CHAPTER VII

CONCLUSIONS AND FUTURE WORK

In this research, we have obtained a large body of results pertaining to the performance of interaction techniques for universal tasks in immersive virtual environments. These results are useful in choosing the appropriate techniques for VE applications, given their interaction requirements. We have also produced several new techniques, using our methodology, that provide more options to VE developers. This work also resulted in a VE application that we have shown to be both useful and usable for environmental designers.

Beyond these immediate results, however, our research has also produced some more abstract and high-level improvements in our understanding of the design and evaluation of VE interaction. Here, we will briefly discuss several of these important contributions.

7.1 VE Interaction Guidelines

In practical reality, few application developers are likely to take the time and effort required to quantify the interaction requirements of their systems, compare these to the results of testbed evaluation, and choose a set of ITs in a systematic fashion. One solution to this would be to create an interactive system that would accept a set of application requirements and automatically suggest possible ITs that match those requirements (discussed in the section on future work, below).

However, there is a well-established tradition in the HCI community of publishing sets of *guidelines* for user interfaces, interaction techniques, and the like. Guidelines are principled, practical aids that help a designer create interaction that is usable and performs well. Guidelines for VE interaction are not new (e.g. Kaur, 1999), but most sets of guidelines have two drawbacks. First, they are too general and subject to interpretation. They do not reduce the space of possible techniques far enough to allow the developer to make an informed decision. Second, guidelines have been simply adapted from 2D HCI guidelines, or they come from experience and intuition only. This does not ensure that the guidelines will be sound or that their use will produce well-performing systems.

Therefore, the VE community needs a set of interaction guidelines that are specific and practical, and which come directly from evaluation of techniques in the laboratory and in deployed systems (hypothesis 2). Our experiments and usability studies are a valuable source of such guidelines, and we present some of them here. Although all of these guidelines can be found elsewhere in the text, it is useful to view them together here.

7.1.1 Generic VE Interaction Guidelines

Do not assume that techniques based on a natural, real-world metaphor will be the most intuitive or that they will have the best performance.

Our initial user study on selection and manipulation techniques showed that techniques closer to the natural mapping often exhibited serious usability problems. Testbed evaluation has confirmed this fact. Therefore, the use of “magic” techniques which differ greatly from the natural mapping (but which may still take advantage of well-developed human skills) is an important principle for VE interaction. Natural interaction techniques may still be useful, especially in situations where the VE is used as training for a real-world task, or where the target user population has no VE experience and will only use the VE for a short time.

Provide redundant interaction techniques for a single task.

One of the biggest problems facing evaluators of VE interaction is that the individual differences in user performance seem to be quite large relative to 2D interfaces. Some users seem to comprehend complex techniques easily and intuitively, while others may never become fully comfortable. Work on discovering the human characteristics that cause these differences is ongoing, but one way to mitigate this problem is to provide multiple interaction techniques for the same task. For example, one user may think of navigation as specifying a location within a space, and therefore would benefit from the use of a technique where the new location is indicated by pointing to that location on a map. Another user may think of navigation as executing a continuous path through the environment, and would benefit from a continuous steering technique. In general, “optimal” interaction techniques may not exist, even if the user population is well known, so it may be appropriate to provide two or more techniques each of which have unique benefits. Of course, the addition of techniques also increases the complexity of the system, and so this must be done with care and only when there is a clear benefit.

7.1.2 Guidelines for the Design of Travel Techniques

Make simple travel tasks simple by using target-specification techniques.

If the goal of travel is simply to move to a new location, such as moving to the location of another task, target-based techniques provide the simplest metaphor for the user to accomplish this task. In many cases, the exact path of travel itself is not important; only the end goal is important. In such situations, target-based techniques make intuitive sense, and leave the user’s cognitive and motor resources free to perform other tasks. The use of target-based techniques assumes that the desired goal locations are known in advance or will always coincide with a selectable position in the environment. If this is not true (e.g. the user wishes to obtain a bird’s-eye view of a building model), target-based techniques will not be appropriate.

Avoid the use of teleportation; instead, provide smooth transitional motion between locations.

Teleportation, or "jumping," refers to a target-based travel technique in which velocity is infinite – that is, the user is moved immediately from the starting position to the target. Such a technique seems very attractive from the perspective of efficiency. However, evaluation (Bowman et al, 1997) has shown that disorientation results from teleportation techniques. Interestingly, all techniques that used continuous smooth motion between the starting position and the target caused little disorientation, even when the velocity was relatively high.

If steering techniques are used, train users in strategies to acquire survey knowledge. Use target-specification or route-planning techniques if spatial orientation is required but training is not possible.

Spatial orientation (the user's spatial knowledge of the environment and her position and orientation within it) is critical in many large-scale VEs, such as those designed to train users about a real world location. The choice of interaction techniques can affect spatial orientation. In particular, evaluation (Bowman, Davis, Hodges, & Badre, 1999) has shown that good spatial orientation performance can be obtained with the use of steering techniques, where the user has the highest degree of control, but only if sophisticated strategies are used (e.g. flying above the environment to obtain a survey view, moving in structured patterns). If such strategies are not used, steering techniques may actually perform worse, because users are concentrating on controlling motion rather than viewing the environment. Techniques where the user has less control over motion, such as target-based and route-planning techniques, provide moderate levels of spatial orientation due to the low cognitive load they place on the user during travel – the user can take note of spatial features during travel because the system is controlling motion.

Constrain the user's travel to fewer than three dimensions if possible to reduce cognitive load.

Our information gathering experiment (Bowman, Koller, & Hodges, 1998) showed that the higher the dimensionality of the path the user travels, the more likely he is to forget information seen along that path. Many VE applications allow the user to fly in three dimensions, even when it is not necessary. A simple constraint that keeps the user on the ground plane should reduce cognitive load. The use of this guideline, however, must be tempered with the fact that 3D flying may also increase spatial orientation if used correctly.

Use non-head-coupled techniques for efficiency in relative motion tasks. If relative motion is not important, use gaze-directed steering to reduce cognitive load.

Relative motion is a common VE task in which the user wishes to position the viewpoint at a location in space relative to some object. For example, an architect wishes to view a structure from the proposed location of the entrance gate, which is a certain distance and direction from the front door – movement must be relative to the door, and not to any specific object. A comparison of steering techniques (Bowman, Koller, and Hodges, 1997) showed that a pointing technique performed much more efficiently on this task than gaze-directed steering, because pointing allows the user to look at the object of interest while moving, while gaze-directed steering forces the user to look in the direction of motion. Gaze-directed steering performs especially badly when motion needs to be in the opposite direction from the object of interest. Thus, techniques that are not coupled to head motion support relative motion tasks. On the other hand, non-head-coupled techniques are slightly more cognitively complex than gaze-directed steering, so it may still be useful if relative motion is not an important task.

7.1.3 Guidelines for the Design of Selection Techniques

Use ray-casting techniques if speed of remote selection is a requirement.

Evaluation (Bowman & Hodges, 1999) has shown that ray-casting techniques perform more efficiently than arm-extension techniques over a wide range of possible object distances, sizes, and densities. This is due to the fact that ray-casting selection is essentially 2D (in the most common implementation, the user simply changes the pitch and yaw of the wrist). Ray-casting includes both the virtual light ray metaphor and image plane techniques such as occlusion and framing.

Ensure that the chosen selection technique integrates well with the manipulation technique to be used.

Selection is most often used to begin object manipulation, and so there must be a seamless transition between the selection and manipulation techniques to be used in an application. Arm-extension techniques generally provide this transition, because the selected object is also manipulated directly with the virtual hand, and so the same technique is used throughout the interaction. As demonstrated by the HOMER technique, however, it is possible to integrate ray-casting techniques with efficient manipulation techniques.

If possible, design the environment to maximize the perceived size of objects.

Selection errors are affected by both the size and distance of objects, using either ray-casting or arm-extension techniques (Bowman & Hodges, 1999). These two characteristics can be combined in the single attribute of visual angle, or the perceived size of the object in the image. Unless the application requires precise replication of a real-world environment, manipulating the perceived size of objects will allow more efficient selection.

7.1.4 Guidelines for the Design of Manipulation Techniques

Reduce the number of degrees of freedom to be manipulated if the application allows it.

Provide general or application-specific constraints or manipulation aids.

These two guidelines address the same issue: reducing the complexity of interaction from the user's point of view. This can be done by considering the characteristics of the application (e.g. in an interior design task, the furniture should remain on the floor), by off-loading complexity to the computer (using constraints or physical simulation), or by providing widgets to allow the manipulation of one or several related DOFs (Mine, 1997).

Allow direct manipulation with the virtual hand instead of using a tool.

Tools, such as a virtual light ray, may allow a user to select objects from great distances. However, the use of these same tools for object manipulation is not recommended, due to the fact that positioning and orienting of the object is not direct – the user must map desired object manipulations to the corresponding tool manipulations. Manipulation techniques that allow the direct positioning and orienting of virtual objects with the user's hand have been shown empirically (Bowman & Hodges, 1999) to perform more efficiently and to provide greater user satisfaction than techniques using a tool. For efficient selection and manipulation, then, we need to combine a 2D selection metaphor such as ray-casting with a hand-centered, direct manipulation technique. This is the basis of techniques such as HOMER and Sticky Finger (Pierce et al, 1997).

Avoid repeated, frequent scaling of the user or environment.

Techniques that scale the user or the world to allow direct manipulation have some desirable characteristics. The user's perception of the scene does not change at the moment of selection, and small physical movements can allow large virtual movements. However, experimental data (Bowman & Hodges, 1999) shows a correlation between the frequent use of such techniques and discomfort (dizziness and nausea) in users. Techniques that scale the user or environment infrequently and predictably should not suffer from these effects.

Use indirect depth manipulation for increased efficiency and accuracy.

Indirect control of object depth, using joystick buttons for example, is not a natural technique (although it borrows from a real-world "fishing reel" metaphor), and requires some training to be used well. However, once this technique is learned, it provides more accurate object placement, especially if the target is far from the user (Bowman & Hodges, 1999). This increased accuracy leads to more efficient performance as well. Moreover, these techniques do not exhibit the arm strain that can result from the use of more natural arm-extension techniques.

7.2 Formal Design & Evaluation Frameworks

A second major contribution of this work is the framework and methodology we proposed and used in all of the design and evaluation components of the research. The methodology includes the use of taxonomy, guided design, multiple performance metrics, consideration of outside factors on performance, and testbed evaluation.

Such a framework has several advantages. First, formalism is a great aid to understanding. In order to create a useful and believable taxonomy, for example, it was necessary to study and consider both the interaction task and the techniques proposed for that task. Second, the use of the methodology in multiple experiments allows us to view all of the results within a common framework. For example, we know many of the relative merits of the common steering techniques gaze-directed steering and pointing due to the multiple evaluations. Third, the framework provides a common ground for discussion among researchers in the field, allowing more precise and well-understood conversations.

Finally, special mention needs to be made of the utility of guided design in creating new interaction technique possibilities. We have shown that combining previously untried sets of components can produce useful and interesting techniques (such as HOMER). Furthermore, there seems to be a slowing of the publication of completely novel interaction techniques and metaphors for immersive environments. It is possible, though certainly not proven, that we have identified many, or most, of the fundamental components of VE interaction for these universal tasks. If so, then guided design becomes the best method for covering the design space.

7.3 Focus on Applications and Usability

A third major contribution of this research has been our focus from the beginning on improving the usability (and more generally, the performance) of immersive VE applications. As we noted in the first chapter, there are very few in-use applications of VEs due to the usability problems associated with high levels of interactivity. Therefore, this research has not been simply an academic exercise. It has had as its goal from the beginning to improve VE interaction for real-world applications. This led us to a methodology that included applications and their requirements explicitly.

There have been other efforts to quantify the performance of VE interaction techniques, but few of them have extended this work to real applications. On the other hand, a large number of applications have been prototyped, but interaction was developed in an ad hoc manner, based on intuition. This work has bridged the gap, providing empirical evidence and practical guidelines for real applications based on formal evaluation.

We believe that this philosophy of research will be fruitful in other types of virtual environments, such as tabletop stereo displays, and in many emerging areas of interactive systems, such as augmented reality, ubiquitous computing, and wearable computing. Because of their newness, such areas need empirical, low-level studies to quantify performance and effectiveness. However, these research areas are also under pressure to produce real applications to prove that the research funding is worthwhile. Using the philosophy embodied in this thesis, which we call "basic research with an applied focus," can allow both of these things to happen in the same research program.

7.4 Future Work

Research in a relatively new area usually raises more questions than it answers, and this work is no exception. There are a multitude of topics in the general area of VE interaction that still need to be explored in depth. In particular, there are four areas directly related to the current work that we claim would be extremely useful.

7.4.1 Automatic Interaction Design and Performance Modeling

Our testbed evaluations and other experiments have produced a large body of empirical results for IT performance on various tasks. However, it is still difficult for application designers to wade through these numbers in order to choose an appropriate set of techniques for a particular system. Therefore, it would be useful to create a tool that automates some of this process for the developer. Such a tool would likely ask the developer a series of questions about the application, including what tasks were involved, what requirements existed for the various aspects of performance, and what devices were available. It could then, based on evaluation results, suggest a set of interaction techniques that would fit the requirements.

This leads to another problem, however, in that such a tool would only be able to suggest the use of techniques that had actually been tested experimentally. It would be more useful if the tool could predict the performance of an untested technique by interpolating the results from related techniques. Fortunately, our taxonomy and framework is set up to allow the creation of these predictive models of performance.

Consider a simple example. A task has two subtasks, each of which has two components. The components are numbered one through four (figure 7.1). An experiment found that technique A, composed of components 1 and 3, scored 5.0 on a certain metric; technique B, composed of components 1 and 4, scored 10.0; and technique C, composed of 2 and 3, scored 8.5. A simple prediction algorithm would guess that component 2 is responsible for 1.5 units more than component 1 (based on the scores of techniques B and C). So, to predict the score for technique D, composed of components 2 and 3, we can take the score for technique A (1 and 3), and add 1.5, for a score of 6.5. The same result would be obtained if we first determined the contribution of component 3 relative to 4 (3.5 units less), and then added this to the technique B score (10.0-3.5 = 6.5).

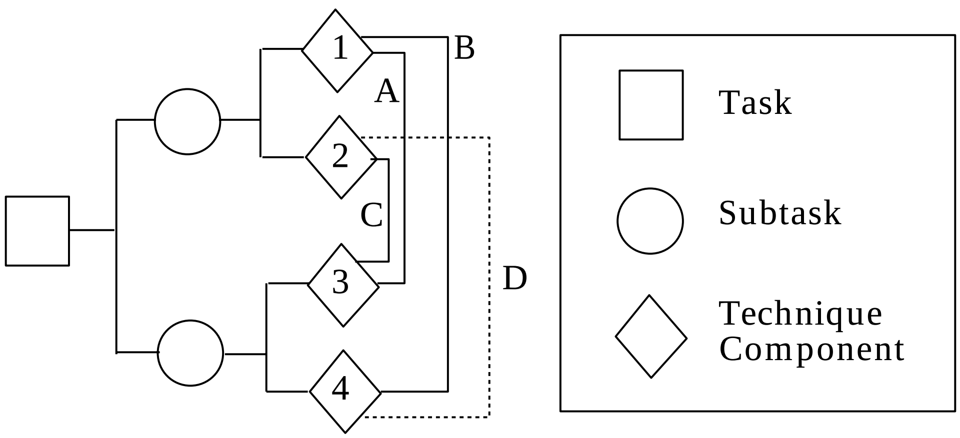

Figure 7.1 Example Taxonomy and Technique Components: If Performance Results for Techniques A, B, and C are Known, the Performance of Technique D can be Inferred

With more complex results, such a simple prediction is not possible, but the same concept holds. Regression or other types of analysis of the experimental data would lead to predictive models that would predict the performance of any technique which falls within the space defined by the techniques actually tested.

7.4.2 Cross-task Interaction Techniques

In this work, we have found a number of times that a technique originally designed for one task is useful for another task, with slight modifications. For example, the route-planning technique for travel actually uses manipulation of objects in a small version of the environment, similar to the World in Miniature (WIM) technique (Pausch et al, 1995).

This concept can be generalized when one realizes that all three of the universal tasks have as their basis the specification of a spatial position and/or orientation. Travel sets the position and orientation of the viewpoint, manipulation does the same for an object, and selection can be thought of as specifying the position of an object as a naming mechanism. This means that we can consider a technique designed for any one of the tasks as a possible technique for any of the others. We call these “cross-task” interaction techniques, because they cross the boundaries between the tasks.

In fact, many such techniques have already been developed, most of which use manipulation techniques to effect travel. There are other possibilities, however, such as using travel for object manipulation (the user "becomes" the object and sets the position and orientation from a first-person point of view), or using object selection for manipulation (place the object in my hand next to the selected object). This analysis also implies that the taxonomies for the three tasks are actually linked together, creating a single unified design space, as shown in figure 7.2.

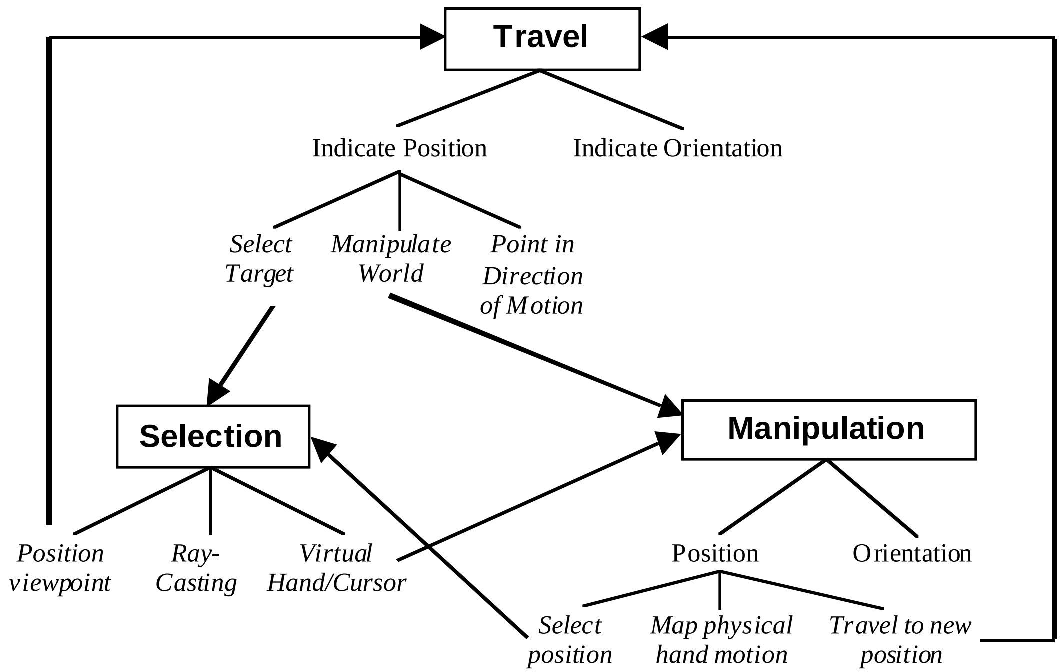

Figure 7.2 Simplified Taxonomies Linked Together by Cross-Task Techniques

We believe that cross-task interaction techniques can be useful and powerful in VEs. In particular, they have the advantage that the same metaphor may be used for multiple tasks, increasing the consistency of the interface and reducing the amount of complexity with which the user has to cope. Further research into such techniques should prove fruitful.

7.4.3 Comparison with Usability Engineering

Our design, evaluation, and application methodology has proven to be useful in increasing our understanding of VE interaction and in increasing the usability and performance of a specific VE application. However, our methodology is not the only way to improve system usability. One particular method that has received attention recently is usability engineering.

Usability engineering has a tradition in 2D HCI research, and has now been applied to VEs (Gabbard & Hix, 1998). The basic approach is centered on a particular VE system, and the iterative design and evaluation of the interaction and interface in that system. Like our methodology, it relies on a formal task and user analysis. Unlike our techniques, it uses more qualitative performance metrics, and performs evaluation within the system rather than in a generalized testbed.

Obviously, these two methodologies each have their advantages and disadvantages, and it is likely that they are complementary techniques. However, it would be instructive to do a controlled comparison of the two to determine where most of the gains in performance and usability come from. We would hazard to guess that neither method alone is sufficient. Usability engineering will not work unless it begins with a set of possible interaction techniques that have good performance characteristics, and our methodology will likely produce an application that would still benefit from iterative design and evaluation.

7.4.4 Interaction in Other Display Modalities

Finally, our work has focused solely on immersive VEs that are implemented using head-mounted displays. While this is still the most common VE display device, it has fallen out of favor in some circles, and other displays such as tabletop stereo displays and spatially immersive displays (e.g. the CAVE™) are being widely tested.

However, the VE community has no notion of how these various display modalities differ or what applications or tasks for which each is appropriate. Some vague notions exist based on intuition and limited experience, but for the most part a given display is used simply because it is available.

The studies we have presented have some generality, and the principles derived from them can be applied in a variety of VEs. On the other hand, interaction in the other display modalities is likely to be somewhat different from interaction in an HMD-based VE, and so further work in this area is needed. In particular, it would be interesting to study whether the relative performance of various ITs changes as we move to a new display modality. A study of task appropriateness in the different modalities would also be instructive.