TRANSITIONING BETWEEN 3D AND AR IN WAYFINDING SYSTEMS

Interactive Systems Spring 2014

GROUP MEMBERS

Ville Niemi
013461649

Lei Wang
014343711

Ayesha Ahmad
014109366

Matti Ruoho
013312730

Contents

1 Introduction

3D-maps and Augmented Reality systems can be used for example when visiting a cultural attraction or location and using a mobile device to navigate at the venue. The Augmented reality interface can be used to get contextual information on various objects at the location, while the 3D maps can be used for simple map navigation. The aim is to improve user experience by allowing switching between 3D and Augmented Reality (AR) views of the application. The type of interactions the user needs to employ to switch views within such a system must be seamlessly integrated into the system in such a way that the user feels the application is natural, intuitive, and effortless to use. The user interface and design of such an application is critical to the user experience; it will determine whether the user finds the system burdensome and a pain to use, or facilitative and easy to employ.

The purpose of this project is to study which are the most suitable interaction techniques for the system, mainly those concerning transitioning between 3D and AR views. These techniques include finger gestures and traditional GUI techniques like button-press. Several interaction techniques have been implemented and tested and two user studies have been performed with a paper and software prototype respectively. The paper prototype was used to determine the interaction techniques favoured by users out of a larger selection of possible interactions. The most preferred interactions were then implemented in the software prototype and users' opinions were collected a second time.

The results of user evaluation of the paper prototype made it quite clear that users prefer well known established gestures to others. As many map navigation applications are already in existence (notably Google Maps[1]) all of our testers were already familiar with map navigation on mobile devices, and none of them had used an AR application. Novel gestures that may have been more efficient were nonetheless rejected by users as not being 'intuitive' or too difficult to use. The final result was that users want to use gestures that are well-known to them. In the opinion of the researchers, it is possible that users may begin to prefer the novel gestures if they were able to use the software for a longer period. This is because these gestures provide shortcuts for longer tasks like

[1] http://maps.google.com

zooming in and out of views which normally require multiple touch inputs.

2 Background & Related Work

Mobile Augmented reality (AR) systems have been much studied and developed in recent years. A common use of these applications is to enable the user to examine their environment through a mobile tablet or smartphone and extract information about objects of interest in the environment. In the AR view the user is usually provided a camera view from their viewpoint with objects of interest visualized on the screen which the user can examine or manipulate by some interaction techniques.

In the 3D-view, the user can examine a 3D-model of the location and navigate in the 3D-map to more closely examine places that are of interest to the user. This requires that a 3D model of the location (this location can be for example a museum, a university campus etc.) is available to the program. These models are commonly obtained from a database, for example from a website. Lots of research has also been done to model the object 'on the fly', that is, the 3D modelling of the location is calculated and rendered dynamically on the scene with some input sensors on the device to calculate a proper 3D model of the environment in question, without downloading an already existing model from a library for example. The overall problems of creating a valid 3D model is also an area of research that has been paid much interest. Common issues related to 3D modelling are the problematic occlusion issues among the objects in the map and specifically dense graphs, that is, where there are lots of objects and they are easily overshadowed by each other or by the user's interaction (like a finger crossing an object). The user may easily have problems selecting a certain object from the map when there are lots of these objects around at the same time.

Commonly the objects of interest on the 3D- or AR- view are visualized by explicit markers (or 'icons') around the objects of interest for example a rectangular shape labelled by the object's name. The user can select the object to get more detailed information, like an address and type of the object (like 'business building', 'restaurant', 'museum' and so on). However, recently the fashion of displaying objects has moved more towards to so-called markerless schemes where there are no explicit markers relating to the objects of interest, instead, the object can be recognized by any feature of the real environment, not by an explicit marker. The motive for this direction is that the previously common markers can be problematic as they may seem intrusive on the 3D- or AR- screen since they are actually not a part of the environment. One down-side of using markerless 3D-AR systems is however that tracking of objects becomes computationally more complex.

User interaction techniques are also an active research area. Conventional techniques are touch-based where the user manipulates the scene, AR or 3D, with finger input gestures. These techniques include commonly tapping, swiping the screen with the finger or dragging objects with a finger (like with mouse on common computer screens). Two-finger techniques have also been proposed, like selecting an object by ray-casting through a midpoint between two fingers. Also augmented hand techniques have been studied where the user moves a hand in front of the device's camera and the hand is visualized in the device screen and the user is able to select and manipulate objects by examining the visualized hand and objects of interest on the screen. The systems where the interaction is by hand gestures in front of the camera, opposed to the more common type of system where the user physically interacts with the 2D screen of the device, are usually defined as vision-based. Whether the interaction happens on the touch-input of the screen or in front of the camera, like the augmented hand techniques, they all however present problematic issues when the user is moving the mobile device and the user wants to access a certain object in a dense map while moving themselves in the real physical world. Solutions to these kinds of issues have also been proposed, like the freeze-technique where the user can 'freeze' the object on the screen and move the viewpoint elsewhere, with the aforementioned object still visible on the screen.

3 Designing the Application

The focus of this project, however, was mainly to study how the user should execute transitions between the AR and 3D views efficiently and easily with a touch-based system. The secondary focus was on how the user should access an object of interest easily, whether in 3D or AR view, and what type of information the user commonly wishes to extract from an object. The software used for implementing the interaction techniques was the 'CultAR' software provided by Konrad Markus from the Helsinki Institute of Information Technology which is described next.

3.1 The CultAR Software

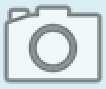
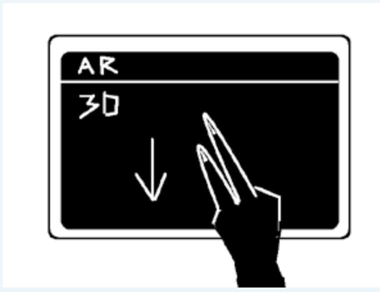
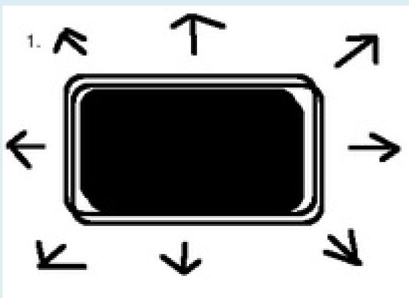
The CultAR software implements a 3D map navigation system with the possibility of transitioning to and from Augmented Reality views. It also allows interaction with objects of interest. The target platform for this software is Android mobile devices and the 3D map model and its navigation are implemented in Native code, while the user interface is implemented in Java. The interaction techniques for transitioning between the 3D and AR views have been added on top of the provided software in Java. This project is has been under development for many years and has recently been ported to the Android platform.

The researchers added to this project using the Eclipse ADT and the mobile device for testing was an ASUS tablet (product ID *tft700*). A screenshot of the original application is shown in Figure 1. The user can drag the ground in any direction by sliding their finger in area '1'. '3' controls the height of the camera, '4' controls the pitch and '2' rotates the view.

Figure 1: A screenshot of the CultAR application.

3.2 Gestures for View Transitioning & Object Interaction

The first step was to identify a set of possible gestures to implement the switching of views and object interaction. For the researchers, the first and simplest method of transitioning between 3D and AR views was pushing a button on the screen. This led to a brainstorming session about the ways in which a user could switch between views and also on how they might like to interact with objects. We considered interaction techniques based on finger gestures (single tap, double tap, long press, swipe), mobile device positioning gestures (shaking device, rolling device), and simply using buttons. For objects of interest a similar list of possible interactions was created. Table 1 summarizes the various interaction techniques that were suggested for both transitions and objects of interest.

Gesture	Description
Tap button 	Transition from 3D – AR and vice versa
Swipe screen ('fling' gesture in Android)	Transition from 3D – AR and vice versa 
Shake tablet	Transition from 3D – AR and vice versa 
Roll tablet	Transition from 3D – AR and vice versa

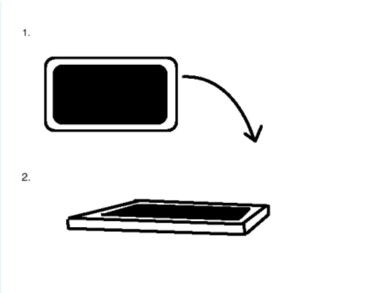
	
Single Tap	Transition from 3D – AR and vice versa
Double Tap	Transition from 3D – AR and vice versa
Tap button	Fly to birds-eye view of user's current location
Tap button	Fly to street-level view of user's current location
Double Tap on Object	Object of Interest Interaction: opens context menu
Long press Object	Object of Interest Interaction: opens context menu
Tap button 	Object of Interest Interaction: opens list menu on side of screen

Table 1: Features Thought Up in Brainstorming Session.

A few other options suggested included giving a spoken command to the application for transitioning, using audio to provide feedback whenever the view changes, and providing some haptic feedback to the user when they touch an object of interest. These options were not taken forward as they could not be added to the project in the limited time.

The list of possible gestures became quite extensive and the paper prototyping and usability evaluation stage which is described next was essential in narrowing down the kind of gestures users were more comfortable with.

3.3 Paper Prototype

A picture of the paper prototypes we constructed is shown in Figure 2. The paper prototypes were used to discover how users intuitively approached the application. A total of five users tested the paper prototype. At first the user was given no cues as to what type of gestures were expected from them, and they were encouraged to 'Think Aloud' about their actions. The users were only told which task they were supposed to accomplish, for example 'switch to AR view', and the researchers observed what gestures the users selected themselves to complete the task. After this 'free session' we then explicitly told the user which interaction to use and then recorded the user's feedback about the gesture and its usefulness. After the test the users were asked for their opinions. The researchers were able to film two user test sessions and they are available on Youtube[2]. The questionnaire we used is available in Appendix A.

Figure 2: The Paper Prototype

3.4 *Usability Testing & Results*

There were four scenarios that were tested: 1) Switch from 3D to AR view 2) Navigate to your street-level view 3) Zoom out to a birds-eye view and finally 4) Interact with an object of interest. The detailed questionnaire is provided in Appendix A.

---

[2] https://www.youtube.com/watch?v=4JczRPLVAVk&feature=youtu.be
https://www.youtube.com/watch?v=HXZ0fC5kkeo&list=UUd9wnLgGzC8NOIICXu7UAyA

It was found that almost all users used the same set of gestures to perform transitions, zooming in or out, and interacting with objects. Button press and swipe won out over all other gestures for transitioning between views. The shake and roll gestures were universally disliked because they were cumbersome, one user mentioned that they did not like that the screen was not visible all the time during the roll gesture. One user preferred to use swipe to transition between views. All users liked to use button presses to perform the transition or zooming between birds-eye and street level views. All the users said that double tap would be used for zoom and not transitioning, but it could be used for interacting with an object of interest. Long-press was not favoured by most users for interacting with an object. Most users liked the feature of a list of visible objects appearing on the Object-list button press. Curiously, single tap was not used or recommended by any user, when one user was queried about this they answered that single tap just does not occur to them for navigating on a map. One user tried to transition between views by pinching to zoom from the street-view – this was an unexpected and good suggestion.

4 Redesigning the Application

As a result of the paper prototype the researchers were able to identify the interaction methods most appreciated by the users and also recorded various other unanticipated suggestions that were provided by the users. These extra suggestions referred to positioning of buttons, the sort of size and graphic that should be on a button, and many suggestions for the general map navigation part of the CultAR application. Based on the user feedback the design of the system of interactions was re-evaluated and a few new features were added like double-tap to zoom and finally transition. Table 2 summarizes the final features selected for development. A description of the software prototype follows.

Gesture	Description
Tap button 	Transition from 3D – AR and vice versa
Swipe screen (‘fling’ gesture in	Transition from 3D – AR and vice versa

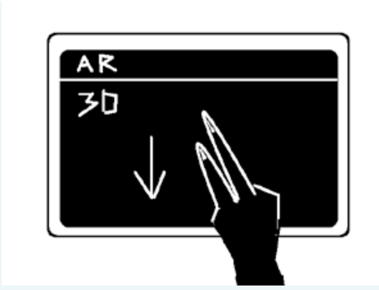
Android)	
Tap button 	Fly to birds-eye view of user's current location
Tap button 	Fly to street-level view of user's current location
Pinch to zoom in/out, then transition	*Could not be implemented because of lack of support in the original CultAR software*
Double Tap to zoom, then transition	For zooming in, when the user zooms-in close enough to the street-view, the 3D view transitions to AR view
Shake tablet	*The users did not like this feature in the paper prototype but the researchers decided to implement it to see if an actual implementation of the feature would get better feedback.*

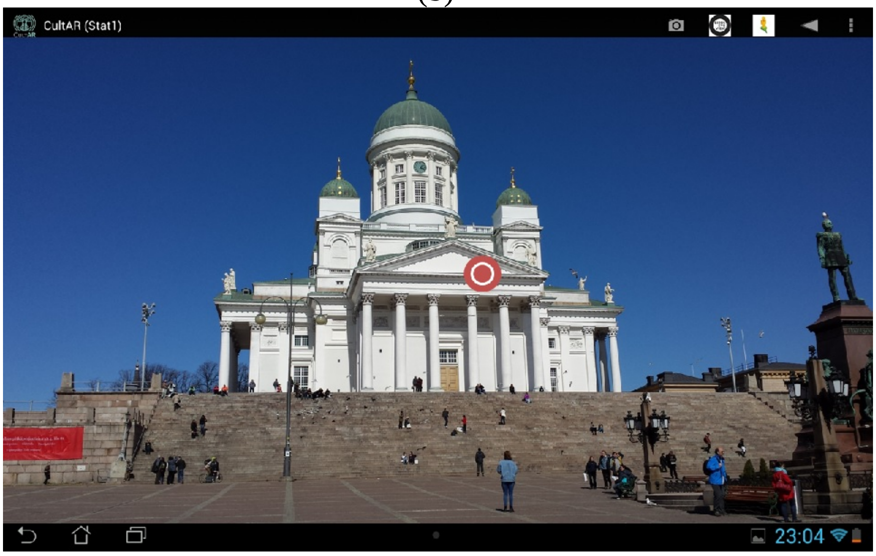
Table 2: Features Selected for Development.

4.1 Software Prototype

Figure 3 shows some screen shots of the developed software. The major changes that have been included are: buttons for performing zooms to birds-eye view and street-view, the addition of ***swipe*** and ***shake*** gestures to transition between AR and 3D views, and ***double tap*** gesture for zooming in and also transitioning. Because of difficulties encountered during development, the secondary objective of interacting with objects could not be fully implemented although the code interface for accessing the objects was added to the project, and tapping on an object does show its name on screen.

(a)

(b)



(c)

Figure 3: Views from the Software Prototype. a) A random street b) 3D map view of cathedral c) AR view of cathedral

To judge the user-friendliness of the software system, another usability test was conducted in which three users participated. The questionnaire is shown in Appendix B. The same scenarios were provided as for the paper prototype, and users were asked to test out each of the added features and tell us their opinion.

4.2 Usability Testing & Results

All users were able to complete the tasks without much prompting. This was different from the paper prototype testing where users needed a lot of explanation to understand how to use the prototype. The software prototype was much easier to use than the paper prototype. This may be an issue with doing paper prototyping with mobile devices and touch input as the paper prototype is not very supportive for the user, and it is difficult for the researcher to provide for all scenarios.

From testing we found that two users liked the ***swipe*** gesture best, while one user liked the **button** best. All users ranked **swipe** and the **button press** as their top two favourite interactions for switching between views. The **shake** gesture had received negative reviews during paper prototype testing, and it received negative reviews once again. The users unanimously disliked this gesture as it was "silly looking" and "took too much time and effort". The **double-tap-to-transition** gesture also caused some confusion as users expected only a zoom functionality with double tap. Overall users were satisfied with the gestures provided and discovered the buttons themselves. They had to be told about the **swipe** and **double-tap-to-transition** gesture as these are not usual gestures in map navigation.

The few changes that users suggested can easily be added to the application. The suggestions were: bigger buttons, converting the 3D-AR transition buttons to a toggle button, putting in some indicator to the view to let the user know they can swipe the screen to switch views and a 'back' button that allows you to go back to the previous state.

The process of paper prototyping and testing followed by developing of the software prototype has proved to be quite effective in figuring out what interactions are in fact the most simple and intuitive. Many good suggestions and ideas have been collected from our testers and there is a lot that can be added to this 3D-AR navigation application to develop a highly user-friendly interface.

5 Conclusions & Future Work

Although the users were presented with a wide variety of choices for transitioning and interacting with objects in the application, users invariably preferred gestures that they were already familiar with. Especially as the users were all familiar with map navigation software, they consistently tried to use swipe for dragging the map view, double tap for zooming in, and shake and roll were considered completely unusable, whereas the researchers were suggesting that these gestures be used for transitioning between views. The researchers concluded that users always favor gestures they already know rather than learn new gestures that might even be more efficient. An example of this behavior is that users preferred to use pinch to zoom-out of a scene (which is a multiple-press operation) rather than using a button (single-press) to zoom to birds-eye view.

The researchers have considered several interesting directions for further research: Firstly, detailed analysis of the time taken by different gestures in a touch-screen environment, similar to 'Key-Press Analysis' for keyboard and mouse input. Timestamps and logs could be used for judging how quickly users completed tasks, and the speed of gestures like swipe versus tap and so on. Secondly, the researchers believe that some of the more unfamiliar gestures like shake and roll might be appreciated by users if they used the application for a longer time. So a longer user study could be more useful in judging the true utility of some gestures. Finally, another possibility for improvement would be creating an interface to the Native library in CultAR that exposes all the information about the map model. For this project there were hardly any interfaces available to the Native library and Konrad Markus very kindly implemented changes for the researchers many times. Exposing the interface to the library would make gesture additions to the user interface much easier.

6 Challenges

Working with an existing project always has its issues and this project was no different. The design of CultAR was not conducive to additions by the researchers as the user interface interaction functionalities were distributed in both the Native and Java code. The researchers had planned and expected to work in the Java code but as it turned out almost all the critical functionality controlling navigation and interpreting gestures was being performed in the Native library. The Native code being inaccessible to the researchers meant that a lot of functionality could not be added or leveraged from the existing code and any small change had to be requested through our CultAR contact.

Another issue was the overlap between gestures already in use by the CultAR application and gestures that the researchers would have liked to use. The zoom functionalities via double tap and zoom were very difficult to implement as the CultAR application has no way of detecting where the user has tapped in world coordinates, and the researchers were only able to implement an approximation of zoom.

Acknowledgements

We would like to thank Diogo Cabral of the University of Helsinki and Konrad Markus from the Helsinki Institute of Information Technology (HIIT) for their considerable help and support during this project. Diogo Cabral provided great help in directing and advising this project with respect to user interactions, user evaluation techniques and overcoming any problems faced during design and development. Konrad Markus provided invaluable help in solving and providing solutions to our many queries, development and coding problems encountered during implementation.

References

Kje13 Kjellmo, I., Educational: 3D Design for Augmented Reality. Proc. on 12th International Conference, ICEC 2013, São Paulo, Brazil, October 16-18, 2013, pages 200-203.

TC13 Telkenaroglu, A. and Capin, T., Dual-finger 3D Interaction Techniques for mobile devices. Personal and Ubiquitous computing, volume 17, issue 7, 2013 october.

CAO13 Crowdhury, S., Arshad, H. and Parhizkar, B., Handheld augmented reality interaction technique. Proc. Third International Visual Informatics Conference, IVIC 2013, Selangor, Malaysia, November 13-15, 2013, pages 418-426.

UNK14 Unuma, Y., Niikura, T. and Komuro, T., See-through Mobile AR System for Natural 3D Interaction. Proceedings of the companion publication of the 19th international conference on Intelligent User Interfaces, Dan Carmel Hotel Haifa, Israel, February 24-27, 2014, pages 17-20.

LLC13 Lee, S., Lim, Y. and Chun, J., 3D interaction in Augmented Reality with stereo-vision technique. Advanced Communication Technology (ICACT), 2013 15th International Conference on, Pyeong, Chang, 27-30 Jan. 2013, pages 401-405.

DJo12 Dolz, J., Markerless Augmented Reality. http://www.arlab.com/blog/markerless-augmented-reality accessed 15.4.2014

LB11 Lee, G. and Billinghurst, M., A User Study on the Snap-to-feature Interaction method. Mixed and Augmented Reality (ISMAR), 2011 10th IEEE International Symposium on, Basel, 26-29 Oct. 2011, pages 245-246.

Kah13 Kahn, S., Reducing the gap between Augmented Reality and 3D modeling with real-time depth imaging. Virtual Reality, volume 17, issue 2, 2013(06), pages 111-123.

Appendix A – Questionnaire for Paper Prototype Usability Testing

Information about User

[Users: University Students]

1. Do you regularly use a smartphone or other touch input device? Please specify which one.

2. Have you used a map navigation tool on a touch input device before? If yes then which app? If no then have you used such an application (like Google Maps/ Google Earth) on a desktop/laptop?

3. Have you used an augmented reality tool on a touch input device before? If yes then which app?

Description:
We are prototyping part of a 3D + Augmented Reality map navigation system. Our prototype is related to switching between the 3D and AR views. We would like to find out what the most user- friendly way would be to switch between views. You will try out different interaction methods to switch between views and rate them on the basis of ease-of-use and efficiency

Instructions:
1. Please feel free to make any comments while you are using the prototype - your input is highly valuable!
2. Think about what you like and dislike while using the app
3. Think about what you wish the app would do
...We will ask you your opinion after the demonstration

Usage Scenario 1: Transitioning from Street-level 3D view to AR view
You are standing outside Kumpula campus and seeing a 3D map view of your surroundings. To get a better view of your surroundings you want to change view to ‘Augmented Reality’. You can see ‘Objects of Interest’ like the buildings chemicum, physicum, exactum. How would you switch to the AR view?

i. Press the 3D <->AR button
ii. Shake tablet
iii. Roll tablet
iv. Swipe down from top of screen
v. Tap / Double Tap
vi. Use a slider --- (not shown on prototype)

How would you switch back to 3D?
0. Repeat same gesture
i. Press the 3D <->AR button
ii. Shake tablet
iii. Roll tablet
iv. Swipe down from top of screen
v. Tap / Double Tap
vi. Use a slider --- (not shown on prototype)

Q: Does it make sense to have multiple ways of changing between views?
Which 3 interaction methods do you like most? Why? (Is it quick / intuitive / familiar / etc.)
Which interaction is most inconvenient for you and why? (Is it too laborious / clumsy / slow / unfamiliar / etc.)

Are any of the interactions strange or unexpected for you to use? Which ones?
Apart from these methods can you suggest a way to change between views?
Are there any other features/interactions you would expect in this application?
Any other comments?

Usage Scenario 2: Reset View to Street level
You are viewing a zoomed out 3D map of Helsinki. To get back to street level view you can:
i. Press the ‘Street View’ button
ii. Shake tablet
iii. Roll tablet

Usage Scenario 3: Zoom out to Bird’s eye view in 3D
You are viewing the 3D map at street level. To to zoom out to a bird’s eye view you can:

i. Press the ‘Birds Eye’ button
ii. Shake tablet
iii. Roll tablet

Which interaction is most convenient for you and why? (Is it quick / intuitive / familiar / etc.)
Which interaction is most inconvenient for you and why? (Is it too laborious / clumsy / slow / unfamiliar / etc.)

Any other comments?

Usage Scenario 4: Interact with an Object of Interest
You are looking at the 3D map with several Objects of Interest visible. How would you interact with the objects?

i. Select from list
ii. Tap on an object
iii. Double tap
iv. Long press
v. [Double tap reserved for transition to 3D view with that particular object zoomed in]

What is an Object of Interest in your opinion? (person / event / building / etc)
What kind of actions do you expect to be able to perform with an Object of Interest?
What kind of information do you expect to see about an Object of Interest?

Any other comments.

Appendix B – Questionnaire for Software Prototype Usability Testing

Information about User

1. Do you regularly use a smartphone or other touch input device? Please specify which one.

2. Have you used a map navigation tool on a touch input device before? If yes then which app? If no then have you used such an application (like Google Maps/ Google Earth) on a desktop/laptop?

3. Have you used an augmented reality tool on a touch input device before? If yes then which app?

Description:
We are prototyping part of a 3D + Augmented Reality map navigation system. Our prototype is related to switching between the 3D and AR views. We would like to find out what the most user- friendly way would be to switch between views. You will try out different interaction methods to switch between views and rate them on the basis of ease-of-use and efficiency

Instructions:
1. Please feel free to make any comments while you are using the prototype - your input is highly valuable!
2. Think about what you like and dislike while using the app
3. Think about what you wish the app would do

...We will ask you your opinion after the demonstration

Usage Scenario 1: Transitioning from Street-level 3D view to AR view
You are standing in Senate Square and viewing a 3D map view of your surroundings. To get a better view of your surroundings you want to change to the camera view. How would you switch to the camera view?

i. Camera button
ii. Fast swipe
iii. Double Tap

How would you switch back to 3D?
0. Back button on tablet
i. Camera/Map button
ii. Double Tap
iii. Fast swipe

Q: Which methods do you like most? Why? (Is it quick / intuitive / familiar / etc.)
Which interaction is most inconvenient for you and why? (Is it too laborious / clumsy / slow / unfamiliar / etc.)

Are any of the interactions strange or unexpected for you to use? Which ones?
Apart from these methods can you suggest a way to change between views?
Are there any other features/interactions you would expect in this application?
Any other comments?

Usage Scenario 2: Reset View to Street level
You are viewing a zoomed out 3D map of Helsinki. To get back to street level view you can:

i. Press the 'Street View' button
ii. Use map navigation dragging

Usage Scenario 3: Zoom out to Bird's eye view in 3D
You are viewing the 3D map at street level. To to zoom out to a bird's eye view you can:

i. Press the 'Birds Eye' button
ii. Navigate using map controls

Which interaction is most convenient for you and why? (Is it quick / intuitive / familiar / etc.)

Which interaction is most inconvenient for you and why? (Is it too laborious / clumsy / slow / unfamiliar / etc.)
Can you think of a better way to change between views?
Are there any other features/interactions you would expect in this application?
Any other comments?