

CSE 165: 3D User Interaction

Lecture #14:
UI Evaluation

Announcements

- Final Project
 - Due Thursday, March 21st at 3pm
 - Two blog entries due 3/11 and 3/18
- Midterm to be returned and reviewed on Tuesday

3D UI Evaluation

Why User Evaluation?

- Need to compare
 - devices
 - interaction techniques
 - Applications
- Problem identification and redesign
- General usability understanding

Some Terminology

- Usability – everything about an artifact and what affects a person's use of an artifact
- Evaluator – person who designs, administers, implements, or analyzes an evaluation
- Subject – person who takes part in the evaluation

Evaluation Tools

- User task analysis
 - generates list of detailed task descriptions, sequences, user work, and information flow
- Scenarios
 - built from task analysis
 - important for experiment design
- Taxonomy
 - science of classification
 - break down techniques into components
 - used in evaluation process
- Prototyping
 - need to have something to test
 - paper-based sketches
 - Wizard of Oz approach

Evaluation Methods

- Cognitive walkthrough
 - Walk through every step of the study with expert
 - Reveals potential usability problems
- Heuristic evaluation
 - Done by user interface experts using a well-defined set of guidelines
 - Experts evaluate the UI by checking whether guidelines are being followed
 - Example guideline “Eliminate extraneous degrees of freedom for a manipulation task”
- Formative evaluation
 - observational user studies
 - questionnaires, interviews
- Summative evaluation: compare various techniques in a single experiment
 - task-based usability evaluation: more structured, aimed at improving the interface
 - formal experimentation: have a formal design, analyzed statistically
- Questionnaires
- Interviews

Evaluation Metrics – System Performance

- Average frame rate (fps)
- Average latency / lag (milliseconds)
- Variability in frame rate / lag
- Network delay
- Image distortion

Evaluation Metrics – Task Performance

- Speed / efficiency
- Accuracy
- Domain-specific metrics
 - education: learning
 - training: spatial awareness
 - design: expressiveness

Evaluation Metrics – User Preference

- Ease of use / learning
- Presence
- User comfort
- Usually subjective (measured in questionnaires, interviews)

User Comfort

- Simulator sickness
 - Kennedy - Simulator Sickness Questionnaire (SSQ)
- Aftereffects of VE exposure
 - Stanney 1998: Aftereffects from virtual environment exposure: How long do they last?
- Arm/hand strain
- Eye strain

Formality of Evaluation

- Informal
 - looser procedure
 - often more qualitative
 - subject comments very important
 - looking for broad usability issues
 - usually done during the design process to inform redesign
- Formal
 - independent & dependent variables
 - statistical analysis
 - strict adherence to a procedure
 - hold constant all other variables
 - usually done to compare multiple techniques or at the end of the design process

What is Being Evaluated?

- 3D application:
 - Prototype - consider fidelity, scope, form
 - Complete working system
 - Controlled experiments are rare
- Interaction techniques / UI metaphors
 - Can still evaluate a prototype
 - More generic context of use
 - Formal experiments more often used
 - Consider “Wizard of Oz” evaluation

Subjects / Participants

- How many people?
- What backgrounds?
 - technical vs. non-technical
 - experience with 3D/virtual reality applications
 - domain experts vs. general population
- What age range?
- Recruiting
 - flyers
 - email/newsgroups
 - psychology, cognitive science departments
 - CS classes

Number of Evaluators

- Roles
 - cable wrangler
 - software controller
 - note taker
 - timer
 - behavior observer

Procedure

- Welcome
- Informed consent
- Demographic/background questionnaire
- Familiarization with equipment
- Exploration time with interface
- Tasks
- Questionnaires
- Interview

Pilot Testing

- Pilot testing should be used to:
 - “debug” your procedure
 - identify variables that can be dropped from the experiment

Instructions

- How much to tell the subject about purpose of experiment?
- How much to tell the subject about how to use the interface?
- Always tell the subject what they should try to optimize in their behavior.
- If using think-aloud protocol, you will have to remind them many times.
- If using trackers, you will have to help users “learn” to move their heads, feet, and bodies – it doesn’t come naturally to many people.
- Remind subjects you are NOT testing THEM, but the interface.

Independent Variables

- Main variable of interest (e.g. interaction technique)
- Manipulated by the experimenter
- Conditions under which the tasks are performed
- The number of different values used is called level
 - Example: background can be blue, green, or white (3 levels)
- Secondary variables
 - task characteristics
 - environment characteristics
 - system characteristics
 - user characteristics

Dependent Variables

- Affected by the independent variables
- **Measured** in the user study
- Objective values: e.g., time to complete a task, number of errors, etc.
- Subjective values: ease of use, preferred option
- They should only depend on the independent variables (conditions)

Formal Experiment Decisions

- Choosing independent variables
- Choosing dependent variables
- Controlling (holding constant) other variables
- Within- vs. between-subjects design
- Counterbalancing order of conditions
- Full factorial or partial designs

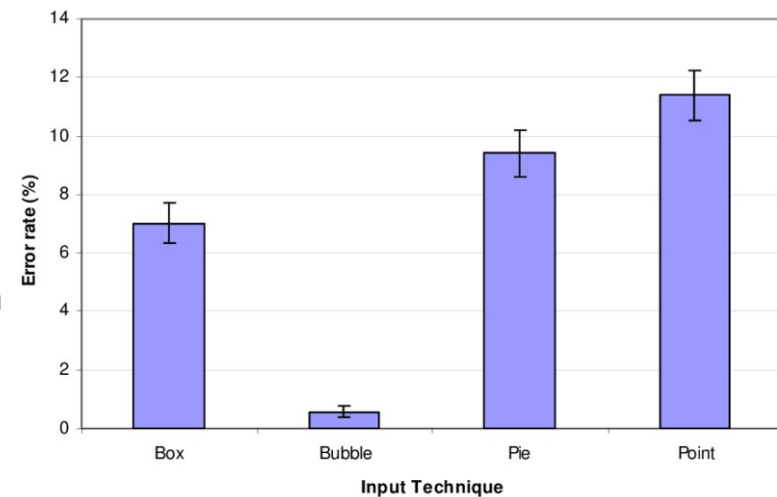
Metrics (dependent variables)

- Task performance time
- Task errors
- User comfort (subjective ratings)
- Observations of behavior (e.g. strategies)
- Spoken subject comments (e.g. preferences)
- Surveys/questionnaires
- Interviews

Data Analysis

- Averages (means) of quantitative metrics
- Counts of errors, behaviors
- Correlate data to demographics
- Analysis of variance (ANOVA)
- Post Hoc analysis (t-tests)
- Visual analysis of trends (i.e., to look for learning effects)

- *Interactions between variables* are often important
- Expect high variance in 3D UI interaction studies



Analysis Tools

- SPSS, SAS, etc.
 - UCSD site license for SPSS available
 - full statistical analysis packages
 - parametric and non-parametric tests
- Excel
 - basic aggregation of data
 - correlations
 - confidence intervals
 - graphs
- Matlab, Mathematica
 - more complex mathematical analysis