

3D Interaction in Mixed Realities  
Computer Science Department - IME/USP

# Selection and Manipulation

Prof. Carlos Hitoshi Morimoto  
[hitoshi@ime.usp.br](mailto:hitoshi@ime.usp.br)

2019 S2

# Readings

“Selection and Manipulation”

chapter 7 of the book “3D User Interfaces -  
Theory and Practice”,

by LaViola, Kruijff, McMahan, Bowman, and  
Poupyrev

# 3D Interaction

Fundamental 3D interaction tasks

Selection and Manipulation

Travel

moving in and around an environment  
motor component of navigation

Wayfinding

cognitive component of navigation

# 3D Manipulation tasks

Selection  
Positioning  
Rotation  
Scaling

# Canonical manipulation tasks

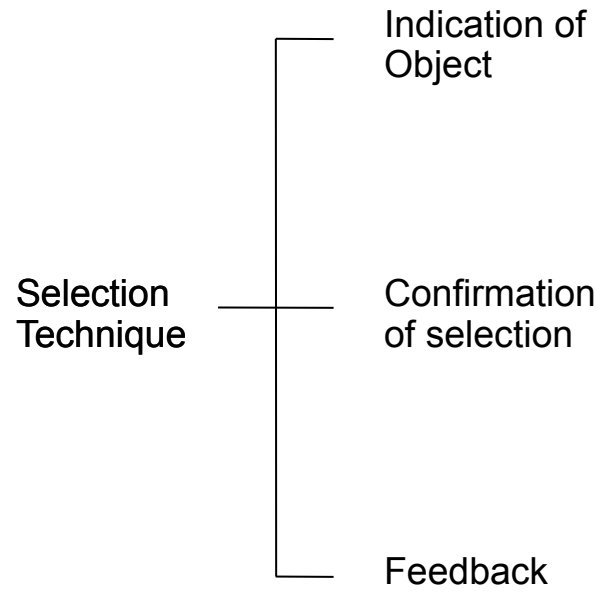
- Selection
  - target acquisition
  - picking an object with a hand
- Positioning
  - change the position of an object
  - Moving an object from X to Y
- Rotation
  - change the orientation of an object
  - Rotate an object to a target orientation
- Scaling
  - change the size of an object

# Task Parameters

# Selection

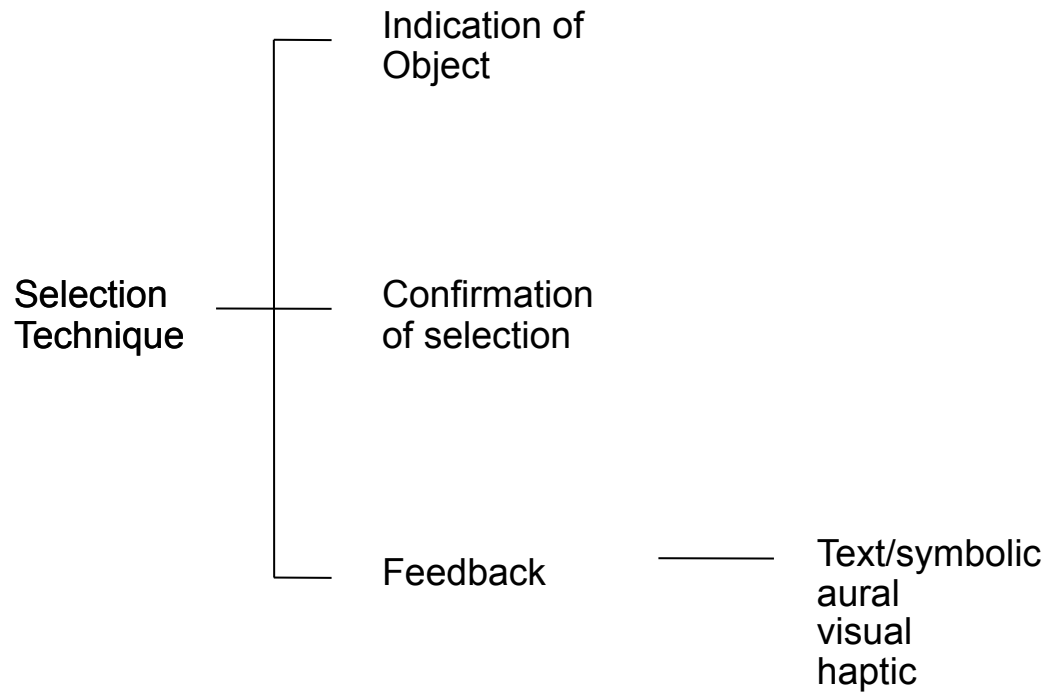
- Distance/direction
- Target size
- Target context (e.g. number/kind of objects near target)
- number of targets
- Occlusion of targets
- Required accuracy/precision

# Selection

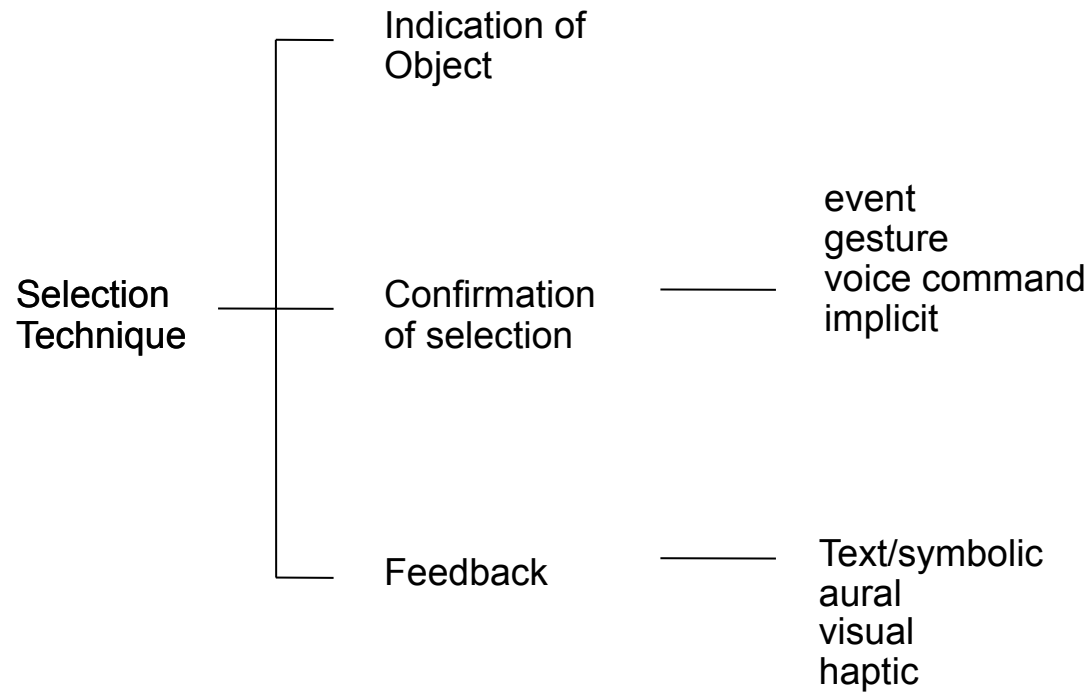




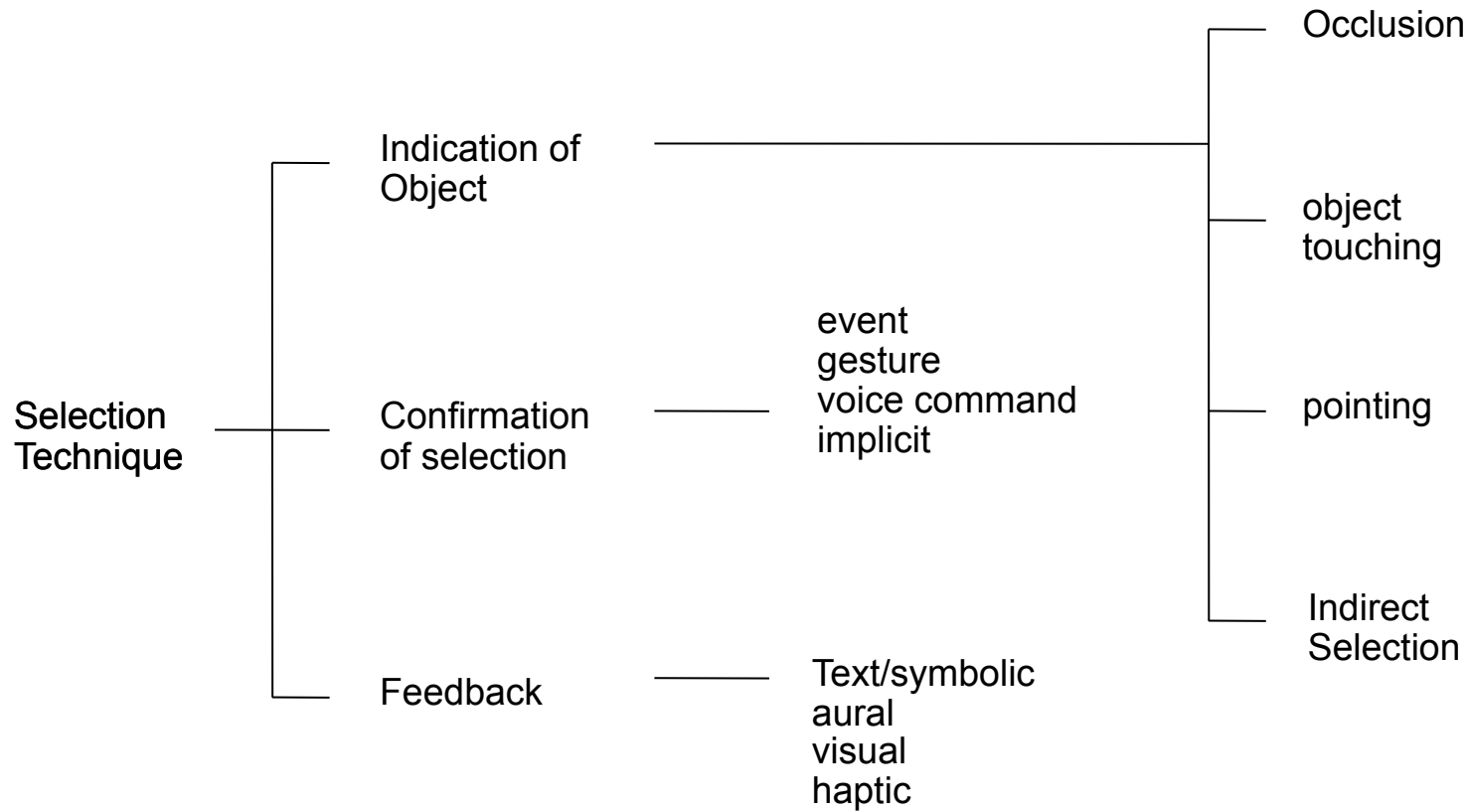
# Selection



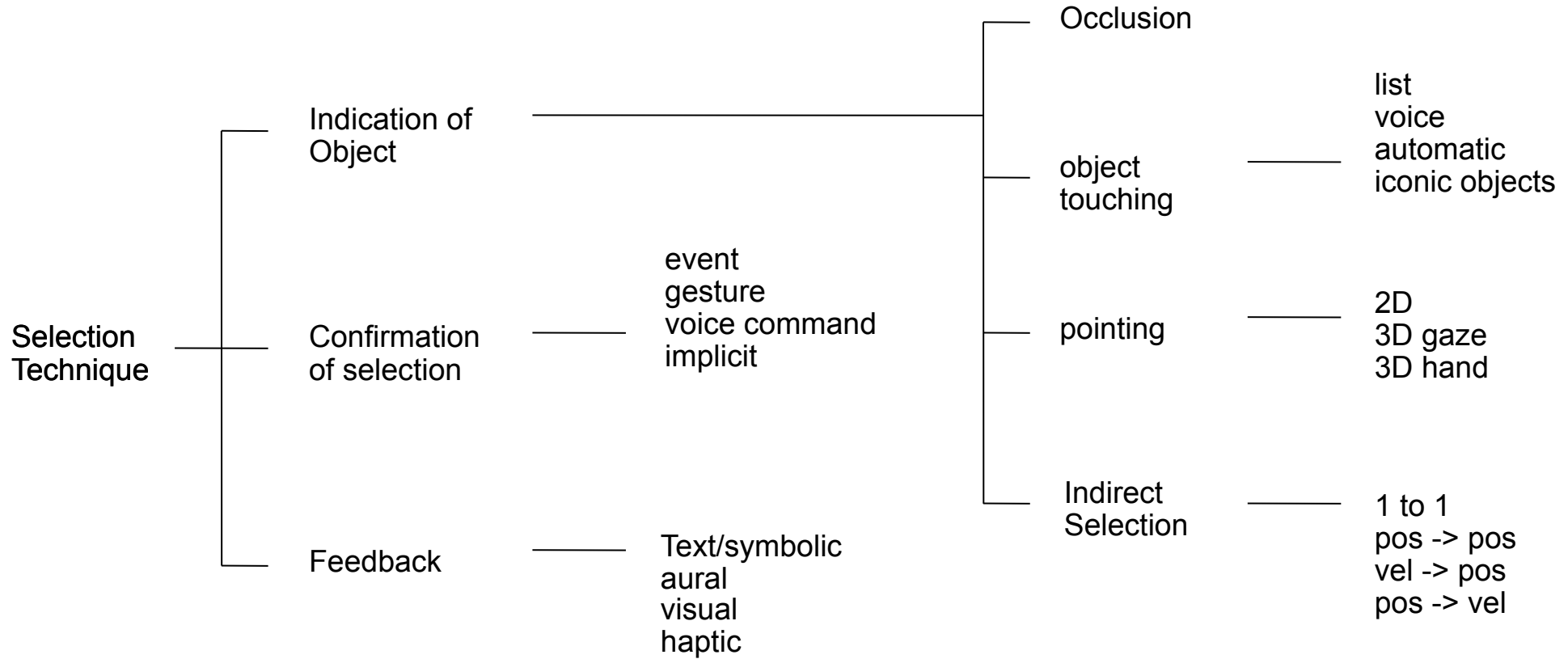
# Selection



# Selection



# Selection



# Position

- Distance/direction to target source position
- Distance/direction to target destination position
- Distance to translate
- Required accuracy/precision

# Rotation

- Distance/direction to target
- Initial orientation
- Final orientation
- Amount to rotate
- Required accuracy/precision

# Scaling

Distance to target

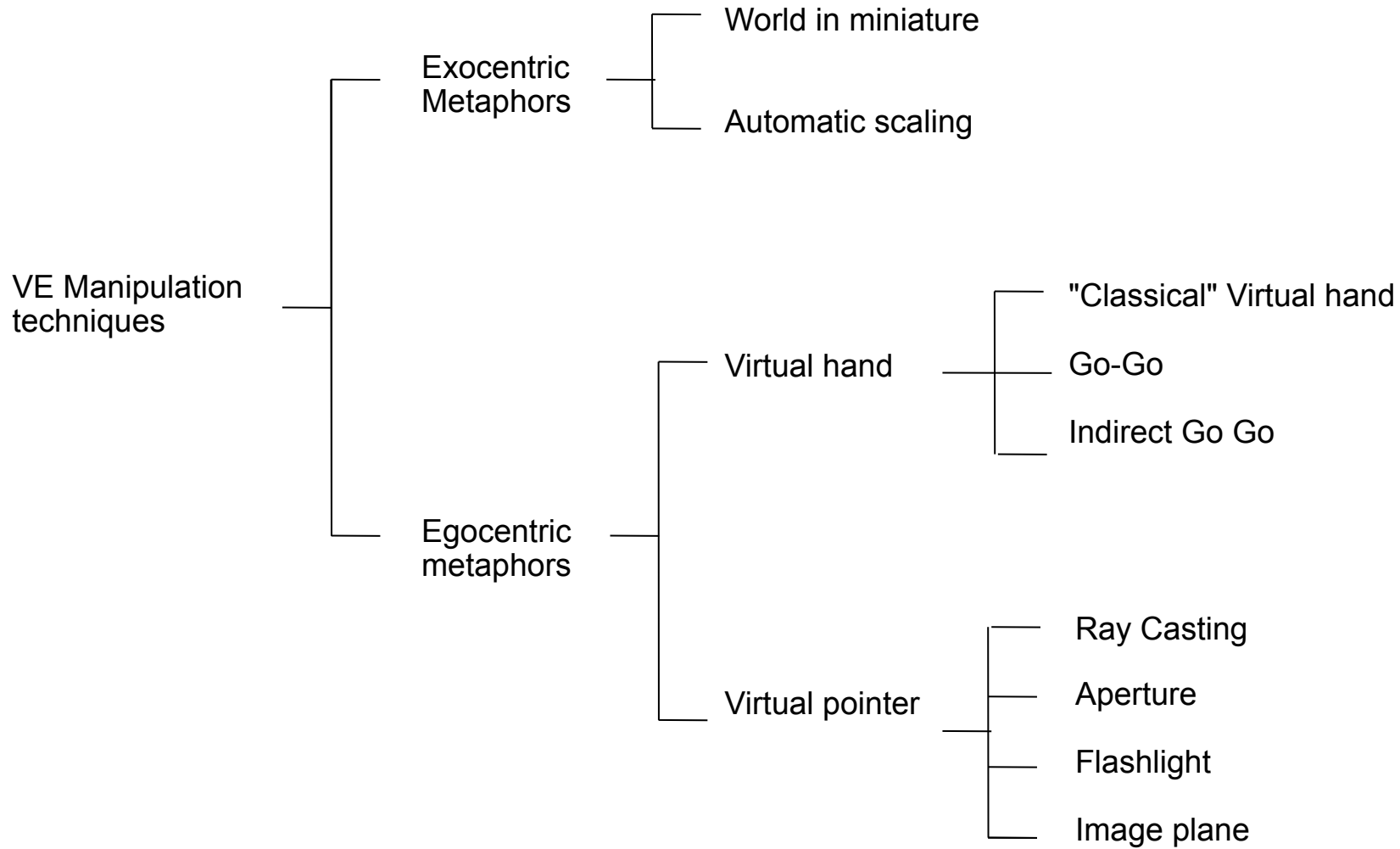
Initial scale

Final scale

Amount of scale

Required precision of scale

# Manipulation by metaphor





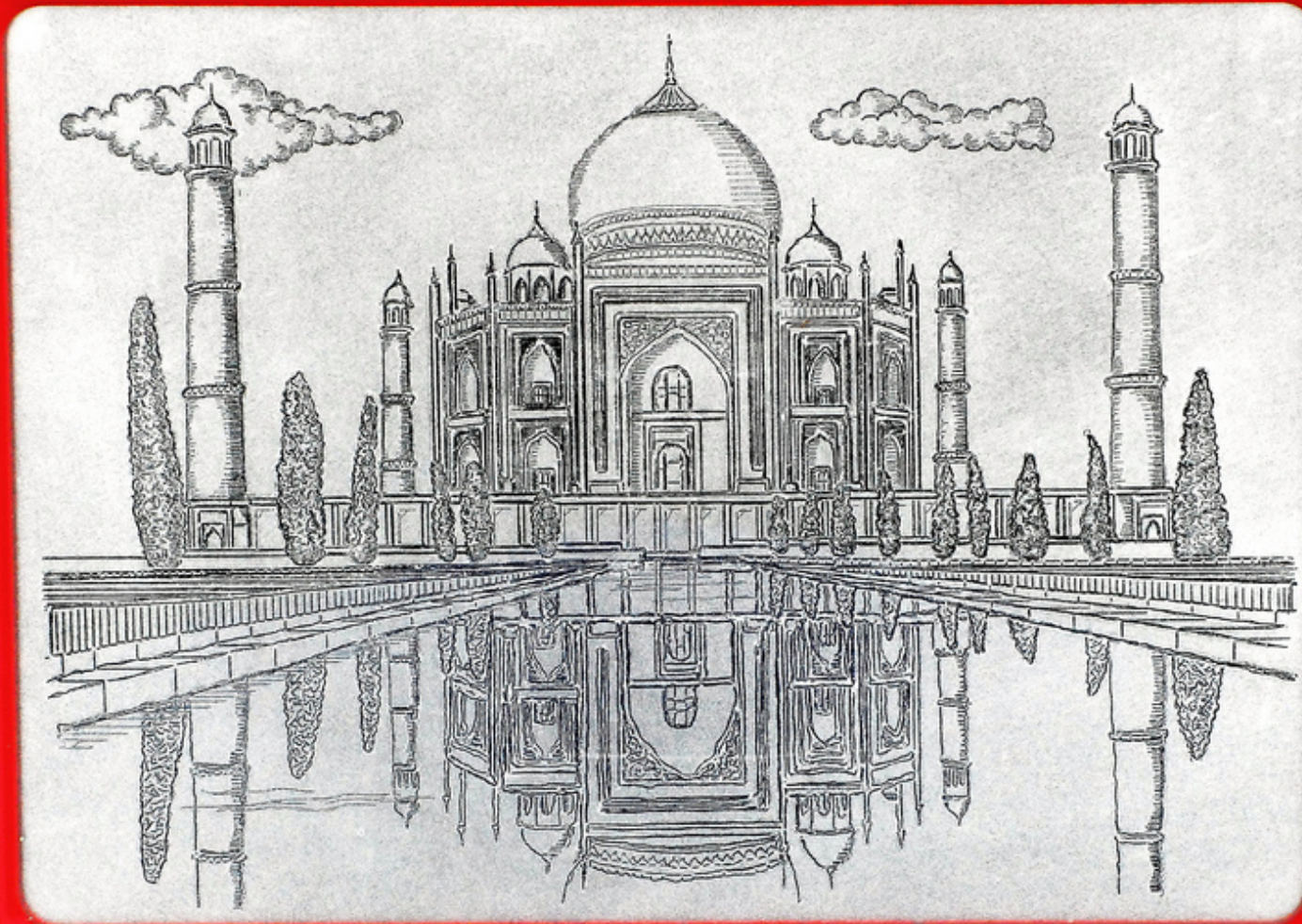
# Device Properties

# Device Properties

- DoF - degrees of freedom
  - Number of DoF
  - Type of DoF
  - Integration of DoF
- Example:
  - Etch-a-sketch
  - Skedoodle



MAGIC *Etch A Sketch*<sup>®</sup> SCREEN



[https://en.wikipedia.org/wiki/Etch\\_A\\_Sketch](https://en.wikipedia.org/wiki/Etch_A_Sketch)

# Skedoodle



- Bill Buxton's Notes
- The Skedoodle is my favourite classic foil vis-à-vis the Etch-a-Sketch. Which is the best for drawing? Besides working that question through, this is also a very good example of the power of layering constraints on a system to enable it to assume attributes that it would not have had otherwise. Try drawing a square with the joystick without using the template.

<https://www.microsoft.com/buxtoncollection/detail.aspx?id=182&from=http%3A%2F%2Fresearch.microsoft.com%2Fen-us%2Fum%2Fpeople%2Fbibuxton%2Fbuxtoncollection%2Fdetail.aspx%3Fid%3D182>

# Device Properties

## Position vs Force Measurement

- Position measurement
  - Isotonic device moves with user's body part
  - Better for position control
  - E.g. PlayStation Move senses position/orientation
- Force measurement
  - Isometric device does not move
  - Elastic device moves slightly (finite resistance increases with distance)
  - Better for rate control
  - E.g., elastic joystick senses force/torque

# PlayStation Move



# SpaceMouse

www.3dconnexion.com

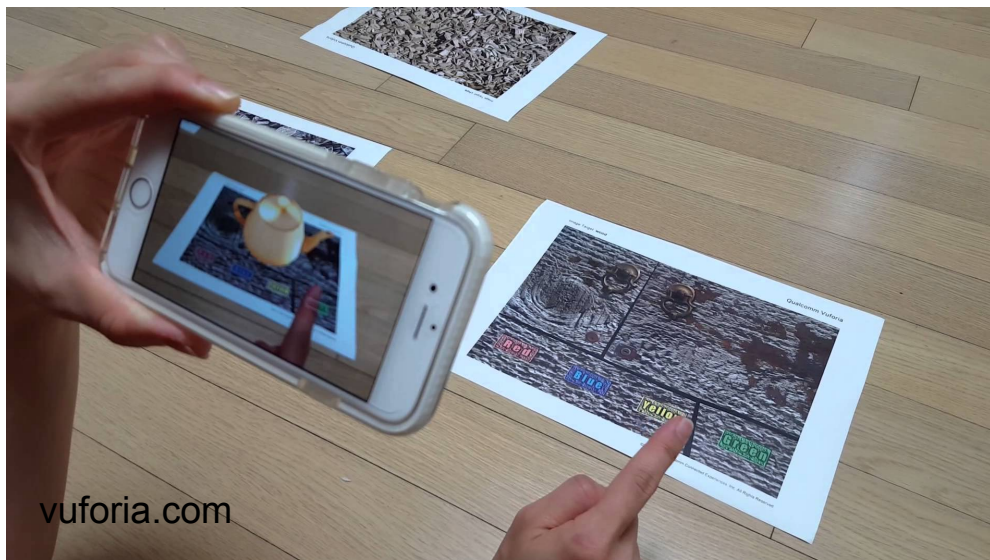


<https://www.youtube.com/watch?v=TYqSljkoWh8>

# Device properties

## Placement / attachment

- Hand-mounted
  - Tracks the hand pose
  - Hand is empty
- hand-held
  - Roughly tracks hand pose
  - Can be manipulated
  - Hand is full





# Device properties

## Form Factor (shape / size)

- Grip
  - precision, comfort, steadiness, range of motion
- Axes of symmetry
  - effect on clutching

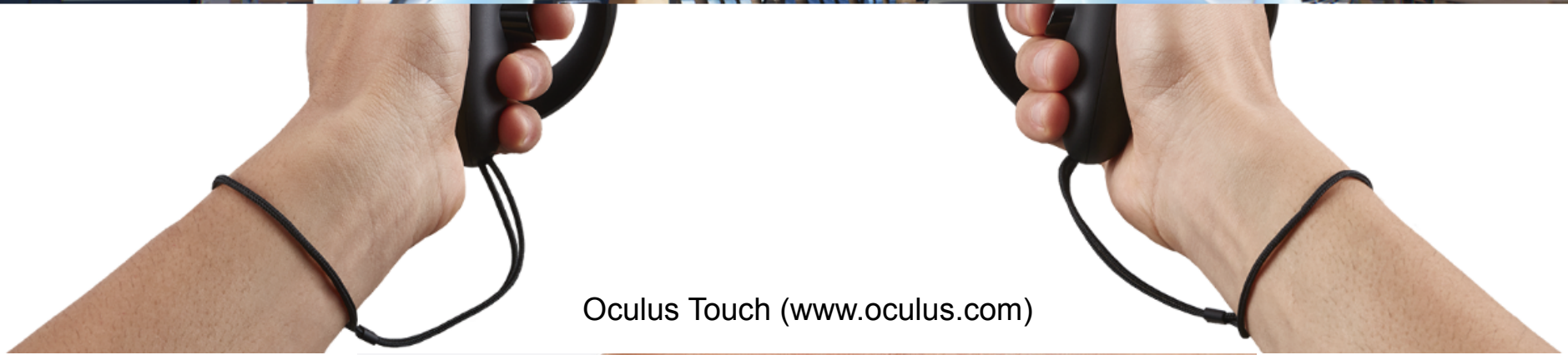


STEM controller



6 axis wireless controller





Oculus Touch ([www.oculus.com](http://www.oculus.com))



# Classification of 3D Manipulation Techniques

by mapping  
by task decomposition  
by metaphor

# Classification by mapping

Mapping real to virtual:

Isomorphic

1:1 -- "natural" physical correspondence

Nonisomorphic

Magical, non-physical mapping

e.g.: rubber arm

# Classification by task decomposition

- Analyze technique in terms of basic component “building blocks”
- E.g., selection decomposed into
  - Indication of object(s)
    - moving body part, manipulating device
  - Confirmation of selected object(s)
    - with explicit trigger action
  - Feedback during task
    - highlighting object

# Classification by Metaphor

- Basic mental model, often based on comparison with some existing concept
  - Constraints
    - What the user cannot do
  - Affordances
    - What the user can do

# Classification by Metaphor

- e.g.: user's relationship to environment classified into
  - exocentric (aka "god's-eye view")
    - user('s head) external to environment
  - egocentric (aka "first-person view", immersive)
    - user('s head) inside the environment
- e.g.: virtual X
  - Virtual pointer
    - user interacts with objects by pointing at them
  - Virtual hand
    - user touches/picks up objects with a virtual hand

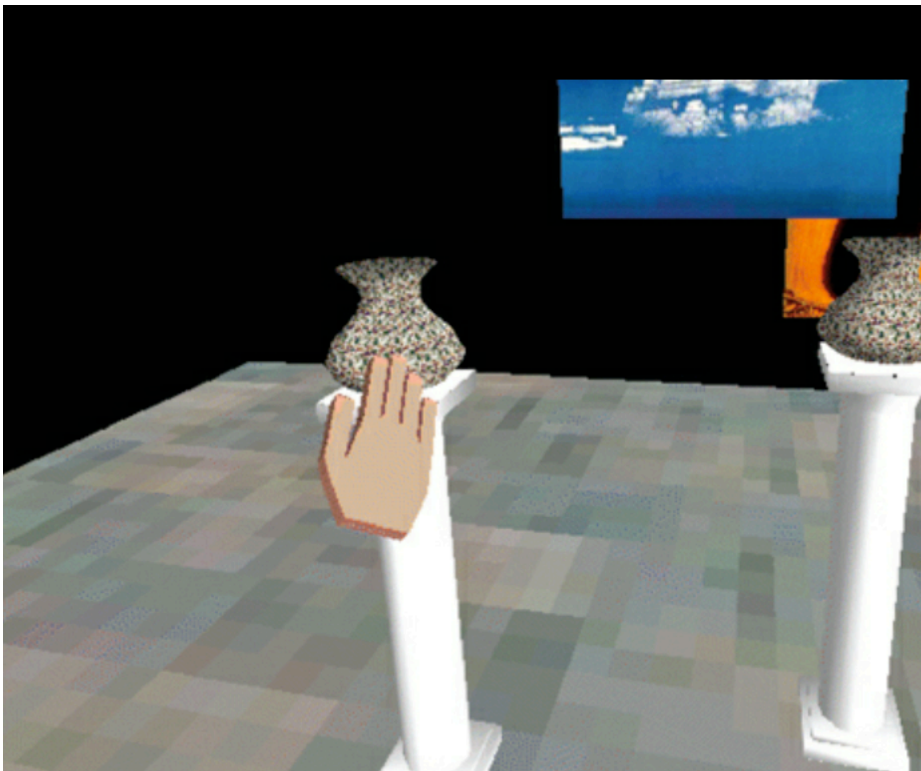
# Grasping

Idea: grasp the object you want to manipulate

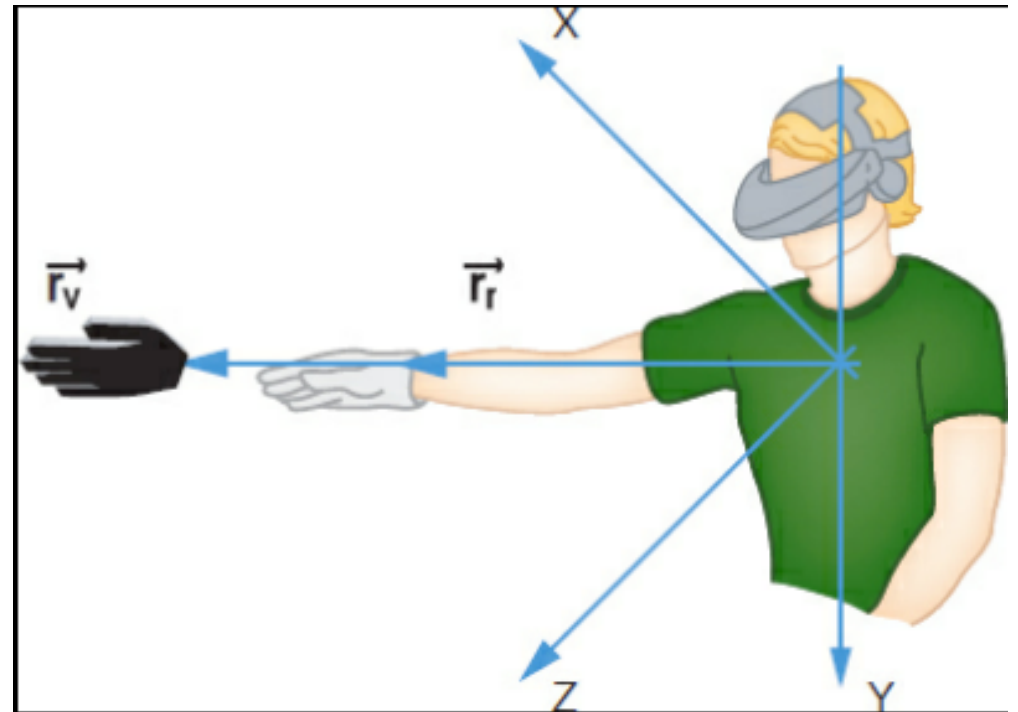


# Hand Grasping

Simple Virtual Hand



Go-Go

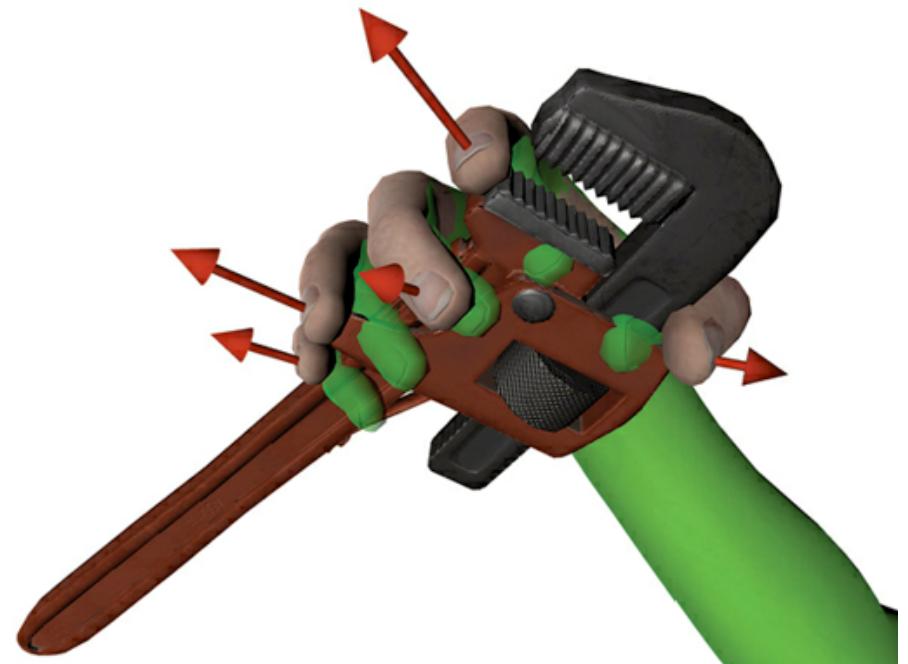


# Finger Grasping

Rigid fingers

Soft (deformable)

God (virtual pad) Fingers



# Pointing

Vector based  
Volume based

# Pointing

- User points with tracked device(s)/body part(s)
- Pointing direction indicates target
- User can confirm desired target by issuing trigger event
  - e.g. voice, button, gesture
- Parameters
  - pointing direction
    - how defined by body and device DoFs?
  - Selection volume: position, size, shape
    - how defined by body and device DoFs?
    - determines visual feedback, number of targets selected

# Pointing

- Advantages
  - Allows selection/manipulation beyond user's reach
  - Requires less physical movement than virtual hand
- Disadvantages
  - Poor positioning except radially around user
  - Rotations easy only around pointer vector

To keep in mind...

There are many ways in which  
physical DoFs can be mapped to  
technique parameters

# Sorcerer's Apprentice

D.Vickers, 72

- 6DoF HMD + "wand"
- Wand tip is tracked in 3D
- Selection
  - 3D vertices
    - Intersection with virtual cube at tip (3D cursors/virtual hand)
    - Buttons move virtual cube along infinite ray from right eye through tip (pointing)
  - Menu items on wall chart
    - intersection with ray from right eye through tip (pointing)

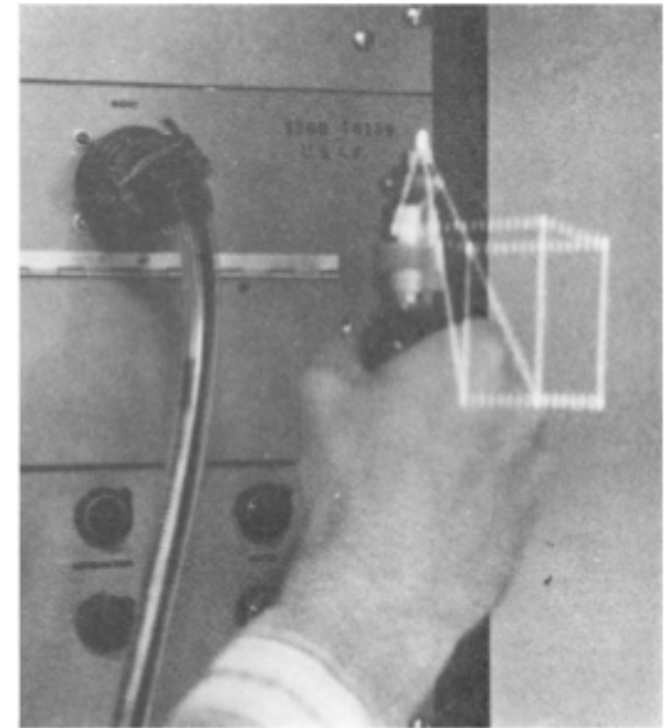
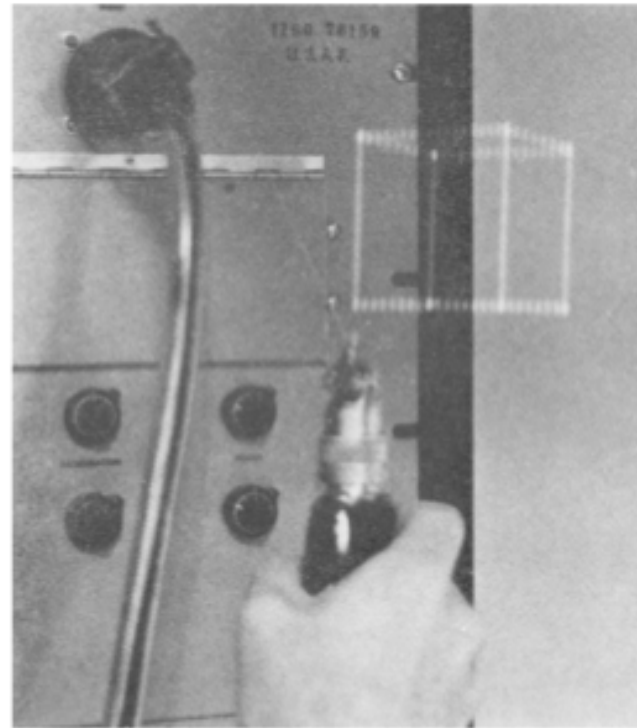
tracked  
wand tip



# Sorcerer's Apprentice

Vertex manipulation using wand (virtual hand)

Physically tracked wand tip  
(Note attached strings)



Greenfield, H., Simulation of atherogenesis by computer display techniques, *Cardiovascular Diseases, Bull. Of the Texas Heart Inst.*, 1(4), 1974  
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC287499/pdf/cardioidis00031-0054.pdf>



# Put That There

R. Bolt, 80

- User points at wall-sized rear projection display with 6DOF tracked hand that defines ray
- Voice commands combined with pointing gesture manipulate objects
  - Create a blue square there
  - Make that green
  - Put that there
- 1st multimodal system combining gesture and speech
  - "That" and "there" are triggers



[http://media.mit.edu/speech/videos/Putthatthere\\_1981\\_edited\\_full.mov](http://media.mit.edu/speech/videos/Putthatthere_1981_edited_full.mov)

# Pointing: Ray Casting using Hand

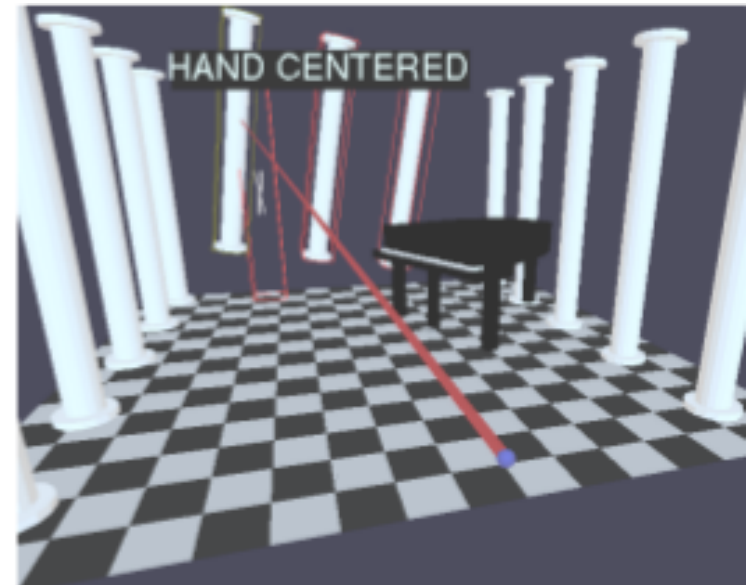
- User points with a virtual ray
  - . defined by 6 DOF-tracked hand (or device)
  - . typically visualized as a line segment attached to hand

$$\mathbf{p}(a) = \mathbf{h} + a \mathbf{p}$$

where  $0 < a < \infty$ ,  $\mathbf{h}$  is the hand position

$\mathbf{p}$  is the ray direction

- Ray can be invisible, short, or infinitely long
- Target disambiguation by selecting closest object intersected by ray
- Hand/tracker 'jitter' amplified by distance makes it difficult to pick distant objects



M. Mine, ISAAC, 1997

# Ray casting using hand and eye

- User points with a virtual ray

Defined by position of 3 DOF-tracked eye

Immersive: position of 3DOF hand/tool

Usually not drawn as a ray, since it is a point as seen from eye

$$\mathbf{p}(a) = \mathbf{e} + a (\mathbf{h} - \mathbf{e})$$

where  $0 < a < \infty$ ,  $\mathbf{h}$  is the hand position, and  $\mathbf{e}$  is the eye position

- Need to designate specific eye in stereo
- Disambiguate target by selecting closest object
- hand/tracker 'jitter' amplified by distance
  - makes it difficult to pick distant objects



D. Vickers, 1972

# Pointing: two handed

- One hand specifies origin and the other the point on ray
- Simple version

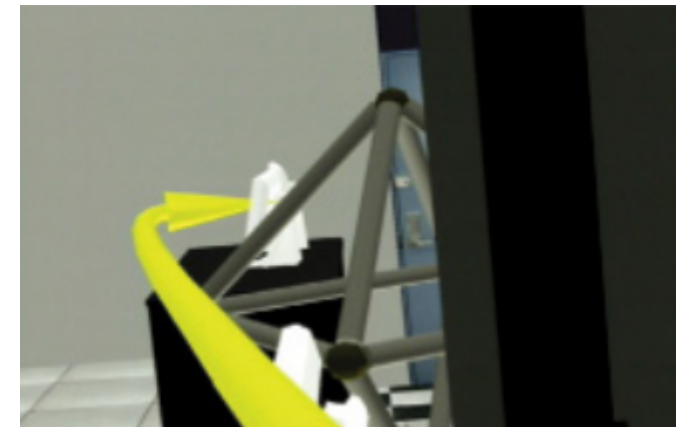
$$p(a) = h_0 + a(h_1 - h_0)$$

where  $0 < a < \infty$ ,  $h_0$  and  $h_1$  are hand positions

- Extensions
  - distance between hands determines length of ray or which object selected
  - Orientation of hands curves the pointer
    - point around objects
    - avoid obscuring important objects



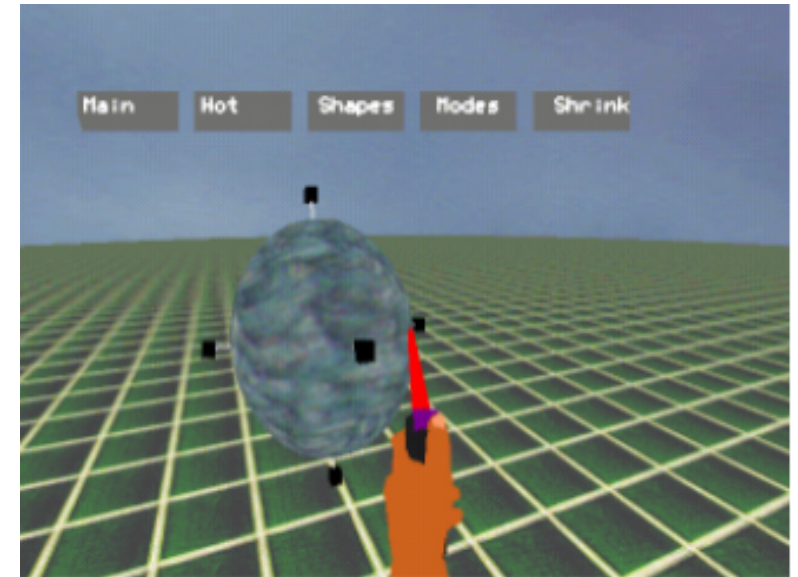
A. Olwal & S. Feiner, The Flexible Pointer, 2003



[www.cs.columbia.edu/graphics/courses/csw4172/resources/olwalFlexPointer.mov](http://www.cs.columbia.edu/graphics/courses/csw4172/resources/olwalFlexPointer.mov)

# Fishing Reel for pointing

- Use additional input to specify length of pointing vector
  - to disambiguate target, "reel" target in and out
    - change its position along the ray
  - separates DOFs, but not necessarily a problem

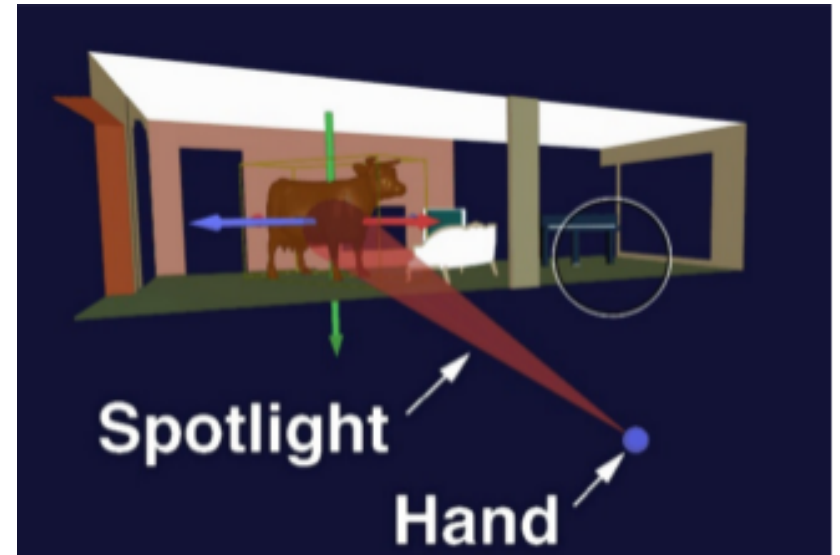


Bowman, 1997

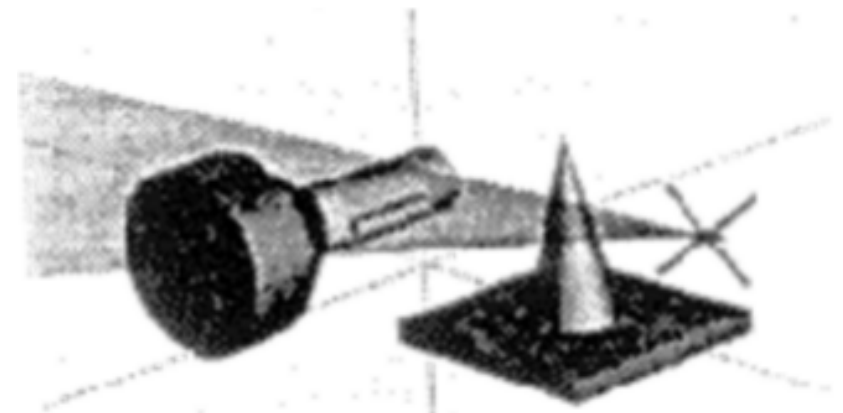
A variant of Vickers' use of buttons to move the selection cube along the pointing ray

# Flashlight (or spotlight)

- Replace ray casting vector with cone having apex at input device
  - easier to select small objects
  - inclusion criteria
    - object entirely inside cone
    - object partially inside cone
  - disambiguation rules for multiple objects in cone
    - Object closest to center line
    - Object closest to apex
      - if rule 1 is not enough
- What is the appropriate cone angle?



M. Mine, CHIMP, 1996



J. Liang & M. Green, JDCAD, 1993

# Aperture

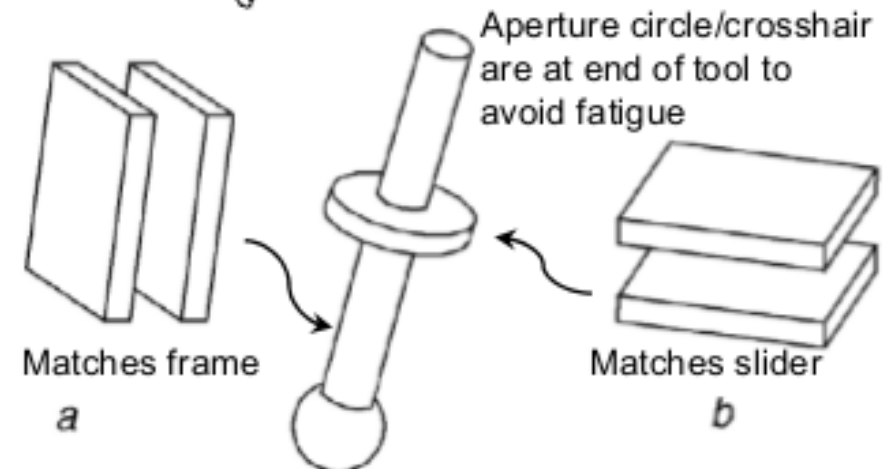
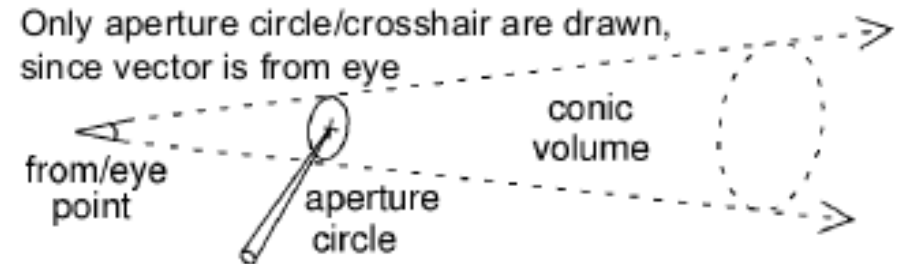
- Flashlight emanating from eye through hand
- Determine flashlight
  - vector from eye to hand (or hand-held tool) position
  - cone angle from distance from eye to hand (tool) position



A. Forsberg, K. Herndon, and R. Zeleznik, 1996

# Aperture with orientation

- Flashlight emanating from eye through hand
- Determine flashlight
  - orientation sensitivity from hand (tool) orientation
  - selected object is the one whose orientation is closest to that of a pair of parallel plates drawn near aperture geometry, controlled by tracker orientation

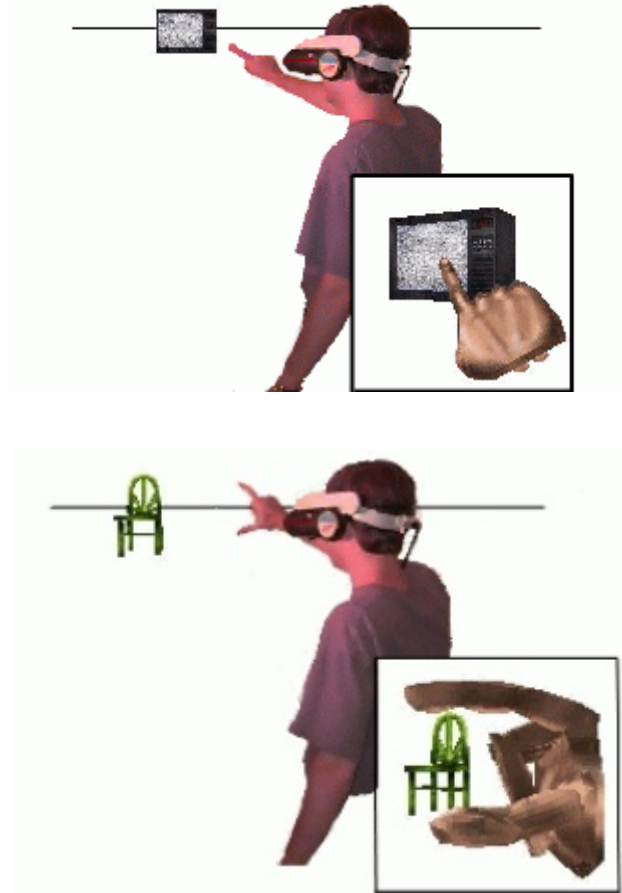


A. Forsberg, K. Herndon, and R. Zeleznik, 1996



# Pointing on an Image Plane

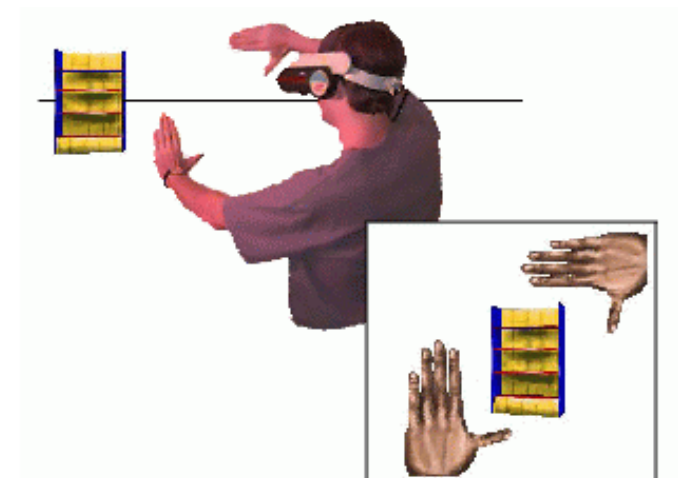
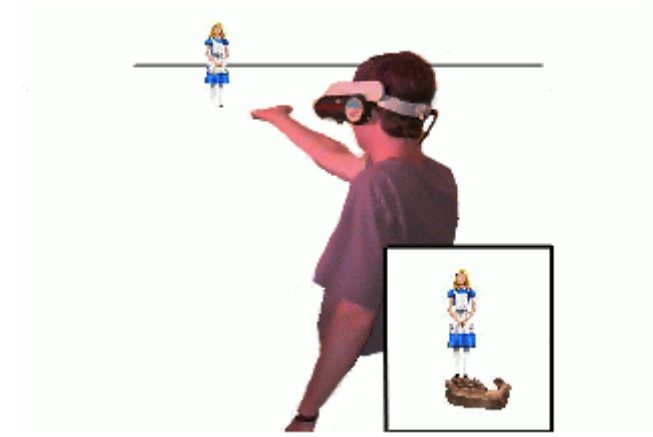
- Like eye-hand ray casting, but
  - All interaction occurs in 2D on the projection plane
  - After selection, to support manipulation, object can be scaled/moved to
    - pick point, maintaining approximate projected size, or
    - canonical size/location,...
  - “Sticky finger”
    - Object under finger is selected
  - “Head crusher”
    - Object framed by tracked fingers is selected
- Note: Need to designate a specific dominant eye in stereo, as with all single eye-defined pointing



J. Pierce et al., 1997

# Image Plane

- Lifting Palm
  - Object is selected by flattening his outstretched hand and positioning his palm below the desired object
- Framing Hands
  - uses both hands to form a "frame" to select objects
- Note: orientation of hands can be used to disambiguate objects as in aperture pointing.



J. Pierce et al., 1997

# Virtual Hand



- User controls 3D "cursor" (e.g. virtual hand model) intersected with objects
- Position/orientation of body part/device are mapped to virtual hand
  - Many versions of the "virtual hand" are really "virtual 3DOF/6DOF mice"
    - i.e. map to a single point, not a set of fingers
- User can confirm desired object by issuing trigger event
  - e.g. voice, button, gesture, ...
- Object is attached to hand for manipulation until released with additional trigger event

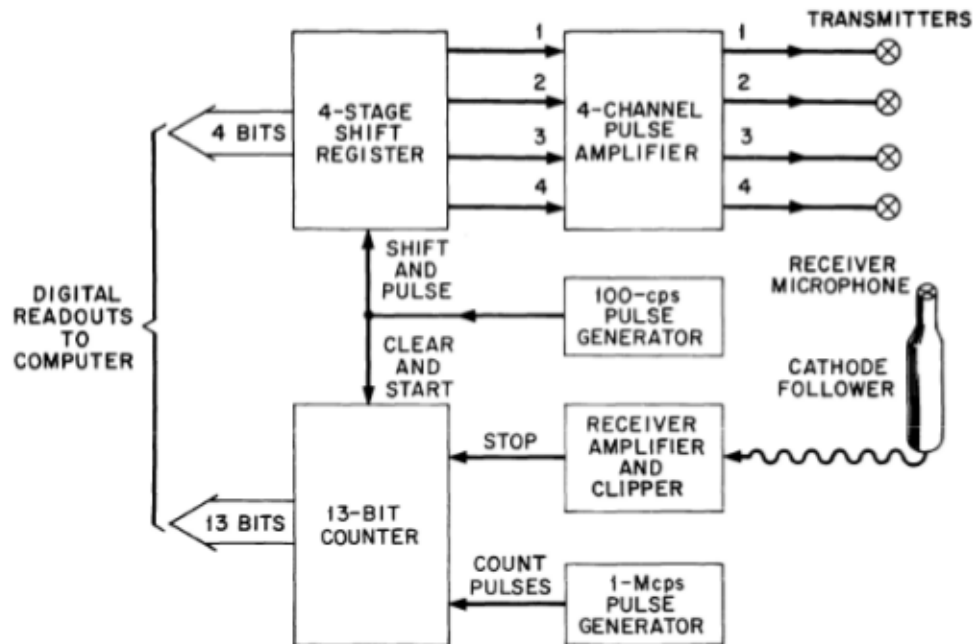
# Virtual Hand: Basic Version

- User's hand/device position/orientation are mapped directly to virtual hand's position/orientation with linear scale

J. Lanier, VPL, mid 80s–early 90s



# Virtual hand: origins



L. Roberts, The Lincoln Wand, 1966

Ultrasonic tracking of 3D wand tip position



M. Krueger, Videoplace, 1970s

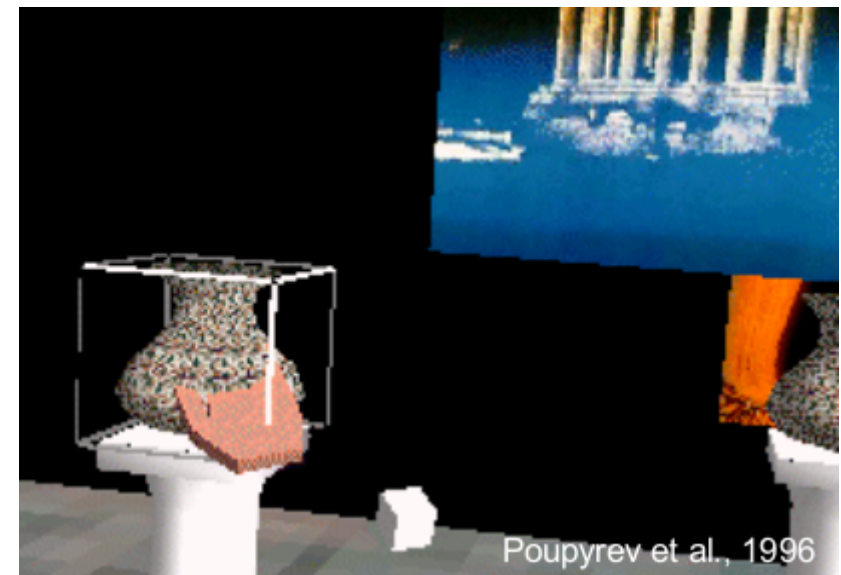
2D projections of real 3D hands are used to interact with 2D objects. Real time processing of thresholded video of user makes it possible to determine interesting properties useful for interaction

# Virtual Hand: limitations

- Isomorphic
  - Intuitive, but...
    - can only select objects within arm's reach
    - must travel to select objects farther away

# Virtual hand: Go-Go

- "Classic" nonisomorphic technique
- User's hand stretches when extended past a limit
  - Mapping function determines C-D (control-display) ratio
  - cube represents real hand
- Typically less effective than pointing for selection, but allows uniform approach to selection and manipulation



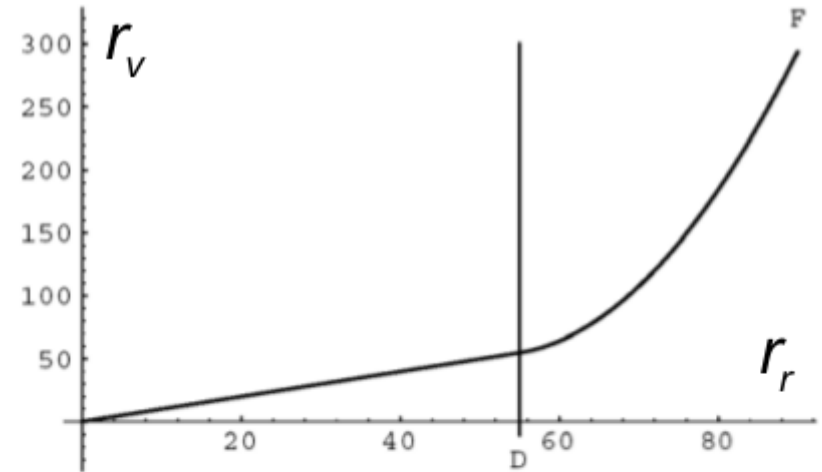
# Virtual hand: Go-Go

- Length of virtual arm

$$r_v = F(r_r) =$$

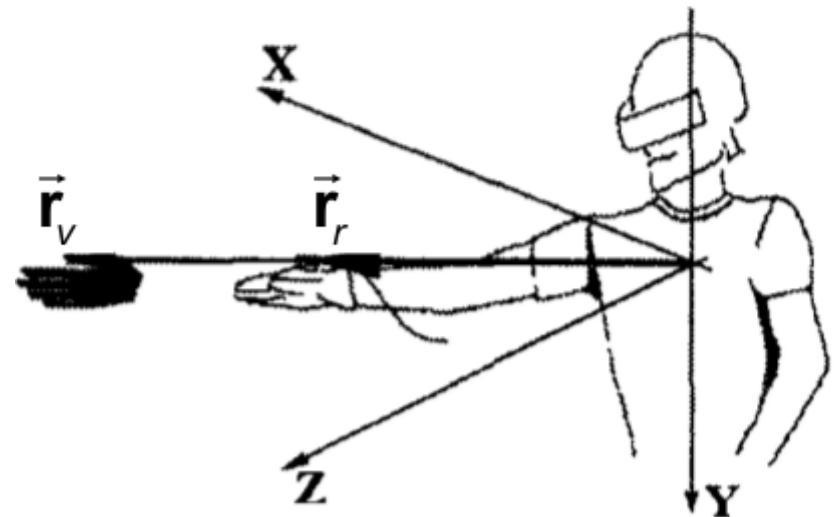
$$r_r \text{ if } r_r \leq D$$

$$r_r + a(r_r - D)^2 \text{ otherwise}$$



Poupyrev et al., 1996

where  $r_r$  is the length of the vector from the user's hand. The hand is in spherical coordinates ( $r_r$ ,  $\phi$ ,  $\theta$ ),  $r_v$  is the length of the virtual arm,  $F$  is the nonlinear mapping function,  $D$  and  $a$  are constants.





# Go-Go Variations

- Stretch go-go (Bowman & Hodges, 97)
  - Distance partitioned in 3 zones.
  - far: stretches at constant rate
  - middle: does not change
  - near: retracts at constant rate
- Indirect go-go
  - Buttons make hand stretch/retract (similar to Vickers)



Gauge at right indicates zones:  
green - stretch  
pink - stable  
blue - retract

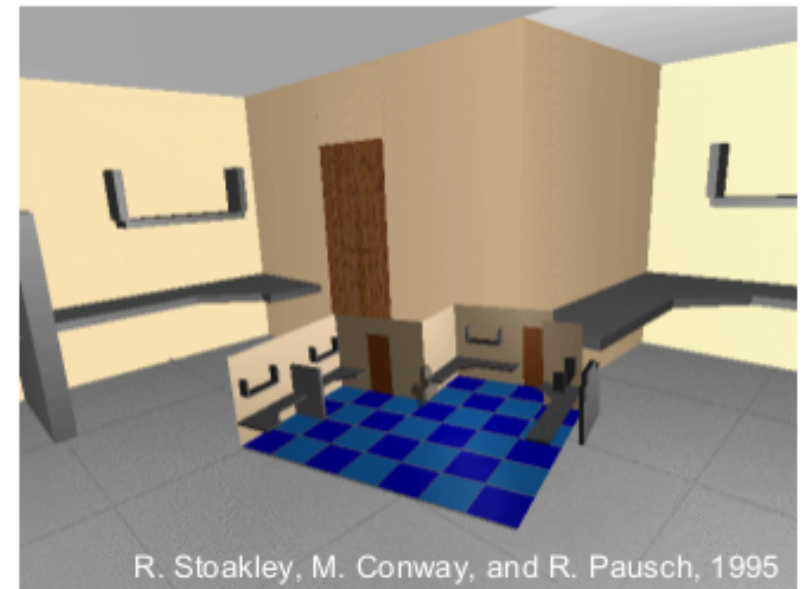
# Virtual Hand: Silk Cursor

- method to precise containment selection
  - emphasizing feedback
- 'Hand' is a 3D box cursor with low opacity ('silk') faces
- Moving cursor changes number of faces through which target is seen
  - this visual cue makes it easier to position box to contain target fully



# Selection aids - World in Miniature (WIM)

- User manipulates small-scale copy of virtual world
  - typically viewed exocentrically
  - can also be used for travel
  - updates are asynchronized between world and WIM
  - larger worlds can use selective display of WIM subset, with scaling and scrolling to select portion
  - need to avoid occlusion by walls
    - hack: model only interior wall surfaces for room and use backface culling to suppress when seen from outside



# Combining selection techniques

- When one technique is insufficient, combine techniques
  - choice
    - user chooses among alternative modes
    - e.g. menus, gestures, aural, ...
  - integration
    - system switches among techniques automatically based on task
    - e.g. switch between selection and manipulation, based on selection and release triggers

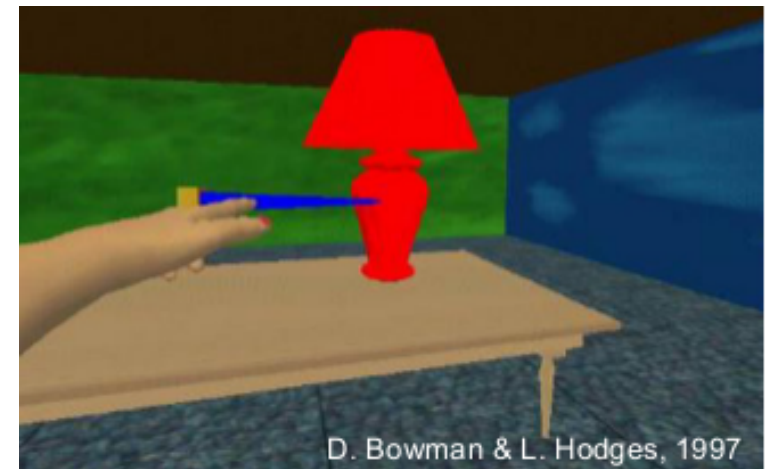
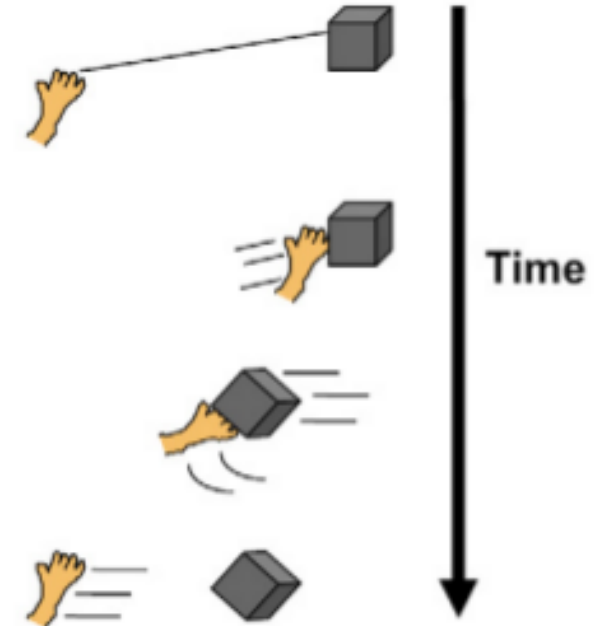
# HOMER (hand centered object manipulation extending ray-casting)

- Selection with pointing (ray-casting) switches to
- manipulation with virtual hand that is automatically moved and attached to object

$$r_v = a_h r_r, \quad a_h = D_0 / D_h$$

where  $a_h$  is a scaling factor defined at time of selection,  $D_0$  is distance to object,  $D_h = r_r$  at time of selection

- hand returns when object is dropped
- distance at which object can be placed depends on  $a_h$ 
  - Asymmetric: easier for moving far object near, than near object far



# HOMER variation

- Indirect HOMER
  - Buttons are used to translate virtual hand
    - like indirect go-go

# Scaled-World Grab

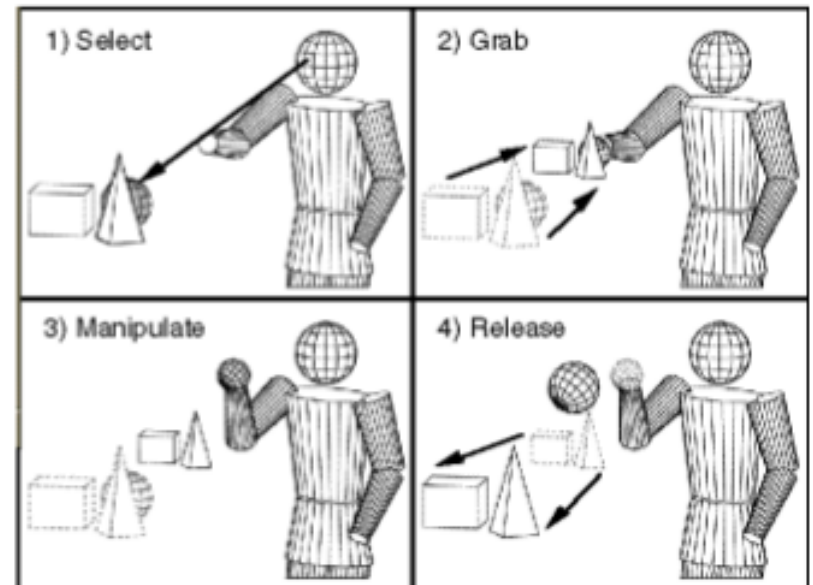
- Selection with image-plane technique causes temporary scaling of entire world around viewpoint by  $a_s$ , so object can be manipulated relative to world with virtual hand

$$a_s = D_v / D_o$$

where  $D_v$  is the distance from virtual viewpoint to virtual hand,  $D_o$  is the distance from virtual viewpoint to object at time of selection.

- User might notice no or little change, depending on locations of other objects, use of stereo
- World rescaled when released
- near-far asymmetry of manipulation because of scaling

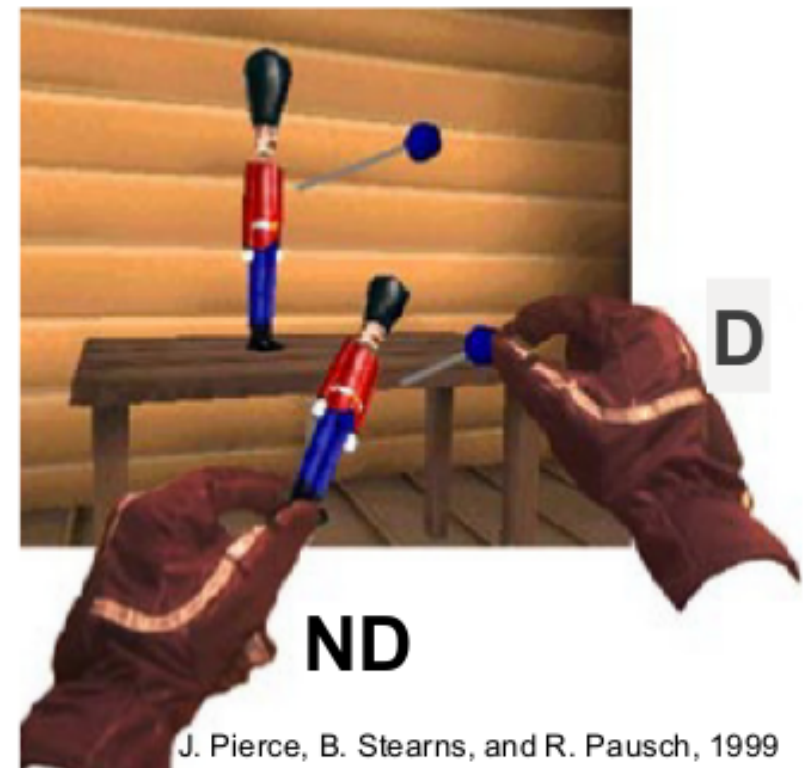
Note: object and world maintain relative spatial relationships unlike image-plane pointing, in which only one object is temporarily moved.



M. Mine, F. Brooks, & C. Sequin, 1997

# Voodoo Dolls

- Requires two 6DOF tracked hands
- Select object with image-plane technique
  - system creates doll (copy) of object in hand
  - pinch to hold the doll
- Dolls can be transferred between hands
- Dolls disappear when released

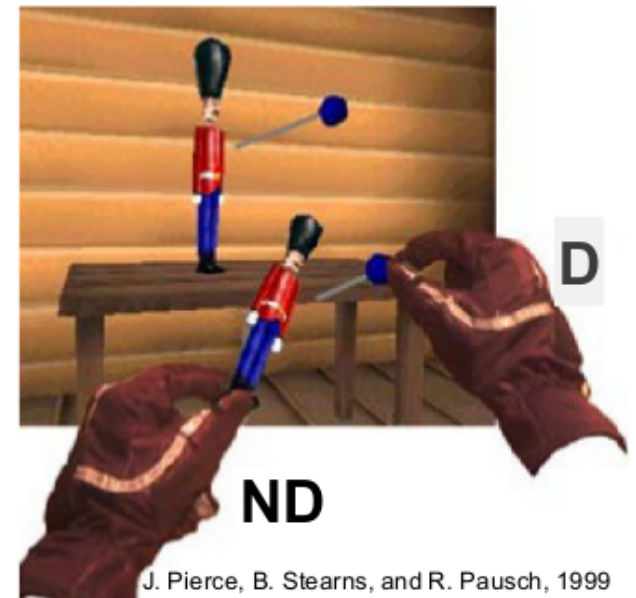


J. Pierce, B. Stearns, and R. Pausch, 1999



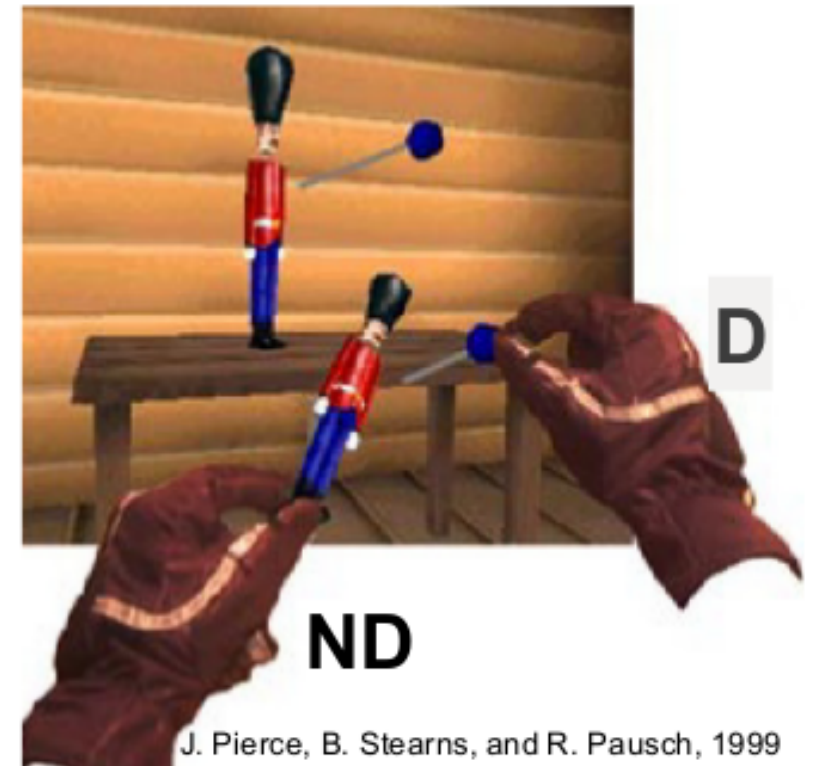
# Doll in nondominant hand

- Doll in nondominant (ND) hand
  - Scale to constant length - 50 cm along longest dimension
  - D doll is scaled proportionally
- Acts as reference frame for D doll (in dominant hand)
- Context: spawns new dolls that can be selected and manipulated by D hand
  - they represent the context around the selected doll
  - context can also be defined by 2 hands (framing)



# Doll in dominant hand

- Controls object's position and orientation relative to the ND doll
- Transformation of D doll relative to ND doll is applied to D doll's scene object
  - Transformation to ND doll does not affect the scene

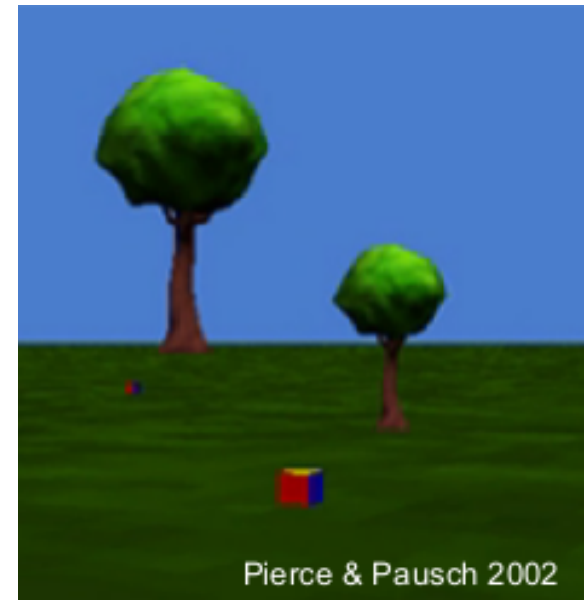


J. Pierce, B. Stearns, and R. Pausch, 1999

# Voodoo Dolls - limitations

- Hard to move D doll to position away from other objects, since ND doll is created by picking an object, and D doll is transformed relative to ND doll
- Extension (Pierce & Pausch, 2002)

selecting a point on the "ground" creates a temporary object represented as multi-colored cube, that can be an ND doll



# Design Guidelines

- 1. Use existing manipulation techniques unless a large amount of benefit might be derived from designing a new, application-specific technique
  - - no need to reinvent the wheel
- 2. Use task analysis when choosing a 3D manipulation technique
  - - how precise? how far? how large?
  - - positioning is required?

# Design Guidelines

- 3. Match the interaction technique to the device
  - how many DOF?
  - what is the shape of the device?
  - accuracy might require grasping
- 4. Use techniques that can help to reduce clutching
  - reducing clutching increases comfort and performance

# Design Guidelines

- 5. Nonisomorphic techniques are useful and intuitive
  - no need to be restricted to real-world, allow “magic” interactions
- 6. Use pointing techniques for selection and virtual hand techniques for manipulation
- 7. Use grasp-sensitive object selection
  - matches device orientation to desired object
- 8. Reduce degrees of freedom when possible

# Design Guidelines

- 9. Consider the tradeoff between interaction technique design and environment design
  - Tune interaction to maximize performance
  - Design your environment so that existing techniques can be used effectively
- 10. There is no single best manipulation technique
  - the ideal 3D manipulation technique has not yet been designed.
  - the best for one might not be for another