Multimodal Interaction Metaphors for Manipulation of Distant Objects in Immersive Virtual Environments

Frank Steinicke
WWU Münster
Institut für Informatik
Einsteinstraße 62
48149 Münster, Germany
fsteini@uni-muenster.de

Timo Ropinski
WWU Münster
Institut für Informatik
Einsteinstraße 62
48149 Münster, Germany
timo@math.uni-muenster.de

Klaus Hinrichs
WWU Münster
Institut für Informatik
Einsteinstraße 62
48149 Münster, Germany
khh@uni-muenster.de

ABSTRACT

In this paper we discuss direct interaction metaphors for selection and manipulation of distant objects in immersive virtual environments and we propose extensions of the improved virtual pointer (IVP) metaphor. In particular, we describe how the process of object selection with the IVP metaphor can be enhanced by modifying the distance calculation used to determine the closest object to be selected. Furthermore we introduce direct 6 DOF manipulations of virtual objects using the IVP metaphor. We demonstrate how the task of object selection can be improved by combining existing interaction metaphors with multimodal feedback.

Keywords

Virtual Reality, Interaction Techniques, Multimodal Interaction Metaphors

1. INTRODUCTION

Virtual environments (VEs) have shown considerable potential as an intuitive and natural form of human-computer interfaces. Many scientific application areas benefit from virtual reality (VR). To improve the acceptance of VR technologies, the most basic interaction techniques need to be optimized to enable efficient human-computer interaction (HCI).

In this paper *direct interaction* metaphors for selection and manipulation of both local as well as distant objects in VEs are discussed and evaluated. In direct interactions the user directly manipulates objects with the input device, whereas in indirect interactions the user performs changes indirectly using menus, icons or widgets.

Before directly interacting with virtual objects, the user needs to specify the target for the interaction from the set of *selectable* objects. After *selecting* a virtual object, the user may manipulate any of the object's attributes, e.g., change the color or add a texture. In this paper we focus on six *degree of freedom* (DOF) manipulations, i.e., changing position and orientation of the virtual object.

Many VR application areas have shown that virtual pointer metaphors are natural and require less effort for both local and distant direct object interaction [Bow97a]. Even though virtual pointer metaphors can be used intuitively, their way of aiming at virtual objects and performing 6 DOF manipulations needs to be improved.

In order to achieve these goals we have proposed the *improved virtual pointer* (IVP) metaphor [Ste04a], which avoids most of the aforementioned disadvantages of current interaction metaphors. Our approach allows the user to select a desired object without requiring an exact hit. A straight ray is used to indicate the direction of the virtual pointer, while an additionally visualized bendable ray points to the closest selectable object (see Figure 1).

The closest selectable object which would be chosen if the user would perform a selection (e.g., by press-

WSCG 2005 SHORT papers proceedings,
ISBN 80-903100-9-5
WSCG'2005, January 31-February 4, 2005
Plzen, Czech Republic.

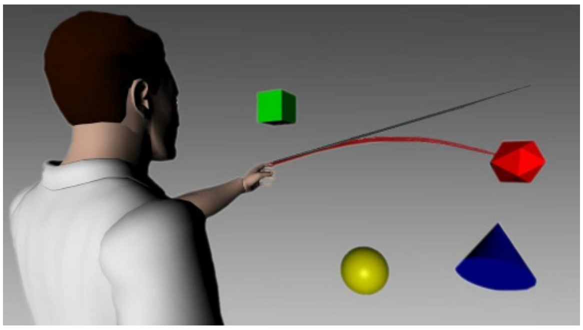

Figure 1: Illustration of the IVP metaphor.

ing a button or by pinching a glove) is called *active object*. After selecting the active object, successive manipulations can be accomplished.

In this paper we extend the previously introduced IVP metaphor by specifying a modified mechanism to determine the closest selectable object and proposing concepts how to use the IVP metaphor to perform manipulations of distant objects. Furthermore we give examples how the interaction process can be extended by giving adequate multimodal feedback.

2. IVP METAPHOR EXTENSIONS

In this section we present extensions of the IVP metaphor concerning selection and manipulation of virtual objects.

Selection of Virtual Objects

With the IVP metaphor a user performs a direct object selection by roughly pointing at the desired object. Thereupon the flexible ray bends to this object if it is the one closest to the straight ray. Determination of this active object is a major task of the IVP metaphor and is achieved by computing the distances of all selectable objects to the virtual ray. The scene graph structure used in most computer graphics systems can be exploited to enhance the performance by calculating all distances during a pre-evaluation phase of the scene graph. Thus interactive frame rates are maintained when using the IVP metaphor. The results of the distance calculations are stored in an ordered list, called the *ActiveObjectList*. This list contains all selectable objects and their distances to the virtual ray; the entries are sorted with respect to increasing distance, the first object with minimal orthogonal distance is the active object. This list provides the possibility to switch between active objects. Thus difficulties occurring during the selection of partially or fully occluded objects can be solved.

2.1.1 Distance Calculation

The distance between the virtual ray and a selectable object may refer to different *reference points* of a virtual object, e.g., the center of its bounding box or the closest edge or vertex. To compute the world coordinate distance between the reference point of a selectable object and the ray, we consider the line perpendicular to the ray which connects the reference point and the ray. Sorting the objects on the basis of this world distance within the ActiveObjectList may result in disadvantages when using perspective projection since the displayed distance may be distorted.

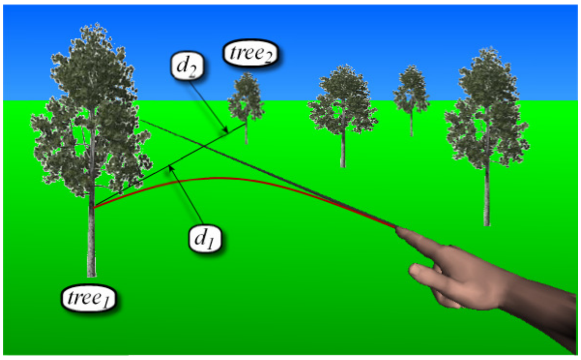

Figure 2: Problem of distant object selection caused by perspective distortion.

Figure 2 illustrates this problem. From the user's point of view, $tree_2$ seems to be closer to the ray. However, $tree_1$ attracts the curve and gets active since the distance d_1 between $tree_1$ and the virtual ray is less than the distance d_2 between the ray and $tree_2$, even though d_1 seems to be larger because of the perspective distortion. To prevent this drawback, we introduce two different approaches.

2.1.2 Image Plane Approach

An obvious approach is to evaluate the distances used for sorting the objects in the ActiveObjectList in image space coordinates. Therefore each world space distance vector d_i is transformed into the corresponding image space distance vector id_i as illustrated in Figure 3.

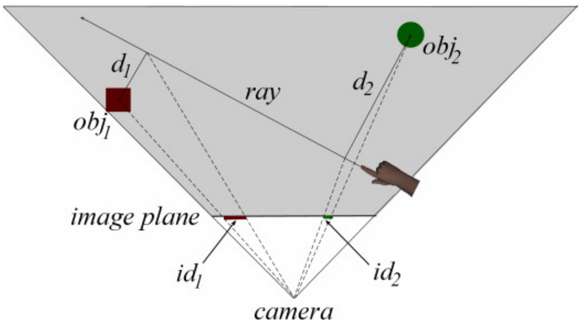

Figure 3: Projection of world space distance vectors onto the image plane.

The resulting lengths of the image space distance vectors are compared, and the objects are sorted accordingly into the ActiveObjectList. A world space distance vector appears shortened after the projection, although it may be quite long in world coordinates. Figure 3 illustrates this characteristic of projective geometry. Although the object obj_2 is located farther from the virtual ray than obj_1, its projection id_2 is clearly shorter than id_1. Thus sorting the ActiveObjectList based on image space distance vectors may lead to inconvenient results.

2.1.3 Distance Scaling Approach

A better approach considers both the world space distance between a particular object and the ray as well as the distance between the virtual input device and the object.

The world space distance is multiplied by a factor s, and the resulting value is used to sort the objects stored in the ActiveObjectList. The factor s is the inverse of the length of the vector from the position of the virtual input device to the position on the ray of the orthogonally mapped reference point of the considered object. Hence, with increasing distance between the virtual input device and a virtual object the value of s decreases, and multiplying the world space distance with this factor yields a smaller value as basis for sorting the ActiveObjectList. Analogously a decreasing distance between input device and object leads to a larger s and therefore a larger value is used for the comparison.

Although both approaches use different values for the sorting process, access to the distances in world space needs to be guaranteed, e.g., to measure distances between virtual objects.

Manipulation of Virtual Objects

For the manipulation of virtual objects we extend the idea of the HOMER technique described in [Bow97a] to provide an intuitive and natural alternative for 6 DOF direct manipulations. In contrast to the HOMER technique, the virtual input device and the selected object both remain at their initial positions after a selection is performed. Afterwards rotations and translations are accomplished as described in the following subsections.

2.1.4 Rotation

Although the virtual input device and the desired object both remain at their initial position, all rotations are implemented by a one-to-one mapping between the rotational movements of the virtual input device and the object. First the manipulated object is translated to the origin, then the rotational components of the transformation matrix of the virtual input device are copied to the corresponding components of the transformation matrix of the object. After finishing all rotational manipulations the object is moved back to its initial position or to a modified position according to the translation carried out during the manipulation process. Thus rotations are applicable in a natural and intuitive way like real world rotations, except that the manipulated object remains at its original position without being relocated into the user's hand. By using this approach manipulations can be accomplished accurately without occluding the desired object by the virtual input device, as it may happen when using the HOMER technique.

2.1.5 Translation

During manipulation of distant virtual objects, perspective distortion may cause a user to perceive the translation of such an object as distorted when compared to the movement of the virtual input device. Therefore, for the translation of a distant virtual object we scale its translational movement by a linear mapping function.

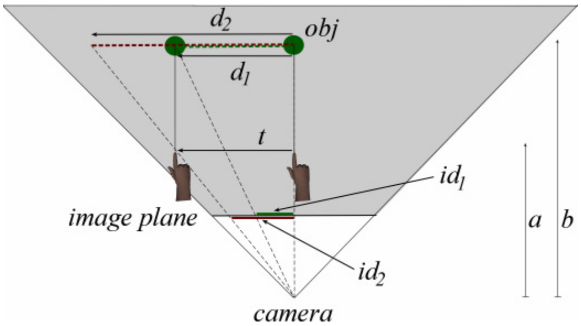

Figure 4: Distant object translation affected by perspective distortion.

Figure 4 clarifies this issue. After the user selects the object *obj*, a translation vector t of the virtual input device is mapped to this object. The projection vector id_1 of the one-to-one mapped translation d_1 of this translation vector t appears curtailed on the image plane (compared to id_2, which is the projection of d_2). Users anticipate that a translation vector t of the virtual input device would result in the scaled translation vector d_2 corresponding to id_2 on the image plane. To obtain the longer vector d_2, all translational movements are scaled with the factor $f = b / a$ where a is the distance between the camera and the virtual input device and b is the distance between the camera and the selected object (see Figure 4). The translation vector t is scaled by the value f and applied to the selected object. Indeed, small and accurate translation of distant objects is complicated by this approach, but existing VR applications have revealed that precise manipulations are accomplished primarily by local interaction within the immediate reach of the user. Direct 6 DOF translations of distant objects are mostly used for moving objects close to the user for exploration or for performing larger translations.

3. MULTIMODAL INTERACTIONS

As described in [Rai99a], multimodal interactions have the potential to enhance HCI and support the user during the manipulation process [Ric94a]. In this section we will describe the adaptation of multimodal interaction concepts to fit the needs of object selection and manipulation in immersive VEs.

Multimodal Input

To improve object selection two-handed interaction can be used. In our responsive workbench environment we use a pinch glove in combination with a haptic input device, which is shown in Figure 5. Position and orientation of both devices are tracked using an optical tracking system. The haptic input device is used to control the virtual input device with the IVP metaphor, simply by pointing at the desired object as described above. The pinch glove is used by the non-dominant hand for accessing menus or to assist the user when performing an object selection, e.g., selection of occluded objects by *tabbing* through the ActiveObjectList.

Multimodal Output

We use multimodal feedback to inform the user about a possible selection, i.e., visual, acoustic and tactile senses are addressed.

To get an adequate visual feedback of a possible selection, we visualize the ray direction vector as well as a beziér curve graph. This curve bends to the active object as described in [Ste04a] and illustrated in Figure 1.

To improve the user's perception of the active object's position, we inform the user acoustically when the active object is changed. By moving the virtual input device a different selectable object gets active, i.e., the beziér curve bends to the new active object, the position and orientation of both the active object and the user are used as parameters for the sonification process. Thus a change of the active object can be emphasized by a gentle sound dispersing from the position of the active object towards the user's position. As a result, the active object can be spatially located more easily.

In addition to the visual and acoustic cues, we added haptic feedback, i.e., the user gets haptic information regarding the active object. During a change of the active object, the user receives a light and short vibration signal emitted by the haptic input device (see Figure 5). Since the signal is emitted using a BlueTooth® connection, no cables constrain the interaction process. The level of vibration can be altered depending on the distance between the virtual input device and the active object. Starting from an initial minimal level of vibration, a decreasing distance between input device and active object results in a higher level of vibration.

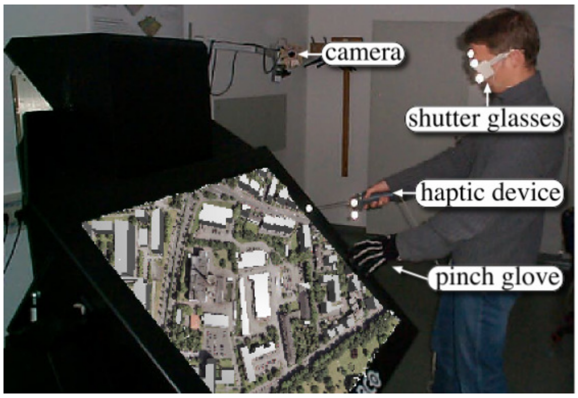

Figure 5. Workbench environment with an optical tracking system.

4. CONCLUSION AND FUTURE WORK

For evaluation we have compared the IVP metaphor to some of the techniques described in [Bow97a]. Our tests have shown that object selection can be accomplished faster with the IVP metaphor, especially when using the scaled distance approach. In a survey users have evaluated the use of multimodal feedback as very helpful, especially the haptic feedback given when a new object becomes active. The questionnaires as well as the detailed results of the user study are available upon request.

To further improve usability of direct interaction metaphors in general appropriate combinations of constraints restricting the available DOF may be useful. Therefore we will examine the possibilities of controlling 3D widgets with the IVP metaphor.

5. REFERENCES

[Bow97a] Bowman, D., and Hodges, L. An evaluation of techniques for grabbing and manipulating remote objects in immersive virtual environments. In Symposium on Interactive 3D Graphics, pp. 35-38, 1997.

[Rai99a] Raisamo, R. Multimodal human-computer interaction: A constructive and empirical study. PhD thesis, University of Tampere, 1999.

[Ric94a] Richard, P., Burdea, G., Gomez, D., and Coiffet, P. A comparison of haptic, visual and auditive force feedback for deformable virtual objects. In Proceedings of the Conference on Artifical Reality and Teleexistence (ICAT94), pp. 49–62, 1994.

[Ste04a] Steinicke, F., Ropinski, T., and Hinrichs, K. Object selection with an improved virtual pointer metaphor. To appear in Proceedings of International Conference on Computer Vision and Graphics (ICCVG), 2004.