Design and User Evaluation of Augmented-Reality Interfaces

Team “Augment”

Members:
Payel Bandyopadhyay
Héctor Martínez

15.04.2014
Outline

- Introduction
- AR Application
  - Description
  - Functionality
- Usability methods
  - Cognitive walkthrough
    - Description
    - Experiment design
    - Result
  - Heuristic evaluation
    - Description
    - Experiment design
    - Result
- Lab observations
  - Description
  - Experiment design
  - Result
- Questionnaire
  - Description
  - Experiment design
  - Result
- Conclusion
Introduction

- What is Augmented Reality (AR) application?

*Any application having the following properties can be classified as an augmented reality application [1]*

- combination of real and virtual objects in a real environment
- interactive and real time operation
- registration (alignment) of real and virtual objects with each other

![Figure 1: Milgram’s reality–virtuality continuum [2]](image-url)
Chosen AR Prototype

- “Augment” - 3D Augmented Reality [3]

Figure 2: Logo of Augment [3]
“Augment is a mobile app that lets you and your customers visualize your 3D models in Augmented Reality, integrated in real time in their actual size and environment. Augment is the perfect Augmented Reality app to boost your sales and bring your print to life in 3 simple steps.” [3]
Software Requirement

- “Augment” is available only on the following OS:
  - iOS
  - Android
- OS used in this project:
  - Android
- Devices used for evaluation:
  - Android version of mobile
    - 4.1.1 (Jelly bean)
  - Android version of tablet
    - 4.0.4 (Ice cream sandwich)
System functionality

- Application can be used for following 2 purposes:
  - Sales and design
  - Interactive print

Figure 3: (a) A screenshot of Scan user interface [3]  (b) A screenshot of 3D Model user interface [3]
Example of usage

- Select 3D model
- Select marker
- See the AR content
- Interact (scale, rotate, take photo, share...)

Figure 4: A screenshot of user interface of Augment
# Usability methods (1/2)

Table 1: Various usability evaluation method and it’s corresponding category [4]

<table>
<thead>
<tr>
<th>Category</th>
<th>Usability evaluation methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inspection methods</td>
<td>Heuristics</td>
</tr>
<tr>
<td>Inspection methods</td>
<td>Cognitive walkthroughs</td>
</tr>
<tr>
<td>Inspection methods</td>
<td>Pluralistic walkthroughs</td>
</tr>
<tr>
<td>Inspection methods</td>
<td>Feature inspections</td>
</tr>
<tr>
<td>Inspection methods</td>
<td>Guideline checklist</td>
</tr>
<tr>
<td>Inspection methods</td>
<td>Perspective-based inspection</td>
</tr>
<tr>
<td>Testing methods</td>
<td>Co-discovery</td>
</tr>
<tr>
<td>Testing methods</td>
<td>Question asking protocol</td>
</tr>
<tr>
<td>Testing methods</td>
<td>Think aloud protocol</td>
</tr>
<tr>
<td>Testing methods</td>
<td>Performance measurement</td>
</tr>
<tr>
<td>Testing methods</td>
<td>Field observation</td>
</tr>
<tr>
<td>Testing methods</td>
<td>Laboratory observation</td>
</tr>
<tr>
<td>User reports</td>
<td>Interview</td>
</tr>
<tr>
<td>User reports</td>
<td>Questionnaire</td>
</tr>
</tbody>
</table>
Usability Methods (2/2)

- Cognitive walkthrough
- Heuristics
- Laboratory observation
- Questionnaire
Cognitive Walkthrough (CW)
What is Cognitive walkthrough?

- Exploratory learning
- Evaluators involved
  - Scribe
  - Facilitator
- CE+ theory [5]
  - The user sets a goal to be accomplished;
  - The user searches the interface for available actions;
  - The user selects an action that seems likely to make progress toward the goal; and
  - The user performs the action and checks to see whether the feedback indicates that progress is being made towards the goal.
CW Experiment design (1/3)

- Preparation phase
  - [link](https://docs.google.com/document/d/1K4wNk3aSBZvpkfKIRI_R1le4H4wDJ0PpCDSXnoF44E4/edit)
CW Experiment Design (2/3)

- Task description

“Select a furniture 3D model and place it in your room in a desired place and save it. The system will be in a state such that someone could immediately start testing.”
a) Open the application  
b) Choose the desired option  
c) Choose the desired 3D model  
   c1) Only for smartphone in vertical position: Swipe to see more buttons  
   c2) Select "create marker"  
   c3) Read and close help window  
   c4) Perform the scan  
d) Place it in your environment in a desired way:  
   i) Turn on flash (if required)  
   ii) Adjust the scale of the 3D model:  
      a) Make it big if required  
      b) Make it small if required  
   iii) Rotate it in your desired location  
e) Take a photo of it and save it

Figure 5 : Correct actions: Part of from cognitive walkthrough start-up sheet
CW Results (1/4)

- Evaluation phase

  - [https://docs.google.com/spreadsheet/ccc?key=0Aq8UBoDGSEVPdHdPLUoxaGpmVThvNExsNGVZUDBUUm&usp=drive_web#gid=0](https://docs.google.com/spreadsheet/ccc?key=0Aq8UBoDGSEVPdHdPLUoxaGpmVThvNExsNGVZUDBUUm&usp=drive_web#gid=0)
## CW Results (2/4)

### Figure 5 (a): Part 1 of cognitive walkthrough evaluation sheet

<table>
<thead>
<tr>
<th>We consider that the user understands the goal of the app and some basic knowledge of how AR works. Will the user try to achieve the right effect?</th>
<th>Step a: Open the application</th>
<th>Step b: Choose the desired option</th>
<th>Step c1: (vertical smartphone only): Swipe to see more buttons</th>
<th>Step c2: Select &quot;Create marker&quot;</th>
<th>Step c3: Read and close help window</th>
<th>Step c4: Perform the scan (1)</th>
<th>Step d: Place it in the right place</th>
<th>Step d: Turn on flash</th>
<th>Step d: Adjust the scale of the 3D model (applied in next columns)</th>
<th>Step d: Make it small if required</th>
<th>Step e: Take a photo of it and save it</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>2</td>
<td>3</td>
<td>5</td>
<td>5</td>
<td>-</td>
<td>No need</td>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

User will probably scroll down and find the 3D model. User may also use the search box for finding the model. It is clear that user needs to select one option when the "create option" appears, the user will probably select it. The user will try to scan the image. User will take the marker to the desired place. No flash is available. User may not understand that he/she is able to scale the model by pinching the screen. User may rotate the marker so that the object will be rotated. User may understand that the photo can be taken directly from this view.

### Will the user notice that the correct action is available?

| | 5 | 4 | 5 | 1 | 5 | 5 | 5 | 5 | 1 | 4 | 4 |

The user may understand that the option is one of the two available options. The option is "hidden" in the tool bar as only 3 icons are visible in the vertical position of the phone. The option for realizing is only explained in the help, so unless the user is using the help feature, the option is not visible. The option to rotate by using the software is not either visible nor intuitive. Icon is available with an image of a camera.
### CW Results (3/4)

<table>
<thead>
<tr>
<th>If the correct action is performed, will the user see that progress is being made towards the solution of the task?</th>
<th>5</th>
<th>4</th>
<th>4</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>4</th>
<th>5</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>If user selects the right option, the selection of models appear which may lead the user to think he/she is in the right direction. However, some users may think that this is not the right way to proceed</td>
<td>5</td>
<td>4</td>
<td>4</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>4</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

**Marking conventions:**
- **Totally wrong**
- **Quite wrong**
- **Confusing**
- **Quite obvious**
- **Totally obvious**

---

**Figure 5 (b): Part 2 of cognitive walkthrough evaluation sheet**
Table 2: Specific issues identified during our cognitive walkthrough. Problems marked with an asterisk (*) indicate problems that were discussed before the CW was done, but were also revealed by the CW.

<table>
<thead>
<tr>
<th>Description</th>
<th>Usability impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>After opening the application user may get confused with 2 options (*)</td>
<td>Serious</td>
</tr>
<tr>
<td>User may get confused in choosing the desired 3D model</td>
<td>Cosmetic</td>
</tr>
<tr>
<td>User may not know that swiping the menu bar will show more options</td>
<td>Critical</td>
</tr>
<tr>
<td>User may not be able to create the marker in a proper way</td>
<td>Critical</td>
</tr>
<tr>
<td>User may not be able to enlarge the 3D model as desired</td>
<td>Critical</td>
</tr>
<tr>
<td>Option to rotate the marker is not visible</td>
<td>Cosmetic</td>
</tr>
<tr>
<td>User may not use the ”help” menu (*)</td>
<td>Serious</td>
</tr>
</tbody>
</table>
Heuristics Evaluation (HE)
What is Heuristic evaluation?

- Evaluator
- Based on Nielsen’s 10 heuristics
<table>
<thead>
<tr>
<th>No</th>
<th>Heuristic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Visibility of system status</td>
<td>The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.</td>
</tr>
<tr>
<td>2</td>
<td>Match between system and the real world</td>
<td>The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.</td>
</tr>
<tr>
<td>3</td>
<td>User control and freedom</td>
<td>Users often choose system functions by mistake and will need a clearly marked &quot;emergency exit&quot; to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.</td>
</tr>
<tr>
<td>4</td>
<td>Consistency and standards</td>
<td>Users should not have to wonder whether different words, situations, or actions mean the same thing. Follow platform conventions.</td>
</tr>
<tr>
<td>5</td>
<td>Error prevention</td>
<td>Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.</td>
</tr>
<tr>
<td>6</td>
<td>Recognition rather than recall</td>
<td>Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.</td>
</tr>
<tr>
<td>7</td>
<td>Flexibility and efficiency of use</td>
<td>Accelerators – unseen by the novice user – may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.</td>
</tr>
<tr>
<td>8</td>
<td>Aesthetic and minimalist design</td>
<td>Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.</td>
</tr>
<tr>
<td>9</td>
<td>Help users recognize, diagnose, and recover from errors</td>
<td>Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.</td>
</tr>
<tr>
<td>10</td>
<td>Help and documentation</td>
<td>Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.</td>
</tr>
</tbody>
</table>
HE Experiment Design

- The members of the group have acted as the evaluators of the heuristics.
- The heuristic evaluation has been based on Nielsen’s 10 heuristics.
- Evaluator 1
  - Smartphone
  - Language of app: Spanish
- Evaluator 2
  - Tablet
  - Language of app: in English
HE Results (1/2)

- Evaluator 1 found the following usability problems
  - [https://docs.google.com/document/d/1WiP7kMHcWfNj_I83uboZ7qCoigMlAIaGKpcEzGmmpJ1z4/edit](https://docs.google.com/document/d/1WiP7kMHcWfNj_I83uboZ7qCoigMlAIaGKpcEzGmmpJ1z4/edit)

- Evaluator 2 found the following usability problems
  - [https://docs.google.com/document/d/1-hKAAWRNbpX3aiYc7iokqwdB6wS-vMgWfHo2eB9D8/edit](https://docs.google.com/document/d/1-hKAAWRNbpX3aiYc7iokqwdB6wS-vMgWfHo2eB9D8/edit)
Table 4: The total number of problems found in Heuristics evaluation phase. Second column shows problems found in “Augment” when installed in Smartphone. Third column shows problems found in “Augment” when installed in Tablet.

<table>
<thead>
<tr>
<th>Heuristic</th>
<th>Evaluator 1, number of problems found (# problem)</th>
<th>Evaluator 2, number of problems found (# problem)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Visibility of system status</td>
<td>10 (3, 4, 6, 7, 8, 10, 11, 13, 19, 29)</td>
<td>2 (1.i, 1.ii)</td>
</tr>
<tr>
<td>2. Match between system and the real world</td>
<td>9 (3, 8, 12, 16, 17, 18, 20, 21, 22)</td>
<td>1 (2)</td>
</tr>
<tr>
<td>3. User control and freedom</td>
<td>1 (2)</td>
<td>1 (3)</td>
</tr>
<tr>
<td>4. Consistency and standards</td>
<td>4 (1, 2, 7, 8)</td>
<td>1 (4)</td>
</tr>
<tr>
<td>5. Error prevention</td>
<td>5 (4, 14, 30, 31, 33)</td>
<td>5 (5.i, 5.ii, 5.iii, 5.iv, 5.v)</td>
</tr>
<tr>
<td>6. Recognition rather than recall</td>
<td>2 (7, 10)</td>
<td>2 (6.i, 6.ii)</td>
</tr>
<tr>
<td>7. Flexibility and efficiency of use</td>
<td>3 (7, 9, 26)</td>
<td>1 (7)</td>
</tr>
<tr>
<td>8. Aesthetic and minimalist design</td>
<td>3 (24, 25, 27)</td>
<td>0</td>
</tr>
<tr>
<td>9. Help users recognize, diagnose, and recover from errors</td>
<td>2 (3, 5)</td>
<td>6 (5.i, 5.ii, 5.iii, 5.iv, 9.i, 9.ii)</td>
</tr>
<tr>
<td>10. Help and documentation</td>
<td>1 (22)</td>
<td>1 (10)</td>
</tr>
</tbody>
</table>
More usability problems in the case of smartphone than in tablet:

- Smartphone is less powerful $\rightarrow$ More errors and crashes
- Smaller screen in smartphone $\rightarrow$ App layout is better designed for big screen $\rightarrow$ Lack of consistency
- Smartphone has been evaluated in Spanish $\rightarrow$ Severe translation problems
Laboratory observations (LO)
What is laboratory observation?

- Conducted in laboratories with test
- Not necessary to take place in dedicated “laboratory” [7]
- Controlled environment
- Can be conducted in various controlled environments [7]
  - office
  - hall way
  - simulator
  - others
- Mimic real life scenario
- Users
- Video/Audio Recordings
LO Experimental Design

- Users recruited
  - 7 users
  - Smartphone users
  - CS students
- Device provided
  - Samsung Galaxy Tab 10.1 (Touchscreen)
  - Model Number: GT-P7500
- Recordings
  - Video
  - Audio

Figure 5: Laboratory set up done
LO Results (1/5)

- Links to the video and audio recordings:
  - User 1: https://www.youtube.com/watch?v=o8DF-CXkX2c
  - User 2: https://www.youtube.com/watch?v=7vORzJ0s55A
  - User 3: https://www.youtube.com/watch?v=yGKHAsY_T_8
  - User 4: https://www.youtube.com/watch?v=Nz-Ciq3XFvc
  - User 5: https://www.youtube.com/watch?v=YnWGwb6_Tn8
  - User 6: https://www.youtube.com/watch?v=PXwrRO4KVPs
  - User 7: https://www.youtube.com/watch?v=1w934zTn0KE
LO Results (2/5)

Figure 7: Graph showing the task completion times of the 7 users. All the 7 users were novice users in terms of using Augmented Reality applications. All the 7 users were shown a live demo of the application and how to perform the task. User number 4 saw the demo once, saw the previous user performing the same task, did a demo herself before performing the experiment. Hence, the task completion time of user 4 was approximately half that of the average of other users but the user made a lot of errors while completing the task.
Figure 7: (a): The blue color represents the percentage of users who could find the desired 3D model very easily. The red color represents the percentage of users who could not find the desired 3D model easily. (b): The blue color represents the percentage of users who could easily create the tracker. The red color represents the percentage of users who could not create the tracker easily.
LO Results (4/5)

Figure 8: Number of users who tried to rotate the 3D model in other axis (x-axis and z-axis) other than the one provided in the current user interface.
Table 2: Number of identified usability problems (total number of individual problems found in all users and total number of common problems experienced by all users)

<table>
<thead>
<tr>
<th>Usability problems</th>
<th>Individual problems</th>
<th>Common problems</th>
</tr>
</thead>
<tbody>
<tr>
<td>Critical problems</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Serious problems</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>Cosmetic problems</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>
Questionnaire Evaluation (QE)
QE Description

- Performed
  - before or
  - after testing methods
- Users
  - User opinion
  - User preferences
- Statistics
QE Experiment Design (1/2)

- Questionnaire contained the following:
  - 13 statements
  - Marks 1-5
  - Comment option

Figure 9: Part of questionnaire
QE Experiment Design (2/2)

https://docs.google.com/viewer?a=v&pid=forums&srcid=MDc1MTI4NzA1MjM0MzEwMjk3MjABMTU5MTA4ODE2OTY5NjkwNTM1NDgBaUpaa0kwamZzamtKATQBAXYy
Figure 10: Results from the questionnaires. The image shows the frequency of every mark for each statement. The reader should note that there were 4 statements (number 4, 5, 8 and 12) that were not answered by all users.
QE Results (2/2)

- Users have rated positively the questionnaire
- They have been more concentrated in the novelty of AR than in the usability
- For future:
  - Tests without pre-instructing users
  - Tests with smartphone
  - Tests with users with previous knowledge of AR
  - Tests in other languages
Conclusions (1/5)

- AR new technology
- Novice users (90%)
- Current Learnability curve
  - high (Figure 11)
  - Learnability curve of user interface of “Augment” should be lowered
- Alternative of help menu
  - none of the users used it
- Informative feedback
  - addition needed
- Usability methods from 3 categories should be employed
  - few unique problems were found in each of 4 methods
Figure 11: Red color represents percentage of users who faced both the problems. Prussian blue color represents percentage of users who did not face any problem. Light blue color represents user who found selecting 3D model a problem but could easily create the tracker. Yellow color represents percentage of users who could easily select a 3D model but could not create the tracker.
Conclusions (2/5)

- Consistency should be added
  - phone version explain why tracker is not working
  - tablet version explain why tracker required
- Rotation button
  - y-axis (available)
  - x-axis (addition required)
  - z-axis (addition required)
- “Email” and share options
  - need to be improved
Conclusions (3/5)

- “Search” button
  - modification required
  - users ignored
- Catalogue of models
  - Presentation modification
- Possibility of changing options
  - “scan”
  - “browse”
Conclusions (4/5)

- “Augment” design guidelines:
  - Follow the standards of Android platform.
  - Provide more intuitive interfaces and a more organized “option dialog”.
  - Improve the manipulation of 3D objects.
  - Provide more information about AR concept and about what a good marker is.
  - Provide help tips in the appropriate context.
  - Provide better translations of the languages.
Conclusions (5/5)

- AR evaluation design guidelines:
  - Combining one usability inspection method with one usability testing method is recommended to obtain a reliable outcome.
  - Using more than one expert in the inspection methods is suggested.
  - If questionnaire method is going to be used, the number of users to fill it should be large enough and contain a variety of users, including AR experts.
References


