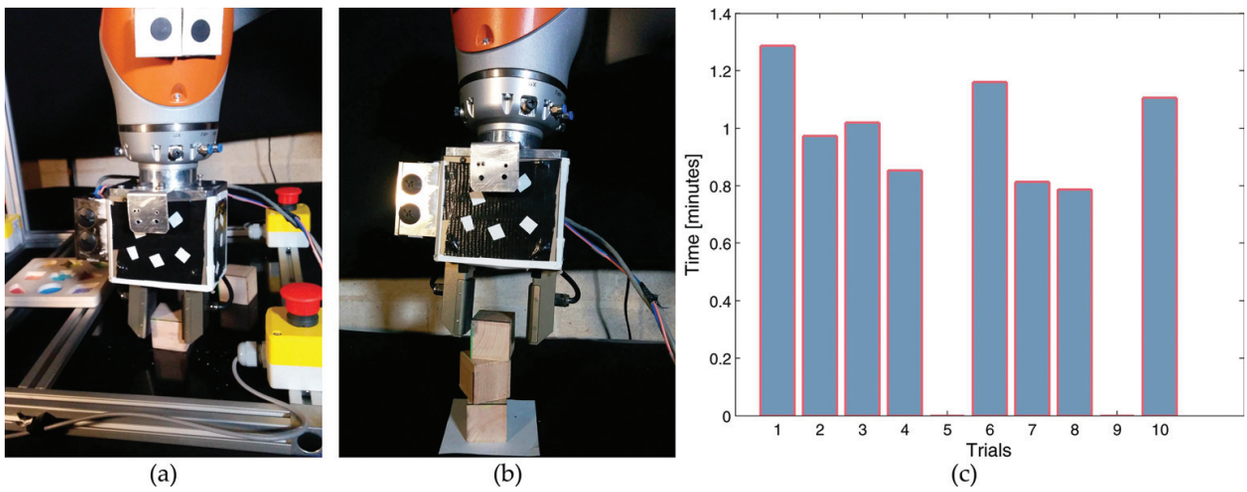
Figure 7. Bar charts illustrating the averaged performance of the participants using self-reported measure, NASA TLK for (a) task 1 and (b) task 2. An additional audio and video impact was evaluated specifically for task 1.

http://dx.doi.org/10.5772/intechopen.69739

sub-scales. This seems to reflect that the tasks required participants to construct the 3D 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. Besides, Task 2 requires more precise movements in handling the objects, grasping them in a suitable position such that they can be stacked one above the other. The task complexity and the lack of experience in using the robotic tools resulted in the participants feeling the impact. On the contrary, the physical demands and frustration were relatively low, suggesting that the tele-manipulation could reduce the physical tiredness for such repetitive tasks. 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. In addition, it can be seen again from **Figure 7(a)** the effect of surrounding audio and video live feed on human operator.

4.1.3. Analysis of semi-autonomous block stacking

These set of experiments are conducted to compare and evaluate the performance of a semi-autonomous system (explained in Section2.3.1). As mentioned before, this task consists of 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 initial poses. Then the robot has been automatically navigated to the pre-grasp pose, which is accomplished by regulating the positional error. **Figure 8(a)** and **(b)** show respectively the robot grasping first object and the final stacked objects. **Figure 8(c)** shows the time taken to stack three objects over 10 trials. On average the system requires 49.3s to stack three objects, which is almost times faster than the time taken by a human operator in accomplishing the same. Similar to the direct human tele-operation, this task has also been monitored for collisions and failures. Even though no collisions are observed, the task failed during 5th and 9th



(a) (b) (c)

Figure 8. Illustration of semi-autonomous system results. (a) Robot grasping the initial object in the task. (b) Final stack of three objects in the specified location (white square area). (c) Overall time taken for semi-automated block stacking during 10 trials.

trials due to tracking error. Hence, the overall performance directly depends on the success of visual tracking system. 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 system performance.

4.2. Robot state estimation results

Two series of experiments are conducted. Firstly, the precision of the implemented chained method in estimating the robot configuration is assessed by commanding the robot in a trajectory where all joints are excited. During which the vision-estimated joint angles are compared to the ground-truth values obtained by reading positional encoders. Next, the vision-derived robot's configuration estimates are used in a kinematic control loop to demonstrate the efficiency in performing Cartesian regulation tasks. For this purpose, a classical kinematic controller of the form given by Eq. (9) has been implemented.

$$\dot{q}_{ref} = J^{\dagger}(q)\left(K_p e\right) - \left(K_D \dot{q}\right) \tag{9}$$

where, $\dot{q}_{ref}$ is the desired/reference velocity, and $J^{\dagger}(q)$ is the pseudo-inverse of the the robot Jacobian computed using our estimated joint configuration. K_p and K_D are proportional and derivative gain matrices, respectively. Since the robot is controlled in positional mode, Eq. (9) is integrated numerically to generate control commands.

4.2.1. Estimating robot's configuration by chained method

Figure 9(a) shows the arbitrarily chosen trajectory to analyse the estimation efficiency. Since only one camera is used to track the robot, the trajectory has been chosen such that the entire

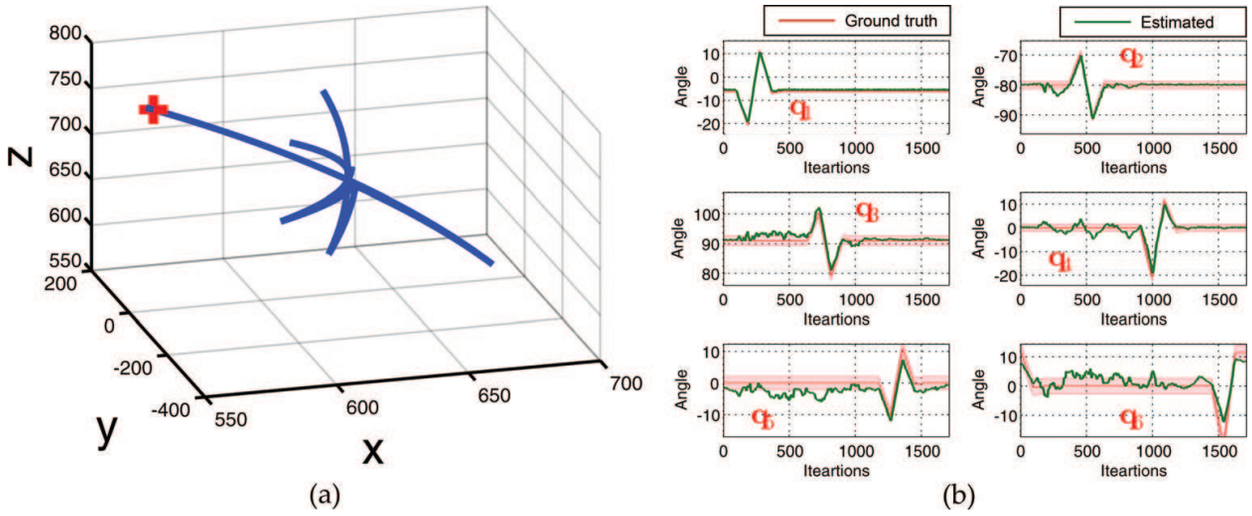


Figure 9. (a) Selected trajectory to evaluate robot state estimation. It has been chosen such that all the joints of the robot are excited. (b) Estimated and ground-truth joint angles during the trajectory (trial 3). Angles are expressed in degrees.

http://dx.doi.org/10.5772/intechopen.69739

tracked robot's links are visible throughout the trajectory. In order to perform the quantitative analysis, the robot has been commanded to execute the trajectory five times, repeatedly. The estimated and ground-truth values during the trajectory (for third trial) are shown in **Figure 9(b)**.

Joint	RMSE	Std.
q_1	0.581	0.269
q_2	0.801	0.777
q_3	1.372	1.054
q_4	1.508	1.504
q_5	2.512	1.634
q_6	4.106	2.636

Table 3. Performance analysis of the developed state estimation schema.

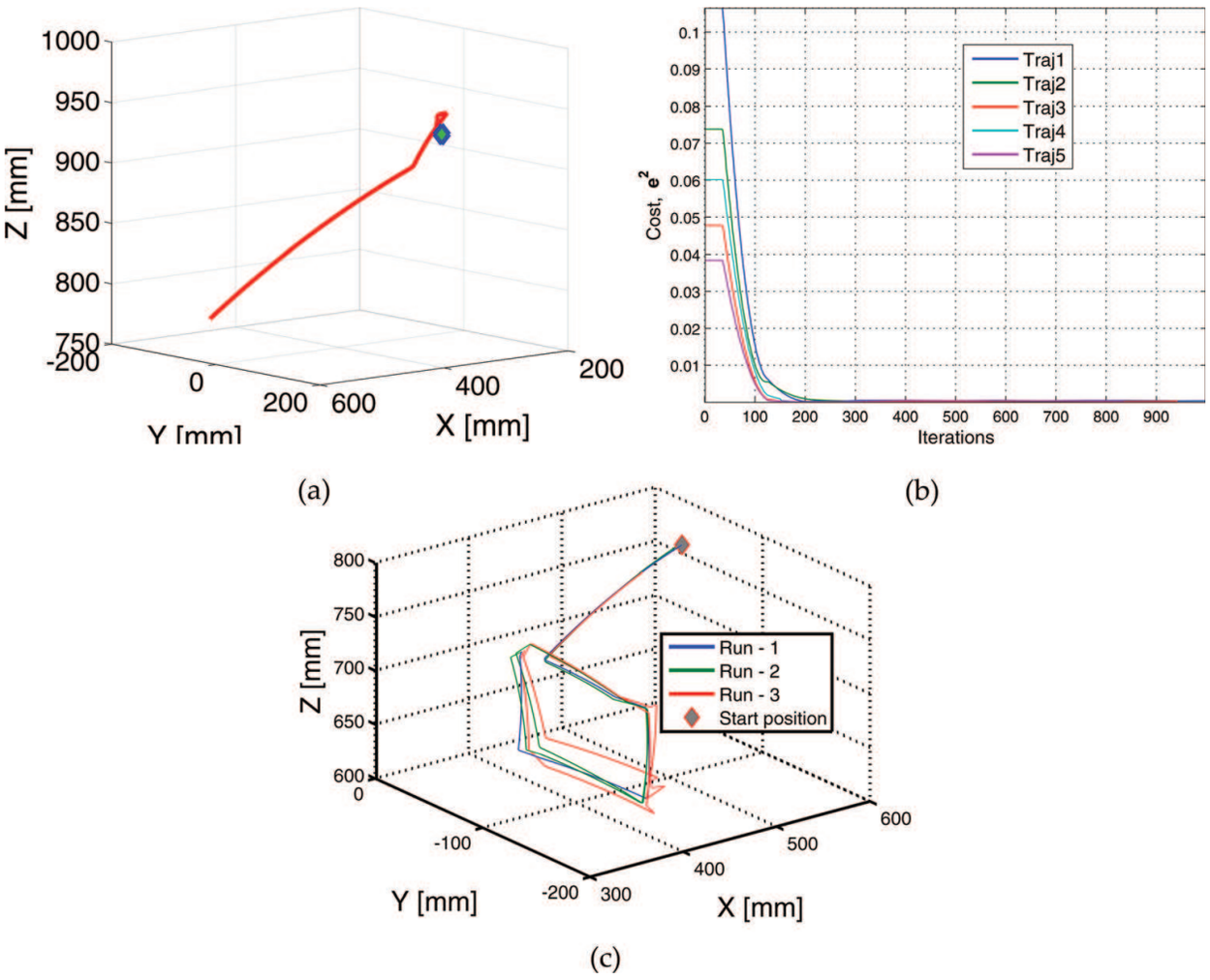


Figure 10. (a) Trajectory followed by the end-effector while reaching first goal position. (b) Evolution of controller costs during all five trajectories. (c) Square-perimeter trajectories followed by the end-effector. Diamond marker represents robot's starting position.

The average RMSE and standard deviation values over all trials are summarised in **Table 3**. On average, the estimation error is less than 4˚, which clearly demonstrates the efficiency of the method in estimating robot's configuration through vision.

4.2.2. Cartesian regulation with vision estimates

Two different experiments are conducted. First, the robot end-effector has to be positioned automatically to five different goal positions using vision estimates and second, the robot was required to move its end-effector along a trajectory tracing out the perimeter of a square. Square corner locations in robot world frame are supplied as targets. **Figure 10(b)** shows the controller [given in Eq. (9)] cost variations while positioning to the five goal positions and **Figure 10(c)** shows the square trajectory followed by the robot in three different runs. These results clearly demonstrate the robustness of the method.

5. Conclusion

This chapter investigated two different concepts in the scope of hazardous environments. At the first hand, human performance was evaluated in executing remote manipulation tasks by tele-operating a robot in the context of nuclear decommissioning. Two commonly performed tasks are studied using which, various measures are analysed to identify the human performance and workload. Later, the human subject performance has been compared with a semi-autonomous system. The experimental results obtained by simulating the tasks at a lab environment demonstrate that the human performance improves with training, and suggest how training requirements scale with task complexity. They also demonstrate how the incorporation of autonomous robot control methods can reduce workload for human operators, while improving task completion time, repeatability and precision. On the other hand, a vision-guided state estimation framework has been presented to estimate the configuration of an under-sensored robot through the use of a single monocular camera. This mainly helps in automating the currently used heavy-duty industrial manipulators.

Author details

Naresh Marturi[1,2*], Alireza Rastegarpanah[1], Vijaykumar Rajasekaran[1], Valerio Ortenzi[1], Yasemin Bekiroglu[1], Jeffrey Kuo[3] and Rustam Stolkin[1]

*Address all correspondence to: nareshmarturi@kuka-robotics.co.uk

1 Extreme Robotics Lab, University of Birmingham, Edgbastion, UK

2 Kuka Robotics UK Ltd., Great Western Street, Wednesbury, UK

3 National Nuclear Laboratory (NNL) Ltd., Birchwood Park, Warrington, UK

References

[1] Nuclear Energy Institute. Nuclear energy around the world [Internet]. [Updated: 2017]. Available from: https://www.nei.org/Knowledge-Center/Nuclear-Statistics/World-Statistics [Accessed: 19 February 2017]

[2] Nuclear Decommissioning Authority. Nuclear Provision: Explaining the Cost of Cleaning Up Britain's Nuclear Legacy [Internet]. 11 February 2015 [Updated: 1 September 2016]. Available from: https://www.gov.uk/ [Accessed: 23 September 2016]

[3] Shaukat A, Gao Y, Kuo JA, Bowen BA, Mort PE. Visual classification of waste material for nuclear decommissioning. Robotics and Autonomous Systems. 2016;**75**:365-378

[4] Jean V. editor. Teleoperation and Robotics: Applications and Technology. 3rd ed. Netherlands: Springer; 1985. 256 p. DOI: 10.1007/978-94-011-6103-9

[5] Nagatani K, Kiribayashi S, Okada Y, Otake K, Yoshida K, Tadokoro S. et al. Emergency response to the nuclear accident at the Fukushima Daiichi nuclear power plants using mobile rescue robots. Journal of Field Robotics. 2013;**30**(1):44-63

[6] Qian K, Song A, Bao J, Zhang H. Small teleoperated robot for nuclear radiation and chemical leak detection. International Journal of Advanced Robotic Systems. 2012;**9**(3):70-79

[7] Iqbal J, Tahir AM, ul Islam R. Robotics for nuclear power plants—Challenges and future perspectives. In: 2nd International Conference on Applied Robotics for the Power Industry (CARPI). IEEE; Zurich, Switzerland. 2012. pp. 151-156

[8] Andreff N, Martinet P. Vision-based self-calibration and control of parallel kinematic mechanisms without proprioceptive sensing. Intelligent Service Robotics. 2009;**2**(2): 71-80

[9] Jorge LCS, editor. Advances in Machine Vision. 1st ed. New York: Springer-Verlag; 1989. p. 420. DOI: 10.1007/978-1-4612-4532-2

[10] Marturi N, Ortenzi V, Xiao J, Adjigble M, Stolkin R, Leonardis A. A real-time tracking and optimised gaze control for a redundant humanoid robot head. In: IEEE-RAS 15th International Conference on Humanoid Robots. IEEE; Seoul, South Korea. 2015. pp. 467-474

[11] Marturi N, Tamadazte B, Dembele S, Piat N. Visual servoing schemes for automatic nanopositioning under scanning electron microscope. In: IEEE International Conference on Robotics and Automation. IEEE; Hong Kong, China. 2014. pp. 981-986

[12] Baek C, An SJ, Kim H, Kwak SW, Chung YH. Development of a pinhole gamma camera for environmental monitoring. Radiation Measurements. 2013;**59**:114-118

[13] Russell RA, Thiel D, Deveza R, Mackay-Sim A. A robotic system to locate hazardous chemical leaks. In: IEEE International Conference on Robotics and Automation. IEEE; Nagoya, Aichi, Japan. 1995. pp. 556-561

[14] Gelhaus FE, Roman HT. Robot applications in nuclear power plants. Progress in Nuclear Energy. 1990;**23**(1):1-33

[15] Bogue R. Robots in the nuclear industry: A review of technologies and applications. Industrial Robot: An International Journal. 2011;**38**(2):113-118

[16] Marturi N, Rastegarpanah A, Takahashi C, Stolkin R, Kuo J, Bekiroglu Y. Towards advanced robotic manipulation for nuclear decommissioning: A pilot study on tele-operation and autonomy. In: IEEE International Conference on Robotics and Automation for Humanitarian Applications. IEEE; Kerala, India. 2016; pp. 1-8

[17] Ortenzi V, Marturi N, Stolkin R, Kuo J, Mistry M. Vision-guided state estimation and control of robotic manipulators which lack proprioceptive sensors. In: IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE; Daejeon, South Korea. 2016. p. 3567-3574

[18] Chaumette F, Hutchinson S. Visual servo control. I. Basic approaches. IEEE Robotics and Automation Magazine. 2016;**13**(4):82-90

[19] Dembele S, Piat N, Marturi N, Tamadazte B. Gluing free assembly of an advanced 3D structure using visual servoing. In: 23rd Micromechanics and Microsystems Europe Workshop; Ilmenau, Germany. 2012: pp. 6-12

[20] Malis E. Improving vision-based control using efficient second-order minimization techniques. In: IEEE International Conference on Robotics and Automation. IEEE; New Orleans, LA, USA. 2004: pp. 1843-1848

[21] Marchand E, Chaumette F, Spindler F, Perrier M. Controlling an uninstrumented manipulator by visual servoing. The International Journal of Robotics Research. 2002;**21**(7): 635-647

[22] Nickels K, Hutchinson S. Model-based tracking of complex articulated objects. IEEE Transactions on Robotics and Automation. 2001;**17**(1):28-36

[23] Ma M, Marturi N, Li Y, Stolkin R, Leonardis A. A local-global coupled-layer puppet model for robust online human pose tracking. Computer Vision and Image Understanding. 2016;**153**:163-178. DOI: http://doi.org/10.1016/j.cviu.2016.08.010

[24] Pellegrini S, Schindler K, Nardi D. A generalisation of the ICP algorithm for articulated bodies. In: Proceedings of the British Machine Vision Conference; BMVA Press; Leeds, UK. 2008. p. 87.1-87.10. DOI: 10.5244/C.22.87

[25] Pauwels K, Ivan V, Ros E, Vijayakumar S. Real-time object pose recognition and tracking with an imprecisely calibrated moving RGB-D camera. In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE; Chicago, Illinois, USA. 2014: pp. 2733-2740

[26] Klingensmith M, Galluzzo T, Dellin CM, Kazemi M, Bagnell JA, Pollard N. Closed-loop servoing using real-time markerless arm tracking. In: IEEE International Conference on Robotics and Automation (Humanoids Workshop). Karlsruhe, Germany. 2013

[27] Bohg J, Romero J, Herzog A, Schaal S. Robot arm pose estimation through pixel-wise part classification. In: IEEE International Conference on Robotics and Automation (ICRA). IEEE; Hong Kong, China. 2014: pp. 3143-3150

[28] Widmaier F, Kappler D, Schaal S, Bohg J. Robot arm pose estimation by pixel-wise regression of joint angles. In: IEEE International Conference on Robotics and Automation. IEEE; Stockholm, Sweden. 2016. pp. 616-623

[29] Marchand E, Spindler F, Chaumette F. ViSP for visual servoing: A generic software platform with a wide class of robot control skills. IEEE Robotics and Automation Magazine. 2005;**12**(4):40-52

IntechOpen

IntechOpen

IntechOpen