Effects of Viewing Conditions and Rotation Methods in a Collaborative Tabletop AR Environment

Sangyoon Lee            Hong Hua
3DVIS Lab, College of Optical Sciences, University of Arizona

Outline

• Introduction
• Related Work
• Viewing Conditions and Rotation Methods
• Test Environment and Apparatus
• Tasks and Metrics
• Participants and Procedure
• Results
• Conclusion and Future Work
Introduction

• Collaborative tabletop AR environments
  - Computer generated virtual imagery is augmented on physical objects or props (unlike from conventional tabletop display environments).
  - Inherently supporting natural face-to-face interactions among users (like the conventional tabletop display environments).

• Important features of collaborative tabletop AR environments
  - Able to easily provide individualized view for particular attributes (such as rendering method, color, scaling factor, rotation angle, etc.).
  - Support direct physical interaction with the digital contents via direct manipulation of the physical objects or props (i.e. interface tangibility).

Effects of Individualized View

• Positive effects
  - Individual view control may allow participants to perform certain tasks in more efficient and parallel manner.
    • E.g. applying an individual rotation angle for the tabletop world allows participants to individually rotate the world without significantly changing their viewpoints and disturbing others’ views.

• Negative effects
  - Applying individualized view may cause confusion in referring-type of user interaction.
    • E.g. when an individual rotation angle (or scaling, translating factor) of the tabletop world is allowed, location mismatch (or confusion) problem can arise.
    • Location mismatch problem: the location at which a user is pointing is shown as a different location in the others’ views.
  - Applying individualized view may compromise the sense of sharing and collaboration.
**Effects of Direct Interaction**

- **Positive effects**
  - Direct manipulation of physical objects/props may be easier and more intuitive to use than indirect interaction methods.
  - It also provides an explicit visual cue on collaborator’s action, which promotes the awareness of collaborator’s intentions and sense of collaboration.

- **Negative effects**
  - It requires users to take turns for access to physical objects/props (if the objects are shared).
  - It usually requires users to make more physical effort to use than indirect interaction methods.

**Purpose of This Study and Considered Factors**

- **Purpose**: investigating the effects of visualization consistency of a shared scene among collaborators and the directness of user interfaces on different types of collaborative tasks in a two-user co-located, tabletop AR environment.

- **Considered factors**:
  - **Viewing condition**: how the manipulation of a tabletop world made by one user is shown in the other users’ views.
    - Consistent (or synchronized) view condition
    - Inconsistent (or non-synchronized) view condition
  - **Rotation method**: what type of input devices is used to rotate the tabletop world for alternative orientations.
    - **Direct turn**: using a physical turntable device
    - **Indirect turn**: using buttons on two wireless mice
Task Types and Our Hypothesis

- Considered task types: orientation-strong (and referring-weak) type and referring-strong (and orientation-weak) type.
  - Orientation-strong type tasks: the tasks requiring relatively many orientation changes of the tabletop world.
  - Referring-strong type tasks: the tasks requiring relatively many references to locations or objects.

- Informally classifying tasks in terms of degree of orientation and reference dependency since we think orientation and reference are the most important factors for collaborative tasks in tabletop environments.

- Our hypothesis: The viewing conditions and rotation methods would have different effects on task performances and collaboration experience according to the task types.

Related Work (1/3)

- Using AR environments for collaborative work
  - Two pioneer studies: Studierstube by Schmalstieg et al. (1996) and Shared space by Billinghurst et al. (1996).
  - Collaborative AR environments have been applied for various application areas such as entertainment (Ohshima et al., 1998), education (Kaufmann and Schmalstieg, 2003), visualization (Fuhrmann et al., 1998), and so on.
• Usability evaluations related to collaborative AR environments
  - Kiyokawa et al. (1999, 2002) showed the superiority of a shared AR environment over a shared VR environment for face-to-face collaboration, and found that both display configurations and task space configurations affected two-user communication behaviors (2002).
  - Wang et al. (2007) showed that collaborative AR environment had several potentials for urban design.
  - Chastine et al. (2007) showed importance of “inter-referential awareness” in a collaborative AR context, through an experiment exploring various referencing techniques in different environments.

• Most of previous studies (except the study by Kiyokawa et al. (2002)) focused on comparing collaborative AR environments with other different environments.

• In this study, we compare different interface configurations “within” a collaborative AR environment.
**Viewing Conditions**

- **Consistent (synchronized) view**
  - Identical to the real-world situation.
  - Changing the rotation angle of the tabletop world by one user also changes the world orientation by the same amount in the other user's view.
  - No location mismatch problem.

- **Inconsistent (non-synchronized) view**
  - The rotation angle change is individually applied.
  - Rotating the tabletop world by one user does not affect the world orientation in the other user's view.
  - May provide better comprehension for the tabletop world, but location mismatch problem can arise.

---

**Location/Object Highlighting**

- The location mismatch (or confusion) may make it extremely difficult for users to understand a location or object referenced by their collaborators.

- To alleviate the location mismatch (or confusion) problem, we use location/object highlighting.
  - Each user is provided with a 3D pointer on which a virtual ray is augmented.
  - When a user points her/his own 3D pointer at a location or object, the intersected one with the virtual ray is highlighted in both users' views.

- In the inconsistent view condition, the highlighted one by a user is not affected by the collaborator's rotating the tabletop world.
Rotation Methods

- **Direct turn**
  - Adopting a tangible input technique where users rotate virtual world by directly turning a physical tabletop surface equipped with a turning motor and an angle measurement device.
  - Orientation angles of the physical turntable are mapped 1:1 to those of the tabletop world.

- **Indirect turn**
  - Adopting a non-tangible UI technique using buttons of two wireless mice (one for each user).
    - Left button: the world is rotated clockwise.
    - Right button: the world is rotated counter-clockwise.
  - The rotation speed was fixed at 50 degrees/sec.

Test Environment and Apparatus

- **Stereoscopic Collaboration in Augmented and Projective Environments (SCAPE), consisting of:**
  - Two polarized Head-Mounted Projection Displays (p-HMPD), one for each user
    - Resolution: 1280x1024 pixels
    - Diagonal FOV: 55 deg.
  - Retro-reflective turntable screen
    - Diameter: 72 cm
  - Various interface devices (such as 3D pointers and wireless mice).
  - HiBalls and Flock of Birds for head and 3D pointer tracking, respectively.
3D Pointers and Wireless Mice

• 3D Pointers: Used for object pointing and manipulation (with a pointing-and-selection metaphor)

• Wireless Mouse

Tasks

• Two tasks considered
  - Lego-like block building task (shortly, Lego task): referring-strong type.
  - Text label selection task (shortly, text label task): orientation-strong type.
  - Both tasks involved two-user collaboration.

• Lego-like block building task (referring-strong type)
  - Task goal: To construct a duplicate of a given end-product model in the working area, block-by-block, as quickly as possible.
  - User roles: One user was a guider, and the other was a builder.
    • The role of guider: To direct how the blocks are assembled and what should be built.
    • The role of builder: To actually place the blocks in the workspace to build the model with the guider’s guidance.
Task 1: Lego-like Block Building Task

- The end-product model was shown only to the guider.
- The 3D pointers provided different functions according to the role.
  - For builder, 3D pointer was used for manipulating blocks and pointing at positions in the working area or specific blocks that have already been laid.
  - For guider, 3D pointer was used only for pointing at positions or blocks.
- Note: Both users could use their 3D pointers with verbal communication for mentioning a position or a block to their collaborators.

Example of Block Pointing and Manipulation

- Red ray on the right controlled by the builder, and blue ray on the left by the guider.
Task 2: Text Label Selection Task

- Designed for the orientation-strong type.
- Task goal: To tag every of 9 text labels with two point values using given keyword lists (tagging stage), and then to select 4 text labels that had the highest sum of two point values (selection stage), as quickly as possible.
- No difference in role between two users.
- An example of a text label

<table>
<thead>
<tr>
<th>garlic</th>
<th>ginger</th>
<th>bass</th>
<th>scallion</th>
<th>salmon</th>
</tr>
</thead>
</table>

- Each user was required to determine a point value and a selection status for each text label, based on words in that text label and words in her/his own keyword list.
  - Left (blue) number and the filling status of left box were determined by one user only; the right counterpart was determined by the other user only.

Keyword List

- For each trial, there were 2 different keyword lists: one for User 0 and the other for User 1 (The 2 keyword lists are not shared).
- Example of keyword list

<table>
<thead>
<tr>
<th>Keywords and their point values for User 0</th>
<th>Keywords and their point values for User 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>cocoyam 1</td>
<td>girasole 1</td>
</tr>
<tr>
<td>fennel 2</td>
<td>scallion 2</td>
</tr>
<tr>
<td>garlic 3</td>
<td>ginger 3</td>
</tr>
</tbody>
</table>

- In each keyword list, words in the left column mean keywords, and the numbers in the right column mean point values for corresponding keywords (e.g. for User 0, cocoyam has a point value of 1, and for User 1, scallion has a point value of 2).
- For all trials, keyword lists have the same format, but word sets in the lists were different across trials. (There were only three point values: 1, 2, and 3.)
- A keyword list was given to each of two users before each trial.
Determining Point Values of a Text Label

- To determine user’s point value of a text label
  - Find all of user’s keywords appearing on the text label, and then sum their corresponding point values. The summed point value is user’s point value for that text label.

- Example

  - For User 0, the point value of the text label is
    \[ 3 \text{ (for garlic)} + 2 \text{ (for fennel)} = 5 \]

  - For User 1, the point value of the text label is
    \[ 3 \text{ (for ginger)} + 2 \text{ (for scallion)} + 3 \text{ (for ginger)} = 8 \]

Text Label Attachment and Examples of Manipulation

- Text label attachment
  - Text labels were attached to 5 virtual Objects or turntable surface.
  - The positions and orientations of text labels, and words shown on the labels were different across trials.

- Examples of text label manipulation
  - Tagging point values (using 1st and 2nd buttons)
  - Selecting a text label (using 3rd button)
Metrics

• For both tasks,
  - Task completion time
  - Turn angle: The sum of the rotation angle changes of the world in both users’ views
  - Subjective questionnaire: Co-presence, Preference, etc.

• Only for Lego task,
  - The number of expressions used by builder for confirming positions
  - The number of different blocks between the builder-created model and the given end-product model (quality of final product)

• Only for text label task,
  - The number of negotiation expressions used for scene-orientation changes
  - The number of text labels of which point values were correctly tagged (quality of final product)
  - The number of correctly-selected text labels (quality of final product)

Participants and Procedure

• 32 participants (16 groups)
  - 7 females and 25 males (ranged from 18 to 39 years).
  - 8 groups were already familiar with their collaborators.

• For each group, 2 sessions run, each session on a separate day.

• In each session,
  - The group involved in only one task (8 groups performed the Lego task first).
  - 2 training trials followed by 4 evaluation trials (1 evaluation trial for each of 4 test combinations): within-subject repeated measures design.
  - The order of presentations for 4 test combinations was arranged with a digram-balanced Latin square design methodology.
Results: Effects of Viewing Conditions (1/5)

- Task completion time

  ![Graph showing task completion time for consistent and inconsistent views for Lego and text label tasks.]

  For Lego task: $F(1,15) = 13.26, p = 0.0024$

  For text label task: $F(1,15) = 8.21, p = 0.0117$

  - Significant differences for both tasks.
  - However, the effects of viewing conditions were opposite: for Lego task, faster in consistent view, but for text label task, faster in inconsistent view.
  - For Lego task, the reason can be explained with the results on the number of expressions used by builder for confirming positions.
  - For text label task, the reason can be explained with the results on turn angle and the number of negotiation expressions.

Results: Effects of Viewing Conditions (2/5)

- For Lego task: The number of expressions used by builder for confirming positions

  ![Graph showing the number of expressions used for confirming location by consistent and inconsistent views.]

  For Lego task: $F(1,15) = 11.38, p = 0.0042$

  - Significant difference.
  - In consistent view, expressions for confirming positions were less used.
    - Because there was no location mismatch problem.
  - For Lego task, using the expressions less led to faster task completion time.
Results: Effects of Viewing Conditions (3/5)

- For text label task: Turn angle & the number of negotiation expressions used for scene-orientation changes
  - Significant differences.
  - In inconsistent view, participants made more turns and negotiated less for making turns.
  \[ \Rightarrow \] Implies that they could more easily change the orientation.
  - For text label task, easy orientation change led to faster task completion time.

Results: Effects of Viewing Conditions (4/5)

- Co-presence (the sense of being together in an environment)
  - Grading: 7-scale Likert-style (1 = co-located with my partner, 7 = staying away from my partner)
  - Significant differences for both tasks.
  - Although the participants knew they were with their collaborators in the same physical space, they felt less co-presence in inconsistent view condition.
  \[ \Rightarrow \] Interesting, but further investigation is required for explaining the reason.
Results: Effects of Viewing Conditions (5/5)

- User Preference

• For Lego task, no significant difference, but for text label task, significant difference.
• Only for text label task, inconsistent view was more preferred by the participants.

For Lego task: $\chi^2(1) = 1.58$, $p = 0.2087$
For text label task: $\chi^2(1) = 27.1$, $p < 0.0001$

Results: Effect of Rotation Methods (1/2)

- Only on turn angle, significant differences were found for both tasks.
- For both tasks, the participants made more turns with indirect turn.
- However, the differences between rotation methods did not have an effect on task completion time. (There was no differences on task completion time.)

For Lego task: $F(1,15) = 23.07$, $p = 0.0002$
For text label task: $F(1,15) = 25.57$, $p = 0.0001$
Results: Effect of Rotation Methods (2/2)

- User Preference

- No significant difference for both tasks.
- Regardless of tasks, two rotation methods were equally preferred.

For Lego task: $\chi^2(1) = 0.50$, $p = 0.4795$
- No significant difference for both tasks.
- Regardless of tasks, two rotation methods were equally preferred.

For text label task: $\chi^2(1) = 0.29$, $p = 0.5900$

Conclusion and Future Work (1/2)

- In this study, we investigated the effects of viewing conditions and rotation methods for the two different types of collaborative tasks.

- The consistent view showed better task performance in the referring-strong type task (i.e. the Lego task), but the inconsistent view showed better performance in the orientation-strong type task (i.e. the text label task).
  - Expecting similar results for the same types of other tasks.

- The rotation method factor had no significant effects on all of the measurements except turn angle.
Conclusion and Future Work (2/2)

- In this study, orientation angle of the tabletop world was considered as a particular attribute for the individual view.

- Our future work
  - To investigate the effects of other attributes such as rendering method, scaling and translating factors, etc, on collaborative task performance.
  - To investigate the effects of viewing conditions on co-presence.

Thank you!
p-HMPD Technology

- Exploiting micro projectors and retro-reflective screens [Zhang and Hua, AO 2009].
  - Light beams emanated from micro projectors pass through an optical system and arrive on a retro-reflective screen.
  - When the light beams strike the screen, the beams are reflected back to the source and reach user's eyes.

Difference between 2 viewing conditions

Consistent view condition

Inconsistent view condition

When you make turns (in your view)

When your partner makes turns (in your view)
Example: Carrying out Lego Task

Selection Stage: Selecting 4 text labels

- After determining and tagging the two point values (one value for each user) on each of 9 text labels, two users was asked to select the 4 text labels that had highest sum of the two point values to finish the task.

- To select the 4 highest-value text labels,
  1. Find and select the text label that has the highest total point value.
  2. Find and select the text label that has the second highest total point value.
  3. Find and select the text label that has the third highest total point value.
  4. Find and select the text label that has the fourth highest total point value.
Example: Carrying out Text Label Task

• Goal: Tag a point value for each of 4 text labels with your own keyword list, and then select the 2 text labels that have highest total point values, as quickly as you can.

• Given the following keyword lists, one for each user
  • The word sets on each text label

<table>
<thead>
<tr>
<th>Text label 1</th>
<th>Text label 2</th>
<th>Text label 3</th>
<th>Text label 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>tuna garlic tuna</td>
<td>garlic ginger</td>
<td>ginger cocoym</td>
<td>taimen fennel</td>
</tr>
<tr>
<td>garlic blenny</td>
<td>bass scallion</td>
<td>salmon girasole</td>
<td>scallion garlic</td>
</tr>
<tr>
<td>girasole cocoyam</td>
<td>ginger salmon</td>
<td>shark cocoym</td>
<td>shark scallion</td>
</tr>
</tbody>
</table>

1st stage: tagging point values for User 0

<table>
<thead>
<tr>
<th>Text label 1</th>
<th>Text label 2</th>
<th>Text label 3</th>
<th>Text label 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>tuna garlic tuna</td>
<td>garlic ginger</td>
<td>ginger cocoym</td>
<td>taimen fennel</td>
</tr>
<tr>
<td>garlic blenny</td>
<td>bass scallion</td>
<td>salmon girasole</td>
<td>scallion garlic</td>
</tr>
<tr>
<td>girasole cocoyam</td>
<td>ginger salmon</td>
<td>shark cocoym</td>
<td>shark scallion</td>
</tr>
</tbody>
</table>

3 (garlic) + 3 (garlic) + 1 (cocoyam) = 7
3 (garlic) = 3
1 (cocoyam) + 1 (cocoyam) = 2
2 (fennel) + 3 (garlic) = 5
1st stage: tagging point values for User 1

- Text label 1
  - tuna garlic tuna
garlic blenny
girasole cocoyam
  - 1 (girasole) = 1

- Text label 2
  - garlic ginger
bass scallion
ginger salmon
  - 3 (ginger) + 2 (scallion) + 3 (ginger) = 8

- Text label 3
  - ginger cocoyam
salmon girasole
shark cocoyam
  - 3 (ginger) + 1 (girasole) = 4

- Text label 4
  - taimen fennel
scallion garlic
shark scallion
  - 2 (scallion) + 2 (scallion) = 4

2nd stage: selecting 2 highest-value text labels

- Result of 1st stage
  - Text label 1: 7 1
  - Text label 2: 3 8
  - Text label 3: 2 4
  - Text label 4: 5 4

- Total point value of each text label
  - For Text label 1, 7 + 1 = 8
  - For Text label 2, 3 + 8 = 11
  - For Text label 3, 2 + 4 = 6
  - For Text label 4, 5 + 4 = 9

- Selecting the 2 highest-point-value text labels
  1. Select the text label that has the highest total point value.
     Text label 2 is the highest.
  2. Select the text label that has the second highest total point value.
     Text label 4 is the second highest.
Result of the task

7 1
tuna garlic tuna
garlic blenny
girasole cocoyam

8 3
garlic ginger
bass scallion
ginger salmon

2 4
ginger cocoyam
salmon girasole
shark cocoyam

4 5
taimen fennel
scallion garlic
shark scallion

Demonstration: Text Label Task