Facilitating Human-Robot Collaboration
Using a Mixed-Reality Projection System

by

Yash K. Rathore

A Thesis Presented in Partial Fulfillment
of the Requirements for the Degree
Master of Science

Approved July 2017 by the
Graduate Supervisory Committee:

Hani Ben Amor, Chair
Robert Atkinson
Brian Nelson

ARIZONA STATE UNIVERSITY

August 2017

ABSTRACT

Human-Robot collaboration can be a challenging exercise especially when both the human and the robot want to work simultaneously on a given task. It becomes difficult for the human to understand the intentions of the robot and vice-versa. To overcome this problem, a novel approach using the concept of Mixed-Reality has been proposed, which uses the surrounding space as the canvas to augment projected information on and around 3D objects. A vision based tracking algorithm precisely detects the pose and state of the 3D objects, and human-skeleton tracking is performed to create a system that is both human-aware as well as context-aware. Additionally, the system can warn humans about the intentions of the robot, thereby creating a safer environment to work in. An easy-to-use and universal visual language has been created which could form the basis for interaction in various human-robot collaborations in manufacturing industries.

An objective and subjective user study was conducted to test the hypothesis, that using this system to execute a human-robot collaborative task would result in higher performance as compared to using other traditional methods like printed instructions and through mobile devices. Multiple measuring tools were devised to analyze the data which finally led to the conclusion that the proposed mixed-reality projection system does improve the human-robot team's efficiency and effectiveness and hence, will be a better alternative in the future.

ACKNOWLEDGMENTS

I would like to express my sincere gratitude to my advisor, Dr. Hani Ben Amor, for the continuous support and guidance he provided me throughout my research. I would like to thank my thesis committee members, Dr. Brian Nelson and Dr. Robert Atkinson, for the valuable advice and mentoring they kindly offered me.

I would like to thank my fellow lab mates at ASU Interactive Robotics Lab: Ramsundar K G, Kevin Luck, Trevor Richardson, Simon Stepputtis, Joe Campbell, Ashish Kumar, Indranil Sur for the stimulating discussions and collaborations. I thank all the volunteers involved in my human subject study for their participation and feedback, which helped me to complete my experiments successfully.

I would like to thank my father Dr. Kanaiyalal Rathore, my mother Ranju Rathore and my sister Dr. Khyati Rathore, and Dr. Parth Shah for believing in me and supporting me during this entire journey.

Thanks to my mini ASU family: Abdul Kalyaniwala, Manoj Kumar, Dhiren Navani, Chintan Khatri, Himanshu Tyagi, Aakash Chaurasia, Sushant Sakolkar, Bhakti Bohara, Akriti Jain, Anusha Rajan for their moral support and encouragement whenever I needed it.

Finally, I would like to thank all my friends and loved ones who have supported me throughout my research.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

The world is advancing and looking forward to a boom in collaborative robots in the industries. One could call it the "Fourth Industrial Revolution"; where robots are being deployed and used for their precision in manufacturing industries. It is essential for these robots to be human-friendly and adaptive to human collaborative environments. Hence, it is a challenge to make these robots more dynamic, as well as indicative for the human partner to excel in performing tasks. Traditionally, robots were meant to boost production speed by leveraging their superior pace and precision. Robots have the ability to execute repetitive tasks faster than an average human. But manufacturing today requires robots to seamlessly work hand-in-hand with human partners to further increase flexibility and efficiency. As mentioned in the Roadmap for U.S. Robotics report, "Understanding the user‘s activity and intent are necessary components of human-machine and thus human-robot interaction, in order to respond appropriately and in a timely and safe fashion" (Christensen *et al.*, 2009). The human-robot interaction remains a major obstacle that leads to reduced trust in the human-robot team. It is essential for a robot to make its intentions clear to its human partner to achieve seamless human-robot collaboration (Dragan *et al.*, 2013; Mainprice *et al.*, 2010; Stulp *et al.*, 2015). Numerous efforts have been made to overcome this challenge by incorporating speech (Tellex *et al.*, 2014; Perera *et al.*, 2016), vision, and other human-like features into the robots, to make their human partner feel more comfortable while working simultaneously. One such efficient method by (Andersen *et al.*, 2016) uses intention projection based systems which can indicate the objectives of the robot. But learning the intention of the human is equally important for the robot to make better decisions.

In this project, I propose a mixed-reality based projection system which is capable of communicating robot's intentions by augmenting graphical visuals onto the environment while taking into account the human actions simultaneously. The main idea is to facilitate interaction between humans and robots without the use of any wearable technology, thereby making the experience more natural. The system can track any 3D-object in the real environment, provided the 3D-CAD model is available. Based on the received pose of the object, it can generate informative visuals to aid the human's understanding of the robot's intention by projecting the same on the working environment. Moreover, the system can perform human skeleton tracking which helps it to acquire more information about the human actions and can, therefore, be more adaptive catering the priorities of the human. A simple example can be seen in Figure 1.1 where the system tries to warn the human partner by projecting a cautionary signal on the floor. The user is not supposed to come close to the object while the robot is performing a manipulation task.

Use of robots is not just limited to industrial manufacturing purposes but the growth can be significantly seen in medical, educational, household environments and various other sectors (Iftikhar *et al.*, 2011; Chen and Wang, 2016; Xiao *et al.*, 2012). Hence, it is important to create a general methodology which can serve multiple objectives. Humans tend to learn faster when provided with visual depictions or drawings for executing tasks, instead of textual or vocal instructions (Verdi *et al.*, 1997). We created a simple visual language which can be easily used to construct a complex set of messages depending on the requirements of the user. More precisely, it is a visual language which serves as a medium for a smooth human-robot interaction.

We hypothesize that the overall team efficiency and effectiveness will be higher when working with our proposed mixed-reality projection system as compared to

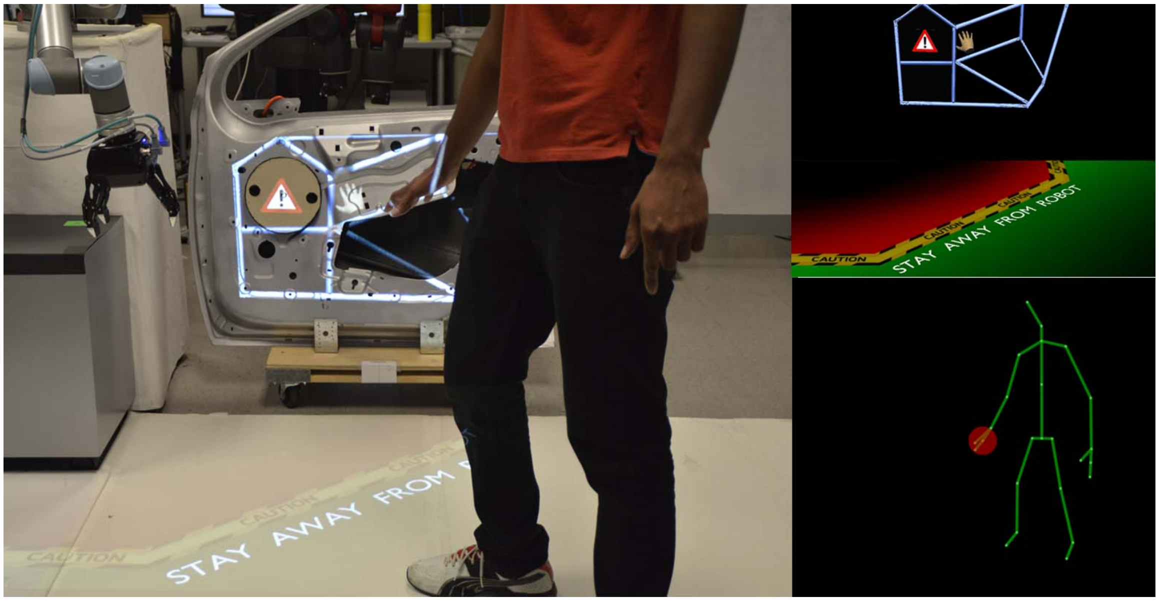


Figure 1.1: Warning Sign Displayed When User Comes Closer To The Object

other traditional methods like use of mobile devices and printed instructions. The increased efficiency and effectiveness will, in turn, improve the human-robot team fluency and trust within the team. To be able to make any deductions, we conducted a user study with 15 participants who were given a set of tasks to assemble a car door with the help of a UR5 robot. The robot had to be handed over tools for completing the tasks. Each participant was required to perform the exercise thrice using our proposed system (projected mode), printed instructions (printed mode) and lastly using instructions on a mobile display(mobile display mode). We evaluated the results using multiple tools for analyzing the data and found significant differences between the outcomes generated from projection mode and other two modes. However, the performance in the mobile mode was comparable to that of projected mode.

Chapter 2

REVIEW OF LITERATURE

2.1 History of Robotics

The field of robotics came into existence not more than seventy years ago. The first industrial robot designed by George Devol in the year 1954, was able to transfer objects within a small distance when the start and end locations are specified. Devon founded a company called Unimation in 1956 to manufacture its first industrial robot, UNIMATE, which was installed by General Motors at its New Jersey plant in 1962 (Ayres *et al.*, 1981). German robotics company KUKA designed Famulus, a robot with six electromechanically driven axes in 1973. Prof. Victor Scheinman developed "The Silver Arm", in 1974, which could perform small-parts assembly jobs utilizing the feedback from pressure and touch sensors. For industrial purposes, Vicarm Inc. developed a mini computer controlled version of the same. The same team along with Unimation and General Motors, also came up with PUMA, in 1978, which was used in assembly lines and is still used by researchers today. In 1975, the world's first fully electrical robot, ASEA IRB, was built by a European company called ASEA. It was also the first microprocessor-controlled robot which worked on Intel's first chipset. Robots with grippers were launched in the market soon beginning from the year 1977. One such robot, Motoman L10, was developed by Yaskawa America Inc. which was capable of moving weights up to 10kgs. The first servo gun technology robot was developed by Nachi Robotics, in 1979, which was used for spot welding. In the same year, OTC Japan introduced the first generation of dedicated arc welding robots.

By this time, the robotic revolution had already paved the way for the future of industrial robots that were made in large numbers to cater the need of the market. After 1980, a new robot was out almost every month. These robots were smarter than their predecessors with a higher degree of operational freedom and were controlled by microprocessors. Robotics was at its zenith, and it was only a matter of time before people realized that these robots do not necessarily have to be just a machine following the human's instructions. As Human-Robot collaboration began in its early days, a need for more human-friendly robots emerged, and people started looking out for ways of interaction between the team.

2.2 Human-Robot Interaction

Due to the emerging growth of robotics, Human-Robot Interaction(HRI) became essential and gained attention over the past years. So what exactly is HRI? According to a definition -

"Human Robot Interaction (HRI) is a field of study dedicated to understanding, designing, and evaluating robotic systems for use by or with humans" - (Goodrich and Schultz, 2007) p. 204

What does HRI try to achieve?

"The HRI problem is to understand and shape the interactions between one or more humans and one or more robots" - (Goodrich and Schultz, 2007) p. 216

Interaction in this context refers to the communication between robots and human. The mode of communication is of high importance and could vary according to different scenarios. (Goodrich and Schultz, 2007) pp 203-275 It can be broadly classified into two types:

- <u>Remote Interaction</u>: In this mode of interaction, the human and robot are not physically co-located i.e. generally separated spatially or even temporarily.

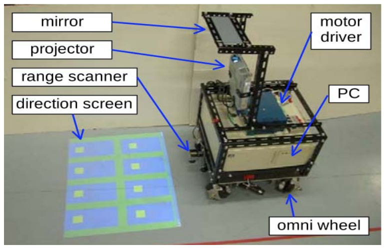
- Proximate Interaction: This mode of interaction requires both the entities to co-exist in the same location, for example, industrials robots working closely with the human workers.

For this project, we are entirely concerned with the proximate interactions as a way to achieve our final goal. Over the past years, people have been using different ways to communicate with robots. Computers, remote controllers, mobile devices, etc are some of the external devices widely used to interact with robots.

2.3 Related Work

With advancing technology, more emphasis is put on improving the human-robot interaction. Augmented Reality(AR) and Mixed Reality(MR) are changing the way we used to interact with physical and virtual objects. Early work by (Sato and Sakane, 2000) shows the use of LCD projector and vision algorithms to detect and track hand gestures. The system projects visual marks which could be used to control and interact with the Robot.

It was soon realized that the use of projection systems in human-robot interaction could open up new possibilities. In 2009, researchers at Dr. Matsumaru's Bio-Robotics and Human-Mechatronics Laboratory worked on interesting human-robot interaction projects. One such project was the robot intention projection, Figure 2.1(a), wherein a projector was mounted on a robot to display the trajectory of its motion. (Matsumaru and Akai, 2009) A concept of Step-On-Interface was introduced which creates an interface on the floor using projections to control the robot's movements. This can be helpful for people with disabilities who cannot use keyboards or other devices.



(a) Step-On-Interface (Matsumaru and Akai, 2009)

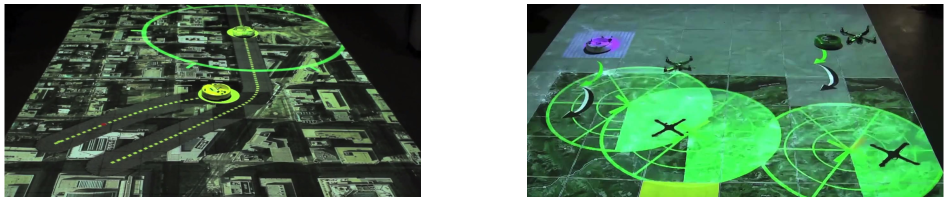
(b) Laser Interface (Ishii *et al.*, 2009)

Figure 2.1: Use Of Projections To Control Robots

A similar approach was used, by Kentaro Ishii and his team for their IGARASHI Design Interface project (Ishii *et al.*, 2009), where they made use of a laser pointer to control a robot. The laser pointer can accurately point to real world locations, Figure 2.1 (b). The basic idea was to apply stroke gesture recognition to laser trajectory. The user can specify the type of task and the target position, as well as target objects.

Recent work at MIT uses an Augmented Reality Room which shows what robots are thinking. The researchers used their AR system to for robots traveling through a virtual city avoiding obstacles like human pedestrians and other vehicles, Figure 2.2 (a). The robots avoid the collision by detecting the obstacles and then computing the optimal route. As the robots did that, a projection system displayed their thoughts on the ground, so that researchers could visualize them in real time. The thoughts consisted of colored lines and dots representing obstacles, possible paths, and the optimal route that were always changing as the robots and pedestrians moved.

(Omidshafiei *et al.*, 2015) demonstrates an advanced projection system, MAR-CPS, which augmented the physical laboratory space with real-time status and intentions of drones and ground vehicles in a cyber physical system, Figure 2.2 (b). Several other studies have also used projection systems to convey information to the user. However, these systems were confined to displaying on flat surfaces and did not



(a) Robots Displaying The Paths (b) MAR-CPS

Figure 2.2: Displaying The Intention of The Robot When Multiple Agents Are Involved (Omidshafiei *et al.*, 2015)

consider the state of physical objects while projecting information.

In contrast to that, (Andersen *et al.*, 2016) demonstrated an early prototype of a projection system that tracks physical objects in real-time and projects visual cues at specific spatial locations. A preliminary usability study demonstrated improved effectiveness and user satisfaction with the projection-based approach in a human-robot collaborative task. However, the experimental study was limited to simple tasks like tracking, moving and rotating a single object on a flat surface, which does not reflect a real-world workspaces. Also, the set of different signals that could be communicated was limited.

(Hrvoje and Andrew et al. 2014) developed a unique Spatial Augmented Reality (SAR) system called, Mano-a-Mano, which combines dynamic projection mapping, multiple perspective views and device-less interaction to support face to face interaction with 3D virtual objects. The system supports different perspectives for two users when they are facing each other and are several feet apart in a room. It renders virtual 3D objects as if there is a hologram in the space between the two users. Moreover, the users can simultaneously interact with those virtual objects. For example, a 3D Earth globe can be rotated by one user, and the other user can also see the proper view of the action, Figure 2.3. Its main advantage over more traditional AR approaches, such as handheld devices with composite graphics or see-through head

Figure 2.3: Mano-A-Mano, Enabling Two Users To Interact With 3D Virtual Objects Simultaneously (Hrvoje and Andrew et al. 2014)

worn displays, was that users were able to interact with 3D virtual objects and each other without cumbersome devices that obstruct face-to-face interaction.

In this research, we describe a novel system that is capable of tracking and projecting information on multiple objects in three dimensions simultaneously. We also present a rich visual language that goes beyond the display of trajectories or distances and allows for complex signaling.

Chapter 3

METHODOLOGY

The proposed mixed-reality system comes with minimal hardware requirements consisting of a RGB camera to track 3D objects, a projector to project the graphical information onto the environment, a Microsoft Kinect sensor to track human movements and a computing device on which the system runs. A schemetic visualization of the arrangement can be seen in Figure 3.1, where the system is shown projecting visuals on a box while the human is around.

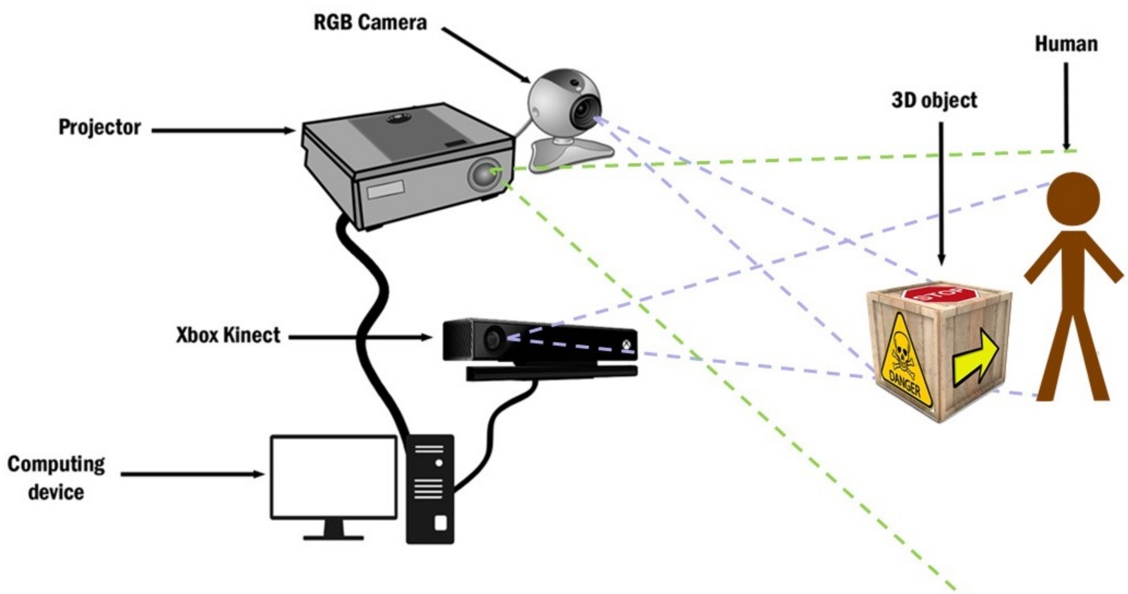


Figure 3.1: Schemetic Representation Of The System

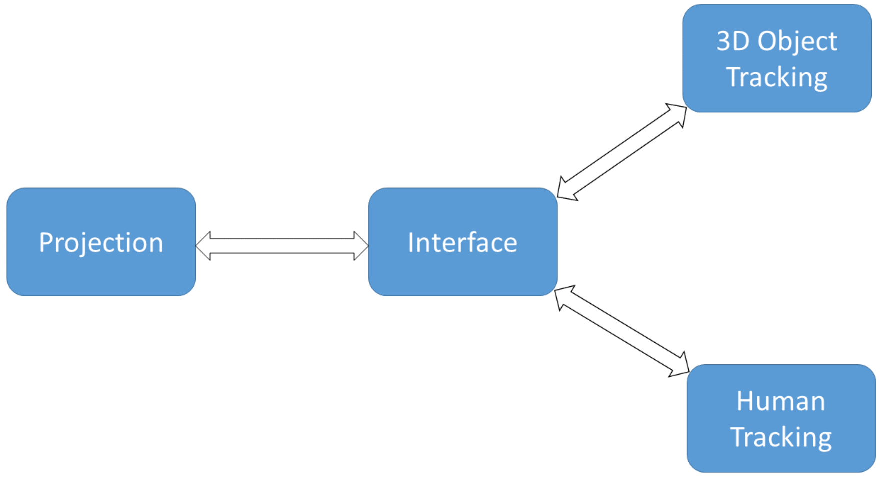


Figure 3.2: High-Level System Architecture.

3.1 Architecture

The mixed reality system is designed to be a generic system which can be easily customized to support various scenarios. Hence it has been modularized into four independent components, as shown in Figure 3.2, that serve as building blocks of a visual signaling framework. My focus was entirely on developing the projection and human tracking components and the rest were handled by a colleague. The objective of this architecture is to provide a medium for the human-robot team to interact using mixed-reality visual cues which are projected onto dynamic environment. The interactive medium permits a bi-directional communication between the team. In this chapter, all the four components are described in detail to get a clear understanding of system's working. Next, we will see how the components are incorporated together to produce visual tokens that are used to create a visual dialect.

3.1.1 3D Object Tracking

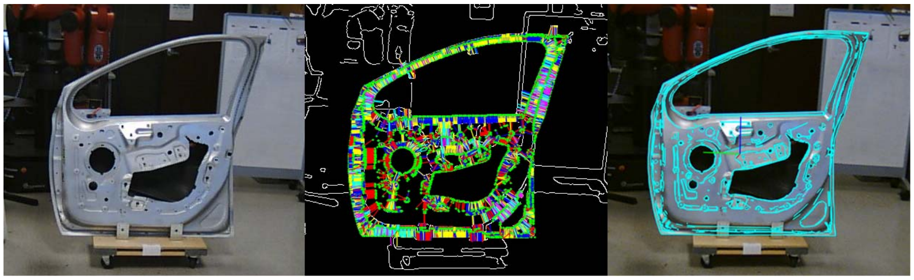
The very first component of the system is a model based 3D object tracker, inspired from the work by (Choi and Christensen, 2010), which uses vision based algorithms to estimate 6-DOF pose of objects in the environment in real time. The tracker expects the 3D CAD model of the object which is mapped onto the image frame consisting of the canny edges (Canny, 1986). The tracking system extracts features from the given model and tries to estimate the 6-DOF pose for the tracked object, by wrapping one of the many possible poses onto the real object to find the closest matching pose. The object boundaries are of more importance than the inner features, which also makes the tracking system occlusion free.

A case circumstance is shown in Figure 3.3 (a) where a car-door is being tracked in the environment. The leftmost image is the input raw frame, center one shows the errors and sample points, and the rightmost one corresponds to the final estimated pose being wrapped onto the object. Figure 3.3 (b) contains similar images, but the car-door is now rotated anti-clockwise at an angle of 30 degrees and is also partially blocked by another object. The algorithm overcomes the occlusion and tracks the object successfully. The tracker can track simple objects like a box which contains minimal feature points (c).

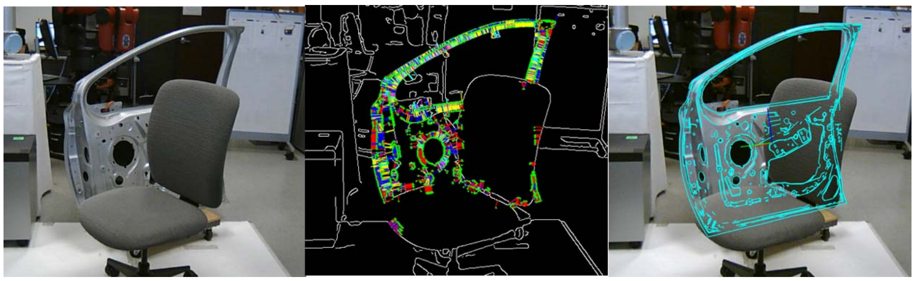
The pose of the tracked object can now provide a fast and continuous stream of tracking data to other components for further computations, using a secured UDP connection.

3.1.2 Projection

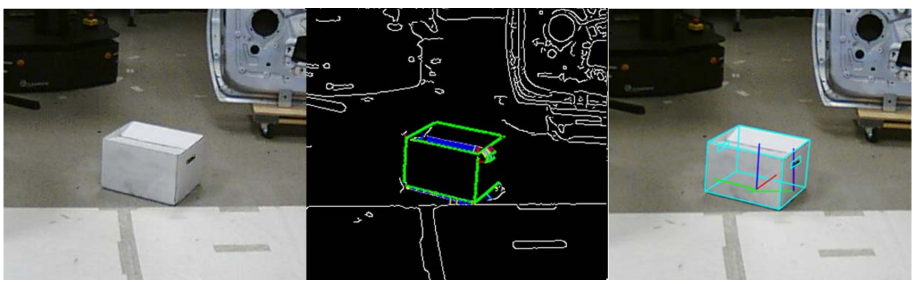
This very component is responsible for producing the graphical elements in the system, that can be easily projected onto the environment using a projector. It takes



(a) Car-Door Tracking At Initial Pose



(b) Car-Door Tracking At Rotated Pose With Occlusion



(c) Box Tracking At Initial Pose

Figure 3.3: 3D Object Tracking

the 3D pose of the objects in focus and generate relevant information for the human using a library called Augmented-Projection-Lib(APL). APL is specifically developed for this module which makes it a simple to utilize library providing multiple structures and functions to generate elementary visual tokens. The graphic development uses a toolkit known as OpenSceneGraph, which is described in the subsections below.

OpenSceneGraph

OpenSceneGraph is an open source high performance 3D graphics toolkit, used by application developers in fields such as visual simulation, games, virtual reality, scientific visualization and modelling. It is written entirely in standard C++ and wrapped around OpenGL, maintaining a similar architecture and hence making it easier for developers who are familiar with OpenGL.

Everything in OpenSceneGraph is a tree; typically drawn schematically with the root at the top, and leaves at the bottom. It all starts with a top-most root node which encompasses your whole virtual world, be it 2D or 3D. The world is then broken down into a hierarchy of nodes representing either spatial groupings of objects, settings of the position of objects, animations of objects, or definitions of logical relationships between objects such as those to manage the various states of a traffic light. The leaves of the graph represent the physical objects themselves, the drawable geometry and their material properties.

This structure makes it easier for scene creations and allows one to control almost everything in it, like the camera positions, respective object's orientations, lightings, etc. This makes it perfect for our requirements since the object positions can be manipulated in real time by literally using a single line of C++ code.

Augmented-Projection-Lib

For simplicity purpose a projection library is created that could be efficiently used to create set of tasks comprising of visual elements. Elements that can be generated using the pose and shape of the tracked objects, which enables it to project right onto the real objects in focus.

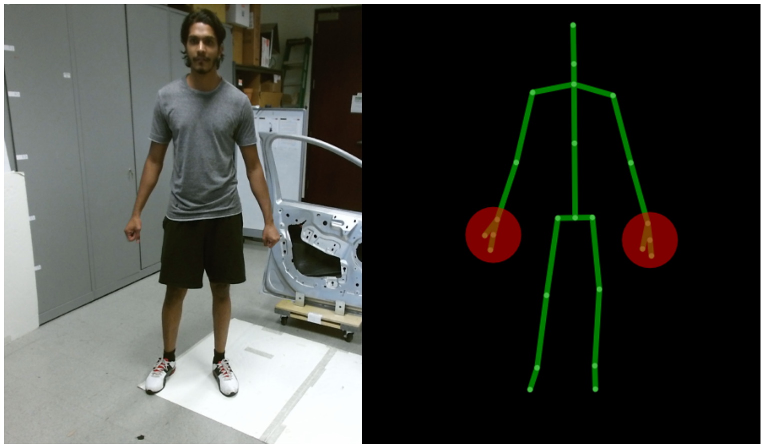


Figure 3.4: Human-Skeleton Tracking

3.1.3 Human Tracking

This component is completely centered around tracking humans in the environment using Microsoft's Kinect Sensor 2.0. The device comprises of an inbuilt 1920x1080 rgb camera, 640x480 depth camera and a 640x480 infrared camera. These configurations makes it possible to perform human skeleton tracking with high accuracy that can be leveraged for making our system aware of the human movements. Microsoft's sdk provides kinect.h library which contains utility functions to implement skeleton tracking. The current device supports tracking for upto 6 persons at the same time and 25 joint angles for each person in the frame at the rate of 30 fps. Figure 3.5 shows an example of human skeleton tracking and the list of possible joint angles that can be tracked. For each joint angle we obtain the transformation matrix and the degree of confidence with which it is tracked. All this data is sent over to other components using a secured UDP connection, which is then manipulated and used by the mixed-reality system.

The human skeleton tracking data is then used by our system to generate more dynamic visuals and become responsive to human movements. It is also possible to

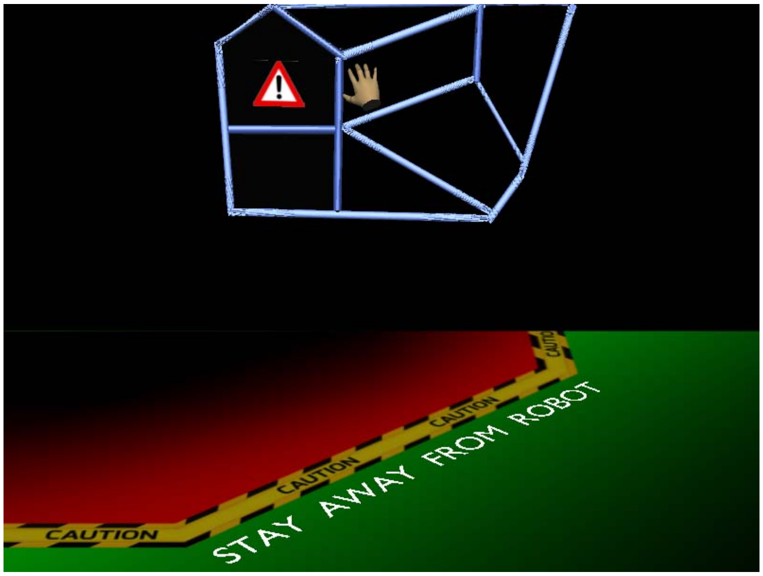


Figure 3.5: Output Generated By Projection Component

implement simple grasping gesture, which can be highly useful to interact with the graphical elements in the environment.

The projection module now obtains all the data from the 3D object tracking component and the human tracking component, using which it generates informative visuals. An example output produces by the projection module can be seen in Figure 3.5, where a graphical car-door frame is rendered corresponding to the received pose of the actual car-door. A warning signal and the region of interest for the robot(area in red) is attached to the car-door frame. Cautionary text and region for safety are also drawn. These rendered images can now directly be seen projected in the real world.

3.1.4 Interface

The Interface component is the spinal cord of this mixed-reality system, in a sense that all the inter-communication between the above three components is being controlled here. The first important role of this component is data passing. It acts as a two-way bridge filtering, ordering, and passing all the data and messages to maintain a smooth communication within the system.

The second important role is to take care of the extrinsics between the physical devices, namely the RGB camera(used for the 3D object tracking), the projector, and the kinect sensor. An extrinsic calibration between the devices is necessary in order to have a single perspective frame. This simplifies the visual creation process as we only need to consider a single world coordinate system. All the vector algebra calculations required for bringing down the three frames perspective to a single one, is performed by this component.

Extrinsic Calibrations

The mixed-reality system uses 3 devices in total and hence it is required to calibrate them in order to achieve a single coordinate model. There are two calibrations involved in the process

- Camera-Projector Calibration: The tracked object pose received by the RGB camera needs to be brought in the projector's coordinate frame and hence a camera-projector calibration is performed. This was accomplished by using Daniel Moreno and Gabriel Taubin's method of camera-projector calibration which gives a reprojection error as low as 0.1447 in ideal conditions.

- Kinect-Projector Calibration: For similar purpose, a kinect-projector calibration is also necessary, but there is no efficient way to do it with the kinect v2 sensor. Hence, an elementary approach based on vector arithmetics is used. Firstly a camera-camera calibration between the kinect and the camera is performed. We already have the camera-projector calibration, which can now be

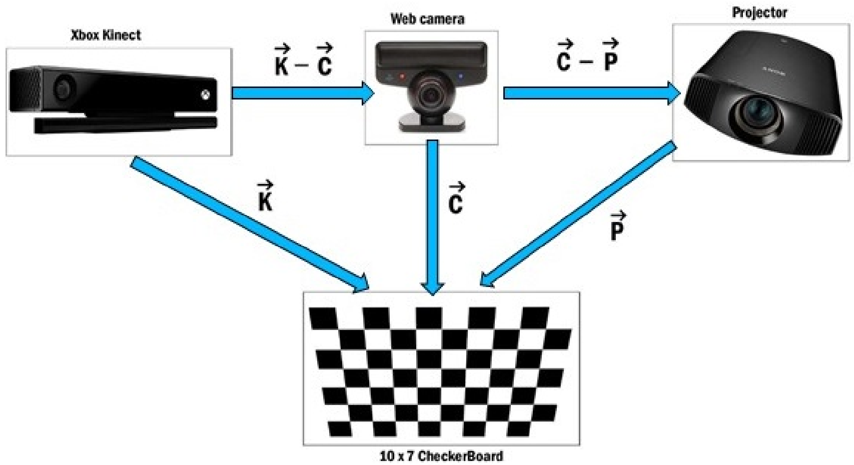


Figure 3.6: Camera-Projector And Kinect-Projector Calibration

used to get the kinect-projector calibration.

The interface component additionaly maintains a record of all the tracking information for both the 3D objects and the human movements. The output is a well structured excel file with each row corresponding to the data recorded at a single timestamp. This information could further be used for creating a smart system that is capable of anticipating the end results for a task based on prior learning.

The four components described above form the integral elements of our mixed-reality system. All these components are closely bind together forming a powerful tool for creating the interaction medium required for the human-robot collaboration.

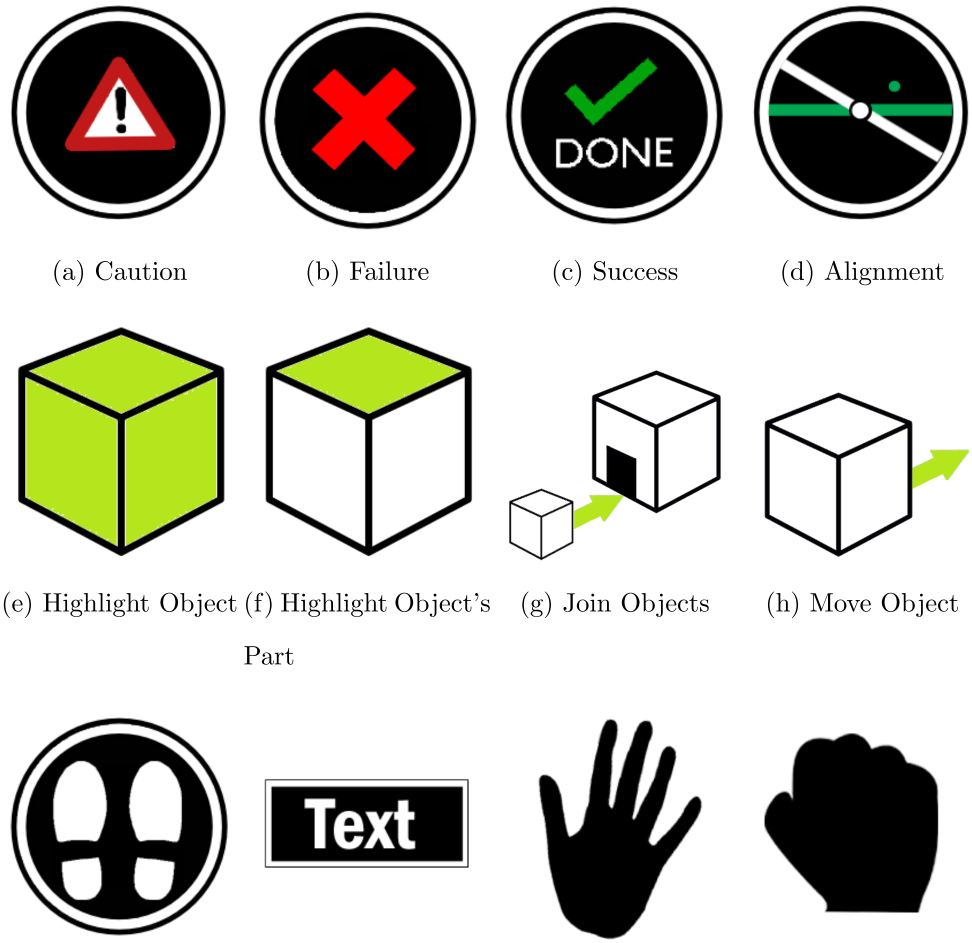
Substantives.	highlight-object(X) highlight-object-part(X,Y) hand-open(X) hand-close(Y)
Verbs.	move-to(X,Y) remove(X,Y) join(X,Y) align(X,Y)
Prepositions.	in-front-of(X,Y) left-of(X,Y) right-of(X,Y) at-position(X,Y) relative-to(X,Y,Z)
Affirmation.	success() failure()
Safety and Hazard.	stop(X) caution(X) robot-work-area()
Text.	text(X) text-flash(X)

Table 3.1: Subset Of Proposed Visual Cues

3.2 Visual Tokens

The intention of this research was to create an extensible visual language that can be used for various industrial scenarios. For this, we coined the idea of creating visual tokens that forms the basis of every visual segment in a mixed-reality system. Just as we use English language to communicate with other humans, we can use these visual tokens to frame a meaningful visual chain of tasks. Hence, we devised a simple set of tokens considering the various possible tasks involved in any human-robot interaction. These tokens can be easily extended according to the requirements of the task.

The basic fragment of visual cues proposed here includes patterns for designating and targeting objects (substantives), indicating positions, relations, and orientations (prepositions), basic movement instructions (verbs), success and failure (affirmation),



(a) Caution (b) Failure (c) Success (d) Alignment

(e) Highlight Object (f) Highlight Object's Part (g) Join Objects (h) Move Object

(i) Stand Here (j) Text (k) Hand Open (l) Hand Closed

Figure 3.7: Visual Tokens

hazards and visualizing the robot work area, as can be seen in Table 3.1. An example set of the visual cues are shown in Figure 3.7 which can help the human partner to understand the intention of the robot. Basic cues can be composed to generate a sequence of instructions or a visual equivalent to a phrase. These, in turn, are translated into a visual message by generating appropriate mixed-reality signals.

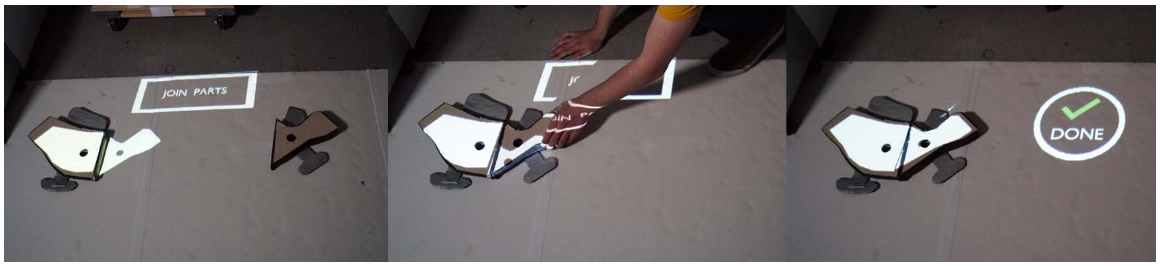
3.3 Visual Planning

Using the basic visual tokens, we can now design and plan a domain-specific meaningful visual language to represent a chain of tasks. This can be of significant help for human-robot collaborative tasks making it easier for the human partner to understand what the robot intends to do or expects the human to do.

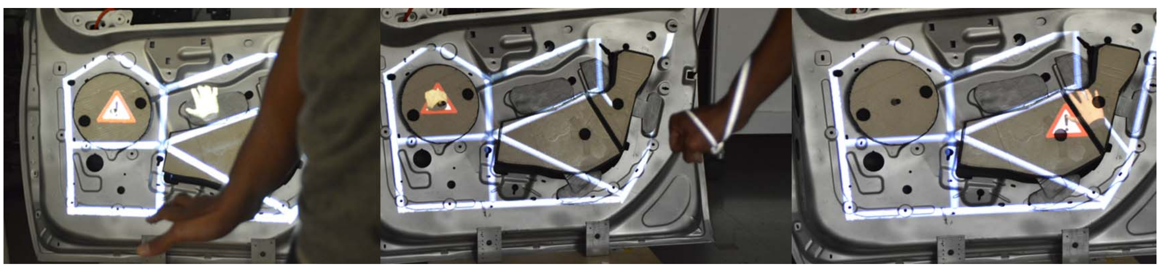
A task can be divided into smaller subtasks, which can be easily described with the visual language. Consider a task of placing two objects and then joining them together. This can be done in a stepwise manner by placing one object at a time and then joining them together, as could be seen in Figure 3.8 (a) (b). The projections clearly simplifies the process by making it more intuitive and easy to understand for the human. Considering another scenario where a human wants to instruct the robot to work on a specific part. The human-tracking module allows the human to control the visual elements which, in turn, can be used to physically control the robot's action. Figure 3.8 (c) shows a human hand projected on the car-door that grabs a warning signal and places it to another location. The new location becomes the next working area for the robot.

A collection of visual cues and interaction metaphors can be used to signal the state of the collaboration, next tasks, etc. For example, the robot can project the boundaries of its work area, communicate information about the success of the current subtask, highlight specific objects, or highlight a particular object part. Similarly, the user may be instructed to move the object to a specified location. In this case, a slider metaphor is used in order to dynamically indicate the remaining amount of translation needed. The robot may also indicate a safe position for the human partner or instruct the user to join specific components. Finally, as can be seen in, the mixed-reality approach also allows us to visualize hidden objects, e.g., the contents

(a) Place Two Objects



(b) Join Two Objects



(c) Grab And Move Warning Sign

Figure 3.8: Visual Plans

of a box. This is particularly helpful in domains where information about content can be derived from bar codes or other types of input that are not human-readable. The approach can easily be applied to different environments and object sets as long as the corresponding 3D models are available. This is, however, typically the case in manufacturing environments.

3.4 Working of the System

The system works in a continuous loop of object tracking, generating visuals based on the tracked object poses, and finally projecting these visuals on the environment in real time. The system keeps tracking objects and projecting information even when the object in focus changes its orientation or position or both.

All the task related visual instructions are created in the projection component which controls other components indirectly by using the inter-component communication framework. For example, one can write simple commands to start or stop tracking a particular 3D object or the human skeleton tracking. This is very useful in generating the visual information when multiple objects are in focus. A simple high level example of the working framework is described in the following manner:

- The projection component requests the interface component for the pose of an object-A, which in turn sends the command to 3D object tracking component to start tracking the object-A.
- Upon receiving the message, the 3D object tracking component starts tracking object-A in the environment and sends the detected 3D pose to the interface component.
- The Interface component checks if there is a pending request for object-A. If yes, a new pose for object-A is computed from the received pose and the extrinsic calibration values. In this way, a continuous stream of pose is sent over to the projection component.
- Based on the final received pose, the projection component generates the visual information and projects it onto the environment.

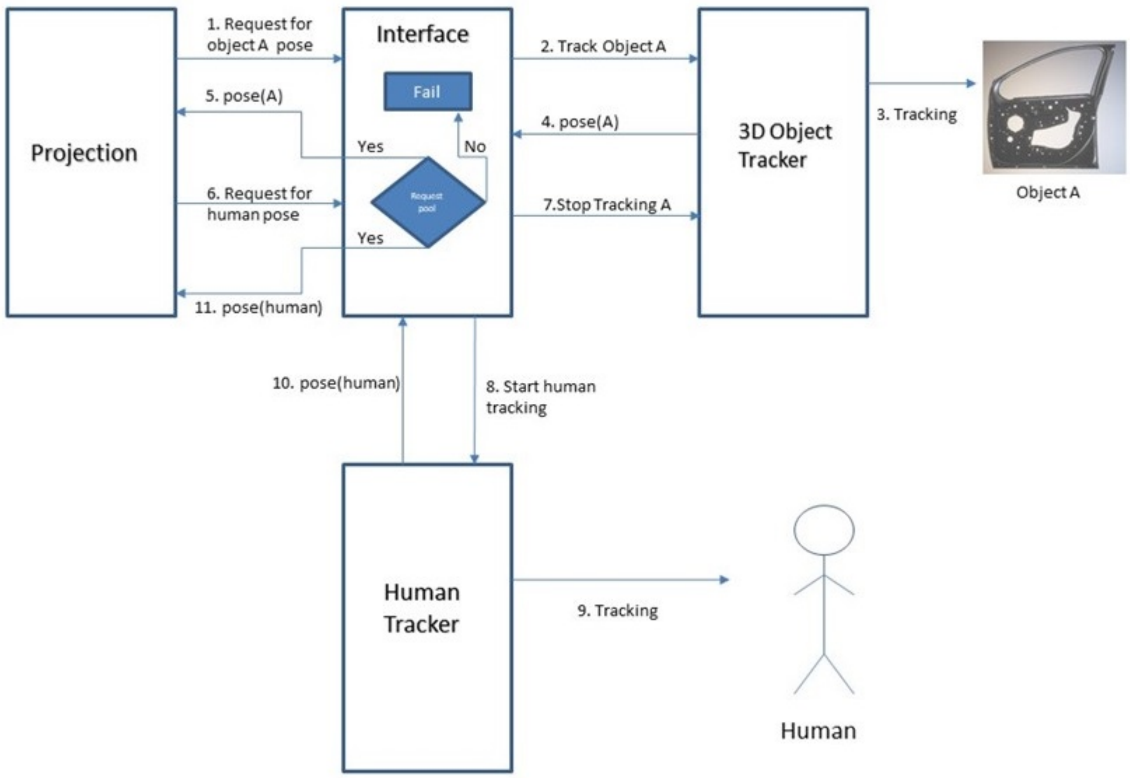


Figure 3.9: A High-Level Working Protocol Of The System

- The projection component can now similarly ask for the pose of a new object-B or a human pose. Hence, a command to stop tracking object-A is sent followed by a start tracking object-B command.

An example test case using the system's protocol is displayed in the figure 3.9, where the the projection component initially requests the pose of object-A and then the human's pose. The idea can be generalized to create complex tasks by tracking one object after the other.

3.5 Salient Features of the System

To this end our mixed-reality projection system is now successfully capable to do the following:

- Track various 3D moving objects in the environment and estimate their correct 6-DOF pose.
- Generate graphical elements and visuals based on the calculated pose and project them onto the environment.
- Provide step by step task instructions to the human in the form of projections.
- Display intentions of the robot and warn the human while the robot is performing its tasks.
- Track human movements in the environment, respond to human actions, and project dynamic visuals.

Chapter 4

EXPERIMENTS

4.1 Objective

It was essential to test the usability and performance of our system with some existing methods prevailing in the industries. Hence, we conducted a human subject experiment to analyze and evaluate how the system's performance when compared to static printed instructions and a mobile display instructions. The aim of this experiment was to measure the efficiency and effectiveness of the human-robot team in all the three scenarios and see which one is the best.

4.2 Modes of Communication

We used three different modes of communication in this experiment which enables us to compare our technology with the existing methods. Detailed step-by-step instructions required to successfully complete the task were provided to the participants in each mode. Participants were asked to carefully read and understand each instruction, one at a time, and only move to the next one after the execution is complete. The three modes are described below:

- **Printed Mode** - Participants were provided with instructions printed on a sheet of paper. The paper was attached on a wall right next to the workspace and was available throughout the experiment. Participants had the option to look back at the previous tasks and make corrections if required. The instructions also contained figures to make it easier for the participants to understand the complex tasks.

- **Mobile Display Mode** - In this mode, the participants were provided with a windows tablet device, Microsoft Surface Pro, which contained graphical instructions in the form of powerpoint presentation slides. Each slide had only one instruction with texts and simple animations which clearly explains the task execution. Few complex tasks also contained demonstration videos to make the understanding easier. Participants had the option to navigate between tasks and go through the instructions multiple times using on screen buttons. Since, this mode uses a mobile device, participants were allowed to carry the device with themselves while executing the tasks.

- **Projected Mode** - This mode uses our mixed-reality system with projected instructions on the workspace. In this way, participants are free from wall-stuck instructions or carry any device around while performing the tasks. This mode provides just-in-time instructions by augmenting the workspace with the visual cues. All that the participants need to do is keep an eye on the projected instructions, understand, execute and wait for the next instruction.

4.3 Method and Procedure

The goal for the participants was to collaborate with a robotic arm (UR5 - Universal Robots) on a simulated assembly task where the human-robot team works on a real car-door. The tasks were designed to resemble a subset of an actual car manufacturing process. The human-robot team performs 12 manipulation tasks involving the car-door and other objects which were to be removed out from 2 toolboxes and placed on the car door after some manipulations. The robot, at some point of time, also uses a drill on the car-door which was handed over to it by the human.

The participants were initially instructed on what the task is all about and how they need to collaborate with the robot. Participants were asked to read, understand and sign a consent form prior to starting the experiments which clearly explains the experiment related information and other necessary details. They had the option to leave the experiment at any point of time if they feel uncomfortable working with the robot. They were made aware about the three different modes and that each mode contains 12 subtasks. 9 out of the 12 tasks were assigned to the participants and the rest were left for the robot. The order of mode of communication was randomized and also the order of tasks within each mode was different than the other modes. This approach makes our analysis and comparison unbiased to any mode. The participants get to take a short break of approximately 10 minutes between each mode during which they fill out a questionnaire related to the most recent performed experiment.

The entire experiment was video and audio recorded which helps us to do the analysis and evaluations. The participants were asked to verbalize their thoughts as they execute the tasks. After completing each mode, participants were required to answer a subjective questionnaire form and also two free response questions about their experience once they are done with complete experiment.

4.4 Experiment Task

The goal of the human-robot team was to assemble a car-door efficiently. The complete task was divided into a set of 12 simple subtasks which were to be performed in a sequential manner. An outline of these tasks could be seen in Table 4.1. The order of these subtasks was varied across the 3 modes to make sure that the participants carefully read and follow the instructions without using much of their prior knowledge from the previous experiment. Few of the subtasks also required the participants to take measurements and thereby, a tape measure was made available.

	Outline of Sub-Tasks
1.	Take out drill from the box and place it on the table.
2.	Take out objects from the box and place it on workspace floor.
3.	Remove the toolboxes out of the workspace.
4.	Move the Car-door.
5.	Wait for the Robot to complete its task.
6.	Join the objects.
7.	Place the joined objects on the Car-door.
8.	Rotate and align Car-door.
9.	Rotate circular object.

Table 4.1: Outline Of Subtasks

4.5 Hypothesis

Efficiency and Effectiveness of a human-robot collaborative team will be greater when the human subjects are provided with just-in-time instructions in the form of augmented visual cues as opposed to printed and mobile display instructions in the form of texts, figures and animations.

Displaying information at the right place and right time is always faster and intuitive. Humans tend to learn better with visuals than from reading texts which can be confusing. A continuous feedback, by the system, on the task execution make humans confident and thereby, increasing the overall efficiency. We defined efficiency as the time taken for the human subjects to complete the task. We defined effectiveness as the accuracy percentage of task completion.

4.6 Measurement Tools

Various tools were used to measure and analyze the team's performance in the experiment. Demographic information was collected from each participant to understand their prior experiences. The time taken for completing each subtask by the human and the robot in each mode, subtask completion, total errors for 4 of the subtasks, were noted to evaluate the team's efficiency and effectiveness of the whole task. After completing each experiment mode, the participants were asked to fill up a questionnaire form which was later utilized for post-task subjective analysis. Each such form contained 17 questions (7-point Likert scale) to be answered based on their experience performing the experiment and also two free-response questions which was analyzed to understand the overall experience of the participants comparing all the 3 modes. Questionnaire items were inspired and adopted from works by (Hoffman, 2013), (Gombolay *et al.*, 2015) and Dragan *et al.* (2015). The form can be found in Appendix B at the end of this report.

Chapter 5

RESULTS AND ANALYSIS

In this chapter, we discuss the post experimental findings and analysis which can lead to approval or disapproval of our hypothesis. We also report statistically significant findings from our experiment. We used a significance level of $\alpha = .05$ for all statistical tests. The measurement tools described in section 4.6 enabled us to derive statistical results which are discussed in the sections below.

5.1 Participants

A total of 15 participants (aged 21 to 48, M = 25.866, SD = 6.424) comprising of undergraduate and graduate engineering students at a substantial urban research university were incorporated into the study. All the volunteers were recruited from the college grounds by means of email and word-of-mouth. Of the 15 participants, 12 were male and the rest were females. 5 participants confirmed that they had experience collaborating with robots in the past and 5 spoke English as a first language. The within-subjects design of the experiment empowered the participants to compare between the three modes of communication. To control learning impact, participants were informed that the three task trials had diverse arrangement of subtasks in assembling the car door, despite the fact that only the order of the subtasks was randomized.

5.2 Efficiency and Effectiveness

To test our hypothesis, section 4.5, which states that the overall efficiency and effectiveness of the human-robot team will be highest in case of the projected mode than in the mobile display mode or printed mode. To measure efficiency of the system, the total task completion time for the human-robot team as well as the total task completion time by the human partner only, between all the three modes, was evaluated. Figure 5.1 illustrates a task completion time versus time comparison between the three modes. It can be inferred that the highest amount of time was taken during the printed mode, and then the mobile display mode, and the least was taken during the projected mode.

An analysis of variance showed statistically significant differences in total task completion times among the different task conditions, $F(2, 42) = 8.07$, $p < 0.01$. Total task completion time in projected condition ($M = 467.73, SD = 135.22$) was lower than the time in the printed condition ($M = 678.60, SD = 165.60$), $t(14) = 8.02, p < 0.00001$ and mobile display condition (M = 606.53, SD = 135.59) , $t(14) = 6.31, p < 0.0001$.

Similarly the subject task completion times between the different conditions showed statistical significance using the one-way ANOVA test, $F(2, 42) = 7.62$, $p < 0.01$. It was noted that, the subject task time in projected condition ($M = 269, SD = 134.23$) was lower than the other two conditions printed condition ($M = 469, SD = 157.14$), $t(14) = 7.61, p < 0.00001$ and mobile display condition ($M = 402.4, SD = 136.12$), $t(14) = 6.66, p < 0.0001$.

To test the effectiveness, we evaluated the percentage of task completion for all the three modes as shown in Figure 5.2. The most number of completed tasks was found in projection mode as compared to others. The percentage of task completion by the human-robot team was computed from the fraction of successfully completed subtasks out of all given subtasks. We compared the three conditions using one-way ANOVA test, and found statistical differences in task completion percentage as a function of the mode of communication, $F(2, 42) = 7.26$, $p < 0.01$. There is not a significant

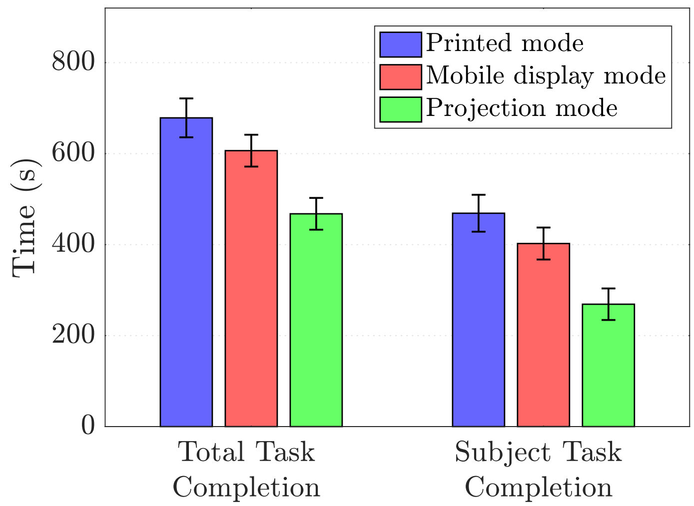


Figure 5.1: Plot For Total Task-Completion And Subject Task-Completion vs Time

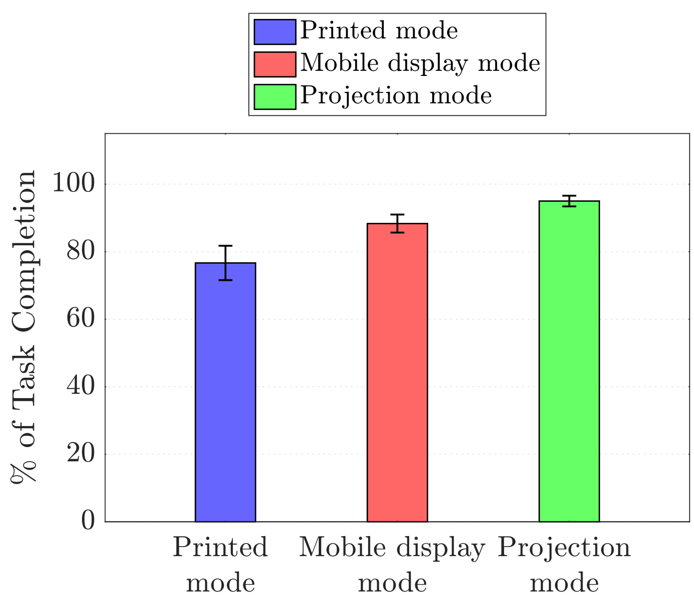


Figure 5.2: Percentage of Task Completion Between All Three Modes

difference between the projected and mobile display mode, however, printed mode performed the worst out of all the three modes.

For relatively complex subtasks such as, car-door alignment and circular-disc rotation we measured the translation and rotation errors which are plotted in Figure 5.3 and Figure 5.4. A significant difference can be seen in the translation errors between projected mode and the other two modes. The error being the lowest in the projected mode. Similar results were found for the rotation errors which also consisted of the circular-disc rotation error. In all the above cases, both translation and rotation errors in projected mode were the least.

Another analysis involved evaluation of the time taken to comprehend each subtask which explains which method is more easy for the participants to understand. Figure 5.5 shows the comparison of understanding time for three subtasks which refers to the car-door alignments. It can be inferred that understanding the first car-door alignment subtask, which does not involve rotation, is easier in case of mobile display mode and projection mode as compared to the printed mode. Whereas, for the subtask 8 (car-door alignment 2) and subtask 9 (car-door alignment 3), a significant difference can be seen in case of the projected mode and the other 2 modes. This means that although it took more time to understand the alignment process in the projected mode, it became relatively simple to perform similar tasks subsequently. This form of objective analysis helps us to justify our hypothesis made in section 4.5.

Subjective analysis was made from the post-task questionnaire and subjective questions which helped us to understand more about the user's state of mind while working with the robot. Overall, participants enjoyed working with the projected mode as it felt game-like to some and also showed likeliness to work with the robot in future, which gives a hint of trust and reliability within the team.

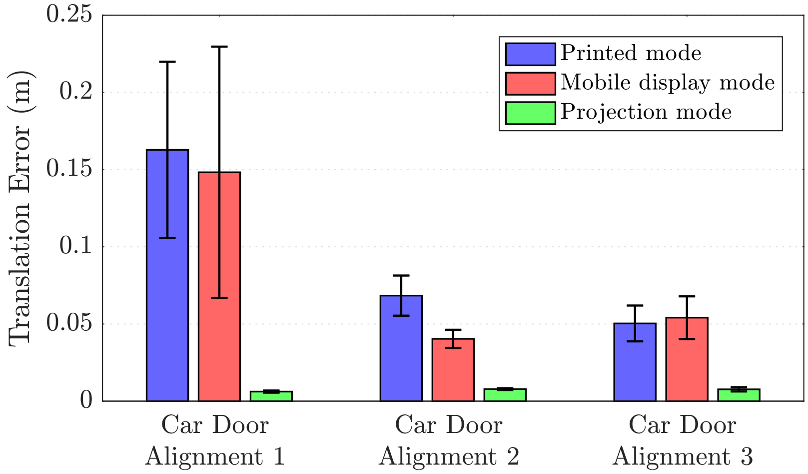


Figure 5.3: Translation Errors For Car-Door Alignment

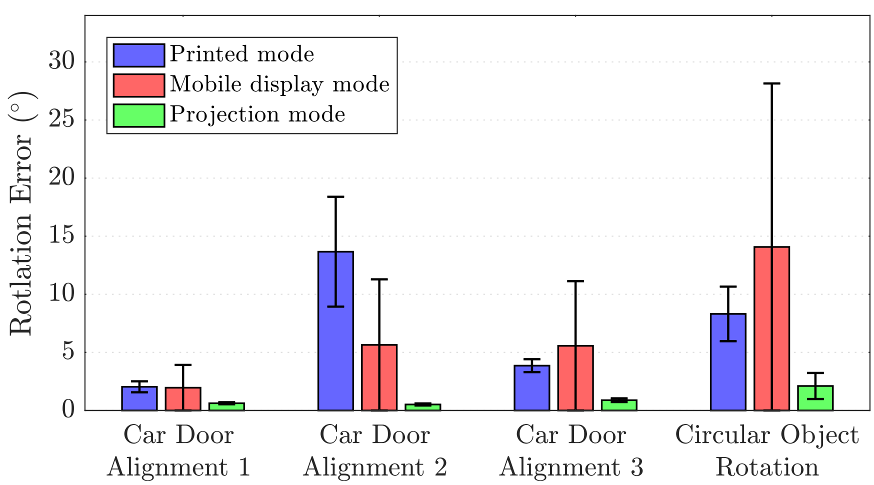


Figure 5.4: Rotation Errors For Car-Door Alignment And Circular Disk Rotation

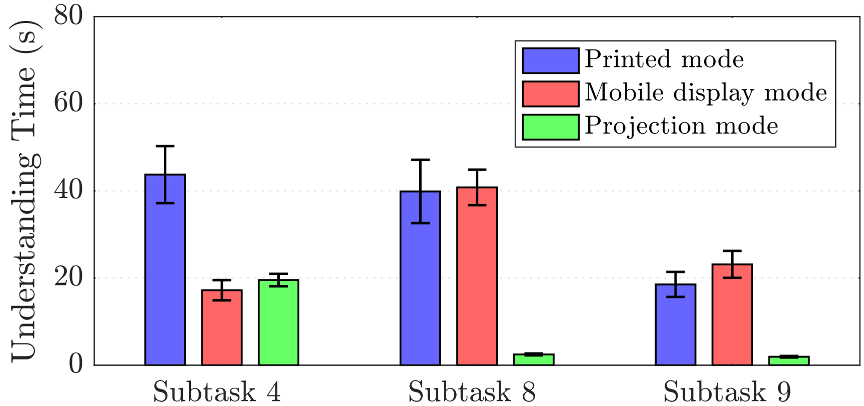


Figure 5.5: Task Understanding Time by Human

Chapter 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In this project, a novel technique for human-robot collaboration has been proposed, which utilizes a mixed-reality system to make the collaboration less demanding and instinctive. The system involves concurrent tracking and projection components that makes it conceivable to utilize real world objects and the environment as the interaction medium for the team to work together flawlessly. The system additionally performs human-tracking which empowers it to adjust and react to human movements with respect to the objects around. Along these lines, a robot can demonstrate its intentions and also take contribution from its human accomplice, thereby, making a superior collaborative environment.

A visual dialect is created which could frame the premise of communication for any mixed-reality framework. A visual signaling dialect is produced utilizing basic visual tokens; practically equivalent to words in English dialect.

A user study was performed to analyze the proposed system. A total of 15 participants volunteered for the examination and every one of them performed 3 tests corresponding to 3 diverse communication modes namely; printed, mobile, and projected. The objective assessment of these tests affirmed our hypothesis, that our mixed-reality framework will result in an increased efficiency and effectiveness in the human-robot team.

From the subjective analysis it can be reasoned that a large portion of the members favored projection mode over the other two modes and contrasted the involvement with game like quality. This perception proposes the chance to investigate further integration of game design concepts to improve the human experience and task performance. The participants experienced higher satisfaction and trust for their robot partner.

It can be concluded that utilization of such an intuitive strategy can be valuable in a human-robot collaborative environment.

6.2 Future Work

The mixed-reality projection system opens up various conceivable outcomes that could help in making it more easy to use and fit for taking care of complex situations. Few of the future improvements I would like to see in this research are as follows:

- Human Action Prediction

  The system is currently fit for tracking human movements and furthermore record the human-tracking information along with the object-tracking data. This information is of high significance since it can be utilized to perform machine learning and identify patterns based on which several predictions related to the completion of the task can be made. Utilizing this, the system could provide additional information about the anticipated success or failure of the given task. Subsequently, the system would become smarter and increase the team's performance.

- Voice Recognition

  It is intriguing to see the integration of voice recognition feature into the system. The user would get easy control over the action of the robot and the robot, on other hand, would have the capability to give audio feedback, which brings the system closer to being a perfect work partner.

- Mobile system

  Our system is currently stationary, which implies that the camera, projector, kinect device, and the robot itself have fixed positions and do not move amid the entire execution, which likewise implies that we get constrained projection area to work in. The system would turn out more robust if the entire setup was mobile. This makes it accessible at various workspaces by continuously following and tracking the human as well as the adjacent objects and projecting useful information around.

REFERENCES

Andersen, R. S., O. Madsen, T. B. Moeslund and H. B. Amor, “Projecting robot intentions into human environments”, in “Ro-man 2016-Proceedings of the 25th Ieee International Symposium on Robot and Human Interactive Communication”, (2016).

Ayres, R., S. Miller, C.-M. U. D. of Engineering, P. Policy and C.-M. U. R. Institute, *The impacts of industrial robots*, AD/A (Dept. of Engineering and Public Policy and the Robotics Institute, Carnegie-Mellon University, 1981).

Canny, J., “A computational approach to edge detection”, IEEE Transactions on pattern analysis and machine intelligence , 6, 679–698 (1986).

Chen, A. T. Y. and K. I. K. Wang, “Computer vision based chess playing capabilities for the baxter humanoid robot”, in “2016 2nd International Conference on Control, Automation and Robotics (ICCAR)”, pp. 11–14 (2016).

Choi, C. and H. I. Christensen, “Real-time 3d model-based tracking using edge and keypoint features for robotic manipulation”, in “Robotics and Automation (ICRA), 2010 IEEE International Conference on”, pp. 4048–4055 (IEEE, 2010).

Christensen, H. I., T. Batzinger, K. Bekris, K. Bohringer, J. Bordogna, G. Bradski, O. Brock, J. Burnstein, T. Fuhlbrigge, R. Eastman *et al.*, “A roadmap for us robotics: from internet to robotics”, Computing Community Consortium and Computing Research Association, Washington DC (US) (2009).

Dragan, A. D., S. Bauman, J. Forlizzi and S. S. Srinivasa, “Effects of robot motion on human-robot collaboration”, in “Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction”, pp. 51–58 (ACM, 2015).

Dragan, A. D., K. C. T. Lee and S. S. Srinivasa, “Legibility and predictability of robot motion”, in “2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)”, pp. 301–308 (2013).

Gombolay, M. C., R. A. Gutierrez, S. G. Clarke, G. F. Sturla and J. A. Shah, “Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams”, Autonomous Robots **39**, 3, 293–312 (2015).

Goodrich, M. A. and A. C. Schultz, “Human-robot interaction: a survey”, Foundations and trends in human-computer interaction **1**, 3, 203–275 (2007).

Hoffman, G., “Evaluating fluency in human-robot collaboration”, in “International conference on human-robot interaction (HRI), workshop on human robot collaboration”, vol. 381, pp. 1–8 (2013).

Iftikhar, M., M. J. Majid, M. Muralindran, G. Thayabaren, R. Vigneswaran and T. T. K. Brendan, “Otorob: Robot for orthopaedic surgeon - roboscope: Non-interventional medical robot for telerounding”, in “2011 5th International Conference on Bioinformatics and Biomedical Engineering”, pp. 1–5 (2011).

Ishii, K., S. Zhao, M. Inami, T. Igarashi and M. Imai, “Designing laser gesture interface for robot control”, in “IFIP Conference on Human-Computer Interaction”, pp. 479–492 (Springer, 2009).

Mainprice, J., E. A. Sisbot, T. Siméon and R. Alami, “Planning safe and legible hand-over motions for human-robot interaction”, IARP workshop on technical challenges for dependable robots in human environments **2**, 6, 7 (2010).

Matsumaru, T. and K. Akai, “Step-on interface on mobile robot to operate by stepping on projected button”, Open Automation and Control Systems Journal **2**, 1, 85–95 (2009).

Omidshafiei, S., A.-A. Agha-Mohammadi, Y. F. Chen, N. K. Ure, J. P. How, J. Vian and R. Surati, “Mar-cps: Measurable augmented reality for prototyping cyber-physical systems”, in “AIAA Infotech@ Aerospace Conference”, (2015).

Perera, V., S. P. Selvaraj, S. Rosenthal and M. Veloso, “Dynamic Generation and Refinement of Robot Verbalization”, in “Proceedings of RO-MAN’16, the IEEE International Symposium on Robot and Human Interactive Communication”, (Columbia University, NY, 2016).

Sato, S. and S. Sakane, “A human-robot interface using an interactive hand pointer that projects a mark in the real work space”, in “Robotics and Automation, 2000. Proceedings. ICRA’00. IEEE International Conference on”, vol. 1, pp. 589–595 (IEEE, 2000).

Stulp, F., J. Grizou, B. Busch and M. Lopes, “Facilitating Intention Prediction for Humans by Optimizing Robot Motions”, in “International Conference on Intelligent Robots and Systems (IROS)”, (2015).

Tellex, S., R. Knepper, A. Li, D. Rus and N. Roy, “Asking for help using inverse semantics”, in “Proceedings of Robotics: Science and Systems”, (Berkeley, USA, 2014).

Verdi, M. P., J. T. Johnson, W. A. Stock, R. W. Kulhavy and P. Whitman-Ahern, “Organized spatial displays and texts”, The Journal of Experimental Education **65**, 4, 303–317, URL `http://dx.doi.org/10.1080/00220973.1997.10806606` (1997).

Xiao, Z., M. Su, Y. Guan, H. Zhang, Q. Yan and Z. Wang, “Wawa: A tumbler-like household robot”, in “2012 IEEE International Conference on Mechatronics and Automation”, pp. 623–628 (2012).

APPENDIX A

DEMOGRAPHIC QUESTIONNAIRE FORM

DEMOGRAPHIC INFORMATION

Please do not write your name on this form. It will be stored separately from any other information that you complete during this study and will not be linked with your responses in any way. The information will allow us to provide an accurate description of the sample.

For the following items, please select the *one* response that is most descriptive of you or fill in the blank as appropriate.

1. Date of Birth (mm/dd/yyyy): ________________

2. Gender:
   - ☐ Female
   - ☐ Male

3. Choose your education level.
   - ☐ Some high school
   - ☐ High school graduate or equivalent
   - ☐ Trade or Vocational degree
   - ☐ Some college
   - ☐ Associate degree
   - ☐ Bachelor's degree
   - ☐ Graduate or professional degree
   - ☐ Prefer not to answer

4. Do you have experience working with a robot?
   - ☐ Yes
   - ☐ No

5. Are you a native English speaker?
   - ☐ Yes
   - ☐ No
   - ☐ Prefer not to answer

APPENDIX B

SUBJECTIVE QUESTIONNAIRE FORM

S no.	Task	Date

1. The human-robot team worked fluently together.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

2. The robot contributed to the fluency of the interaction.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

3. I felt uncomfortable with the robot. (reverse scale)

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

4. I was confident the robot will not hit me as it is moving.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

5. I felt safe working next to the robot.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

6. I trusted the robot to do the right thing at the right time.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

7. I was able to clearly understand robot's intentions and actions.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

8. How satisfied you feel about executing the whole task?

○	○	○	○	○	○	○
Extremely dissatisfied	Dissatisfied	Somewhat dissatisfied	Neither satisfied nor dissatisfied	Somewhat satisfied	Satisfied	Extremely satisfied

9. I was comfortable in interpreting the instructions. The instructions were clear and easy to understand.

○	○	○	○	○	○	○
Strongly Disagree	Disagree	Somewhat Disagree	Neutral	Somewhat Agree	Agree	Strongly Agree

10. I feel that I accomplished the task successfully.

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

11. I was able to assist the robot in completing its task successfully.

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

12. The robot/system provided me with necessary feedback in order to complete the task.

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

13. I would work with the robot the next time the tasks were to be completed.

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

14. How was your attitude towards the task while you were performing it?

Strongly negative — Negative — Somewhat negative — Neutral — Somewhat positive — Positive — Extremely positive

15. The task was mentally demanding (e.g., thinking, deciding, remembering, looking, searching, etc.).

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

16. The task was physically demanding. I had to put a lot of physical effort to complete the task.

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

17. I never felt discouraged, irritated, stressed or frustrated at any point of time during the task execution.

Strongly Disagree — Disagree — Somewhat Disagree — Neutral — Somewhat Agree — Agree — Strongly Agree

S no.	Task **Open response questions**	Date

1. Which form of instruction (Printed/ Mobile/ Projected) will you prefer if you were to collaborate with the robot on a similar task and why?

2. Explain your overall experience working on the collaborative task in all the three scenarios (Printed, Mobile and Projected)?

APPENDIX C

ASU IRB HUMAN SUBJECTS RESEARCH DOCUMENTS

APPROVAL: MODIFICATION

Hani Ben Amor
Computing, Informatics and Decision Systems Engineering, School of (CIDSE)
-
Hani.Benamor@asu.edu

Dear Hani Ben Amor:

On 7/8/2017 the ASU IRB reviewed the following protocol:

Type of Review:	Modification
Title:	Visual Augmented Reality Signals for Human-Machine Collaboration
Investigator:	Hani Ben Amor
IRB ID:	STUDY00005767
Funding:	None
Grant Title:	None
Grant ID:	None
Documents Reviewed:	• Subjective_questionnaire.pdf, Category: Recruitment Materials; • visual-benamor-2-5.docx, Category: IRB Protocol; • Informed_Consent-1-2.pdf, Category: Consent Form; • Demographic_Information.pdf, Category: Recruitment Materials; • Recruitment Flyer-1-2.pdf, Category: Recruitment Materials;

The IRB approved the modification.

When consent is appropriate, you must use final, watermarked versions available under the “Documents” tab in ERA-IRB.

In conducting this protocol you are required to follow the requirements listed in the INVESTIGATOR MANUAL (HRP-103).

Sincerely,

IRB Administrator

cc:
Yash Kanaiyalal Rathore
RAMSUNDAR KALPAGAM GANESAN

Informed Consent

Visual Augmented Reality Signals for Human-Machine Collaboration

INTRODUCTION

We invite you to take part in a research study because we would like to analyze how healthy users between 20 and 50 years react to a novel augmented reality method. The purposes of this form are to provide you (as a prospective research study participant) information that may affect your decision as to whether or not to participate in this research and to record the consent of those who agree to be involved in the study.

Investigator

Principal Investigator: Heni Ben Amor, PhD, Assistant Professor in the Ira A. Fulton School of Engineering at Arizona State University (ASU).

PARTICIPATION REQUIREMENTS

In order to participate, you must be between the ages of 20 and 50 years, have no current arm impairment, have no known neurological disorders, be in general good health, and be proficient in English. If you do not meet these criteria, please inform the researcher.

STUDY PURPOSE

In this study we investigate a novel methodology for human-machine collaboration that uses augmented reality information that is projected into the environment. We are investigating the benefits of this approach when compared to traditional written or oral instructions.

DESCRIPTION OF RESEARCH STUDY

If you decide to participate, then as a study participant you will join a study involving research on human-machine interaction. You will be asked questions about your general health. You are encouraged to notify a researcher immediately if the experimental set-up is uncomfortable at any time so the problem can be fixed. The study session will be recorded through a video camera for a later analysis of your movements and responses. The video will capture your entire body, including posture and face. However, the subsequent analysis will only address your posture and uttered words and questions.

If you say YES, then your participation will last for approximately 45 minutes for today's session in the Centerpoint Building, Room 203-27. Approximately 20 subjects will be participating in this study.

RISKS

Potential risks include temporary fatigue of the arm or the feet. This may occur during the experiment and last for approximately 5 minutes after completion of the experiment. There are no long-term risks to participants.

BENEFITS

Although there may be no direct benefits to you, the results of this study may enhance the scientific understanding of the interaction between robots and humans, that can enhance our understanding about robot-human cooperation, as well as create communication methods for robotic systems to better aid humans. This knowledge may benefit the fields of Human-Machine Interaction, Computer Graphics, and Robotics and has practical applications for the advancement of the control of robotic systems.

NEW INFORMATION

If the researchers find new information during the study that would reasonably change your decision about participating, then they will provide this information to you.

CONFIDENTIALITY

All information obtained in this study is strictly confidential unless disclosure is required by law. The results of this research study may be used in reports, presentations, and publications, but the researchers will not identify you. In order to maintain confidentiality of your records, Dr. Ben Amor or a member of his research team will assign you a random participant ID. Your anonymity is guaranteed by the use of the random code, which will be used to anonymize all data and data collection forms.

This informed consent form will be stored in a locked filing cabinet in Dr. Ben Amor's office. Data files will be stored on computers in secure folders, accessible only by authorized researchers. Data will be retained for 2 years, after which, paper documents will be shredded and electronic documents will be deleted.

WITHDRAWAL PRIVILEGE

It is OK for you to say NO. Even if you say YES now, you are free to say NO later, and withdraw from the study at any time without penalty. Your decision will not affect your relationship with ASU or otherwise cause a loss of benefits to which you might otherwise be entitled. Participation in this study is entirely voluntary and nonparticipation or withdrawal from the study will not affect your grades or employment status.

COSTS AND PAYMENTS

The researchers want your decision about participating in the study to be absolutely voluntary. Should you have any concerns, do not hesitate to talk to Dr. Ben Amor.

COMPENSATION FOR ILLNESS AND INJURY

If you agree to participate in the study, then your consent does not waive any of your legal rights. While no funds have been set aside to compensate you in the event of injury, the researchers do not foresee any risk of injury to you in this study.

VOLUNTARY CONSENT

Any questions you have concerning the research study or your participation in the study, before or after your consent, will be answered by the Principal Investigator: Heni Ben Amor, PhD, Assistant Professor in the Ira A. Fulton School of Engineering at ASU, Centerpoint, Room 203-07, (404) 234-8507.

If you have questions about your rights as a subject/participant in this research, or if you feel you have been placed at risk; you can contact the Chair of the Human Subjects Institutional Review Board, through the ASU Office of Research Integrity and Assurance, at (480) 965-6788.

This form explains the nature, demands, benefits and any risk of the project. By signing this form you agree knowingly to assume any risks involved. Remember, your participation is voluntary. You may choose not to participate or to withdraw your consent and discontinue participation at any time without penalty or loss of benefit. In signing this consent form, you are not waiving any legal claims, rights, or remedies. A copy of this consent form will be given (offered) to you.

Your signature below indicates that you consent to participate in the above study.

___________________________	_________________________	____________
Subject's Signature	Printed Name	Date
___________________________	_________________________	____________
Signature of Investigator	Printed Name	Date