Master’s Thesis

Manipulation Techniques for Image-based Representations of 3D Objects in a 3D Scene

Juan Carlos Yu Yu

March 16, 2012

Department of Information Systems
Graduate School of Information Science
Nara Institute of Science and Technology
A Master’s Thesis
submitted to Graduate School of Information Science,
Nara Institute of Science and Technology
in partial fulfillment of the requirements for the degree of
MAster of ENGINEERING

Juan Carlos Yu Yu

Thesis Committee:
Professor Hirokazu Kato (Supervisor)
Professor Tsukasa Ogasawara (Co-supervisor)
Assistant Professor Sei Ikeda (Co-supervisor)
Assistant Professor Yuki Uranishi (Co-supervisor)
Manipulation Techniques for Image-based Representations of 3D Objects in a 3D Scene*

Juan Carlos Yu Yu

Abstract

The field of 3D computer graphics has received much attention in the recent decade. Nowadays, 3D computer graphics allow rendering of complex and realistic 3D models. However, using the geometry-based approach to model real world scenes or objects is a tedious process; as scene complexity or level of detail increases, modeling time increases. Also, the geometry-based approach achieves limited levels of detail and realism. Recently, image-based rendering techniques have been considered as a viable alternative because scene complexity or level of detail does not affect modeling time. Moreover, by sampling the real world through photographs, a higher degree of detail and realism is achievable compared to the geometry-based approach.

Similar to the geometry-based approach, image-based rendering techniques has its limitations. When using a geometry-based 3D model, the scene or object can be viewed at arbitrary viewpoints. The main limitation of using a discrete set of images is the limited viewpoints available. Research that use image-based rendering techniques focus on this problem.

Majority of research address the problem of limited viewpoints through image-based modeling or through image-based rendering. Image-based modeling reconstructs a 3D model of a scene or object through an image-data set and its related geometric parameters. On the other hand, image-based rendering focuses on view synthesis, which generates new viewpoints from a limited image data set.

Few research have considered the 3D manipulation techniques that will allow users to interact with an image-data set without view synthesis or 3D reconstruction. This is because after 3D reconstruction or view synthesis, conventional 3D manipulation tools can be used to interact with the object or scene. In cases where the image data set is not processed for 3D reconstruction or view synthesis, manipulation techniques can provide a solution for the problem of limited viewpoints.

This thesis proposes two 3D manipulation techniques for objects represented as multi-viewpoint images. We do not perform view synthesis or 3D reconstruction using the available images but rather, we present a technique that will allow the user to manipulate the object at any orientation despite the limited viewpoints. We implemented a prototype system on a mobile device that features these proposed manipulation tools. The design of the prototype system presents a unique problem setting: combining an actual 3D model of a scene and a variable object represented through multi-viewpoint images. We designed the prototype system in such a way that it can be used as a building block for a more elaborate application. We evaluate the proposed manipulation techniques through the prototype using two user studies. Through these user studies, we found that there was no significant difference in performance between the proposed manipulation tools and conventional 3D manipulation tools. Aside from this finding, we also found that it takes the same amount of time to be proficient with the proposed manipulation tools as compared to the conventional 3D manipulation tools. Lastly, we were able to determine that we could effectively increase the usability of our system by decreasing the number of images needed for object representation without affecting its performance.

Keywords:

3D interaction technique, image-based rendering technique, user studies
3Dシーンにおける3Dオブジェクトの多視点画像切り替え手法*

Juan Carlos Yu Yu

内容梗概

近年、三次元コンピュータグラフィックス（3DCG）技術は複雑で写実的な描画を可能としている。グラフィックス描画技術の向上によって、計算コストと引き換えに、実世界シーンやオブジェクトの三次元形状モデルを用いて高精細・現実感を表現できる。一方で、処理に長時間を費やさず、写実的な描画を可能とする画像ベースの手法が注目されている。特に現実世界の実写画像を用いるため、形状ベースに比べて高い写実性を有している。しかしながら、形状ベースの手法では任意の地点からの見方を描画することは容易であるが、画像ベースの手法では視点が限定されているという制約がある。この視点制御問題を解決するために、画像ベースモデリングや画像ベースレンダリングなどの手法が提案されている。画像ベースモデリングでは、画像データから推定されるパラメータによってシーンおよびオブジェクトの三次元を再構成する。それに対して、画像ベースレンダリングは画像データセットから中間視点画像を生成する。従来、画像ベースの任意視点画像生成の研究は行われてきただが、画像生成を行わずに画像データセットのみで表現される3DCGを操作する手法については十分な研究がなされていない。これまでは任意視点画像を生成することで、シーンおよびオブジェクトとのインタラクションは可能であったため、画像データセットのみで表現される3DCGの扱いに着目されなかったと考えられる。そのため、画像データセットのみの場合における三次元シーンおよびオブジェクトに対する操作技術は開発されず、また

* 奈良先端科学技术大学院大学情報科学研究科情報処理学専攻修士論文、NAIST-IS-MT1051139、2012年2月2日
ユーザインタフェース設計に関しても指針が無い。また、今日ではモバイル機器上での三次元オブジェクトを扱える環境が整いつつあるが、計算コストの低い画像ベース表現は親和性が高いと考えられる。これらを踏まえ、本論文では、複数視点画像で表現されたオブジェクトの三次元操作として二つの手法を提案する。そして、それらをタブレット機器上で実装する。本手法では任意視点画像生成はせず、画像データセットに限定された視点位置画像によって回転するオブジェクトを表現する。さらに、モデルで表現される三次元シーンと複数視点画像で表現されるオブジェクトを組み合わせた実用的な問題設定を行った。2種類の被験者実験を通じて、従来手法と同等の描画表現および学習のしやすさがあると示した。

キーワード

3Dユーザインタフェース、イメージベースドレンダリング、ユーザスタディ
Contents

1 Introduction 1
  1.1. Overview ......................................................... 1
    1.1.1 Traditional 3D Computer Graphics and Image-based Render-
         ing .......................................................... 1
    1.1.2 3D Interaction Techniques ................................... 2
  1.2. Problem Setting .................................................. 4
  1.3. Motivation and Research Theme .................................... 4
  1.4. Thesis Overview .................................................. 7

2 Related Literature 8
  2.1. Image-based Rendering and Manipulation Techniques ............... 8
    2.1.1 Rendering with Explicit Geometry and Implicit Geometry 8
    2.1.2 Rendering with No Geometry ................................ 12

3 Methodology 14
  3.1. Prototype Implementation .......................................... 14
    3.1.1 Target Object ............................................... 16
    3.1.2 Target Scene ............................................... 16
  3.2. Manipulation Tools ............................................... 19
    3.2.1 Proposed Manipulation Techniques .............................. 19
    3.2.2 Conventional Manipulation Tools .............................. 23

4 User Study 1: Performance and Preference 31
  4.1. Current Prototype State .......................................... 31
  4.2. Participants .................................................... 32
6 Conclusion and Discussion

6.1. Proposed Manipulation Techniques’ Performance ....... 61
6.2. Users’ Perception of Proposed Manipulation Techniques ..... 61
6.3. Learnability of Proposed Manipulation Techniques ........ 62
6.4. Conclusion ............................................ 62

Acknowledgements ............................................ 64

References ..................................................... 66

Appendix ....................................................... 69

A. User Study 1 ................................................. 69
   A.1 Experiment 1 ............................................. 69
   A.2 Experiment 2 ............................................. 69
B. User Study 2 ................................................ 70
List of Figures

1.1 2D input devices for 3D manipulation .......................... 3
1.2 3D Widgets and ARCBALL function ............................ 4
1.3 Medical application using 3D manipulation techniques. .......................... 5
1.4 Education application using 3D technology .......................... 5
1.5 System diagram .................................................. 6

2.1 Image-based modeling .............................................. 9
2.2 The Phototourism system .......................................... 10
2.3 View morphing ..................................................... 11
2.4 Yu et al.’s [1] research overview .................................. 12

3.1 Key elements of the prototype system ............................ 15
3.2 The overall prototype screenshot .................................. 17
3.3 Using Yu et al.’s [1] approach in our thesis ........................ 17
3.4 Teapot spritesheet used in the prototype .......................... 18
3.5 The modelview matrix .............................................. 18
3.6 An illustrative explanation of the main idea behind the two proposed techniques .......................... 20
3.7 Four snapshots of an object taken around the Y-axis at 90° angles. .......................... 20
3.8 Step by step explanation of proposed tools: Step 1 .......................... 24
3.9 Step by step explanation of proposed tools: Step 2 .......................... 24
3.10 Step by step explanation of proposed tools: Step 3 .......................... 25
3.11 Step by step explanation of proposed tools: Step 4 .......................... 25
3.12 Step by step explanation of proposed tools: Step 5 .......................... 26
3.13 Shows the coordinate system used for all interactions. .......................... 26
3.14 Shows the next image of swipe to the right for PT1 and PT2 .......................... 27
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1</td>
<td>Implemented tools for the prototype during User Study 1</td>
<td>32</td>
</tr>
<tr>
<td>4.2</td>
<td>The tools for each condition in User Study 1: Experiment 1</td>
<td>35</td>
</tr>
<tr>
<td>4.3</td>
<td>Independent variables for Experiment 2 tasks</td>
<td>38</td>
</tr>
<tr>
<td>4.4</td>
<td>Average completion time for all conditions under Experiment 1: Task 1</td>
<td>41</td>
</tr>
<tr>
<td>4.5</td>
<td>Average completion time for all conditions under Experiment 1: Task 2</td>
<td>42</td>
</tr>
<tr>
<td>4.6</td>
<td>Average completion time for all conditions under Experiment 2: Task 1</td>
<td>43</td>
</tr>
<tr>
<td>4.7</td>
<td>Average difficulty rankings for all conditions under Experiment 2: Task 1</td>
<td>44</td>
</tr>
<tr>
<td>5.1</td>
<td>The list of implemented manipulation techniques during the second user study.</td>
<td>50</td>
</tr>
<tr>
<td>5.2</td>
<td>Conventional tool set: The complete list of the available manipulation tools with the gestures associated for each tool.</td>
<td>55</td>
</tr>
<tr>
<td>5.3</td>
<td>PT1 tool set: The complete list of the available manipulation tools with the gestures associated for each tool.</td>
<td>55</td>
</tr>
<tr>
<td>5.4</td>
<td>Average number of tests taken by participants for each condition in User Study 2</td>
<td>57</td>
</tr>
<tr>
<td>5.5</td>
<td>Average confidence rating of participants when answering the first test for each condition under User Study 2</td>
<td>57</td>
</tr>
<tr>
<td>5.6</td>
<td>Average difficulty rankings both conditions under User Study 2</td>
<td>57</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

1.1. Overview

1.1.1 Traditional 3D Computer Graphics and Image-based Rendering

Extensive research have been done in the past to develop new ways of representing and modeling 3D scenes and objects. Image-based modeling (IBM) and Image-based rendering techniques (IBR) have recently been given a lot of attention as an alternative to the conventional 3D computer graphics approach in which 3D geometry of the scene or object is known.

Image-based rendering focuses on the problem of synthesizing new views of a scene or objects from a collection of images [2].

Image-based modeling is an example of an IBR technique that reconstructs three-dimensional models from a collection of images [2].

Image-based rendering techniques are classified into three categories depending on how much geometry information is available [3]. The three categories are:

1. Rendering with no geometry
2. Rendering with implicit geometry
3. Rendering with explicit geometry

A common challenge when using IBR techniques is producing intermediate views from the discrete sample of images available, otherwise known as view synthesis. IBM and view interpolation are two different methods that produce intermediate views from a set of images. For the purposes of this paper, we call the set of images used for IBR techniques as multi-viewpoint image collection (MVIC).

1.1.2 3D Interaction Techniques

With the technology behind 3D user interfaces (3DUIs) fast maturing, previous technologies that were limited to laboratories and research facilities are starting to proliferate the consumer market for public use. This led to an increase in 3D applications that cater to a broader audience.

This advancement in 3D technology is an opportunity to develop and research on more effective and intuitive 3D user interfaces and 3D interaction techniques. 3D user interface is defined as a user interface that involves 3D interaction. 3D interaction on the other hand is defined as:

"Human-computer interaction in which the user’s tasks are performed directly in a 3D spatial context. Interactive systems that display 3D graphics do not necessarily involve 3D interaction; for example if a user tours a model of a building on her desktop computer by choosing viewpoints from a traditional menu, no 3D interaction has taken place. On the other hand, 3D interaction does not necessarily mean that 3D input devices are used; for example, in the same application, if the user clicks on a target object to navigate to that object, then the 2D mouse input has been directly translated into a 3D location, and thus 3D interaction has occurred [4]."

Input devices play an important role in designing 3D manipulation techniques. Two important characteristics of input devices that greatly affect manipulation tasks are (1) the number of control dimensions (number of degrees of freedom or DOF) the device has and (2) the integration of these control dimensions [4]. The
Figure 1.1: (a) the 2-DOF movement of the desktop mouse is depicted. In contrast, a popular gaming input device, the Wiimote, is shown in (b) has 6-DOF. It has translational components (X, Y, and Z) and also rotational components, which are roll, pitch, and yaw.

PlayStation Move and Wiimote (Figure 1.1(b)) are examples of input devices that provide 6-DOF for interacting with 3D user interfaces. However, many systems still use 2-DOF input devices. This makes designing interaction techniques that allow the user to control the 3D position and orientation of virtual objects using 2-DOF input devices an important consideration when designing 3DUIs [4]. An example of a commonly used interaction technique used for 2-DOF input devices is 3D widgets [7]. 3D widgets are 3D graphical tools that are directly placed inside the 3D scene (Figure 1.2(a)). These widgets control only a specific number of DOF. Other examples of 3D manipulation techniques for 2D input devices are the Virtual Sphere [8] and the ARCBALL [9] technique. Both techniques place the object to be rotated inside an imaginary sphere which can be freely rotated around its center point (Figure 1.2(b)).

Various fields are starting to incorporate 3D technology. Medical [12] and educational [13] fields are just some of the fields that are now incorporating 3D technology to support their work.

The proliferation of 3D technology in our modern society gives way to new problems and problem settings. These problems present opportunities for re-
searchers to develop 3DUIs and 3D manipulation techniques.

1.2. Problem Setting

Traditionally, in 3D computer graphics, scenes or objects are represented as either, purely 3D models or through objects using IBR techniques. We explore the idea of combining a 3D model of a scene with a variable object represented through a collection of images. This kind of approach for a system could be useful in object arranging applications similar to My.Ikea [14].

1.3. Motivation and Research Theme

In our research we implemented a system that combines a variable object represented through MVIC and a 3D scene. The current prototype is the main component to a conceptual application. This application is a hand-held system that allows a user to view a desired target object in a 3D model of a target scene. The conceptual application should have three main parts. First, the target scene
Figure 1.3: In Tawara and Ono’s work [12], they use Augmented Reality (AR) technology together with a Wiimote that allows a two-handed direct manipulation system that can achieve volume segmentation of CT/MRI data in real 3D space.

Figure 1.4: A screenshot of the Tangible Chemical Reactions project by Maier, Tonnis, and Klinker. By using AR technology, it allows user to dynamically visualize 3D models of molecules [13].

will be modeled using a modeling software that can be directly loaded on to the application. Second, the user will need the ordered image collection of the target object. The collection of images should be acquirable solely with the mobile device without the need for additional apparatus. For the system to be usable, we aim to keep the number of images in the MVIC to a minimum. Lastly, a user interface that allows the user to effectively interact with the 3D scene and target object. This is depicted in Figure 1.5.

In this thesis, we propose two 3D manipulation techniques. These manipulation techniques do not provide intermediate viewpoints to the user. However, these techniques allow the user to manipulate the target object at any orientation despite the limited viewpoints available. These two techniques are implemented in a user interface system that allows users to interact with the combined target
Figure 1.5: The user interface will combine both the reconstructed 3D model of the target scene and the image-based representation of the target object allowing the user to interact with these two components using a set of manipulation techniques.
object and 3D scene.

Without performing view synthesis, our work does not completely fall under the category of IBR techniques. We instead use MVIC and propose two manipulation techniques that address the issue of having limited viewpoints when using MVIC. We call our approach image-based representation (IBREP).

1.4. Thesis Overview

This thesis is organized as follows: Chapter 2 presents different research that are closely related to our work. Chapter 3 discusses the two proposed techniques and the implementation of the prototype of the system. To evaluate the proposed techniques, we conduct two user studies which we discuss in Chapter 4 and Chapter 5 respectively. Lastly, Chapter 6 provides a discussion on the results of both user studies and concludes this thesis.
Chapter 2

Related Literature

This chapter provides a brief discussion on research topics that are related to this thesis. We also discuss the feasibility of applying their methods in our system. We then discuss the difference between their approach and our approach.

2.1. Image-based Rendering and Manipulation Techniques

We focus our discussion on IBR techniques and their associated manipulation techniques. IBR techniques use MVIC, which limits the available viewpoints. Because of the limited viewpoints, normal 3D manipulation techniques cannot be used. We discuss how different research topics provided a solution for the limited viewpoint constrain. We then examine how their solution affects the 3D manipulation technique used for their system.

2.1.1 Rendering with Explicit Geometry and Implicit Geometry

Mentioned in Chapter 1, IBR can be classified into three categories. We first discuss research categorized under Rendering with explicit geometry and Rendering with implicit geometry. Rendering with explicit geometry requires encoding of 3D information along with the image data set. On the other hand rendering
with implicit geometry are techniques that compute the 3D information using projection calculations [3].

Pollefeys and Van Gool’s approach in generating intermediate views uses the IBM technique [15]. They use geometric data and MVIC for 3D reconstruction of a real world object or scene. Similar to [15], [16] uses IBM for view synthesis. In their work they connect several iPhones in a peer-to-peer network to take synchronized images of a real world object or scene. Each phone’s orientation is also recorded and sent to a server along with the MVIC. The server handles the 3D reconstruction based on the uploaded images. Both [15] and [16] rely on explicit geometric data for IBM.

Instead of 3D reconstruction, Scharstein [17] builds on Seitz and Dyer’s [18] approach to generate intermediate views. [17] and [18] use view morphing techniques to provide new viewpoints from the MVIC. Similarly, [2] uses view interpolation techniques to smoothly transition between images in the MVIC. [17], [18], and [2] are examples of IBR techniques that use implicit geometry to provide a
Figure 2.2: Phototourism: from a large input of images, either from personal photo collections or Internet photo-sharing sites (a), a sparse 3D model (b) is computed along with each photo’s viewpoint. Image (c) allows the user to interactively move about the 3D space through view interpolation between photographs [2].

solution for the limited viewpoint constraint.

[15], [16], [17], [18], and [2] solve the constraint of limited viewpoints when using IBR techniques. Their work provides intermediate viewpoints that are unavailable to the user. After generating intermediate views, they rely on conventional manipulation techniques to interact with the reconstructed 3D model or move between viewpoints. Their approach focuses on view synthesis and not the manipulation technique.

An example of a 3D manipulation technique that is commonly used is the ARCBALL technique [9]. As mentioned in Chapter 1, this technique maps input from a 2D device into 3D space. The ARCBALL technique uses quaternions
to avoid *hysteresis* which occurs when closed loops of 2D motion do not produce closed loops of rotation. Imagine rotating an object in the following order: 1. $x = +90^\circ$, 2. $y = +90^\circ$, 3. $x = -90^\circ$, 4. $y = -90^\circ$. This kind of closed loop rotation will not bring the rotated object back to its original orientation. Also, by using quaternions, the ARCBALL technique eliminates the gimbal lock problem.

The constraints imposed by our system design refrain us from using the mentioned techniques. First, in Patro *et. al*’s [16] work there is a need for additional hardware (multiple mobile phones) and an internet connection that will upload the image data set to a server for reconstruction. Without uploading the images to a server, 3D reconstruction is not feasible on a mobile device. The overall system design mentioned in Chapter 1 hinders us from using such an approach because the system will not be fully functional without the use of additional apparatus. View synthesis through view morphing or view interpolation on the other hand require heavy processing to efficiently calculate the geometric data necessary for such techniques to work. On current mobile devices this approach is not feasible because it is computationally expensive, which may lead to a delay in rendering the output based on user interaction leading to a loss in the overall 3D experience.

Figure 2.3: From two images of different viewpoints, view morphing produces the illusion of physically moving a camera [18].
2.1.2 Rendering with No Geometry

Another category that classifies IBR techniques is rendering without geometry. Apple’s QuickTime VR [19] is an example of an IBR technique that is rendered without geometry. QuickTime VR first requires the user to gather numerous snapshots of the target object at precise angle increments by mounting the camera on a special apparatus. Ideally, the angle increment in between each successive snapshot should be 10°. These images are then organized into a structured image collection. Instead of performing view synthesis, their work relies in the number of images taken. With images taken at low angle increments, the users are given the illusion that they are interacting with a 3D model. This approach is an example of IBREP. However, two problems arise when using their system: (1) to gather the images, a special apparatus is needed and (2) to produce a 3D experience the angle increments should be as close to the ideal value as possible. Since view synthesis is not performed, high angle increments would eliminate the illusion of interacting with a 3D model.

![Figure 2.4](image)

Figure 2.4: (a) shows the available image from the data set. Initially, the MVIC allows for rotation in the Pitch and Yaw axis. In [1], they allow rotations about the Roll axis by rotating images before they are presented on screen (b). This allows for rotations in an arbitrary axis.

Most of the techniques for 3D geometry based models are not applicable to systems that use IBREP for objects or scenes. For example, the ARCBALL
technique requires an object or a scene to be viewed at arbitrary viewpoints.

Aside from Yu et. al [1], few have explored manipulation techniques for IBREP objects. In their work, they provided a manipulation technique that allows the user to rotate about an arbitrary axis providing more viewpoints without view synthesis. Instead, the authors adjusted the orientation of the displayed image. In the case of Quicktime VR, rotations are only done in the $x$ and $y$-axis. By changing the orientation of the displayed image, Yu et. al was able to add $z$-axis rotations therefore increasing the number of viewpoints.

Our work is different since the manipulation techniques we propose will allow users manipulate the object at any orientation by taking advantage of the orientation relationship between the scene and the variable object. We also explore the possibility of combining the IBREP 3D object with an actual 3D scene.
Chapter 3

Methodology

In this chapter, we discuss the prototype system implementation. We first explain how we simulated some of the key elements of the conceptual program and how we implemented the prototype. Lastly, we focus our discussion on the different manipulation techniques that we implemented.

We give the following descriptions to key elements of the system.

**Target object** defines the variable object to be placed inside the *target scene*. This target object is represented as a collection of images.

**Target scene** pertains to the 3D model of the scene without the *target object*.

**Combined scene** is the resulting scene when the target object is placed inside the target scene.

### 3.1. Prototype Implementation

The system prototype for this research was implemented on an iPad2 tablet. It has a 9.7 inch (diagonal) multi-touch display with a 1024x768 pixel resolution at 132 pixels per inch (ppi). Its graphics chip allows for maximum texture sizes of 2048x2048. It has a 1GHz dual-core Apple A5 low power system-on-a-chip as a processor [20].
Figure 3.1: The key elements of the system are depicted above. The target object (a) uses IBREP. The target scene (b) meanwhile is an actual 3D model. When the target object is placed inside the target scene, they make up the combined scene (c).
For our purpose of building the user interface (Figure 3.2) and testing the proposed manipulation techniques, the target object and the target scene was represented and modeled respectively using different programs.

### 3.1.1 Target Object

The conceptual application design requires a MVCI of the target object. We simulated this design detail by using OpenGL to take snapshots of a 3D teapot. We took snapshots of the teapot after being rotated in specific angle increments around the Y-axis and Z-axis. This simulates the action of photographing the teapot while it is in its natural orientation.

In order to reduce the number of images, we used Yu et. al’s [1] approach mentioned in Section 2.1.2 and rotated some of the images 180° (Figure 3.3). The modelview matrix associated with each snapshot was also indexed. The snapshots were gathered and then converted into a spritesheet using Zwoptex [21]. The generated spritesheet, together with the text file containing the model view matrices were then loaded into the iPad 2.

**Modelview Matrix**

is an orthogonal matrix that represents the current transformations. It combines the viewing matrix which represents the camera transformations and modeling matrix which represents the scene transformations. The three matrix elements on the 4th column represent translation transformations. The 3 element sets, \((m_0, m_1, m_2)\): +X-axis, left vector, \((m_4, m_5, m_6)\):+Y-axis up vector, and \((m_8, m_9, m_{10})\):+Z axis, forward vector represent affine transforms such as rotations [22] (Figure 3.5).

### 3.1.2 Target Scene

We simulated an actual 3D scene by modeling the target scene using various 3D modeling software. We used the 3D authoring tools Google Sketchup together with Autodesk 3DS Max to construct a 3D scene. The constructed scene was then converted to a VRML file. We used a VRML loader to load the VRML models into the iPad 2.
Figure 3.2: The images show the overall prototype system and some of the interface elements available for the user.

Figure 3.3: Instead of loading these two images as two separate files, the image on the left could be rotated and used as a replacement for the image on the right.
Figure 3.4: An example spritesheet of a teapot with snapshots taken at 45 degree angle increments.

Figure 3.5: The modelview matrix is decomposed into its translational components and the affine transform components \((X,Y,Z)\) [22]. Also, the coordinate system followed by the modelview matrix is shown.
3.2. Manipulation Tools

We classify the manipulation tools into two categories: (1) proposed manipulation techniques and (2) conventional manipulation techniques.

**Proposed manipulation techniques**
are the two manipulation techniques that are the focus of this thesis.

**Conventional manipulation techniques**
are various manipulation techniques that are commonly found in 3D geometry based user interfaces.

3.2.1 Proposed Manipulation Techniques

One disadvantage of representing the target object through IBREP is the inherent lack of viewpoints available. When the target object is placed inside the target scene, this limits the number of orientations a user can use to view the combined scene. We propose the following 3D manipulation tools that will allow the user to manipulate the object at any orientation producing more orientations than previously possible.

When the target object is rotated, its orientation with respect to the target scene changes. However, if we cannot rotate the target object due to the lack of available viewpoints, the target scene can be rotated instead. The target scene is rotated with the same angle as the expected rotation angle of the target object, but in the opposite direction. This produces the same orientation as if the target object was the one rotated.

As an example, let Figure 3.7 be the available images of the target object. The initial orientation of the combined scene is shown in Figure 3.6(a). After the user executes a rotation command of $45^\circ$ swiping right, the supposed output orientation of the combined scene should result to Figure 3.6(b). However, as we can see from Figure 3.7 we do not have the associated image for this type of rotation. The manipulation technique will instead rotate the target scene $-45^\circ$ resulting in an output of Figure 3.6(c). This is the key concept followed for both proposed techniques.
Figure 3.6: Suppose our image collection consists of all the images in Figure 3.7. From the initial orientation (a), the user places a rotate command that typically would result in (b). However, we see from Figure 3.7 that we do not have the necessary target object image for this kind of rotation. Using the proposed technique tool, the target scene is adjusted to accommodate the same orientation despite the lacking image as shown in (c).

Proposed Manipulation Techniques ($R_{\text{input}}$)

receives rotation input $R_{\text{input}}$ and compares it with a threshold $T$ – dependent on the angle spacing for each successive image and which Proposed Technique – to determine whether the target object image will be changed or the target scene will be rotated with a rotation command $R_{\text{scene}}$.

Proposed Manipulation Technique 1 (PT1)

The threshold $T$ for PT1 is half of the angle spacing for each successive image: $T \leftarrow \text{anglespacing}/2$.

Figure 3.7: Four snapshots of an object taken around the Y-axis at 90° angles.
We examine the step-by-step process of PT1 and how $R_{input}$ is interpreted using two cases.\(^1\)

1. Receive the rotation input, $R_{input}$, and set up storage variables.

2. Evaluate $R_{input}$ and other variables to determine the course of action the program will perform (2 cases):

   case 1: If $R_{total} <= -T$,
   
   i. Change the target object image and adjust the target scene rotation.

   case 2: else if $R_{total} > -T$
   
   i. Adjust the target scene rotation.

We now examine these steps in detail.

1. **Receive the rotation input $R_{input}$ and set up convenience variables.**

   The main goal of this step is to set up the $R_{input}$ to load all the necessary values required for processing in step 2. First, we store the current rotation $R_{total}$ value of the target scene in a separate variable $R_{prev}$ which is the current rotation before processing. Once the previous value of $R_{total}$ is stored, we load the new value of $R_{total}$ for further examination.

   1: Set $R_{prev} \leftarrow R_{total}$
   2: Set $R_{total} \leftarrow R_{total} - R_{input}$

   These steps will prepare $R_{total}$ for processing.

2. **Evaluate $R_{input}$ and other variables to determine the course of action the program will perform (2 cases):**

   With the new $R_{total}$ ready, we examine which course of action to take.

   Case 1: $R_{total} <= -T$, Change target object image and adjust the target scene rotation. We first compute for another convenience variable, $R_{excess}$.

\(^1\)For brevity, the following literate code is only valid for positive $R_{input}$s and is simplified.
i. Set $R_{excess} \leftarrow R_{total} + T$.

$R_{excess}$ is used to determine the amount of rotation for the current orientation ($R_{tscene}$) as well as its effect on the accumulated rotation ($R_{total}$) of the target scene. To determine the amount of rotation to be applied on the current orientation we

ii. Set $R_{tscene} \leftarrow -R_{prev} + T + R_{excess}$.

We proceed by preparing the variables for the next set of rotation input from the user by

iii. Set $R_{total} \leftarrow T + R_{excess}$

Case 2: $R_{total} > -T$, Adjust the target scene rotation.

i. Set $R_{tscene} \leftarrow -R_{input}$

ii. and do not change the current displayed image of the target object.

Proposed Manipulation Technique 2 (PT2)

The threshold $T$ for PT2 is the angle spacing for each successive image:

$T \leftarrow \text{angle spacing}$.

There are 2 main cases for PT2 and an additional 2 subcases for the second case. These cases determine how $R_{input}$ is interpreted. These cases are stated as follows.\footnote{For brevity, the following literate code is only valid for positive $R_{input}$ and is simplified.}

Case 1: The current $R_{total}$, without the new rotation command, is equal to 0.

Case 2: The current $R_{total}$, without the new rotation command, is not equal to 0.

subcase 1: The modified $R_{total}$, with the new rotation command, is less than 0.

subcase 2: The modified $R_{total}$, with the new rotation command, is greater than or equal to 0.
We now discuss these cases in detail. First, we examine the current value of $R_{total}$.

**Case 1:** If the current $R_{total}$, without the new rotation command, is equal to 0.

1: Change the current image to the next image (Figure 3.14).

After changing the current image, we then rotate the scene and prepare the variables for the next rotation command by

2: Set $R_{tscene} \leftarrow T - R_{input}$ and

3: Set $R_{total} \leftarrow R_{tscene}$.

**Case 2:** Otherwise, if the current $R_{total}$, without the new rotation command, is not equal to 0. We then prepare the necessary convenience variables for further processing by setting $R_{prev} \leftarrow R_{total}$ and setting $R_{total} \leftarrow R_{total} - R_{input}$. This saves a copy of the current $R_{total}$ before further processing. We proceed by evaluating the two cases.

subcase 1: If the modified $R_{total}$, with the new rotation command, is less than 0. We set

i. $R_{total} \leftarrow 0$

We then perform a rotate command on the target scene by

ii. $R_{tscene} \leftarrow -R_{prev}$

subcase 2: Else if the modified $R_{total}$, with the new rotation command, is greater than or equal to 0. We simply perform a rotate command on the target scene by

i. $R_{tscene} \leftarrow -R_{input}$

For further explanations we provide a graphical step-by-step explanation through Figure 3.8 through Figure 3.12.

### 3.2.2 Conventional Manipulation Tools

For all interactions we follow the coordinate system shown in Figure 3.13.
Figure 3.8: Left: Starting orientation for PT1 rotation. Right: Starting orientation for PT2 rotation.

Figure 3.9: User input: $R_{input} = 10^\circ$; Left: Output orientation for PT1 rotation (Case 2). Right: Output orientation for PT2 rotation (Case 1).
Figure 3.10: **User input:** $R_{\text{input}} = 45^\circ$; **Left:** Output orientation for PT1 rotation (Case 1). **Right:** Output orientation for PT2 rotation (Case 2: Subcase 2).

Figure 3.11: **User input:** $R_{\text{input}} = 90^\circ$; **Left:** Output orientation for PT1 rotation (Case 2). **Right:** Output orientation for PT2 rotation (Case 2: Subcase 2).
Figure 3.12: **User input**: $R_{input} = 100^\circ$; **Left**: Output orientation for PT1 rotation (Case 2). **Right**: Output orientation for PT2 rotation (Case 2: Subcase 1).

**Combined Scene: Rotate**

With this tool, the user can rotate the whole scene. The scene can be rotated only in the $0^\circ$ to $-180^\circ$ around the X-axis and $360^\circ$ degrees around the Y-axis. The rotation of the whole scene is constrained by the available viewpoints of the MVIC of the target object.

**Combined Scene: Up and Down Rotate**

We implemented a constrained version of the Combined Scene: Rotate manipulation tool; allowing rotations about the X-axis only. When the intention of the user is to rotate about the X-axis only, this tool is particularly helpful since it eliminates the chance of accidentally rotating about the Y-axis.

Figure 3.13: Shows the coordinate system used for all interactions.
Combined Scene: Left and Right Rotate

Another constrained version of the Combined Scene: Rotate manipulation tool, the Combined Scene: Left and Right Rotate allows rotations about the Y-axis only. Again, by imposing this constraint, we eliminate the rotation about the X-axis when the user wants to rotate solely around the Y-axis.

Combined Scene: Arcball

For the earlier versions of the prototype, we built on the ARCBALL implementation to suit an IBREP object instead of a 3D model. The conventional ARCBALL implementation converts the 2D input (swipe or mouse movement) into quaternions. The quaternions are then changed into a rotation matrix. If a 3D object is used for our system the next step would be to apply this rotation matrix to the current modelview matrix to rotate it accordingly. However, in our case, the quaternion output might be a rotation that will lead to an orientation where our system does not have the corresponding viewpoint image.

We solved this problem by determining the closest available viewpoint to the intended viewpoint orientation. We accomplished this by adding 4 additional steps after the new rotation matrix is computed by the ARCBALL function. The 4 additional steps are discussed below.

1: First we set a temporary matrix ($R_{temp}$) as a storage variable for the product of the output of the ARCBALL function ($R_{ab}$) and the affine transformation component of the associated modelview matrix ($R_{curr}$) of the current displayed image.

Figure 3.14: Shows the next image of swipe to the right for PT1 and PT2
1: Set $\mathbf{R}_{\text{temp}} \leftarrow \mathbf{R}_{\text{curr}} \times \mathbf{R}_{\text{ab}}$

2: Next, we compare the intended viewpoint to the closest available viewpoint. We first provide a reference vector from which we will compare the available viewpoints. To accomplish this, we could transpose $\mathbf{R}_{\text{temp}}$ and multiply it with $-1$ then get the third column of $\mathbf{R}_{\text{temp}}^T$ and store it in a vector. Instead of transposing, we proceeded with storing the third row of $\mathbf{R}_{\text{temp}}$ and multiplying it with $-1$. We then store the value to a vector $\mathbf{V}_{\text{ref}}$. This vector $\mathbf{V}_{\text{ref}}$ points directly to the forward vector of the camera for the intended output viewpoint. This will be used to compare the intended output viewpoint with available viewpoints.

3: From the current image, we determine the next image to display. We have 8 possible images to choose from as the next image, shown in Figure 3.15. These 8 possible images are the 8 nearest neighbors of the current image. Each image has an indexed modelview matrix. For each of the 8 nearest neighbors, we access the modelview matrix and perform the same process in step 2. We store the resulting vectors in $\mathbf{V}_1, \mathbf{V}_2...\mathbf{V}_8$ for comparison in step 4.

4: We determine which of the 8 vectors from step 3 is closest to $\mathbf{V}_{\text{ref}}$ by getting the angle between each of the 8 vectors and $\mathbf{V}_{\text{ref}}$. The smallest angle represents the closest match to the output of the ARCBALL function. The associated image for that vector will be the next image displayed.

We encountered erratic behavior when using this extended ARCBALL function. This can be attributed to the approximations we needed to perform for the ARCBALL function to accommodate an IBREP object being rotated.

**Combined Scene: Zoom**

With this tool, the user can zoom in and out the combined scene. Using a pinch-in command, the user can have a wider view of the overall combined scene. With a pinch-out command, the user can zoom in and focus on some details of the combined scene.
Figure 3.15: From the middle image, the next image could be any of the 8 images around it. We refer to these images as the 8 nearest neighbors.

**Target Object: Rotate**

The current target object can be rotated only around its Y-axis. The possible rotations that a target object can have are the ones that will adhere to the natural constrains imposed by the scene. For example, a target object placed on top of a table should not be rotated around its X-axis since this will make the object go inside the table as shown in Figure 3.16.

**Target Object: Translate**

The target object can only move along the surface (i.e. on top of the table) on which it’s placed. This 2-DOF constrain is placed to reduce the cognitive load produced when the user has to worry that the object will be positioned in an unnatural state (i.e. floating objects).

**Target Scene: Rotate**

Using this tool, the entire target scene is rotated while the target object preserves its current orientation. This technique is limited only to rotations around the
target scene’s Y-axis. The reason behind this constraint is similar to Target Object: Rotation technique’s constraint.

This tool allows for intermediate rotations or finer adjustments in terms of the orientation of the target scene relative to the target object. This is possible because this is the only rotation command that is not dependent on the number of available viewpoints in the image collection set.
Chapter 4

User Study 1: Performance and Preference

We conducted two user studies to evaluate our proposed methods. In this chapter we will discuss the details of the first user study.

The first user study is comprised of two experiments. The first experiment evaluates the performance of the two proposed manipulation tools compared to the similar conventional manipulation tools namely (1) Target scene: Rotate and (2) Target object: Rotate. The second experiment is designed to determine the effect on performance when reducing the number of images in the MVIC. In both experiments, the participants asked to recreate a scene orientation under different conditions.

4.1. Current Prototype State

Table 4.1 shows the manipulation tools that have been implemented as part of the prototype.

In Section 3.2.2, we discussed our extended version of the ARCBALL function. We found the behavior of the extended ARCBALL function to be erratic. Hence, we decided not to include it in the user study. The Combined scene: Zoom function was not yet implemented when user study 1 was conducted.

The target scene at this juncture of the research consisted of a table and several cups modeled in 3D (Figure 4.1). We have yet to model a complete 3d
Table 4.1: The list of implemented manipulation techniques during the first user study. Manipulation tools marked with an asterisk (*) are not used in this user study.

<table>
<thead>
<tr>
<th>Combined Scene</th>
<th>Target Scene</th>
<th>Target Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotate</td>
<td>Rotate</td>
<td>Rotate</td>
</tr>
<tr>
<td>Arcball*</td>
<td>——</td>
<td>PT1</td>
</tr>
<tr>
<td>Up and Down Rotate*</td>
<td>——</td>
<td>PT2</td>
</tr>
<tr>
<td>Left and Right Rotate*</td>
<td>——</td>
<td>Translate</td>
</tr>
</tbody>
</table>

environment for the target scene at this point in the research.

4.2. Participants

There was a total of 24 participants (16 males and 8 females) for the first user study. Their ages ranged from 24 to 37 years old (average age of 26.75). All participants had prior experience in using a touch screen device. A total of 14 participants stated that they had prior experience of using 3D authoring applications such as Google SketchUp and/or AutoCAD. Four participants gave themselves a ranking of more than 5 (10 being extremely proficient) when using the 3D authoring applications.

4.3. Procedure

We began the user study by asking the participant to accomplish a demographics survey (Appendix A). The moderator then used a visual aid to give an overview of the user study and the system.

After the overview, the participant was given a hands-on tutorial on how to use the system using another prototype. Using a visual aid and instructions from the moderator, the participant is trained on all the manipulation tools and the necessary gestures needed to use them. The moderator was very careful in his explanations regarding the different functions of the manipulation tools. He did
not explain any technical details about any of the manipulation tools to allow
the participants to learn the manipulation tools on their own.

After the tutorial, the instructions for the entire session were discussed and
explained to the participants. The moderator stressed that he would not be
able to provide help for the participant once a specific task has started. He also
mentioned that additional discussions regarding the manipulation tools will not
be given. The moderator emphasized that the participant was not being tested
but rather the system to alleviate any pressure or uneasiness the participant could
be feeling. The participants were told that when the task became too difficult,
they had the liberty to stop at any time.

For all the tasks in both experiments, we showed the participant a goal oriene-
tation image. This is final orientation that the participant must recreate using
the given set of conditions. For each task, the moderator observed for certain
behaviors or comments from the participants.

After each task was performed for each condition, the participant filled out a questionnaire about their experience with the specific condition (Appendix A). They were also asked to list down any problems they encountered during the task. At the same time, the moderator recorded the task success and the time of completion for each task on the moderator’s package (Appendix A). After a task was completed in all available conditions, the participant was asked to fill out an additional questionnaire that would compare all the conditions for that specific task shown in Appendix A.

After both experiments were completed, the participants were asked to fill out a brief post-session questionnaire (Appendix A) that evaluated their entire experience in interacting with the prototype more specifically with the manipulation tools. After filling out the post-session questionnaire, the moderator asked the participants for any comments, suggestions or questions the participants might have regarding the system and the user study.

4.4. Experiment 1: Manipulation Tools Test

Experiment 1 focused on comparing the two proposed manipulation tools with a set of conventional tools when performing orientation tasks. Two tasks comprise Experiment 1.

4.4.1 Experiment 1 Conditions

In order to recreate the given orientation for each task, the participant was given a specific set of manipulation tools. Each set was a different condition for the task. For each task there was only one variable condition. The list of manipulation tools associated with each condition are listed in Table 4.2

4.4.2 Experiment 1 Design

We employed a within-subjects design for Experiment 1. There was a single independent variable (Condition A, B, and C) and two dependent variables (Task
Table 4.2: The list of available manipulation techniques for each tool set/condition in Experiment 1.

<table>
<thead>
<tr>
<th>Target Element</th>
<th>Conventional</th>
<th>PT1</th>
<th>PT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined Scene</td>
<td>Rotate</td>
<td>Rotate</td>
<td>Rotate</td>
</tr>
<tr>
<td>Target Scene</td>
<td>Rotate</td>
<td>——</td>
<td>——</td>
</tr>
<tr>
<td>Target Object</td>
<td>Translate</td>
<td>Translate</td>
<td>Translate</td>
</tr>
<tr>
<td></td>
<td>Rotate</td>
<td>PT1</td>
<td>PT2</td>
</tr>
</tbody>
</table>

success and Completion time\(^1\)). We recorded completion times based on the assumption that a good manipulation technique will allow the user to recreate an orientation as fast and precise as possible.

In order to reduce a learning effect when performing the tasks, the order of using each manipulation tool set to accomplish the task was completely balanced for all six possible orderings. Each participant was given a random order among the six possible orderings.

Task success was binary, either a pass or fail. To pass a task, the participant must successfully recreate the goal orientation using the given manipulation tool set. To fail a task, the participant must inform the moderator that the task cannot be completed. If a participant perceives that he/she has successfully completed a task, the moderator will evaluate the orientation accomplished by the participant on the mobile device. If the participant has indeed recreated the orientation of the entire scene, he/she will be given a binary 1 for task success, otherwise the moderator informs the participant that he/she has not yet recreated the entire scene and must continue performing the task.

Completion time was dependent on task success. If a participant fails to complete a task, their time was discarded. Only the time of successfully completed tasks were recorded. The questionnaires and answer sheets given to the participant are listed in Appendix A.

\(^1\)Throughout this paper, the reported completion times are normalized over a 60-second range; e.g., a completion time of 90 seconds is reported as 90/60 = 1.50.
4.4.3 Task 1: Orientation that Does Not Require Minor Table Rotations

The specific goal orientation (Figure 4.2(b)) for all conditions under task 1 did not require the participant to make small rotations on the table. We are evaluating the performance of the two proposed manipulation techniques when presented with a task whose goal does not require small table rotations.

Our hypotheses for Experiment 1: Task 1’s outcome were as follows:

- H1.1: Using the two proposed methods for the first time, participants will take a significantly longer time in completing the task as compared to the Standard Tool set.

- H1.2: The Conventional tool set will be preferred by the participants since they have a higher degree of familiarity with its tools.
Figure 4.3: (a) shows the starting orientation for all conditions in Experiment 1: Task 2. (b) is the orientation that the participants must reproduce using the specific set of manipulation tools available.

4.4.4 Task 2: Orientation that Requires Minor Table Rotations

Task 2’s goal orientation (Figure 4.3(b)) required the participant to make small adjustments to the table to bring out some details in the scene (Figure 4.4). This task was designed in order to evaluate the performance of the two proposed manipulation techniques compared to conventional techniques when performing a task that requires small table adjustments.

- H1.3: Gaining familiarity with the tools, there will not be a significant difference between all three conditions.

- H1.4: The Conventional tool set will still be preferred by the participants because of their moderate understanding or familiarity of the two proposed tool sets.
### Table 4.3: Independent variables for Experiment 2 tasks

<table>
<thead>
<tr>
<th>Condition</th>
<th>No. of Images</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
</tr>
<tr>
<td>4</td>
<td>16</td>
</tr>
</tbody>
</table>

#### 4.5. Experiment 2: Performance Under Different Settings

An important design goal of our system is maintaining functionality even without additional apparatus. Part of the process in using the overall system is image acquisition. Without additional apparatus, acquiring a large image data set is cumbersome. We designed Experiment 2 to determine the least number of images that will not affect user performance and preference.

Similar to Experiment 1, the participant is asked to recreate the orientation (Figure 4.5(b)) in the presented image using a single manipulation tool set. The manipulation tool set will be determined in accordance to the answer the participant provides at the end of Experiment 1: "Which of the two proposed manipulation techniques would you prefer to use?"

#### 4.5.1 Experiment 2 Conditions

The participant was given different prototypes for each condition. The independent variable for Experiment 2 (Table 4.3) was the number of images taken from different angles of rotation around the Y-axis. These are the conditions under which the participant will perform the orientation task.

#### 4.5.2 Experiment 2 Design

Experiment 2 follows the within-subjects design of Experiment 1. There was a single independent variable (Number of images: 4, 8, 12 and, 16) and two
Figure 4.4: The image shows the level of detail that must be produced by the participant. This level of detail can not be achieved by using Combined Scene: Rotate tool alone.

(a) Experiment 2: Task 1 starting orientation
(b) Experiment 2: Task 1 goal orientation

Figure 4.5: (a) shows the starting orientation for all conditions in Experiment 2: Task 1. (b) is the orientation that the participants must reproduce using the specific set of manipulation tools available.
dependence of variables (Task success and Completion time). The order in which all conditions were tested was completely balanced for all 24 possible orderings.

Similar to Experiment 1, task success was binary, either a pass or fail. To pass a task, the participant must successfully recreate the goal orientation using the given manipulation tool set. A task will be considered a failure if the participant tells the moderator to stop the current task. If a participant perceives that he has successfully completed a task, the moderator will evaluate the orientation accomplished by the participant on the mobile device. If the participant has recreated the orientation of the entire scene, he/she will be given a binary 1 for task success, otherwise the moderator informs the participant that he/she has not yet recreated the entire scene and must continue with the task.

Completion time was dependent on task success. If a participant fails to complete a task, their time was discarded. Only the time of successfully completed tasks were recorded. The various questionnaires and answer sheets are shown in Appendix B.

4.5.3 Task 1: Different Image Collection Sizes’ Effect on Performance

We are evaluating the impact on performance when we change the number of available images. Like the previous experiment, we also gathered preference data from the participants regarding which condition they preferred to use to accomplish orientation tasks.

Our hypothesis for Experiment 2’s outcome were as follows:

- H2.1: Participants will find the 4-image Condition difficult to use because of the inherent lack of available viewpoints.
- H2.2: Participants will find the 16-image Condition easiest to use because of more available viewpoints.

4.6. User Study 1 Results

We performed statistical analysis on the task success, completion time and the participants’ subjective rating of the difficulty of completing each task using the
given set of interaction controls. For the difficulty rating participants were asked to give feedback on a seven point Likert scale [23], with 7 being very easy and 1 being very hard.

### 4.6.1 Experiment 1: Task 1

#### Post Condition Results

Using a one-way ANOVA test, we found a significant difference in average task completion time across manipulation tool set conditions ($F(2,66) = 5.40$, $P = .0067$) and the difficulty rating ($F(2,66) = 9.143$, $P < 0.001$) (Table 4.4). We used the Bonferroni test for further post hoc analysis to identify significance between conditions. For the task completion time we found that PT2 and Conventional Tool sets were both significantly faster than the PT1 condition with $P$ values of $P = 0.012$ and $P = 0.024$ respectively. On the other hand, there was statistical difference between the PT1 and Conventional tool set completion times. Likewise, after each condition was performed the subjects reported their perceived difficulty of performing the task under that condition. Performing a similar statistical analysis on this data showed that PT1 tool set was significantly more difficult to use to complete the task compared to the Conventional tool set ($P = 0.00012$). There were no significant differences between other condition pairs.

Table 4.4: Average completion time for all conditions under Experiment 1: Task 1.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Avg. Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>1.16</td>
</tr>
<tr>
<td>PT1</td>
<td>2.33</td>
</tr>
<tr>
<td>PT2</td>
<td>1.24</td>
</tr>
</tbody>
</table>
Table 4.5: Average completion time for all conditions under Experiment 1: Task 2.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Avg. Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>1.31</td>
</tr>
<tr>
<td>PT1</td>
<td>1.73</td>
</tr>
<tr>
<td>PT2</td>
<td>3.04</td>
</tr>
</tbody>
</table>

Post Task Questionnaire

After performing all the conditions, the participants were asked to choose which of the three manipulation tool sets were the easiest and the hardest to work with. The results showed that 66.7% of the participants chose Conventional tool set as the easiest tool set to use (PT1 tool set: 12.5% and PT2 tool set: 20.8%). While 58.3% of the participants picked PT1 tool set as the hardest tool set to work with (Conventional tool set: 0% and PT2 tool set 41.7%).

4.6.2 Experiment 1: Task 2

Post Condition Results

There was a significant difference in average completion times depending on the manipulation tools used (one-way ANOVA $F(2,59) = 7.59, P < 0.05$) (Table 4.5). Using Bonferroni, we found that there was a significant difference in average completion time between the Conventional tool set and PT2 tool set ($P = 0.001$), and also between PT1 tool set and PT2 tool set ($P = 0.016$). Similarly there was also a significant difference in the average perceived difficulty of using the manipulation tools ($F(2,59) = 12.41, P < 0.001$). Using Bonferroni, we also found that the participants perceived using the Conventional tool set was significantly easier compared to both PT1 tool set ($P = 0.011$) and PT2 tool set ($P = 0.001$).
Table 4.6: Average completion time for all conditions under Experiment 2: Task 1.

<table>
<thead>
<tr>
<th>No. of Images</th>
<th>Avg. Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1.72</td>
</tr>
<tr>
<td>8</td>
<td>1.42</td>
</tr>
<tr>
<td>12</td>
<td>1.65</td>
</tr>
<tr>
<td>16</td>
<td>1.54</td>
</tr>
</tbody>
</table>

Post Task Questionnaire

Similar to Experiment 1: Task 1, the participants were asked to pick the easiest condition and the hardest condition. Approximately 70% of the participants picked the Conventional tool set as the easiest (PT1 tool set: 12.5% and PT2 tool set: 16.67%). While 75% of the participants picked PT2 tool set as the hardest condition (Conventional tool set: 8.3% and PT1 tool set: 16.67%).

4.6.3 Experiment 2: Task 1

Post Condition Results

Using one-way ANOVA to analyze the average task completion time for varying conditions (4, 8, 12, and 16 image conditions), we found that there was no significant difference ($F(2, 59) = 7.59, P < 0.05$)(Table 4.6) across all conditions. We also found that there was no significant difference in the perceived difficulty of the task when the number of images is varied $F(3, 111) = .379, P = .769$.

Post Task Questionnaire

At the end of Experiment 2: Task 1, participants were asked to rank (Hardest = 4 and Easiest = 1) the four conditions that they worked with (Table 4.7). Using a one-way ANOVA to analyze the difficulty ratings, we found that there was a significant difference ($F(3, 92) = 2.611, P = 0.0561$). Accumulating all the scores, Condition 1 totaled 73 points while the three other conditions had an almost equal difficulty rating (56, 55, and 56 respectively). The post-hoc analysis
Table 4.7: Average difficulty rankings for all conditions under Experiment 2: Task 1.

<table>
<thead>
<tr>
<th>No. of Images</th>
<th>Avg. Difficulty</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3.04</td>
</tr>
<tr>
<td>8</td>
<td>2.33</td>
</tr>
<tr>
<td>12</td>
<td>2.29</td>
</tr>
<tr>
<td>16</td>
<td>2.33</td>
</tr>
</tbody>
</table>

using Bonferroni showed that participants found that the 4-image condition was significantly harder than the rest. We did not find any significant difference in the difficulty rating across the other conditions.

4.7. User Study 1 Discussion

4.7.1 Experiment 1

Overall, the Conventional tool set was the preferred manipulation tool set by the participants in performing orientation tasks as can be inferred from the Post Task Questionnaire of Experiment 1. With the Conventional tool set, the interaction and system was straightforward enough for the participants, which confirms H1.2 and H1.4. These manipulation tools can be found in most 3D applications requiring very little learning to effectively use it. Despite this result, according to the interview conducted after the entire session, some participants suggested that after learning to use the two proposed manipulation techniques, they did not feel much difference between the three tool sets.

Experiment 1: Task 1

Contrary to our initial hypothesis for Experiment 1: Task1 (H1.1), PT1 tool set had a significant difference as compared to the Conventional and PT2 tool sets in average Completion time. We attribute this result to the implementation of PT1 manipulation tool under the PT1 tool set. When a rotate command is issued by the user, the target scene is adjusted in small increments first before the target
object and the target scene are rotated together after reaching a specific threshold as discussed in Section 3.2.1. To achieve the goal orientation, the participant had to perform small target scene rotations first in order to realize the goal orientation. On the other hand, the PT2 manipulation tool under PT2 tool set rotates the target scene and the target object first to the nearest viewpoint. The succeeding rotation commands will slowly adjust the target scene until it reaches a specified threshold as discussed in Section 3.2.1. Using the PT2 tool set, the participant did not need to adjust the table in small increments because this is already achieved in the first rotation command.

An explanation as to why there was no significant effect using the Conventional tool set and PT2 tool set on time completion could be the goal orientation of the task. The goal orientation did not require small rotation increments. Based on the explanation above, the participant did not need to go through the extra step of adjusting the Target Scene in small increments to achieve the goal orientation. Using the PT2 manipulation tool under PT2 tool set, the target object was rotated to the desired orientation in the first step alone.

Although there was no significant difference in task completion time between Conventional tool set and PT2 tool set, the participants did not fully understand how to use the proposed manipulation tool. This explains the significant difference found in the perceived difficulty rating for the Conventional tool set and PT2 tool set despite no difference in their average completion times.

**Experiment 1: Task 2**

Despite using the same manipulation tool sets for Task 1, the statistical analysis showed that the PT2 tool set was significantly slower in accomplishing the goal orientation compared to the other tool sets. The task at hand required the participant to make small target scene adjustments to achieve the goal orientation. Again, due to the implementation of PT2 manipulation tool under the PT2 tool set, the participants were confused ("unexpected flipping of the target scene when rotating") because the target scene and target object rotated to the available viewpoints first before the target scene could be adjusted in small increments therefore skipping the goal orientation. The moderator observed that after seeing the target scene and target object farther from the intended orientation, some
participants made a swipe in the opposite direction to correct their perceived mistake. For some participants, it took a while for them to realize that the next step in the PT2 manipulation tool was the adjustment of the target scene in small angle increments. For some participants, they did not figure this next step and gave up. Only 16 participants completed Task 2 under PT2 tool set.

On the other hand, using either the PT1 or Conventional tool set did not significantly affect the Completion times. This can be attributed to the implementation of PT1 manipulation tool as discussed in Section 4.7.1. Using PT1 manipulation tool, the participant was able to see quickly the needed action of small target scene increments since it is the first step in the implementation of the manipulation technique. This led to less confusion and faster completion times for the participants.

Similar to Task 1 of Experiment 1, the participants were observed to be in the process of learning the proposed tools. Although the task completion time difference was insignificant, this could explain the significant difference in the perceived difficulty between the PT1 tool set and the Conventional tool set.

4.7.2 Experiment 2

The inherent lack of viewpoints when using the 4-image Condition (Condition 1) for Experiment 2 was the major reason why most participants ranked it as the hardest condition to work with. This proves our initial hypothesis H2.1. However, there was no difference in the average difficulty rating with the rest of the conditions, which disproves our initial hypothesis H2.2. The reason for this is that the angle increments between Conditions 2 to 4 were not noticeable. However, for Condition 1, the participants noticed the angle increments. That could be the reason for the statistically higher difficulty rating for Condition 1.

Experiment 2: Task 1

By this time, some participants have developed a sequence of actions to accomplish an orientation task. This caused the majority of the participants to have an easier time in completing the goal orientation even with varying conditions imposed on the task. Through observation, those participants who found the task
in this experiment difficult have not yet fully developed their sequence of actions or have not yet fully understood the how the manipulation techniques work. This caused even more confusion as the each condition changed the number of images available for them.

4.8. Summary of User Study 1

Although the majority of the participants preferred to use the Conventional tool set, the performance of both PT1 and PT2 tool set, when performing tasks 1 and 2 respectively, show that the proposed manipulation techniques can perform as fast as the Conventional tool set. Also, some participants suggested that through learning, the difference between the conventional manipulation techniques and the two proposed manipulation techniques would be negligible in performing orientation tasks.

The results of Experiment 2 provide us with information that reducing the number of images needed to represent the target object will not affect the efficiency of the system to perform orientation tasks. By keeping the number of images needed to a minimum, we do not burden the users to gather large amounts of data before the system can be used therefore effectively increasing the usability of our prototype system.
Chapter 5

User Study 2: Learnability

The second user study aims to build on the results of the first user study. The collected data from the first user study suggests that the proposed manipulation techniques have no significant difference with the conventional tools in terms of performance. However, the preference data from the initial user study indicated that the participants preferred to use the conventional tools over the proposed manipulation tools. From the user comments of the first study, the participants suggested that the reason behind their low ratings of the proposed manipulation tools was the unfamiliarity with the proposed techniques. The moderator also observed that some participants were able to accomplish the goal orientation through trial and error. We conducted this study in order to measure how easily the users can learn the proposed manipulation technique. We also designed this user study in order to eliminate the possibility of success through trial and error.

We will continue the evaluation of Proposed Technique 1 and discontinue the evaluation of Proposed Technique 2. We base this decision on the less than satisfactory performance of Proposed Technique 2 in the initial user study.

5.1. Current Prototype State

The first user study helped us understand what improvements can be made for the prototype. We studied the user comments and the observed behavior of the participants from the first user study to further improve the program. The main improvement of the prototype since the first user study was the representation
of the target scene. In the first user study, we modeled only a table with some cups to represent the target scene Figure 3.1b. For the second user study we modeled a more elaborate target scene (Figure 5.1). We used Google SketchUp and Autodesk 3DS Max to model it. From 3DS Max, we exported the 3D model into a WRL format, the format that can be loaded into our system.

Aside from this, we also improved the layout and fixed problem areas in the user interface. The available manipulation techniques at this juncture of the research are listed in Table 5.1.

5.2. Participants

For the second user study, we had a total of 12 participants (6 male, 6 female). An important design parameter for this user study is to gather participants that were also part of the 1st user study. This is to ensure that all participants have the same experience of using the system. These participants were chosen at random from the 24 participants of the initial user study.
Table 5.1: The list of implemented manipulation techniques during the second user study. Manipulation tools marked with an asterisk (*) are not used in this user study.

<table>
<thead>
<tr>
<th>Combined Scene</th>
<th>Target Scene</th>
<th>Target Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rotate</td>
<td>Rotate</td>
<td>Rotate</td>
</tr>
<tr>
<td>Arcball*</td>
<td>——</td>
<td>PT1</td>
</tr>
<tr>
<td>Up and Down Rotate*</td>
<td>——</td>
<td>PT2*</td>
</tr>
<tr>
<td>Up and Down Rotate*</td>
<td>——</td>
<td>Translate</td>
</tr>
<tr>
<td>Zoom*</td>
<td>——</td>
<td>——</td>
</tr>
</tbody>
</table>

5.3. Procedure

User Study 2 spans two days for each participant. During the first day, the moderator discussed the current version of the prototype with the participant. After the discussion, the user was given 10 minutes to explore and familiarize themselves with the new prototype and the various manipulation tools that will be used for the user study. We employ this process to ensure that all participants have the same exposure in using the prototype and the manipulation techniques.

On the second day, the procedure and the instructions for the entire experiment were explained to the participant. Questions related to the experiment procedure were answered by the moderator. Questions related to the manipulation techniques however were not answered by the moderator. After all the necessary instructions were discussed, the participant was given the first written test under a specified condition. All participants answered the tests for both conditions. In order to minimize learning, the conditions were balanced for all participants.

The participant was asked to answer the first test in the first set without interacting with the system. The test contains 8 numbers. Each number has two images. Figure 5.2a is the starting orientation and Figure 5.2b is the goal orientation. The participant was given a list of manipulation tools and their associated gestures depending on the current condition. The participant chose which manipulation tool and associated gesture will change the starting orientation to
Figure 5.2: An example test question for User Study 2. Figure 5.2a shows the starting orientation and Figure 5.2b is the goal orientation. The correct answer for this question would be: “Object Translate: Left”.

the goal orientation. All goal orientations presented in the experiment can be achieved by using just a single manipulation technique and gesture. Also, part of each question is a confidence value. The participants were asked to provide a confidence rating for each of their answers. The answer sheet given to the participant is shown in image Figure 5.3.
Figure 5.3: The answer sheet given to the participant for the Conventional tool set exams.
If the participant answers a question incorrectly, the participant will be asked to review their mistake using the prototype. The participant will be given 2 minutes to interact with the prototype. After which, the participant will be asked to answer another test. All the questions in the succeeding test are similar in nature to the first test. The process will continue until the participant answers all 8 questions correctly. Once the participant answers all questions correctly for a particular test in the first condition, the participant will be asked to answer the Post Condition Questionnaire for their feedback on the first test.

After finishing the first condition, the participant will be given the next test set corresponding to second condition. The same process as the first test set applies. Once the participant completes the second test set, another Post Condition Questionnaire will be given. After answering the Post Condition Questionnaire, the participant will be interviewed by the moderator for any additional comments.

During the entire test, the moderator observes the behavior of participants that may seem relevant to the goal of this user study.

5.4. Experiment 1: Proposed Manipulation Technique Learnability

This experiment focuses on the learnability of the Proposed Manipulation Technique 1 compared to a similar conventional technique. The whole user study will have one task performed for two conditions.

Learnability measures how much time is needed for a user to become proficient in using a product [23].

The task requires the participant to answer all questions in a test correctly. If the participant makes a mistake, the participant will be asked to answer another test for that given condition. The participant will be asked to do this until all questions are answered correctly. We record how many tests for a specific condition the participant needs to take before getting all answers correctly. We associated the number of tests taken with learnability because if a participant fully grasps the manipulation techniques then the participant can use these techniques correctly to achieve a goal orientation.
From the first task discussed in Section 4.4, some users arrived at the goal orientation through trial and error. Because of this, we were not able to gauge whether the participants really understood the proposed manipulation technique. We address that issue and focus on how well the participant understands the manipulation techniques by asking the participant to repeat the test even with a single mistake.

Our hypotheses for this experiment are as follows.

- **H3.1**: Participants will repeat the test the same number of times with the Conventional tool set and the PT1 tool set. This is because of the increasing familiarity with the PT1 manipulation technique.

- **H3.2**: Participants will be more confident answering with the Conventional tool set as compared to the PT1 tool set because of the amount of familiarity they have with the Conventional tool set.

- **H3.3**: There would be a significant difference between the difficulty ratings across conditions because the level of proficiency the participants have with PT1 is not yet comparable to the conventional manipulation tools.

### 5.4.1 Experiment 1 Conditions

As mentioned in Section 5.3, to answer the questions in the test the participant has to pick a specific manipulation tool and an associated gesture from a list to answer each question. The list depends on the current condition the participant is using. The answer list for each condition is shown in Table 5.2 and Table 5.3.

### 5.4.2 Experiment 1 Design

We employed a within-subjects design for this experiment. There was a single independent variable (Condition A and B) and a single dependent variable (Test Repetition).

In order to reduce a learning effect when performing the tasks, the order of each test set was completely balanced for both orderings. A total of 6 participants started with Condition A while the remaining 6 participants started with Condition B.
Table 5.2: Conventional tool set: The complete list of the available manipulation tools with the gestures associated for each tool.

<table>
<thead>
<tr>
<th>Conventional tool set</th>
<th>Manipulation tools</th>
<th>Gestures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scene Rotation</td>
<td>up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
<tr>
<td></td>
<td>Object Translation</td>
<td>up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
<tr>
<td></td>
<td>Object Rotation</td>
<td>——</td>
</tr>
<tr>
<td></td>
<td></td>
<td>——</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
<tr>
<td></td>
<td>Table Rotation</td>
<td>——</td>
</tr>
<tr>
<td></td>
<td></td>
<td>——</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
</tbody>
</table>

Table 5.3: PT1 tool set: The complete list of the available manipulation tools with the gestures associated for each tool.

<table>
<thead>
<tr>
<th>PT1 tool set</th>
<th>Manipulation tools</th>
<th>Gestures</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Scene Rotation</td>
<td>up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
<tr>
<td></td>
<td>Object Translation</td>
<td>up</td>
</tr>
<tr>
<td></td>
<td></td>
<td>down</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
<tr>
<td></td>
<td>PT1</td>
<td>——</td>
</tr>
<tr>
<td></td>
<td></td>
<td>——</td>
</tr>
<tr>
<td></td>
<td></td>
<td>left</td>
</tr>
<tr>
<td></td>
<td></td>
<td>right</td>
</tr>
</tbody>
</table>
5.5. User Study 2 Results

We performed statistical analysis on the number of tests taken, confidence values for each question in the first test and the participants' subjective rating of the difficulty level of answering each test using the given set of manipulation tools. We used a Likert scale for the difficulty rating where participants were asked to provide a rating of 5 (very easy and) to 1 (very difficult).

5.5.1 Experiment 1

Post Condition Results

Using a t-test [23] on the repetition values, we found no significant difference in the average number of tests taken of the participants across manipulation tool set conditions. Using a paired t-test we found \( t(11) = -0.24, p = 0.81 \) (Table 5.4). To further support the null hypothesis, we computed for the t-tabulated value (2.26) proving no significant difference across conditions \((t - \text{computed} : 0.8118 < t - \text{tabulated} : 2.200985)\).

We also performed statistical analysis on the confidence values provided by the participants for the first tests of each condition. By performing the statistical analysis on these confidence values, we gain understanding on how confident the users were with their understanding of the manipulation tools in both conditions. Performing a t-test on the reported confidence values, we find that the participants felt as confident with their answers in PT1 tool set as with their answer in the Conventional tool set \( t(11) = 0.57, p = 0.58 \) (Table 5.5). Performing the t-tabulated test shows that the null hypothesis is acceptable \((t - \text{computed} : 0.5836 < t - \text{tabulated} : 2.200985)\).

Post Task Questionnaire

However, when we performed the one-tailed t-test on the difficulty rating of each condition, we found that it was approaching significance, \( t(11) = 1.90 \) and \( p = 0.08 \). For further analysis we performed a two-tailed t-test that showed a significant difference in the average degree of difficulty reported by the subjects between conditions, \( t(11) = 2.1, p < 0.05 \) (Table 5.6). However, we also received
Table 5.4: Average number of tests taken by participants for each condition in User Study 2.

<table>
<thead>
<tr>
<th>Tool set</th>
<th>Avg. Exams Taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>2.58</td>
</tr>
<tr>
<td>PT1</td>
<td>2.75</td>
</tr>
</tbody>
</table>

Table 5.5: Average confidence rating of participants when answering the first test for each condition under User Study 2.

<table>
<thead>
<tr>
<th>Tool set</th>
<th>Avg. Confidence Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>4.08</td>
</tr>
<tr>
<td>PT1</td>
<td>3.95</td>
</tr>
</tbody>
</table>

Table 5.6: Average difficulty rankings both conditions under User Study 2.

<table>
<thead>
<tr>
<th>Tool set</th>
<th>Difficulty Avg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Conventional</td>
<td>3.67</td>
</tr>
<tr>
<td>PT1</td>
<td>3.13</td>
</tr>
</tbody>
</table>
mixed reactions when it comes to the conventional manipulation tool set with some participants saying that "This tool set (conventional) was harder because it is confusing." While some participants reported that they did not see any significant difference in the difficulty of using both tool sets.

5.6. User Study 2 Discussion

We hypothesized (H3.1) that the participants will take the test the same number of times with the Conventional tool set as with the PT1 tool set. This is because of their repeated exposure to the PT1 tool set makes them more familiar with the usage of PT1 manipulation tools. This is confirmed by the statistical analysis we performed on the gathered data for this user study. A reason behind this may be because, unlike the first user study, the participants were given the chance to use the prototype before the test to examine how the PT1 manipulation tool and other tools work. This gives the participant time to familiarize and understand how PT1 works.

On the other hand, since this is a relatively new manipulation technique compared to the conventional tools, we assumed that the participants will still be uneasy when being tested on their understanding of this tool (H3.2). However, from the statistical data, we see that the participants’ confidence in using both condition tool sets does not have any significance. A reason could be that the participants’ exposure to the prototype and the manipulation tools gave them an added confidence in using all these manipulation tools. Although some participants repeated the test several times, providing a high rating in their confidence with their answer meant that their understanding of the manipulation technique might be invalid.

Lastly, we expected the participants to have a moderately harder time in using the PT1 tool (H3.3) set because it required them to imagine a new manipulation technique’s function without visual help. Also, we assumed that the level of proficiency the participants have with using the PT1 manipulation is not yet comparable to the conventional manipulation tools. The statistical analysis performed on the difficulty ratings reported by the participants confirms our initial hypothesis.
5.7. Summary of User Study 2

Based on the first user study, we wanted to determine how easily the proposed manipulation technique 1 (PT1) can be learned by users. This allows us to gain a better understanding as to why the users preferred the Conventional tool set over the PT1 tool set in the first user study. Based on this second user study, we can conclude that PT1 manipulation tool set can be learned as easily as the conventional tool set. As the proficiency of the users with the PT1 manipulation tool set increases, the users experience lesser difficulty in completing tasks using PT1. Also, with this increasing proficiency with PT1, the users feel as confident using this tool compared to the conventional tools. For our study, the level of proficiency of the participants was not yet satisfactory leading to a significant difference in the reported difficulty between conditions. In addition, the PT1 manipulation was used in an environment setting that did not allow the participant to visually translate their thoughts about PT1 in to actions leading to a slightly more difficult task.
Chapter 6

Conclusion and Discussion

In our work, we presented a unique problem setting of placing a target object represented as a collection of images into an actual 3D model of a target scene. By representing this target object as a collection of images, we are limited to the number of available viewpoints the combined scene can be viewed in. Conventionally, research based on this technique focused on providing new views based on the MVIC. Our approach is different because we do not propose a technique that will present new views to the user, but a technique that will allow the user to correctly manipulate the target object at any orientation despite the lack of viewpoints.

Following the design of a conceptual application, we developed a prototype that will allow us to test the proposed manipulation tool. This prototype along with the various manipulation techniques implemented on it will serve as the basic building block for the conceptual application and could also be used in other applications.

We conducted two user studies that examined the two proposed manipulation techniques. From the data in these two user studies, we gained valuable information that allowed us to improve our prototype. In the succeeding sections of this chapter, we give a brief summary of our findings from the two user studies we conducted. Lastly, we conclude by providing some insight on possible improvements on the manipulation tools and the prototype as well.
6.1. Proposed Manipulation Techniques’ Performance

We conducted experiments that focused on PT1 and PT2’s performance in orientation tasks. We provide a basis for these findings by comparing both techniques with a conventional technique. Statistical analysis on the data we gathered from the experiments showed that there was no significant difference between the conventional technique and PT1 in terms of task completion times. Comparing both PT1 and PT2’s performance, we found that PT1 performed better. Based on this result, we decided to focus our study on PT1.

Aside from this, we also examined the performance of PT1 and PT2 when the number of images to represent the target object was changed. An important design parameter for the conceptual application is to represent the target object with the least amount of images possible without sacrificing functionality. Based on our experiments, there is no difference when representing the target object with 16 images or 8 images taken around a Y-axis rotation. By reducing the number of images in half, the system becomes more usable by not requiring the users to gather a large amount of images.

6.2. Users’ Perception of Proposed Manipulation Techniques

The first user study showed us that the users prefer the conventional manipulation techniques compared to PT1 and PT2. This can be attributed to the fact that it was the first time for the participants to use these techniques. Generally, there is a need to learn manipulation techniques first before a user feels comfortable using it. Users are more inclined to use techniques that they are familiar with as compared to other techniques. This is especially true when the manipulation technique is compared to frequently used techniques. The participants’ unfamiliarity with PT1 and PT2 lowered their proficiency. The lack of proficiency could be a source of frustration that led to the low ratings of PT1 and most especially PT2. We examined next the learnability of PT1. The significance of the reported difficulty
rating in User Study 2 shows us that although the participants are becoming more proficient with using PT1, their level of proficiency with the PT1 manipulation tool is not yet comparable to the conventional manipulation tools.

6.3. Learnability of Proposed Manipulation Techniques

In order to gauge how fast the user becomes proficient with PT1, we conducted another user study that focuses on how easily the proposed manipulation technique can be learned. To provide a baseline for the experiment, we performed the same experiment on the Conventional tools. The results showed that there was no significant difference between PT1 and the conventional manipulation tools in terms of how easily users learned to use it. This result shows that even though the participants initially preferred the conventional manipulation techniques, becoming easily proficient with PT1 manipulation techniques would increase users’ inclination to use it.

6.4. Conclusion

We present a manipulation technique that allows a user to manipulate an object placed inside a 3D model of scene at any orientation despite the limited viewpoints associated with IBR techniques. From the two user studies we conducted, we were able to draw a conclusion that PT1 is comparable in performance with the conventional manipulation techniques. In our second user study, our data suggested that PT1 can be easily learned by users thus allowing the users to become proficient with PT1 easily. This increase in proficiency will increase the preference of users to use PT1.

This manipulation technique was implemented on a prototype that would serve as a building block for a more complete system such as the conceptual application mentioned in this paper. To realize this conceptual application, we need to improve the user interface of our system. One improvement could be providing visual guides to the users that will assist them in using the prototype
and its tools. Another issue to address is how to make smoother transitions in between images without performing view synthesis. A possible approach for this issue is to change the orientation of the current displayed image in such a way that it provides an illusion that the object is being rotated.

Aside from these possible improvements, there is also a need to conduct more user studies. All the participants in these user studies had prior experience with touch screen interfaces. Also, most of the participants have had prior experience working with 3D authoring tools. Another user study should be conducted with the participants having very little experience with touch screen interfaces and 3D authoring tools. This will allow us to improve the user interface and the manipulation technique so that it will cater to a broader audience. Also, the two user studies used a computer generated teapot. Future user studies could be conducted to include different objects other than a teapot and also real world objects. This would provide a deeper understanding on how different types of objects affect the performance of the manipulation techniques.

Once a stable and usable prototype is created, we can then build on an image-acquisition system for the target object and also a 3D modeler for the target scene. These two new programs should then be seamlessly integrated with the prototype system. The resulting application would allow the users to visualize real world objects inside a 3D scene. This kind of application is useful especially when designing and visualizing a 3D space with new objects.
Acknowledgements

First and foremost, I would like to extend my sincere gratitude to my professor and thesis adviser, Professor Hirokazu Kato. He has welcomed me in the Interactive Media Design Laboratory and helped me achieve my dream. With his guidance and unwavering patience, I was able to complete my thesis. Through his example and constant encouragement, I was able to grow as a researcher and as a person. Without him, my thesis would not have been a success.

I would also like to extend my deepest gratitude to Professor Tsukasa Ogasawara, Assistant Professor Sei Ikeda, and Assistant Professor Yuki Uranishi. Their insightful comments and suggestions allowed me to further improve my research. With their help, I was able to improve the quality of my work.

I would like to also thank Professor Mark Billinghurst. I was lucky enough to have met him in his visit here in NAIST. Despite a difference in time zones, he would always find time to accommodate me. Without his generosity, I would not have been able to produce quality research.

Thank you also to the other members of the Interactive Media Laboratory. First, their insightful comments and advice made a big difference in my work. Secondly, their companionship throughout the three years helped create a healthy working environment.

To my friends here in NAIST, they are my second family. Being away from home and faced with so many challenges in a foreign country, they were there to support me. The joy of having met such great people in my short stay here is such a blessing. This journey would not have been the same without these people.

I would like to thank my family back in the Philippines. You have been very understanding and supportive of my endeavors. Even though I am miles away, I
could feel the love and support that you have for me. To my parents who have made all of this possible, I can’t thank you enough.

To the Panasonic Scholarship, thank you for giving me a great opportunity. You have given me a chance to pursue my dreams. Without your support and guidance I would not be able pursue my dream of quality education, thank you.

I believe that I would not have achieved anything in my life without You.
References


Appendix

A. User Study 1

In this section, we present the various materials we used for the user study. In order to conduct the user study smoothly we prepared two packages both for the participant and the moderator. Before the actual experiments, we asked the participants to fill out a demographic survey shown in Figure 6.1

A.1 Experiment 1

All tasks under Experiment 1 were performed in all three conditions shown in Table 4.2. After performing a task under the Conventional tool set, the participants were asked to fill up Figure 6.2. On the other hand, after performing a task under any of the Proposed tool set conditions, the participants were asked to fill up Figure 6.3. Experiment 1 comprised of two tasks. After task1 was performed under all conditions the participants were asked to fill up Figure 6.4. After task 2 was performed under all conditions, the participants were asked to fill up Figure 6.5.

A.2 Experiment 2

Experiment 2 under User study 2 had four conditions. After performing a task under each condition, the participants were asked to fill up Figure 6.6. After the task had been completed under all four conditions, the participant was then asked to fill up Figure 6.7. Lastly, to conclude the user study, participants were asked to fill out Figure 6.8.
B. User Study 2

For the second user study, there were two conditions. For each condition, the participant was handed a specific answer sheet (Figure 6.9. An answer sheet for the Conventional tool set condition only contained conventional tools as choices for the answer. This is the same for the PT1 condition. After answering all questions correctly in a test for a specific condition, the participant was given Figure 6.10.

Demographics Survey

Thank you for agreeing to participate in this user study.

Please answer this questionnaire until you encounter a command “Stop”. The moderator will ask you to continue afterwards.

Participant#______
Male□ Female □
Age:___________

Your answers to the following questions will help us study the test results better.

1. Have you used a 3D visualization software before? (Autocad, Google Sketchup etc)
   Yes □ No □
   ***If “Yes”, answer question numbers 2 & 3, If “No”, proceed to question 4.

2. How often do you use it? (Please circle one.)
   Always    Sometimes    Rarely

3. On a scale of 1 to 10 (10 being the highest), how familiar are you with such applications? ______

4. Have you used a touch screen device?
   Yes □ No □
   ***If “Yes”, answer question 5. If “No”, then you are done!

5. How often do you use it? (Circle one)
   Always    Sometimes    Rarely

Figure 6.1: Before the start of the actual experiments, participants are asked to fill up a demographics survey.
Post-task 1 questionnaire: (Conventional tool set)

1. Overall, this task was:

<table>
<thead>
<tr>
<th>Difficult</th>
<th>Very</th>
<th>Neutral</th>
<th>Very</th>
<th>Neutral</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. I could easily rotate the teapot to match the intended orientation:

<table>
<thead>
<tr>
<th>Difficult</th>
<th>Very</th>
<th>Neutral</th>
<th>Very</th>
<th>Neutral</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. The interaction control produced my intended outcome

<table>
<thead>
<tr>
<th>Difficult</th>
<th>Very</th>
<th>Neutral</th>
<th>Very</th>
<th>Neutral</th>
<th>Very</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Problems arising from using the Object: Rotate tool:

Figure 6.2: After performing a task under the Conventional tool set, a participant is asked to provide feedback regarding the task.
**Post-task 1 questionnaire: (PT1 and PT2 tool set)**

1. Overall, this task was:

<table>
<thead>
<tr>
<th>Very difficult</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Very easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. I could easily rotate the teapot to match the intended orientation:

<table>
<thead>
<tr>
<th>Very difficult</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Very easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. The interaction control produced my intended outcome

<table>
<thead>
<tr>
<th>Very difficult</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Very easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Problems arising from using the Object: Rotate + table adjustment 1 tool:

5. Did you feel that you were controlling the teapot with the Object: Rotate + table adjustment 1 tool?

Figure 6.3: After performing a task under any of the Proposed tool set, a participant is asked to provide feedback regarding the task.
Post-experiment 1 questionnaire

1. Which interaction tool set gave you the easiest time to accomplish the task?
   - First set [ ] Second set [ ] Third set [ ]

2. Why did you find this technique easier as compared to the others?

3. Which interaction tool set gave you the hardest time to accomplish the task?
   - First set [ ] Second set [ ] Third set [ ]

4. Why did you find this technique harder as compared to the others?

Figure 6.4: After completing Experiment 1: Task 1 under all conditions, participants are asked to fill up Figure 6.4.
Post-experiment 1 Task 2 questionnaire

1. Which interaction tool set gave you the easiest time to accomplish the task?
   - First set □  Second set □  Third set □

2. Why did you find this technique easier as compared to the others?

3. Which interaction tool set gave you the hardest time to accomplish the task?
   - First set □  Second set □  Third set □

4. Why did you find this technique harder as compared to the others?

5. Choose which technique you’d rather use:
   - Second set □  Third set □

Figure 6.5: After completing Experiment 1: Task 2 under all conditions, participants are asked to fill up Figure 6.5.
Post-task 3 questionnaire: __ images

1. Overall, this task was:

<table>
<thead>
<tr>
<th>Very difficult</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Very easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

2. I could easily rotate the teapot to match the intended orientation:

<table>
<thead>
<tr>
<th>Very difficult</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Very easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3. The interaction control produced my intended outcome

<table>
<thead>
<tr>
<th>Very difficult</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Very easy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Neutral</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Problems arising from using the Object: Rotate + table adjustment 1 tool:

5. Problems arising from this given condition? (# of images)

Figure 6.6: After completing Experiment 2: Task 1 under each condition, participants are asked to fill up Figure 6.6.
**Post-task 3 questionnaire (Condition 1)**

1. Rank the sets in terms of how hard it was to use them? *(1 easiest to 4 hardest)*
   
   ____ First set
   ____ Second set
   ____ Third set
   ____ Fourth set

2. Why did you find that set as the easiest as compared to the other technique?

3. What problems did you encounter with the hardest set?

4. Did the number of images affect your performance? If yes, how?

Figure 6.7: After completing Experiment 2: Task 1 under all conditions, participants are asked to fill up Figure 6.7.
Figure 6.8: After completing Experiment 2: Task 1 under all conditions, participants are asked to fill up Figure 6.8.
<table>
<thead>
<tr>
<th>Answer</th>
<th>Guess</th>
<th>Neutral</th>
<th>Sure</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>1  2</td>
<td>1  4</td>
<td>1  5</td>
<td></td>
</tr>
</tbody>
</table>

**CHOICES FOR MANIPULATION TECHNIQUES AND GESTURES**

<table>
<thead>
<tr>
<th>a</th>
<th>Scene Rotation</th>
<th>up</th>
<th>i</th>
<th>Object Rotation</th>
<th>left</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>Scene Rotation</td>
<td>down</td>
<td>j</td>
<td>Object Rotation</td>
<td>right</td>
</tr>
<tr>
<td>c</td>
<td>Scene Rotation</td>
<td>left</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>Scene Rotation</td>
<td>right</td>
<td>b</td>
<td>Table Rotations</td>
<td>left</td>
</tr>
<tr>
<td>e</td>
<td>Object Translation</td>
<td>up</td>
<td>i</td>
<td>Table Rotation</td>
<td>right</td>
</tr>
<tr>
<td>f</td>
<td>Object Translation</td>
<td>down</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>g</td>
<td>Object Translation</td>
<td>left</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td>Object Translation</td>
<td>right</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 6.9: Each participant will be handed one of these two answer sheets depending on the current condition under which the task is being performed.
Figure 6.10: The participants will be asked to fill up this form to provide feedback on the previously completed task.

| 1. How hard was it for you to understand the manipulation tools for this condition? |
| --- | --- | --- | --- |

2. Which manipulation tools were hard to understand? (You can list more than 3)

3. Among them, which is the hardest manipulation tool to understand? Why? (e.g., Translate tool - unresponsive or not intuitive)

4. What suggestions do you have to make this tool easier to understand?