Faculty of Science - University of Zurich
Department of Geography
Geographic Information Visualization & Analysis (GIVA)

Master thesis - Geo511

Efficiency of linking techniques and individual differences in navigation in 2D – 3D side-by-side views

April 2014

Author
Matthias Kuhn
06-736-920

Supervisor
Dr. Arzu Çöltekin

Faculty Member
Prof. Dr. Sara I. Fabrikant

Contents

1. Introduction **1**

2. Research Questions **3**

2.1. Research Question 1 3
2.1.1. Hypothesis 1.A 3
2.1.2. Hypothesis 1.B 3
2.1.3. Hypothesis 1.C 3
2.2. Research Question 2 3
2.2.1. Hypothesis 2.A 3
2.2.2. Hypothesis 2.B 4
2.3. Research Question 3 4
2.3.1. Hypothesis 3.A 4
2.3.2. Hypothesis 3.B 4

3. Theoretical Background **5**

3.1. Extending the Dimensionality 5
3.1.1. Depth Cues 7
3.2. Assessing the Problem 8
3.3. Navigation 11
3.4. Multiple Linked Views 14
3.4.1. Brushing 16
3.4.2. Proxying 16
3.5. Eye-Tracking 18
3.6. Spatial Ability 21

4. Methods **23**

4.1. Participants 23
4.2. Session 24
4.2.1. General Conditions 24
4.2.2. Procedure 24
4.2.3. Pre Questionnaire/ Eye-Tracker Calibration 24
4.2.4. Training 26

4.3. Experimental Design 26
4.3.1. Independent Variables 28
4.3.2. Dependent Variables 30
4.3.3. Counterbalancing 31
4.4. Materials 33
4.4.1. Online Questionnaire 33
4.4.2. 3D geobrowser 34
4.4.3. PyQGIS 37
4.4.4. Experiment Plugin 37
4.4.5. Data 38
4.5. Gaze Data 38
4.5.1. Geocoded Gaze Data 38
4.6. Statistics 40
4.6.1. Data Pre-Processing 40
4.6.2. Analysis of Variance 41
4.6.3. Regression Analysis 43
4.7. Assumptions for Navigation Analysis 43
4.7.1. Search stages 43
4.7.2. View Alignment 46

5. Results 47
5.1. Main Effects 47
5.1.1. Spatial Ability 47
5.1.2. Interaction designs 47
5.1.3. Subtasks 48
5.1.4. Scene Layouts 48
5.2. Accuracy 48
5.3. Learning Effect 49
5.4. Satisfaction 50
5.5. Interaction Designs 52
5.5.1. Map Alignment 57
5.6. Spatial Ability 58
5.6.1. Map Alignment 58

6. Discussion **61**
6.1. Learning Effect . . . 61
6.2. Accuracy . . . 63
6.3. Scene Layout . . . 63
6.4. Subtasks . . . 64
6.5. Navigation in 3D Environments and Multiple Linked Views . . . 65
6.5.1. Linking Techniques . . . 65
6.5.2. View Alignment . . . 67
6.5.3. Combining Linking Techniques . . . 71
6.6. Satisfaction . . . 72
6.7. Individual Differences in Spatial Ability . . . 73
6.7.1. Map Alignment . . . 74
6.8. Gaze Geo-Referencing . . . 74
6.8.1. Feature Referencing . . . 75
6.9. Limitations . . . 77

7. Conclusion **79**

A. Personal declaration **81**

Glossary **90**

Acronyms **91**

B. Experiment Material **99**

C. Additional Figures **103**

D. Statistical Tables **109**

E. Comments **113**

1. Introduction

Three-dimensional interactive representations of spatial data became increasingly available in recent years. With computers getting more affordable and powerful, a wide variety of applications in the gaming industry and in the geo-spatial domain got published. While 3D representations are often appreciated for their "nifty" appearance, they also offer challenges because of properties like occlusion (Dodge et al., 2008) which require special attention. To deal with this particular problem, rotation is a tool that can be – and often is – made available to users. Shepherd (2008) comments on 3D views that such representations often tend to impose severe interaction demands on users. To facilitate the interaction process, a user should be provided with effective overview tools to support rapid situation awareness, something that has already been highlighted as important before by Shneiderman (2003).

Overviews make use of a concept called multiple linked views (MLV), a presentation approach well-known in human-computer interaction (HCI) with a large amount of research attached to it. By offering a second, two-dimensional view of the same scene we can overcome some of the criticized points of 3D visualizations as well as increasing a users situation awareness with an overview and detail approach. Such overviews are also known from traditional 2D mapping in combination with linking techniques such as a *you are here* icon or as an *overview map* depicting the bounding box of a detail map, in such a way trying to support the user in the process of map matching. These tools cannot be transferred just like that to interactive 3D views and need a critical re-evaluation for their applicability.

The possibility to perform rotational operations on a 3D view can be accompanied by doing the same rotation on a 2D view (a technique known as track-up display) but keeping the orientation of the 2D map stable (north-up display) has its advantages as well, like supporting a user to build a cognitive map of the scene (Aretz, 1991). With a north-up 2D view a user has the possibility to align the 3D view to show the same orientation as the 2D view. Map alignment has been identified as a research topic in several contexts such as hardcopy maps (Liben, Myers & Christensen, 2010) and mental maps (M. Jeanne, 1987). Research is therefore obligated to get insights into how users deal with a combination of a north-up map and an interactive, freely rotatable 3D view.

The ability to perform mental rotation of the environment and objects is embedded in a wider domain differs between individuals. This determines to a large degree a persons spatial ability, a personal property which can be measured with various tests. A well-known test is the Vandenberg's mental rotation test (MRT) which is able to measure somebody's ability to rotate an object by imagination. It is known that the ability to perform object-based mental rotations is closely linked to the ability of performing ego-centric rotations of the environment (Hegarty, 2004). A possible connection between map alignment and spatial ability seems obvious as they both deal with rotation. Liben et al. (2010) discussed this topic in real-world environments.

In this study we aim to contribute to the growing body of knowledge of navigation in 3D views and MLV with a particular focus on map alignment and individual differences in spatial ability. In a user study, we asked participants to perform several map matching, navigation and rotation tasks on 2D – 3D side-by-side views and analyzed the outcome. Different linking techniques, particularly a simple form of brushing identification highlight (IH) and a novel tool we refer to as frustum highlight (FH) are assessed in terms of supporting the users efficiency in map matching, navigation and rotation. The FH draws a highlight on the 2D view of the frustum/terrain intersection determined from the 3D view. A noticeable efficiency improvement could be found with IH and a major problem in map alignment introduced by FH has been identified and discussed.

Furthermore a method to perform feature-based eye-tracking on MLVs has been implemented, assessed and discussed. While this method was not used for data-analysis in this thesis, the ideas behind it and the possible benefits offered for future research vindicate the critical review of the method to perform feature-based eye-tracking.

2. Research Questions

2.1. Research Question 1

How well do people perform with linked 2D – 3D views side-by-side while conducting visuo-spatial tasks? Specifically, what are the effects of alternative linking techniques for map matching, navigation and rotation tasks?

2.1.1. Hypothesis 1.A

Highlighting the frustum/terrain intersection on the 2D view: frustum highlight (FH) improves user performance.

2.1.2. Hypothesis 1.B

Single feature brushing: identification highlight (IH) improves user performance.

2.1.3. Hypothesis 1.C

The two linking techniques FH and IH will additionally increase user performance when offered in combination compared to their single occurrence.

2.2. Research Question 2

The problem of *getting lost* in virtual 3D environments is connected to the possibility to rotate the view. We are interested if in the context of a north-up and therefore not necessarily aligned overview map a users' efficiency is influenced by linking techniques. Does an unaligned map decrease a users' performance? Is map alignment related to linking techniques available, especially as FH gives an indication about view direction and therefore makes it easier to re-align the view?

2.2.1. Hypothesis 2.A

When users work with unaligned views, user performance is compromised.

2.2.2. Hypothesis 2.B

Frustum highlight (FH) leads to a decrease in time people work with unaligned views.

2.3. Research Question 3

Does spatial ability show a positive effect on the performance using interactive multiple linked 2D and 3D views?

2.3.1. Hypothesis 3.A

People with higher spatial ability show an increased efficiency in navigating and map matching on linked 2D – 3D side-by-side views.

2.3.2. Hypothesis 3.B

People with higher spatial ability spend less time with unaligned views than people with lower spatial ability.

3. Theoretical Background

3.1. Extending the Dimensionality

Everyday people are confronted with spatial information and more and more visualizations are produced and presented, on paper, screens or other media, static, dynamic or interactive. This spatial information may be visualized as realistically as possible or contain symbolized information which often would not be visible otherwise. Maps as spatial representations are mostly presented on two-dimensional media as opposed to the three-dimensional real world we live in (Chen, 2006). There are ways to encode height information inside a two-dimensional map by means of contour lines or color coding and similar techniques. Such representations of height require the user to mentally translate this information into a three-dimensional image of space (Rase, 2003; Taylor, 2005). It is also possible to generate the impression of three-dimensional space by offering viewers *depth cues* and therefore helping them to understand the extension of things in three dimensions by making use of their everyday experience and the training they naturally receive by looking at the world surrounding them day by day.
Rase (2003) claims that especially less experienced users less in map-reading can benefit from three-dimensional maps due to the fact that the encoding of height information has not to be translated first but can intuitively be realized thanks to the experience people have of this kind of visual input in their everyday life.

Gore (1998) gave a speech in the California Science Center where he described a vision of the *Digital Earth* as a multi-dimensional representation of the planet to access the ever-growing amount of available information in a geo-referenced manner. This should be connected to archives of digital knowledge to get an overview and allow better description and understanding of human activities and the system earth. His description of a fully immersive system allows to travel through space and time and explicitly highlights that this system has to present information in a way that the human brain can handle it and at the also emphasizes that such a project will "require the grassroots efforts of hundreds of thousands of individuals, companies, university researchers, and government organizations" [1].

[1] No page number (speech)

With our study we aim to understand better how the human visual system (HVS) deals with three-dimensional representations and therefore to contribute also to Gore's Digital Earth vision.

A year later MacEachren et al. (1999) list three key-challenges to be assessed for a successful GeoVirtual Environment:

- determining the appropriate balance of realism and abstraction for different geospatial application domains, different users, and different tasks
- developing new, innovative, methods for interaction with the spatial, temporal, and attribute components of geospatial information, separately and together
- developing approaches that take advantage of VE's potential to facilitate both same time – same place and same time – different place collaboration in research, learning, and decision-making that involves geospatial data.

Since the start of the new millennium, work on interactive three-dimensional geobrowsers has experienced acceleration and the progress in capability of hardware and its low price have ultimately led to an immense advance in technology and availability of frameworks capable of presenting information in a three-dimensional form on the one hand and data prepared for such a representation on the other hand (OpenWebGlobe [2]; GoogleEarth [3]; World Wind [4]; you may also read the Editorial by Çöltekin and Clarke (2011) for more thoughts about virtual globes).

It seems that people like to use three-dimensional representations. A certain "Wow"-effect is reported, it is engaging and aesthetic, Kray, Elting, Laakso and Coors (2003) state among others that even when 3D did not always improve performance it was fun to use. However, there is a prevalent feeling and a recently started debate that advances in this area have more often been driven by technological motivation rather than user demand and often peculiarities of particular implementations are tested rather than trying to understand the underlying problems (Shepherd, 2008).

[2] http://www.openwebglobe.org
[3] http://earth.google.com
[4] http://worldwind.arc.nasa.gov

Shneiderman (2003) cited in Chen (2006) lists three important key-challenges for a "good" 3D geo-browser:

- rapid situation awareness through effective overviews
- prompt, meaningful feedback for user actions.
- reduced number of actions to accomplish tasks

These theoretical recommendations are rarely empirically tested (if ever). In this project, we consider all three of the requirements listed above in an empirical user study, and particularly focus on the first item as we experiment with whether people can get an effective overview of the situation (thus eventually facilitate rapid situation awareness).

Kraak (1993) presents a very systematic analysis of three-dimensional map design. While outlining a lot of potential problems with 3D he concludes that "when well designed and properly used [a 3D map] can be a powerful cartographic representation to communicate spatial data".

3.1.1. Depth Cues

In three-dimensional views the effect of depth has to be generated by the means of *depth cues*. To understand the way a picture gets transformed into a three-dimensional image by the HVS we want to give a short non-exhaustive introduction into the different types of cues.

We can categorize these cues according to if they can be visualized by a static picture *(monocular static)*, a moving picture *(monocular dynamic)* or if they require a different picture for the two eyes *(binocular)* (Ware, 2004). With the recent advent of screens capable of displaying binocular cues these are becoming more popular but still the majority of representations make use of monocular cues as they can be represented on a broader range of media and so does the system we used for our study.

Another approach for classification of depth cues is to separate them into *physiological* and *psychological* cues. Physiological cues can be monocular or binocular, on the other hand, psychological cues are obtained from the retinal image and are all monocular (Okoshi, 1976). Kirschenbauer (2005) further describes a concept called *True 3D* "[...] when perceiving the visualized third dimension, either it lays behind and/or in front of the display plane [...]" [5], a property she defines based on a combination of physiological and psychological cues.

3.2. Assessing the Problem

The usefulness of 3D displays is a controversial issue (Chen, 2006; Bleisch and Nebiker, 2008; Shneiderman, 2003; Shepherd, 2008), while some studies report better performance with 3D others do not find any difference or report worse performance (Wickens, Liang, Prevett & Olmos, 1996).

Bertin (1983) published a well-known classification introducing the topic of visual variables and the possibilities to use these for communication including, but not limited to, cartography. Kraak (1993) argues that these variables only have a meaning in two-dimensional graphics and have to be reassessed for their use in three-dimensional maps. As has been discussed before (See section 3.1.1), three-dimensional maps often use psychological depth cues to create the sensation of depth. By their nature, these cues use visual variables to achieve their goal and therefore may create confusion between their meaning as depth cue and their value in communicating other content. Figure 1 shows the relation of the visual variables according to Bertin to their usage as psychological depth cues as developed by Kraak (1993). The last item in the list addresses the occlusion of objects which is an often reported problem of three-dimensional displays (see e.g. Dodge et al., 2008).

Shepherd, 2008 lists a number of different possibilities to deal with the problem of occlusion.

Object culling refers to the process of removing objects with the main drawback of removing contextual information.

Object minimization helps to reduce visual clutter by re-sizing unimportant objects.

[5] Page 363

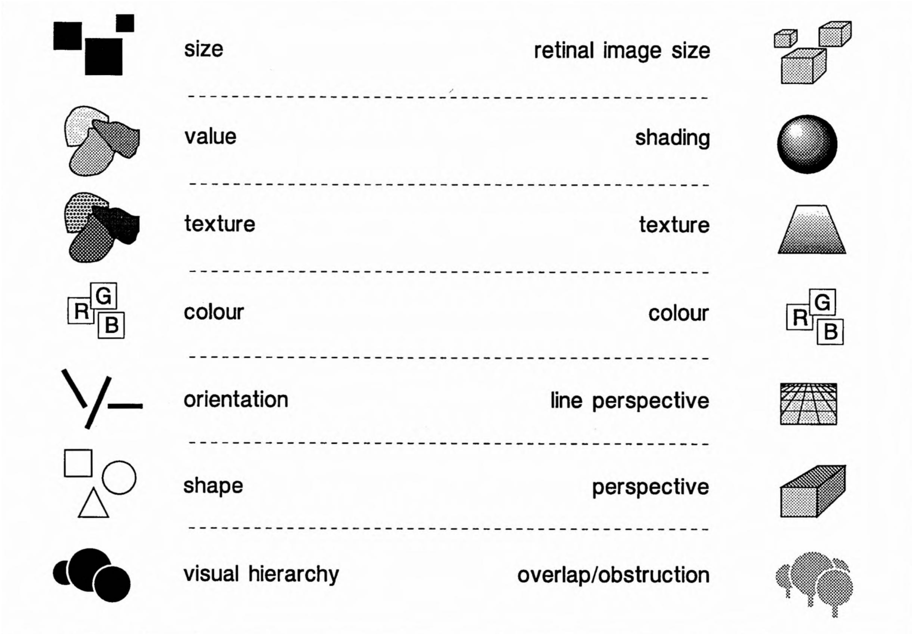


Figure 1: The possible relation between the visual variables and psychological depth cues. (Kraak, 1993)

Object displacement is the movement of objects to reduce clutter but they do so at the expense of loosing a degree of spatial relation.

View distortion Modifies the geometry of objects and increases the size of objects close to the user's view-point. There exist several different possibilities to implement this approach. Rosen and Popescu (2011) assessed a multiperspective image (MPI) framework called *graph camera* which offers an approach to split, bend and merge multiple conventional planar pinhole cameras into a single image which is locally distortion free and avoids occluders. The problems introduced by all techniques are basically the same as for *object displacement*.

Rotation or viewer moment can only be used in interactive systems and improves at the same time the effect of depth. The main drawback of this solution is that it may lead to unpleasant user side-effects and requires adaptedness in navigating within 3D scenes.

Symbol transparency can be used to see through objects to reveal what's behind.

There are the possibilities to have all objects transparent or only a subset of objects.

Symbol shadows help to improve the perceive the spatial distribution in all three axes of space. For example in a cube containing a 3D scatter-plot it is often hard to tell the exact location of a point. Shadows projected onto a plane help to improve the spatial distribution.

Multiple linked views ofter another viewpoint to appreciate the whole scene at once. This method is discussed in section 3.4.

Our study intends to fill a gap in the area of *rotation or viewer moment* and *multiple linked views*. The topic of *rotation or viewer moment* as used in the categorization above actually refers to two different actions, the first one is to rotate an object on the view, the second one is to move the viewpoint. Of these two we only look at the second which is also applicable to objects with a defined spatial orientation which it leaves unchanged in their spatial context.

From our everyday experience living in a three-dimensional world we are used that objects which are further away appear smaller than objects which are close. This effect is used as a psychological depth cue by perspective projections. Shepherd (2008) concludes that "the pseudo-3D versions of bar and pie graphs which appear increasingly in research publications introduce apparently unnoticed distortions into the messages conveyed to readers" [6]. When designing a 3D visualization we therefore have to keep this effect in mind – while it helps to improve the effect of depth by manipulating the objects' size this is done at the expense of comparability. We might wonder if the human brain compensates for this effect, but results from Wanger, Ferwerda and Greenberg (1992) suggest the contrary. In a size scaling experiment participants' accuracy decreased when presented perspective projections compared to orthographic projection.

Norman, Todd, Perotti and Tittle, 1996 have found evidence that the perceived length depends not only on the physical length of a line but also on the distance to the viewer and the orientation in space. With increasing distance to the viewer distances in depth are perceived compressed whereas they increase slightly or stay

[6] Page 202

constant in the frontoparallel plane. Therefore they conclude that the relationship between physical and perceived space cannot be described as euclidean.

In our study we use a perspective projection because it is a very effective and often-used cue (see Okoshi, 1976). The potential issue of bad comparability was dealt with by choosing height differences of objects in a way that the difference was big enough on the one hand and their spatial distance close enough on the other hand that the comparison itself would not lead to misjudgment.

3.3. Navigation

When we talk about the process of navigation we think of coordinated and goal-directed movement through the environment (Montello, 2005). This task is something we do as humans (along with a very wide range of animals) on a daily basis and while some struggle more and sometimes it is harder than other times everybody is generally capable of mastering the required skills as the fact that we find the way to work, the cupboard or simply send our hand towards a spoon proves.

Most of us stay on the ground while navigating, we walk, bike or drive a car. Some may learn to handle a helicopter or an aeroplane which offer different possibilities but also different challenges to navigate. When navigating in a virtual environment the concepts are often the same. Google Street View [7] lets you navigate on the ground while a geo-browser typically lets you fly around, Fuhrmann and MacEachren (2001) found the metaphor of a unidentified flying object (UFO) or flying saucer the best description for this navigation concept. Shepherd and Bleasdale-Shepherd (2009) compared geographic information systems (GIS) to video games and describe this navigation concept with “By and large, the analyst stays outside the scene, keeping his or her distance, and essentially playing god” [8].

While we navigate, we must maintain a sense of orientation and location relative to our goal, places and obstacles while we move (Montello, 2005). Sometimes we may get lost and in this case we require reorientation and reestablishing travel towards a destination (Loomis, Klatzky, Golledge and Philbeck, 1999). To do so we may use landmarks or maps, which can be visible or mental or we can

[7] https://www.google.com/maps/views/home
[8] Page 22

resort to a technique called *path integration.* Loomis et al. (1999) describes path integration as follows: Imagine navigating from a known starting point in a city with very narrow streets and tall buildings on both sides, effectively preventing any kind of self-location. The only method left to keep track of one's location is path integration.

The problem of *getting lost* was also investigated by Carlson, Hölscher, Shipley and Dalton (2010) in the context of architecture and indoor routing. They particularly stressed the importance of the cognitive map a (building) user has, its completeness and its correspondence to the buildings spatial structure as well as the compatibility of the between the building and the strategies and individual abilities of a user.

A nice — although not very scientific — illustration concerning the *getting lost* problem in a geo-browser is visible on figure 2. While it effectively offers a solution to this problem (the scroll-wheel), it also suggests that the user in question does not want to use this solution. In the strip's context it is considered to be cheating, in other situations it might be because a user is lazy and does not want to interact more than required, it might be because it interrupts the workflow or because with this action intended to *gain* context one could actually end up *losing* the context because the user loses track of the location where the zoom action has been initiated and finding the same location again proves more difficult than expected.

Taylor (2005) also looks at the alignment of maps with the surrounding. She starts with a review of literature on static maps where Levinew, Marchon and Hanley (1984) found that *you are here* maps contraligned with the surrounding often result in people heading in the wrong direction. She goes on to electronic map displays where there exists the possibility to rotate the map to align it with the environment. There is a trade-off shown in previous studies, indicating that north-up orientation improve the building of a cognitive map, while track-up maps, which are aligned to the user's orientation in space increase immediate performance (Aretz, 1991).

Wilkening and Fabrikant (2013) investigated the effect of time pressure on task solving in a 3D geobrowser. In their research participants were confronted with four cartometric tasks [9] in a time pressure (limit of two minutes for solving four

[9] Identification of elevation at two given points, Selection of the highest point along a given path,

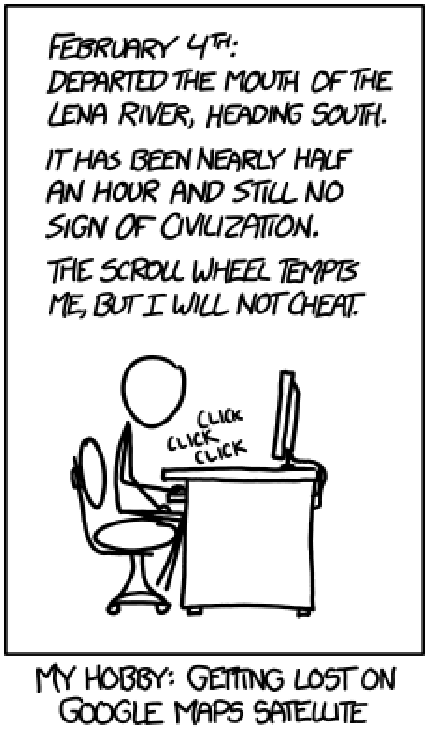


Figure 2: How zoom-levels can lead to lost sense of location. (Munroe, 2013)

tasks) and a no time pressure situation. They found that the interaction tools most used in a 3D geobrowser are the same tools which are available in 2D map viewers as well (panning and zooming) while tools which are typical for a 3D geobrowser (tilting and rotating) are not used as much. Apart from tilting, they found that all map interaction tools were used significantly less when under time pressure.

3.4. Multiple Linked Views

One of the most common operations in cartography is the task of generalization. This process is used to separate the important information concerning a certain topic from useless information or, in other words, to separate the signal from the noise. While generally cartographers claim that the task of generalization requires human judgment usually a certain degree of classification can also be achieved with filtering by logical criteria, what is necessary as we cannot afford to have a cartographer doing the generalization work for every GIS professional which is exploring a new data-set. Therefore the cartographic quality requirements are relaxed compared to traditional, high-quality paper maps in favor of (near) real-time behavior (Weibel and Burghardt, 2008).

Regardless of *how* the data gets generalized, this process always involves a certain amount of data reduction. To get more information than can be displayed in one single view the data is often distributed into several assumingly easy-to-understand and simple views, each focused on particular aspects of the underlying data. Such views require not to be regarded in isolation but to be linked with each other so users can make connections between them and integrate the insights they get from different views (Buja, McDonald, Michalak and Stuetzle, 1991). While Buja et al. (1991) also define temporal separation of views as a form of *linked views*, like for example when rotating a three-dimensional point-cloud, we use this term only referring to spatial separation without any implications for the temporal dimension. When highlighting objects with a color or another highlighting method, a user expects this to happen on the datasource level and therefore that the same action happens to the same object on other views as well according to McDonald,

Selection of the steepest slope based on three given locations and Qualitative description of the terrain between two given points

Stuetzle and Buja (1990).

MLV techniques has often been used for linking different plots with each other including or not views containing spatial representations of the data (E.g. Voigt, 2002; Brunsdon, 2001; Monmonier, 1989; N. Andrienko and G. Andrienko, 2003). At the same time MLVs are often used for spatial representation without involving non-spatial representations. A very typical example is the overview map which tells map users, where the spatial extent they are currently looking at is situated in a larger areal context. This can be seen in static as well as in dynamic maps. Ware (2004) states that the great advantage of the multiple window technique over others is that it does not distort and is able to show focus and context simultaneously.

Another reason for using multiple linked views is that different representation of the same data may overcome shortcomings of a certain representation. As we have seen in section 3.2, three-dimensional representations come with a number of caveats which two-dimensional views do not suffer from but on the other hand offer certain properties which two-dimensional representations lack. We can therefore assume that if we provide users with both kinds of views side-by-side at the same time, they will be able to take the best of both worlds. Bleisch and Nebiker (2008) implemented and tested side-by-side 2D and 3D MLVs employing information visualization or geovisualization techniques such as maps or bar charts. To connect the two views they mainly use brushing, but mention that "other techniques of interaction between the visualizations or combined navigation might prove useful and effective" [10].

MLV may also help to overcome the problems of focus and context, outlined in section 3.3. A linked overview may provide additional context information which helps to better understand an associated detail view. North and Shneiderman (2000) found a 30% to 80% increase in user performance (depending on the task) by using linked overview and detail views.

Ware and Lewis (1995) have looked into the subject of loosing context when zooming into images before. They describe a system called *DragMag* where it is possible to open new windows showing zoomed areas of a source picture connecting the edges with lines to the source (in a later paper Plumlee and Ware (2003) called this tethering). This allows a user to zoom and keep the context of the zoomed

[10] Page 5

area with respect to an overview. Cockburn, Karlson and Bederson (2009) have reviewed several different approaches of overview and detail approaches in single views, multiple views and temporal separation and discussed the tradeoffs that they have compared to each other.

3.4.1. Brushing

When a subset of objects is selected with an input device in a visual manner and highlighted according to this input we refer to brushing. This is most interesting in the context of linked views as it allows to make inferences of multiple views not apparent in a single one. (Voigt, 2002; Roberts and Wright, 2006)

Brushing does not have to be limited to the main visualizations but can also be extended to other parts of a user interface (UI). Roberts and Wright, 2006 discuss the concept of ubiquitous brushing which allows brushing to be applied to various elements of meta-information such as legends, menus or axis to improve dynamic filtering.

In this work we assess a very simple form of single feature brushing which we refer to as IH. We introduced brushing for two reasons. First, brushing is an often-used technique which deserves assessment. Second, we want to have a well-established reference for the second linking technique outlined below.

3.4.2. Proxying

As has been discussed in the section 3.3 getting lost is a known problem in interactive views. For example, when a user zooms into a subpart of the display the context of how the zoomed area fits in with the whole is lost. (Roberts and Wright, 2006)

The very same problem was found by Fuhrmann and MacEachren (2001) [11] where they assessed a head-up display (HUD) for navigation in a 3D geo-browser which, while it differs in the navigation controls, is in the freedom of movement comparable to our study. In a focus group following a test of the implementation users gave the feedback that "If you are very close to the surface, it would be still nice to know where you are on the landscape. [...] Why not having a little map in

[11] when continuing their work originally initiated in Fuhrmann and MacEachren (1999)

the corner with a position icon?" which was generally accepted to be a good idea, as one participant stated, "I was moving over the topo-map and it made me lose a sense where I was. It did not provide me [with] enough information to know where to navigate". They argued that orientation and wayfinding in desktop geovirtual environments requires some kind of orientation facilitation with real-time user positioning and suggests a moving *you are here*-symbol in an overview map. The subsequent implementation of such a symbol was greatly approved by a second focus group.

Plumlee and Ware (2003) tested a number of linking techniques between multiple views. The three different techniques [12] they assessed are:

View proxy Explicit representation of one view (or point of interest) within another.

Tethers Explicit lines connecting one view to another. This mainly makes sense when there are more than two views available and the connection between those needs to be explicitly visualized.

Orientation coupling Implicit aid that keeps two views oriented in similar directions.

They performed a within-subject study concerning a multi-perspective identification task where a camera followed a path in a virtual environment with several distractor objects and a target object which had to be noticed on one view and then selected on another view.

Both studies by Plumlee and Ware (2003) and Fuhrmann and MacEachren (2001) provide an overview with a *you are here*-icon which is a point in overview's coordinates (typically x and y). Plumlee and Ware (2003) additionally project the camera frustum to the terrain plane in the overview's coordinate system, what results in an additional information about the heading and aperture of the camera. This works fine as long as the camera is situated close to the ground. However, when the camera can also be freely moved in the direction orthogonal to the overview (usualy z) and tilted, it is not easy for a user to infer from the

[12] Plumlee and Ware (2003) used the term *linking aids* instead of *linking techniques*

given information about the camera on the overview to the objects visible on the detail view. Imagine you are sitting in an airplane and you are looking out of the front window, while having a map with an icon of your plane's location and heading. If you want to know which lake you are seeing out of the window, it is still a complicated job, as you will start searching on the map starting from the *you are here*-icon – if given a heading in the indicated direction – until you find the first point which is visible through the window and on the map, because the close surroundings of the *you are here*-icon will be covered from your view by the airplane's floor. The same counts for a camera, in fact, if it is tilted more than its aperture the camera's position marked by an icon on the overview map will no longer be visible on the detail view.

In response to this problem we propose to highlight the intersection of the camera's frustum with the terrain on the overview map. It is very common in two-dimensional maps to offer an overview map highlighting the extent of the detail view but this technique is not being applied for three-dimensional maps. As long as the camera is oriented strictly top-down, this works as known from two-dimensional maps. When the camera is tilted and there is a perspective projection in place distances the edge of the picture closer to the camera will depict less distance than the edge further from the camera. Thus, the resulting projection to a flat terrain results in an isosceles trapezoid. The shape of this trapezoid can even help a user to determine the current heading due to the fact that the shortest side on the overview is always the closest to the camera and the longest side points away from the user. Figure 3 visually offers a visual explanation of the concept.

In this work we test FH as described above for overcoming the problem of *getting lost* in a three-dimensional detail view.

3.5. Eye-Tracking

> Simply put, we move our eyes to bring a particular portion of the visible field of view into high resolution so that we may see in fine detail whatever is at the central direction of gaze. Most often we also divert our attention to that point so that we can focus our concentration (if only for a very brief moment) on the object or region of interest. Thus,

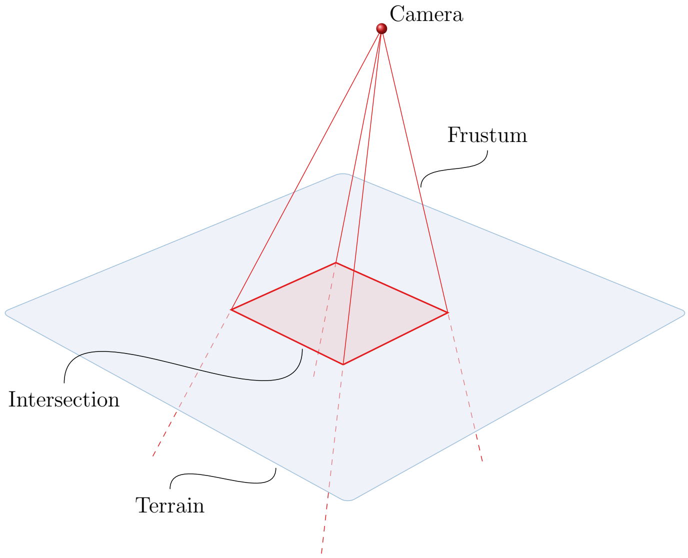


Figure 3: Intersection of the camera frustum with the terrain.

> we may presume that if we can track someone's eye movements, we can follow along the path of attention deployed by the observer. This may give us some insight into what the observer found interesting that is, what drew their attention, and perhaps even provide a clue as to how that person perceived whatever scene she or he was viewing. (Duchowski, 2007) [13]

A number of studies have used eye-tracking as input device for HCI by either explicit or implicit usage of the gaze. Salvucci and Anderson (2000) evaluated the usage of an eye-tracker in addition to other standard input devices. The eye-tracker could be used analogue to a mouse. Giannopoulos, Kiefer and Raubal (2012) discussed a system called GeoGazemarks where spots which previously received the user's attention and therefore are considered part of their mental map are highlighted to support the user's orientation while navigating on a map.

Eye-tracking can also be used in research to evaluate systems and concepts and performing usability studies. Much like the initial quote by Duchowski (2007) points out, this can give a researcher insights into a user's interest. Sometimes the reason for keeping one's gaze at a particular spot is not interest, but instead the

[13] Page 3

very reverse that a user is challenged to make sense of a particular element of a UI or whatever object may be situated in the center of the gaze (Poole & Ball, 2006). Tory, Atkins, Kirkpatrick, Nicolaou and Yang (2005, 2006) have researched combinations of 2D and 3D representations with displays called *ExoVis* and *Orientation Icon*. In the Tory et al., 2005 study they made use of an eye-tracker. They divided the screen into areas of interest (AOIs) bounding their 2D and 3D representations and measured the time people spent looking at each of them and counted the number of gaze switches between these. They also evaluated strategies by studying a single subject with a good performance [14] and described the approach taken by this subject. They also investigated error trials in two ways:

- Error size (small errors where assumed to be caused by different failures than big errors).
- Visual examination of eye-gaze patterns.

When working with geo-spatial visualizations the coordinate system of the data coming from an eye-tracker (often in screen coordinates) does not match the coordinate system of the geographical data. Still a researcher is often interested in the geographical coordinates rather than the screen coordinates. Geographical data also in many cases depicts features situated at certain positions, and in such cases the researcher may be interested in knowing, which feature the user has been looking at and not only the geo-coordinates. Kiefer and Giannopoulos, 2012 present an approach to match a users caze to line features (proposed to analyze bike routes) via post-processing incorporating a hidden markov model.

Most studies involving eye-tracking are done on static displays and maps. Digital maps are often interactive introducing the additional challenge that the projection from screen- to geo-coordinates change in the course of an experiment. Data may also be visible based on zoom-level or dynamically loaded and therefore be loaded and shown incrementally.

We will present and discuss a method we implemented to perform real-time gaze geo-referencing on multiple interactive views. While originally planned to identify landmark features used for a user's orientation and navigation in the context of

[14] After an initial learning curve

MLVs this plan had to be abandoned due to the data not being precise enough to be used. We therefore included the methodological basics in this thesis and review the merits as well as the challenges to be addressed in order to improve this system.

3.6. Spatial Ability

Spatial ability is the ability of an individual to manipulate or transform the image of spatial patterns into other arrangements. It is an ability which every individual possesses and without which we would not be able to manage even some of our simplest everyday tasks. However, this ability differs strongly between individuals, an effect which has been investigated repeatedly and which can be measured with various tests (Vandenberg and Kuse, 1978; Ekstrom and Harman, 1976; Hegarty and Waller, 2005).

Mental rotation ability is one of the few domains, where there is strong evidence for a sex-difference which was repeatedly reported with males achieving higher scores than females [15]. Women seem to be compensating this in real-world situation in routing and navigation tasks with other strategies (Malinowski, 2001).

Other work concerning spatial ability covers the connection to neuronal processes. Wolbers and Hegarty (2010) assessed spatial and navigational abilities and isolated cerebral areas which are particularly involved in different stages of the navigation process. They also concluded that people refer to different strategies to maintain orientation and to infer spatial relationships. While some people prefer featural cues to maintain orientation, others focus on geometric properties such as the layout of an environment.

Malinowski (2001) found a weak but significant correlation of MRT performance and real-world situations. This interest in in porting findings from paper-and-pencil tests to large-scale environments has experienced increasing interest. Hegarty (2004) used the term environmental spatial ability[16] to refer to this kind of spatial ability and performed several tests to differentiate between *egocentric*

[15] Lippa, Collaer and Peters (2010) offer an extensive overview of previous research concerning the relation between spatial ability and gender, social role and stereotypes. They also report a large international study taking 90'000 women and 111'000 men into account.

[16] Also in Hegarty, Montello, Richardson, Ishikawa and Lovelace (2006)

spatial transformations, meaning to change one's egocentric frame of reference with respect to the environment and the ability to make *object-based* transformations, meaning to transform an object's orientation while maintaining the own spatial orientation in the environment. They found a dissociation between the two spatial abilities, while they were still correlating to a large degree and was therefore probably not found earlier.

Research by Liben et al. (2010) suggests that aligning a map to its physical surroundings improves navigation precision. However, it has to be taken into account that the authors explicitly point out their uncertainty about cause and effect, meaning that it could as well be that individual differences induce both, the action of map rotation and navigation and orientation ability.

M. Jeanne (1987) have much earlier assessed the orientation of mental maps and the effect of alignment. In particular, when learning from a map, a person creates a mental image with a preferred orientation. Just like most people will think of Scandinavia as being "Up" on a map of Europe because most maps are drawn with a north-up orientation. In her experiment people were asked to point into the direction of an object which was not visible from their location. When facing the direction of the learned map, people were considerably faster in pointing than when facing the opposite direction. This indicates that a mental map does have a preferred orientation and that people need to mentally rotate it to the environment.

In this work we will use the Vandenberg's mental rotation test (MRT) introduced by Vandenberg and Kuse (1978) and refer to it as spatial ability when not stated otherwise.

4. Methods

To research the questions we developed a an experiment which requires users to map match objects in the context of 2D – 3D side-by-side views. We have then asked people to take part in the experiment in a one hour session including a task to examine their spatial ability and questionnaires assessing their demographics, skills and opinions about the stimuli. On the following pages the design of the experiment as well as the questionnaires and the spatial ability test are explained in detail. The experimental design will be outlined and the intentions it is based upon will be introduced. There will also be a part that describes the technical process that was required to design the experiment as well as the data processing and statistical methods used to analyze the data.
Finally, an experimental implementation is presented that is able to perform geo-referencing and feature-detection on multiple views based on real-time eye-tracking gaze-data although the data collected with this method was not analyzed for this work. The reasons for this are discussed in more detail in section 6.8.

4.1. Participants

A total of 30 participants took part in the experiment. They were personally asked to attend and not selected by any specific criterion. Most of them are geography students or staff of the department of geography of the university of Zurich. 11 participants were female and 19 male, all between 22 and 32 years old. Participation was not connected to any promised benefit.

	count	mean	std	min	25%	50%	75%	max
Age	30	26.700000	2.053592	22	25	26.5	28	32
Carto/GIS Experience	30	2.200000	1.214851	0	1	3.0	3	4
3D Games Expereince	30	1.033333	1.066200	0	0	1.0	2	3
3D GIS Experience	30	2.266667	0.784915	1	2	2.0	3	4

Table 1: Descriptive statistics of the participants who took part in the experiment.

4.2. Session

4.2.1. General Conditions

All the sessions took place in a laboratory at the University of Zurich on the Irchel campus in a windowless office specifically designed to conduct controlled experiments. For every session the room lighting was kept at the same level. Apart from a short chat when the participants arrived or direct requests for translations the language was consistently kept in English. All information and training was given on the computer screen to reduce any possible bias introduced by the experimenter's behavior.

4.2.2. Procedure

The whole session followed an experiment protocol which was developed, piloted and adjusted to correct some minor flaws. The experiment protocol is attached in appendix B.

To investigate the research questions we designed an experiment which involved two linked interactive views. One of them shows a 2D map of a city the other one a 3D visualization showing the same scene, but with the buildings extruded. Participants were then asked to solve tasks which required navigating the 3D view and optionally the 2D map.

The duration of the entire experiment, including MRT, was between 40 and 60 minutes. After completion of the experiment, participants were debriefed and offered chocolate in return for their participation.

4.2.3. Pre Questionnaire/ Eye-Tracker Calibration

Participants first were asked to fill a questionnaire about their personal details concerning demographics and their spatial skills, training and experience.

After the participants had filled the pre questionnaire, the eye-tracker was calibrated. For this, participants were asked to keep their head in the same position, and told to keep it stable throughout the whole experiment. Then they were told to follow a point with their eyes moving on the screen, stopping at 9 different locations. If there were bad calibration points, calibration was repeated for these

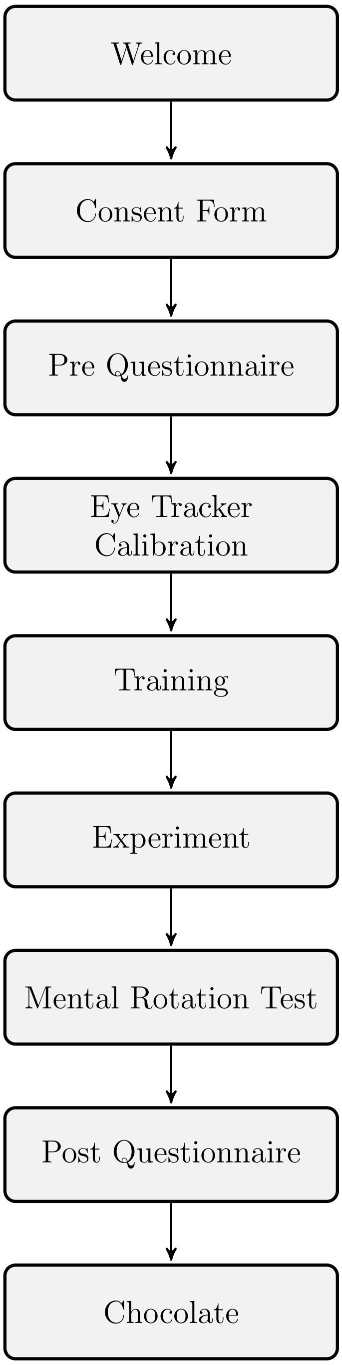


Figure 4: General session procedure

but it was never required to repeat the process more than once.

4.2.4. Training

Before the experiment started, participants were exposed to a training session. The training was built up in such a way that participants were incrementally introduced into the navigation concept and interaction tools. Throughout the training participants were encouraged to interact with the views and available tools and were given small tasks in order to practice and to reduce the learning effort in the controlled experiment.

Participants were guided through the training by the same plugin as they were guided through the main experiment, therefore the consistency of the training and experiment was very high.

The training started with an introduction to the basic navigation tools on the 2D view, panning and zooming. Then they were introduced to the basic navigation tools on the 3D view, panning, zooming, tilting and rotating. Participants were asked to zoom to objects first and then to rotate around it.

After that, they learned how to interact with the objects on the maps. First they were shown how to select an object as solution on the 2D view.

Finally the linking techniques were introduced individually. The IH was presented first after which the FH was also presented.

In the end of the training session both linking techniques were available at the same time to allow the participants to test them once again before the start of the real experiment.

4.3. Experimental Design

The main experiment consists of sixteen tasks, each divided into two subtasks resulting in 32 trials in total.

The tasks always involve the same assignment. The participants are presented two objects which we refer to as *target buildings* in blue on the 2D view. They then have to locate these two objects on the 3D view, compare their height on the 3D view and finally select the taller of the two on the 2D view.

Each pair of subtasks takes place in the same scene.

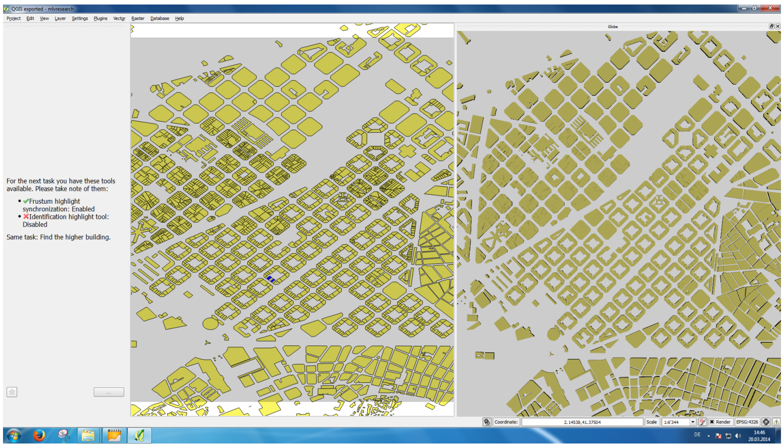


Figure 5: The initial state of the screen for subtask A. The two views show an almost identical extent and the target buildings are visible on both views, although highlighted only on the 2D view.

Subtask A On the 2D view, the scene is shown without highlighted buildings. An explanation of the assignment is shown and the available linking techniques are listed. The participants are asked to take note of the linking techniques available for this task and press the [Start]-button when they are ready.

As soon as the [Start]-button is clicked, the 2D view and the 3D view are set to the same extent (with minor differences due to different projections) and the two target buildings are highlighted in blue on the 2D view. (See figure 5)

The participants now need to locate the two buildings on the 3D view and navigate in the 3D view to a position where they are able to tell which of the two buildings is higher.

They then needs to click the appropriate building on the 2D view, what leads to a change of color of the building on the 2D view. This can still be changed until the [Confirm]-button is clicked.

Subtask B The tool availability for subtask B is always the same as in the preceding subtask A. In contrast to subtask A the two views are not initialized with the same extent, but instead the views are left in the state where they have been at the end of the subtask A. For reference, the solution highlight of subtask A is left visible (still only on the 2D view). Two new buildings are highlighted in the 2D view and the participant should immediately start to search for these. (See figure 6)

The main difference to subtask A is that there is no guarantee (in fact it is almost never the case) that the target buildings are visible on the 3D screen when Subtask B starts. This changes the possibilities to locate the target buildings substantially, as it is not possible to keep them visible throughout the whole search process. Instead there are basically two possibilities available to find the new target buildings:

- Zoom out in the 3D view until the scale is small enough to see the new target buildings (and most likely keep the old target buildings visible for reference)
- Keep a large scale and navigate to the new target buildings without losing reference.

A major difference is that there is no guarantee for the 3D view to be north-oriented and therefore aligned to the 2D view. It is one of the main expectations that participants will need to rotate the view at the end of subtask A in order to get the view into a state where they can compare the height of the two target buildings. It is further expected that at least some participants will in at least some of the tasks not be aware of the rotation they introduced themselves in the aforementioned action. The result of this is that the 3D view is not aligned to the 2D view.

4.3.1. Independent Variables

Linking techniques were varied systematically in such a way that every participant was exposed to every linking technique the same number of times. They were enabled and disabled systematically, such that they appeared in any combination. These combinations which we refer to as **interaction designs** are thus also systematically varied as a logical consequence. See table 2 for a systematic mapping of possible combinations.

Subtasks A and B differ in the navigation state of the 3D view as has been

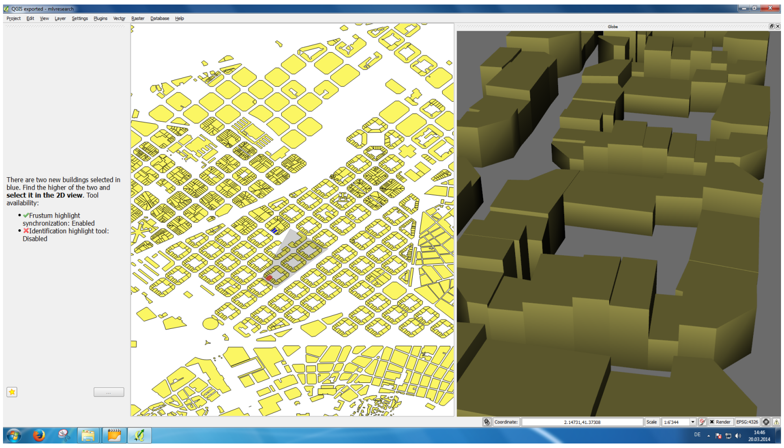


Figure 6: The initial state of the screen for subtask B. The 3D view is still left in the navigation state that was the final state of subtask A. The building selected as solution is still left highlighted in the 2D view (red). The new target buildings are highlighted in blue and are both not yet visible on the 3D view. In this case the heading of the 3D view deviates by slightly less than 45° from north and is therefore considered as aligned as will be discussed in section 4.7.2.

	Identify disabled	**Identify enabled**
Frustum enabled	frustum highlight (FH)	combined linking techniques (CH)
Frustum disabled	no linking technique (NH)	identification highlight (IH)

Table 2: The different combinations of the linking techniques in the top and left header and their reflection as interaction design in the table content.

explained before.

Scenes are the different neighborhoods where the assignments took place and are defined by their spatial extent and mostly recognizable by their particular building configuration. We used four different scenes in which the tasks were designed (See figure 7). The scenes were layouted such that the objects they contain are arranged in different ways: linearly, circularly and scattered. There are two areas which contain a scattered layout with a slight overlap.

Spatial ability is treated as a personal property and was determined with the MRT. As such the variable could not be controlled, but the effects on other variables could be analyzed.

4.3.2. Dependent Variables

Time was measured as the time from the task start to the time the participant clicked the [Confirm]-button. The raw time data was cleaned before being statistically analyzed as described in section 4.6.1.

Accuracy is a binary variable. Either a task was correctly solved (target building found and selected) or not. Aggregated on a participant (or other) level this becomes a number of correctly solved tasks.

Pitch is the angle between the surface (reference ellipsoid) and the camera direction. It ranges from $0°$ to $-90°$ with the former being a horizontal view and the latter being a vertical (top-down) view. With an empirical assessment of the data we found that a threshold of $-60°$ can be used to separate the two search stages *localization* and *comparison* (see section 4.7.1).

Zoom is the camera height in meters above the reference ellipsoid.

Context switches are extracted from the gaze data. Every time the users' gaze switched from the 2D to the 3D view or vice versa a new context switch is counted. Hence for each task *context switches* is a ratio scaled number. Context switches are extracted from the raw gaze-data.

We also measured a number of variables which are available for subsequent studies but that have not been analyzed in the current study.

Viewpoint The WGS 84 coordinates of the point on the reference ellipsoid where the camera was pointing to at a given moment.

Heading The angle in [°] relative to north. This can be used to analyze rotation movement.

Geo-referenced gaze-coordinates The WGS 84 coordinates of the location on the 2D and 3D screen where the user has been looking at.

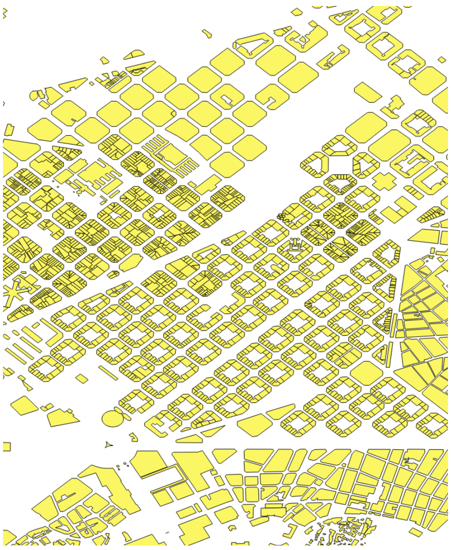
Feature-referenced gaze-data Ids of the features which the user has looked at have been recorded, sampled to five per second. For an assessment of data quality refer to section 6.8.

4.3.3. Counterbalancing

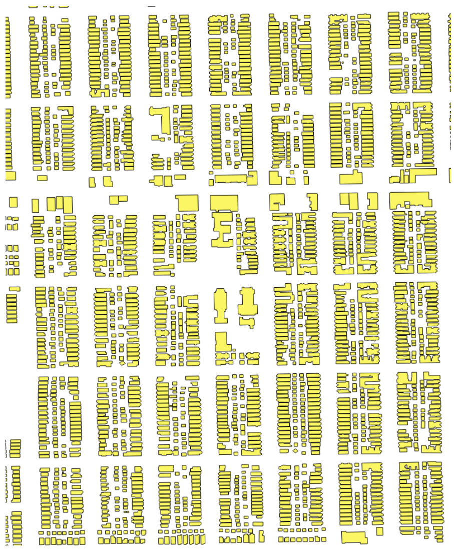
With the different independent variables *scene* and *interaction design* attention has to be payed not to mistake the effect of one for the other. To minimize this risk as well as the peril of a bias by an expected learning effect (see also section 6.1) the independent variables have been systematically varied and counterbalanced.

First, the interaction design was rotated on the scene, such that the first participant was solving the first scene/target building combination with interaction design NH, the second participant solved the same with FH, the third with IH and the fourth with CH, the second scene target building combination for the first participant was done with FH, for the second with IH and so on. This way, after every fourth participant, every scene/target building combination was presented once in combination with every interaction design.

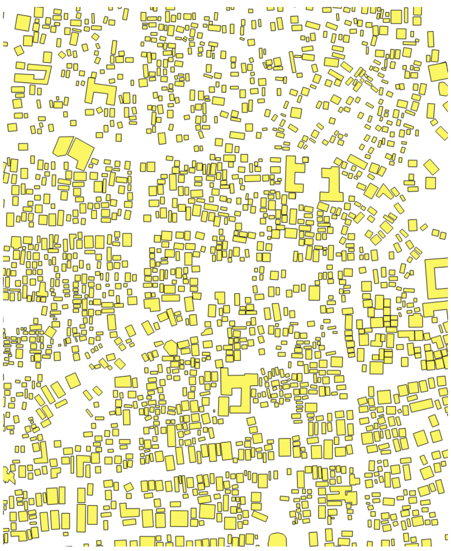
The combinations of independent variables, in such a way systematically prepared were then finally shuffled and presented in a random order, so that no learning effect for any of the independent variables could be expected.



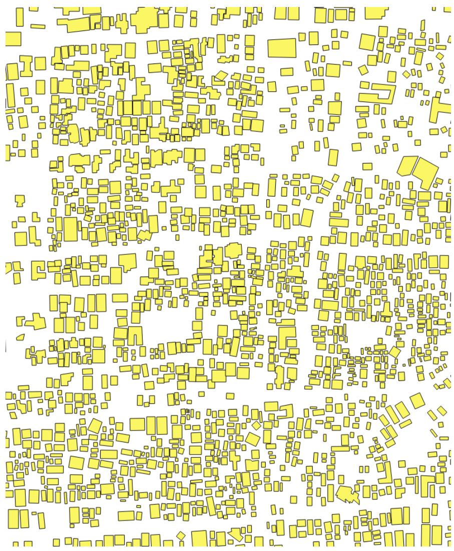
(a) Circular layout



(b) Linear layout



(c) Scattered layout A



(d) Scattered layout B

Figure 7: The different scenes in which the users had to solve the tasks. For each scene there are four times two (subtasks) pairs of target buildings defined (Not visible in this figure).

4.4. Materials

4.4.1. Online Questionnaire

The online questionnaire was created on *Surveymonkey* [17]. This platform offers an easy and intuitive way to create questionnaires and to download the results in a tabular form. Additionally there were questionnaires of previous studies present which could be used as a starting point.

Pre questionnaire The pre questionnaire consisted of questions regarding demographics [18], questions regarding potential problems with the eye-tracker or experiment setup [19] and training or expertise in the field of cartography and 3D visualization [20].

Post questionnaire The post questionnaire consisted of questions to assess the quality of the linking techniques. First, participants had to sort the interaction designs in order of preference. Then they had to individually score the tools considering ease-of-use, confidence, confusion and learnability. They were also offered the possibility to leave comments.

Vandenberg's Mental Rotation Test To determine the participants' spatial ability we included a Vandenberg's mental rotation test (MRT). The instructions and samples normally used in the paper-and-pencil test were integrated in the online platform. The test consists of twenty items such as the one visible in Figure 8. Each item contains a reference figure on the left and four candidate figures on the right. Two of the candidate figures depict the reference figure in a rotated position, the other two are different figures. Participants are asked to select the two rotated figures. In comparison to Vandenberg and Kuse (1978) who conducted the test in two sessions of ten items and three minutes each, we conducted the test in a single session of six minutes to solve all the twenty items. Points were counted according to Vandenberg and Kuse

[17] http://www.surveymonkey.com
[18] age and gender
[19] prescription glasses/contact lenses, imperfect color vision, physical and mental condition, use of medical drugs and English language competence
[20] Cartography/GIS systems experience, spatial data experience, 3D gaming experience, virtual 3D geo-browser experience

(1978). For each correct selection a point was given, therefore a minimum of zero points[21], a score of one point[22] and a maximum of two points[23] was possible for each item, resulting in a possible overall score reaching from zero to forty points with zero reflecting a very low spatial ability and twenty reflecting a very high spatial ability.

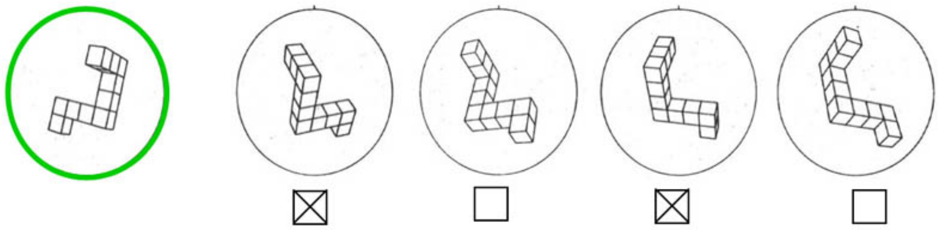


Figure 8: Example for a MRT assignment. On the left a reference figure, on the right four candidate figures of which the two selected represent a rotated state of the reference figure and thus are the correct answers.

4.4.2. 3D geobrowser

The implementation of the experiment was completely done in QGIS. The usage of an open-source GIS as foundation for the experiment enables us to access a well-established framework for displaying, querying and modifying spatial data with the possibility to adjust it to our particular needs wherever required.

In order to conduct the experiment we modified QGIS to suit our needs. QGIS already contains a *Globe*-plugin which enables anything that is rendered on the 2D map canvas to be rendered on a 3D digital globe surface. This was previously used by Bernasocchi, Çöltekin and Gruber, 2012 to research the visualization of multivariate spatio-temporal data on steep slopes. While their main focus was the terrain with draped textures, we required to extrude objects on the surface. The globe plugin is based on OsgEarth[24] which in turn is based on the OpenSceneGraph (OSG) library[25]. This offers a cross-platform 3D graphics toolkit which is used

[21] No selection made
[22] One rotated figure correctly selected
[23] Both rotated figures correctly selected
[24] http://osgearth.org
[25] http://www.openscenegraph.com

in different fields of visualization and simulation. OsgEarth and OSG offer a convenient highlevel application programming interface (API) for the implementation of 3D visualizations on top of Open Graphics Library (OpenGL) which itself is an abstraction layer to communicate with the graphics processing unit (GPU).

OsgEarth offers the possibility to extrude features based on their attributes. However, the QGIS globe plugin was only able to pass rendered and thus raster images to osgEarth to be used as textures. We therefore took previous work by Oslandia, 2012 that makes it possible to expose vector features provided by PostGIS as such to osgEarth and adapted it to be able to work for any vector layer source which QGIS supports and to the most recent QGIS API which had changed meanwhile. We extended their work to not only expose the features' geometries but also their attributes and implemented the necessary bits to leverage osgEarth's extrusion symbols.

To be able to offer the user the linking techniques two additional tools have been implemented.

Identification highlight All features visible on the globe are indexed with their feature id and whenever a single mouse-click is performed on the globe one of two callbacks is invoked. The first one accepts a feature id and is called whenever the mouse-click was actually performed on a feature. The geometry of this feature is then highlighted on the map canvas and any previous feature highlight is removed. The second one does not accept any parameters and is called whenever the user clicks on the 3D view but not on a feature. In this case any feature highlight is removed. See figure 9 for an example.

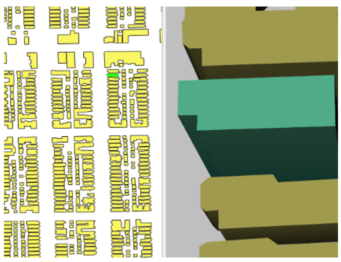


Figure 9: Linking technique: identification highlight (IH)

Frustum highlight Whenever a navigation event occurs on the globe a callback is invoked. This callback takes the four corner points of the view and converts them into geo-coordinates. With these geo-coordinates the four corner points of a quadrilateral is created which is then painted on the 2D map canvas. See figure 10 for an example.

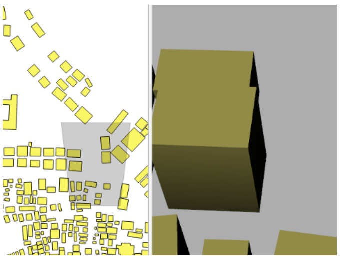


Figure 10: Linking technique: frustum highlight (FH). Only part of the display is visible on this screenshot. Therefore not the whole 3D view can be seen. See figure 6 for an example with a fully visible 3D view.

The goal of the experiment was to study the participant's ability to search, navigate and mentally link between the two views and explicitly not to study this particular implementation's details. We therefore tried to keep the number of available navigation tools as small, intuitive and consistent as possible.

In order to make the navigation between the different participants comparable we removed redundant functionality so it would be used by all participants in the same way. For navigation purpose the *Globe*-plugin offers buttons to zoom, pan, rotate and tilt the view. As these functions are also available via mouse interaction we removed the buttons completely. Instead participants received an initial training to get used to the navigation tools (See section 4.2.4).

As the participants had to navigate on two different views and thus to learn a number of navigation and interaction concepts, we tried to keep the learning process required for this as easy as possible. In order to do so, the navigation of the two views was made as similar as possible. The pan functionality already worked the same way on the 2D and the 3D map from a usage perspective. A minor discrepancy was that the reference for panning on the 3D view was the reference ellipsoid, what caused the cursor to have a slight offset after pan operations [26]. The zoom tool on the globe was adjusted to work the same way as on the 2D map (scrolling up is zooming in, scrolling down is zooming out).

[26] This was also noted by a participant. See section E

4.4.3. PyQGIS

PyQGIS offers an exhaustive API to QGIS functionalities for the python scripting language[27]. With this API it is possible to control large parts of a running QGIS instance. In contrast to C++ plugins such as the globe plugin (See section 4.4.2) it is possible to change a python plugin without compiling and test commands in a running QGIS environment. This makes it much easier and faster to develop PyQGIS plugins.

We extended the PyQGIS API to be able to also expose methods implemented by C++ plugins. We used this new possibility to make a couple of methods available for python plugins. In particular, we made it possible to synchronize the visible extent of the globe with the extent of the 2D map canvas[28] and to enable and disable the IH and FH.

4.4.4. Experiment Plugin

To conduct the experiment a PyQGIS (See section 4.4.3) plugin was created. The purpose of this plugin was to guide the participant through the experiment and to log appropriate information for later analysis.

This plugin defines all the tasks. Each task consists of the following information:

- Initial extent
- Two target buildings for subtask A
- Two target buildings for subtask B
- Linking techniques
- Task ID (1–16)

Each of the four comparison buildings is situated inside the initial extent.

When started, this plugin asks for the participant id and as soon as this is entered it creates a new folder with the participant's id as name where all logged data will be saved to.

[27] http://www.python.org
[28] Done on a best-effort basis as the projections do not necessarily match.

4.4.5. Data

The data used for this analysis consists of footprints of buildings acquired from three cities [29] from OpenStreetMap (OSM). In this dataset the height information was not available for each building. In order to get a decent virtual city, heights were assigned as random numbers between 5 m and 30 m to each building.

We aimed for an appropriate balance between ecological validity [30] and internal validity [31].

4.5. Gaze Data

We have recorded the users gaze for the whole experiment including training and task-solving, but excluding any questionnaires and the MRT with a *Tobii TX300* eye-tracker running at a sampling frequency of 60 Hz. The on-screen part of the entire session was conducted on a 23" display with a resolution of 1920 x 1080 pixels. The screen is optimized for eye-tracking with the given device and can be mounted on the eye-tracking device itself, therefore offering an optimal physical setup. Participants were put in a distance of approximately 65 cm from the device.

4.5.1. Geocoded Gaze Data

We implemented a methodology to perform geo-referenced gaze tracking. This was realized as a QGIS C++ plugin. The implementation is based on a two-level approach. The first level has the job of delegating incoming gaze data to second-level modules. For every view that should be tracked, a new module has to be implemented. Such a module offers basically two interfaces: One to get the rectangle on the screen, which the attached view occupies. The other one is a callback method that will be called, when gaze data is received which intersects with the rectangle and therefore the gaze is on this view. This method then takes the gaze coordinates and geo-references them in real-time, based on the currently visible extent. It may then enrich the received gaze data with further information.

[29] Barcelona, Jakarta and Boston
[30] I.e. building distribution in a real city and interactions with a 3D geo-browser
[31] I.e. adjusting building heights for comparability and reducing the contents of the maps to a basic set of visual variables. Both measures are taken to reduce bias.

There are three fields which a module can populate with custom data:

- Longitude (X)
- Latitude (Y)
- A feature ID (FID)

This data is then logged, along with the gaze data received from the eye-tracker and the ID of the view, which had the gaze and thus has done the geo-referencing. So there is one additional field

- Display ID (DID)

With this modular system, it is possible to track a users gaze while working with multiple independent or linked views side-by-side.

Modules have been implemented for the two views used in the experiment.

2D View Translating from screen coordinates to geo-coordinates is a rather simple and straightforward task for which QGIS already offers the required APIs. This then queries the data-source for features in a rectangular buffer around the current coordinates of the gaze. The implementation then takes the nearest neighbor as current feature.

3D View The 3D view introduced some additional challenges:

Asynchronous loading of features and terrain makes it more difficult to know if a feature is loaded at a given point in time and at which altitude and therefore point on the screen it is being rendered.

Pitching of the view is possible. If the view is pitched, the intersection of the gaze with the terrain will often be behind a building and not on its footprint, therefore making it impossible to locate a feature without knowing the exact state variables of the camera at the time of processing.

This profits from the implementation as real-time modules, as then the camera state variables are present and also the buildings are either present in the model or not at the exact time of processing. In this case, the 3D engine can easily be queried to do the intersection and return a feature ID (FID) for a given screen coordinate. As we made sure that all the objects currently being shown on the screen have the FID indexed in their geodes this process can also entirely be performed in memory.

Downsampling The feature-referencing should be considered an experimental by-product of this thesis. To make sure it does not affect the performance of the rest of the system we decided to limit feature-detection to a maximum of 10 per second and for the rest of the gaze events we only calculated the geo-coordinates.

4.6. Statistics

4.6.1. Data Pre-Processing

All data processing has been performed in python [32]. For statistical analysis which was performed in R we used the rpy2 package to act as a bridge between the two applications. To create plots we used the matplotlib package.

The task times as they were recorded were biased by some undesired effects which we assessed and corrected as outlined below.

Subtask A When the participants clicked on [Start] the 2D view was already set to the extent which it was supposed to be. To prevent users from starting to map match the two views before the timer started, the 3D view was only synchronized to this extent after the click. The 3D view was synchronized with an animation (fly-to) which lasted for one second. Thus participants actually were only able to start with map matching one second after the click, so we corrected the measured times consistently by one second.

Subtask B In contrast to subtask A users did not have to click on a [Start]-button in subtask B which was hence disabled. This was done to prevent users from navigation while time was not measured. As the participants were exposed to the two subtasks alternating sometimes they were waiting for this button

[32] For processing we used the packages scipy, numpy, statsmodels and pandas

to become active and this waiting time was included in the time data. We used the gaze-data to correct for this issue by scanning it for the first three consecutive seconds in which the participant uninterruptedly has not been looking at the task description view which contains the [Start]-button. The start of this interval was defined as the real task start time.

A learning effect was expected to happen and could be confirmed by looking at the data (See figure 15). As we can see, the learning effect is most intense over the first two tasks after which it is still present, but not as strong as before. For the statistical analysis these two tasks have therefore been removed in order to get a more stable data-set. The remaining learning effect is assessed in more detail in section 6.1.

4.6.2. Analysis of Variance

ANOVA To analyze the variance in the data and find group differences we used ANOVA for repeated measures as required for within-subject designs. The random effect was specified to originate from a per participant basis. Order and scene were not specified as random effects.

Pre-tests for ANOVA included test for normality and test for homoscedasticity as reported below.

Test for normal distribution was performed with the Shapiro-Wilk test for normality. H_0 for this test is that the data is distributed normally. We work with a significance level of 0.05 for this work and therefore will accept any p value > 0.05 to be distributed normally. Any data we work with is within the range of N=3 and N=5000. (Refer to Royston (1995) for more details)

Time was not distributed normally, but instead closely follows a *log*-normal distribution. Therefore we always worked with *log*(time) whenever an analysis involved working with mean values. As we can assume that the mean values in *log*-transformed domain match the median values and we are able to safely compare mean values in *log*-transformed domain and additionally the median value is robust against *log*-transformation we can deduce that comparing median values in the untransformed domain is a safe thing to do. We therefore resort to boxplots for comparing *log*-normal distributed data as boxplots are using quantiles for visualizing data and the effects outlined above apply not only to the median but to any quantile.

To be able to perform ANOVA we need to make sure that the variance in the groups is equal (homoscedastic). In order to do so we executed Levene's test for homogeneity of variance. H_0 for this test is that the population variances are equal. We test this on a significance level of 0.05. If the p-value is below this threshold, we will reject H_0 and assume that the difference of the variances is too big to be introduced by chance. If H_0 is above this level we assume homoscedasticity. (For more information see Fox and Weisberg, 2011)

We also conducted ANOVA when the assumption of normality was not met, as ANOVA can be run on unaggregated data (see section 4.6.2) and to improve comparability of different analyses of variance. Although often discouraged, ANOVA is fairly robust against the constraint of normality. As (a) the prerequisites for ANOVA is not that the *data* is normally distributed but that the *means of the residuals* are normally distributed and (b) the central limit theorem states that the means for large samples sizes tend against a normal distribution we can conclude that for a large enough sample size the distribution of the means is approximately normal. Minimum sample size estimated to be large enough are reported as 30 in literature (see Iman and Conover, 1983 reported in Helsel, 1987; and Wooldridge, 2012), the data analyzed in this thesis contains 104–106 values per group. For increased trust we also conducted a Friedman rank sum test (See section 4.6.2).

ANOVA was performed using the R functions lme and anova from the package nlme.

More information about ANOVA can be found in Chambers and Hastie (1991).

Tukey HSD Post-hoc testing When significant differences were found using ANOVA, post-hoc testing was performed with Tukey honestly significant differences (HSD) test using the R functions glht and mcp from the multcomp package.

Friedman Rank Sum Test Whenever the data being analyzed was not distributed normally we conducted a Friedman rank sum test in addition to the ANOVA for increased confidence in the results as they are not based on any assumptions about the distribution of data. This test requires an unreplicated complete block design, something we cannot provide because we have up to four measures per participant and stimulus, possibly reduced by the pre-processing outlined in section 4.6.1. Thus, the data needs to be aggregated, which was done by calculating the median in log-transformed domain to account for outliers. Although the assumption of normality is discarded, data still follows almost a log-normal distribution and aggregation therefore is still better performed in log-transformed domain. Due to this aggregation the sample size N is decreased and significance therefore reduced compared to ANOVA which is able to work on the complete sample size.

These tests were conducted using the R function friedman.test.

Wilcoxon Post-hoc testing If significant differences were found using Friedman rank sum test, post-hoc testing was performed with a Pairwise Wilcoxon test with Bonferroni correction using the R function pairwise.wilcox.test.

4.6.3. Regression Analysis

For regression analysis the ordinary least squares (OLS) method was used.

Regression analysis was performed using the python library statsmodel.api.OLS.

The model premises for linearity and homoscedasticitiy are checked by visual inspection of the scatterplots. The requirement not to be auto-correlated is based on the Durbin-Watson test which needs to be close to 2.

4.7. Assumptions for Navigation Analysis

4.7.1. Search stages

In order to research behavior in and effects related to navigation there is the need to first assess the navigation in connection to this particular experiment and specify the assumptions required for the subsequent analyses. This is particularly important for the 3D view with its increased degree of navigational freedom. We will only refer to navigation actions in the 3D view in this section.

For the next paragraph, recall that the values for pitch range from $0°$ to $-90°$ with the second one being a top-down perspective (For an explanation refer to section 4.3.2).

We expected that the participants would use the navigation tools in a given order. Our assumption is that they would zoom first and use tilting in the final *comparison* stage. A visual inspection of a series of sample tasks confirmed this assumption (Some examples included in figure 11). The expected behavior is visible in almost all trials for subtask A (figure 11: first column) with the lag between zooming and pitching being of different length, most likely caused by interposed panning. For subtask B (figure 11: second column) the interaction sequence is slightly more complicated as it involves navigation actions to restore the view to a state where it can be used for localization (Often involves decreasing pitch and zooming out) but in general the stages are comparable.

Based on this observation we divided the task into a *localization* and a *comparison* stage. We defined a threshold of a pitch of $-60°$ where larger values are treated as *comparison* and smaller values as *localization* stage. This threshold seems to be a good trade-off between a value which is conservative in terms of not being an accidental pitch operation but at the same time is exceeded in 96.2% of the trials. However, it still leaves possible sources for errors where (a) a user did never pitch back to less than the threshold level (see e.g. 6/2/1 figure 11), (b) the navigation has not been restored yet at the beginning of subtask B or (c) the user was able to solve the task with less intense pitching. We refer to a pitch level of more than $-60°$ as a *tilted* view.

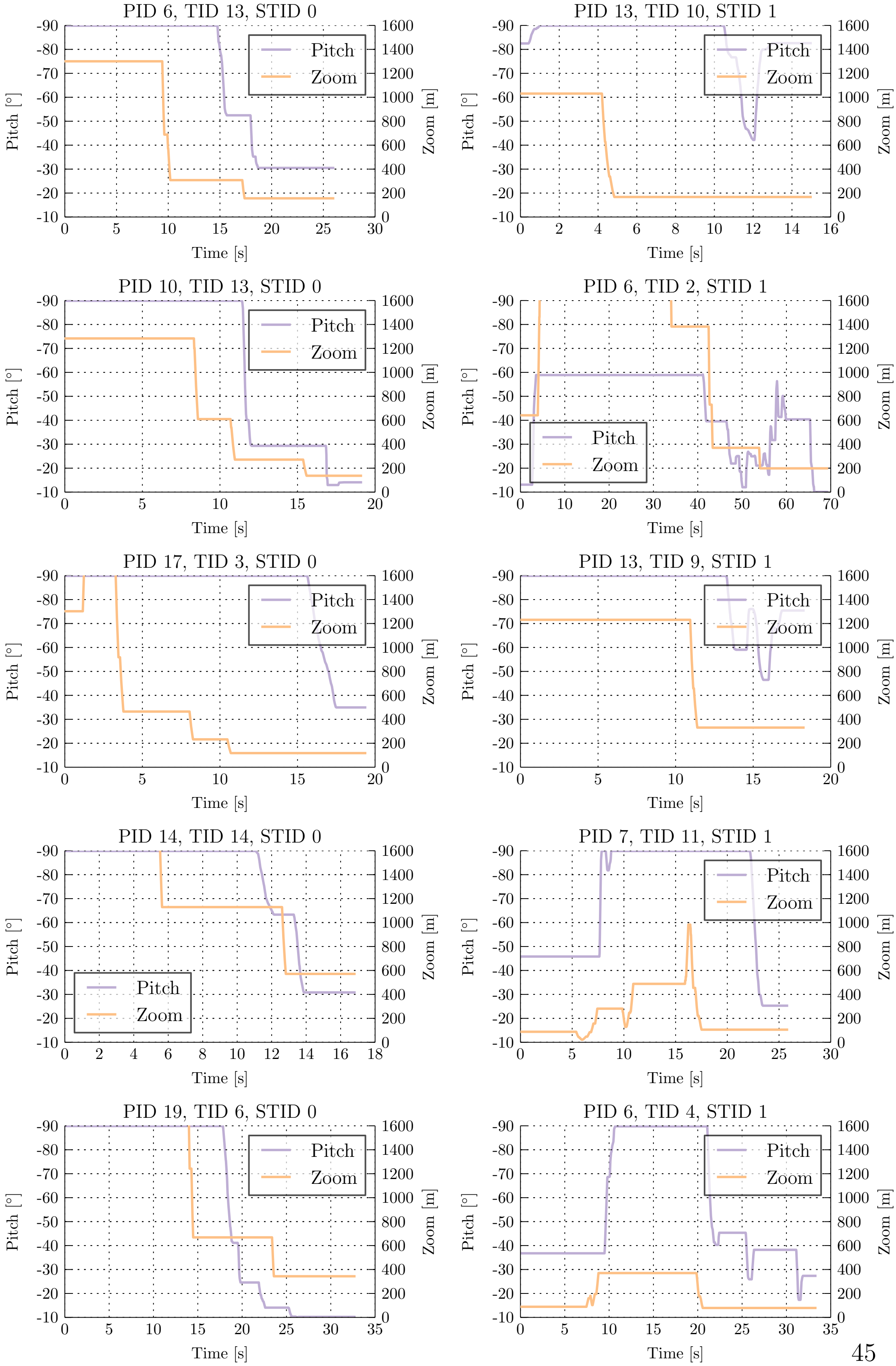


Figure 11: A number of samples showing the zoom and pitch behavior of participants while solving tasks.

4.7.2. View Alignment

In subtask B we tried to increase the risk of loosing the sense of direction, because in contrast to subtask A there was no guarantee that the initial orientation of the 3D view would point towards north and therefore be aligned to the 2D view. We therefore assessed view alignment only for subtask B.

The assessment of view alignment was performed on numbers operationalized based on two conditions:

- The 3D view is unaligned to the 2D view with a margin of 45° on both sides.
- The view has to be be tilted to a pitch of less than −60°. This constraint was introduced to remove the comparison stage, as in the comparison stage participants are expected to rotate in order to find the sweet spot to compare building heights. Rotation under this circumstances is considered mandatory and only optional or unintentional rotation was subject to research.

Figure 12 graphically depicts the two described constraints.

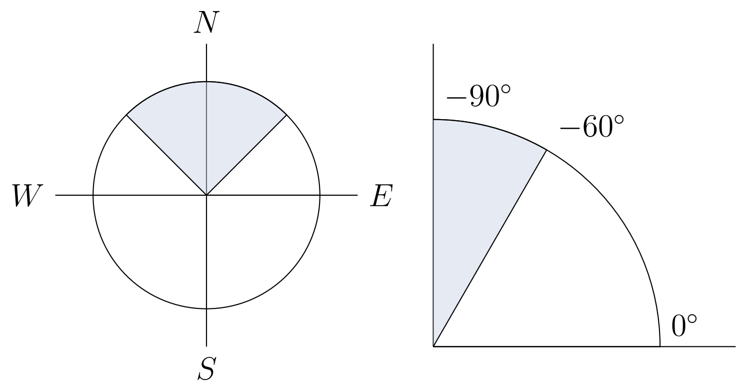


Figure 12: Camera direction considered when analyzing unaligned views. On the left the rotational component, 3D view rotation facing the white parts of the circle are considered unaligned. On the right the inclination component, if the pitch value is not in the blue part, the alignment is not further checked because we assume that the participant is in the comparison stage.

5. Results

5.1. Main Effects

To get a first impression of the data that was collected, these have been aggregated on a coarse level to examine the main effects. We performed this step on yet unfiltered data, therefore the first two tasks which in subsequent results are being considered training tasks are still present in the data presented in this section.

5.1.1. Spatial Ability

Spatial ability is the only independent variable which was measured and not self-reported (for the latter see table 1). Table 3 lists descriptive statistics of the results of the MRT, showing that the participants scored from very low to very high with the mean and median being situated relatively close to each other around 22.

	count	mean	std	min	25%	50%	75%	max
MRT Score	30	23.066667	6.90294	6	18.25	21.5	28.25	39

Table 3: Descriptive statistics of spatial ability of all participants.

5.1.2. Interaction designs

The main independent variable which has been tested is interaction design. Median times differ by a maximum of almost 10 s between CH and NH. FH and IH are in between with the latter having 2.5 s lower median time.

	count	mean	std	min	25%	50%	75%	max
CH	240	32.6	19.8	10.9	20.8	25.6	36.0	122.0
FH	240	38.6	24.0	11.2	22.9	32.7	48.1	214.7
IH	240	34.6	20.2	9.9	22.7	30.2	38.6	202.7
NH	240	40.7	23.3	13.1	27.1	36.0	49.4	242.5

Table 4: Descriptive statistics of the interaction designs.

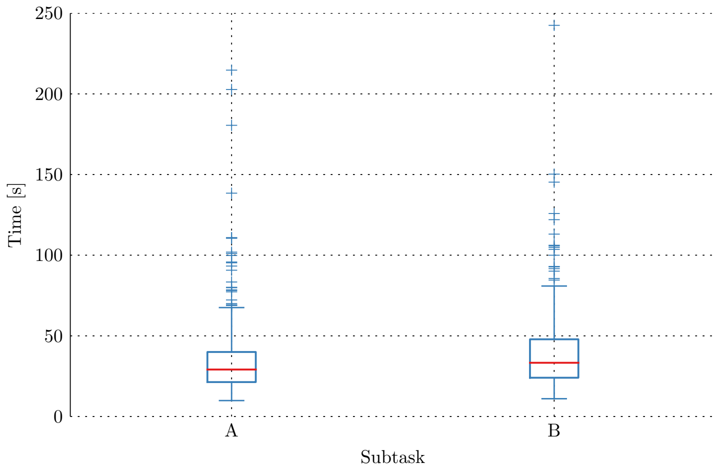


Figure 13: Overall comparison of efficiency in subtasks.

5.1.3. Subtasks

The two subtasks, which differ in their initial navigation state. Subtask A was solved overall in about 4 s less time than subtask B.

	count	mean	std	min	25%	50%	75%	max
A	480	34.6	22.1	9.9	21.3	29.1	40.0	214.7
B	480	38.6	21.9	11.0	24.0	33.3	47.9	242.5

Table 5: Descriptive statistics of subtasks.

5.1.4. Scene Layouts

The different tested layouts of the scenes showed some difference in the time that the participants required to solve them. While the minimum time required was between 9.9 s and 12.5 s we see a discrepancy of almost 7 s in the median times.

5.2. Accuracy

Looking at the accuracy we notice that the overall accuracy was very high. All participants were able to solve between 28 and all 32 tasks while half of the people

	count	mean	std	min	25%	50%	75%	max
Circular	240	36.0	17.4	10.9	24.5	31.9	43.3	138.4
Linear	240	39.6	22.0	11.0	25.1	34.3	48.3	202.7
Scattered A	240	35.9	26.0	9.9	20.9	27.4	42.0	242.5
Scattered B	240	35.1	22.0	12.5	22.1	29.5	39.5	214.7

Table 6: Overall comparison of efficiency in different scenes.

had a maximum of one error (See Table 7).

	count	mean	std	min	25%	50%	75%	max
Task Score	30	30.7	1.1	28	30	31	32	32

Table 7: Description of overall accuracy of the participants. Numbers refer to the amount of correctly solved tasks.

We further investigated if there is a connection between the number of correctly solved tasks and the amount of time required to solve the task. Therefore we aggregated the mean log(time) and the accuracy on a participant level to check if some participants traded speed for accuracy. The regression between the two variables showed a R^2 of 0.003.

5.3. Learning Effect

The tasks which the participants had to solve were always the same. The time in which they interacted with the experiment views (from the start of the training session to confirming the last task) lasted between 19 and 41 minutes. To analyze the effect of the experience that user gained while using the experiment we sorted all tasks by their appearance and then took the median time which should match $e^{mean(log(x))}$ for a value of x which is log-normal distributed. Figure 15 shows the median time that it took the participants to solve the tasks. We can clearly see that there is a strong learning effect observable ranging from slightly over 50 s average in the first task to 20 s – 30 s in the last tasks. We can also see that the learning effect is strongest in the first two tasks.

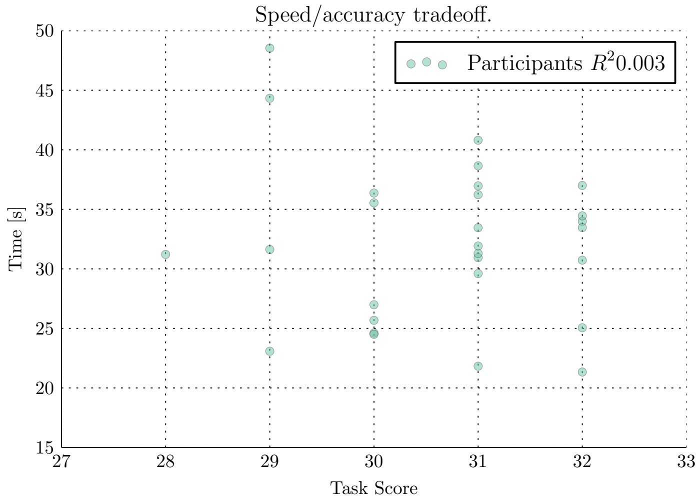


Figure 14: Accuracy (Task Score) against Time ($e^{mean(log(time))}$). There is no connection at all between participants' accuracy and their required time for solving the task. The y axis is corrected for the log-normal distribution of the data to calculate the means, while preserving an understandable representation of the numbers by transferring them back into the normal time domain.

5.4. Satisfaction

Analysis of satisfaction was performed on data collected with the post-questionnaire. In particular, participants were asked to order the interaction designs in order of preference. Numbers have been assigned from 0 (least preferable) to 3 (most preferable). It was not allowed to give two interaction designs the same score. CH was the most preferred choice by more than 75 % of the participants, NH was the least preferred choice by more than 75 % of the participants. For the two linking techniques in their single appearance, IH has been slightly preferred to FH. See table 8.

Participants were also asked to score the interaction designs for *ease of use*. The possible choices ranged from *very hard* (0) and *hard* (1) over *normal* (2) and *easy* (3) to *very easy* (4). In this question, scores could be assigned to every interaction design individually, therefore allowing the same score to be attributed to several interaction designs. Nobody found CH or IH hard to use. Less than 25 % of the participants found FH hard to use, nobody very hard and 50 % or more easy or

Figure 15: Learning Effect

	count	mean	std	min	25%	50%	75%	max
FH	30	1.40	0.67	0	1	1	2	3
IH	30	1.67	0.66	1	1	2	2	3
CH	30	2.60	0.89	0	3	3	3	3
NH	30	0.33	0.88	0	0	0	0	3

Table 8: Descriptive statistics of interaction design preference. The different interaction designs had to be arranged in order of preference. Scores range from 0 (least preferable) to 3 (most preferable).

very easy. NH was judged as hard or very hard by 50 % or more of the participants. See table 9.

Participants were also asked for feedback in terms of confidence. The possible choices ranged from *very unsure* (0) and *unsure* (1) over *undecided* (2) and *confident* (3) to *very confident* (4). In this question, scores could be assigned to every interaction design individually, therefore allowing the same score to be attributed to several interaction designs. 50 % or more of the participants stated to be very confident with IH and 75 % or more stated to be very confident with CH. 75 % or more of the participants stated to be at least confident with FH and 50 % stated to be at least confident with NH while less than 25 % stated to be very confident with FH and NH.

	count	mean	std	min	25%	50%	75%	max
FH	30	2.50	0.94	1	2	3	3	4
IH	30	3.23	0.77	2	3	3	4	4
CH	29	3.45	0.74	2	3	4	4	4
NH	30	1.60	0.93	0	1	1	2	4

Table 9: Descriptive statistics of interaction design ease of use. Each interaction design had to be judged individually in terms of *ease of use*. Scores ranged from 0 (very hard) to 4 (very easy) with 2 being normal.

	count	mean	std	min	25%	50%	75%	max
FH	30	3.0	0.53	1	3.00	3	3	4
IH	30	3.7	0.53	2	3.25	4	4	4
CH	30	3.9	0.31	3	4.00	4	4	4
NH	30	2.6	0.72	1	2.00	3	3	4

Table 10: Descriptive statistics of interaction design confidence. Each interaction design had to be judged individually in terms of *confidence*. Scores ranged from 0 (very unsure) to 4 (very confident) with 2 being undecided.

5.5. Interaction Designs

The influence of the different linking techniques on the task performance was one of our main interests. In order to assess the impact the variation of the linking techniques has, we analyzed the variance in times required to solve the tasks based on the interaction design. Time did not show a *log*-normal distribution for FH and CH for both subtasks. We therefore conducted a Friedman test besides the ANOVA. Friedman reported differences on a $p < 0.01$ level for both subtasks and so did the ANOVA (See tables 11 and 14). A post-hoc pairwise comparison of subtask A with a Tukey HSD test revealed differences between interaction designs FH and CH, CH and NH and IH and NH, all on a level of significance of $p < 0.01$. Post-hoc Wilcoxon testing of subtask A reported significant differences between FH and CH, CH and NH and IH and NH, all on a level of significance of 0.01, therefore confirming the results found with Tukey HSD. Post-hoc testing for subtask B did not differ much from post-hoc testing for subtask A. The only difference is, that the difference between FH and CH is slightly over a level of significance of 0.01, while with Tukey HSD reports the difference also on this level of significance. The results are still treated as significant, as we decided for a level of significance of 0.05 beforehand. The changes in probability are likely introduced by the data aggregation necessary for Friedman/Wilcoxon test and the therefore smaller sample size N.

	numDF	denDF	F-value	p-value
(Intercept)	1	300	6237.075093	0.000000
desgn	3	87	11.033247	0.000003

Table 11: ANOVA for subtask A. Comparing the *log*-transformed time grouped by interaction design.

	T Statistics	P Values
FH - CH	3.331109	0.004813
IH - CH	1.478777	0.450334
NH - CH	5.424542	0.000000
IH - FH	-1.865663	0.242800
NH - FH	2.083448	0.158503
NH - IH	3.962307	0.000414

Table 12: Tukey HSD for subtask A. Interaction design differences in efficiency.

	Combined	Frustum	Identify
Frustum	0.006229	-	-
Identify	0.230511	1.00000	-
None	0.000055	0.32953	0.001877

Table 13: Pairwise Wilcoxon test for subtask A. Interaction design differences in efficiency.

We further investigated the differences in more detail by assessing them graphically (See Figure 16) and comparing their mean values (calculated in *log*-transformed domain and then transferred back). The strongest difference can be observed between combined linking techniques and no linking techniques, were the corrected average was 25 s compared to 34 s which means that the time decreased by

	numDF	denDF	F-value	p-value
(Intercept)	1	300	6459.049848	0.000000
desgn	3	87	9.521414	0.000017

Table 14: ANOVA for subtask B. Comparing the *log*-transformed time grouped by interaction design.

	T Statistics	P Values
FH - CH	3.292135	0.005546
IH - CH	1.050831	0.719422
NH - CH	4.832784	0.000005
IH - FH	-2.255995	0.108520
NH - FH	1.528743	0.420136
NH - IH	3.798522	0.000890

Table 15: Tukey HSD for subtask B. Interaction design differences in efficiency.

26 %. The uncombined linking techniques led to a corrected average of 30 s (-12 %) for FH and 27 s (-21 %) for IH.

	Combined	Frustum	Identify
Frustum	0.010312	-	-
Identify	0.166449	0.439464	-
None	0.000010	1.000000	0.001877

Table 16: Pairwise Wilcoxon test for subtask B. Interaction design differences in efficiency.

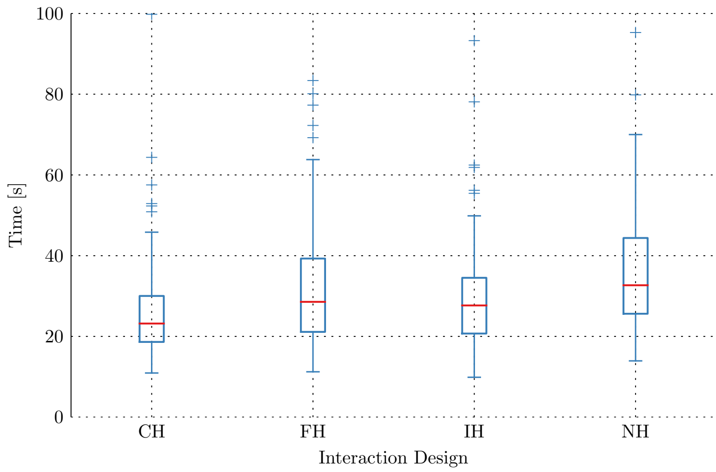


Figure 16: Time versus interaction design in Subtask A. Y axis limited to 100 s.

	$e^{mean(log(time))}$	%
desgn		
CH	24.868308	72.93
FH	29.887601	87.65
IH	27.179083	79.71
NH	34.098716	100.00

Table 17: Comparison of mean time required to solve subtask A. Means are calculated in *log*-transformed domain and then transferred back.

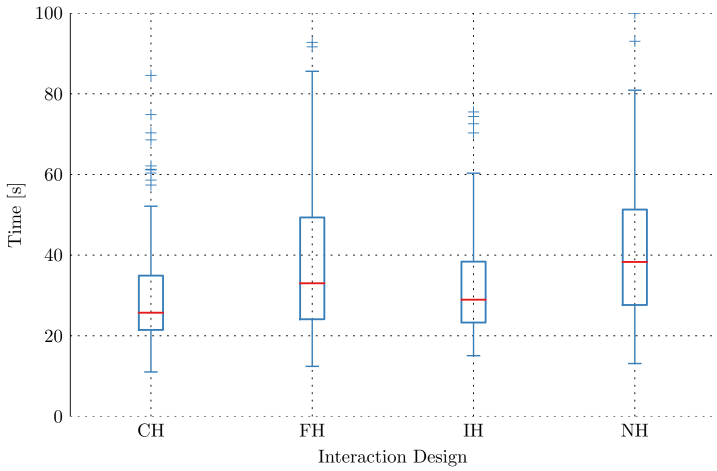


Figure 17: Time versus interaction design in Subtask B. Y axis limited to 100 s.

	$e^{mean(log(time))}$	%
desgn		
CH	28.961591	77.16
FH	34.337633	91.48
IH	30.936308	82.42
NH	37.534915	100.00

Table 18: Comparison of mean time required to solve subtask B. Means are calculated in *log*-transformed domain and then transferred back.

5.5.1. Map Alignment

The time spent with the 3D view in a navigation state meeting the criteria explained in section 4.7.2.

Table 19 lists the cumulated time that participants spent with a rotated and untilted view. Especially remarkable is that the time for the two interaction designs involving FH exceeds the time for the other two interaction designs by a factor of three to four.

	Time with unaligned view
Combined	399.14 s
Frustum	476.94 s
Identify	120.60 s
None	125.68 s

Table 19: Time spent with unaligned views, grouped by interaction design.

For the subsequent tests, the time spent with the 3D view unaligned and untilted was subtracted from the task solving time. The same analysis as in section 5.5 was repeated. The repeated check for normal distribution this time conceded this distribution to the times for every interaction design (See table 27). A subsequent ANOVA (See table 20) revealed differences between the interaction designs on a level of significance of 0.01. A post-hoc performed Tukey HSD test reported differences on a level of significance of 0.01 in the time required to solve the task for any pair of interaction designs but FH/IH (Table 21). In average, CH improved performance compared to NH by 30 %. FH, FH by 16 % and IH by 18 % (Table 22).

	numDF	denDF	F-value	p-value
(Intercept)	1	300	6727.529857	0.000000e+00
desgn	3	87	15.482445	3.742988e-08

Table 20: ANOVA results for the influence of interaction design on task solving times in subtask B, after subtracting the time spent in unaligned views. The test shows a highly significant effect on a level of significance of 0.01.

	T Statistics	P Values
FH - CH	3.391595	4.087239e-03
IH - CH	2.827409	2.441474e-02
NH - CH	6.778603	2.718015e-11
IH - FH	-0.583726	9.370193e-01
NH - FH	3.370656	4.187760e-03
NH - IH	3.974732	4.377898e-04

Table 21: Tukey HSD post-hoc testing for the influence of interaction design on task solving times in subtask B, after subtracting the time spent in unaligned views. The test shows significant differences on a level of significance of 0.01 between all interaction designs but the single occurence of FH and IH

	$e^{mean(log(time))}$	%
desgn		
CH	25.793433	70.45
FH	30.661617	83.75
IH	29.987836	81.91
NH	36.612888	100.00

Table 22: Comparison of mean (calculated in *log*-domain) time after subtracting the time spent in unaligned views used to solve tasks with the different interaction designs.

5.6. Spatial Ability

We were interested, if people who show a good spatial ability in the MRT also perform faster in the experiment. We found that there is a significant connection between the two results as shown in Figure 18. With the mental rotation ability it is possible to explain 38% of the time differences in subtask A and even 45% in subtask B with a statistically very significant level, as the probability of the F-value is below 0.01 (See Tables 29 and 31).

5.6.1. Map Alignment

To check if the time spent with unaligned views differs by the spatial ability of people, we ran a regression analysis on this data. 13 % of the time spent in a

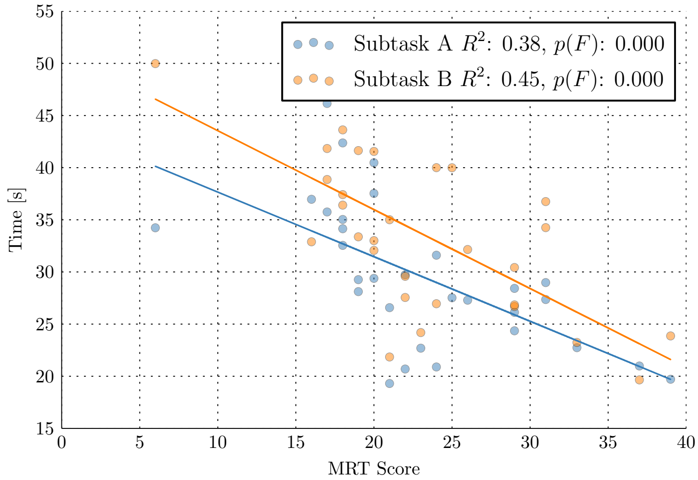


Figure 18: Mental rotation test score vs $e^{mean(log(time))}$.

non-aligned state can be explained by a participant's spatial ability with a level of significance of almost 0.05. Homoscedasticity can be considered critical for this data-set. This is further elaborated in the discussion (section 6.7.1).

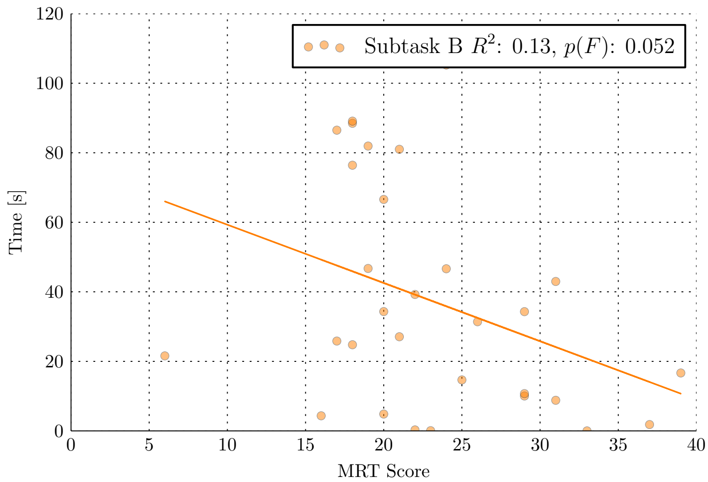


Figure 19: Time spent with unaligned views compared to spatial ability as found by MRT. OLS regression showing a small and borderline significant (on a level of significance of 0.05) correlation.

6. Discussion

We have performed an experiment to investigate users' behavior in a 3D browser and a linked 2D overview in connection with different linking techniques and conducted a MRT to determine the users' spatial ability. We have statistically analyzed the data collected in this process.

In this section we will review the data collected in this way, assess its uncertainties and its potential as well as embed it in a scientific research context and discuss what the impact of our results is and outline some ways forward.

Further we have also presented, implemented and shortly evaluated an approach to perform geo-referenced eye-tracking on multiple linked interactive views. There is an explanation of our contribution, the challenges of such a system as well as possible future improvements that have been identified.

Before the collected data is being numerically and statistically assessed, it is important to first take a step back and recall a statement by N. Andrienko et al. (2002): geovisualization researchers try to test a certain concept, but to do so they need to implement this concept and investigate this implementation. Implementations always comes with their own particulars which influence the results as a bias. The design of the UI, performance characteristics or occasional bugs may have a strong impact on the outcome. In this thesis, we want to test the concept of interaction designs but to do this we had to implement an experiment and had to make choices on its design, sometime consciously, sometimes not. While we tried wherever possible to reduce the bias introduced in such a way, it is unquestionable that it cannot be totally circumvented and therefore the results need to be critically judged like any results obtained from designed experiments.

6.1. Learning Effect

By giving the participants first an exhaustive training session we tried to minimize the learning effect throughout the experiment itself. However due to the limitations of having only a short time for the experiment and also because the repetitive nature of the tasks allows to develop a strategy, we expected a learning effect as it is known also from other studies (e.g. Tory et al., 2005). We could confirm this expectation, leading to an overall reduction of the time from the first to the last task of almost 50 %.

As possible sources for this learning effect there are a number of different factors which could be considered. As mentioned earlier, participants had received training covering the navigation and the linking techniques in an introduction session directly preceding the experiment, therefore the only first-time challenge for them was the task itself which they have not carried out before. The peak in the very first task (see figure 15) is most likely caused by the understanding of the task. There are a number of other factors which can potentially act as sources for a learning effect. We suspect that a combination of three factors listed below in addition to the one explained before are responsible for the performance improvement observed over all tasks:

Strategies
: Participants most likely developed strategies for task solving. For example they could have noticed that landmarks are an important reference to search for the target buildings and have started to look for these earlier.

Linking techniques
: While using the linking techniques, participants could carry experiences made with these from one task to the next, resulting in more educated usage.

Navigation
: The longer participants were using the system, the more they have used the navigation tools and were exposed to the navigation concept. For example, the amount of zooming induced by a given movement on the mouse-wheel. We can therefore assume that participants increasingly got used to the navigation and could improve their performance by applying the gathered experience.

One may be concerned about the influence of these results on our other findings. To deal with this, we considered to calculate the average learning effect and correct the raw task times by the expected learning factor. We eventually decided against this because there is still a big amount of uncertainty in this learning effect which can be observed by the variation in figure 15 as well as we expect significant differences in its magnitude between participants. Therefore the correction would have been very imprecise. Instead we treated the first two tasks that people solved as training tasks and removed them from the analysis. With this measure, the reduction of the median time reduced to approximately 35 %. As we have systematically randomized the order in which the tasks appeared, this affects all stimuli by the same amount and we therefore are safe to do so. Furthermore, the randomization helps to distribute the learning effect present in the remaining tasks over all stimuli which renders it safe to leave these numbers uncorrected despite the observed effect.

6.2. Accuracy

Overall participants had a very good accuracy, a single participant had only 28 of 32 subtasks solved correctly, all other participants had between 29 and 32 tasks correct. We can therefore conclude that the participants generally understood the task and that they all were able to solve it. However, we can not conclude that the task was easy, as the time required to solve the task sometimes adds up to several minutes indicating that an effort was required to solve the task.

Looking at Figure 14 and regarding the R^2 value of 0.003 for the regression of time versus task score, we can not see any indication that participants performed more accurately when they took more time. This further validates our findings, as we can compare the time which participants required to solve a task without worrying for the effect of a speed/accuracy trade-off.

The absence of a noticeable speed/accuracy tradeoff contrasts with non-spatial research (E.g. Pew, 1969; further examples can also be found in Wickelgren, 1977) but lines up with the findings by Wilkening and Fabrikant (2013) where they assessed the usage of a 3D geo-browser.

6.3. Scene Layout

The descriptive statistics for the different layouts show a difference between the time required to solve tasks in the different layouts. The fastest times could be measured on the two scattered layouts followed by circular and slowest in linear layout. This is possibly caused by the fact that it is more demanding for the HVS to visually search an object among similar objects than among different objects, thus additional effort can be expected in order to search objects in repetitive layouts like circular and linear.

However, these layouts were quite particular for this test and for any other experiment different layouts could be taken, possibly resulting in different results. Apart from the arrangement of the buildings, the layouts differed in the number of buildings, footprint area, distances between the buildings, shape of the buildings and likely further bias. A number of studies are available investigating the topic of visual search, layout and context of an object. These suggest that there is a connection between layout and visual search (e.g. Neider and Zelinsky, 2006). However, these results should be treated very carefully and empirically verified with an experiment eliminating any possible bias and designed for this particular purpose.

6.4. Subtasks

The two subtasks were, while being identical in the exercise assigned to the participant, quite different in the steps involved to solve them. The tasks were designed in a way that would force participants to re-orient themselves in the virtual scene (i.e. they would potentially "get lost") more often in subtask B compared to subtask A. We expected this to happen more often because

- often the 3D view was left unaligned as a result from the comparison stage of subtask A
- the target buildings are in almost every case not visible on the screen

We therefore predicted that the user performance will be worse for subtask B than for subtask A.

The numbers confirm this expectation (See figure 13). We see the expected performance difference from the minimum to the maximum time showing up on every quantile. We further could observe that the *lost sense of orientation*-problem was more present in subtask B due to observations made in the participant sessions. We could repeatedly watch participants navigate apparently lost, often with the

3D view not being north-oriented after having it left in this orientation at the end of subtask A.

6.5. Navigation in 3D Environments and Multiple Linked Views

Navigation in 3D virtual environments is a complex topic as it involves a broad range of involved processes ranging from the HVS over questions concerning HCI and spatial ability. With MLV yet another research domain on interaction designs enters the discussion. It is not easy and not even always possible to single out concepts as they depend on and influence each other. Brushing for example can be used to assist a user in linking certain objects mentally between different views, at the same time it can increase situation awareness and therefore help in navigation and it requires a user interaction to define the affected features. The same goes for MLVs while they are often researched in context of exploring information they can as well be used for overview and focus to embrace navigation. We therefore argue that an integrated approach is required to investigate problems in this domain and it not only fair, but required to transfer knowledge between the different research areas.

6.5.1. Linking Techniques

That viewing the data from multiple perspective through linked views requires linking to reveal its full potential has been brought up numerous times and can be treated as fact (e.g. Buja et al., 1991). The question therefore is not *if* but *how*. We have assessed a very simple version of brushing which highlights a single feature on demand on both linked views referred to as identification highlight (IH). We also proposed a new linking technique called frustum highlight (FH) which implicitly links a 2D and a 3D view of the same scene. We also assessed the combined availability of the two at the same time. IH and combined linking techniques (CH) revealed significant speed improvements compared to no linking technique (NH).

The strongest improvement introduced by a single technique was observed with the IH almost 20 % compared to NH in both conditions (subtask A and subtask B). We attribute this result to the fact that the IH could facilitate several stages of the task solving process. A highlighted feature can act as a reference in a large- or a small-scale view [33]. Participants are also trained to use brushing in the context of MLV as it is a widely known and used linking technique.

With FH we could also observe a speed improvement, although it is only around 10 % compared to NH less strong and not statistically significant. This will not be discussed further here, as it gets investigated in more detail in section 6.5.2.

These improvements in efficiency show that linking techniques can support the HVS when linking MLVs and when navigating in virtual 3D environments. These findings therefore underline the call for tools that support rapid situation awareness through effective overviews in Shneiderman (2003).

The two linking techniques cannot be only compared at a numerical level. The following list discusses specific properties which might add additional benefit to one or the other depending on the setting in which they appear:

Explicity/implicity
: FH is an implicitly available technique. It can be enabled by default or enabled on demand, but once it is enabled it does not require any further interaction but those performed for navigation. Brushing in contrast requires manual or semi-automatic selection of features and therefore explicit interactions to be updated and usable.

Visual Variables
: Both linking techniques require visual variables to be modified. Brushing modifies the appearance of features, the tested version of FH modifies the base map, which also may contain information, although this was not the case in our experiment. In any case, any change in visual variables should be carefully considered to not overexert the HVS.

Collision of Requirements
: Based on the previous two points we can deduce a potential collision of requirements. As brushing can be used for visualizing information of the data and therefore can often be changed while investigating different aspects of a data-set, its ephemeral appearance may reduce its aptitude for orientation.

[33] Up to the point of the natural limit where the scale becomes so small that the feature is no longer discernible

Applicability

Both linking techniques have their limitations in applicability. Both need special attention at small scales as they may be too small if rendered without any precautions just like that. FH additionally is most powerful in steep inclination (top-down), as when used in flat inclinations the upper parts of buildings close to the camera that may be in the 3D view may still be outside the highlighted area in the 2D view.

View impact

IH works on both views synchronously, a feature highlighted on one view is also highlighted on the other view. FH in contrast only adds additional information onto the 2D view. The highlighting of the IH can therefore be used as a visual anchor in navigation in the 3D view as long as the highlighted feature is visible. The FH in contrast leaves the 3D view unchanged, what leaves the visual variables unaffected, therefore decreasing the risk of visual clutter on the 3D view and leaving the available visual variables available for other purposes.

Thus, a designer of an application must consider carefully if his use-case is affected by some of these points and adjust linking techniques to his particular requirements.

6.5.2. View Alignment

Taking the results of the subsequent analysis of view alignment into account, the effects introduced by the FH can be explained from another perspective as well. This analysis was only performed for subtask B because experimental design ensures that this task was started with unaligned views.

After removing the time spent with an unaligned 3D view in the localization stage, the times for all interaction designs show a *log*-normal distribution. This indicates that the time spent in an unaligned view was causing outliers which prevented the times from showing a *log*-normal distribution. Combining this observation with the result, that participants spent three to four times as long with unaligned views with the two interaction designs containing FH [34]. The subsequently conduc-

[34] I.e. in the conditions FH and the condition CH which is a combination of FH and IH.

ted ANOVA with post-hoc Tukey HSD testing strengthens these observations by attributing significance also to FH compared to NH and an increased performance-boost for CH compared to NH (See table 22 and table 18). With the alignment correction taken into account the performance improvement compared to NH for the FH (16 %) is comparable to the one with IH (18 %) and the significance of the difference between FH and IH lost strength and remains only significant at a level of 0.1.

This suggests two things:

- Users are able to take advantage of the additional information they receive from the FH
- The FH can introduce confusion regarding map alignment and can greatly degrade the performance when this happens.

Thus, in order to get the full potential of the FH the problem of confusion needs to be addressed. To understand the problem, it is first important to be aware of the fact that participants did not get any additional help to realize their misalignment when they had no FH. This in turn means that they had all the necessary information as well with the FH enabled. On the contrary, it was even possible to query the screen for orientation information by tilting the view and observing the deformation of the FH on the 2D view and taking into account that the shortest edge is always the closest to the camera. Regardless, participants did not react to any of this information with rotating the view back to north-orientation, as they would without the FH.

These results contradict strongly with what we have expected. Instead of leading the users to align the view with the indication of direction they can receive from the FH, it leads to even less alignment. We can imagine different explanations for this behavior.

Ignorance

Participants may have been aware of the fact that the 3D view was in a rotated state or at least accepted this possibility. It is possible to navigate with a rotated view and the FH could help in such a situation as it gives an immediate feedback about the direction of a pan operation on the 2D view, allowing to inspect if the operation is properly directed to the target. The fact that a rotated view introduced many outliers indicates that in general the participants were not able to increase their performance with such a strategy.

Confusion

If participants were not aware of the fact that they are working with a rotated view it raises the question "*why?*". The most simple answer to this would be the lack of information. But that is no option due to reasons outlined above. The second answer to this would be confusion or mental overload. Participants have not been assessing the problem "what's wrong with the orientation" to which they could have reacted, but rather been busy with a question like "what's wrong with the system" or "what's wrong with me". Roughly speaking they have been struggling with strange reactions of the system which did not meet their expectations but they did not have an explanation for this.

In previous research we have evidence for both cases. Looking at Liben et al. (2010), it seems that in her study users were well aware of the fact that they were working with the map unaligned to the environment and mentally rotated the scene, possibly inducing a performance degradation. This conscious decision or at least awareness of alignment discrepancy has to be distinguished from the findings of Levinew et al. (1984). In their study people headed off into the wrong direction after being presented unaligned maps, suggesting a complete unawareness of orientation. While we suspect that in our experiment the first type (navigation with unaligned maps with conscience about it) happened, we are convinced that we were also confronted with the second type (users were unaware of the alignment). We suspect this because of observations of people apparently being lost with unaligned views who eventually figured the non-alignment out and rotated their view to north-up again.

To overcome these problems there are a couple of possible solutions at hand:

Training

By means of training, users could be made aware of this potential problem and given strategies to encounter it. If the problem is confusion, the awareness alone might help already. In the other case, they would need to accept that they are faster with both views aligned or to be taught how to increase efficiency with unaligned views. Apart from an organized training there is also the factor of experience. If it is considered that FH is new to the users, but that other interaction designs are not new, we can expect that a certain learning process will lead to increased awareness and response strategies to this problem. An informal feedback of a participant supports this theory when she reported that "towards the end of the session I realized, that the FH could indicate the view direction when tilted" (sic). Last but not least there are studies that suggest that perspective transformation could be a partially trainable skill (Darken & Cevik, 1999).

Compass

The most basic and obvious help in terms of orientation would be a compass. The integration of such an orientation aid was even proposed by a participant (See appendix E). While this device is undoubtedly one of the most helpful orientation aids in real-world navigation with tools that work without power supply (it is not by accident part of the basic equipment of mountain guides), it is just one of several options in virtual navigation.

Implicit techniques

Besides training and compass, there are additional possibilities available to support the users' sense of orientation.

Track-up map (orientation coupling)
This was proposed and is implemented in several systems. This focuses on the egocentric reference frame in the 3D view rather than the world reference frame with north orientation by adjusting the 2D map orientation along the 3D map orientation (Darken & Cevik, 1999; Plumlee & Ware, 2003).

Reference edge highlighting
It would also be possible to integrate orientation help into the frustum highlight itself to make its possible use as a compass more evident. By highlighting one of its edges differently, a user could potentially easier realize the

possibility to use the frustum highlight as an indicator for direction. In comparison to the implicit qualification of the frustum highlight for indication of direction, this has the additional advantage of also being noticeable in pure top-down view, when the frustum highlight is an exact rectangle.

6.5.3. Combining Linking Techniques

Another noticeable result is that the CH offers an additional increase in user performance compared to a single linking technique. Corrected for the time spent with unaligned views, in average performance improves by 30 % compared to NH.

This could be caused by two different effects. Either certain participants were able to improve their performance with one technique and other participants were able to improve their performance with the other technique, resulting in an increased improvement in the aggregated average [35]. Or all participants were able to increase their performance to a similar degree using both techniques and could additionally increase their performance with both techniques available in parallel at the same time. This second explanation would then suggest that both linking techniques support different stages of the search. As already stated earlier, we can very well imagine that FH improves the localization stage more while IH as a form of brushing improves the comparison stage more.

For now we can accept the fact that the combined appearance supports an increased efficiency and it should be carefully considered by designers of spatial MLVs to offer multiple linking techniques at the same time. This will either assist different users or facilitate different tasks, regardless of the exact effect, in average the efficiency should increase if our findings hold true for the task in question. It will be of great value to have further research performed that explains where exactly the different linking techniques can support a user's navigation and linking cognitive processes and the effects on the HVS. Once such evidence is available there will be a broader decision base available for developers of 3D geo-browsers with overview maps to decide what kind of linking techniques should be integrated.

[35] Instead of participants it would theoretically also be possible that a combination with certain scenes would add more benefit to one highlighting technique over the other. However, we could not see any indication for this by ourselves nor did a participant give a feedback pointing into this direction.

6.6. Satisfaction

After having conducted the experiment and the MRT, participants were asked to fill a post-questionnaire assessing their satisfaction with the different interaction designs. Thereby, preference, ease of use and confidence were assessed.

Preference for CH was clearly highest and IH was overall slightly preferred to FH, while NH was the least preferred interaction design. This matches very well with the efficiency which has been measured.

Ease of use analysis revealed similar results as preference analysis and efficiency. CH was judged as the easiest to use with IH being slightly easier to use than FH and NH was hardest. IH was never judged hard or very hard, FH has been judged to be hard by some participants but less than 25 % and never very hard. Both linking techniques in their single occurrence have been judged as easy or very easy by more than half of the participants.

Confidence rating revealed that with no interaction design, nobody was ever very unsure about the answer. This matches pretty well with the results of accuracy, which have been discussed in section 6.2. The tasks were all chosen and designed in a way that participants are able to solve them. The confidence results match the ease of use, preference and efficiency results to a certain degree. CH left the participants most confident, followed by IH and FH. NH was attributed the least confidence. The difference between confidence rating and the other ratings is however, that for the confidence rating there is only a small gap between CH and IH and a bigger gap between IH and FH, while for the other ratings and efficiency it was vice versa. This effect is likely to be caused by the last few seconds of a task, where the camera is tilted and the users are looking for one single building which they have chosen as solution. While IH in this case can precisely indicate if two features of the two views represent the same object, FH gives only approximate results and leaves room for interpretation to the user.

Overall, acceptance of and satisfaction with the tested linking techniques can be treated as good based on the received feedback. It can also be noticed that the feedback matches to a high degree the measured performance on an aggregated level. This could also be analyzed on a participant level to see if performance matches confidence.

6.7. Individual Differences in Spatial Ability

We have predicted that we would find a correlation between participants' spatial ability and participants' performance in the experiment. We found a correlation between the results of the MRT and the time required by participants to solve the tasks. With this measure of spatial ability we can explain 38 % of the variation of user performance in subtask A and even 45 % of the variation in subtask B. Thus, psychometric tests can be good predictors for localization, map matching and navigation tasks involving 2D – 3D side-by-side views.

Subtask B is considered to be more demanding as it requires to navigate to a point invisible on the 3D display at the beginning and is likely to involve starting with an unaligned display. The higher prediction level observed for subtask B could well be introduced by these additional challenges. Participants with better mental rotation abilities can be assumed to have less problems with not north-oriented displays by either keeping track of their rotational navigation movements previously made by path integration (Loomis et al., 1999) and reversing them or realizing faster that this happened. As the MRT score has also been shown to correlate with several other spatial abilities we could also imagine an improved situation awareness to help.

Our results tie in with previous research by Malinowski (2001), Hegarty (2004) and Hegarty et al. (2006) that found mental rotation abilities to correspond to environmental spatial ability. The indication found by Hegarty and Waller (2005) that the abilities required to perform mental object-based rotation differ from subject-based rotation are quite interesting for our case, as the MRT measures object-based rotation while the experiment requires skills in the domain of subject-based rotation. As they already stated themselves previously that only part of the variation is dissociated and a large portion of the involved processes is shared and therefore correlates (Hegarty, 2004), the correlation showing up in our results can neither confirm nor contradict their findings.

6.7.1. Map Alignment

The results for the test of time spent with the maps in non-aligned state show a small and borderline-significant [36] (on a level of significance of 0.05) correlation. This indicates that people with good spatial abilities spend less time with their views non-aligned.

One has to be cautious with interpreting these results. Visual inspection of the scatterplot on figure 19 shows that the main differences in non-aligned time are found around a MRT score of around 20 ("medium" spatial ability), thus the detected regression is mainly influenced by these measurements. While it can be easily verified on this plot that participants scoring higher than 25 show lower values of non-alignment, we can also see that we find quite a few examples of participants keeping their views well-aligned in the range of participants with MRT scores less than 25.

Therefore, we are only able to give a very vague answer to our hypothesis 3.B where we predicted that low-spatial people tend to work more often with unaligned views. We can confirm a slight indication into this direction, or, more precisely, the results suggest that while high-spatial people tend to work more with the views aligned, low- to medium-spatial people do not show a particular tendency.

6.8. Gaze Geo-Referencing

We have designed and implemented a system that is able to geo-reference gaze-data. In contrast to existing approaches this …

- … is done in real time
- … is performed on interactive views
- … can be applied to multiple views in parallel
- … was implemented for a 3D view

[36] Meaning that it is well possible that the results become significant with an increased number of participants.

In this section we will shortly discuss this implementation, the benefits it offers and the reasons why the data collected was not considered trustworthy enough to be used for the analysis in our work. In the end there will be a short discussion of potential improvements.

We implemented a two-level approach. The first level is responsible for the proper assignment of a gaze-point to a certain view which then – if applicable – forwards the gaze-package to an appropriate second level, view-specific module which invokes the required more specific code. This offers a performance aware solution as the CPU intensive calculations for geo-coding are only run for the view where the gaze really is directed to and where therefore meaningful results will be calculated.

The abstraction of the generated geo-coded data from the view offers the possibility to process the data in a unified way. If all view modules generate the data in the same way, it is easy to link the data between the MLVs without parsing all the data differently. For a future generic implementation we would, however, recommend to also introduce the possibility to record additional, view-dependent data. For example, it might be required to record the height value of the terrain at the position at the time of detection which may be determined by the status of a not yet fully-loaded network-based tile service, delivering the digital terrain model (DTM).

6.8.1. Feature Referencing

There was an implementation realized to allow the identification of features. Such a method allows to be used in different ways, either to give immediate or subsequent feedback to the user or in subsequent analysis of gaze-data. Our aim was originally to use it the second way to identify landmark buildings but we refrained from such a usage because due to lack of trust in the data collected this way.

We identified two main challenges which shall be shortly discussed here in order to help future researchers to address these issues.

Spatial resolution versus tracker resolution

It is typical for a interactive view that one can zoom and this action leads often to a change in item density [37]. Unfortunately the eye-tracker is not able to increase its precision the same way. The eye-tracker used in our experiment reports an accuracy of 0.5° in ideal conditions, which is about 0.7 cm in the corner of our 23" screen for a distance of 64 cm. Given that we had no ideal conditions (no fixed head, probably imperfect lighting conditions) this is rather optimistic. Therefore it is necessary to consider the involved uncertainty and take appropriate measures. We propose different possibilities to encounter this problem:

- In simple cases (like the 2D view) query for a set of features in a buffer around the gaze point and save all FIDs which are candidates. Along with the detected geo-referenced coordinates, one should save the buffer size or the coordinates of the bounding box of the buffer.
- In advanced cases it should be considered to do multiple intersections in the buffered area (in screen coordiantes). For the geo-coordinates, the buffer size corrected by the distance to the camera, the pitch of the camera and the camera aperture should be saved.
- Advance in technology could also improve eye-tracker accuracy and thus remediate the problem by improved devices.
- If there is additional information available about a users' possible gaze-patterns, advanced algorithms can be incorporated like the one proposed by Kiefer and Giannopoulos (2012) where a hidden markov model makes use of heuristics based on the knowledge that a user will follow a bike trail.

Intersection algorithm

In the evaluation of our implementation, false results with the intersection algorithm were revealed. We did not exactly investigate where this problems have their roots. So we can only speculate that either some inconsistencies in favor of speed were taken into account, or that these inconsistencies were introduced unintentionally [38]. Regardless of what caused these inconsistencies,

[37] It is possible to filter or aggregate data as well as to invoke some kind of decluttering algorithm.

[38] One explanation for certain false reports would be that the near and far clipping plane were not properly taken into account.

they would not heavily reduce user experience when used for their original purpose of feature selection on screen. As we could observe with the IH, which works with the same algorithm, participants just clicked a feature a second time or even rotated the view when it did not work immediately. For an analysis of gaze-data to identify features, a user has looked at this reduces trust in the data. We therefore conclude that requirements for user interaction tools with immediate feedback and the possibility for a user to detect problems as well as to take appropriate action are different from requirements for a tool, acting as a sensor to record data which will be used for research. This leads us to advice future researchers in this area to carefully evaluate and review any involved algorithms precociously.

6.9. Limitations

The experiment has been limited to flat areas. The appropriateness of the tested linking techniques on rough terrain and especially in high mountain areas is likely to be different and needs to be assessed separately. The same applies for features which are detached from the terrain, as all features which have been tested have been extruded 2.5D geometries.

The experiment which has been executed was based on very simple maps, making very scarce use of visual variables. More precisely, no colors have been used apart from those used to distinguish the buildings from the white background map, the depth cues in the 3D view and the colors introduced by the linking techniques. It is unlikely that the linking techniques would have shown the same effect in different contexts like for example in combination with a basemap or different content and styling of vector layers.

It further remains open, what tasks exactly benefit from which kind of linking techniques and how big the influence of individual differences on linking techniques is. Particularly, it would be very interesting to know whether spatial ability has an effect on how people benefit from different linking techniques. Such insights could then be used to design UIs which help to diminish the difference in efficiency of using MLVs in combination with spatial, 3D or any other kind of visualization.

The assumptions on which some of our findings base constrain the reliability of the results and need to be accepted in order to attribute plausibility. Therefore it would for example be very interesting to validate whether the assumption of the search stages and the execution of navigation in a given sequence can be verified. A possible approach for this could be to adapt the methodology used by Çöltekin, Fabrikant and Lacayo (2010) for finding sequences in gaze-tracks to navigation data.

7. Conclusion

A map matching, navigation and rotation experiment involving the usage of two side-by-side 2D and 3D views was conducted on 30 participants. Two different linking techniques have been assessed in their effect on participants' task solving efficiency. It has been shown that linking techniques are able to increase user performance and therefore to support the HVS. This highlights the importance of research in the area of situation awareness and effective overviews as has been asked for by Shneiderman (2003).

The strongest impact by a single linking technique could be observed with IH, which is based on the concept of brushing. This indicates that brushing, which is a widely-spread linking technique, is very well able to support mental linking in the context of a MLVs. This technique is already built into numerous applications and our research supports this trend.

FH, which has to our knowledge been assessed for the first time, was not able to improve efficiency significantly. The main reason for this was identified to stem from an increased usage of unaligned views, an effect contrary to our expectations. The results suggest that additional visual information requires to be decoded by the HVS and if it fails to successfully do this, confusion may arise and impact the user's performance. In such cases we assume that additional or modified tools may help to early prevent a user from such pitfalls. Further research for this is expected to reveal prospective insights.

The analyzed data further contains evidence that the combination of different linking techniques may increase a user's performance when combined. This indicates that different linking techniques support different tasks. Most likely certain techniques support panning and zooming operations, while others support pitching and rotating operations. These insights, combined with a more general discussion of benefits and drawbacks of the linking techniques that has been presented in section 6.5.1, should be considered when developing MLVs.

It has been highlighted once again that psychometric tests can be good predictors for spatial tasks (Liben et al., 2010; Francelet & Çöltekin, 2014; Malinowski, 2001; Hegarty, 2004). The difference in people's individual abilities of working with MLV and interacting with 3D views has impacts on education and application design.

Researchers, teachers, designers and GIS professionals have to be aware that not everybody is able to work and interact with spatial representations the same way and special care has to be taken to support people with lower spatial ability to have access and make the best use of visualizations and 3D geo-browsers.

Connecting eye-tracker data to geo-spatial vector data in MLVs is expected to give interesting insights into how people interact and guide their attention while working with geo-spatial representations. More work in this area will help future researchers to create knowledge about the way that the HVS deals with such representations for example by, but not limited to, identifying landmark buildings used in linking and navigation. A framework which facilitates the collection of such information could help to boost this research area. A number of challenges have been discussed and we hope that the identification of these challenges as well as the ideas presented to encounter them will provide to an increased research in this area.

A. Personal declaration

I hereby declare that the submitted thesis is the result of my own, independent, work. All external sources are explicitly acknowledged in the thesis.

References

Andrienko, N. & Andrienko, G. (2003). Informed spatial decisions through coordinated views. *Information Visualization*, *2*(4), 270–285.

Andrienko, N., Andrienko, G., Voss, H., Bernardo, F., Hipolito, J. & Kretchmer, U. (2002). Testing the usability of interactive maps in commongis. *Cartography and Geographic Information Science*, *29*(4), 325–342.

Aretz, A. J. (1991). The design of electronic map displays. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *33*(1), 85–101.

Bernasocchi, M., Çöltekin, A. & Gruber, S. (2012). An open source geovisual analytics toolbox for multivariate spatio-temporal data for environmental change modeling. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *1-2*, 123–128.

Bertin, J. (1983). *Semiology of graphics*. University of Wisconsin Press.

Bleisch, S. & Nebiker, S. (2008). Connected 2D and 3D visualizations for the interactive exploration of spatial information. In *Proc. of 21th isprs congress, beijing, china* (pp. 1037–1042).

Brunsdon, C. (2001). The comap: exploring spatial pattern via conditional distributions. *Computers, Environment and Urban Systems*, *25*(1), 53–68. {GISRUK} 2000. doi:http://dx.doi.org/10.1016/S0198-9715(00)00042-9

Buja, A., McDonald, J. A., Michalak, J. & Stuetzle, W. (1991). Interactive data visualization using focusing and linking. In *Proceedings of the 2nd conference on visualization '91* (pp. 156–163). VIS '91. San Diego, California: IEEE Computer Society Press. Retrieved from http://dl.acm.org/citation.cfm?id=949607.949633

Carlson, L. A., Hölscher, C., Shipley, T. F. & Dalton, R. C. (2010). Getting lost in buildings. *Current Directions in Psychological Science*, *19*(5), 284–289. doi:10.1177/0963721410383243. eprint: http://cdp.sagepub.com/content/19/5/284.full.pdf+html

Chambers, J. M. & Hastie, T. J. (1991). *Statistical models in s*. CRC Press, Inc.

Chen, C. (2006). *Information Visualization* (2nd). Springer-Verlag.

Cockburn, A., Karlson, A. & Bederson, B. B. (2009, January). A review of overview+detail, zooming, and focus+context interfaces. *ACM Comput. Surv. 41*(1), 2:1–2:31. doi:10.1145/1456650.1456652

Çöltekin, A. & Clarke, K. (2011). A representation of everything. *Geospatial Today, January 2011*, 26–28.

Çöltekin, A., Fabrikant, S. I. & Lacayo, M. (2010). Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Science*, *24*(10), 1559–1575.

Darken, R. P. & Cevik, H. (1999). Map usage in virtual environments: orientation issues. In *Virtual reality, 1999. proceedings., ieee* (pp. 133–140). IEEE.

Dodge, M., Mcderby, M., Turner, M., Goodchild, M. F., Roberts, J. C., Harrower, M., . . . Orford, S. (2008). *Geographic Visualization: Concepts, Tools and Applications*. doi:Doi10.1080/13658810802687335$\backslash$nPii913688706

Duchowski, A. (2007). *Eye tracking methodology: theory and practice*. Methods in molecular biology. Springer. Retrieved from http://books.google.ch/books?id=27jCNmafYU4C

Ekstrom, R. & Harman, H. (1976). *Manual for kit of factor-referenced cognitive tests, 1976*. Kit of Factor-referenced Cognitive Tests. Educational Testing Service. Retrieved from http://books.google.ch/books?id=vgklAQAAIAAJ

Fox, J. & Weisberg, S. (2011). *An r companion to applied regression*. Sage.

Francelet, R. & Çöltekin, A. (2014). *Realism and Individual Differences in Route-Learning* (Doctoral dissertation, Universität Zürich).

Fuhrmann, S. & MacEachren, A. M. (1999). Navigating Desktop GeoVirtual Environments. In *Ieee information visualization symposium, late breaking hot topics*.

Fuhrmann, S. & MacEachren, A. M. (2001). Navigation in desktop geovirtual environments: usability assessment. In *Proceedings of the 20th international cartographic conference* (pp. 2444–2453).

Giannopoulos, I., Kiefer, P. & Raubal, M. (2012). Geogazemarks: providing gaze history for the orientation on small display maps. In L.-P. Morency, D. Bohus, H. K. Aghajan, J. Cassell, A. Nijholt & J. Epps (Eds.), *Icmi* (pp. 165–172). ACM.

Gore, A. (1998, January 31). The Digital Earth: Understanding our planet in the 21st Century. Retrieved from http://portal.opengeospatial.org/files/?artifact%5C_id=6210

Hegarty, M. (2004, April). A dissociation between mental rotation and perspective-taking spatial abilities. *Intelligence*, *32*(2), 175–191. doi:10.1016/j.intell.2003.12.001

Hegarty, M., Montello, D. R., Richardson, A. E., Ishikawa, T. & Lovelace, K. (2006). Spatial abilities at different scales: individual differences in aptitude-test performance and spatial-layout learning. *Intelligence*, *34*(2), 151–176. doi:http://dx.doi.org/10.1016/j.intell.2005.09.005

Hegarty, M. & Waller, D. (2005). Individual differences in spatial abilities. In *The cambridge handbook of visuospatial thinking* (pp. 121–169). Cambridge Handbooks in Psychology. Cambridge University Press. Retrieved from http://books.google.ch/books?id=m91B8zm%5C_1qgC

Helsel, D. R. (1987). Advantages of nonparametric procedures for analysis of water quality data. *Hydrological Sciences Journal*, *32*(2), 179–190.

Iman, R. L. & Conover, W. J. (1983). *A modern approach to statistics*. Wiley New York.

Kiefer, P. & Giannopoulos, I. (2012). Gaze map matching: mapping eye tracking data to geographic vector features. In I. F. Cruz, C. A. Knoblock, P. Kröger, E. Tanin & P. Widmayer (Eds.), *Sigspatial/gis* (pp. 359–368). ACM.

Kirschenbauer, S. (2005). Applying "True 3D" Techniques to Geovisualization : An Empirical Study. In *Exploring geovisualization* (November 2004, pp. 363–387).

Kraak, M.-J. (1993). Three-dimensional map design. *The Cartographic Journal*, *30*(2), 188–194. doi:10.1179/caj.1993.30.2.188. eprint: http://www.maneyonline.com/doi/pdf/10.1179/caj.1993.30.2.188

Kray, C., Elting, C., Laakso, K. & Coors, V. (2003). Presenting route instructions on mobile devices. In *Proceedings of the 8th international conference on intelligent user interfaces* (pp. 117–124). IUI '03. Miami, Florida, USA: ACM. doi:10.1145/604045.604066

Levinew, M., Marchon, I. & Hanley, G. (1984). The placement and misplacement of you-are-here maps. *Environment and Behavior*, *16*(2), 139–157.

Liben, L. S., Myers, L. J. & Christensen, A. E. (2010). Identifying locations and directions on field and representational mapping tasks: predictors of success. *Spatial Cognition & Computation*, *10*(2-3), 105–134. doi:10.1080/13875860903568550. eprint: http://www.tandfonline.com/doi/pdf/10.1080/13875860903568550

Lippa, R. A., Collaer, M. L. & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior*, *39*(4), 990–997.

Loomis, J. M., Klatzky, R. L., Golledge, R. G. & Philbeck, J. W. (1999). Human navigation by path integration. In R. Golledge (Ed.), *Wayfinding behavior: cognitive mapping and other spatial processes* (Chap. 5). Wayfinding Behavior. Johns Hopkins University Press. Retrieved from http://books.google.ch/books?id=TjzxpAWiamUC

M. Jeanne, S. (1987). Cognitive maps as orienting schemata. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *13*(4), 615–628. doi:10.1037/0278-7393.13.4.615

MacEachren, A. M., Edsall, R., Haug, D., Baxter, R., Otto, G., Masters, R., ... Qian, L. (1999). Virtual environments for geographic visualization: potential and challenges. In *Proceedings of the 1999 workshop on new paradigms in information visualization and manipulation in conjunction with the eighth acm internation conference on information and knowledge management* (pp. 35–40). NPIVM '99. ACM. doi:10.1145/331770.331781

Malinowski, J. C. (2001). Mental rotation and real-world wayfinding. *Perceptual and Motor Skills*, *92*(1), 19–30.

McDonald, J. A., Stuetzle, W. & Buja, A. (1990). Painting multiple views of complex objects. In *Acm sigplan notices* (Vol. 25, *10*, pp. 245–257). ACM.

Monmonier, M. (1989). Geographic brushing: enhancing exploratory analysis of the scatterplot matrix. *Geographical Analysis*, *21*(1), 81–84. doi:10.1111/j.1538-4632.1989.tb00879.x

Montello, D. R. (2005). Navigation. In *The cambridge handbook of visuospatial thinking* (pp. 257–294). Cambridge Handbooks in Psychology. Cambridge

University Press. Retrieved from http://books.google.ch/books?id=m91B8zm%5C_1qgC

Munroe, R. (2013). Expedition. Retrieved March 27, 2014, from http://xkcd.com/1169/

Neider, M. B. & Zelinsky, G. J. (2006). Scene context guides eye movements during visual search. *Vision research*, *46*(5), 614–621.

Norman, J. F., Todd, J. T., Perotti, V. J. & Tittle, J. S. (1996). The visual perception of three-dimensional length. *Journal of experimental psychology. Human perception and performance*, *22*, 173–186. doi:10.1037/0096-1523.22.1.173

North, C. & Shneiderman, B. (2000). Snap-together visualization: a user interface for coordinating visualizations via relational schemata. In *Proceedings of the working conference on advanced visual interfaces* (pp. 128–135). ACM.

Okoshi, T. (1976). Physiology and psychology of depth perception. In T. Okoshi (Ed.), *Three-dimensional imaging techniques* (Chap. 3, pp. 43–59). Academic Press. doi:http://dx.doi.org/10.1016/B978-0-12-525250-8.50008-6

Oslandia. (2012). Postgis 3d demo. Retrieved March 14, 2014, from http://http://vimeo.com/54776907

Pew, R. W. (1969). The speed-accuracy operating characteristic. *Acta Psychologica*, *30*, 16–26. doi:http://dx.doi.org/10.1016/0001-6918(69)90035-3

Plumlee, M. & Ware, C. (2003). An evaluation of methods for linking 3D views. *Proceedings of the 2003 symposium on Interactive 3D graphics - SI3D '03*, 193. doi:10.1145/641514.641517

Poole, A. & Ball, L. J. (2006). Eye tracking in hci and usability research. *Encyclopedia of Human-Computer Interaction, C. Ghaoui (ed.)*

Rase, W.-D. (2003). Von 2D nach 3D - perspektivische Zeichnungen, Stereogramme, reale Modelle. *Kartographische Schriften*, *7*(Visualisierung und Erschließung von Geodaten), 13–24.

Roberts, J. & Wright, M. (2006). Towards Ubiquitous Brushing for Information Visualization. *Tenth International Conference on Information Visualisation (IV'06)*. doi:10.1109/IV.2006.113

Rosen, P. & Popescu, V. (2011, June). An Evaluation of 3-D Scene Exploration Using a Multiperspective Image Framework. *The Visual computer*, *27*(6-8), 623–632. doi:10.1007/s00371-011-0599-2

Royston, P. (1995). Remark as r94: a remark on algorithm as 181: the w-test for normality. *Applied Statistics*, 547–551.

Salvucci, D. D. & Anderson, J. R. (2000). Intelligent gaze-added interfaces. In *Proceedings of the sigchi conference on human factors in computing systems* (pp. 273–280). ACM.

Shepherd, I. D. H. (2008). Travails in the third dimension: a critical evaluation of three-dimensional geographical visualization. In M. Dodge, M. McDerby & M. Turner (Eds.), *Geographic visualization: concepts, tools and applications* (pp. 199–210). Wiley.

Shepherd, I. D. H. & Bleasdale-Shepherd, I. D. (2009). Videogames: the new gis? In H. Lin & M. Batty (Eds.), *Science press, beijing*.

Shneiderman, B. (2003). *Leonardo's laptop: human needs and the new computing technologies*. MIT Press.

Taylor, H. A. (2005). Mapping the understanding of understanding maps. *The Cambridge handbook of visuospatial thinking*, 295–333.

Tory, M., Atkins, M. S., Kirkpatrick, A. E., Nicolaou, M. & Yang, G.-Z. (2005). Eyegaze analysis of displays with combined 2d and 3d views. *Visualization Conference, IEEE*, 66. doi:http://doi.ieeecomputersociety.org/10.1109/VIS.2005.37

Tory, M., Kirkpatrick, A. E., Atkins, M. S. & Moller, T. (2006, January). Visualization task performance with 2d, 3d, and combination displays. *IEEE Transactions on Visualization and Computer Graphics*, *12*(1), 2–13. doi:10.1109/TVCG.2006.17

Vandenberg, S. G. & Kuse, A. R. (1978, October). Mental Rotations, a Group Test Of Three-Dimensional Spatial Visualization. *Perceptual and Motor Skills*, *47*(2), 599–604. doi:10.2466/pms.1978.47.2.599

Voigt, R. (2002). *An Extended Scatterplot Matrix and Case Studies in Information Visualization* (Doctoral dissertation, Hochschule Magdeburg-Stendal).

Wanger, L. C., Ferwerda, J. A. & Greenberg, D. P. (1992). Perceiving spatial relationships in computer-generated images. *IEEE Computer Graphics and Applications*, *12*(3), 44–51.

Ware, C. (2004). *Information visualization: perception for design*. Interactive Technologies. Elsevier Science. Retrieved from http://books.google.ch/books?id=ZmG%5C_FiqqyqgC

Ware, C. & Lewis, M. (1995). The dragmag image magnifier. In *Conference companion on human factors in computing systems* (pp. 407–408). ACM.

Weibel, R. & Burghardt, D. (2008). Generalization, on-the-fly. In *Encyclopedia of gis* (pp. 339–344). Springer US. doi:10.1007/978-0-387-35973-1_450

Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta psychologica*, *41*(1), 67–85.

Wickens, C. D., Liang, C.-C., Prevett, T. & Olmos, O. (1996). Electronic maps for terminal area navigation: effects of frame of reference and dimensionality. *The international journal of aviation psychology*, *6*(3), 241–271.

Wilkening, J. & Fabrikant, S. (2013, January). How users interact with a 3D geo-browser under time pressure. *Cartography and Geographic Information Science*, *40*(1), 40–52. doi:10.1080/15230406.2013.762140

Wolbers, T. & Hegarty, M. (2010). What determines our navigational abilities? *Trends in Cognitive Sciences*, *14*(3), 138–146. doi:http://dx.doi.org/10.1016/j.tics.2010.01.001

Wooldridge, J. (2012). *Introductory econometrics: a modern approach*. Cengage Learning.

Glossary

context switch describes a change of the gaze from one view to the other. Is normally used as cumulated count over a trial..

feature denotes an object originating from a vector data source. It may be drawn on a map and have attributes attached to it..

osgEarth is a 3D mapping SDK for OpenSceneGraph applications and used by the QGIS globe plugin..

PostGIS adds support for geographic objects to the PostgreSQL object-relational database. In effect, PostGIS "spatially enables" the PostgreSQL server, allowing it to be used as a backend spatial database for geographic information systems (GIS), much like ESRI's SDE or Oracle's Spatial extension. PostGIS follows the OpenGIS "Simple Features Specification for SQL" and has been certified as compliant with the "Types and Functions" profile..

QGIS is a free and open source geographic information system to create, edit, visualize, analyze and publish geospatial information on a wide variety of platforms..

Acronyms

AOI area of interest.

API application programming interface.

CH combined linking techniques.

DTM digital terrain model.

FH frustum highlight.

FID Feature ID (Unique).

GIS geographic information system.

GPU graphics processing unit.

HCI human-computer interaction.

HUD head-up display.

HVS human visual system.

IH identification highlight.

MLV multiple linked views.

MPI multiperspective image.

MRT Vandenberg's mental rotation test.

NH no linking technique.

OLS ordinary least squares.

OpenGL Open Graphics Library.

OSG OpenSceneGraph.

OSM OpenStreetMap.

UFO unidentified flying object.

UI user interface.

List of Figures

1. Possible relation between visual variables and depth cues 9
2. How zoom-levels can lead to lost sense of location. (Munroe, 2013) 13
3. Intersection of the camera frustum with the terrain. 19
4. General session procedure 25
5. Initial screen state: subtask A 27
6. Initial screen state: subtask A 29
7. Different scenes used for the tasks 32
   (a). Circular layout 32
   (b). Linear layout 32
   (c). Scattered layout A 32
   (d). Scattered layout B 32
8. Example MRT assignment 34
9. Linking technique: IH 35
10. Linking technique: FH 36
11. Zoom/pitch samples 45
12. Camera direction considered when analyzing unaligned views 46
13. Overall comparison of efficiency in subtasks. 48
14. Speed/accuracy trade-off 50
15. Learning Effect 51
16. Time versus interaction design in Subtask A. Y axis limited to 100 s. 55
17. Time versus interaction design in Subtask B 56
18. Mental rotation test score vs $e^{mean(log(time))}$. 59
19. Alignment compared to spatial ability 60
20. Sample plot: gaze x coordinate/time 104
21. Self-reported cartography experience versus user performance 104
22. Self-reported cartography experience versus MRT Score 105
23. Self-reported 3D gaming experience versus user performance 105
24. Self-reported 3D gaming experience versus MRT Score 106
25. Self-reported GIS experience versus user performance 106
26. Self-reported GIS experience versus MRT Score 107
27. Self-reported GIS experience versus MRT Score 107

List of Tables

1. Descriptive statistics of the participants who took part in the experiment. 23
2. Linking techniques combinations and interaction designs 29
3. Descriptive statistics of spatial ability of all participants 47
4. Descriptive statistics of the interaction designs 47
5. Descriptive statistics of subtasks 48
6. Overall comparison of efficiency in different scenes. 49
7. Description of overall accuracy of the participants 49
8. Descriptive statistics of interaction design preference 51
9. Descriptive statistics of interaction design ease of use 52
10. Descriptive statistics of interaction design confidence 52
11. ANOVA for subtask A. Comparing the *log*-transformed time grouped by interaction design. 53
12. Tukey HSD for subtask A. Interaction design differences in efficiency. 53
13. Pairwise Wilcoxon test for subtask A. Interaction design differences in efficiency. 53
14. ANOVA for subtask B. Comparing the *log*-transformed time grouped by interaction design. 54
15. Tukey HSD for subtask B. Interaction design differences in efficiency. 54
16. Pairwise Wilcoxon test for subtask B. Interaction design differences in efficiency. 55
17. Comparison of mean time required to solve subtask A 55
18. Comparison of mean time required to solve subtask B 56
19. Time spent with unaligned views 57
20. ANOVA results for interaction design versus efficiency after correction for unaligned view 57
21. Tukey HSD results for interaction design versus efficiency after alignment correction 58
22. Mean times per interaction design after correction for unaligned view 58
23. Shapiro-Wilk test for normality: Subtask A 109
24. Levene's test for homoscedasticity: Subtask A 109

25. Shapiro-Wilk test for normality: Subtask B 109
26. Levene's test for homoscedasticity: Subtask B 110
27. Shapiro-Wilk test for normality: Alignment corrected data 110
28. Levene's test for homoscedasticity: Alignment corrected data 110
29. OLS results for MRT compared to efficiency in solving subtask A. . 110
30. OLS conditions for MRT compared to efficiency in solving subtask A.111
31. OLS results for MRT compared to efficiency in solving subtask B. . 111
32. OLS conditions for MRT compared to efficiency in solving subtask B.111
33. OLS results for MRT compared to time spent with unaligned views. 111
34. OLS conditions for MRT compared to time spent with unaligned views. 112

B. Experiment Material

A number of on-screen materials have been used which cannot be made available in a printed version. Wherever appropriate we have illustrated the thesis itself with screenshots to support the reader in understanding what exactly has been done. The following pages contain the forms which have been either used by the experimenter to perform the experiment or handed out to to user on paper.

Instructions for running the experiment Evaluating multiple linked views with 2 and 3 dimensions: A Case Study with Eye Movement Analysis
January, 2014

Preparation:

QGIS

1. Connect to eye tracker
2. Start plugin
3. Enter participant nr (check for collisions)

tobii Studio

1. Connect to eye tracker
2. Start recording
3. Enter participant nr

Web Browser

1. Open Pre questionnaire
2. Open MRT Intro
3. Open MRT
4. Open Post questionnaire

Other

1. Switch cell phone to silent/airplane mode
2. Check battery and plug in if required
3. Start up stop timer app

When the subjects arrive:

Important: Never send a subject away! The exception to this rule is when you have to reschedule because of unexpected delays etc. Better inform them before they come. If not, make sure to apologize

1) Welcome them, use a few minutes to see if they need to talk a little bit ask how they are, if they found the lab easily (small talk for 2 minutes).
2) Explain always with same sentences what this is about
   (We are trying to get insight into how people mentally link objects in different views. If you are willing to participate, you will be given a series of tasks on a computer screen. While you are working on these tasks, you will be recorded by a camera, a microphone and an eye tracker. The eye tracker works with infrared light. This will neither be uncomfortable, nor do we know about harm caused by this. Please resolve the tasks as appropriate and fast as possible.).
3) Assign their participant number. Have them sign the consent form.
4) Fill the ‘pre-questionnaire’
5) Calibrate and start recording in Studio
6) Switch to QGIS

The University of Zurich - **Participant Information Statement and Consent Form** **Evaluating multiple linked views with 2 and 3 dimensions: A Case Study with Eye Movement Analysis**
January, 2014
Participant No:

Purpose of study

You are invited to participate in a study regarding an evaluation of a interactive 2d and 3d maps. We hope to learn more about the perception of and interaction on multiple linked views with different numbers of dimensions.

Description of study and risks

If you decide to participate, we will ask you to begin by filling out a short background questionnaire including demographic information. This will be followed by a session at the computer where you will be asked to use a map interface. During this process we will record your interactions with the computer using a webcam, audio recorder and eye tracking. The eye tracking device is non-contact, uses near infrared light and should not cause any discomfort. After the experiment we will ask you to fill out a second questionnaire.

The whole procedure should take approximately 60 minutes and there are no particular risks or benefits to you from participating in this experiment.

Confidentiality and disclosure of information

Any information and that can be identified with you in connection with this study will remain confidential and will be disclosed only with your permission. If you give us permission by signing this document, we plan to publish the results of this research in scientific publications. In any publication, information will be provided in such a way that you cannot be identified.

Compensation

We do not provide any compensation for your participation in this experiment, nor are there any costs for you for your participation.

Feedback to participants

If you would like to be kept informed about the results of this research, please leave your name and contact details with the experiment leader. A copy of publications resulting from this research will be sent to you when available.

Your consent

Your decision whether or not to participate will not prejudice your future relations with University of Zurich. If you decide to participate, you are free to withdraw your consent and to discontinue participation at any time without prejudice.

If you have any questions, please feel free to ask us. If you have any additional questions later, Matthias Kuhn (076 435 67 63, mku@access.uzh.ch) or Dr. Arzu Coltekin (044 635 54 40, arzu@geo.uzh.ch) will be happy to answer them.

You will be given a copy of this form to keep.

You are making a decision whether or not to participate. Your signature indicates that, having read the information provided above, you have decided to participate.

……………………………………. ………………………………

Signature of Research Participant Signature of Experimenter

……………………………………. ………………………………

Please PRINT name Please PRINT name

…………………………………….

Date and Place

The University of Zurich - **Participant Information Statement and Consent Form (continued)** **Evaluating multiple linked views with 2 and 3 dimensions: A Case Study with Eye Movement Analysis**
January 2014
Participant No:

REVOCATION OF CONSENT

Evaluating Interface Design for Interactive Geovisualizations: A Case Study with Eye Movement Analysis

I hereby wish to WITHDRAW my consent to participate in the research proposal described above and understand that such withdrawal WILL NOT jeopardize any treatment or my relationship with The University of Zurich.

……………………………………. ………………………………

Signature Date

…………………………………….

Please PRINT name

This section of Revocation of Consent should be forwarded to Dr. Arzu Coltekin, Geographic Information Visualization and Analysis, Dept. of Geography, University of Zurich, CH-8057, Zurich.

C. Additional Figures

In the course of analyzing the data collected for this thesis, a number of plots have been created which could not be discussed in detail and are therefore excluded from the results section 5. It may however be very interesting to skim through these to get ideas about how participants interacted with the experiment implementation, what kind of participants participated or what kind of relations of the data we have thought could be of interest but then refrained from further analyzing in favor of other more promising results.

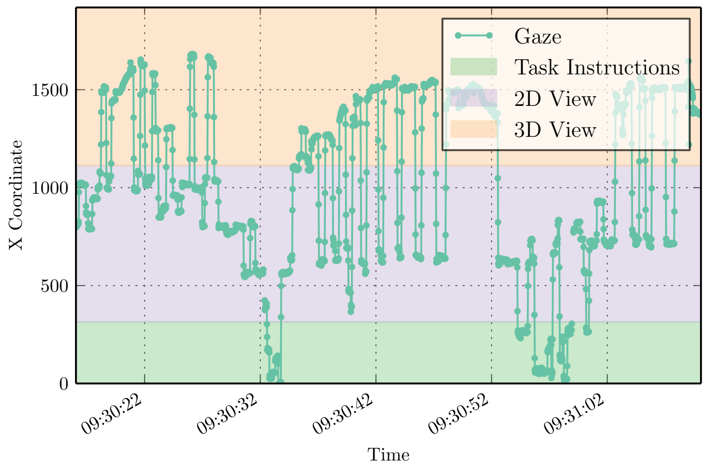


Figure 20: For a random sample task (Participant 5, Task 5, Subtask 1) the x coordinate of the user's gaze was plotted against time. This illustrates how the user guided his attention between the two views in the course of this task. It can further be observed, that the user was inspecting the task instructions while working on the task and that the eye-tracker lost the signal for a short time twice.

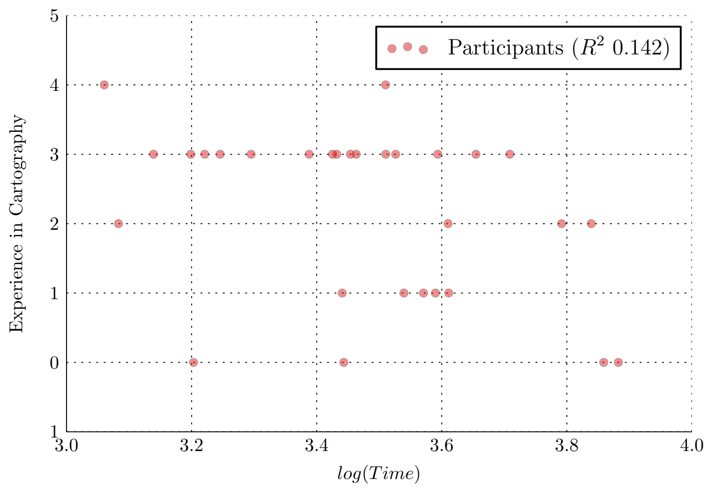


Figure 21: Self-reported cartography experience versus user performance

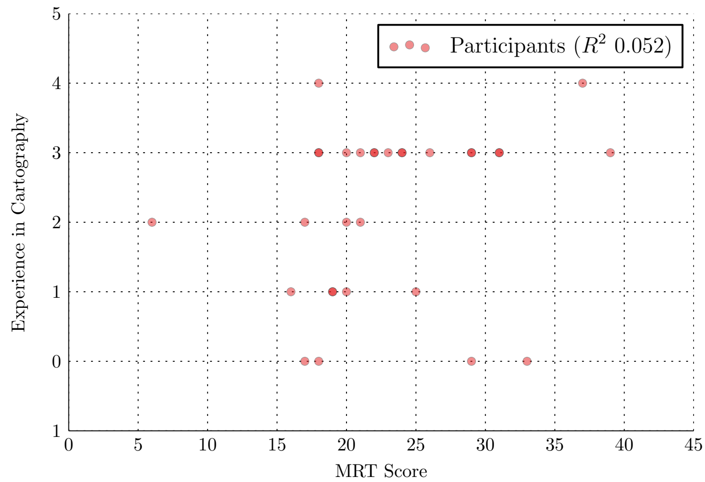


Figure 22: Self-reported cartography experience versus MRT Score

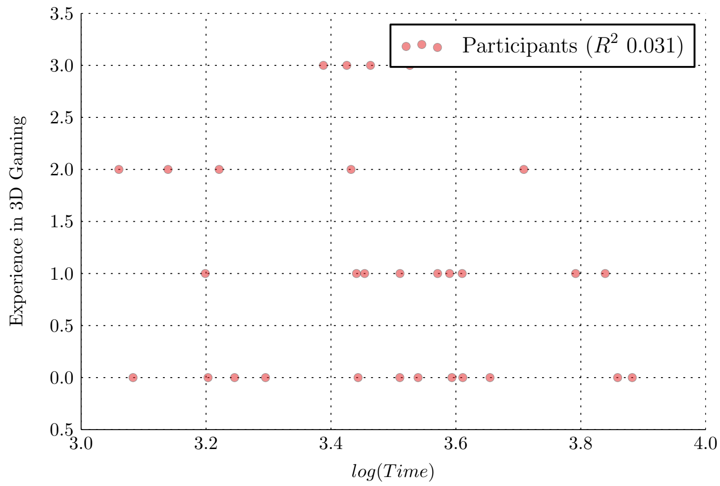


Figure 23: Self-reported 3D gaming experience versus user performance

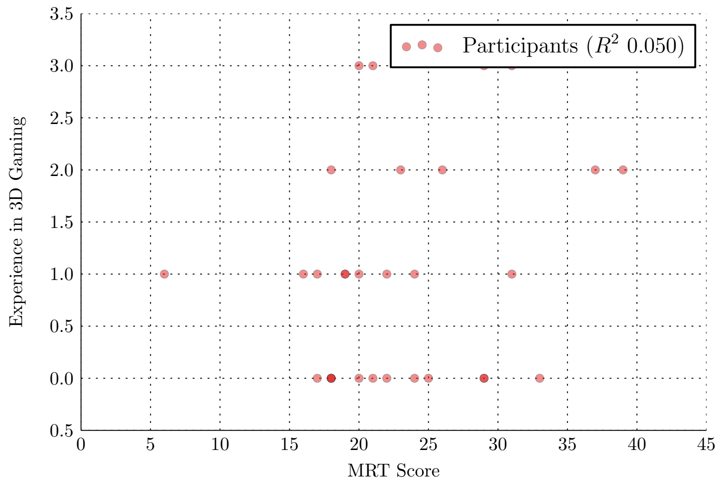


Figure 24: Self-reported 3D gaming experience versus MRT Score

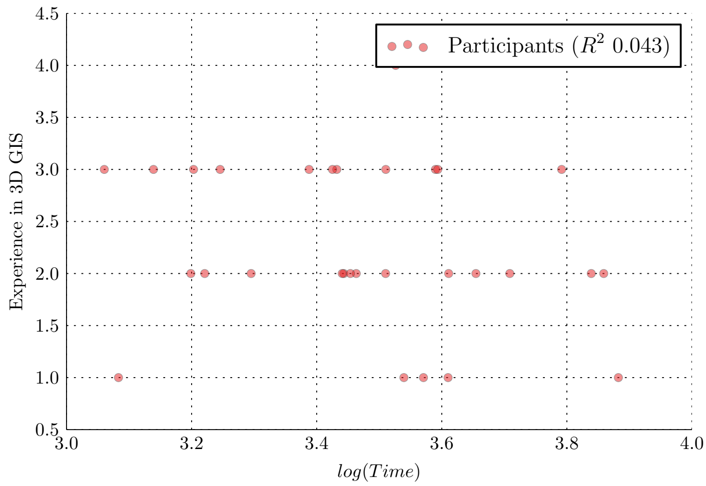


Figure 25: Self-reported GIS experience versus user performance

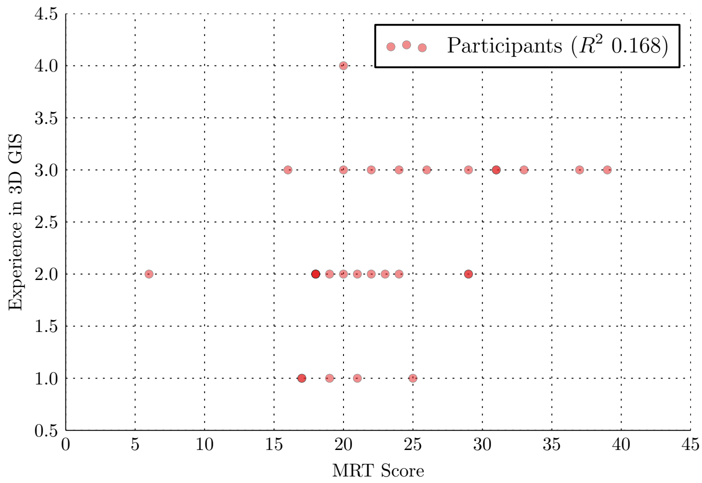


Figure 26: Self-reported GIS experience versus MRT Score

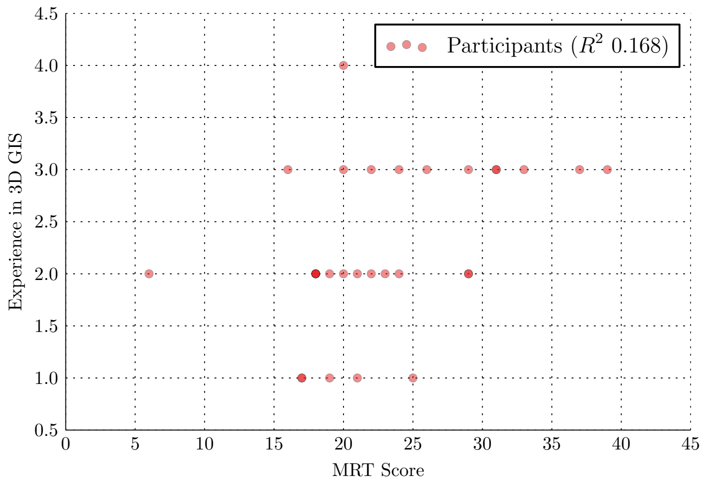


Figure 27: Self-reported GIS experience versus MRT Score

D. Statistical Tables

	P value	W statistic
CH	0.000014	0.923516
FH	0.189491	0.982588
IH	0.276991	0.984950
NH	0.596551	0.989582

Table 23: Shapiro-Wilk test for normal distribution for subtask A. Not all interaction desgns follow a *log*-normal distribution on a level of significance of 0.05.

	Df	F value	Pr(¿F)
group	3	1.545907	0.202038
	416	NaN	NaN

Table 24: Levine's test for homoscedasticity for subtask A. Equality of variance can be assumed.

	P value	W statistic
CH	0.002313	0.958463
FH	0.418464	0.987142
IH	0.005778	0.964118
NH	0.952764	0.994517

Table 25: Shapiro-Wilk test for normal distribution for subtask B. Not all interaction desgns follow a *log*-normal distribution on a level of significance of 0.05.

	Df	F value	Pr(¿F)
group	3	2.178657	0.089942
	416	NaN	NaN

Table 26: Levine's test for homoscedasticity for subtask B. Equality of variance can be assumed.

	P value	W statistic
CH	0.089289	0.978712
FH	0.295628	0.985076
IH	0.203831	0.983265
NH	0.882097	0.993176

Table 27: Shapiro-Wilk test for normal distribution. The times is from subtask B after subtracting the time spent on unaligned views and after applying a *log*-function. All data is considered to show a normal distribution on a level of significance of 0.05.

	Df	F value	Pr(¿F)
group	3	1.709148	0.164464
	416	NaN	NaN

Table 28: Levine's test for homoscedasticity for subtask B corrected for unaligned navigation. Equality of variance can be assumed.

Dep. Variable:	logttc	**R-squared:**	0.381
Model:	OLS	**Adj. R-squared:**	0.359
Method:	Least Squares	**F-statistic:**	17.24
Date:	Thu, 24 Apr 2014	**Prob (F-statistic):**	0.000280
Time:	17:25:56	**Log-Likelihood:**	-92.915
No. Observations:	30	**AIC:**	189.8
Df Residuals:	28	**BIC:**	192.6
Df Model:	1		

Table 29: OLS results for MRT compared to efficiency in solving subtask A.

Omnibus:	0.675	**Durbin-Watson:**	1.910
Prob(Omnibus):	0.714	**Jarque-Bera (JB):**	0.112
Skew:	0.113	**Prob(JB):**	0.945
Kurtosis:	3.197	**Cond. No.**	85.3

Table 30: OLS conditions for MRT compared to efficiency in solving subtask A.

Dep. Variable:	logttc	**R-squared:**	0.452
Model:	OLS	**Adj. R-squared:**	0.433
Method:	Least Squares	**F-statistic:**	23.13
Date:	Thu, 24 Apr 2014	**Prob (F-statistic):**	4.67e-05
Time:	17:25:56	**Log-Likelihood:**	-94.519
No. Observations:	30	**AIC:**	193.0
Df Residuals:	28	**BIC:**	195.8
Df Model:	1		

Table 31: OLS results for MRT compared to efficiency in solving subtask B.

Omnibus:	0.166	**Durbin-Watson:**	2.014
Prob(Omnibus):	0.920	**Jarque-Bera (JB):**	0.353
Skew:	-0.126	**Prob(JB):**	0.838
Kurtosis:	2.533	**Cond. No.**	85.3

Table 32: OLS conditions for MRT compared to efficiency in solving subtask B.

Dep. Variable:	olost	**R-squared:**	0.128
Model:	OLS	**Adj. R-squared:**	0.097
Method:	Least Squares	**F-statistic:**	4.115
Date:	Thu, 24 Apr 2014	**Prob (F-statistic):**	0.0521
Time:	17:26:30	**Log-Likelihood:**	-144.27
No. Observations:	30	**AIC:**	292.5
Df Residuals:	28	**BIC:**	295.3
Df Model:	1		

Table 33: OLS results for MRT compared to time spent with unaligned views.

Omnibus:	1.638	**Durbin-Watson:**	1.879
Prob(Omnibus):	0.441	**Jarque-Bera (JB):**	1.452
Skew:	0.411	**Prob(JB):**	0.484
Kurtosis:	2.303	**Cond. No.**	85.3

Table 34: OLS conditions for MRT compared to time spent with unaligned views.

E. Comments

In this appendix you will find a selection of comments left in the post-questionnaire. (sic!)

How easy was it to use...

- The frustum highlight was difficult to use in the beginning, but after you get used to it it becomes easy
- if I only navigated with the identification highlight, I didn't pay any attention to the surroundings. it was just a random clicking until I selected the closest building.
- the frustrum highlight was helpful as a first step in navigation, once zoomed in it was difficult to use
- it's very easy to get to know the tools
- shadowing in 3d-Mode and lack of Northing/ reset perspective made it harder without anything
- Mouse cursor should stay at the building when draging the map, then it would be also 'very easy' without any instrument

How confusing was the usage of...

- frustum highlighting was difficult for the first time used
- it would be nice if highlithing would work on both layers

How much learning effort was inolved

- It took a little time to understand that you have to click at the bottom of the building in 3D
- maybe the unusual name can be confusing - frustum :)
- Frustum doesn't take buildings into account... Thus, 3D-View and Frustum are not equal.

General feedback

- If I had the identification tool, I randomly clicked on a building, that I belived to be close. In that case, I didnt' pay any attention to buildings with a special shape. I think these tools make it quicker to find the right building, you don't use your sense of orientation and your map-skills anymore... It was an interesting study :).
- The chocolate was very yummy
- :-)
- It was very cool+ good luck
- would be easier if frustum highlight would have something like north-south compass
- BIG UUUUUP
- thank you
- Have fun with the video...
- well done!