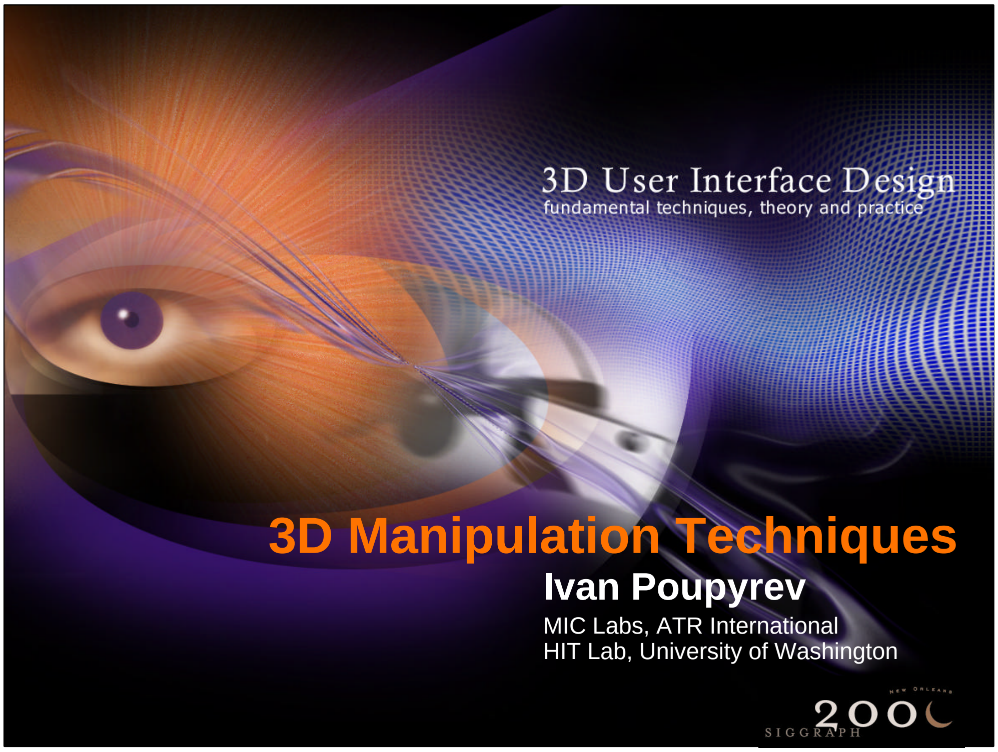
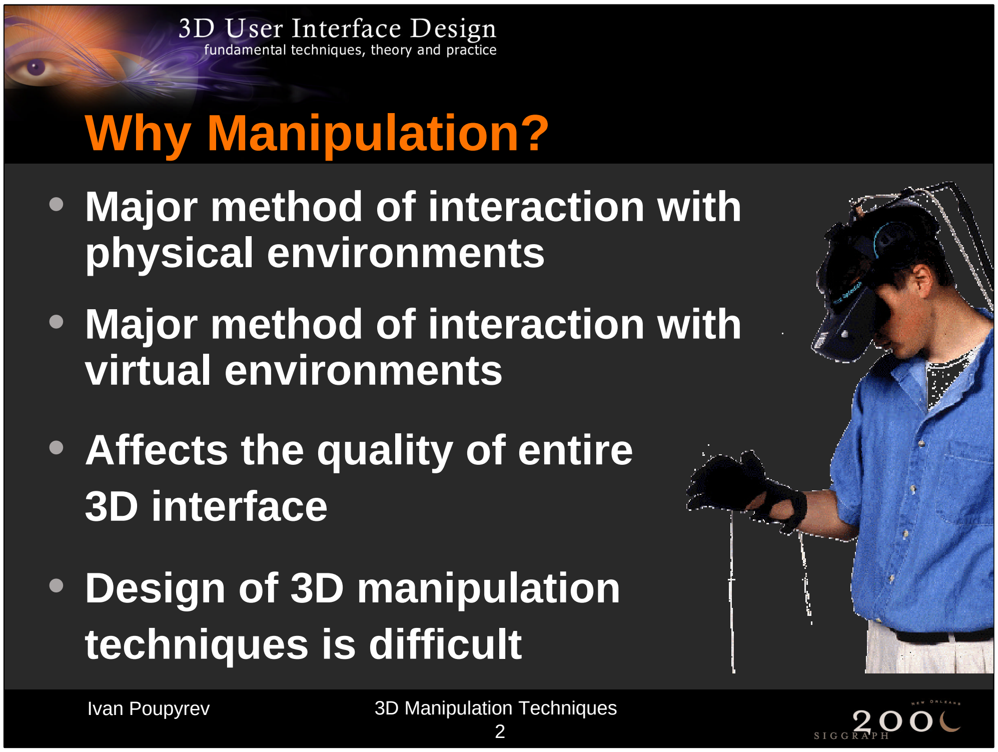


3D Manipulation Techniques

Ivan Poupyrev
Researcher

Media Integration and Communication Research Laboratories (MIC Labs),
Advanced Telecommunication Research Institute International (ATR),
Kyoto, Japan

http: http://www.mic.atr.co.jp/~poup
email: poup@mic.atr.co.jp

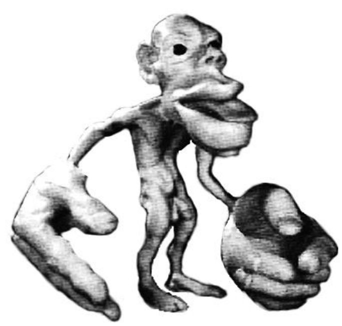


Why manipulation techniques?

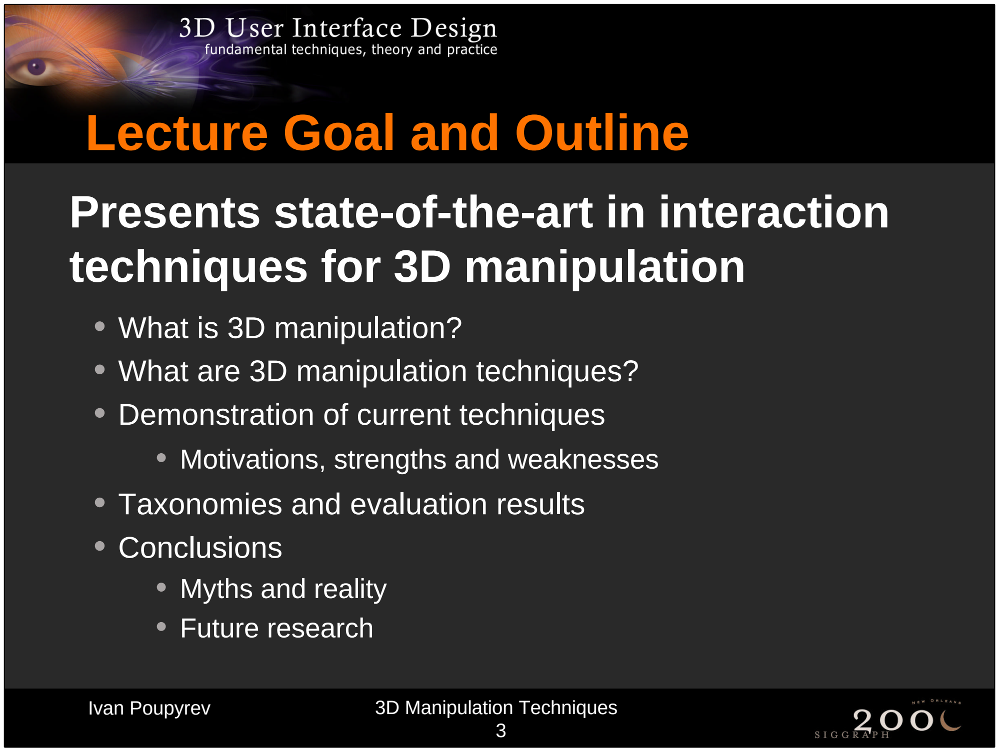
Touching, picking, and manipulating objects is the main way for humans to affect their surrounding physical world. From the moment we are born we learn to manipulate things around us using our hands and by adulthood we get to the point where we do not need any conscious attention to perform extremely complex manipulation. The areas of the human brain that allow us to control hands are highly developed, occupying a major portion of the human motor cortex, which is often demonstrated using motor cortex homunculus cartoon, designed by Wilder Penfield (see below, also Zhai, Milgram, Buxton 1996).

Naturally, hand manipulation is the major method of interaction in virtual worlds. While voice, gaze and movement of other other body parts are also used, direct hand manipulation remains the most natural and efficient input method for humans in 3D user interfaces. That is why, 3D manipulation techniques have a profound impact on the quality of the whole 3D interface: if the user cannot efficiently manipulate objects in virtual environments, then other high-level tasks simply cannot be accomplished. Hence, understanding techniques for direct 3D manipulation is an important and necessary step toward developing effective VR applications.

The development of effective direct manipulation techniques, however, has turned out to be a difficult problem (see Mine, Brooks and Sequin, 1998, for discussions on some sources of this difficulty). In this part of the course, I survey interaction techniques for virtual manipulation,discuss their strengths and weaknesses, and investigate design issues involved in using manipulation techniques in VE.



The motor cortex homunculus model maps the amount of the brain's motor cortex devoted to the control of particular body part into the size of this body part.



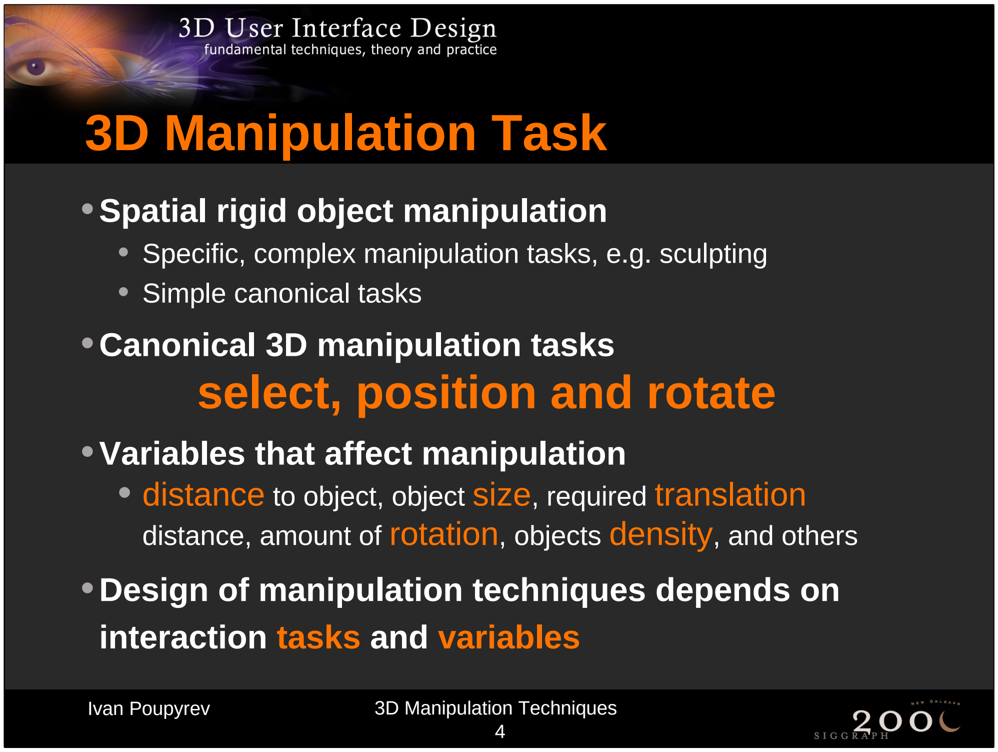
The lecture goal and outline

Manipulation in VR is a vast topic. The present discussion is limited only to *interaction techniques* for 3D manipulation, many of them reported recently. The large variety of techniques that have been reported, on one hand, present the interface developer with a choice of techniques that can be effectively used in designing 3D user interfaces. On the other hand, the lack of guidelines, taxonomies and formal experimental evaluations makes informed design decisions difficult.

The goal of this talk is to present the state of the art in interaction techniques for 3D manipulation. It starts with a discussion of the 3D manipulation task and its properties, since the nature of the task directly affects the design of techniques. The place of interaction techniques in 3D user interface is discussed after that.

I then overview some classical and recent techniques for virtual manipulation. For each technique, I present the motivation for its development, the ideas behind the techniques, demonstrate how techniques work using video or illustration, and finally, discuss their strengths and weaknesses. I will also mention several attempts to build aggregative techniques, discuss classifications of the manipulation techniques, and present some results of their experimental evolution.

In conclusions, I will discuss some myths and realities in the design of manipulation techniques, as well as future research directions.



What is a manipulation task?

The design of interaction techniques is driven by the desire to maximize user performance and comfort in a manipulation task. That is why the understanding of a manipulation task and its properties is important for the effective design of 3D manipulation dialogs and techniques.

The word “manipulation” itself is greatly overused. In colloquial language, it often refers to any task involving or resembling manipulation, from *physically* manipulating objects to changing, or manipulating, any other properties of objects, such as color. Here, I use a narrow definition of a manipulation task as spatial rigid object manipulation. This follows from the definition of manipulation in classical studies on the manipulation used in human factors and motion analysis literature (McCormick, 1970, Mundel, 1978).

Manipulation task analysis: two approaches

Even within this narrow definition, there is virtually an infinite variety of possible manipulations that users may want to accomplish. Hence, the design of tools that allow the user to accomplish one manipulation task may not be appropriate for another task: consider, for example, hammering down a nail or writing a letter, both are examples of complex manipulation tasks. That is why, in order to design and evaluate manipulation techniques, first we have to analyze 3D manipulation tasks and their properties. There are two approaches to doing this.

The first approach is to analyze a very specific, usually complex manipulation task. The main applications of this approach are in performance critical applications, such as industry, military, medicine, and other applications with few very well defined manipulation sequences. The major drawback is that techniques developed for some of these specific task might not be usable in other, even very similar, tasks. Hence, each time we want to develop the technique for a some task, we must from scratch.

Continued on the next page

The second approach to manipulation task analysis is based on the assumption that all human motions are composed of the simple manipulation tasks, which are basic building blocks for complex task scenarios (Philips, 1988). Although emergent properties and context effects must be accounted for, the development of interaction techniques for these elementary tasks can potentially result in techniques that be able to deal with *any* manipulation task. The task decomposition approach was pioneered by Gilbreth and Gilbreth early in this century (1917). In one classic example, they analyzed the task of brick laying and broke it down into 13 basic motion elements, which they called “therbligs”; by studying and optimizing therbligs, they improved the brick layers' performance by a factor of three.

This approach’s generality is the main advantage but also its main drawback. The generic techniques designed for simple canonical tasks may be ineffective for certain specific tasks and generally do not allow for the maximization of user performance. Furthermore, there is no assurance that all important manipulation task conditions have been accounted for. Still, the task decomposition approach is suitable for developing techniques that can be used across many application domains. In addition, the techniques developed for generic tasks can be easily tailored and optimized for specific tasks in performance critical applications.

Canonical 3D manipulation tasks

What are the canonical tasks for rigid object manipulation? Following previous work on 2D user interfaces (Foley et al., 1984; Grissom, 1995) as well as other categorizations proposed for 3D interfaces (Mine, 1995; Poupyrev, 1997) we suggest three basic manipulation tasks:

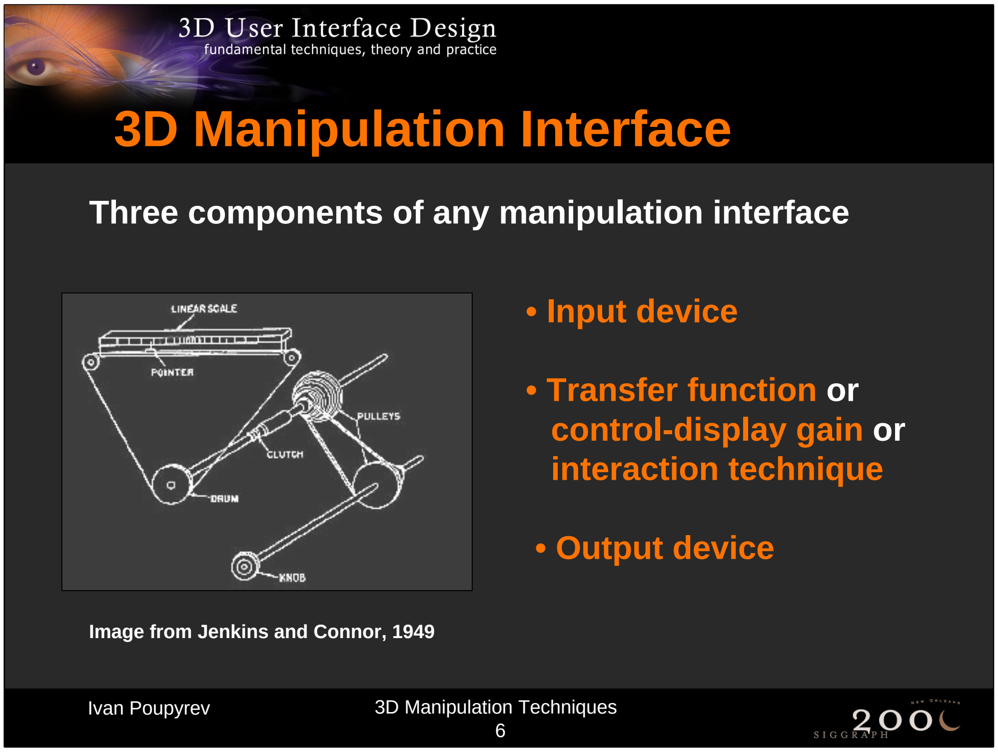
- *position*, the task of positioning an object from an initial to a final, terminal, position,
- *selection*, the task of *manually* identifying an object, and
- *rotation*, the task of rotating an object from an initial to a final orientation.

While scaling has been suggested as a basic manipulation task (Mine, et al. 1998), we do not include it here since it is not a rigid object manipulation task.

Variables in manipulation tasks

As I noted before, the main criteria in techniques design is maximizing user performance. For each task, there is a number of variables that routinely affect user performance. For example, *distance* to the objects is one such variable: not only a change in the distance to the manipulated object affects the user performance, but also different techniques are needed for local manipulation and manipulation at-a-distance (Mine, 1995; Hand, 1997). The other important variables of interest are object *size*, object *density*, *translation distance* in the positioning task, and the *amount of rotation* in 3D rotation task.

A design of 3D manipulation techniques and manipulation dialogs can only be done relative to the certain task conditions defined by the manipulation tasks and values of variables affecting user performance.



Manipulation techniques in 3D interface

Virtual manipulation as well as manipulation in the natural world are examples of a continuous control task; the basic elements of continuous control systems were first outlined in the early 50s. According to it any interface between humans and machines that uses continuous manual control includes three basic components: 1) the input devices, capturing user actions; 2) the display devices, presenting the effect of these actions back to the user; and 3) the *transfer functions*, often referred to as *control-display mappings* or *interaction techniques*, which map the device's movements into the movements of system's or interface's elements. These three components are represented on the slide's figure as knob, pointer and pulleys, respectively. The goal of interface design is to design input devices, displays and transfer functions that facilitate high user performance and comfort, while diminishing the impact from both human and hardware limitations.

The manual control interfaces for 3D interactive computer graphics applications follow the same principles (Zhai, 1996, Latta, et al., 1994): the user interacts with VE applications in a closed-loop system by applying motor stimuli to input devices and receiving sensory feedback through display devices. The interaction techniques map the user input captured by devices, such as the positions of head and hands, into the corresponding actions and commands within the virtual environment.

The design of interaction techniques depends on properties of the input and output devices, such as whether the device is isotonic or isometric, the device's form factor, and others (e.g. Zhai, Milgram, Buxton 1996). Furthermore, in many cases, interaction techniques designed for a certain class of input devices will not work for a different class of devices.

3D User Interface Design
fundamental techniques, theory and practice

Techniques Design Objectives

- **Design 3D manipulation techniques**
  - that conform to used input and output devices
  - that are effective in desired task conditions
    - object distance: within/outside the reach
    - object sizes and shape: small, large, flat objects
    - object density
    - accuracy
  - that allow high user performance and comfort
  - that are easy to learn
  - that conform to external constraints, e.g., price

Ivan Poupyrev | 3D Manipulation Techniques 7 | SIGGRAPH 2000

Objective in manipulation techniques design

The objective is to design interaction techniques that satisfy a number of properties. First, the techniques should work well with the designated 3D input and output devices. For example, different techniques and manipulation scenarios should be developed for an HMD-based VR system or for a desktop 3D application.

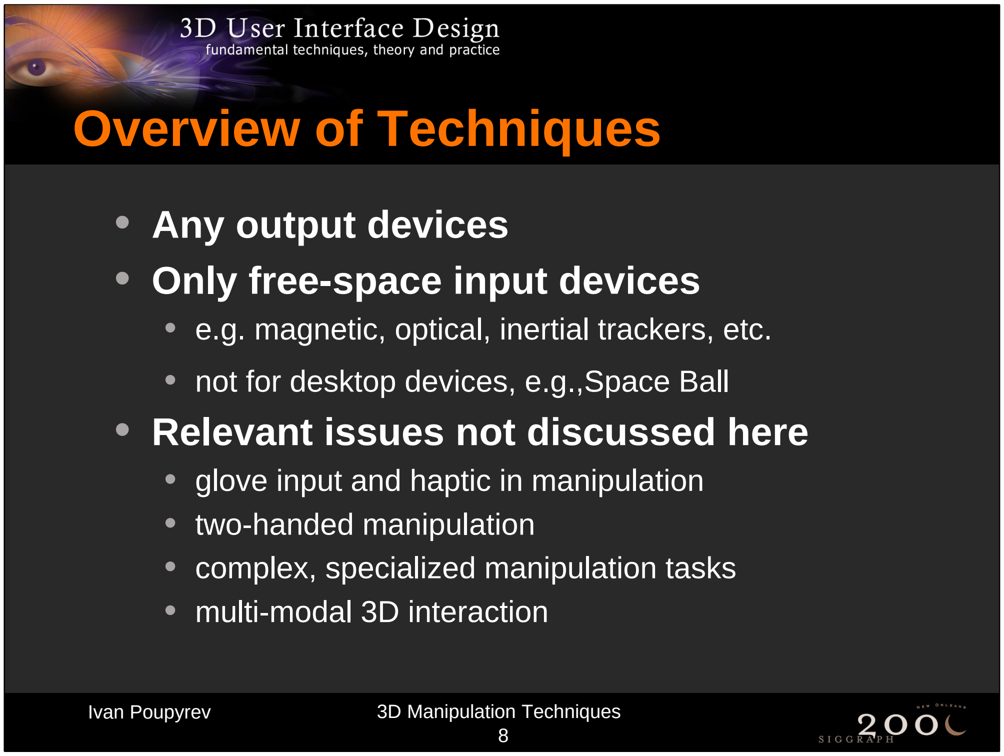
Second, the techniques should perform well in the required task conditions. The typical questions that 3D interface designers should ask are:

- Does the application require manipulation at a distance or not?
- What are the typical sizes of the virtual objects that will be manipulated?
- How many objects are there and how precise should the manipulation be?

Third, the manipulation techniques should allow for high user performance and comfort and should not depend on the particular user characteristics, such as experience, left-handedness or right-handedness, age and so on.

Fourth, the manipulation dialog should be easy to learn and master.

Fifth, the interface design process should conform to a variety of external constraints such as whether the user can move or not (in other words how much physical space is available) or how much effort (and money) can be to spent on the system development and so on.



Overview of 3D manipulation techniques

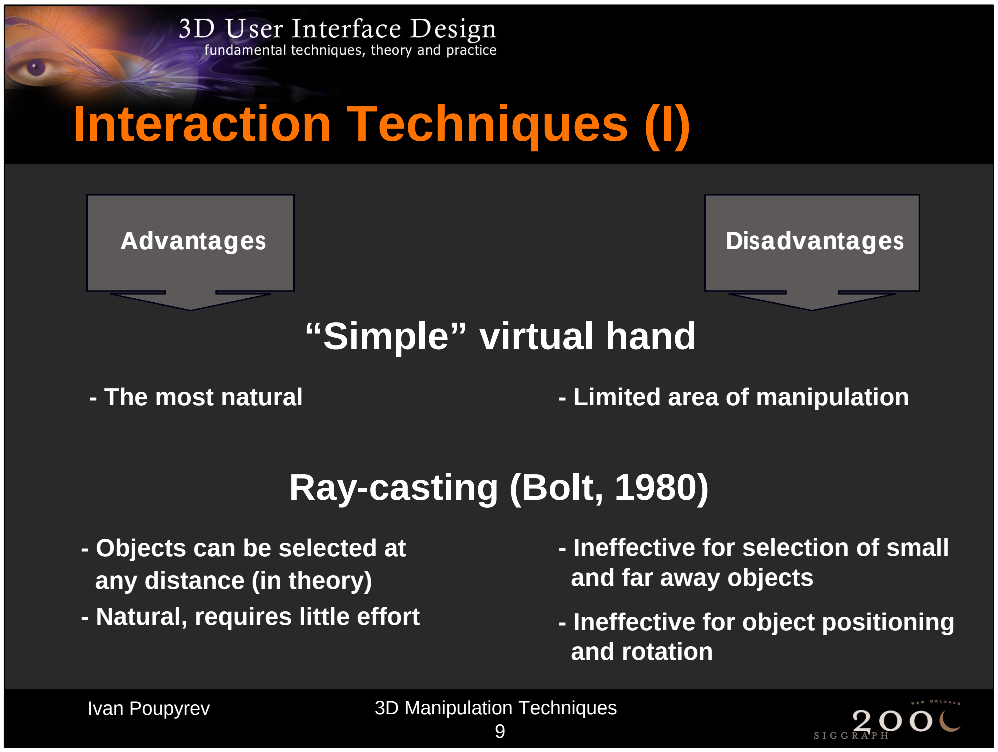
Before starting discussion of interaction techniques for 3D manipulation, I have to introduce certain limits because it is impossible to present every possible interaction technique in one lecture.

First, while techniques I review in this lecture can be used with almost any type of output devices, though with different performance implications, there are some restrictions on the input devices: only free-space isotonic devices, such as magnetic, ultrasound, and optical trackers, can be used. Most of the techniques cannot be used with desktop isometric devices, such as SpaceBall, and 2D input devices, such as mouse.

Second, some important and relevant issues will not be discussed in this lecture:

- whole hand input, i.e., gloves and gestures, as well as use of haptic;
- two-handed manipulation;
- techniques for complex, specialized tasks, such as sculpting;
- multimodal interaction.

Each of these topics could be a theme for a separate lecture.

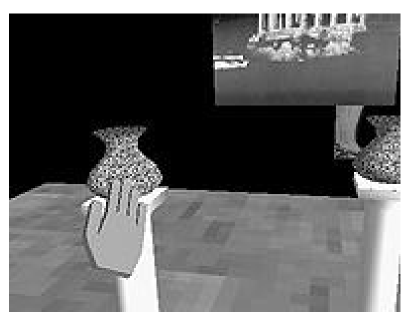


Virtual hand and virtual pointer

The *virtual hand* and *virtual pointer* (ray-casting, laser ray) are the most basic techniques for 3D manipulation in VEs. One of the earliest implementation of these techniques has been described in the "Put-that-There" interface developed by Bolt at MIT in the early 80s. The "Put-That-There" featured the 6DOF magnetic sensor for selecting objects by pointing and manipulating them. Since then, countless variations of ray-casting have been developed and reported (Jacoby et.al., 1994; Mine, 1995; Bowman and Hodges, 1997).

The *virtual hand* technique is a direct mapping of the user hand motion into the affected motions in a virtual environment, typically linearly scaled to establish the correspondence between the device and environment coordinate systems. The user is provided with a virtual "hand" - a 3D cursor, often shaped like a human hand, whose movements correspond to the movements of the tracker worn on the hand or held in the user's fingers. To select an object, the user simply intersects the virtual hand with the target, and presses a trigger (or issues a voice command or a hand gesture) to pick it up. The object is then attached to the virtual hand and can be easily translated and rotated within the VE. The technique is rather intuitive; one problem is that only those objects that are within the area of reach can be picked up, and this significantly limits the technique's applicability.

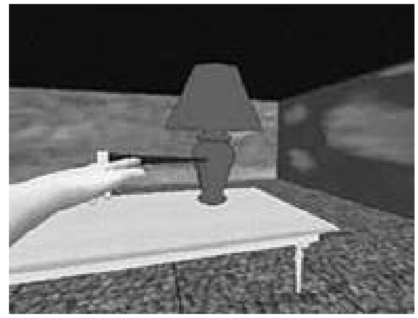
Continued on the next page



Virtual hand (from Poupyrev et al., 1996)

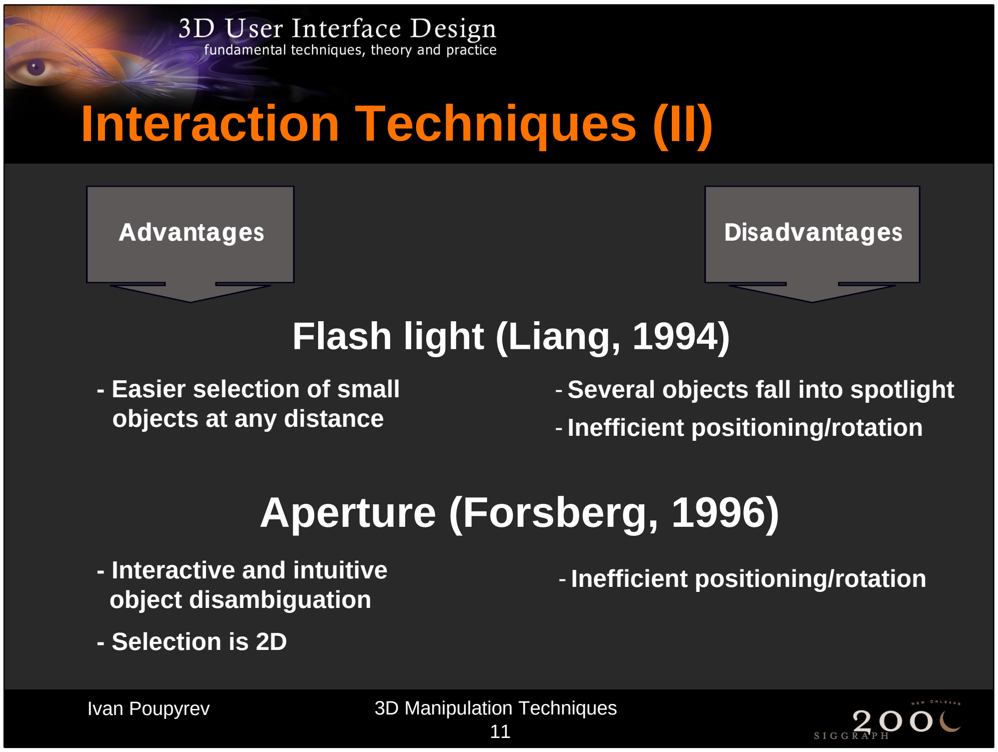
The motivation behind the *ray-casting* technique was to allow the user to select and manipulate objects beyond the area of normal reach. The user points at objects with a virtual ray emanating from a virtual hand and then objects intersecting with the virtual ray can be selected, attached, and manipulated.

While this technique is very easy to use, several studies and informal observations have shown that selecting small objects (i.e., when high selection accuracy is required) or those-at-a distance is difficult with ray-casting (Poupyrev, et al. 1998); in fact, even at close distances, selecting with high accuracy might be more efficient with a virtual hand.



Virtual pointer (from Bowman et al, 1997)

Moreover, object manipulation can be efficiently accomplished only in radial movements around the user (perpendicular to the ray direction) and rotations only around the ray axis. Full 6DOF manipulation with ray-casting is impossible. Hence, even though ray-casting is seemingly a 6DOF manipulation technique, there are natural constraints that limit the user manipulations degrees of freedom.

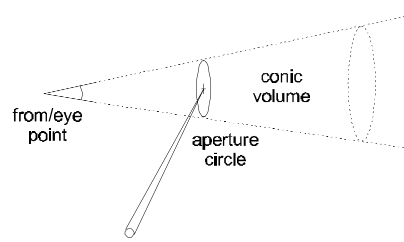


Improvements in ray-casting: Flashlight and Aperture techniques

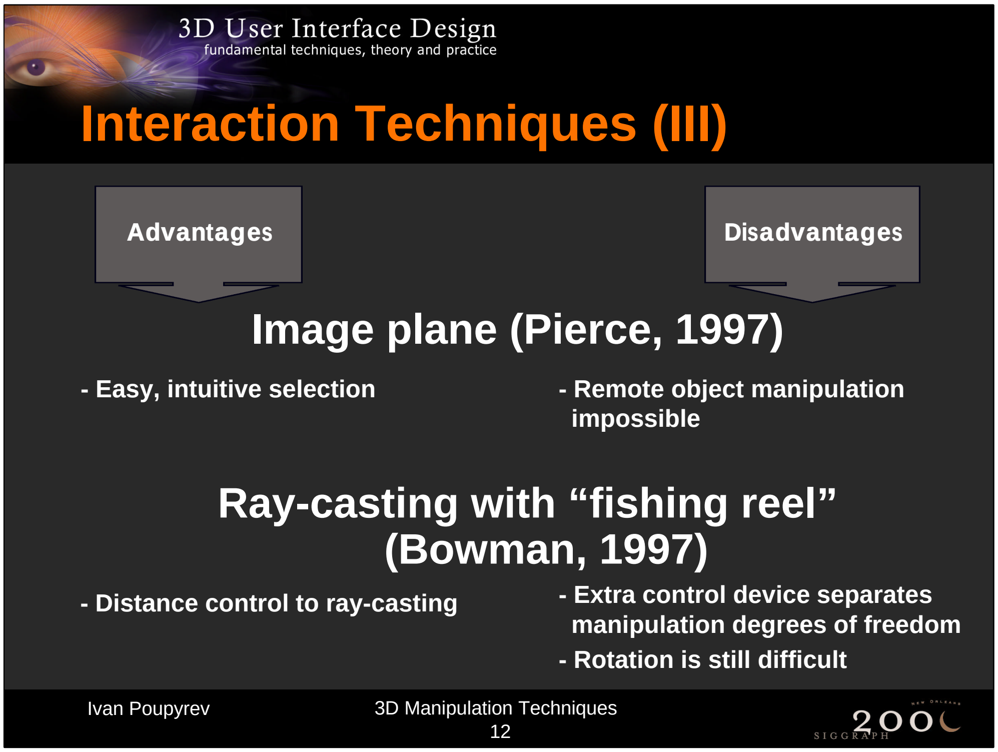
Both the classical version of the virtual hand and the virtual pointer have weaknesses that are disturbing enough to warrant improvements.

The *spotlight* or flashlight technique (Liang, et al. 1994) uses a conic selection volume, so that objects falling within the cone are selected. The technique imitates pointing at objects with a flashlight and allows the user to easily select small objects no matter how far they are from the user. The shortcoming of this technique is that more than one object can fall into the spotlight, especially with increased distance to the object. This requires techniques for further disambiguation of the target objects. Furthermore, all limitations on object manipulation that have been discussed for ray-casting techniques are also true for spotlight.

The *aperture* technique (Forsberg et al., 1996) is a modification of the spotlight technique that allow the user to interactively control the selection volume. The conic pointer direction is defined with the location of the user's eye, which is estimated from the tracked head location and the location of a hand sensor represented as an aperture cursor within the VE (see figure below). The user can control the diameter of selection volume by bringing the hand sensor closer or moving it farther away. The aperture technique, thus, perfects the spotlight technique by providing an interactive mechanism of object disambiguation within the conic volume.

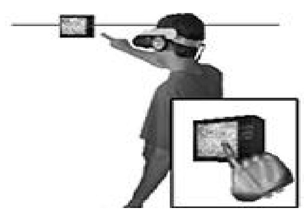


Aperture technique
(Forsberg, et al. 1996)



The Image Plane techniques

The *Image Plane* family of techniques (Pierce et al., 1997) provides a way to simplify the object selection task by reducing a number of degrees of freedom from 3D to 2D. The techniques allow the user to select and manipulate objects by interacting with their 2D projections on an image plane in front of the user: the user selects objects by simply touching their projection on an image plane.



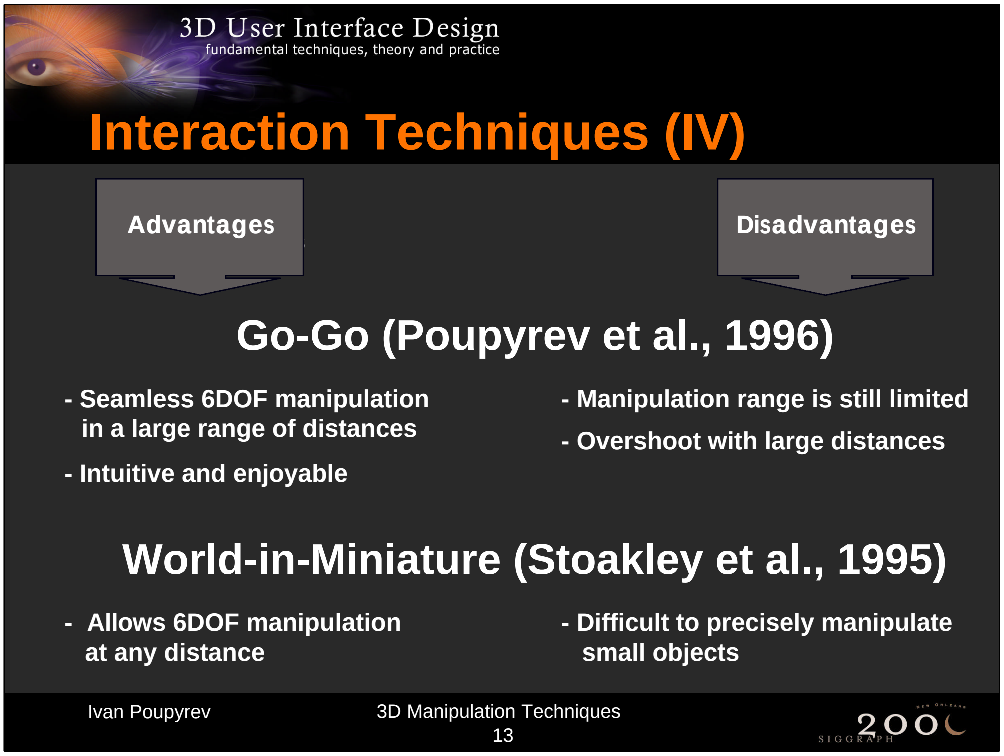
Sticky finger (Pierce, et al. 1997)

The object underneath the user’s finger is selected by first casting a vector from the user’s eye-point through the finger, and then finding an object intersecting with this vector (figure on the right). Alternatively, the user can select objects using two fingers by positioning his hand so that his thumb and index finger are directly positioned on the image plane below and above the target object (Head Crusher technique).

After selection, the object “projection” can be manipulated: the object is scaled down and brought within the user reach, so the user in some sense can manipulate the projection of the object. However, the distance to the object cannot be directly controlled. Mine’s world-scale grab technique (1997) addresses this problem.

Fishing reel

Difficulties with the control of object distance are an attribute of all pointing techniques. One possible way to address this is to supply the user with a direct method of controlling distances via a *fishing reel*. After selecting an object, the user can reel it back and forth using the mouse (Bowman, 1997) or some special purpose device. The fishing reel lets the user control the distance, but it separates the manipulation’s degrees of freedom. The ray direction is controlled by the 6DOF movements of the user’s hand, while distance is controlled by a separate controller. This technique also requires an extra input device for the control of the ray length.

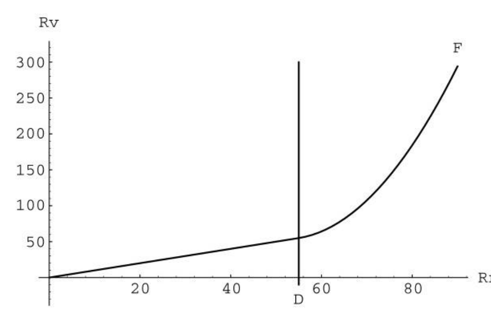


The Go-Go technique

The *Go-Go* technique (Poupyrev et al., 1996), flexibly extends the virtual hand's reaching distance by using a non-linear mapping function applied to the user's real hand extension. The space around the user is split into two concentric regions. While the user's real hand is within the first closest region around the user, that is, the distance to the hand is smaller then some threshold distance D, the mapping is one-to-one and the movements of the virtual hand correspond to the real hand movements (see figure below). However, as the user extends her/his hand beyond D, the mapping becomes non-linear and the virtual arm "grows", thus permitting the user to access and manipulate remote objects.

Different mapping functions can be used to achieve a different control-display gain between real and virtual hands (Fast and Stretch Go-Go techniques of Bowman, et al. 1997).

Go-Go technique allows direct seamless 6DOF object manipulation both close to the user and at-a-distance. Some evaluations (Bowman et al., 1997; Poupyrev, et al. 1998) have shown that the technique is natural and intuitive; allowing full 6DOF of object manipulation and rotations. The maximum reaching distance afforded by the technique, however, is limited; as the distance increases, the technique maps small movements of the user hand into the large movements of the virtual hand, which complicates precise manipulation at a distance.



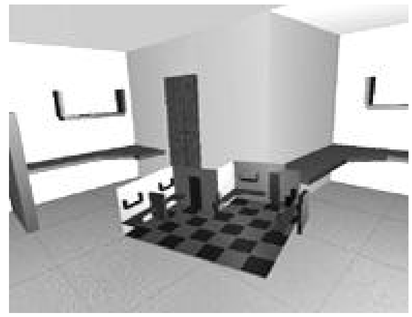
Mapping function for the Go-Go (Poupyrev, et al. 1996)

Continued on the next page

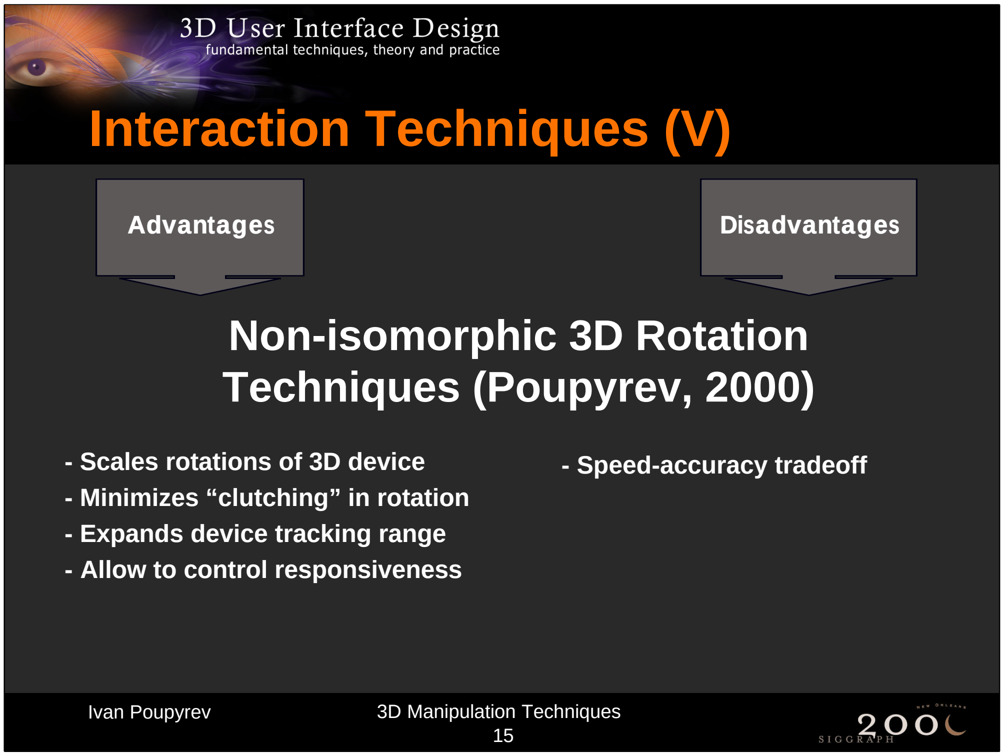
World-in-Miniature

An alternative to extending the length of the user arm is to scale the entire world and bring it within the user reach. The World-In-Miniature (WIM) technique (Stoakley et al., 1995) provides the user with a miniature hand-held model of the VE, which is scaled down using some constant coefficient (see right figure). The user can then indirectly manipulate virtual objects by interacting with their representations in the WIM.

The WIM technique is a powerful technique allowing easy object manipulation both within and outside of the area of user reach. It also can combine navigation with manipulation since the user can easily move his or her own representation on the WIM. There is a downside to the technique. When scaling a large environment results in very small representations of objects in the WIM, accurate manipulation of small objects might become difficult. A technique that would choose the part of the environment within the WIM could overcome this problem.



World-in-Miniature
(Stoakley, et al. 1995)



Non-isomorphic 3D rotational techniques

All of the techniques discussed above mostly dealt with object selection and translation tasks. 3D rotational techniques have been neither designed nor investigated until very recently (Poupyrev, et al. 2000).

3D rotational techniques that interactively amplify device rotations can provide the following advantages 1) minimize need for clutching during a large range of rotations, 2) extend tracking range of input devices, and 3) control of the sensitivity of spatial rotation allowing for more or less responsive user input in 3D interaction dialogs. All of these problems have been discussed before in the literature (e.g. Zhai, et al. 1996).

One of the difficulties in designing effective 3D rotational techniques is that a space of rotations is not a vector space, but a closed and curved surface, a manifold, in 4 dimensions, which can be represented as a 4D sphere. All possible 3D rotations of the rigid body are distributed on the surface of this sphere. Therefore, to design the interaction techniques that would transform the rotation of the input device, we have to operate on the surface of this sphere, i.e. use arcs instead of vectors.

Nuts and bolts of amplifying rotations of 3D input devices

The most basic operation on the 3D rotations of the input device is to amplify its rotation, which would result in more or less sensitive responses to the rotations of the input device. We can show that:

1) a basic equation of rotation amplification is a power function of the device rotation relative to a reference point:

$$R_V = R_R^k$$

where *Rr* is the physical rotation of the device, *Rv* is the virtual rotation in a 3D environment, and *k* is a coefficient of the amplification

2) Unlike in translations, in rotations scaling an absolute device orientation, as measured on each cycle of the simulation loop, is *different* from scaling the incremental device rotation on each cycle of the simulation loop. In fact, this would result in two different interaction techniques, each with very different properties.

a) The *absolute* amplification of the 3D device rotations preserves the *nulling consistency* of rotations, i.e., rotating the physical device in the initial orientation returns the virtual object to the initial orientation. However, this mapping does not necessarily preserve the *directional consistency* of rotations (stimulus-response correspondence), i.e., the rotation of the device would not always rotate the virtual object in the same direction as the device.

b) The relative amplification of the device rotation shows opposite properties: the relative amplification does not preserve the nulling consistency of the device rotations, but it does preserve the directional consistency of the rotations, i.e. the virtual object would *always* rotate in the same direction as the physical device.

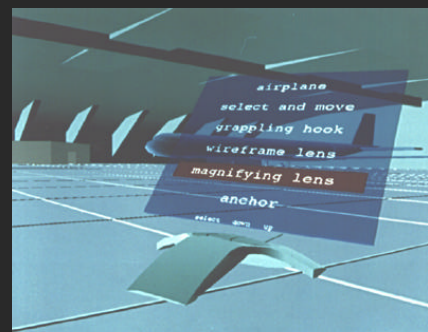
3) The choice of the mapping depends on the choice of the form factor of the input device: if the user does not have a constant feedback on the device orientation, e.g. when the device is a sphere, then nulling compliance is not important, since the user cannot perceive the mismatch between the initial orientation of the device and initial orientation of the virtual object. Amplified techniques, in this case, can provide better user performance.

For details on designing 3D rotational techniques refer to Poupyrev et al., 2000.

3D User Interface Design
fundamental techniques, theory and practice

Technique Combination

- **Combining techniques**
  - Aggregation
  - Hybrid techniques
- **Aggregation**
  - Virtual Tricorder (Wloka, 1995)
  - Tool belt (Mine et al., 1997)
- **Hybrid techniques**
  - HOMER (Bowman, 1997)
  - World-scale grab (Mine, 1997)
  - Voodoo Dolls (Pierce, 1999)



Virtual Tricorder

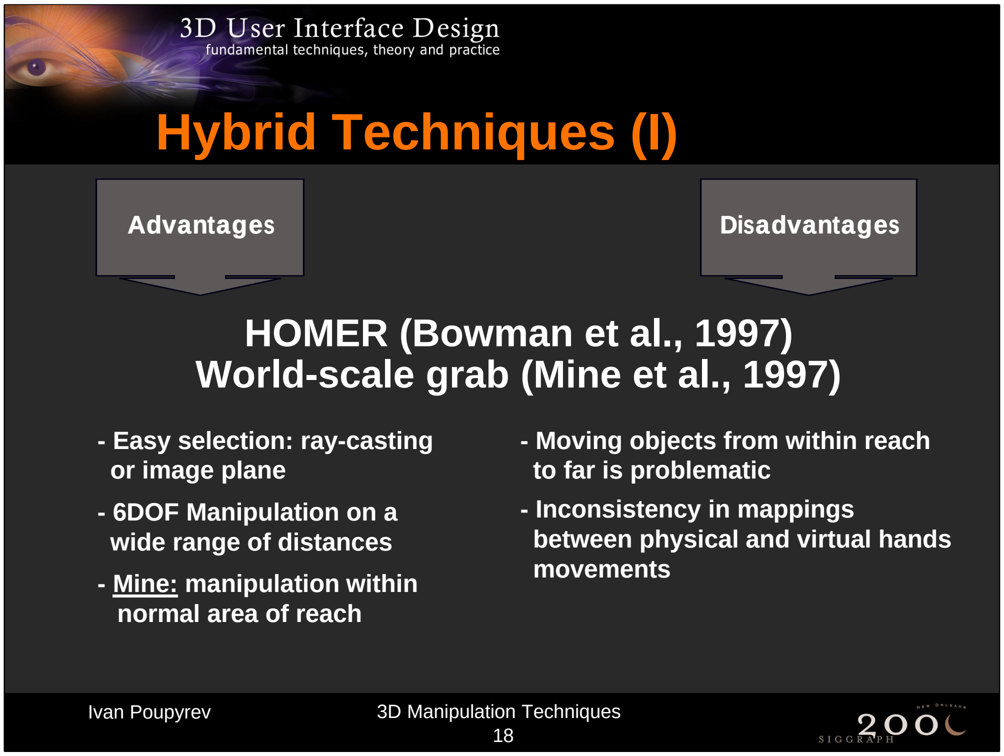
Ivan Poupyrev | 3D Manipulation Techniques 17 | SIGGRAPH 2000

Techniques combination

None of the interaction techniques are effective in every imaginable manipulation scenario. Hence, there have been a few attempts at combining interaction techniques.

Two basic approaches have been investigated. The first is a simple aggregation of techniques, where the user is provided with an explicit mechanism for choosing the desired technique. Examples include the techniques tool belt (Mine et al., 1997) or a universal virtual controller, the Virtual Tricorder, a term first suggested by Henry Sowizral.

A second approach is to attempt a seamless combination of interaction techniques , in which the interfaces switches transparently between the interaction techniques depending on the current context of the task. The idea is to provide the user with the best manipulation technique at any given moment of manipulation. Three manipulation techniques have been reported here: HOMER technique (Bowman et al, 1997), world-scale grab (Mine et al., 1997) and Voodoo Dolls (Pierce et al., 1999)



HOMER and world-scale grab

Both techniques are based on the simple observation that selection and manipulation are essentially sequential tasks; indeed, an object has to be selected before it can be manipulated. Hence, selection and manipulation can be completed using different techniques: the interface simply switches from the selection to the manipulation technique after an object is selected by the user, and switches back to the selection mode after the user releases the manipulated object. Theoretically, the techniques can be optimized to achieve an overall best performance in each mode.

HOMER

HOMER (Bowman et al., 1997) stands for Hand-centered Object Manipulation Extending Ray-casting technique. The user selects an object using ray-casting and after that, instead having the object attached to the ray, the user's virtual hand instantly moves to the object and attaches to it. The technique switches to the manipulation mode in which the technique interactively amplifies the user physical reach, i.e., the user-to-hand distance, with coefficient *K*, which is defined as a the distance to the virtual object devide by the distance to the real hand at the moment when the virtual object is selected. The rotations of the object are controlled independently.

With HOMER, the user can easily re-position a virtual object within the area between a virtual object and the user, no matter how far the object is at the moment of selection. The user can also push an object further away from its initial distance but the distance where the user can reposition it depends on ratio of distance to the virtual object at the moment of selection to how far the user's physical hand is stretched during selection.

Continued on the next page

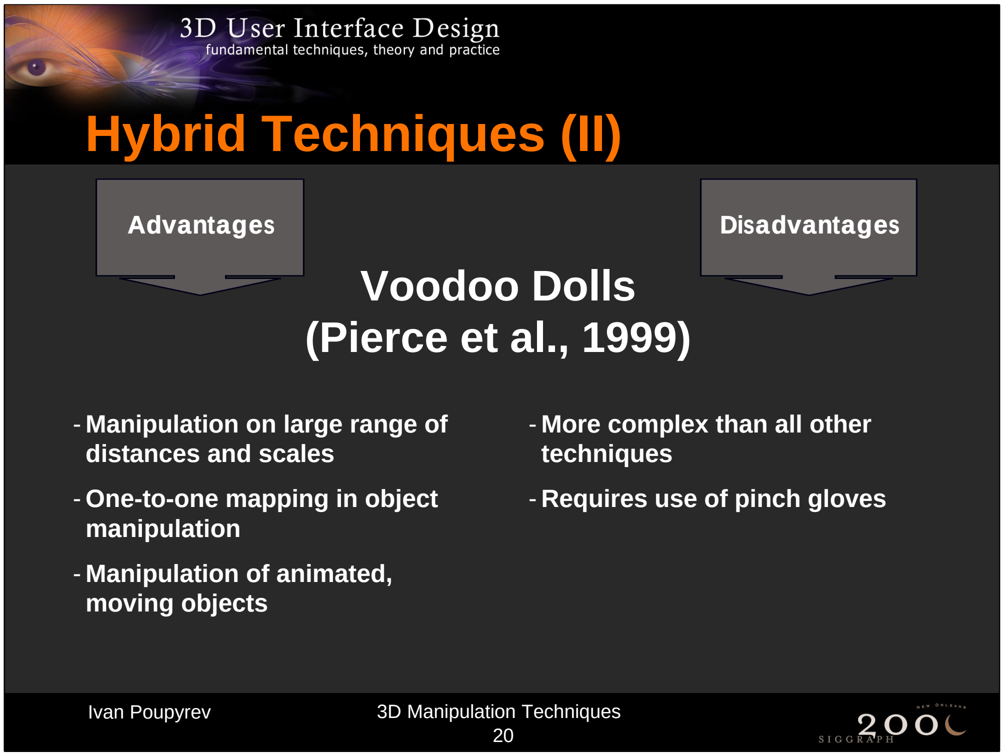
Scaled-world grab

A technique related to HOMER is the scaled-world grab (Mine et al., 1997) by which the user first selects an object using an image-plane technique (Pierce et al., 1997). After an object is selected, instead of scaling up the user's reach as in HOMER, the technique scales down the whole virtual environment around the user's virtual viewpoint, so that objects are brought within the user reach and are manipulated using the simple virtual hand. The scaling coefficient is calculated so that the *visual size* of objects in the environment remains unchanged. As a result, the user does not even notice that scaling actually takes place, since the world does not change visually.

Design issues in HOMER and Scaled-world grab

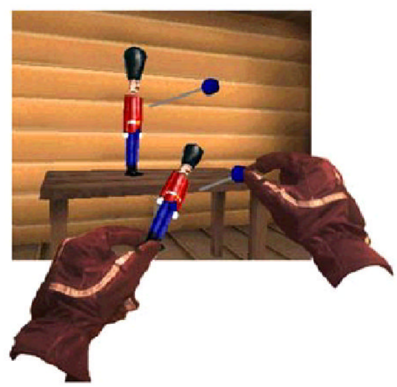
The scaling coefficient for both techniques is not constant but varies depending on the distance to a virtual object at the moment of selection: the further an object is when the user picks it up, the larger the coefficient of amplification is and, therefore, the larger the range of manipulation. While this quality allows a user to access and reposition objects at any distance from him/herself, there are certain problems. First, a problem happens when the user wants to perform an operation opposite to the manipulation at-a-distance, i.e., when instead of manipulating an object located far away, the user wants to pick up an object located *within* the area of user reach, close to the his or her virtual body, and move it far away, at-a-distance. In this case, the coefficient of scaling becomes close to one and it becomes difficult to move an object away.

The other limitation to these techniques is that none of them preserve the consistent kinesthetic correspondence between the user's physical hand movements and virtual hand movements in the virtual space. Indeed, every time the user selects a different object the scale coefficient changes and the same displacement of the hand results in a different displacement of the virtual hand. The inconsistent visual feedback in motor movements usually results in a decrease in the operator performance, since the user can not effectively built a "kinesthetic model" of the hand motion (Smith and Smith, 1987). Whether this will result in a decrease in user manipulation performance in some conditions of virtual manipulation remains to be seen.



Voodoo Dolls

The Voodoo Dolls technique (Pierce, et al. 1999) is a two-handed interaction technique for manipulating objects at a distance in immersive virtual environments. The technique combines and builds upon a number of other techniques, such as Image Plane (Pierce et al., 1997) and WIM (Stoakley et al., 1995). Voodoo Dolls uses a couple of pinch gloves to allow the user to switch seamlessly between different modes of manipulation. It aims to provide an easy method of interacting with objects of widely varying sizes and at different distances.



The technique is based on several ideas. First, to start object manipulation the user dynamically creates dolls: temporary, miniature, hand-held copies of objects. Similar to the WIM technique, the user can interact with objects in the environment by manipulating these dolls instead of directly manipulating the objects so that manipulated virtual objects can be at any distance, size and state of occlusion.

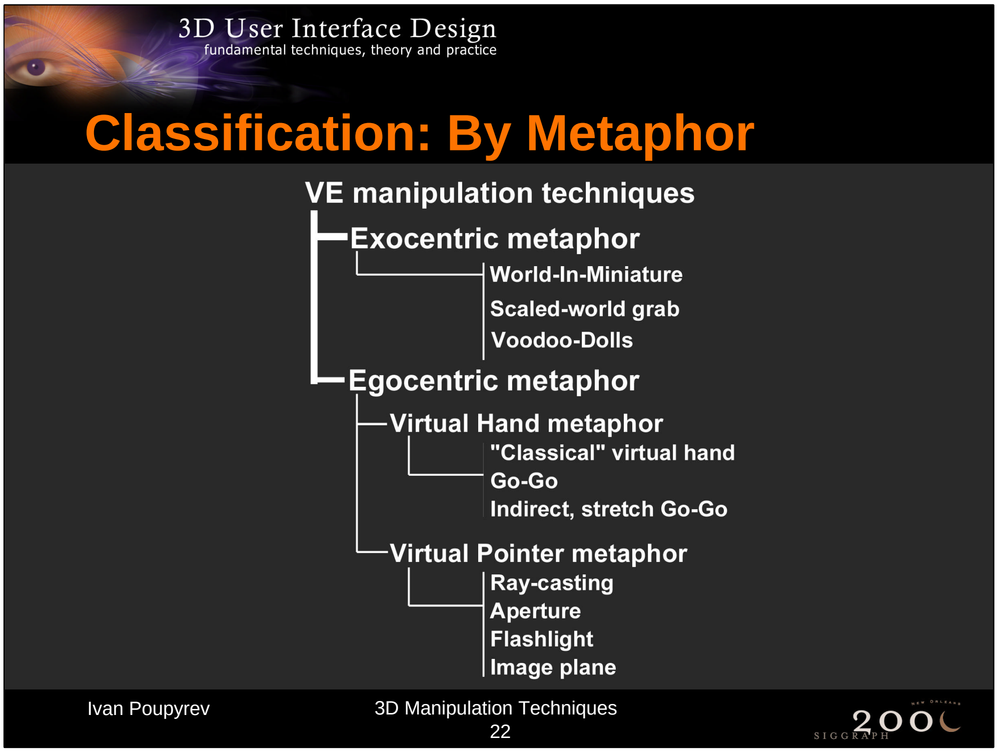
Second, the technique allows the user to explicitly and interactively specify a frame of reference for manipulation. The doll that the user holds in the non-dominant hand represents a stationary frame of reference, and the corresponding virtual object does not move when the user moves this doll.

Continued on the next page

The doll that the user holds in the dominant hand represents a manipulated object; as the user manipulates this doll relative to the doll in the other hand, the corresponding virtual object moves to the same position and orientation relative to the virtual object represented by the doll in non-dominant hand (see figure above). Since the spatial frame of reference for manipulation is defined explicitly by the user, Voodoo Doll scales down/up dolls to a convenient size, thus overcoming one of the limitations of the WIM technique.

Finally, the Voodoo Doll technique also separates the selection mode from the manipulation mode: the user first selects an objects by using an Image Plane technique, then the technique switches to the manipulation mode, and switches back to the selection mode when the dolls are released. The technique also provides a number of heuristics that allow the user to select context of different sizes, varying from a single object to whole parts of the environment.

Voodoo Dolls allows the user to achieve some interesting effects that are difficult to achieve with other techniques. For example, the user can easily manipulate parts of animated objects that are freely moving in the environment. This sophistication might also be a source of complexity and difficulties for the user since it takes some time to understand and master this technique. It also requires the use of pinch gloves, which could be yet another constraint in designing the manipulation interface.



Classification of manipulation techniques

While the variety of the reported techniques can be overwhelming for the developer, many techniques are apparently related to each other and share many common properties. For example, there are more similarities between ray-casting and flashlight techniques than there are between ray-casting and the Go-Go. A taxonomy of techniques, that is, a classification according to their common properties, can be instrumental in understanding relations between techniques and can help in designing effective manipulation interfaces.

Metaphor-based taxonomy of techniques

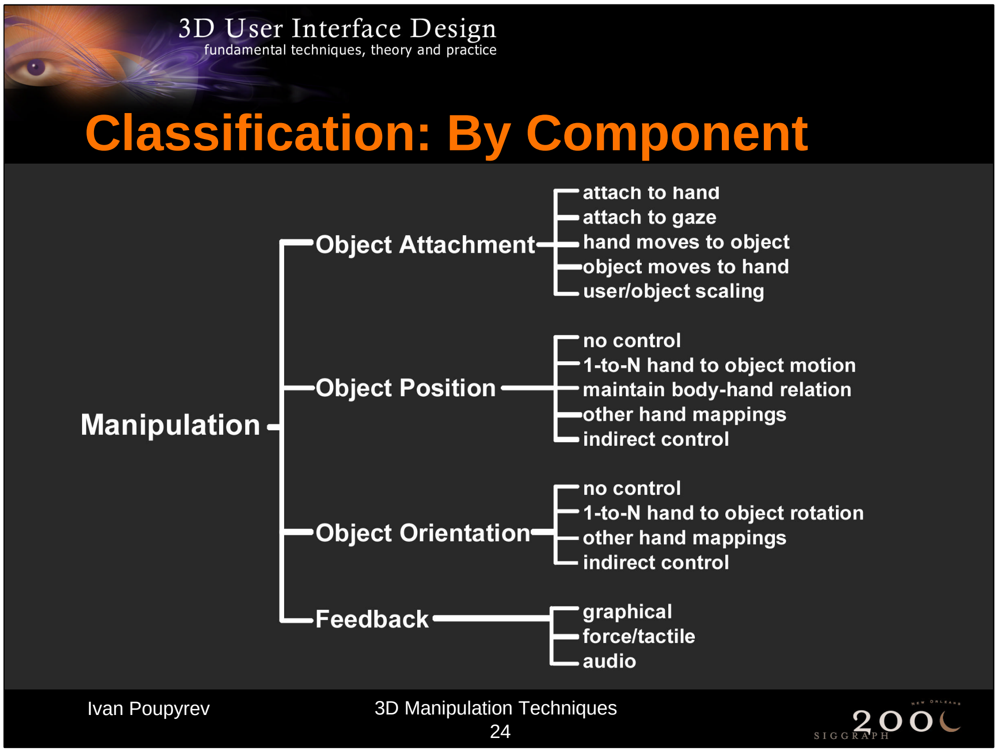
An analysis of current VE manipulation techniques suggests that most are based on a few interaction metaphors or their combinations (Poupyrev et al., 1998). Each of these metaphors forms the fundamental mental model of a technique -- a perceptual manifestation of what users can do (affordances), and what they cannot do (constraints) by using the technique. Particular techniques, therefore, can be considered as implementations of these basic metaphors.

This slide presents a simple metaphor-based taxonomy of current VE manipulation techniques, categorized first into *exocentric* or *egocentric* techniques. These two terms originated in the studies of cockpit displays , and are now used to distinguish between two fundamental styles of interaction within VEs. In exocentric interaction, also known as the God's eye viewpoint, users interact with VEs from the outside (the outside-in world referenced display); an example is the World-In-Miniature technique. In egocentric interaction, which is the most common in immersive VEs, the user interacts from inside the environment, i.e., the VE embeds the user.

Continued on the next page

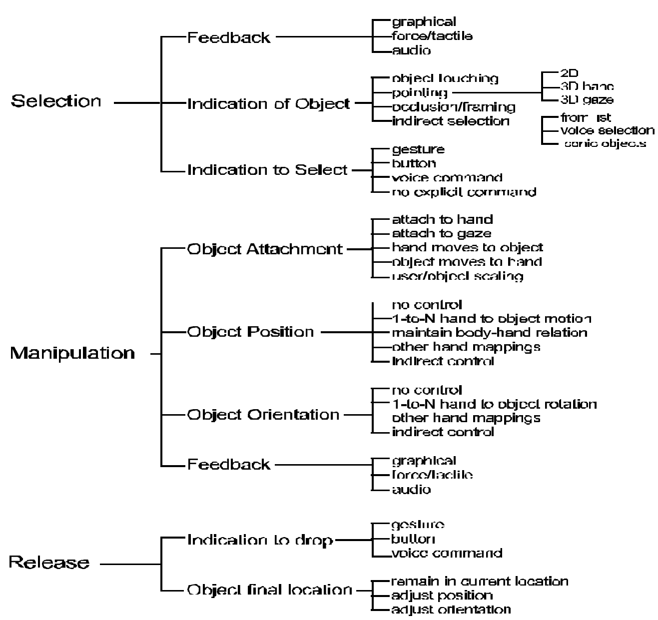
There are currently two basic metaphors for egocentric manipulation: virtual hand and virtual pointer. With the techniques based on the virtual hand metaphor, users reach and grab objects by “touching” and “picking” them with a virtual hand. The choice of the mapping function discriminates techniques based on the virtual hand metaphor.

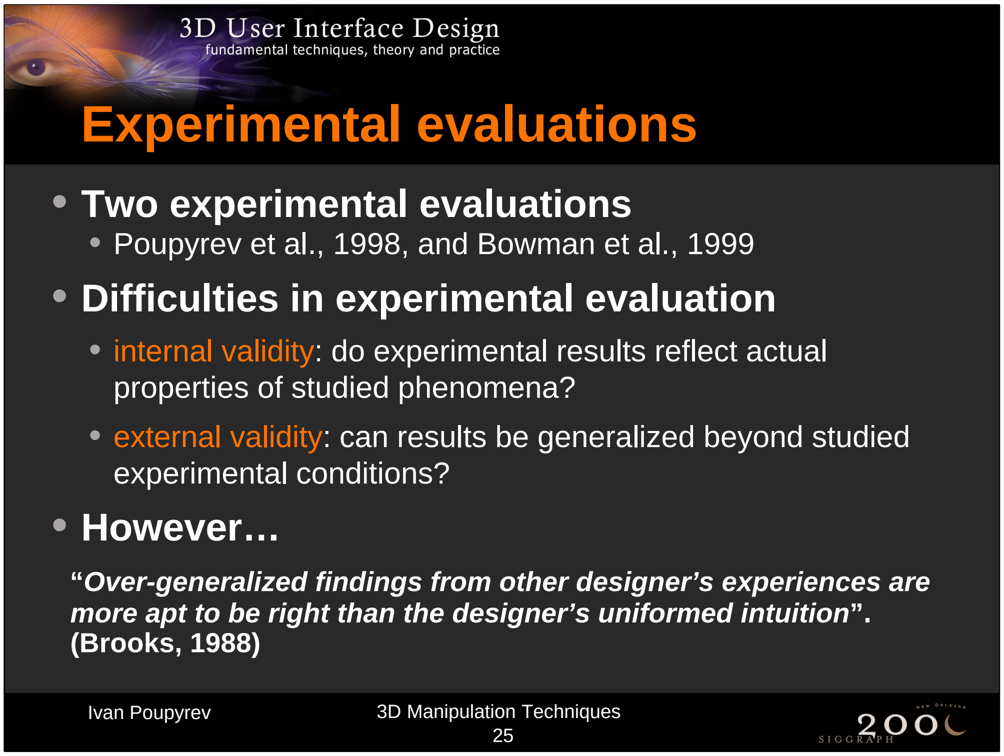
With techniques based on the virtual pointer metaphor, the user interacts with objects by pointing at them. The selection volume (i.e. conic or ray), definition of ray direction, and disambiguation function are some of the design parameters for such techniques.



Component-based taxonomy of techniques

The component-based taxonomy is based on the observation that all techniques consist of several basic components with similar purposes (Bowman, 1999). For example, in a manipulation task, the object can be positioned, rotated, and the feedback should be provided, etc. Then, instead of classifying the techniques as a whole, we only classify components that accomplish these subtasks within techniques. As a result, theoretically, any technique can be constructed out of these components simply by picking the appropriate components and putting them together.





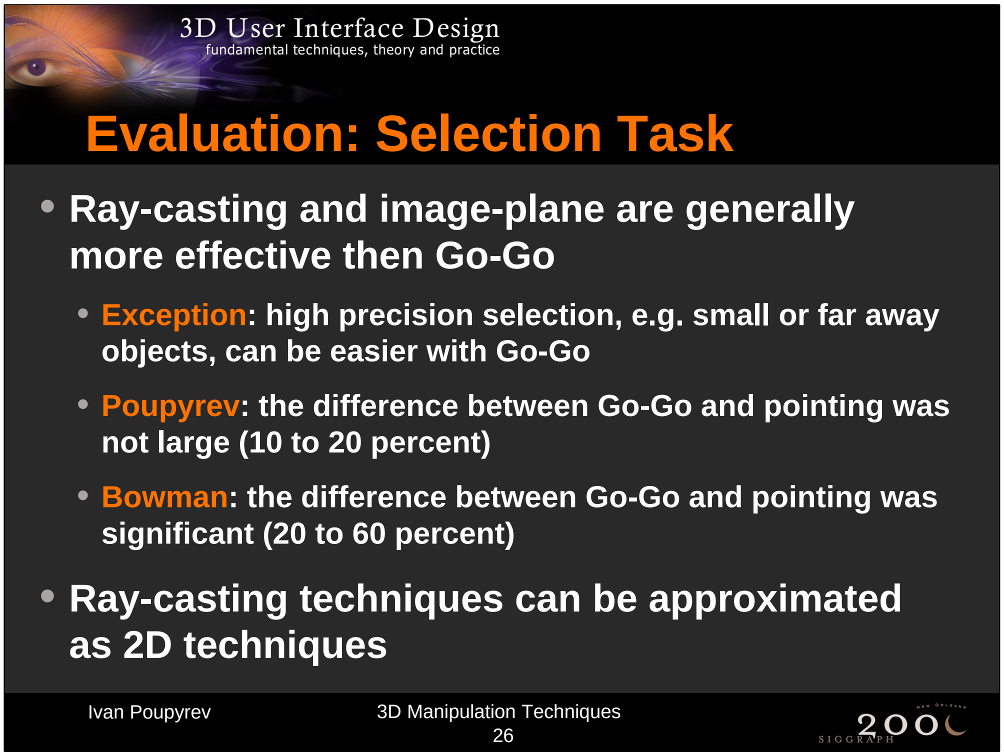
Experimental evaluations of manipulation techniques

There have been only a few formal studies that have attempted to compare and evaluate interaction techniques for 3D manipulation: Poupyrev et al. (1998) and Bowman et al. (1999). Bowman also has reported an informal evaluation in 1997. Both studies have just barely scratched the surface and there is a lot left to be done. Conducting experimental evaluations of techniques, however, is a challenging task because of the difficulties in ensuring the internal and external validity of experiments:

Internal validity means that the experimental results reflect the actual properties of the investigated phenomena. For example, in experiments comparing different methods of object selection, a slight modification of the interaction technique which might go unnoticed by the experimenter, e.g. changing the length of the virtual ray in the ray-casting technique, can alter user performance. Such modification makes it difficult to understand whether the performance differences are the result of inherited properties of the technique or a side effect of the implementation peculiarities. A large variety of other factors can confound the results, such as skill transfer, order effect, etc.

External validity means that the experimental results can be generalized beyond the conditions of particular experiments. Suppose, for example, that the study of variable A finds that user performance does not change significantly for two levels of A: A_1 and A_2. However, concluding that variable A does not effect the user performance may not be a proper generalization. Indeed, it simply might be that the difference between A_1 and A_2 is not large enough to detect the influence of A. If experiments demonstrate that technique T_1 is significantly better then technique T_2, can we be sure that this would be true for every possible variation of the experimental task? In which task conditions this result is true and in which conditions it is not applicable? These questions are not easy to answer, so this limits usefulness of the techniques experimental evaluations.

However, even though the results of experiments are often prone to problems and multiple interpretations, their importance can not be argued. As Prof. Brooks noticed once: "*Over-generalized findings from other designer's experiences are more apt to be right than the designer's uniformed intuition*" (Brooks, 1988)



Interaction techniques in selection task

This slide presents some of the results from the experiments conducted by Poupyrev et al. (1998) and Bowman et al. (1999). Poupyrev et al. evaluated two interaction techniques: ray-casting and Go-Go interaction technique with 5 distance levels and 3 object size levels; within subject design and 12 participants were in the experiments.

Bowman evaluated three interaction techniques: the ray-casting, image plane, and Go-Go techniques with 3 distance levels and 2 object sizes levels; between subject design with 5 participants for each technique was used in the experiments.

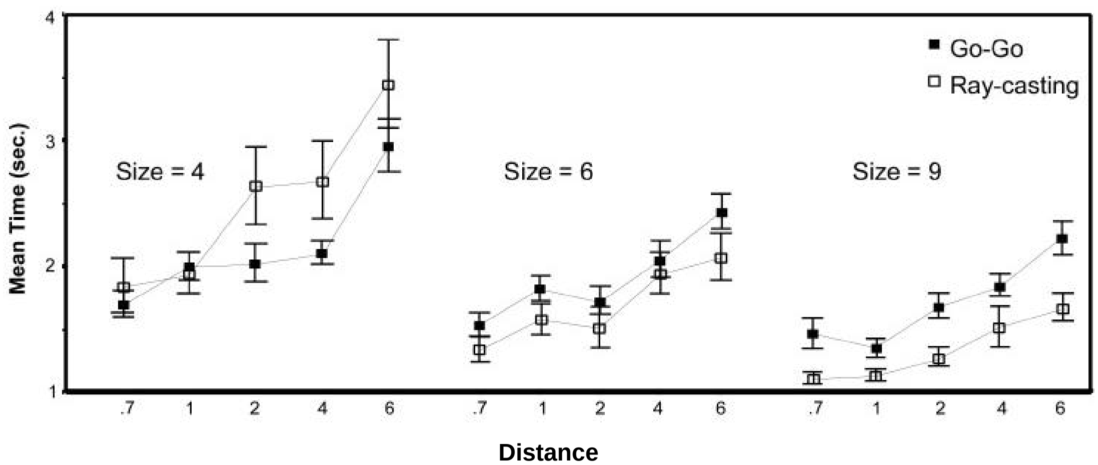
The experiments have shown that in selection task pointing techniques generally demonstrated better performance then Go-Go since they required less physical movement from the user. The only exceptions reported by Poupyrev were when a very high precision of selection was required, e.g. for small objects (about 4 degrees of field of view) and those objects located far away. In this case pointing, was not as effective as Go-Go, which is an intuitive observation: pointing at small objects is especially difficult when they are further away. This replicates the observations of other researchers (Liang, 1994; Forsberg, 1996); the Fitt's law provides a theoretical explanation for this phenomena. In Bowman's experiments, however, the size and the distance did not significantly affect the user performance of pointing techniques.

There were differences in user performance in these two experiments. This was because slightly different implementations of the interaction techniques were used in each experiments. For example, in Poupyrev experiments, the ray-casting was implemented as a short ray by which the user pointed at objects. Bowman, on the other hand, implemented ray-casting with infinite ray so the user could actually use it to touch the object, which might be a more effective implementation. Furthermore, different settings in the Go-Go technique can also greatly affect user performance.

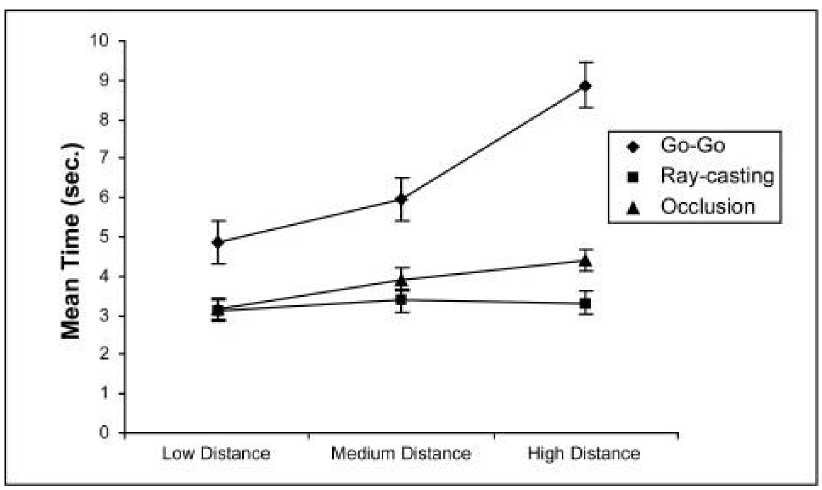
Continued on the next page

For example, in the Poupyrev experiments, the non-linear threshold was re-adjusted individually for each user, while in the Bowman experiments it was not.

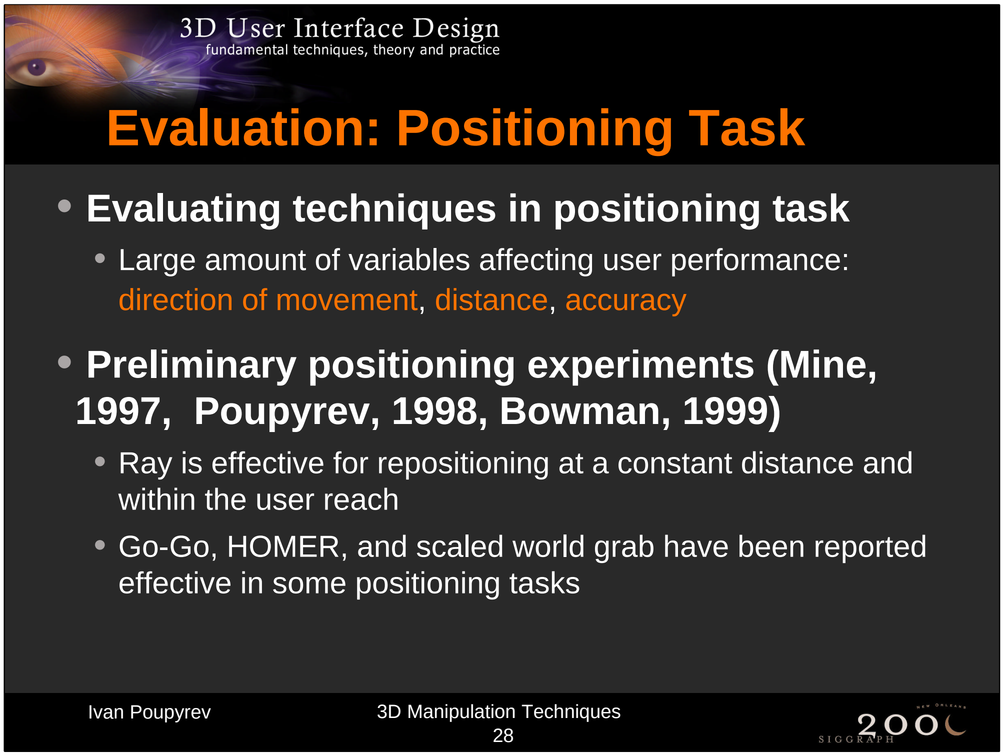
Poupyrev et al., (1998) suggested that all ray-casting techniques, e.g., virtual ray, image-plane and other techniques, can be approximated as 2D techniques (see also a taxonomy introduced earlier). Bowman partially supported this observation, who found no performance difference between ray-casting and image-plane.



Experimental results comparing Go-Go and Ray-casting techniques (Poupyrev et al., 1998)



Experimental results comparing Go-Go, ray-casting and occlusions techniques (Bowman et al., 1999)



Interaction techniques in positioning tasks

Evaluating techniques for positioning tasks is difficult because of the vast amount of variables affecting user performance during the manipulation task: the user performance can differ drastically depending of the direction of positioning (e.g., inward or outward, left to right, constant distances from the user or changes in distance, etc.), travel distance, required accuracy of manipulation as well as degrees of freedom. One of the reasons behind these differences is the use of different groups of muscles that are active in different task conditions.

Several evaluations, exploratory in nature, have been reported (Mine et al., 1997, Poupyrev, et al. 1998, Bowman et al. 1999). For example, although the ray-casting was found to be not very efficient for positioning with a change in distance from the user, it can be very efficient when the starting and target positions of the object are located at a constant distance.

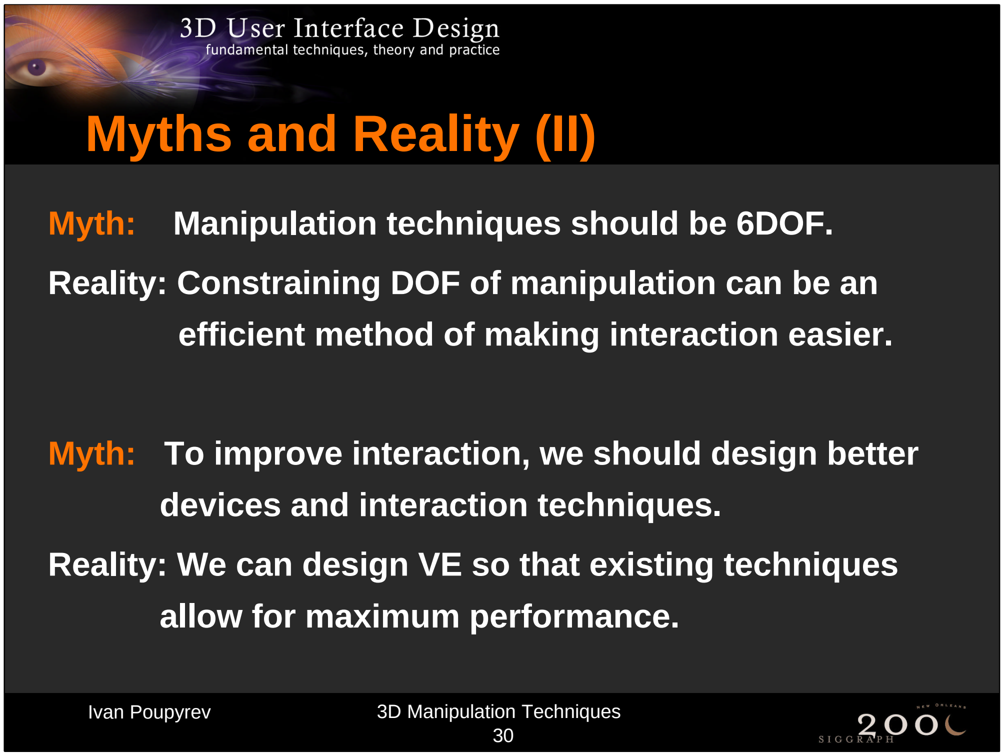
The Go-Go, HOMER, and scaled-world grab (Mine, 1998) have been found effective in certain manipulation task conditions, however, it is difficult to compare them because these experiments evaluated techniques in a few conditions and results were somewhat inconclusive. For example, HOMER and scale-world grab have not been evaluated in outward positioning of objects, i.e when the user wants to pick an object that is already located within reach and move it away at-a-distance. These are conditions that are unfavorable for both techniques and it is not clear what is be the user performance in these conditions. Significantly more systematic experimental evaluations of techniques are needed to understand their mutual performance characteristics.

Myths and realities of techniques design

Manipulation techniques have been actively researched, and a large variety of techniques have been created and reported. There are several conclusions that we can make from an analysis of these techniques.

First, the early notion that interaction in VE should be an exact imitation of our interaction in the real world does not hold. In fact, most manipulation techniques depart from real world interaction to a greater or lesser degree by allowing “magical” interactions with virtual environments.

Second, the single best interaction technique has not been designed yet. Each technique has its weaknesses and strengths. I believe that each technique has its niche and those particular task conditions in which it would result in better user performance than other techniques.



Third, manipulation techniques do not necessarily need to be 6DOF ones. In fact, some very effective selection techniques, such as Image Plane techniques, provide very effective interaction by restricting degrees of freedom and essentially making selection a 2D task.

Finally, there is no one universal interaction technique, so the design of the manipulation interfaces in VE is a tradeoff between tasks that need to be performed and the affordances of the techniques. Two strategies should be applied in designing manipulation interfaces: tuning techniques so that they maximize user performance in the target task conditions, and at the same time designing virtual environments so that these techniques can be used in the best possible way.

3D User Interface Design
fundamental techniques, theory and practice

Conclusions and further work

- **Further evaluations of techniques**
- **Techniques for new input/output devices**
- **Re-evaluation of today's techniques for new input/output devices**
- **Techniques integration**
- **Multiple modalities**
- **Techniques for complex tasks**
- **Standards and guidelines**

Ivan Poupyrev 3D Manipulation Techniques 31 SIGGRAPH 2000

Bibliography

For most of the references see the "20 Century 3DUI Bib". Below are only those references that are not in the main bibliography.

Gilbreth, F. and Gilbreth, L. (1917). Applied motion study. New York: MacMillian.

Jenkins, W. and Connor, M., Some design factors in making settings on a linear scale. Journal of Applied Psychology, 1949. 33(4): pp. 395-409.

McCormick, E. (1970). Human factors engineering. (3 ed.): McGraw-Hill.

Mundel, M. (1978). Motion and time study. (5 ed.). Englewood Cliffs, NJ: Prentice-Hall.

Philips, M., Bashinski, H., Ammerman, H. (1988). A task analytic approach to dialogue design. In M. Helander (Ed.), Handbook of Human-Computer Interaction (pp. 835-857): Elsevir Science Publisher B.V.

Smith, T. and Smith, K., Feedback-control mechanisms of human behavior. In Handbook of human factors, G. Salvendy, Editor. 1987, John Wiley and Sons. pp. 251-293.