Augmented Reality for Feedback in a Shared Control Spraying Task

Joshua Elsdon[1] and Yiannis Demiris[1]

***Abstract*— Using industrial robots to spray structures has been investigated extensively, however interesting challenges emerge when using handheld spraying robots. In previous work we have demonstrated the use of shared control of a handheld spraying robot to assist a user in a 3D spraying task. In this paper we demonstrate the use of Augmented Reality Interfaces to increase the user's progress and task awareness. We describe our solutions to challenging calibration issues between the Microsoft Hololens system and a motion capture system without the for well defined markers or careful alignment on the part of the user. Error relative to the motion capture system was shown to be 10mm after only a 4 second calibration routine. Secondly we outline a logical approach for visualising liquid density for an augmented reality spraying task, this system allows the user to see target regions to complete, areas that are complete and areas that have been overdosed clearly. Finally we produced a user study to investigate the level of assistance that a handheld robot utilising shared control methods should provide during a spraying task. Using a handheld spraying robot with a moving spray head did not aid the user much over simply actuating spray nozzle for them. Compared to manual control the automatic modes significantly reduced the task load experienced by the user and significantly increased the quality of the result of the spraying task, reducing the error by 33-45%.**

I. INTRODUCTION

Handheld robots offer the possibility of leveraging the human user's ability to move around the environment and control the progress of the task at hand, whilst allowing the robot to provide the final actuation to the target and make use of task specific data. This gives system designers the opportunity to lower the implementation costs of the robotic system, as locomotion systems are no longer needed. This forms a collaborative system, as neither entity is capable of completing the task without the assistance of the other. A good example of an application area that could find this technique useful is that of skin medicine application, where for certain treatments, accuracy in dosage and in placement is important. The system presented in this paper is one that is designed to apply liquid to the surface of a human analogue. This robot relies on the user for both locomotion and for local movement to complement it's single degree of actuation. The details of the development of the shared control algorithm are summarised in our previous work [1].

Due to the required shared understanding of the task, the robotic system must be able to indicate the current status of the task such that the user can collaborate effectively. In our system this is achieved using an augmented reality (AR) headset, specifically the Hololens by Microsoft. This is a binocular AR headset, and therefore can indicate 3D information to the user. In order to visualise holograms in the users vision, the headset tracks the users position using a visual odometry system. Whilst this works well for visualisations where positioning is not critical, it does suffer from drift as the Hololens adjusts its internal map or loses trackable features in the scene. Finding a way to quickly register the shifting Hololens coordinate system with the ground truth of the motion capture system would be of great use to system designers. They can make use of the reliability and the ability to track secondary objects that a camera based motion capture system offers. We offer our solution to this problem in Section III.

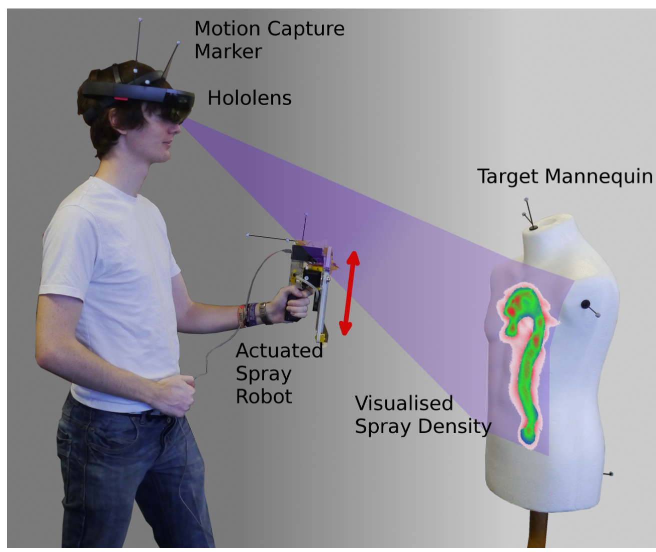


Fig. 1: The experimental setup used in this work. The Hololens, spraying robot and mannequin all are tracked with retro-reflective markers. The visualisation of the spray density is only visible to the user of the system as it is displayed to them through the Hololens. The spraying robot's head can be actuated up and down to assist with the spraying task.

Finding a way to logically provide task specific information to the user of a shared control system is critical as the task cannot be completed without the cooperation of the human user. We outline a simple colour scheme that we believe to be a logical way of representing targets, completed areas, and overdosed areas in a spraying task. This is detailed in Section IV.

[1]Joshua Elsdon, je310@imperial.ac.uk, and Yiannis Demiris, y.demiris@imperial.ac.uk both from the Personal Robotics Laboratory, Imperial College London www.imperial.ac.uk/PersonalRobotics

The benefit of a shared control system is that the parts of the task that require access to rich information or precise timing can be offloaded to the robotic part of the system. Providing some assistance is likely to reduce task load on the user, though knowing exactly how much extra help still adds utility to the system can allow designers to optimise other aspects of the system, such as size or cost. In Section V we evaluate three levels of assistance for a spraying task, considering the users performance at applying liquid, the speed in which they do the task and their subjective task loading.

II. Background

Here we will summarise some of the key works on which this one was based. There are three primary areas of work that each relate to one contribution of this paper. Firstly, performing calibration of augmented reality devices. Secondly, methods of visualising task progress in a spraying task. Lastly, work relating to assistance in shared control handheld robots.

A. Calibrating Augmented Reality Devices

Tuceryan et al. [2] proposed a system of calibration called ‘Single Point Active Alignment Method’ or SPAAM. In their system the position of the augmented reality headset is reported by a magnetic tracking system, the offset between the magnetic marker affixed to the headset and the optical centre of each eye is not known an is to be calibrated. The offset between the magnetic tracking system and the world coordinate system is analogous to the Hololens world frame and the world coordinate system in this work. The calibration of the tracker system was described in their previous work [3], they used a pointer object that was tracked using the same magnetic system to index known locations in the world coordinate system. The offset between the eye and the headset marker was found by having the user align a single known location in the environment up with a cursor that is displayed through the headset.

Tuceryat et al’s approach seems to be simple and effective, though is not directly applicable to the the hardware we are using. The primary problem is that the high frequency and low latency measurement of the Hololens is calculated internally, and there is no way to have a secondary object in the same frame used for indexing the world coordinate frame. In their work the headset and the wand were tracked with the same system, allowing them to match real world points with hologram locations directly. We would have to use the wand to index virtual images, which is difficult to do accurately. Also there is no attempt at recalculating any of the offsets in an online manner, due to the magnetic tracker’s coordinate system not shifting over time, however the Hololens does shift it’s coordinate system due to adjustments in the map used for the visual odometry.

Gilson et al [4] take a different approach that does not involve the user performing manual alignment tasks. They mount a camera inside the head mounted display such that it can see through the display into the scene. They can then put a motion tracked object into the scene and locate it also in the view of the camera. Then they can display a know grid visualisation on the head mounted display, and find that in the camera view. By finding correspondence between these two measurements the position of the marker attached to the head mounted display can be found relative to the optical centre of the eye. This approach has the advantage that it does not rely on the user to perform any alignment task, though at the cost of requiring a fairly elaborate calibration setup. A mannequin head mounted with cameras, special motion tracked and optically marked calibration boards add significantly to the complexity. Further it is likely that the mounting of the cameras in the head will effect the calibration, and should at least somewhat match the position of the end users eyes.

B. Visualising Task Progress in Augmented Reality

The set of work by Gregg-Smith et al [5, 6] is a key precursor to this work, as it investigates both augmented/virtual reality displays and the utility of handheld robotics. Their investigation into spacial guidance [5] compares 4 methods of user feedback, using both a handheld robot and a non-actuated wand. The feedback methods include a 7 inch display, a virtual reality system, a monocular augmented reality system and a gesture based method where the handheld robot points to it’s next waypoint. This final method was not implemented on the wand. They found that the actuation of the handheld robot helped the users perform the 5 degree of freedom positioning task much faster than using just the wand. Though they also found that while visual feedback was better than the gesturing, none of three methods was significantly better than the other. They comment that a significant issue with their monocular augmented reality display was the lack of depth perception giving the users difficulty. We wish to address this criticism by using a binocular augmented reality system.

One of the primary conclusions of Gregg-Smith et al was that perhaps only very rough feedback is necessary for the user to complete their task, and they conclude that perhaps lower quality systems than theirs could be viable for real world use cases.

A significant difference between this work and that of Gregg-Smith et al is that their handheld robot had sufficient degrees of freedom to complete the 5 degree of freedom task independently, and only required the user as transport to the area of interest. Our handheld robot is much less capable, having only one degree of freedom and the task is more nuanced, as such the level of cooperation with the user should be much higher, and quality of feedback is likely to be significant.

Yang et al. [7] designed a similar system for training technicians to use a spray gun to paint ship parts. They experimented with a large monitor emulating a surface, and a head mounted display to allow 3D manoeuvring during the painting task. They discuss the importance of having access to the paint thickness measurement in judging the quality of the trainees work, though they do not describe how, if at all, they display this to the user mid task. Further their work was for training purposes and as such some of the techniques used would not be appropriate if the guns were loaded with paint. For example using the monitor as a method of feedback emulates spraying through a window onto a surface, if the surface was contoured then the spray paint landing on the screen would not match that landing on the simulated 3D geometry. The head mounted display does not suffer from this problem however.

Kim et al. [8] developed a system specifically to demonstrate the thickness of paint coating to a user who is training in a spray painting task. They present 3 methods for giving feedback to the user regarding paint thickness: A numerical readout of the thickness of paint at the ray cast directly out of the spray gun; A graph showing the paint thickness along a line horizontal to the ray intersection; and 2D visualisation placed on the object itself. They give no in depth information regarding the 2D visualisation, though it seems to be a single colour modulated by the thickness, and has no indication of the required paint thickness at different locations. This is likely due to the application that they were targeting requiring uniform painting on all surfaces of and object.

C. Assistance Levels in Handheld Robots

A key piece of work was presented by Gregg-Smith et al. [9], this work presents 3 levels of assistance for a placement task and a simplified painting task. Manual mode allows the user full control of the trigger, which either 'sprays' paint or applies suction to the picking task, further there is no automatic articulation of the head. Semi autonomous mode actuates the trigger for the tasks automatically, though the head is still not articulated. Autonomous mode automates both the trigger and the head articulation, the head moves to help the user complete the task. They found that the larger amount of assistance reduced the amount of task load, as defined by the NASA TLX survey [10], especially in the painting task. They also measured a significant reduction in task completion time, again especially so for the painting task. Interestingly the effects were not as pronounced on the placement task, which seemed to be the one where even the most automated mode still relied on the user to be quite involved in the task.

In this work we are following a similar paradigm as Gregg-Smith et al, though we are extending the line of enquiry to a task that is closer to a realistic task, in our case the paint behaves realistically, and perfectly meeting the dose is impossible. Further the robot that we are using is significantly simpler, leaving more of the task to the user, and the feedback via the headset necessarily contains richer information that that in [5].

III. Finding Correspondence Between Hololens Frame and World Frame

Due to there being no well defined datum to use as a reference on the Microsoft Hololens we do not know the location of any of its features in 3D. Therefore any marker attached to it has to be assumed to be arbitrary, containing no helpful information to find the points of interest on the headset.

The Hololens does provide its location within it's own coordinate frame that was generated with visual odometry. The challenge is the calculate the offset between the Hololens world frame and the global world frame, in our case defined by a motion capture system. This section is not necessary for the understanding of the user study investigating shared control assistance level in Section V, though at the time of writing is necessary for it's implementation.

A. Method

In contrast to the techniques introduced in Section II-A the presented method does not rely on user alignment tasks or on well defined marker setups, though it must be noted the the difference in aims of the calibrations. We will be implicitly relying on the built in calibration of the Hololens and the inter-pupillary distance calibration routine provided by Microsoft, which can be replaced with a manual inter-pupillary distance measurement. Also the other methods of calibration can correct for some other modes of error such as display distortion, which we will take for granted.

In order for the visualisation to be rendered smoothly to the users, the Hololens calculates the movement of the users head in real time on the headset itself. Whilst we can override this, and set the users virtual position based on external measurements, we have found that the additional latency makes for an unacceptable experience. Also if tracking from an external system is lost, such as that from our marker based motion capture system the user would notice the lack of position updates. Therefore we propose a system that only acts on changing the location of the Hololens origin point, then allows the Hololens to maintain high frequency updates of the the users position in that frame.

In order to gain information about the Hololens' location relative to the world frame, defined by the motion capture system, we attached a rigid body of reflective markers to the headset. Due to the fact that the optical centre of the hololens depends on the particular calibration to the current user we cannot easily know the offset between the attached marker and the current optical frame. Thus we have two unknowns in the system, offset between motion tracking origin and Hololens world origin and the offset between the headset marker position and the optical centre of the Hololens. With two unknowns it is impossible to solve for both with only one measurement, and thus to discover these unknowns we optimise for them over a time series of observations. Further we must continually adjust for the drift introduced by the Hololens visual odometry system.

We can formalise Figure 2 using matrix equations where each matrix is representing a rigid rotation and translation. In Equation 1 T_{ho} is a rigid transform from the Hololens frame to the motion capture frame, T_{ro} is the rigid transform that describes the offset between the marker on the headset and the optical centre of the headset. P_h is the pose of the Hololens as reported by the Hololens's internal visual odometry, in it's own frame of reference. P_r is the measured pose of the marker attached to the Hololens in the motion capture frame.

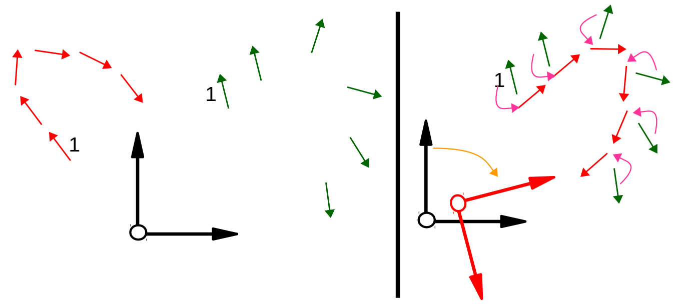


Fig. 2: Summary of the transforms involved. Red: Hololens position in Hololens frame of reference. Green: Rigid body marker position in the motion capture frame of reference. Pink: marker offset from optical centre. Orange: transform between Hololens frame of reference and the motion capture frame of reference. The '1' indicates that these are matched measurements, the other pairs are matched but are not numbered for clarity. It can be seen that there is one transform that will align the Hololens reported values (red) with the motion capture reported values (green) in such a way that there is a unique transform after this coordinate shift that links the optical centre and the headset mounted marker. This unique transform describes the transform between the marker on the Hololens and the optical centre of the Hololens.

$$T_{ho}P_h = P_r T_{ro} \tag{1}$$

$$T_{ho}, T_{ro}, P_h, P_r \in \mathbb{R}^{4\times4} \tag{2}$$

It is important to note that given one observation of P_h and P_r finding the appropriate transforms is impossible as there are infinite solutions. Therefore any solution must solve over a time series of such observations. We achieved this by forming an optimisation that was solved numerically, using a method leveraging singular value decomposition which allows for fast convergence of the optimisation. This method was described in [**Ho2013**].

First the centroids were found by averaging the translations defined by the poses measured in each frame.

$$C_h = \frac{1}{N}\sum_{i=0}^{N-1} \text{trans}(P_h^i) \tag{3}$$

$$C_r = \frac{1}{N}\sum_{i=0}^{N-1} \text{trans}(P_r^i) \tag{4}$$

Next the relative rotation to best align the data sets was found, here R is a rotation matrix. U,S,V are the typical outputs of the singular value decomposition (SVD).

$$H = \sum_{i=0}^{N-1} (P_h^i - C_h)(P_r^i - C_r)^T \tag{5}$$

$$[U, S, V] = \text{svd}(H) \tag{6}$$

$$R = UV^T \tag{7}$$

Using the offset between the centroids C_h and C_r, and the rotation matrix we now have the two series of poses roughly aligned, this gives us a starting point for the transform T_{ho}. Next we use a numerical optimisation library to minimise Equation 8, which can be thought of as a circular set of transforms, if the two transforms being calculated are correct applying them in this order should give the identity matrix.

$$\min_{T_{ro},T_{ho}} \sum_{i=0}^{N-1} P_r^i T_{ro} {P_h^i}^{-1} {T_{ho}}^{-1} - I \tag{8}$$

Now we can apply the transform ${T_{ho}}^{-1}$ to an object we would like to render and it will appear in the correct place within the motion capture arena. We can make the assumption that the marker offset transform will remain static for each use of the system. Importantly however, moving around will cause the Hololens's visual odometry to drift over time, or a loss of tracking could be slow to recover from. In this case, objects that were previously well registered to the real world environment will drift away from their intended position. Therefore we must continue to calculate the offset between the Hololens world and the motion capture world (T_{ho}).

Now that there is only one unknown, T_{ho}, we can solve for this using only one data point.

$$T_{ho} = {P_h}^{-1} P_r T_{ro} \tag{9}$$

This is not advised however as the noise in measuring P_r, the marker in the motion capture arena, causes large jagged movements of all rendered items. This is especially a problem if the noise is primarily in the rotation of the marker. Therefore it is useful to use a method that smooths the result of this calculation. A simple infinite impulse response filter is adequate for this task.

$$t_{\text{filtered}} = (1-f)t_{\text{old}} + f t_{new} \tag{10}$$

$$Q_{\text{filtered}} = \text{slerp}(Q_{\text{old}}, Q_{\text{new}}, f) \tag{11}$$

Where t represents the translation of the transform and Q represents the rotation quaternion. f represents the filtering coefficient, nominally $f = 0.01$, smaller f leads to smoother filtering with more latency. Slerp is the spherical linear interpolation between the given quaternions, where f denotes the interpolation coefficient.

B. *Validation*

To collect the data for P_h and P_r their timestamps are matched and a hold-off of 5mm is used. The hold-off represents the minimum distance required to move from the previous sample before a new sample is collected. To the user this sequence is a short walk around the motion capture arena whilst rotating their head, the time of which depends on the number of samples required, 500 samples takes around 4 seconds to collect.

Unfortunately due to the nature of this calibration it is difficult to demonstrate the accuracy of the alignment without experiencing it first hand. As such it is necessary to have a human in the testing loop, and the results will be somewhat dependant on their perceptions and care in completing the tasks.

To validate the calibration the user must place a known model such that it lines up well with the rendering that they see through the augmented reality system. From each placement iteration we will be able to collect the true location of the placed object, as measured by the motion capture system and the intended position of the hologram. Ideally the location that the user places the object exactly lines up with where the hologram was intended to be rendered, relative to the real world.

C. Results

The method described previously was tested for 10 iterations using different numbers of calibration samples. The results can be seen in Figure 3. After about 500 samples there is little increase in accuracy, and the best achieved accuracy is about 10mm position error and 20mRad angular offset (1.15 degrees). It is likely that the users ability to accurately perceive the depth information and manoeuvre the real world tracked object sets a floor on the error.

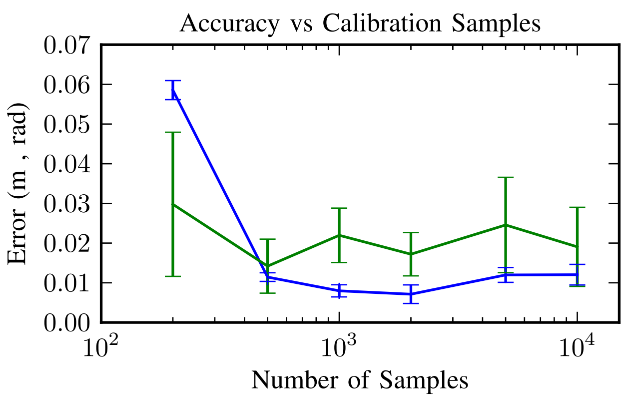


Fig. 3: The error between the measured location of a motion tracked marker and intended position against length of calibration series. Error bars represent the standard deviation in the measurement. N=10 for each experiment. After 500 samples there is no discernible improvement.

IV. Augmented Reality Feedback for Spraying Task

Now that we have a system where physical objects can be tracked reliably within the same frame of reference as the Hololens location, we can visualise interactive tasks to the user. In this section we will use this system to investigate an augmented reality spraying task. In order to complete a painting task with a handheld robot there needs to be some means of feedback to the user, otherwise they will have little idea of what areas still need further attention. With the cost of augmented reality and virtual reality and their quality reaching mass market standards, these technologies are a viable alternative to monitors for feedback.

A. Proposed Visualisation

Drawing on the work of Kim et al. [8] we decided that a 2D representation mapped directly to the surface of the 3D object using augmented reality was the most intuitive. With the task we present there is an additional complication, there is no implication that the paint should be evenly applied everywhere, there will instead be target regions. This means we must not only feedback what the user has already completed but must also distinguish between areas that still need to be painted and areas that should never be painted. This visualisation should also be natural to understand.

We propose that regions that require spraying, target regions, should be blue. As the area becomes filled this will transition to green. If the area becomes overfilled then the area will transition through a gradient to red. This is shown, along with the indication of reward in Figure 4. Traditionally green is seen as ‘good’ and red as ‘bad’, therefore the user will see that their job is to make all blue areas green, whilst making as few areas red as possible. Areas that were never intended to receive spray start white, due to this being the colour of the real world model they are working with, and will shift through a gradient to red as the area receives more spray. Full red in this case would be reached when the overdosing is in proportion the the typical correct dose within the target regions.

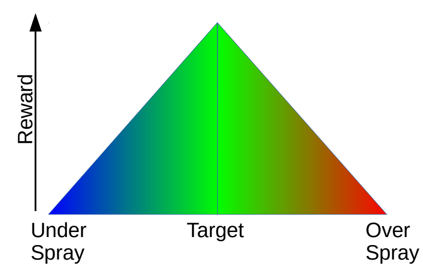


Fig. 4: The colour scheme for a region that requires spraying. Blue would indicate to the user that they should spray the area more, green that the area is correctly sprayed and red that it has become over sprayed. The vertical height in the image refers to the reward that is applied to the whole spraying job at this location, when the score becomes negative the colour will stay red.

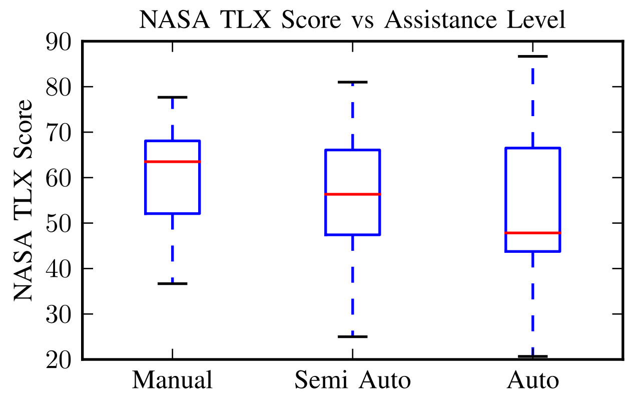
V. Contrasting Different Levels of Robotic Assistance in Spraying Task

Now that the Hololens, robotic spray gun and target mannequin can be located within the same coordinate system and we have outlined a sensible visualisation scheme, robotic spraying tasks can now be carried out in augmented reality. With our system we wish to investigate the extent to which robotic assistance could help in a 3D spraying task. Using the feedback mechanisms described in Section IV the user would have access to rich task specific information. This knowledge makes it possible to form a shared control behaviour with the robot, this could potentially allow the robot to be less complicated as the user can take control of some elements of the task and the robot others. In our case the human will be directing the high level flow of the task and the locomotion of the robot by manoeuvring it to the next target region.

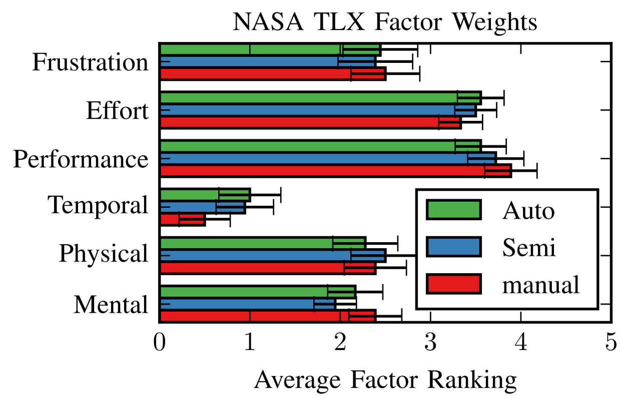
A. Task Outline

The task that the user will be attempting to complete will be to spray virtual liquid onto a mannequin on the zones indicated to them on the augmented reality headset. Three modes will be tested, in a similar paradigm to Greg-Smith et al [9]. Manual gives the user full control of the trigger, which releases the virtual paint from the nozzle, and nozzle remains stationary on the robot. Semi automatic mode activates the trigger on the users behalf when the simulation, described in [1], finds that spraying will increase the task score, as defined in Section IV-A. Finally automatic mode will also activate the trigger, though it will have the capability to slide the nozzle up and down a short gantry mounted on the handheld robot. The movement of the nozzle is calculated in a receding horizon manner using the current velocity and position to predict future positions over the next one second. This is described fully in our previous work [1].

The user will partake in 3 painting experiments per assistance level. After each assistance level they complete the NASA TLX survey [10] using the official NASA IPad application. They will be wearing the Microsoft Hololens which will overlay on the mannequin the target areas and their current paint status as described in Section IV-A, the setup can be seen in Figure 1. The user will communicate to the assistant when they feel that they can no longer make reasonable progress on the task. By the end of the experiment they will have filled the 3 different target patterns with each of the three assistance levels. Before each new mode the users were given the opportunity to test the mode and to ask any clarifying questions that they had.



(a) The combined TLX scores by mode.



(b) The factor weighting by mode.

Fig. 5: A summary of the NASA TLX scores, it can be seen that increasing the assistance level reduces task load, though does not change the relative factor importance much.

B. Results

There are three key areas of measurement that we will present here, NASA TLX [10] that indicate task load, task completion time, and accuracy of paint placement. In total there were 18 participants, 4 female and 14 male, aged between 20 and 34.

1) NASA TLX: It can be seen from Figure 5a there is a marked decrease in the median task load as the level of automation increases from 63.5 in manual to 47.835 in automatic mode. The factors involved in task load are shown in Figure 5b, over all there is no significant change in the relative importance of each factor across the experiments. Though the median of the task load decreased, as automation increased, the automated modes were not favoured by all, in fact both the highest and the lowest task load score was registered for the automatic mode.

2) Completion Time: Across nearly all of the trials manual was the quickest mode, with little difference between the automatic modes. Though this does require some careful qualification. Due to the nature of an automatic mode not allowing the participant to make a mistake, users often become quite perfectionist, hunting for the smallest area of improvement. In the manual mode however they get nervous that they will ruin what they have achieved and may stop early to avoid incurring negative scores due to overdosed regions. This effect could likely be removed easily by instructing the user the level of coverage the task requires, such that they do not waste their time, and further practice to remove their nerves in manual mode. To help remove this effect in this study, we will look to see the times taken to get from 10% to 90% of final coverage for that attempt, to remove the thinking time at the beginning and the hunting behaviour observed in the automatic modes at the end. The results can be seen in Figure 6, the time taken in the different modes did not vary significantly. Only in round 2 was there any appreciable difference, though this is explainable by the fact the perfectionist attitude of the users is applied per sub-patch, which can be seen in Figure 9, which is not removed by the filtering described above.

3) Accuracy: To effectively compare performance across trials the mean squared error (MSE) per pixel is a good measure as it is agnostic to the dose level (which happens to be constant) and the total area required to be covered. As can be seen from Figure 7 the MSE is best in the automatic

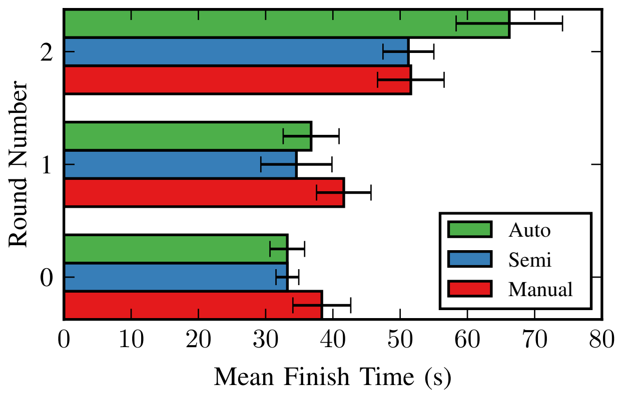
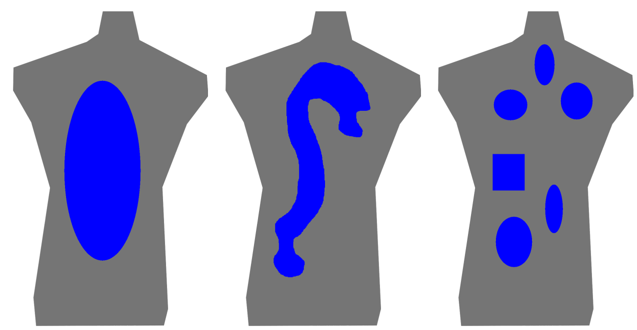


Fig. 6: The time to complete the spraying task. Time is taken from the interval between 10% and 90% of the maximum score for that run to remove unproductive time at the beginning and end of each run. The error bars represent the standard error of the mean (SEM). There is little difference in completion time with different assistance levels.

(a) Typical result of manual spraying (b) Typical result of semi auto spraying (c) Typical result of auto spraying

Fig. 8: A typical example of the 3 modes used on a given target. Green sections are well painted, blue is in need of paint, and red areas are over painted. It can be seen that in manual mode there is significant over spraying.



(a) Target 0 (b) Target 1 (c) Target 2

Fig. 9: These were the three patterns the users were asked to spray in each mode. 9a is designed to test area fill, 9b for detailed strips, and 9c for small isolated areas. Further the boundary to area ration increases through the levels

modes and worst in manual mode. There is between a 33% and 45% reduction in MSE between the automatic modes and manual mode. There is no significant difference between the automatic modes. In Figure 8 you should be able to notice that in manual mode there is significant amounts of paint outside of the intended boundary, whereas the automatic modes have a similar kind of paint distribution.

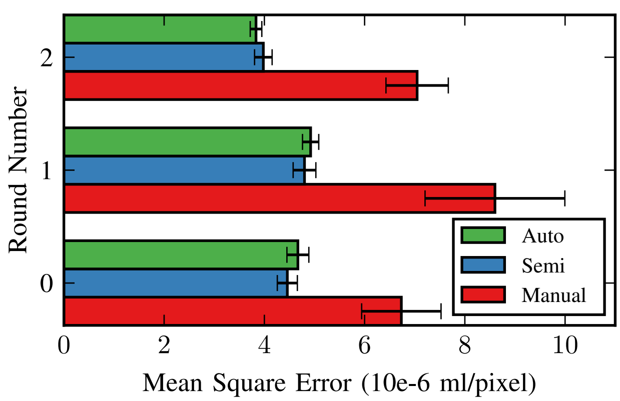


Fig. 7: The Mean Squared Error (MSE) per pixel in the bounding box surrounding the target patch. It can be seen that the manual attempts are much more error prone, also the user variation is large with the manual mode, and small with the automatic modes. Variation between rounds is also small in the automatic modes. Error bars represent the Standard Error of the Mean (SEM).

VI. Conclusion

In this work we demonstrated three contributions. Firstly we demonstrated a method of calibrating the Hololens's internal coordinate system to a frame defined by a motion capture system. This method does not require well defined markers or any careful interaction from the user. We showed that the error compared to the motion capture system can be as low as 10mm and that a calibration taking 4 seconds is enough to reach this. This kind of calibration could be useful when a researcher has a system that reports it's location in it's own coordinate frame and there is no well defined datum to help convert this to a motion capture or other global frame. Our second contribution was outlining a logical system for visualising liquid density in an augmented reality spraying task. This was implemented on with the Microsoft Hololens such that users can see target regions to spray on a mannequin that is tracked by the motion capture system. Our third contribution is to analyse which level of assistance is best used in the augmented reality spraying task. We demonstrated that whilst a pure manual spray gun can often be quicker, most users report that the automatic modes put them under less load. Further we have showed that the performance of the users in the spraying task was much improved with the automatic modes, with the mean square error being 33-45% less for the automatic modes compared to the manual mode. Though automatic and semi-automatic mode performed similarly for the most part the automatic version is significantly more complicated, which for this particular use case is unlikely to be worth while given the marginal benefits.

References

[1] Joshua Elsdon and Yiannis Demiris. "Assisted painting of 3D structures using shared control with a hand-held robot". In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*. 2017, pp. 4891–4897.

[2] Mihran Tuceryan, Yakup Genc, and Nassir Navab. "Single-Point Active Alignment Method (SPAAM) for Optical See-Through HMD Calibration for Augmented Reality". In: *Presence: Teleoperators and Virtual Environments* 11.3 (2002), pp. 259–276.

[3] M. Tuceryan et al. "Calibration requirements and procedures for a monitor-based augmented reality system". In: *IEEE Transactions on Visualization and Computer Graphics* 1.3 (1995), pp. 255–273.

[4] Stuart J Gilson, Andrew W Fitzgibbon, and Andrew Glennerster. "Spatial calibration of an optical see-through head-mounted display." In: *Journal of neuroscience methods* 173.1 (2008), pp. 140–6.

[5] Austin Gregg-Smith and Walterio W. Mayol-Cuevas. "Investigating spatial guidance for a cooperative hand-held robot". In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*. 2016, pp. 3367–3374.

[6] G. Klein and D. Murray. *Mixed and Augmented Reality, 2007. ISMAR 2007. 6th IEEE and ACM International Symposium on*. 2007.

[7] Ungyeon Yang et al. "Virtual Reality based Paint Spray Training System". In: *2007 IEEE Virtual Reality Conference*. 2007, pp. 289–290.

[8] Daeseok Kim et al. "Visualizing Spray Paint Deposition in VR Training". In: *2007 IEEE Virtual Reality Conference*. 2007, pp. 307–308.

[9] Austin Gregg-Smith and Walterio W. Mayol-Cuevas. "The design and evaluation of a cooperative handheld robot". In: *2015 IEEE International Conference on Robotics and Automation (ICRA)*. 2015, pp. 1968–1975.

[10] Sandra G Hart and Lowell E Staveland. "Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research". In: *Advances in Psychology* 52 (1988), pp. 139–183.