OPEN ACCESS

University of BRISTOL

Gregg-Smith, A., & Mayol-Cuevas, W. W. (2016). Investigating spatial guidance for a cooperative handheld robot. In 2016 IEEE International Conference on Robotics and Automation (ICRA16): Proceedings of a meeting held in Stockholm, Sweden, 16-21 May 2016. (pp. 3367-3374). Institute of Electrical and Electronics Engineers (IEEE). DOI: 10.1109/ICRA.2016.7487512

Peer reviewed version

Link to published version (if available):
10.1109/ICRA.2016.7487512

Link to publication record in Explore Bristol Research
PDF-document

This is the author accepted manuscript (AAM). The final published version (version of record) is available online via IEEE at http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7487512. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available:
http://www.bristol.ac.uk/pure/about/ebr-terms.html

Investigating Spatial Guidance For a Cooperative Handheld Robot

Austin Gregg-Smith and Walterio W. Mayol-Cuevas

***Abstract*— In this paper we address the question of how to provide feedback information to guide users of a handheld robotic device when performing a spatial exploration task. We consider various feedback methods for communicating a five degree of freedom target pose to a user including a stereoscopic VR display, a monocular see-through AR display and a 2D screen as well as robot arm gesturing. The spatial exploration task with the handheld robot arm was compared against a baseline of a passive handheld wand. We compared the performance of each of the methods with 21 volunteers using a repeated measures ANOVA experimental design, and recorded users' opinions via a NASA Task Load Index Survey. The robot assisted reaching feedback methods significantly outperform manual reaching with the wand for all three visual feedback methods. However, there is little difference between each of the three visual feedback methods when using the robot. The completion time of the task varies with changing difficulty when using the wand but remains stable when assisted by the robot. These results convey useful information for the design of cooperative handheld robots.**

I. Introduction

Handheld robotics [1] aims to benefit from the mechanical precision, speed and extra degrees of freedom of a robot that has task knowledge in order to cooperatively solve tasks with users (figure 1).

Being able to simply pick up a handheld robot that knows where and what to do can empower inexperienced users and extend and maintain repeatability of experts. This novel type of cooperative robot calls for an investigation on how best to interact with the user to achieve this goal. In our previous work [1] we studied user interaction issues as different degrees of autonomy change, however the way in which the robot communicates instructions to the user was not considered. This is important for better bringing together the skills and features of both robot and user and to reduce frustration at points of divergent intentions.

When doing a job where the robot is tracking task progression, there are times when the robot needs to communicate with the user to move to another location to continue the task. For example, in an inspection task, the robot can guide the inspector towards the next inspection location once it confirms that the current area has been properly examined. In a cleaning application, if the user has missed cleaning a spot, the robot should tell the user to move the robot within range of the dirty surface so that the robot complete the cleaning. There are several ways in which the robot can transmit this information to the user, each of which has its own advantages, disadvantages, implementation complexity, saturation of cognitive channels, and cost. In this paper we explore four feedback method types in a spatial guidance task and analyse their relative performance through quantitative measures and user feedback.

Department of Computer Science, University of Bristol, UK
agreggsmith@gmail.com, wmayol@cs.bris.ac.uk

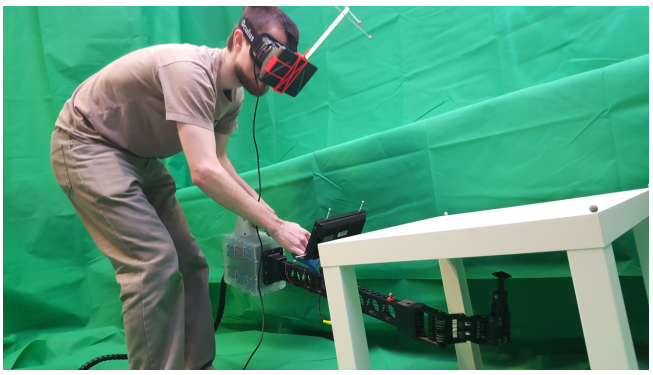


Fig. 1: A cooperative handheld robot assisting a user in a simulated inspection task using a virtual reality headset for feedback.

The following section reviews the related work in the areas of reaching and human computer interaction for spatial positioning tasks. Section III explains the spatial reaching task and IV covers the details of each of the feedback methods in more detail. Section V explains the experimental design and method. The results are explained in section VI and discussed further in section VII.

II. Related Work

Cooperative task solving with handheld robots is intrinsically different from that with robots that are not manually manipulated by humans. The proximity to the user and the fact that it is in what we may call the *tool space* makes it more closely related to studies on human factors. In the literature of human computer interaction, there are examples of studies of how people point or reach certain spatial locations. When discussing reaching, Fitts's law is a well known model for estimating the time to reach an area of known distance and size [2], [3] . It was originally formulated for 2D tasks and since been extended to 3D [4], [5].

Liang and Green[6] introduced a 5 degree of freedom (5-DoF) selection tool "laser gun" that is part of the

family of ray casting techniques. The user holds a 6-DoF input device and a ray is cast from the current pose of the user's hand in the direction they want. They can select object by moving their hand so the ray intersects with it. They found a 5x speed increase in CAD modelling time using this input method.

The ray casting technique has been extended to the "spotlight" method where the ray is expanded to a conic section to make it easier to select small objects. Another variation uses a cylinder instead of a cone, but both of these extensions still only use 5-DoF to select objects. A survey of 3D and 5D pointing techniques can be found in [7].

Zhai and Milgram [8] evaluated the performance of four manipulation schemes to match a 6-DoF pose in a virtual environment with eight subjects. They explored proportional and non-proportional velocity and position control and found that proportional position control was the best method for matching pose.

Ranasinghe et al [9] use haptic feedback to guide users in low visibility environments. The authors performed an experiment where a participant would guide a user along a path by pulling on a rigid link that they both held. Experimental systems identification was used to determine that the guider used a 3rd order autoregressive predictive control policy while the follower used a 2nd order reactive control policy. A controller was designed to match the human guider's control policy and implemented on a robot that guided users via the same rigid link. The control policy was designed to modulate the pulling force in response to the level of trust the user placed in the guidance of the robot.

Erden and Billard [10] studied the difference in end-point impedance between novice and expert users in a welding task. A welding torch was attached to a KUKA LWR robot and an admittance controller was used to counteract the mass of the tool so that it behaved as a free floating body. The users performed welding while the robot simulated random disturbances of the tool. Expert users provided more resistance to the disturbances than novice users. Erden and Billard suggests that measuring the impedance values that user provides could be used to estimate their skill in the task. The estimate of user ability could in turn be potentially factored into the level of guidance a handheld robot provides during a task.

The above examples are mostly constrained to a reduced operational space. One of the few examples of guiding users spatially around real objects is the work of Echtler *et al* [11] who designed and tested an *intelligent* welding gun for constructing experimental vehicles. The device aims to guide users to the weld locations by displaying visual instructions on a screen attached to the gun. They compared the accuracy and physical characteristics of several feedback methods including projector based augmented reality, head-mounted displays, and finally chose a tool-mounted 2D screen as most suitable for the task. Working within an optical tracking system with retro reflective markers, workers quadrupled the speed in which they completed the task using the augmented welding gun compared with the control group. In this paper, we explore cases where the tool has more degrees of freedom and can be more proactive in helping the user with the task. This calls for an investigation of the best ways to guide people under these different conditions.

III. Spatial Reaching Task

Several of the applications we consider suitable for handheld robots are tasks related to maintenance; inspection, cleaning and construction: e.g, non-destructive-testing, spot welding, bolting and drilling. These tasks have generic requirements such as having to place a probe or drill bit in the right location, spray a substance on the right spot or apply some pressure or friction in the right place. These tasks are all rotationally invariant along one axis and so only require 5-DoF positioning.

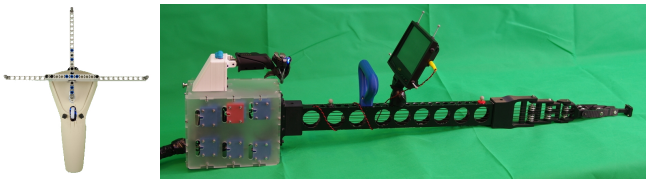


Fig. 2: The two devices used to reach 5D points on a table: A handheld wand on the left and a handheld 6-DoF robot.

We are thus exploring a generic 5-DoF reaching task where the user has to reach a series of target points around an object as quickly as possible while maintaining a specified desired accuracy.

In our study we use our 6 degree of freedom (6-DoF) cable-driven handheld robot with one redundant joint. Due to the redundant joint, the end effector has 5-DoF positioning capability so can perform tasks that are rotationally invariant. The robot is 1.19m long including the 0.3m 6-DoF actuator and weighs 3.5kg. This robot is offered as an open platform and further details are available at [12], [13].

We explored seven different modes of operation for a 5-DoF reaching task. In each task, a series of 10 random 5D targets on an Ikea side table (model LACK), were chosen and the user was instructed to reach those points as quickly as possible using either a handheld robot or a handheld wand used as a baseline alternative (figure 2). In each task, the user has to move the tip of the device (wand/robot) to within 5mm and 5° of the desired pose and hold it there for at least 200ms. This aims to emulate some of the required specifications for an inspection task.

We tested three screen-based feedback methods for both the robot and the wand. The information displayed

on the screen is a wire-frame schematic view of the table and the target point that needs to be reached on the table. Only the gesturing feedback mode does not use a screen to display the target information, but instead the robot points the end effector towards the target. This is a list of all the seven conditions considered:

1) The robot gestures by pointing towards the target point (figure 3).
2) A 7" screen is mounted on the robot which shows the wire-frame model of the table and where the target is (figure 4b).
3) The user holds the 7" screen in one hand and the wand in the other (figure 4c).
4) The user wears a see-through augmented reality display and uses the handheld robot (figure 5b).
5) The user wears a see-through augmented reality display and holds the wand in their hand (figure 5c).
6) The user wears the stereo virtual reality (VR) headset and uses the robot arm (figure 6b).
7) The user wears the stereo VR headset and holds the wand in their hand (figure 6c).

We use an eight camera, infra-red, optical tracking system (Optitrack Flex 3) with retro-reflective markers to give 6-DoF pose information for all of the objects in the task. The tracked objects are the robot arm, handheld wand, table, augmented reality headset, stereo VR, and 7" screen.

IV. Design of Feedback Methods

A. Hardware Gesturing

Gestures, such as pointing, are a natural interaction method widely used to communicate. During daily living tasks, gestures during cooperative tasks such as parking or moving large furniture in confined spaces are common. We wanted to explore this natural method of communication and see if it could be applied to the robot so that users would intuitively understand robots instructions without special training.

Robot gesturing is an attractive option as it does not require any extra hardware for feedback, thus reducing both the potential cost and complexity of the system. When the robot wants to indicate a new area of interest, the arm points toward the right location to indicate to the user where to move the robot's base.

1) Gesturing Algorithm: The robot indicates a target pose by pointing towards the position of the target. Once the end effector is in range of the target, it moves to match its 5-DoF pose. If there is no clear line of sight between the current pose of the robot and the desired pose (e.g. in the presence of obstacles), a path through free space must be found first.

The pose of the robot in world space has 6-DoF and the joint angles of the robot have another 6-DoF. To find a collision free path from the current 12-DoF pose to a valid 12-DoF target pose is currently too computationally expensive to solve dynamically in real time using standard methods. For high dimensional path planning tasks, the execution time is in the order of minutes [14], instead of the required sub-second times for interactive responses. Instead, we use the high-dimensional path-planning ability of the human operator to reduce the path-planning problem to 3-DoF. A human holding a tool is able to negotiate complex obstacles intuitively and so can aid the robot in a task which is otherwise an open problem.

Rapidly exploring random trees (RRT) are a family of sample based probabilistic path planners that are widely used in robotics [15]. Karaman et al [16] exploited their ability to find an initial solution quickly and then improve on that solution incrementally given more computation time.

We wrote a variant of informed RRT*[17] to solve the path-planning problem.The primary objective is to calculate the shortest route from the robot to the target point and then move the end effector to point along that path. For the type of task we are looking at, there are generally two important frames of reference: the item of interest and the robot's frame. In most inspection cases the item of interest will be stationary, while the user moves the robot around in free space. When building an RRT, the root of the tree is the current state of the problem, and the branches extend towards desired states. Given that the robot is held by the user the root of the tree will become invalid as soon as the user moves the robot. The item of interest will stay stationary most of the time so we choose to put the root of the tree there.

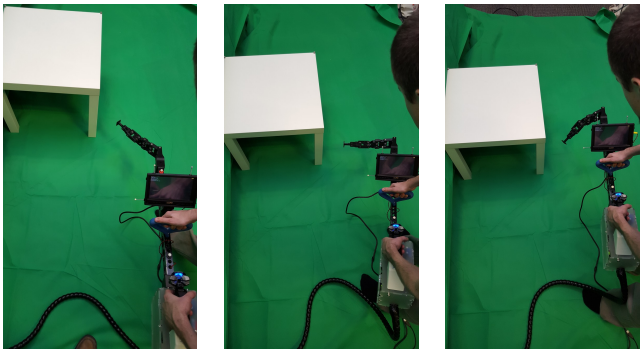
When choosing where to point, the robot solves a 3D path planning problem to find the shortest route from the robot's current position to the target position. The robot then searches for the furthest point along this 3D trajectory that has a 5-DoF inverse kinematics solution. This means that the robot end effector moves to a point on the trajectory linking the robot to the target position while pointing in the same direction as the trajectory at that point.

A disadvantage of reducing the problem to 3-DoF is that the robot will only attempt to minimise the 3D trajectory. In some cases this will return paths that are not optimal given the 12-DoF state space of the robot. However, users quickly notice when this is the case and are able to correctly move the robot to the optimal place.

Figure 3 shows the behaviour of the gesturing feedback. When the user moves the robot around it compensates and keeps the end effector pointing towards the target. It is not necessary to make these movements however as the user can move the robot directly along the line defined by the end effector's pose.

Note that the gesturing algorithm is used in all evaluations where the robot is used, regardless of the screen based feedback method. It ensures that the robot's

joints are configured in a way that allows the user to easily access the desired points because the robot will attempt to minimise the distance to the target while simultaneously avoiding obstacles in real time.



(a) (b) (c)

Fig. 3: Right to left: The robot points toward the same location while the user moves the robot around.

B. Graphical Feedback

The three visual feedback methods we investigated were 3D rendered graphics showing the current state and the desired state in a virtual world. An advantage of this approach is flexibility in the types of information one can transmit.

The desired end effector pose is shown as a white arrow on the display and the current position of the end effector is shown as a yellow arrow. These 3D objects are rotationally invariant about their axis so unambiguously describe the 5-DoF pose as shown in figure 7. Other information is shown in white with a black background so that it is visible on the see-through head-mounted display. The back background does not show up at all, and only the bright pixels can be seen.

We explore three different types of screen for giving feedback: a monocular optical see-through head-mounted display, a 7" LCD monitor, and a stereo virtual reality headset. The graphical information displayed on each device is the same regardless of the device but the inherent properties of the device and their proximity to the user are different.

1) Hand Held Display: Figure 4a shows the handheld display (HHD), a Liliput 7" LCD screen with a 16:9 aspect ratio and 854x480 pixels. The viewpoint rendered on the screen is not an augmented reality view in that the screen does not act like a virtual window the user can look though. This is due to the position of the display relative to the user, the field of view visible through the "window" would be very small and not be able to display enough relevant information. Instead, a virtual viewpoint is rendered with a field of view of 67.5°. Figure 4b shows a first person view of a user holding the robot with the display attached. Figure 4c shows the user holding the wand and display in one hand each.

(a) Liliput 7" display with tracking markers on the outer edge (b) Robot with handheld display (c) Wand and handheld display

Fig. 4: The 7" LCD Handheld display (HHD)

2) See-through Augmented Reality: The augmented reality head-mounted display (AR HMD) is a Trivsio 800x600 pixel LCD screen projected onto a half mirror that the user looks through (figure 5a). The user can see the real world through the half mirror, but sees computer generated images superimposed onto it. Both the AR HMD and HHD we used are non-stereoscopic display devices. For the user to estimate the pose and depth of objects they have to use parallax and shading information. Figures 5(b,c) show the user using the robot and wand with the AR HMD.

(a) Trivsio see-through HMD (b) Robot with see-through HMD (c) Wand and see-through HMD

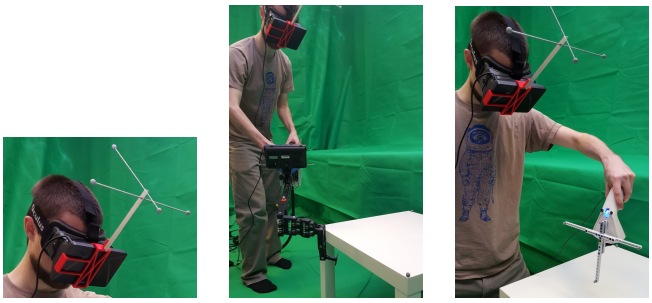
Fig. 5: See-through augmented reality display (AR HMD)

In order to display an image that correctly overlays virtual objects onto their real world counterparts we need to calculate the position of the user's eye relative to the display. To display an overlaid image a virtual camera is set to the position and field of view of the user's eye. When the image from this virtual camera is displayed on the half mirrored lens, virtual objects will appear to overlay the real ones.

As the position of the user's eye will change slightly every time they put the AR HMD on, we need an efficient calibration procedure to calculate the parameters for the virtual camera. The calibration procedure we used is similar to SPAAM [18] in that a single known 3D location is used as a reference point and the users move their head during calibration.

3) Stereoscopic VR Display (VR HMD): The virtual reality head-mounted display is an Oculus Rift development kit 1 with a resolution of 1280x800 pixels (640x800 per eye). This provides 3D stereo information to communicate the target's 5D pose information. Figure 6 shows a user completing the task while wearing the VR headset. Unlike augmented reality their eyes are completely covered by the VR HMD so they are not able

to see any objects in the world. In this task the floor, walls, table, robot and wand are rendered so users don't accidentally collide with anything in the real world.



(a) Oculus Rift virtual reality headset (b) Robot and Rift (c) Wand and Rift

Fig. 6: Oculus Rift stereoscopic display (VR HMD)

TABLE I: Summary of feedback methods

Method	Effective FOV	Resolution	Stereo	Set up time
Robot Gesturing	$\approx 170°$	N/A	yes	0s
Handheld Screen	67.5°	854x480	no	0s
See-through AR	$\approx 23°$	800x600	no	$\approx 120s$
Stereo VR	110°	1280x800	yes	$\approx 10s$

4) Further information and display summary: Each display has a finite field of view so they need to have a mechanism for indicating that the region of interest is outside the screen's display frustum. We use an approach found in flight simulators where arrow indicators track points of interest when they are out of the field of view of the cockpit. The 3D target position is projected into 2D screen coordinates. If the screen coordinates lie outside the target's viewable area, a chevron is drawn at the edge of the screen as close as possible to the target, as shown in figure 7. Table I shows a summary of the technical specifications of all four feedback methods.

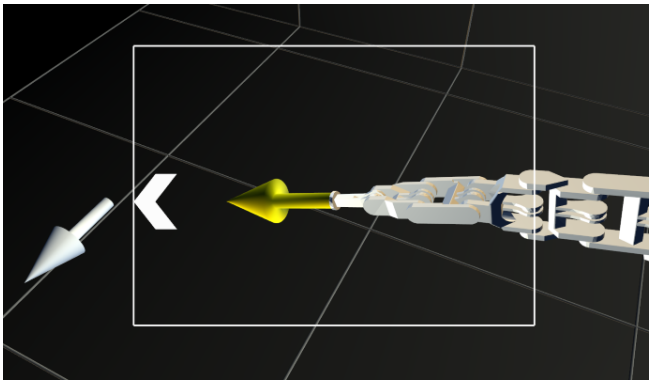


Fig. 7: An example of a 5 DoF target (white 3D arrow), including a chevron to indicate the target is outside the field of view of the display (represented by the white rectangle).

V. Experimental Method

We used a within subjects two-way repeated measures ANOVA experimental design. The independent variable was the method of feedback and the dependent variables were completion time, and NASA TLX scores.

We involved 21 Participants (8 female, 13 male) aged 21-56 in the experiment, with each participant taking on average one hour to complete all seven modes. One participant dropped out due to nausea problems from the Oculus Rift so their data was discarded. 10 out of 21 participants had corrected vision via glasses or contact lenses during the experiment and did not remove them while wearing the augmented reality or VR displays.

The experimental procedure is applied to each of the seven feedback methods in III.

A. Experimental Procedure

1) A random feedback method is chosen (robot gesture, robot handheld display, wand handheld display, robot augmented reality, wand augmented reality, robot VR, wand VR)
2) User is shown how the task works for that method:
   a) The user holds the device (robot/wand) in a neutral position.
   b) The word "next" is shown on the display.
   c) The user clicks a button to indicate they are ready to move to the next target point.
   d) A random point on the table is randomly selected and a timer is started.
   e) The user moves the robot/wand to match the target pose.
   f) The timer stops when the target pose is held for at least 200ms with 5mm positional accuracy and 5°angular error.
3) The user practices reaching ten target points per feedback method.
4) The user is recorded reaching another ten points.
5) The user is asked if she/he wants to rest.
6) The user is told they can make notes of their experience for each mode directly after each attempt, but they should only fill out the TLX survey once they have completed all the modes.
7) Another random feedback mode is chosen until all seven modes have been completed.
8) User fills out NASA TLX survey anonymously.

VI. Results

A within subjects repeated measures two-way ANOVA experimental design with the Greenhouse-Geisser correction was used. The independent variables were the inspection device and method of feedback. The dependent variables were completion time, and NASA TLX scores.

The device has a significant effect on the completion time ($F(1,20) = 168.8, p < 3.3.4 \times 10^{-11}, \eta^2 = 0.89$) and TLX scores ($F(1,20) = 58.7, p < 4.4 \times 10^{-7}, \eta^2 = 0.73$).

TABLE II: Bonferroni corrected ANOVA pairwise comparisons of completion times between feedback modes

	Robot Gesture	Robot HHD	Robot VR	Robot AR	Wand HHD	Wand VR	Wand AR
Robot Gesture	X	p<0.001	p<0.001	p<0.001	p<0.001	**p>0.05**	p<0.01
Robot HHD	p<0.001	X	**p>0.05**	**p>0.05**	p<0.001	p<0.001	p<0.001
Robot VR	p<0.001	**p>0.05**	X	**p>0.05**	p<0.001	p<0.001	p<0.001
Robot AR	p<0.001	**p>0.05**	**p>0.05**	X	p<0.001	p<0.001	p<0.001
Wand HHD	p<0.001	p<0.001	p<0.001	p<0.001	X	p<0.001	**p>0.05**
Wand VR	**p>0.05**	p<0.001	p<0.001	p<0.001	p<0.001	X	p<0.01
Wand AR	p<0.01	p<0.001	p<0.001	p<0.001	**p>0.05**	p<0.01	X

TABLE III: Bonferroni corrected ANOVA pairwise comparisons of NASA TXL scores between feedback modes

	Robot Gesture	Robot HHD	Robot VR	Robot AR	Wand HHD	Wand VR	Wand AR
Robot Gesture	X	**p>0.05**	p<0.01	**p>0.05**	p<0.01	**p>0.05**	**p>0.05**
Robot HHD	**p>0.05**	X	**p>0.05**	**p>0.05**	p<0.001	**p>0.05**	p<0.001
Robot VR	p<0.01	**p>0.05**	X	**p>0.05**	p<0.001	p<0.001	p<0.001
Robot AR	**p>0.05**	**p>0.05**	**p>0.05**	X	p<0.001	**p>0.05**	p<0.001
Wand HHD	p<0.01	p<0.001	p<0.001	p<0.001	X	p<0.01	**p>0.05**
Wand VR	**p>0.05**	**p>0.05**	p<0.001	**p>0.05**	p<0.01	X	p<0.05
Wand AR	**p>0.05**	p<0.001	p<0.001	p<0.001	**p>0.05**	p<0.05	X

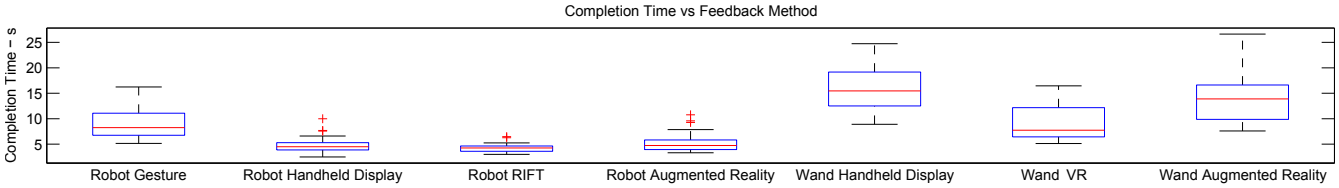


Fig. 8: Box and whisker plot of completion time

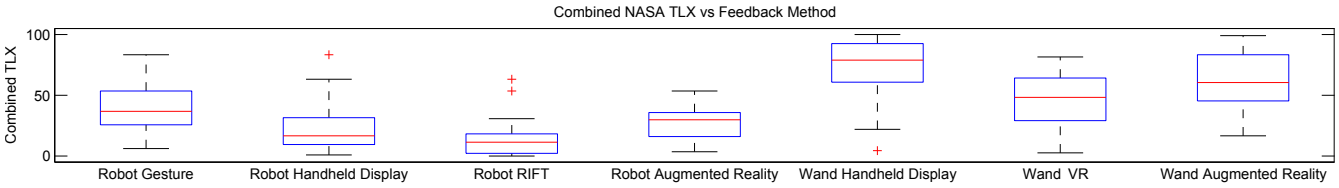


Fig. 9: Box and whisker plot of the combined NASA TLX results

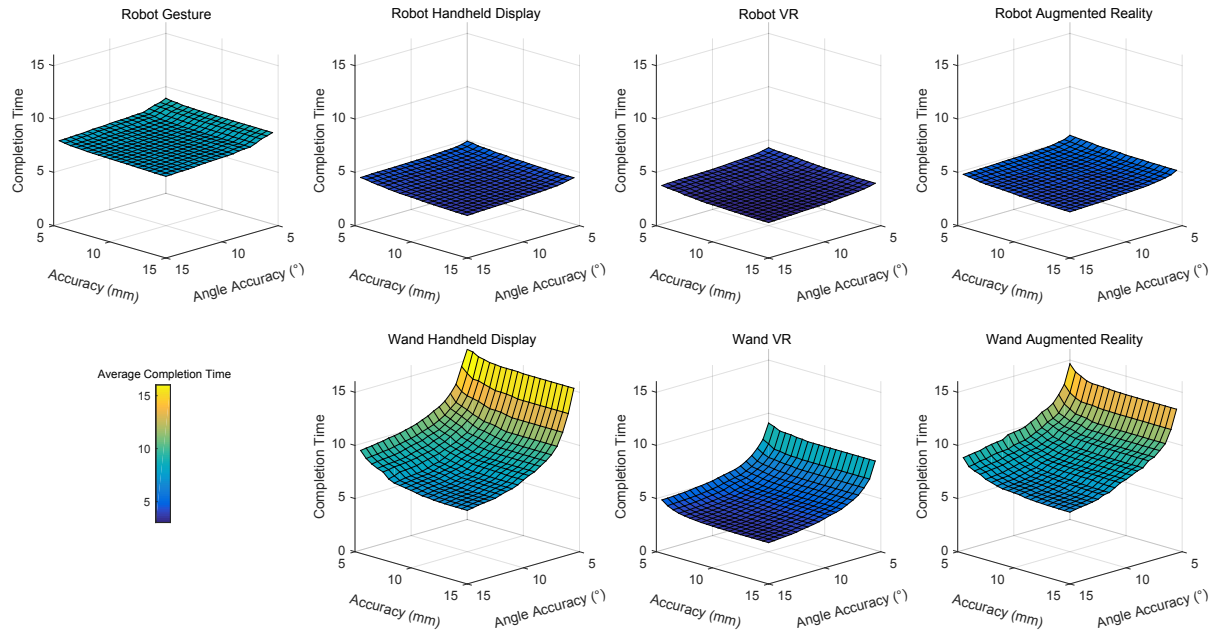


Fig. 10: Surface plot of completion time vs required positional and angular accuracy for the reaching task

The feedback mode has a significant effect on the the completion time ($F(1.64, 32.7) = 22.3, p < 3.0 \times 10^{-7}, \eta^2 = 0.53$) and TLX scores ($F(1.56, 31.4) = 11.5, p < 1.1 \times 10^{-4}, \eta^2 = 0.37$).

There is also a significant interaction between the device and feedback method for completion time ($F(1.76, 35.3) = 28.8, p < 1.8 \times 10^{-8}, \eta^2 = 0.59$) and TLX scores ($F(1.46, 29.3) = p.14, p < 0.002, \eta^2 = 0.31$).

Table II shows the pairwise comparisons of the seven test conditions results for task completion time. In order to reduce type 1 errors from repeated pairwise comparisons, the confidence intervals have been reduced using the Bonferroni correction. The statistics package reported the p values to three decimal places so the reason many values are quoted as $p < 0.001$ is because the value was too small to display. Pairwise comparisons where there is no significant difference between the means are highlighted in bold.

Table III shows pairwise comparisons of the ANOVA of the combined NASA TLX results using the Bonferroni correction. All of the insignificant results from the completion times table are also insignificant for the TLX results and the results that are insignificant for TLX but not the completion times are underlined.

Figure 8 is a box and whisker plot of the completion time for each of the seven modes and figure 9 is a box and whisker plot of the combined NASA TLX results.

We recorded the age, gender and average number of hours of video games that the participants played per week. There was no statistically significant correlation between any of these measures and the average performance over the trials. There was also no correlation between these measures and the combined TLX scores.

Figure 10 is a surface plot of the task completion time vs reaching accuracy. The point on the far back corner of each plot (5mm, 5°) is the accuracy that participants were asked to complete while the rest of each graph is generated from their recorded movements. That is, we simulate more relaxed accuracy requirements as explained below. However, note that the results in the rest of tables and figures relate to the strictest accuracy case only (5mm, 5°).

During the trial, a complete recording of the experimental variables were recorded including desired target information, 6-DoF wand and robot pose, robot joint angles and user button presses. This information was processed in real time during the trial to drive the robot path planning logic, visualisations and to calculate if the user had reached the desired pose with the required accuracy. After the experiments had been completed the resulting 147 recordings of the trials were re-analysed with lowered difficulty settings. This method of analysis works well because if the user was able to match the most difficult conditions, they must be able to reach the relaxed accuracy settings. Post-processing this information should give very similar results to actually running the trial with reduced difficulty settings because the user behaviour up until the reaching of the point is identical.

VII. Discussion

There is no significant difference in completion time between any of the visual feedback methods using the robot (table II), but there *is* a significant difference between all of the visual robot feedback modes and the robot gesture condition. The lack of effect of visual feedback mode on completion time when using the robot is an interesting result given that the visual feedback methods were so different. The differences in field of view and stereo information had no effect on task performance, indicating that users ascertain the approximate location of the target point from visual feedback but rely on the robot for accurate matching. One can theorise that the visual feedback could be of even worse quality, as long as it is enough for the user to put the robot in the general vicinity of the target. This result gives designers of handheld robots flexibility when choosing a feedback method.

The robot gesturing completion time was not significantly different to the wand VR performance but was significantly different to all the other modes, with gesturing performing the same or better than all the wand feedback methods.

When using the wand, there is no significant difference in completion time between the handheld display and augmented reality, but both perform significantly worse than when the VR HMD is used. This is probably due to the stereo information that gives a better representation of the desired poses than the 2D projections. Users also spent less time searching for the target when it was out of their field of view compared to the other feedback methods.

Figure 10 shows the robot performance is stable compared to the wand across varying accuracy requirements. The plots for the robot are relatively flat, and this is constant across all four robot modes. Even with the gesturing feedback where the mean time is significantly higher than for the other robot visual feedback methods, the overall surface still stays flat. This is in sharp contrast to the shape of the plots for the wand reaching, where the completion time increases quickly with greater required angular accuracy and less so with positional accuracy.

The trend of angular accuracy dominating the difficulty of the wand task stays constant across all three visual feedback modes. The overall difficulty is reduced for the VR HMD due to the presence of stereo depth information, but the richer visual feedback is not enough to overcome the physical difficulty of matching the required angle.

A curved difficulty response is an undesirable characteristic when performing a task as it reduces the predictability of the overall performance. Even when using the information-poor gesturing mode, the performance

is much more consistent than the best-wand reaching mode using the VR HMD.

Figure 8 and 9 show that the TLX results broadly match the completion times. The Pearson correlation coefficient between actual performance and the TLX results is 0.69 ($p < 0.0001$) which implies there is a strong correlation between them. However, the variance in user perception was larger than the completion time and the pairwise comparisons between TLX score showed fewer statistically significant differences.

The wand HHD has significantly worse TLX scores than all other modes except for the wand AR. The wand AR has no significant difference to the robot gesturing but is worse than all other methods. Both the HHD and AR feedback methods were difficult to use with the wand because of the lack of depth and the users found the task frustrating.

There is no significant difference in perception between the robot gesturing, robot HHD, robot AR, and wand VR. The stereo information in the wand VR condition provided enough information to the users so they found the task approximately as difficult as non-stereo feedback with robotic assistance.

There is no significant difference in TLX score between any of the robot visual feedback methods. Some users preferred the VR HMD because of the stereo information, while others found it made them motion sick. Similarly, some users found the HHD the most intuitive to use while others found looking back and forth between the HHD and task, frustrating.

In summary, handheld robots do not need sophisticated visual feedback as the cooperation between user and robot can be supported by minimal guidance and even in the absence of any screen. Simple gesturing still allows for consistent performance over a range of accuracy requirements.

VIII. Conclusions

In this paper we investigated how best to provide feedback information to guide users of a handheld robotic device when performing a spatial exploration task. Our study compares four feedback methods for communicating a five degree of freedom target pose to a user. We investigated three visual feedback methods and one that used mechanical gestures. Our robotic arm is held by the user to reach a target pose in space as fast and as accurately as possible. As a baseline, we also evaluated robot assisted reaching performance against a handheld wand. We compared the performance of each of the methods using a repeated measures ANOVA experimental design, as well as recording users' opinions via a NASA TLX. The robot-assisted reaching significantly outperform manual reaching with the wand for all three visual feedback methods. However there is no significant difference between any of the three visual feedback methods when benefiting from the cooperation with the robot. We investigated how the performance of the user changed when the difficulty of task was altered in terms of spatial and orientational accuracy and show that non-robot assisted performance dropped rapidly with increasing task difficulty while the robot-assisted task performance remained stable despite the different feedback methods.

References

[1] A. Gregg-Smith and W. W. Mayol-Cuevas, "The Design and Evaluation of a Cooperative Handheld Robot," in *ICRA*, 2015.

[2] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of experimental psychology*, vol. 47, no. 6, 1954.

[3] R. W. Soukoreff and I. S. MacKenzie, "Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI," *International Journal of Human Computer Studies*, vol. 61, 2004.

[4] X. Zeng, A. Hedge, and F. Guimbretiere, "Fitts' Law in 3D Space with Coordinated Hand Movements," in *Human Factors and Ergonomics Society Annual Meeting*, 2012.

[5] Y. Cha and R. Myung, "Extended Fitts' law for 3D pointing tasks using 3D target arrangements," *International Journal of Industrial Ergonomics*, vol. 43, no. 4, 2013.

[6] J. Liang and M. Green, "JDCAD: A highly interactive 3D modeling system," *Computers & Graphics*, vol. 18, 1994.

[7] N. T. Dang, "A survey and classification of 3D pointing techniques," *RIVF*, 2007.

[8] S. Zhai and P. Milgram, "Human performance evaluation of manipulation schemes in virtual environments," *Virtual Reality Annual International Symposium*, 1993.

[9] A. Ranasinghe, N. Sornkarn, P. Dasgupta, K. Althoefer, J. Penders, and T. Nanayakkara, "Salient Feature of Haptic-Based Guidance of People in Low Visibility Environments Using Hard Reins," in *Cybernetics, IEEE Trans on*, 2015.

[10] M. S. Erden and A. Billard, "End-point impedance measurements at human hand during interactive manual welding with robot," in *ICRA*, 2014.

[11] F. Echtler, F. Sturm, K. Kindermann, G. Klinker, J. Stilla, J. Trilk, and H. Najafi, "The intelligent welding gun: Augmented reality for experimental vehicle construction," *Virtual and Augmented Reality Applications in Manufacturing*, 2004.

[12] "http://www.handheldrobotics.org."

[13] A. Gregg-Smith and W. W. Mayol-Cuevas, "Inverse Kinematics and design of a Novel 6-DoF Coupled Handheld Robotic Arm," in *ICRA*, 2016.

[14] J. D. Gammell, S. S. Srinivasa, and T. D. Barfoot, "BIT *: Batch Informed Trees for Optimal Sampling-based Planning via Dynamic Programming on Implicit Random Geometric Graphs," 2015.

[15] S. M. Lavalle and J. J. Kuffner, "Randomized kinodynamic planning," *IJRR*, vol. 20, no. 5, may 2001.

[16] S. Karaman, M. R. Walter, A. Perez, E. Frazzoli, and S. Teller, "Anytime Motion Planning using the RRT*," *ICRA*, 2011.

[17] J. Gammell, S. Srinivasa, and T. Barfoot, "Informed RRT*: Optimal incremental path planning focused through an admissible ellipsoidal heuristic," *IROS*, 2014.

[18] M. Tuceryan and N. Navab, "Single point active alignment method (SPAAM) for optical see-through HMD calibration for AR," in *ISMAR*, 2000.