A Review on Object Manipulation Techniques in Virtual Environments and their Performance Metrics

Jaemoon Jung, Hanjun Park, Dongwook Hwang, Minseok Son, Donghyun Beck, Jungmin Park, and Woojin Park

Department of Industrial Engineering, Seoul National University, Seoul, 151-744

ABSTRACT

Objective: The aim of this paper is to classify and characterize the existing interaction methods regarding object manipulation in virtual reality environments. Background: Over the past few decades, the field of virtual reality has gone through rapid advances and various virtual object manipulation techniques have been developed. However, an established classification of the existing object manipulation techniques does not seem available at this point. A systematic review of the advances during the last three decades and a classification of the existing object manipulation techniques based on it would help researchers prospect the future of human-machine interaction in virtual environments and further formulate future research directions for virtual object manipulation. Method: We collected and examined studies that are related to virtual reality and virtual object manipulation. The collected studies were reviewed in chronological order, and virtual object manipulation techniques were identified. Then, the existing virtual object manipulation techniques were sorted into categories according to similarities and each category was characterized. Additionally, this study reviewed the existing performance metrics for evaluating virtual object manipulation techniques. Application: This paper will aid in ameliorating the understanding of the existing virtual object manipulation techniques, and furthermore, offer directions to the future researchers in this field.

Keywords: Virtual Reality, 3D Interaction, Object Manipulation, Human-computer Interaction, Performance Metrics

1. Introduction

For the last few decades, the rapid expansion of high-end computer technologies and development of 3D-graphic displays have allowed the virtual environments (VEs) interfaces to become feasible in many different fields (Göbel, 1996). The term VEs refers to the environment in which interactive and virtual image displays are enhanced by special processing and by nonvisual display modalities, such as auditory and haptic, to convince users that they are immersed in a synthetic space (Ellis, 1994). In other words, users can use their eyes, ear, and hands just like they do in the real world, performing basic humans-environments reactions. Among such basic interactions, object manipulation is particularly important, because it is the most fundamental interactions between humans and their environment (Poupyrev, et al., 1999), with a potential to allow users to perform high-level tasks and activities.

Despite its deliberate acceptance and interest in the fields of research, there seem to have been very few studies that attempted to perform a systematic review and classification of virtual object manipulation techniques. Also, the number of classification scheme itself is very scant at this point. A study conducted by Poupyrev, et al. (1999) categorized the existing virtual interaction techniques by their interaction metaphors in a well-defined, systematic procedure.

An absence of understanding in the existing virtual objects manipulations techniques and lack of their classification may hinder researchers in developing enhanced and creative technologies. Thus, the goals of this study are as follows:
1) Conduct a survey of the existing virtual object manipulation techniques and related research studies.
2) Identify existing classification schemes of virtual object manipulation techniques.
3) Develop an up-to-date classification of virtual object manipulation techniques.

2. Literature Review

2.1 Background

In the early days of virtual reality, dating back as far as 1960s, two-dimensional (2D) input systems were mainly used for research, because 3D input systems were either too expensive, or simply did not exist at all at that time. Therefore, physical input devices, like mice and joysticks, or rudimentary 3D input methods using light rays or strings were integrated into a simulated 3D input system. Such systems limited users’ abilities in VEs to simple manipulations, like select, rotate, and move. Some of the examples of such 2D to 3D conversion systems are ray-casting (Roth, 1982), SPIDAR (Hirata, et al., 1992), and Bat (Ware, 1990).

Since then, advances in technologies related to virtual reality, such as head-mounted displays (HMDs) and motion capture cameras, expanded the horizons and allowed much more complex systems to be utilized in researches regarding virtual object manipulation. Rather than using physical devices as pointers or as “self” in VE, gestures and motions became the standard input methods, thus making the manipulation techniques much more immersive. For example, the Go-Go (Poupyrev, et al., 1996) and its derivatives, CAVE (Audio Visual Experience Automatic Virtual Environment) (Cruz-Neira, et al., 1992), ISAAC (Immersive Simulation Animation And Construction) (Mine, 1995), World-In-Miniature (WIM) (Stoakley, et al., 1995), and HOMER (Hand-centered Object Manipulation Extending Ray-casting) (Bowman, et al., 1997) are products of virtual object manipulation techniques with 3D input.

A considerable amount of research has been done on virtual object manipulation techniques, and in the following section, we categorize the existing virtual object manipulation techniques according to a novel classification scheme.

2.2 Classification of Virtual Object Manipulation Techniques

There have been a few surveys done on the existing virtual object manipulation techniques. Bowman, et al. (1997) has classified methods for grabbing and manipulating remote objects into two categories, “arm-extension” and “ray-casting”. This study defined arm-extension as those techniques that use virtual arms which can grow to the lengths desired by the users, so that the users can manipulate the objects out of reach. Ray-casting, on the other hand, makes use of a virtual light ray to grab or manipulate objects, where the ray’s direction is specified by the users. Poupyrev, et al. (1999) developed taxonomy of the virtual object manipulation techniques and empirically evaluated them. This study classified the techniques as either “egocentric” or “exocentric”, as shown in Figure 1, depending on the position of the user experiencing the virtual reality. Here, exocentric viewpoint is when the user is interacting with VEs from the outside, and egocentric viewpoint is when the user is interacting inside the VEs. Hand (1997) examined some of the object manipulation techniques in many different perspectives, such as type of input devices, feedback methods, viewpoint manipulation, and frames of reference.

![Taxonomy of virtual object manipulation techniques](image-url)
In this study, we used a classification scheme derived from the previous egocentric/exocentric method, which Poupyrev, et al. (1999) and a number of other authors have used, as it seems to be the most widely referred scheme as of today. In addition to that, we added new sub-categories to more clearly define the subtle differences between the virtual object manipulation techniques. Thus, we classified the techniques into 2 main categories, “Immersive” and “Non-immersive”, then into 3 sub-categories, “Direct Manipulation”, “Physical Control”, and “Virtual Control”. Table 1 shows the techniques and the related literature classified into these categories. For those techniques without particular names, the names of the devices used are listed instead.

### 2.3 Performance Metrics

The expansion of the number of virtual objects manipulation techniques has naturally raised a question on how to evaluate performances of those techniques.

<table>
<thead>
<tr>
<th>Main Category</th>
<th>Sub-category</th>
<th>Techniques or Devices Used</th>
<th>Literature (Author, Year)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Immersive</strong></td>
<td>Direct Mani...</td>
<td>HMDs only</td>
<td>(Kallmann, et al., 2001), (Lee, et al., 2002), (Kaiser, et al., 2003)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HOMER</td>
<td>(Bowman, et al., 1997)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Tilt, electric compass, and flex sensors</td>
<td>(Lee, et al., 2005)</td>
</tr>
<tr>
<td>Physical Control</td>
<td>Mice or joysticks</td>
<td></td>
<td>(Ware, 1990), (Mizell, 2000)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Aperture</td>
<td>(Forsberg, et al., 1996)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Interactive stereoscopic computer graphic workspace</td>
<td>(Schmandt, 1983)</td>
</tr>
<tr>
<td>Virtual Control</td>
<td>Go-Go</td>
<td></td>
<td>(Poupyrev, et al., 1998), (Poupyrev, et al., 1996), (Bowman, et al., 1997)</td>
</tr>
<tr>
<td></td>
<td>ISAAC</td>
<td></td>
<td>(Mine, 1995)</td>
</tr>
<tr>
<td></td>
<td>Virtual hand</td>
<td></td>
<td>(Tomozoe, et al., 2004), (Kiyokawa, et al., 2005), (Duval, et al., 2006)</td>
</tr>
<tr>
<td></td>
<td>Voodoo Dolls</td>
<td></td>
<td>(Pierce, et al., 1999), (Pierce, et al., 2002)</td>
</tr>
<tr>
<td><strong>Non-immersive</strong></td>
<td>Direct Mani...</td>
<td>Scaled-world grab</td>
<td>(Mine, et al., 1997)</td>
</tr>
<tr>
<td>Physical Control</td>
<td>World-In-Miniature</td>
<td></td>
<td>(Stoakley, et al., 1995)</td>
</tr>
<tr>
<td></td>
<td>Reconfigurable Tangible Device</td>
<td></td>
<td>(Aguerreche, et al., 2010)</td>
</tr>
<tr>
<td>Virtual Control</td>
<td>CAVE</td>
<td></td>
<td>(Mizell, et al., 2000), (Cabral, et al., 2005)</td>
</tr>
<tr>
<td></td>
<td>Virtual control panels</td>
<td></td>
<td>(Su, et al., 1993)</td>
</tr>
</tbody>
</table>

In this study, we used a classification scheme derived from the previous egocentric/exocentric method, which Poupyrev, et al. (1999) and a number of other authors have used, as it seems to be the most widely referred scheme as of today. In addition to that, we added new sub-categories to more clearly define the subtle differences between the virtual object manipulation techniques. Thus, we classified the techniques into 2 main categories, “Immersive” and “Non-immersive”, then into 3 sub-categories, “Direct Manipulation”, “Physical Control”, and “Virtual Control”. Table 1 shows the techniques and the related literature classified into these categories. For those techniques without particular names, the names of the devices used are listed instead.

### 2.3 Performance Metrics

The expansion of the number of virtual objects manipulation techniques has naturally raised a question on how to evaluate performances of those techniques.

---

**Figure 2.** Taxonomy of selection/manipulation techniques (Bowman, et al., 1999)
Quantifying the performance of virtual interaction techniques, however, is not an easy task since performance is not well defined (Bowman, et al., 1999). Aware of this fact, Bowman, et al. (1999) defined taxonomy of selection/manipulation techniques in order to quantify performance measures (Figure 2).

The performance metrics defined by Bowman, et al. (1999) includes a very broad definition of performance and most precedent evaluation studies fall into one of these performance variables. Zhai and Milgram (1993) evaluated human performance of manipulation schemes in virtual environments. As for the performance criteria, they defined three taxonomic elements for manipulating 3D objects: mapping relationship, sensing mode, and degree of integration. These three elements measure changes in position and time as subject moves an object from one position to other. The subject’s performance is then determined by how precise and fast certain object manipulation task was performed based on task completion time and position placement. Similarly, Poupyrev, et al. (1993) evaluated the performance of ray-casting and Go-Go techniques through empirical experiments. Completion time, number of iterations it took to complete positioning, and “net” manipulation time were set as performance criteria.

The choice of interaction techniques for selection, manipulation, and release were considered to be important factors that affects human performance during virtual object manipulation. However, more recently, studies tend to focus more on the issue of lag, the delay between input action and output response, and use it as the determinant of human performance (Liang, et al., 1991; MacKenzie, 1993; Watson, et al., 1997; Ware and Balakrishnan, 2000; Ha and Woo, 2010). It is obvious that the presence of lag in the environment causes more time to complete virtual object manipulation tasks, and consequently, increases the rate of error (MacKenzie, 1993). In order to measure the lag in the environment, Mackenzie used a revised version of Fitt’s Law (Figure 3).

\[ M' = a + b + c \times LAG \times ID \]

ID: index of difficulty; 
A: amplitude or distance to the target; 
W: width or size of the target 
\( a, b, c \): constant values
\( c \) : the weighting for the lag \( \times ID \)

**Figure 3.** Revised Fitt’s Law (MacKenzie, 1993)

In this revised law, index of difficulty (ID) is measured to determine the difficulty of a task based on two factors. If the distance between the subject and the target object is longer (higher \( A \)), or the target width is narrower (lower \( W \)), the ID of the task is higher (Ha and Woo, 2010).

### 3. Discussion

It is evident that VEs are becoming more immersive. As we can see in Table 1, the techniques are increasingly leaning towards direct manipulation and virtual controls in immersive environment. This is natural as in order for the techniques to maximize their interactivity, sensory vividness, and the overall human performance, high level of immersion is necessary. (Biocca, et al., 1995) In the past, due to the complex hardware systems required to emulate such highly immersive VEs, extremely high cost and low diffusion rate have been the biggest challenges that people faced in conducting research. Nevertheless, rapid advances in technology have made it possible to set up VEs easily at an affordable cost. So the next generation of challenges we face related to virtual object manipulation will be how to find the optimal techniques for it. Numerous virtual object manipulation methods have been proposed, yet not much evaluation has been done regarding their efficiency, accuracy, intuitiveness, learnability, physical or mental stress and many other human factors related issues. Therefore, future studies will have to pay more attention on the ways to measure performance of the techniques, and the criteria behind them.

### References


