The Pennsylvania State University
The Graduate School

APPLY HUMAN-COMPUTER INTERACTION PRINCIPLES TO AN AUGMENTED REALITY APP FOR HUMAN MOTION SEQUENCE OBSERVATION

A Thesis in
Computer Science and Engineering
by
Haonan Cao

Submitted in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science

August 2021

The thesis of Haonan Cao was reviewed and approved by the following:

Yanxi Liu
Professor of Computer Science and Engineering
Thesis Advisor

Wang Chein-Lee
Asscoiate Professor of Computer Science and Engineering

Chitaranjan Das
Distinguished Professor of Computer Science and Engineering
Head of the Department of Computer Science and Engineering

Abstract

User-centered design is a corner stone of mobile application research. However, the evaluation of User Experience for mobile Augmented Reality(AR) remains a widely unexplored field. Mobile distribution channels such as Apple's App Store and Google's Android Market have created the opportunity to study the behaviors of end-users for improving user-centered design and guidelines. In order to identify the potential problems and possible improvements involving the end-users while using AR applications, it is important to evaluate the user-centered design of current mobile AR service. This thesis focuses on evaluating an existing AR app called AR Taiji with three modifications of the current version following principles in Human-Computer Interaction(HCI). Based on feedback from 20 human raters, we identify factors that are accounted for the effectiveness of developing an AR application in human motion sequence observation. We also report a cross-sectional questionnaire survey with 41 questions for evaluating HCI factors such as satisfaction, efficiency and complexity in the modified versions. Quantitative and qualitative analysis related to user experience is acquired to articulate and understand various aspects of HCI principles. The responses to the questionnaire show a positive attitude of users towards applying HCI principles to AR Taiji. According to the questionnaire survey on three modifications comparing with the baseline, user satisfaction on user interface and functionality experience has an average increase of 13.3%. Score rated on Efficiency has an increase of 22% on average. These results show a positive trend in the additional functions tested.

Contents

List of Figures

List of Tables

Chapter 1 Introduction

In current mobile Augmented Reality(AR) applications research [22] [26] [6] [19], end-users are usually not integrated into the design process of the system. Most of the effort is focusing on optimizing technical functions([20]), developing software or hardware tools [6] or improving related algorithms([26]) rather than User Interface(UI) design or User Experience(UX). As a consequence, mobile applications are still generally a product of function-based design, and the features implemented in terms of interactions have virtually no impact on the design of their functionality, even if they are directly related to the user experience of the system.

This problem is magnified when considering mobile AR applications. With the recent technological developments in computing, Human-Computer interaction(HCI) has changed significantly. HCI, as an academic discipline involving human and technology, is widely concerned with the study of design practices, as well as tools and techniques used [1]. HCI is a primary study on the way people interact with computer technology. With the development of mobile devices recently, HCI is no longer limited to desktop computers. Nowadays mobile devices are capable of launching an AR app with their high-resolution camera module, higher RAM, and speedy CPUs. The importance of algorithm optimization is decreasing, while HCI principles and user-based design are highly valued. HCI design is an activity that explores value and meaning of information technology, based on understanding, insight and imagination. This advanced way of design thinking is the source of innovation. Thus, the knowledge base of general Human-Computer Interaction could be applied to a mobile AR app as well.

Due to the characteristics of mobile use, people hope that mobile apps have a simple and direct way of interaction. Researchers have been trying to introduce more natural interaction methods into the design of mobile phones, such as gesture interaction, expression interaction, and voice interaction. Among them, voice interaction has always been a hot topic in the computer field. Compared with desktop systems, mobile devices are suitable for integrating voice control in a complex environment due to the fact that microphones are common on mobile devices.

In an augmented reality system, the real world remains central to the experience and is enhanced by virtual elements that generally interact with it. For an AR system to reach its full potential, what is now required are "new paradigms that support heads-up information presentation and interaction, seamlessly integrated with viewing and interacting with the real world" [15].

Additionally, today's AR system is more than showing floating objects on the screen. Features like plane detection and world tracking excel in the AR field by making the virtual world more believable. World tracking aligns the pose of an AR camera that renders virtual objects with the device's camera, making virtual objects feel as if they are part of the real world. Plane detection is commonly used in the initialization step where the user drops virtual content on a carpet or the surface of a desk.

This thesis applies HCI and user-centered design to an existing AR Taiji [28] app, where several useful AR technologies are being used. For example, the baseline app of this thesis work, AR Taiji, embodied a powerful plugin for IOS called ARKit. In the remaining part of this chapter, we will present a brief introduction to those technologies and methods.

1.1 ARKit

ARKit is a new framework added to the IOS 11 system released by Apple on June 6, 2017. ARkit helps IOS developers implement AR technology functions in the simplest and fastest way on mobile devices. The ARKit framework provides two AR technologies, one is augmented reality based on 3D scene (SceneKit) and the other is augmented reality based on 2D scene (SpriktKit).

Tracking: Tracking is the core function of an AR system, which means that devices can be tracked in real time. World tracking can provide the relative position of the device in the physical environment. With the help of the recently developed motion tracking algorithm Visual-Inertial Odometry (VIO) [17], an accurate view of the position and the orientation of the device could be easily obtained by developers. The VIO uses camera images and device motion data. More importantly, there is no need for peripherals, no need to know the environment in advance, and no additional sensors are required.

Scene Understanding: The upper layer of tracking is scene understanding, which means determining the attributes or characteristics of the surrounding environment of the device. It will provide functions such as plane detection. Plane detection can determine the surface or plane in the physical environment e.g., the floor or the table. In order to place custom virtual objects on a surface, a hit test function is implemented in the plane detection. This function can obtain the intersection point with the real world topology in order to place virtual objects in the physical world accurately.

Additionally, scene understanding can be used for light estimation. Light estimation is used to correctly illuminate the virtual geometry to match the physical world. Combined with the above functions, virtual content can be seamlessly integrated into the physical environment.

Rendering: The last layer of ARKit is rendering. The continuous camera image stream, tracking information, and scene understanding can be imported into any rendering program. It's simple for the developer to custom any desired AR view using rendering layer.

1.2 Motion Capture(Mocap)

Motion capture(Mocap) refers to the technology that records and processes the actions of people or other objects. It is widely used in the fields such as entertainment, medical applications, computer vision and robotics. In the field of App development, it usually records the actions of human actors(Figure 1.1(a)) and converts them into digital model(Figure 1.1(b)), and generates two-dimensional or three-dimensional computer animated motion sequence.

(a) Recordings of human actions

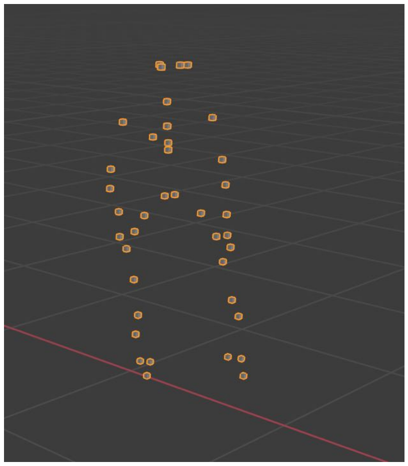
(b) Digital animation of Mocap data

Figure 1.1. Converting motion captured human actions(a) into animations(b)

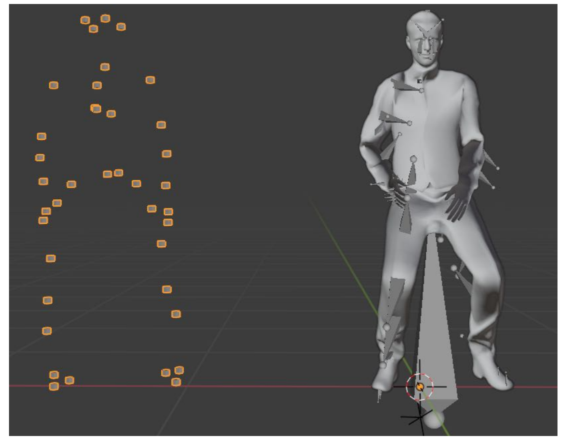
The greatest benefit of Mocap is that it expresses human movement in a physically accurate manner. Because various elements such as muscle, human frame, and placement of the center of gravity have an influence on human motion, models animated manually often make an unnatural impression [12]. An AR app has the merit of creating a realistic world by rendering virtual content as if it was in the real world. Merging AR and Mocap would favorably increase authenticity in user experience, which has already been done in AR Taiji app [28].

1.2.1 Animations

To convert complex Mocap data generated in the lab to animated virtual characters in an AR app has to be done meticulously and carefully. The challenge of this task is that, since the body joints on raw Mocap data has a different range of motion from those of the virtual characters, the range of motion on the virtual character must be manually defined. Thus a proper body joints alignment is required to maintain the quality of the model. Figure 1.2(a) shows the positions of 39 markers placed on a human subject (yellow dots), which provides precise information about poses and motions. Figure 1.2(b) displays a mapping from raw Mocap data into the virtual character with proper rotation, translation and scale.



(a) Raw Mocap data [16]



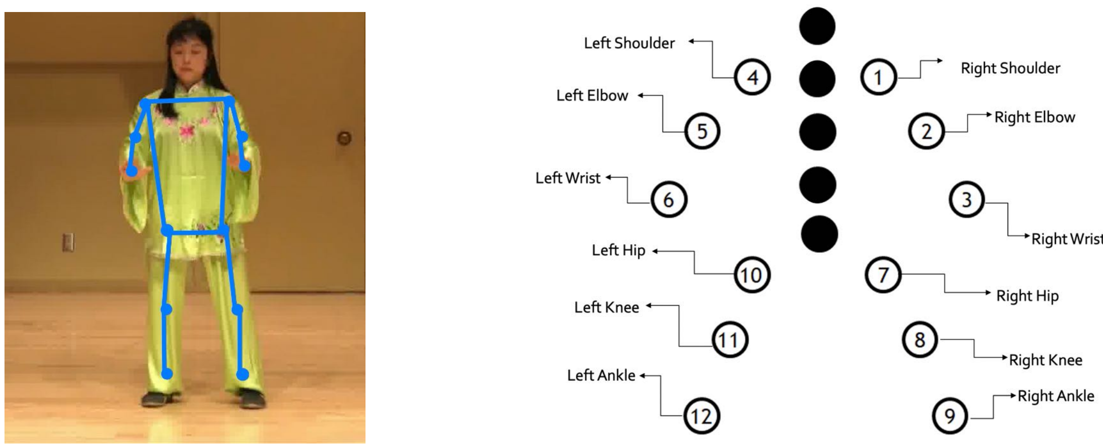
(b) Align raw data with a model [7]

Figure 1.2. Animate virtual character using raw Mocap data

12 Body joints list is as follows:

1. Right Shoulder
2. Right Elbow
3. Right Wrist
4. Left Shoulder
5. Left Elbow
6. Left Wrist
7. Right Hip
8. Right Knee
9. Right Ankle
10. Left Hip
11. Left Knee
12. Left Ankle

Figure 1.3(a) shows those 12 joints on a human who is performing the 24-form simplified Taiji Quan [21]. All 24 key poses of the 24-form simplified Taiji Quan are shown in Figure 1.4. Figure 1.3(b) demonstrates the positions of 12 body joints and their corresponding names.



(a) 12 Joints on human body [16]

(b) 12 body joints with labeled names

Figure 1.3. 12 body joints list on human body with labeled body joints names

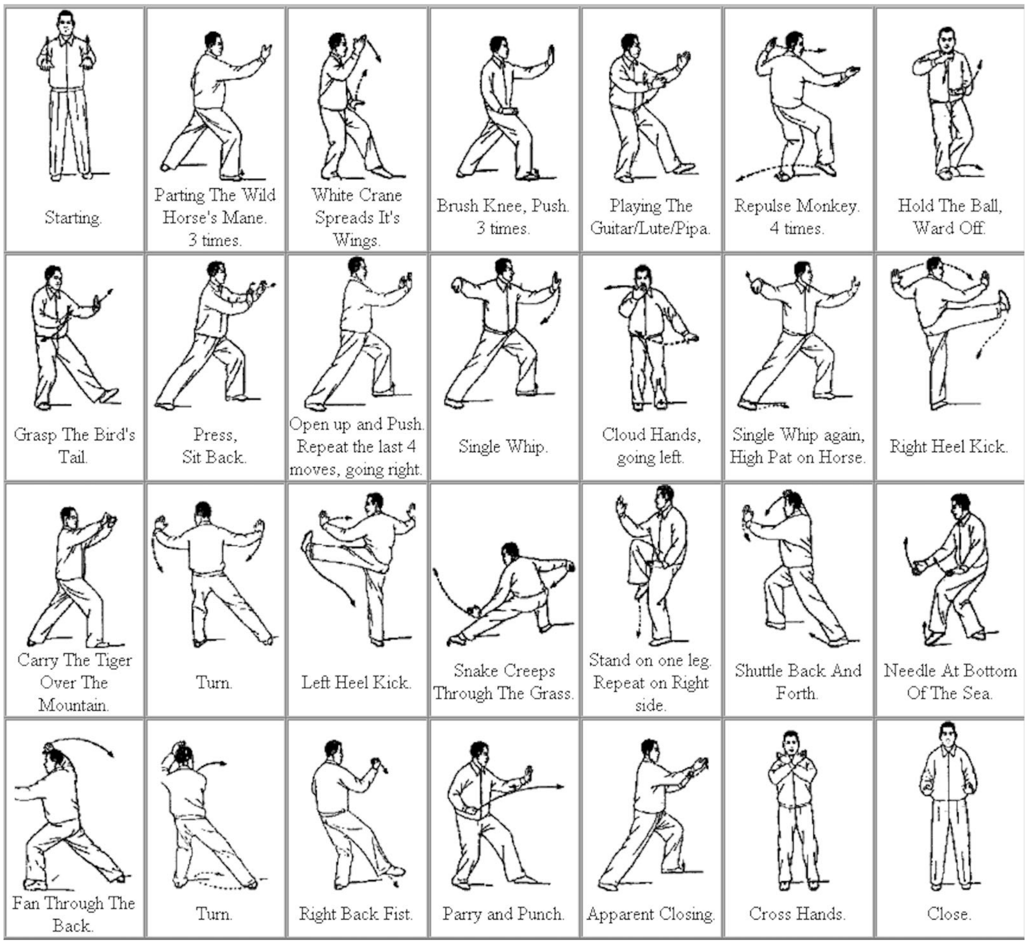


Figure 1.4. Key poses of 24-form Taiji [9]

In order to address these issues mentioned previously within mobile AR apps in this thesis, we propose a set of design principles and guidelines for mobile AR systems evaluation, using an AR app with human motion sequences as apparatus. We have divided our experiment into four phases, with each phase corresponding to one specific version of AR Taiji app. The contributions of this thesis work is: (1)Explore various HCI principles in an AR app called AR Taiji, make three modified versions of AR Taiji that implement advanced layouts, new virtual characters, voice control and an VR option. After three modified versions, user satisfaction on user experience and functional experience has an average increase of 15.2%. After three modifications, score rated on Efficiency has an increase of 22% on average. Score rated on Efficiency also has an 27% increase with the addition of Voice Control. Screen Area Division for virtual content has a 20% increase. User satisfaction on screen area division increases by 60%. (2)Design a questionnaire that evaluates the effectiveness of developing an AR application in human motion sequence observation. Conduct distinguished analysis based on gender, first language and screen size (3)Pave the way for the development of AR Taiji for a better user experience. Specify the direction for functions to be iterated in the future.

In the remainder of this thesis, we first present related work in the field of mobile AR app and HCI principles. Then, we describe our experimental design and survey mythology. Before the final conclusion and future work, we present quantitative evaluation results from our survey and statistically analysis of the results.

Chapter 2 Related Work

Most research on App development focuses on optimizing technical issues rather than UI design and user experience [22] [26] [20] [6] [19]. These technical issues includes optimizing technical functions([20]), developing software or hardware tools [6] and improving AR-related algorithms([26]). Therefore, an AR application is usually a product of function-centered design. Even if those functions are directly related to user experience and satisfaction, the functions implemented in terms of interaction do not actually have any effect on its functional design. [25] and [4] introduce two common models for modern design process of an application. Waterfall [25], also referred to as a linear-sequential life cycle model, divides the process of software development into separate phases(see Figure 2.1). Typically, the outcome of one phase acts as the input for the next phase sequentially. User experience is not integrated in any of those stages, thus eventually leading to a function-based product.

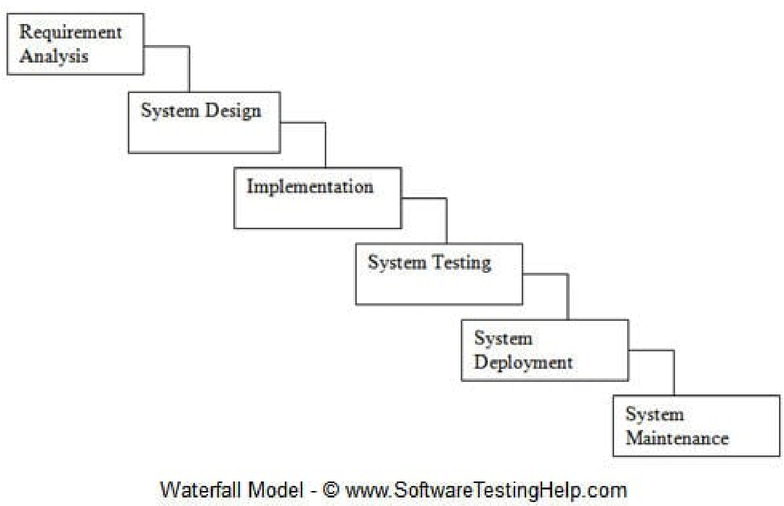


Figure 2.1. General process of Waterfall model [25]

The Spiral Model [4] is a refinement of the classical Waterfall [25] model(see Figure 2.2). The advantage of this model is that it combines iterative development with controlled aspects of water fall model, but with continuous refinement and implementation.

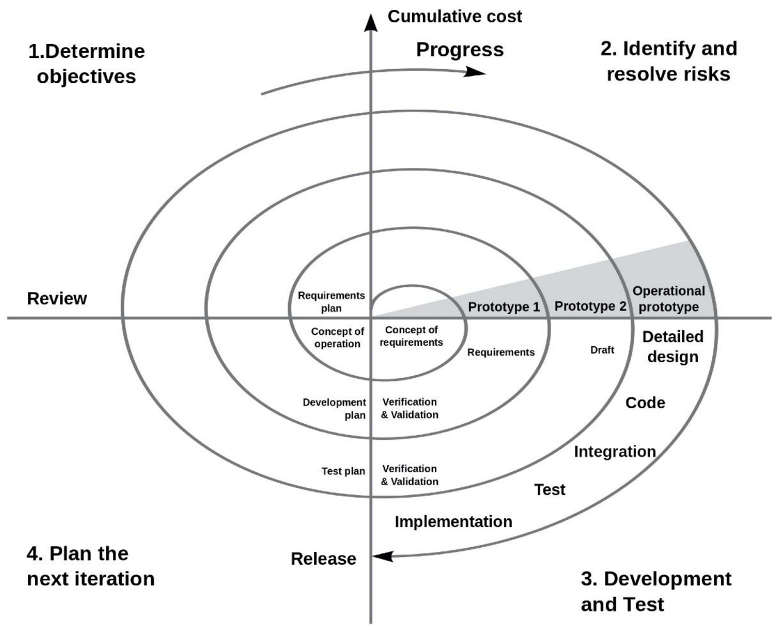


Figure 2.2. General process of the Spiral model [4]

2.1 User-based Design and Principles

However, these two models [25] [4] only pay attention to function-centered principle rather than user-based design. Due to the characteristics of mobile use, nowadays the demand of Human Computer Interaction(HCI) and User Experience(UX) is growing, especially in AR service that is usable for people with a varied range of abilities and expertise.

In 2007, [3] introduced some useful user-centered design principles and related them to the demands of AR system. This becomes an initial attempt to fill the gap that currently exists in this area. [3] also provides a general review on HCI principles that are important, e.g., Flexibility, Affordance, User satisfaction and User privacy. Affordance means that design must conform to users' needs based on users' physical capabilities. For example, A menu button using familiar shapes of the real props is subject to Affordance. Flexibility means different preferences on the input command method e.g., voice recognition or gesture detection(Figure 2.3). Users may have different preference on the input method thus it's vital to follow Flexibility in the app design. User privacy is the principle that protects users' privacy from being violated e.g, camera usage or location usage. Thus designer should always ask for the permission of camera in the very first launching of the app.

Figure 2.3. Input command method: voice recognition

Moore [5] makes an attempt to test the usability of a tangible AR system with a heuristic evaluation. [6] reviews the research and development of augmented reality applications in design and manufacturing. It mainly focuses on hardware, software and technical challenges in AR, as well as design and manufacturing issues in industry. Our research mainly focuses on HCI principles in each revision in additional to an existing AR app.

[10] has demonstrated how rich UIs, such as augmented reality systems, may be integrated into efficient, business-oriented methods without compromising on either the quality of the functionalities or the system's usability. [10] also demonstrates a new design life cycle called Symphony development cycle. Symphony is split into three design branches: Functional, Technical and Central(see Figure 2.4). The approach is directed by four core principles, which are: (1) a seamless integration of HCI practices; (2) collaborative design activities; (3) traceable and consistent models for collaboration and communication; and (4) a pervasive tooling support.

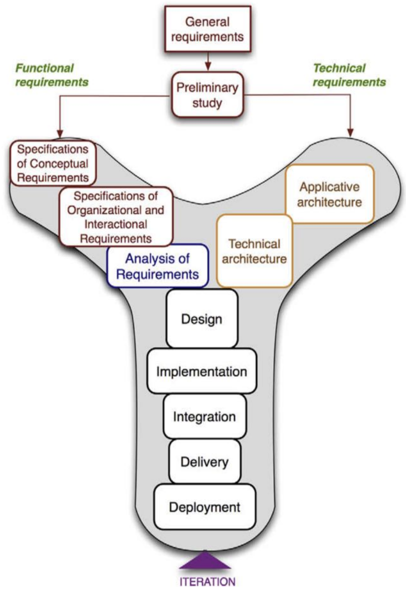


Figure 2.4. Symphony development cycle [10]

2.2 Experiments with Human Survey

Similarly to our survey mythology, [23] and [14] report several experiments that they had conducted by publishing apps in the Market(Figure 2.5). The apps from [14] are freely available and have been installed more than 30,000 times. The outcomes of the experiments range from failure to valuable insights. Based on these outcomes they identified factors that account for the success of experiments using mobile application stores. The differences between [23] and this paper is that [23] investigates on the user experience of 5 totally different AR apps while this paper explores only one AR app and three enhanced versions. [23] and [14] study on various themes of AR app, such as AR map, AR game. Yet AR Taiji only focuses on human motion sequence observation in 3D. Current effort in this thesis mainly focuses on applying HCI principles for functional variation evaluation.

Figure 2.5. Screenshots of one of the AR apps in [23]

[14] has some similarity to this work, which embodies a pilot survey study to evaluate several Augmented Reality applications, which are specially designed on mobile devices. A mix of quantitative and qualitative UX evaluation and the opinion of respondents about their MAR experience is the key contribution of this study.

The baseline app of this thesis work, AR Taiji [28], is currently available for free download on App Store. A detailed description of current functions is introduced in Chapter 3.

Chapter 3 Evaluation Design for AR Taiji App and Its Modifications

In order to investigate different features of HCI in a mobile AR app, we divide the evaluation into four versions, with AR Taiji app [28] as a baseline model. In each version except version 1 - baseline, we make some modifications to AR Taiji with respect to the corresponding principles of HCI. Feedback is collected by conducting an online questionnaire based on the modified content of each version of App, where we combine quantitative analysis to AR Taiji with qualitative research strategies.

3.1 Version 1: AR Taiji App (Baseline)

In Augmented Reality, users have the ability to build a virtual world using object detection and world-tracking. To make virtual characters in a real environment more vivid human, motion capture animations become one of the best choices. Mocap technology allows the recreation of complex movement and immersive physical interactions. The AR Taiji Mocap data is a publicly available dataset from PSU LPAC (Laboratory for Perception, Action and Cognition) motion capture lab [16], containing scripted motion sequences of 24-form Taiji (Figure 1.4). The reference Mocap model contains a time series of 12 labeled 3D joint positions (Figure 1.3) providing information about position of body segments and movement of each labeled joint with a joint position accuracy of less than 0.5mm [16]. There are 12 joints in each frame, which are mapped into body joints on a virtual character model(Figure 1.2).

The baseline model is called AR Taiji [28], which uses SceneKit that is based on 3D scene for virtual content implementation. The AR system of AR Taiji [28] can be divided into three layers: Tracking, Scene Understanding and Rendering as described in Chapter

1. Users are able to watch the whole sequence and every gesture of 24-form Taiji in a spatially accurate manner from this 3-layer structure, without loss of 3D realism(see Figure 3.1).

Figure 3.1. The baseline version of ARTaiji [28]

Compared to watching a video which could only see one specific view of a Taiji teacher, users can learn 24-form Taiji from all angles around a virtual character in AR Taiji app. This App makes it possible for users to learn from the best angle, and to pause the virtual character at a specific pose while imitating it frame by frame. Instead of only seeing the teacher from a certain angle, users of AR Taiji are able to learn Taiji from any angle that they desire at any time (see Figure 3.2). Beside the viewing angle, the position and size of virtual character can also be controlled easily by using Translation and Scaling functions (Figure 3.3).

Front view Rotate by 90° Rotate by 180° Rotate by 270°

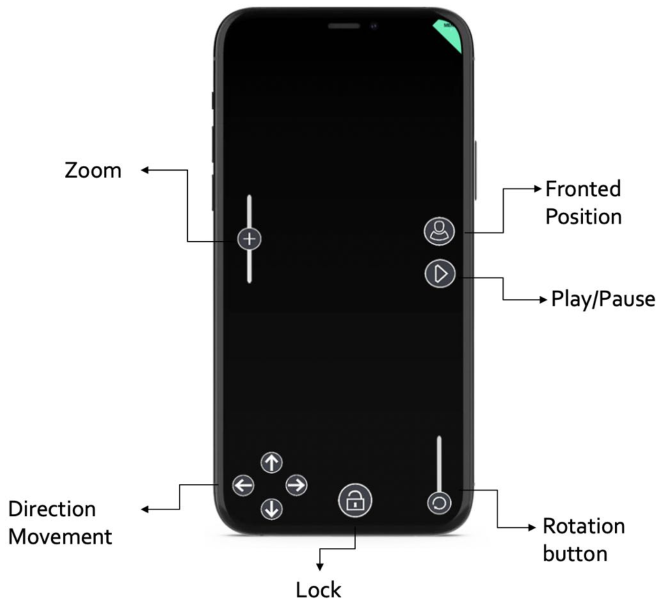
Figure 3.2. A sequence of 3D views with clockwise rotation(top-down view) about the z-axis in AR Taiji app

(a) Original position

(b) Translate and scale the virtual character to a desired position

Figure 3.3. A sequence of operations: Translation and Scaling

We use the current version (currently available for free download on App Store [28]) as a baseline model. This model integrates a 2D graphic control panel (see Figure 3.4.), and only asks for permission of the camera and microphone at the first launching of the App.



(a) Control panel of AR Taiji

(b) 24 forms selection menu (topright green spot)

Figure 3.4. Control panel and form selection of Baseline model

By clicking anywhere on the screen, users can drop a virtual character on any planes or surfaces in the real world(Initialization, see Figure 3.5). While watching 24-form Taiji sequences, users can easily control the position and direction of the virtual character by using the control panel on the screen(see Figure 3.4(a)). Click the top-right green menu gives a list of 24-form Taiji that the user can let the virtual character perform the corresponding form(see Figure3.4(b)). Here all operating functions are displayed straightforwardly to users. The effect of operations e.g. Transformation, Rotation is showed instantly on the character.

(a) Initialization on the surface of a desk

(b) Initialization on the floor

Figure 3.5. Drop the virtual character on different planes(Initialization)

Users can perform various operations on the virtual character. After enabling the character by dropping it in the real world, users can perform a sequence of internal activities on AR character, e.g., Rotation, Scaling and Translation. Such sequences of activities that involve position-tracking and coordinate alignment are all carried out by the Transformation manager (see Figure 3.6).

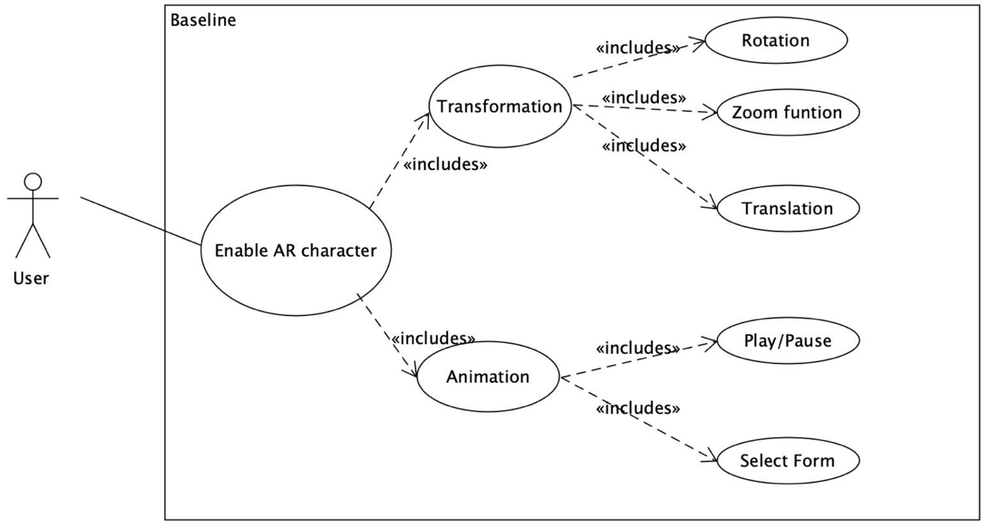
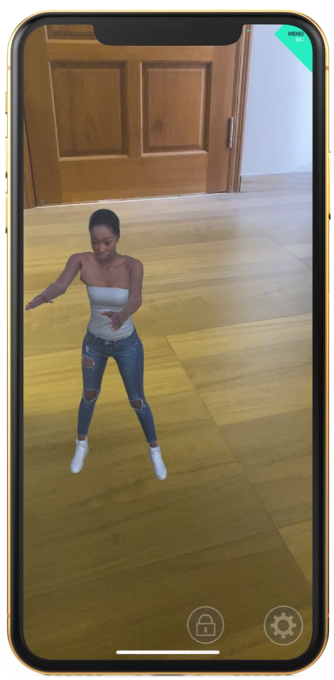
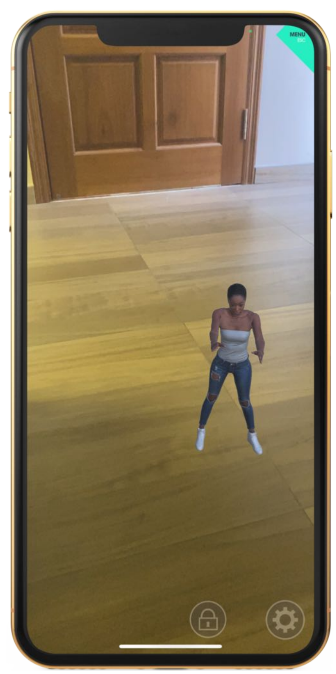


Figure 3.6. Use cases of baseline model

Figure 3.2 shows how to rotate the virtual character for a better angle. Figure 3.7 illustrates how translation and scaling functions work on the virtual character.



(a) Original position



(b) Translate and scale the virtual character to a desired position

Figure 3.7. A sequence of operations: Translation and Scaling

3.2 Version 2: Appearance Modifications

Affordance is one of the most vital principles in mobile HCI [3]. The concept of Affordance suggests that there should be an inherent connection between a user interface and its functional and physical properties [3]. As we know, User Interfaces in mobile AR are more tangible than those in desktop computers since an AR system not only augments physical properties through its screen, but also increases information perceived by users. Thus applying this design principle to an AR system can take advantage of the Affordance for direct 3D manipulation in AR systems [11]. Therefore in Version 2, (1)an implicit layout is designed for AR Taiji app (Figure 3.8), which potentially enhances the connection between UI and its functions. (2) Two new virtual characters(Figure 3.10, Figure 3.11) with different features are introduced, which increase diversity of the app.

3.2.1 Part 1: User Interface Layout

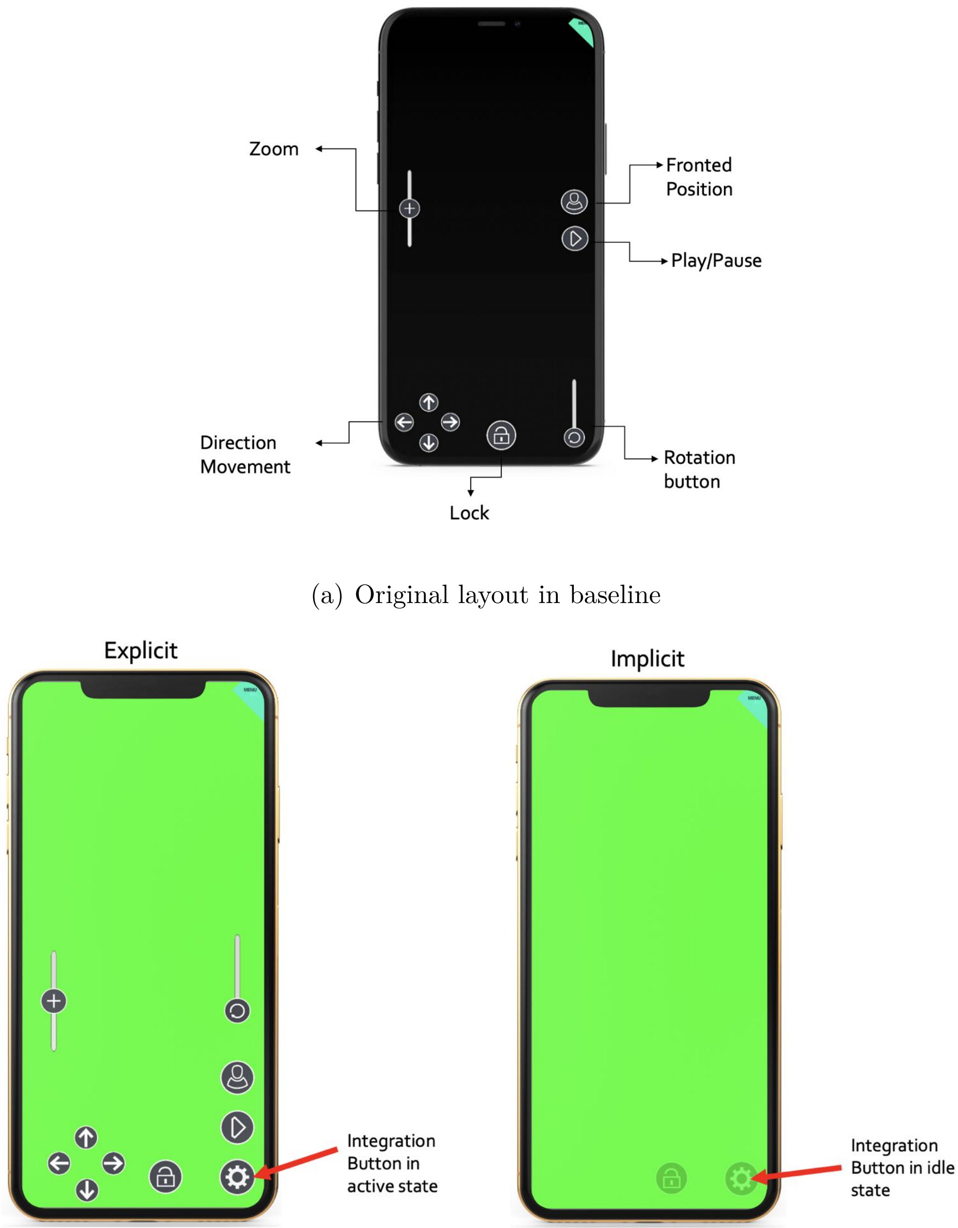
The first objective on Version 2 is to explore the layout of UI in a mobile AR app. In the baseline, nearly 25% of the screen is covered by an explicit UI(Figure 3.4), sacrificing a significant amount of user perspective. The Advantages of such an explicit layout are clear: (1)This layout could maximize the efficiency of the UI system. Any interactions resulting from users' operations can be reacted instantly. (2)Since this layout requires less steps of operations, such layout could also reduce additional operations, thus reduce the possibility of fatigue.

However, since too much screen is covered(Figure 3.4(a)), such an explicit layout increases the possibility of unintended control button touch as well. Since the buttons are arranged on both sides of the screen, the user has a high probability of accidentally touching buttons near the edge. Such layout could also lead to a reduction in the immersion of the AR world.

Hence now there are two different design ideas in UI. One is an implicit layout which maximizes user immersion by saving the space (Figure 3.8(c)). The other is an explicit layout that covers most of the screen area(Figure 3.8(b)). These two kinds of layouts can be considered as a trade-off between user efficiency and user satisfaction. In order to investigate this trade-off, we make two modifications on the baseline as the experiment of Version 2 Part 1.

1. The first modification is an implicit layout with an integration button that folds all other control units(Figure 3.8(c)).

2. Secondly, after one second of idle time, the integration button becomes translucent, which only takes about 5% of screen space. Users can expand it by simply click the gear button (Figure 3.8(b)).



(a) Original layout in baseline

(b) Explicit layout

(c) Implicit layout

Figure 3.8. Add an integration button that integrates all other buttons

Such modifications involving extra physical efforts, i.e., click the integration button to go from implicit layout to explicit layout, which may potentially increase cognitive overhead, which is contrary to the universal principle of HCI. Therefore, It is especially significant to investigate whether users are biased towards efficiency(Version 1 baseline Figure 3.4) or an implicit control panel(Version 2 layout Figure 3.8).

3.2.2 Part 2: Animated Characters Option

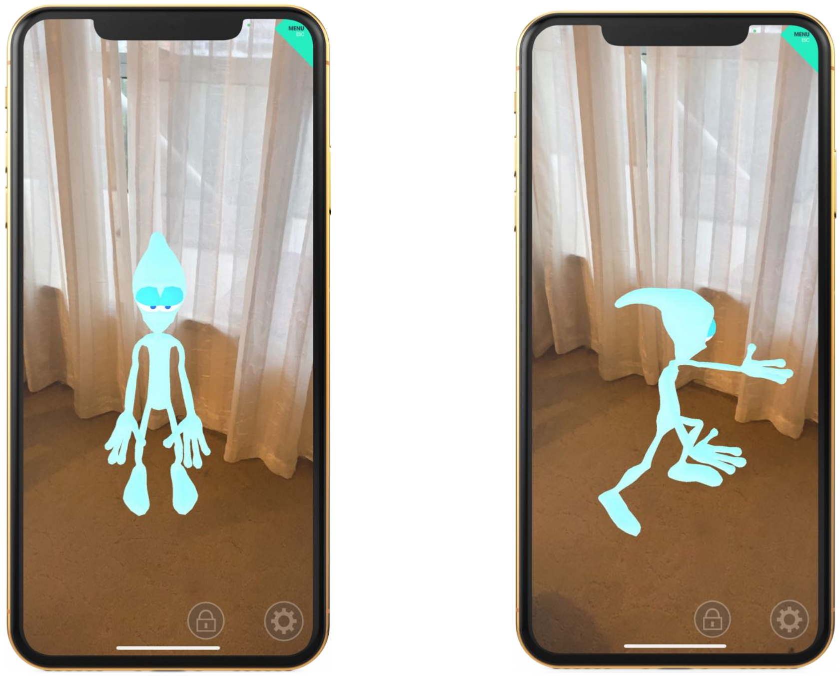
One merit of applying Mocap to the virtual characters is that Mocap technology allows multiple models to share one Mocap sequence. As long as the virtual models have the same number of joints, one set of Mocap data can be applied to different characters with different appearances.

The second goal is to study the influence of virtual characters' appearance. Since AR Taiji app can be possibly used by Taiji learners of all ages, cartoon characters have a promising attraction to young users, especially to children. Meanwhile, since this app includes human motion sequence from Mocap data, an angular model that reveals the contours and body joints may demonstrate the poses and movement of 24-form Taiji clearly, thus increasing the learning efficiency. Therefore, in the Version 2 model, we provide a list for character selection, where users can choose a character with preferred features as their Taiji teacher. See the difference in Figure 3.9

(a) A female character in starting pose

(b) Female character in key pose #2(see more key poses in Figure 1.4)

Figure 3.9. Two poses of the original character on AR Taiji: Female character



(a) Alien in starting pose

(b) Alien in key pose #2

Figure 3.10. Two poses of the first new character: Alien

(a) A male character in starting pose

(b) Male character in key pose #2

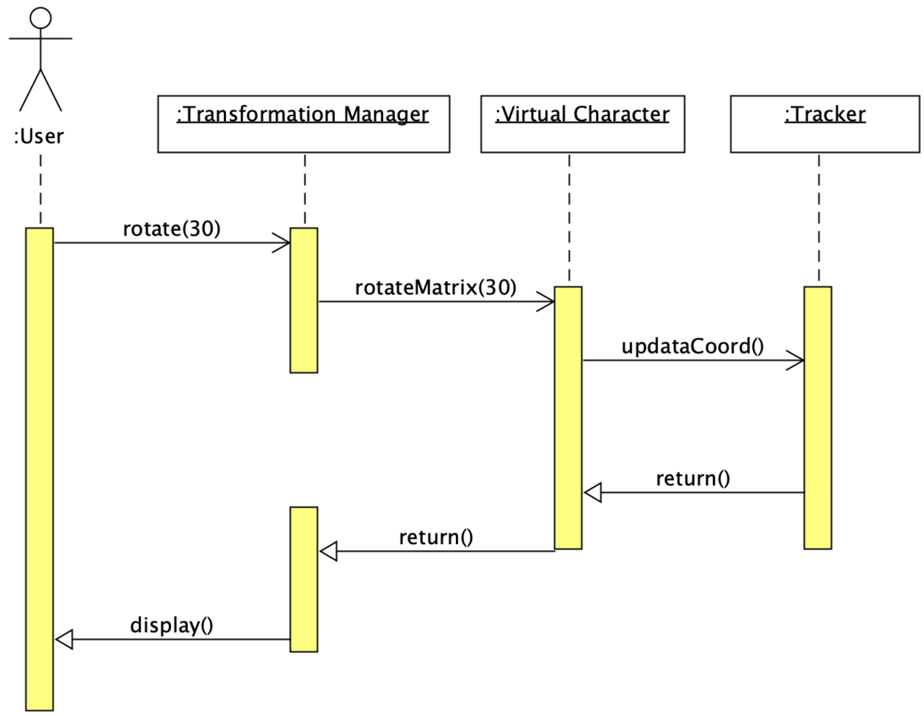
Figure 3.11. Two poses of the second new character: Male character

Figure 3.9 is the animation of a female character. It shows the human body joints and posture more clearly, which would help with human motion sequence learning. On the contrary, the cartoon-like character in Figure 3.10 sacrifices the layered lines on contours in exchange for a more adorable appearance. Figure 3.11 shows another animated character that provides a clear view of joints and poses. We are to evaluate which model is more attractive to users in this enhancement of AR Taiji.

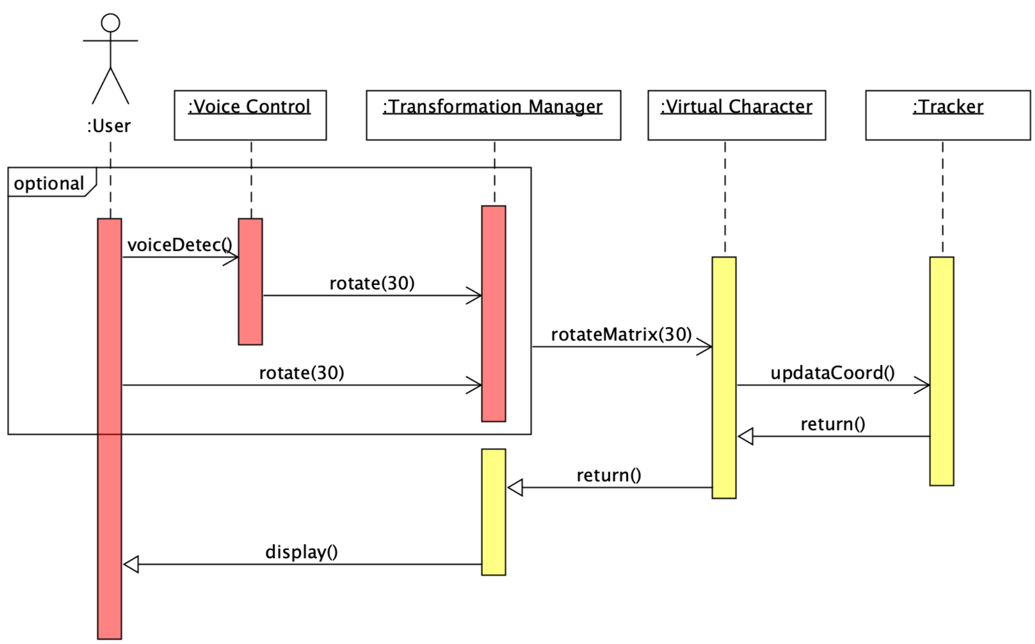
3.3 Version 3: Voice Control

Another important principle in Human-Computer Interaction that we want to investigate in AR Taiji app is **Flexibility**. This principle takes both user preference and ability into consideration [3]. In a mobile AR app, some users may prefer controlling everything with both hands, while others have different preferences on the input command method. The main idea of applying Flexibility into an AR app is to accommodate individual preferences by providing alternative options for users. Therefore, in the third version, we add voice recognition, which allows users to navigate and interact with AR Taiji by simply speaking out the commands.

In the baseline, operation logic is explicit and straightforward. Suppose a user wants to rotate the virtual character by 30 degrees, the only approach to do this is turning the spin button(see Figure 3.2). This operation will then be sent to a unit called *Transformation Manager* which handles all transformation operations and converts everything into matrices. After that, the coordinate stored by the character and *Tracker* will be updated simultaneously. The Sequence Diagram for this process is shown in Figure 3.12(a). *Tracker* is a unit used for world-tracking. It allows the AR camera to memorize the location of the virtual character so that even if it is removed from the field of view, the virtual character would remain where it was.



(a) Before modification: Sequence Diagram of baseline



(b) After modification: Sequence Diagram of Version 3

Figure 3.12. Sequence Diagrams for Version 3

Time sequence becomes different after adding a voice control unit, see Figure 3.12(b). Suppose we want to rotate the virtual character by 30 degrees as well. There is a new modality provided before *Transformation Manager*, which is called Voice Control. This is the place where Flexibility is applied: users can either rotate the virtual character by giving an order to Voice Control first, or they can spin the button directly.

Integrating various kinds of input methods is an interesting feature of AR app. Therefore, accommodating different user preferences plays a significant role in mobile HCI. Users are able to free their hands by speaking instructions directly. This means that they are able to observe the human motion sequence without touching the screen. Voice control complements the current visualization control of the human motion sequence by providing a new input modality.

3.4 Version 4: VR Environment Option

The management of personally identifiable information is considered as a key factor both in personal privacy and in the perceived risks which influence a user's digital privacy preferences [13]. It is essential in a mobile AR app to protect users' privacy from being violated, since it needs users to authorize several permissions, e.g., camera and microphone usage. The most vital one in AR Taiji is the permission to use the camera since the whole AR system is built upon the camera capturing physical environments. In some situations, users may not be willing to expose the surrounding environment through the camera due to privacy or sensitivity. But they still have the need to use AR Taiji.

Thus, we provide an option with virtual-reality(VR) environment to users. In Version 4, we investigate the possibility to merge AR and VR into AR Taiji. If users choose to run the app without camera, they could place the virtual character in a preset virtual environment, running the app in VR mode, see Figure 3.13.

(a) Starting pose with real environment

(b) Starting pose with VR environment

Figure 3.13. Running ARTaiji in non-VR mode/VR mode

This mode provides an preset animated virtual environment instead of the real environment in AR mode. It allows users to only focus on the virtual content, which in turn will enable the virtual characters made from Mocap data to achieve their full potential. Currently we implement a virtual background that contains white clouds and black soil, yet it is not complicated to bring in more realistic backgrounds, e.g. a beach or a valley.

There is another advantage of adding a VR option. AR has critical challenges on power consumption, compute resources and memory, which are valuable on mobile devices. The pipeline to render an AR scene is shown in Figure 3.14. The first step is to capture the real world by the camera for world tracking and plane detection. These are essential factors in Scene Understanding. The second step is to render the virtual and real world by merging scene understanding and motions captured by world tracking. After the rendering process, the resulting AR world is displayed on the screen that allows interactions from the user.

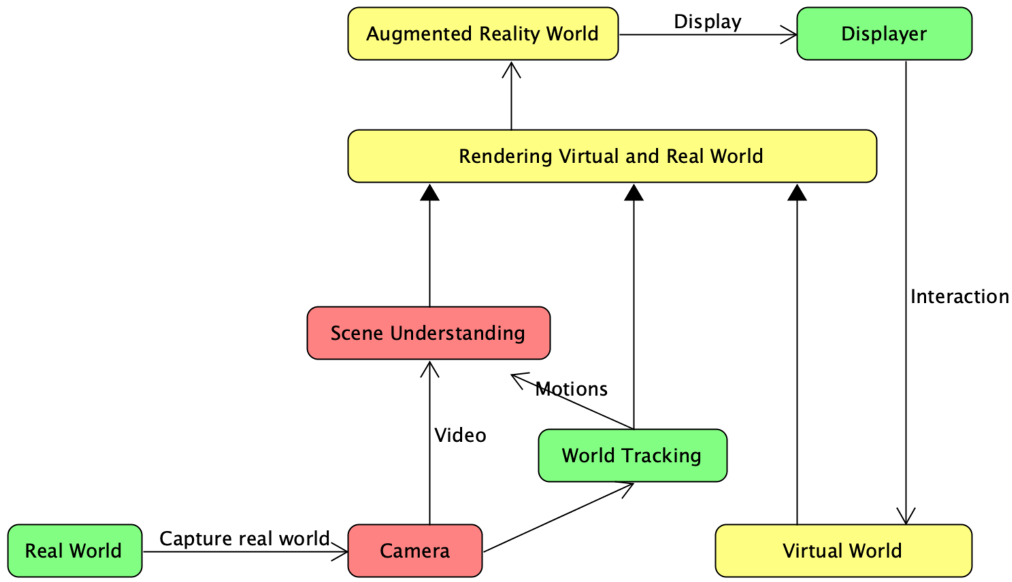


Figure 3.14. Pipeline of AR rendering

All the green blocks in Figure 3.14 remain unchanged after enabling a VR mode. In other words, world tracking and displaying method will not be affected by the new option. All red blocks are removed by VR since it does not require a camera to do scene understanding anymore. Yellow blocks are partially affected by the modification. We no longer need to render the real world, thus the original augmented reality world has changed into a virtual reality world. Therefore, a VR option that can render virtual content without a camera saves a mass of rendering resources.

Chapter 4
Human Survey Methodology

The primary purpose of this study is to evaluate user experience towards modifications and HCI principles applied in each version. Updates are made based on the baseline AR Taiji app, thus the corresponding questions are mainly emphasizing the comparisons of user-centered experience among updated versions and the baseline model. Firstly, we conduct this survey by letting participants explore three modified versions of AR Taiji one after another. Different HCI principles has been applied to UI design, augmented system and privacy properties. As mentioned in Chapter1, mobile AR is a relatively new field that requires experiments and investigations on various aspects. Hence, it is significant to build sufficient understandings of key factors like consistency, system complexity, field size and flexibility in use. We hope that after exploring all modified versions of AR Taiji, participants will likely to have an enhanced HCI experience and a unique insight into the research.

Secondly, we further conduct the survey by giving an online questionnaire to every participant who have explored the app for 5 minutes. Five minutes are long enough for users to explore all functions in any version. To quantitatively measure the effects of different HCI principles and design, we use scores of four factors primarily: Satisfaction, Efficiency, Complexity and Privacy. They are queried about those four factors by scoring questions. Through the scoring discrepancies on different implementations of the same factor, we are able to get a clearer picture of user attitudes and feedback on those modifications in different versions.

4.1 Primary HCI Factors

In each question, at least one factor will be explored. Participants are required to rate factors like efficiency, satisfaction and complexity according to their user experience.

Likert is used to evaluate the attitude of humans. A Likert (or Likert scale) is a bipolar scaling method, measuring either positive or negative response to a statement [2]. Each answer in the Likert has corresponding points. For instant, on Satisfaction, the scale is: Extremely dissatisfied(1 point); Somewhat dissatisfied(2 points); Neither satisfied nor dissatisfied(3 points); Somewhat satisfied(4 points); Extremely satisfied(5 points).

4.1.1 Satisfaction

Satisfaction is defined in this thesis as: a primary measure to the **overall** quality of one modification to the baseline. It shows the subjective user perception of interacting, which measures the performance of all HCI principles described in Chapter 3(**Affordance**, **Flexibility**, **Privacy**) and the user experience.

In Version 2 Part 1, we apply **Affordance** to AR Taiji, resulting in an implicit layout with an integration button. User can score on Satisfaction to imply the overall impact they experience from the new implicit layout. A higher score represents a better user experience for the new layout and the Affordance. In Version 2 Part 2, several new characters are added according to Affordance. A higher score reflects on a better quality and a more immersive user experience with virtual characters. In version 3 with a voice control unit, Satisfaction score tries to measure the effect of HCI principle called **Flexibility**. The more satisfied the users are with voice control, the higher the satisfaction score will be in version 3. Moreover, in version 4 with a VR environment, Satisfaction score reflects on the user experience in VR mode. The better user experience in VR mode, the higher the satisfaction score will be.

At the beginning of each section of the questionnaire i.e., Question 4, Question 14, Question 17 and Question 26(Appendix A), we first ask the participants about their user satisfactions to obtain a general impression on this version. For example, by giving a satisfaction score on different virtual characters(Question 23), we are able to know which model is more favored by users. Scores of Satisfaction are also being used to evaluate whether the effect of the modification is positive or negative. If the satisfaction score increases greatly on one function from baseline to an updated version, it is implying that the enhancement on this function is strong.

4.1.2 Efficiency

Efficiency is defined in this thesis as: a measure to the quality of a **single** function. It measures the additional amount of physical efforts to use one function after the modification.

We ask participants to provide an efficiency score on a function to determine whether adding this function has a positive or negative impact on the HCI process(Question 8, 15, 18). Moreover, we use efficiency scores on the same function in different versions as comparisons of different HCI principles. For example, by giving efficiency scores before(Question 8) and after the implementation of Version 3 voice recognition(Question 18), participants could show their level of preference for this new function.

Efficiency scores are only queried in Version 2 Part 1 and Version 3. In version 2 Part 1 with an implicit layout, efficiency score represents how many physical efforts are required in the new layout to finish one operation(Question 15). With principle **Affordance** applied, a higher efficiency score means that users do not think that many physical steps are required for the operations. In version 3, efficiency scores represents how many steps are saved by the voice control(Question 18). With principle **Flexibility** applied, a higher efficiency score means that voice control saves many physical efforts compared to version 1.

4.1.3 Complexity

Complexity is defined in this thesis as: to measure the level of being complicated to use of one function. It is a measure of additional interventions required by one function.

We query about system complexity in Question 11, 19. In Version 1 baseline which uses the original layout, we ask about complexity scores to assess how complex the system is(Question 11). A higher complexity scores means the original layout is more difficult for users to use. In version 3, complexity score reflects how hard it is to use the newly added voice recognition for the basic operations like Transformation and Rotation, in case that the disadvantages of adding a voice control outweigh the advantages.

4.1.4 Privacy

Privacy is defined in this thesis as: a measure of level to which user privacy is protected.

The HCI principle, **Privacy**, is only investigated in baseline and Version 4 as a control group(Question 13, 28). By comparing a privacy score in baseline(without VR) and Version 4(with VR option), we are able to explore the contribution of the VR mode in improving the privacy protection rate. A higher privacy score from user represents a stronger sense of privacy protection in user experience. For example, in the new VR mode which will not show the real environment on the screen, user would feel a stronger sense of privacy protection(Question 28).

4.2 Survey Structure and Design

In this section, we present how we design the survey structure. The cross-sectional questionnaire is intended to discover which kinds of design principles have better effects on the target users. The primary requirements for the questionnaire survey are as follows: (1)The function-oriented questionnaire should include all modifications that appear in four versions. (2)Each group of questions must have a corresponding control group which reflects the effect on applying HCI principles. There are three steps in the questionnaire design procedure.

1. An initial draft of the questionnaire is designed to evaluate the difference in user satisfaction towards HCI principles. Each modification has a scoring question on that function and a control group of questions on the baseline model.
2. The initial draft of the questionnaire is pre-tested and evaluated by an expert regarding its scoring scales, structure and question description.
3. Finally, an experimental study is conducted on two selected participants in order to determine the reliability of the analytical methods.

4.2.1 Basic information of participants

There are 20 participants in total, including 12 males and 8 females(Figure 4.1). The human subjects are asked whether they had any previous experience with AR apps. 18 out of 20 people(90%) have the experience of using an AR App before the survey. Among 20 participants, the first language of 9 participants is English, while the other 12 people is Mandarin.

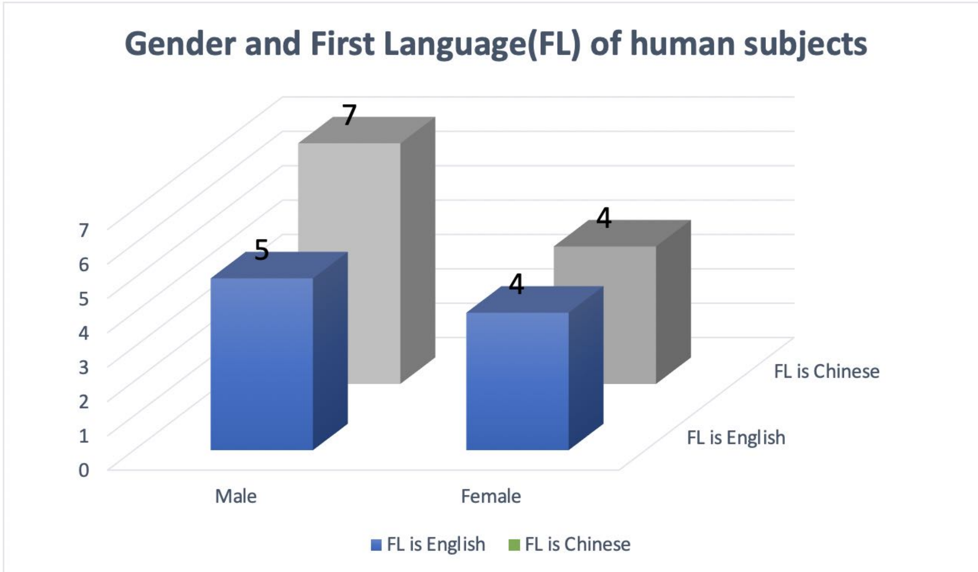


Figure 4.1. Gender and First Language

We have also paid attention to the device model running AR Taiji since screen size is involved in this survey. 60%(12) of participants are running AR Taiji app with a 5.8-inch or 6-inch screen, which are the screen size of iPhone11/12/X. Four users run AR Taiji and its enhanced versions on iPhone 7 and iPhone 7 plus that have a screen size under 5.5 inches, while the other four users run it on a screen larger than 6.5 inches. Detailed statistics are shown in Figure 4.2.

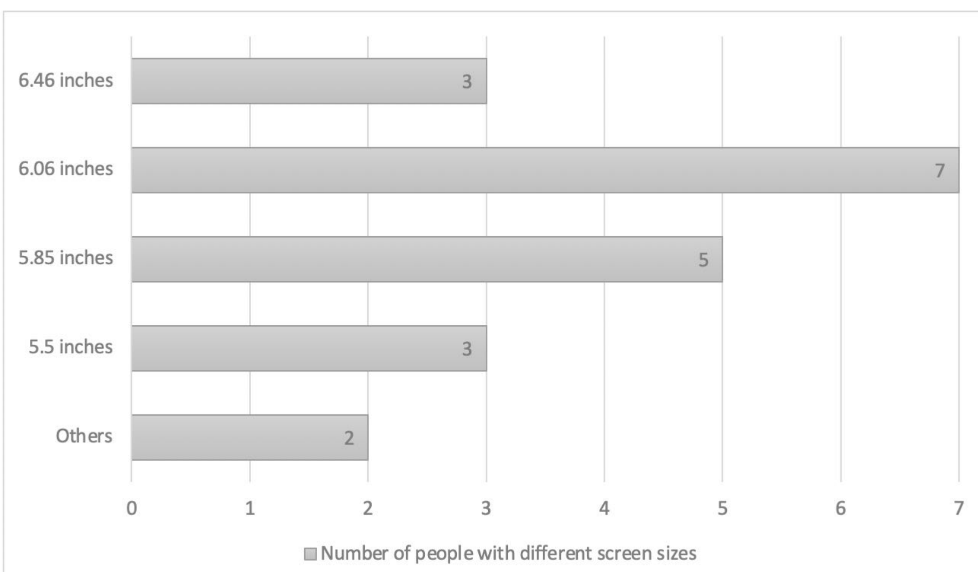


Figure 4.2. Number of responds with different screen sizes

The average completion time of the questionnaire is 7 minutes and 20 seconds. This is a cross-sectional questionnaire (Appendix A) with a total of 41 questions that are divided into five blocks: Demographic questions, Baseline, User Interface, Voice control and VR option. Each block contains a minimum of 4 and a maximum of 10 questions, see the number of questions corresponding to each factor in Table 4.1. Most questions are assessed according to a scale of five levels in a matrix form of Likert, in order to ensure consistency. They are:

1. Extremely dissatisfied(1 point)

2. Somewhat dissatisfied(2 points)

3. Neither satisfied nor dissatisfied(3 points)

4. Somewhat satisfied(4 points)

5. Extremely satisfied(5 points)

Factors	Version 1	Version 2	Version 3	Version 4	Mean	Std
Satisfaction	6	6	2	3	4.25	1.79
Efficiency	3	3	3	0	2.25	1.30
Complexity	2	0	1	0	0.75	0.83
Privacy	2	0	2	2	1.5	0.87

Table 4.1. Number of questions corresponding to each type of evaluation

Chapter 5
Result Evaluation and Analysis

In this chapter, we present the results and statistical analysis of the survey. To control variables, the survey is conducted version by version. In each version we only make modifications based on related HCI principles and ask questions about corresponding factors. Results with the same objective will be presented together in this chapter.

5.1 Overall Satisfaction

We begin with providing the results of overall user satisfaction on each of the four versions(Table 5.1). Participants are asked to rate the overall satisfaction in each version. A higher score means higher user satisfaction, with a scoring scale of Satisfied/Dissatisfied Likert: 1(strongly unsatisfied), 2(somewhat unsatisfied), 3(neither unsatisfied nor satisfied), 4(somewhat satisfied) and 5(strongly satisfied). Here we show mean, standard deviation, increases and p-values in Table 5.1.

Versions	Mean/Std	Min	Max	Increase	p-value
V1(baseline)	3.6/1.02	2	5	NA	NA
V2 P1(Implicit Layouts)	4.35/0.73	3	5	+0.75(20.8%)	**0.002**
V2 P2(New Characters)	3.65/**0.64**	1	5	+0.05(1%)	0.5
V3(Voice Control)	**4.5**/0.81	2	5	**+0.9(25.0%)**	**0.004**
V4(VR Mode)	3.4/1.02	1	5	-0.2(6%)	0.223

Table 5.1. Overall satisfaction in all versions. P values are a comparison with V1

From Table 5.1, user satisfaction on user interface(Version 2 Part 1) and functionality experience(Version 3 and Version 4) has an average increase of 13.3%((20.8%+25.0%-6%)/3=13.3%). From the baseline model, we can see that user satisfaction had greatly improved on Version 2 Part 1 with an implicit layout(+0.75, Figure 3.8(c)) and Version 3(+0.9). The best average score of satisfaction was achieved in Version 3. Although we received mixed reviews towards modifications in all enhanced versions, there is an undeniable increase from baseline to Version 2 Part 1 and Version 3. Note that in Version 2 Part 1, we want to explore the trade-off between efficiency and user satisfaction brought by an implicit layout, which has an increase of 20.8%. On Version 2 Part 2(Figure 3.10), we investigate on the user satisfaction on different virtual character. Therefore, loss of efficiency from the implicit layout would potentially lower user expectations, giving rise to mixed responses. But eventually the enhanced layout with a larger screen area in Version 2 that increases user satisfaction has made more contributions to the increment of scoring. We provide more details in the following section.

Version 3 has an approximate 25% increase in user satisfaction from the addition of voice recognition. A lower Std from baseline also means the majority are holding a positive review towards this enhanced function. In fact, 65%(13 people) of participants stated they are extremely satisfied with Version 3, while 5 people(25%) agreed they are somewhat satisfied, see the comparison between baseline and Version 3 on overall satisfaction in Figure 5.1.

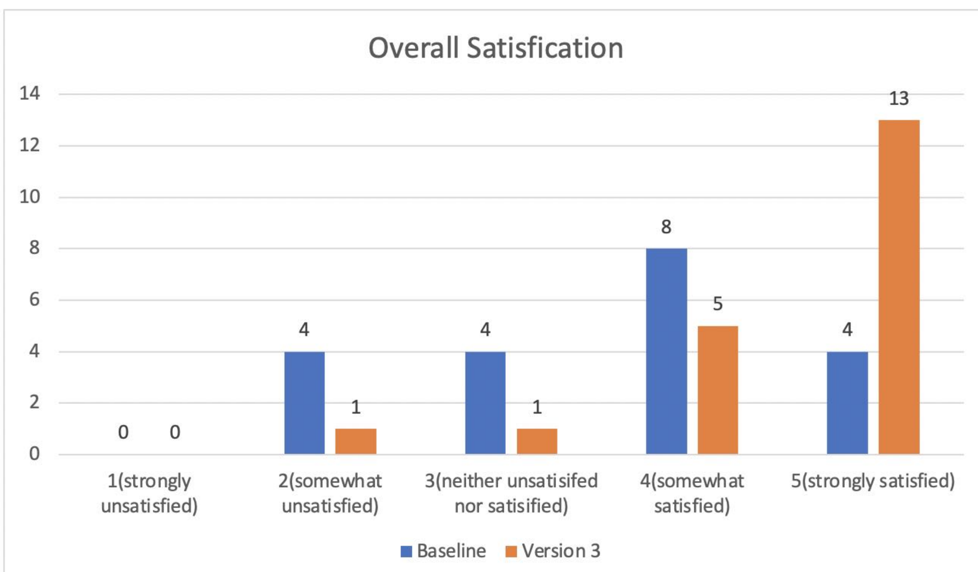


Figure 5.1. Total number of Responds of overall satisfaction on baseline and Version 3: 20(subjects) * 2(scores) = 40

To quantitatively analyze the results, we apply a statistical method called *The Paired Samples t-test*. The purpose of t-test is to quantitatively check if there exists statistical evidence that the mean difference between paired samples on measurement is significantly different from zero.

Before we apply this method, the first step is to make two hypotheses. Like many statistical procedures, the paired sample t-test has two competing hypotheses, the null hypothesis and the alternative hypothesis [18].

1. **Null hypothesis H_0:** The null hypothesis should be that baseline has no significant improvement than Version 3.

2. **Alternative hypothesis H_1:** The alternative hypothesis is that version 3 is better than the baseline.

The alternative hypothesis can take one of several forms depending on the expected outcome. If we don't care about the direction of differences, a two-tailed hypothesis is used. Otherwise, an upper-tailed or lower-tailed hypothesis can be used to increase the power of the test [18]. No matter what direction we choose in alternative hypothesis, the null hypothesis remains the same for each type of alternative hypothesis.

Second, we will apply the equation shown in Eq(5.1) to the result.

$$t = \frac{\bar{x}_{diff}}{s_{\bar{x}}} \tag{5.1}$$

where

$$s_{\bar{x}} = \frac{s_{diff}}{\sqrt{n}}$$

$\bar{x}_{diff}$ is the sample mean of the differences, n is the number of observations and s_{diff} is the sample standard deviation of the differences. Calculated result value t is then compared with tables of the t-distribution. This will give a p-value for the paired samples t-test. If $p < 0.05$, we have 95% confidence to accept the alternative hypothesis and conclude that on average, the updated version has a significant improvement from baseline(result is significant). Otherwise, there is no statistical evidence to the alternative hypothesis. We have to accept the null hypothesis when $p > 0.05$, which means that results are not significant.

The resulting t value of Eq(5.1) between Version 1 and Version 3 is -2.93, meaning this is a correct lower-tailed t-test($t < 0$), with a corresponding $p = 0.004 < 0.05$. Thus we can reject null hypothesis and conclude that the alternative hypothesis is critically supported. There exists strong evidence that Version 3 has a great improvement over the baseline. If we apply the same method to Version 2, we have a resulting p value $= 0.0019 < 0.05$. It also indicates that Version 2 has a significant improvement than the baseline.

Nevertheless, the average score on satisfaction remains nearly unchanged in Version 4(Figure 3.13), only decreases by 0.2(-6%). The specific numbers of responses are shown in Figure 5.2.

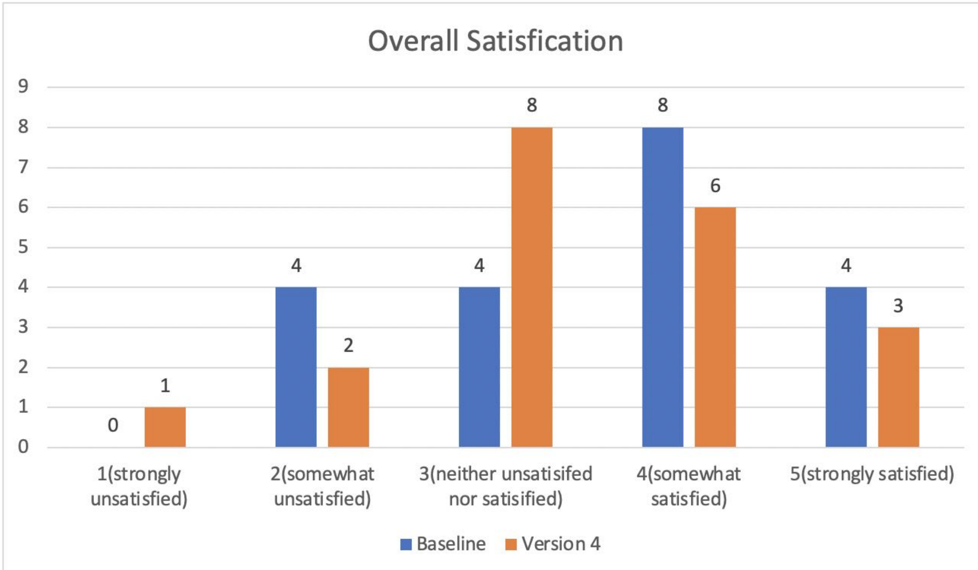
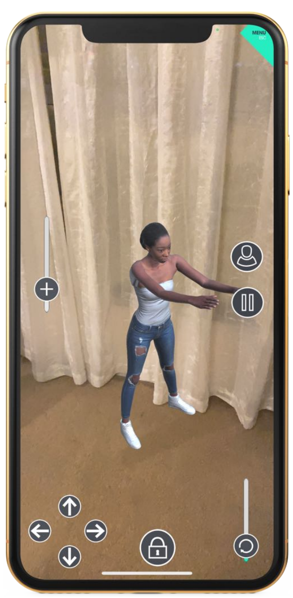


Figure 5.2. Total number of Responds of overall satisfaction on baseline and Version 4(Figure 5.3): 20(subjects) * 2(scores) = 40



(a) Starting pose with real environment

(b) Starting pose with VR environment

Figure 5.3. Running ARTaiji in non-VR mode(baseline) and VR mode(version 4)

We can see that most of the participants(40%) maintain a neutral attitude towards Version 4(see Figure 5.2), while originally 40% of users are somewhat satisfied with baseline. Intuitively the baseline has a higher user satisfaction score given the same Std in both baseline and Version 4. To quantitatively prove this, we apply *The Paired Samples t-test* to the result as well. However under this circumstance, we are not quite sure about the direction of the differences, thus two-tailed hypotheses are applied here.

1. H_0: Satisfaction scores had not changed significantly from baseline to Version 4.

2. H_1: Satisfaction scores had changed significantly from baseline to Version 4.

Here the resulting t = 1.71 with a corresponding value p = 0.103 > 0.05. This is not a significant result, which means that we have no choice but to accept the null hypothesis. Therefore modifications in Version 4 did not lay a perceptible impact on overall satisfaction.

5.2 Version 2 Part 1: layout of UI

In Version 2 Part 1, we add an integration button and implement an implicit layout. Several factors related to those features were investigated with the help of scoring questions. As described in Chapter 3, user can control the view of virtual character by several basic operations:

1. Initialization: By clicking anywhere on the screen after launching AR Taiji, users could drop virtual characters on a surface or a plane.

2. Transformation: General term for translation, scaling (see Figure 3.3) and rotation (see Figure 3.2)

3. Play/Pause: User can use the Play/Pause button (see Figure 3.4(a)) to play/pause the virtual character in the middle of a human motion sequence.

4. Form Selection: Click the top-right green menu gives a list of 24-form Taiji that the user can let the virtual character run the corresponding form (see Figure3.4(b)).

Since an integration button would possibly increase the physical efforts of several operations, resulting in reduced efficiency, therefore we asked participants to score from 1(strongly inefficient) to 5(strongly efficient) on two fundamental activities that were affected directly by the integration button: Transformation and Initialization functions. See the number of responses in Figure 5.4.

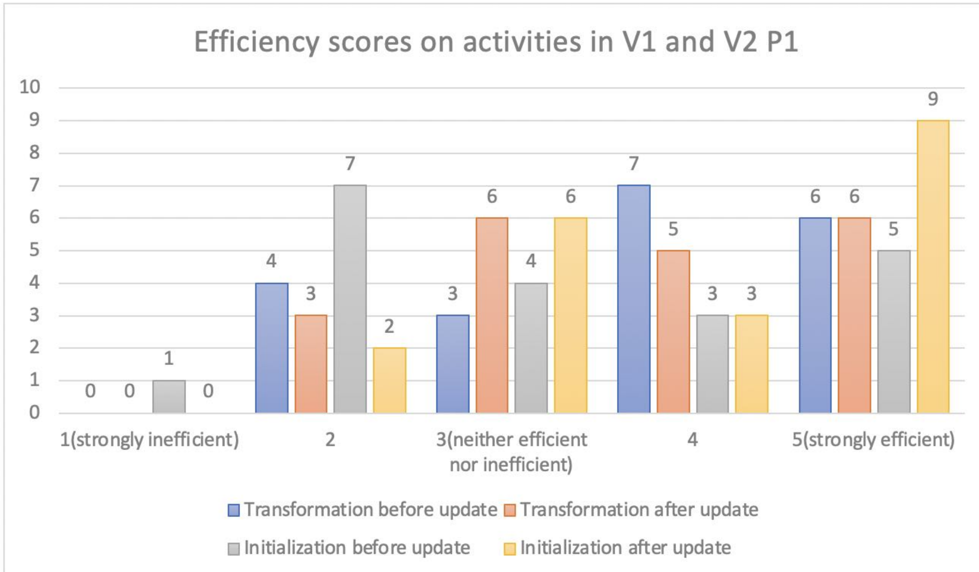
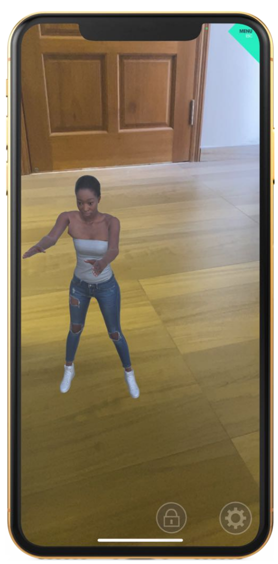
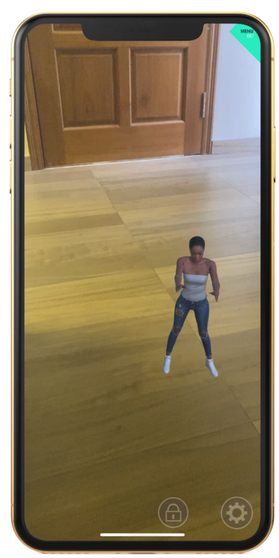


Figure 5.4. Total number of responds of efficiency on baseline and Version 2 Part 1(Figure 5.5 and Figure 5.6): 20(subjects) * 4(scores) = 80



(a) Original position



(b) Translate and scale the virtual character to a desired position

Figure 5.5. A sequence of operations: Translation and Scaling(Transformation)

(a) Initialization on the surface of a desk

(b) Initialization on the floor

Figure 5.6. Drop the virtual character on different planes(Initialization)

From 5.4 we could find that after modifications on Version 2 Part 1, the discrepancies of Efficiency scores on the Transformation function are modest, but are huge on the Initialization function. The two sets of control data are blue/red histograms and gray/yellow histograms, representing the changes of user satisfaction in those two functions from Version 1 baseline to Version 2 Part 1 respectively. Blue and red are approximate evenly distributed at each level of satisfaction with small differences, most of which(30%) are concentrated in *5.Strongly Satisfied* level. Gray and yellow histograms are unevenly distributed. 35% of gray histograms concentrated in level 2(Somewhat dissatisfied), yet after the enhancement, 45% gave a *Strongly satisfied* response. With an enhancement on the layout that increases screen area materially, there is a visible increase on the efficiency score of Initialization function.

So far the change of adding an integration button has not had much impact on the efficiency score of the Transformation function. This is the result we expect to see since we are interested in the trade-off between efficiency and user satisfaction. If such modification does not affect the efficiency of functions largely, we could conclude that at least it is not incorrect of using such a two-layer implicit layout with an integration button. On the contrary, we are pleased to see a positive feedback on the Initialization function since the initial intend was to save screen space and increase efficiency on Initialization process.

Thus the quantitative analysis is required. We first show the average score/std with p-values in Table 5.2.

Functions	Mean	Std	p-value
V1 Transformation	3.75	1.09	NA
V2 Transformation with an integration button	3.7	1.00	0.44
V1 Initialization	3.2	1.29	$\approx \mathbf{0}^*$
V2 Initialization with an integration button	3.95	1.07	$\approx \mathbf{0}$

Table 5.2. Scores rated from Efficiency of Transformation and Initialization on V1 and V2. P values are a comparison between V1 and V2. P value* is a comparison of Transformation and Initialization on V1

Table 5.3 shows the efficiency scores of Transformation and Initialization split into two different genders. We can see that only on Version 2 Initialization, Female users give two Efficiency scores that are significant different with each other. There is a huge increase of female Efficiency score from V1 to V2 on Initialization.

Functions	Mean/Std(Male)	p-value	Mean/Std(Female)	p-value
V1 Transformation	**3.67**/0.94	NA	3.88/1.27	0.13*
V2 Transformation	**3.67**/0.94	0.5	3.75/1.20	0.43
V1 Initialization	3.08/1.11	NA	3.38/1.49	0.21*
V2 Initialization	**3.67**/1.11	0.07	**4.38**/0.86	**0.02**

Table 5.3. Scores rated from Efficiency of Transformation and Initialization(Gender difference). P values are a comparison between V1 and V2. P values* are a comparison of male and female users on V1

Furthermore, we investigate the performance of different sizes of screen. According to Figure 4.2, screen sizes can be divided into three categories: 6.5 inches, 6 inches and 5.5 inches. We summarized mean, standard deviation and p-value from V1 to V2 in Table 5.4

Functions	Mean/Std(5.5 in)	p-value	Mean/Std(6 in)	p-value
V1 Transformation	**4.67**/0.47	NA	3.42/1.11	$\approx \mathbf{0}^*$
V2 Transformation	3.67/0.94	0.11	3.58/1.19	0.38
V1 Initialization	3.00/1.41	NA	3.25/1.36	0.63^*
V2 Initialization	3.67/0.94	0.09	3.92/1.19	**0.05**
Functions	Mean/Std(6.5 in)	p-value		
V1 Transformation	4.00/0.82	$\approx \mathbf{0}^*$		
V2 Transformation	**4.00**/0.5	0.21		
V1 Initialization	**3.67**/0.94	$\approx \mathbf{0}^*$		
V2 Initialization	**4.00**/0.21	0.5		

Table 5.4. Scores rated from Efficiency of Transformation and Initialization(screen size difference). P values are a comparison between V1 and V2. P values* are a comparison with 5.5 in screen

From Version 1 baseline to Version 2 Part 1, the average score of Efficiency on transformation function remains nearly unchanged(-0.05). But it has increased by 0.75(23%) for the Initialization function. For the Transformation function, the total change on Std is much less than 0.1. For such a close result, we could apply t-test to determine if there is a significant improvement in the two versions. As usual, we first make two hypotheses.

1. H_0: The average score on Efficiency has not changed significantly for the Transformation function.

2. H_1: The average score on Efficiency has changed negatively for the Transformation function.

Apparently those are one-tailed hypotheses with a certain directional assumptions. We apply t-test to the results to obtain a p-value $= 0.44 > 0.05$. Therefore we can accept the null hypothesis and conclude that the impact of adding an integration button to Transformation function is minimal to overall efficiency.

Similarly, the average efficiency on the Initialization function increases by 0.75(23%) on the Initialization function. The total change on Std is 0.22 lower. The resulting p-value is 0.0000037 « 0.05. Thus we can also draw a conclusion that adding an integration button has not caused any adverse effects on the Transformation function, but seriously improves the efficiency of the Initialization process.

5.2.1 Screen area Division

Meanwhile, user satisfaction with the size of the screen area has a significant increase(p value ≈ 0 in Table 5.5) On iPhone 11 that has a 6-inch screen, User Interface occupies approximately 25% of screen area, with only 75% of the screen left for displaying virtual content. By adding an integration button, screen area increases from 75% of screen to 95% of screen (see Figure 3.8(c)). We asked participants how satisfied are the integration buttons with the impact of screen area, see the number of responses in Figure 5.7.

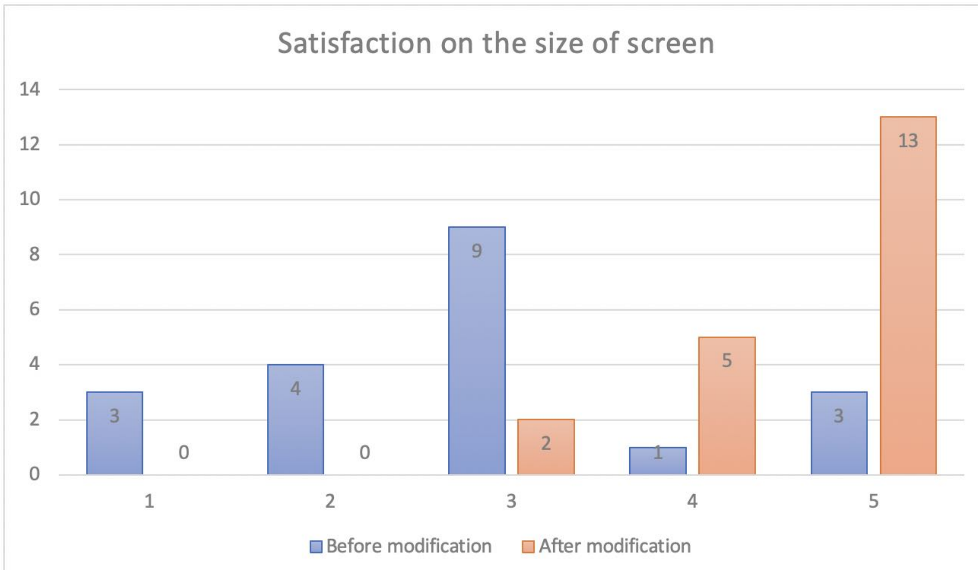


Figure 5.7. Total number of responds on size of screen area: 20(subjects) * 2(scores) = 40

As to the distribution of satisfaction on screen area division, a clear pattern can be discerned. The spike of blue histograms representing the satisfaction scores before modifications has generally translated to the right after the modifications. In the baseline, 9 people(45%) held a neutral view on screen area division while 13 participants(65%) felt extremely satisfied with the screen area division after the enhancement on the layout. From Table 5.5, we can see the average score has gone up by 1.7(60%) and std has decreased by 0.52. The results depict that the appearance of an integration button greatly increases the user's satisfaction towards the screen area division.

Size of screen area	Mean	Std	p-value
Before updated layout	2.85	1.19	NA
After updated layout	4.55	0.67	$\approx$ **0**

Table 5.5. User attitudes towards screen area division on V1 and V2

Table 5.6 gives user attitudes towards screen area division on different gender. It clearly demonstrates that both male users and female users have a significant increase on user attitude after the modification.

Size of screen area	Mean/Std(Male)	p-value	Mean/Std(Female)	p-value
V1 baseline layout	2.75/0.93	NA	3.00/2.57	0.35*
V2 implicit layout	4.42/0.63	**0.0005**	4.75/0.21	**0.006**

Table 5.6. User attitudes towards screen area division on V1 and V2(Gender difference). P values are a comparison between V1 and V2. P value* is a comparison of male and female users on V1

We provide user attitudes towards screen area division on Version 2 Part 1 with different gender and First Language(FL) in Table 5.7. Female users with FL in English have the highest satisfaction score. Among four groups, only Male in Chinese has a significant difference with Male in Female in English, with p value < 0.05.

p values	Mean/Std	Male FL En	Female FL En	Male FL Ch	Female FL Ch
Male FL En	4.8/0.4	NA	0.19	0.07	0.21
Female FL En	**5.0**/0.0	0.19	NA	**0.02**	0.09
Male FL Ch	4.14/0.83	0.07	**0.02**	NA	0.22
Female FL Ch	4.5/0.5	0.21	0.09	0.22	NA

Table 5.7. User attitudes towards screen area division on Version 2(FL difference).

By combining the results of Table 5.2 and Table 5.5, importing such a layout will not critically reduce the efficiency of operations, but will greatly increase the screen area, thereby increasing user satisfaction on related functions. It is the first time that we realize how important the screen area division is for an AR App, especially with an initialization process to drop an virtual character. For a better user experience, it is vital for an AR app to take the size of field of view into consideration.

However, since AR Taiji runs on screens of different sizes(see Chapter 4), screen size is also a vital factor on the score of screen area division. According to Figure 4.2, screen sizes can be divided into three categories: 6.5 inches, 6 inches and 5.5 inches. We summarized mean, standard deviation and p-value from V1 to V2 in Table 5.8.

Screen Size Division	Mean/Std(5.5 in)	p-value	Mean/Std(6 in)	p-value
V1 baseline layout	**4.33**/0.94	NA	2.50/1.19	$\approx \mathbf{0}^*$
V2 implicit layout	**4.67**/0.47	0.21	4.50/0.76	$\approx \mathbf{0}$
Screen Size Division	Mean/Std(6.5 in)	p-value		
V1 baseline layout	3.00/0.00	$\approx \mathbf{0}^*$		
V2 implicit layout	**4.67**/0.47	**0.02**		

Table 5.8. User attitudes towards screen area division(screen size difference). P values are a comparison between V1 and V2. P values* are a comparison with 5.5 in screen

According to Table 5.8, users with a 6-inch screen have the lowest average score on baseline layout. On the contrary, users with a 5.5-in screen are satisfied with baseline layout. With the new implementation of an implicit layout, users with 6-in and 6.5-in screen both have significant increases on their user satisfaction towards screen area division, with both p-value smaller than the threshold.

5.2.2 Version 2 Part 2: Change of characters

Another new feature provided is a list of characters for users to choose from in Version 2 Part 2. Currently there are three distinct characters with different model appearances. The contours and joints of Woman(see Figure 3.9) and Guy(see Figure 3.11) are succinct, which can precisely show the exchange of force. On the contrary, Alien(see Figure 3.10) is a cute model that only shows the outline of the action. We have summarized the user's attitude towards these three models in Figure 5.8.

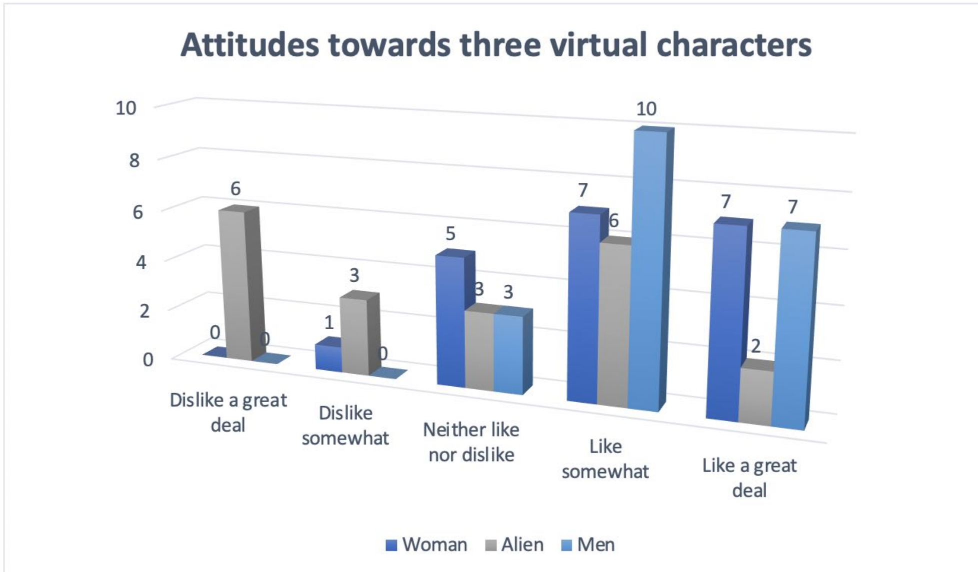
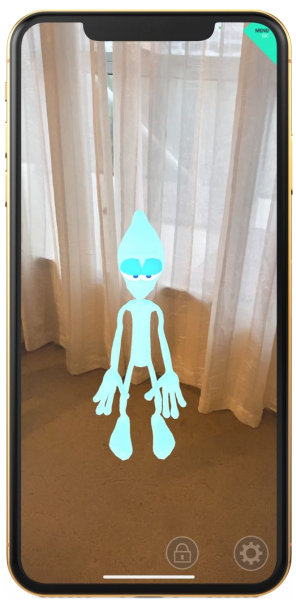
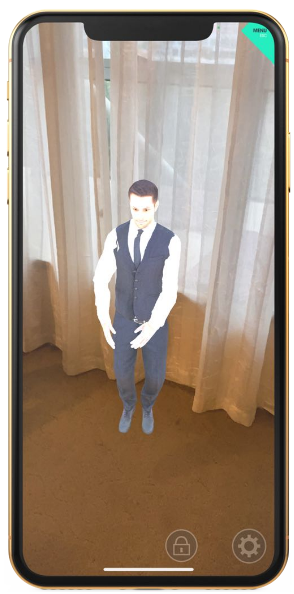


Figure 5.8. Attitudes towards three virtual characters(Figure 5.9): Number of total ratings = 20(subjects) * 3(scores) = 60



(a) Alien character



(b) Man character

Figure 5.9. New virtual characters in version 2 part 2

We are able to clearly distinguish that rates on Alien concentrate mainly on the left three levels(Dislike a great deal, Dislike somewhat and Neither like nor dislike). While ratings on Woman and Men characters concentrate on the right two levels(Like somewhat and Like a great deal). Most of ratings on Alien model are concentrated in "Dislike(1 point)" and "Somewhat like(4 points)" areas. Most of the mass of two blue triangles lay in "Like somewhat(4 points)" and "Like a great deal(5 points)" areas. In the Woman model, 70%(14 people) of participants gave at least four points. This data in the Men model is 85%(17 people).

The statistical analysis of user attitude towards the appearance of character shows this dividing line more straightforward(see Table 5.9). The average rate of user satisfaction has a 1.25 gap between Woman and Alien. A higher Std in Alien depicts that participants held a negative view regarding this as well.

Characters	Mean	Std	Woman p-value	Alien p-value	Men p-value
Woman	4.0	0.89	NA	$\approx \mathbf{0}$	0.08
Alien	2.75	1.41	$\approx \mathbf{0}$	NA	$\approx \mathbf{0}$
Men	4.2	0.68	0.08	$\approx \mathbf{0}$	NA

Table 5.9. Comparison on satisfaction of Woman, Alien and Men

This result is not beyond the expectations. Since this is an AR App with human motion sequence learning from MoCap, an angular model that reflects the body shape and movements of the recorder will definitely be more popular.

Further more, the Satisfaction score between different genders is list in Table 5.10. Female users score 17% higher on the Woman character than male users. Additionally, female users also score 7% higher on the Men character than male users, both of which are very similar models to real human. However on the Alien character, the cartoon character receives a score that is 95% higher from male users to female users. We can conclude that Female users have a preference on models that are similar to real human than male users. Yet Male users prefer cartoon character much more than female users.

Characters	Mean/Std(Male users)	Mean/Std(Female users)
Woman	3.75/0.92	**4.38**/0.70
Alien	**3.42**/1.19	1.75/1.19
Men	4.08/0.64	**4.38**/0.70

Table 5.10. Comparison on satisfaction of Woman, Alien and Men(Different Gender)

Table 5.11 gives detailed data analysis on three virtual characters. For both the Woman character and the Man character, female users with FL English give the highest satisfaction score, but give the lowest score on Alien character. On the contrary, male users give a highest score on Alien character.

Characters	Mean/Std(Male English)	Mean/Std(Male Chinese)
Woman	4.00/0.89	3.57/0.90
Alien	**3.8**/0.75	3.14/1.36
Men	4.20/0.40	4.00/0.76
Characters	Mean/Std(Female English)	Mean/Std(Female Chinese)
Woman	**4.75**/0.43	4.00/0.71
Alien	**1.25**/0.43	2.25/1.30
Men	**4.75**/0.43	4.00/0.71

Table 5.11. Comparison on satisfaction of Woman, Alien and Men(Different First Language). Female FL English has the lowest average score(marked in red)

Table 5.12 and Table 5.13 further analyze p values among Male/Female users with different first language. For the Alien character, Male FL English are significantly different from Female FL English, with p value $< 0.05(0.04)$. Meanwhile, for the Men character, Male FL Chinese are significantly different from Female FL English, with p value $= 0.04 < 0.05$.

P values	Mean/Std	Male FL En	Female FL En	Male FL Ch	Female FL Ch
Male FL En	**3.8**/0.75	NA	**0**	0.17	0.07
Female FL En	**1.25**/0.43	**0**	NA	**0.01**	0.14
Male FL Ch	3.14/1.36	0.17	**0.01**	NA	0.19
Female FL Ch	2.25/1.30	0.07	0.14	0.19	NA

Table 5.12. P values on satisfaction of virtual character: Alien. Female FL English has the lowest average score(marked in red)

P values	Mean/Std	Male FL En	Female FL En	Male FL Ch	Female FL Ch
Male FL En	4.20/0.40	NA	0.07	0.30	0.34
Female FL En	**4.75**/0.43	0.07	NA	**0.04**	0.09
Male FL Ch	4.00/0.76	0.30	**0.04**	NA	0.5
Female FL Ch	4.00/0.71	0.34	0.09	0.5	NA

Table 5.13. P values on satisfaction of virtual character: Men

5.3 Version 3 Voice Control

In Version 3, the exclusive feature we have updated is a voice control unit. We explored various factors related to these modifications e.g. complexity and consistency. The most concerning issue is a soaring complexity resulting from the cognitive difficulty brought by voice recognition.

Participants were required to score on the complexity of the AR system with and without voice recognition based on their experiences. In the baseline version, 35% of the participants agreed it's simple to use and nearly half(45%) remained indifferent. After modifications, nearly half(45%) of the participants think a version with voice recognition is simple enough to use. We have summarized the user's attitude towards system complexity of voice control in Figure 5.10.

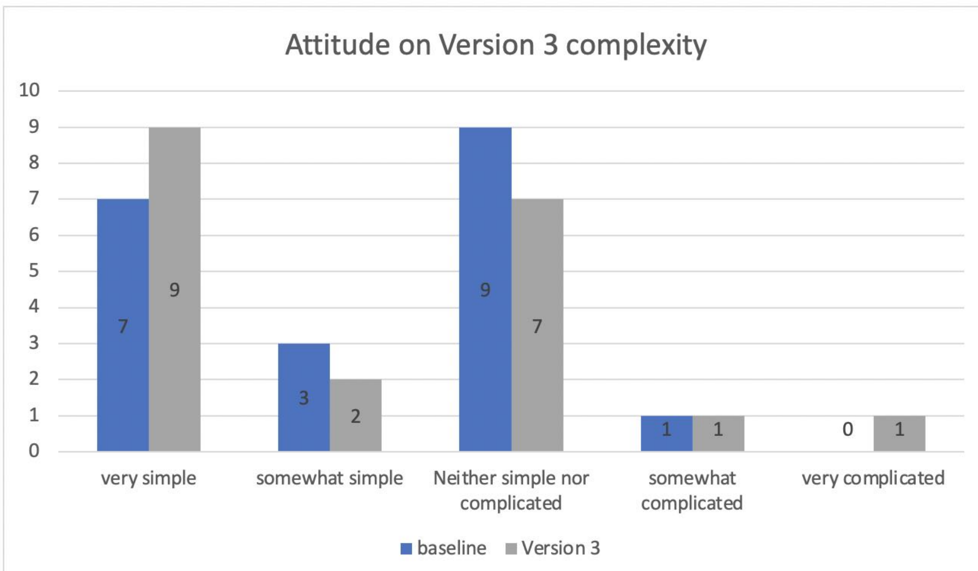


Figure 5.10. Attitudes towards complexity: Number of total ratings = 20(subjects) * 2(scores) = 40

Only 1 response falls in *Very complicated* and 2 fall in *somewhat complicated.* A detailed analysis is shown in Table 5.14.

Versions	Mean	Std	p-value
V1 Baseline	2.2	0.98	NA
V3 Voice Control	2.15	1.19	0.43

Table 5.14. Comparison of system complexity in V1 Baseline and V3 Voice Control

The average score and Std on system complexity have hardly changed before and after modifications. This means that the users potentially do not feel any increase of system complexity. To further prove this, we can make two hypotheses and apply t-test to them.

1. H_0 System complexity has changed significantly.

2. H_1 System complexity has not changed significantly.

p value is $0.43 > 0.05$, thus we could reject the null hypothesis and conclude that complexity has not changed significantly based on user experience.

We further study whether there is a difference in complexity between different genders(Table 5.15). The average score on complexity has no change from V1 to V3 in male users, meaning adding a voice control does not cause any negative impact or significant change, with p-value $= 0.5 > 0.05$. Female users has a same situation, with p-value $= 0.39 > 0.05$. Thus adding a voice control does not increase system complexity, whether it's a male user or a female user.

Versions	Mean/Std(Male users)	p-value	Mean/Std(Female users)	p-value
V1 Baseline	2.33/0.85	NA	2.00/1.12	0.55*
V3 Voice Control	2.33/1.31	0.5	1.88/0.93	0.39

Table 5.15. Comparison of system complexity in V1 and V3(Gender difference). P values are a comparison between V1 and V3. P value* is a comparison of male and female users on V1.

Table 5.16 further splits the result by different FL. Among four groups, Female FL English gives lowest complexity scores on both V1 and V3. All p values are greater than 0.05, which means there is no significant change from version 1 to version 3.

Versions	Mean/Std(Male English)	p-value	Mean/Std(Male Chinese)	p-value
V1 Baseline	2.40/0.80	NA	2.28/0.90	0.42*
V3 Voice Control	2.60/1.49	0.41	2.14/1.48	0.37
Versions	Mean/Std(Female English)	p-value	Mean/Std(Female Chinese)	p-value
V1 Baseline	**1.75**/0.83	NA	2.25/1.30	$\approx \mathbf{0}$*
V3 Voice Control	**1.5**/0.87	0.20	2.25/0.83	0.5

Table 5.16. Comparison of system complexity on Version 1 and Version 3(Different Gender/FL). P values are a comparison between V1 and V3. P values* are a comparison of male and female users on V1.

Table 5.17 shows all p values among four groups on Version 3 complexity scores. All of the p values are greater than the threshold. We can conclude that there is no significant difference among four groups on Version 3 complexity score.

P values	Mean/Std	Male FL En	Female FL En	Male FL Ch	Female FL Ch
Male FL En	**2.60**/1.49	NA	0.13	0.31	0.35
Female FL En	1.5/0.87	0.13	NA	0.19	0.16
Male FL Ch	2.14/1.48	0.31	0.19	NA	0.44
Female FL Ch	2.25/0.83	0.35	0.16	0.44	NA

Table 5.17. P values of system complexity score on Version 3(Different Gender/FL)

5.3.1 Satisfaction on Voice Control

The second question we asked is how users were satisfied with the experience with and without voice recognition. The feedback is recorded in Figure 5.11.

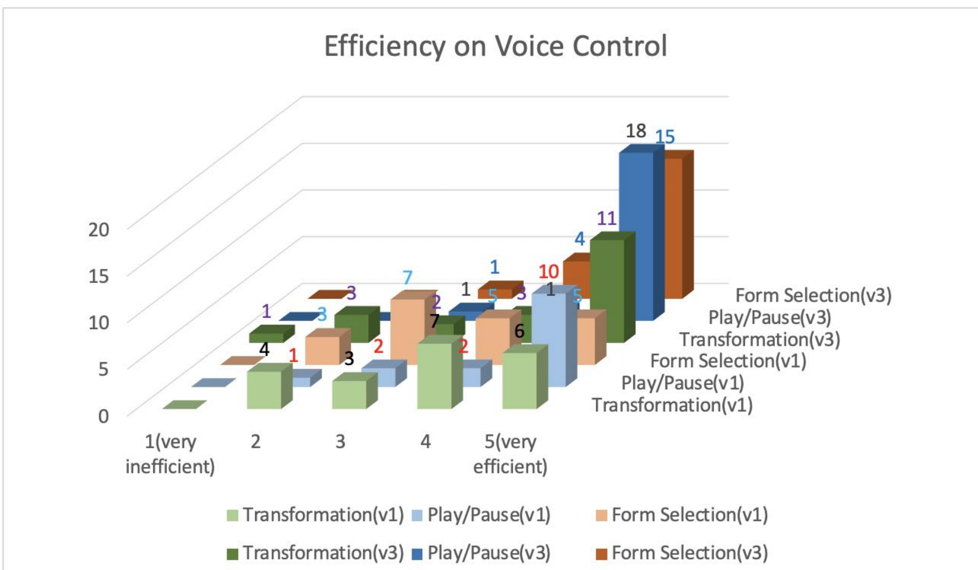


Figure 5.11. Efficiency with and without voice control: Number of total ratings = 20(subjects) * 6(scores) = 120

We can find that after adding a voice recognition, the histogram has a tendency to shift to a higher score integrally. This means that participants believe that overall efficiency has been significantly improved not only on Transformation functions, but also on functions like Form Selection and Play/Pause, see Table 5.18.

Functions	Mean	Std	Increase	p-value
V1 Transformation	3.75	1.09	NA	NA
V1 Play/Pause	3.8	1.33	NA	NA
V1 Form Selection	3.6	1.02	NA	NA
V3 Transformation	4.0	1.30	+0.25(+6.7%)	0.27
V3 Play/Pause	4.85	0.48	+1.05(+28%)	$\approx \mathbf{0}$
V3 Form Selection	4.7	0.56	+1.1(+31%)	$\approx \mathbf{0}$

Table 5.18. Efficiency of operations in V1 baseline and V3 voice control. P values are a comparison between V1 and V3.

From Table 5.18, score rated on Efficiency increase by 31%(+1.1) on Form Selection, 28%(+1.05) on Play/Pause and 6.7%(+0.25) on Transformation. In average, score rated on Efficiency has an increase of 22%((31%+28%+6.7%)/3=22%).

Voice control has a prominent help for the play and pause of the virtual character. The average score increases from 3.8 to 4.85(+28%), while Std decreases 0.85(-64%). This result strongly proves that voice recognition is perfectly suitable for AR Taiji, which is not beyond the expectations since users will naturally put their devices on a desk and imitate character movements. We can conclude that the ability to interact with virtual content by voice commands would reach the potential of human motion sequence learning. Basically, this tells us that it's easy to have good results to reduce interaction complexity on the most frequently used functions.

Moreover, the average score and std of Form selection changed 31%(+1.1) and 45%(-0.46) respectively. Voice recognition saves the trouble of having to click the menu bar multiple times, thus also contributes to the overall efficiency.

The average score has a small increase on Transformation, from 3.75 to 4.0(+6.7%). Similarly, we apply one-tail t-test to the data. Resulting $t = -0.62$ and a corresponding $p = 0.27 > 0.05$, which means there is no significant change in Efficiency of Transformation. The reason behind this could be that Transformation is an activity that need to be operated multiple times in a short period of time. Due to the inevitable response time of voice control, voice control would not have benefited it a lot.

5.4 VR option

The primary objectives of importing a VR option are (1)Enhancing user privacy protection and (2)Improving user immersion. As mentioned in3, it is essential in a mobile

AR App to protect users' privacy from being violated. We want to investigate how VR can help in protecting privacy. Therefore, we first asked whether participants were willing to give camera permission. Participants gave their responses based on the scale of 1(Strongly disagree), 2(Somewhat disagree), 3(Neither agree nor disagree), 4(Somewhat agree) and 5(Strongly agree). Results are shown in Table 5.19.

Versions	Mean	Std	p-value
V1 Baseline	3.5	0.67	NA
V4 VR	4.45	0.67	$\approx \mathbf{0}$

Table 5.19. Comparison of attitudes of giving camera permission on V1 and V4

All responses from Version 1 and Version 4 are somewhere in between *Somewhat agree* and *Strongly agree*. It's normal for people to acquiesce in AR App obtaining camera permissions, and this situation will not be changed by a new VR option.

On this basis, we continue to verify whether it will increase the feeling that user privacy is well protected. The second question is: Do you think your privacy is under protected? This question has the same scales on Agree/Disagree with as previous question. Results are shown in Figure 5.12

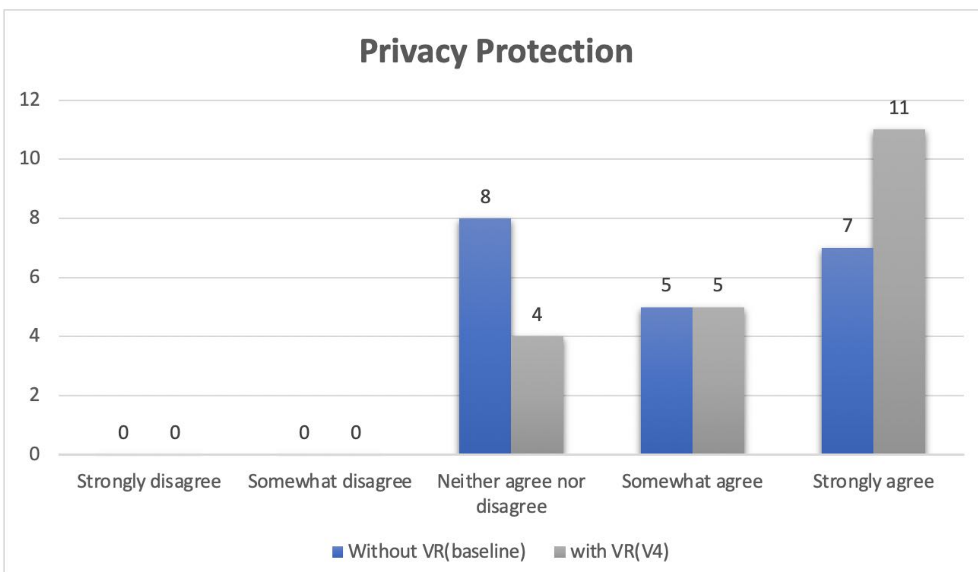


Figure 5.12. Privacy score with and without VR: Number of total ratings = 20(subjects) * 2(scores) = 40

In baseline model, only 7 participants strongly agreed they have a sense of protection but 8 participants remained indifferent. After adding a VR feature in Version 4, only 4 people felt indifferent to the privacy protection question. Yet 11 responses(55%) strongly agreed that their privacy is well protected. The average score increases from 3.5 to 4.45, which proves that a VR option did give users a sense that their privacy is protected from being violated.

Versions	Mean	Std	p-value
V1 Baseline	3.95	0.86	NA
V4 VR	4.35	0.79	**0.004**

Table 5.20. Comparison of attitudes of privacy protection in V1 and V4

Versions	Mean/Std(Male)	p-value	Mean/Std(Female)	p-value
V1 Baseline	3.75/0.75	NA	4.25/0.78	**0.02****
V4 VR	4.25/0.57	**0.01**	4.5/0.85	0.09

Table 5.21. Comparison of attitudes of privacy protection in V1 and V4(Gender difference). P values are a comparison between V1 and V4. P value* is a comparison of male and female users on V1.

5.4.1 Satisfaction on VR

Additionally, we explore how this new VR environment option contributes to user satisfaction. From 3 we learned that a VR option not only increases user engagement, but also saves rendering cost. Thus participants were quired to score satisfaction in terms of these two factors. Participants gave their responses based on the scale of 1(Extremely dissatisfied), 2(Somewhat dissatisfied), 3(Neither satisfied nor dissatisfied), 4(Somewhat satisfied) and 5(Extremely satisfied). Results are shown in Figure 5.13.

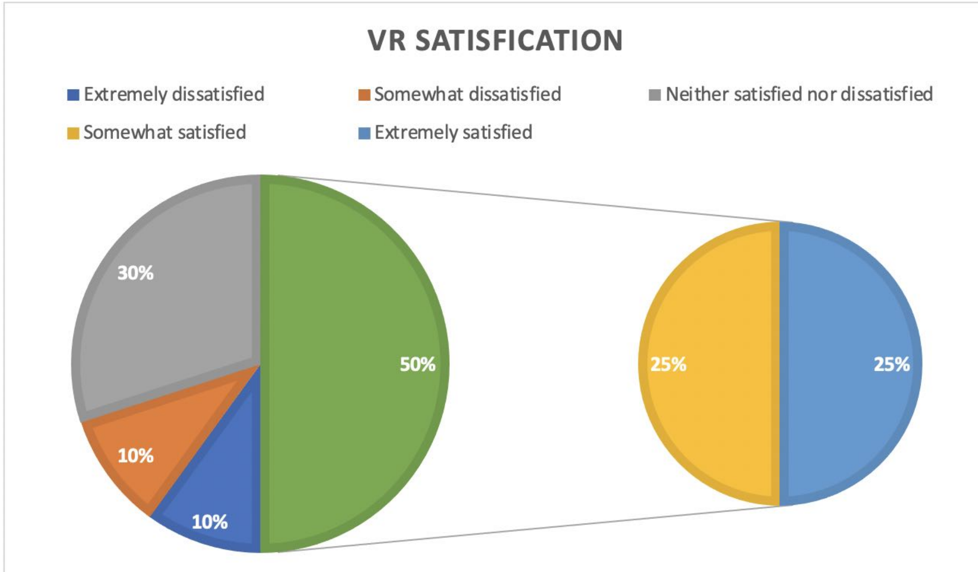


Figure 5.13. User Satisfaction on VR option: Number of total ratings = 20(subjects) * 1(scores) = 20

Most of responses(50%) were extremely or somewhat satisfied with the brand new experience it brought. 80% of participants' attitudes are better than or equal to the third indifferent level(Neither dissatisfied nor satisfied). Two users felt extremely dissatisfied and two users felt somewhat dissatisfied. The average score is 3.45 and Std is 1.24. We have to admit that this new feature to run AR Taiji with VR mode is not as popular as voice recognition or other features since it has a rather low average satisfaction. One of the reasons may be that the VR background(see Figure 3.13) we provide is too monotonous and dull. The other reason may be that AR users prefer to render the real physical environment instead of the virtual environment.

Chapter 6 Conclusions and Future Work

In this thesis research, we report evaluation on four versions of AR Taiji app and corresponding experiments conducted.

6.1 Summary on Satisfaction

According to the questionnaire survey on three modifications comparing with the baseline, user satisfaction on user interface and functionality experience has an average increase of 13.3%(Table 5.1). Score rated on Efficiency has an increase of 22% on average(Table 5.18).

In Version 2 Part 1, we implement a new layout with an integration button(Figure 3.8). In Part 2, a list of virtual characters for users to choose from is added(Figure 3.10 3.11). The overall satisfaction on Part 1 increases by 0.75(Table 5.1). Yet the overall satisfaction on Part 2 increases by only 0.05(Table 5.1), which is not a significant improvement. With further analysis, participants continue to provide a positive review on the increased screen area brought by the new layout(Table 5.5, Table 5.7, Table 5.6, Table 5.8). Adding such an implicit layout increases physical efforts for functions such as Transformation. By definition of **Efficiency** in Chapter 4, this may potentially reduce the score rated on Efficiency. But according to Table 5.2, Table 5.3 and Table 5.4, responses from users are not changed significantly from V1 to V2(p values > 0.05). Therefore a feasible design is to increase screen area as much as possible with an implicit layout.

In Version 3, voice recognition is introduced as a user-centered design to provide flexibility. Instead of interacting with virtual world by touching User Interface, users are able to free their hands by speaking directly to the App. This means that they are able to imitate and learn human motion sequences without touching the screen. The overall satisfaction has increased by 25% from baseline(Table 5.1), giving rise to the highest level of satisfaction score. The complexity of voice recognition is negligible to most of the participants, compared to higher efficiency and engagement it brings(Table 5.14). The function that has the most noteworthy growth under voice control among all functions in user satisfaction is Play/Pause, which has an approximately 30% increase(Table 5.18). Due to the fact that Play/Pause is the most commonly used and most useful function, it's not unexpected to see such increase of satisfaction in the Play/Pause function.

In Version 4, we implemented a VR option for a user to launch the app without a camera on. The original intention is to protect privacy since camera usage may leak the user's environment. Statistical analysis shows that most users agree on giving camera permissions to AR Taiji app(Table 5.19). They either strongly agreed or somewhat agreed to give camera permission to AR Taiji app. VR mode is not easily accepted by all users since the average satisfaction score on VR is 3.7, which is slightly lower than the overall satisfaction of baseline. (Table 5.1)

In summary, an implicit layout with an integration button(Figure 3.8), a list of new virtual characters(Figure 3.9, Figure 3.10, Figure 3.11) and a voice control unit(Figure 3.12) improve user HCI experience.

6.2 Study on Efficiency

Efficiency is one of the most important factors in HCI. Objectively speaking, modifications like voice control and integration button would definitely involve additional interventions by users, thus increase the chances of fatigue. However, since we are exploring user-centered design, we investigate users' attitudes as measurement of the efficiency of different versions. A summary table is shown in Table 6.1. Note that modifications on Version 4 are not efficiency-oriented, thus we only listed baseline, Version 2 and Version 3 on the table.

Functions	Version 1	Version 2 Part 1	p-value	Version 3	p-value
Initialization	3.2	**3.95**	$\approx$ **0.00**	3.2	0.5
Transformation	3.75	3.7	0.44	**4**	0.27
Play/Pause	3.8	3.8	0.5	**4.85**	**0.002**
Form Selection	3.6	3.95	**0.045**	**4.7**	$\approx$ **0.00**

Table 6.1. Comparisons of scores rated from efficiency for all functions in four versions. P values are a comparison with V1.

1. **Initialization:** Efficiency scores have not changed seriously on Version 3, but it has a huge improvement on Version 2 Part 1. The highest score on the Initialization function was achieved on Version 2 Part 1 with an integration button. It makes sense because a wider screen area evoked by the new layout will contribute to an easier Initialization process.

2. **Transformation:** There is no improvement in Transformation functions in Version 2 P1 but a small increase in Version 3. Efficiencies of the Transformation function become a little bit lower in Version 2 P1, affected by the implicit layout that involved additional operations.

3. **Play/Pause:** Play/Pause functions had a huge change on efficiency score in Version 3. As previously mentioned in Chapter 3, Play/Pause are the most commonly used functions, which require a large amount of physical efforts and interactions. Voice recognition would critically prevent unnecessary activities, thus leading to an increase in Efficiency score.

4. **Form Selection:** The highest Efficiency score on the Form Selection function is 4.7 from Version 3. By voice command, users could effortlessly select favorable forms from the list by simply saying the form number.

6.3 Study on Complexity

Another factor of HCI that we have investigated is complexity. Any modifications that potentially increase system complexity should be carefully reviewed because they may lead to a higher learning cost. In baseline, 40% of participants had trouble with interactions and required tech support. In order to allow all users to fully experience functions, we provided instruction on those functions in place.

In Version 3, we added voice recognition which required additional learning overhead. However, responses agreed that the complexity of voice recognition is not high. The most straightforward reason is that speech recognition technology is common in daily life, such as Google Voice, Amazon Alexa, Microsoft Cortana, and Apple's Siri. Those experiences largely reduce the cognitive overhead of using voice control. It's a successful HCI design to implement a voice recognition in an AR App with human motion sequence learning.

6.4 Future Work

As people having growing expectations towards AR technology, it is vital to evaluate the user-centered design and experience of AR systems progressively. We received mixed opinions towards AR Taiji app modifications on annoying issues such as an unstable initialization function and a low quality virtual character. We also receive compliments for the strengths like the voice control and an immersive user experience. Some pointed out unreasonable design in the system e.g., pause button will only stop the sequence of virtual character instead of pausing in the middle of the movement. Others gave their opinions on what can be improved in the future, such as the latency of voice control and a believable VR environment in Version 4. Additionally, one user suggests that the rotation function can be divided into higher granularity when voice control is activated. We could say "Leap Rotate" to drastically change the orientation of virtual character or just say "Rotate" to slightly change the orientation of virtual character. Another user suggests that it would be better if they can lock the explicit layout by double-clicking the gear button, which is a great idea.

In the future, we plan to prepare a more detailed and comprehensive survey on HCI principles that involve at least 30 participants to be statistically well-justified. Besides functions like voice control and VR, we plan to explore more functionalities for a better learning effect.

Appendix | Questionnaire

Default Report

Apply HCI

June 30, 2021 10:01 AM MDT

1 - Gender

#	Field		Choice Count
1	Male	60.00%	12
2	Female	40.00%	8
3	Non-binary / third gender	0.00%	0
4	Prefer not to say	0.00%	0
5	Not listed	0.00%	0
			20

Showing rows 1 - 6 of 6

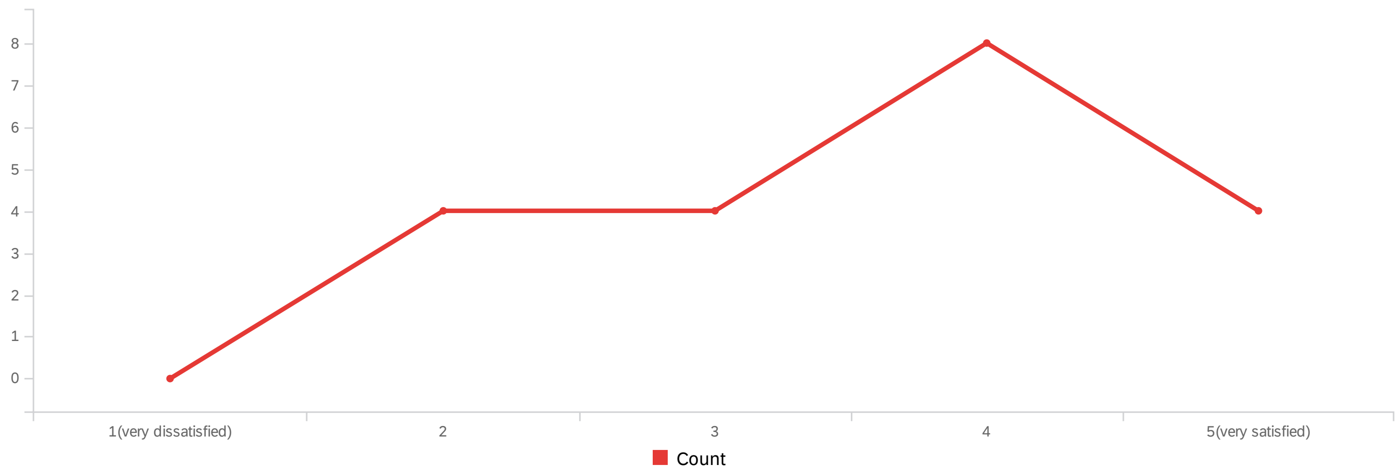
2 - First Language

3 - Have you ever tried an AR App before?

#	Field		Choice Count
1	Yes	90.00%	18
2	No	10.00%	2
			20

Showing rows 1 - 3 of 3

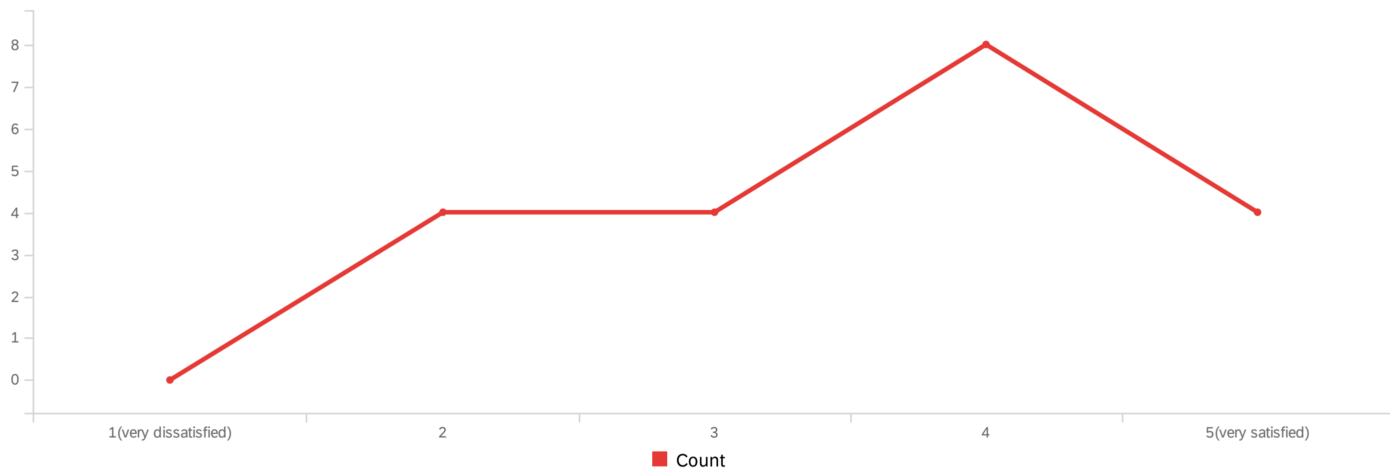
4 - What is your overall satisfaction over the baseline model on a scale from 1(lowest) to 5(highest)?



#	Field	Choice Count	
1	1(very dissatisfied)	0.00%	0
2	2	20.00%	4
3	3	20.00%	4
4	4	40.00%	8
5	5(very satisfied)	20.00%	4
			20

Showing rows 1 - 6 of 6

5 - What is your overall satisfaction over the baseline model on a scale from 1(lowest) to 5(highest)?



#	Field	Choice Count	
1	1(very dissatisfied)	0.00%	0
2	2	20.00%	4
3	3	20.00%	4
4	4	40.00%	8
5	5(very satisfied)	20.00%	4
			20

Showing rows 1 - 6 of 6

6 - Do you need any tech support while using the App?

#	Field	Choice Count	
1	Definitely not	10.00%	2
2	Probably not	25.00%	5
3	Might or might not	25.00%	5
4	Probably yes	5.00%	1
5	Definitely yes	35.00%	7
			20

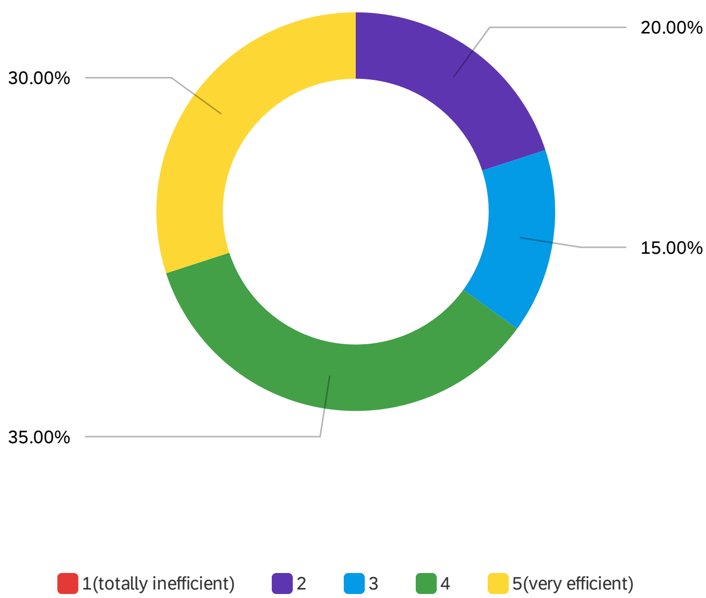
Showing rows 1 - 6 of 6

7- Do you think the functions of this system is consistent?

#	Field	Choice Count	
1	Definitely not	0.00%	0
2	Probably not	10.00%	2
3	Might or might not	20.00%	4
4	Probably yes	35.00%	7
5	Definitely yes	35.00%	7
			20

Showing rows 1 - 6 of 6

8 - How efficient do you feel when using the Version 1 model doing the following activities?



#	Field	1(totally inefficient)		2		3		4		5(very efficient)		Total
1	Rotation, Translation, Scaling	0.00%	0	20.00%	4	15.00%	3	35.00%	7	30.00%	6	20
2	Select forms from menu button	0.00%	0	15.00%	3	35.00%	7	25.00%	5	25.00%	5	20
3	Place virtual character on the screen	5.00%	1	35.00%	7	20.00%	4	15.00%	3	25.00%	5	20

Showing rows 1 - 3 of 3

9 - What do you think of the initial process(place the virtual character on the screen)?

#	Field	Choice Count	
1	1(very hard to use)	25.00%	5
2	2	30.00%	6
3	3	5.00%	1
4	4	20.00%	4
5	5(very easy to use)	20.00%	4
			20

Showing rows 1 - 6 of 6

10 - What do you think of the size of the screen's visual field

#	Field		Choice Count
1	1(view is too narrow)	15.00%	3
2	2	20.00%	4
3	3(enough view)	45.00%	9
4	4	5.00%	1
5	5(view is wide)	15.00%	3
			20

Showing rows 1 - 6 of 6

11 - What do you think of system complexity of version 1 baseline model？

#	Field	Choice Count	
1	1(very simple)	35.00%	7
2	2	15.00%	3
3	3(normal)	45.00%	9
4	4	5.00%	1
5	5(very complicated)	0.00%	0
			20

Showing rows 1 - 6 of 6

12 - How do you feel about the interaction with this App. Is it easy to use?

#	Field	Choice Count	
1	Strongly disagree	0.00%	0
2	Somewhat disagree	5.00%	1
3	Neither agree nor disagree	20.00%	4
4	Somewhat agree	30.00%	6
5	Strongly agree	45.00%	9
			20

Showing rows 1 - 6 of 6

13 - What do you feel about User Privacy in the following situations?

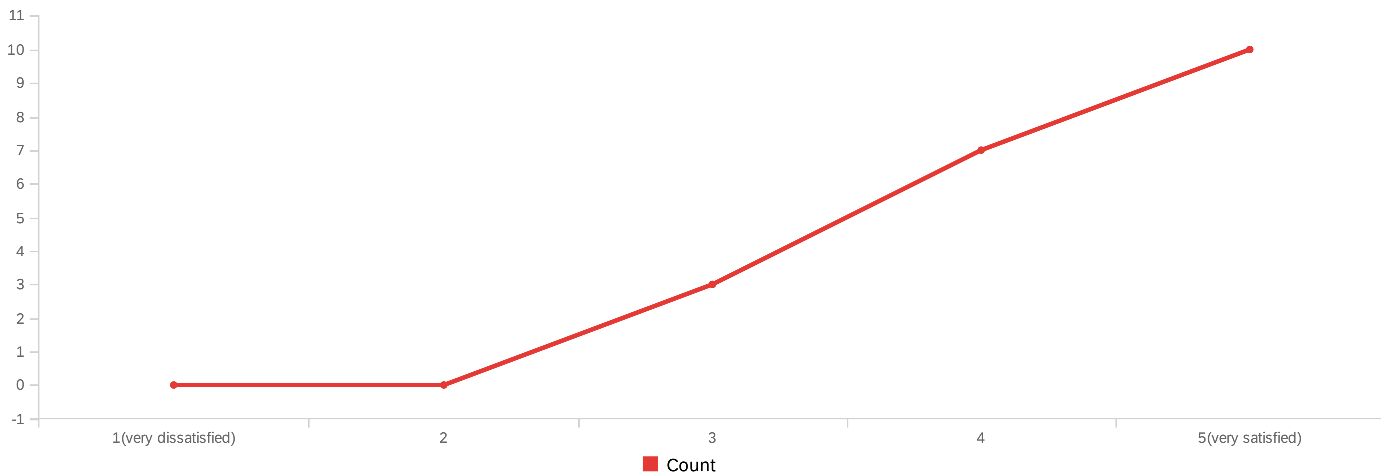
#	Field	Strongly disagree		Somewhat disagree		Neither agree nor disagree		Somewhat agree		Strongly agree		Total
1	I am willing to give permission to use camera	0.00%	0	5.00%	1	20.00%	4	35.00%	7	40.00%	8	20
2	I think my privacy is well-protected	0.00%	0	0.00%	0	40.00%	8	25.00%	5	35.00%	7	20

Showing rows 1 - 2 of 2

14 - What is your overall satisfaction over the Version 2 model on a scale from 1(lowest) to 5(highest)?

#	Field	Choice Count	
1	1(very dissatisfied)	0.00%	0
2	2	0.00%	0
3	3	15.00%	3
4	4	35.00%	7
5	5(very satisfied)	50.00%	10
			20

Showing rows 1 - 6 of 6



15 - How efficient do you feel in the following situations of Version 2 model on a scale from 1 to 5?

#	Field	1(inefficient at all)		2		3		4		5(very efficient)		Total
1	Rotation, Translation, Scaling	0.00%	0	10.53%	2	31.58%	6	26.32%	5	31.58%	6	19
2	Select forms from menu button	0.00%	0	0.00%	0	35.00%	7	35.00%	7	30.00%	6	20
3	Place virtual character on the screen	0.00%	0	10.00%	2	30.00%	6	15.00%	3	45.00%	9	20

Showing rows 1 - 3 of 3

16 - How do you feel about the AR view in the following situations of Version 2 model on a scale from 1 to 5?

#	Field	1(view is too narrow)		2		3(enough view)		4		5(view is wide)		Total
1	Fold up the integration button	0.00%	0	0.00%	0	10.00%	2	25.00%	5	65.00%	13	20
2	Unfold the integration button	0.00%	0	15.00%	3	65.00%	13	5.00%	1	15.00%	3	20

Showing rows 1 - 2 of 2

17 - What is your overall satisfaction over the version 3 model on a scale from 1(lowest) to 5(highest)?

#	Field	Choice Count	
1	1(very dissatisfied)	0.00%	0
2	2	5.00%	1
3	3	5.00%	1
4	4	25.00%	5
5	5(very satisfied)	65.00%	13
			20

Showing rows 1 - 6 of 6

18 - How efficient do you feel in the following situations of version 3 model on a scale from 1 to 5?

#	Field	1(inefficient at all)		2		3		4		5(very efficient)		Total
1	Voice command for basic transformation	5.00%	1	15.00%	3	10.00%	2	15.00%	3	55.00%	11	20
2	Voice command for PLAY/PAUSE	0.00%	0	0.00%	0	5.00%	1	5.00%	1	90.00%	18	20
3	Voice command for form selection	0.00%	0	0.00%	0	5.00%	1	20.00%	4	75.00%	15	20

Showing rows 1 - 3 of 3

19 - How do you feel about the complexity of Voice Control

#	Field	Choice Count	
1	1(very simple)	45.00%	9
2	2	10.00%	2
3	3(normal)	35.00%	7
4	4	5.00%	1
5	5(very complicated)	5.00%	1
			20

Showing rows 1 - 6 of 6

20 - Did you give the permission to us to use camera?

#	Field	Choice Count	
1	Yes	100.00%	20
2	No	0.00%	0
			20

Showing rows 1 - 3 of 3

21 - What do you feel about giving the permission

#	Field	Choice Count	
21	Extremely unhappy	0.00%	0
22	Somewhat unhappy	0.00%	0
23	Neither happy nor unhappy	60.00%	12
24	Somewhat happy	30.00%	6
25	Extremely happy	10.00%	2
			20

Showing rows 1 - 6 of 6

22 - What do you think of the virtual character in our App?

#	Field	Choice Count	
16	Dislike a great deal	0.00%	0
17	Dislike somewhat	15.00%	3
18	Neither like nor dislike	40.00%	8
19	Like somewhat	35.00%	7
20	Like a great deal	10.00%	2
			20

Showing rows 1 - 6 of 6

23 - What do you think of the three virtual characters in this version?

#	Field	Dislike a great deal		Dislike somewhat		Neither like nor dislike		Like somewhat		Like a great deal		Total
1	Woman	0.00%	0	5.00%	1	25.00%	5	35.00%	7	35.00%	7	20
2	Alien	30.00%	6	15.00%	3	15.00%	3	30.00%	6	10.00%	2	20
3	Guy	0.00%	0	0.00%	0	15.00%	3	50.00%	10	35.00%	7	20

Showing rows 1 - 3 of 3

24 - What do you think of respondence time of voice control.

#	Field	Choice Count	
1	Extremely dissatisfied	0.00%	0
2	Somewhat dissatisfied	25.00%	5
3	Neither satisfied nor dissatisfied	25.00%	5
4	Somewhat satisfied	20.00%	4
5	Extremely satisfied	30.00%	6
			20

Showing rows 1 - 6 of 6

25 - What do you feel about User Privacy in the following situations?

#	Field	Strongly disagree		Somewhat disagree		Neither agree nor disagree		Somewhat agree		Strongly agree		Total
1	I am willing to give permission to use microphone	0.00%	0	0.00%	0	15.00%	3	30.00%	6	55.00%	11	20
2	I think my privacy is well-protected	0.00%	0	0.00%	0	20.00%	4	25.00%	5	55.00%	11	20

Showing rows 1 - 2 of 2

26 - What is your overall satisfaction over the Version 4 model on a scale from 1(lowest) to 5(highest)?

#	Field	Choice Count	
1	1	5.00%	1
2	2	10.00%	2
3	3	40.00%	8
4	4	30.00%	6
5	5	15.00%	3
			20

Showing rows 1 - 6 of 6

27 - What do you think of the VR option to start without camera.

#	Field	Choice Count	
1	Extremely dissatisfied	10.00%	2
2	Somewhat dissatisfied	10.00%	2
3	Neither satisfied nor dissatisfied	30.00%	6
4	Somewhat satisfied	25.00%	5
5	Extremely satisfied	25.00%	5
			20

Showing rows 1 - 6 of 6

28 - What do you feel about User Privacy in the following situations?

#	Field	Strongly disagree		Somewhat disagree		Neither agree nor disagree		Somewhat agree		Strongly agree		Total
1	I'm willing to give permission to use camera	0.00%	0	0.00%	0	10.00%	2	35.00%	7	55.00%	11	20
2	I think my privacy is well-protected	0.00%	0	0.00%	0	20.00%	4	25.00%	5	55.00%	11	20

Showing rows 1 - 2 of 2

Bibliography

[1] Adele Botha, Darelle Greunen *Mobile human-computer interaction perspective on mobile learning.*, 2010.

[2] Allen Elaine, Seaman Christopher *"Likert Scales and Data Analyses". Quality Progress. pp. 64–65.*, 2007.

[3] Andreas D., R. Grasset *Applying HCI principles to AR systems design. 2nd International Workshop on Mixed Reality User Interfaces*, 2007

[4] Ann Nosseir, Derek Flood *Mobile Development Process Spiral. Seventh International Conference on Computer Engineering & Systems. IEEE*, 2013

[5] Antoni Moore *A tangible augmented reality interface to tiled street maps and its usability testing. In 12th International Symposium on Spatial Data Handling, Vienna, Austria, pages 12-14.*, 2006.

[6] AYC Nee, SK Ong *Augmented reality applications in design and manufacturing. CIRP Annals, 1(2)*, 2012

[7] cgtrader *"Free 3D models", Website. Link: https://https://www.cgtrader.com/free-3d-models/*, 2021.

[8] Donald McMillan, Alistair Morrison *Categorised Ethical Guidelines for Large Scale Mobile HCI*, 2013

[9] Grush B *"An Introduction to 24 Posture Yang Tai Chi", Website. Link: http://taichisnob.blogspot.com/2013/03/ an-introduction-to-24-posture-yang-tai.html*, 2013.

[10] Guillaume Godet-Bar, Dominique Rieu *HCI and business practices in a collaborative method for augmented reality systems*, 2010

[11] Gun A. Lee, Claudia Nelles, Mark Billinghurst, and Gerard Jounghyun Kim *Immersive Authoring of Tangible Augmented Reality Applications. In ACM International Symposium on Mixed and Augmented Reality 2004 (ISMAR 2004), Arlington, VA, USA*, 2004.

[12] Hiromitsu Sato, Michael Cohen *Using Motion Capture for Real-time Augmented Reality Scenes.*, 2010.

[13] Iachello, G., Hong, J. *End-User Privacy in Human-Computer Interaction. FNT in HCI 1, 1 , 1–137.*, 2007.

[14] Irshad, Shafaq, et al *User experience evaluation of mobile AR services." Proceedings of the 12th International Conference on Advances in Mobile Computing and Multimedia.*, 2014.

[15] J. Edward Swan II, Joseph L. Gabbard *Survey of User-Based Experimentation in Augmented Reality*, 2005

[16] Jesse Scott, Robert Collins, Christopher Funk and Yanxi, Liu. *4D Model-Based Spatiotemporal Alignment of Scripted Taiji Quan Sequences. 10.1109/ICCVW.2017.99.*, 2017.

[17] Mingyang Li, Anastasios I Mourikis *High-precision, consistent EKF-based visual-inertial odometry. The International Journal of Robotics Research 32, 6 (2013), 690–711.*, 2013.

[18] Manova *"Statistics Solutions", Website. Link: https://www.statisticssolutions.com/manova-analysis-paired-sample-t-test/*, 2021.

[19] Mekni, Mehdi, and Andre Lemieux. *"Augmented reality: Applications, challenges and future trends." Applied Computational Science 20 (2014): 205-214.*, 2014.

[20] Merino, Leonel, Alexandre Bergel, and Oscar Nierstrasz. *"Overcoming issues of 3D software visualization through immersive augmented reality." 2018 IEEE Working Conference on Software Visualization (VISSOFT). IEEE.*, 2018.

[21] M. P. Garofalo. *"Taijiquan 24 form Yang style.* [Online; accessed August 6, 2017].

[22] N. Nazri and D. Rambli *Current limitations and opportunities in mobile augmented reality applications. International Conference on Computer and Information Sciences (ICCOINS), pp. 1-4, doi: 10.1109/ICCOINS.2014.6868425.*, 2014.

[23] Niels Henze, Martin Pielot, et al *My app is an experiment: Experience from user studies in mobile app stores." International Journal of Mobile Human Computer Interaction (IJMHCI) 3.4 (2011): 71-91.*, 2011.

[24] Regina Connolly, Frank Bannister *Consumer Trust in Internet Shopping in Ireland: Towards the Development of a More Effective Trust Measurement Instrument. Journal of Information Technology, 22(2):102-118*, 2007

[25] Royce, W. *Managing the development of large software systems: Concepts and techniques. Proceedings of the 9th International Conference on Software Engineering, 26, 328-338*, 1970

[26] Tan Qing, Chang, William *Location-Based Augmented Reality for Mobile Learning: Algorithm, System, and Implementation. Electronic Journal of e-Learning, v13 n2 p138-148*, 2015.

[27] *https://data.library.virginia.edu/using-and-interpreting-cronbachs-alpha/*

[28] *https://apps.apple.com/us/app/ar-taiji/id1495725552*