



FAKULTÄT FÜR INFORMATIK  
DER TECHNISCHEN UNIVERSITÄT MÜNCHEN

Bachelorarbeit in Informatik

**Investigating Viewpoint Control Metaphors  
for hand-held devices in Virtual Environments**

Sandro Weber





FAKULTÄT FÜR INFORMATIK  
DER TECHNISCHEN UNIVERSITÄT MÜNCHEN

Bachelorarbeit in Informatik

**Investigating Viewpoint Control Metaphors  
for hand-held devices in Virtual Environments**

**Untersuchung von Navigationsmetaphern  
für tragbare Geräte in virtuellen Umgebungen**

Editor: Sandro Weber  
Supervisor: Prof. Dr. Gudrun Klinker  
Advisor: Dr. Marcus Tönnis  
Submission Date: 17.10.2011



I assure the single handed composition of this bachelor thesis only supported by declared resources.

Munich, 17. Oktober 2011

Sandro Weber

## Table of contents

1 Introduction.....	8
1.1 Premise.....	9
1.2 Intention.....	10
2 Related works.....	11
2.1 Intuitive spatial interfaces.....	11
2.1.1 Virtual Balance and EFS.....	11
2.1.2 Space Actor.....	13
2.2 Design Issues.....	16
2.3 Preceding Work.....	18
3 Navigation Interface.....	19
3.1 General idea.....	19
3.2 Functionality and metaphors.....	20
3.3 Problems.....	22
4 Technical realisation.....	24
4.1 Hardware.....	24
4.1.1 Tablet.....	24
4.1.2 ART tracking system.....	25
4.1.3 FRAVE.....	26
4.2 Software.....	26

---

Table of contents

---

4.2.1 Rendering software and system.....	26
4.2.2 Android client.....	27
4.2.3 Porting to FRAVE setup.....	28
5 User study.....	29
5.1 Primary goals.....	29
5.2 Choice of probands.....	29
5.3 Design process.....	30
5.3.1 First approach.....	30
5.3.2 Second approach.....	31
5.3.3 Third approach.....	32
5.4 General adjustments to the interface.....	34
5.5 Results.....	35
5.5.1 Analyzed data.....	35
5.5.2 Analyzation program.....	35
5.5.3 Evaluation.....	36
5.6 Conclusions.....	43
6 Discussion.....	44
6.1 Nature of the metaphorical conceptions.....	44
6.2 Future work.....	44
6.2.1 Use of head tracking.....	44
6.2.2 Dependency of switch between airplane and window frame .....	45
6.2.3 Strafing position changes.....	45

---

Table of contents

---

7 Appendix.....	46
7.1 List of references.....	46
7.2 Schedule for user study.....	47

# **Acknowledgement**

I want to thank Prof. Dr. Gudrun Klinker and Dr. Marcus Tönnis for the chance to work on such a big, cooperative and interesting project and for their continuous support.

I also want to thank my parents for giving me every opportunity to realize my ambitions and desires.

Thanks to all the friends making me lose and keep my sanity.

## 1 Introduction

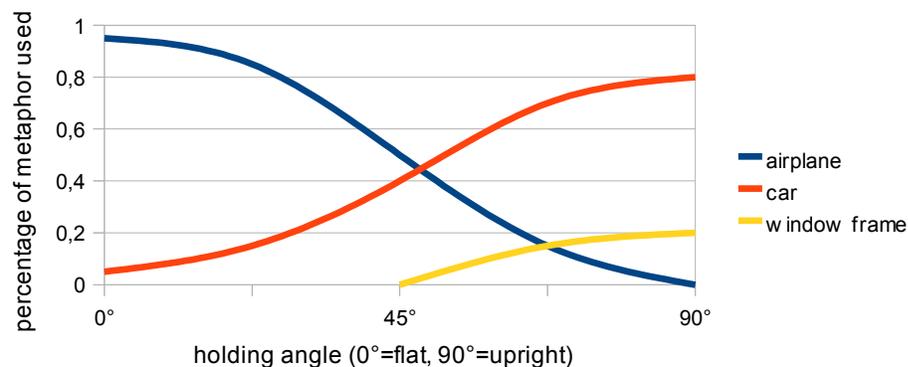
Today two dimensional map data is available almost everywhere. While this kind of data is well tried and proven to be useful in many situations where an overview suffices, diving into three dimensional virtual space is becoming more and more aspired. Sometimes more visual landmarks like in google street view - which fakes the impression of being three dimensional - are a great help. And in case of 3D simulation or prototyping scenarios the need to navigate all three dimensions is a given. Unfortunately, navigating three dimensional virtual space is a lot more complex and leaves a lot more options than the two dimensional case. Due to the amount of different setups of screens and input devices there are no standard interaction methods yet and it is questionable if there ever will be.

Still there are some basic taxonomies for user interfaces. In a virtual environment, ways of movement are not bound to physical limitations, however the user needs enough cues to his position and orientation and how to manipulate them in order to easily find his way around. This usually involves a mental metaphor of motion that is quickly grasped by the user and shows parallels to real world situations, such as driving a car using a steering wheel. There are numerous devices developed for this task, usually trying to depart from the classical mouse/keyboard setup enabling gestural input while standing free. But navigation is often only one of the tasks a user needs to handle. When the necessity to easily access system tasks arises, it poses additional requirements to the navigation interface/device layout.

## 1.1 Premise

The interaction device used in the following studies is a hand-held tablet PC. It is tracked and can therefore be used as a way of gestural input while offering additional information and options on its screen. In a former diploma thesis by Nicolas Heuser [4], two steering metaphors have already been implemented and tested. The first is the before mentioned car metaphor, which lets the user turn exactly the way a steering wheel in a car would. The second metaphor is more analogous to an airplane. Holding the device flat before you, turning is performed by tilting it sideways like the wings of an airplane. Details of the control will be explained in chapter 3.2. To avoid influencing any conception of steering metaphors the users could have beforehand, both metaphors should also be usable simultaneously with a seamless transition between both modes of motion.

User tests showed that in addition to these two metaphors, a third one is intuitively used during “car mode” - with the device held upright - that could be described as a window frame that is aligned with the desired perspective. Here, the sideways tilting used during “airplane mode” is anticipated to perform a turn into the exact opposite direction as it would in airplane mode. The intention is then to figure out at which point it is best to transition between the metaphors, making a switch between them a fluent experience.



*Fig 1.1 : hypothesis on metaphors used*

Fig. 1.1 illustrates the hypothesis about when the metaphors take effect. The assumption is that while holding the device flat the airplane metaphor will be dominant. While in the upright position the device will be used like a steering wheel by most users with some users preferring to use said window frame metaphor or a mix of both. In any case, airplane and window frame can never be used at the same time since they both rely on turning the device around the same axis but are otherwise diametrical to each other.

## **1.2 Intention**

The objective of this work is to improve on the existing navigation method with its two metaphors. To do this, a user study will help to find out at which point the mental model flips from airplane to window frame depending on the holding angle of the device.

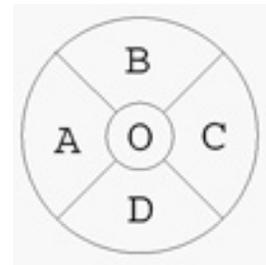
## 2 Related works

### 2.1 Intuitive spatial interfaces

#### 2.1.1 Virtual Balance and EFS

A work on intuitive navigation in virtual environments [7] by Yelistratov et.al. presents two interfaces.

The first is called “Virtual Balance” and places the user standing on a platform, maneuvering by taking steps in different directions. The system functions through three weight sensors placed under the border of the platform in equal distance to each other. This allows for a detection of the point to which the user's mass is applied on the disk. The platform was then split into five areas as seen in fig.2.1.



*Fig. 2.1 : areas of Virtual Balance disk*

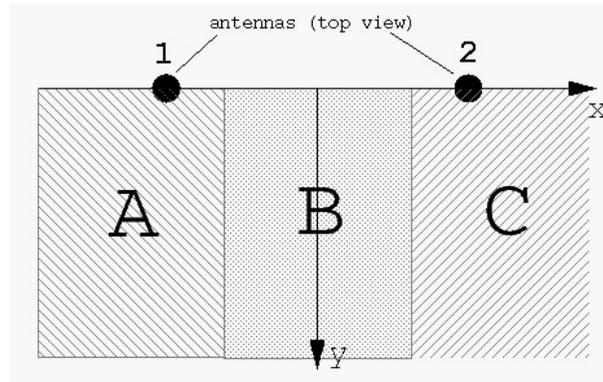
The zone in the middle represents a neutral “resting” area where no reaction was caused while standing in it. Turning is performed by stepping into area A or C located respectively at the left and right side of the disk. Leaning further away from the center of the platform linearly decreases the radius  $r$  of the turn at a given speed  $v$  and angular velocity  $\omega$  in accordance with the formula known from physics  $v = [\omega, r]$  with  $v$  and  $r$  of course being vectors. In addition the user's perspective also rotates around the speed vector by an angle with a similar linear dependency on the distance to the center. This enhances the feel of leaning into the curve and the metaphor of balancing on the disk. To adjust the viewpoint's height the user moves his weight to areas B or D.

Again, the distance to the center determines at what rate the perspective changes. Area B to the front results in a decent and also rotates the perspective to look down. Area D increases the perspective's height combined with an upwards rotation analogously.

Since all four areas are now taken for turning and height adjustment the speed of motion has to be changed differently. The first concept has the speed depend on the height of the viewpoint. Since interesting information is usually placed on the ground in a virtual environment it makes sense to travel with slower speed in proximity to the ground and increase the speed when a perspective higher up with better overview is taken. The second solution to speed control uses a combination of area B and D. Defining a time window of one second (or a similarly small range) that the consecutive actions had to be undertaken in, the sequence of leaning to one area followed by immediately moving to the other one makes a change of speed possible. If started by leaning back to area D then going to B the speed would increase as long as the user stayed in area B. After leaving B all the areas resume to their previously described mapping of turn and height. Speed is decreased by reverse order of actions traversing from B to D. Studies showed that the second approach is not as intuitive and requires some time to adapt to.

The second interface uses electrostatic field sensors (EFS) to recognize movements of the user's hand that is put between them.

Without going into too much details about the physical principles of the system, it is based on the Theremin, a musical instrument invented by Lev Termen in 1919-1920. The frequency of an oscillator is changed by bringing any mass into the vicinity of a spherical antenna linked to the oscillator. The mass changes the electrostatic field of the antenna depending on how far it is located away from the center of the sphere. By using two antennas and analyzing the shift in frequencies of both circuits the position of the mass – or in this case the user's hand – can be computed as an intersection of isolines.



*Fig 2.2 : areas of the EFS*

The position data received that way was split into an  $x$  and a  $y$  component and the working area of the antennas was also split into three zones (see fig.2.2). The  $y$  coordinate was solely used to determine the speed of motion – becoming linearly faster as the the hand approaches  $y = 0$  (plane through the two antennas). Turning is distributed to zones A and C. Again, the distance to the border with area B measured by  $x$  decreases the radius in a linear matter.

In testing, both interfaces showed very good acceptance through all ages. Users would grasp the concept within a minute.

### 2.1.2 Space Actor

The Space Actor project of the University of Arts and Design Aargau, Switzerland, resulted in two navigation interfaces for virtual environments [3] when it was presented to industrial design students.

One primary focus of the design process was to comprehend the input devices developed as an extension of the human body leading to so called extended body schema, a term to describe habitual non-conscious body actions that – after being learned – feel most natural. Each input device also offers two modes of motion with a fluent and integrated transition between them. The modes primarily change with the way the user holds and moves the devices. Tracking of the devices is done via ART infrared tracking returning 6 degree of freedom (DOF) poses. The devices themselves were also accompanied by a head tracking target.

The *NaviStick* consists of two handles with 6DOF tracking targets attached to each. The two types of travel techniques offered are a flying motion and a motorcycle driving motion that follows the terrain. Switching between them is based on the distance between the grips and their angles to the world coordinates.

Taking one handle into each hand, the flying mode is activated by spreading them apart more than 0,6 meters. This somewhat reminds of the motion of spreading wings levitating the user off the virtual ground. In this mode, speed is controlled by how far the handles are extended away from each other. A bigger extension increases the flight speed whereas bringing them together reduces speed. The heading of the viewpoint reacts to the heading of the axis through both handles. Therefore turning the body translates to turning the perspective. In order to look up or down, both grips need to be pitched up or down respectively.

To assume the driving mode the grips are held together like holding the handle bar of a motorcycle. This also drops the viewpoint to the virtual surface. This time the travelling speed is controlled by the relative rotation angle of the right handle compared to the left handle, again enforcing the metaphor of a motorcycle by imitating a gas knob. By leaning to one side, thereby rolling the axis described by the handles against the horizontal plane, a turn to that side is performed. To again improve the feel for the metaphor the change of heading during a turn is bound to a rolling motion against the speed vector as a motorcyclist would experience it when leaning into a curve.

The *Kwisath* resembles a 1,8 meter long walking stick or lance. Like the *NaviStick* it features a flying and a ground based travel mode. This time a button is placed in a convenient position that is used to initiate a starting position for flying mode.

In flight mode the *Kwisath* is balanced in a horizontal manner besides the body, just like a lance. When the button is pressed the movement of the front tip with the tracking target is mapped to changes of the viewpoint. Lifting the tip lets the user increase height, sideways rotation turns the heading of the viewpoint and pitching the tip up or down translates to the virtual pitch. Pushing it forward increases speed the further it is pushed.

Walking mode is enabled as soon as the end tip touches the ground which also serves as an initialisation for the null position. This mode is again limited to movement on the terrain. With the stick balanced on the ground it acts like a huge joystick. Tilting it in any direction results in a translation in that direction including sideways strafing – the higher the tilting angle, the higher the speed. Heading of the virtual camera is changed by turning the *Kwisath* around its axis.

The fluent transition between two metaphors is something that connects the concepts of *NaviStick* and *Kwisath* with the interface presented here. However, the separation between both modes is more present with the *NaviStick* or *Kwisath*. In the case of driving/walking mode the user is bound to the terrain and only three DOF of the virtual world can be influenced. In contrast, the car and airplane metaphor of the interface discussed here influence the same amount of DOF and the separation of the two modes is not well defined but instead a mix of both is often used at the same time.

The comparison of the airplane metaphor to a new window frame metaphor that seems to be necessary is even harder since they work against each other without yet a clear cut between them.

## 2.2 Design Issues

In the work “A Survey of Design Issues in Spatial Input” [5] Ken Hickley et.al. examined several 3D space input interfaces and put together some design concepts that would improve the users ability to experience virtual space.

The first point they mention is *spatial references*. This refers to the concept of a real-world object serving as a reference point for gestures. Giving the user a real object that reflects a virtual counterpart should make the task of manipulation “natural and effortless”. For the work on hand, the user's own body could be seen as a real-world representation of the virtual camera that is manipulated and in combination with the tablet held in his/her hands the gestures performed with in relation to it.

Concerning *relative vs. absolute gestures*, the investigation suggests that relative gestures show better results and are easier to handle. This is reflected in the tablet interface as well since the gestures are relative to a starting pose defined by the user. People have a very good feeling of where their hands are positioned relative to their body. So moving from an experienced starting pose or coming back to it should be very natural to the user and should need minimum cognitive effort.

By using *both hands*, the interface further improves the feel of presence for the user. It is directly stated that “... in essence the user's own body becomes a spatial reference”. They concluded that interfaces requiring both hands also make it less likely for the user to become disoriented compared to one-handed interfaces.

*Multisensory feedback* also plays a role in understanding virtual space. This is a point not yet examined in the interface presented here and could be the topic of further investigations. The possibility of vibration could be used to give haptic feedback to the user about how fast he/she is travelling.

*Physical constraints and affordances* make it easier for the user to identify the possible means of interaction. In our case, the physical layout of the tablet already suggests how to hold it with both hands. Together with physical limitations of the body and wrist joints it leaves only a certain amount of convenient movements that can be undertaken. A navigation interface with steering metaphors that are easy to grasp should cover all of these possibilities in a meaningful way without including counter-intuitive actions or artificial limitations through the software.

*Head tracking* is a good way to make the interface even more user centered. This is not yet realized in our interface but is well imaginable. With head tracking the tablet could not only serve as an input method in relation to the user's body but – vice versa – the body could be used in relation to the tablet.

Another proposition is the *reduction of degrees of freedom* where they are not needed. In this case, rolling in the local coordinate system of the camera can be considered unnecessary. It offers no benefits to the task of navigation since losing orientation is quite easy when unintentionally turning upside-down, even for experienced users. Another interesting point is the use of sideways and height translation and how they affect performance in different situations. Two such distinct cases could be close range movement and examination versus free travel flight and exploration.

They also mention the importance of a *recalibration mechanism*. A user shouldn't have to constantly concern himself/herself with fitting into a global coordinate frame provided by every tracking system on the lowest level. Instead the user should be able to define his/her own logical and comfortable starting point. Our interface realizes this concept through the necessity of pushing a button on the tablet in order to initiate a steering maneuver. The button can be released any time resulting in a complete stop. Then a new posture can be taken to continue working from there.

## 2.3 Preceding Work

This work expands on the work done by Nicolas Heuser in his diploma thesis [4].

In his work, he pursued the idea of a hand-held device (a Samsung Q1 Ultra PDA) and using it for gestural spatial input to control the camera in a 3D virtual environment. Users would operate standing in front of a screen or projection image with comfortable space around them. The PDA is held in both hands. It is tracked by an ART infrared tracking system that provides 6DOF – position as well as rotation – in a global coordinate system.

Since metaphors procure simple and intuitive understanding without much explanation and a chance to relate to already trained every day behaviours, skills and observations, it seemed logical and beneficial to base the interface on fitting metaphors. The two metaphors of a car's steering wheel and the wings of an airplane were found to fit the layout of the device and the way it could be held. A transformation from the global room coordinate system of the ART system to a local coordinate system fit to the screen of the PDA was implemented to make the use of device centered metaphors available.

All changes to the virtual viewpoint were mapped to movements of the device relative to a starting position. To initialize the starting position a button on the device had to be pressed. As long as the button was held down the controls remained active.

During evaluation of the interface, the question occurred if the two metaphors couldn't be used in combination without the user having to explicitly decide which one he wants to use at the time. The target for this was to not influence the user in his pre-existing mental models and let him/her freely choose the way of steering best suited to him/her.

The details of this navigation interface will be discussed in the following chapter 3.

## 3 Navigation Interface

### 3.1 General idea

Like Weiser predicted in 1993 [6] smartphones with touch surfaces (his *pads*) are nowadays very common. The same development is just now starting for tablets with products from Apple, Asus and Samsung to name some. As tablets become articles of daily use they will be quite affordable and lots of people will already be familiar with them. Since they are designed to be held in hand and carried around with wireless network they easily lend themselves to be used as an input device for a spatial 3D navigation interface.

A goal for almost every novel spatial input interface is to give best possible freedom to the user. A tablet PC light enough to be held comfortably even for longer periods of time realizes this intention very well. It can even be put aside when not needed and quickly picked up again.

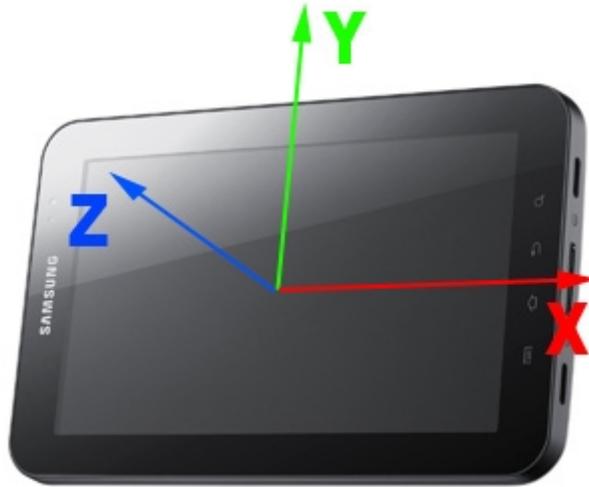
Another benefit of the tablet is the availability of a 2D multi-touch screen. Any necessary system tasks can be easily realized in a traditional WIMP fashion. This should be much easier to use than a 3D menu with ray-casting selection or similar techniques and in the case of virtual simulation environments changing numerical variables is no problem at all.

### 3.2 Functionality and metaphors

As mentioned in chapter 2.3 the tablet is equipped with a target for ART optical tracking. Through a button-like press command the control is initialized and the pose at that moment is saved as the starting position. Virtual movement happens by dragging the tablet away from the starting position. Releasing the button forces an immediate halt of any virtual motion. The steering interface as a whole is a relative system as the user needs to be able to face the screen in front at all times. The metaphors established only concern the rotation part of the controls.

In order to change the three DOF of translation for the virtual perspective (up/down, left/right, forward/backwards) the position change of the tablet in the global coordinate system is analyzed. That way, independent of the rotation of the device, the position change is held consistent in its relation to the display screen. While standing directly in front of the screen, pushing the device forward and towards the screen moves the camera forward in the viewing direction. The further the device is pushed away from the starting position, the faster the movement becomes. Height and sideways movement are controlled in much the same way along the up-axis of the room (height) and the axis parallel to the screen plane and orthogonal to the up-axis (sideways drifting) respectively. The *speed* of motion for each component is determined by a quadratic equation including the position change along the corresponding axis (*axisTranslation*), a factor increasing the speed with the height of the virtual view (*fVirtHeight*) and a sensitivity factor (*fSensitivity*):

$$speed = fSensitivity \cdot fVirtHeight \cdot (axisTranslation)^2$$



*Fig. 3.1 : local coordinate system of the tablet*

As for the heading and pitch of the virtual camera, the local coordinate system of the tablet as seen in fig.3.1 is considered. The origin lies in the middle of the display. The pitch around the x-axis of the tablet is mapped to the pitch of the virtual camera. The heading of the camera is influenced by both the heading around the y-axis and the roll around the z-axis of the tablet. This is where the metaphors of steering wheel and airplane wings take effect. The heading part is interpreted as the airplane metaphor while the roll translates to the car metaphor. As mentioned before, both metaphors should be usable in combination. Two different methods of bringing them together exist.

One has the effective parts of the metaphors linearly depend on the angle between the tablet's local z-axis and the up-vector of the global coordinates. Holding the device horizontally flat the angle  $\alpha$  is  $0^\circ$ . In this posture the influence of the *airplane* metaphor predominates for most users ( $0^\circ =$  "full airplane") and therefore the tablet's rotation around its y-axis should have the bigger influence on the virtual heading.

In an ideal upright holding position  $\alpha$  is  $90^\circ$  and the *car* association more present.

Using a tilting factor  $t$  that maps  $\alpha$  to  $[0;1]$  the equation then looks like this:

$$\textit{heading} = t \cdot \textit{car} + (1-t) \cdot \textit{airplane} \quad \textit{with} \quad t = \frac{\alpha}{90^\circ}$$

The other one simply adds both dues of *airplane* and *car* metaphor:

$$\textit{heading} = \textit{car} + \textit{airplane}$$

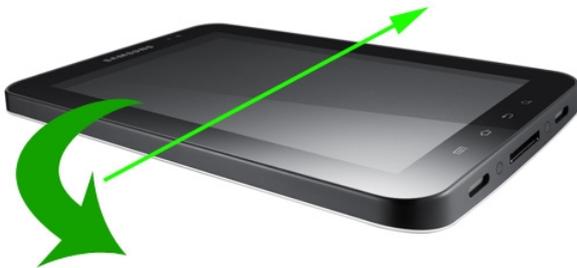
This second approach has proven to be very robust and comfortable for the user.

So the 6 DOF of the physical tablet were mapped to only 5 DOF of the virtual space. This makes sense since, as stated in chapter 2.3, the degree of freedom of rolling around the forward vector of the virtual camera results in unnecessary cognitive load for the users of systems similar to this one. It is rather cumbersome for them to have to manually align with the horizon.

### 3.3 Problems

After testing the first unintuitive reaction of the controls that could immediately be identified concerns the airplane metaphor. In the case of holding the control device in the position horizontally flat “meant” for airplane mode it works fine. When holding the device upright however, the user's expectation and his metaphorical understanding changes.

The concept of linear combination of car and airplane metaphor would partly avoid this problem since the airplane part loses effect the more the device nears the upright position. The metaphors are less likely to counteract each other. But this also means that the holding angle of the device would exclude the metaphors to some extent up to disabling them. Moreover, the additive approach proved convenient. It doesn't constrain the user and the metaphors can support each other to perform sharper turns without limiting the possibility for vernier adjustments.



*Fig. 3.2 : flat airplane*



*Fig. 3.3 : upright airplane*

Fig.3.2 and fig.3.3 shall illustrate the problem of introducing a necessary third metaphor. According to the current implementation of the airplane metaphor holding the tablet flat and turning it counter-clockwise around the y-axis results in a left turn (fig.3.2), as expected of a plane leaning its left side downwards during a left turn. In the upright case, the same motion of rotating counter-clockwise around the y-axis would mean pulling the right side towards the user while pushing the left side away (fig.3.3). This has no metaphorical reference to a left turn but much rather the opposite. The device is imagined to be like a window frame looking into the desired direction of movement. Such a rotation would then suggest a right turn instead of a left one.

## 4 Technical realisation

During this work, the setup of hardware and software was changed.

The first concepts were designed under the version Nicolas Heuser already used for his work. This is a Windows version running with DirectX in the laboratory of the Chair for Augmented Reality of the TU Munich. Later on the work was ported to the FRAVE system where the software is running under Linux and OpenGL. The tracking system for the FRAVE also differed from the one used in the laboratory.

The plan was to be able to do the user study in the FRAVE setup. Besides being more attractive the FRAVE is one example of an environment where the interface is ultimately supposed to be used in. Unfortunately, due to persistent graphic and networking issues outside the scope of this work the plan had to be given up.

### 4.1 Hardware

#### 4.1.1 Tablet

The tablet used here is a Samsung Galaxy Tab. It features a 7 inch touchscreen with a resolution of 1024x600, wireless network and runs on Android 2.2. The size is 190 x 121 x 11,98 mm and it weighs 380g.



*Fig. 4.1 : tablet with ART target*

#### 4.1.2 ART tracking system

Tracking systems are used to measure the position of bodies in a defined space.

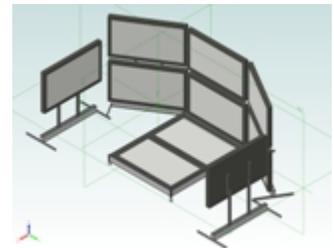
ART's (Advanced Realtime Tracking GmbH) optical tracking system [1] uses two or more cameras covering a tracking volume. The cameras send out flashes of infrared light. These flashes are retroreflected back at the cameras by small spherical markers attached to the body that needs to be tracked. With the cameras catching the reflected flashes a computer synchronizing the timing of the data can determine 2D positions of markers for each camera picture. Through a room calibration the configuration of the cameras in position to each other is known. By merging the data of all pictures a 3D position of markers can be derived. With three or more markers building an ART target a 6DOF pose is then calculated for the target. To counteract possible occlusions of markers more than three should be preferred.

Advantages of ART tracking are a high precision (in the range of millimeters) and minimum lag (running at 60Hz).

The ART system installed in the laboratory consists of four cameras running with version 1 of ART's software. For the FRAVE eight cameras are installed and software version 2 is employed. The tablet's ART target is composed of five spheres and mounted on a commercial casing for the Galaxy Tab increasing the weight marginally. It was attached so that it would not hinder the movement of the user.

### 4.1.3 FRAVE

The FRAVE is a highly configurable VR setup developed at the Chair for Augmented Reality of the TU Munich as an alternative to a traditional CAVE. It is composed of up to ten screens all used to make multiple views of the same virtual environment available. The OpenGL library equalizer makes the parallelisation of image data possible.



*Fig. 4.2 : screens of the FRAVE*

## 4.2 Software

### 4.2.1 Rendering software and system

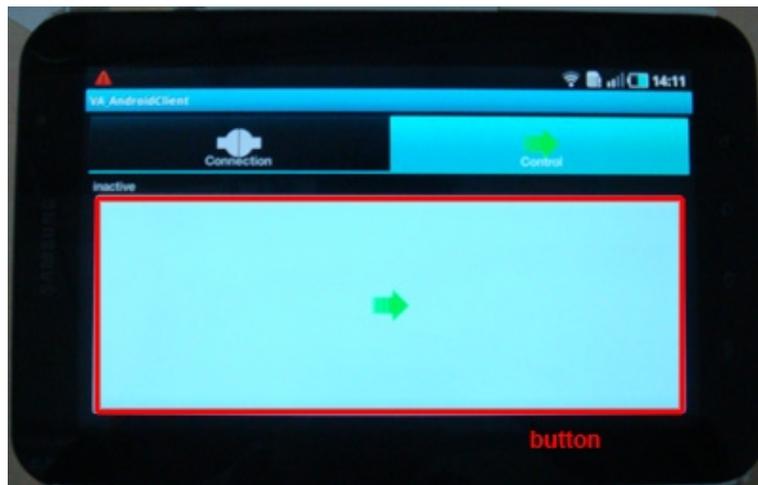
The software called Terrain3D used for rendering the terrain is developed by the Chair for Computer Graphics at the TU Munich. The terrain data itself is a part of The US state Utah generously provided by the same chair.

The software was run on a Windows 7 machine with an Intel Core2 6600 at 2,4 GHz, 2 GB RAM and an ATI Radeon 2900 XT graphics card.

Displaying the virtual environment happened with a projector on a resolution of 1280x1024. The projected screen was approximately 2,20m high and 2,75m wide. The user was placed 2,70m away from the screen and had the horizon of the virtual landscape at eye level. This provided a convenient viewing experience.

#### 4.2.2 Android client

In its first iteration the Android application for the Samsung Galaxy Tab used to connect with the visualization software was developed for the DirectX version of Terrain3D. The interface was held very simple.



*Fig. 4.3 : Android client interface*

The former concept of a physical button being pressed to initialize the steering was replaced with a relatively large area on the touchscreen that registered a touch event to induce the same effect (red box in fig.4.3). The thumb(s) can be placed anywhere in that area to activate the navigation. On activation the button also changes appearance.

Above the button area a status text is shown giving feedback on the wireless connection to Terrain3D and also on the activation.

A second iteration of the Android client application was developed to be consistent with the Terrain3D version running on the FRAVE setup. This iteration of Terrain3D had an Android server class that managed communication with another navigation interface developed in parallel to this one by Mohamed Ashry [2]. It uses a smartphone held in one hand and a combination of phone based accelerometers for rotation and touchscreen for translation. Aiming to later compare these two the second Android client software merged the smartphone approach with the ART based input method so the mode of input could be chosen on the device.

### 4.2.3 Porting to FRAVE setup

The software components managing the ART data on the Terrain3D end also had to be ported and were included in the Android server class to encapsulate tracking and communication with the navigation device. With every update of the virtual camera an update of tracking data is also called and the Android server decides through the input method chosen on the tablet/smartphone which data he provides to the camera class.

## **5 User study**

### **5.1 Primary goals**

The closing user study in the work of Nicolas Heuser showed that the two metaphors implemented and their additive combination did not result in a satisfactory control interface for all users in terms of intuitivity. The goal of the user study and for this work was to identify when and how the controls reacted differently from the expected. First of all, this meant examining where the airplane metaphor transitions into the window frame metaphor and pinpointing the angle at which the tablet was held at that time.

### **5.2 Choice of probands**

There were 12 probands in total. The test users were aged between 17 and 27 years with an average of 22,08. Three of them were female. All but one had a driving license and almost everyone had experience with 3D computer graphics and videogames and nine had experienced 6DOF input devices before.

## 5.3 Design process

All conceptions discussed in the following are based on the task of following a path in virtual space. This path was at first visualized in the form of a blue band. To make the band more viewable the plane through the camera's position and a section of the path was used. The band could then be placed along the plane's normal vector in that section. This way the band would always “face” the camera.

### 5.3.1 First approach

The first idea was to use the same test track already designed in [4]. The course features several curves of different radii together with parts that required flying up and down.

The first effects of interest, the point of flipping between airplane and window frame metaphor, would occur specifically during a left/right turns. However, the complexity of the test track would make it hard to conclude whether any curve executed corresponded to a right or a left turn on the track. Any unexpected reaction of the controls that diverted the user away from the desired path enough so he/she would have to make an effort to return to the track would include additional unintended turning. Considering the length of the track, this would confuse any actual matching of intended curve and actual reaction of the steering system.

A similar problem would occur if the user took shortcuts or even just cut turns. This could happen after a failed steering maneuver that left the user closer to a different part of the track or the whole setup seeming like a time trial.

A more simplistic approach was needed so the specifics of a turn (direction and the course of height level) could be known during its passing.

### 5.3.2 Second approach

The second approach would feature the original test track only for training purposes so the user could familiarize with the controls. The assessment of the controls would then happen with single curves in different variations. The curves would alternate between left and right turns with the assumption of this having insignificant effect on the way users steered but having the advantage of showing more variation to the the user and limiting monotony of the task. The curves would vary in their size of radius but otherwise be flat. In order to get data for all possible ways of holding the tablet, the range of holding angles between horizontal/flat and vertical/upright position was split into three intervals. The angle between the z-axis of the tablet's coordinate system pointing out of the screen and the up vector of the room calibration was computed and the range of below 33 degree was taken as *flat* holding position. An angle of 33 to 66 degree meant *tilted* position and over 66 degree was considered to be *upright*. The users would then undergo three runs, each with one fixed starting position. During these runs the control would not activate if the device wasn't held in the given range of degree. The task for the users consisted of completing a randomized iteration of all curves for every starting position.

However, there was another element still missing in this approach. To be able to determine whether the metaphor changed “on the fly” during up and down curves, there needed to be curves with enough up or down pitch so the user would start in one position but then change the angle of holding before going into a curve and keeping this angle throughout a sufficient part of the curve. This would give a metaphor different from the one usually associated with the initial starting position a chance to occur. For example, a user would start in the *flat* position. A left curve

directed upwards should then first require the user to rotate the tablet into the range considered *tilted* followed by a steering maneuver to the left while staying in the *tilted* position.

### 5.3.3 Third approach

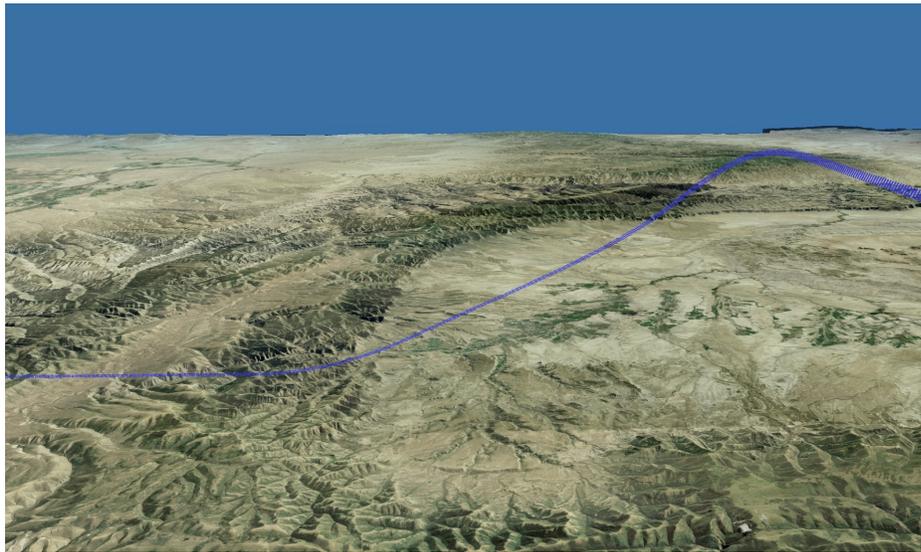
Modifying the second design, the study started with a test and training track seen in fig.x. The first four phases all used the training path and had the intention of familiarizing with the initiation of steering and the general concept of the navigation.



*Fig. 5.1 : training path*

For the main part the users were introduced to the concept of fixed starting positions. The experiment was split into three big cycles, one for each starting position (*flat*, *tilted* and *upright*). These cycles were permuted in equipartition over all test subjects to avoid learning effects. If the user tried to start the controls in a pose not matching the required one for this cycle the system gave a short acoustic feedback and the controls did not activate. In a prior training run the users completed three curves – one flat, one upwards and one downwards – for each cycle.

Having made sure all concepts were introduced and the users had enough time to familiarize with the controls, they were then presented with the full range of curves for each starting position. The cycles for fixed initiation positions came in the same order as for the training curves. The first two curves were always flat – one with a wider and one with a smaller radius. The remaining ones were randomized in their sequence.



*Fig. 5.2 : downwards left curve*

Trying to investigate if the change of holding angle during flight had any influence on the used metaphor, two more variations of curves starting with an up or down pitch of 30 degrees were included. A combination of a) small or big radius and b) amplitude of pitch (flat curve, light or steep ascent and light or steep descent) defined a curve, so there were ten curves in total.

As the users navigated through the different phases notes were taken whenever an unintended steering maneuver occurred. This allowed for future reference with the recorded data.

## **5.4 General adjustments to the interface**

The overall sensitivity of the interface for translating as well as rotating motions was set to 70% of the usual values as probands of the previous iterations complained about the interface being too hectic. After finishing the experiment in its final version some users wished for a more sensitive steering. This shows that the usual behaviour fits nicely for trained users but for the sake of easing into the experiment the factor of 0,7 was kept.

Since the primary goal was to explore the turning behaviour of users any sideways or height translations were disabled for the course of the experiment. Only the forward and backwards translation for de-/acceleration was left so the users had to rotate the device in order to successfully make a turn. If users would have been able to complete an upwards directed curve by elevating without adjusting the pitch of the physical device the purpose of the study would have been defeated. The same goes for sideways drifting motions.

One final issue concerned the visualization of the paths. As the interface provides only visual cues to the current travelling speed through the landscape and the sky has just one color and no clouds it was sometimes hard for the users to tell how fast they moved or if they moved at all. This especially was the case for curves directed upwards where for most part of the curve the perspective would solely show the blue of the sky. Other means for motion cues had to be found and finally the appearance of the path was changed from a continuous band to a segmented version with lines across the direction of the path (see fig.5.2). This helped the users judge their own speed.

## **5.5 Results**

### **5.5.1 Analyzed data**

To make sure all the data necessary and of interest was captured the recorder component of Terrain3D implemented by Nicolas Heuser was readjusted to include:

- the current angle between device z-axis and room up-vector,
- the initial angle to the up-vector at the time of the button press,
- the relative definition of the three components of heading, pitch and roll of the tablet around its local coordinate axes (reflecting airplane metaphor, pitch and car metaphor respectively)
- the fixed starting pose defined at that moment

The absolute pose of the steering device in the global coordinate system as well as the pose of the tablet relative to its starting position were also recorded with regard to further analysis but could not be considered here.

### **5.5.2 Analyzation program**

The program written to help sum up and analyze the recordings concentrates the data into steering activities.

One steering activity lasts from the next occurrence of a button press to the release of the button. Over that period the two rotation angles representing the car and the airplane metaphor are each divided into their components that result in the camera's perspective turning to the left or right and summed up. The resulting four values ( $sumAirplaneLeft$ ,  $sumAirplaneRight$ ,  $sumCarLeft$ ,  $sumCarRight$ ) are saved as percentage of the total steering activity for the metaphors.

$$sumOverall = |sumAirplaneLeft| + |sumAirplaneRight| + |sumCarLeft| + |sumCarRight|$$

$$percentageAirplaneLeft = \frac{|sumAirplaneLeft|}{sumOverall}$$

In the code it is also left if the steering activity happened during a left or right curve of the experiment. In this manner, a comparison between supposed turning direction and actual turns allows a categorization into percentages for each metaphor that went in the wrong direction and those that had the correct effect.

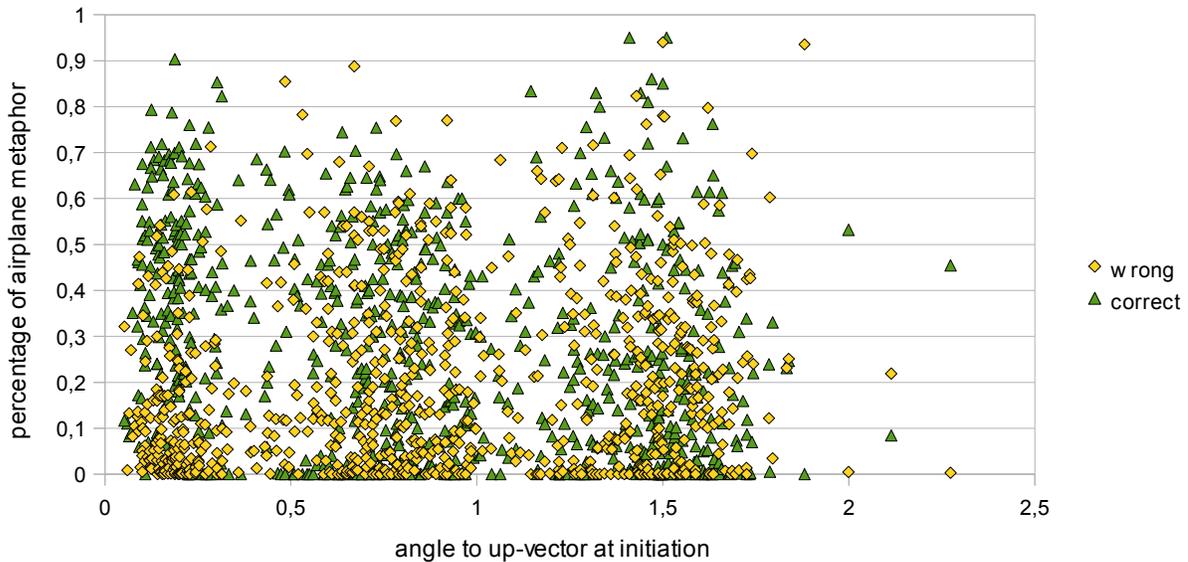
The pitch value is summed up and saved in a similar fashion so ascending or descending curves can be identified.

As a last point of consideration the time the steering activity lasted is saved.

### 5.5.3 Evaluation

First it has to be noted that the car metaphor never resulted in unexpected changes of the viewpoint when utilized. Therefore the airplane metaphor will have the focus in the following evaluation of gathered data.

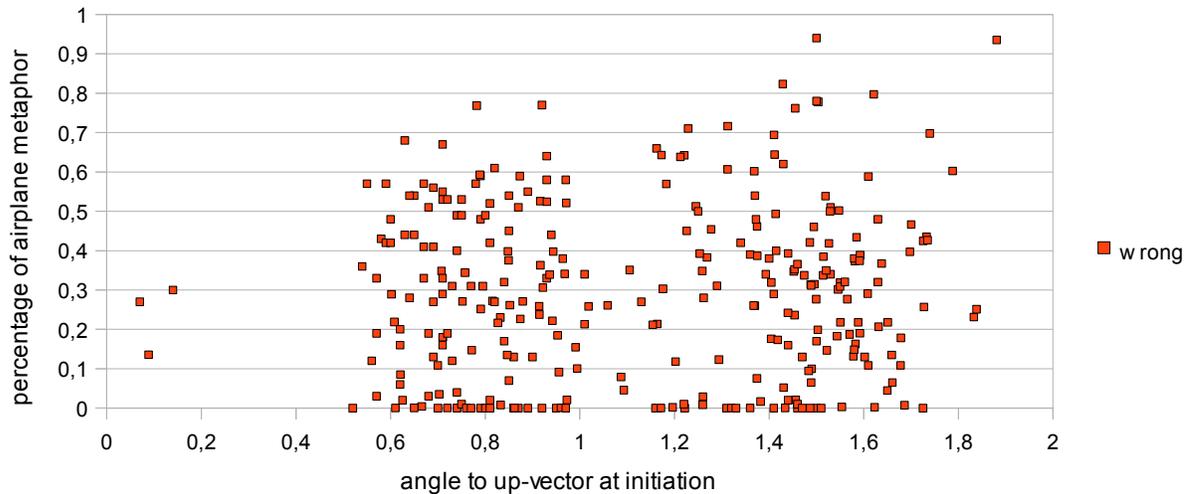
If not specified otherwise all angles given are in radians.



*Fig. 5.3 : overall statistics for correct and incorrect airplane*

A first look on an aggregation over all participants and steering activities of the experiment (fig.5.3) does not yet reveal too much. The noise from small corrective activities that aren't necessarily headed in the wrong direction to a considerable amount seems to be too much. Such activities – usually very short in time – can still show a high percentage for wrong direction as the proportion is calculated within every single activity. They would need to be weighted with their duration. This however would also influence steering maneuvers that really do go in the wrong direction but are immediately aborted by the user. If at all it can be noted that correct responses of the airplane metaphor mass at an angle below 0,5 radians whereas the number and intensity of inappropriate reactions seem to grow with the angle the device is held at.

As the phases explicitly containing unwanted reactions by the interface were noted on the basis of observation throughout the tests a down-focus on these phases seems more promising.



*Fig. 5.4 : incorrect airplane of phases noted to be problematic*

This approach (fig.5.4) shows a clearer picture. It can be seen that starting from an angle of 0,5 (~29°) unwanted influences of airplane rotations appear in extensive amounts.

The next question is whether this distribution stays true for all users or if there are differences in the way the airplane metaphor is handled.

For the reason of a clearer overview the next three statistics on the following pages split the probands into groups of four.

Looking at the data, the conclusion can be drawn that the conception of where the airplane metaphor ends differs with the users. For proband 1 the line seems to be at the angle of 0,5. Proband 3 and 10 seem to agree on a range of 0,6 to 0,7 (34° - 40°). User 5 and 7 are closer to 0,8 (~46°) while proband 6 only starts at 1,2 (~69°) and 2, 8 and 12 no sooner than an angle of 1,4 (~80°). Test person 4, 9 and 11 avoid the use of airplane in its divergent form altogether.

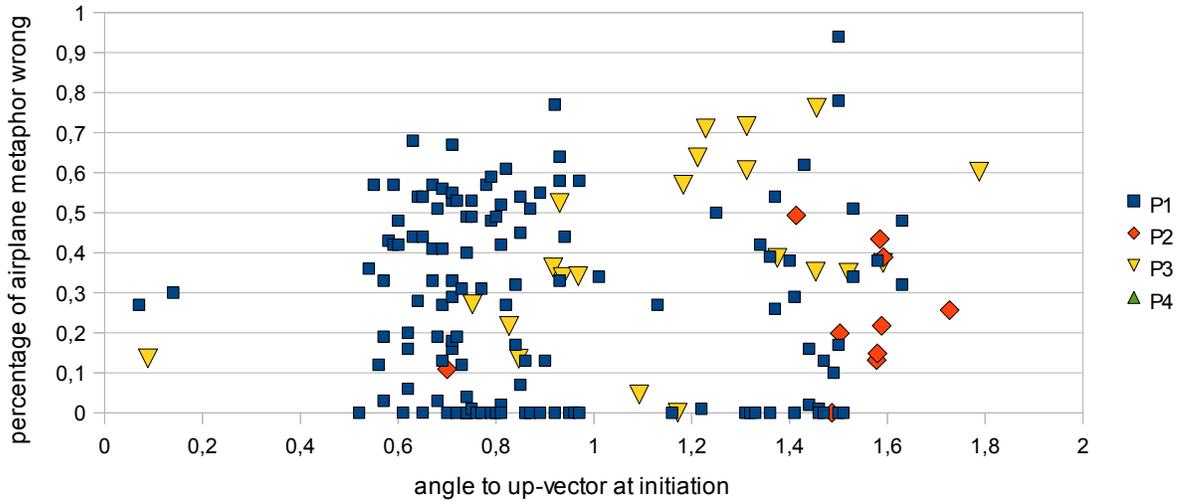


Fig. 5.5 : probands 1 through 4

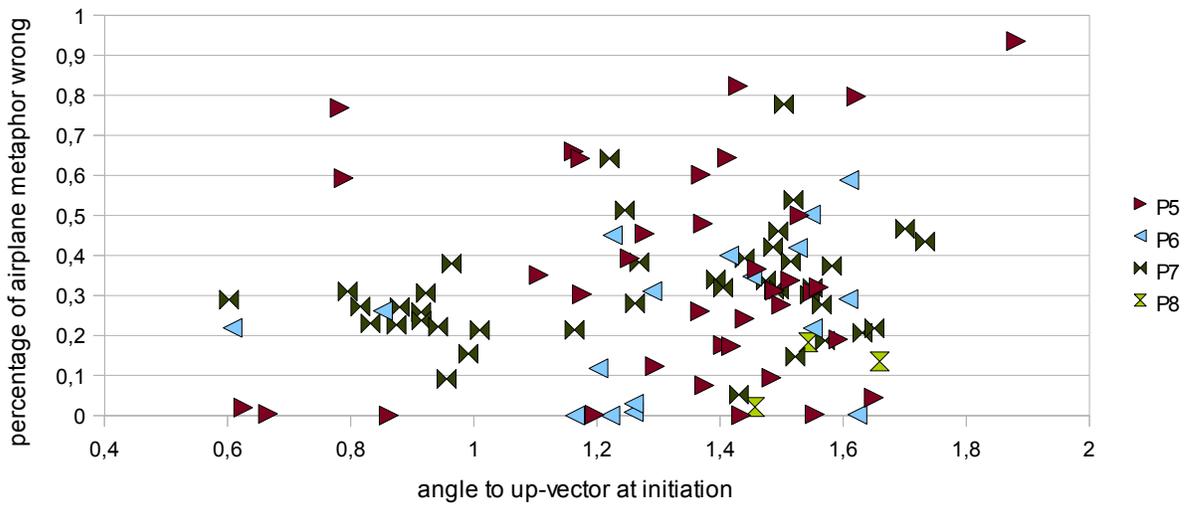


Fig. 5.6 : probands 5 through 8

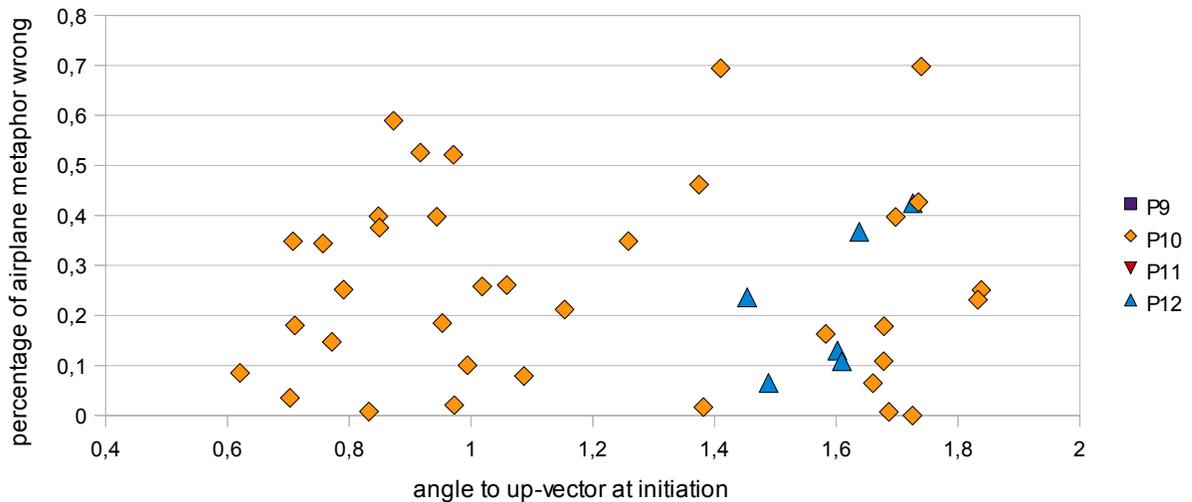


Fig. 5.7 : probands 9 through 12

The next question must be up to which point the airplane metaphor is used in a beneficial way by the individual user and if this concludes to sharp lines of division with no real converging point over all users.

In this respect the cases of proband 2, 8 and 12 with whom the failure of the airplane metaphor occurs at a relatively high angle are now examined. If they show to use airplane steering with success well over the range of angle where the metaphor starts failing for other users then the plan of devising a threshold for the flip of metaphors has failed.

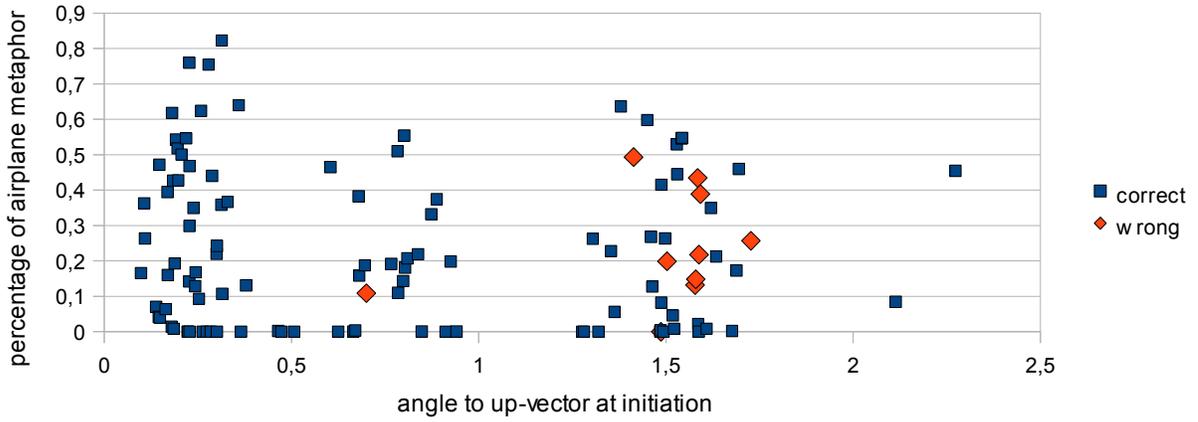


Fig. 5.8 : proband 2

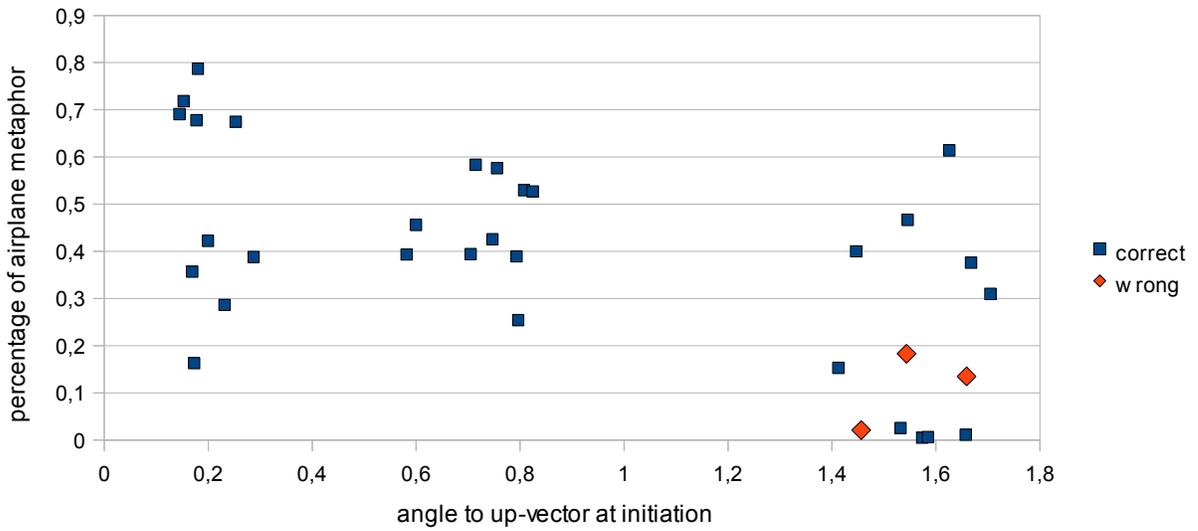


Fig. 5.9 : proband 8

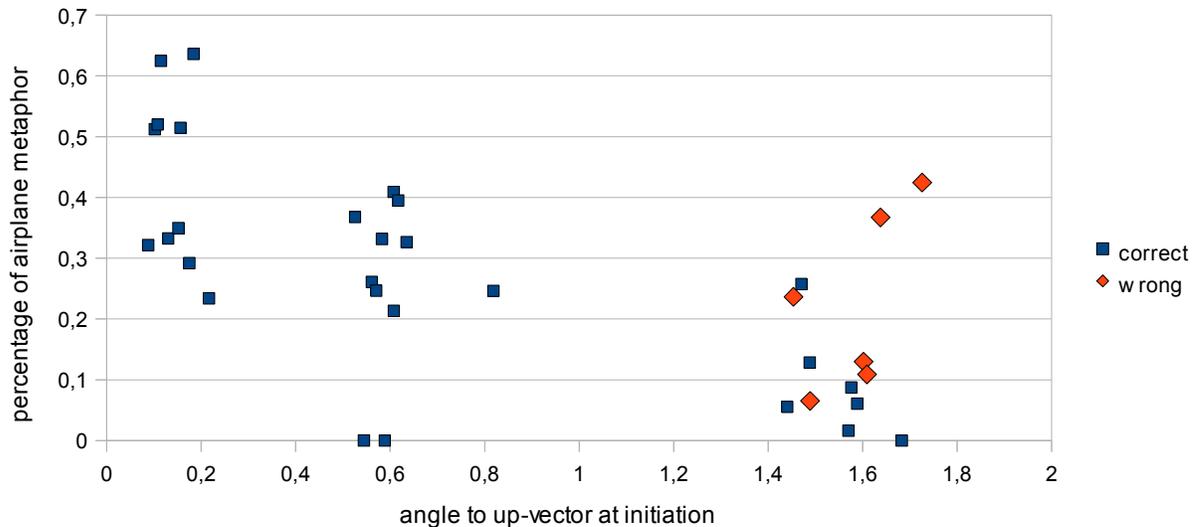


Fig. 5.10 : proband 12

Interestingly enough all three of them seem to use airplane mode with effectiveness up to an angle of 0,8 to 0,9 followed by a gap of airplane usage. The parallel occurrence of both positive and negative employment of the metaphor can be safely interpreted as the user realizing the unexpected inversion of the metaphor and executing a counterrotation in order to compensate. This has been observed for almost every user in the study and phases in which the user intentionally used the inverted airplane steering to perform the right turn because he learned to work with it – which did not happen often but did on a few occasions – were excluded here.

With these results placing the parting line between airplane metaphor and a new window frame metaphor at an angle around 0,8 radians should improve the experience of navigating for most of the users.

## 5.6 Conclusions

Working with the above results a third combination of metaphors was implemented. It acts as a merge of the additive variant and a sharp transition between airplane and window frame. The new formula for the virtual view's heading is this:

$$\begin{aligned} \textit{heading} &= \textit{car} + \textit{airplane} , \text{ if } \textit{initialAngleToUp} \leq 0,8 \\ &\textit{car} - \textit{airplane} , \text{ if } \textit{initialAngleToUp} > 0,8 \end{aligned}$$

## **6 Discussion**

### **6.1 Nature of the metaphorical conceptions**

During this work the question arose if a combined use of the three metaphors of car, airplane and window frame is approaching one single metaphor. It could be called flying carpet or glider and resembles the way children would fly around a toy held in their hands. The change of viewing direction would depend on a vector pointing from the center of the tablet in the viewing direction to the middle of the virtual perspective. It would not depend on the angle the device is held. An initialisation pose would still be used to calculate any rotations of the device and the vector would rotate with the device, as to point at the desired new viewing direction. This would be much like a raycasting metaphor. All mappings of the rotation would of course still be relative. A comparison would be interesting.

### **6.2 Future work**

#### **6.2.1 Use of head tracking**

The addition of head tracking could be especially useful for close examinations where the user wants to move around objects of interest, a scenario where the current navigation interface still lacks. If the the interface featured a selection and “grabbing” command, the user could then use the tablet as a spatial reference to the grabbed virtual object and his tracked movements relative to the tablet could change the view on the object.

### 6.2.2 Dependency of switch between airplane and window frame

A question unsettled is whether the transition from airplane to window frame metaphor depends on the initial starting position and its angle to the up-vector (staying in one metaphor for the duration of a steering activity) or on the up-angle at execution time (solely depending on the angle the device is held at that moment).

### 6.2.3 Strafing position changes

Another interesting point of investigation is the way in which sideways/up-down translation is expected to affect movement. It could either be interpreted as a horizontal/vertical strafing motion (fit for close range examination tasks) or as a means for the user to describe a desired curve in the air by turning his body (fitting travel flight), much in consens with the flying carpet metaphor.

## 7 Appendix

### 7.1 List of references

- [1] Advanced Realtime Tracking GmbH, <http://www.ar-tracking.de/>, last visited on 16.10.2011
- [2] Ashry Mohamed, “*Development of Android touch interface for navigation in an immersive environment such as CAVE*”, bachelor thesis, TU Munich, 2010
- [3] Bowman D.A., Kruijff E., LaViola J.J., Poupyrev I., “*3D User Interfaces*”, Addison Wesley, 2005
- [3] Doulis, M., V. Zwimpfer, J. Pfluger, A. Simon, C. Stern, T. Haldimann and C. Jenni, “*SpaceActor - Interface Prototypes for Virtual Environments*”, 3D User Interfaces (3DUI'06)
- [4] Heuser Nicolas, “*Methoden und Metaphern zur Navigation in virtuellen Landkarten*”, diploma thesis, TU Munich, 2010
- [5] Ken Hinckfey, Randy Pausch, John C. Goble, and Neal F. Kassel, “*A Survey Of Design Issues In Spatial Input*”, *Proceedings of UIST '94*, November 1994
- [6] Weiser M., “*Some computer science issues in ubiquitous computing*”, *Commun. ACM* 36, 7 (July) 1993, 75–84
- [7] Yelistratov V., Strauss W., Fleischmann M., “*Two approaches for intuitive navigation in virtual environments*”, *Proceedings of GraphiCon-99, the Ninth International Conference on Computer Graphics and Vision, Moscow, Russia, 26 August–1 September 1999*

## 7.2 Schedule for user study

# Ablaufplan

- Ziel der Studie Ansagen: „Steuermechanismen für 3D Anwendung entwickeln“
- Demographischer Fragebogen
- Aufbau erklären
  - Projektor
  - 3D Welt
  - ART-Kameras, auf Trackingvolumen hinweisen, roter Startpunkt
  - Steuergerät
- Genereller Standort  
Versuchsperson auf Ausgangspunkt stellen, Blick zur Leinwand – Ansage: „Für jeden Versuch ist das NUR der Startpunkt. Wir gewährleisten damit, dass Sie in jeder Richtung Platz für Bewegung haben. Sie dürfen nach dem Start jederzeit Ihre Position verändern, z.B. Schritt vor, zur Seite“. Versuchsperson einige Schritte durchführen lassen!
- Erster Durchgang: Pfad folgen ohne Kopplung mit Steuergerät
  - NUM 0 (neues Experiment)
  - NUM 7 = Starthaltung auf FREI
  - NUM 1 (Durchgang auf festem Pfad)
  - Ansage: „Sie sehen den angezeigten Pfad. Sobald Sie den Knopf drücken ist die automatische Steuerung aktiv. Es geht SEHR langsam los, wird dann aber schneller“
  - Vorführen (PHASE 1)
  - Versuchsperson vertraut machen lassen mit Knopf drücken, halten und loslassen –automatische Bewegung

- Reset zu Ausgangsposition – Aufzeichnung starten
- Ansage: „Sie sollen nun das Steuergerät so bewegen, als ob Sie selbst entlang des Pfades steuern. Wollen Sie eine Pause machen, lassen Sie einfach den Knopf los, Sie bleiben dann sofort stehen. Der Durchgang dauert ca. 1 min.“
- NUM 1 (Durchgang auf festem Pfad)
- Durchgang 1 (PHASE 2 – A)
- Pause/Fragen
- Zweiter Durchgang: Freie Steuerung
  - NUM 2 (Durchgang mit freier Steuerung)
  - Ansage: „Im zweiten Durchgang sollen Sie nun selbst steuern. Wie im ersten Durchgang müssen sie den Knopf drücken und gedrückt halten. Nur wenn der Knopf gedrückt ist, steuern Sie. Bewegen Sie nun das Steuergerät z.B. nach etwas nach vorne, bewegen Sie sich langsam vorwärts. Je weiter Sie sich vom Ausgangspunkt entfernen, desto schneller werden Sie. Abbremsen und stehenbleiben tun Sie indem Sie das Steuergerät wieder zu dem Punkt zurück führen an dem Sie den Knopf gedrückt haben. Lassen Sie den Knopf los, bleiben Sie sofort stehen.“
  - Vorführen – nur Translation (PHASE 3)
  - Ausprobieren lassen: Vorwärts – nur Translation!
  - Ansage: „Kurvenbewegungen liegt dasselbe Prinzip zugrunde wie den geraden Bewegungen. Je stärker Sie eine Kurve steuern, desto enger ist die Kurve.“
  - Ausprobieren lassen (immer noch PHASE 3)
  - Ca. 5 min Eingewöhnungsphase
  - Ansage: „Bewegen Sie sich jetzt wieder entlang des Pfades. Der Durchgang dauert ca. 2 min.“
  - NUM 2 (Durchgang mit freier Steuerung)
  - Durchgang 2 (PHASE 4 – A)
  - Pause/Fragen

- Dritter bis fünfter Durchgang: Ungefähr vorgegebene Starthaltungen
  - Ansage: „Im dritten Durchgang möchten wir Ihnen die Haltung des Steuergerätes zum Zeitpunkt des Knopfdrückens vorgeben. Sobald Sie den Knopf gedrückt haben, können Sie aber wieder vollkommen frei und selbstständig steuern. Folgen Sie bitte wieder dem vorgegebenen Pfad.“
  - Vorgegebene Haltungen: Horizontal, 45 Grad, Vertikal – diese permutiert GLEICHVERTEILT über alle Versuchspersonen (NUM 4 = FLACH / NUM 5 = 45 GRAD / NUM 6 = AUFRECHT)
  - NUM 2 (Durchgang mit vorgegebener Starthaltung) für jede Phase
  - Ausprobieren lassen für alle drei Starthaltungen an jeweils 3 Kurven (mittlerer Radius + flach, mittlerer Radius + Aufstieg, scharfer Radius + Abstieg (PHASE 5 bis 13)
  - Pro Haltung ein Durchgang (PHASE 14 bis 43 - A)  
Durchgang1: Phase 14-23, D2: Phase 24-33, D3: Phase 34-43
- Vierter Durchgang: Nochmal Freie Steuerung (Aufbau gleich dem Zweiten Durchgang) (NUM 2; NUM 7; PHASE 44 – A)

VP	1. Durchgang	2. Durchgang	3. Durchgang
1	Horizontal	Schräg	Vertikal
2	Horizontal	Vertikal	Schräg
3	Schräg	Horizontal	Vertikal
4	Schräg	Vertikal	Horizontal
5	Vertikal	Horizontal	Schräg
6	Vertikal	Schräg	Horizontal
7	Horizontal	Schräg	Vertikal
8	Horizontal	Vertikal	Schräg
9	Schräg	Horizontal	Vertikal
10	Schräg	Vertikal	Horizontal
11	Vertikal	Horizontal	Schräg
12	Vertikal	Schräg	Horizontal