CHI 2008
art.science.balance.

3D UIs in a nutshell

Doug Bowman

The goal of this lecture is to provide a foundation for the rest of the course. It will provide a whirlwind overview of research on 3D UIs to date, using our book 3D User Interfaces: Theory and Practice as a guide. Given the limited time, we'll just present a few highlights, so that those not familiar with 3D UIs can understand the topics and issues presented in the rest of the course.

What are 3D UIs?

3D UIs | displays | input devices | basic techniques | evaluation | open questions

CHI 2008
art.science.balance.

3D interaction: Human-computer interaction in which the user's tasks are carried out in a 3D spatial context
  ► 3D input devices
  ► 2D input devices with direct mappings to 3D

3D user interface (3D UI): A UI that involves 3D interaction

3D interaction technique: A method (hardware and software) allowing a user to accomplish a task in a 3D UI

doug bowman | 3D UIs in a nutshell   3

Our definitions of 3D UI and related terms.

Why 3D Interaction?

3D UIs | displays | input devices | basic techniques | evaluation | open questions

CHI 2008
art.science.balance.

3D applications should be useful
- ▶ immersion
- ▶ natural skills
- ▶ immediacy of visualization

But, applications in common use have low complexity of interaction

More complex applications have serious usability problems

Technology alone is not the solution!

doug bowman | 3D UIs in a nutshell   4

Three dimensions and virtual environments intuitively make sense for a wide range of applications, because of the characteristics of the tasks and their match with the characteristics of these environments. Immersion is the feeling of "being there" (replacing the physical environment with the virtual one), which makes sense for applications such as training and simulation. If a user is immersed *and* can interact using natural skills, then the application can take advantage of the fact that the user already has a great deal of knowledge about the world. The immediacy characteristic refers to the fact that there is a short "distance" between a user's action and the system's feedback that shows the result of that action. This can allow users to build up complex mental models of how a simulation works, for example.

Most applications in common use (e.g. walkthroughs, psychiatric treatment, entertainment, and training) contain user interaction which is not very complex. Other types of applications (e.g. immersive design, education, complex scientific visualizations) are for the most part still stuck in the research lab, often because they have usability problems that limit their usefulness.

Better technology is not the only answer - 30 years of VE technology research have not ensured that today's VEs are usable - we must also focus on the design of interaction for VEs.

Therefore, we feel that 3D interaction is a vital topic for all 3D/VE developers, designers, and evaluators to understand.

Goals of interaction design

3D User Interfaces: From Lab to Living Room

CHI 2008
art.science.balance.

3D UIs | displays | input devices | basic techniques | evaluation | open questions

Performance
- ▶ efficiency
- ▶ accuracy
- ▶ productivity

Usability
- ▶ ease of use
- ▶ ease of learning
- ▶ user comfort

Usefulness
- ▶ interaction helps meet system goals
- ▶ interface relatively transparent so users can focus on tasks

doug bowman | 3D UIs in a nutshell   5

We will try to keep in mind as we discuss ways to accomplish 3D interaction tasks that we want to design for performance, usability, and usefulness.

Performance relates to quantitative measures indicating how well the task is being done by the user and the system in cooperation. This includes standard metrics like efficiency and accuracy.

Usability refers to the ease of communicating the user's intentions to the system, and the qualitative experience of the user.

Usefulness implies that the system is actually helping the user perform work or meet his/her goals, without being hindered by the interface.

All three of these goals must be considered together, as all are essential. A system will not be used if users become frustrated after five minutes of usage (usability) even if it's been shown to aid the user in getting work done in a new way. A business will not adopt a system that is incredibly easy to use but decreases productivity (performance).

Display devices for 3D UIs

Some characteristics of visual displays

3D UIs | **displays** | input devices | basic techniques | evaluation | open questions

CHI 2008
art.science.balance.

field of view (FOV)—the size of the visual field (in degrees of visual angle) that can be viewed instantaneously

field of regard (FOR)—the total size of the visual field (in degrees of visual angle) surrounding the user

display size

display resolution

stereoscopy

refresh rate

doug bowman | 3D UIs in a nutshell   7

Although this is not an exhaustive list, it gives a sense of the ways that visual displays for 3D UIs can be characterized. It also provides a more or less standardized way to compare visual displays that are very different.

Visual displays for 3D UIs

3D UIs | **displays** | input devices | basic techniques | evaluation | open questions

CHI 2008
art.science.balance.

Standard monitor (mono/stereo)
Head-mounted/head-referenced
Projected (usually stereo)
   ▶ single-screen
   ▶ multiple, surrounding screens
Retinal display
Volumetric displays

doug bowman | 3D UIs in a nutshell   8

We'll summarize the pros and cons of a few of the more common and/or interesting visual displays for 3D UIs.
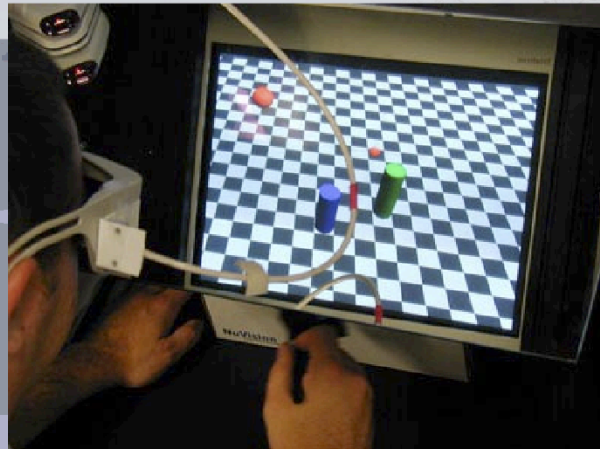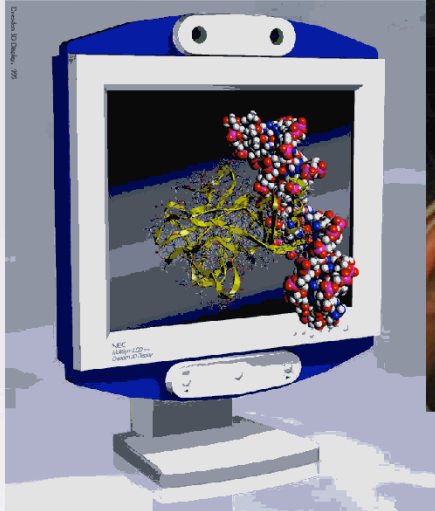
3D UIs on the desktop are easier to achieve now than ever before. There are commercially-available autostereoscopic displays, making 3D viewing without glasses feasible. Adding a head tracker produces so-called "fishtank VR," and a handheld tracking device allows 3D input as well.

Head-mounted displays (HMDs)

CHI 2008
art.science.balance.

3D UIs | **displays** | input devices | basic techniques | evaluation | open questions

+ full surround (FOR=360)
+ simple stereo

- cumbersome
- small FOV
- no peripheral vision
- single user

doug bowman | 3D UIs in a nutshell   10

One of the most common display devices used for virtual environment applications is the head mounted display (HMD). With a tracking device attached to the device, it produces a stereoscopic view that moves relative to the user's head position and orientation. Although traditionally the user cannot naturally see the real world, cameras are sometimes mounted on the HMD which allows it to display both real world video and graphical objects. In addition, some HMDs offer see-through options. This type of technology is used in augmented reality systems.

Since each eye is presented with one screen, HMDs allow for good stereoscopic viewing. These two screens are very close to the user's eyes (1 to 2 inches). As a result, all viewable objects are behind the screen so any object clipping will appear to the user as being outside his/her field of view. A big disadvantage of HMDs is that can get heavy very quickly and, unfortunately, the higher the HMD's quality, the heavier it usually is. Although HMDs are still used in many VR labs and entertainment centers, researchers and practitioners are rapidly moving towards projection-based display devices especially when high-resolution graphics are required.
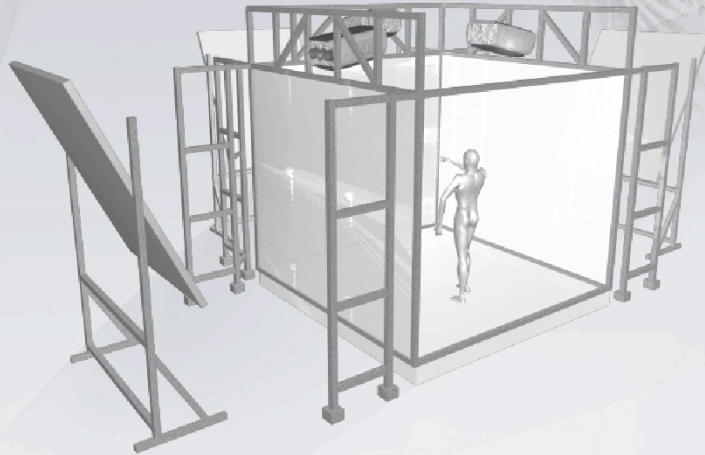
Recently a high-resolution and wide FOV HMD came onto the market (www.sensics.com). It remains to be seen whether this will cause some high-end applications to return to HMDs.

Surround-screen displays, such as the CAVE™ are also extremely popular. Instead of attaching the displays to the user, they place the displays in the world. Such displays are typically rear-projected, stereoscopic, and head tracked. They range from two-screen L-shaped configurations to semi-cylindrical displays to spherical displays.

Six-sided CAVE

CHI 2008
art.science.balance.

3D UIs | **displays** | input devices | basic techniques | evaluation | open questions

DiVE at Duke University

doug bowman | 3D UIs in a nutshell   12

Traditionally, HMDs have one big advantage over surround-screen displays - a 360-degree field of regard (i.e., the graphics appear around the user in every direction). But this advantage was eliminated with the advent of fully-surrounding surround-screen displays, such as the six-sided DiVE at Duke University.

These devices are usually single screen and go by many different names such as Fakespace's Immersadesk and Immersive WorkBench and VersaBench (pictured on the left), the Barco Baron, and the ITI VisionMaker Digital Desk. In some cases just a single vertical screen is used. The second picture to the left shows Fakespace's Mini Workbench. A pressure sensitive display surface for 2D input is an optional feature with the Workbench. The TAN Holobench, shown in the two pictures on the right, is an L-shaped desk which provides a holographic impression to the user since objects appear to be raised above the it.

In general, workbenches provide high resolution displays, make for an intuitive display for certain types of applications (i.e 3D modeling and drafting, virtual surgery), and can be shared by several users. However, due to technological limitations, at most two users can be head tracked and these devices suffer from the same stereo problems that all rear-projected devices do.

The Virtual Retinal Display (VRD) was invented at the Human Interface Technology Lab in 1991. It is based on the idea that images can be directly displayed onto the retina. With a VRD, a photon source is used to generate a coherent beam of light which allows the system to draw a diffraction limited spot on the retina. The light beam is intensity modulated to match the intensity of the image being rendered. The beam is then scanned to place each image point, or pixel at the proper position of the retina. VRDs are commercially available from Microvision and are used in augmented reality systems.

Volumetric displays produce a "truly 3D" image by actually illuminating locations in physical 3D space. The display shown here, from Actuality Systems, uses a rotating transparent display enclosed in a glass dome.

These displays solve a problem common to all other 3D display types - the accommodation-convergence mismatch. Accommodation is an oculomotor depth cue based on the depth of focus of the eye, while convergence, also an oculomotor cue, is based on the rotation of the eyes to look at a single object. In 3D displays that project stereoscopic images on a flat screen, accommodation and convergence are always in conflict (unless the object is at the depth of the screen). Volumetric displays provide correct accommodation and convergence cues.

3D auditory displays

CHI 2008
art.science.balance.

3D UIs | **displays** | input devices | basic techniques | evaluation | open questions

Technologies:
- ▶ Speaker-based
- ▶ Headphone-based

Uses:
- ▶ Virtual objects emitting sound
- ▶ Sensory substitution

doug bowman | 3D UIs in a nutshell   16

There are a number of different ways in which a 3D auditory system can be set up. A simple setup is to use stereo head phones. However, this restricts usage to only one person at a time. Another setup is to place speakers in certain logistic areas around the environment. This setup allows for more than one user to take part in the experience but is somewhat more complicated to setup and write software for.

There are two different ways, localization and sonification, in which sound can be used as an output medium in virtual environment applications. In localization, the goal is to generate three dimensional sound. In sonification, the goal is to turn certain types of information into sounds.

Haptics represents a critical component in virtual environment interaction. Allowing a user to touch and feel in the virtual world in the same way that they do in the physical world is extremely powerful. Unfortunately, haptic and tactile output device research has not made rapid progress.

There are essentially four different methods in which haptic and tactile feedback is generated. The first method is ground-referenced feedback which creates a physical link between the user and ground with the feedback relative to a single contact point. An example is the Sensable Phantom. The second method is body-referenced feedback which places a device on some part of the user's body. An example of a body-referenced haptic device is Virtual Technologies' CyberGrasp which is shown in the top picture. The third method for generating feedback is tactile which uses some type of oscillatory or vibrating device to stimulate the user's tactile sense. Finally, the last method of generating feedback is via dermal tactile which stimulates the user's nerves in the fingertips.

References:

www.sensable.com

www.immersion.com

A simpler way to provide haptic feedback is the use of props - physical objects that represent their virtual counterparts. This is also called "near-field haptics" or "passive haptics." This has been an extremely important idea historically in 3D UIs.

Hinckley, K., Pausch, R., Goble, J. and Kassell, N., Passive Real-World Interface Props for Neurosurgical Visualization. in *CHI: Human Factors in Computing Systems*, (1994), 452-458.

Schell, J. and Shochet, J. Designing Interactive Theme Park Rides. *IEEE Computer Graphics & Applications*, *21* (4). 11-13.

Input devices for 3D UIs

Input device characteristics

CHI 2008
art.science.balance.

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions

Degrees of Freedom (DOFs) & *DOF composition* (integral vs. separable)

Type of electronics: Digital vs. analog

Range of reported values: discrete/continuous/hybrid

Data type of reported values: Boolean vs. integer vs. floating point

User action required: active/passive/hybrid

Method of providing information: "push" vs. "pull"

Intended use: locator, valuator, choice, …

Frame of reference: relative vs. absolute

Properties sensed: position, motion, force, …

doug bowman | 3D UIs in a nutshell   20

There are *many* different ways to characterize input devices to be used in 3D UIs, some of which are shown here. In the 3D UI community, researchers often focus on degrees of freedom. But other characteristics can also be important. For example, a typical position tracker provides absolute position information. A Wii remote, which some have seen as a replacement for position trackers, provides relative position information. This difference completely changes the way these devices are used in 3D interaction techniques.

Practical classification system

CHI 2008
art.science.balance.

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions

Desktop devices
Tracking devices
3D mice
Special-purpose devices
Direct human input

For simplicity, in this lecture, we use a more practical classification system for 3D input devices.

Desktop devices: 6-DOF devices

CHI 2008
art.science.balance.

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions

6 DOFs without tracking

Often isometric

Examples:
- ▶ SpaceBall
- ▶ SpaceMouse
- ▶ SpaceOrb

doug bowman | 3D UIs in a nutshell   22

In the category of desktop devices, the most popular 3D input devices are those that provide six degrees of freedom, such as the SpaceMouse shown here. It allows the user to push/pull/twist the device to specify 3D translation and rotation directly.
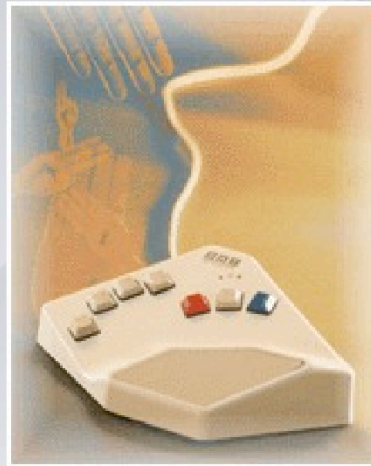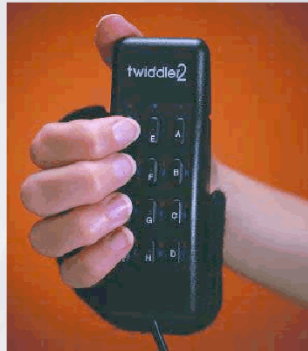
Keyboard input (for text or numeric entry) is not often needed in 3D UIs, but when it is, traditional keyboards are often not practical to use. Thus, 3D UIs often make use of handheld or wearable keyboards, that may use chords instead of individual button presses since they have fewer physical buttons. Soft keyboards, such as those on a TabletPC, may also be used.

Tracking devices: position trackers

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions

Measure position and/or orientation of a sensor

Degrees of freedom (DOFs)

Most VEs track the head
- ► motion parallax
- ► natural viewing

Track hands, feet, etc.
- ► "whole body" interaction
- ► motion capture application

Correspondence between physical/virtual objects
- ► Props
- ► spatial input devices

Position trackers are on of the most fundamental input devices for 3D UIs. In VEs, they are most often used to track the head and hand(s). But they can also be used to track physical objects that are used as props or spatial input devices (e.g., a physical paintbrush used to paint virtual objects).

Hybrid tracking

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions

Intersense IS-600 / 900
inertial (orientation)
acoustic (position)
highly accurate
complexity, cost

doug bowman | 3D UIs in a nutshell   25

One popular type of position tracking today uses a hybrid of inertial tracking for orientation and acoustic (ultrasonic) tracking for position. Such trackers have good accuracy and low latency, and can be wireless. The Intersense IS-900 is a common tracking system of this type.

Optical/vision-based trackers

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions
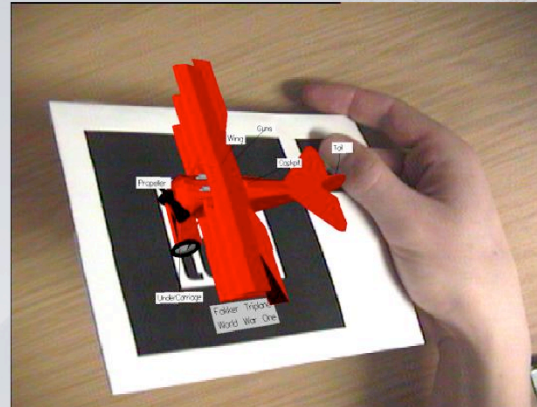
CHI 2008
art.science.balance.

Examples: Vicon, HiBall, ARToolkit
Advantages
- ► accurate
- ► can capture a large volume
- ► allow for untethered tracking

Disadvantages
- ► complex vision techniques
- ► occlusion problem

doug bowman | 3D UIs in a nutshell   26

Another popular tracking type for 3D UIs is vision-based tracking. Vicon trackers, which are often used for offline motion capture, can also be used for real-time position tracking. A much lower-cost option is the ARToolkit, which does 6-DOF vision-based tracking using standard webcams and printed tracking markers, as shown in the picture.

Tracking devices: bend-sensing gloves

CHI 2008
art.science.balance.

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions
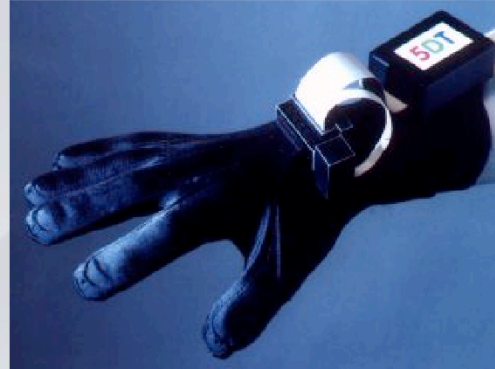
CyberGlove, 5DT glove

Reports hand posture

Gesture:
- ▶ single posture
- ▶ series of postures
- ▶ posture(s) + location or motion

Calibration issues

Lack of knowledge on gestural interfaces

doug bowman | 3D UIs in a nutshell   27

Data gloves measure finger movement of the hand by using various kinds of sensor technology. These sensors are embedded in the glove or placed on top of the glove, usually on the back of the hand. The number of sensors in the glove depends on the manufacturer. Virtual Technologies' CyberGlove has either 18 or 22 sensors which can measure at least 2 joints in each finger, wrist roll and yaw, and others. These types of gloves are commonly used for hand gesture and posture recognition which can be applied to a variety of different interface techniques in virtual environments. Fifth Dimension Technologies (5DT) offers gloves that have either 5 sensors, one for each fingertip or 16 sensors, 2 for each finger and abduction between fingers. 5DT also has wireless versions of each glove.

Pinch gloves are a much simpler and more robust glove-based input device for 3D UIs. They do not sense finger movements or postures; rather, they sense when two or more fingers are touching ("pinch gestures"). A large number of gestures are possible, and the gloves can also be tracked to allow spatial input. Pinch gloves are often a good replacement for tracked button devices (flying mice), since the gloves allow many more discrete inputs and don't require the user to hold a device - the hand becomes the device.
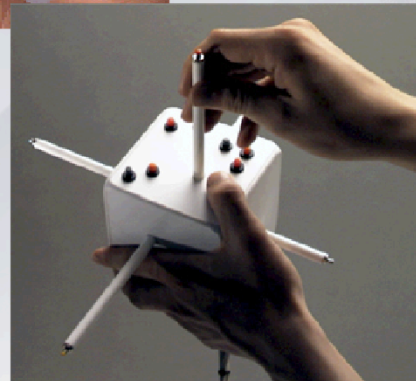
The Ring Mouse (top picture) is a small device worn on the user's finger which uses ultrasonic tracking. It also has two buttons for generating discrete events. The main advantages of this device is that it is wireless and inexpensive. The Fly Mouse is a 3D mouse that also uses ultrasonic tracking. This device has five buttons instead of two and also can be used as a microphone.

The Cubic Mouse (shown in the figure on the right) is an input device developed at GMD that allows users to intuitively specify three-dimensional coordinates in graphics applications. The device consists of a box with three perpendicular rods passing through the center and buttons for additional input.
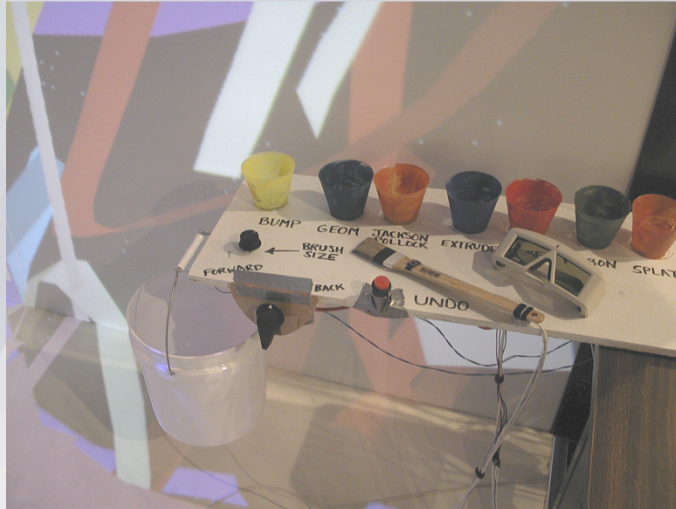
The Painting Table is another example of a special-purpose input device that is used in the CavePainting application, a system for painting 3D scenes in a virtual environment. The device uses a set of conductive cloth contacts as well as traditional buttons and digital sliders. Users can dip the paint brush prop into the colored cups to change brush strokes. The bucket is used to throw paint around the virtual canvas.

References:

Keefe, D., Acevedo, D., Moscovich, T., Laidlaw, D., and LaViola, J. "CavePainting: A Fully Immersive 3D Artistic Medium and Interactive Experience", Proceedings of the 2001 Symposium on Interactive 3D Graphics, 85-93, 2001.

The human body and brain are also sources of input for 3D UIs. In particular, brain-computer interfaces (BCIs) have great potential for 3D UI input.

TabletPC as ideal 3D input device?

3D UIs | displays | **input devices** | basic techniques | evaluation | open questions

CHI 2008
art.science.balance.

Offers both input and output

Has on-board memory

Wireless communication

Portable, (somewhat) light, robust

Allows text / number input

Can be tracked to allow spatial input

But, can still cause fatigue, and requires multiple networked applications

doug bowman | 3D UIs in a nutshell   32

Many researchers have used PDAs or tabletPCs for input in 3D UIs, for the reasons shown. They provide several advantages, and overcome some of the common usability problems in 3D UIs (e.g., it's difficult to provide menus or readable text on 3D displays).

Guidelines for choosing displays & devices

3D UIs | **displays** | **input devices** | basic techniques | evaluation | open questions

CHI 2008
art.science.balance.

Consider combination of input devices, display devices, and interaction techniques

Stereo often not necessary

BUT, the combination of wide FOR, stereo, and head tracking is very powerful

Several specialized input devices vs. one general device

Free moving 6-DOF input for speed and ease of learning

Constrained 6-DOF input for precision and comfort

doug bowman | 3D UIs in a nutshell   33

Choosing displays and input devices for 3D UIs is difficult because of the wide range of technologies available, and the lack of standards. In addition, since input devices don't determine interaction techniques, the techniques must also be considered when choosing devices.

Basic 3D interaction techniques

Universal 3D interaction tasks

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

CHI 2008
art.science.balance.

Navigation
- ► Travel: motor component of viewpoint motion
- ► Wayfinding: cognitive component; decision-making

Selection: picking object(s) from a set
Manipulation: modifying object properties (esp. position/orientation)

System control: issuing a *command* to change system state or mode

doug bowman | 3D UIs in a nutshell   35

We'll be discussing techniques for four basic 3D interaction tasks that are found in most complex 3D applications Obviously, there are other tasks which are specific to an application domain, but these are some basic building blocks that can often be combined to create a more complex task.

Navigation is the most common VE task, and is actually composed of two tasks. Travel is the motor component of navigation, and just refers to the physical movement from place to place. Wayfinding is the cognitive or decision-making component of navigation, and it asks the questions, "where am I?", "where do I want to go?", "how do I get there?", and so on.

Selection is simply the specification of an object or a set of objects for some purpose. Manipulation refers to the specification of object properties (most often position and orientation, but also other attributes). Selection and manipulation are often used together, but selection may be a stand-alone task. For example, the user may select an object in order to apply a command such as "delete" to that object.

System control is the task of changing the system state or the mode of interaction. This is usually done with some type of command to the system (either explicit or implicit). Examples in 2D systems include menus and command-line interfaces. It is often the case that a system control technique is composed of the other three tasks (e.g. a menu command involves selection), but it's also useful to consider it separately since special techniques have been developed for it and it is quite common.

Common travel techniques

CHI 2008
art.science.balance.

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

Pointing
"Grabbing the air"
Locomotion devices

doug bowman | 3D UIs in a nutshell   36

We'll discuss three common techniques.

Pointing technique

CHI 2008
art.science.balance.

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

A "steering" technique

Use hand tracker instead of head tracker

Slightly more complex, cognitively, than gaze-directed steering

Allows travel and gaze in different directions – good for relative motion

doug bowman | 3D UIs in a nutshell   37

Pointing is a steering technique (where the user continuously specifies the direction of motion). In this case, the hand's orientation is used to determine direction. This technique is somewhat harder to learn for some users, but is more flexible than gaze-directed steering.

See: Mine, M. (1995). *Virtual Environment Interaction Techniques* (Technical Report TR95-018): UNC Chapel Hill CS Dept., and

Bowman, D. A., Koller, D., & Hodges, L. F. (1997). *Travel in Immersive Virtual Environments: an Evaluation of Viewpoint Motion Control Techniques.* Proceedings of the Virtual Reality Annual International Symposium, 45-52.

Grabbing the air technique

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

CHI 2008
art.science.balance.

Use hand gestures to move yourself through the world

Metaphor of pulling a rope

Often a two-handed technique

May be implemented using Pinch Gloves

The "grabbing the air" technique uses the metaphor of literally grabbing the world around you (usually empty space), and pulling yourself through it using hand gestures. This is similar to pulling yourself along a rope, except that the "rope" exists everywhere, and can take you in any direction.

This technique may be done with one or two hands, and is often implemented using Pinch Gloves™.

See: Mapes, D., & Moshell, J. (1995). A Two-Handed Interface for Object Manipulation in Virtual Environments. *Presence: Teleoperators and Virtual Environments, 4*(4), 403-416.

Instead of relying solely on common input devices and software-based interaction techniques, locomotion devices are special-purpose devices specifically designed for the task of travel. These can range from simple exercise bikes, to omni-directional treadmills, like the Torus treadmill shown in the picture.

Classification of travel and locomotion

CHI 2008
art.science.balance.

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

	Virtual turning	Real turning
Virtual translation	Desktop VEs Vehicle simulators CAVE wand	Most HMD systems Walking in place VMC
Real translation	Stationary cycles Treadport Biport	Wide-area tracking UNIPORT ODT

A useful way to classify locomotion devices and other travel techniques is their use of virtual and physical movements - both translation and rotation. We know that physical movements can be helpful in helping users maintain spatial orientation, although providing both real translation and real turning can be costly and difficult.

Travel design guidelines

Make simple travel tasks simple (target-based techniques for motion to an object, steering techniques for search).

Provide multiple travel techniques to support different travel tasks in the same application.

Use graceful transitional motions if overall environment context is important.

Train users in sophisticated strategies to help them acquire survey knowledge.

Consider integrated ("cross-task") ITs if travel is used in the context of another task (e.g. manipulation).

Most travel tasks are simple in the mind of the user – they just want to change their location while focusing on something else. Thus, you should use a technique that meets the requirements of the task: e.g. use a target-based technique if the only goal is to move between known objects - don't put unnecessary cognitive load on the user.

Remember the differences between tasks such as exploration and primed search – you may need more than one technique. There is a tradeoff between the specificity of the technique and the amount of learning load you want to put on the user. In many cases, multiple techniques requiring a bit more learning time may be much more efficient in the long run.

Many applications require the user to be aware of their location within the space, have an overall survey knowledge of the space, etc. (see the lecture on wayfinding). In these cases it is important to use transitional motion between locations, even if it is fast, in order to maintain awareness of the space. (A good use of this concept in a desktop system is Mackinlay, Card, and Robertson, Rapid controlled movement through a virtual 3D workspace, SIGGRAPH '90, 171-176.)

Strategies (how the user uses the technique) are as important as the technique itself, especially in tasks requiring spatial knowledge. Therefore, you should provide training, instructions, and guidance to help the user take advantage of the technique.

Cross-task ITs can be useful if travel is not the main interaction, but is only used, for example, to gain a better viewpoint on a manipulation task. Remember that such motion can be tiring, however, and should not be used for very long exposure period applications.

Common selection techniques

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

CHI 2008
art.science.balance.

Simple virtual hand
Ray-casting
Occlusion
Go-go (arm-extension)

doug bowman | 3D UIs in a nutshell   42

We'll discuss four selection techniques.

Simple virtual hand technique

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

CHI 2008
art.science.balance.

One-to-one mapping between physical and
    virtual hands

Object can be selected by "touching" or
    intersecting virtual hand with object

"Natural" mapping

The most common technique is the simple virtual hand, which does "real-world" selection via direct "touching" of virtual objects. In the absence of haptic feedback, this is done by intersecting the virtual hand (which is at the same location as the physical hand) with a virtual object.

Implementing this technique is simple, provided you have a good intersection/collision algorithm. Often, intersections are only performed with axis-aligned bounding boxes or bounding spheres rather than with the actual geometry of the objects.

Another common technique is ray-casting. This technique uses the metaphor of a laser pointer – an infinite ray extending from the virtual hand. The first object intersected along the ray is eligible for selection. This technique is efficient, based on experimental results, and only requires the user to vary 2 degrees of freedom (pitch and yaw of the wrist) rather than the 3 DOFs required by the simple virtual hand and other location-based techniques.

See: Mine, M. (1995). *Virtual Environment Interaction Techniques* (Technical Report TR95-018): UNC Chapel Hill CS Dept.

Next, we'll cover the occlusion technique (also called the "sticky finger" technique). This technique works in the plane of the image – that is, you select an object by "covering" it with the virtual hand so that it is occluded from your point of view. Geometrically, this means that a ray is emanating from your eye, going through your finger, and then intersecting an object.
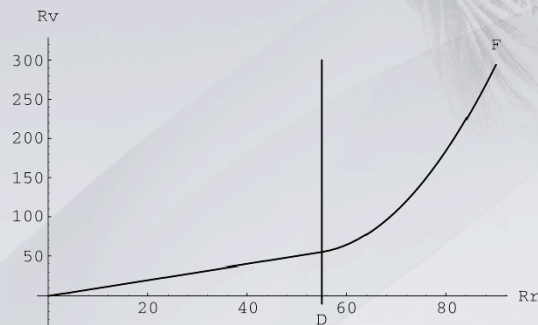
See: Pierce, J., Forsberg, A., Conway, M., Hong, S., Zeleznik, R., & Mine, M. (1997). *Image Plane Interaction Techniques in 3D Immersive Environments.* Proceedings of the ACM Symposium on Interactive 3D Graphics, 39-44.

The Go-Go technique is based on the simple virtual hand, but it introduces a non-one-to-one mapping between the physical hand and the virtual hand, so that the user's reach is greatly extended. This is called an arm-extension technique.

The graph shows the mapping between the physical hand distance from the body on the x-axis and the virtual hand distance from the body on the y-axis. There are two regions. When the physical hand is at a depth less than a threshold 'D', the one-to-one mapping applies. Outside D, a non-linear mapping is applied, so that the farther the user stretches, the faster the virtual hand moves away.

See: Poupyrev, I., Billinghurst, M., Weghorst, S., & Ichikawa, T. (1996). *The Go-Go Interaction Technique: Non-linear Mapping for Direct Manipulation in VR.* Proceedings of the ACM Symposium on User Interface Software and Technology, 79-80.

We'll discuss four 3D object manipulation techniques.

Simple virtual hand technique

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

CHI 2008
art.science.balance.

Simply attach object to virtual hand, move it directly

doug bowman | 3D UIs in a nutshell   48

We already saw the simple virtual hand technique for selection. When this technique is used for object manipulation, the implementation is quite easy. It simply involves making a change to the scene graph by attaching the selected object to the virtual hand. Then, as the virtual hand moves and rotates, the selected object will inherit those transformations. When the object is released, it should just be reattached to its earlier location in the tree.

The only tricky issue here is that you must ensure when grabbing or releasing the object that it does not move (in the world CS). If you simply make the object a child of the hand, it may move since its position is now being interpreted relative to a new CS (the hand's). To be completely general, then, you must get the object's position p in the world CS first, then do the attachment, then calculate p's location in the hand CS, then move the object to that position (relative to the hand). The opposite transformation is done upon release.

This same basic procedure works for other techniques that simply attach the object to the selector, like Go-Go and ray-casting.
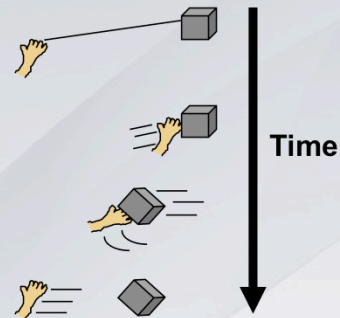
The HOMER technique uses ray-casting for selection and then moves the virtual hand to the object for hand-centered manipulation. The depth of the object is based on a linear mapping. The initial torso-physical hand distance is mapped onto the initial torso-object distance, so that moving the physical hand twice as far away also moves the object twice as far away. Also, moving the physical hand all the way back to the torso moves the object all the way to the user's torso as well.

See: Bowman, D., & Hodges, L. (1997). *An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments.* Proceedings of the ACM Symposium on Interactive 3D Graphics, 35-38.

Often used w/ occlusion

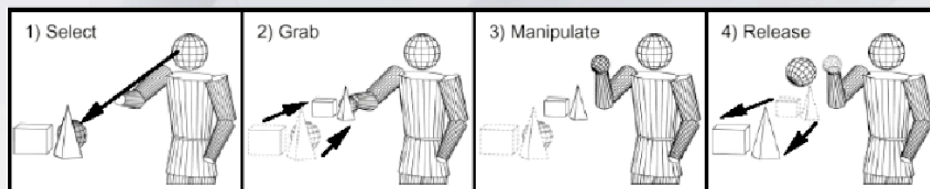At selection, scale user up (or world down) so that virtual hand is actually touching selected object

User doesn't notice a change in the image until he moves

1) Select   2) Grab   3) Manipulate   4) Release

doug bowman | 3D UIs in a nutshell   50

The scaled-world grab technique is often used with occlusion selection. The idea is that since you are selecting the object in the image plane, you can use the ambiguity of that single image to do some magic. When the selection is made, the user is scaled up (or the world is scaled down) so that the virtual hand is actually touching the object that it was occluding. If the user doesn't move (and the graphics are not stereo), there is no perceptual difference between the images before and after the scaling. However, when the user starts to move the object and/or his head, he realizes that he is now a giant (or that the world is tiny) and he can manipulate the object directly, just like the simple virtual hand.

See: Mine, M., Brooks, F., & Sequin, C. (1997). *Moving Objects in Space: Exploiting Proprioception in Virtual Environment Interaction.* Proceedings of ACM SIGGRAPH, 19-26, and

Pierce, J., Forsberg, A., Conway, M., Hong, S., Zeleznik, R., & Mine, M. (1997). *Image Plane Interaction Techniques in 3D Immersive Environments.* Proceedings of the ACM Symposium on Interactive 3D Graphics, 39-44.

The world-in-miniature (WIM) technique uses a small "dollhouse" version of the world to allow the user to do indirect manipulation of the objects in the environment. Each of the objects in the WIM is selectable using the simple virtual hand technique, and moving these objects causes the full-scale objects in the world to move in a corresponding way. The WIM can also be used for navigation by including a representation of the user, in a way similar to the map-based travel technique, but including the 3rd dimension.

See: Stoakley, R., Conway, M., & Pausch, R. (1995). *Virtual Reality on a WIM: Interactive Worlds in Miniature.* Proceedings of CHI: Human Factors in Computing Systems, 265-272, and

Pausch, R., Burnette, T., Brockway, D., & Weiblen, M. (1995). *Navigation and Locomotion in Virtual Worlds via Flight into Hand-Held Miniatures.* Proceedings of ACM SIGGRAPH, 399-400.

Manipulation design guidelines

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

Match the interaction technique to the device

Use techniques that can help to reduce clutching

Use pointing techniques for selection and virtual hand techniques for manipulation

Reduce degrees of freedom when possible

Common system control techniques

CHI 2008
art.science.balance.

Virtual menus
Tool selectors (belts, palettes, chests)
Speech commands
Pen & tablet technique

For the most part, these only require a selection technique
Good visual feedback is necessary

System control is a wide-ranging topic, and there are many different techniques, some of which are listed here. For the most part, these techniques are not difficult to implement, since they mostly involve selection, which we've already covered. For example, virtual menu items might be selected using ray-casting. For all of the techniques, good visual feedback is required, since the user needs to know not only what he is selecting, but what will happen when he selects it.

Pen & tablet technique

CHI 2008
art.science.balance.

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

Create
Tree Rock Grass

doug bowman | 3D UIs in a nutshell   54

I only want to touch on one system control technique, because of its widespread use. The pen & tablet technique uses a physical pen and tablet (see left image). In the virtual world, the user sees a virtual pen and tablet, and a 2D interface on the surface of the virtual tablet (right image). The physical devices provide near-field haptics and constraints that make such an interface easy to use.

As we mentioned in the section on input devices, the same effect (and more) can be achieved with a tabletPC, but this only works if your display device allows the user to see the physical world (i.e., it wouldn't work with an HMD).

See: Angus, I., & Sowizral, H. (1995). *Embedding the 2D Interaction Metaphor in a Real 3D Virtual Environment.* Proceedings of SPIE, Stereoscopic Displays and Virtual Reality Systems, 282-293, and

Schmalsteig, D., Encarnacao, L., & Szalzvari, Z. (1999). *Using Transparent Props For Interaction with The Virtual Table.* Proceedings of the ACM Symposium on Interactive 3D Graphics, 147-154.

System control design guidelines

CHI 2008
art.science.balance.

3D UIs | displays | input devices | **basic techniques** | evaluation | open questions

Don't disturb flow of action
Use correct spatial reference
Allow multimodal input
Structure available functions
Prevent mode errors by giving feedback

Be creative, and test!

Extracted from the descriptions of system control techniques, several important design guidelines can be stated. Due to the relative lack of formal evaluations, these guidelines are primarily based on tendencies described by researchers and personal experience.

System control is often integrated within another universal interaction task. Due to this integration, we should avoid disturbing the flow of action of an interaction task. The user should stay focused on the task. "Modeless" interaction (where the mode changes are very natural) is ideal. One way of supporting the user to easily access a system control interface is by using a correct spatial reference. This guideline is of course mostly applicable to graphical menus, but tools also benefit from a strong spatial reference. Another method to allow a more seamless integration of system control into a flow of action is to use a multimodal, or hybrid, system control interface. Multimodal interfaces can increase the performance of issuing a command, and may allow multiple channels to access the system control interface. However, keep in mind that multimodal system control is not always suitable or applicable.

After the user has accessed a system control interface, he/she has to select an item from a set: when this set is large, i.e. when a large number of functions are available, one needs to structure the items. As stated in the guidelines on graphical menus, this might be achieved by methods like using context-sensitivity, or by clearly communicating the hierarchy of items and (sub)menus.

Finally, always try to prevent mode errors by providing the user with appropriate feedback during and after selection of a command. Mode errors can be highly disturbing and they interrupt the flow of action in an application.

CHI 2008
art.science.balance.

3D UI evaluation

Doug Bowman

How 3D UI evaluation is different

3D UIs | displays | input devices | basic techniques | **evaluation** | open questions

Physical issues
▶ User can't see world in HMD
▶ Think-aloud and speech incompatible

Evaluator issues
▶ Evaluator can break presence
▶ Multiple evaluators usually needed

User issues
▶ Very few expert users
▶ Evaluations must include rest breaks to avoid possible sickness

Evaluation type issues
▶ Lack of heuristics/guidelines
▶ Choosing independent variables is difficult

3D USER INTERFACES: FROM LAB TO LIVING ROOM

CHI 2008
art.science.balance.

doug bowman | 3D UIs in a nutshell   57

There are a number of ways in which evaluation of 3D UIs is different from traditional user interface evaluation.

First, there are physical issues. For example, in an HMD-based VE, the physical world is blocked from the user's view. This means that the evaluator must ensure that the user does not bump into objects or walls, that the cables stay untangled, and so on. Another example involves a common method in traditional evaluation called a "think-aloud protocol". However, many VE applications use speech input, which of course is incompatible with this evaluation method.

One of the most important evaluator issues is that an evaluator can break the user's sense of presence during an evaluation. If the sense of presence is considered important to the task/application, the evaluator should try to avoid contact with the user during the tests. Another example is that multiple evaluators are usually needed, because 3D systems are so complex (hardware and software) and because users have more freedom and input bandwidth than users of a traditional UI.

One problem related to users is the lack of users who can truly be considered "experts" in 3D UI usage. Since the distinction between expert and novice usage is important for interface design, this makes recruiting an appropriate subject pool difficult. Also, 3D UI systems have problems with simulator sickness, fatigue, etc. that are not found in traditional UIs. This means that the experimental design needs to include provisions like rest breaks and the amount of time spent in the VE needs to be monitored.

Heuristic evaluation can be problematic, because VE interfaces are so new that there is not a large body of guidelines from which to draw, although this is changing. Also, if you are doing a formal experiment, there are a huge number of factors which might affect performance. For example, in a travel task, the size of the environment, the number of obstacles, the curviness of the path, the latency of the system, and the spatial ability of the user all might affect the time it takes a user to travel from one location to the other. Choosing the right variables to study is therefore a difficult problem.

Testbed evaluation framework

3D UIs | displays | input devices | basic techniques | **evaluation** | open questions

CHI 2008
art.science.balance.

Main independent variables: ITs

Other considerations (independent variables)
► task (e.g. target known vs. target unknown)
► environment (e.g. number of obstacles)
► system (e.g. use of collision detection)
► user (e.g. VE experience)

Performance metrics (dependent variables)
► Speed, accuracy, user comfort, spatial awareness…

Generic evaluation context

doug bowman | 3D UIs in a nutshell   58

An evaluation testbed is a generalized environment in which many smaller experiments or one large experiment can be run, covering as much of the design space as you can. Like other formal experiments, you're evaluating interaction techniques (or components), but you also include other independent variables that could have an effect on performance. These include characteristics of the task, environment, system, and user.

You also measure multiple dependent variables in such experiments to try to get a wide range of performance data. Here we use performance in the broader sense, not just meaning quantitative metrics. The more metrics you use, the more applications can use the results of the experiment by listing their requirements in terms of the metrics, then searching the results for technique(s) that meet those requirements.

Doug Bowman performed such evaluations in his doctoral dissertation, available online at: http://www.cs.vt.edu/~bowman/thesis/. A summary version of these experiments is in this paper:

Bowman, Johnson, & Hodges, Testbed Evaluation of VE Interaction Techniques, Proceedings of ACM VRST '99

Also see: Poupyrev, Weghorst, Billinghurst, and Ichikawa, A Framework and Testbed for Studying Manipulation Techniques, Proceedings of ACM VRST '97.
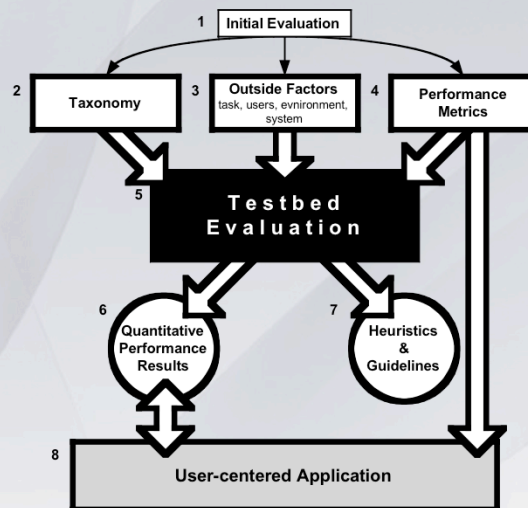
In terms of our three issues, testbed evaluation involves users, produces quantitative (and perhaps qualitative) results, and is done in a generic context.

This figure shows the process used in testbed evaluation. Before designing a testbed, one must understand thoroughly the task(s) involved and the space of interaction techniques for those tasks. This understanding can come from experience, but it's more likely to come from some initial (usually informal) evaluations. This can lead to a taxonomy for a task, a set of other factors that are hypothesized to affect performance on that task, and a set of metrics (discussed later).

These things are then used to design and implement a testbed experiment or set of experiments. The results of running the testbed are the actual quantitative results, plus a set of guidelines for the usage of the tested techniques. The results can be used many times to design usable applications, based on the performance requirements of the application specified in terms of the performance metrics.

User comfort

CHI 2008
art.science.balance.

3D UIs | displays | input devices | basic techniques | **evaluation** | open questions

Simulator sickness
Aftereffects of VE exposure
Arm/hand strain
Eye strain

The other novel user preference metric for VE systems is user comfort. This includes several different things. The most notable and well-studied is so-called "simulator sickness" (because it was first noted in things like flight simulator). This is similar to motion sickness, and may result from mismatches in sensory information (e.g. your eyes tell your brain that you are moving, but your vestibular system tells your brain that you are not moving). There is also work on the physical aftereffects of being exposed to VE systems. For example, if a VE mis-registers the virtual hand and the real hand (they're not at the same physical location), the user may have trouble doing precise manipulation in the real world after exposure to the virtual world. More seriously, things like driving or walking may be impaired after extremely long exposures (1 hour or more). Finally, there are simple strains on arms/hands/eyes from the use of VE hardware.

User comfort is also usually measured subjectively, using rating scales or questionnaires. The most famous questionnaire is the simulator sickness questionnaire (SSQ) developed by Robert Kennedy. Kay Stanney has attempted some objective measures in her study of aftereffects – for example by measuring the accuracy of a manipulation task in the real world after exposure to a virtual world.

Guidelines for 3D UI evaluation

3D UIs | displays | input devices | basic techniques | **evaluation** | open questions

CHI 2008
art.science.balance.

Begin with informal evaluation

Acknowledge and plan for the differences between traditional UI and 3D
UI evaluation

Choose an evaluation approach that meets your requirements

Use a wide range of metrics – not just speed of task completion

doug bowman | 3D UIs in a nutshell   62

Here are a set of guidelines to be used in any type of evaluation of 3D UIs.

Informal evaluation is very important, both in the process of developing an application and in doing basic interaction research. In the context of an application, informal evaluation can quickly narrow the design space and point out major flaws in the design. In basic research, informal evaluation helps you understand the task and the techniques on an intuitive level before moving on to more formal classifications and experiments.

Remember the unique characteristics of 3D UI evaluation from the beginning of this talk when planning your studies.

There is no optimal evaluation technique. Study the classification presented in this talk and choose a technique that fits your situation.

Remember that speed and accuracy do not equal usability. Also remember to look at learning, comfort, presence, etc.

Open questions

Remove encumbrances completely?

CHI 2008
art.science.balance.

3D UIs | displays | input devices | basic techniques | evaluation | **open questions**

Approaches to input/tracking
- ▶ Wireless transmission
- ▶ Vision-based

Approaches to output/display
- ▶ Displays in the world
- ▶ Displays and display generators worn by user

Autostereoscopic displays: no glasses

doug bowman | 3D UIs in a nutshell   64

A common issue in 3D UIs is that the user is often encumbered with wires, wearable computers or devices, awkward goggles or glasses, etc. Is it possible to completely remove these encumbrances and allow the user to interact with 3D UIs as they interact with the real world?

Appropriate mappings?

CHI 2008
art.science.balance.

3D UIs | displays | input devices | basic techniques | evaluation | **open questions**

Mappings between display devices, input devices, interaction techniques, tasks, and applications

Huge potential for empirical research

Examples from our group:

► How should the WIM technique change to be used in the CAVE?
► What navigation strategies do users adopt in fully- vs. semi-immersive displays?
► Which VE displays are the most useful for decision-making tasks in the construction industry?

doug bowman | 3D UIs in a nutshell   65

A large unanswered question in 3D UIs is how to determine the appropriate mappings between displays, input devices, interaction techniques, tasks, and applications. 3D UIs have such a large design space and lack of standards that developers often start from scratch when designing each new 3D UI.
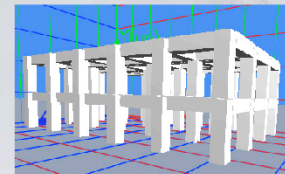
The 3D interaction techniques we've covered are generic; that is, they can be applied to any application domain because they were designed for generic tasks like selection and manipulation. Of course, this makes these techniques reusable. But are they usable enough for real-world applications? One approach is to design domain-specific 3D UIs that try to strike a balance between usability and reusability.

Chen, J. and Bowman, D., Effectiveness of Cloning Techniques for Architectural Virtual Environments. in *IEEE Virtual Reality*, (Alexandria, Virginia, 2006), IEEE CS Press.

Developing 3D UIs is still a major challenge. Often developers must code everything from scratch, using toolkits designed for 3D graphics programming, not UI programming. Some researchers are addressing the area of 3D UI development tools by designing 3D UI description languages or dataflow architectures for 3D interaction techniques.

Wingrave, C. and Bowman, D., CHASM: Bridging Description and Implementation of 3D Interfaces. in *IEEE VR 2005 Workshop on New Directions in 3D User Interfaces*, (Bonn, Germany, 2005), Shaker Verlag, 85-88.

Ray, A. and Bowman, D. Towards Reusable 3D Interaction Techniques *ACM Symposium on Virtual Reality Software and Technology*, Newport Beach, California, 2007, 187-190.

Evaluation of AR/MR?

3D UIs | displays | input devices | basic techniques | evaluation | **open questions**

With optical see-through displays, it's impossible for the evaluator to see exactly what the user sees!

Possible solutions:
- ► Video see-through
- ► Evaluator has own tracked display
- ► Simulate AR with VR

An open question related to evaluation relates to the evaluation of 3D UIs in augmented or mixed reality systems, where optical see-through displays are popular (the user sees the real world directly, with virtual overlays). In such systems, the real world view is not captured directly, so the evaluator can't tell exactly what the user is seeing.

Will there ever be a standard 3D UI?

3D UIs | displays | input devices | basic techniques | evaluation | **open questions**

CHI 2008
art.science.balance.

Standards would increase usability between applications

Standards may actually decrease the potential usability of individual applications

Depends on "mass market" applications

► If VEs are only used in niche applications, no need to standardize

► If VEs are to be used by the general population, standardization will probably follow

doug bowman | 3D UIs in a nutshell   69

Finally, we note that we are still not close to a standard 3D UI. There is considerable disagreement among 3D UI researchers about the answer to this question.